N° d'ordre: 3877

THÈSE

Présentée

devant l'Université de Rennes 1

pour obtenir

le grade de : DOCTEUR DE L'UNIVERSITÉ DE RENNES 1
Mention INFORMATIQUE

par

Mohand Yazid SAIDI

Équipe d'accueil : IRISA
École Doctorale : Matisse
Composante universitaire : IFSIC

Titre de la thèse :

Méthodes de contrôle distribué du placement de LSP de secours pour la protection des communications unicast et multicast dans un réseau MPLS

Soutenue le 28 novembre 2008 devant la commission d'examen

M. :	Raymond	MARIE	Président
M. :	Jean-Jacques	PANSIOT	Examinateur
M. :	Guy	LEDUC	Rapporteur
M. :	Maurice	GAGNAIRE	Rapporteur
M. :	Bernard	COUSIN	Directeur de thèse
M. :	Jean-Louis	LE ROUX	Membre

Table des matières

Introduction

Avec l'explosion du nombre d'applications temps réel déployées dans les réseaux (comme la voix sur IP, la TV, les jeux en réseau, etc.), la récupération rapide lors des pannes est de plus en plus désirée pour assurer la continuité des services de communications. Différentes techniques de résistance aux pannes ont été développées pour éviter et/ou réduire le temps de coupure des communications. Ces techniques ont pour rôle de déterminer des chemins de secours capables de recevoir et de re-router le trafic des communications affectées par une panne. Elles peuvent être classées en deux catégories : *restauration* (ou *protection réactive*) et *protection* (ou *protection proactive*).

Avec la restauration [RM99b], aucun calcul de chemin de secours n'est accompli avant l'occurrence d'une panne. Le processus de récupération, déclenché après la survenue d'une panne, consiste à déterminer et à configurer de nouveaux chemins contournant les composants défaillants et capables de router le trafic des communications affectées par la panne. En plus de sa capacité à résister aux changements de la topologie du réseau, la restauration a l'avantage de décroître les coûts de maintenance et de calcul. Cependant, elle n'assure aucunement la disponibilité des ressources (particulièrement la bande passante) lors d'une panne et son délai de récupération est souvent élevé et inacceptable pour beaucoup de types d'applications réseau (VOIP, TV, jeux en réseau, etc.).

Pour éviter les inconvénients précédents, la tendance actuelle est d'utiliser le second type de techniques de résistance aux pannes qui est la protection [MVDBD04, RM99a]. Avec ce type de technique (la protection), le délai de récupération est significativement réduit car les chemins de secours sont pré-calculés et souvent pré-configurés avant l'occurrence des pannes. Concrètement, pour pallier une panne donnée, il suffira de basculer le trafic des communications affectées par la panne vers les chemins de secours qui les protègent.

Selon la localisation du routeur responsable du basculement du trafic vers les chemins de secours, nous distinguons deux principales classes de protection : protection globale et protection locale. Avec la protection globale, ce sont les routeurs sources des communications affectées par la panne qui basculent le trafic des communications affectées vers leurs chemins de secours. Pour ce faire, un plan de contrôle notifiant la panne aux routeurs sources des communications affectées est nécessaire et est prévu. Concrètement, tout routeur détectant une panne construit un message de notification de la panne qu'il envoie à tous les routeurs sources des communications affectées. A cause du délai induit par la notification de la panne, la protection globale peut présenter des

délais de récupération élevés (surtout si la panne est loin de la source) et indésirables pour certains types d'applications[1]. Afin de minimiser le délai de récupération, le choix de la protection locale qui élimine le plan de contrôle notifiant la panne s'impose dans les réseaux en général et particulièrement dans les plus larges. Avec cette classe de protection, ce sont les routeurs détectant la panne (ou les routeurs très proches de la panne) qui réagissent à la défaillance en basculant localement le trafic des communications affectées vers les chemins de secours permettant la récupération. En plus des avantages de l'élimination des messages de notification de la panne et de la réduction du délai de récupération, la protection locale évite la perte d'un grand nombre de paquets après une panne.

Avec l'arrivée de MPLS (MultiProtocol Label Switching) [RVC01] dans la dernière décennie, la protection locale a été améliorée et rendue plus efficace. Grâce à sa grande flexibilité pour le choix des routes et à sa capacité de réserver explicitement les ressources (plus particulièrement la bande passante) et à pré-configurer des chemins de secours locaux, MPLS permet de (1) réduire sensiblement les délais de récupération (jusqu'à 50 millisecondes), (2) assurer la disponibilité des ressources après une panne et (3) optimiser l'utilisation des ressources. Typiquement, la pré-configuration de chemins de secours locaux permet de réduire les délais de récupération en anticipant les pannes et les traitements associés. Ainsi, pour pallier toute panne, il suffira à l'ensemble des routeurs détectant une panne (i.e. routeurs voisins du composant en panne) de basculer localement le trafic des communications affectées vers leurs chemins de secours déjà configurés. En conséquence, aucun délai supplémentaire n'est requis puisqu'aucun message (ou peu de messages) de signalisation ou de configuration n'est envoyé dans le réseau pour permettre la récupération. De plus, grâce à sa capacité à pré-réserver les ressources, MPLS peut garantir des ressources suffisantes sur les chemins de secours afin d'écouler le trafic des communications affectées suite à une panne. Enfin, MPLS permet d'optimiser l'utilisation des ressources grâce à sa capacité à configurer et à explorer toutes les routes possibles dans un réseau pour le routage.

Contrairement à la protection locale de niveau *Liaison de données* (comme la protection par *p-cycles* décrite dans [GS98]), la protection locale sous MPLS permet de tenir compte des caractéristiques et propriétés du trafic réel protégé. Par exemple, pour mieux utiliser les ressources sous MPLS, il est possible de répartir les flux transportant le trafic en plusieurs classes de différentes priorités et de ne protéger que les communications transportant des flux de haute priorité.

Selon le nombre de pannes simultanées[2]que doivent traiter avec succès les techniques de protection employées, la quantité de bande passante (ressource) allouée aux chemins de secours peut être plus au moins élevée. En effet, le nombre de pannes simultanées détermine tous les scénarios de pannes possibles, qui à leur tour, contrôlent le nombre et les structures des chemins de secours fournissant la protection. A cause de la rareté

[1]La protection globale reste intéressante pour protéger les communications dont la source et la destination sont proches.

[2]Par abus de langage, nous disons que deux pannes successives sont simultanées si la deuxième panne survient avant la récupération complète de la première panne. Elle permet de diminuer le nombre de structure de secours à administrer.

des pannes multiples et à la difficulté de protéger contre ce type de panne, beaucoup de travaux dans la littérature ne considèrent que les pannes simples. Avec ce dernier type de panne, la quantité de bande passante allouée aux chemins de secours est sensiblement diminuée grâce au *partage de bande passante*. Pratiquement, comme il ne peut y avoir qu'une seule panne à la fois dans un réseau, les ensembles de chemins de secours qui protègent contre des pannes différentes ne peuvent être actifs en même temps : ils peuvent donc *partager leur allocation de bande passante* sur les liens en commun pour optimiser la quantité de bande passante disponible.

En conséquence, pour minimiser le taux de rejet des requêtes de protection, le partage de bande passante doit être pris en compte lors du calcul des chemins de secours. Dans cette thèse, nous nous intéressons au calcul en ligne et distribué des chemins de secours vérifiant les contraintes de bande passante tout en la partageant dans une plateforme de type MPLS. Pour des raisons liées aux performances, nous avons choisi le mode de calcul en ligne qui accélère les calculs et évite les reconfigurations des chemins déjà établis. Nous avons aussi opté pour un modèle de calcul distribué afin d'augmenter la réactivité des serveurs de calcul et pour éviter un allongement sensible des délais de calcul après une panne. Enfin, pour des raisons de facilité de déploiement, nous avons tenu compte de deux contraintes :

- La quantité de trafic de contrôle circulant dans le réseau afin de permettre le placement des LSP de secours doit être réduite et acceptable.
- La mise en œuvre des méthodes de placement de LSP de secours doit être facile (i.e. des extensions légères des protocoles existants doivent suffire pour le déploiement des méthodes proposées).

Cette thèse est organisée en deux parties : la partie unicast et la partie multicast. Comme son nom l'indique, la première partie est consacrée au placement des chemins de secours vérifiant les contraintes de bande passante et protégeant des communications unicast. La seconde partie quant à elle, aborde le placement des chemins de secours vérifiant les contraintes de bande passante et protégeant efficacement des communications de type point-à-multipoint.

Dans le premier chapitre de cette thèse, nous décrivons les outils et mécanismes permettant d'établir et protéger localement des communications unicast sous MPLS. Nous commençons par une introduction à la technologie MPLS tout en nous focalisant sur deux principales fonctionnalités de MPLS qui sont l'ingénierie du trafic et la protection. La première fonctionnalité permet un contrôle fin des chemins suivis par les paquets dans le réseau, pour optimiser les ressources et assurer la qualité de service désirée. Elle est rendue possible grâce à la capacité de MPLS de définir et configurer des routes explicites. La seconde fonctionnalité permet quant à elle de réduire le temps de coupure des communications lors d'une panne afin de ne pas perturber les services supportées. Tous les types de risques de pannes simples sont gérés et pris en compte dans cette thèse. Ces risques sont classés dans trois groupes : lien, nœud et SRLG (Shared Risk Link Group). Ce dernier risque est lui-même formé d'un ensemble de liens partageant une ressource physique commune et donc, pouvant tomber en panne simultanément. Pour des raisons liées à la pratique, nous décrivons aussi trois protocoles permettant de mettre en œuvre le calcul des LSP primaires et de secours sous MPLS : les deux

premiers, qui sont OSPF-TE et ISIS-TE, permettent de collecter et distribuer l'information nécessaire au calcul des routes explicites ; le dernier protocole (RSVP-TE) est utilisé pour configurer les LSP primaires et leur secours.

Dans le deuxième chapitre, nous nous concentrons sur le partage de ressources et particulièrement sur le partage de la bande passante. Nous verrons qu'un placement distribué des chemins de secours, qui tient compte du partage de la bande passante, passe souvent par une distribution d'une grande quantité d'informations dans le réseau, ce qui le surcharge et détériore considérablement les performances. Pour remédier à cela, nous avons envisagé différentes solutions consistant, par exemple, à optimiser le placement des entités de calcul des chemins de secours et/ou à agréger l'information nécessaire au calcul des chemins de secours avant sa transmission dans le réseau.

Le troisième chapitre décrit l'algorithme TDRA (Targeted Distribution of Resource Allocation) que nous avons développé pour permettre un calcul distribué des chemins de secours sans inonder le réseau. Nous précisons que l'algorithme TDRA fournit les moyens permettant de maximiser le partage de la bande passante sur tous les liens du réseau avec une distribution ciblée et optimisée de l'information requise au calcul des chemins de secours. Concrètement, avec TDRA, l'information requise pour le calcul des chemins de secours est organisée en plusieurs structures, chacune de ces dernières n'est envoyée qu'à l'ensemble des entités de calcul capables de l'exploiter.
Pour agréger et réduire notablement l'information requise au calcul des chemins de secours, nous proposons et décrivons dans ce même chapitre deux heuristiques DBSH (Distributed Bandwidth Sharing Heuristic) et PLRH (PLR-based Heuristic). Bien que ces deux dernières utilisent des mécanismes très similaires pour approximer l'information requise au calcul des chemins de secours, leurs puissances et performances diffèrent très sensiblement (essentiellement) à cause de la différence des localisations des entités de calcul des LSP de secours. Alors que dans DBSH, les calculs des chemins de secours protégeant contre la panne d'un nœud donné (resp. d'un lien unidirectionnel donné) sont centralisés sur le nœud lui même (sur le nœud sortant du lien unidirectionnel), avec PLRH, tous les calculs de chemins de secours sont effectués par les routeurs de tête de ces chemins de secours. L'heuristique DBSH permet de réduire la taille de l'information diffusée dans le réseau par rapport à PLRH au détriment d'une asymétrie de distribution de cette information et d'un déploiement plus coûteux.

Dans le quatrième chapitre, nous proposons un nouvel algorithme exploitant les structures des SRLG pour améliorer le taux de protection et augmenter la disponibilité des ressources. En observant que certains chemins de secours actifs après une panne d'un SRLG ne peuvent recevoir du trafic, notre algorithme restreint l'ensemble des risques protégés par un chemin de secours à ceux dont la panne induit son opération (i.e. dont la panne induit la réception du flux de la communication protégée.)

Dans le cinquième chapitre de cette thèse, nous nous intéressons au placement des chemins de secours protégeant des communications point-à-multipoint. Nous décrirons dans ce chapitre différentes méthodes de protection proactive que nous combinons avec différentes stratégies de partage de la bande passante pour le placement les chemins de secours.

Dans le dernier chapitre de la thèse, nous nous focalisons sur le placement des

chemins de secours locaux qui protègent des communications point-à-multipoint, dans un réseau de type MPLS. Nous verrons que tous les algorithme et heuristiques proposés pour l'unicast sont extensibles et modifiables pour permettre le placement des chemins de secours protégeant des communications point-à-multipoint. De plus, afin d'étudier l'impact du choix de la stratégie de partage de la bande passante sur les performances des mécanismes de placement des chemins de secours point-à-multipoint locaux, nous effectuons une étude comparative des performances de deux stratégies de partage de la bande passante.

Contributions

Dans cette thèse, nous avons élaboré et proposé différents mécanismes permettant le calcul distribué des chemins de secours locaux (point-à-point et point-à-multipoint) dans une infrastructure de communication de type MPLS. Nos propositions ont pour avantage d'être faciles à déployer (quelques légères extensions des protocoles de routage et/ou de signalisation sont suffisantes) et elles réduisent la taille de l'information transmise dans le réseau pour permettre le placement des chemins de secours. Nous avons aussi proposé durant cette thèse une méthode de protection des communications multicast et avons étudié l'impact du choix des stratégies de partage de bande sur les performances des mécanismes de placement des chemins de secours.

Nos principales contributions, ayant fait l'objet pour la plupart de publications dans des conférences et journaux internationaux, sont résumées ci après :

Algorithme TDRA [SCLR08] : TDRA permet de maximiser le partage de bande passante sur les liens en réduisant la quantité d'informations transmises dans le réseau. Pour ce faire, TDRA effectue une distribution ciblée et optimisée de l'information requise pour le calcul distribué des chemins de secours. L'implantation de cet algorithme ne requiert que de légères extensions des protocoles de signalisation (et des extensions minimes des protocoles de routage).

Heuristique DBSH [SCLR07] : DBSH permet le placement des chemins de secours en ne diffusant qu'une information partielle et agrégée dans le réseau. Avec cette heuristique, les calculs des chemins de secours protégeant contre la panne d'un nœud donné (resp. d'un lien unidirectionnel donné) sont effectués sur le nœud lui même (resp. sur le nœud sortant du lien). Cette heuristique réduit considérablement la taille de l'information diffusée dans le réseau et elle est facile à déployer puisqu'elle ne requiert que de légères extensions aux protocoles de routage et aux protocoles de signalisation.

Heuristique PLRH [SCLRb, SCLR09a] : PLRH permet le placement des chemins de secours en ne diffusant qu'une information partielle et agrégée dans le réseau. Avec cette heuristique, tous les calculs des chemins de secours sont effectués par les routeurs de tête de ces chemins. Cette heuristique a pour avantages d'être symétrique (toutes les entités de calcul disposent de la même information) et facile à déployer (elle

ne nécessite que de très légères extensions aux protocoles de routage).

Exploitation des structures des SRLG pour un placement efficace des chemins de secours [SCLR09b, SCLRa] : Pour améliorer le taux de protection et augmenter la disponibilité des ressources, nous proposons d'exploiter les structures des SRLG. En constatant que certains chemins de secours actifs après une panne d'un SRLG ne peuvent recevoir du trafic, nous proposons de restreindre l'ensemble des risques protégés par un chemin de secours à ceux dont la panne induit la réception du trafic.

Protection multicast par forêt duale [SCM06b, SCM06a, SCM06c] : Pour protéger les communications multicast, nous proposons d'utiliser une structure de secours formée d'une forêt reliant les nœuds feuilles de l'arbre multicast. Cette méthode de protection a l'avantage de réduire la taille de la structure de secours. Elle n'est décrite que très brièvement dans cette thèse.

Impact du choix de la stratégie de partage sur les performances des mécanismes de placement des chemins de secours [SC08] : Nous étudions l'impact du choix de la stratégie de partage de bande passante employée sur les performances des mécanismes de placement des chemins de secours. Deux stratégies de partage de bande passante sont ainsi comparées : (1) partage de la bande passante entre les chemins de secours uniquement et (2) partage de la bande passante entre les chemins de secours et le LSP primaire. Dans notre étude, nous employons un algorithme de calcul basé sur les plus courts chemins en termes de nombre de sauts.

Chapitre 1

MPLS et la protection

1.1 Introduction

Avec la croissance des besoins des applications en ressources et pour assurer des temps de réponse limités, il devient incontournable de doter les réseaux de fonctionnalités d'ingénierie de trafic et de protection. La première fonctionnalité permet de maximiser la quantité de trafic susceptible de traverser un réseau tout en assurant le respect de contrats liés aux ressources disponibles. La seconde fonctionnalité, quant à elle, permet d'assurer la continuité des services, même après une panne.

Avec l'arrivée de MPLS durant la dernière décade, les fonctionnalités d'ingénierie de trafic et de protection ont été améliorées. Grâce à la flexibilité dans le choix des routes et à la possibilité de pré-configurer des chemins, MPLS permet de :

1. optimiser l'utilisation des ressources, ce qui augmente la quantité de trafic susceptible de traverser un réseau,
2. réduire les délais de récupération après une panne pour assurer la continuité des services.

Dans ce chapitre, nous nous intéressons aux mécanismes et outils permettant d'établir des connexions protégées sous MPLS. Nous commençons par une introduction à MPLS et aux techniques de commutation d'étiquettes. Puis, nous décrivons les protocoles de routage et de signalisation permettant le calcul, la configuration et la réservation des ressources sur les chemins primaires transportant le trafic des communications en l'absence de pannes. Ensuite, nous nous focalisons sur la protection (essentiellement la protection locale) et expliquons son utilité et ses principes. Enfin, nous terminons par une description des extensions apportées au protocole de signalisation (RSVP) pour la mise en œuvre de la protection locale sous MPLS.

1.2 MPLS

Pour optimiser l'utilisation des ressources, fournir la protection et offrir différents niveaux de services, le protocole MPLS est de plus en plus adopté par les opérateurs

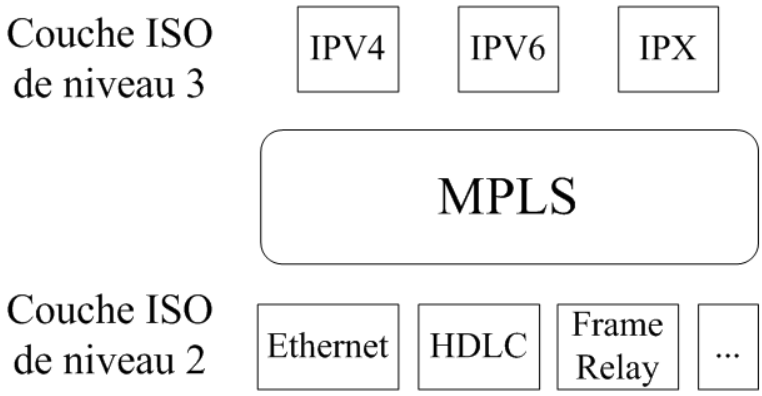


FIG. 1.1 – MPLS dans le modèle ISO

de télécom, surtout dans le cœur du réseau. Ce protocole, qui est défini et normalisé par l'IETF (Internet Engineering Task Force), regroupe un ensemble de spécifications améliorant le routage par l'insertion d'étiquettes MPLS dans les entêtes des paquets transmis dans le réseau. En fonction, non pas de l'adresse *Réseau* de destination (adresse de destination de niveau 3 de la couche ISO) mais de l'étiquette MPLS, les paquets sont aiguillés et transmis aux prochains routeurs permettant d'atteindre la destination.

Comme son nom l'indique, MPLS est multi-protocole : il n'est donc restreint à aucune couche de niveau 2 du modèle ISO et il devrait fonctionner sur tous les types de protocole permettant l'acheminement des paquets au niveau 3 de la couche ISO. Souvent, l'étiquette MPLS est placée entre l'entête de niveau 2 et l'entête de niveau 3 de la couche ISO, c'est pourquoi on dit de MPLS qu'il est de niveau 2,5 (cf. figure 1.1).

1.2.1 Acheminement des paquets par commutation d'étiquettes

Contrairement à IP qui utilise l'adresse de destination pour router les paquets saut par saut, MPLS définit des circuits permettant d'acheminer les paquets sur des mêmes chemins unidirectionnels dits LSP (Label Switched Paths)[1]. Chaque LSP interconnecte un nœud source à un nœud de destination et est formé d'une succession de LSR (Label Switched Routers) permettant aux paquets de suivre le même itinéraire pour atteindre leur destination.

Pour déterminer le LSP que doit emprunter un paquet (afin d'atteindre sa destination), MPLS utilise des étiquettes. Chaque étiquette identifie sur tout LSR un seul LSP[2]et est associée à groupe de paquets (FEC ou *Forwarding Equivalence Class*) devant être transmis de manière identique sur le réseau MPLS.

[1]Le routage saut par saut existe aussi sous MPLS (cf. section 1.2.5).

[2]Lorsque l'allocation d'étiquettes est globale au routeur MPLS, une étiquette MPLS est suffisante pour identifier un LSP. Par contre, lorsque l'allocation d'étiquettes est locale à chaque interface d'un LSR, c'est le couple, étiquette MPLS et interface du LSR, qui identifie un LSP.

Ainsi, à la réception d'un paquet par un routeur frontière d'entrée LER (Label Edge Router) à un domaine MPLS, ce dernier déduit la FEC associée au paquet à partir d'informations contenues dans son entête (comme l'adresse de destination, le numéro de port, etc.) et consulte sa table d'étiquettes LFIB (*Label Forwarding Information Base*) pour déduire l'étiquette et l'interface de sortie permettant l'acheminement du paquet. Ensuite, le LER d'entrée ajoute l'étiquette déterminée au paquet (opération *push*) avant de l'envoyer sur l'interface de sortie déduite précédemment et permettant d'atteindre le prochain LSR.

Une fois que l'étiquette MPLS est insérée dans le paquet envoyé, l'acheminement se fait par *commutation d'étiquettes* (opération *swap*) le long du LSP associé à la FEC. Typiquement, tout LSR recevant le paquet consulte sa table LFIB et déduit l'étiquette et l'interface de sortie à partir de l'étiquette d'entrée (resp. à partir de l'étiquette d'entrée et de l'interface d'entrée) lorsque l'allocation d'étiquettes est globale au LSR (resp. lorsque l'allocation d'étiquettes est locale à cette interface). Le paquet est ensuite envoyé sur l'interface de sortie déterminée précédemment en substituant l'étiquette d'entrée par l'étiquette de sortie. Lorsque le paquet arrive au routeur d'extrémité du LSP (ou parfois au routeur aval du routeur d'extrémité du LSP) associé à la FEC, l'étiquette est supprimée (opération *pop*).

Sur la figure 1.2, Le LER d'entrée *A* reçoit un paquet destiné à une adresse appartenant au sous-réseau 131.3.0.0/16. En consultant sa table LFIB, le routeur *A* détermine la FEC associée au sous-réseau 131.3.0.0/16 et déduit ensuite l'étiquette *L1* et l'interface de sortie *P1* correspondant à la FEC 131.3.0.0/16. Après cette recherche, le routeur *A* insère l'étiquette *L1* dans le paquet reçu (opération *push*) et le transmet sur son interface de sortie *P1* déterminée précédemment. Le LSR *B* qui reçoit le paquet consultera à son tour sa table LFIB et substituera l'étiquette *L2* à l'étiquette *L1* (opération *swap*) avant d'envoyer le paquet obtenu sur l'interface *P3*. A la réception du paquet par le LSR *C*, ce dernier supprime l'étiquette *L2* du paquet (opération *pop*) avant de l'envoyer sur

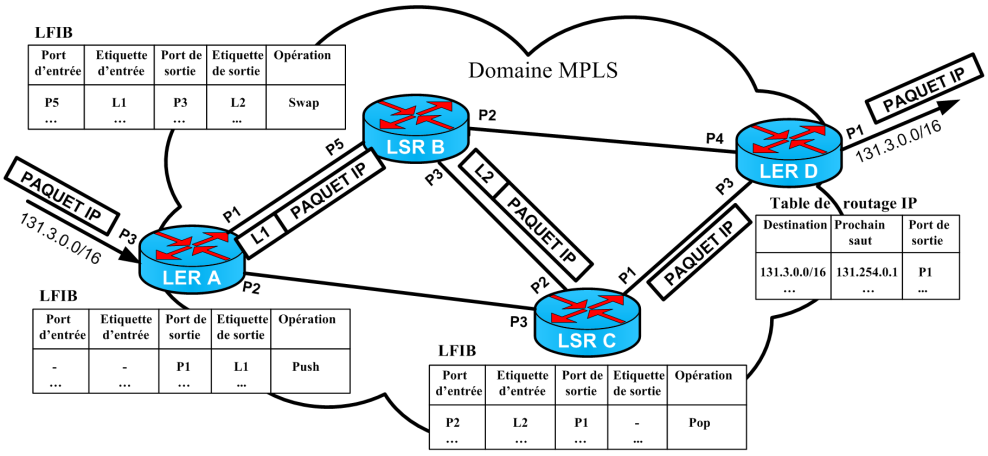


FIG. 1.2 – Commutation d'étiquettes sous MPLS

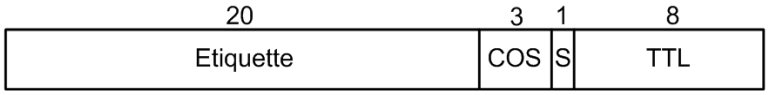


FIG. 1.3 – Structure d'un entête MPLS élémentaire

l'interface *P1* permettant d'atteindre un LER de sortie (D) au domaine MPLS. Pour acheminer le paquet vers le prochain routeur (131.254.0.1) situé en dehors du domaine MPLS, le routeur D utilisera sa table de routage IP.

1.2.2 Structure d'un entête MPLS

Pour bénéficier des fonctionnalités rendues par MPLS, les paquets transmis dans le réseau doivent contenir au moins un entête de taille fixe (32 bits), appelé entête MPLS. Cet entête est composé des champs suivants (figure 1.3) :

1. **Etiquette** (ou ***label***) : c'est un identifiant utilisé pour aiguiller le paquet sur un routeur MPLS. Elle est codée sur 20 bits.
2. **COS (*Class Of Service*)** : ce champ permet d'inclure des informations sur la classe de service. Il a un rôle similaire au champ TOS (*Type Of Service*) de l'entête IP. Il est codé sur 3 bits dont la sémantique exacte reste expérimentale.
3. **S (*Bottom of Stack*)** : ce champ est constitué d'un seul bit et permet d'établir une hiérarchie dans la pile d'étiquettes (plusieurs étiquettes successives). Il n'est positionné que pour l'entête situé au fond de la pile MPLS. Il est surtout utile pour agréger le trafic de plusieurs LSP dans un seul tunnel *TE-trunk* [AMA+99] et pour acheminer le trafic des VPN (*Virtual Private Networks*).
4. **TTL (Time To Live)** : ce champ permet de limiter l'effet des boucles et sert aussi à mettre à jour le champ TTL du paquet IP lors de sa sortie du domaine MPLS. Il est codé sur 8 bits.

1.2.3 Suppression d'étiquettes MPLS

L'étiquette MPLS n'est utile et significative qu'à l'intérieur d'un domaine MPLS pour router du trafic sur un LSP donné. En conséquence, toute étiquette MPLS doit être supprimée, au plus tard, à l'arrivée au routeur de sortie du LSP correspondant.

Selon les capacités du routeur de sortie d'un LSP (et de ses souhaits) et du LSR situé en aval, la suppression d'étiquette MPLS est effectué par le LSR aval au routeur de sortie du LSP ou sur le routeur de sortie du LSP lui-même.

Lorsque l'étiquette est supprimée par le LSR en aval du routeur de sortie, l'opération est appelée PHP (*Penultimate Hop Popping*). Cette opération est souvent utile pour éviter au routeur de sortie d'effectuer deux recherches dans ses tables de routage. En effet, puisque le but de l'insertion d'une étiquette MPLS dans un paquet est de permettre son acheminement d'un routeur source à un routeur de destination en suivant le LSP

associé à l'étiquette, il résulte que le routeur de sortie n'a besoin d'aucune étiquette pour router le paquet sur le LSP précédent (la destination du LSP est atteinte). En conséquence, le PHP permet d'éviter des recherches supplémentaires dans les tables LFIB et autorise l'établissement de LSP dont les routeurs de sortie ne sont pas MPLS.

Lorsque le LSR situé en aval du routeur de sortie n'est pas capable de dépiler l'étiquette ou lorsque le routeur de sortie ne désire pas le PHP, l'étiquette est supprimée par le routeur de sortie du LSP. Comme nous allons le voir plus loin dans cette thèse, la suppression de l'étiquette au routeur de sortie du LSP et non pas à son LSR aval est parfois nécessaire (ex. protection multicast par tunnels point à multipoint).

1.2.4 Distribution d'étiquettes

Pour assurer la distribution d'étiquettes, différents protocoles (protocoles de distribution d'étiquettes) ont été définis. Ces derniers sont constitués de l'ensemble des procédures permettant à un LSR d'informer un autre des associations étiquette/FEC qu'il a mis en place (ces associations sont sauvegardées dans les tables LFIB). Les protocoles de distribution d'étiquettes englobent aussi toutes les négociations engagées entre deux routeurs MPLS pour apprendre les capacités MPLS de l'un ou de l'autre.

Pour offrir plus de flexibilité, l'architecture MPLS **ne fait aucune hypothèse concernant le protocole de distribution d'étiquettes employé**. Plusieurs protocoles ont été normalisés :

1. LDP (*Label Distribution Protocol*) [ADF+01] : ce protocole englobe les procédures et messages permettant aux LSR d'établir des LSP dans un réseau en se basant sur l'information transmise par la couche *Réseau*.
2. CR-LDP (*Constraint-Based LSP Setup using LDP*) [JAC+02] : ce protocole étend LDP pour permettre l'ingénierie du trafic. Bien que CR-LDP ait été standardisé par l'IETF, son déploiement est très limité.
3. RSVP-TE (Resource reSerVation Protocol-Traffic Engineering) : initialement, RSVP était utilisé comme protocole de signalisation de la qualité de service (QoS) [BZB+97].

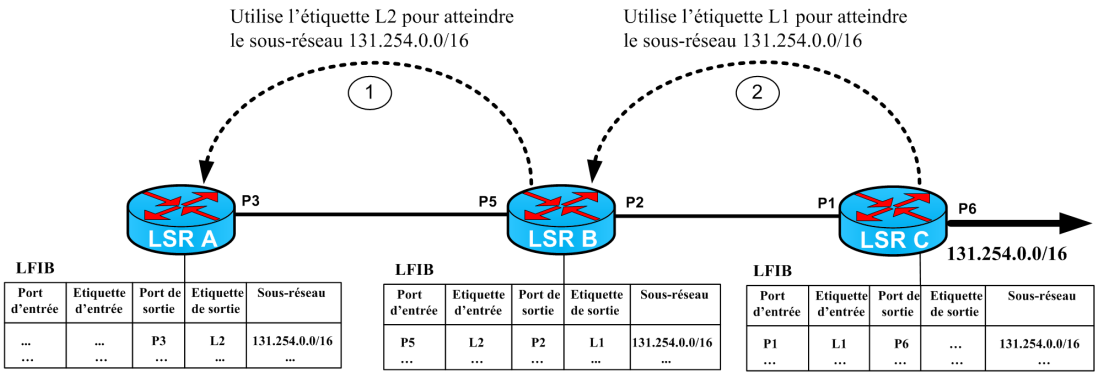


FIG. 1.4 – Distribution d'étiquettes avec le mode *Unsolicited Downstream*

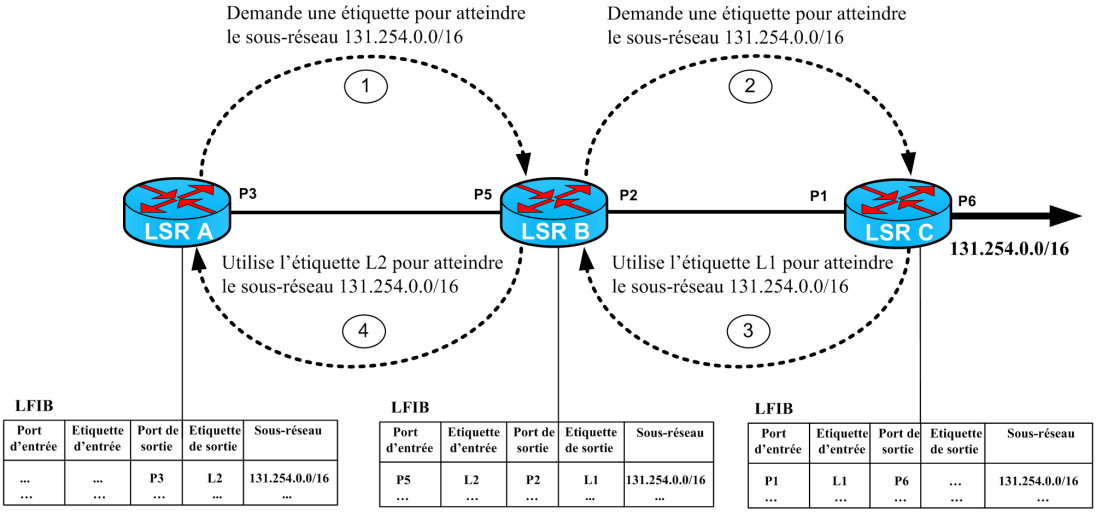


FIG. 1.5 – Distribution d'étiquettes avec le mode *Downstream-on-Demand*

Il a été étendu par [ABG+01, KL04, FPVA06, LFDC07, FAV08] pour distribuer les étiquettes et permettre l'ingénierie du trafic.

Deux modes de distribution d'étiquettes sont définis : *Unsolicited Downstream* et *Downstream-on-Demand*.

Unsolicited Downstream

Dans le mode *Unsolicited Downstream*, les LSR distribuent les associations étiquette/FEC aux autres LSR sans que ces derniers ne les demandent explicitement (figure 1.4). Les étiquettes sont toujours allouées par les LSR avals et transmises aux LSR amonts qui leurs sont voisins. Ce mode est surtout utilisé par les LSR pour configurer une interface permettant d'atteindre une destination nouvellement découverte.

Sur la figure 1.4, le LSR *C* alloue l'étiquette *L1* (sur l'interface *P1*) pour permettre à son LSR voisin *B* d'atteindre le réseau 131.254.0.0/16 qu'il vient de découvrir. Ensuite, il met à jour sa table LFIB en insérant une entrée contenant l'étiquette *L1* et la nouvelle adresse de réseau 131.254.0.0/16. Puis, il envoie l'étiquette *L1* et l'adresse de réseau 131.254.0.0/16 à son LSR amont *B*. Lorsque ce dernier reçoit le couple (*L1*, 131.254.0.0/16), il vérifie qu'il ne dispose pas d'entrée correspondant à l'adresse 131.254.0.0/16 avant d'allouer une nouvelle étiquette *L2* et une nouvelle entrée dans sa table LFIB composée du quintuple (*P5*, *L2*, *P2*, *L1*, 131.254.0.0/16)[3]. Ensuite, l'étiquette *L2* ainsi que l'adresse de réseau 131.254.0.0/16 seront envoyées à tous les LSR en amont du LSR *B* (différents

[3]En réalité, deux entrées sont insérées dans la table LFIB du LSR *B* : une première entrée associant au sous-réseau 131.254.0.0/16 une FEC donnée (qui est un pointeur vers une entrée de la LFIB permettant d'acheminer le trafic vers le sous-réseau 131.254.0.0/16) et une deuxième entrée associant à la FEC précédente les étiquettes et les interfaces permettant l'aiguillage des paquets destinés au sous-réseau précédent.

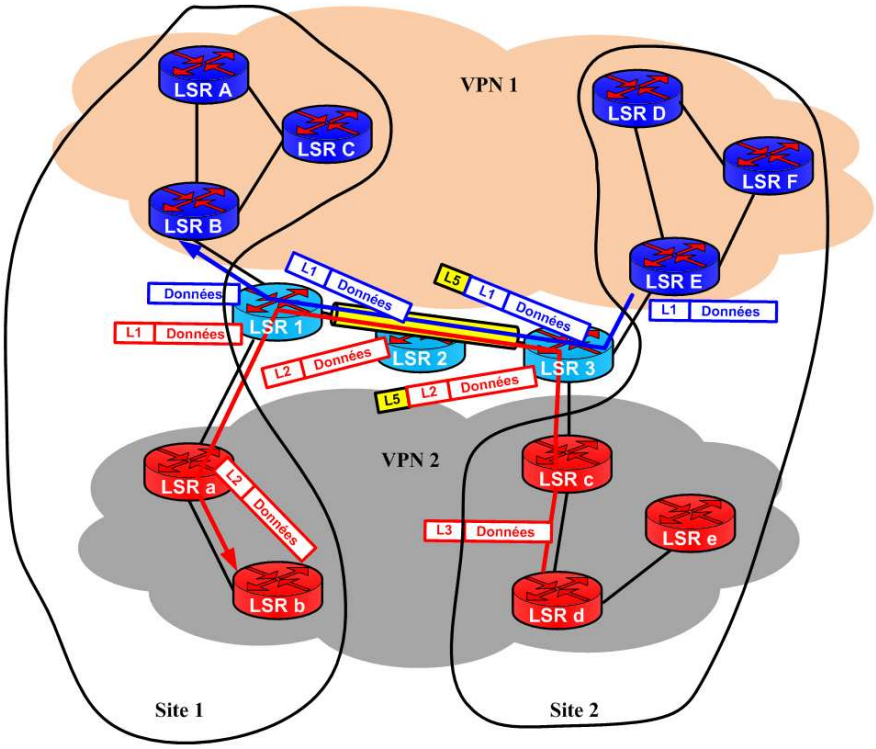


FIG. 1.6 – LSP hiérarchiques

du LSR *C* et voisins à *B*). Si un LSR reçoit (de la même interface) un couple d'étiquette d'entrée et d'adresse de réseau déjà existants dans sa table LFIB, il n'effectuera aucun traitement particulier (pas d'envoi de message aux LSR en aval).

Downstream-on-Demand

Dans le mode *Downstream-on-Demand*, la distribution d'étiquettes est effectuée par les LSR en aval à la suite de requêtes explicites provenant des LSR en amont. Ce mode est utilisé par un LSR qui veut atteindre une destination donnée (qui n'est pas encore présente dans sa table LFIB).

Sur la figure 1.5, le LSR *A* déclenche la procédure de distribution d'étiquettes afin d'atteindre une adresse située dans le réseau 131.254.0.0/16. Pour cela, il détermine son prochain nœud en utilisant sa table IP (ou en utilisant un algorithme de calcul de routes différent) et envoie une requête de demande d'étiquette à son prochain LSR *B*. Ce dernier effectuera le même traitement que le LSR *A* et enverra une requête de demande d'étiquette au LSR *C*. Lorsque le LSR *C* reçoit la requête, il s'aperçoit qu'il est le LSR de sortie du LSP permettant d'atteindre le réseau 131.254.0.0/16 et alloue alors l'étiquette *L1* qu'il transmettra au LSR *B* dans un message de réponse à la requête de demande d'étiquette reçue. De même, le LSR *B* allouera l'étiquette *L2* qu'il enverra ensuite au

LSR *A* en guise de réponse au message de demande d'étiquette reçu précédemment.

1.2.5 Séparation des plans d'usager et de contrôle

Pour fournir plus de flexibilité et afin de permettre le support de différents protocoles, l'architecture MPLS sépare le plan usager du plan de contrôle.

Au niveau du plan usager, les données sont acheminées en utilisant les mêmes techniques d'aiguillage basées sur le paradigme de la commutation d'étiquettes. C'est grâce au niveau du plan de contrôle que le trafic est séparé en plusieurs flux de classes de service différentes. C'est ce niveau aussi qui permet de spécifier les routes et le protocole de niveau 3 à utiliser[4].

1.2.6 Hiérarchie MPLS

Grâce à l'empilement des étiquettes, MPLS permet d'établir des LSP de différents niveaux. Chaque LSP de niveau p est défini comme une séquence de LSR commençant par un LSR (LSR de tête du LSP) insérant une étiquette de niveau p et se terminant par un LSR (routeur de sortie du LSP) qui achemine le paquet sur la base d'une pile de profondeur strictement inférieure à p. Cette hiérarchie de niveaux MPLS à deux avantages principaux : (1) l'agrégation du trafic de plusieurs flux et (2) la garantie d'un routage de bout en bout transparent à chaque niveau.

Sur la figure 1.6, deux réseaux privés (VPN 1 et VPN 2) localisés chacun sur deux sites différents sont représentés. Afin de permettre le passage à l'échelle et pour des raisons de sécurité, les LER (*LSR 1* et *LSR 3*) du réseau de transit (*LSR 1-LSR 2-LSR 3*) ne distribuent aux LER des réseaux VPN 1 et VPN 2 que des informations de routage relatives à leurs réseaux privés. En conséquence, pour établir un LSP du *LSR E* (resp. *LSR c*) vers le *LSR B* (resp. *LSR a*), le *LSR E* (resp. *LSR c*) ne se servira que des informations de routage concernant son réseau VPN (i. e. étiquette de premier niveau dans la pile MPLS) : le passage par le routeur *LSR 2* du cœur du réseau est complètement transparent aux routeurs *LSR E* et *LSR c*.

Ainsi, pour permettre au *LSR E* (resp. *LSR c*) d'atteindre le *LSR B* (resp. *LSR a*), le *LSR 3* lui transmet l'étiquette *L1* (resp. *L2*).

Pour ne pas encombrer le *LSR 2* qui est interne au cœur du réseau de transit (*LSR 1-LSR 2-LSR 3*), le *LSR 3* établit un tunnel identifié par l'étiquette *L5* entre lui et le *LSR 1*. De cette manière, tous les paquets reçus par le *LSR 3* et provenant du *LSR E* avec l'étiquette *L1* ou du du *LSR c* avec l'étiquette *L2* seront acheminés sur le tunnel ou le LSP de deuxième niveau (*LSR 3-LSR 2-LSR 1*). Bien évidemment, la gestion et la maintenance des LSP sont simplifiées sur le routeur *LSR 2* puisque tout se passe, pour lui, comme s'il n'était traversé que par un seul LSP (un seul flux).

Notons que l'établissement de tunnels a des incidences sur l'allocation et la distribution des étiquettes. Concrètement, sur la figure 1.6, les routeurs *LSR 1* et *LSR 3*

[4]Par exemple, pour permettre au routeur de sortie d'un domaine MPLS de déterminer le protocole à utiliser en dehors du domaine, RSVP-TE inclut dans les messages d'établissement d'un LSP un objet (LABEL_REQUEST) spécifiant le protocole de niveau 3.

sont *adjacents sur le LSP de premier niveau* (i.e. le *LSR 1* et le successeur immédiat du *LSR 3* dans les deux LSP de niveau 1). En conséquence, c'est le *LSR 1* qui alloue l'étiquette *L1* transmise par le *LSR 3* au *LSR E* (resp. au *LSR c*) pour atteindre le *LSR B* (resp. le *LSR a*).

Comme nous allons le voir dans la section 1.4.4.4, la hiérarchie MPLS est utilisée par la protection par tunnel de secours, qui agrège plusieurs LSP dans un même tunnel, pour réduire le trafic de contrôle et diminuer le nombre d'étiquettes allouées dans le réseau.

1.2.7 Choix des routes

Afin de choisir le LSP que doivent suivre les paquets associés à une FEC donnée, MPLS offre deux options [RVC01] : (1) routage saut par saut et (2) routage explicite.

Avec le premier type de routage, la route est construite saut par saut, chaque nœud du LSP est responsable et libre de choisir son prochain saut dans le LSP. Ce type de routage est le mode usuellement adopté par les réseaux IP existants.

Avec le second type de routage, le passage d'un LSP par certains LSR est imposé et est dicté par un (ou plusieurs) LSR du LSP ou par l'administrateur du réseau. Souvent, c'est le routeur de tête (i.e. le routeur source du LSP) ou le routeur de sortie d'un LSP (i.e. le routeur de destination du LSP) qui détermine et choisit certains (ou tous) LSR du LSP.

Selon que c'est l'ensemble des LSR qui forment un LSP qui est spécifié ou pas, nous distinguons deux types de LSP routés explicitement : les LSP stricts (*strict LSPs*) et les LSP libres (*loose LSPs*). Dans un LSP strict, la séquence des LSR le formant est fixée et est choisie statiquement par configuration ou dynamiquement à chaque LSR. Dans un LSP libre par contre, l'ensemble des LSR spécifiés et devant être traversés par le LSP ne permet pas de définir entièrement le LSP (i.e. certains LSR du LSP sont libres de choisir leur prochain saut).

Bien évidemment, la simplicité et l'efficacité du routage explicite sous MPLS a été un facteur déterminant dans le déploiement des fonctionnalités d'ingénierie de trafic, de différenciation de service et de protection. Sous IP, le routage explicite est très coûteux. Il peut être implanté de deux façons : par encapsulation (surcoût très élevé) ou en utilisant l'option *source routing*. Avec cette dernière, le routeur source énumère les adresses IP de tous les routeurs intermédiaires de la route explicite et les insèrent dans l'entête de chaque paquet de données transmis. Sous MPLS, la spécification des routeurs intermédiaires d'un LSP n'est effectué que lors de la configuration de ce LSP. Pour faire correspondre un paquet de données à sa route explicite, MPLS ne se sert que de l'étiquette MPLS transmise dans le paquet envoyé et renseignée lors de la configuration de la route.

1.2.8 Gestion de la QoS et de l'ingénierie du trafic

La gestion de la QoS et de l'ingénierie du trafic est réalisée sous MPLS grâce au plan de contrôle qui fournit une granularité suffisante pour la partition du trafic (FEC) et

offre des moyens permettant d'associer à chaque partition du trafic un LSP établi avec ou sans réservation de ressource et non contraint par le protocole de routage employé (grâce au routage explicite). Ainsi, l'ingénieur réseau peut optimiser son réseau en répartissant le trafic en plusieurs classes. chaque classe assure une certaine QoS en utilisant un LSP qui réduit la consommation des ressources.

1.2.9 Routage par contrainte MPLS-TE

Pour assurer le respect des contraintes d'ingénierie de trafic (TE), de qualité de service et des ressources disponibles dans la topologie TE du réseau, un mécanisme de routage par contrainte MPLS-TE est défini. Ce dernier permet le calcul et l'établissement des chemins, dits TE-LSP, en se basant sur trois fonctions principales [LR06] :

Découverte de la topologie TE

Cette fonction a pour rôle de transmettre une vision actualisée de la topologie du réseau ainsi que des paramètres TE associés (comme la bande passante réservable sur chaque lien, les quantités de bande passante résiduelles sur les différents niveaux de priorité pour tout lien, groupes administratifs de tout lien, etc.) à tous les routeurs du réseau. La topologie TE du réseau est enregistrée par chaque routeur dans une base de données appelée TED (TE Database).

Nous notons que la fonction de découverte de la topologie TE est fournie par un protocole IGP-TE (*Interior Gateway Protocol-TE*) qui correspond à un protocole IGP à état de liens (OSPF [Moy98b, Moy98a] ou ISIS [Ora90]) étendu pour permettre l'annonce des paramètres TE dans le réseau. Deux protocoles IGP-TE sont employés pour assurer cette fonction dans un réseau MPLS : OSPF-TE [KR05b] et ISIS-TE [KR05a]. Ils sont décrits plus en détails dans les sections suivantes.

Calcul des chemins

Cette fonction permet la détermination des chemins vérifiant les contraintes TE des TE-LSP (bande passante, groupes à inclure/exclure, etc.) ainsi que les contraintes de la topologie TE du réseau (quantités de bande passante résiduelles sur chaque niveau pour tout lien, groupes administratifs des liens, etc.). Elle est assurée par un ou plusieurs routeurs qui exécutent un algorithme de calcul de routes, souvent basé sur l'algorithme de Dijkstra (comme CSPF ou *Constrained Shortest Path First*) ou utilisant la programmation ILP (*Integer Linear Programming*).

Deux modes de calcul de chemins sont à distinguer : *le mode en ligne* et *le mode hors ligne*. Dans le premier mode, les chemins sont calculés au fur et à mesure qu'ils arrivent, sans connaissance préalable de l'ordre d'arrivée des futures demandes d'établissement de TE-LSP. Ce mode a l'avantage de réduire les délais d'établissement des TE-LSP mais il n'est souvent pas optimal. Avec les second mode de calcul, les requêtes de calcul des chemins sont traitées ensemble, indépendamment de leur ordre d'arrivée. Ce mode induit des temps de latence élevés (temps d'établissement des TE-LSP élevés) et n'est utile que pour router des des matrices de trafic statique et pour ré-optimiser les ressources des TE-LSP déjà établis.

Signalisation des TE-LSP

Cette dernière fonction intervient pour configurer les routes explicites déterminées par la fonction de calcul des chemins. Elle est rendue grâce au protocole RSVP-TE (le déploiement de CR-LDP étant très limité) qui assure en plus le maintien, la suppression des LSP et la notification d'erreurs. Le protocole RSVP-TE est décrit plus en détail dans les sections suivantes.

1.3 Mise en œuvre du mécanisme de placement des TE-LSP

Un mécanisme de placement des TE-LSP regroupe toutes les procédures permettant le calcul et la configuration des TE-LSP. Différents protocoles IGP et de signalisation ont été étendus pour permettre la prise en charge des contraintes d'ingénierie de trafic et de la qualité de service lors de l'établissement des TE-LSP. Dans cette section, nous décrivons les protocoles OSPF-TE [KR05b], ISIS-TE [KR05a]) et RSVP-TE [ABG+01] qui sont standardisés par l'IETF et déployés sur les plate-formes MPLS.

1.3.1 Le protocole OSPF-TE

OSPF-TE (*Open Shortest Path First-TE*) est un protocole IGP à état de liens étendu pour transmettre des paramètres liés à l'ingénierie de trafic et à la qualité de service. Comme son prédécesseur (OSPF), le protocole OSPF-TE utilise essentiellement des diffusions sûres des champs LSA (*Link State Advertisement*) pour établir une vision commune et synchronisée de la topologie TE du réseau sur tous les routeurs OSPF-TE. Chaque LSA est constitué de deux parties : entête et corps du LSA.

L'entête d'un LSA a une taille fixe et est constitué des champs suivants :

1. *Link State ID*, *Advertising router* et *LS Type* : permettant d'identifier le LSA ainsi que son type.
2. *LS Sequence Number* : utilisé pour différencier les différentes instances d'un LSA.
3. *LS Age* : permettant de cadencer la fréquence de rafraîchissement du LSA dans des circonstances normales.
4. *LS Checksum* : permettant de se prémunir contre la corruption du LSA.
5. *Options* : servant à indiquer que le LSA requiert un traitement spécial lors de la diffusion ou du calcul des routes.
6. *Longueur* : contenant la longueur totale du LSA (somme de la taille de l'entête et du corps du LSA).

Le corps d'un LSA transmet, en plus de l'information permettant de déduire la topologie IP du réseau (constituées des routeurs et des liens du réseau) et la métrique associée à chaque lien (cette information est annoncée dans OSPF [Moy98b, Moy98a]), des champs liés à l'ingénierie du trafic et à la qualité de services :

bande passante maximale : ce champ détermine la quantité de bande passante maximale pouvant être utilisée sur un lien. En général, elle correspond à la capacité physique du lien en termes de bande passante.

bande passante réservable : c'est la quantité de bande passante maximale qui peut être réservée sur un lien pour l'ensemble des TE-LSP qui le traversent. Cette quantité peut être supérieure (*overbooking*) ou inférieure (*underbooking*) à la bande passante maximale du lien.

bandes passantes résiduelles : il s'agit des quantités de bande passante réservables et disponibles sur chacun des huit niveaux de priorité sur un lien.

groupes administratifs (couleurs) : dans ce champ, tous les groupes administratifs auxquels appartient un lien sont dénombrés. Nous notons qu'un lien peut appartenir à un ou plusieurs groupes administratifs parmi les trente-deux groupes pouvant être définis. Les groupes administratifs peuvent être utilisés comme contraintes pour inclure/exclure un lien lors du calcul des routes.

métrique TE : cette métrique permet d'inclure les contrainte TE lors du calcul des chemins. C'est une métrique qui peut être combinée à la métrique IGP pour déterminer par exemple un plus court chemin vérifiant un ensemble de contraintes TE (délai, bande passante, etc.).

Suite à la diffusion des différents LSA, chaque routeur construit une table TED contenant la topologie TE du réseau et applique un algorithme de calcul des chemins pour satisfaire les futures requêtes de demande d'établissement de LSP. Souvent, les routeurs exécutent l'algorithme CSPF (*Constrained Shortest Path First*) qui est résumé par les deux étapes suivantes :

1. Éliminer de la topologie TE du réseau les liens ne vérifiant pas les contraintes TE (utilisation de la métrique TE).
2. Calcul du plus court chemin en appliquant l'algorithme SPF, basé lui-même sur l'algorithme de Dijkstra, à la topologie du réseau obtenue après l'exécution de (1).

1.3.2 Le protocole ISIS-TE

Le protocole ISIS-TE (*Intermediate System to Intermediate System-TE*) est aussi un protocole IGP à état de liens qui étend le protocole ISIS pour prendre en charge l'ingénierie du trafic et offrir la qualité de service. Comme OSPF-TE, le protocole ISIS-TE permet de construire une base données TED commune aux routeurs grâce à la diffusion des LSPDU (Link State Protocol Data Unit). Ces derniers transmettent des informations semblables à celles envoyées dans OSPF-TE, à savoir :

1. la topologie IP du réseau,
2. la bande passante maximale,
3. la bande passante réservable,
4. les bandes passantes résiduelles sur les huit niveaux de priorité,
5. les groupes administratifs (couleurs),

6. la métrique TE et métrique IGP.

Bien évidemment, après l'établissement de la base TED sur les routeurs, ces derniers s'occuperont de satisfaire les différentes requêtes de calcul des LSP reçus en exécutant un algorithme de calcul des chemins vérifiant les contraintes TE et offrant la qualité de service désirée.

Comme on le constate, les deux protocoles OSPF-TE et ISIS-TE rendent à peu près les mêmes fonctions en annonçant dans les réseaux des informations très semblables. Pour des raisons de généricité, nous utiliserons les deux protocoles (OSPF-TE et ISIS-TE) afin de rendre la fonctionnalité de protection (cf. section 1.4).

1.3.3 Le protocole RSVP-TE

RSVP-TE est un protocole de signalisation destiné à réserver les ressources pour les flux de données des applications dans un réseau MPLS. Il permet ainsi d'obtenir différentes qualités de service pour les flux de données, de router explicitement les LSP avec ou sans réservation de ressources, de rerouter rapidement les LSP, de préempter les ressources, de détecter les boucles, etc.

Dans la pile protocolaire ISO, RSVP-TE utilise le protocole IP. Il est référencé par ce dernier en utilisant l'identificateur de protocole 46. Cependant et lorsque le système ne permet pas l'utilisation de services réseau directement (pas de mode *raw*), les paquets RSVP sont encapsulés dans des paquets UDP.

Comme nous allons le voir dans la suite de ce chapitre, RSVP-TE permet aux nœuds de disposer de différents paramètres utiles au contrôle d'admission et au calcul distribué des LSP.

Dans la suite de cette section, nous nous intéressons aux états de réservation et aux types de messages RSVP. Nous décrirons aussi les tunnels RSVP qui sont constitués de plusieurs LSP avant de nous intéresser de près aux objets transmis dans les messages RSVP. Nous finirons en montrant comment établir et supprimer un LSP avec le protocole RSVP.

1.3.3.1 États et messages RSVP-TE

Afin de gérer l'établissement et la suppression des LSP ainsi que les ressources qui leur sont allouées, RSVP-TE garde des informations d'état pour un LSP sur tous les routeurs traversés. Cette information est appelée *état RSVP* et contient entre autres :

1. l'adresse source et l'adresse de destination ;
2. les numéro de tunnel RSVP-TE et de LSP (cf. section 1.3.3.2) ;
3. la bande passante ;
4. le nœud précédent et le nœud suivant sur le LSP ;
5. la route explicite ;
6. l'étiquette d'entrée et l'étiquette de sortie.

Les états RSVP-TE sont souples (*soft states*) et sont donc rafraîchis périodiquement, au déclenchement du *temporisateur de rafraîchissement* R, pour pallier les pertes d'éventuels messages RSVP-TE. Si un état RSVP-TE n'est pas rafraîchi au bout d'une certaine période de temps déterminée par le *temporisateur d'expiration* L (par défaut, $L = 4 \times R$), il est supprimé. Contrairement au mode dur (*hard state*) ou un état n'est supprimé qu'explicitement, le mode souple convient bien à RSVP-TE puisque :

1. il diminue les traitements effectués lors d'un changement de route. En effet, pour supprimer une route (ou une partie de la route), RSVP-TE peut se contenter de ne pas rafraîchir les états RSVP-TE correspondants.
2. il permet de toujours supprimer les états de réservation RSVP (ce qui permet de récupérer les ressources), même si une panne déconnecte l'émetteur du récepteur.

Pour établir, maintenir et supprimer un LSP (ou une réservation de ressources), confirmer une réservation ou notifier des erreurs et garder l'adjacence entre routeurs RSVP-TE voisins, le protocole RSVP-TE utilise neuf types de messages :

Path Ce message est responsable de l'établissement et du maintien d'un LSP dans le sens descendant. Il est envoyé par l'émetteur au(x) récepteur(s) pour déterminer la liste des routeurs du chemin qui sera utilisée pour l'envoi des données. Ce message permet aussi de spécifier d'autres informations comme les caractéristiques du flux de données supporté.

Resv Ce message a pour rôle d'établir et maintenir un LSP dans le sens montant. C'est un message de réponse à la requête transmise dans le message *path* du même LSP afin d'effectuer la réservation de ressources.

PathErr Ce message est retourné à l'émetteur pour lui signaler une erreur dans le sens descendant (exemple : boucle ou chemin inexistant).

ResvErr Ce message indique une erreur dans le sens montant (exemple : ce message peut signaler un manque de ressources ou un refus de droits d'accès aux ressources).

PathTear La réception de ce message a pour effet d'effacer les états RSVP-TE concernant le LSP associé. Cela implique l'annulation des ressources utilisées par le LSP précédent.

ResvTear Ce message indique aux routeurs qui le reçoivent d'annuler les états de réservation montants. L'information d'état RSVP-TE créée par les messages *path* est conservée.

ResvConf (optionnel) C'est un message envoyé par un routeur au demandeur de la réservation. La réception d'un tel message indique que la réservation est réussie avec une probabilité élevée (ce message n'est sûr à 100 % que quand il est émis par le dernier routeur sur le chemin vers la station émettrice).

Srefresh Ce message rafraîchit un ensemble de sessions RSVP-TE.

Hello (optionnel) Ce message gère l'adjacence entre deux routeurs voisins RSVP-TE.

1.3.3.2 Tunnel RSVP-TE et LSP

Pour des considérations liées à l'ingénierie de trafic, le protocole RSVP fait une distinction entre la notion de *tunnel RSVP-TE* et la notion de LSP [LR06] :

1. un tunnel RSVP-TE[5]est une entité de routage unidirectionnelle (unicast ou multicast), avec des contraintes TE. Il est instancié par un ou plusieurs LSP, chacun correspondant à un chemin particulier du tunnel RSVP-TE.
2. un *LSP* est un chemin MPLS caractérisé par une distribution d'étiquettes. C'est une instance du tunnel RSVP-TE.

Au sein d'un tunnel RSVP-TE, plusieurs LSP peuvent être actifs simultanément mais un seul peut être utilisé à la fois (à tout instant). Le nombre d'instances de LSP associés à un tunnel RSVP-TE est variable et peut changer au cours du temps.

Un tunnel RSVP-TE à instances multiples a plusieurs applications :

1. il permet d'optimiser les ressources par changement de route du tunnel RSVP-TE, sans perte de trafic et sans compter deux fois la bande passante sur les liens communs au nouveau et ancien LSP. Cette procédure de changement de route d'un tunnel RSVP-TE est appelée aussi *make before break* et elle consiste à établir un nouveau LSP partageant la bande passante avec l'ancien LSP (sur les liens qui leurs sont communs) et reliant les mêmes nœuds source et destination avant de basculer le trafic de l'ancien LSP vers le nouveau.
2. il facilite la configuration des LSP de secours en permettant de les associer à un LSP primaire (cf. section 1.5.3).

1.3.3.3 Objets RSVP-TE

Un message RSVP-TE est constitué d'un entête commun à tous les messages RSVP, suivi du corps du message qui consiste en un nombre variable d'objets dépendant du type du message. Ces objets RSVP-TE sont eux même composés d'un entête et d'un corps. L'entête inclut la classe, la sous-classe et la longueur de l'objet alors que le corps contient les valeurs de l'objet RSVP.

Plusieurs objets RSVP-TE sont définis, parmi lesquels nous citons les principaux :

SESSION Cet objet permet d'identifier une *session* qui correspond à *un flux de données* routé sur un tunnel RSVP-TE. Cet objet inclut un identifiant, l'adresse source[6], l'adresse de destination du tunnel RSVP-TE et permet d'identifier sans ambiguïté le tunnel RSVP-TE dans le réseau.

[5]La notion d'un *tunnel RSVP-TE* est indépendante de la notion de *tunnel TE-trunk* défini dans [AMA+99] pour l'agrégation d'un ensemble de LSP qui peuvent transporter plusieurs flux de données simultanément.

[6]Souvent, c'est l'adresse source du tunnel RSVP-TE qui transmise dans le champ *Extended Tunnel ID* de l'objet SESSION. Cela permet d'avoir un identifiant de tunnel RSVP-TE dont la porté est propre à la paire constituée de la source et de la destination du tunnel RSVP-TE (utile pour assurer l'unicité des identifiants des tunnels RSVP-TE lorsque les objets SESSION sont construits par les routeurs de têtes des tunnels RSVP-TE).

SENDER_TEMPLATE et FILTER_SPEC Ces deux objets ont la même structure : le premier est transmis dans les messages *path* et le second est envoyé dans les messages de réponse *resv*. Ils contiennent l'adresse de la source du tunnel RSVP-TE ainsi qu'un identifiant de LSP. La combinaison de l'un de ces deux objets avec avec l'objet SESSION permet d'identifier d'une manière unique et non ambiguë un LSP.

SENDER_TSPEC Cet objet définit les caractéristiques du flux de données de l'émetteur. Il est requis dans les messages *path*.

FLOWSPEC Cet objet définit la QoS désirée et il est transmis dans les messages *resv*.

LABEL_REQUEST Cet objet indique qu'une allocation d'étiquette est requise et il spécifie le protocole de niveau 3 dans le modèle ISO qui est utilisé (ce protocole est souvent IP mais cela n'est pas obligatoire).

LABEL Il s'agit de l'étiquette allouée sur un LSR à un LSP appartenant à une session donnée.

EXPLICIT_ROUTE (ERO) Cet objet contient la liste des routeurs (stricts) que doit traverser le LSP en cours d'établissement. Il n'est transmis que dans les messages *path*.

RECORD_ROUTE (RRO) Cet objet contient la liste des routeurs traversés par un LSP et optionnellement les étiquettes associées sur chaque routeur. Cet objet peut être envoyé dans un message *path* auquel cas il inclut la liste des routeurs situés entre le nœud source et le nœud qui le reçoit ou bien il peut être envoyé dans un message *resv* auquel cas il liste les routeurs situés entre la destination et le nœud recevant le message.

RSVP_HOP Dans un message *path* reçu, cet objet renseigne l'adresse IP du routeur RSVP précédent sur le LSP alors que dans un message *resv*, cet objet indique l'adresse IP du routeur suivant sur le LSP. C'est grâce à cet objet que l'itinéraire emprunté par les messages *path* et *resv*, associés à un LSP, est fixé et est déterminé.

SESSION_ATTRIBUTE Cet objet contient des informations sur les priorités de préemption et de maintien d'un LSP, des affinités (lien à inclure ou à exclure d'un LSP) et des drapeaux pour indiquer si la protection et l'enregistrement des étiquettes sont requis.

1.3.3.4 Établissement d'un LSP

L'établissement d'un LSP avec RSVP-TE passe par deux phases : la phase descendante et la phase montante. Dans la première phase, le LSR de tête du LSP envoie un message *path* au LSR de sortie pour fixer la route et transmettre l'ensemble des paramètres TE (source et destination du LSP, identifiants du tunnel et du LSP, bande passante, affinités, etc.). Dans la deuxième phase, un message *resv* est envoyé par le LSR de sortie au LSR de tête afin de réserver les ressources (bande passante et étiquettes) sur tous les liens du LSP en cours d'établissement.

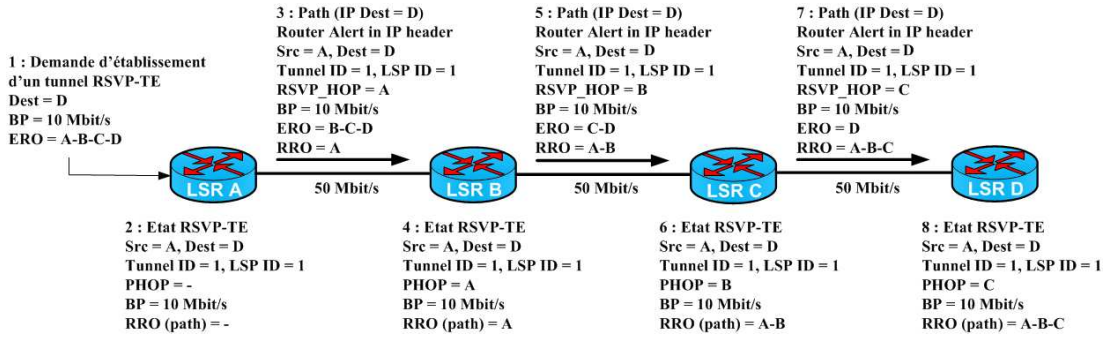


FIG. 1.7 – Propagation du message *path* et création d'états RSVP

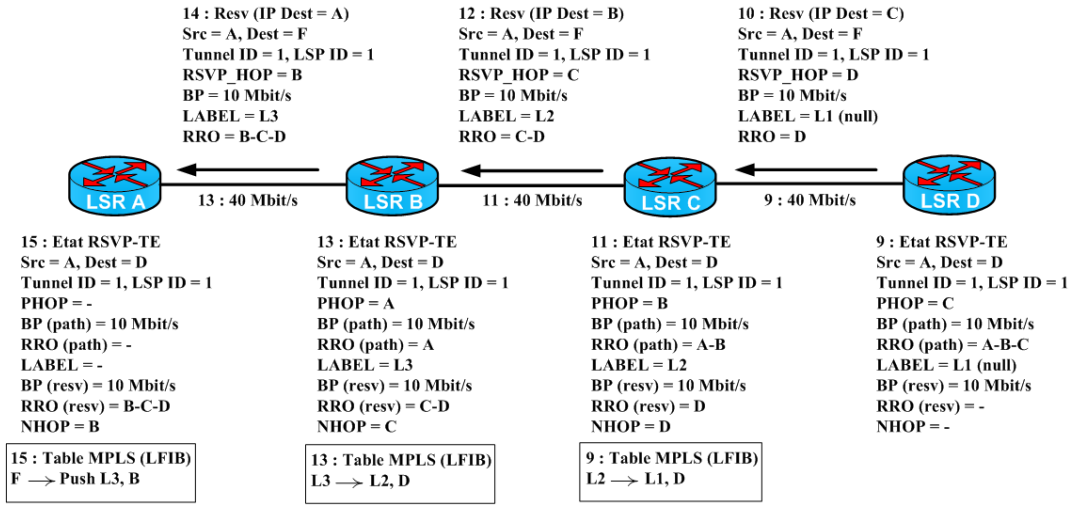


FIG. 1.8 – Propagation du message *resv* et mise-à-jour des états RSVP

Message *path*

Afin d'initier l'établissement d'un LSP (dans le sens descendant), le LSR de tête construit un message *path* (cf. figure 1.7). Ce message contient des informations concernant la session (identifiant du tunnel et adresse de la destination), l'émetteur (adresse du LSR de tête et identifiant du LSP), les caractéristiques du trafic (bande passante et un ensemble de paramètres de classe de service), la structure de la la route choisie et ses propriétés (optionnellement la route explicite et les liens à inclure ou à exclure) ainsi que le protocole de niveau 3 du modèle ISO acheminé. Ce message *path* inclut aussi des informations utiles comme la route sélectionnée (objet RRO) qui sera communiquée à tous les LSR du LSP, la porté des étiquettes allouées (étiquette globale au LSR ou locale à une interface d'un LSR) et des indications concernant la protection (protection locale fournie ou pas, protection en cours d'utilisation ou pas, etc.). Avant d'envoyer le

message au prochain LSR permettant d'atteindre la destination, le LSR de tête crée un état RSVP constitué d'informations déduites du message *path* précédent (typiquement les identifiants du LSP et du tunnel RSVP-TE, les adresses source et destination, le prochain et le précédent nœud, la bande passante réclamée, etc.).

Le message *path* transite de proche en proche du LSR de tête jusqu'à atteindre le LSR de sortie. Chaque LSR de transit qui le reçoit le traite. Si aucune erreur n'est détectée, le LSR créé un état RSVP déduit du message reçu avant de transmettre le message *path* traité au prochain routeur, déterminé à partir de l'objet ERO s'il y en un (et si le prochain routeur est strict), sinon à partir d'informations communiquées par le protocole de routage (souvent c'est la table IP qui sera consultée). Si par contre, une erreur est détectée (par exemple, la route est invalide), un message d'erreur (message *resvErr*) est envoyé à l'émetteur. La propagation du message *path* ainsi que les états RSVP créés sur les LSR du LSP sont illustrés sur la figure 1.7.

Message *resv*

Pour réserver les ressources, le LSR de sortie répond au LSR de tête (après la réception du message *path*) par un message *resv* qui transitera de proche en proche, en empruntant le chemin inverse du message *path* correspondant. Pour éviter des traitements supplémentaires, la destination IP de chaque message *resv* sur un LSR est extraite de l'objet RSVP_HOP de l'état RSVP créé lors de la réception du message *path* associé (en conséquence, les routeurs non RSVP parcourus par le message *resv* ne transmettront pas le paquet au driver RSVP).

Tous les routeurs du LSP (sauf le routeur de tête) effectuent un contrôle d'admission sur leur lien précédent. Si ce contrôle d'admission réussit, de la bande passante et une étiquette MPLS[7]seront allouées, ce qui induit une mise-à-jour de l'état RSVP associé au LSP, de la table de commutation MPLS (LFIB) et de la bande passante résiduelle disponible sur le lien[8]. Après traitement et modification de certains objets (objets ERO, RRO, etc.) du message *resv* reçu, ce dernier est envoyé au LSR en amont (cf. figure 1.8). Si, par contre, le contrôle d'admission ou l'allocation d'étiquette échoue, un message d'erreur (message *resvErr*) est envoyé au LSR de sortie du LSP.

A la réception du message *resv* associé au LSP, le routeur de tête met-à-jour son état RSVP et ses tables de routage IP et LFIB (cf. figure 1.8). A ce moment, le LSP sera prêt pour acheminer le trafic.

1.3.3.5 Suppression d'un LSP

Un LSP peut être supprimé explicitement ou implicitement avec RSVP-TE. Pour supprimer explicitement un LSP, un message *pathTear* est envoyé de la source vers la

[7]Dans les premières versions de RSVP-TE, seul le mode *downstream-on-demand* était préconisé pour l'allocation d'étiquette. De nouvelles extensions ont été apportés à RSVP-TE par [Ber03] pour permettre une allocation d'étiquettes dictée par le nœud en amont. Ainsi, grâce à l'objet "Suggested Label", le nœud en amont est désormais capable de proposer une étiquette permettant d'aiguiller le trafic d'un LSP vers un nœud aval.

[8]La bande passante résiduelle disponible sur le lien est diminuée de la bande passante allouée au LSP sur tous les niveaux de priorité inférieure ou égale à celle du LSP.

destination. Ce message permet de détruire l'ensemble des états RSVP sur les LSR du LSP. Un autre message *resvTear*, envoyé de la destination vers la source du LSP, peut aussi détruire le LSP. Cependant et contrairement au premier message (*pathTear*), le message *resvTear* ne supprime des états RSVP que l'information de réservation de ressources (typiquement, les étiquettes MPLS et les bandes passantes allouées sur les liens). Afin de détruire complètement le LSP, le message *resvTear* doit être suivi d'un message (*pathTear*) pour effacer totalement les états RSVP associés au LSP détruit.

1.4 Protection contre les pannes

1.4.1 Introduction

Avec l'augmentation et la multitude des types d'applications temps réel (TV, VoIP, visioconférence, télémédecine, etc.) déployées sur l'Internet, la récupération rapide lors des pannes touchant les composants du réseau devient de plus en plus désirée pour satisfaire les fortes contraintes de temps de ces applications. Pour ce faire, diverses techniques de résistance aux pannes ont été développées. Ces dernières peuvent être classées en deux catégories : restauration (protection réactive) et protection (protection proactive).

Avec la restauration [RM99b, MK98], aucun traitement n'est effectué avant l'occurrence d'une panne. Pour pallier une panne, de nouveaux chemins de secours contournant le composant défaillant sont calculés et configurés afin de router le trafic des communications affectées. Cette catégorie de techniques de résistance aux pannes a l'avantage de la simplicité et de la flexibilité vis à vis des changements de la topologie du réseau. Cependant, elle induit des délais de récupération élevés et inacceptables pour beaucoup d'applications temps réel comme la voix ou la TV sur IP.

Pour diminuer les délais de récupération et satisfaire ainsi les fortes contraintes de temps des applications répandues sur l'Internet, la protection est souvent préférée à la restauration. Avec la protection [RM99a, MVDBD04, MFB99], les chemins de secours sont calculés et souvent préconfigurés avant l'occurrence des pannes. Concrètement, pour rétablir des communications affectées suite à une panne, il suffit de basculer le trafic de ces communications sur les chemins de secours qui leur sont associés. Les délais de restauration sont ainsi diminués puisqu'aucun calcul de chemins (et souvent pas de configuration de chemins) n'est effectué après l'occurrence de la panne.

Selon le niveau de la couche (Liaison de donnée ou Réseau) assurant la protection, nous distinguons trois classes principales de techniques de protection : protection de niveau Liaison de données, protection de niveau Réseau et protection mixte.

Avec les techniques de la première classe [GS98, LCC, Moy04, YJ05], la protection des communications est réalisée au niveau de la couche Liaison de données, ce qui permet de réduire considérablement les délais de récupération à moins de 100 ms. Cette classe de techniques de protection a l'avantage de réduire les états d'acheminement à maintenir (puisque tous les flux de données sont agrégés sur les liens) mais elle présente divers inconvénients parmi lesquels nous citons la non prise en charge de la QoS et de la matrice de trafic réelle et la sous-utilisation de la topologie du réseau. Différentes

méthodes appartenant à cette classe sont connues comme :

1. Protection par anneau dans SDH : avec cette méthode, le trafic est dupliqué et est envoyé sur deux anneaux de sens inverses. Le signal de meilleure qualité est sélectionné à la réception. Lors de l'occurrence d'une panne, il suffira de sélectionner le seul signal reçu pour pallier la panne. Cette méthode de protection permet d'obtenir des délais de récupération inférieurs à 50 ms mais en consommant le double en bande passante, c'est pourquoi on l'appelle la *protection dédiée 1+1*.
2. P_cycle : cette méthode de protection est proposée par Grover et al. dans [GS98]. Elle consiste à pré-établir des cycles dans un réseau. Lorsqu'un lien ou un nœud d'un cycle tombe en panne, le trafic est basculé localement sur l'autre partie du cycle qui n'est pas affectée. Cette méthode de protection diminue la consommation de la bande passante mais elle ne peut être appliquée que hors ligne sur des réseaux stables.

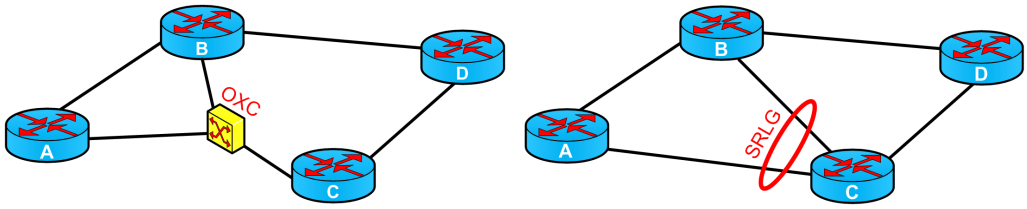
La seconde classe de techniques de protection permet de tenir compte des changements de la matrice de trafic et de la topologie du réseau tout en assurant la QoS désirée après une panne. Cela a un coût au niveau du plan de contrôle puisqu'il faudra collecter, stocker et maintenir l'information concernant le trafic, la topologie du réseau et la QoS. Les méthodes de protection appartenant à cette classe sont souvent implantées dans des plate-formes de types MPLS [PSA05].

La troisième classe de techniques de protection [BCG+07] combine les deux premières classes afin de maximiser la protection et réduire les délais de protection.

Dans cette section, nous nous intéressons qu'aux techniques de protection de la seconde classe (i.e. protection de niveau Réseau) vu leur flexibilité à prendre en compte les changements de la matrice de trafic et leur capacité à offrir la QoS désirée.

1.4.2 Notion de risque de panne

Pour pallier les pannes des composants d'un niveau quelconque (*niveau physique*) à un niveau supérieur (*niveau logique*), trois types de risques de panne ont été définis [KKL+01, VCLF+04, KR05c] : lien, nœud et groupe de liens (*Shared Risk Link Group* ou SRLG). Le premier risque correspond à la panne d'un lien logique due à la défaillance d'un composant physique exclusif au lien. Sur une topologie d'un réseau, tous les liens logiques correspondent à des risques de panne de type lien.



(a) Topologie physique (b) Topologie logique

FIG. 1.9 – Correspondance entre topologies physique et logique d'un même réseau

Comme pour le premier risque de panne, nous définissons le second comme un risque de panne d'un nœud logique lié à la défaillance d'un composant physique exclusif au nœud. Sur une topologie logique d'un réseau, tous les nœuds correspondent à des risques de panne de type nœud.

Enfin, pour traiter les pannes simultanées de plusieurs liens logiques dues à une défaillance d'un seul composant physique, le troisième type de risque de panne (SRLG) a été défini [KR05c]. Ce dernier est constitué d'un ensemble de liens logiques partageant un composant physique dont la panne affecte tous les liens formant le SRLG.

Sur la figure 1.9 sont illustrées deux topologies correspondant à un même réseau : la première topologie de réseau (figure 1.9 (a)) montre les composants de niveau Liaison de données du réseau représenté alors que la seconde topologie du réseau (figure 1.9 (b)) ne montre que les composants de niveau Réseau. Les liens de niveau Liaison de données *A*-*B*, *B*-*D* et *C*-*D* relient directement des routeurs visibles (*A*, *B*, *C* et *D*) au niveau Réseau : ce sont donc des liens de niveau Réseau aussi. Par contre, les liens *A*-*OXC*, *B*-*OXC* et *OXC*-*C* ne sont pas visibles au niveau Réseau puisque le commutateur *OXC* n'est pas un routeur. Ce sont des liens de niveau Liaison de données (figure 1.9 (a)) utilisés pour interconnecter le routeur *C* au routeurs *A* et *B* : ils sont donc remplacés par les deux liens de niveau Réseau *C*-*A* et *C*-*B* dans la topologie du réseau illustrée sur la figure (figure 1.9 (b)).

Pour faciliter le traitement des pannes simples des composants de niveau Liaison de données (figure 1.9 (a)), les liens et nœuds de la topologie de niveau Réseau (figure 1.9 (b)) seront groupés dans 3 classes de risques de pannes :

1. risques de panne de type lien : incluant tous les liens de la topologie logique du réseau (*A*-*B*, *A*-*C*, *B*-*C*, *B*-*D* et *C*-*D*),
2. risques de panne de type nœud : incluant tous les routeurs du réseau (*A*, *B*, *C* et *D*),
3. risque de panne de type SRLG : constitué des liens de niveau Réseau (*A*-*C* et *B*-*C*) qui partagent un composant de niveau Liaison de donnée (le commutateur *OXC* ou le lien *OXC*-*C*).

Malgré la présence des SRLG dans beaucoup de réseaux actuels, les travaux de recherche qui leur sont consacrés sont rares. Dans cette thèse, nous pencherons de près sur le problème de protection des SRLG afin de proposer des mécanismes génériques de placement distribué des LSP de secours.

Notons que les deux protocoles OSPF-TE [KR05b] et ISIS-TE [KR05a] ont été étendus pour transporter l'information sur les risques de panne de type SRLG. De plus, pour traiter efficacement les pannes multiples de liens, des *SRLG virtuels* peuvent être définis. Ces derniers sont constitués de tuples de liens pouvant/risquant de tomber en panne simultanément.

1.4.3 Phases de récupération

Lors de la récupération d'une panne, différentes opérations sont effectuées dans chacune des phases suivantes [MP06] :

Détection d'une défaillance
C'est lors de cette phase qu'une défaillance est détectée par un composant (souvent par un nœud du réseau) qui la transmettra à sa couche IP/MPLS s'occupant de la récupération.

Localisation (et isolation) de la panne
Cette phase permet à l'entité s'occupant de la récupération de localiser et peut-être d'identifier le risque défaillant. Pour des raisons de performances, seule la localisation de la panne est fournie : la distinction entre les trois risques de pannes définis précédemment n'est souvent pas rendue.

Notification de la panne
Dans cette phase, la panne est signalée à l'entité s'occupant de la récupération. Cette phase n'est pas apparente lorsque c'est la protection locale qui est employée puisque l'entité détectant la panne est la même que celle qui s'occupe de la récupération.

Récupération
Dans cette phase, un ou plusieurs LSP de secours sont choisis (et peut être signalisés) pour aiguiller le trafic pendant la panne. Les tables MPLS et les tables IP sont ainsi modifiées pour tenir compte des changements de routes.

Normalisation
Après la réparation de la panne, la normalisation, qui consiste à rebasculer le trafic vers l'ancien LSP primaire, peut ou ne pas être effectuée. Cette phase est préconisée pour ré-optimiser les ressources et n'est pas nécessaire pour réaliser la récupération.

A l'exception de la première phase, toutes les autres phases sont complètement effectuées par le plan de contrôle. En conséquence, l'anticipation et la minimisation des traitements effectués par le plan de contrôle (lors des phases de localisation, notification et récupération) permettrait de diminuer sensiblement les délais de récupération. Cette approche d'anticipation et de minimisation des traitements effectués par le plan de contrôle sera d'ailleurs adoptée dans tous les chapitres suivants.

1.4.4 Protection sous MPLS

MPLS permet d'implanter les deux types de protection : globale et locale.

1.4.4.1 Protection globale

Avec la protection globale, un message de notification de la panne est envoyé au routeur de tête de chaque LSP primaire affecté qui basculera ensuite le trafic vers le(s) chemin(s) de secours. Différentes techniques de protection globale existent sous MPLS, parmi lesquelles nous citons :

Protection de bout-en-bout (end-to-end protection) 1+1 ou 1:1
Deux chemins reliant la source à la destination de la connexion à protéger sont établis : le premier est un chemin primaire et le second est son chemin de secours.

Pour une meilleure résistance contre les pannes (simples) de liens (resp. de liens et nœuds), les deux chemins précédents ne doivent pas partager de liens (resp. de liens et nœuds).
Deux versions de la protection de bout-en-bout existent : 1+1 (protection dédiée) et 1:1 (protection partagée). Dans la protection 1+1, le trafic est dupliqué à la source pour être envoyé ensuite sur les deux chemins primaire et de secours. Pour des considérations d'optimisation, les deux chemins primaire et de secours correspondent aux deux chemins les plus courts [Bha99]. Nous notons qu'avec la protection 1+1, le chemin de secours est complètement dédié à la protection et ne peut être utilisé pour la transmission d'un flux autre que celui qui est protégé. Dans la protection 1:1, le trafic de la connexion protégée n'est transmis que sur un seul chemin. En l'absence de pannes, le chemin de secours peut être utilisé pour le transport d'autres flux de données.

Protection 1:N (N>1)
N chemins primaires reliant la même source à la même destination sont protégés par un chemin de secours qui interconnecte les mêmes nœuds. Pour une meilleure résistance contre les pannes (simples) de liens (resp. de liens et nœuds), les N chemins primaires précédents ne doivent pas partager de liens (resp. de liens et nœuds). Lors d'une panne touchant un (ou plusieurs) LSP primaire, un seul LSP primaire est choisi pour être restauré.
En l'absence de pannes, le chemin de secours peut être utilisé pour le transport d'autres flux que ceux protégés.

Protection M:N (N>M)
Un ensemble de M chemins de secours est utilisé pour protéger un ensemble de N chemins primaires. Bien évidemment, tous les chemins primaires et leurs secours relient le même nœud source au même nœud de destination. Au plus M chemins primaires peuvent être restaurés par les chemins de secours. De plus, tous les chemins de secours disponibles (non utilisés pour la récupération) peuvent être utilisés pour le transport d'autres flux que ceux protégés.

Malgré sa simplicité, la protection globale souffre de deux principaux problèmes réduisant son intérêt dans la pratique :

1. Délais de récupération non bornés : la notification de la panne doit aller du nœud ayant détecté la panne jusqu'au nœud source du LSP primaire. Cela allonge les délais de récupération, surtout dans les réseaux larges où la panne risque d'être loin de la source.

2. NP-difficulté de la recherche de deux chemins SRLG disjoints : avec l'introduction des SRLG dans les réseaux d'aujourd'hui, le problème de détermination d'au moins deux chemins SRLG disjoints (i.e. des chemins ne partageant aucun lien appartenant à un SRLG) devient NP-difficile [Hu03].

1.4.4.2 Protection locale

La protection locale MPLS-TE, appelée aussi *MPLS fast reroute*, est définie dans [PSA05]. Elle permet de garantir des délais de récupération inférieurs à 100 ms par l'élimination de la phase de notification de la panne. Ainsi, avec la protection locale MPLS-TE, tout LSR (excepté le LSR de sortie) du LSP primaire configure un LSP de secours permettant d'atteindre le LSR de sortie par une route contournant le prochain lien (ou prochain routeur) du LSP primaire. De cette manière, lors de la détection d'une panne, c'est le LSR amont du lien (ou nœud) affecté, appelé aussi PLR (*Point of Local Repair*), qui réagit localement et rapidement à la panne en basculant le trafic vers le LSP de secours permettant la récupération.

Deux techniques de protection locale sont largement déployées dans les réseaux MPLS : la protection par LSP de détour et la protection par tunnels de secours. Dans la première technique, chaque LSR construit un LSP de secours séparé, dit *LSP de détour*, pour chacun des LSP primaires qu'il protège. Dans la deuxième technique par contre, un LSR utilise un même LSP de secours, dit *tunnel de secours*, pour protéger plusieurs LSP primaires traversant les deux nœuds d'extrémités du tunnel de secours.

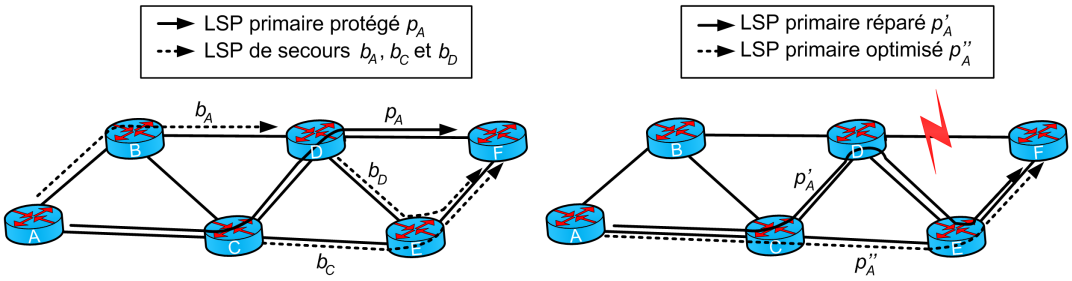
La communauté n'a pas convergé vers l'une des deux techniques ci-dessus car les performances que l'on peut attendre de chacune d'elles sont très liées à la topologie du réseau. Pour des raisons d'interopérabilité, les constructeurs fournissent en général des implantations supportant les deux techniques de protection précédentes.

Bien que toutes les techniques de protection 1+1, 1:1, 1:N ou M:N peuvent être combinées avec la protection locale, nous nous intéressons ici qu'à la protection locale 1:1 qui permet de réduire la consommation des ressources.

1.4.4.3 Protection par LSP de détour (*one-to-one backup protection*)

Dans cette méthode de protection, chaque LSR du LSP protégé (excepté le LSR de sortie) établit un LSP de secours (LSP de détour) protégeant contre son prochain lien et éventuellement contre son prochain nœud sur le LSP primaire. Pour ce faire, deux types de LSP de secours sont utilisés [PSA05] : LSP NNHOP (Next Next HOP LSP) et LSP NHOP (Next HOP LSP). Un LSP NNHOP est un chemin de secours établi entre un PLR (LSR du LSP primaire protégé) et un LSR situé en aval de son prochain LSR sur le LSP primaire protégé. Il permet de protéger contre les pannes du prochain lien et du prochain nœud du PLR. De la même manière, un LSP NHOP est défini comme un chemin de secours reliant un PLR à un LSR situé à son aval sur le LSP primaire protégé. Ce type de LSP de secours permet de protéger contre la panne du prochain lien du PLR. Bien entendu, les deux types de LSP de secours doivent croiser le LSP primaire en un nœud de fusion situé en aval du PLR et appelé MP (*Merge Point*).

Afin de protéger localement un LSP primaire de N (N>1) LSR, N-1 LSP de secours sont requis ($Max(N-2,0)$) LSP NNHOP reliant les $Max(N-2,0)$ premiers PLR au LSR de sortie et 1 LSP NHOP reliant le LSR aval au LSR de sortie à ce dernier). Sur la figure 1.10(a), un LSP primaire p_A passant successivement par les quatre nœuds (A, C, D et F) est établi et est protégé localement en utilisant trois LSP de détour b_A, b_C et b_D. Le premier (resp. second) LSP de secours b_A (resp. b_C), qui est formé de la



(a) Avant la panne (b) Pendant et après la panne

FIG. 1.10 – Protection locale par LSP de détour

séquence des nœuds $A \rightarrow B \rightarrow D$ (resp. $C \rightarrow E \rightarrow F$), est de type NNHOP puisqu'il protège contre la panne d'un nœud (et d'un lien), en l'occurrence le nœud C (resp. le nœud D). Le dernier LSP de secours b_D ($D \rightarrow E \rightarrow F$) est établi entre le nœud PLR D situé en aval du LSR de sortie F et le LSR de sortie F : il est de type NHOP puisqu'il ne peut protéger que contre la panne du dernier lien sur le LSP primaire.

Pour le calcul des LSP de détour vérifiant un ensemble de contraintes, l'algorithme CSPF suivant pourrait être adopté :

1. Éliminer de la topologie du réseau le nœud et le lien à protéger (situés en aval du PLR sur le LSP primaire) ainsi que tous les liens appartenant à un même SRLG que le lien à protéger.
2. Éliminer de la topologie du réseau tous les liens ne vérifiant pas l'ensemble des contraintes.
3. Exécuter l'algorithme SPF pour déterminer un chemin reliant le PLR au LSR de sortie. Pour augmenter le taux de protection, l'arbre des plus courts chemins qui est utilisé et qui est enraciné à la destination, doit inclure la partie du chemin primaire située entre le prochain routeur du prochain routeur du PLR et le LSR de sortie si le LSP de secours recherché est un LSP NNHOP. Si par contre, le LSP de secours recherché est de type NHOP, c'est toute la partie du chemin primaire située entre le prochain routeur du PLR et le LSR de sortie qui sera ajoutée à l'arbre des plus courts chemins à l'initialisation.

Pour minimiser le coût de l'établissement et du maintien des LSP de secours, ces derniers sont souvent fusionnés entre eux sur un routeur commun s'ils suivent la même route vers la destination. Ils sont aussi fusionnés avec le LSP primaire protégé sur un routeur commun s'ils ont la même interface de sortie et le même prochain nœud. Sur l'exemple de la figure 1.10 (a), le LSP de détour permettant de protéger contre la panne du nœud C (et du lien A-C) et retourné par l'algorithme CSPF ci-dessus est $A \rightarrow B \rightarrow D \rightarrow F$. Ce chemin croise le LSP primaire au LSR D et a un même interface de sortie ainsi qu'un même prochain nœud de sortie avec le LSP primaire sur le LSR D : il est donc fusionné avec le LSP primaire pour obtenir le LSP de secours b_A $A \rightarrow B \rightarrow D$. De la même manière, les deux LSP de secours b_C et b_D se croisent au LSR E où ils ont le même interface de sortie et le même prochain LSR : ils peuvent donc être fusionnés

(agrégés) au LSR D (une seule étiquette MPLS est allouée pour ces deux LSP de secours sur le lien $E \rightarrow F$).

Lorsqu'une panne est détectée le long d'un LSP protégé, le PLR redirige le trafic localement vers son LSP de détour. Par exemple, si le lien $D \rightarrow F$ tombe en panne dans la figure 1.10 (a), le PLR D basculera le trafic du LSP primaire vers le LSP de secours b_D, comme illustrée sur la figure 1.10 (b) (LSP p'). Pour réoptimiser les ressources après l'occurrence de la panne du lien $D \rightarrow F$, le LSR de tête A du LSP primaire pourrait décider de recalculer un nouveau chemin ($p"_A$) et d'utiliser la procédure *make before break* de RSVP pour le configurer (sans perte de trafic et sans double allocation de la bande sur les liens en commun).

1.4.4.4 Protection par tunnel de secours (*facility backup protection*)

Avec cette technique de protection, au lieu de consacrer un LSP séparé pour chaque LSP protégé sur un PLR donné, un seul LSP de protection, dit *tunnel de secours*, est utilisé et est partagé entre plusieurs LSP protégés. En d'autres termes, le tunnel de secours est construit pour protéger plusieurs LSP primaires en même temps. Le tunnel de secours doit commencer en un nœud PLR et se terminer en un nœud du LSP protégé, dit MP ou *Merge Point*. Cela réduit l'ensemble des LSP primaires pouvant partager un tunnel de secours à ceux passant par le PLR et le MP. Comme dans la protection par LSP de détour, deux types de tunnels de secours sont employés : NNHOP et NHOP. Un tunnel NNHOP relie un nœud PLR au prochain LSR (dit MP) d'un prochain LSR r du PLR : il permet donc de protéger contre les pannes du nœud r et du lien *PLR-r*. Un tunnel NNHOP relie un nœud PLR à un prochain LSR (dit MP) du PLR : il permet de protéger contre la panne du lien *PLR-MP*.

La technique de protection par tunnels de secours est rendue possible sous MPLS grâce à la hiérarchie qui permet d'empiler plusieurs étiquettes dans un même paquet de données. Ainsi, pour contourner une panne donnée, deux étiquettes sont insérées dans les piles MPLS des paquets appartenant aux différents LSP protégés par le tunnel de secours qui contourne la panne : l'étiquette en tête de pile MPLS permet aux paquets de suivre le chemin correspondant au tunnel de secours activé pour pallier la panne et l'étiquette qui la suit permet d'aiguiller les différents paquets vers leur LSP primaires à la sortie du tunnel.

Sur la figure 1.11, un tunnel de secours b_B ($B \rightarrow G \rightarrow H \rightarrow D$) est construit par le routeur B pour pallier une éventuelle panne du lien *B-C* ou du nœud C. Ce tunnel de secours peut être utilisé pour protéger n'importe quel LSP traversant successivement les nœuds B, C et D. Typiquement, sur la figure 1.11, le tunnel de secours b_b est employé pour protéger les trois LSP primaires p_A ($A \rightarrow B \rightarrow C \rightarrow D \rightarrow E$), p_B ($B \rightarrow C \rightarrow D$) et p_F ($F \rightarrow B \rightarrow C \rightarrow D \rightarrow H$) qui traversent le segment de chemin $B \rightarrow C \rightarrow D$.

En l'absence de pannes, le trafic émis par les LSR A, B et F est acheminé sur les trois LSP primaires p_A , p_F et p_B en utilisant des étiquettes différentes sur les liens en commun *B-C* et *C-D* (cf. figure 1.12 (a)). Lorsqu'une panne du LSR C est détectée (figure 1.12 (b)), les flux des trois LSP primaires (p_A , p_F et p_B) qui le traversent sont agrégés et basculés vers le tunnel de secours b_B. Dans le cas où les étiquettes sont

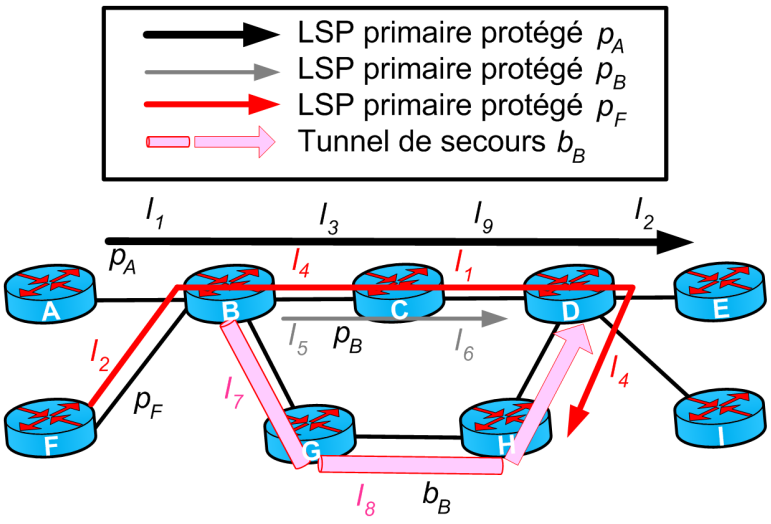


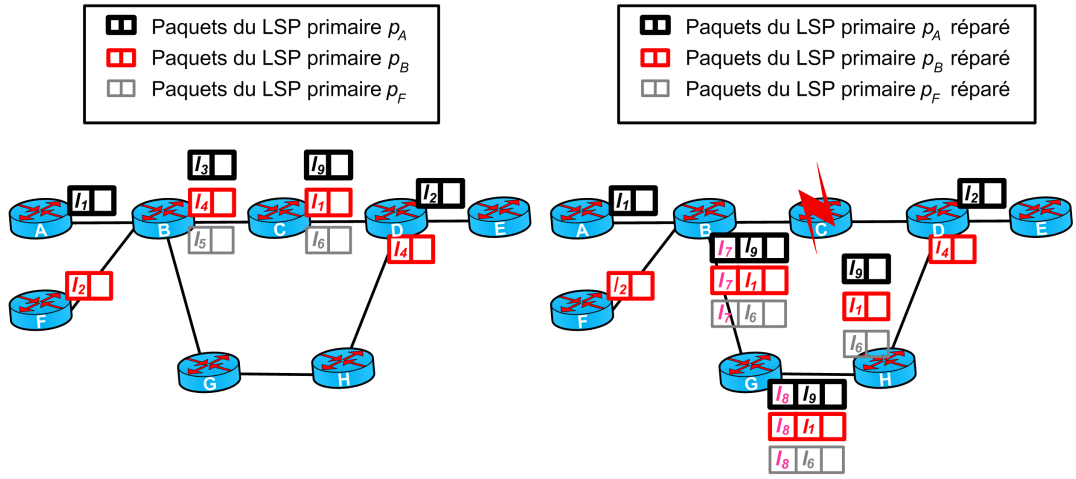
FIG. 1.11 – Protection par tunnel de secours

globales au LSR D (cas de la figure 1.12 (b)), le PLR B échangera l'étiquette l_1 (resp. l'étiquette l_1) associée au LSP primaire p_A (resp. p_F) par l'étiquette l_2 (resp. l'étiquette l_4) allouée par le MP D sur le lien $C \rightarrow D$ pour le LSP primaire p_A (resp. p_F). Nous notons que les deux étiquettes l_2 et l_4 ne seront utilisées qu'à la sortie du tunnel de secours b_B (i.e. sur le MP D) pour séparer les flux des LSP primaires protégés. Après la première opération d'empilement, le PLR D insère en tête de la pile MPLS une deuxième étiquette l_7 permettant l'acheminement sur le tunnel de secours b_B.

Lorsque le PLR est un routeur de tête du LSP protégé, la première opération effectuée n'est pas un échange d'étiquette mais une insertion d'étiquette. Concrètement, pour rétablir le LSP primaire p_B lors de la panne du LSR C, le PLR B insère dans un premier temps l'étiquette l_6 allouée par le MP D au LSP protégé sur le lien $C \rightarrow D$ (pour la séparation des flux sur le MP D) puis une deuxième étiquette l_7 permettant le ré-acheminement du trafic sur le tunnel de secours b_B (cf. figure 1.12 (b)).

Lorsque l'allocation des étiquettes n'est pas globale au MP D, ce dernier devra allouer une étiquette différente (située derrière l'étiquette identifiant le tunnel de secours) à chaque LSP primaire protégé pour permettre la séparation du trafic. Cette opération d'allocation d'étiquette peut se faire par signalisation après l'occurrence d'une panne (délai de récupération allongé) ou avant (pour minimiser le délai de récupération) .

Bien que le nombre de tunnels de secours nécessaires à la protection d'un LSP primaire ne soit pas différent du nombre de LSP de détour requis à la protection du même LSP primaire (N-1 tunnels de secours pour un LSP protégé de N nœds), la technique de protection par tunnels de secours réduit considérablement le nombre total de LSP (de secours) dans le réseau. En effet, grâce au partage d'un même tunnel de secours pour la protection de plusieurs LSP primaires, le nombre de LSP est diminué, ce qui permet une meilleure résistance au passage à l'échelle. Cependant, il faut noter



(a) Avant la panne (b) Pendant la panne

FIG. 1.12 – Acheminement des paquets dans un réseau protégé par tunnels de secours

que cette diminution du nombre de LSP se fait au détriment d'une agrégation des flux à l'intérieur du tunnel (même QoS pour tous les flux traversant le tunnel de secours), d'un délai d'acheminement des paquets souvent plus élevé et d'une consommation de bande passante supérieure (voir l'exemple de la figure 1.12 où les paquets du LSP p_F sont transmis sur le tunnel de secours du LSR H au LSR D avant que ce dernier ne les renvoie au LSR H en utilisant la partie du LSP primaire non affectée).

1.5 Extensions de RSVP pour la protection locale

Dans la section 1.3.3, nous avons vu que RSVP-TE permet d'établir des LSP primaires avec ou sans réservation de ressources. Afin de gérer la protection locale sous MPLS, [PSA05] a étendu le protocole RSVP en y ajoutant et définissant de nouveaux objets et drapeaux (*flags*). Pour saisir les mécanismes permettant aux nœuds d'effectuer les tâches de contrôle d'admission et afin de comprendre les différentes extensions de protocoles proposées dans cette thèse, nous présentons brièvement, dans ce qui suit, les objets et drapeaux utilisés pour fournir la protection, avant de nous intéresser aux comportements et tâches des différents LSR, à l'identification et à la signalisation des LSP de secours, à la fusion des LSP, etc.

Pour comprendre

1.5.1 Nouveaux objets et drapeaux RSVP pour la protection locale

Deux nouveaux objets RSVP-TE (FAST_REROUTE et DETOUR) transportés dans les messages *path* ont été définis et deux autres objets (SESSION_ATTRIBUTE et RECORD_ROUTE) ont été étendus pour permettre l'établissement et la configuration des LSP de secours sous RSVP-TE.

1.5.1.1 Drapeaux de l'objet SESSION_ATTRIBUTE

L'objet SESSION_ATTRIBUTE définit deux nouveaux drapeaux pour permettre de demander explicitement la protection contre la panne d'un nœud et pour solliciter la bande passante. Ces drapeaux sont :

Bandwidth protection desired Ce drapeau indique aux PLR du LSP primaire protégé qu'une bande passante garantie est désirée pour les LSP de secours. La quantité de bande passante sollicitée est celle du LSP protégé si le message *path* ne contient pas d'objet FAST_REROUTE. Dans le cas contraire, cette quantité est déterminée par l'objet FAST_REROUTE (cf. section 1.5.1.3).

Node protection desired Ce drapeau indique aux différents PLR du LSP primaire protégé (excepté pour le PLR en aval du LSR de sortie) que les chemins de secours doivent contourner le prochain LSR sur le LSP primaire protégé.

Deux autres drapeaux, définis dans [ABG+01], sont utilisés par la protection locale :

Local protection desired Ce drapeau est positionné pour demander aux LSR de fournir la protection locale.

Label recording desired Ce drapeau sollicite l'insertion des étiquettes dans l'objet RECORD_ROUTE. Il peut être utile pour éviter des signalisations supplémentaires de LSP (nécessaires à la découverte des étiquettes en aval du MP) lors de l'utilisation de la protection par tunnels de secours.

1.5.1.2 Drapeaux de l'objet RECORD_ROUTE

Pour indiquer si la protection contre la panne d'un nœud et/ou la bande passante est garantie, deux drapeaux ont été nouvellement définis. Les différents drapeaux de l'objet RECORD_ROUTE qui sont utiles à la protection locale sont :

Local protection available Indique que le prochain lien (du LSR inclus dans cet objet) est protégé par un mécanisme de protection locale.

Local protection in use Indique que le mécanisme de protection local est en cours d'utilisation.

Bandwidth protection Le PLR positionne ce drapeau pour indiquer que son LSP de secours dispose d'une bande passante suffisante. Le positionnement de ce drapeau n'est obligatoire que si le drapeau *Bandwidth protection desired* de l'objet SESSION_ATTRIBUTE est positionné et que la bande passante est garantie.

Node protection Le PLR positionne ce drapeau pour indiquer que son LSP de secours contourne son prochain routeur sur le LSP protégé. Le positionnement de ce drapeau n'est obligatoire que si le drapeau *Node protection desired* de l'objet SESSION_ATTRIBUTE est positionné et que le chemin de secours offre une protection contre la panne du prochain routeur au PLR. Dans le cas où seule la protection contre le lien en aval (du PLR) est fournie (pas de protection contre la panne du prochain nœud), le drapeau *Local protection available* est positionné mais le drapeau *Node protection* ne l'est pas.

1.5.1.3 FAST_REROUTE

Cet objet est transporté dans les messages *path* établissant un LSP primaire protégé. Il est utilisé pour le contrôle des LSP de secours d'un LSP protégé. Concrètement, il permet de spécifier les priorités des LSP de secours à préempter ou à restituer les ressources, de spécifier les filtres (i.e. les liens à inclure ou à exclure dans les LSP de secours) et la bande passante à utiliser par les LSP de secours.

Hormis les champs permettant d'identifier l'objet et de déterminer sa longueur, l'objet FAST_REROUTE contient :

Setup Prio
: C'est la priorité du LSP de secours à réquisitionner des ressources (entre 0 et 7). Cette priorité est utilisée pour déterminer si cette session peut préempter une ressource détenue par une autre session de moindre priorité, lors de son établissement.

Hold Prio
: C'est la priorité du LSP de secours à conserver ses ressources (entre 0 et 7). Cette priorité est utilisée pour déterminer si cette session peut être préemptée par une autre session plus prioritaire.

Hop-limit
: C'est le nombre maximum de nœuds (intermédiaires) constituant le LSP de secours. Le PLR et le MP ne sont pas comptés.

Drapeaux
: Ils permettent de spécifier le type de protection locale : le drapeau *One-to-one backup desired* indique que la protection locale par LSP de détour est réclamée alors que le drapeau *Facility backup desired* est utilisé pour demander une protection locale par tunnel de secours.

Bandwidth
: Il permet d'estimer la quantité de bande passante nécessaire pour les LSP de secours.

Exclude-any, Include-any et Include-all
: Ce sont des filtres permettant de spécifier les liens à exclure ou à inclure lors de la protection du LSP primaire.

1.5.1.4 L'objet DETOUR

Cet objet est transporté dans les messages *path* établissant un LSP de détour. Il permet d'identifier les LSP de détour pour offrir une meilleure flexibilité lors de leur fusion. Le contenu de cet objet est déterminé par l'objet RECORD_ROUTE (du message *resv* reçu par le LSR de tête) ou par l'objet EXPLICIT_ROUTE (du message *path*) dans le cas où tous les nœuds sont stricts.

Hormis les champs permettant d'identifier l'objet et de déterminer sa longueur, l'objet DETOUR énumère tous les PLR ainsi que leur prochains routeurs sur le LSP primaire à protéger.

Dans les sous-sections suivantes, nous allons nous intéresser aux différentes tâches effectuées par les routeurs RSVP afin d'assurer et rendre efficace la fonction de protection.

1.5.2 Rôles des LSR

Avec RSVP, différents LSR interviennent dans la construction, le contrôle, la signalisation et la fusion des LSP de secours. Trois rôles sont à distinguer : le routeur de tête du LSP primaire à protéger s'occupe essentiellement de la tâche de contrôle en spécifiant les différents paramètres nécessaires à la configuration des LSP de secours. Le PLR est responsable de la construction, la signalisation des différents LSP de secours (LSP de détour ou tunnel de secours). Enfin, le MP a pour rôle de fusionner les différents LSP de secours entre eux et/ou avec le LSP primaire protégé.

1.5.2.1 Rôle du LSR de tête du LSP protégé

C'est le LSR de tête qui construit les objets SESSION_ATTRIBUTE et FAST_REROUTE. Par conséquent, c'est ce LSR qui réclame la protection du LSP primaire en positionnant le drapeau *Local protection desired*. C'est lui aussi qui spécifie la bande passante dédiée aux LSP de secours, la technique de protection locale à employer (protection par LSP de détour ou protection par tunnels de secours), les liens à inclure ou à exclure lors des calculs des LSP de secours, etc. (cf. sections 1.5.1.1 et 1.5.1.2). Par exemple, si le LSR d'entrée désire une protection locale en utilisant des LSP de détour, il devrait inclure l'objet FAST_REROUTE et positionner le drapeau *One-to-one backup desired*. Le LSR de tête d'un LSP primaire protégé doit aussi positionner le drapeau *Label recording desired* de l'objet SESSION_ATTRIBUTE. Cela permet de découvrir l'étiquette en amont du MP sur le LSP primaire, ce qui facilite l'utilisation de la protection par tunnels de secours.

1.5.2.2 Comportement du PLR

Le PLR détermine le type de technique de protection locale (protection par LSP de détour ou protection par tunnels de secours) à utiliser lorsque l'objet FAST_REROUTE est inclus dans le message *path*. Dans le cas contraire et si le drapeau *Local protection desired* de l'objet SESSION_ATTRIBUTE est positionné, le PLR peut utiliser l'une ou l'autre des deux techniques de protection locale. Le PLR doit aussi fournir une protection contre la panne du prochain nœud si le drapeau *Node protection desired* de l'objet SESSION_ATTRIBUTE est positionné. Si cela n'est pas possible, le PLR essayera de protéger uniquement contre la panne du prochain lien. Si le drapeau *Bandwidth protection desired* est positionné, le PLR devra réclamer une bande passante suffisante sur tous les liens du LSP de secours protégeant le LSP primaire.

Le PLR s'occupe aussi de la suppression des LSP de secours dont il est le routeur de tête ainsi que de la propagation des messages d'erreurs reçus sur le LSP primaire protégé.

1.5.2.3 Comportement du MP

Un MP est un LSR qui reçoit un message *path* pour le LSP primaire protégé et au moins un autre message *path* pour un LSP de secours. Tout LSR du LSP protégé doit supposer qu'il est un MP puisque la signalisation des LSP de secours avant la panne n'est pas obligatoire dans le cas de la protection par tunnels de secours.

Un MP a pour rôle de fusionner les messages *path* des LSP de secours entre eux et/ou avec le message *path* du LSP primaire protégé lorsque cela est possible. Pour ce faire, il consulte les objets contenus dans les messages *path* reçus afin d'identifier les LSP correspondants et leur chemins (cf. section 1.5.3), puis détermine si les conditions de fusion sont satisfaites. Si c'est le cas, le MP sélectionne un message (le message *path* du LSP primaire protégé étant prioritaire) parmi tous les messages *path* à fusionner, le met-à-jour et l'envoie ensuite au LSR en aval (voir section 1.5.7 pour plus de détails).

Bien évidemment, le MP doit aussi gérer la libération des ressources sur les différents LSP de secours dont il MP. C'est lui aussi qui détermine et décide de propager ou pas, sur les LSP de secours, les messages d'erreurs liés à la réservation de ressources sur le LSP primaire protégé.

Après cette brève description des rôles des LSR dans la protection, nous expliquons ci-après comment assurer ces rôles.

1.5.3 Identification des LSP

Il existe deux approches pour identifier d'une manière unique un LSP de protection : identification par la source et identification par chemin.

1.5.4 identification par la source (*Sender Template-Specific*)

Dans cette approche, l'objet SESSION et l'identifiant du LSP de l'objet SENDER_TEMPLATE seront copiés du LSP primaire protégé. Seule l'adresse source, présente dans l'objet SENDER_TEMPLATE, sera différente de celle correspondant au LSP protégé. Cette dernière contiendra une adresse du PLR différente. De cette manière, l'association entre le LSP primaire protégé et ses LSP de secours se fera grâce aux deux objets SESSION et SENDER_TEMPLATE qui ne diffèrent que dans l'adresse source contenue dans l'objet SENDER_TEMPLATE.

Dans le cas d'un LSR de tête qui est aussi un PLR, ce dernier doit utiliser une adresse source différente de celle utilisée dans l'objet SENDER_TEMPLATE du LSP primaire protégé.

1.5.5 identification par chemin (*Path-Specific*)

Avec cette approche, tous les LSP de secours et leur LSP primaire protégé utilisent les mêmes objets SESSION et SENDER_TEMPLATE. Afin de différencier les messages des LSP de secours de ceux correspondant à leur LSP primaire protégé, l'approche *Path-Specific* se base sur la présence ou l'absence de l'objet DETOUR dans les messages *path* associés. Concrètement, les messages du LSP primaire sont identifiés par l'absence du

l'objet DETOUR : pour indiquer que le LSP primaire requiert une protection, l'objet FAST_REROUTE est inséré et/ou le drapeau *Local protection desired* de l'objet SESSION_ATTRIBUTE est positionné dans le message *path*. A l'opposé, pour identifier un message *path* correspondant à un LSP de secours, un objet DETOUR est construit et est inséré dans le message.

Contrairement à la première technique qui permet de distinguer le LSP primaire protégé de tous ses LSP de secours, la seconde technique ne renseigne que le type du LSP (LSP primaire ou LSP de secours). Cela a une incidence sur la flexibilité du choix des LSP de secours puisque la fusion de ces derniers est parfois imposée. Typiquement, tous les LSP de secours (protégeant un même LSP primaire) ayant la même interface de sortie et le même prochain nœud doivent être fusionnés (cf. section 1.5.7.2). Cela réduit le nombre d'états RSVP au détriment d'une fusion de LSP ayant des objets ERO différents et d'une consommation de bande passante potentiellement supérieure.

1.5.6 Détermination des étiquettes en aval du MP

Comme nous l'avons déjà expliqué en section 1.4.4.4, le PLR doit connaître les étiquettes permettant au MP d'associer les différents flux reçus d'un tunnel de secours vers leurs LSP primaires protégés. Concrètement, le PLR doit insérer dans chaque paquet, envoyé sur le tunnel de secours, une première étiquette identifiant le LSP primaire protégé (associé au paquet) sur le MP, avant d'empiler une deuxième étiquette, placée au sommet de la pile MPLS et servant à l'acheminement du paquet sur le tunnel de secours.

Selon que l'espace d'allocation d'étiquettes est global ou pas au MP, la découverte de la première étiquette insérée par le PLR dans les paquets redirigés vers un tunnel de secours peut s'effectuer avec ou sans signalisation.

1.5.6.1 Approche sans signalisation

Cette approche (sans messages *path*/*resv*) n'est applicable que si l'espace d'attribution d'étiquettes sur le MP est global. Dans ce cas, le PLR utilisera la même étiquette transmise par le MP à son routeur en aval sur le LSP primaire protégé (cf. figure 1.12). Cette étiquette sera découverte en consultant l'objet RRO lors de la configuration du LSP primaire protégé. En effet, cet objet contient (obligatoirement) l'étiquette ainsi que le type d'allocation d'étiquettes utilisé par le MP.

1.5.6.2 Approche avec signalisation

Lorsque le MP utilise un espace d'étiquettes local à chacune de ses interfaces, une signalisation supplémentaire sera nécessaire afin de déterminer les étiquettes permettant de séparer, sur le MP, les flux transportés sur un tunnel de secours. Concrètement, avant de basculer le trafic sur un tunnel de secours, le PLR doit demander au MP une étiquette pour chaque LSP protégé par ce tunnel de secours. Ceci se fait en envoyant un message *path* au MP pour chaque LSP protégé. Ce message sera identifié avec l'approche *Sender Template-Specific* et contiendra un objet ERO commençant par l'adresse du

MP[9]. Lorsque le MP reçoit un tel message *path*, il alloue une nouvelle étiquette (comme pour les LSP de détour) qu'il transmet au PLR dans un message de réponse *resv*.

1.5.7 Fusion des LSP de secours

La fusion permet de réduire le nombre d'étiquettes allouées tout en diminuant le nombre d'états RSVP à gérer et à maintenir. Cette fusion s'effectue souvent au détriment d'un routage moins flexible et d'une consommation de la bande passante qui peut être supérieure (cas de fusion de LSP sur un LSR dont les messages *path* contiennent des objets ERO différents).

A cause de l'existence de deux approches distinctes pour l'identification des LSP de secours, les règles de fusion des LSP de secours diffèrent.

1.5.7.1 Fusion de LSP identifiés par la source

Avec cette approche, tous les messages *path* des LSP de secours et de leur LSP primaire protégé sont éligibles à la fusion dès qu'ils ont la même interface de sortie ainsi que le même prochain nœud. Cependant et pour être conforme avec, seuls les messages *path* reçus par un LSR avec des objets ERO équivalents (i.e. même route vers le LSR de sortie) peuvent être réellement fusionnés.

Lors de la fusion d'un ensemble de LSP de détour sur un LSR, un seul message *path* sera envoyé au prochain nœud pour l'ensemble des LSP fusionnés. Ce message doit correspondre à celui du LSP protégé si un des LSP fusionnés est le LSP primaire protégé ; sinon, un des messages *path* des LSP fusionnés est choisi pour être transmis au prochain LSR sur la route explicite. Bien évidemment, l'objet DETOUR doit être modifié pour inclure tous les LSR que doivent contourner les LSP de secours fusionnés.

1.5.7.2 Fusion de LSP identifiés par chemin

La fusion dans ce cas est plus stricte et est obligatoire dès lors que deux messages *path*, ayant les mêmes objets SESSION et SENDER_TEMPLATE, ont la même interface de sortie et le même prochain nœud. Dans ce cas, le MP exécute les étapes suivantes afin de créer le message path final (fusionné) à transmettre au prochain nœud :

- Si un ou plusieurs messages *path* correspondent à un LSP primaire protégé (message avec un objet FAST_REROUTE ou message sans objet DETOUR), alors un de ces messages est sélectionné pour être transmis au prochain nœud.
- Sinon, éliminer de l'ensemble des messages *path* des LSP de secours fusionnés ceux traversant des nœuds que d'autres veulent contourner. S'il reste des messages *path* qui ne sont pas éliminés, alors un message est choisie suivant une politique locale. Sinon, le MP devrait essayer de déterminer une nouvelle route évitant tous les nœuds que les différents LSP fusionnés doivent contourner. Un nouvel objet DETOUR et un objet ERO seront déterminés afin de construire le message *path* final.

[9] L'objet ERO de ce message est déduit du même objet du message *path* du LSP protégé en supprimant tous les nœuds précédents le MP et en les remplaçant par l'adresse du MP.

Lorsque les messages *path* ne contiennent ni d'objet FAST_REROUTE, ni d'objet DETOUR, la fusion n'est pas requise et les messages devraient être traités suivant le RFC [ABG+01].

1.5.8 Traitement des pannes

Lorsqu'une panne est détectée par un LSR, ce dernier met-à-jour sa table de commutation LFIB[10] pour rerouter les LSP primaires protégés et affectés sur leurs LSP de secours. De ce fait, l'étiquette et l'interface de sortie de chaque LSP protégé et affecté par la panne est écrasée par l'étiquette (resp. les étiquettes) et l'interface de sortie correspondant au LSP de détour (resp. au tunnel de secours) permettant la restauration.

Lorsque l'étiquette en aval du MP sur le LSP de secours est découverte sans signalisation, le MP est contraint de conserver l'étiquette associée au LSP primaire protégé jusqu'à ce que le LSP de secours soit signalisé. Pour ce faire, il réinitialise les *temporisateurs de rafraîchissement* associés aux messages *path* et *resv* du LSP primaire protégé. Cela permet de conserver ces états jusqu'au prochain déclenchement du temporisateur de rafraîchissement, c'est-à-dire jusqu'à ce que son LSP de secours soit signalisé.

1.5.9 Ré-optimisation des LSP

RSVP permet la ré-optimisation des LSP primaires et des LSP de secours en utilisant la procédure *make before break*. Cette dernière permet de configurer un nouveau LSP optimisé qui remplacera un ancien LSP sans perte de messages et sans compter deux fois la bande passante sur les liens communs aux deux LSP. Pour ce faire, le LSR de tête copiera les objets SESSION et SENDER_TEMPLATE des messages associés à l'ancien LSP avant de changer l'adresse source contenue dans l'objet SENDER_TEMPLATE par une adresse locale et différente.

1.6 Conclusion

Dans ce chapitre, nous nous sommes concentrés sur la description de la technologie MPLS ainsi que sur les fonctionnalités principales qu'elle permet de rendre comme l'ingénierie de trafic, la séparation du trafic en plusieurs classes et la protection.

Nous avons ainsi vu que MPLS est capable de partitioner le trafic en plusieurs flux, chacun est identifié par une FEC. A chaque FEC est associé un chemin unidirectionnel constitué d'une séquence fixe de routeurs MPLS (LSR), appelé LSP. Ce dernier est utilisé pour acheminer le flux associé à sa FEC, le long des LSR qui le forment, en se basant sur le paradigme de commutation d'étiquettes. Concrètement, après l'insertion d'une étiquette MPLS dans l'entête d'un paquet (à l'entrée d'un domaine MPLS), l'acheminement se fait sur un circuit virtuel complètement identifié, sur chaque LSR, par l'étiquette transmise dans le paquet. Une table de commutation d'étiquettes LFIB est

[10] Lorsque le LSR est un routeur de tête d'un LSP primaire protégé et affecté, sa table IP est aussi mise-à-jour.

utilisée sur chaque LSR afin de déterminer la prochaine interface de sortie et l'étiquette à substituer à celle transmise dans le paquet reçu.

Bien évidemment, la flexibilité dans le choix de la séquence de LSR formant un LSP et la faculté de MPLS à partitioner le trafic en plusieurs flux facilite l'implantation des mécanismes d'ingénierie de trafic (TE) et permet de rendre la qualité de service désirée par chaque flux. Concrètement, tout flux est identifié par une FEC qui permet son acheminement sur un LSP établi explicitement afin de vérifier les contraintes de qualité de service (délai, bande passante, etc.) associées au flux et pour optimiser les ressources du réseau.

Grâce à la hiérarchie de LSP, MPLS peut aussi agréger et acheminer les flux de plusieurs LSP dans un seul tunnel. Pour ce faire, il suffit d'insérer, après la première étiquette utilisée pour séparer les flux à la sortie du tunnel, une deuxième étiquette permettant l'aiguillage des paquets le long du tunnel. Cette hiérarchie MPLS a deux principaux avantages : (1) elle facilite l'implantation et le déploiement des réseaux VPN et (2) elle permet une meilleure résistance au passage à l'échelle grâce à la réduction du nombre de LSP traversant un LSR.

Pour doter MPLS de fonctions permettant de calculer, configurer et réserver dynamiquement les ressources, différents protocoles de routage et de signalisation ont été développés. Dans ce chapitre, nous nous sommes focalisés sur la description des protocoles IGP-TE (OSPF-TE et ISIS-TE) et RSVP-TE qui sont les plus en vogue pour rendre la fonctionnalité d'ingénierie de trafic. Nous avons ainsi vu que les protocoles IGP-TE ont pour rôle de collecter et distribuer l'information liée à la topologie du réseau et et aux paramètres d'ingénierie de trafic. Cette information est exploitée par le module de calcul des chemins qui permet de déterminer les routes explicites vérifiant les contraintes de ressources et les contraintes TE. Ensuite, le protocole RSVP-TE intervient pour configurer ces différentes routes tout en réservant ou pas les ressources.

En plus des contraintes liées aux ressources et à l'ingénierie de trafic, les communications d'aujourd'hui doivent souvent être sûres et résistantes aux pannes. Pour ce faire, différentes techniques de protection ont été développées et décrites dans ce chapitre. Nous avons ainsi évoqué brièvement les techniques de protection globale (sous MPLS) qui ne peuvent pas assurer, dans les réseaux larges, des délais de récupération acceptables pour certains types d'applications temps réel comme la voix sur IP. Puis, nous nous sommes intéressés de près à la protection locale qui réduit considérablement les délais de récupération par la suppression de la phase de notification de la panne. Ainsi, avec ce type de protection, c'est le nœud (PLR) détectant la panne qui réagit en basculant localement et rapidement le trafic des communications affectées vers leurs chemins de secours.

Deux méthodes de protection locale sous MPLS ont été largement expliquées dans ce chapitre : protection par LSP de détour et protection par tunnels de secours.

Avec la protection par LSP de détour, un chemin de secours séparé est calculé et est configuré sur chaque LSR du LSP primaire protégé (excepté le LSR de sortie). Par sa grande flexibilité dans le choix des LSP de secours, cette méthode de protection peut réduire la consommation de bande et offrir une meilleure qualité de service lors d'une panne. Par contre, elle augmente le nombre de LSP dans le réseau, ce qui pourrait poser

des problèmes lors du passage à l'échelle.

Avec la protection par tunnels, le nombre de chemins de secours traversant un LSR est sensiblement diminué grâce au partage d'un même chemin de secours pour la protection de plusieurs chemins primaires. Cette méthode de protection utilise le hiérarchie MPLS pour agréger le trafic de plusieurs LSP primaires vers un seul tunnel de secours lors d'une panne. Cela a l'avantage de réduire les traitements sur les LSR intermédiaires du tunnel au détriment de l'utilisation de la même qualité de service pour tous les LSP primaires protégés par un même tunnel (lors d'une panne) et d'une consommation de bande qui peut être supérieure.

Enfin, pour mettre en œuvre la protection locale sous MPLS, nous avons décrit quelques extensions apportées au protocole de signalisation RSVP. Comme nous l'avons vu, ces extensions emploient de nouveaux objets et drapeaux qui permettent d'identifier les LSP de secours et les associer à leur LSP primaire protégé, d'indiquer la méthode de protection utilisée, de faciliter la gestion de la fusion, de déterminer la quantité de ressources à allouer au LSP de secours, etc.

Dans la suite de cette thèse, nous allons retrouver tous les outils et protocoles décrits dans ce chapitre. Typiquement, nous utiliserons la protection locale afin de diminuer les délais de récupération et nous adopterons un environnement MPLS (MPLS, IGP-TE et RSVP-TE) pour la flexibilité de routage et l'optimisation des ressources. De plus, pour faciliter le déploiement des mécanismes de placement distribué des LSP de secours, nous tâcherons d'éviter autant que possible de définir de nouveaux protocoles ou de modifier excessivement ceux existants. Pour ce faire, nous emploierons et exploiterons au mieux l'information transmise par les protocoles IGP-TE et RSVP-TE, que nous étendrons le plus légèrement possible lorsque cela est nécessaire.

Chapitre 2

Partage de ressources et placement local de LSP de secours unicast

2.1 Introduction

La technologie MPLS (cf. chapitre 1) permet une récupération rapide des pannes, la durée de cette récupération est de l'ordre de quelques millisecondes (50 à 100 ms). Cela est rendu possible grâce à la pré-configuration de LSP de secours permettant de contourner localement et rapidement les différentes pannes affectant les LSP primaires protégés.

Pour assurer la disponibilité des ressources après une panne, les LSP de secours pourraient réserver leurs ressources à l'avance. Cependant et contrairement aux LSP primaires, les LSP de secours n'utilisent effectivement leurs ressources que lors de la panne du risque (nœud, lien ou SRLG) protégé. Étant donné qu'une récupération d'une panne n'induit pas l'activation de tous les LSP de secours, il est alors envisageable et est souhaitable de partager les ressources entre les LSP pour optimiser l'utilisation des ressources. En effet, sous l'hypothèse pratique d'une seule panne à la fois (probabilité très réduite qu'une deuxième panne survienne avant que la première panne soit complètement traitée), deux LSP de secours protégeant contre les pannes de deux composants différents ne peuvent réclamer leurs ressources en même temps puisque les deux LSP de secours ne peuvent pas être actifs aux mêmes moments. En conséquence, au lieu d'assigner de nouvelles ressources non attribuées à un LSP de secours en cours d'établissement, ce dernier pourra utiliser des ressources déjà détenues par un autre LSP de secours qui protège contre la panne d'un composant différent. Cette technique de réservation multiple des mêmes ressources par des LSP de secours différents est dite *partage de ressources de protection* : elle permet d'optimiser l'utilisation des ressources dans le réseau et elle n'est applicable qu'entre des LSP de secours qui ne pourront pas être actifs en même temps.

Bien que le partage de ressources soit une politique locale aux LSR du réseau et non pas aux protocoles, certains protocoles l'imposent (dans certains cas de figure) et le prennent en compte. Avec RSVP-TE par exemple, le partage de ressources (étiquettes

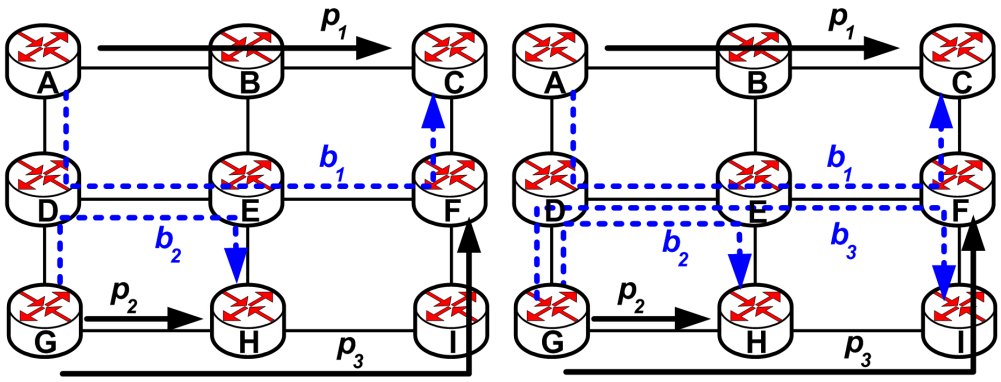
MPLS et bande passante) est implicitement effectué entre les différents LSP de détour protégeant le même LSP primaire grâce à la fusion (cf. chapitre 1).

Pour optimiser davantage l'utilisation des ressources tout en garantissant leur disponibilité après une panne, il serait nécessaire et judicieux d'étendre le partage de ressources aux LSP de secours protégeant différents LSP primaires. Dans ce chapitre, nous nous concentrons sur le problème de partage de ressources, particulièrement de la bande passante, entre LSP de secours protégeant différents LSP primaires. Nous verrons d'abord comment déterminer l'ensemble des ressources allouées pour les LSP de secours sur un lien donné, puis comment établir en ligne les LSP de secours de manière à partager et à optimiser l'ensemble des ressources allouées. Dans le cas d'un serveur centralisé, ce dernier a connaissance de la topologie du réseau et de tous les LSP déjà construits. Par conséquent, le problème peut être ramené à une optimisation combinatoire. Dans le cas distribué par contre, les nœuds risquent de ne pas disposer de toute l'information concernant les LSP construits (car la distribution de cette information est très coûteuse), déterminer donc une heuristique permettant d'approcher la meilleure solution est primordial.

La suite de ce chapitre est organisée comme suit. Après l'introduction des notations utilisées dans la suite de cette thèse, nous décrivons et expliquons en section 2.3 le principe de partage des ressources, en mettant l'accent sur les différents types de ressources. En section 2.4, nous définissons quelques concepts et outils permettant le calcul de la quantité de ressources partagées entre un groupe de LSP de secours. Dans la section suivante, nous modéliserons mathématiquement le problème de placement des LSP de secours réduisant la quantité de bande allouée en utilisant le partage de ressources. Dans la section 2.6, nous expliquons comment implanter les solutions au problème de placement de LSP de secours dans des environnements centralisé et distribué. Dans la section 2.7 (resp. section 2.8), nous citons différents algorithmes (resp. heuristiques) en ligne et distribués permettant de déterminer des solutions exactes ou approchées au problème de placement de LSP de secours. Les performances de chaque algorithme et/ou heuristique en termes de taux de partage et/ou de quantité d'informations circulant dans le réseau sont étudiées. Dans la section qui suit, nous nous intéressons à l'impact du choix du mécanisme de différenciation des types de panne sur la quantité de bande passante allouée et sur le délai de récupération. Dans la section 2.10, nous décrirons d'autres types de partage de ressources (partage entre LSP primaires et LSP de secours). La section 2.11 sera consacrée à la conclusion.

2.2 Notation

Dans la suite de cette thèse, nous modélisons le réseau par un graphe $G = (V, E, Rs)$ constitué de n nœuds ($n = |V|$), m arêtes ($m = |E|$) et d'un ensemble de risques de panne Rs. Toute arête α-β du graphe supporte deux arcs de sens opposés $\alpha \rightarrow \beta$ et $\beta \rightarrow \alpha$ et de capacités positives correspondant à $C^{\alpha \rightarrow \beta}$ et $C^{\beta \rightarrow \alpha}$ (une arête dont la capacité est nulle pour un sens donné est considérée comme unidirectionnelle). Par abus de langage, nous utilisons l'expression $\alpha \rightarrow \beta \in E$ pour indiquer que l'arête α-β est dans l'ensemble



(a) Partage total des resources (b) Partage partiel des resources

FIG. 2.1 – Partage de resources

E.

L'ensemble des risques de panne Rs contient un ensemble de SRLG (les SRLG sont formés de sous-ensembles d'arêtes du graphe), en plus de l'ensemble $V \cup E$.

2.3 Partage de ressources

Lorsque deux LSP de secours protègent contre les pannes de composants différents et ne risquant jamais de tomber en panne simultanément, ils pourront partager leur réservation de ressources. En effet, les deux LSP de secours n'utilisent effectivement leurs ressources que lors d'une panne affectant le composant protégé. Comme il est infiniment improbable que les composants protégés par les deux LSP de secours tombent en panne simultanément, au plus un des deux LSP de secours peut être actif à tout moment. En conséquence, chacun des deux LSP de secours pourra utiliser et donc partager ses ressources avec l'autre LSP de secours sur toutes les parties de chemins qui leur sont communes.

Sur la figure 2.1 (a), deux LSP primaires p_1 ($A \rightarrow B \rightarrow C$) et p_2 ($G \rightarrow H$) ont été établis. Afin de protéger le LSP primaire p_1 contre la panne du nœud B (et du lien A-B), le LSP de secours b_1 est configuré. De même pour protéger le LSP primaire p_2 contre la panne du lien G-H, le LSP de secours b_2 est déterminé et est configuré.

Sous l'hypothèse de pannes simples où les composants (nœud B, lien A-B et lien G-H) ne tombent jamais en panne simultanément, le LSP de secours b_1 ne peut pas être actif au même moment que le LSP de secours b_2. Par conséquent, les deux LSP de secours b_1 et b_2 pourront utiliser et partager les mêmes ressources sur tous les liens en communs.

Sur la figure 2.2, les ressources du lien D-E ainsi que les ressources allouées aux LSP de secours b_1 et b_2 sont représentées. Les ressources réservées sont représentées par des carrés colorés (pleins) alors que les ressources libres sont représentées par des carrés blancs. Ainsi, le lien D-E commun à b_1 et b_2 dispose d'un pool de 9 ressources de

type T_1 et d'un pool de 8 ressources de types T_2. Bien que b_1 ait réservé 5 ressources de chaque type, b_2 disposera de toutes les ressources (i.e. 9 ressources de type T_1 et 8 ressources de types T_2) du lien *D-E* lors de son placement. En effet, les ressources réservées par b_1 peuvent être partagées et affectées à b_2 vu que b_1 ne peut pas les réclamer en même temps que b_2 (un LSP de secours au plus, parmi b_1 et b_2, peut être actif à tout moment). Pour optimiser la disponibilité des ressources du réseau, le LSP de secours b_2 doit commencer par réserver les ressources à partir du pool de ressources allouées à b_1. Concrètement, si b_2 nécessite 6 ressources de type T_1 et 4 ressources de type T_2, alors b_2 ne réservera qu'une seule ressource additionnelle de type T_1 (la ressource numéro 6 de type T_1 sur la figure 2.2) à partir du pool de ressources libres du lien *D-E*. Les autres ressources (5 ressources de type T_1 et 4 ressources de type T_2) seront réservées sur les ressources déjà allouées à b_1. En conséquence, les ressources 1, 2, 3, 4 et 5 de type T_1 et les ressources 1, 2, 3, et 4 de type T_2 sont donc partagées entre les deux LSP de secours b_1 et b_2.

Si l'on construit maintenant un nouveau LSP primaire p_3 et un LSP de secours b_3 permettant de le protéger contre les pannes du nœud *H* et du lien *G-H* (figure 2.1 (b)), le partage des ressources sera possible mais restreint aux ressources de b_1 non utilisées par b_2. En effet, b_2 et b_3 protègent les LSP primaires (p_2 et p_3) contre la panne d'un même composant (lien *G-H*) et ne pourront donc pas partager leurs ressources. Par contre, les ressources allouées à b_1 peuvent être ré-utilisées par b_2 vu que les deux LSP de secours protègent contre des risques de panne différents (les pannes des composants qui provoqueront l'activation de b_1 et b_2 ne se présentent jamais au même moment). De ce fait, on conclut que b_3 ne peut utiliser que les ressources réservées à b_1 et qui ne sont pas utilisées par b_2. Ainsi, si b_3 nécessite 2 ressources de chaque type, il réservera 2 ressources de type T_1 et 1 ressource de type T_2 à partir du pool des ressources libres du lien *D-E*. En effet, seule 1 ressource de type T_2 réservée par b_1 n'est pas partagée avec b_2 et pourra donc être attribuée et partagée entre b_1 et b_3. En conséquence, 8 ressources de type T_1 et 6 ressources de type T_2 seront allouées aux trois LSP de secours sur le lien *D-E* grâce au partage (cf. figure 2.3). Notons bien que sans partage de ressources, b_2 n'aurait pas pu être configuré vu que le nombre de ressources de type T_1 disponibles (9) est inférieur au nombre de ressources réclamées par les trois LSP de

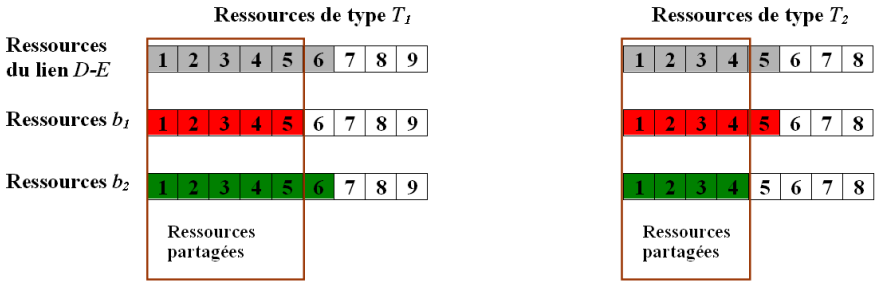


FIG. 2.2 – Ressources partagées entre deux LSP de secours sur un lien

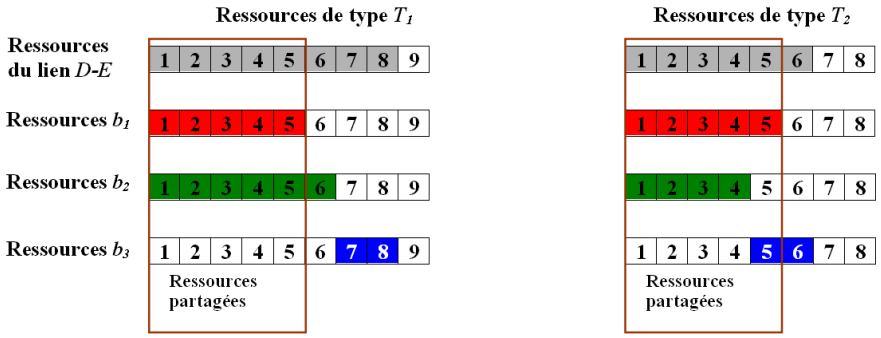


FIG. 2.3 – Ressources partagées entre trois LSP de secours sur un lien

secours (11).

Comme nous le constatons, une ressource n'est partagée sur un lien qu'entre des LSP de secours protégeant des composants ne risquant pas de tomber en panne simultanément. Ceci a pour but d'éviter que la ressource soit utilisée effectivement et simultanément par plus d'un LSP de secours actif. Implicitement, bien qu'une ressource peut être réservée par plusieurs LSP de secours, l'accès à cette ressource quant à lui est exclusif : un seul LSP de secours ayant déjà réservé la ressource peut l'utiliser effectivement (et donc un seul LSP ayant réservé la ressource peut être actif à tout moment). Ceci est le cas de la plupart des types de ressources réseau d'un lien comme la bande passante, les instants de transmission, etc.

D'autres types de ressources existent. On peut citer les niveaux de priorité, le taux d'erreur, le coût financier, les étiquettes MPLS, les états de signalisation, etc. Ces ressources peuvent être à accès exclusif ou non, partageables ou non, avec ou sans contraintes. Une ressource à accès exclusif est une ressource qui ne peut être accédée et utilisée effectivement que par un seul LSP actif (transmettant effectivement les paquets). Une ressource partageable est une ressource qui peut être accédée et utilisée par plusieurs LSP de secours en même temps. Enfin, une ressource avec contraintes est une ressource qui ne peut être utilisée que par des LSP vérifiant une certaine condition dictée par les contraintes de la ressource.

Ainsi, les niveaux de priorité par exemple peuvent être considérés comme des ressources partageables à accès non exclusif qui ne peuvent être attribuées qu'à un certain type de LSP de secours (dépendant du type du LSP par exemple). En d'autres termes, le partage de ressources dans ce cas ne tiendra pas compte des composants protégés mais uniquement de la contrainte d'attribution des priorités aux LSP de secours (i.e. tous les LSP de secours ayant des priorités égales ou compatibles partagent leur ressources). Une autre politique de gestion de ressources peut considérer chaque niveau de priorité comme un pool de ressources d'un type donné à accès exclusif. Dans ce cas, les LSP de secours ne pourront demander que les ressources de niveaux de priorité compatibles (par exemple, un LSP de secours de priorité 1 ne peut réclamer que des ressources de niveau 0 ou 1). Bien évidemment, les mêmes principes de partage que ceux appliqués à

l'exemple des figures 2.2 et 2.3 pourront être utilisés pour ce type de ressource.

D'autres types de ressource peuvent être partageables, à accès non exclusif et sans contraintes. Par exemple, le coût financier (les entrées dans les table IP et LFIB le sont aussi) peut être partagé à tout moment par tous les LSP de secours (indépendamment des composants protégés). En effet, le coût financier d'un LSP de secours sur un lien peut dépendre uniquement du nombre de LSP passant par le lien ou il peut être aussi constant (indépendant des composants protégés). Le taux d'erreur est une autre ressource qui est partageable, à accès non exclusif et qui ne dépend que du support du lien (indépendant des composants protégés).

Enfin, un dernier type de ressources à accès exclusif et non partageables (avec ou sans contraintes) existe. Ainsi, si l'on interdit la fusion des LSP, les étiquettes MPLS allouées sur un lien doivent être différentes pour chaque LSP de secours. En conséquence, ces ressources ne sont pas partageables (tant que la fusion de LSP de secours est interdite).

Dans ce chapitre, nous nous intéressons uniquement aux ressources partageables et à accès exclusif. C'est le seul type de ressource où le partage doit vérifier certaines conditions pour éviter les conflits d'accès aux ressources. En effet, l'accès à des ressources à accès non exclusif ne pose aucun problème puisque tous les LSP pourront utiliser effectivement la ressource en même temps. Idem pour les ressources non partageables où chaque LSP ne peut disposer que des ressources non réservées par les autres LSP. Par contre, l'introduction des contraintes pour le type de ressource traité ici ne change pas la nature du partage puisque ce sont les mêmes principes de partage qui seront appliqués aux LSP vérifiant les contraintes.

Pour simplifier et pour mieux illustrer les principes de partage, nous utilisons ici la bande passante comme ressource partageable et à accès exclusif. Cette ressource est prise en compte par les protocoles de signalisation comme RSVP-TE et CR-LDP et par les différents protocoles IGP-TE comme OSPF-TE ou ISIS-TE (cf. chapitre 1). Les mêmes principes de partage de ressources peuvent être appliqués à toutes les autres ressources du même type.

2.4 Concept de partage de la bande passante entre LSP de secours

La bande passante peut être partagée sur un arc appartenant à plusieurs LSP si ces derniers ne peuvent pas être actifs au même moment. Ceci est le cas des LSP de secours protégeant contre les pannes de risques *indépendants* (i.e. risques de panne ne pouvant pas être affectés par une même et unique panne).

Pour faciliter le placement en ligne des LSP de secours, un ensemble de risques de panne protégés (*Protection Failure Risk Group* ou PFRG, [LRC02]) est associé à chaque LSP. Cet ensemble est constitué de tous les risques dont la panne induit l'activation du LSP de secours associé (pour la récupération). Ainsi, tous les risques de panne (lien, nœud ou SRLG) contenant le lien protégé (resp. le lien et le nœud protégés) d'un LSP NHOP (resp. d'un LSP NNHOP) doivent appartenir à son ensemble PFRG. Formellement, le PFRG d'un LSP de secours b_1 de type NHOP (resp. d'un LSP de

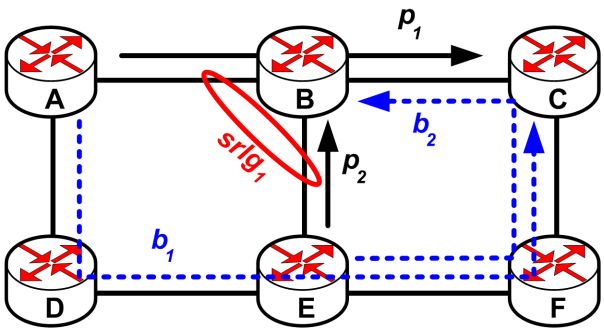


FIG. 2.4 – Protection locale et contournement des groupes de risques de panne protégés (PFRG)

secours b_2 de type NNHOP) protégeant contre la panne d'un lien λ (resp. d'un lien λ et d'un nœud n) est calculé comme suit :

$$PFRG\,(b_1) \;=\; \{r \setminus r \;\textit{contient}\; \lambda \wedge r \in Rs\}$$

$$PFRG\,(b_2) \;=\; \{r \setminus (r \;\textit{contient}\; \lambda \vee r = n) \wedge r \in Rs\}$$

Bien évidemment, un LSP de secours doit contourner tous les risques appartenant à son ensemble PFRG.

De la même manière que pour un LSP de secours, un PFRG d'un lien unidirectionnel (ou d'un arc) est défini comme l'ensemble des risques de panne protégés par les LSP de secours le traversant. Formellement, le PFRG d'un arc λ est calculé comme suit :

$$PFRG\,(\lambda) \;= \cup_{(b \setminus \lambda \in b)}\, PFRG\,(b)$$

Exemple : Sur la figure 2.4, deux LSP primaires p_1 ($A{\rightarrow}B{\rightarrow}C$) et p_2 ($E{\rightarrow}B$) sont établis. Afin de protéger le LSP primaires p_1 (resp. p_2) contre la panne du lien A-B et du nœud B (resp. contre la panne du lien E-B), un LSP de secours b_1 (resp. b_2) de type NNHOP (resp. NHOP) est calculé et est établi. A chacun des deux LSP b_1 ($A{\rightarrow}D{\rightarrow}E{\rightarrow}F{\rightarrow}C$) et b_2 ($E{\rightarrow}F{\rightarrow}C{\rightarrow}B$) est associé un groupe de risques de panne protégés consistant en :
$PFRG\,(b_1) \;=\; \{A\text{-}B,\; B,\; srlg_1\}$
$PFRG\,(b_2) \;=\; \{B\text{-}E,\; srlg_1\}$
Chaque arc traversé par au moins un LSP de secours a lui aussi un groupe de risques de panne protégés contenant au moins un élément. Sur l'exemple de la figure 2.4, nous déterminons les groupes de risques de panne protégés sur les liens $D{\rightarrow}E$, $E{\rightarrow}F$ et $F{\rightarrow}C$ comme suit :
$PFRG\,(D \rightarrow E) \;= PFRG\,(b_1) \;=\; \{A\text{-}B,\; B,\; srlg_1\}$
$PFRG\,(E \rightarrow F) \;= PFRG\,(b_1) \cup PFRG\,(b_2) \;=\; \{A\text{-}B, B\text{-}E,\; B,\; srlg_1\}$
$PFRG\,(F \rightarrow C) \;= PFRG\,(b_2) \;=\; \{B\text{-}E,\; srlg_1\}$

2.4.1 Conditions d'indépendance des LSP de secours

Les LSP de secours (b_1, b_2, .., b_n) sont dits *indépendants* si et seulement si leurs groupes de risques de panne protégés sont disjoints deux-à-deux (i.e. $\forall i, j \in [1-n]$: $i \neq j \Rightarrow PFRG(b_i) \cap PFRG(b_j) = \phi$). En conséquence, lors d'une panne simple, au plus un LSP de secours est actif (à la fois) parmi tous les LSP indépendants.

Si l'on définit $bw\ (b_i)$ comme la quantité de bande passante réclamée par le LSP de secours b_i, alors la bande passante $bw\ (\lambda)$ allouée sur un arc λ supportant uniquement des LSP de secours indépendants (b_1, b_2, .., b_n) est déterminée comme suit :
$bw(\lambda) = Max_{i=1}^{n}\, bw(b_i)$

Évidemment, une ressource allouée à un ensemble de LSP de secours indépendants peut toujours être réservée à un autre LSP de secours si ce dernier est indépendant avec tous les LSP de secours de l'ensemble.

2.4.2 Prix de protection d'un risque de panne sur un arc

Le *prix de protection* d'un risque r sur un arc λ, noté δ_r^λ, correspond à la quantité de la bande passante à allouer aux LSP de secours sur l'arc λ afin d'assurer la disponibilité de la bande passante suite à une panne du risque r. Il est déterminé comme la bande passante cumulée de tous les LSP de secours passant par l'arc λ et protégeant contre le risque de panne r. Formellement :

$$\delta_r^\lambda = \sum_{b \setminus \lambda \in b \wedge r \in PFRG(b)} bw(b) \tag{2.1}$$

Par convention, nous considérons que le coût de protection d'un risque r sur un arc λ est infini si :

- le risque r est un lien supportant l'arc λ,
- le risque r est un SRLG contenant un lien supportant l'arc λ,
- le risque r est un nœud adjacent à l'arc λ.

2.4.3 Coût de protection sur un arc

Un *coût de protection* sur un arc est toujours associé à un lien ou un nœud du réseau. Il correspond à la quantité de bande passante nécessaire sur un arc pour faire face à n'importe quelle panne affectant le lien ou le nœud associé au coût de protection. En conséquence, pour un lien u-v, le coût de protection sur un arc λ est déterminé comme le plus grand prix de protection de tous les risques de panne contenant le lien u-v. Pour un nœud u, le coût de protection sur n'importe quel arc correspond exactement au prix de protection du même nœud u sur le même arc. Concrètement :

$$\theta_c^\lambda = Max_{r \setminus c \in \{r\} \wedge r \in Rs}\, \delta_r^\lambda \tag{2.2}$$

Par extension, nous définissons le coût de protection $\theta^\lambda(b)$ d'un LSP de secours b sur un arc λ comme le prix de protection le plus élevé de tous les risques appartenant à $PFRG\ (b)$:

$$\theta^\lambda(b) = Max_{r \in PFRG(b)}\, \delta_r^\lambda \tag{2.3}$$

Comme nous le constatons, l'information sur les coûts de protection n'est qu'une information agrégée (compressée) sur les prix de protection.En fonction de l'algorithme de calcul distribué des LSP de secours utilisé, il pourrait être plus efficace de distribuer directement dans le réseau les coûts de protection au lieu des prix de protection.

2.4.4 Quantité minimale de bande passante de secours à allouer sur un arc

Pour assurer la disponibilité de la bande passante lors d'une panne, il est nécessaire de pré-allouer une certaine quantité de bande passante pour les LSP de secours. Cette quantité correspond à *la bande passante de secours*. Elle doit être supérieure ou égale à toutes les quantités de bande passante des LSP de secours qui peuvent être actifs en même temps. Cependant et afin d'optimiser la disponibilité des ressources dans le réseau, il est souhaitable que la quantité de bande passante de secours réservée, pour assurer le succès lors de la récupération de n'importe quelle panne, soit minimale. De ce fait, nous calculons la quantité de bande passante de secours (minimale) G^{λ} à allouer sur un arc λ pour faire face à n'importe quel risque de panne comme suit :

$$G^{\lambda} = Max_r\left(\delta^{\lambda}_{r \setminus r \in Rs}\right) = Max_c\left(\theta^{\lambda}_{c \setminus c \in (V \cup E)}\right) \tag{2.4}$$

Exemple : Afin de déterminer la quantité de bande passante de secours minimale à allouer sur l'arc $E \rightarrow F$ de la figure 2.4, il suffira de déterminer les prix de protection des risques appartenant à $PFRG$ $(E \rightarrow F)$ ou de calculer les coût de protection des liens et nœuds protégés par des LSP de secours traversant cet arc (i.e. les liens A-B et B-E ainsi que le nœud B).

Ainsi, nous avons :

$\delta^{E \rightarrow F}_{A\text{-}B} = bw\,(b_1)$

$\delta^{E \rightarrow F}_{B} = bw\,(b_1)$

$\delta^{E \rightarrow F}_{B\text{-}E} = bw\,(b_2)$

$\delta^{E \rightarrow F}_{srlg_1} = bw\,(b_1) + bw\,(b_2)$

et :

$\theta^{E \rightarrow F}_{A\text{-}B} = Max\,(\delta^{E \rightarrow F}_{A\text{-}B},\, \delta_{srlg_1}) = bw\,(b_1) + bw\,(b_2)$

$\theta^{E \rightarrow F}_{B} = \delta^{E \rightarrow F}_{B} = bw\,(b_1)$

$\theta^{E \rightarrow F}_{B\text{-}E} = Max\,(\delta^{E \rightarrow F}_{B\text{-}E},\, \delta_{srlg_1}) = bw\,(b_1) + bw\,(b_2)$

En conséquence, la bande passante de secours minimale à allouer sur l'arc $E \rightarrow F$ est déduite comme suit :

$G^{\lambda} = Max\,(\delta^{E \rightarrow F}_{A\text{-}B},\, \delta^{E \rightarrow F}_{B},\, \delta^{E \rightarrow F}_{B\text{-}E},\, \delta^{E \rightarrow F}_{srlg_1}) = Max\,(\theta^{E \rightarrow F}_{A\text{-}B},\, \theta^{E \rightarrow F}_{B},\, \theta^{E \rightarrow F}_{B\text{-}E}) = bw\,(b_1) + bw\,(b_2)$

2.4.5 Respect des contraintes de bande passante

Pour assurer le respect des contraintes de bande passante, la quantité de bande passante utilisée sur un arc doit être constamment inférieure à la capacité de l'arc, c'est-à-dire l'invariant (2.5) doit être valide à tout moment[1] :

$$G^{\lambda} + F_{\lambda} \leq C^{\lambda} \tag{2.5}$$

En conséquence, pour établir un nouveau LSP de secours b, seuls les arcs λ vérifiant l'inégalité (2.6) peuvent être utilisés :

$$\theta^{\lambda}(b) + F_{\lambda} + bw(b) \leq C^{\lambda} \tag{2.6}$$

2.4.6 Surcoût d'un LSP de secours

Avec le partage de bande passante, l'établissement d'un LSP de secours traversant un arc λ n'induit pas systématiquement une augmentation de la bande passante de secours de l'arc λ d'une valeur correspondante à la bande passante du LSP construit. En effet, la quantité de bande passante de secours additionnelle à allouer sur un arc pour un LSP de secours, dite *surcoût du LSP*, dépend uniquement de la différence entre le plus grand coût de protection sur l'arc (déterminé avant l'établissement du LSP) et le coût de protection du LSP de secours (calculé après l'établissement du LSP). Formellement, le surcoût d'un LSP de secours b sur un arc λ est déterminé comme suit :

$$\gamma^{\lambda}(b) = \begin{cases} \infty \ si \ \theta^{\lambda}(b) + bw(b) + F_{\lambda} > C^{\lambda} \\ Max\left(\theta^{\lambda}(b) + bw(b) - G^{\lambda}, 0\right) \ sinon \end{cases} \tag{2.7}$$

Par extension, nous pouvons associer un surcoût de protection à un lien ou à un nœud. Ainsi, nous définissons le surcoût d'un lien ou nœud (c) sur un arc λ pour le support d'un LSP de secours de bande bw comme suit :

$$\gamma_c^{\lambda}(bw) = \begin{cases} \infty \ si \ \theta_c^{\lambda} + bw + F_{\lambda} > C^{\lambda} \\ Max\left(\theta_c^{\lambda} + bw - G^{\lambda}, 0\right) \ sinon \end{cases} \tag{2.8}$$

2.5 Modélisation ILP du problème de partage de la bande passante

2.5.1 Introduction

Dans cette section, nous modélisons le problème de placement d'un ensemble de LSP de secours minimisant une métrique donnée (bande passante additionnelle ou délai) en utilisant la programmation linéaire par entiers (ILP). Le but de cette modélisation est de montrer la complexité de la résolution du problème de placement des LSP de secours en ILP et la détermination de l'information requise au calcul des LSP de secours.

[1] La notation F_{λ} correspondant à la quantité de bande passante primaire, qui traverse le risque de panne λ, est volontaire (λ est mis comme indice car il correspond à un risque de panne).

2.5.2 Modèle

Étant donnés un réseau MPLS représenté par un graphe $G = (V, E, Rs)$. Des requêtes de demande d'établissement (ou de suppression) de LSP arrivent vers les différents LSR (ou nœuds) qui les traitent localement. Ces LSR déterminent les chemins qui devraient supporter les LSP primaires en utilisant les différentes informations (annoncées dans l'IGP-TE ou transmise par le protocole de signalisation) dont ils disposent. Ils configurent ensuite ces LSP (routés explicitement) en utilisant un protocole de signalisation tels que RSVP-TE [ABG+01] ou CR-LDP [JAC+02]. Chaque LSP primaire dispose d'un ensemble de LSP de secours permettant sa protection contre les pannes. Tous les LSP de secours sont configurés avec une quantité de bande passante égale à celle utilisée par leur LSP primaire pour rendre une même qualité de service après une panne[2]. Bien évidemment, toute nouvelle réservation de bande sur un arc $\alpha \rightarrow \beta$ doit être effectuée de telle sorte que la bande passante cumulée des LSP, pouvant être actifs au même moment et traversant l'arc, soit inférieure à la capacité de l'arc $C^{\alpha \rightarrow \beta}$.

Pour assurer le respect des contraintes de bande passante et/ou pour optimiser la quantité de bande passante allouée dans le réseau tout en garantissant la même qualité de service après une panne, les nœuds peuvent partager la bande passante entre les LSP de secours si ces derniers sont indépendants.

Nous nous plaçons délibérément dans le contexte en ligne où un certain nombre de LSP primaires et de secours sont déjà établis et qu'une nouvelle requête de demande de calcul et d'établissement d'un LSP primaire protégé est reçue. Ce mode en ligne nous semble effectivement plus réaliste puisqu'il diminue souvent les calculs (i.e. moins de calculs sont effectués avec le mode en ligne) et il conserve les configurations des LSP déjà établis. Avec ce mode, les requêtes sont traitées rapidement, au fur et à mesure qu'elles arrivent.

Bien que le mode en ligne ne permette pas d'optimiser la bande passante consommée dans le réseau, diverses heuristiques ont été élaborée et proposée pour augmenter la disponibilité de la bande passante. Par exemple, pour réduire la consommation de la bande passante avec le mode en ligne, il pourrait être judicieux de minimiser la bande passante additionnelle allouée à chaque nouveau LSP. D'autres métriques (comme le délai) pourraient être optimisées sous différentes contraintes (exemple : diminution du délai de récupération tout en évitant la violation des contraintes de la bande passante).

Supposons que $k\text{-}1_{(k>0)}$ LSP primaires sont déjà établis et qu'à chaque LSP primaire $p_{i\,(i \leq k)}$ un ensemble S_i de LSP de secours lui est associé pour le protéger. Les différents LSP de secours appartenant à l'ensemble S_i sont configurés par leur LSR de tête qui correspondent tous à des PLR.

A chaque requête d_i de demande d'établissement d'un LSP primaire p_i et de tous ses LSP de secours est associée une valeur bw_i représentant la quantité de bande passante réclamée par p_i. Pour des soucis de simplification, nous ne nous intéressons dans cette section qu'aux risques de panne de type lien. Les autres types de risques peuvent être

[2]On peut imaginer un facteur réduisant la quantité de ressources réservées pour les LSP de secours en supposant qu'en cas de panne, les liens survivant peuvent tolérer une légère surcharge qui durera le temps de la panne.

traités de la même manière (voir les sections suivantes). De plus, nous supposons que chaque lien logique α-β est composé de deux arcs logiques $\alpha \rightarrow \beta$ et $\beta \rightarrow \alpha$ partageant les mêmes composants physiques. En d'autre terme, la panne de l'arc $\alpha \rightarrow \beta$ implique forcément la panne de l'arc $\beta \rightarrow \alpha$ et vice versa. Dans le cas d'arcs logiques $\alpha \rightarrow \beta$ et $\beta \rightarrow \alpha$ n'utilisant pas les mêmes composants physiques, chaque arc sera considéré comme un risque de panne à part (de type arc) et sera traité indépendamment de l'autre.

Définissons les trois ensembles suivants : $A_{u\text{-}v}$, $B^{\alpha \rightarrow \beta}$ et $\xi_{u\text{-}v}^{\alpha \rightarrow \beta}$. L'ensemble $A_{u\text{-}v}$ est constitué de toutes les requêtes satisfaites, dont le LSP primaire traverse le lien u-v (i.e. le LSP primaire traverse l'arc $u \rightarrow v$ ou l'arc $v \rightarrow u$). L'ensemble $B^{\alpha \rightarrow \beta}$ est constitué de toutes les requêtes satisfaites et utilisant l'arc $\alpha \rightarrow \beta$ pour protéger les LSP primaires qui leur sont associés. Nous rappelons qu'une requête concerne un seul LSP primaire et tous les LSP de secours qui le protègent. Enfin, l'ensemble $\xi_{u\text{-}v}^{\alpha \rightarrow \beta}$ est constitué de toutes les requêtes utilisant au moins un LSP de secours passant par l'arc $\alpha \rightarrow \beta$ et protégeant contre le risque de panne de type lien u-v.

Si nous notons par $F_{u\text{-}v}$ la quantité de bande passante allouée sur le lien u-v pour les LSP primaires déjà construits et si nous définissons $G^{\alpha \rightarrow \beta}$ comme étant la quantité de bande passante effectivement allouée sur l'arc $\alpha \rightarrow \beta$ pour les différents LSP de secours déjà établis, afin d'assurer la disponibilité de la bande passante lors d'une panne, nous obtiendrons :

$$F_{u\text{-}v} = \sum_{i=(1,n) \setminus d_i \in A_{u\text{-}v}} bw_i$$

$$G^{\alpha \rightarrow \beta} \leq \sum_{i=(1,n) \setminus d_i \in B^{\alpha \rightarrow \beta}} bw_i \qquad \text{(l'inégalité est due au partage)}$$

Avec les définitions des sections 2.4.2, 2.4.3 et 2.4.4, nous obtenons aussi :

$$\delta_{u\text{-}v}^{\alpha \rightarrow \beta} = \theta_{u\text{-}v}^{\alpha \rightarrow \beta} = \sum_{i=(1,n) \setminus d_i \in \xi_{u\text{-}v}^{\alpha \rightarrow \beta}} bw_i$$

$$G^{\alpha \rightarrow \beta} = Max_{[(u,v) \setminus u \rightarrow v \in E \vee v \rightarrow u \in E]} (\delta_{u\text{-}v}^{\alpha \rightarrow \beta})$$

A l'arrivée d'une nouvelle requête k réclamant une quantité de bande passante égale à bw_k (ou bw pour simplifier), un nouveau LSP primaire p_k est déterminé. Définissons PLR_k comme l'ensemble des PLR activés sur le LSP primaire p_k. L'ensemble de tous les LSP de secours associés au LSP primaire p_k est noté S_k. A chaque $plr_k^h \in PLR_k$, est associé un LSP de secours $s_k^h \in S_k$ dont le LSR de tête est plr_k^h et qui joint le LSP primaire protégé en un nœud MP, noté m_k^h, en aval du risque (de type lien) à protéger. L'ensemble de tous les nœuds du LSP p_k qui sont en aval du risque (lien) à protéger est noté M_k^h ($m_k^h \in M_k^h$).

Pour satisfaire la requête d_k, un ensemble de LSP de secours protégeant les différents risques de panne (de type lien) du LSP primaire p_k seront établis. Grâce au partage,

la quantité de bande passante de secours additionnelle à allouer sur les différents arcs supportant ces LSP de secours sera inférieure ou égale à la bande passante du LSP primaire p_k. Pour chaque LSP de secours b de bande passante bw, cette quantité (qui correspond aussi au surcoût du lien) est déterminée comme suit :

$$\gamma_{u\text{-}v}^{\alpha\rightarrow\beta}(bw) = \begin{cases} 0, \ si \ (\delta_{u\text{-}v}^{\alpha\rightarrow\beta} + bw \leq G^{\alpha\rightarrow\beta}) \wedge (\alpha\text{-}\beta \neq u\text{-}v) \\ \delta_{u\text{-}v}^{\alpha\rightarrow\beta} + bw - G^{\alpha\rightarrow\beta}, \ si \ (\delta_{u\text{-}v}^{\alpha\rightarrow\beta} + bw > G^{\alpha\rightarrow\beta}) \wedge (u\text{-}v \in PFRG(b)) \\ \qquad\qquad\qquad\qquad \wedge (\delta_{u\text{-}v}^{\alpha\rightarrow\beta} + bw + F_{\alpha\rightarrow\beta} \leq C^{\alpha\rightarrow\beta}) \wedge (\alpha\text{-}\beta \neq u\text{-}v) \\ \infty, \ sinon \end{cases}$$

Différentes métriques peuvent être sélectionnées pour être optimisées. Par exemple, [KL03] propose un système ILP permettant d'optimiser conjointement la bande passante primaire allouée au LSP protégé ainsi que la bande passante de secours additionnelle allouée à son LSP de secours (en utilisant de la protection de bout en bout). Une autre étude décrite dans [XCX+06] propose et prouve la NP-difficulté du problème de détermination d'une paire de chemins disjoints (un chemin primaire et un chemin de secours) de telle sorte que la longueur du chemin primaire est minimale.

Ici, nous nous focalisons uniquement sur la protection locale et nous adoptons une approche successive où le LSP primaire est calculé avant tous ses LSP de secours. Deux métriques sont choisies pour être optimisées[3] : la bande passante de secours additionnelle allouée sur les arcs du réseau et le délai sur chaque LSP de secours (ou bien la longueur de chaque LSP).

2.5.3 Optimisation de la bande passante de secours additionnelle

En définissant la variable $z^{\alpha\rightarrow\beta}$ comme la quantité de bande passante de secours additionnelle allouée sur l'arc $\alpha\rightarrow\beta$ pour l'établissement de l'ensemble des LSP de secours protégeant le LSP primaire p_k (déjà établi), nous obtenons une solution optimisant la bande passante de secours additionnelle en minimisant la fonction d'objectif suivante :

$$Min \sum_{\alpha\rightarrow\beta \in E} z^{\alpha\rightarrow\beta}$$

Notons que $z^{\alpha\rightarrow\beta}$ peut aussi s'exprimer par la différence entre la bande passante de secours minimale allouée à l'étape k avec celle obtenue à l'étape *k-1*.

Si l'on pose :

$$y_k^{\alpha\rightarrow\beta} = \begin{cases} 1, \ si \ (\exists b \in S_k : \alpha\rightarrow\beta \in b) \\ 0, \ sinon \end{cases}$$

Le problème d'optimisation de la bande passante additionnelle allouée à un ensemble de LSP de secours S_k, protégeant un LSP primaire p_k qui réclame une quantité de bande passante bw, sera modélisé par le système suivant (constitué d'une fonction d'objectif

[3] Les systèmes ILP décrits dans cette section sont facilement adaptables pour l'optimisation d'autres métriques.

et de six contraintes) :

Min : $\sum_{\alpha\to\beta} z^{\alpha\to\beta}$

$$\begin{cases} \sum_{\beta\in V} y_h^{\alpha\to\beta} - \sum_{\beta\in(V\setminus M_k^h)} y_h^{\beta\to\alpha} = 0,\ \forall h\in[1,|PLR_k|]\ \wedge\ \forall\alpha\ \in V\setminus(\{plr_k^h\}\cup M_k^h) \\ \sum_{\beta\in V} y_h^{plr_k^h\to\beta} + \sum_{\beta\in V} y_h^{\beta\to plr_k^h} = 1,\ \forall h\in[1,|PLR_k|] \\ \sum_{m\in M_k^h} y_h^{\alpha\to m} + \sum_{m\in M_k^h} y_h^{m\to\alpha} = 1,\ \forall h\in[1,|PLR_k|]\ \wedge\ \forall\alpha\ \in V \\ \sum_{\beta\in V} y_h^{\alpha\to\beta} \le 1,\ \forall h\in[1,|PLR_k|]\ \wedge\ \forall\alpha\ \in V \\ \gamma_{u\text{-}v}^{\alpha\to\beta}(bw).y_h^{\alpha\to\beta} - z^{\alpha\to\beta} \le 0,\ \forall h\in[1,|PLR_k|]\ \wedge\ \forall\alpha\to\beta\ \in\ E\ \wedge\ \forall u\text{-}v\ \in PFRG(s_k^h) \\ y_h^{u\to v} + y_h^{v\to u} = 0,\ \forall\, u\text{-}v\ \in PFRG(s_k^h) \\ y_h^{\alpha\to\beta}\ \in\ \{0,1\},\ z_h^{\alpha\to\beta}\ \in\ [0,C^{\alpha\to\beta}],\ \forall h\in[1,|PLR_k|]\ \wedge\ \forall\,\alpha\to\beta\ \in E \end{cases} \tag{2.9}$$

Les quatre premières contraintes du système (2.9) permettent d'assurer que chaque LSP de secours interconnecte un nœud PLR à un nœud MP situé en aval du lien à protéger. La cinquième contrainte définit la quantité de bande passante additionnelle à allouer pour l'ensemble des LSP de secours protégeant un même LSP primaire. Cette quantité est déterminée sur tout arc comme le plus grand surcoût de protection de tous les LSP de secours traversant l'arc et protégeant le LSP primaire. La sixième contrainte impose la disjonction entre les différents LSP de secours calculés et leur LSP primaire. L'ajout de cette sixième contrainte n'est pas obligatoire puisque la disjonction entre le LSP primaire et ses LSP de secours est garantie grâce aux valeurs infinies des surcoûts de protection sur les arcs du LSP primaire. Enfin, la dernière contrainte donne les ensembles de définition des variables $z_{\alpha\to\beta}$ et $y_h^{\alpha\to\beta}$ utilisées pour indiquer sur tout arc la quantité de bande passante additionnelle et totale à allouer à l'ensemble des LSP de secours ainsi que les arcs formant les LSP de secours déterminés. La combinaison de la cinquième contrainte avec la septième contrainte permet aussi d'assurer le respect des contraintes de la bande passante puisque la bande passante additionnelle et totale allouée aux LSP de secours doit être inférieure à la capacité de l'arc.

Notons que même si les ensembles de LSP de secours protégeant les *k-1* premiers LSP primaires ont été établis et placés d'une manière optimale, la détermination de l'ensemble S_k (S_k contient l'ensemble des LSP de secours protégeant le LSP primaire p_k et minimisant la bande passante additionnelle) n'optimise pas forcément la totalité de la bande passante allouée à tous les LSP de secours.

2.5.3.1 Solutions du problème d'optimisation de la bande passante de secours additionnelle

La résolution du système (2.9) permet de déterminer un ensemble de LSP de secours fournissant la protection au nouveau LSP primaire (p_k), tout en minimisant la quantité de bande passante additionnelle qui leur sera allouée. Comme tous les LSP de secours protégeant un même LSP primaire sont indépendants ici (puisque l'ensemble des risques de panne Rs ne contient aucun SRLG), nous déduisons que la bande passante additionnelle allouée pour l'ensemble des LSP de secours sur tout arc est égale au maximum des surcoûts de protection de chaque LSP sur le même arc (cf. cinquième contrainte du système (2.9)).

Si l'on suppose que les surcoûts de protection de tous les LSP de secours sont égaux sur tous les arcs (ce qui est le cas lors de la protection du premier LSP primaire), nous nous apercevons que la solution du système (2.9) passe par la détermination d'un arbre (unidirectionnel) de Steiner (sur la topologie du réseau diminuée des arcs du LSP primaire à protéger) contenant tous les PLR activés[4] ainsi que la destination. En effet, la solution du système (2.9) doit vérifier les conditions suivantes :

1. contenir le nœud de destination : au moins un LSP de secours (celui qui protège contre la panne lien en aval de la destination) doit contenir le nœud de destination,
2. contenir tous les PLR,
3. une seule réservation de la bande passante de secours est requise sur chaque arc traversé par au moins un LSP de secours,
4. absence de boucles : deux LSP de secours protégeant un même LSP primaire ne doivent pas interconnecter deux nœuds du réseau en utilisant deux chemins différents car cela augmentera la bande passante additionnelle inutilement.

En conséquence, le problème d'optimisation de la bande passante de secours additionnelle est un problème NP-difficile. De ce fait, nous déduisons que la résolution du système (2.9) en un temps polynomial par rapport à la taille du problème ne peut pas être garantie.

2.5.3.2 Information requise pour résoudre le problème d'optimisation de la bande passante de secours additionnelle

Le système (2.9) permet de déduire l'information requise pour le calcul des LSP de secours protégeant un LSP primaire et minimisant la bande passante additionnelle allouée sur les arcs du réseau. Cette information est composée de :

1. la topologie du réseau $G = (V, E, Rs)$ ainsi que des capacités des arcs du réseau,
2. la structure du LSP primaire à protéger ainsi que sa bande passante,
3. les quantités de bande passante cumulées de tous les LSP primaires sur chaque arc (pour permettre le calcul des surcoûts de protection),

[4]Nous considérons que tous les nœuds du LSP primaire, excepté le nœud de destination, sont des PLR.

4. les quantités de bande passante de secours (minimales) allouées sur chaque arc,
5. l'ensemble de PLR activés sur le réseau $G = (V, E, Rs)$,
6. les prix de protection de tous les risques appartenant au LSP primaire à protéger (pour le calcul des surcoûts des différents LSP de secours).

Dans un environnement distribué, l'efficacité d'une solution de placement de LSP de secours dépend fortement de la qualité du mécanisme distribuant l'information ci-dessus. Nous nous pencherons plus en détails, par la suite sur l'étude de ces mécanismes (sections 2.7 et 2.8 et chapitre 3).

2.5.4 Optimisation du délai (optimisation de la longueur des LSP de secours)

Pour optimiser toute autre métrique différente (que la bande passante additionnelle), le système (2.9) peut être adapté. Ainsi, si nous voulons minimiser le délai sur chaque LSP de secours, il suffira de modifier la fonction d'objectif. Concrètement, si nous définissons $d\ (\alpha{\rightarrow}\beta)$ comme une fonction retournant le délai sur l'arc $\alpha{\rightarrow}\beta$, nous obtiendrons le système (2.10) qui minimise le délai[5] :

Min : $\sum_{(\alpha\rightarrow\beta,h)}\ d\ (\alpha{\rightarrow}\beta) \times y_h^{\alpha\rightarrow\beta}$

$$
\begin{cases}
\sum_{\beta\in V} y_h^{\alpha\rightarrow\beta} - \sum_{\beta\in (V\setminus M_k^h)} y_h^{\beta\rightarrow\alpha} = 0,\ \forall h \in [1,|PLR_k|] \wedge \forall\alpha\ \in V\setminus(\{plr_k^h\}\cup M_k^h) \\
\sum_{\beta\in V} y_h^{plr_k^h\rightarrow\beta} + \sum_{\beta\in V} y_h^{\beta\rightarrow plr_k^h} = 1,\ \forall h \in [1,|PLR_k|] \\
\sum_{m\in M_k^h} y_h^{\alpha\rightarrow m} + \sum_{m\in M_k^h} y_h^{m\rightarrow\alpha} = 1,\ \forall h \in [1,|PLR_k|] \wedge \forall\alpha\ \in V \\
\sum_{\beta\in V} y_h^{\alpha\rightarrow\beta} \leq 1,\ \forall h \in [1,|PLR_k|] \wedge \forall\alpha\ \in V \\
\delta_{u\text{-}v}^{\alpha\rightarrow\beta}.y_h^{\alpha\rightarrow\beta} + bw + F_{u\rightarrow v} \leq C^{\alpha\rightarrow\beta},\ \forall h \in [1,|PLR_k|] \wedge \forall\alpha{\rightarrow}\beta\ \in\ E \wedge \forall u\text{-}v\ \in PFRG\,(s_k^h) \\
y_h^{u\rightarrow v} + y_h^{v\rightarrow u} = 0,\ \forall\, u\text{-}v\ \in PFRG\,(s_k^h) \\
y_h^{\alpha\rightarrow\beta}\ \in\ \{0,\,1\},\ \forall h \in [1,|PLR_k|] \wedge \forall\,\alpha{\rightarrow}\beta\ \in E
\end{cases}
\tag{2.10}
$$

Si nous comparons les deux systèmes (2.9) et (2.10), nous nous apercevons que seule la fonction d'objectif, la cinquième et la septième contraintes ont été modifiées. La modification de la fonction d'objectif est nécessaire et s'explique par le changement de la métrique optimisée. Pour les cinquième et septième contraintes, leur modification n'est pas nécessaire pour optimiser le délai. Nous les avons modifié pour éliminer les variables $z^{\alpha\rightarrow\beta}$ et $\gamma_{u\text{-}v}^{\alpha\rightarrow\beta}$ dans le nouveau système (2.10).

[5]Bien évidemment, l'algorithme de Dijkstra est moins complexe et est plus adéquat que l'ILP pour la recherche de l'ensemble des LSP de secours minimisant le délai.

2.5.4.1 Solutions du problème d'optimisation du délai

La résolution du système (2.10) permet de déterminer un ensemble de LSP de secours minimisant le délai entre le PLR ayant détecté la panne et la destination. Les contraintes du système permettent d'éliminer de la topologie du réseau tous les arcs ne vérifiant pas les contraintes de bande passante. La fonction d'objectif est utilisée ensuite pour déduire les LSP de secours minimisant le délai.

Évidemment et pour des considérations de performances, il sera mieux d'appliquer l'algorithme de Dijkstra (au lieu de l'ILP) pour calculer les LSP de secours sur la topologie du réseau diminuée des arcs ne vérifiant pas les contraintes de bande passante. Cela garantit la détermination d'une solution en un temps polynomial par rapport à la taille du problème d'optimisation du délai.

Notons que même si la métrique optimisée n'est pas la bande passante, nous nous efforçons d'allouer un minimum de bande passante de secours sur chaque arc, en utilisant le partage de la bande passante, afin d'éviter le rejet inutile de certaines requêtes de protection.

2.5.4.2 Information requise pour minimiser le délai

Mis à part les quantités de bande passante de secours allouées sur les arcs, l'optimisation de la métrique du délai requiert la connaissance des mêmes informations que celles qui sont requises pour l'optimisation de la bande passante additionnelle. En effet, l'optimisation de la métrique du délai (ou de toute autre métrique) nécessite la vérification des contraintes de bande passante, ce qui implique la connaissance des informations suivantes :

1. La topologie du réseau $G = (V, E, Rs)$ ainsi que les capacités des arcs du réseau.
2. La structure du LSP primaire à protéger ainsi que sa bande passante.
3. Les quantités de bande passante cumulées de tous les LSP primaires sur chaque arc.
4. L'ensemble de PLR activés sur le réseau $G = (V, E, Rs)$.
5. Les prix de protection de tous les risques appartenant au LSP primaire à protéger.
6. le délai sur chaque arc de la topologie du réseau (ou l'information permettant de déterminer la valeur de la métrique associée à chaque arc).

Comme nous le constatons, indépendamment de la métrique optimisée, le placement des LSP de secours requiert la distribution d'une information semblable à celle permettant l'optimisation de la bande passante additionnelle, pour éviter la gaspillage de la bande passante et vérifier ses contraintes sur chaque arc.

2.6 Partage et environnement

En fonction de l'environnement adopté (environnement centralisé ou environnement distribué), la tâche de placement des LSP de secours peut être plus ou moins coûteuse et efficace (passage à l'échelle). Dans cette section, nous verrons comment tenir compte

du partage et réaliser le contrôle d'admission dans un environnement centralisé ainsi que dans un environnement distribué.

2.6.1 Partage de la bande passante dans un environnement centralisé

Dans un environnement centralisé, un seul serveur s'occupe du placement des LSP de secours. Par conséquent, toute l'information nécessaire au placement des LSP de secours (cf. sections 2.5.3.2 et 2.5.4.2) doit être transmise au serveur. Dans la pratique, seule la topologie du réseau et les risques associées sont transmis au serveur. Les autres informations nécessaires pour le placement des LSP de secours (prix de protection, bandes passante primaires cumulées, etc.) sont déduites des requêtes envoyées au serveur et des chemins calculés et configurés par le serveur.

Bien que la solution centralisée soit simple à mettre en œuvre, son utilisation induit certains inconvénients comme :

1. Le serveur constitue un goulot d'étranglement et donc un obstacle pour le passage à l'échelle. De plus, la liaison entre le serveur et le reste du réseau constitue un point de panne critique.
2. Nécessité d'un protocole de communication pour l'envoi des requêtes de calcul (ou de suppression) des LSP vers le serveur et la réception des configurations en réponse.
3. Délais de réponse des requêtes très allongés lors d'une panne.

A cause de l'allongement des délais de réponse après une panne, l'hypothèse de pannes simples adoptée dans cette thèse pour simplifier la procédure de protection, est remise en cause. En effet, suite à une panne, le serveur centralisé doit entrer dans une période de calculs intensifs afin de protéger entièrement le réseau. Pendant toute cette période de calcul qui est assez longue, l'occurrence d'une nouvelle panne pourra être considérée comme une panne survenue simultanément avec la première panne, ce qui ne garantit pas le succès de la récupération.

2.6.2 Contrôle d'admission dans un environnement centralisé

Le contrôle d'admission a pour tâche de prémunir la violation des contraintes de bande passante sur tous les arcs de la topologie du réseau. Pour ce faire, à chaque calcul et/ou configuration d'un nouveau LSP, le mécanisme de contrôle d'admission doit vérifier que la bande passante disponible sur les différents arcs formant le LSP est suffisante pour transporter son flux.

Dans un environnement centralisé, la tâche de contrôle d'admission peut être assurée par le serveur centralisé. En effet, ce dernier peut déduire les valeurs de tous les prix de protection sur tous les arcs ainsi que les bandes passantes primaires cumulées sur les arcs à partir des structures et propriétés des LSP qu'il a calculés. Les contraintes de bande passante sont ensuite vérifiées sur tous les arcs afin d'assurer que seuls les arcs disposant d'une bande passante suffisante sont sélectionnés pour être dans le prochain LSP calculé.

2.6.3 Partage de la bande passante dans un environnement distribué

Avec le partage de bande passante dans un environnement distribué, plusieurs modules sont dédiés au placement distribué des LSP de secours. Ces modules se chargent de la collecte et du traitement de l'information nécessaire au placement des LSP de secours (cf. sections 2.5.3.2 et 2.5.4.2). Les nœuds supportant ces modules devraient être sélectionnés de manière à optimiser le coût du traitement et surtout le coût de la distribution de l'information requise au placement des LSP de secours.

Constatant que certains paramètres de l'information requise pour le placement des LSP de secours sont relativement stables et/ou changent moins fréquemment que d'autres, nous concluons qu'il est utile de partitionner l'information en plusieurs segments, chacun contenant des paramètres ayant des propriétés de variabilité (ou stabilité) semblables. Ainsi et concernant la topologie du réseau et ses risques (l'information $G = (V, E, Rs)$) qui sont relativement stables, il est judicieux de ne les distribuer qu'à l'initialisation du réseau et à des intervalles de temps assez long (quelques minutes pour OSPF par exemple). Les protocoles IGP, en général, permettent de découvrir et de transmettre la topologie du réseau ainsi que les risques associés à l'ensemble des nœuds du réseau. De la même manière, les bandes passantes primaires cumulées $F_{\alpha\rightarrow\beta}$ et éventuellement les capacités $C^{\alpha\rightarrow\beta}$ des arcs peuvent aussi être diffusés efficacement dans le réseau grâce aux protocoles IGP-TE. Pour les structures des LSP primaires et leurs quantités de bande passante, il est plutôt intéressant de limiter leur distribution aux nœuds qui les constituent. Cela réduit significativement le nombre de messages transmis dans le réseau mais induit une nouvelle contrainte qui consiste en l'instanciation d'au moins un module de calcul sur chaque nœud du LSP primaire.

Concernant le dernier paramètre (prix de protection[6]) de l'information requise au placement des LSP de secours, il est moins stable et dépend de tous LSP de secours déjà établis. Tout nœud doit recevoir les valeurs de ce paramètre sur tous les arcs et pour tous les risques dont il est responsable de la protection contre les pannes.

Comme nous l'apercevons, le coût de la distribution de l'information (requise au placement des LSP de secours dépend très étroitement du coût de la distribution des prix de protection des risques sur les différents arcs de la topologie du réseau. En conséquence, la détermination d'un algorithme efficace de distribution de l'information requise pour le placement des LSP de secours aux différents modules de calcul[7] passe par une transmission ciblée et intelligente des prix de protection aux nœuds.

Différentes techniques ont été élaborées afin de permettre un partage maximal de la bande passante de secours sur les arcs, tout en assurant le respect de ses contraintes et éventuellement en optimisant différentes métriques (comme la bande passante de secours additionnelle ou le délai). Certaines techniques se basent sur une optimisation du coût de la distribution des prix de protection, d'autres sur l'envoi d'une information

[6] les prix de protection combinés aux quantités de bande passante primaires sur les arcs, aux quantités de bande passante de secours sur les arcs et aux capacités des arcs permettent aussi de déduire les surcoûts de protection, comme le montre les formules (2.2), (2.3), (2.4) et (2.8).

[7] Bien évidemment, les modules de calcul doivent être eux aussi placés efficacement sur les nœuds de la topologie du réseau.

équivalente ou agrégée, permettant de déduire les valeurs de ces prix de protection.

Dans la section 2.7, nous décrivons les méthodes exactes et les heuristiques distribuées qui minimisent la quantité de bande passante de secours allouée sur chaque arc. Les méthodes exactes de partage permettent de déterminer la quantité exacte de bande passante de secours qui peut être partagée pour chaque risque sur tout arc avant de placer les LSP de secours alors que dans les heuristiques de partage, la quantité de bande passante de secours partageable sur un arc utilisée dans le placement des LSP de secours n'est qu'une valeur approchée.

2.6.4 Contrôle d'admission dans un environnement distribué

Dans un environnement distribué, il est possible d'implanter le mécanisme de contrôle d'admission sans aucun surcoût. Avec IGP-TE et RSVP-TE par exemple, tous les nœuds du réseau connaissent les LSP qui les traversent, leurs quantités de bande passante, les risques de panne protégés, ainsi que les structures de ces différents risques et les capacité des arcs. De cette information, chaque nœud est capable de déduire les différents paramètres de bande passante permettant d'effectuer le contrôle d'admission sur tous les arcs qui lui sont adjacents.

2.7 Méthodes exactes et distribuées de placement des LSP de secours

Afin de placer efficacement les LSP de secours dans un environnement distribué, les nœuds supportant les modules de calcul ont besoin de connaître certaines informations liées à la topologie, au trafic (les LSP et leur quantité de bande passante) et aux risques de panne (groupes de risques de panne et leurs structures). Comme nous l'avons expliqué précédemment, seule la distribution des prix de protection peut être coûteuse.

Dans les méthodes présentées ci-dessous, nous nous concentrons plus sur les mécanismes de distribution des prix de protection. Pour un calcul efficace, rapide et en ligne des LSP de secours, nous utilisons l'algorithme CSPF qui détermine des plus courts chemins sur la topologie du réseau restreinte aux arcs vérifiant les contraintes de la bande passante (cf. section 2.4.5). La métrique à optimiser peut être la bande passante additionnelle (cf. section 2.5.3) ou une toute autre métrique (cf. section 2.5.4).

Dans la suite de cette section, nous présentons deux méthodes exactes et distribuées de placement des LSP de secours basées sur les diffusions et une autre méthode réduisant la quantité d'informations transmises dans le réseau par un placement intelligent des entités de calcul des LSP de secours dans le réseau.

2.7.1 Diffusion de tous les prix de protection dans le réseau

Dans une première approche, [KKL+01] propose de diffuser directement les valeurs des prix de protection de tous les risques et les quantités de bande passante résiduelles sur tous les arcs à tous les nœuds du réseau (i.e. $\forall \lambda \in E$: annoncer la quantité $R^{\lambda} =$

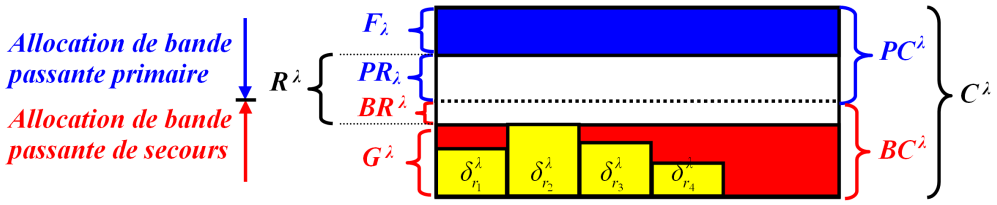


FIG. 2.5 – Partition de la capacité d'un arc λ en deux pools disjoints (pool primaire et pool de secours)

C^λ - F_λ - G^λ). De cette manière et comme dans le cas centralisé, chaque nœud dispose de toute l'information requise au placement des LSP de secours.

Cette méthode nécessite la diffusion d'une quantité élevée d'informations puisqu'elle requiert de distribuer, pour chaque arc λ d'un nouveau LSP de secours un vecteur de prix de protection contenant jusqu'à $|Rs|$ composantes (où Rs est l'ensemble de tous les risques de panne).

Bien qu'une simple extension des états de liens des protocoles IGP-TE (extension des LSPDU pour ISIS-TE et extension des LSA pour OSPF-TE) permette la diffusion des valeurs des prix de protection, cette méthode de placement de LSP de secours souffre de trois problèmes limitant son utilisation :

- augmentation significative de la charge du réseau,
- fréquence de diffusion très élevée puisque tout placement d'un nouveau LSP de secours induit une diffusion de plusieurs vecteurs de prix de protection,
- taille d'un état de lien pouvant être très élevée (taille plus grande que les messages IGP-TE) puisqu'elle dépend de la taille de l'ensemble des risques de panne.

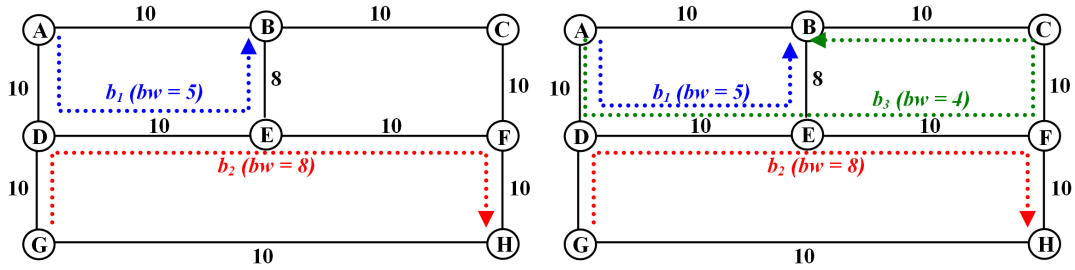
2.7.2 Diffusion des structures et des quantités de bande passante des tunnels de secours

Pour remédier aux problèmes de la fréquence élevée de diffusion et de la grande taille des vecteurs des prix de protection annoncés dans le réseau, [LRC02] propose d'adopter la protection par tunnel de secours et de regrouper toutes les informations concernant des tunnels de secours d'un même PLR dans un même message qui sera ensuite diffusé dans le réseau. Ce message doit spécifier pour chaque tunnel de secours :

- son type (NHOP ou NNHOP) et le groupe de risques de panne qu'il protège,
- le chemin le supportant,
- sa bande passante de secours cumulée.

En combinant l'information sur les tunnels de secours avec les différents paramètres transmis par les protocoles IGP-TE et les protocoles de signalisation, chaque PLR est capable de calculer les prix de protection (ainsi que les surcoûts de chaque nouveau tunnel de secours) qu'il stockera dans une base de données.

Pour faciliter le placement des LSP et afin de simplifier la gestion de la bande passante tout en améliorant son utilisation, la capacité de chaque arc λ est divisée en deux pools disjoints : pool de bande passante primaire de capacité PC^λ et pool de bande



(a) Base de données disponible sur chaque nœud après la création de deux tunnels de secours

(b) Base de données disponible sur chaque nœud après la création de trois tunnels de secours

FIG. 2.6 – Placement des tunnels de secours par construction de bases de données identiques sur chaque nœud

passante de secours de capacité BC_λ (cf. figure 2.5). Le premier pool (constitué de la bande passante primaire cumulée F_λ et de la bande passante primaire résiduelle PR_λ) est utilisé pour allouer la bande passante aux LSP primaires. Le second pool (constitué de la bande passante de secours G^λ allouée sur l'arc λ et de la bande passante de secours résiduelle[8] BR^λ) est employé pour réserver la bande passante aux tunnels de secours. En plus des avantages cités plus haut, la partition de la bande passante en deux pools disjoints permet d'assurer une protection complète (avec une garantie de la bande passante) de tous les LSP de primaires qui vérifient les contraintes de la bande passante primaire [Gro04].

Sur la figure 2.6 (a), deux tunnels de secours b_1 ($A \rightarrow D \rightarrow E \rightarrow B$) et b_2 ($G \rightarrow D \rightarrow E \rightarrow F \rightarrow H$) ont été établis. Le premier tunnel de secours réclame une quantité de bande passante bw (b_1)= 5 unités et protège contre la panne du lien A-B. Le second tunnel de secours réclame une quantité de bande passante bw (b_2) = 8 unités et protège contre la panne du lien G-H. Les deux liens A-B et G-H n'appartiennent pas à un même SRLG. Tous les arcs de la topologie ont une capacité de secours égale à 10 unités (i.e. $\forall\ \lambda \neq E \rightarrow B$: BC^λ = 10 unités) sauf l'arc $E \rightarrow B$ qui n'en dispose que de 8 (i.e. $BC^{E \rightarrow B}$ = 8 unités).

Après l'annonce des structures et propriétés des deux tunnels de secours (b_1 et b_2) et grâce à l'information transmise par les protocoles IGP-TE (topologie du réseau, risques, capacités des arcs, bandes passantes primaires cumulées sur chaque arc), chaque nœud sera en mesure de calculer les valeurs des prix de protection en appliquant la formule 2.1. Pour les arcs $D \rightarrow E$, $E \rightarrow B$ et $E \rightarrow F$, nous avons :

$$\delta_{A\text{-}B}^{D \rightarrow E} = 5\ unités,\ \delta_{G\text{-}H}^{D \rightarrow E} = 8\ unités,\ \delta_{A\text{-}B}^{E \rightarrow B} = 5\ unités,\ \delta_{G\text{-}H}^{E \rightarrow F} = 8\ unités.$$

Pour assurer le respect des contraintes de bande passante sur un arc λ dont la capacité est partitionnée en deux pools disjoints (pool primaire et pool de secours),

[8] La somme de la bande passante primaire résiduelle PR_λ et la bande passante de secours résiduelle BR^λ sur un arc λ est égale à la bande passante résiduelle R^λ.

l'invariant suivant doit être valide :

$$G^{\lambda} \leq BC^{\lambda} \tag{2.11}$$

En conséquence, pour établir un nouveau LSP de secours b (lorsque la capacité de chaque arc est divisé en deux pools), seuls les arcs λ vérifiant l'inégalité (2.12) peuvent être utilisés :

$$\theta^{\lambda}(b) + bw(b) \leq BC^{\lambda} \tag{2.12}$$

Si l'on désire, par exemple, placer un nouveau LSP de secours b_3 de bande passante égale à 4 unités et protégeant contre la panne du lien A-B, l'arc $E \rightarrow B$ ne pourra pas être utilisé puisque : $\theta^{E \rightarrow B}(b_3) + bw(b_3) = \delta_{A\text{-}B}^{E \rightarrow B} + bw(b_3) = 5 + 4 > BC^{E \rightarrow B} = 8$. Par contre, les deux arcs $D \rightarrow E$ et $E \rightarrow F$ pourront appartenir au nouveau LSP de secours b_3 puisque : $\theta^{D \rightarrow E}(b_3) + bw(b_3) = \delta_{A\text{-}B}^{D \rightarrow E} + bw(b_3) = 5 + 4 \leq BC^{E \rightarrow B} = 10$ et $\theta^{E \rightarrow F}(b_3) + bw(b_3) = \delta_{A\text{-}B}^{E \rightarrow F} + bw(b_3) = 0 + 4 \leq BC^{E \rightarrow B} = 10$.

Pour le calcul du surcoût d'un LSP (lorsque la capacité d'un arc est subdivisée en deux pools disjoints) :

$$\gamma^{\lambda}(b) = \begin{cases} \infty \ si \ \theta^{\lambda}(b) + bw(b) > BC^{\lambda} \\ Max\left(\theta^{\lambda}(b) + bw(b) - G^{\lambda},\, 0\right) \ sinon \end{cases} \tag{2.13}$$

De même, le surcoût de protection d'un lien ou d'un nœud c sur un arc λ pour le support d'un LSP de secours de bande bw est déterminé comme suit :

$$\gamma_c^{\lambda}(bw) = \begin{cases} \infty \ si \ \theta_c^{\lambda} + bw > BC^{\lambda} \\ Max\left(\theta_c^{\lambda} + bw - G^{\lambda},\, 0\right) \ sinon \end{cases} \tag{2.14}$$

Pour optimiser la quantité de bande passante additionnelle induite par l'établissement du tunnel de secours b_3, il suffit de calculer les surcoûts de ce tunnel sur tous les arcs de la topologie du réseau et d'appliquer ensuite l'algorithme de Dijkstra. Ainsi, nous avons :
$\gamma^{A \rightarrow B}(b_3) = \gamma^{B \rightarrow A}(b_3) = \infty$ par convention,
$\gamma^{E \rightarrow B}(b_3) = \infty$ car $\theta^{E \rightarrow B}(b_3) + bw(b) = 5 + 4 > BC^{E \rightarrow B} = 8$,
$\gamma^{D \rightarrow E}(b_3) = Max\left(\theta^{D \rightarrow E}(b_3) + bw(b_3) - G^{D \rightarrow E},\, 0\right) = 5 + 4 - 8 = 1$,
$\gamma^{G \rightarrow D}(b_3) = \gamma^{E \rightarrow F}(b_3) = \gamma^{F \rightarrow H}(b_3) = Max\left(\theta^{G \rightarrow D}(b_3) + bw(b_3) - G^{D \rightarrow E},\, 0\right)$
$= Max(0 + 4 - 8, 0) = 0$,
$\forall \lambda \notin \{A \rightarrow B,\, B \rightarrow A,\, E \rightarrow B,\, G \rightarrow D,\, D \rightarrow E,\, E \rightarrow F,\, F \rightarrow H\} : \gamma^{\lambda}(b_3) = 4$.

En conséquence, le LSP de secours b_3 qui minimise la bande passante de secours additionnelle (surcoût global égal à 13 unités) sera constitué des arcs : $A \rightarrow D$, $D \rightarrow E$, $E \rightarrow F$, $F \rightarrow C$ et $C \rightarrow B$ (cf. figure 2.6 (b)).

Cette méthode de placement de tunnels de secours nécessite que les nœuds du réseau connaissent les structures et les propriétés de tous les tunnels de secours établis. La distribution et la maintenance de cette information requiert la diffusion de messages de

contrôle périodiquement et à chaque établissement ou suppression d'un LSP primaire protégé, un pour chaque PLR. Cette information peut être diffusée dans les messages IGP-TE mais cela augmente considérablement la taille de ces messages (augmentation de la taille moyenne des messages OSPF-TE de 35% dans le cas de la protection locale contre les pannes de nœuds pour un réseau de degré 4).

2.7.3 Placement des LSP de secours par PCE

Afin d'éliminer les diffusions de messages permettant la déduction des prix de protection, [VCLF+04] propose de centraliser tous les calculs de LSP de secours protégeant un même ensemble de risques de panne Rs_i ($Rs_i \in Rs$) sur une même entité PCE_i. Nous rappelons qu'un PCE (*Path Computation Element*) est une entité capable de calculer des chemins, avec ou sans contraintes, dans un réseau [FVA06].

Sous l'hypothèse de pannes simples et en partitionnant la capacité de chaque arc en deux pools disjoints[9] (pool primaire et pool de secours), tout PCE_i est capable de calculer les LSP de secours vérifiant les contraintes de la bande passante et protègeant contre les pannes des risques qui lui sont associés (Rs_i). En effet, pour assurer le respect des contraintes de la bande passante, il suffira d'assurer la validité de l'invariant suivant :

$$G^{\lambda} \leq BC^{\lambda}$$

Or

$$G^{\lambda} = Max_{r \in Rs}(\delta_r^{\lambda})$$

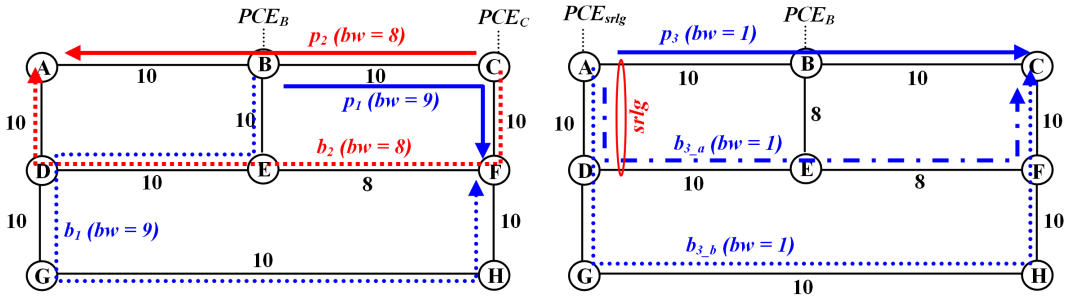
Pour protéger contre les pannes de risques appartenant à Rs_i tout en garantissant le respect des contraintes de bande passante, il suffira alors d'assurer sur chaque PCE_i la validité de l'invariant suivant :

$$Max_{r \in Rs_i}(\delta_r^{\lambda}) \leq BC^{\lambda}$$

Comme tous les chemins de secours protégeant contre les risques de panne appartenant à Rs_i sont effectués par la même entité PCE_i, nous déduisons que tous les prix de protection associés aux risques contenus dans Rs_i sont connus du PCE_i. En conséquence, ce dernier pourra déterminer tous les arcs pouvant supporter un nouveau LSP de secours lorsque ce dernier ne protège que contre des risques appartenant à Rs_i.

Pour assurer la tâche de calcul des LSP de secours sans la transmission des prix de protection (ou des structures et propriétés des LSP de secours) aux différents PCE, l'ensemble des risques de panne Rs doit être partitionné en plusieurs sous ensembles $Rs_{i0<i<n}$ de telle sorte que l'ensemble des risques de panne protégés par tout nouveau LSP de secours soit inclus dans au moins un sous-ensemble Rs_i. Formellement :
$\forall\ b \setminus b$ est un LSP de secours, $\exists i$: $PFRG$ (b) $\subseteq Rs_i$.

[9] Sans la subdivision de la capacité de chaque arc en deux pools disjoints (pool primaire et pool de secours), il sera nécessaire de connaître les quantités de bande passante de secours (minimales) allouées sur les arcs pour assurer le respect des contraintes de la bande passante.



(a) Un LSP de secours pour la protection d'un nœud et de son lien en aval (réseau sans SRLG)

(b) Un LSP de secours par lien et/ou nœud protégé (réseau avec SRLG)

FIG. 2.7 – Placement de LSP de secours par PCE

En l'absence de SRLG dans le réseau, le placement de LSP de secours de type NNHOP[10] requiert le regroupement de tout nœud avec ses liens adjacents. En conséquence, tout sous-ensemble Rs_i doit contenir au moins un nœud et tous les liens qui lui sont adjacents. Pour ce faire, les structures et propriétés des LSP de secours protégeant contre la panne d'un lien doivent être partagées par les deux nœuds adjacents au lien (car un lien peut être protégé par deux PCE, chaqu'un le protège dans un sens). Cela se fait sans surcoût lors de l'envoi des configurations de LSP de secours calculés par le PCE vers le PLR. Pour une partition de Rs minimisant la communication entre PLR et PCE, [VCLF+04] propose d'associer à chaque nœud du réseau un PCE permettant de protéger contre les pannes du nœud lui-même et tous ses liens adjacents.

Sur la figure 2.7 (a), deux LSP primaires p_1 ($B \rightarrow C \rightarrow F$) et p_2 ($C \rightarrow B \rightarrow A$) traversant le même lien B-C sont établis. Afin de protéger le LSP primaire p_1 contre les pannes du lien B-C et du nœud C, un LSP de secours b_1 ($B \rightarrow E \rightarrow D \rightarrow G \rightarrow H \rightarrow F$) est calculé par l'entité PCE_C supportée par le nœud C. Pour configurer le LSP de secours b_1, la structure et les propriétés de ce LSP sont envoyées au nœud PLR B. Ce dernier supporte une entité de calcul de LSP de secours protégeant contre les pannes du nœud B lui-même et de ses liens adjacents. Grâce aux configurations de LSP de secours protégeant contre la panne du lien C-B reçues du nœud C et aux LSP de secours calculés localement, l'entité PCE_B supportée par le nœud B, sera capable de calculer un LSP de secours b_2 ($A \rightarrow D \rightarrow G \rightarrow H \rightarrow F \rightarrow C$) protégeant le LSP primaire p_2 contre les pannes du nœud B et du lien C-B.

Lorsque l'ensemble des risques de panne Rs contient des risques de type SRLG, l'établissement de LSP de type NNHOP nécessite des échanges d'informations entre certains PCE. Ainsi, pour distribuer les calculs de LSP de secours sur plusieurs PCE, [VCLF+04] propose de restreindre l'ensemble des risques protégés par un LSP de type NNHOP à un seul nœud (i.e. un LSP NNHOP ne doit pas protéger contre la panne d'un lien). En conséquence, pour protéger un LSP de N nœuds, *N-2* LSP de secours NNHOP et *N-1* LSP NHOP pourraient être nécessaires.

[10] le placement d'un LSP NHOP n'induit aucune contrainte particulière sur la manière de partitionner l'ensemble Rs.

Sur la figure 2.7 (b), un LSP primaire p_3 ($A \rightarrow B \rightarrow C$) traversant un lien A-B appartenant au SRLG *srlg* (*srlg* = (A-B, D-E)) est établis. Afin de protéger ce LSP primaire contre les pannes du nœud B et du lien A-B, deux LSP de secours sont nécessaires : un LSP de secours b_{3_a} ($A \rightarrow D \rightarrow E \rightarrow F \rightarrow C$) protégeant contre la panne du nœud B et un LSP de secours b_{3_b} ($A \rightarrow D \rightarrow G \rightarrow H \rightarrow F \rightarrow C$) protégeant contre la panne du lien A-B (ainsi que contre la panne de tous les SRLG qui le contiennent). De cette manière, lors de panne du nœud B, seul le LSP de secours b_{3_a} de la figure 2.7 (b) sera activé. De même, lors de la panne du lien A-B ou de la panne du SRLG *srlg*, seul le LSP de secours b_{3_b} de la figure 2.7 (b) sera activé.

Notons que pour optimiser la bande passante additionnelle de secours allouée sur les arcs, cette méthode de placement des LSP de secours requiert la diffusion des quantités de bande passante de secours (minimales) allouées sur les arcs (i.e. $\{G^{\lambda}\}_{\lambda \in E}$).

Bien que cette technique de placement des LSP de secours avec PCE élimine complètement la diffusion de messages permettant de déduire les prix de protection, elle présente divers inconvénients restreignant son utilisation :

- Elle nécessite communication entre PLR et nœuds supportant des PCE. Cela requiert l'implantation d'un protocole gérant cette communication.
- Le placement d'un LSP de secours protégeant contre la panne d'un SRLG $srlg_1$ doit être effectué par un même PCE que celui qui place les LSP de secours protégeant contre la panne d'un SRLG $srlg_2$ ($srlg_1 \neq srlg_2$) si les deux SRLG $srlg_1$ et $srlg_2$ ne sont pas disjoints (i.e. $srlg_1$ et $srlg_2$ partagent au moins un lien commun). Ceci induit la centralisation du placement de tous les LSP de secours, protégeant contre les pannes de SRLG non disjoints, dans un même PCE, ce qui induit les mêmes problèmes que ceux rencontrés dans les systèmes centralisés.
- L'utilisation de deux LSP de secours pour protéger contre les pannes d'un nœud et de son lien en aval nécessite un mécanisme permettant une distinction des pannes des nœuds de celles des liens. Bien que cette distinction de pannes peut être effectuée relativement vite avec [VC02, AKNS08], elle allonge significativement les délais de récupération puisqu'aucun LSP de secours ne peut être activé avant. De plus, l'utilisation de deux LSP pour la protection contre les pannes d'un nœud et de son lien en aval induit une augmentation du nombre de LSP de secours établis. Cela accroît le nombre de messages envoyés dans le réseau pour l'établissement, le maintien, la mise-à-jour et la suppression de ces LSP.

2.7.4 Constats

Bien que les méthodes exactes de placement des LSP de secours puissent permettre d'optimiser la quantité de bande passante additionnelle de secours allouée sur les arcs ou toute autre métrique en assurant le respect des contraintes de bande passante, elles induisent différents désavantages comme :

- la surcharge du réseau dans le cas de diffusion des prix de protection (ou des structures de LSP de secours) à tous les nœuds du réseau. En effet, à chaque création, modification ou suppression d'un LSP de secours ou à chaque déclenchement d'un temporisateur, un message transportant les nouvelles valeurs des prix

de protection est diffusé.
- la centralisation des calculs de LSP de secours protégeant contre des risques de type SRLG non disjoints et la nécessité de mécanismes rapides de distinction des pannes des nœuds de celles des liens pour diminuer la charge du réseau (placement par PCE).

Pour faire face aux inconvénients précédents, différentes heuristiques réduisant la quantité d'informations envoyées dans les réseau ont été développées. Elles sont décrites ci-après.

2.8 Heuristiques distribuées pour le placement des LSP de secours

Pour pallier ces désavantages et pour diminuer la taille et la fréquence d'envoi des messages transportant les prix de protection, des heuristiques peuvent être envisagées. Pour ce faire, deux stratégies peuvent être adoptées :

1. Agréger (avec ou sans perte) les coûts et/ou les prix de protection.
2. Estimer statistiquement la quantité de bande passante partageable sur un arc ainsi que les prix et (sur)coûts de protection des risques sur les différents arcs de la topologie du réseau.

Avec la première stratégie, au lieu de diffuser tous les prix permettant la protection contre les différents risques de panne sur un arc donné, seul un sous-ensemble de valeurs de ces prix (ensemble agrégé des prix de protection sur un arc) est diffusé par arc (agrégation basée sur les risques de panne). De même, au lieu de diffuser tous les prix de protection permettant la protection contre un risque donné sur tous les arcs de la topologie du réseau, il serait intéressant de ne diffuser les valeurs de ces prix de protection que pour un sous-ensemble d'arcs de la topologie du réseau (agrégation basée sur les arcs). Cela permet de diminuer la fréquence d'envoi puisque les sous-ensembles agrégés changent moins souvent que l'ensemble des prix de protection par arc ou par risque. De plus, la taille de l'information diffusée sera considérablement diminuée.

Dans la deuxième stratégie, il s'agit d'estimer statistiquement les prix et/ou les (sur)coûts de protection en se basant sur des informations incomplètes (valeurs maximales des prix/coûts de protection, moyennes et médianes des prix/coûts de protection, prix/coût de protection des risques qui sont le plus souvent protégés, etc.).

Dans ce qui suit, seules les heuristiques utilisant la première stratégie seront décrites. La deuxième stratégie n'étant pas actuellement explorée.

Comme pour les méthodes exactes de placement des LSP de secours, nous supposons ici que les nœuds connaissent la topologie du réseau et les structures des risques de panne, les LSP de secours auxquels ils appartiennent, les quantités de bande passante associées, les risques de panne protégés ainsi que les chemins supportant ces LSP de secours. Ces informations peuvent être obtenues d'une manière efficace grâce aux protocoles IGP-TE et aux protocoles de signalisation. Ainsi, chaque nœud est capable d'effectuer un contrôle d'admission local et il peut calculer la quantité de bande passante de secours à allouée sur les arcs qui lui sont adjacents en tenant compte du partage.

Deux principales heuristiques sont décrites ci-après. La première utilise uniquement les quantités de bande passante résiduelles sur les arcs pour approximer les prix de protection alors que la seconde combine les quantités de bande passante primaire cumulées et les bandes passantes résiduelles sur les arcs pour mieux approximer les prix de protection.

2.8.1 Heuristique basée sur la bande passante résiduelle

Cette heuristique a l'avantage d'être simple et elle ne nécessite aucune extension aux protocoles IGP-TE ou aux protocoles de signalisation pour son implantation. Elle est basée sur les informations recueillies par chaque nœud qui connaît les prix de protection de tous les risques de panne sur tous les arcs $\alpha \rightarrow \beta$ qui lui sont adjacents (grâce à l'information transmise par les protocoles de signalisation et IGP-TE). Ainsi, au lieu de diffuser les prix de protection de tous les risques sur les arcs adjacents, un nœud α ne diffuse qu'une seule valeur agrégée, correspondant à la bande passante résiduelle $R^{\alpha \rightarrow \beta}$, pour tout arc $\alpha \rightarrow \beta$. Avec cette méthode de placement de LSP de secours, le prix de protection $\delta_r^{\alpha \rightarrow \beta}$ de n'importe quel risque r (ou le coût de protection $\delta_c^{\alpha \rightarrow \beta}$ de n'importe quel lien ou nœud c) sur un arc $\alpha \rightarrow \beta$ est approximé par $C^{\alpha \rightarrow \beta}$ - $R^{\alpha \rightarrow \beta}$ - $F_{\alpha \rightarrow \beta}$ ($C^{\alpha \rightarrow \beta}$ - $R^{\alpha \rightarrow \beta}$ - $F_{\alpha \rightarrow \beta} = Max_{r \in Rs}\, \delta_r^{\alpha \rightarrow \beta} = G^{\alpha \rightarrow \beta}$).

Le choix de la diffusion des quantités de bande passante résiduelle au lieu des quantités de bande passante de secours allouées sur les arcs est dicté par des raisons de facilité de déploiement et d'efficacité. En effet, il existe actuellement plusieurs protocoles IGP-TE (OSPF-TE et ISIS-TE) diffusant d'une manière efficace les quantités de bande passante résiduelles ainsi que les capacités et les bandes passantes primaires cumulées) alors que les quantités de bande passante de secours allouées sur les arcs ne sont pas encore prises en compte.

En remplaçant la valeur du coût de protection dans la formule (2.6), nous déduisons qu'un arc $\alpha \rightarrow \beta$ peut être utilisé pour établir un LSP de secours b si et seulement si :

$$G^{\alpha \rightarrow \beta} + F_{\alpha \rightarrow \beta} + bw\,(b) \leq C^{\alpha \rightarrow \beta}$$

Et comme :

$$C^{\alpha \rightarrow \beta} - G^{\alpha \rightarrow \beta} = R^{\alpha \rightarrow \beta}$$

Nous déduisons que seuls les arcs $\alpha \rightarrow \beta$ vérifiant la condition suivante peuvent être utilisés dans le calcul d'un nouveau LSP de secours b :

$$bw\,(b) \leq R^{\alpha \rightarrow \beta} \tag{2.15}$$

Dans cette heuristique, le partage de bande passante de secours n'est effectué que lors de la configuration des LSP de secours. Les PLR n'ont pas connaissance des possibilités de partage de la bande passante entre le LSP de secours en cours de calcul avec ceux déjà calculés. Ils risquent ainsi de rejeter des requêtes de protection susceptibles d'être satisfaites (grâce au partage de la bande passante) uniquement sur la base de la bande passante résiduelle. La probabilité de blocage[11] est donc élevée lorsque la bande passante

[11] Une requête de protection est dite *bloquée* lorsqu'elle ne peut pas être satisfaite.

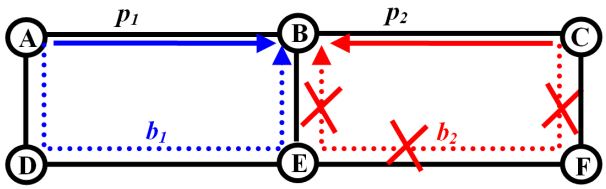


FIG. 2.8 – Blocage par erreur lors placement du LSP de secours b_2

résiduelle est petite bien qu'il y ait suffisamment de bande passante disponible et de possibilité de partage à ce moment.

Sur la figure 2.8 où tous les arcs disposent d'une capacité de bande passante égale à 10 unités ($\forall\lambda : C^{\lambda} = 10$), deux LSP primaires p_1 ($A \rightarrow B$) et p_2 ($C \rightarrow B$) de bande passante égale à 10 unités ont été établis. Pour protéger le LSP primaire p_1 contre la panne du lien A-B, le LSP de secours b_1 ($A \rightarrow D \rightarrow E \rightarrow B$) de bande égale à 10 unités a été configuré. Cependant et bien que la condition suffisante pour établir un autre LSP de secours b_2 de quantité de bande passante égale 10 unités et protégeant p_2 contre la panne du lien B-C, soit vérifiée sur le chemin $C \rightarrow F \rightarrow E \rightarrow B$, l'heuristique ne permet pas de le déterminer car la bande passante résiduelle sur l'arc $E \rightarrow B$ est plus petite que la quantité de bande passante réclamée par le LSP de secours b_2 (i.e. $10 > R^{E \rightarrow B} = 0$).

2.8.2 Heuristique de Kini et al.

Dans le cas de la panne d'un risque donné r, seul le trafic le traversant sera redirigé vers les LSP de secours. En conséquence, un arc appartenant à un LSP de secours protégeant le risque r peut au plus recevoir une quantité de trafic égale à celle traversant le risque r. Pour diminuer la probabilité de blocage, [KKL+01, KL01] proposent d'approcher le prix de protection $\delta_r^{\alpha \rightarrow \beta}$ de chaque risque r sur l'arc $\alpha \rightarrow \beta$ par le minimum entre F_r[12] et $G^{\alpha \rightarrow \beta}$.

En conséquence, un arc $\alpha \rightarrow \beta$ peut être utilisé pour établir un nouveau LSP de secours b si et seulement si :

$$Min\,(F_r, G^{\alpha \rightarrow \beta}) + F_{\alpha \rightarrow \beta} + bw\,(b) \leq C^{\alpha \rightarrow \beta} \tag{2.16}$$

Pour déterminer les surcoûts des LSP de secours établis en ligne, il suffira de remplacer la quantité $\delta_r^{\alpha \rightarrow \beta}$ par $Min\ (F_r,\ G^{\alpha \rightarrow \beta})$ dans la formule (2.3) et d'appliquer ensuite la formule (2.7). Ainsi, nous avons :

$$\gamma_c^{\lambda}\,(bw) = \begin{cases} \infty \ si\ Min\,(F_r, G^{\alpha \rightarrow \beta}) + bw + F_{\lambda} > C^{\lambda} \\ Min\,(F_r, G^{\alpha \rightarrow \beta}) + bw - G^{\lambda}\ sinon \end{cases} \tag{2.17}$$

Cette heuristique permet une meilleure prise en charge de la topologie du réseau et un meilleur partage par rapport à l'heuristique précédente. En effet, l'ensemble des arcs

[12] F_r correspond à la bande passante cumulée de tous les LSP primaires traversant un composant du risque r.

utilisables par un PLR pour construire un LSP de secours est étendu aux arcs vérifiant les inégalités $R^{\alpha\rightarrow\beta} < b \leq R^{\alpha\rightarrow\beta} + G^{\alpha\rightarrow\beta}$ - F_r. De plus, cette heuristique permet une meilleure approximation de la quantité de bande passante de secours additionnelle par rapport à l'heuristique précédente (cf. formule (2.17)). Cependant, cette heuristique présente divers inconvénients comme :

- Performances médiocres et proches de celles de l'heuristique basée sur la bande passante résiduelle, surtout dans le cas de risques de type nœud et SRLG. En effet, pour un risque r de type nœud ou SRLG, F_r est souvent supérieur à $G^{\alpha\rightarrow\beta}$ sur tout arc $\alpha\rightarrow\beta$.
- Le partage est sous-optimal et la probabilité de blocage reste assez élevée car les nœuds (PLR) tendent à surestimer les valeurs des prix et coûts de protection.
- nécessité d'extension des protocoles IGP-TE pour la diffusion des bandes passantes de secours allouées sur les arcs.

2.8.2.1 Heuristique de Kini et al. améliorée (HKA)

Pour contrôler la quantité de bande passante dédiée à la protection, nous avons légèrement modifié l'heuristique de Kini et al., en partitionnant la capacité de chaque arc en deux pools disjoints : le pool primaire et le pool de secours (comme dans la section 2.7.2). De cette manière, un arc $\alpha\rightarrow\beta$ ne peut être utilisé pour établir un nouveau LPS de secours b que si :

$$Min\,(F_r, G^{\alpha\rightarrow\beta}) + bw\,(b) \leq BC^{\alpha\rightarrow\beta}$$

2.9 Types de panne, bande passante de secours et délai de récupération

Afin de diminuer la quantité de bande passante de secours allouée sur les arcs, les nœuds PLR doivent être en mesure de distinguer les pannes des nœuds de celles des liens. En effet, avec l'utilisation de LSP de secours NHOP (les LSP de secours NNHOP ne posent aucun problème puisqu'ils protègent simultanément contre les pannes des prochains nœuds et de leurs liens en aval), un nœud PLR doit distinguer la panne d'un lien de celle d'un nœud afin de décider de l'activation ou pas des LSP de secours NHOP. Concrètement, un nœud PLR détectant une panne d'un lien doit impérativement activer tous ses LSP de secours NHOP pour récupérer de la panne alors que lors d'une panne d'un nœud, aucun LSP de secours NHOP ne doit être activé afin d'éviter la violation des contraintes de bande passante.

Bien qu'il existe aujourd'hui différents mécanismes permettant la distinction des types de pannes [VC02, AKNS08], leur utilisation induit une augmentation conséquente du délai de récupération. Un compromis entre la quantité de bande passante de secours allouée sur les arcs et les délais de récupération doit être établis et défini. Deux alternatives assurant le respect des contraintes de bande passante se présentent :

Récupération rapide

Dans cette alternative, le calcul des LSP NHOP est coordonné avec le calcul des

LSP de type NNHOP. Toute panne est ainsi considérée comme une panne d'un lien, ce qui induit l'activation de tous les LSP de secours (NHOP et NNHOP) sur chaque nœud PLR détectant la panne.

Cette solution minimise le délai de récupération puisque le trafic des communications affectées par une panne est instantanément redirigé vers les LSP de secours, dès la détection de la panne. Par contre, elle induit une surconsommation de la bande passante puisque lors de la détection d'une panne correspondant à la défaillance d'un nœud, un PLR active inutilement tous ses LSP de secours de type NHOP (ce qui consomme inutilement la bande passante).

Diminution de la consommation de bande par distinction des pannes
Avec cette alternative, un mécanisme de distinction des pannes des liens de celles des nœuds est nécessaire. Ce mécanisme indique le type de la panne survenue aux différents PLR, ce qui leur permettra d'activer uniquement les LSP NNHOP (resp. les LSP NHOP et NNHOP) dans le cas d'une panne d'un nœud (resp. dans le cas d'une panne d'un lien).

Dans une implantation, le choix de l'une ou de l'autre alternative dépend essentiellement des contraintes de temps des applications supportées et des ressources (capacités des arcs) disponibles. Avec l'apparition, ces dernières années, du mécanisme BFD (*Bidirectional Forwarding Detection*) [AKNS08] qui permet une distinction assez rapide des pannes des nœuds de celles des liens, la tendance est d'adopter la seconde alternative.

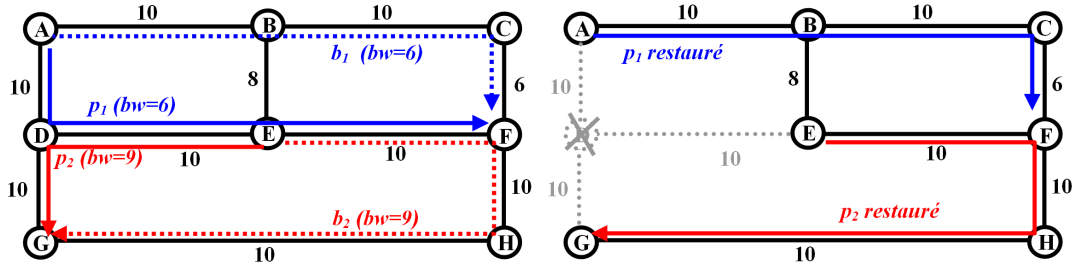
2.10 Extension du partage aux LSP primaires

De la même manière qu'il est possible de partager la bande passante entre les LSP de secours (qui ne pourraient être actifs au même moment), il est envisageable de partager la bande passante entre les LSP primaires et les LSP de secours. En effet, lors d'une panne touchant un LSP primaire protégé, la bande passante allouée sur la partie du LSP primaire située entre le composant en panne et le MP peut être utilisée par les autres LSP de secours protégeant le même composant. De plus, lors d'une panne d'un nœud, tous les LSP primaires commençant ou finissant en ce nœud sont interrompus définitivement puisqu'il n'est pas possible de protéger le nœud source ou le nœud de destination d'une communication avec un LSP de secours.

Dans les deux sous-sections suivantes, nous expliquons les principes du partage de la bande passante entre les LSP primaires et les LSP de secours.

2.10.1 Partage global de la bande passante

Lors d'une panne affectant un LSP primaire, toute la partie de ce LSP allant du composant défaillant jusqu'au MP du LSP de secours permettant la récupération (ou jusqu'à la destination du LSP primaire si le composant défaillant n'est pas protégé) ne sera plus utilisée. Pour une meilleure disponibilité de la bande passante, [MBL03, BML06] proposent de récupérer *la bande passante libérée* sur cette partie du LSP primaire et de l'allouer aux LSP de secours qui seront activés suite à la panne.



(a) Participation de l'arc $E \rightarrow F$ à l'établissement du LSP de secours b_2 grâce au partage de bande entre les LSP primaires et les LSP de secours

(b) Pas de violation des contraintes de bande passante après la récupération de la panne du nœud D

FIG. 2.9 – Partage de la bande passante entre le segment $E \rightarrow F$ du LSP primaire p_1 et le LSP de secours b_2

Sur la figure 2.9 (a), deux LSP primaires p_1 ($A \rightarrow D \rightarrow E \rightarrow F$) et p_2 ($E \rightarrow D \rightarrow G$) sont établis et protégés par deux LSP de secours b_1 ($A \rightarrow B \rightarrow C \rightarrow F$) et b_2 ($E \rightarrow F \rightarrow H \rightarrow G$) contre la panne du nœud D. Les LSP p_1 etb_1 réclament une quantité de bande passante égale à 6 unités alors que les LSP p_2 etb_2 réclament une quantité de bande passante égale à 9 unités. Tous les arcs de la topologie du réseau sont de capacité égale à 10 unités sauf les arcs $E \rightarrow B$ et $B \rightarrow E$ qui n'en disposent que de 8 et les arcs $C \rightarrow F$ et $F \rightarrow C$ qui n'en disposent que de 6.

Sans partage de la bande passante entre les LSP primaires et les LSP de secours, le calcul du LSP de secours b_2 n'aurait pas pu aboutir puisqu'il n'existe pas de chemin alternatif disposant d'une quantité de bande passante suffisante (les LSP p_1, p_2 et b_1 sont supposés configurés avant le placement du LSP b_2). Typiquement, tous les algorithmes et heuristiques présentés dans les sections précédentes ne pourront pas déterminer un LSP de secours permettant de protéger le LSP primaire p_2 contre la panne du nœud D. En effet, l'arc $E \rightarrow B$ ne peut pas être utilisé pour le calcul de b_2 puisque sa capacité est inférieure à la bande passante réclamé par b_2. De même, l'arc $E \rightarrow F$ ne peut non plus appartenir au LSP de secours b_2 car $\delta_D^{E \rightarrow F} + F_{E \rightarrow F} + bw\ (b_2) = 0 + 6 + 9 > C^{E \rightarrow F} = 10$.

En appliquant par contre le partage de bande passante entre les LSP primaires et les LSP de secours, nous constatons que la partie du LSP primaire p_1 allant du nœud A (LSR de tête du LSP de secours protégeant p_1 contre la panne du nœud D) jusqu'au nœud F (LSR de sortie du LSP de secours protégeant p_1 contre la panne du nœud D) sera remplacée par le LSP de secours b_1 (figure 2.9 (b)) lors de la panne du nœud D. En conséquence, la bande passante allouée sur cette partie du LSP primaire sera de nouveau libre après la panne du nœud D. Pour augmenter la disponibilité de la bande passante et pour éviter les situations de blocage, la bande passante libérée sur les arcs du LSP primaire p_1 peut être réattribuée à n'importe quel LSP de secours activable après la panne du nœud D. Sur l'exemple de la figure 2.9 (a), le LSP primaire p_1 libérera une quantité de bande égale à 6 unités sur l'arc $E \rightarrow F$ suite à la panne du nœud D. Cette quantité de bande passante pourrait donc être exploitée et réservée au nouveau LSP de

secours b_2 qui ne pourra s'en servir effectivement qu'après la panne du nœud D (voir figure 2.9 (b)).

Pour optimiser la quantité de bande passante de secours additionnelle ou toute autre métrique en respectant les contraintes de bande passante, il suffira de modifier la formule calculant les prix de protection (formule (2.1)) afin qu'elle tienne compte de la bande passante primaire libérée sur les arcs. Formellement, si l'on note par L_r^λ la quantité de bande passante primaire cumulée et libérée sur un arc λ suite à la panne du risque r, le prix de protection (*effectif*) $(\delta*)_r^\lambda$ du risque r sur l'arc λ sera déterminé comme suit :

$$(\delta*)_r^\lambda = \sum_{b \setminus \lambda \in b \wedge r \in PFRG(b)} bw(b) - L_r^\lambda \tag{2.18}$$

De même, la quantité de bande passante de secours $(G*)^\lambda$ (*effectivement*) allouée sur un arc λ sera calculée comme suit :

$$(G*)^\lambda = Max_r\left((\delta*)^\lambda_{r \setminus r \text{ est un risque}}, 0\right) = Max_c\left((\theta*)^\lambda_{c \setminus c \text{ est un lien ou un nœud}}, 0\right) \tag{2.19}$$

Ce type de partage de la bande passante peut être appliqué à tous les algorithmes et heuristiques de placement des LSP de secours vus dans les sections précédentes, à condition d'appliquer la formule (2.18) pour le calcul des prix de protection.

Comme pour le cas du partage de la bande passante restreint au LSP de secours (sections précédentes), le choix d'un LSP de secours protégeant un risque donné est important pour agencer (et augmenter ainsi) la quantité de bande passante libérée. Sur la figure 2.9 (a) par exemple, deux chemins $A{\rightarrow}B{\rightarrow}C{\rightarrow}F$ et $A{\rightarrow}B{\rightarrow}E$ permettant de protéger le LSP primaire p_1 contre la panne du nœud D peuvent être choisis pour supporter le LSP de secours b_1. Le premier chemin (sélectionné pour l'exemple de la figure 2.9 (a)) permet de libérer une quantité de bande passante égale à 6 unités sur l'arc $E{\rightarrow}F$ dans le cas de la panne du nœud D, ce qui permet d'établir ensuite le LSP de secours b_2, alors que le second ne libère aucune ressource suite à la panne du nœud D (ce qui induit un échec lors du placement du LSP b_2). En conséquence, pour augmenter la disponibilité de la bande passante, le calcul du surcoût de protection (*effectif*) d'un risque r sur un arc λ doit tenir compte non seulement de la bande passante primaire libérée sur l'arc λ (cf. formule (2.18)) mais aussi de la bande passante primaire libérée sur le risque r :

$$(\gamma*)^\lambda(b) = \begin{cases} \infty \text{ si } (\theta*)^\lambda(b) + bw(b) + F_\lambda > C^\lambda \\ Max\left((\theta*)^\lambda(b) + bw(b) - (G*)^\lambda - d_\lambda.\epsilon, 0\right) \text{ sinon} \end{cases} \tag{2.20}$$

où :

$$d_\lambda = \begin{cases} 0 \text{ si } \lambda \text{ n'est pas adjacent à un nœud en aval du lien protégé sur le LSP primaire} \\ \text{distance en nombre de sauts entre le PLR et l'extrémité primaire de } \lambda \text{ sinon} \end{cases}$$

Bien que cette technique de partage permette d'augmenter la disponibilité de la bande passante, elle présente beaucoup de désavantages comme :

- L'ordre d'établissement des LSP n'est pas symétrique. Pour l'exemple de la figure 2.9 (a), l'établissement des LSP p_2, p_1, b_2 puis b_1 dans cet ordre permet toujours de protéger les LSP primaires p_1 et p_2 contre la panne du nœud D (ce qui n'est pas le cas si l'ordre d'établissement des LSP est p_1, p_2, b_1 puis b_2) . Dans la pratique, une telle asymétrie peut générer des échecs lors de la phase de normalisation exécutée après la réparation du composant défaillant (cf. section 1.4.3 du chapitre 1).
- Le recalcul et/ou la reconfiguration d'un LSP de secours (dû à une augmentation de la quantité de bande passante de son LSP primaire par exemple) peut entraîner des recalculs en cascade. Sur la figure 2.9 par exemple, l'augmentation de la bande passante du LSP primaire p_1 de 2 unités provoquera le recalcul du LSP de secours b_1 puisque l'arc $C \rightarrow F$ ne dispose pas d'assez d'unités de bande passante pour supporter ce LSP de secours. Le PLR A cherchera alors à déterminer un nouveau LSP de secours b'_1 permettant de protéger p_1 contre la panne du nœud D et déterminera le chemin $A \rightarrow B \rightarrow E$. Ce chemin ne permet pas de libérer de la bande passante sur l'arc $C \rightarrow F$ suite à la panne du nœud D, ce qui provoquera le recalcul du LSP de secours b_2.
- Lorsque les valeurs des prix de protection, des quantités de bande passante primaire libérées et des quantités de bande passante de secours allouées sur les arcs ne sont pas diffusées dans le réseau, les LSR ne pourront pas préempter efficacement les LSP puisqu'ils ne sont pas capables de déterminer la quantité de bande passante récupérée suite à la pré-emption d'un LSP.
- La surcharge du réseau avec la transmission de nouveaux paramètres de bande passante (les quantités cumulées de bande passante primaire libérées suite à une panne) et la nécessité de définir de nouveaux protocoles (ou d'étendre les protocoles existants) pour la distribution de ces paramètres aux entités de calcul.

Nous notons que les performances de ce dernier type de partage sont étudiées et comparées à celles du partage de la bande passante restreint aux LSP de secours dans le chapitre 6.

2.10.2 Partage étendu de la bande passante

Comme il n'est jamais possible de protéger les LSR frontières (LSR d'entrée ou LSR de sortie) d'un LSP primaire en utilisant des LSP de secours de type NHOP ou NNHOP, la panne d'un de ces LSR coupe définitivement la connexion supportée par le LSP primaire correspondant. [ARUK06] propose de récupérer la bande passante des LSP primaires interrompus suite à la panne de leur LSR d'entrée ou LSR de sortie[13] pour la réattribuer aux LSP de secours protégeant contre la panne de tels LSR.

Cette technique de partage de bande passante n'introduit aucune nouveauté par

[13] En réalité, seule la bande passante libérée par les LSP primaires affectés à la source peut être exploitée et réattribuée aux LSP de secours qui protègent contre la panne de ce nœud source. En effet, dans une connexion affectée à la destination, le trafic continue à circuler jusqu'à ce que la notification de la panne atteigne le nœud source. Ceci peut provoquer la violation des contraintes de la bande passante si les LSP de secours locaux sont activés avant que la source ne supprime la connexion affectée.

rapport à la technique de partage décrite dans la section précédente (elle correspond à un cas spécial de la technique de partage global de la bande passante décrite dans la section précédente). Cependant et contrairement à la technique de partage global de la bande passante, cette technique de partage ne requiert aucune extension ou modification des protocoles existants pour son implantation. En effet, les quantités de bande passante des LSP primaires interrompus suite à une panne d'un LSR d'entrée ou d'un LSR de sortie peut être obtenue sur chaque PLR (et donc tout nœud du LSP primaire) en consultant les messages des protocoles de signalisation.

2.11 Conclusion

Le partage de la bande passante entre les LSP de secours sous MPLS est très important puisqu'il permet d'augmenter la disponibilité de la bande passante et donc, d'établir plus de LSP dans un réseau. Différentes méthodes distribuées de placement de LSP de secours tenant compte du partage de la bande passante ont été élaborées. Ces méthodes peuvent être groupées dans deux classes : les méthodes exactes et les méthodes heuristiques.

Avec les méthodes exactes, les fuites (ou pertes) de bande passante de secours sont optimisées puisque la quantité de bande passante de secours allouée sur chaque arc est minimisée. Le placement des LSP de secours pourrait ainsi être effectué de manière à optimiser la bande passante additionnelle ou toute autre métrique tout en assurant un partage optimal de la bande passante. Malgré leur efficacité dans la réduction de la quantité de bande passante allouée aux LSP de secours, ces méthodes exactes ne sont pas pratiques. En effet, elles conduisent souvent à une surcharge du processus de commande du réseau avec la diffusion de messages transportant l'information sur les LSP de secours déjà établis et/ou elles induisent la quasi-centralisation des calculs dans le cas de SRLG non disjoints.

Pour pallier ces désavantages, diverses heuristiques ont été mises en œuvre. Ces dernières permettent de diminuer considérablement la taille de l'information diffusée dans le réseau au détriment d'un partage moins efficace. Deux stratégies peuvent être utilisées à cet effet :

- agrégation avec perte d'information (avant la diffusion),
- estimation des quantités de bande passante partageables en adoptant des approches probabilistes.

Actuellement, seule la première stratégie est explorée. Ainsi, différentes heuristiques de placement de LSP de secours (heuristique basée sur la bande passante résiduelle et heuristique de Kini et al.) tenant compte du partage de la bande passante ont été développées. Bien que ces heuristiques résolvent le problème de la surcharge du réseau, leur application dans la pratique reste limitée à cause de l'inefficacité des approches d'agrégation adoptées (perte d'une quantité élevée d'informations lors de l'agrégation, ce qui détériore les possibilités de partage de la bande passante).

Pour améliorer le partage de la bande passante, nous pensons qu'il est judicieux de combiner les deux stratégies précédentes, c'est pourquoi nous comptons explorer la

seconde stratégie dans les chapitres suivants. Ainsi, il serait intéressant d'annoncer les $n_{(n>1)}$ plus grandes valeurs des prix de protection au lieu d'une seule (exemple : les cinq prix de protection les plus élevés par arc), les seuils de bande passante, les moyennes des prix de protection, leur variance, etc. Ces paramètres peuvent n'être diffusés que pour des risques particuliers (exemple : risques de type SRLG) et vérifiant certaines conditions (exemple : risques induisant des échanges de messages pour leur protection).

Sur un autre volet et afin d'améliorer la disponibilité de la bande passante dans le réseau, d'autres techniques de placement des LSP de secours, qui étendent le partage de la bande passante entre les LSP primaires et les LSP de secours, peuvent être adoptées. Deux classes de ces techniques sont présentées dans ce chapitre : le partage global de la bande passante et le partage étendu de la bande passante. Alors que les techniques de la première classe permettent de récupérer la bande passante allouée sur les arcs du LSP primaire suite à n'importe quel type de panne, le techniques de la seconde classe n'exploitent que la bande passante libérée suite à la panne des nœuds d'extrémité des LSP primaires. Bien qu'elles puissent facilement être combinées avec les techniques de placement des LSP de secours qui restreignent le partage de la bande aux LSP de secours, ces techniques peuvent entraîner de nouveaux problèmes comme : (1) la reconfiguration en cascade des LSP suite à un changement ou à un recalcul d'un LSP de secours, (2) la complexification des mécanismes de placement des LSP et (3) l'augmentation de la charge du réseau avec la distribution de nouveaux paramètres de la bande passante (sans compter que cette distribution requiert de nouvelles extensions aux protocoles existants).

Dans le prochain chapitre, nous présenterons un algorithme et deux heuristiques permettant de placer efficacement les LSP de secours dans un environnement distribué. Ces algorithme et heuristiques traitent tous les risques de pannes et réduisent significativement la quantité d'informations distribuées dans le réseau tout en diminuant les taux de rejet des LSP de secours.

Chapitre 3

Algorithmes et heuristiques améliorant le placement local des LSP de secours

3.1 Introduction

Après avoir donné un large aperçu des techniques de placement des LSP de secours existantes (cf. chapitre précédent), nous proposons dans ce chapitre de nouveaux algorithme et heuristiques permettant d'améliorer, d'éviter ou de résoudre les différents problèmes auxquels sont confrontées les techniques de placement de LSP de secours existantes (essentiellement la quantité élevée d'informations transmises dans le réseau ou le taux élevé des requêtes de placement de LSP de secours bloquées).

Comme pour le chapitre précédent, nous ne traiterons ici que les cas de pannes simples. De plus, nous assumerons que chaque nœud de la topologie du réseau est responsable de la gestion des ressources (typiquement du contrôle d'admission) sur tous les liens qui lui sont adjacents (cf. section 2.6.4 du chapitre 2). De cette manière, tout nœud disposera de toutes les valeurs des prix de protection sur ses arcs adjacents, ce qui lui permet de les annoncer, au besoin, aux autres nœuds du réseau.

La suite de ce chapitre sera consacrée à la description des algorithme et heuristiques distribués que nous proposons pour le placement en ligne des LSP de secours. Dans ce chapitre, nous nous concentrons sur l'optimisation du délai tout en assurant le respect des contraintes de bande passante. Le comportement des nœuds du réseau ainsi que les extensions à apporter aux protocoles IGP-TE et aux protocoles de signalisation (plus spécialement au protocole RSVP-TE) pour implanter ces algorithme et heuristiques seront détaillés dans les différentes sections qui suivent. Concrètement, nous présenterons en section 3.2 l'algorithme de distribution ciblée des prix de protection (*Targeted Distribution of Resource Allocations* ou TDRA) qui cible les nœuds auxquels il envoie l'information requise pour le placement des LSP de secours. Avec cet algorithme, seuls les nœuds susceptibles de protéger contre la panne d'un risque reçoivent l'information requise à sa protection. Comme nous allons le voir, cet algorithme permet une excellente

prise en charge des possibilités de partage puisqu'il est capable de déterminer tous les arcs pouvant supporter un LSP de secours en cours de calcul. L'algorithme TDRA nécessite néanmoins d'apporter certaines extensions aux protocoles de signalisation pour partager les valeurs des prix de protection des SRLG. Il est plutôt destiné aux larges réseaux contenant un petit nombre de SRLG qui sont de taille limitée. En section 3.3, nous présentons l'heuristique de partage efficace et distribué de la bande passante (*Distributed Bandwidth Sharing Heuristic* ou DBSH) qui est basée sur la diffusion, pour chaque arc du réseau, d'un vecteur réduit de prix de protection de SRLG. Cette heuristique partage la charge de calcul des LSP de secours équitablement sur les nœuds et diminue considérablement la quantité d'informations diffusées dans le réseau. Pour son implantation, elle nécessite de légères extensions aux protocoles IGP-TE et de signalisation. Dans la section 3.4, nous présentons l'heuristique de placement de LSP de secours basés sur les PLR (*PRH-based Heuristic* ou PLRH) qui ne diffuse, pour un arc donné, qu'un nombre limité de prix de protection (les plus élevés). Cette heuristique permet de coordonner le calcul des LSP de secours avec le LSP primaire, ce qui améliore le taux de protection et facilite le calcul de LSP de secours multi-domaines. De plus, c'est une heuristique qui ne nécessite que de très légères extensions aux protocoles IGP-TE pour son implantation.

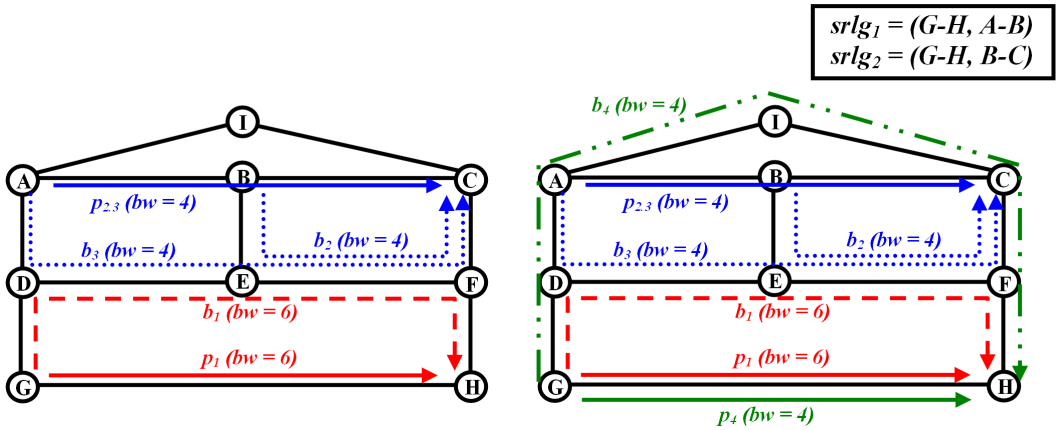
3.2 Algorithme de la distribution ciblée des prix de protection (TDRA)

Avec l'algorithme TDRA (*Targeted Distribution of Resource Allocations*) [SCLR08], nous associons à tout nœud du réseau une entité de calcul des chemins de secours (*Backup Path Computation Element* ou BPCE). Cette entité sera responsable du calcul des LSP de secours protégeant contre les pannes du nœud la supportant et de tous les liens qui lui sont adjacents. Par exemple, l'entité $BPCE_\alpha$ tournant sur le nœud α, dont les nœuds voisins appartiennent à l'ensemble $voisins\ (\alpha)$, se charge du calcul des LSP de secours protégeant contre les pannes du nœud α et de tous les liens unidirectionnels[1] $(\beta{\rightarrow}\alpha)_{\beta \in voisins\,(\alpha)}$.

Comme nous l'avons expliqué dans la section 2.7.3 du deuxième chapitre, ce type de distribution des entités de calcul sur les nœuds du réseau permet de protéger contre les pannes des nœuds et des liens *indépendants*[2] sans aucun échange de messages entre les différentes entités de calcul des LSP de secours et sans restreindre l'ensemble des risques protégés par un LSP NNHOP à un seul nœud. Par contre, lorsqu'un lien à protéger β-α apparaît dans au moins un SRLG, les entités $BPCE_\alpha$ ne seront pas en mesure de déterminer localement (sans échange d'information avec les autres $BPCE$) et efficacement tous les LSP de secours (particulièrement le LSP NNHOP) permettant de

[1] Par abus de langage, nous définissons un LSP de secours protégeant contre la panne du lien unidirectionnel $\beta{\rightarrow}\alpha$ comme un LSP de secours protégeant contre la panne du lien α-β et dont le LSR d'entrée doit être le nœud β. Par contre, nous désignerons un LSP de secours protégeant contre la panne du lien α-β comme un LSP dont le LSR d'entrée est α ou β.

[2] Un lien est dit *indépendant* s'il n'apparaît dans aucun SRLG.



(a) Protection de deux LSP primaires avec respect des contraintes de bande passante

(b) Ajout d'un LSP primaire et d'un LSP de secours le protégeant

FIG. 3.1 – Placement des LSP de secours avec l'algorithme TDRA

protéger contre la panne de ce lien. En effet, l'entité $BPCE_\alpha$ ne connaît que les prix de protection des liens qui sont adjacents au nœud α. Or, pour le calcul des LSP de secours protégeant contre la panne d'un lien apparaissant dans un (ou plusieurs) SRLG, il est nécessaire de connaître les prix de protection de tous les liens composant ce(s) SRLG (cf. section 2.4 du deuxième chapitre). Pour tenir compte des risques de type SRLG lors du calcul des LSP de secours, nous proposons ici l'algorithme TDRA qui partage les structures et les propriétés des LSP de secours protégeant contre la panne d'un lien avec tous les nœuds d'extrémité des liens partageant au moins un même SRLG avec le lien protégé. Par exemple, après le placement d'un LSP de secours b protégeant contre la panne du lien $\beta \rightarrow \alpha$ qui appartient à deux SRLG $srlg_1$ et $srlg_2$ ($srlg_1$=(α-β, α-ϑ) et $srlg_2$=(α-β, μ-ν)), l'entité $BPCE_\alpha$, exécutant l'algorithme TDRA, envoie aux nœuds d'extrémité (β, ϑ, μ et ν) des liens appartenant aux SRLG $srlg_1$ et $srlg_2$ les structures et les propriétés (type et quantité de bande passante) du LSP b.

	$A \rightarrow D$	$B \rightarrow E$	$D \rightarrow E$	$E \rightarrow F$	$F \rightarrow C$	$G \rightarrow D$	$F \rightarrow H$	$A \rightarrow B$	$B \rightarrow C$	$G \rightarrow H$	autres
δ_H	0	0	0	0	0	0	–	0	0	–	0
δ_{G-H}	0	0	6	6	0	6	6	–	–	–	0
δ_{F-H}	0	0	0	0	0	0	–	0	0	0	0
δ_{A-B}	4	0	4	4	4	0	0	–	0	–	0
δ_{B-C}	0	4	0	4	4	0	0	0	–	–	0

TAB. 3.1 – Prix de protection sur un nœud H

3.2.1 Description de l'algorithme TDRA

Afin de faciliter la compréhension de l'algorithme TDRA, nous expliquons ses principes par un exemple. Sur la figure 3.1 (a), deux LSP de secours b_1 ($G \rightarrow D \rightarrow E \rightarrow F \rightarrow H$) et b_2 ($B \rightarrow E \rightarrow F \rightarrow C$) de type NHOP et un LSP de secours b_3 ($A \rightarrow D \rightarrow E \rightarrow F \rightarrow C$) de type NNHOP ont été établis pour protéger les deux LSP primaires p_1 ($G \rightarrow H$) et $p_{2.3}$ ($A \rightarrow B \rightarrow C$). Concrètement, b_1 (bw (b_1) = 6 unités) protège le LSP primaire p_1 contre la panne du lien G-H, b_2 (bw (b_2) = 4 unités) protège le LSP primaire $p_{2.3}$ contre la panne du lien B-C et b_3 (bw (b_3) = 4 unités) protège le LSP primaire $p_{2.3}$ contre les pannes du nœud B et du lien A-B. Comme nous l'avons explicité plus haut, chaque

---

Algorithme 1 Calcul d'un LSP de secours b vérifiant les contraintes de bande passante sur un graphe $G = (V, E)$

1.
$E' \leftarrow E$
pour chaque arc $\lambda \in E$ **faire**
  $coût\_protection\_LSP\_secours^{\lambda} \leftarrow 0$
  pour chaque risque $r \in$ ***PFRG*** *(b)* **faire**
    $\delta_r^{\lambda} \leftarrow$ ***deduire_prix_protection*** (λ, r) {déduire le prix de protection du risque r sur l'arc λ à partir de la table des prix de protection}
    si $coût\_protection\_LSP\_secours^{\lambda} < \delta_r^{\lambda}$ **alors**
      $coût\_protection\_LSP\_secours^{\lambda} \leftarrow \delta_r^{\lambda}$
    finsi
  fin pour
  si $coût\_protection\_LSP\_secours^{\lambda} +$ ***bw*** $(b) + F_{\lambda} > C^{\lambda}$ **alors**
    $E' \leftarrow E' \setminus \{\lambda\}$
  finsi
fin pour
2.
$b =$ ***calculer_LSP_secours***(***PLR*** *(b)*, $G'(V, E')$, ***LSP_primaire*** *(b)*) {***PLR*** *(b)* retourne le nœud source de b et ***LSP_primaire*** *(b)* retourne le LSP primaire protégé par b}
3.
mettre_à_jour_prix_protection (b, ***bw*** *(b)*) {augmenter le prix de protection du lien et du nœud protégés par b de ***bw*** *(b)* unités}
4.
$N \leftarrow$ ***nœuds_extrémités_liens_SRLG*** (***PFRG*** *(b)*) {déterminer l'ensemble des nœuds d'extrémité de liens appartenant aux SRLG de ***PFRG*** *(b)*}
envoyer_structure_et_propriétés *(b, N)* {envoyer la structure, le lien et le nœud protégés par b ainsi que la bande passante allouée à b à tous les nœuds appartenant à l'ensemble N}
5.
retourner b

	$A \rightarrow D$	$B \rightarrow E$	$D \rightarrow E$	$E \rightarrow F$	$F \rightarrow C$	$G \rightarrow D$	$F \rightarrow H$	$A \rightarrow B$	$B \rightarrow C$	$G \rightarrow H$	autres
δ_{G-H}	0	0	6	6	0	6	6	–	–	–	0
δ_{srlg_1}	4	0	10	10	4	6	6	–	–	–	0
δ_{srlg_2}	0	4	6	10	4	6	6	–	–	–	0
$\theta^{\cdot}(b_4)$	4	4	10	10	4	6	6	–	–	–	0
$BC. - bw$	6	6	6	6	6	6	6	6	6	6	6

TAB. 3.2 – Détermination des arcs vérifiant les contraintes de bande passante

nœud α du réseau supporte une entité $BPCE_\alpha$ dont le rôle est de calculer les LSP de secours protégeant contre les pannes du nœud α et de tous les liens qui sont adjacents. De ce fait, nous concluons que le LSP de secours b_1 a été calculé par $BPCE_H$, b_2 est déterminé par $BPCE_C$ et enfin b_3 est calculé par $BPCE_B$.

Pour assurer le respect des contraintes de bande passante (tout en optimisant une métrique donnée), chaque $BPCE_\alpha$ construira une table contenant les prix de protection des risques (nœud et liens) dont la panne activera les LSP de secours qu'il a calculés

Algorithme 2 Mise-à-jour des prix de protection lors de la réception de la structure et des propriétés d'un nouveau LSP de secours b

1.
$message \leftarrow$ ***attendre_ réception_ message*** ( ) {le processus s'endort jusqu'à la réception d'un message contenant la structure et les propriétés d'un LSP de secours créé ou supprimé}
2.
$arc\_LSP \leftarrow$ ***déduire_ arc_ LSP_ secours*** ($message$) {déduire à partir du message reçu les arcs formant le LSP de secours reçu}
$l \leftarrow$ ***déduire_ lien_ protégé*** ($message$) {déduire le lien protégé par le LSP de secours contenu dans le message reçu}
$BW \leftarrow$ ***déduire_ bande_ passante*** ($message$) {déduire la bande passante du LSP de secours contenu dans le message reçu}
si ***création_ LSP_ secours*** ($message$) **alors**
  $op \leftarrow 1$ {c'est une création d'un LSP de secours}
sinon
  $op \leftarrow (-1)$ {c'est une suppression d'un LSP de secours}
finsi
3.
pour chaque arc $\lambda \in arc\_LSP$ **faire**
  $\delta_l^\lambda \leftarrow \delta_l^\lambda + op \times bw$
fin pour
4.
aller à 1 :

ou a reçus. En d'autres termes, chaque $BPCE_\alpha$ maintiendra une table contenant les prix de protection de deux ensembles de risques de panne : ensemble de risques locaux (ERL_α) et ensemble de risques distants (ERD_α). L'ensemble ERL_α d'un nœud α est composé du nœud α lui-même ainsi que des liens adjacents au nœud α. L'ensemble ERD_α d'un nœud α est formé quant à lui de tous les liens qui appartiennent à des SRLG qui contiennent au moins un lien adjacent au nœud α et qui ne sont pas dans l'ensemble ERL_α. Par exemple, pour le nœud H de la figure 3.1, son ensemble de risques locaux (ERL_H) est constitué du nœud H et de ses liens adjacents G-H et F-H alors que son ensemble de risques distants (ERD_H) est formé des deux liens A-B (partageant $srlg_1$ avec le lien G-H) et B-C (partageant $srlg_2$ avec le lien G-H). Les valeurs des prix de protection des risques appartenant aux deux ensembles ERL_H et ERD_H sont initialisées au démarrage du $BPCE_H$ à zéro et elles sont mises-à-jour à chaque création d'un LSP de secours protégeant un lien ou un nœud appartenant à l'un ou l'autre des ensembles ERL_H ou ERD_H. Ainsi et à la création du LSP de secours b_1, toutes les valeurs des prix de protection sur les arcs de b_1 sont augmentées de 6 unités. De même, à la réception de la structure et des propriétés du LSP de secours b_2 (resp. de b_3) par le nœud H, ce dernier augmentera les valeurs des prix de protection du SRLG $srlg_1$ (resp. $srlg_2$) sur les arcs de b_2 (resp. de b_3) de 4 unités (cf. table 3.1[3]).

Afin de partager les structures et propriétés des LSP de secours qui ne sont pas indépendants, les entités $BPCE_{\alpha\,(\alpha\in V)}$ exécutent les deux algorithmes 1 et 2. Ainsi et lors de la réception d'une requête de demande de calcul d'un LSP de secours par un nœud, ce dernier transmet la requête à son BPCE qui exécutera l'algorithme 1. Pour l'exemple de la figure 3.1, après l'établissement du LSP primaire p_4 ($G\rightarrow H$), l'entité $BPCE_H$ recevra une requête de demande de calcul d'un LSP de secours b_4 ($bw\ (b_4) = 4$) protégeant p_4 contre la panne du lien G-H. Dans une première étape, $BPCE_H$ copiera l'ensemble des arcs de la topologie du réseau dans un nouvel ensemble *E'*, puis il déterminera le coût de protection du LSP de secours b_4 en utilisant les formules (2.2) et (2.2) (cf. ligne 3 de la table 3.2). Après cela, l'entité $BPCE_H$ éliminera de l'ensemble *E'* tous les arcs qui ne disposent pas d'une quantité suffisante de bande passante pour supporter le LSP de secours b_4. Ces derniers correspondent aux arcs $\{\lambda\}_{\lambda\in E}$ qui ne vérifient pas la formule $\theta^\lambda(b_4) + bw(b_4) + F_\lambda > BC^\lambda$ (arcs situés dans les cases grises de la table 3.2). A cette fin, $BPCE_H$ calculera pour chaque arc λ du réseau la valeur $C^\lambda - F_\lambda - bw(b_4)$ (dernière ligne de la table 3.2)) et la comparera avec le coût de protection du LSP b_4 (troisième ligne de la table 3.2)). Ainsi, les arcs $D\rightarrow E$, $E\rightarrow F$, $A\rightarrow B$, $B\rightarrow C$ et $G\rightarrow H$ seront éliminés de l'ensemble *E'*. A l'étape 2 de l'algorithme 1, $BPCE_H$ lance l'algorithme de calcul des chemins entre le PLR (nœud G) et la destination (nœud H) du LSP p_4 sur la topologie de réseau restreinte aux arcs appartenant à l'ensemble *E'*. Il déterminera alors l'unique chemin $G\rightarrow D\rightarrow A\rightarrow I\rightarrow C\rightarrow F\rightarrow H$ permettant de protéger contre la panne du lien G-H. Ensuite (à l'étape 4 de l'algorithme 1), $BPCE_H$ mettra-à-jour sa table de prix de protection en augmentant[4] les prix de protection du lien G-H

[3]Il n'est pas utile de calculer le prix de protection d'un risque r sur un arc λ contenu dans le risque r (utilisation du caractère '-' dans la table 3.1). En effet, l'arc λ ne peut pas être utilisé pour protéger contre la panne du risque r.

[4]Dans le cas d'une suppression d'un LSP de secours, les prix de protection sont diminués.

de 4 unités sur tous les arcs ($G \rightarrow D$, $D \rightarrow A$, $A \rightarrow I$, $I \rightarrow C$, $C \rightarrow F$ et $F \rightarrow H$) appartenant au LSP b_4 déterminé. A l'étape 5, $BPCE_H$ envoie aux nœuds d'extrémité (G, A, B et C) des liens, formant les SRLG contenant le lien protégé, la structure et les propriétés du LSP b_4 calculé. Dans la dernière étape de l'algorithme 1, le LSP de secours b_4 sera retourné.

Lorsqu'une entité $BPCE_\alpha$ reçoit un message contenant la structure (arcs formant le LSP de secours) et les propriétés (bande passante, lien protégé) d'un LSP de secours créé ou supprimé, elle met-à-jour sa table de prix de protection. Pour une création (resp. suppression) de LSP de secours, l'entité $BPCE_\alpha$ augmente (resp. diminue) les prix de protection du lien protégé de la valeur de la bande passante du LSP reçu sur tous les arcs de ce derniers (cf. algorithme 2). Pour l'exemple de la figure 3.1, les nœuds G, A, B et C recevront la structure et les propriétés du LSP b_4 et augmenteront les prix de protection du lien G-H sur tous les arcs de b_4. Formellement, chaque nœud recevant un message contenant la structure et les propriétés d'un LSP de secours exécutera l'algorithme résumé dans les étapes de l'algorithme 2.

3.2.2 Extensions des protocoles de signalisation et IGP-TE pour implanter l'algorithme TDRA

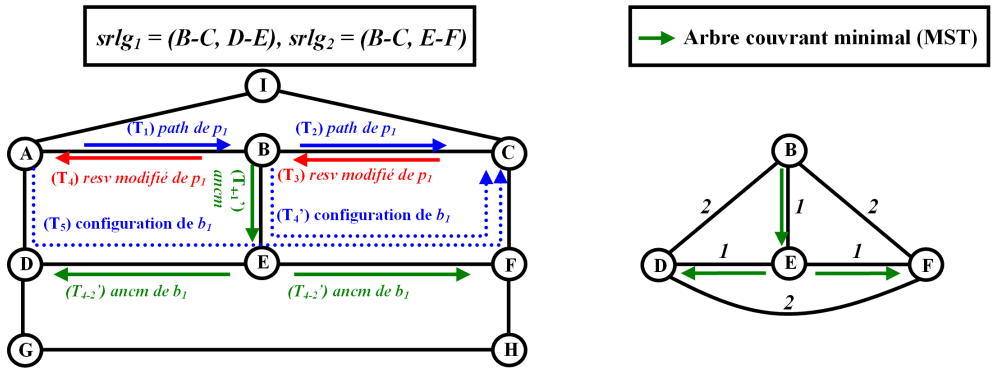
Pour implanter l'algorithme TDRA, deux approches peuvent être adoptées : (1) définir de nouveaux protocoles pour permettre la communication entre les entités de calcul des LSP de secours ou (2) étendre les protocoles existants. Vu la légèreté des extensions nécessaires aux protocoles de signalisation et IGP-TE pour implanter TDRA, nous optons pour la seconde approche.

Ainsi, afin de transmettre les structures et les propriétés des LSP de secours aux nœuds d'extrémité des liens appartenant à l'ensemble des risques distants, nous proposons d'étendre les protocoles de signalisation. La sous-section 3.2.2.1 décrit ces extensions pour le protocole RSVP-TE. Dans la sous-section 3.2.2.2, nous proposons d'autres extensions aux protocoles IGP-TE (OSPF-TE et ISIS-TE) afin de transporter l'information sur la bande passante de secours (les quantités de bande passante de secours allouées sur chaque arc et/ou éventuellement les capacités de secours).

3.2.2.1 Extensions du protocole RSVP-TE

Pour permettre l'envoi des structures et des propriétés des LSP de secours à l'ensemble des *Nœuds d'Extrémité des Liens formant les SRLG contenant le Lien Protégé* (NELSLP), nous ajoutons au protocole RSVP-TE un nouveau type de message *ancm* (announcement messages) ainsi qu'un nouvel objet appelé BACKUP_PATH. L'objet BACKUP_PATH, contiendra la structure et le type d'un LSP de secours, le lien protégé ainsi que les propriétés du LSP de secours (identifiant, bande passante, etc). Il est transporté dans les messages *resv* et *path* de RSVP-TE et dans les messages *ancm* (announcement messages) définis ici à cette fin.

Pour solliciter la protection d'un LSP primaire, le LSR de tête (du LSP primaire) doit positionner le drapeau *local protection desired* et/ou inclure l'objet FAST_REROUTE



(a) Séquencement des messages RSVP-TE lors de l'établissement d'un LSP primaire et des LSP de secours le protégeant

(b) MST couvrant les nœuds de graphe de distances

FIG. 3.2 – Configuration d'un LSP primaire et de ses LSP de secours

dans le message *path* du protocole RSVP-TE [PSA05]. Lorsque ce message atteint le nœud de destination *dest* du LSP primaire (figure 3.2 (a)), ce dernier effectue un contrôle d'admission sur son lien en amont pour vérifier qu'il peut supporter le trafic du LSP primaire en cours de configuration. Si tel est le cas[5], l'entité de calcul $BPCE_{dest}$ tournant sur le même nœud *dest* est contactée pour calculer un LSP de secours protégeant contre la panne du lien en amont sur le LSP primaire. Après le calcul de ce LSP de secours, le nœud *dest* crée un objet BACKUP_PATH contenant la structure et les propriétés du LSP de secours déterminé, construit un nouveau message *resv* incluant l'objet BACKUP_PATH et envoie le message à son nœud en amont sur le LSP primaire.

Lorsqu'un nœud du LSP primaire reçoit un message *resv* contenant l'objet BACKUP_PATH, il effectuera les mêmes traitements que ceux accomplis par le nœud *dest* à la différence que le LSP de secours calculé est de type NNHOP. En plus de ces traitements, le nœud extrait l'objet BACKUP_PATH du message *resv* reçu et configure le LSP de secours contenu dans cet objet. Si le lien protégé apparaît dans un SRLG, un message *ancm* contenant l'objet BACKUP_PATH sera transmis à tous les nœuds de NELSLP, excepté le nœud supportant le BPCE ayant calculé le LSP de secours. Ce dernier sera informé de la réussite de l'établissement du LSP de secours par l'inclusion de l'objet BACKUP_PATH dans les messages *path* rafraîchissant le LSP primaire.

Afin de diminuer le nombre de messages transmis dans le réseau pour informer les nœuds appartenant à NELSLP des structures et propriétés des LSP de secours établis, nous proposons d'utiliser une structure d'arbre pour le routage de ces messages. Ainsi, un arbre couvrant minimal (*minimal spanning tree*) incluant tous les nœuds d'extrémité des liens, formant les SRLG contenant le lien protégé, sera déterminé et utilisé par chaque nœud envoyant ou recevant les messages *ancm*. Pour tout nœud, l'arbre couvrant minimal utilisé pour le transport des messages *ancm* sera calculé sur le graphe des distances qui ne contient que les nœuds d'extrémité des liens formant les SRLG

[5]Si la réservation de bande échoue, la configuration du LSP primaire pourrait être abandonnée.

incluant le lien protégé. Dans ce graphe, les nœuds sont interconnectés entre eux en utilisant des arcs dont le coût est égal à la plus petite distance (en nombre de sauts par exemple) entre ces nœuds dans la topologie originale du réseau. Après la déduction du graphe des distances, les nœuds exécutent l'algorithme de Prim [Pri57] ou l'algorithme de Kruskal [Kru56] afin de déterminer l'arbre couvrant minimal. C'est ce dernier arbre qui sera utilisé pour le routage des messages *ancm*. Chaque nœud de l'arbre ne transmet les messages *ancm* reçus qu'à ses nœuds enfants.

Exemple

Sur la figure 3.2 (a), le nœud A reçoit une requête de demande d'établissement d'un LSP primaire interconnectant le nœud A au nœud C. A ce moment, le nœud A détermine le plus court chemin $A \rightarrow B \rightarrow C$ reliant A à C et envoie un message *path* à son prochain nœud B afin de configurer ce chemin (instant T_1). Lorsque le nœud B reçoit ce message *path*, il effectue les traitements décrits dans [PSA05] (vérification d'absence de boucles, création d'un état RSVP, etc.) et redirigera le message vers le nœud C à l'instant T_2 ($T_2 > T_1$). A la réception de ce message par le nœud C, ce dernier analyse le message et déduit que le LSP primaire en cours d'établissement requiert une protection locale. En conséquence, le nœud C calcule un LSP de secours b_1 ($B \rightarrow E \rightarrow F \rightarrow C$) protégeant le LSP primaire contre la panne du lien $B \rightarrow C$ et construira un objet BACKUP_PATH incluant la structure et les propriétés de b_1. Après la vérification des contraintes de bande passante (contrôle d'admission) sur l'arc $B \rightarrow C$, le nœud C enverra un message *resv* contenant l'objet BACKUP_PATH au routeur B à l'instant T_3 ($T_3 > T_2$).

Lorsque le nœud B reçoit le message *resv* envoyé par le nœud C, il effectue un contrôle d'admission sur l'arc $A \rightarrow B$ et calcule un LSP de secours b_2 ($A \rightarrow D \rightarrow E \rightarrow F \rightarrow C$) protégeant contre les pannes du nœud B et du lien A-B. Après cela, le nœud B construit un objet BACKUP_PATH contenant la structure et les propriétés du LSP b_2 qu'il envoie au nœud A dans un message *resv* (à l'instant $T_4 > T_3$). Ensuite, le nœud B extrait de l'objet BACKUP_PATH transmis dans le message *resv* reçu (du nœud C) la structure et les propriétés du LSP de secours b_1 qu'il configure à l'instant $T_4' > T_3$.

Comme le lien B-C protégé par le LSP de secours b_1 apparaît dans des SRLG ($srlg_1$ = (B-C, D-E) et $srlg_2$ = (B-C, E-F)), le nœud B devra informer les nœuds d'extrémité D, E et F des liens apparaissant dans les SRLG ($srlg_1$ et $srlg_2$) de la structure et des propriétés de b_1. Pour ce faire, le nœud B calculera l'arbre MST couvrant ces nœuds ainsi que le nœud B lui-même. De ce fait, le nœud B déduit le graphe des distances couvrant les nœuds B, D, E et F à partir de la base de données topologique, puis détermine l'arbre couvrant minimal MST (figure 3.2 (b)). Ensuite, le nœud B construit un message *ancm* contenant l'objet BACKUP_PATH transmis dans le message *resv* reçu du nœud C et l'envoie à son unique nœud fils E dans l'arbre MST à l'instant $T_{4\text{-}1}'$ ($T_{4\text{-}1}' > T_4'$). Ce nœud E mettra-à-jour alors sa table de prix de protection et redirigera le message à ses deux nœuds fils dans l'arbre MST à l'instant $T_{4\text{-}2}'$ ($T_{4\text{-}2}' > T_{4\text{-}1}'$). Ces deux derniers nœuds mettront-à-jour, à leur tour, leur table de prix de protection avant de détruire les messages reçus.

Enfin, lorsque le nœud A reçoit le message *resv* envoyé par le nœud B, il configure le LSP de secours b_2 (à l'instant $T_5 > T_4$) contenu dans l'objet BACKUP_PATH.

3.2.2.2 Extensions aux protocoles OSPF-TE et ISIS-TE

Avec l'algorithme TDRA, deux stratégies peuvent être adoptées pour allouer la bande passante aux LSP de secours : stratégie à un seul pool et stratégie à deux pools disjoints. Dans la première stratégie, les allocations de bande passante pour les LSP primaires et les LSP de secours sont effectuées à partir d'un même pool de bande passante qui correspond à la capacité de l'arc. Dans la seconde stratégie, la capacité d'un arc est divisée en deux pools disjoints : pool primaire et pool de secours. Toute allocation de bande passante pour un LSP primaire (resp. LSP de secours) ne peut être effectuée qu'à partir du pool primaire (resp. du pool de secours).

Selon la stratégie d'allocation de bande passante employée, différents paramètres doivent être annoncés par les protocoles IGP-TE afin d'implanter l'algorithme TDRA. Concrètement, il est nécessaire d'annoncer les quantités de la bande passante de secours allouées sur les arcs (resp. les capacités de secours des arcs) si l'on utilise la première stratégie (resp. la seconde stratégie).

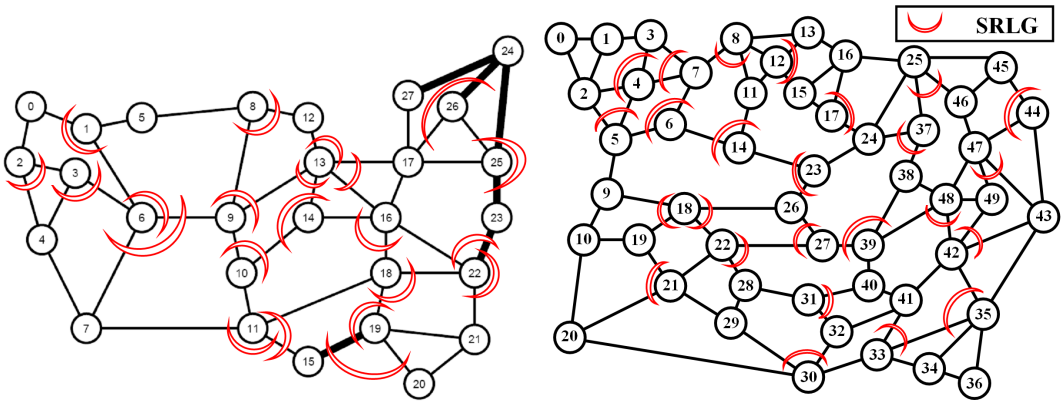
Pour des raisons de généricité, nous proposons ici d'ajouter à l'IGP-TE deux nouveaux champs afin d'annoncer les deux paramètres précédents[6] (les capacités de secours et les quantités de bande passante de secours) pour tout arc de la topologie du réseau. Pour ce faire, nous proposons d'utiliser les paramètres de l'ingénierie de trafic définis dans OSPF-TE et dans ISIS-TE. Ainsi, nous suggérons d'utiliser un nouveau champ au format TLV associé à chaque arc. Ce champ, chargé de transmettre les valeurs de la capacité de secours et de la bande passante de secours allouées sur chaque arc, sera transporté dans les *LSA* pour OSPF-TE (cf. section 1.3.1) et dans les *LSPDU* pour ISIS-TE (cf. section 1.3.2).

3.2.3 Évaluation des performances

Afin de mesurer les performances de l'algorithme TDRA, nous l'avons comparé à l'heuristique de Kini et al. améliorée que nous notons dans la suite de cette thèse par HKA (cf. section). Notre choix s'est porté sur cette heuristique de Kini vu ses nombreux avantages comme :

1. sa facilité de déploiement : de très légères extensions aux protocoles IGP-TE permettent sa mise-en-œuvre.
2. sa rapidité de restauration : cette heuristique n'exige pas la distinction de la panne d'un nœud de celle d'un lien (comme dans [VCLF+04]), ce qui réduit le délai de recouvrement.
3. elle n'inonde pas le réseau avec les messages transportant l'information nécessaire au calcul des LSP de secours (comme dans l'algorithme diffusion de tous les prix de protection décrit en section 2.7.1).

[6] Il est judicieux de ne transmettre qu'un seul paramètre dans l'IGP-TE bien que nous ayons défini deux champs. Ainsi, si l'on utilise la première stratégie, seules les quantités de bande passante de secours allouées sur les liens sont annoncées. Par contre, si l'on utilise la seconde stratégie, il est préférable de ne transmettre que les capacités de secours des liens.



(a) Taille moyenne (95 risques) (b) Taille large (162 risques)

FIG. 3.3 – Topologies de tests

Pour nos simulations, nous avons utilisé deux topologies de réseau (cf. figure 3.3) : une topologie de réseau de taille moyenne et une topologie de réseau de grande taille. La première topologie de réseau (illustrée sur la figure 3.3 (a)) est constituée 28 nœuds, 45 liens bidirectionnels (i.e. 90 arcs) et 22 SRLG (représentés par des croissants sur la figure 3.3). La seconde topologie de réseau (illustrée sur la figure 3.3 (b)) est constituée 50 nœuds, 87 liens bidirectionnels (i.e. 174 arcs) et 25 SRLG (représentés par des croissants sur la figure 3.3). La bande passante disponible sur chaque arc de ces topologies de réseau est divisée en deux pools : pool primaire d'une capacité infinie (i.e. capacité suffisante pour router tous les LSP primaires sans blocage) et pool de secours d'une capacité égale à 300 unités sur les liens en gras et à 100 unités sur tous les autres arcs. Pour la consistance de nos résultats, nous avons effectué, dans [SC07], d'autres tests sur d'autres topologies de réseau présentant des caractéristiques différentes par rapport à celles illustrées sur la figure 3.3.

La matrice de trafic est générée aléatoirement et est formée de LSP réclamant des quantités de bande passante uniformément distribuées sur l'intervalle [1, 10]. Ainsi, pour générer un LSP primaire, deux nœuds (source et destination du LSP) sont choisis aléatoirement parmi les nœuds du réseau puis reliés par un LSP primaire calculé en utilisant l'algorithme des plus courts chemins de Dijkstra. A chaque LSP primaire généré, l'algorithme TDRA et l'heuristique HKA sont appliqués afin de déterminer les LSP de secours permettant la protection de tous les liens et nœuds du LSP primaire (l'algorithme CSPF est utilisé pour le calcul de ces LSP de secours).

3.2.3.1 Métriques

Pour comparer l'algorithme TDRA à l'heuristique HKA, nous avons choisi d'utiliser les trois métriques suivantes : taux de rejet des demandes d'établissement de LSP de secours (TR), nombre moyen de messages transmis dans le réseau pour permettre le calcul des LSP de secours (NMM) et taux d'utilisation de la bande passante de secours

(*TUBS*).

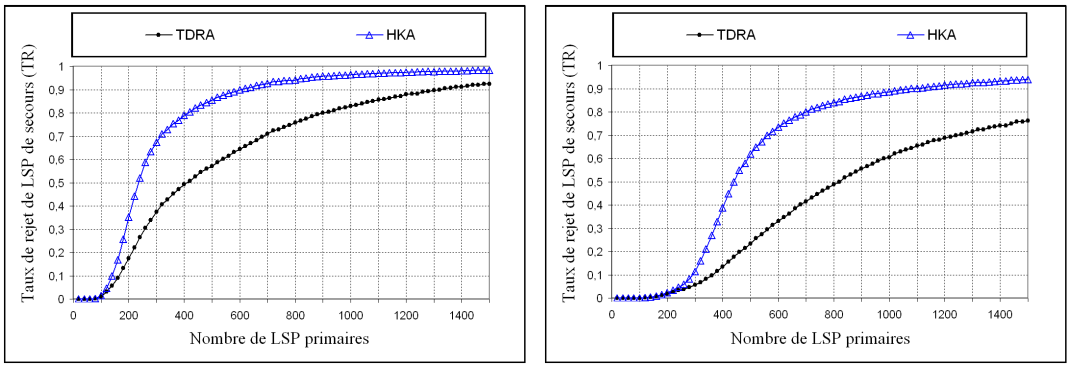
La première métrique (*TR*) mesure le taux de LSP de secours rejetés pour manque de bande passante. Cette métrique correspond au rapport entre le nombre de LSP de secours rejetés sur le nombre total de demandes d'établissement de LSP de secours. Plus cette métrique est élevée, moins la technique de calcul des LSP de secours est efficace.

La seconde métrique (*NMM*) mesure le nombre moyen de messages transmis sur les liens du réseau, à chaque calcul/établissement d'un nouveau LSP de secours, pour mettre-à-jour l'information nécessaire au placement des LSP de secours. Pour l'algorithme TDRA, cette métrique mesure le nombre moyen de messages *ancm* transmis sur les liens du réseau après l'établissement d'un seul LSP de secours. Concernant l'heuristique HKA, cette métrique mesure le nombre moyen de messages nécessaires pour annoncer dans l'IGP-TE les nouvelles valeurs de la bande passante de secours sur les arcs (les messages annonçant les quantités de bande passante primaire ou les capacités de secours ne sont pas comptabilisés ici puisqu'ils sont communs et transmis aux mêmes instants avec les deux techniques de calcul comparées) après l'établissement d'un seul LSP de secours.

La troisième métrique (*TUBS*) mesure l'efficacité des procédés adoptés pour l'allocation de la bande passante aux LSP de secours. C'est une métrique permettant d'estimer efficacement le taux de partage de bande passante de secours sur les différents arcs de la topologie du réseau. Elle est déterminée comme le rapport entre la moyenne des quantités de bande passante cumulée des LSP de secours sur les arcs et la capacité de secours moyenne. Formellement (α-β et λ ne doivent partager aucun SRLG) :

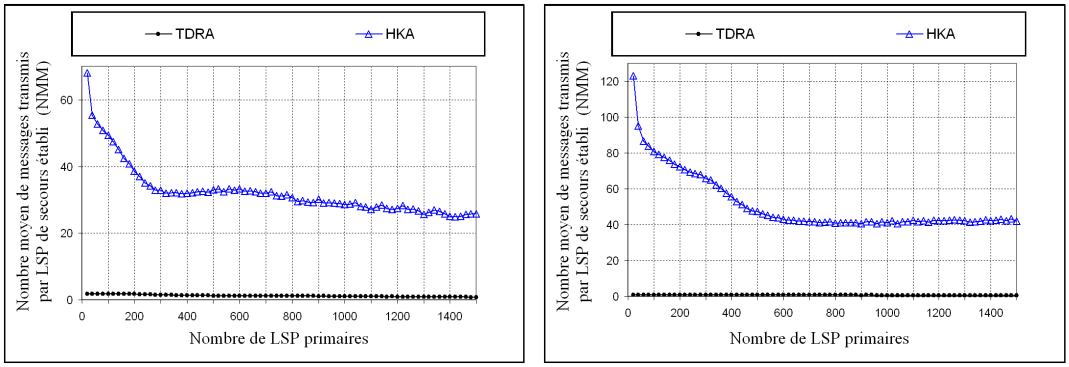
$$TUBS = \sum_{(\lambda, \alpha\text{-}\beta) \setminus (\lambda \in E) \wedge (\alpha \rightarrow \beta \in E \setminus \{\lambda\} \wedge \beta \rightarrow \alpha \in E \setminus \{\lambda\})} \delta^{\lambda}_{\alpha\text{-}\beta} \; / \sum_{\lambda \in E} BC^{\lambda}$$

Afin de montrer l'impact de la charge du réseau sur la quantité d'informations diffusées dans le réseau, nous mesurons les valeurs des trois métriques précédentes pour différentes charges du réseau. Ainsi et à chaque établissement de 20 LSP primaires, les valeurs des métriques *TR*, *NMM* et *TUBS* sont déterminées et retournées pour l'algo-



(a) Topologie du réseau de 95 risques (b) Topologie du réseau de 162 risques

FIG. 3.4 – Evolution du taux de rejet des LSP de secours (TR)



(a) Topologie du réseau de 95 risques (b) Topologie du réseau de 162 risques

FIG. 3.5 – Evolution du nombre moyen de messages transmis dans le réseau pour l'établissement d'un LSP de secours (NMM)

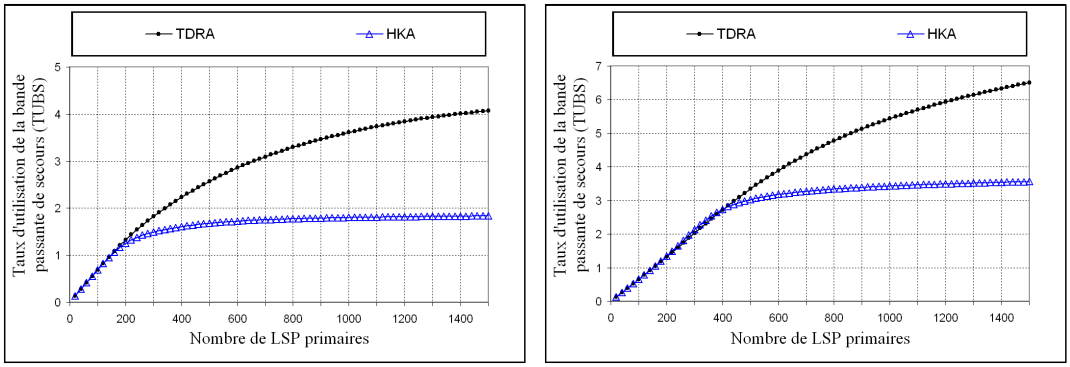
rithme TDRA et l'heuristique HKA. Tous les résultats illustrés dans la section 3.2.3.2 correspondent aux valeurs moyennes obtenues après 1000 exécutions et tirages aléatoires.

3.2.3.2 Résultats et analyse

La figure 3.4 illustre l'évolution du taux de rejet des LSP de secours en fonction du nombre de LSP primaires établis. Sur cette figure, il apparaît clairement que l'algorithme TDRA a un taux de rejet de LSP de secours toujours inférieur et meilleur à celui de l'heuristique HKA. Ceci s'explique par la disponibilité de toute l'information nécessaire au placement des LSP de secours avec l'algorithme TDRA (ce qui permet un partage maximal de la bande passante sur tout arc) alors que l'heuristique HKA n'utilise qu'une information partielle et incomplète (i.e. l'heuristique HKA surestime les prix de protection alors que TDRA a une connaissance des valeurs des prix de protection).

Pour les faibles charges du réseau (nombre de LSP primaires inférieur à 140 dans la figure 3.4 (a) et inférieur à 240 dans la figure 3.4 (b)), nous remarquons que l'heuristique HKA a un taux de rejet des LSP de secours assez proche (légèrement supérieur) de celui qui est obtenu avec l'algorithme TDRA. Ceci s'explique par les valeurs des quantités de bande passante de secours qui sont petites (inférieures aux capacités de secours des liens moins la quantité maximale de bande que peut réclamer un LSP), ce qui permet d'inclure presque tous les arcs lors du placement d'un nouveau LSP de secours. En effet, pour ces charges du réseau, les taux de rejet des LSP de secours des deux techniques de calcul des LSP de secours (TDRA et HKA) sont quasi-nuls.

Lorsque la charge du réseau augmente, la différence entre le taux de rejet des LSP de secours de l'heuristique HKA et celui de l'algorithme TDRA devient plus apparente et plus grande. Alors que l'algorithme TDRA dispose de toute l'information permettant la sélection ou le rejet d'un arc lors du calcul d'un nouveau LSP de secours, l'heuristique HKA adopte une approche pessimiste surestimant les prix de protection. En consé-



(a) Topologie du réseau de 95 risques (b) Topologie du réseau de 162 risques

FIG. 3.6 – Evolution du taux d'utilisation de la bande passante de secours (TUBS)

quence, un nombre important d'arcs est rejeté par erreur lors du calcul d'un nouveau LSP de secours, surtout lorsque les valeurs des quantités de bande passante de secours sont élevées sur les arcs (charges élevées).

Concernant la seconde métrique *NMM* (figure 3.5), nous apercevons que l'heuristique HKA transmet en moyenne beaucoup plus de messages dans le réseau par rapport à l'algorithme TDRA (bien que l'information transmise par HKA change moins souvent que l'information transmise par TDRA). Cela s'explique essentiellement par la différence des mécanismes utilisés par les deux techniques comparées pour la transmission de l'information nécessaire au placement des LSP de secours. Alors que l'heuristique HKA utilise la diffusion qui est très coûteuse en termes de messages transmis, l'algorithme TDRA cible les envois et ne transmet de messages qu'aux nœuds d'extrémité de liens appartenant aux mêmes SRLG que le lien protégé (toute diffusion dans le réseau illustré sur la figure 3.3 (a) nécessite l'envoie de 27 messages et toute diffusion dans le réseau illustré sur la figure 3.3 (b) nécessite l'envoie de 49 messages).

Concernant la dernière métrique *TUBS*, la figure 3.6 montre que le taux d'utilisation de la bande passante de secours de l'algorithme TDRA est assez proche de celui qui correspond à l'heuristique HKA lorsque la charge réseau est faible (i.e. le nombre de LSP primaires protégés est inférieur à 200 dans le réseau de la figure 3.3 (a) et inférieur à 400 dans le réseau de la figure 3.3 (b)). Ceci s'explique essentiellement par les taux de rejet de LSP de secours qui sont faibles pour ces charges, ce qui permet d'accepter, à peu près, les mêmes LSP de secours.

Notons tout de même que sur la figure 3.6 (b), les taux d'utilisation de la bande passante de secours des deux techniques comparées sont assez semblables, même pour des charges réseau induisant des taux de rejet de LSP de secours élevés. Par exemple, pour 400 LSP primaires établis, nous apercevons sur la figure 3.4 (b) que les taux de rejets de TDRA et HKA sont supérieurs à 48% alors que leur taux d'utilisation de bande passante de secours sont quasi-identiques. Ceci s'explique par l'allongement, avec HKA, de la taille des LSP de secours pour éviter certains blocages. En effet, la surestimation

des valeurs des prix de protection avec HKA induit le rejet par erreur de certains arcs lors du placement des LSP de secours, ce qui réduit la flexibilité dans le choix des LSP de secours.

Pour des charges réseau élevées, la figure 3.6 montre que le taux d'utilisation de la bande passante de secours de TDRA est meilleur et est plus élevé que celui de HKA. Ceci est dû à la surestimation des prix de protection avec HKA, ce qui induit des fuites de la bande passante. Avec TDRA par contre, l'information nécessaire au calcul des prix de protection est entièrement disponible sur les nœuds plaçant les LSP de secours, ce qui diminue la quantité de bande passante de secours réservées aux LSP de secours et donc augmente le taux d'utilisation de la bande passante de secours.

3.2.4 Conclusion

L'algorithme TDRA permet un placement efficace des LSP de secours tout en distribuant la charge du calcul équitablement sur les nœuds du réseau. Chaque nœud du réseau, dans TDRA, supporte une entité BPCE s'occupant du calcul des LSP de secours (de type NHOP ou NNHOP) protégeant contre les pannes du nœud lui-même ainsi que de ses liens rentrants. Contrairement aux algorithmes classiques de placement des LSP de secours qui ne traitent pas les pannes des SRLG, TDRA a été conçu pour faire face à tous les types de risques de panne, notamment les SRLG.

ainsi, pour gérer les liens appartenant à des SRLG, l'algorithme TDRA transmet la structure et les propriétés de chaque LSP de secours à tous les nœuds d'extrémité des liens appartenant à des SRLG contenant le lien protégé. Pour les autres types de risque de panne, seule la structure du LSP de secours est transmise au nœud PLR par le nœud voisin qui supporte l'entité BPCE ayant calculé le LSP de secours. Cette distribution ciblée de l'information requise au placement des LSP de secours a l'avantage de réduire considérablement le trafic de contrôle dans le réseau.

Avec de très légères extensions aux protocoles de signalisation et protocoles IGP-TE, l'algorithme TDRA peut être déployé. En effet, l'ajout d'un seul objet (BACKUP_PATH) au protocole RSVP-TE et la définition d'un nouveau champ TLV dans l'IGP-TE permettra le transport de toute l'information nécessaire au calcul des LSP de secours.

Enfin, la comparaison de l'algorithme TDRA avec l'heuristique améliorée de Kini montre que l'algorithme TDRA réduit les risques de blocage, augmente l'utilisation de la bande passante de secours et passe mieux à l'échelle puisqu'il ne nécessite que la transmission d'un nombre limité de messages dans le réseau.

3.3 Partage efficace et distribué de la bande passante (DBSH)

Pour diminuer la taille et la fréquence d'envoi de l'information requise au calcul des LSP de secours, nous proposons l'heuristique DBSH (*Distributed Bandwidth Sharing Heuristic*) [SCLR07] qui réduit et agrège cette information avant de la diffuser dans

le réseau. Ainsi, avec la transmission d'un vecteur de $x^{\lambda}_{(0<x^{\lambda}<\infty)}$[7] valeurs de prix de protection (vérifiant certaines contraintes) par arc λ, il est possible de déterminer, avec une probabilité élevée de succès, les arcs susceptibles de supporter les LSP de secours en cours de placement. Cette heuristique repose sur les deux observations suivantes :

1. Certaines valeurs des prix de protection sur un arc sont très petites. Approximer ces valeurs par leur maximum décroît la quantité d'informations transmises dans le réseau sans détériorer les possibilités de partage de la bande passante.
2. Si l'on fixe un seuil de confiance au dessous duquel toute requête de demande d'établissement de LSP de secours est toujours satisfaite, il sera inutile de transmettre les valeurs des prix de protection inférieures à ce seuil.

3.3.1 Principes de l'heuristique DBSH

Comme dans l'algorithme TDRA, chaque nœud α du réseau supporte une entité $BPCE_{\alpha}$, exécutant l'heuristique DBSH, dont le rôle est de calculer les LSP de secours protégeant contre les pannes du nœud α lui-même et de tous ses liens rentrants. De ce fait, tout calcul de LSP de secours (de type NHOP ou NNHOP) protégeant contre la panne d'un lien indépendant se fait sans distribution, dans le réseau, d'informations sur les prix de protection. Cependant, lorsque le lien à protéger est dans au moins un SRLG, la transmission d'une information permettant de déterminer les prix de protection des SRLG contenant le lien à protéger est nécessaire. Pour ce faire, l'approche triviale, consistant à diffuser les prix de protection des SRLG contenant le lien à protéger, peut être adoptée et utilisée dans les réseaux de très petite taille et/ou incluant un nombre faible de SRLG.

Cependant, dans le cas de réseaux larges contenant un nombre relativement élevé de SRLG, la diffusion des prix de protection de tous les SRLG (sur tous les arcs) inonde le réseau et détériore les performances. Pour éviter cette situation, nous proposons d'adopter l'heuristique DBSH qui agrège les prix de protection avant de les diffuser dans le réseau. Dans cette heuristique, un vecteur, appelé $x^{\lambda}\_vecteur_{(0<x^{\lambda}<\infty)}$, contenant les x^{λ} prix de protection de SRLG les plus élevés (et leur SRLG correspondants) sont diffusés dans le réseau pour chaque arc λ. La taille réelle d'un tel vecteur peut être variable d'un arc à un autre et elle dépend du seuil de confiance Sc^{λ} fixé sur chaque arc λ et de la valeur du paramètre x^{λ} appelé aussi *la taille maximale* du vecteur. Typiquement, si le nombre de SRLG, dont le prix de protection sur l'arc λ est supérieur au seuil de confiance Sc^{λ}, est égal à y^{λ} alors la taille réelle du vecteur $x^{\lambda}\_vecteur$ sera égale à $Min\ (x^{\lambda}, y^{\lambda})$.

Pour optimiser le partage et diminuer la quantité d'informations diffusées dans le réseau, le seuil de confiance Sc^{λ} (constante) doit être fixé de telle sorte que la quantité maximale de bande passante que peut réclamer tout nouveau LSP de secours soit égale à BC^{λ} - Sc^{λ}. De même, la valeur du paramètre x^{λ} sur un arc λ doit être choisie suffisamment élevée pour maximiser le partage de la bande passante mais assez petite

[7] Évidemment, la valeur réelle de x^{λ} ne peut être infinie. Sa valeur maximale est égale au nombre de SRLG du réseau.

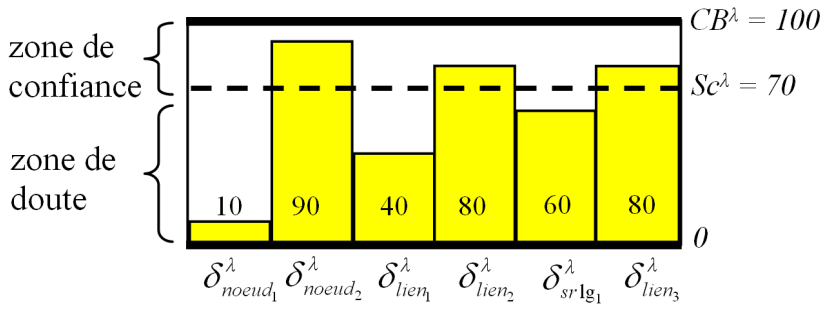


FIG. 3.7 – Prix de protection des SRLG sur un arc λ

pour limiter la taille de l'information diffusée dans le réseau. Par conséquent, le choix de la valeur de x^{λ} sur chaque arc λ est primordial pour réguler le degré de partage et la quantité d'informations diffusées dans le réseau.

Afin de simplifier la compréhension de l'heuristique DBSH, pour contrôler la quantité de bande passante dédiée à la protection et pour séparer la tâche de calcul des LSP primaires de celle déterminant les LSP de secours, nous divisons la capacité de tout arc λ en deux pools disjoints[8] (comme dans la section 2.7.2) : pool primaire et pool de secours. Le pool primaire, de capacité PC^{λ}, est utilisé pour allouer la bande passante aux LSP primaires et le pool de secours, de capacité BC^{λ}, est employé pour allouer la bande passante aux LSP de secours. Ce dernier pool est lui-même structuré en deux zones disjointes et séparées par un seuil de confiance Sc^{λ} : zone de confiance et zone de doute (figure 3.7). La première zone, qui couvre l'intervalle allant de 0 jusqu'à Sc^{λ}, contient des valeurs de bande passante suffisamment petites pour être ignorées dans le processus de calcul des LSP de secours vérifiant les contraintes de bande passante. La seconde zone, s'étalant de Sc^{λ} jusqu'à BC^{λ}, contient des valeurs de bande passante (ou prix de protection) élevées et importantes pour la décision d'inclusion ou d'exclusion de l'arc λ lors d'un calcul d'un nouveau LSP de secours.

Sur la figure 3.7 est illustré un pool de secours d'un arc λ supportant des LSP protégeant contre les pannes de liens appartenant aux SRLG suivants : $srlg_1$, $srlg_2$, $srlg_3$, $srlg_4$, $srlg_5$, ou $srlg_6$ (les prix de protection des SRLG nuls ne sont pas représentés sur la figure). Étant donné que la capacité de secours de l'arc λ est égale à 100 unités et que tout nouveau LSP ne peut réclamer une quantité de bande passante supérieure à 30 unités, le seuil de confiance sera fixé à 70 unités (70 = 100 - 30). Pour utiliser l'heuristique DBSH, tous les nœuds du réseau participant au calcul des LSP de secours doivent exécuter les algorithmes 3 et 4[9].

Concrètement, chaque nœud d'extrémité sortante o^{λ} à un arc λ détermine les $x^{\lambda}_{(0<x<\infty)}$ plus grands prix de SRLG (sur l'arc lui-même) et construit un nouveau $x^{\lambda}\_vecteur$ constitué de ces prix de protection et des SRLG qui leur correspondent. Si ce vecteur est différent de l'ancien $x^{\lambda}\_vecteur$ transmis dans le réseau, le nœud o^{λ}

[8] Cette subdivision de la capacité d'un arc en deux pools disjoints n'est pas obligatoire.
[9] Une description plus détaillée des algorithmes 3 et 4 peut-être trouvée dans [SCLR07].

Algorithme 3 Routine exécutée par tout nœud d'extrémité sortante o^λ d'un arc λ

*ancien_*x^λ*_vecteur* $\leftarrow 0$
tantque *vrai* **faire**
  Affecter au vecteur *nouveau_*x^λ*_vecteur*, associé à l'arc λ, les x^λ *plus grands prix de protection de SRLG* ainsi que les identifiants de SRLG associés
  si le $(x^\lambda+1)^{eme}$ *plus grand prix de protection de SRLG* est supérieur au seuil Sc^λ **alors**
    Remplacer, dans le vecteur *nouveau_*x^λ*_vecteur*, l'identifiant du $(x^\lambda)^{eme}$ plus grand *prix de protection de SRLG* par l'identifiant (spécial) *SRLG_générique*
  finsi
  si *nouveau_*x^λ*_vecteur* $\neq$ *ancien_*x^λ*_vecteur* **alors**
    Diffuser le vecteur *nouveau_*x^λ*_vecteur*
    *nouveau_*x^λ*_vecteur* $\leftarrow$ *ancien_*x^λ*_vecteur*
  finsi
fin tantque

le diffuse dans le réseau.

En fonction de la différence entre le seuil de confiance Sc^λ et de la valeur du $(x+1)^{eme}$ plus grand prix de SRLG (notée x^λ*_plus_un*), nous distinguons deux situations : situation de doute (Sc^λ - x^λ*_plus_un* < 0) et situation sûre (Sc^λ - x^λ*_plus_un* ≥ 0). Dans la situation de doute, l'arc λ peut être *rejeté par erreur* lors d'un calcul d'un nouveau LSP de secours alors que dans une situation sûre, l'arc λ est accepté ou rejeté *sans erreur*.

Dans l'exemple de la figure 3.7, nous nous retrouvons dans la situation de doute d'inclusion/exclusion de l'arc λ lors d'un calcul d'un nouveau LSP de secours si la taille maximale du vecteur x^λ*_vecteur* est inférieure ou égale à 2 (i.e. $x^\lambda \leq 2$). Par contre, lorsque la taille maximale du vecteur x^λ*_vecteur* est strictement supérieure à 2 (i.e. $x^\lambda > 2$), l'inclusion/exclusion de l'arc λ lors d'un calcul d'un nouveau LSP de secours est sûre.

3.3.1.1 Situation de doute ($x^\lambda \leq 2$)

Avec $x^\lambda = 2$ (et même pour $x^\lambda = 1$), le nœud d'extrémité sortante o^λ de l'arc λ détermine les deux plus grands prix (de protection) de SRLG qui sont supérieurs au seuil de confiance Sc^λ. Ensuite, le nœud o^λ construit un *2_vecteur* contenant les deux prix de protection déterminés précédemment ainsi que les SRLG qui leur correspondent (cf. algorithme 3). Par exemple, le nœud d'extrémité sortante de l'arc λ de la figure 3.7 diffusera le *2_vecteur* suivant : [($srlg_2$, 90), (*SRLG_générique*, 80)]. L'identifiant *SRLG_générique* est inclus dans le *2_vecteur* pour indiquer aux nœuds qui le reçoivent que le troisième plus grand prix de protection de SRLG est strictement supérieur au seuil de confiance Sc^λ (i.e. lorsque l'identifiant *SRLG_générique* est inclus dans le vecteur x^λ*_vecteur* diffusé, cela indique une situation de doute).

A la réception du x^λ*_vecteur* par un nœud α différent de l'extrémité entrante de l'arc

Algorithme 4 Routine exécutée par tout nœud recevant un vecteur $x^{\lambda}\_vecteur$ associé à l'arc λ

si $SRLG\_générique \in x^{\lambda}\_vecteur.\textbf{\textit{index}}\ ()$ **alors**
  $prix\_min \leftarrow x^{\lambda}\_vecteur\ [SRLG\_générique]$
sinon
  $prix\_min \leftarrow 0$
finsi
pour chaque identifiant id d'un SRLG **faire**
  si $id \in x^{\lambda}\_vecteur.\textbf{\textit{index}}\ ()$ **alors**
    $table\_locale\_prix\_protection\ [\lambda][id] \leftarrow x^{\lambda}\_vecteur\ [id]$
  sinon
    $table\_locale\_prix\_protection\ [\lambda][id] \leftarrow prix\_min$
  finsi
fin pour

λ, ce dernier met-à-jour sa table de prix de protection des SRLG suivant l'algorithme 4. Concrètement, le nœud α copie les valeurs des prix de protection et les identifiants de SRLG contenus dans le vecteur $x^{\lambda}\_vecteur$ reçu dans sa table locale de prix de protection. Pour les autres SRLG, le nœud α leur affecte le prix de protection associé à l'identifiant $SRLG\_générique$. Par exemple, lors de la réception du $2\_vecteur$ associé à l'arc λ de la figure 3.7 par un nœud non adjacent à l'arc λ, ce dernier met-à-jour sa table de prix de protection en approximant les valeurs des prix des SRLG comme suit : $(\delta^{\lambda}_{srlg_2} = 90) \wedge (\forall\, srlg_i \neq srlg_2 : \delta^{\lambda}_{srlg_i} = 80)$

A cause de la localisation du $(2+1)^{eme}$ plus grand prix de protection de SRLG dans la zone de doute (présence de l'identifiant $SRLG\_générique$ dans le $2\_vecteur$), une entité $BPCE_{\alpha}$ recevant le $2\_vecteur$ précédent, peut refuser, par erreur, d'inclure l'arc λ dans le calcul d'un nouveau LSP de secours. Typiquement, tout calcul d'un nouveau LSP de secours NHOP protégeant contre la panne d'un lien apparaissant dans un SRLG n'appartenant pas à l'ensemble $\{srlg_2, srlg_4, srlg_6\}$ exclut l'arc λ par erreur si la quantité de bande passante demandée est supérieure à 20 unités (i.e. la quantité de bande passante réclamée appartient à l'intervalle $]BC^{\lambda}$ - 80, BC^{λ} - $Sc^{\lambda}]$). Dans tous les autres cas, l'entité $BPCE_{\alpha}$ exclura ou inclura sans erreur l'arc λ dans le calcul du prochain LSP de secours NHOP.

Exemple :

Considérons des requêtes d'établissement de LSP de secours de type NHOP telles que :

1. Les quantités de bande passante demandées par chaque requête sont uniformément distribuées sur l'intervalle [1, 30].
2. Toute requête concerne la protection contre la panne d'un lien appartenant à un seul SRLG. Ce lien (à protéger) est choisi aléatoirement parmi les liens appartenant à un SRLG.
3. Le nombre de SRLG est égal à $nbSRLG$.

La probabilité $P_{exc\backslash e}$ (λ) d'exclure par erreur l'arc λ de la figure 3.7 est déterminée comme suit :
$P_{Exc\backslash e}$ (λ) = [(30 - 20) / (30 - 1 + 1)] $\times$ ($nbSRLG$ - *cardinal* ({$srlg_2$, $srlg_4$, $srlg_6$}) / $nbSRLG$) = ($nbSRLG$ - 3) / (3 $\times$ $nbSRLG$).
En fonction du nombre de SRLG considéré, la probabilité d'exclure par erreur l'arc λ peut aller de 16,67% ($nbSRLG$ = 6) jusqu'à 33,33% pour un nombre infini de SRLG.

Notons enfin que cette probabilité d'exclusion par erreur sera réduite sensiblement si l'on considère les cas de LSP protégeant contre les pannes de liens indépendants et les LSP de type NNHOP.

3.3.1.2 Situation sûre ($x^\lambda > 2$)

Avec $x^\lambda = 3$ (et même pour $x^\lambda > 3$), le nœud d'extrémité sortante o^λ de l'arc λ détermine les trois plus grands prix (de protection) de SRLG qui sont supérieurs au seuil de confiance Sc^λ. Ensuite, le nœud o^λ construit un *3_ vecteur* contenant les trois prix de protection déterminés précédemment ainsi que les SRLG qui leur correspondent (cf. algorithme 3). Par exemple, le nœud d'extrémité sortante o^λ de l'arc λ de la figure 3.7 diffusera le *3_ vecteur* suivant : [($srlg_2$, 90), ($srlg_4$, 80), ($srlg_6$, 80)]. L'identifiant *SRLG_ générique* est absent dans le vecteur *3_ vecteur* transmis, ce qui indique que le quatrième plus grand prix de protection de SRLG est inférieur au seuil de confiance Sc^λ (situation sûre).

A la réception du x^λ*_ vecteur* par un nœud α différent de l'extrémité entrante de l'arc λ, ce dernier met-à-jour sa table de prix de protection des SRLG suivant l'algorithme 4. Concrètement, le nœud α copie les valeurs des prix de protection et les identifiants de SRLG contenus dans le vecteur x^λ*_ vecteur* reçu dans sa table locale de prix de protection. Pour les autres SRLG, le nœud α leur affecte un prix de protection nul. Par exemple, lors de la réception du *3_ vecteur* associé à l'arc λ de la figure 3.7, par un nœud non adjacent à l'arc λ, ce dernier met-à-jour sa table de prix de protection des SRLG en approximant les valeurs des prix des SRLG comme suit :

$$\begin{cases} (\delta^\lambda_{srlg_2} = 90) \wedge (\delta^\lambda_{srlg_4} = 80) \wedge (\delta^\lambda_{srlg_6} = 80) \\ \forall\, srlg_i : (srlg_i \neq srlg_2 \wedge srlg_i \neq srlg_4 \wedge srlg_i \neq srlg_6) : \delta^\lambda_{srlg_i} = 0 \end{cases}$$

A cause de la localisation du *(3+1)^{eme}* plus grand prix de protection de SRLG dans la zone sûre (absence de l'identifiant *SRLG_ générique* dans le *3_ vecteur*), une entité $BPCE_\alpha$ recevant le *3_ vecteur* précédent, déduira sans erreur si l'arc λ peut être sélectionné ou pas lors d'un calcul d'un nouveau LSP de secours. En effet, pour un LSP de secours protégeant contre la panne d'un lien α-β qui apparaît dans un SRLG de l'ensemble {$srlg_2$, $srlg_4$, $srlg_6$}, l'entité $PBCE_\alpha$ pourra déterminer le coût de protection du lien α-β sur l'arc λ puisqu'elle connaît le plus grand prix des SRLG contenant le lien α-β sur l'arc λ (ce coût sera égal à 90 unités si le lien α-β appartient à $srlg_2$, 80 unités sinon). Pour un LSP de secours b protégeant contre la panne d'un lien n'appartenant à aucun SRLG de l'ensemble {$srlg_2$, $srlg_4$, $srlg_6$}, $PBCE_\alpha$ sélectionnera sans erreur l'arc λ lors du calcul de ce LSP b puisque :

$$\left.\begin{array}{l} Max_{srlg \notin \{srlg_2,\, srlg_4,\, srlg_6\}}\, \delta^\lambda_{srlg} \leq Sc_\lambda) \\ Sc_\lambda \leq BC_\lambda - bw\,(b)\ (hypothèse) \end{array}\right\} \Rightarrow Max_{srlg \notin \{srlg_2,\, srlg_4,\, srlg_6\}}\, \delta^\lambda_{srlg} \leq BC_\lambda - bw\,(b)$$

Comme nous le constatons, la diffusion d'un vecteur x^{λ}_ *vecteur* d'une taille maximale égale à 3 est suffisante pour décider sans erreur si l'arc λ peut être sélectionné ou pas dans le calcul d'un nouveau LSP de secours. En conséquence, toute augmentation du paramètre x^{λ} est inutile (i.e. elle n'améliore pas la qualité du prochain LSP de secours calculé) et ne provoque aucune augmentation de la taille réelle du x^{λ}_ *vecteur* diffusé.

Concernant l'information diffusée dans le réseau, elle est restreinte aux différents vecteurs $\{x^{\lambda}$_ *vecteurs*$\}_{\lambda \in E}$. Comme les valeurs de ces derniers ne changent pas systématiquement, à chaque établissement d'un nouveau LSP de secours, la fréquence de diffusion est diminuée. De même, comme la taille de tout x^{λ}_ *vecteur* est limitée à x^{λ} couples au plus, la taille de l'information diffusée dans le réseau est aussi réduite considérablement (en d'autres termes, les tailles des états de liens sont petites, cf. section 3.3.2.2).

Pour obtenir de meilleures performances avec cette heuristique DBSH, il serait intéressant d'attribuer aux paramètres $\{x^{\lambda}\}_{\lambda \in E}$ des valeurs dépendant de la charge des liens $\{\lambda\}_{\lambda \in E}$ associés. Ainsi, un arc λ très utilisé pour la protection de différents SRLG peut attribuer une valeur élevée au paramètre x^{λ} alors qu'un arc λ' peu utilisé dans la protection contre les pannes de SRLG peut fixer la valeur de son paramètre $x^{\lambda'}$ à une valeur très basse. Ces valeurs (paramètres $\{x^{\lambda}\}_{\lambda \in E}$) ne sont d'ailleurs pas statiques et peuvent donc changer d'une période à une autre, pour mieux s'adapter à la matrice de trafic. Cependant, pour déterminer toujours et sans erreur les arcs à inclure ou à exclure lors d'un calcul d'un nouveau LSP de secours, les paramètres $\{x^{\lambda}\}_{\lambda \in E}$ devraient être fixés à l'infini[10].

Enfin, nous notons que, lorsqu'une requête d'établissement d'un nouveau LSP de secours ne peut pas être satisfaite pour manque de bande passante, l'heuristique DBSH permet de préempter et/ou de réorganiser les LSP de secours afin de libérer de la bande passante et accepter la nouvelle requête.

3.3.2 Extensions des protocoles de signalisation et IGP-TE pour le support de l'heuristique DBSH

Pour permettre l'implantation de l'heuristique DBSH, nous proposons ici de légères extensions aux protocoles de signalisation et aux protocoles IGP-TE. Comme pour l'algorithme TDRA, nous ne décrirons dans cette thèse que les modifications à apporter au protocole de signalisation RSVP-TE et aux deux protocoles OSPF-TE et ISIS-TE pour supporter l'heuristique DBSH.

3.3.2.1 Extensions du protocole RSVP-TE

Pour placer un LSP de secours b protégeant contre la panne d'un lien $\beta \rightarrow \alpha$, une communication entre le nœud α supportant l'entité de calcul $BPCE_{\alpha}$ et le nœud PLR β est nécessaire pour la transmission des requêtes de calcul et la réception des résultats.

[10] Pour le seuil de confiance Sc^{λ}, il peut prendre n'importe quelle valeur inférieure ou égale à la capacité de secours de l'arc λ moins la quantité maximale de bande passante que peut réclamer un LSP.

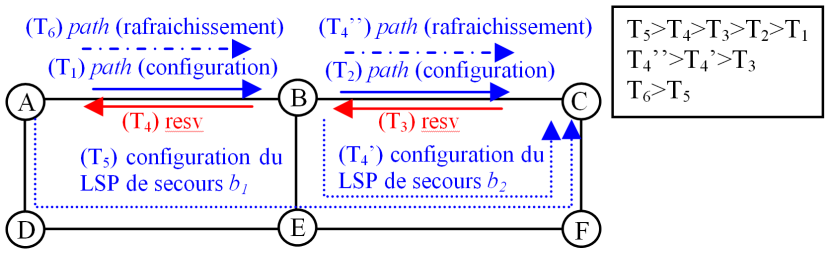


FIG. 3.8 – Configuration d'un LSP primaire et de ses LSP de secours avec l'heuristique DBSH

Typiquement, lorsque le nœud PLR β reçoit une requête d'établissement d'un nouveau LSP de secours protégeant contre la panne du lien $\beta \rightarrow \alpha$, il la redirige vers le nœud α. Ce dernier fait appel à l'entité $BPCE_\alpha$ qui calcule le LSP de secours et envoie ensuite les résultats au nœud PLR β. Ce dernier configure alors le LSP de secours déterminé et envoie ensuite une indication de la réussite ou de l'échec de la procédure de configuration au nœud α.

Avec de légères extensions du protocole de signalisation RSVP-TE, cette communication entre BPCE et PLR peut être gérée efficacement et sans introduction de nouveaux messages. Pour cela, nous proposons d'ajouter à RSVP-TE un nouvel objet BACKUP_PATH (comme pour TDRA) qui sera transporté dans les messages *path* et *resv* de configuration du LSP primaire protégé. Ainsi, tout nœud du LSP primaire (différent de la source) calculera un LSP de secours protégeant contre la panne de son lien en amont (et éventuellement contre la panne de son nœud en amont) et construira un objet BACKUP_PATH contenant la structure et les propriétés du LSP de secours déterminé. Cet objet sera inséré dans le message *resv* qui sera envoyé au nœud en amont sur le LSP primaire.

Lorsqu'un nœud PLR reçoit un message *resv* contenant un objet BACKUP_PATH, il accomplit les traitements standards décrits dans premier chapitre. Ensuite (si tout se passe bien), ce nœud PLR extrait du message l'objet BACKUP_PATH et initie la configuration du LSP de secours contenu dans cet objet. Au prochain rafraîchissement du LSP primaire qui suit la fin de la configuration du LSP de secours, le nœud PLR insérera l'objet BACKUP_PATH relatif au LSP de secours configuré dans le message *path* relatif au LSP primaire, afin d'indiquer la réussite de la procédure de configuration du LSP de secours (autrement, le nœud PLR n'insère aucun objet BACKUP_PATH). Bien évidemment, les valeurs des prix de protection des liens et nœuds ne sont mise-à-jour qu'après la réception d'un message *path* contenant l'objet BACKUP_PATH.

Sur la figure 3.8 sont illustrés les différents messages échangés entre les nœuds du réseau pour configurer un LSP primaire p_1 ($A \rightarrow B \rightarrow C$) et ses deux LSP de secours b_1 ($A \rightarrow D \rightarrow E \rightarrow F \rightarrow C$) et b_2 ($B \rightarrow E \rightarrow F \rightarrow C$). Initialement, le nœud A déclenche la procédure de configuration du LSP primaire protégé p_1 en envoyant un message *path* au nœud B (ne contenant pas d'objet BACKUP_PATH). Ce dernier traite (modifie) et redirige le message *path* au nœud C qui est le nœud de destination du LSP primaire p_1. A la réception du message *path* par le nœud C, ce dernier effectue un contrôle

d'admission sur l'arc $B \rightarrow C$, calcule un LSP de secours b_2 protégeant contre la panne du lien $B \rightarrow C$ et construit ensuite un message *resv* incluant un objet BACKUP_PATH (cet objet contient la structure et les propriétés du LSP de secours b_2). Ce message sera envoyé au nœud B. A la réception du message par le nœud B, ce dernier calcule un LSP de secours b_1 protégeant contre la panne du nœud B lui-même et du lien $A \rightarrow B$ et construit un nouvel objet BACKUP_PATH contenant la structure et les propriétés du LSP de secours b_1. Ce dernier objet sera inséré dans le prochain message *resv* qu'il enverra au nœud A. En plus de ces traitements, le nœud B déduira du message *resv* reçu la structure et les propriéts du LSP de secours b_2 qu'il configurera par la suite. Au prochain rafraîchissement du LSP primaire, le nœud B inclura l'objet BACKUP_PATH dans les messages *path* pour indiquer la réussite de la configuration du LSP de secours b_2. De la même manière, le nœud A configurera le LSP de secours b_1 contenu dans l'objet BACKUP_PATH reçu et inclura cet objet dans les messages *path* (rafraîchissant le LSP primaire) suivant la réussite de la configuration du LSP de secours b_1.

3.3.2.2 Extensions des protocoles OSPF-TE et ISIS-TE

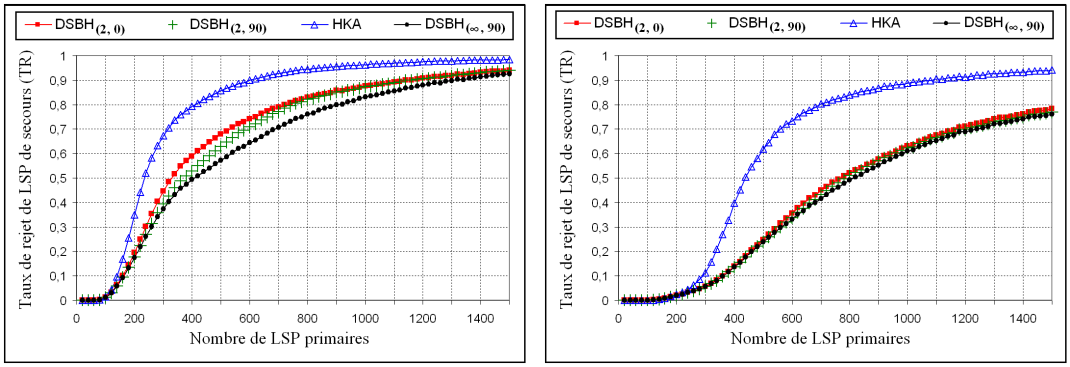
Afin d'annoncer les capacités de secours[11] et les vecteurs $\{x^{\lambda}\_\ vecteurs\}_{\lambda \in E}$, nous proposons d'étendre les protocoles IGP-TE. Comme dans la section 3.2.2.2, nous proposons d'utiliser les paramètres de l'ingénierie de trafic définis dans OSPF-TE et dans ISIS-TE pour annoncer les valeurs de ces deux paramètres. Pour cela, nous suggérons de définir un nouveau champ au format TLV associé à chaque arc. Ce champ sera transmis dans les *LSA* pour OSPF-TE et dans les *LSPDU* pour ISIS-TE.

3.3.3 Évaluation des performances

Afin de mesurer les performances de l'heuristique DBSH, nous l'avons comparé à l'heuristique HKA (nous avons choisi HKA pour les mêmes raisons que dans la section 3.2.3). Pour ces simulations, nous avons adopté le même environnement que celui que nous avons décrit dans la section 3.2.3. Typiquement, nous avons sélectionné les mêmes topologies de réseau pour nos tests et nous avons généré les matrices de trafic en adoptant le même procédé que dans la section 3.2.3.

Afin de mesurer l'impact du choix des paramètres x^{λ} et Sc^{λ} sur les performances de l'heuristique DBSH, nous avons sélectionné trois variantes de DBSH, à savoir $DBSH_{(\infty,90)}$, $DBSH_{(2,0)}$ et $DBSH_{(2,90)}$. La première variante ($DBSH_{(\infty,90)}$) emploie sur chaque arc λ un seuil de confiance égal à 90% de la capacité de l'arc (i.e. $\forall \lambda : Sc^{\lambda} = 0.9 \times BC^{\lambda}$) et elle utilise un $x^{\lambda}\_\ vecteur$ potentiellement infini (i.e. $\forall \lambda : x^{\lambda} = \infty$). Cette variante permet d'inclure (resp. d'exclure) sans erreur tous les arcs vérifiant (resp. ne vérifiant pas) les contraintes de bande passante lors du calcul d'un nouveau LSP de secours. La seconde variante ($DBSH_{(2,0)}$) n'emploie pas de seuil de confiance (seuils de confiance nuls) et elle utilise un $x^{\lambda}\_\ vecteur$ d'une taille maximale égale à 2 (i.e. $\forall \lambda : x^{\lambda} = 2$).

[11] L'annonce de la capacité de secours d'un arc λ n'est nécessaire que si les capacité de l'arc λ est divisée en deux pools disjoints (pool primaire et pool de secours).



(a) Topologie du réseau de 95 risques (b) Topologie du réseau de 162 risques

FIG. 3.9 – Evolution du taux de rejet des LSP de secours (TR)

Cette variante peut être utile lorsque nous nous disposons d'aucune information concernant la quantité maximale que peuvent réclamer les LSP de secours. Enfin, la dernière variante ($DBSH_{(2,90)}$) emploie sur chaque arc λ un seuil de confiance égal à 90% de la capacité de l'arc (i.e. $\forall\lambda : Sc^{\lambda} = 0.9 \times BC^{\lambda}$) et elle utilise un $x^{\lambda}\_vecteur$ d'une taille maximale égale à 2 (i.e. $\forall\lambda : x^{\lambda} = 2$). La comparaison des performances de cette variante avec la variante $DBSH_{(2,0)}$ permet de quantifier le gain obtenu en renseignant la valeur du seuil de confiance lors du placement des LSP de secours avec DBSH.

3.3.3.1 Métriques

Les mêmes métriques que celles choisies dans la section 3.2.3.1 sont utilisées pour comparer DBSH à HKA, à l'exception de la métrique *NMM* qui est remplacée par la métrique *NVD*. Cette dernière métrique correspond au nombre moyen de vecteurs $\{x^{\lambda}\_vecteur_{\lambda \in E}\}$ (i.e. nombre d'état de lien) diffusés dans le réseau à chaque établissement d'un nouveau LSP de secours. Bien évidemment, plus la valeur de la métrique *NVD* est grande, plus est élevé le volume du trafic de contrôle nécessaire au placement des LSP de secours.

Comme dans la section 3.2.3.1, nous mesurons les valeurs des trois métriques précédentes pour différentes charges du réseau. Ainsi et à chaque établissement de 20 LSP primaires, les valeurs des métriques *TR*, *NVD* et *TUBS* sont déterminées et retournées pour les trois variantes de DBSH ainsi que pour l'heuristique HKA. Tous les résultats illustrés dans la section 3.3.3.2 correspondent aux valeurs moyennes obtenues après 1000 exécutions et tirages aléatoires.

3.3.3.2 Résultats et analyse

La figure 3.9 illustre l'évolution du taux de rejet (*TR*) en fonction du nombre de LSP primaires déjà établis dans le réseau. Comme nous le constatons, l'heuristique DBSH a toujours un meilleur et plus petit taux de rejet par rapport à l'heuristique HKA. Cela s'explique par la qualité de l'approximation des prix de protection qui est meilleure

avec DBSH (par rapport à HKA). Alors que l'heuristique HKA approxime, sur tout arc, les prix de protection des risques de type lien ou nœud par le maximum des prix de protection, l'heuristique DBSH n'utilise aucune approximation pour ce type de risque (puisqu'elle dispose des valeurs exactes de ces prix de protection). Pour les risques de type SRLG, la qualité de l'approximation utilisée dans DBSH est toujours meilleure que celle employée dans HKA puisque le x^{eme} plus grand prix de protection de SRLG (valeur approximée d'un prix de protection de SRLG dans DBSH) est toujours inférieur au maximum des prix de protection (valeur approximée d'un prix de protection dans HKA) sur tout arc.

Bien évidemment et comme expliqué dans la section 3.2.3.2, l'heuristique HKA et l'heuristique DBSH disposent de taux de rejet très similaires pour les petites charges du réseau (nombre de LSP primaires inférieur à 140 dans la figure 3.9 (a) et inférieur à 240 dans la figure 3.9 (b)). Ceci est dû aux valeurs des quantités de bande passante de secours allouées sur les arcs qui sont très petites (c'est-à-dire les valeurs des plus grands prix de protection sur les arcs sont petites), ce qui entraîne la sélection de (presque) tous les arcs lors du calcul des LSP de secours (taux de rejet proche de zéro).

Lorsque la charge du réseau augmente, le taux de rejet des LSP de secours de l'heuristique HKA croît plus rapidement que ceux correspondant aux différentes variantes de DBSH, à cause de la qualité d'approximation des prix de protection employée par DBSH qui est meilleure que celle adoptée par HKA (cf. premier paragraphe de cette section).

Concer nant les trois variantes $DBSH_{(2,0)}$, $DBSH_{(2,90)}$ et $DBSH_{(\infty,90)}$, nous apercevons qu'elles disposent de taux de rejet très proches les uns des autres. Ceci est dû à l'algorithme de calcul des LSP de secours qui tend naturellement à utiliser des liens proches du composant à protéger, ce qui réduit le nombre de prix de protection supérieurs au seuil sur les arcs (asymétrie dans la distribution des prix de protection). Dans notre modèle, l'envoi de, uniquement, deux prix de protection de SRLG par arc suffit pour réduire significativement la probabilité d'erreur lors de la décision d'inclusion/exclusion d'un arc dans le calcul d'un nouveau LSP de secours. Néanmoins, de légères différences existent entre les trois variantes de DBSH. Ainsi, la variante $DBSH_{(\infty,90)}$ a un taux de rejet de LSP de secours légèrement inférieur et meilleur que ceux des deux autres variantes puisque $DBSH_{(\infty,90)}$ dispose de toute l'information nécessaire à la décision d'inclusion/excluions, sans erreur, de tout arc lors du calcul d'un nouveau LSP de secours. De même, $DBSH_{(2,90)}$ a un taux de rejet de LSP de secours légèrement inférieur à celui de $DBSH_{(2,0)}$ puisque l'approximation utilisée dans $DBSH_{(2,90)}$ est meilleure lorsque le troisième plus grand prix de SRLG est inférieur au seuil alors que le deuxième plus grand prix de SRLG est supérieur au seuil.

L'autre point important que nous observons sur la figure 3.9 concerne la différence de distance entre les trois courbes des variantes de DBSH selon que c'est la première (figure 3.4 (a)) ou deuxième (figure 3.4 (b)) topologie de réseau qui est employée. Concrètement, la distance entre les courbes des variante de DBSH est plus petite sur la topologie du réseau (figure 3.4 (b)) pour deux raisons essentielles :

1. La densité des SRLG est plus petite dans le réseau de la figure 3.4 (b).

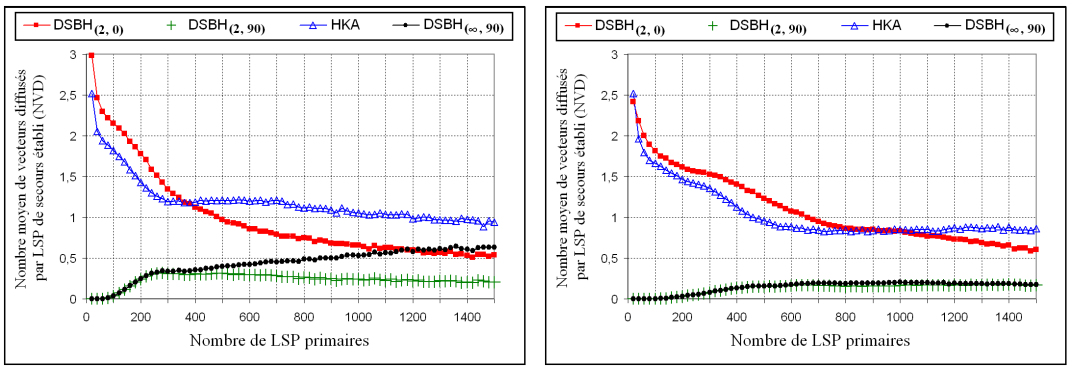
2. Les taux de rejet de LSP de secours sont plus petits sur la topologie de la figure 3.4 (b) (le degré moyen des nœuds 3.21 dans la première topologie de réseau est inférieur à celui 3.48 qui correspond à la seconde topologie de réseau). Ceci induit une diminution légère des prix de protection et réduit en conséquence le nombre de prix de SRLG supérieurs aux seuils.

A propos de la seconde métrique *NVD*, nous observons sur la figure 3.10 que les nombres moyens de vecteurs diffusés dans le réseau avec les deux variantes $DBSH_{(\infty,90)}$ et $DBSH_{(2,90)}$ sont souvent inférieurs à ceux de $DBSH_{(2,0)}$ et HKA. Ceci s'explique par l'utilisation d'un seuil de confiance élevé dans les deux premières variantes ($DBSH_{(\infty,90)}$ et $DBSH_{(2,90)}$) , ce qui permet d'éliminer la diffusion d'un grand nombre de prix de protection de SRLG (ceux inférieurs aux seuils).

Lorsque le seuil de confiance n'est pas utilisé (seuil de confiance nul) comme dans $DBSH_{(2,0)}$, la fréquence d'envoi des vecteurs $\{x^{\lambda}\_vecteur_{\lambda \in E}\}$ augmente sensiblement mais reste comparable à celle correspondant à l'heuristique HKA. Cela veut dire que, dans notre modèle de simulation, les deux plus grands prix de protection de SRLG changent en moyenne aussi vite que le maximum de tous les prix de protection.

Notons aussi la similitude des valeurs de *NVD* des deux variantes $DBSH_{(\infty,90)}$ et $DBSH_{(2,90)}$ (superposition des courbes de *NVD* sur la figure 3.10 (a) lorsque le nombre de LSP primaires est inférieur à 400 et superposition des courbes de *NVD* sur la figure 3.10 (b) pour toutes les charges du réseau). Ceci explique la ressemblance des taux de rejet correspondant à ces deux variantes sur la figure 3.9. En effet, les informations diffusées avec les deux variantes $DBSH_{(\infty,90)}$ et $DBSH_{(2,90)}$ sont assez semblables, ce qui permet de calculer presque les mêmes LSP de secours.

Concernant la dernière métrique *TUBS*, la figure 3.11 montre que les taux d'utilisation de la bande passante de secours des variantes de DBSH sont assez proches de celui qui correspond à l'heuristique HKA lorsque la charge réseau est faible (i.e. le nombre de LSP primaires protégés est inférieur à 200 dans le réseau de la figure 3.3 (a) et inférieur à 400 dans le réseau de la figure 3.3 (b)). Ceci s'explique essentiellement par les taux



(a) Topologie du réseau de 95 risques (b) Topologie du réseau de 162 risques

FIG. 3.10 – Evolution du nombre moyen de vecteurs transmis dans le réseau pour l'établissement d'un LSP de secours (NVD)

de rejet de LSP de secours qui sont faibles pour ces charges, ce qui permet d'accepter, à peu près, les mêmes LSP de secours.

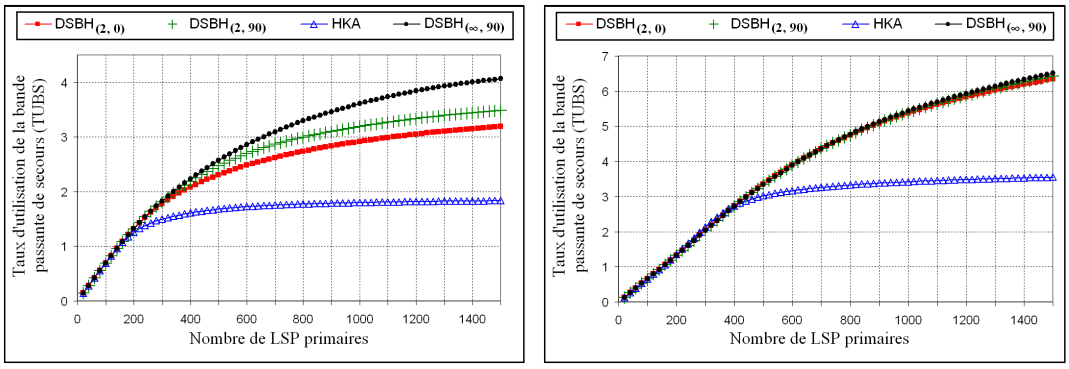
Pour les charges élevées du réseau, la figure 3.11 montre que les taux d'utilisation de la bande passante de secours des variantes de DBSH sont meilleurs et plus élevés que celui de HKA. Ceci est dû à la qualité d'approximation des prix de protection qui est meilleure avec DBSH par rapport à HKA.

A cause de l'insuffisance de la diffusion de deux prix de protection de SRLG par arc (voir la figure 3.10 (a)), la figure 3.11 (a) montre que les taux d'utilisation de la bande passante de secours des variantes $DBSH_{(2,90)}$ et $DBSH_{(2,0)}$ sont plus petits que celui qui correspond à la variante $DBSH_{(\infty,90)}$, lorsque le nombre de LSP primaires dépasse 400. Sur la figure 3.11 (b) par contre, nous apercevons que les taux d'utilisation de la bande passante de secours des trois variantes de DBSH sont quasi-identiques. Ceci s'explique par la suffisance de la diffusion des deux plus grands prix de SRLG par arc, comme c'est illustré sur la figure 3.10 (b).

3.3.4 Conclusion

L'heuristique DBSH permet un placement distribué des LSP de secours avec la diffusion d'un $x^{\lambda}\_vecteur$ pour tout arc λ de la topologie du réseau. Ce $x^{\lambda}\_vecteur$, qui contient les x^{λ} plus grand prix de SRLG, permet souvent de bien estimer les différents prix de SRLG sur les arcs, ce qui réduit la probabilité de blocage. De plus, les valeurs des $\{x^{\lambda}\_vecteur\}_{\lambda \in E}$ sont relativement stables (par rapport aux vecteurs contenant tous les prix de protection) et d'une taille limitée, ce qui réduit la fréquence et la quantité d'informations diffusées dans le réseau.

Pour permettre le déploiement de l'heuristique DBSH, quelques modifications des protocoles existants peuvent être nécessaires. Premièrement, afin de gérer la communication entre les entités de calcul des LSP de secours et les différents PLR, nous avons proposé d'étendre le protocole RSVP-TE. Ainsi, un nouvel objet (BACKUP_PATH) est ajouté dans les messages *resv* (resp. *path*) de configuration du LSP primaire afin de



(a) Topologie du réseau de 95 risques (b) Topologie du réseau de 162 risques

FIG. 3.11 – Evolution du taux d'utilisation de la bande passante de secours (TUBS)

communiquer aux PLR les structures des LSP de secours calculés par les BPCE (resp. afin de permettre aux BPCE de recevoir des PLR les confirmations d'établissement des LSP de secours). Deuxièmement, pour distribuer les prix de protection de SRLG aux entités de calcul, nous avons proposé d'ajouter, aux états de liens des protocoles OSPF-TE ou ISIS-TE, un champ TLV contenant les vecteurs $\{x^{\lambda}\_vecteur\}_{\lambda \in E}$.

Enfin, la comparaison de l'heuristique DBSH avec l'heuristique de Kini améliorée montre que la première heuristique offre de meilleures capacités de partage (taux de rejet des LSP de secours faible et taux d'utilisation de la bande passante de secours plus élevé) et une fréquence de diffusion de messages IGP-TE plus petite notamment lorsque les valeurs des paramètres ($\{(x^{\lambda}, Sc^{\lambda})\}_{\lambda \in E}$) de l'heuristique DBSH sont bien choisis.

3.4 Heuristique de placement de LSP de secours basée sur les PLR (PLRH)

Grâce à une répartition asymétrique et/ou partielle des prix de protection sur plusieurs entités de calcul (i.e. les entités de calcul ne disposent pas de la même information concernant les prix de protection), l'algorithme TDRA et l'heuristique DBSH diminuent sensiblement la taille de l'information transmise dans le réseau pour le placement des LSP de secours. Bien qu'elle soit pratique, cette répartition asymétrique et/ou partielle des prix de protection sur plusieurs entités de calcul induit divers inconvénients comme :

1. nécessité d'une communication supplémentaire entre PLR et BPCE pour la transmission des requêtes de placement des LSP de secours et pour la réception des résultats,
2. pas de possibilité de calcul conjoint de l'ensemble des LSP de secours protégeant un même LSP primaire, ni même d'un calcul conjoint d'un LSP primaire avec ses LSP de secours. Cela se traduit par l'incapacité de fournir un routage primaire protégé sûr (i.e. aucune garantie de détermination de LSP primaires protégés à 100%).
3. difficulté accrue de protection inter-domaine. Typiquement, la protection contre la panne d'un risque nécessitant l'établissement d'un LSP de secours multi-domaines n'est souvent pas possible (les entités de calcul n'ont connaissance que de la topologie du réseau de leur domaine).

Pour faire face aux inconvénients ci-dessus, nous proposons, dans cette section, l'heuristique PLRH (*PLR-based Heuristic*) [SCLRb, SCLR09a] qui permet à tous les nœuds du réseau de disposer de la même vision des prix de protection. Concrètement, en étendant l'ensemble des prix de protection diffusés dans le réseau à ceux correspondant à tout type de risque de panne, il sera possible d'estimer tous les prix de protection des risques sur les différents arcs du réseau, en utilisant le même mécanisme d'approximation que celui adopté par DBSH.

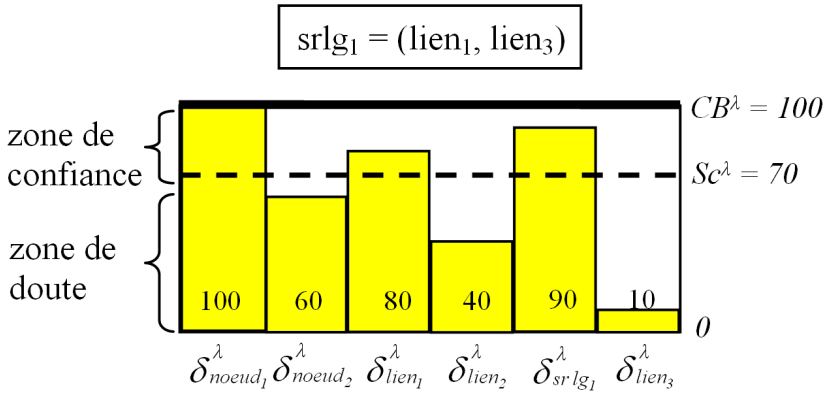


FIG. 3.12 – Prix de protection des risques de panne sur un arc λ

3.4.1 Principes de l'heuristique PLRH

L'heuristique PLRH utilise un procédé très proche de celui adopté par l'heuristique DBSH afin de déduire les valeurs des prix de protection. Précisément, dans l'heuristique PLRH, les prix de protection sont estimés grâce à la diffusion de vecteurs, dits $y^{\lambda}\_vecteurs$, pour chaque arc λ du réseau. Ces vecteurs contiennent $y^{\lambda}_{(0<y^{\lambda}<\infty)}$ couples (au plus) formés des prix de protection les plus élevés sur l'arc λ et des risques correspondants.

Comme dans l'heuristique DBSH, le vecteur diffusé pour tout arc donné λ ne doit contenir que des prix de protection dont la valeur est supérieure au seuil de confiance Sc^{λ}. Ce dernier doit être déterminé de la même manière que dans DBSH, c'est-à-dire il doit être fixé de telle sorte que la quantité maximale de bande passante que peut réclamer tout LSP de secours soit inférieure ou égale à la quantité BC^{λ} - Sc^{λ}.

Sur la figure 3.12 sont illustrés les prix de protection de tous les risques protégés en utilisant l'arc λ. Afin de déterminer le vecteur $y^{\lambda}\_vecteur$ à diffuser par le nœud d'extrémité sortante o^{λ} de l'arc λ, le nœud o^{λ} construit d'abord une liste initiale Li^{λ} ($Li^{\lambda} = [(100, nœud_1), (60, nœud_2), (80, lien_1), (40, lien_2), (90, srlg_1), (10, lien_3)]$) formée de couples contenant les différents prix de protection des risques protégés en utilisant l'arc λ ainsi que les identifiants associés à ces risques. Pour ce faire, le nœud o^{λ} consulte la liste des LSP de secours traversant l'arc λ (information déduite à partir des données transmises par les protocoles de signalisation).

De la liste Li^{λ}, le nœud o^{λ} supprime tous les couples associés à des liens appartenant à des SRLG. Sur l'exemple de la figure 3.12, les deux liens $lien_1$ et $lien_3$ appartiennent au SRLG $srlg_1$. En conséquence, les deux couples ($lien_1$, 80) et ($lien_3$, 10) seront supprimés de la liste initiale Li^{λ}. Ensuite, le nœud o^{λ} ordonnera la liste Li^{λ} suivant les valeurs décroissantes des prix de protection. Ainsi la liste finale Lf^{λ} sera la suivante : [($nœud_1$, 100), ($srlg_1$, 90), ($nœud_2$, 60), ($lien_2$, 40)]. C'est cette liste finale qui sera exploitée par l'algorithme 5 afin de construire le vecteur $y^{\lambda}\_vecteur$ diffusé dans le réseau. Nous notons que l'algorithme 5 est quasi-identique à l'algorithme 3 : il ne lui diffère que

Algorithme 5 Routine PLRH exécutée par tout nœud d'extrémité sortante o^λ d'un arc λ

*ancien_*y^λ*_vecteur* $\leftarrow 0$
tantque *vrai* **faire**
  Affecter au vecteur *nouveau_*y^λ*_vecteur*, associé à l'arc λ, les y^λ *plus grands prix de protection* ainsi que les identifiants de risques de panne associés
  si le $(y^\lambda+1)^{eme}$ *plus grand prix de protection* est supérieur au seuil Sc^λ **alors**
    Remplacer, dans le vecteur *nouveau_*y^λ*_vecteur*, l'identifiant du $(y^\lambda)^{eme}$ plus grand *prix de protection* par l'identifiant (spécial) *risque_générique*
  finsi
  si *nouveau_*y^λ*_vecteur* $\neq$ *ancien_*y^λ*_vecteur* **alors**
    Diffuser le vecteur *nouveau_*y^λ*_vecteur*
    *nouveau_*y^λ*_vecteur* $\leftarrow$ *ancien_*y^λ*_vecteur*
  finsi
fin tantque

dans la liste des prix de protection gérés et diffusés dans les vecteurs $\{y^\lambda\_vecteur\}_{\lambda \in E}$. Alors que l'algorithme 3 ne tient compte que des risques de panne de type SRLG, l'algorithme 5 explore tous les risques de panne, quel que soit leur type.

Lorsqu'un nœud non adjacent à un arc λ reçoit un vecteur $y^\lambda\_vecteur$, il exécutera l'algorithme 6 afin d'approximer les prix de protection des risques sur cet arc λ. Nous notons que cet algorithme est quasi-identique à l'algorithme 4. La seule différence entre ces deux algorithmes réside dans l'ensemble de risques traités par l'algorithme 6 qui inclut en plus les risques de type nœud et les risques de type lien.

Exemple :

Sur l'arc λ de la figure 3.12, la diffusion d'un $y^\lambda\_vecteur$ de 2 couples ($y^\lambda\_vecteur$ = [($nœud_1$, 100), ($srlg_1$, 90)]) est suffisante pour que tous les nœuds puissent décider

Algorithme 6 Routine PLRH exécutée par tout nœud recevant un vecteur $y^\lambda\_vecteur$ associé à l'arc λ

si *risque_générique* $\in$ y^λ*_vecteur.**index** ()* **alors**
  prix_min $\leftarrow$ y^λ*_vecteur* [*risque_générique*]
sinon
  prix_min $\leftarrow 0$
finsi
pour chaque identifiant *id* d'un risque **faire**
  si *id* $\in$ y^λ*_vecteur.**index** ()* **alors**
    table_locale_prix_protection [λ][*id*] $\leftarrow$ y^λ*_vecteur* [*id*]
  sinon
    table_locale_prix_protection [λ][*id*] $\leftarrow$ *prix_min*
  finsi
fin pour

sans erreur de l'inclusion ou de l'exclusion de cet arc λ dans le calcul d'un nouveau LSP de secours. En effet, tout nœud recevant le vecteur [($n\text{œ}ud_1$, 100), ($srlg_1$, 90)] approximera les prix de protection de tous les risques comme suit (cf. algorithme 6) :
$\begin{cases} (\delta^{\lambda}_{n\text{œ}ud_1} = 100) \wedge (\delta^{\lambda}_{srlg_1} = 90) \\ \forall r\,(r \neq n\text{œ}ud_1 \wedge r \neq srlg_1) : \delta^{\lambda}_{r} = 0 \end{cases}$

En conséquence, tout calcul d'un LSP de secours protégeant contre la panne du nœud $noeud_1$ rejettera (exclura) sans erreur l'arc λ. De même, lors du calcul d'un LSP de secours réclamant une quantité de bande passante supérieure à 10 unités et protégeant contre la panne du $lien_1$ ou $lien_3$ (liens appartenant dans $srlg_1$), l'arc λ sera rejeté sans erreur. Cependant, tout autre calcul de LSP de secours (ne vérifiant pas les conditions précédentes) inclura sans erreur l'arc λ.

Comme il est possible d'estimer les valeurs de tous les prix de protection sur tous les arcs grâce à l'information (les vecteurs $\{y^{\lambda}\_\, vecteur\}_{\lambda \in E}$) diffusée dans le réseau, le placement des LSP de secours pourrait être effectué par n'importe quel nœud du réseau. Cela à divers avantages comme :

1. élimination de la communication entre les BPCE et les PLR puisque les calculs pourront être effectués par les PLR eux-mêmes,
2. augmentation du taux de protection des LSP primaires (les nœuds ont la faculté de choisir les LSP primaires qui maximisent la protection),
3. possibilité de protection inter-domaines (il suffira de partager les structures des risques de panne pour permettre la protection inter-domaine).

Notons enfin qu'il est possible de transmettre, dans les vecteurs $\{y^{\lambda}\_\, vecteur\}_{\lambda \in E}$, les coûts de protection des liens et nœuds au lieu des prix de protection. Cela pourrait permettre d'améliorer les performances du mécanisme de placement des LSP de secours, notamment lorsque la densité des SRLG dans le réseau est très élevée. En effet, le nombre de prix de protection définis sur un arc croît avec l'augmentation du nombre de SRLG (puisqu'il est égal à la somme des nombres de SRLG, de liens et de nœuds du réseau) alors que le nombre de coûts de protection définis sur un arc reste constant avec l'augmentation du nombre de SRLG (puisqu'il est égal au nombre de liens et nœuds du réseau).

3.4.2 Extensions des protocoles IGP-TE pour l'implantation de l'heuristique PLRH

Un des avantages majeurs de l'heuristique PLRH est sa facilité à être déployée. En effet, cette heuristique ne nécessite que de très légères extensions aux protocoles IGP-TE existants (OSPF-TE et ISIS-TE) pour sa mise en œuvre.

Comme pour DBSH, nous proposons de définir un nouveau champ au format TLV pour le transport des vecteurs de prix de protection ($\{y^{\lambda}\_\, vecteur\}_{\lambda \in E}$) et éventuellement des capacités de secours. Ce champ sera transmis dans les *LSA* pour OSPF-TE et dans les *LSPDU* pour ISIS-TE.

Notons qu'à cause de l'élimination de la communication entre BPCE et PLR, aucune extension des protocoles de signalisation n'est nécessaire.

3.4.3 Évaluation des performances

Pour mesurer les performances de l'heuristique PLRH, nous l'avons comparée à l'heuristique HKA en adoptant le même modèle et le même environnement de simulation que celui de la section 3.3.3 (pour les mêmes raisons que dans les sections 3.2.3 et 3.3.3, le même procédé de génération de la matrice de trafic et les mêmes topologies de réseau ont été sélectionnés).

Comme dans la section 3.3.3, nous avons sélectionné diverses variantes de l'heuristique PLRH afin de mesurer l'impact du choix des paramètres y^λ et Sc^λ sur les performances de l'heuristique. Dans la première variante, notée $PLRH_{(\infty,90)}$, nous avons utilisé des vecteurs de taille illimitée et nous avons fixé le seuil de confiance Sc^λ sur chaque arc λ à 90% de la capacité de l'arc (i.e. $\forall\lambda : Sc^\lambda = 0.9\times BC^\lambda$). Dans la seconde variante, notée $PLRH_{(\infty,0)}$, nous avons utilisé des vecteurs de taille illimité et nous avons choisi un seuil de confiance nul sur tous les arcs. Dans la troisième (resp. quatrième, cinquième) variante, notée $PLRH_{(2,0)}$ (resp. $PLRH_{(5,0)}$, $PLRH_{(5,90)}$), nous avons limité la taille maximale des vecteurs à 2 (resp. 5, 5) et nous avons fixé le seuil de confiance à 0 (resp. 0, 90% de la capacité de l'arc associé).

Nous soulignons que les variantes ($PLRH_{(\infty,90)}$ et $PLRH_{(\infty,0)}$) utilisant des vecteurs de taille illimitée ont une connaissance complète de l'information requise au placement des LSP de secours. En conséquence, ces variantes permettent d'éviter les fuites de bande passante en maximisant le partage de bande sur les arcs. De même, nous notons que les variantes ($PLRH_{(\infty,0)}$, $PLRH_{(2,0)}$ et $PLRH_{(5,0)}$) n'employant pas le seuil de confiance (i.e. seuil de confiance nul) ne sont utiles que lorsque nous nous disposons pas d'informations concernant la quantité maximale que peut réclamer un LSP de secours. Dans le cas contraire, les deux autres variantes $PLRH_{(\infty,90)}$ et $PLRH_{(5,90)}$ sont mieux adaptées pour réduire la quantité d'informations diffusées dans le réseau.

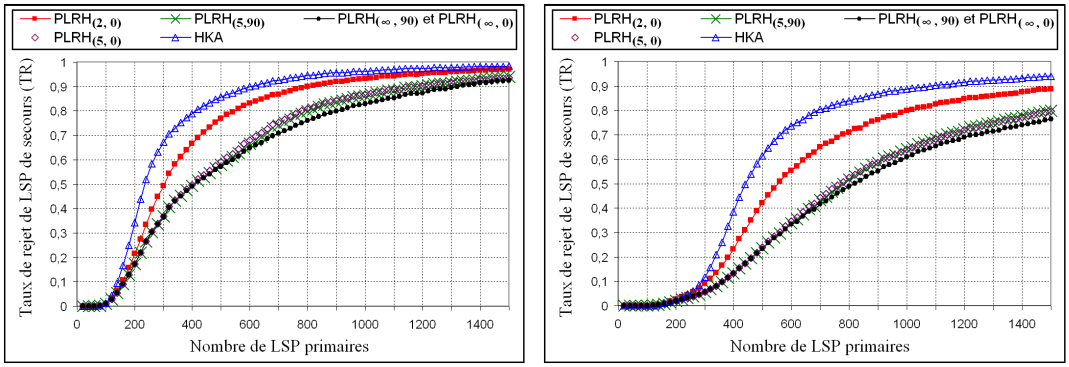
3.4.3.1 Métriques

Les mêmes métriques que celles choisies dans la section 3.3.3.1 sont ré-utilisées ici pour comparer l'heuristique PLRH à l'heuristique HKA.

Comme dans les sections 3.2.3.1 et 3.3.3.1, nous mesurons les valeurs des trois métriques *TR*, *NVD* et *TUBS* pour différentes charges du réseau (i.e. à chaque établissement de 20 LSP primaires). Tous les résultats illustrés dans la section 3.4.3.2 correspondent aux valeurs moyennes obtenues après 1000 exécutions et tirages aléatoires.

3.4.3.2 Résultats et analyse

La figure 3.13 montre l'évolution du taux de rejet des LSP de secours (*TR*) en fonction du nombre de LSP primaires déjà établis dans le réseau. Comme nous le constatons, les cinq variantes de l'heuristique PLRH ont toutes des taux de rejet meilleurs et plus petits que celui de l'heuristique HKA. Ceci s'explique par l'inclusion de l'information transmise avec KHA dans celle diffusée dans les vecteurs de PLRH. En effet, l'unique



(a) Topologie du réseau de 95 risques (b) Topologie du réseau de 162 risques

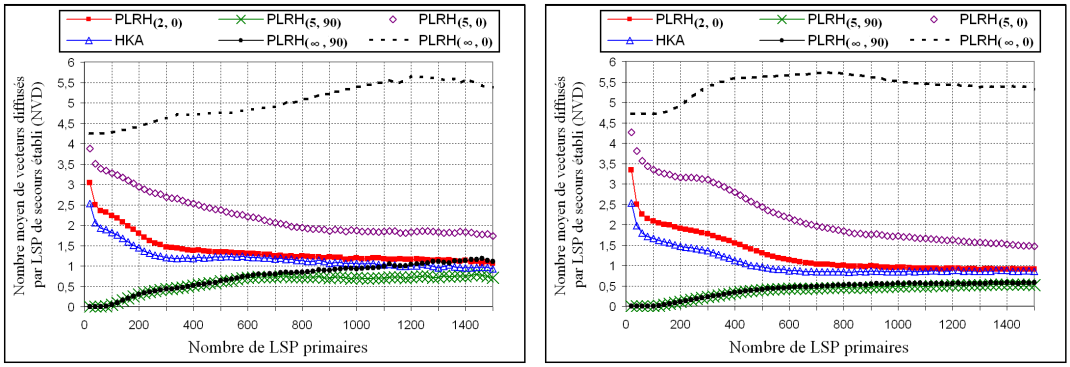
FIG. 3.13 – Evolution du taux de rejet des LSP de secours (TR)

prix de protection ($Max_{(r \in R)}\, \delta_r^{\lambda}$) qui est diffusé dans KHA pour chaque arc λ est toujours inclus dans le premier couple du $y^{\lambda}\_vecteur$ annoncé dans le réseau avec PLRH.

Pour les petites charges du réseau (nombre de LSP primaires inférieur à 140 sur la figure 3.13 (a) et inférieur à 240 sur la figure 3.13 (b)), nous observons que les variantes de PLRH ont un même taux de rejet des LSP de secours qui est lui-même très proche de zéro. Ceci est dû aux valeurs des quantités de bande passante de secours allouées sur les arcs qui sont très petites (c'est-à-dire les valeurs des plus grands prix de protection sur les arcs sont petites), ce qui entraîne la sélection de presque tous les arcs lors du calcul des LSP de secours.

Lorsque la charge du réseau augmente, le taux de rejet des LSP de secours de l'heuristique HKA croît plus vite et se distingue de ceux correspondant aux différentes variantes de PLRH. Ceci s'explique par la qualité d'approximation des prix de protection de PLRH qui est meilleure à celle adoptée par HKA. Alors que l'heuristique HKA approxime tous les prix de protection sur un arc λ par le prix de protection le plus élevé ($Max_{(r \in R)}\, \delta_r^{\lambda}$), l'heuristique PLRH les approxime par un prix de protection inférieur ou égal à celui de HKA, ce qui lui procure plus de flexibilité dans le choix des arcs formant un LSP de secours.

Concernant la comparaison entre les cinq variantes de PLRH, nous signalons que les taux de rejet des LSP de secours obtenus avec $PLRH_{(2,0)}$ sont inférieurs et plus mauvais que les taux correspondant aux variantes ($PLRH_{(\infty,90)}$ et $PLRH_{(\infty,90)}$) qui utilisent des vecteurs de taille illimitée. Cela nous mène à déduire que la diffusion de deux prix de protection (et leur risques associés) par arc n'est pas suffisante pour obtenir un taux de rejet de LSP de secours proche de celui obtenu avec la connaissance complète de l'information requise au placement des LSP de secours (i.e. $PLRH_{(\infty,90)}$ et $PLRH_{(\infty,90)}$). Lorsque la taille maximale des vecteurs diffusés dans le réseau augmente et atteint cinq, nous apercevons sur les figures 3.13 (a) et 3.13 (b) que les taux de rejet de LSP de secours décroissent et atteignent des valeurs proches des taux de rejet des variantes $PLRH_{(\infty,90)}$ et $PLRH_{(\infty,90)}$. Ainsi, nous déduisons que la diffusion de cinq plus grands prix de protection par arc (sur les topologies de réseau de la figure 3.3) est



(a) Topologie du réseau de 95 risques (b) Topologie du réseau de 162 risques

FIG. 3.14 – Evolution du nombre moyen de vecteurs transmis dans le réseau pour l'établissement d'un LSP de secours (NVD)

suffisante pour obtenir un taux de rejet de LSP de secours proche du taux obtenu avec une connaissance complète de l'information requise au placement des LSP de secours. Ceci s'explique par deux raisons :

1. Localité de la protection : la localité de l'algorithme de calcul des LSP de secours assure que le nombre de prix de protection qui atteignent des valeurs élevées et plus grandes que le seuil de confiance est très réduit. Ce nombre dépend significativement du voisinage immédiat du PLR.
2. Hétérogénéité des valeurs des prix de protection sur les arcs : Cette hétérogénéité dans les prix de protection assure la décroissance de la fonction $f\ (y) = \ (y^{\lambda})^{eme}$ *plus grand prix de protection sur l'arc* λ. Dans la pratique, le nombre de prix de protection supérieurs au seuil de confiance dans des réseaux maillés est souvent réduit lorsque le taux de rejet des LSP de secours est acceptable (exemple : pour un taux de rejet de LSP inférieur à 50% sur la figure 3.13 (a), la diffusion de 5 prix de protection les plus élevés par arc évite le blocage, par erreur, des LSP de secours).

A propos de la seconde métrique *NVD*, nous observons sur la figure 3.14 que le nombre moyen de vecteurs diffusés dans le réseau avec les trois variantes de PLRH n'employant pas le seuil de confiance (i.e. $PLRH_{(\infty,0)}$, $PLRH_{(5,0)}$ et $PLRH_{(2,0)}$) sont supérieurs à celui correspondant à HKA. Ceci est évident puisque la seule composante du vecteur diffusé avec HKA est toujours contenue dans les vecteurs diffusés par les variantes $PLRH_{(\infty,0)}$, $PLRH_{(5,0)}$ et $PLRH_{(2,0)}$. De même, pour les mêmes raisons, nous constatons sur la figure 3.14 que le nombre moyen de vecteurs diffusés dans le réseau avec $PLRH_{(2,0)}$ est inférieur à celui de $PLRH_{(5,0)}$ qui est lui-même inférieur à celui de $PLRH_{(\infty,0)}$. Ainsi, lorsque nous disposons d'aucune indication sur la quantité maximale de bande passante que peut réclamer un LSP de secours, l'utilisation d'un vecteur d'une taille maximale de 5 (au lieu de vecteurs de tailles illimitées) permet de réduire le nombre moyen de vecteurs diffusés dans le réseau de plus de 20%.

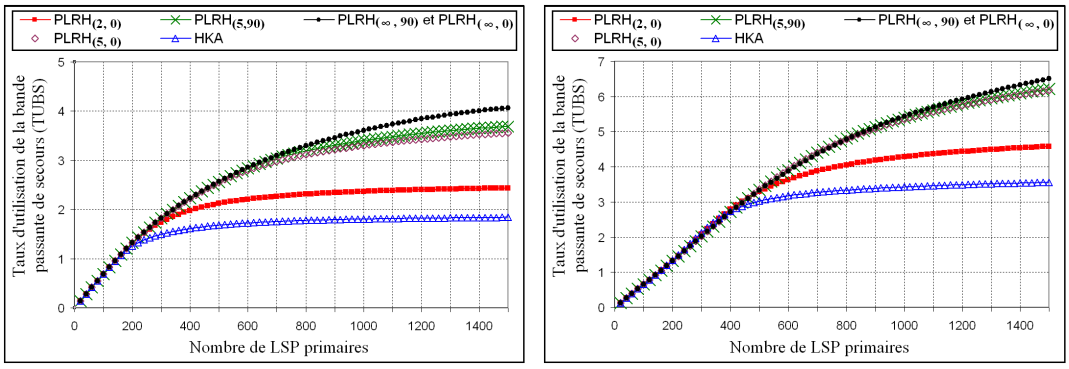
Lorsque le seuil de confiance est utilisé (variantes $PLRH_{(\infty,90)}$, $PLRH_{(5,90)}$) , la

fréquence d'envoi des vecteurs $\{y^{\lambda}\_vecteur\}_{\lambda \in E}$ diminue. Typiquement, pour un seuil de confiance égal à 90% de la capacité de secours, les figures 3.14 (a) et 3.14 (b) montrent que les nombres moyens de vecteurs diffusés avec les variantes $PLRH_{(\infty,90)}$, $PLRH_{(5,90)}$ sont souvent inférieurs aux nombres moyens de vecteurs diffusés avec l'heuristique HKA et très inférieurs à ceux des variantes n'employant pas le seuil de confiance. En effet, l'utilisation d'un seuil de confiance élevé élimine la diffusion d'un grand nombre de vecteurs surtout lorsque la charge du réseau est relativement faible.

L'autre constat important qui ressort des figures 3.14 (a) et 3.14 (b) est la similitude entre les nombres moyens de vecteurs diffusés avec les variantes $PLRH_{(\infty,90)}$ et $PLRH_{(5,90)}$. Ceci explique la ressemblance des taux de rejet de ces deux variantes sur les figures 3.13 (a) et 3.13 (b) et permet de déduire que la diffusion de vecteurs de cinq composantes est suffisante pour placer efficacement les LSP de secours (i.e. l'augmentation de la taille maximale des vecteurs diffusés dans le réseau n'apporterait aucune amélioration au mécanisme de placement des LSP de secours).

Concernant la dernière métrique *TUBS*, la figure 3.15 montre que l'heuristique HKA et toutes les variantes de PLRH ont des taux d'utilisation de la bande passante très proches les un des autres, lorsque la charge du réseau est faible (i.e. le nombre de LSP primaires protégés est inférieur à 200 dans le réseau de la figure 3.15 (a) et inférieur à 400 dans le réseau de la figure 3.15 (b)). Ceci s'explique essentiellement par les taux de rejet de LSP de secours qui sont faibles pour ces charges, ce qui permet de satisfaire, à peu près, les mêmes requêtes de protection.

Lorsque la charge du réseau augment et atteint une valeur relativement élevée, nous apercevons sur les figure 3.15 (a) et 3.15 (b) que les taux d'utilisation de la bande passante de secours des variantes $PLRH_{(\infty,90)}$, $PLRH_{(5,90)}$ et $PLRH_{(5,0)}$ sont supérieurs ou égaux (meilleurs) à ceux de la variante $PLRH_{(2,0)}$ qui sont eux-mêmes supérieurs ou égaux (meilleurs) à ceux de HKA. Ceci s'explique par la qualité d'approximation des prix de protection qui est meilleure avec la diffusion d'un plus grand nombre de prix de protection supérieurs aux seuils de confiance.



(a) Topologie du réseau de 95 risques (b) Topologie du réseau de 162 risques

FIG. 3.15 – Evolution du taux d'utilisation de la bande passante de secours (TUBS)

3.4.4 Conclusion

L'heuristique PLRH permet un placement efficace des LSP de secours avec la diffusion d'un vecteur de taille limitée pour chaque arc du réseau. Ce vecteur est constitué, pour un arc λ, de y^{λ} couples contenant les y^{λ} valeurs de prix de protection les plus élevées ainsi que les risques de panne associés. Grâce à ces vecteurs diffusés, tout nœud du réseau est capable d'estimer les valeurs des prix de protection sur tous les arcs du réseau, pour placer ensuite efficacement les LSP de secours.

Afin de réduire l'information diffusée dans le réseau, PLRH définit un seuil de confiance au delà duquel un prix de protection peut être diffusé dans le réseau. Pour des raisons d'efficacité, le seuil de confiance doit être fixé de telle sorte que tout LSP de secours dont le coût de protection est inférieur au seuil de confiance, sur un arc λ, pourra inclure cet arc lors de son calcul.

A cause de la localité de la méthode de protection et de l'hétérogénéité de la distribution des prix de protection sur un arc, le nombre de risques dont les prix de protection induisant le rejet d'un arc lors du calcul d'un nouveau LSP de secours, est généralement faible. En effet, la diffusion de huit prix de protection par arc est souvent suffisante pour éviter le rejet par erreur de la quasi-totalité des requêtes de protection. Ceci permet de garantir l'efficacité du mécanisme de diffusion adopté par PLRH, puisque même sans utilisation du seuil de confiance, le nombre de messages diffusant les vecteurs de prix de protection ($\{y^{\lambda}\_\, vecteur\}_{\lambda \in E}$) est faible mais souvent suffisant pour éviter le rejet par erreur d'une grande portion des requêtes de protection.

L'heuristique PLRH a plusieurs avantages. Premièrement, elle est symétrique et est complètement distribuée (tous les nœuds du réseau disposent de la même information permettant le placement des LSP de secours). Deuxièmement, elle permet de réduire le trafic de contrôle en éliminant la communication entre les BPCE et les PLR (lorsque les calculs de LSP de secours sont effectués sur les PLR eux-mêmes). Troisièmement, PLRH facilite le calcul coordonné des LSP, ce qui facilite la protection multi-domaines et améliore le taux de protection. Enfin, l'heuristique PLRH est simple à déployer puisqu'elle ne requiert que de très légères extensions aux protocoles IGP-TE (pour la diffusion des vecteurs de prix de protection).

Nos simulations montrent qu'avec la diffusion de vecteurs $\{y^{\lambda}\_\, vecteur\}_{(\lambda \in E)}$ de faible taille (par rapport aux nombre de risques), le taux de partage de l'heuristique PLRH atteint des valeurs élevées et proches de l'idéal (i.e. les fuites de bandes passante sont quasi-nulles). De plus, avec l'utilisation d'un seuil de confiance assez élevé, la fréquence de diffusion des messages est réduite sensiblement.

3.5 Conclusion générale

Dans ce chapitre, nous avons proposé trois techniques de placement distribué des LSP de secours qui tiennent compte du partage de la bande passante : l'algorithme TDRA, l'heuristique DBSH et l'heuristique PLRH. Contrairement à la plupart des méthodes de placement des LSP de secours, nos techniques traitent tous les types de risques de panne, particulièrement les SRLG. elles sont aussi facile à déployer puisque

de légères extensions aux protocoles de signalisation et/ou IGP-TE permettant leur déploiement.

Dans la première technique de placement de LSP de secours proposée (l'algorithme TDRA), sur chaque nœud α du réseau tourne une entité de calcul de LSP de secours ($BPCE_\alpha$). Cette entité $BPCE_\alpha$ est chargée de la détermination des LSP de secours protégeant contre la panne du nœud α et de tous les liens qui lui sont adjacents. Pour permettre une protection efficace contre les risques de type SRLG, les nœuds d'extrémité aux liens, appartenant aux SRLG contenant le lien protégé, partagent les prix de protection de ces SRLG. Pour ce faire, nous avons proposé d'étendre les protocoles de signalisation afin qu'ils transportent les structures et les propriétés des LSP de secours.

Bien que cet algorithme évite les diffusions (transmission ciblée des prix de protection vers les entités de calcul des LSP de secours), ses performances peuvent se dégrader si le nombre de LSP établis et le nombre de SRLG sont très élevés. Il est plutôt destiné aux larges réseaux contenant un nombre petit ou moyen de SRLG.

Dans la seconde technique de calcul de LSP de secours proposée (l'heuristique DBSH), sur chaque nœud α du réseau tourne une entité de calcul de LSP de secours ($BPCE_\alpha$). Comme dans l'algorithme TDRA, cette entité $BPCE_\alpha$ est chargée de la détermination des LSP de secours protégeant contre la panne du nœud α et de tous ses liens adjacents. Cependant et contrairement à l'algorithme TDRA, cette heuristique utilise la diffusion de vecteurs de prix de SRLG de taille limitée pour approximer les prix de protection des SRLG.

Cette heuristique DBSH diminue sensiblement la quantité d'informations diffusées dans le réseau (comparativement à l'algorithme de Kini). Pour son déploiement, elle nécessite de très légères extensions aux protocoles de signalisation (pour la communication entre les entités de calcul des LSP de secours et les nœuds PLR) et aux protocoles IGP-TE (pour la transmission des prix de protection des SRLG).

Enfin, dans la troisième technique de placement des LSP de secours proposée (l'heuristique PLRH), plus de flexibilité est offerte pour le choix et le placement des entités calculant les LSP de secours. Ainsi, avec PLRH, tout nœud du réseau est capable de calculer tous ou n'importe quel LSP de secours. Pour ce faire, l'heuristique propose de diffuser, pour chaque arc λ, un vecteur de $(y^\lambda)_{0<y^\lambda<\infty}$ composantes au plus. Toute composante doit être formée d'un prix de protection supérieur à un seuil de confiance (utilisé pour diminuer la fréquence de diffusion) et du risque associé. De plus, les composantes transmises dans un vecteur diffusé ne doivent contenir que les prix de protection les plus élevés sur l'arc (pour améliorer la qualité d'approximation des prix de protection).

En plus de l'avantage de la réduction de la quantité d'informations diffusées dans le réseau, l'heuristique PLRH élimine la communication entre les nœuds supportant les entités de calcul et les PLR (à condition que le calcul de LSP de secours soit effectué par leurs nœuds de tête). De plus, cette heuristique permet de coordonner le calcul des LSP de secours avec leur LSP primaire (pour améliorer la protection et faciliter le calcul des LSP de secours multi-domaines) et elle ne nécessite que de très légères extensions aux protocoles IGP-TE pour son implantation.

Après avoir proposé différents mécanismes distribués de placement des LSP de se-

cours capables de faire face à tous les types de risques de panne (notamment les SRLG) et tenant compte du partage de la bande passante, nous montrerons dans le chapitre suivant qu'il est possible d'offrir plus de flexibilité dans le choix des LSP de secours d'un côté, et améliorer les allocations de la bande passante de secours d'un autre côté.

Chapitre 4

Exploiter les structures des SRLG pour améliorer le placement des LSP de secours

4.1 Introduction

Dans ce chapitre, nous proposons d'*exploiter les structures des SRLG pour améliorer le placement* (*centralisé et distribué*) *des LSP de secours* (ESSAPL) [SCLR09b, SCLRa].

Contrairement à la protection globale où l'activation d'un seul LSP de secours suffit pour récupérer d'une panne affectant un LSP primaire, la protection locale peut nécessiter l'activation simultanée de plusieurs LSP de secours pour restaurer une seule communication. En effet, lors d'une panne d'un SRLG affectant plusieurs liens d'un même LSP primaire, tous les LSP de secours protégeant contre la panne d'un lien appartenant au SRLG affecté sont activés par leur LSR de tête (PLR) pour faire face à la panne. En constatant que certains LSP de secours activés, après une panne d'un SRLG, ne participent pas réellement au processus de récupération (ils ne reçoivent aucun flux de données), nous augmentons la disponibilité de la bande passante en restreignant la concurrence pour les allocations de la bande passante (de secours) aux LSP de secours recevant effectivement du trafic après la panne du SRLG (c'est-à-dire, seuls les LSP de secours recevant du trafic après une panne d'un SRLG réservent la bande passante). Par ailleurs, nous obtenons plus de flexibilité pour le choix des chemins de secours protégeant contre les liens appartenant à des SRLG en réduisant le nombre de LSP de secours dédiés à la protection d'un SRLG. Ainsi, il ne sera plus nécessaire de contourner systématiquement tous les liens partageant un même SRLG avec le lien protégé lors du placement d'un nouveau LSP de secours.

La suite de ce chapitre est organisée comme suit. Dans la section 4.2, nous introduisons et illustrons la différence entre les LSP de secours *actifs* (i.e. LSP prêts à router les paquets reçus) et les LSP de secours *opérationnels* (i.e. LSP actifs et routant effectivement du trafic). Nous verrons ainsi qu'en exploitant les structures des SRLG, il sera possible de déterminer si un LSP de secours actif est opérationnel ou pas après une

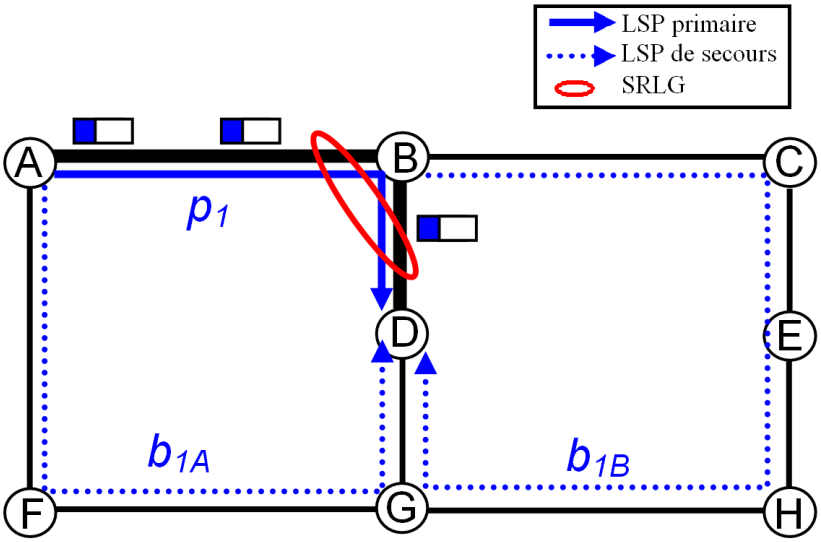


FIG. 4.1 – Protection locale contre les pannes de tous les risques d'un LSP primaire

panne d'un SRLG. Afin d'offrir plus de flexibilité pour le choix des LSP de secours et pour diminuer la probabilité de blocage des requêtes de protection, nous suggérons en section 4.3 d'étendre le partage de bande passante en réduisant l'ensemble des risques de panne protégés par un LSP de secours. Dans notre proposition, les quantités de bande passante allouées aux LSP de secours sont diminuées en restreignant la concurrence pour les allocations de bande passante aux LSP de secours opérationnels (sous-ensemble des LSP actifs). De plus, la topologie du réseau est mieux exploitée puisqu'il n'est plus nécessaire de contourner systématiquement tous les liens appartenant aux mêmes SRLG que ceux contenant le lien protégé. En section 4.5, nous présentons et analysons quelques résultats de simulation. La dernière section sera consacrée à la conclusion.

4.2 LSP de secours actif vs. LSP de secours opérationnel

A cause de la difficulté de distinguer rapidement les types de panne (nœud, lien ou SRLG), tout routeur détectant une panne sur une interface de sortie *active* localement tous ses LSP de secours qui protègent des LSP primaires traversant l'interface défaillante. Ainsi, toutes les étiquettes de sortie associées aux LSP primaires affectés par la panne seront remplacées par les étiquettes correspondant à leur LSP de secours. De cette manière, tout paquet reçu du LSP primaire sera aiguillé vers le chemin de secours permettant la récupération de la panne.

Comme une panne physique peut affecter plusieurs liens logiques à la fois (cf. panne d'un SRLG) et puisque un LSP de secours n'est conçu et établi que pour pallier la panne d'un seul lien (et éventuellement contre la panne d'un seul nœud), plusieurs LSP de secours protégeant un même LSP primaire pourraient être *actifs* en même

temps. Dans certains cas de figure, le LSR de tête d'un LSP de secours actif b_1 est contourné par un autre LSP de secours b_2 activé pour récupérer d'une panne affectant le même LSP primaire que celui protégé par b_1 (c'est-à-dire, le LSR de tête du LSP b_1 est situé entre les nœuds d'extrémité du LSP de secours b_2). Dans un tel cas, le LSP de secours *actif* b_1 ne reçoit et n'achemine aucun trafic après la récupération de la panne ; il est considéré comme *inopérationnel* puisqu'il n'utilise pas réellement ses ressources (particulièrement la bande passante). En conséquence, la bande passante réservée pour un tel LSP de secours pourrait être libérée afin d'être allouée à d'autres LSP de secours. Contrairement au LSP de secours b_1, l'autre LSP de secours b_2 participe réellement au processus de récupération puisqu'il reçoit et achemine le trafic de son LSP primaire affecté. Ce chemins de secours *actif* (b_2) sera considéré comme *opérationnel*. Ses ressources (particulièrement la bande passante) ne pourront pas être ré-allouées à d'autres LSP.

Sur la figure 4.1, deux LSP de secours b_{1A} ($A \rightarrow F \rightarrow G \rightarrow D$) et b_{1B} ($B \rightarrow C \rightarrow E \rightarrow H \rightarrow G \rightarrow D$) sont établis pour protéger le LSP primaire p_1 ($A \rightarrow B \rightarrow D$) contre les risques de panne suivants : le nœud B, le lien A-B, le lien B-D et le SRLG $srlg$ = (A-B, B-D).

Lorsque le LSR A (resp. le routeur B) détecte une panne de son interface menant vers le nœud adjacent B (resp. le nœuds adjacent D), il active localement le LSP de secours b_{1A} (resp. b_{1B}) qui protège l'unique LSP primaire p_1 traversant l'interface défaillante. Concrètement, lors d'une panne affectant uniquement le risque de type lien A-B ou le nœud B (resp. une panne affectant uniquement le risque de type lien B-D), le trafic du LSP primaire p_1 sera aiguillé vers l'unique LSP de secours b_{1A} (resp. b_{1B}), activé pour faire face à la panne. Ainsi, pour la panne de tout risque de type lien ou nœud, tous les LSP de secours activés reçoivent et acheminement réellement le trafic de leurs LSP primaire affectés ; ils sont donc *opérationnels*.

Avec une panne d'un risque de type SRLG par contre, certains LSP de secours actifs ne reçoivent aucun trafic après la récupération. Par exemple, lorsque le SRLG *srlg* de la figure 4.1 tombe en panne, les deux nœuds d'extrémité (A et B) des liens appartenant au SRLG *srlg* détectent une panne sur leur interface de sortie. En conséquence, tous les LSP de secours protégeant contre la panne d'un lien appartenant au SRLG *srlg* et dont le LSR de tête est A ou B seront activés. Typiquement, sur la figure 4.2, les deux LSP de secours b_{1A} et b_{1B} seront activés. Comme l'aiguillage du trafic d'un LSP primaire affecté vers un LSP de secours b (*actif* et *opérationnel*) induit le contournement des nœuds du LSP primaire situés entre les extrémités du LSP de secours b, il en résulte que tout LSP de secours (actif ou pas actif) dont le LSR de tête est situé entre les nœuds d'extrémité du LSP b ne peut pas recevoir du trafic. Sur la figure 4.2, le ré-acheminement du trafic du LSP primaire p_1 sur le LSP de secours b_{1A} (suite à la panne du SRLG *srlg*), induit le contournement du nœud B. En conséquence, le LSP de secours b_{1B} ne reçoit aucun trafic (*inopérationnel*) après la récupération de la panne du SRLG *srlg*, bien que ce LSP soit actif. En effet, après la récupération de la panne du SRLG *srlg*, le trafic de la communication supportée par le LSP primaire p_1 sera acheminé sur le chemin $A \rightarrow F \rightarrow G \rightarrow D$ qui contourne le LSR de tête B du LSP de secours b_{1B}.

Nous notons que la distinction de la notion d'activité (actif ou inactif) d'un LSP de secours de la notion d'opération (opérationnel ou inopérationnel) d'un LSP de secours,

proposée dans [SCLR09b, SCLR09b], est nouvelle. En effet, dans la littérature, la plupart des travaux se consacrent à la protection globale [YJ05, XCX+06] ou ne considèrent que des risques de panne de type nœud ou lien [MBL03, BML06, KL01, KL03, KKL+01]. Dans de tels cas, l'ensemble des LSP actifs après une panne est exactement le même que l'ensemble des LSP opérationnels après la même panne. La distinction de la notion d'activité d'un LSP de secours de la notion d'opération d'un LSP de secours n'est donc pas utile.

Cependant, avec une protection locale tenant compte des risques de type SRLG, la dissociation de la notion d'activité d'un LSP de secours de la notion d'opération d'un LSP de secours est intéressante et est recommandée. Ainsi, pour améliorer le placement des LSP de secours, il serait intéressant de prendre en compte l'état d'opération des LSP de secours (au lieu de ne tenir compte que de l'état d'activité des LSP de secours) avant leur calcul. La sous-section suivante sera consacrée à la description du processus de détermination de l'état d'opération d'un LSP de secours après une panne.

4.2.1 Détermination de l'état d'opération d'un LSP de secours

Un LSP de secours est opérationnel après la panne d'un risque donné si et seulement si ce LSP est actif et il participe réellement à l'acheminement du trafic de la communication protégée, après la récupération de la même panne. De même, tout LSP de secours inactif ou ne recevant aucun trafic après une panne est considéré comme inopérationnel.

En fonction du risque de panne défaillant, la détermination de l'ensemble des LSP de secours opérationnels peut être plus ou moins simple. Ainsi, avec un risque de panne

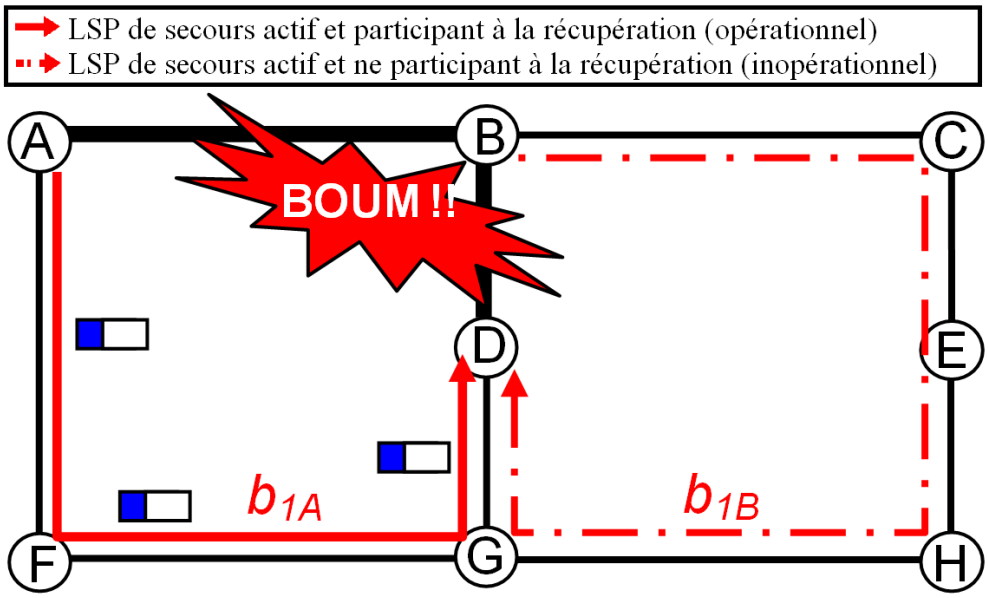
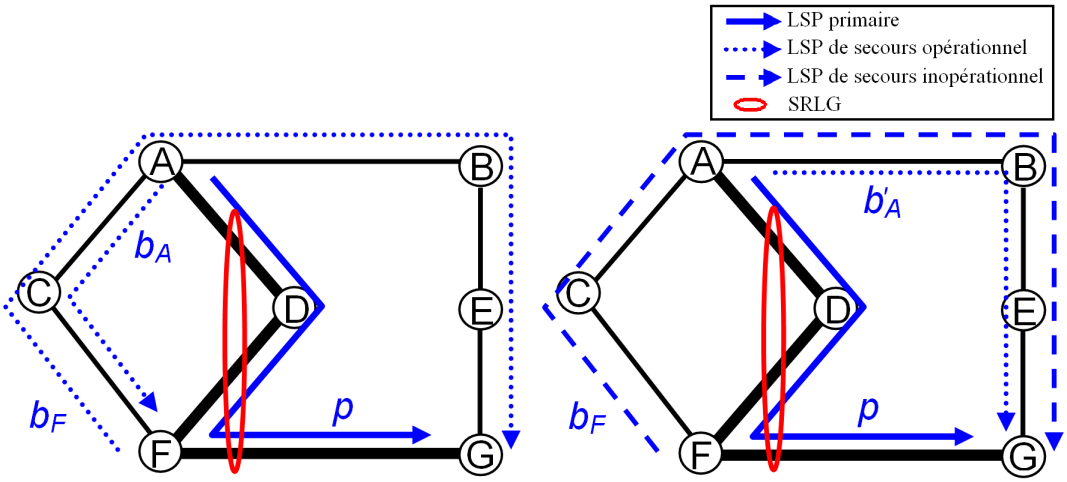


FIG. 4.2 – LSP de secours actif et LSP de secours opérationnel



(a) Aucun LSP de secours inopérationnel (b) Un LSP de secours inopérationnel

FIG. 4.3 – Opération des LSP de secours

simple (lien ou nœud), l'ensemble des LSP de secours opérationnels correspond exactement à l'ensemble des LSP de secours actifs. Avec une panne d'un risque composé (i.e. SRLG) par contre, un LSP de secours b protégeant un LSP primaire p est opérationnel après la récupération d'une panne d'un SRLG *srlg* si et seulement s'il vérifie les conditions suivantes :

1. le LSP de secours b protège contre la panne d'un lien appartenant au SRLG (*srlg*) en panne,
2. il n'existe aucun autre LSP de secours b' ($b' \neq b$) tel que :
   - b' protège le LSP primaire p contre la panne d'un lien appartenant au SRLG *srlg*,
   - le segment primaire situé entre les extrémités du LSP de secours b' contient, comme routeur de transit, le nœud source du LSP b.

4.2.1.1 Exemple

Pour faciliter la compréhension du processus de détermination de l'état d'opération d'un LSP de secours après la panne d'un SRLG, nous considérons un exemple. Sur la figure 4.3, un LSP primaire p ($A \rightarrow D \rightarrow F \rightarrow G$) traversant des liens de l'unique SRLG *srlg* = (A-D, D-F, F-G) du réseau est établi. Pour protéger ce LSP primaire contre la panne du lien F-G, nous avons configuré un même LSP de secours $b_F = F \rightarrow C \rightarrow A \rightarrow B \rightarrow E \rightarrow G$ dans les sous-figures 4.3 (a) et 4.3 (b). Pour protéger p contre les pannes du nœud D et du lien A-D, nous avons utilisé dans la figure 4.3 (a) (resp. figure 4.3 (b)) un LSP de secours $b_A = A \rightarrow C \rightarrow F$ (resp. un LSP de secours $b'_A = A \rightarrow B \rightarrow E \rightarrow G$).

Lors d'une panne du SRLG *srlg* (*srlg* = (A-,D, D-,F, F-,G), les deux nœuds A et F activent les deux LSP de secours b_A et b_F dans la figure 4.3 (a) (resp. les deux LSP de secours b'_A et b_F dans la figure 4.3 (b)) pour récupérer rapidement de la panne.

Sur la figure 4.3 (a), les deux LSP de secours b_A et b_F seront opérationnels après la récupération. En effet, le segment primaire situé entre les nœuds d'extrémité du LSP de secours b_A (resp. b_F) ne contient aucun nœud source d'un autre LSP de secours protégeant le même LSP primaire p. Sur la figure 4.3 (b) cependant, seul le LSP de secours b'_A sera opérationnel après la panne du SRLG *srlg* (pour les mêmes raisons que b_A dans la figure 4.3 (a)). Le second LSP de secours b_F est donc inopérationnel après la panne du SRLG *srlg* car :

1. b'_A protège le LSP primaire p (c'est-à-dire le même LSP primaire protégé par le LSP de secours b_F) contre la panne d'un lien (A-D) appartenant au SRLG *srlg*,
2. le segment primaire $A \rightarrow D \rightarrow F \rightarrow G$ situé entre les nœuds d'extrémité (A et G) du LSP de secours b'_A contient, comme routeur de transit, le LSR de tête (F) du LSP de secours b_F.

Comme seuls les LSP de secours opérationnels utilisent réellement leurs ressources (particulièrement la bande passante), nous montrons dans la section 4.3 que le placement des LSP de secours pourrait être amélioré en restreignant la concurrence pour l'allocation de la bande passante de secours aux LSP opérationnels et en réduisant l'ensemble des risques protégés par un LSP de secours à ceux provoquant son opération.

4.3 Améliorer le placement des LSP de secours

L'exploitation des structures des SRLG permet de restreindre l'ensemble des risques de panne protégés par un LSP de secours à ceux provoquant son opération après une panne. Cela permet de :

1. réduire la quantité de bande passante allouée aux LSP de secours par la restriction de la concurrence pour l'allocation de la bande passante de secours aux LSP opérationnels,
2. mieux exploiter la topologie du réseau puisqu'il n'est plus nécessaire de contourner tous les liens appartenant à des SRLG incluant le lien protégé.

4.3.1 Restreindre la concurrence pour l'allocation de la bande passante de secours aux LSP opérationnels

En substituant l'ensemble des LSP de secours actifs par le sous-ensemble des LSP opérationnels (pour toute panne), nous réduisons la quantité de bande passante allouée aux LSP de secours tout en garantissant le respect des contraintes de bande passante.

Pour déduire l'état d'opération des LSP de secours, nous définissons une nouvelle fonction Op () comme suit :

$$Op : BPaths \times Rs \rightarrow \{0,1\}$$

$$(b,r) \mapsto y = \begin{cases} 1 \text{ si } b \text{ est opérationnel après la récupération de la panne} \\ \qquad \text{du risque } r \\ 0 \text{ sinon} \end{cases}$$

où : $BPaths$ est l'ensemble de tous les LSP de secours et Rs est l'ensemble des risques de panne.

Afin de tenir compte de l'état des LSP de secours lors de leur placement, nous associons à chaque LSP de secours b un *ensemble de risques de panne réellement protégés* (*RPFRG* ou *Really Protection Failure Risk Group*). Cet ensemble, noté *RPFRG* (b), est constitué des risques réellement protégés par le LSP de secours b (i.e. *RPFRG* (b) est composé des risques dont la panne rend opérationnel le LSP de secours b). Il est déterminé comme suit :

$$RPFRG\,(b) = \{r \backslash r \in Rs \wedge Op\,(b,\, r) = 1\} \tag{4.1}$$

Comme seuls les LSP de secours opérationnels après une panne peuvent être en concurrence pour l'allocation de bande passante, le prix *effectif* de protection $(\delta')_r^\lambda$ de tout risque r sur un arc λ sera diminué (car l'ensemble des LSP de secours opérationnels est inclus dans l'ensemble des LSP de secours actifs, après n'importe quelle panne, c'est-à-dire *RPFRG* (b) $\in$ *PFRG* (b)) et est calculé comme suit :

$$(\delta')_r^\lambda = \sum_{b \in BPaths \backslash\, \lambda \in b} Op\,(b,\, r) \times bw\,(b) \tag{4.2}$$

En substituant le couple (*PFRG (b)*, δ_r^λ) par le couple (*RPFRG (b)*, $(\delta')_r^\lambda$) dans les formules (2.2), (2.3), (2.4) et (2.7), nous obtenons le coût effectif de protection $(\theta')_c^\lambda$ d'un lien u-v sur un arc λ, le coût effectif de protection $(\theta')^\lambda\,(b)$ d'un LSP de secours b sur un arc λ, la quantité effective de bande passante de secours $(G')^\lambda$ allouée sur un arc λ et le surcoût effectif de protection $(\gamma')_r^\lambda$ d'un risque r sur un arc λ comme suit :

$$(\theta')_c^\lambda = Max_{r \backslash\, c \in \{r\} \wedge r \in Rs}\,(\delta')_r^\lambda \tag{4.3}$$

$$(\theta')^\lambda\,(b) = Max_{r \in RPFRG\,(b)}\,(\delta')_r^\lambda \tag{4.4}$$

$$(G')^\lambda = Max_r\,((\delta')_{r \backslash\, r \in Rs}^\lambda) = Max_c\,((\theta')_{c \backslash\, c \in (V \cup E)}^\lambda) \tag{4.5}$$

$$(\gamma')^\lambda\,(b) = \begin{cases} \infty \;si\; (\theta')^\lambda\,(b) + bw\,(b) + F_\lambda > C^\lambda \\ Max\,((\theta')^\lambda\,(b) + bw\,(b) - (G')^\lambda,\, 0) \;sinon \end{cases} \tag{4.6}$$

En conséquence, pour assurer le respect des contraintes de la bande passante, l'arc λ ne pourra être traversé par un nouveau LSP de secours b que s'il vérifie l'inégalité suivante :

$$(\theta')^\lambda\,(b) + F_\lambda + bw\,(b) \leq C^\lambda \tag{4.7}$$

Bien évidemment, l'inclusion de l'ensemble des LSP opérationnels dans l'ensemble des LSP actifs après une panne implique une diminution des allocations de la bande passante de secours sur tous les arcs. Ainsi, nous avons :
$\forall\,(\lambda,\, r,\, b) \in (E,\, Rs,\, BPaths) : ((\delta')_r^\lambda \leq \delta_r^\lambda) \wedge ((G')^\lambda \leq G^\lambda) \wedge ((\theta')^\lambda\,(b) \leq \theta^\lambda\,(b))$

4.3.1.1 Exemple

Sur la figure 4.3 (b), nous calculons la bande passante de secours allouée sur l'arc $A \rightarrow B$ comme suit (conformément à la formule (2.4)) :
$G^{AB} = Max(\delta_{A-D}^{A \rightarrow B}, \delta_{D}^{A \rightarrow B}, \delta_{F-G}^{A \rightarrow B}, \delta_{srlg}^{A \rightarrow B}) = \delta_{srlg}^{AB} = 2 \times bw\,(p)$

En exploitant les structures des SRLG (permettant de déterminer l'état d'opération des LSP de secours), nous diminuons et calculons la quantité de bande passante allouée aux LSP de secours comme suit :
$(G')_{AB} = Max((\delta')_{A-D}^{A \rightarrow B}, (\delta')_{D}^{A \rightarrow B}, (\delta')_{F-G}^{A \rightarrow B}, (\delta')_{srlg}^{A \rightarrow B}) = bw\,(p) = G^{A \rightarrow B}/2$

En supposant que $C^{A \rightarrow B} = 2 \times bw\,(p)$, il en résulte que dans le premier cas (où il n'y a pas d'exploitation des structures des SRLG), aucun nouveau LSP primaire et aucun LSP de secours protégeant contre les risques appartenant à l'ensemble {*A*-*D*, *D*, *F*-*G*, $srlg_1$}) ne peut traverser l'arc $A \rightarrow B$ alors que dans le second cas (où les structures des SRLG sont exploitées pour placer les LSP de secours), l'arc $A \rightarrow B$ pourrait être traversé par n'importe quel nouveau LSP primaire ou LSP de secours réclamant une quantité de bande passante inférieure ou égale à $C^{A \rightarrow B}/2$.

4.3.2 Mieux exploiter la topologie du réseau

En plus de la diminution de la quantité de bande passante allouée aux LSP de secours, l'exploitation des structures des SRLG permet une meilleure prise en compte de la topologie du réseau lors du placement des LSP de secours. Ainsi, un LSP de secours *b* ne doit plus contourner systématiquement tous les liens appartenant aux SRLG contenant le lien protégé mais uniquement les liens appartenant aux SRLG de son ensemble de risques de panne réellement protégés (i.e. *RPFRG* (*b*).

Comme tout calcul d'un nouveau LSP de secours requiert la connaissance de tous les liens (et nœuds) qui doivent être contournés, nous concluons qu'il est nécessaire de déterminer l'ensemble des risques de panne réellement protégés par tout LSP de secours avant son calcul. Pour ce faire, il doit être possible de déduire l'état d'opération d'un LSP de secours avant son calcul.

En analysant les conditions nécessaires et suffisantes pour la détermination de l'état d'opération d'un LSP de secours (cf. section 4.2.1), nous déduisons que les arcs traversés par un LSP de secours *b* n'ont aucune incidence sur son état d'opération après une panne. En effet, seuls (1) le lien et le nœud protégés, (2) le LSR de tête du LSP *b* et (3) les LSP de secours protégeant le même LSP primaire que *b*, contre la panne d'un lien partageant un même SRLG et situé en amont du lien protégé par *b*, sont utiles pour la détermination de l'état d'opération du LSP de secours *b*.

4.3.2.1 Exemple

Sur la figure 4.3 (b), n'importe quel LSP de secours b'_D protégeant le LSP primaire *p* contre la panne du lien $D \rightarrow F$ est inopérationnel après la panne du SRLG *srlg*. En effet, après une telle panne, le trafic de la communication affectée et supportée par le LSP *p* est aiguillé par le routeur *A* vers le LSP de secours b'_A qui ne joint le LSP primaire affecté *p* qu'au nœud *G*, situé en aval du LSR de tête *D* du LSP b'_D. Ainsi, indépendamment

des arcs qui le constituent, le LSP de secours b'_D sera toujours inopérationnel après la panne du SRLG *srlg* (car il ne recevra aucun flux après la récupération de la panne du SRLG *srlg*).

4.3.3 Algorithme de placement des LSP de secours exploitant les structures des SRLG

Pour améliorer la qualité de la protection, nous proposons un nouvel algorithme (cf. algorithme 7) exploitant les structures des SRLG lors du placement d'un nouveau LSP de secours b.

Dans la première étape de notre algorithme, l'ensemble des risques de panne réellement protégés (i.e. $RPFRG$ (b)) par le LSP de secours b est déterminé (cf. formule (4.1)).

Dans la deuxième étape de l'algorithme 7, nous supprimons de la topologie du réseau tous les arcs et nœuds supportés (contenus) par les risques de l'ensemble $RPFRG$ (b). Cela garantit la disjonction des LSP de secours des risques induisant leur opération.

Ensuite, pour assurer le respect des contraintes de la bande passante, nous éliminons dans la troisième étape de l'algorithme 7, tous les liens ne vérifiant pas la formule (4.7). Cela garantira une bande passante suffisante pour tous les LSP de secours opérationnels après n'importe quelle panne. Nous notons que la formule (4.7) n'applique le partage de la bande passante qu'entre les LSP de secours. Pour étendre le partage de la bande passante aux LSP primaires (cf. section 2.10), une légère modification de l'algorithme, permettant de tenir compte de la bande passante primaire libérée, est nécessaire et est décrite dans [SCLR09b].

Enfin, dans la dernière étape de notre algorithme, nous appliquons n'importe quel algorithme de calcul de chemins (comme SPF) et n'importe quelle méthode de protection locale (protection par LSP de détour ou protection par tunnel de secours) sur la topologie du réseau réduite pour calculer le nouveau LSP de secours b.

4.3.3.1 Exemple

Pour faciliter la compréhension de l'algorithme 7, nous considérons l'exemple de la figure 4.3 (b) où tous les arcs sont supposés avoir une capacité d'une seule unité (les quantités de bande passante des LSP p, b'_A et b_F sont aussi égales à une unité).

Pour calculer un LSP de secours b'_D protégeant le LSP primaire p contre les pannes du lien D-F et du nœud F, il suffira de déterminer un chemin reliant le nœud D au nœud G, évitant les risques réellement protégés et ne traversant aucun lien ne disposant pas d'une bande passante suffisante.

Avec une approche classique de placement des LSP de secours [KKL+01, LRC02, VCLF+04, SCLR07, SCLR08, SCLRb], aucun LSP de secours (b'_D) ne pourra être déterminé puisque ces approches imposent au LSP b'_D de contourner tous les liens (A-D, D-F, F-G) du SRLG *srlg*. En effet, après l'élimination des liens (A-D, D-F, F-G) du SRLG *srlg* (car le SRLG *srlg* est dans l'ensemble $PFRG$ (b'_D)), aucun chemin ne pourra relier le nœud D au nœud G.

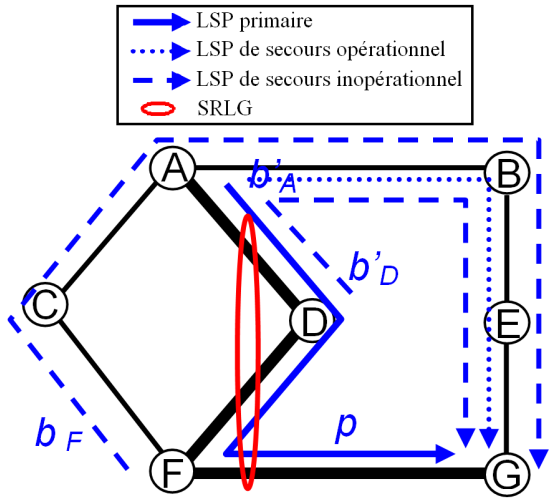


FIG. 4.4 – Un LSP de secours traversant un lien d'un SRLG contenant le lien protégé

Algorithme 7 Calcul d'un LSP de secours b

entrées

Un graphe $G = (V, E, Rs)$ correspondant à la topologie du réseau et à ses risques de panne

sortie

Un LSP de secours b fournissant la protection et assurant le respect des contraintes de la bande passante

début_algorithme

1. {Détermination de l'ensemble $RPFRG\ (b)$ constitué des risques dont la panne rend opérationnel le LSP de secours b}

$RPFRG\ (b) \leftarrow \{r \setminus r \in Rs \wedge Op(b, r) = 1\}$

2. {Détermination des liens devant être contournés par le LSP de secours b}

$E" \leftarrow \{\lambda \in E \setminus \exists\, r \in RPFRG\ (b) : \lambda\ est\ supporté\ par\ le\ risque\ r\}$

{Détermination des nœuds devant être contournés par le LSP de secours b}

$V" \leftarrow \{n \in V \setminus \exists\, r \in RPFRG\ (b) : n \in \{r\}\}$

3. {Détermination de l'ensemble des arcs vérifiant les contraintes de bande passantes}

$E' \leftarrow \{\lambda \setminus \lambda \in E \wedge F_\lambda + Max_{r \in RPFRG(b)}\left((\delta')_r^\lambda\right) + bw(b) \leq C^\lambda\}$

4. {Détermination du chemin traversé par b}

Utiliser n'importe quelle méthode de protection locale (protection par LSP de détour ou protection par tunnel de secours) et n'importe quel algorithme de calcul de chemins (comme SPF) pour déterminer le LSP de secours b sur le graphe réduit $G' = (V \setminus V", E' \setminus E")$

fin_algorithme

Avec l'exploitation des structures des SRLG par contre, la probabilité de succès du placement du LSP de secours b'_D est augmentée car l'ensemble des risques de panne que doit contourner le LSP de secours b'_D (cf. étape 1 de l'algorithme 7) est réduit au sous-ensemble $RPFRG\ (b'_D)$ ($RPFRG\ (b'_D) \subseteq PFRG\ (b'_D)$). Typiquement, le lien A-D pourra être traversé par le LSP de secours b'_D car le SRLG *srlg* n'appartient pas à l'ensemble $RPFRG\ (b'_D)$ (en effet, $RPFRG\ (b'_D) = \{D\text{-}F, F\}$). Après l'élimination des lien (D-F) et nœud (F) composant l'ensemble $RPFRG\ (b'_D)$ (étape 2 de l'algorithme 7), nous supprimons à l'étape 3 de l'algorithme 7 le seul arc $A \rightarrow D$, de la topologie du réseau de la figure 4.3 (b), qui ne vérifie pas les contraintes de bande passante ($A \rightarrow D$ ne vérifie pas la formule (4.7)). A l'étape 4 de l'algorithme 7, nous exécutons l'algorithme SPF, sur la topologie du réseau réduite et obtenue après l'exécution des trois premières étapes de l'algorithme 7, pour déterminer l'unique chemin $D \rightarrow A \rightarrow B \rightarrow E \rightarrow G$ interconnectant le nœud D au nœud G (voir la figure 4.4).

Bien qu'ils protègent contre les pannes de liens appartenant au même SRLG, les trois LSP de secours b'_A, b_F et b'_D de la figure 4.4 sont considérés comme indépendants (au sens de l'opération après une panne) et partagent donc entièrement leur bande passante sur les liens en commun. Nous notons que ce type de partage de la bande passante n'induit aucune violation des contraintes de la bande passante puisque les trois LSP de secours b'_A, b_F et b'_D ne peuvent être opérationnels en même temps.

Après avoir montré que l'exploitation des structures des LSP de secours améliore le placement des LSP de secours, nous nous intéressons dans la section qui suit à l'implantation de notre approche dans des environnements centralisés ainsi que dans des environnements distribués.

4.4 Implantation d'un mécanisme de placement de LSP de secours exploitant les structures des SRLG

En fonction de l'environnement d'exécution choisi, l'implantation d'un mécanisme de placement de LSP de secours exploitant les structures des SRLG peut-être plus ou moins simple.

Ainsi, dans un environnement centralisé, le serveur de calcul peut mémoriser la topologie du réseau, les structures des risques de panne et les propriétés et structures de tous les LSP calculés. A partir de ces informations, le serveur déduira tous les paramètres de bande passante (prix effectifs de protection, quantités effectives de bande passante de secours, bandes passantes primaires, surcoûts effectifs, etc.) et les ensembles de risques de panne réellement protégés par un LSP de secours, ce qui lui permettra ensuite de placer en ligne les nouveaux LSP de secours.

Pour améliorer l'efficacité du mécanisme de placement des LSP de secours, le serveur de calcul doit optimiser et réduire au minimum l'ensemble des risques de panne réellement protégés par un LSP de secours avant son calcul. Cela lui offrira plus de flexibilité pour le choix du chemin supportant cet LSP de secours puisque l'ensemble des risques de panne à contourner est minimisé. Ainsi, le serveur centralisé doit établir un ordre de calcul des LSP de secours protégeant un même LSP primaire. En effet, pour déterminer

l'état d'opération final d'un LSP de secours b (cf. section 4.2.1), le serveur centralisé est amené parfois à consulter les structures des LSP de secours (typiquement, leur MP) protégeant contre les pannes de liens situés en amont du LSR de tête du LSP b. En conséquence, le serveur centralisé devrait calculer les LSP de secours en commençant par ceux dont le LSR de tête est le plus proche (en termes de nombre de sauts) du LSR de tête du LSP primaire protégé.

Avec une implantation distribuée du mécanisme de placement des LSP de secours, l'envoi aux entités de calcul d'une information semblable à celle transmise dans les approches décrites dans [KKL+01, LRC02, MBL03, VCLF+04, BML06, SCLR07, SCLR08, SCLRb] est suffisante pour réduire les allocations de la bande passante lors du placement des LSP de secours. En effet, les approches précédentes (mise-à-part l'approche décrite [VCLF+04] qui centralisent le calcul des LSP de secours protégeant contre les risques de types SRLG sur un même serveur) supposent que le nœud sortant de tout arc connaît les structures de tous les LSP de secours qui le traversent (cette information est fournie par le protocole RSVP). En conséquence, ce nœud peut transmettre cette information (ou une information agrégée ou partielle) aux entités de calcul afin de placer efficacement les LSP de secours.

Cependant, pour éviter le contournement inutile de certains liens (lors du calcul d'un nouveau LSP de secours), un ordre de calcul des LSP de secours protégeant un même LSP primaire doit être établi. Pour ce faire, les LSP de secours dont le LSR de tête est le plus proche du LSR de tête du LSP primaire protégé devraient être déterminés en premier (comme dans le cas d'un seul serveur). Diverses approches induisant différentes extensions aux protocoles existants et/ou définissant de nouveaux protocoles peuvent être adoptées pour fixer cet ordre. Par exemple, il est possible d'établir un ordre de calcul des LSP de secours en introduisant de très légères extensions au protocole de signalisation RSVP-TE. Ainsi, avec l'ajout d'un objet BYPASSED_SEGMENT (contenant les extrémités de chaque LSP de secours établi) dans les messages *path* rafraîchissant le LSP primaire protégé, il sera possible de fixer un ordre de calcul des LSP de secours. Concrètement, tout PLR ayant configuré son LSP de secours doit insérer dans le message *path* relatif au LSP primaire protégé les deux nœuds d'extrémité de son LSP de secours. Bien évidemment, l'ordre de calcul des LSP de secours est garanti en obligeant tout PLR à attendre la réception d'un objet BYPASSED_SEGMENT, contenant les nœuds d'extrémité de tous les LSP de secours dont le LSR de tête est situé à son amont, avant de commencer le calcul de son LSP de secours. Nous notons que cette idée d'extension du protocole RSVP-TE avec l'ajout de l'objet BYPASSED_SEGMENT est décrite en détail dans la section 6.2.1 de cette thèse. C'est une idée qui permet d'ailleurs de combiner l'approche exploitant les structures des SRLG avec le partage de la bande passante élargi aux LSP primaires.

4.5 Évaluation des performances

Afin de mesurer les performances de notre approche qui exploite les structures des SRLG pour améliorer le placement des LSP de secours (ESSAPL), nous l'avons com-

paré à l'algorithme TDRA (approche classique n'exploitant pas les structures des SRLG dans le placement des LSP de secours) [SCLR08] et à l'heuristique de Kini (HKA). Nous notons qu'avec l'algorithme TDRA, les entités de calcul disposent de toute l'information nécessaire à la détermination de l'ensemble des LSP actifs, ce qui nous permet de quantifier le gain en performances obtenu en substituant l'ensemble des LSP actifs par l'ensemble des LSP opérationnels. De plus, cet algorithme a un avantage important et consistant à réduire de façon significative le nombre moyen de messages envoyés dans le réseau pour la collecte des informations requises au placement des LSP de secours. Nous avons aussi sélectionné dans nos tests l'heuristiques HKA, pour les mêmes raisons que celles évoquées dans la section 3.2.3.

Pour nos simulations, nous avons opté pour le même modèle et le même environnement que celui décrit dans la section 3.3.3 (même procédé de génération de la matrice de trafic et mêmes topologies de réseau). De plus, comme dans la section 3.3.3, nous illustrons dans les résultats de simulations les valeurs moyennes correspondant à 1000 tests.

4.5.1 Métriques

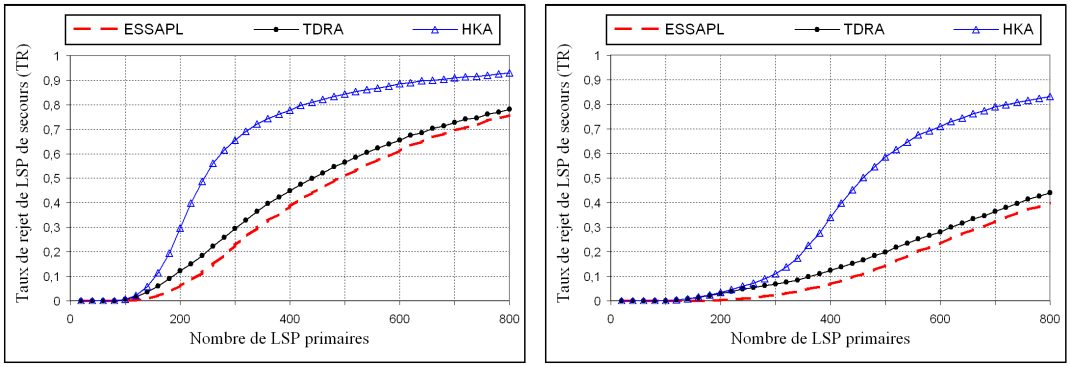
Trois métriques fondamentales ont été choisies pour comparer notre approche ES-SAPL à l'algorithme TDRA et à l'heuristique HKA : le taux de rejet des demandes d'établissement de LSP de secours (TR), le gain relatif dans le rejet des LSP de secours (GRR) et le nombre moyen de messages transmis dans le réseau pour permettre le calcul des LSP de secours (NMM).

La première métrique (TR) détermine le taux de requêtes de protection rejetées. Cette métrique est décrite auparavant dans la section 3.3.3.

La seconde métrique (GRR) permet de déduire le gain dans le taux de rejet des LSP de secours obtenu en adoptant une nouvelle méthode de calcul des LSP de secours à la place d'une ancienne méthode de calcul des LSP de secours. Formellement, le gain relatif dans le rejet des LSP de secours obtenu en utilisant une nouvelle méthode de calcul *nouvMéth* au lieu d'une ancienne méthode *ancMéth* est déterminé comme suit :
GRR (*nouvMéth*, *ancMéth*) = (TR (*ancMéth*) - TR (*nouvMéth*)) / TR (*ancMéth*)
Bien évidemment, plus cette métrique est élevée, plus le taux de rejet de la nouvelle méthode de placement (*nouvMéth*) est faible et est intéressant par rapport à celui de l'ancienne méthode de placement (*ancMéth*).

La troisième métrique NMM compte le nombre moyen de messages transmis dans le réseau pour transmettre aux entités de calcul l'information requise au placement de LSP de secours. Cette métrique est décrite avec détails dans la section 3.3.3.

Contrairement aux valeurs des deux premières métriques TR et GRR, celles de la dernière métrique NMM dépendent fortement du type d'implantation (implantation centralisée ou implantation distribuée) et du mécanisme employé (diffusion ou distribution ciblée) pour la distribution de l'information requise pour le placement des LSP de secours. Dans un environnement centralisé, toute nouvelle requête est envoyée au serveur centralisé qui la traite et envoie les résultats ensuite au routeur ayant initié la requête. Par conséquent, indépendamment de l'approche employée pour le placement des LSP de secours (exploitation ou pas des structures des SRLG), le nombre de mes-



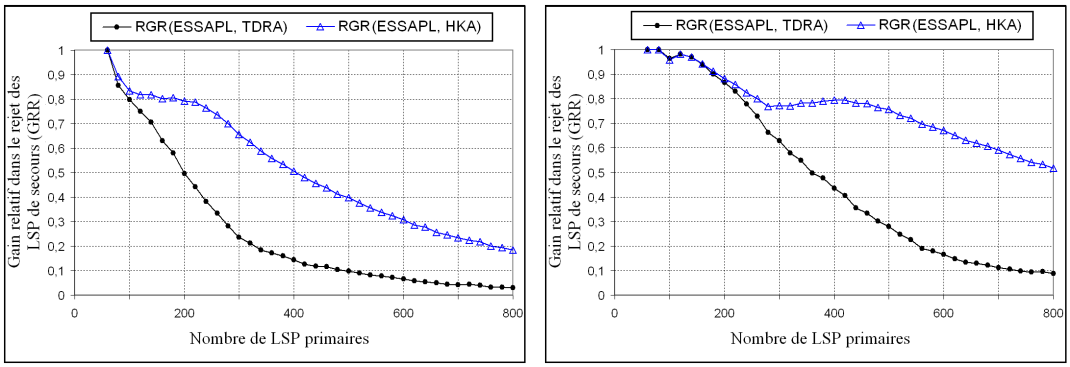
(a) Topologie du réseau de 95 risques (b) Topologie du réseau de 162 risques

FIG. 4.5 – Evolution du taux de rejet des LSP de secours (TR)

sages de contrôle envoyés dans le réseau pour permettre le placement des LSP de secours est le même. De ce fait, il en ressort que la comparaison d'une approche exploitant les structures des SRLG avec une approche centralisée classique (i.e. approche n'exploitant pas les structures des SRLG lors du placement de LSP de secours) n'a aucun intérêt et n'est pas utile. Dans un environnement distribué par contre, le placement des LSP de secours requiert en général un échange d'informations entre les différentes entités de calcul. Deux mécanismes de communication permettent cet échange d'informations : la diffusion ou la distribution ciblée. Dans nos simulations, nous avons opté pour la distribution ciblée car elle réduit la taille de l'information transmise dans le réseau.

4.5.2 Résultats et analyse

La figure 4.5 (resp. figure 4.6) illustre l'évolution du taux de rejet des LSP de secours (resp. le gain relatif dans le rejet des LSP de secours) en fonction du nombre de LSP primaires protégés. La figure 4.5 montre clairement que les taux de rejet des LSP de



(a) Topologie du réseau de 95 risques (b) Topologie du réseau de 162 risques

FIG. 4.6 – Gain relatif dans le rejet des LSP de secours (*GRR*)

secours de l'approche ESSAPL sont meilleurs et plus petits (sauf pour les 100 premiers LSP primaires dans la figure 4.5 (a) et les 140 premiers LSP de primaires dans la figure 4.5 (b)) que ceux de l'algorithme TDRA, qui sont eux même meilleurs et plus petits que ceux de HKA.

Pour les faibles charges du réseau (nombre de LSP primaires inférieur à 100 dans la figure 4.5 (a) et inférieur à 140 dans la figure 4.5 (b)), nous remarquons que l'approche ESSAPL a des taux de rejet des LSP de secours très proches de ceux de l'algorithme TDRA et de l'heuristique HKA. Ceci s'explique par les valeurs des quantités de la bande passante de secours allouées sur les arcs qui sont petites (inférieures aux capacités de secours des arcs moins la quantité maximale de bande que peut réclamer un LSP), ce qui permet d'éviter le rejet de presque toutes requêtes de protection. En effet, pour ces charges faibles du réseau, nous apercevons sur la figure 4.5 que les TR des trois méthodes de placement comparées sont quasi-nuls.

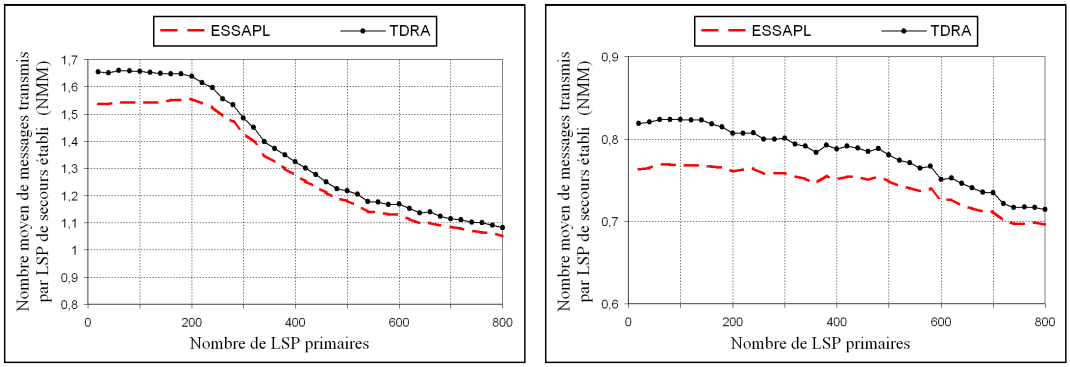
Lorsque la charge du réseau augmente, les TR des trois méthodes de placement des LSP de secours comparées ici se distinguent les uns des autres. Ainsi, l'approche ESSAPL présente des taux de rejet inférieurs à ceux de l'algorithme TDRA qui sont eux même inférieurs à ceux de l'heuristique HKA.

La différence de valeurs entre les taux de rejet des LSP de secours de l'algorithme TDRA (ou de l'approche ESSAPL aussi) et de l'heuristique HKA est due à l'utilisation d'une information complète dans TDRA pour déterminer l'ensemble des LSP actifs alors que HKA n'emploie qu'une information partielle et incomplète. Dans cette section, nous ne nous intéressons pas davantage à la comparaison des performances de l'heuristique HKA à l'algorithme TDRA. De plus amples détails sur cette comparaison (notamment la comparaison des taux de rejet des LSP de secours) sont fournis dans la section 3.2.3.2.

Concernant la différence des taux de rejet des LSP de secours de l'algorithme TDRA avec l'approche ESSAPL, la figure 4.5 (a) (resp. la figure 4.5 (b)) montre que l'approche ESSAPL permet de diminuer les taux de rejet des LSP de secours de 4,3% à 7.5% (resp. de 2.5% à 6%) lorsque le nombre de LSP primaires varie entre 200 et 600 (resp. entre 200 et 800). Cette différence s'explique par l'exploitation des structures des SRLG dans l'approche ESSAPL (et pas dans l'algorithme TDRA), ce qui permet de réduire l'ensemble des liens à contourner et réduit les allocations de la bande passante de secours.

Par ailleurs, la comparaison des résultats des taux de rejet obtenus sur le réseau de la figure 3.3 (a) avec ceux du réseau de la figure 3.3 (b) montre que la distance entre les valeurs des taux de rejet de l'algorithme TDRA et l'approche ESSAPL est plus élevée dans le premier réseau. Ceci s'explique essentiellement par la densité des SRLG qui est plus élevée dans le premier réseau. En effet, les deux méthodes de placement des LSP de secours TDRA et ESSAPL ne se distinguent que par les ensembles de SRLG réellement protégés.

Bien que la baisse du taux de rejet des LSP de secours apportée par la substitution de ESSAPL à TDRA semble petite et limitée par rapport au nombre totale de requêtes de protection, elle reste intéressante et très désirée, surtout lorsque le blocage des requêtes de protection n'est pas autorisé ou doit être très restreint. Par exemple, lorsque le taux de rejet des LSP de secours n'est pas autorisé, la figure 4.6 (a) (resp. la figure 4.6 (b)) montre que l'adoption de l'approche ESSAPL au lieu de l'algorithme



(a) Topologie du réseau de 95 risques (b) Topologie du réseau de 162 risques

FIG. 4.7 – Evolution du nombre moyen de messages transmis dans le réseau pour l'établissement d'un LSP de secours (NMM)

TDRA permet d'augmenter le nombre de LSP primaires établis et protégés de 40 à 60 LSP (resp. de 40 à 80), soit une augmentation de 50% (resp. 100%) du nombre total de LSP entièrement protégés. Lorsque le rejet des requêtes de protection est autorisé, les figures 4.5 (a) et 4.6 (a) (resp. 4.5 (b) et 4.6 (b)) montrent que pour des taux de rejet faibles et pratiques (situé entre 0 et 0.1), la sélection de l'approche ESSAPL à la place de l'algorithme TDRA (ou de l'heuristique HKA) augmente le gain dans le rejet des requêtes de protection de plus de 75% (resp. de près de 95%). Cela montre que l'exploitation des SRLG pour le placement des LSP de secours est surtout importante pour éviter les blocages des requêtes de protection ou pour assurer un taux de rejet des requêtes de protection petit et pratique. La conception de réseaux tolérants aux pannes doit donc prendre en compte les structures des SRLG pour minimiser les coûts de déploiement et améliorer la protection.

En ce qui concerne la dernière métrique *NMM*, la figure 4.7 montre que le nombre moyen de messages envoyés dans le réseau avec l'approche ESSAPL est plus petit que celui transmis avec l'algorithme TDRA. Cela est dû à la réduction, avec ESSAPL, de l'ensemble des risques protégés par un LSP de secours à ceux provoquant uniquement son opération. Alors que l'algorithme TDRA envoie systématiquement les structures et propriétés des LSP de secours protégeant contre la panne d'un lien à tous les nœuds d'extrémité des SRLG contenant le lien protégé, l'approche ESSAPL n'envoie les structures et propriétés des LSP de secours déterminés qu'aux nœuds d'extrémité de liens appartenant aux SRLG réellement protégés par le LSP de secours (l'ensemble des SRLG réellement protégés par un LSP de secours est toujours inclus dans l'ensemble des SRLG contenant le lien protégé).

4.6 Conclusion

Dans ce chapitre, nous avons montré que certains LSP de secours activés après une panne d'un SRLG n'utilisent pas effectivement leurs ressources, puisqu'ils ne reçoivent

aucun flux de données après la récupération. Ces LSP sont *inopérationnels* et sont complètement déterminés en exploitant les structures des SRLG.

Afin d'améliorer le placement des LSP de secours, nous avons proposé une nouvelle approche ESSAPL exploitant les structures des SRLG. Constatant que seuls les LSP *opérationnels* (LSP actifs et recevant un flux de données) après une panne utilisent effectivement leurs ressources, ESSAPL réduit la quantité de bande passante de secours réservée sur tout lien en limitant la concurrence pour les allocations de la bande passante aux LSP opérationnels. De plus, pour réduire les blocages des requêtes de placement de LSP de secours, ESSAPL permet de mieux explorer la topologie du réseau en construisant des LSP de secours ne contournant pas systématiquement tous les liens appartenant à des SRLG contenant le lien protégé.

En plus de sa facilité à mettre en œuvre et à déployer dans des environnements centralisés ou distribués, notre approche est utile pour mieux concevoir et dimensionner les réseaux (elle tient compte des structures des SRLG lors du dimensionnement d'un réseau). Les simulations montrent que l'approche ESSAPL réduit le taux de rejet des LSP de secours et améliore le gain relatif dans le rejet des LSP de secours. Par ailleurs, l'implantation distribuée de l'approche ESSAPL permet de diminuer le nombre de messages de contrôle nécessaires à la transmission de l'information requise au placement des LSP. En effet, avec ESSAPL, l'ensemble des risques de panne protégés par un LSP de secours est réduit et ne contient que les risques dont la panne rend opérationnel ce LSP de secours.

Chapitre 5

Partage de ressources et placement local de LSP de secours multicast

5.1 Introduction

Le développement et la diversité des technologies (wifi, lignes spécialisées, etc.) permettant l'accès à l'Internet, ainsi que les récentes avancées dans le domaine des transmissions numériques (par ex. utilisation des fibres optiques) ont permis aux applications gourmandes en bande passante de voir le jour et de se répandre largement. Parmi les applications qui sont très populaires chez les utilisateurs, quelques unes (comme la vidéo à la demande, la vidéo-conférence, les jeux en réseau, enseignement à distance, les enchères en direct, etc.) sont de type temps réel et font intervenir plusieurs participants. Afin d'offrir une plate-forme d'exécution propice à ces applications, deux procédés doivent être fournis au niveau du réseau. Ces deux procédés sont le multicast et la protection.

Le multicast [Dee91] permet d'envoyer à un groupe de destinataires (membres du groupe multicast) des données de manière efficace, c'est-à-dire sans que ces données ne soient dupliquées sur les liens du réseau. Pour ce faire, un arbre couvrant tous les membres multicast est généralement utilisé pour router le trafic. Une quantité élevée de ressources est alors économisée en partageant les ressources sur les portions communes des routes composant l'arbre multicast. Deux modèles de communication multicast existent : SSM (Source Specific Multicast) où une seule station est autorisée à envoyer le trafic aux membres multicast et ASM (Any Source Multicast) où n'importe quelle station peut émettre le trafic vers les membres multicast (modèle proposé originellement par Deering).

La protection (cf. section 1.4 et [MVDBD04]), quant à elle, a pour but d'éviter ou de réduire le temps de coupure des communications suite aux pannes. C'est une technique pré-calculant et souvent pré-configurant des routes de secours capables de recevoir et de rerouter le trafic des communications primaires affectées par une panne.

Bien qu'il existe aujourd'hui plusieurs méthodes permettant de protéger efficacement les communications unicast (communication avec une seule source et une seule destination), la recherche de nouvelles méthodes de protection propres au multicast est

nécessaire et a une grande importance. En effet, l'application des techniques de protection unicast pour protéger les chemins primaires d'un arbre multicast augmente le volume de l'information gérée et maintenue pour assurer la restauration, peut diminuer les possibilités de partage des ressources et introduit souvent de nouveaux inconvénients comme le risque d'apparition de boucles ou de duplication de trafic sur certains liens du réseau [SCM06b].

Avec l'avènement de MPLS (cf. chapitre 1), la fonctionnalité de protection a été améliorée grâce à la grande flexibilité dans le choix des chemins. Ainsi, les chemins de secours peuvent être choisis de manière à minimiser la quantité de ressources (particulièrement la bande passante) qui leur est allouée (grâce à l'utilisation du partage des ressources). De plus, l'utilisation de LSP permet d'éliminer les boucles par l'utilisation d'étiquettes différentes pour identifier des chemins de secours et primaires différents.

Afin de permettre l'établissement, sous MPLS, d'arbres multicast basés à la source ou des LSP point à multipoint (*P2MP LSPs*), [APY07] a étendu le protocole RSVP-TE décrit dans [ABG+01]. Ainsi, lorsqu'un nouveau nœud adhère à un groupe multicast, le nœud source de la session correspondante est informé. Afin d'alimenter le nouveau membre multicast, le nœud source configure un nouveau LSP unicast (S2L LSP ou *Source to Leaf* LSP) permettant d'atteindre le nouveau membre à partir de la source multicast et le rajoute ensuite à l'ancien LSP P2MP. Bien évidemment, le LSP S2L sera fusionné avec l'ancien P2MP de manière à obtenir un nouveau LSP P2MP évitant la duplication inutile des données sur ses liens.

Dans ce chapitre, nous allons étudier dans les réseaux MPLS les techniques et les mécanismes permettant le calcul en ligne des routes offrant la protection multicast et minimisant l'utilisation de la bande passante par le partage. Nous donnons un aperçu des différentes méthodes de protection distribuées qui pourraient être appliquées au multicast (protection par chemins disjoints, protection par arbre redondant, protection par forêt duale, protection multicast un-à-un, protection par tunnel P2P de secours et protection multicast par tunnel P2MP de secours). Pour chaque méthode, nous spécifions la structure de secours (ou bien les routes de secours) employée et les procédés permettant son calcul de manière à maximiser la disponibilité de la bande passante (sur chaque lien) en utilisant deux différentes stratégies de partage de la bande passante : *le partage restreint de la bande passante* (*PRB*) et *le partage global de la bande passante* (*PGB*). Nous notons qu'avec la première stratégie, le partage de la bande passante est restreint aux LSP de secours. Avec la seconde stratégie, le partage de la bande passante est étendu et est appliqué entre les LSP de secours d'un coté et entre les LSP primaires et LSP de secours d'un autre coté (i.e. la bande passante libérée par les LSP primaires suite à une panne peut être ré-allouée à des LSP de secours). Pour des raisons liées à MPLS, nous adoptons l'approche décrite dans [APY07], basée sur le modèle SSM (*Source-Specific multicast*), afin d'établir les chemins supportant les communications multicast (point à multipoint).

Dans la suite de ce chapitre, nous décrivons dans la section 5.2, le contexte et les outils qui nous permettent de modéliser et comprendre le problème de calcul des routes de secours maximisant la disponibilité de la bande passante par le partage. Ensuite, nous étudions brièvement en section 5.3 les méthodes de protection multicast de niveau

global. Puis nous nous concentrons sur les méthodes de protection de niveau local en section 5.4. Après la description de la méthode de protection par forêt duale (section 5.4.1) qui nécessite une signalisation des chemins de secours après la panne, nous nous concentrons sur les méthodes de protection minimisant les délais de récupération. Ainsi, nous introduisons, dans les sections 5.4.2 et section 5.4.3, les méthodes de protection locale inspirées de l'unicast et permettant de protéger les communications multicast au détriment d'un gaspillage de la bande passante ou de la croissance du volume des tables d'étiquettes sur les routeurs MPLS. Ensuite, nous nous intéressons (section 5.4.4) à la protection locale par tunnels P2MP de secours qui est développée exclusivement pour la protection multicast. La dernière section sera consacrée aux conclusions.

5.2 Configuration des LSP de secours, contexte et modélisation

Avant d'introduire le problème de calcul des LSP protégeant une session multicast, nous décrivons brièvement ci-après le protocole RSVP-TE et les extensions introduites par [APY07] afin de permettre l'établissement de LSP P2MP sous MPLS. Cette description est nécessaire et est utile pour comprendre les différentes extensions que nous proposons pour introduire le partage de la bande passante lors du placement des LSP de secours protégeant des sessions multicast.

5.2.1 LSP point à multipoint sous MPLS

Dans le premier chapitre, nous avons décrit le protocole RSVP-TE qui permet d'établir des routes explicites, avec ou sans réservation de ressources, pour le support et la protection des communications unicast. Pour prendre en compte les communications point à multipoint, Aggarwal et al. ont étendu, dans [APY07], ce protocole RSVP-TE et ont défini de nouveaux objets permettant la gestion de plusieurs destinations (membres multicast). Ainsi, de nouvelles améliorations ont été apportées à RSVP-TE pour diminuer la taille du trafic de contrôle, réduire le nombre et la taille des états RSVP et éviter la duplication (inutile) du trafic sur les liens.

Pour distinguer et supporter le flux de données destiné à un groupe multicast, [APY07] a défini la notion de tunnels P2MP (extension de la notion de tunnel unicast). Ainsi, un tunnel P2MP (tunnel RSVP-TE) est composé d'un ensemble de LSP P2MP et est identifié grâce à l'objet SESSION. Chaque LSP P2MP est lui-même identifié grâce à l'objet SESSION qui détermine le tunnel LSP P2MP associé et à l'objet SENDER_TEMPLATE (ou FILTER_SPEC) qui contient la source et l'identifiant du LSP P2MP. Ce dernier (LSP P2MP), est constitué d'un ensemble de LSP point à point, dits LSP S2L.

Tous les LSP S2L, appartenant à un même LSP P2MP, partagent la même source et sont configurés suivant des procédures très similaires à celles établissant des LSP unicast sous RSVP. Cependant et afin de diminuer le trafic de contrôle, un mécanisme de compression de routes est utilisé. Ce mécanisme subdivise un LSP P2MP en plu-

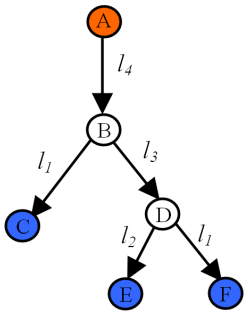


FIG. 5.1 – Arbre point à multipoint

sieurs segments de chemins ne partageant aucun lien. Sur la figure 5.1 par exemple, le nœud A, qui est la source multicast, envoie à son prochain nœud B une structure compressée du LSP P2MP permettant d'alimenter les trois membres multicast C, E et F. Cette structure peut consister en : $B \rightarrow C$, objet S2L_SUB_LSP de C, $B \rightarrow D \rightarrow E$, objet S2L_SUB_LSP de E, $D \rightarrow F$, objet S2L_SUB_LSP de F.
Nous notons qu'un objet S2L_SUB_LSP identifie un nœud membre multicast et que la combinaison de cet objet avec les objets SESSION, SENDER_TEMPLATE (ou FILTER_SPEC) permet d'identifier un seul LSP S2L.

Pour établir le LSP P2MP illustré sur la figure 5.1, le nœud A construit un message *path* incluant des objets permettant d'identifier tous les LSP S2L (SESSION, SENDER_TEMPLATE et la liste des objets S2L_SUB_LSP) et d'autres informations contenant entre autres les propriétés du trafic (SENDER_TSPEC), le nœud RSVP en amont, etc. Ce message sera reçu et traité (création d'état RSVP) par le nœud B qui en déduit et construit deux autres messages *path* destinés à ses deux prochains nœuds C et D. De la même manière que le nœud B, le nœud D traitera le message reçu, en déduira et redirigera deux autres messages *path* vers ses deux prochains nœuds E et F.

Lorsqu'un nœud feuille d'un LSP S2L reçoit un message *path*, il effectue un contrôle d'admission puis réserve les ressources (la bande passante et une étiquette) sur son lien en amont. Après cela, le nœud feuille envoie à son nœud en amont un message *resv*. Ce dernier contient diverses informations comme les objets permettant d'identifier le LSP S2L concerné, l'étiquette allouée, la route supportant le LSP S2L, les priorités, les politiques sélectionnées, etc. Pour l'exemple de la figure 5.1, le nœud C alloue l'étiquette l_1 sur le lien $B \rightarrow C$ (après la vérification du respect des contraintes de la bande passante) et envoie cette étiquette dans un message *resv* à son nœud en amont B. De la même manière, le nœud E (resp. F, D, B) alloue l'étiquette l_2 (resp. l_1, l_3, l_4)[1] sur le lien $D \rightarrow E$ (resp. $D \rightarrow F$, $B \rightarrow D$ et $A \rightarrow B$) et l'envoie à son nœud en amont D (resp. D, B, A). La configuration de LSP P2MP s'achève après la réception du nœud A d'un message *resv*

[1] Lorsque le *penultimate hop* est activé dans un réseau MPLS, aucune étiquette n'est allouée sur les liens en aval des nœuds feuilles d'un LSP P2MP. Dans l'exemple de la figure 5.1, nous avons choisi de désactiver le *penultimate hop* afin de simplifier le discours et faciliter la compréhension du processus d'échange d'étiquettes.

contenant tous les objets S2L_SUB_LSP.

Nous notons que les différents messages *resv* associés aux LSP S2L composant le LSP P2MP suivent les chemins inverses de ceux parcourus par les messages *path* correspondants. De plus, les messages *path* et *resv* peuvent être fragmentés en plusieurs messages RSVP-TE afin de traiter les cas où il n'est pas possible d'inclure tous les S2L dans un seul message IP. Pour cela, [APY07] propose d'utiliser le couple (*Sub-Group Originator ID*, *Sub-Group ID*), transporté dans l'objet SENDER_TEMPLATE[2], afin de distinguer les différents messages RSVP-TE associés à un même LSP P2MP. Le premier champ est un identificateur du nœud ayant effectué la fragmentation et le second est un identifiant permettant de distinguer les différents messages *path* ou *resv* associés à un même LSP P2MP.

Pour protéger un LSP P2MP primaire, plusieurs LSP P2MP de secours (composés chacun de plusieurs LSP S2L) pourraient être nécessaires. Deux méthodes permettent d'identifier les messages associés à ces LSP de secours (cf. section 1.5.3) : identification par la source ou identification par chemin. Avec la première méthode d'identification, les LSP de secours utilisent les mêmes objets SESSION et S2L_SUB_LSP ainsi que le même identificateur de LSP (*LSP ID*) que ceux transmis dans les messages correspondant au LSP primaire S2L protégé ; seule l'adresse de la source du LSP (*ipv4/ipv6 tunnel sender address*) est modifiée et est remplacée par une adresse du PLR (adresse différente de celle transmise dans les messages associés au LSP primaire protégé). Avec la seconde méthode d'identification, les objets SESSION et SENDER_TEMPLATE des messages configurant le LSP primaire sont copiés dans les messages correspondant à leurs LSP de secours ; pour différencier les messages de ces deux types de LSP (primaire et secours), l'objet DETOUR, défini dans [PSA05], est inclus dans les messages configurant les LSP de secours.

Concernant le plan de données, [APY07] n'introduit pas de grands changements (au modèle SSM du multicast). Ainsi, les paquets sont dupliqués sur chaque routeur de branchement du LSP P2MP pour être acheminés vers tous les nœuds membres du groupe multicast. Sur la figure 5.1 par exemple, l'envoi d'un paquet aux membres du groupe multicast C, E et F passe par les étapes suivantes : le nœud A envoie le paquet au nœud B en insérant l'étiquette l_4 dans le message. Le nœud B recevant le message de A le duplique, envoie la première copie du message au nœud C après avoir échangé l'étiquette l_4 par l'étiquette l_1 et envoie la seconde copie du message au nœud D après avoir échangé l'étiquette l_4 par l'étiquette l_3. Un traitement similaire, à celui qui est accompli par le nœud B, est effectué sur le nœud D. Lorsque le nœud C (resp. E, F) reçoit le message émis par le nœud B (resp. D, D), il lui retire l'étiquette et l'envoie ensuite à la station membre.

[2]L'objet SENDER_TEMPLATE contient en plus un champ appelé *sender template-specific identification*.

5.2.2 Contexte et problème de partage de la bande passante entre les sessions multicast

Étant donné un réseau MPLS représenté par un graphe $G = (V, E, R)$ où V est l'ensemble des LSR, E est l'ensemble des liens (arêtes) et R est l'ensemble des risques. A chaque arc du graphe est associé un poids représentant sa capacité en bande passante. Des demandes d'établissement de LSP unicast (point à point ou P2P) et multicast (P2MP) arrivent en ligne sans connaissance préalable, ni de leur ordre d'arrivée, ni de la quantité de la bande passante réclamée par chacun d'eux. Quelques requêtes exigent une protection proactive contre les pannes simples et d'autres pas.

L'objectif consiste à déterminer un routage distribué permettant d'alimenter tous les membres multicast en utilisant une route minimisant la consommation de la bande passante[3] et réduisant ainsi la probabilité de blocage (rejet) des requêtes formulées. Lorsque la requête réclame la protection, un routage de secours sûr (assurant la disponibilité de la bande passante après n'importe quelle panne) et optimisant la quantité de bande passante (additionnelle) allouée dans le réseau doit être prévu. Sous MPLS, notre objectif se traduit par la recherche de procédures permettant le placement distribué d'un LSP P2MP (resp. P2P) primaire et de l'ensemble des LSP de secours permettant sa protection (lorsque cela est exigé), pour toute nouvelle requête multicast (resp. unicast). Les différents LSP (primaire et de secours) doivent vérifier les contraintes de bande passante et appliquer le partage de ressources (particulièrement le partage de la bande passante).

Pour atteindre l'objectif ci-dessus, il est nécessaire de définir deux procédés : le procédé de calcul des LSP et le procédé de distribution des données nécessaires au calcul. Le premier procédé consiste à définir les algorithmes permettant le calcul des LSP primaires ainsi que des LSP de secours. Deux approches peuvent être adoptées pour rendre les fonctionnalités du premier procédé : l'approche successive (*non-joint approach*) et l'approche conjointe (*joint approach*). Dans la première approche, deux calculs différents et successifs sont effectués pour satisfaire une requête d'établissement d'un LSP primaire nécessitant une protection : un premier calcul du LSP primaire qui minimise la bande passante primaire, puis un deuxième calcul des LSP de secours protégeant le LSP primaire et minimisant la bande passante de secours. Dans la seconde approche, le LSP primaire et l'ensemble des LSP de secours qui le protègent sont calculés conjointement, de manière à minimiser la consommation de la bande passante. La première approche a l'avantage d'être facile à implanter dans des environnements distribués alors que la seconde permet une meilleure optimisation de la bande passante. Concernant le second procédé, il a pour tâche de déterminer les données ainsi que les procédures de distribution de ces données pour permettre le calcul des routes (effectué en utilisant le premier procédé). Concrètement, les nœuds rendent les fonctionnalités du second procédé en stockant l'information sur la topologie et les LSP établis, et en transmettant/diffusant cette information aux nœuds calculant les routes.

[3] Afin d'optimiser la bande passante allouée dans le réseau, la tendance est de minimiser la quantité de bande passante additionnelle réservée à tout nouveau LSP. Dans un réseau dédié au support d'applications temps réel, la métrique de délai est souvent plus importante que la métrique mesurant la quantité de bande passante allouée dans le réseau.

Dans ce chapitre, nous ne nous intéresserons pas aux algorithmes de calcul des LSP (car de nombreux algorithmes, comme [Dij59, TM80, KMB81], existent) mais seulement aux méthodes de protection employées dans le multicast, aux algorithmes de calcul de la bande passante de secours allouée sur les liens et aux procédés de distribution de l'information requise au placement distribué des LSP de secours.

5.2.3 Modélisation et grandeurs caractéristiques

Comme dans la section précédente, nous modélisons un réseau MPLS par un graphe $G = (V, E, R)$ où V est l'ensemble des LSR, E est l'ensemble des liens (arêtes) et R est l'ensemble de tous les risques de panne du réseau. Sur le réseau, un certain nombre k ($k \geq 0$) de requêtes d'établissement de sessions multicast protégées ou non protégées (une session unicast peut être considérée comme une session multicast avec une source et une seule destination) ont été déjà satisfaites avec la configuration de k LSP primaires et k' (sans perte de généralité, nous pouvons considérer ici que $k' = k$) structures de secours (une structure de secours est un ensemble de LSP P2MP de secours protégeant un même LSP primaire). Tous les LSP dédiés à une session multicast i réclament la même quantité de bande passante F^i. Le LSP primaire routant le trafic de la session multicast i est noté P_i et la structure de secours qui le protège est notée B_i. A chaque session multicast i est associés un nœud source s_i et un ensemble de nœuds membres multicast (destinations) M^i. De cette manière, un LSP primaire P2MP P_i sera composé d'un ensemble $\{p_{j,i}\}_{(0<j\leq|M^i|)}$ de LSP S2L.

Afin de gérer la bande passante et le partage, nous définissons les grandeurs suivantes :

- F^i : quantité de bande passante réclamée par P_i.
- F^i_λ : lorsque l'arc λ appartient au LSP P2MP P_i, F^i_λ est égale à la quantité de bande passante réclamée par P_i ; sinon, cette quantité est nulle.
- F^i_r : quantité de bande passante cumulée de tous les LSP S2L de P_i qui passent par le risque r.
- F_λ (k) : quantité de bande passante cumulée et allouée sur l'arc λ pour les k premiers LSP P2MP primaires :

$$F_\lambda(k) = \sum_{i=1}^{k} F^i_\lambda$$

- F_r (k) : quantité de bande cumulée et allouée sur le risque r pour les LSP S2L des k premiers LSP primaires :

$$F_r(k) = \sum_{i=1}^{k} F^i_r$$

- $L^{\lambda,\,i}_r$: quantité de bande passante libérée par le LSP P2MP primaire P_i sur l'arc λ suite à la panne du risque r.

- $L_r^{\lambda}\ (k)$: quantité de bande passante libérée, sur l'arc λ, par les k premiers LSP P2MP primaires, suite à la panne du risque r. Cette quantité est déterminée comme suit :

$$L_r^{\lambda}\,(k) = \sum_{i=1}^{k} L_r^{\lambda,\,i}$$

- $G^{\lambda}\ (k)$: quantité de bande passante de secours allouée pour la protection des k premiers LSP de secours. Cette quantité dépend des types de signalisation (cf. sections 5.3.1.1 et 5.3.1.2) et de stratégie de partage employés.

Lorsque le partage de la bande passante est restreint, nous définissons :

- $\delta_r^{\lambda,\,i}$: prix de protection du risque r sur l'arc λ pour la session i. Ce prix correspond à la bande passante cumulée de tous les LSP P2MP de secours protégeant la session i et activés pour faire face à la panne du risque r.
- $\delta m_r^{\lambda,\,i}$: prix *réduit* de protection du risque r sur l'arc λ pour la session k. Ce prix est égal à la bande passante de P_k s'il existe un LSP P2MP de secours appartenant à B_k, protégeant contre le risque de panne r et traversant l'arc λ ; autrement, $\delta m_r^{\lambda,\,i}$ est nul.
- $\delta_r^{\lambda}\ (k)$: prix de protection du risque r sur l'arc λ pour les k premières sessions. Ce prix est calculé comme suit :

$$\delta_r^{\lambda}\,(k) = \sum_{i=1}^{k} \delta_r^{\lambda,\,i}$$

- $\delta m_r^{\lambda}\ (k)$: prix de protection *réduit* du risque r sur l'arc λ pour les k premières sessions. Ce prix est calculé comme suit :

$$\delta m_r^{\lambda}\,(k) = \sum_{i=1}^{k} \delta m_r^{\lambda,\,i}$$

- $\theta^{\lambda}\ (b)$: coût de protection du LSP P2MP de secours b sur l'arc λ. Ce coût correspond à la quantité de bande passante cumulée et maximale de tous les LSP P2MP de secours qui pourraient être activés sur l'arc λ suite à la panne de n'importe quel risque induisant l'activation du LSP de secours b. Par conséquent, le coût de protection, sur l'arc λ, d'un LSP P2MP de secours b protégeant la session multicast *k+1* est calculé comme suit :

$$\theta^{\lambda}\,(b) = Max_{r\,\in\,PFRG\,(b)}\,\delta_r^{\lambda}\,(k)$$

- $\theta m^{\lambda}\ (b)$: coût *réduit* de protection du LSP P2MP de secours b sur l'arc λ. Ce coût *réduit* correspond au prix *réduit* de protection (de toutes les sessions établies et protégées) le plus élevé des risques protégés par le LSP de secours b. En conséquence, le coût de protection *réduit*, sur l'arc λ, d'un LSP P2MP de secours b protégeant la session multicast *k+1*, est calculé comme suit :

$$\theta m^{\lambda}\,(b) = Max_{r\,\in\,PFRG\,(b)}\,\delta m_r^{\lambda}\,(k)$$

Lorsque le partage de la bande passante est global, nous avons :

- $(\delta*)_r^\lambda\ (k)$: prix de protection du risque r sur l'arc λ pour les k premières sessions. Ce prix est calculé comme suit :

$$(\delta*)_r^\lambda\ (k) = \sum_{i=1}^{k} (\delta_r^{\lambda,\,i} - L_r^{\lambda,\,i})$$

- $(\delta*)m_r^\lambda\ (k)$: prix de protection *réduit* du risque r sur l'arc λ pour les k premières sessions. Ce prix est calculé comme suit :

$$(\delta m*)_r^\lambda\ (k) = \sum_{i=1}^{k} (\delta m_r^{\lambda,\,i} - L_r^{\lambda,\,i})$$

- $(\theta*)^\lambda\ (b)$: coût de protection du LSP P2MP de secours b sur l'arc λ. Ce coût de protection, sur un arc λ d'un LSP P2MP de secours b protégeant la session multicast *k+1*, est calculé comme suit :

$$(\theta*)^\lambda\,(b) = Max_{r\,\in\,PFRG\,(b)}\,(\delta_r^\lambda\,(k) - L_r^\lambda\,(k))$$

- $(\theta m*)^\lambda\ (b)$: coût *réduit* de protection du LSP P2MP de secours b sur l'arc λ. Ce coût *réduit*, sur un arc λ d'un LSP P2MP de secours b protégeant la session multicast *k+1*, est calculé comme suit :

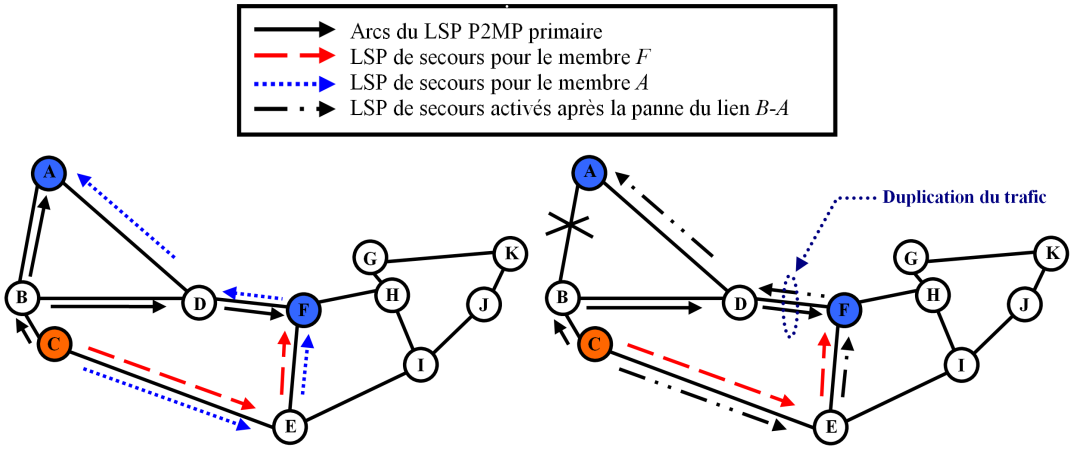
$$(\theta * m)^\lambda\,(b) = Max_{r\,\in\,PFRG\,(b)}\,(\delta_r^\lambda\,(k) - L_r^\lambda\,(k))$$

Après la définition des grandeurs nous permettant de spécifier et calculer les quantités de bande passante allouées sur les arcs, nous décrivons ci-après les différentes méthodes de protection multicast existantes. Pour chacune de ces méthodes, nous précisons l'information nécessaire au calcul des LSP de secours, donnons un ou plusieurs algorithmes transmettant cette information aux entités de calcul et déterminons enfin le degré de partage qui pourrait être atteint sous MPLS avec cette méthode. Pour des raisons évidentes, liées à la minimisation du délai de récupération, nous ne nous intéressons dans la suite de cette thèse qu'aux méthodes de protection proactive (globale ou locale).

5.3 Protection proactive globale

Avec la protection multicast proactive globale, le LSP P2MP primaire est protégé par une structure de secours reliant la source aux différents nœuds membres multicast. Lorsqu'un nœud du LSP P2MP primaire détecte une panne, il la signale au nœud source en lui envoyant un message de notification de la panne. A la réception de ce message de notification par le nœud source multicast, ce dernier active toute ou une partie de la structure de secours afin d'atteindre de nouveau les membres multicast affectés suite à la panne.

Nous distinguons essentiellement deux méthodes de protection multicast proactive globale : protection par chemins disjoints et protection par arbre redondant.



(a) Situation avant la panne du lien B-A

(b) Situation après la détection et la récupération de la panne du lien B-A

FIG. 5.2 – Protection par chemins disjoints

5.3.1 Protection par chemins disjoints (*end-to-end protection*)

Cette méthode de protection [MVDBD04, RM99a], qui provient de l'unicast, subdivise le LSP P2MP primaire en l'ensemble des LSP S2L permettant d'atteindre chaque membre multicast à partir de la source. Chaque LSP S2L (primaire) est protégé par un LSP de secours disjoint (pas de nœud, lien ou SRLG en commun) et reliant ses deux nœuds d'extrémité.

Sur la figure 5.2 (a), un LSP P2MP permettant de relier la source multicast C à deux nœuds membres multicast A et F est établi. Afin de protéger ce LSP P2MP primaire avec la méthode de protection par chemins disjoints, le nœud source déduit les LSP S2L primaires permettant d'atteindre les membres multicast à partir de la source multicast. Pour chacun de ces LSP S2L, le nœud source calcule et souvent configure un LSP P2P de secours disjoint et reliant les extrémités du LSP S2L protégé. Pour l'exemple de la figure 5.2 (a), le nœud source C déduit les deux LSP S2L primaires $C \rightarrow B \rightarrow A$ et $C \rightarrow B \rightarrow D \rightarrow F$ permettant d'atteindre à partir de la source multicast C les deux membres multicast A et F. Ensuite, le nœud source C calcule deux LSP de secours : le premier LSP $C \rightarrow E \rightarrow F \rightarrow D \rightarrow A$, protégeant le nœud membre A, relie le nœud source C au nœud membre A par un chemin de secours disjoint du LSP S2L primaire $C \rightarrow B \rightarrow A$; le second LSP $C \rightarrow E \rightarrow F$, protégeant le nœud membre F, relie le nœud source C au nœud membre F par un chemin disjoint du LSP S2L primaire $C \rightarrow B \rightarrow D \rightarrow F$.

Lorsqu'une panne est détectée, un message de notification de la panne est envoyé au nœud source multicast. A la réception de ce message, le nœud source détermine l'ensemble des nœuds membres affectés par la panne, signalise éventuellement les LSP de secours protégeant les membres affectés et envoie le trafic sur la partie résiduelle du LSP P2MP primaire ainsi que sur les LSP de secours activés pour faire face à la panne. Sur la 5.2 (b), le nœud source multicast C déduit que seul le membre multicast A est affecté suite à la réception d'un message notifiant la panne du lien B-A. Afin de

pallier cette panne, le nœud source C activera[4] le LSP de secours $C \rightarrow E \rightarrow F \rightarrow D \rightarrow A$ protégeant le nœud membre A en envoyant le trafic sur le LSP P2MP primaire résiduel et sur le LSP de secours nouvellement activé. Ainsi, le nœud membre F sera desservi en utilisant le LSP P2MP résiduel (C→B→D→F) alors que le nœud membre A l'est en utilisant le LSP de secours (C→E→F→D→A). Comme c'est illustré sur la 5.2 (b), la récupération de la panne du lien B-A induit une duplication du trafic sur le lien D-F, ce qui gaspille la bande passante.

Concernant les possibilités de partage de la bande passante, elles dépendent étroitement du moment où les LSP de secours sont signalisés. Deux types de signalisation sont à distinguer : signalisation après la panne et signalisation avant la panne.

5.3.1.1 Signalisation après la panne

Avec la signalisation après la panne, les LSP de secours sont pré-calculés mais aucune étiquette n'est allouée avant l'apparition d'une panne. Ce n'est qu'au moment où la source multicast est informée de la panne que les étiquettes sont allouées aux chemins de secours afin de pallier la panne (les chemins de secours qui ne participent pas à la récupération ne seront pas signalisés). Bien évidemment, ce type de signalisation permet d'obtenir un degré de partage de la bande passante plus élevé. En effet, la signalisation après la panne permet d'appliquer *la fusion de flux* qui consiste à agréger les flux de plusieurs LSP sur les liens en commun et de n'en transmettre qu'un. Typiquement, si nous combinons la signalisation après la panne avec le partage restreint de la bande passante, nous déterminons la quantité de bande passante de secours $G^{\lambda}(k)$ allouée sur un arc λ pour pallier la panne de n'importe quel risque comme suit :

$$G^{\lambda}(k) = Max_{r \in R}(\sum_{i=1}^{k} Max(\delta m_r^{\lambda,i} + L_r^{\lambda,i} - F_{\lambda}^{i}, 0))$$

Si nous appliquons maintenant le partage global de la bande passante, nous obtenons la quantité de bande de secours $G^{\lambda}(k)$ comme suit :

$$G^{\lambda}(k) = Max_{r \in R}(\sum_{i=1}^{k} [Max(\delta m_r^{\lambda,i} - F_{\lambda}^{i}, 0) - L_r^{\lambda,i}], 0)$$

Afin de permettre aux nœuds sources multicast de calculer les LSP de secours, partageant la bande passante et vérifiant ses contraintes, les nœuds sources multicast devraient pouvoir déduire les valeurs des paramètres $\{\delta m_r^{\lambda,i}, F_{\lambda}^{i}, L_r^{\lambda,i}\}_{r,0<i<k}$. Cela peut être obtenu en annonçant ces paramètres dans le réseau ou en diffusant les structures et propriétés de tous les LSP (primaires et de secours). Bien évidemment ces solutions sont toutes les deux très coûteuses en termes de messages de contrôle circulant dans le réseau.

[4] En l'absence de la signalisation des LSP de secours avant la panne, le nœud source devra effectuer la réservation de ressources et d'étiquettes pour un LSP de secours avant de l'activer.

La combinaison de la protection par chemins disjoints et de la signalisation après la panne est surtout étudiée dans le cas de réseaux optiques où les chemins sont relativement courts. Un premier exemple formulant en ILP le problème de détermination des structures de secours[5] protégeant un ensemble d'arbres multicast avec des chemins disjoints est illustré dans [SOM06]. Un second exemple comparant deux heuristiques de calcul (en ligne) des arbres primaires et des chemins disjoints les protégeant peut être trouvé dans [SM03a].

Cette technique de signalisation des chemins de secours après la panne augmente significativement les délais de récupération, surtout dans les réseaux MPLS étendus. En conséquence, ce type de signalisation n'est pas désirable si les contraintes de temps des applications supportées sont dures. Le but de l'étude de ce type de signalisation, dans ce chapitre, est de situer le degré maximal de partage qui pourrait être atteint avec la protection par chemins disjoints. Nous ne nous intéresserons plus à ce type de signalisation dans la suite de ce chapitre.

5.3.1.2 Signalisation après la panne

Avec la signalisation après la panne, les délais de récupération sont plus petits mais le degré de partage de la bande passante est nettement diminué. Ainsi, la fusion de flux n'est pas possible avec la signalisation après la panne puisqu'une étiquette (sur une interface donnée) n'identifie qu'un seul LSP (primaire ou de secours). La quantité de bande passante de secours allouée sur un arc λ ne dépend que du nombre maximal de LSP (LSP P2MP primaires et LSP P2P de secours) qui seront activés sur l'arc λ suite à n'importe quelle panne simple.

Par conséquent, si l'on fait abstraction des possibilités de partage entre les structures de secours et les LSP P2MP primaires (le partage de la bande passante est restreint), nous déterminons la quantité de bande passante de secours devant être allouée sur chaque lien λ afin de pallier n'importe quel risque de panne r comme suit :

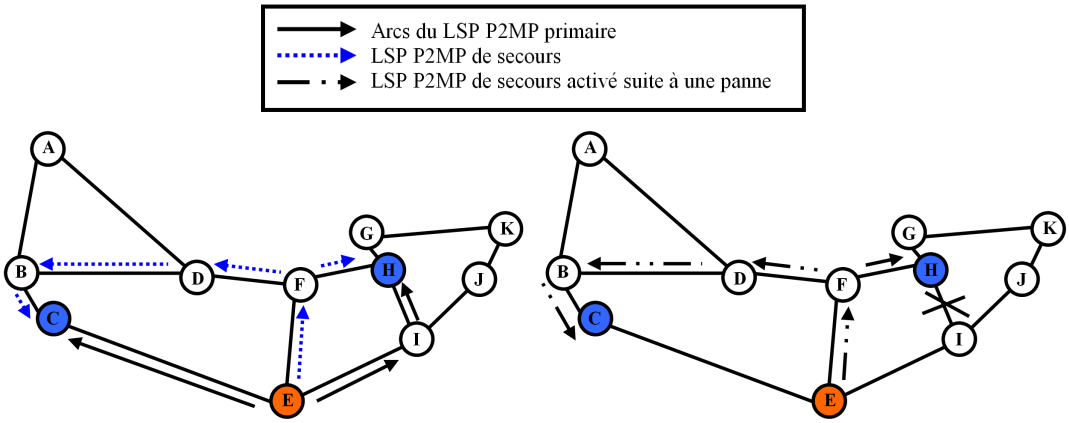
$$G^{\lambda}(k) = Max_{r \in R}(\sum_{i=1}^{k} \delta_r^{\lambda,i}) = Max_{r \in R}(\delta_r^{\lambda}(k))$$

Si nous appliquons maintenant le partage global de la bande passante, la quantité de bande passante de secours devant être allouée sur l'arc λ sera calculée comme suit :

$$G^{\lambda}(k) = Max_{r \in R}(\sum_{i=1}^{k} (\delta_r^{\lambda,i} - L_r^{\lambda,i}), 0) = Max_{r \in R}(\delta_r^{\lambda}(k) - L_r^{\lambda}(k), 0)$$

De la même manière que pour la signalisation après la panne, le calcul des LSP de secours protégeant une session multicast *k+1* par son nœud source requiert la transmission d'une information permettant la déduction des paramètres $\{\delta m_r^{\lambda,i}, F_{\lambda}^{i}, L_r^{\lambda,i}\}_{r,0<i<k}$. Cela peut être obtenu en annonçant ces paramètres dans le réseau ou en diffusant les structures et propriétés de tous les LSP (primaires et de secours). Comme cela a été

[5] Hors ligne et sans récupération de la bande passante primaire libérée.



(a) Situation avant l'occurrence d'une panne affectant le LSP P2MP primaire

(b) Situation après la détection et la récupération d'une panne touchant le LSP P2MP primaire

FIG. 5.3 – Protection par arbre redondant

signalé plus haut, ces solutions sont toutes les deux très coûteuses en termes de messages de contrôle circulant dans le réseau.

5.3.2 Protection par arbre redondant

Dans la méthode de protection par arbre redondant, le LSP P2MP primaire est protégé par un autre LSP P2MP de secours couvrant la source et tous les membres multicast et disjoint partout ailleurs du LSP P2MP primaire (hormis les nœuds source et membres multicast, aucun nœud, lien ou SRLG est commun aux arbres primaire et redondant).

Sur la figure 5.3 (a), un LSP P2MP primaire ($E \rightarrow C$, $E \rightarrow I$, $I \rightarrow H$) est établi afin de servir les deux membres multicast C et H à partir de la source multicast E. Afin de protéger ce LSP P2MP primaire, la source multicast calcule un LSP P2MP de secours ($E \rightarrow F$, $F \rightarrow H$, $F \rightarrow D$, $D \rightarrow B$, $B \rightarrow C$) couvrant les trois nœuds E (source multicast), C et H (membres multicast) et disjoint du LSP primaire partout ailleurs. Ainsi, hormis les membres et la source multicast, le LSP primaire et son LSP secours ne partagent aucun nœud, lien ou SRLG.

Lorsqu'une panne touchant un nœud non membre ou un lien du LSP primaire est détectée, un message de notification (de la panne) est envoyé à la source multicast. Pour pallier la panne, cette dernière se contentera alors de basculer le trafic du LSP P2MP primaire vers son LSP P2MP de secours. Par exemple, sur la figure 5.3 (b), le nœud source E bascule le trafic du LSP primaire vers le LSP de secours dès la réception d'un message notifiant la panne du lien I-H.

Comme nous le constatons, un LSP primaire ne peut pas partager la bande passante avec son LSP de secours vu que ces deux LSP ne partagent aucun lien (par définition). Par contre, il est possible de partager une grande quantité de bande entre les LSP de secours, et (surtout) entre les LSP de secours et les LSP primaires de sessions différentes.

Ainsi, si l'on applique la stratégie de partage restreint de la bande passante, nous déterminons la quantité minimale de bande passante de secours à allouer sur un arc λ pour faire face à n'importe quel risque de panne comme suit (indépendamment du type de signalisation utilisée) :

$$G^{\lambda}(k) = Max_{r \in R}(\sum_{i=1}^{k} \delta_r^{\lambda,i}) = Max_{r \in R}(\sum_{i=1}^{k} \delta m_r^{\lambda,i}) = Max_{r \in R}(\delta_r^{\lambda}(k))$$

Avec ce type de partage, une quantité élevée de la bande passante est gaspillée. En effet, lors d'une panne affectant un lien (resp. un nœud), une grande quantité de bande passante sera libérée sur les liens appartenant aux arbres primaires contenant le lien (resp. le nœud) en panne. Cette quantité de bande passante libérée n'est pas exploitée par les entités calculant les LSP P2MP de secours. Dans [SSM03], les auteurs montrent que ce type de partage a des performances médiocres et inférieures à celles obtenues avec la protection par chemins disjoints.

Si nous utilisons maintenant la stratégie de partage global de la bande passante, la disponibilité de la bande passante croit significativement. Ainsi, la quantité de bande passante minimale à allouer sur un arc λ pour faire face à n'importe quel risque de panne diminue et sera déterminée comme suit :

$$G^{\lambda}(k) = Max_{r \in R}(\sum_{i=1}^{k}(\delta_r^{\lambda,i} - L_r^{\lambda,i}), 0) = Max_{r \in R}(\delta_r^{\lambda}(k) - L_r^{\lambda}(k), 0)$$

Notons qu'avec la protection par chemins disjoints, le prix de protection $\delta_r^{\lambda}(i)$ d'un risque r sur un arc λ pour une session i peut être élevé puisqu'il peut atteindre la valeur $|M^i| \times F^i$). Avec la protection par arbre redondant, ce prix ne peut excéder la quantité F^i) puisque le flux de données routé sur un LSP P2MP de secours n'est pas dupliqué.

Comme pour la protection par chemins disjoints, le calcul de LSP P2MP de secours dans un environnement distribué en appliquant la protection par arbre redondant exige la connaissance, sur chaque arc, de la bande passante primaire cumulée, des prix de protection et des quantités de bande passante libérées suite à la panne d'un risque. Cela peut être obtenu en diffusant les valeurs de ces paramètres ou en annonçant les structures et propriétés de tous les LSP P2MP aux entités chargées du calcul des LSP de secours.

Bien que la protection par arbre redondant soit simple et malgré son taux élevé de partage (lorsque le partage de la bande entre les LSP primaires et les LSP de secours est adopté), elle n'est pas très utilisée puisqu'elle nécessite un degré de redondance de la topologie du réseau très élevé pour être applicable. Par exemple, il n'est pas possible de déterminer un couple de LSP P2MP disjoints (pas de lien en commun) et couvrant la source et les membres multicast de la figure 5.2.

5.3.3 Constats

Bien que le principe de récupération utilisé dans la protection de niveau global soit simple (récupération effectuée toujours par la source multicast), ce type de protection

souffre de plusieurs problèmes parmi lesquels nous citons : le faible taux de protection, le gaspillage de la bande passante, les délais de récupération élevés ainsi que la génération d'un nombre élevé de messages de contrôle lors d'une panne.

Faible taux de protection
La protection de niveau global nécessite en général une topologie de réseau ayant une forte redondance des liens et nœuds pour permettre une protection de tous les risques de panne. Par exemple, la protection par chemins disjoints peut être inefficace dans un réseau disposant de risques de types SRLG (i.e. la protection par chemins disjoints peut être incapable de protéger certains LSP primaires bien que la topologie du réseau le permette) et la protection par arbre redondant exige un taux de redondance de la topologie du réseau très élevé, ce qui restreint et limite son champ d'application.

Gaspillage de la bande passante
Une quantité élevée de bande est gaspillée avec la protection de niveau global lorsque la bande passante primaire libérée suite à une panne n'est pas exploitée par les entités de calcul des LSP de secours. De plus, avec les méthodes de protection de niveau global inspirées de l'unicast (comme la protection par chemins disjoints), une grande quantité de bande passante est dissipée sur les parties communes à plusieurs LSP de secours d'une même session à cause de la séparation des flux de ces LSP de secours. De la même manière, la condition de disjonction (dans la protection par redondant) du LSP P2MP primaire avec son LSP P2MP de secours conduit à une mauvaise exploitation de la bande passante.

Délais de récupération élevés et génération de messages de contrôle
La protection de niveau global exige la notification de la panne depuis les LSR ayant détecté la panne vers les sources multicast de tous les flux affectés, avant de basculer le trafic vers les LSP de secours. Cela augmente considérablement les délais de récupération et génère (après la panne) un nombre élevé de messages de contrôle dans les réseaux étendus (notamment dans les réseaux multi-domaines).

5.4 Protection proactive locale

En raison de sa rapidité de restauration, sa facilité de déploiement et son taux élevé de protection, la protection proactive locale est souvent préférée à la protection proactive globale. Avec la protection proactive locale, le trafic de tout LSP primaire affecté par une panne est basculé vers les chemins de secours rapidement et localement par le nœud en amont et le plus proche du composant défaillant. Cela réduit considérablement le nombre de messages de notification de panne envoyés dans le réseau puisqu'il n'est plus nécessaire d'envoyer un message de notification pour chaque LSP primaire et/ou source multicast et il est possible de ne notifier que les nœuds voisins et/ou très proches du composant en panne.

Comme pour la protection de niveau globale, il existe diverses méthodes locales permettant de protéger les communications multicast. Quelques unes de ces méthodes de protection locales subdivisent le LSP P2MP primaire en segments de petites tailles,

d'autres le divisent en sous-arbres avant d'appliquer la protection sur les parties du LSP primaire obtenues.

Dans cette section, nous nous concentrons sur l'étude de méthodes de protection proactive locale assurant une récupération rapide des pannes. De ce fait (à l'exception de la méthode de protection par forêt duale), nous ne présentons pas ici les méthodes de protection multicast exigeant une signalisation après la panne [MFB99, SM03b] ou subdivisant le LSP P2MP primaire en segments de tailles variables et non bornées [SM03a]. En effet, ces méthodes ne permettent pas de borner les délais de récupération. A l'inverse, les méthodes de protection calculant des LSP de secours dont les LSR de tête sont adjacents aux composants protégés [APY07, LRAVV07], permettent une redirection rapide du trafic vers les LSP de secours lors d'une panne, ce qui diminue considérablement et limite les délais de récupération.

Quatre principales méthodes de protection proactive locale sont décrites ci-après : la protection par forêt duale, la protection multicast un-à-un (*one-to-one protection*), la protection par tunnels P2P de secours et la protection par tunnels P2MP de secours. La première méthode de protection a l'avantage de réduire la taille de la structure de secours au prix d'une gestion plus complexe du mécanisme de récupération alors que les trois dernières méthodes réduisent significativement les délais de récupération au détriment d'une augmentation du nombre de messages nécessaires à la configuration des structures de secours. Nous notons que la protection multicast un-à-un utilise les mêmes principes et mécanismes que ceux adoptés par la protection unicast un-à-un, alors que les deux dernières méthodes de protection multicast décrites ici emploient un procédé similaire à celui qui est appliqué par la protection unicast par tunnels de secours (*facility backup*).

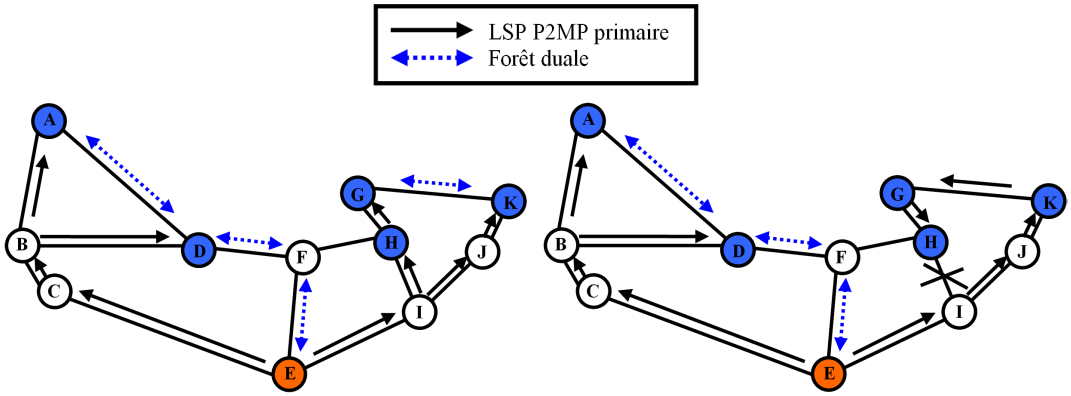
5.4.1 Protection par forêt duale

Pour améliorer la tolérance aux pannes de la méthode de protection par arbre redondant, nous avons proposé, dans [SCM06b, SCM06a], d'utiliser une structure de forêt ne contenant aucun nœud ou lien internes du LSP P2MP primaire pour le protéger. Dans notre approche, nous relions les nœuds feuilles du LSP P2MP primaire et la source en utilisant des arbres de Kou-Markowsky-Berman (arbres KMB) [KMB81].

Sur la figure 5.4 (a), un LSP primaire ($E \rightarrow I$, $I \rightarrow J$, $J \rightarrow K$, $I \rightarrow H$, $H \rightarrow G$, $E \rightarrow C$, $C \rightarrow B$, $B \rightarrow A$, $B \rightarrow D$) est protégé par une forêt composée de deux arbres KMB bidirectionnels (G-K) et (E-F, F-D, D-A).

Lors de l'occurrence d'une panne, le nœud en aval X du composant défaillant cherche à déterminer un chemin dans la forêt duale permettant d'atteindre un nœud X_1 du sous-arbre enraciné à X (i.e. sous-arbre affecté) à partir d'un autre nœud X_2 n'appartenant pas au sous-arbre enraciné à X (i.e. nœud d'un sous-arbre non affecté). Ensuite, le nœud X envoie au nœud X_2 un message de signalisation permettant d'ajouter le chemin $X_2 \rightarrow .. \rightarrow X_1 \rightarrow .. \rightarrow X$ au LSP P2MP primaire et de supprimer de ce LSP la partie du chemin allant du du nœud X au nœud X_1 (i.e. le chemin $X \rightarrow .. \rightarrow X_1$).

Lors de la détection de la panne du lien I-H sur la figure 5.4 (b), le nœud H initie le processus de récupération afin de rétablir la communication multicast affectée. Ainsi,



(a) Situation avant l'occurrence d'une panne affectant le LSP P2MP primaire

(b) Situation après la récupération de la panne du lien I-H

FIG. 5.4 – Protection par forêt duale

le nœud H déduit que seul le chemin $K \rightarrow G$ de la forêt duale permet de relier le sous-arbre primaire affecté (i.e. $H \rightarrow G$) à un nœud du LSP primaire non affecté. De ce fait, il envoie au nœud K un message de signalisation pour ajouter le chemin $K \rightarrow G \rightarrow H$ et supprimer le chemin $H \rightarrow G$ du LSP primaire affecté.

Bien que la protection par forêt duale fournisse une meilleure tolérance aux pannes par rapport à la protection par arbre redondant, son taux de protection reste inférieur à l'idéal (i.e. bien que la topologie du réseau le permette, certains risques ne peuvent être protégés en utilisant une forêt duale). Concrètement, sur la figure 5.4 (a), la protection par forêt duale ne permet pas de protéger contre la panne du lien E-I bien que la topologie du réseau le permette. Pour augmenter le taux de protection, nous avons amélioré la protection par forêt duale dans [SCM06c] en ajoutant à la forêt duale des routes augmentant la tolérance aux panne. Par exemple, l'ajout du lien F-H à la forêt duale de la figure 5.4 (a) permet de fournir une protection de tous les liens et nœuds du LSP primaire (excepté les nœuds membres et la source multicast).

En plus de la nécessité de connaître la structure du LSP primaire, la protection par forêt duale présente deux inconvénients majeurs qui limitent son utilisation :

1. Augmentation des délais de récupération à cause de la nécessité de signalisation du(es) chemin(s) de secours après l'occurrence d'une panne.
2. Augmentation du nombre de messages de contrôle envoyés dans le réseau puisqu'il est nécessaire d'envoyer un ou plusieurs messages de signalisation par LSP P2MP primaire affecté.

A l'opposé, pour protéger un petit nombre de communications desservant des groupes multicast denses, la protection par forêt duale peut s'avérer très performante. En effet, dans un arbre multicast desservant un groupe dense, le nombre de nœuds feuilles est souvent élevé, ce qui diminue la distance entre ces nœuds feuilles et offre plus de flexibilité pour les relier en utilisant une forêt duale.

5.4.2 Protection multicast un-à-un

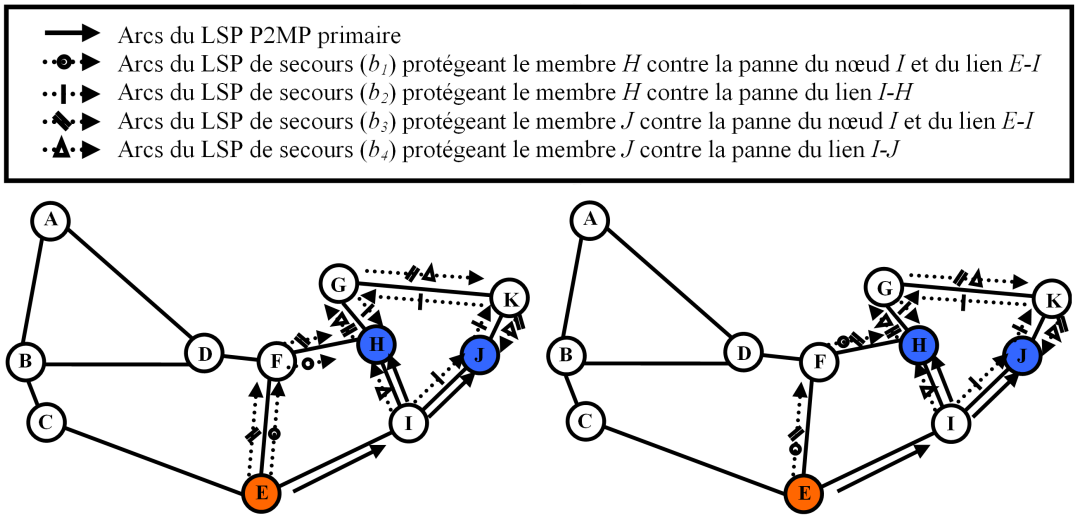
Avec la protection multicast un-à-un, le LSP P2MP primaire est décomposé en l'ensemble de ses LSP S2L qui relient le nœud source aux différents nœuds membres multicast. Chaque LSP S2L est protégé par un ensemble de LSP P2P de secours (détours) calculés selon la méthode de protection un-à-un ([PSA05, APY07]).

Pour baisser le trafic de contrôle, diminuer le nombre d'états de signalisation (états RSVP) et éventuellement réduire les réservations de la bande passante[6], la protection multicast un-à-un autorise *la fusion de LSP* définie initialement dans [PSA05] pour la protection unicast. Cette opération de fusion consiste à agréger des LSP P2P d'une même session multicast entre un nœud commun (dit nœud de fusion) et le nœud de destination des LSP agrégés (les LSP agrégés doivent avoir le même nœud de destination). De cette manière, une seule étiquette est allouée à l'ensemble des LSP agrégés entre le nœud de fusion et leur nœud de destination (un seul état RSVP est géré entre le nœud de fusion et le nœud de destination pour l'ensemble des LSP agrégés).

Pour mieux éclaircir ce concept de la fusion de LSP, nous considérons la session multicast illustrée sur la figure 5.5. Cette session a comme source le nœud E et comme membres multicast les nœuds H et J. Elle utilise le LSP P2MP primaire composé de deux LSP S2L ($E \rightarrow I \rightarrow H$ et $E \rightarrow I \rightarrow J$) pour router le trafic en l'absence de pannes.

Si l'on applique la protection unicast un-à-un pour protéger les LSP S2L primaires les uns indépendamment des autres (cf. figure 5.5 (a)), nous obtenons quatre LSP de secours (P2P) : $b_1 = E \rightarrow F \rightarrow H$, $b_2 = I \rightarrow J \rightarrow K \rightarrow G \rightarrow H$, $b_3 = E \rightarrow F \rightarrow H \rightarrow G \rightarrow K \rightarrow J$

[6] la réduction des allocations de la bande passante n'est possible avec la fusion de LSP que si le partage de la bande passante est restreint.



(a) Protection unicast un-à-un des S2L du LSP P2MP primaire indépendamment les uns des autres

(b) Protection multicast un-à-un du LSP P2MP primaire optimisant la bande passante de secours

FIG. 5.5 – Protection un-à-un unicast et multicast

et $b_4 = I \rightarrow H \rightarrow G \rightarrow K \rightarrow J$. Hormis les LSP de secours b_1 et b_3 qui protègent contre les mêmes risques de panne, tous les autres LSP sont indépendants et peuvent donc partager les ressources sur leurs arcs en commun (i.e. les arcs $H \rightarrow G$, $G \rightarrow K$ et $K \rightarrow J$). Par ailleurs, tous LSP de secours (P2P) protégeant un même LSP primaire S2L peuvent être fusionnés[7] sur un nœud donné si leurs routes respectives à partir de ce nœud jusqu'à la destination sont les mêmes. Ainsi, sur la figure 5.5 (a), les LSP de secours b_3 et b_4 sont fusionnés sur leur partie commune du chemin ($H \rightarrow G \rightarrow K \rightarrow J$) entre les nœuds H et J (un seul état RSVP est géré par les nœuds G, K et J pour les LSP de secours b_3 et b_4). Une seule étiquette est alors allouée pour les deux LSP de secours b_3 et b_4 sur chacun des liens qui leur sont en commun (i.e. liens $H \rightarrow G$, $G \rightarrow K$ et $K \rightarrow J$).

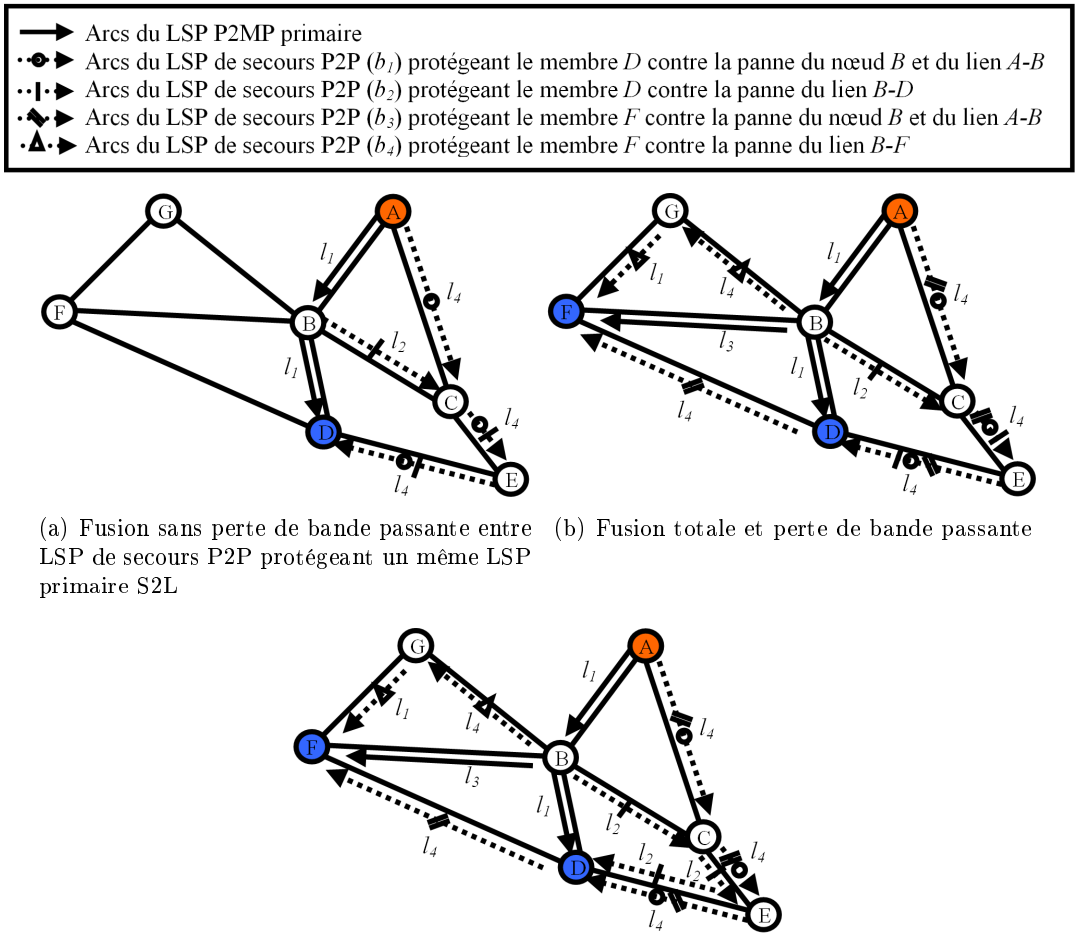
Bien que cette méthode de protection unicast un-à-un permette de réduire le nombre d'états RSVP et de sauver une certaine quantité de bande passante, elle est sous-optimale (en termes de partage et de nombre d'états RSVP) car elle restreint la fusion aux LSP de secours protégeant un même LSP S2L primaire (i.e. même destination) et elle limite le partage aux LSP de secours protégeant contre des ensembles de risques de panne disjoints. Pour améliorer le partage de ressources (particulièrement la bande passante et les étiquettes MPLS), la protection multicast un-à-un [APY07] suggère d'agréger en plus les flux des LSP P2P de secours qui protègent un même LSP P2MP primaire contre les mêmes ensembles de risques de panne[8]. Cela permet d'obtenir un seul flux, routé sur un même LSP P2MP de secours. Ce type d'agrégation sera dénommé par *la fusion de flux* dans la suite de cette thèse.

Sur la figure 5.5 (a), les deux LSP P2P de secours b_1 et b_3 protègent le seul LSP primaire P2MP établi contre le même ensemble de risques de panne ($\{E\text{-}I, I\}$) ; ces deux LSP de secours sont alors combinés, avec la protection multicast un-à-un, dans un seul LSP P2MP de secours (figure 5.5 (b)). Cela a pour conséquences de réduire le nombre d'étiquettes allouées pour les deux LSP b_1 et b_3 sur les nœuds E et F de 50% et de diminuer la quantité de bande passante allouée sur les arcs $E \rightarrow F$ et $F \rightarrow H$ de la moitié.

Notons que si la fusion de LSP P2P d'une même session multicast permet toujours de réduire le nombre d'étiquettes (elle réduit aussi le nombre d'états RSVP), elle conduit parfois à une surconsommation de la bande passante. Considérons le scénario suivant : une session multicast protégée s est établie entre un nœud source A et un membre multicast D (figure 5.6 (a)). A ce moment, la session s utilise le LSP primaire composé d'un seul LSP S2L ($A \rightarrow B \rightarrow D$) et de deux LSP de secours P2P ($b_1 = A \rightarrow C \rightarrow E \rightarrow D$ et $b_2 = B \rightarrow C \rightarrow E \rightarrow D$) fusionnés sur la partie du chemin ($C \rightarrow E \rightarrow D$) allant du nœud C jusqu'au nœud D. Bien évidemment, la fusion des LSP de secours b_1 et b_2 permet ici de ne maintenir qu'un seul état RSVP sur les nœuds D et E. Par conséquent, une seule

[7] [APY07] recommande vivement l'utilisation de l'identification par chemin pour configurer les LSP de détour. Cette méthode rend obligatoire la fusion de LSP P2P de secours protégeant un même LSP S2L primaire si ces LSP de secours partagent le même prochain saut (et même interface).

[8] Dans [APY07], la fusion de LSP de secours P2P d'une même session est effectuée si les messages *path* correspondant à ces LSP contiennent le même objet DETOUR (i.e. les LSP de secours protègent contre la panne d'un même lien). Cela pose le problème de fusion de LSP de secours de types différents (LSP NNHOP avec LSP NHOP).



(a) Fusion sans perte de bande passante entre LSP de secours P2P protégeant un même LSP primaire S2L

(b) Fusion totale et perte de bande passante

(c) Fusion partielle et sans perte de bande passante entre LSP de secours P2P protégeant des LSP primaires S2L différents

FIG. 5.6 – Fusion de LSP de secours et perte de la bande passante

réservation de bande passante et d'étiquettes MPLS est nécessaire et est effectuée sur la partie commune ($C \rightarrow E \rightarrow D$) aux deux LSP b_1 et b_2. Lorsqu'un deuxième membre F rejoint le groupe multicast correspondant à la session s, un second LSP S2L primaire et un autre ensemble de LSP P2P de secours le protégeant seront calculés et configurés afin d'alimenter et protéger le nouveau membre F. Si nous supposons que le LSP S2L primaire alimentant le membre F est supporté par le chemin $A \rightarrow B \rightarrow F$ et que les deux LSP de secours b_3 et b_4 le protégeant sont supportés par les chemins $A \rightarrow C \rightarrow E \rightarrow D \rightarrow F$ et $B \rightarrow G \rightarrow F$, alors deux possibilités nous sont offertes pour configurer le LSP de secours b_3 : configuration avec *une fusion totale* des LSP de secours ou configuration avec *une fusion partielle* des LSP de secours (i.e. avec reconfiguration de certains LSP de secours).

5.4.2.1 Fusion totale

Avec la fusion totale, tous les LSP de secours vérifiant les conditions nécessaires à la fusion de LSP ou à la fusion de flux sont combinés et agrégés. Sur la figure 5.6 (b), les deux LSP de secours b_1 et b_2, établis et fusionnés précédemment sur la partie commune de leurs chemins $C \rightarrow E \rightarrow D$, sont maintenus (en fusion) puisque ces deux LSP de secours vérifient les conditions de fusion de LSP (même nœud de destination et protection d'une même session multicast). De même, les deux LSP b_1 et b_3 vérifient les conditions de fusion de flux (protection de LSP S2L primaires d'une même session contre des ensembles de risques identiques) et sont donc combinés dans un seul LSP de secours P2MP. En conséquence, une seule étiquette MPLS sera allouée sur tout arc traversé par les trois LSP de secours b_1, b_2 et b_3 afin de router leur trafic en cas de panne. Cela permet de réduire le nombre d'étiquettes MPLS allouées mais risque de provoquer une perte de bande passante. Par exemple, lors de la panne du lien B-D de la figure 5.6 (b), le nœud B bascule les paquets émis sur l'arc $B \rightarrow D$ vers l'arc $B \rightarrow C$. Ces paquets transitent alors par les nœuds C et E avant d'atteindre le nœud membre D, ce qui permet d'accomplir le processus de récupération. Cependant et au lieu de détruire les paquets reçu du nœud E après qu'il les ait consommés, le nœud D les redirige inutilement vers le nœud F : on parle alors de *fuite de la bande passante*. Cette fuite de la bande passante est inévitable dans l'exemple de la figure 5.6 (b) car le nœud D ne dispose d'aucune information lui permettant de distinguer entre la panne du nœud B ou du lien A-B (où cette redirection est nécessaire afin de pallier la panne) et la panne du lien B-D (où la redirection est inutile et provoque une fuite de la bande passante). En effet, l'étiquette MPLS l_4 partagée par les trois LSP de secours (b_1, b_2 et b_3) sur les arcs $C \rightarrow E$ et $E \rightarrow D$ ne permet pas de séparer les flux de ces trois LSP de secours. Notons que la fusion totale ne peut pas être évitée[9] si les LSP de secours adoptent la méthode d'identification par chemin puisque cette dernière impose la fusion de LSP et que [APY07] rend obligatoire la fusion de flux lors de la configuration des LSP de secours.

5.4.2.2 Fusion partielle

Avec la fusion partielle, les LSP de secours sont configurés de manière à éviter les fuites de la bande passante. Pour ce faire, au plus un type de fusion[10] (fusion de LSP ou fusion de flux) est appliqué à tout LSP de secours sur chaque arc. Lorsque les deux types de fusion sont possibles, la priorité est donnée à la fusion de flux (obligatoire dans [APY07]) puisque cette dernière permet de diminuer la consommation de la bande passante en évitant sa fuite. Dans l'exemple de la figure 5.6 (c), le LSP de secours b_1 peut être fusionné avec le LSP de secours b_2 en utilisant la fusion de LSP et il peut aussi être combiné avec le LSP de secours b_3 en appliquant la fusion de flux. Afin d'éviter la

[9] Nous pouvons par contre prévenir la perte de la bande passante lorsque la fusion totale est appliquée. Cela se fait en choisissant les LSP de secours de telle sorte qu'au plus, un seul type de fusion peut être appliqué à tout LSP de secours. Cela a l'inconvénient de diminuer le taux de protection et de limiter les possibilités de partage (augmentation de l'utilisation des ressources).

[10] La fusion de flux des LSP P2P de secours protégeant deux LSP S2L d'une même session où l'un des deux LSP S2L est contenu dans l'autre n'est pas considéré.

fuite de la bande passante, le nœud D (resp. E et C) annule la fusion de LSP appliquée jusque-là aux LSP de secours b_1 et b_2 (en envoyant un message RSVP *path* pour chacun de ces deux LSP) pour permettre la fusion de flux entre les LSP de secours b_1 et b_3 sur l'arc $E \rightarrow D$ (resp. $C \rightarrow E$, $A \rightarrow C$). En conséquence, une seule étiquette MPLS sera allouée pour les deux LSP de secours b_1 et b_3 sur chaque arc commun à ces LSP et une autre étiquette MPLS sera allouée sur chacun des arcs formant le LSP b_2. De cette manière, le nœud D saura distinguer le trafic des LSP de secours b_1 et b_3 activés suite à la panne du lien A-B ou du nœud B de celui correspondant au LSP de secours b_2 activé suite à une panne affectant le lien B-D. Ainsi, le nœud D ne redirigera au nœud F que les paquets des LSP b_1 et b_3 reconnus grâce à l'étiquette l_4 transmise dans ces paquets (les paquets du LSP de secours b_2 atteignent le nœud F avec l'étiquette l_2).

5.4.2.3 Types de fusion, ressources et passage à l'échelle

Les deux techniques de fusion (fusion totale et fusion partielle) offrent des mécanismes permettant le calcul de la bande passante nécessaire sur un arc donné afin d'éviter la perte de la bande passante (fusion partielle) ou afin de diminuer le nombre d'étiquettes MPLS allouées aux LSP de secours (fusion totale). Si la fusion partielle est indispensable au passage à l'échelle, la fusion totale ne l'est pas. En effet, la fusion totale conduit à une saturation rapide des capacités des arcs (manque de la bande passante sur les arcs), ce qui décroît le nombre de LSP pouvant être établis dans le réseau. Par ailleurs, le nombre d'étiquettes MPLS allouées peut être diminué significativement, sans perte de la bande passante, en utilisant d'autres techniques d'agrégation comme le tunneling (basé sur la hiérarchie MPLS). De ce fait, seule la fusion partielle sera considérée, dans la suite de cette section, pour le calcul des LSP de secours.

5.4.2.4 Calcul des LSP P2MP de secours optimisant la bande passante

Dans un environnement MPLS distribué où les requêtes d'établissement de sessions multicast protégées arrivent en ligne, le calcul du LSP P2MP primaire et de ses LSP secours en adoptant l'approche successive (*non-joint approach*) est généralement préférée pour diminuer les délais de réponse des requêtes. Avec cette approche, le LSP P2MP primaire est d'abord calculé (puis configuré) avant que les différents nœuds chargés de sa protection ne se lancent dans les calculs des LSP P2MP de secours qui le protègent. Cette approche successive a plusieurs avantages : elle facilite les calculs, répartit la charge de calcul des LSP de secours sur plusieurs nœuds et utilise les meilleurs chemins pour le routage du trafic en l'absence de pannes.

Pour calculer les LSP de secours, deux grandes approches peuvent être adoptées : approche intégrante et approche pas-à-pas. Dans l'approche intégrante, tous les LSP de secours sont calculés conjointement de manière à protéger tous les LSP primaires déjà établis et de telle sorte que la quantité totale de la bande passante allouée à ces LSP de secours est minimisée. Ce type d'approche ne convient pas pour protéger en ligne les LSP primaire ; il est plutôt destiné à ré-optimiser la quantité de la bande allouée à des LSP de secours déjà configurés. Dans la seconde approche, à chaque création d'un nouveau

LSP primaire, un ensemble de LSP de secours le protégeant est calculé et est configuré. Il est à noter que cet ensemble de LSP de secours doit vérifier les contraintes de bande passante et souvent, il est calculé de manière à minimiser la quantité additionnelle de bande passante de secours allouée ou de telle sorte à minimiser le délai. Cette seconde approche est généralement préférée à la première approche pour permettre un calcul et une protection rapides des LSP primaires établis en ligne.

Indépendamment du type d'approche utilisée pour le calcul (approche pas-à-pas ou approche intégrante), nous annonçons que le problème de recherche des LSP de secours optimisant la quantité (totale ou additionnelle) de la bande passante de secours est NP-complet. En effet, le sous-problème consistant à déterminer les LSP de secours protégeant un seul LSP primaire S2L est NP-complet (cf. sections 2.5.3.1 et 2.5.4.1). Pour modéliser et résoudre ce problème, nous pouvons utiliser la programmation linéaire en nombres entiers (ILP) ou les heuristiques.

Formulation ILP et optimisation de la bande passante de secours

Dans un premier temps, nous modélisons en ILP le problème de recherche des LSP de secours protégeant un ensemble de LSP P2MP primaires et minimisant la quantité de bande passante de secours allouée. Dans notre modélisation, nous adoptons l'approche intégrante (car c'est une modélisation plus générale et facile à étendre à l'approche pas-à-pas) et nous ne tenons compte (pour simplifier le discours et sans perte de généralité) que des pannes simples de liens.

En plus des paramètres de la bande passante définis dans de la section 5.2.3, nous ajoutons et définissons les variables et constantes ci-après :

- k : nombre de LSP P2MP à protéger.
- PLR_m^i : ensemble des nœuds ancêtres du nœud membre multicast m dans le LSP P2MP primaire P_i.
- $P_{p,m}^{u \rightarrow v,\, i}$: égal à *1* si l'arc $u \rightarrow v$ est sur le chemin reliant le PLR $p \in PLR_m^i$ au nœud membre multicast $m \in M^i$. Sinon $P_{p,m}^{u \rightarrow v,\, i} = 0$.
- $P_p^{u \rightarrow v,\, i}$: égal à *1* si l'arc $u \rightarrow v$ est sur un chemin du LSP P2MP primaire P_i reliant le nœud p à un nœud feuille. Autrement, $P_p^{u \rightarrow v,\, i} = 0$.
- $B_{p,m}^{u \rightarrow v,\, i}$: égal à *1* si l'arc $u \rightarrow v$ appartient au chemin de secours qui protège le LSP S2L primaire, alimentant le membre m ($m \in M^i$), contre la panne du lien situé en aval du PLR p ($p \in PLR_m^i$). Autrement, $B_{p,m}^{u \rightarrow v,\, i} = 0$.
- $B_{p\text{-}q}^{u \rightarrow v,\, i}$: égal à *1* s'il existe un LSP P2MP de secours passant par l'arc $u \rightarrow v$ et protégeant le LSP P2MP primaire P_i contre la panne du lien p-q (avec $p \rightarrow q$ appartenant à P_i). Autrement, $B_{p\text{-}q}^{u \rightarrow v,\, i} = 0$.

  Nous notons que les informations fournies par ces variables permettent de construire tous les LSP P2MP de secours optimisant la bande passante de secours et protégeant tous les LSP P2MP primaires.

Le problème d'optimisation de la bande passante des LSP de secours protégeant un ensemble de k LSP P2MP primaires déjà établis peut être modélisé comme suit, lorsque le partage de la bande passante est restreint :

$$\text{Min} : \sum_{uv} G^{u\rightarrow v}$$

$$\begin{cases}
\sum_u B_{p,m}^{p\rightarrow u,i} = 1 \ \forall i \in [1,k], \forall m \in M^i, \forall p \in PLR_m^i \\
\sum_u B_{p,m}^{u\rightarrow m,i} = 1 \ \forall i \in [1,k], \forall m \in M^i, \forall p \in PLR_m^i \\
\sum_v B_{p,m}^{u\rightarrow v,i} - \sum_v B_{p,m}^{v\rightarrow u,i} = 0 \ \forall i \in [1,k], \forall m \in M^i, \forall p \in PLR_m^i, \forall u \in V \backslash \{p,m\} \\
\sum_v B_{p,m}^{u\rightarrow v,i} \leq 1 \ \forall i \in [1,k], \forall m \in M^i, \forall p \in PLR_m^i, \forall u \in V \\
B_{p,m}^{p\rightarrow u,i} \leq 1 - P_{p,m}^{p\rightarrow u,i} \ \forall i \in [1,k], \forall m \in M^i, \forall p \in PLR_m^i, \forall u \in V \\
B_{p,m}^{u\rightarrow v,i} - B_{p\text{-}q}^{p\rightarrow u,i} \leq 1 + P_{p,m}^{u\rightarrow v,i} - P_{p,m}^{p\rightarrow q,i} \ \forall i \in [1,k], \forall m \in M^i, \forall p \in PLR_m^i, \forall u, \forall v \\
\sum_{j=1}^{k} F^j . B_{p\text{-}q}^{u\rightarrow v,j} - G^{u\rightarrow v} \leq 0 \ \forall i \in [1,k], \forall m \in M^i, \forall p \in PLR_m^i, \forall (u,v,q) \in V^3 \\
G^{u\rightarrow v} \leq C^{u\rightarrow v} - F_{u\rightarrow v}(k) \ \forall i \in [1,k], \forall m \in M^i, \forall p \in PLR_m^i, \forall (u,v) \in V^2 \\
(B_{p,m}^{u\rightarrow v,i}, B_{p\text{-}q}^{u\rightarrow v,i}) \in \{0,1\}^2, G^{u\rightarrow v} \in [0, C^{u\rightarrow v} - F_{u\rightarrow v}(k)] \ \forall i \in [1,k], \forall m \in M^i, \\
\qquad\qquad \forall p \in PLR_m^i, \forall (u,v,p,q) \in V^4
\end{cases} \tag{5.1}$$

Explication des équations.

La fonction d'objectif consiste à minimiser la bande passante de secours allouée sur tous les arcs de la topologie du réseau. Nous notons que $G^{u\rightarrow v}$ correspond à la quantité minimale de bande passante de secours qui devra être allouée sur l'arc $u\rightarrow v$ pour tous les LSP de secours protégeant les k LSP P2MP primaires établis ($G^{u\rightarrow v} = G^{u\rightarrow v}(k)$).

Concernant les contraintes, elles sont exprimées à travers neuf équations. Typiquement, les quatre premières contraintes correspondent aux équations de conservation de flux des LSP de secours. Ainsi, pour chaque LSP S2L du LSP P2MP primaire, un ensemble de LSP de secours, protégeant tous ses liens, est déterminé. La cinquième contraintes garantit que tout chemin de secours protégeant contre la panne d'un lien ne passe pas par ce dernier. La sixième contrainte détermine les arcs appartenant à la structure de secours qui est elle-même composée d'un ensemble de LSP P2MP de secours (chaque LSP P2MP de secours protégeant une session multicast i contre la panne du lien $p\rightarrow q$ est identifié par l'ensemble des variables $\{B_{p\text{-}q}^{u\rightarrow v,i}\}_{u-v\in E}$). Ainsi, la fusion de flux est appliquée entre tous les LSP P2P de secours protégeant une même session multicast contre le même risque de panne (lien). Cela permet de réduire les allocations de la bande passante. De plus, comme la fusion de LSP n'intervient dans la réduction de la bande passante de secours allouée dans le réseau que lorsqu'un des LSP fusionnés est primaire, nous ne l'avons appliquée qu'entre les LSP P2MP de secours et le LSP P2MP primaire (un LSP P2MP de secours est fusionné avec un LSP P2MP primaire si et seulement si toutes leurs routes vers les destinations sont les mêmes). La septième contrainte définit la quantité minimale de bande passante de secours $G^{u\rightarrow v}$ à allouer sur chaque arc $u\rightarrow v$ lorsque le partage de la bande passante est global (i.e. $G^{u\rightarrow v}$

$= Max_{p-q \in E} \sum_{i=1}^{k} F^i.B_{p\text{-}q}^{u\to v,i} = Max_{p-q \in E}\, \delta_{p\text{-}q}^{u\to v}(k)$). La huitième contrainte garantit le respect des contraintes de la bande passante et la neuvième contrainte définit les domaines de définition des variables ILP.

Lorsque le partage de la bande passante est global, nous obtenons le système suivant :

$$\text{Min} : \sum_{uv} G^{u\to v}$$

$$\begin{cases} \sum_u B_{p,m}^{p\to u,i} = 1 \ \forall i \in [1,k], \forall m \in M^i, \forall p \in PLR_m^i \\ \sum_u B_{p,m}^{u\to m,i} = 1 \ \forall i \in [1,k], \forall m \in M^i, \forall p \in PLR_m^i \\ \sum_v B_{p,m}^{u\to v,i} - \sum_v B_{p,m}^{v\to u,i} = 0 \ \forall i \in [1,k], \forall m \in M^i, \forall p \in PLR_m^i, \forall u \in V \backslash \{p,m\} \\ \sum_v B_{p,m}^{u\to v,i} \leq 1 \ \forall i \in [1,k], \forall m \in M^i, \forall p \in PLR_m^i, \forall u \in V \\ B_{p,m}^{p\to u,i} \leq 1 - P_{p,m}^{p\to u,i} \ \forall i \in [1,k], \forall m \in M^i, \forall p \in PLR_m^i, \forall u \in V \\ B_{p,m}^{u\to v,i} - B_{p\text{-}q}^{p\to u,i} \leq 1 - P_{p,m}^{p\to q,i} \ \forall i \in [1,k], \forall m \in M^i, \forall p \in PLR_m^i, \forall u, \forall v \\ \sum_{j=1}^{k} F^j (B_{p\text{-}q}^{u\to v,j} - P_q^{u\to u,j}) - G^{u\to v} \leq 0 \ \forall i \in [1,k], \forall m \in M^i, \forall p \in PLR_m^i, \forall (u,v,q) \in V^3 \\ G^{u\to v} \leq C^{u\to v} - F_{u\to v}(k) \ \forall i \in [1,k], \forall m \in M^i, \forall p \in PLR_m^i, \forall (u,v) \in V^2 \\ (B_{p,m}^{u\to v,i},\ B_{p\text{-}q}^{u\to v,i}) \in \{0,1\}^2,\ G^{u\to v} \in [0, C^{u\to v} - F_{u\to v}(k)] \ \forall i \in [1,k], \forall m \in M^i, \\ \qquad\qquad \forall p \in PLR_m^i, \forall (u,v,p,q) \in V^4 \end{cases} \tag{5.2}$$

Explication des équations.

La fonction d'objectif ainsi que les contraintes (1), (2), (3), (4), (5), (8) et (9) du système (5.2) sont les mêmes que celles qui leur correspondent dans le système (5.1) ; seules les contraintes (6) et (7) ont été modifiées afin de tenir compte de la bande passante primaire libérée sur les arcs suite à une panne (chaque LSP P2MP de secours protégeant une session multicast i contre la panne du lien $p{\to}q$ est identifié par l'ensemble des variables $\{B_{p\text{-}q}^{u\to v,i}\}_{u-v\in E}$). Pour minimiser la quantité de bande passante de secours allouée sur les arcs du réseau, nous n'appliquons ici que la fusion de flux (la fusion de LSP ne peut être appliquée qu'après la détermination des LSP P2MP de secours afin de réduire le nombre d'état RSVP). En effet, la quantité de bande passante qui serait récupérée sur un arc suite à la fusion d'un ensemble de LSP P2MP de secours avec leur LSP primaire est entièrement exploitée et récupérée avec PGB lorsque ces LSP ne sont pas fusionnés. La septième contrainte définit la quantité minimale de bande passante de secours $G^{u\to v}$ à allouer sur chaque arc $u{\to}v$ lorsque le partage de la bande passante est global (i.e. $G^{u\to v} = Max_{p-q\in E} \sum_{i=1}^{k} F^i(B_{p\text{-}q}^{u\to v,i} - P_q^{u\to u,i}) = Max_{p-q\in E}\,(\delta*)_{p\text{-}q}^{u\to v}(k)$).

Bien que les deux systèmes ILP (5.1) et (5.2) permettent de minimiser la consommation de la bande passante de secours (avec l'adoption de l'approche successive), leur utilisation n'est préconisée que pour le mode hors ligne. Ce mode permet de ré-optimiser

les LSP de secours par la recherche et la reconfiguration de nouveaux chemins de secours. Pour un calcul en ligne des LSP de secours (évitant les reconfigurations de LSP), les deux systèmes précédents pourraient être transformés. Ainsi, à chaque étape $i_{i>0}$, les valeurs des variables sont remplacées par celles obtenues à l'étape *i-1*.

Malgré l'optimalité des solutions obtenues avec la programmation ILP, cette dernière ne peut être utilisée et n'est pas appropriée au calcul en ligne des LSP de secours. En effet, la programmation ILP ne garantit pas la détermination d'une solution (optimale ou approchée) en un temps polynomial, en fonction de la taille du problème. Typiquement, aucun système ILP n'est sûr de trouver une solution optimale pour le problème de recherche des LSP de secours minimisant la bande passante de secours en un temps raisonnable (sur toute topologie du réseau) vu que ce problème est NP-complet. De plus, l'objectif consistant à minimiser la bande passante allouée aux LSP de secours ne garantit aucunement la minimisation de la probabilité de blocage (probabilité qu'un LSP soit rejeté).

Heuristiques de recherche des LSP de secours minimisant la bande passante de secours

Afin d'approcher les solutions optimales du problème de recherche des LSP de secours en un temps polynomial (en fonction de la taille du problème), diverses heuristiques peuvent être adoptées. Ces dernières pourraient subdiviser l'arbre P2MP primaire en ses LSP S2L avant de protéger chacun par un ensemble de chemins de secours ou protéger chaque risque de panne du LSP P2MP primaire par un LSP P2MP de secours. Pour des raisons d'optimalité, les chemins de secours (resp. les arbres de secours) utilisés devraient correspondre aux plus courts chemins [Dij59] (resp. à des arbres approchés de Steiner [TM80, KMB81]).

Comme pour les algorithmes de recherche des LSP de secours qui minimisent la bande passante de secours (exemple de l'ILP), les heuristiques utilisant la protection multicast un-à-un devraient appliquer et privilégier la fusion de flux à tout autre type de fusion. De plus, si le partage de la bande passante est global, il sera vivement recommandé de ne pas tenir compte de la fusion des LSP lors des calculs.

Concernant les mécanismes de distribution de l'information requise au placement des LSP de secours, les heuristiques adoptent souvent des approches de diffusion. Comme nous allons le voir dans le chapitre suivant, il sera possible d'étendre et de réutiliser les mêmes mécanismes de diffusion que ceux décrits dans les sections 2.8, 3.3 et 3.4, pour le calcul des LSP P2MP de secours.

5.4.2.5 Information annoncée dans le réseau pour permettre le calcul des LSP de secours

Pour éviter de bloquer les requêtes de protection par erreur, les entités calculant les LSP de secours doivent disposer, pour tout arc, de toute l'information nécessaire à la détermination des quantités cumulées de la bande passante primaire et des prix de protection de tous les risques de panne. Pour ce faire, les trois approches décrites dans

la section 2.7 ([KKL+01, LRC02, VCLF+04]) pourraient être adaptées à la protection multicast.

Dans la première approche (inspirée de [KKL+01]), tout nœud u du réseau diffuse pour chaque arc $v \rightarrow u_{v \in V}$ dont il est le nœud sortant, les prix des risques protégés en utilisant cet arc, la quantité cumulée de bande passante allouée et la capacité de l'arc. De cette manière, tout nœud vérifiera si un arc λ peut être utilisé afin d'établir un nouveau LSP P2MP de secours b ($b \in B_{k+1}$) en appliquant les formules suivantes :
$Max_{r \in PFRG(b)} \delta_r^{\lambda}(k) + bw(b) + F_{\lambda}(k) \leq C^{\lambda}$ si le partage est restreint ;
$Max_{r \in PFRG(b)} (\delta*)_r^{\lambda}(k) + bw(b) + F_{\lambda}(k) \leq C^{\lambda}$ si le partage est global.

Bien évidemment, cette première approche ne passe pas à l'échelle puisque la fréquence et la taille des messages diffusés dans le réseau peuvent être très élevées.

Dans la deuxième approche (inspirée de [LRC02]), ce sont les structures et propriétés des LSP de secours qui sont diffusées dans le réseau (en plus des quantités cumulées de la bande passante primaire et des capacités des arcs). Cela permet de déduire l'information requise au calcul des LSP de secours et par conséquent d'accepter ou de refuser sans erreur l'arc $u \rightarrow v$, lors du calcul d'un nouveau LSP de secours.

Cette approche s'adapte mal à la protection multicast un-à-un puisqu'elle génère une très grande quantité de trafic de contrôle (quantité dépendant du nombre de LSP de secours établis). C'est une technique qui ne passe pas à l'échelle.

Dans la troisième et dernière approche décrite ici (inspirée de [VCLF+04]), une entité de calcul de LSP de secours est associée à chaque risque de panne. En mémorisant les structures et les propriétés des LSP primaires traversant le nœud qui la supporte et en stockant les structures et propriétés des LSP de secours qu'elle calcule, chaque entité sera capable de déterminer l'ensemble des arcs pouvant supporter un nouveau LSP de secours si ce dernier protège contre le risque de panne géré par l'entité.

Cette dernière approche souffre des mêmes inconvénients que ceux rencontrées dans l'unicast. Concrètement, à cause de la nécessité d'employer un mécanisme de différenciation des pannes des nœuds de celles des liens [VC02, AKNS08], cette approche induit une augmentation significative des délais de récupération.

5.4.2.6 Constats

La protection multicast un-à-un protège efficacement les communications multicast. Typiquement, elle augmente le taux de protection, réduit considérablement la consommation de la bande passante de secours et permet de diminuer les délais de récupération.

Bien que les extensions introduites dans [PSA05] et [APY07] permettent une configuration facile des LSP primaires et de leurs LSP de secours, des précautions devraient être prises, notamment pour le choix d'identification de LSP de secours. Ainsi et contrairement aux recommandations de [APY07], l'identification par la source devrait être préférée à l'identification par chemin pour deux raisons essentielles :

1. L'identification par chemin réduit la flexibilité de choix des LSP de secours, ce qui pourrait diminuer le taux de protection. Typiquement, l'utilisation de l'identification par chemin pour protéger un réseau contenant des SRLG réduit souvent le taux de protection.

2. L'identification par chemin se combine mal avec la fusion de flux. En effet, à cause de l'obligation de la fusion de certains LSP avec l'identification par chemin, de la bande passante peut être gaspillée. En effet, la combinaison de la fusion de LSP avec la fusion de flux peut mener à une fuite de la bande passante comme nous l'avons expliqué en section 5.4.2.1).

Enfin, si la protection multicast un-à-un permet d'optimiser ou de réduire la consommation de la bande passante de secours, elle le fait au détriment d'une utilisation d'algorithmes de calcul plus complexes, une transmission d'une grande quantité d'informations dans le réseau et l'emploi d'un nombre assez élevé d'étiquettes MPLS. Afin d'atténuer les effets des deux premiers désavantages et pour faciliter l'application de la protection multicast un-à-un, des heuristiques de calcul et de diffusion peuvent se substituer aux algorithmes de calcul et techniques de diffusion décrites dans la section 5.4.2.5 (cf. chapitre suivant).

Concernant le nombre d'étiquettes MPLS, il peut être réduit considérablement avec l'utilisation de l'agrégation de LSP (en utilisant la hiérarchie MPLS). Ceci nous mène aux second et troisième types de protection multicast locale présentés dans ce chapitre.

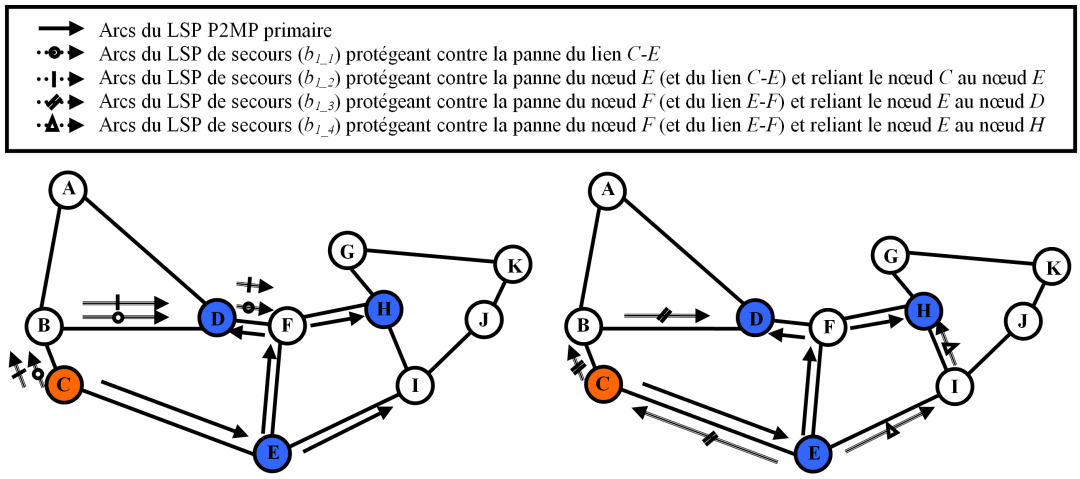
5.4.3 Protection multicast par tunnel P2P de secours

Avec ce type de protection, un seul tunnel MPLS est établi pour protéger un ensemble de LSP S2L primaires (appartenant à plusieurs LSP P2MP primaires) contre la panne d'un lien (resp. contre les pannes d'un lien et d'un nœud) qui leur est commun (resp. leur sont communs). Bien évidemment, un tunnel de secours établi entre un nœud PLR et un nœud MP et contournant un composant donné du réseau (le lien et éventuellement le nœud protégés), ne peut être utilisé que pour la protection des LSP S2L primaires traversant le PLR, le composant protégé et le nœud MP dans cet ordre.

Pour protéger un LSP P2MP primaire contre la panne d'un lien $u \rightarrow v$ situé en aval du PLR u, l'entité de calcul des tunnels de secours vérifie d'abord si le PLR u a déjà configuré au moins un tunnel NHOP permettant de protéger contre la panne du lien $u \rightarrow v$. Si un tel tunnel existe et dispose d'une quantité de bande passante suffisante, il sera sélectionné pour protéger le LSP P2MP primaire contre la panne du lien $u \rightarrow v$. Si, par contre, aucun tunnel n'a été établi pour protéger le lien $u \rightarrow v$ ou que tous les tunnels protégeant ce lien ne disposent pas d'une quantité de bande suffisante, alors un nouveau tunnel de secours protégeant le lien $u \rightarrow v$ et vérifiant les contraintes de bande passante est calculé et est créé. De même, pour protéger contre la panne d'un nœud ayant un seul fils dans un LSP P2MP primaire, il suffit d'utiliser un tunnel NNHOP contournant le nœud protégé et reliant ses nœuds père et fils dans le LSP P2MP primaire.

Lorsque le composant à protéger est un nœud de branchement, la protection multicast par tunnel P2P de secours cherche à déterminer autant de tunnels NNHOP que de nœuds fils du nœud de branchement. Ces tunnels doivent relier le nœud PLR (nœud père du nœud protégé) à tous les nœuds fils du nœud protégé.

Sur la figure 5.7, une session multicast s ayant comme source le nœud C et comme membres les nœuds D, E et H est établie. Cette session multicast utilise, pour le routage



(a) Un seul tunnel P2P de secours pour protéger contre la panne d'un lien ou d'un nœud ayant un seul fils

(b) Plusieurs tunnels P2P de secours pour protéger contre la panne d'un nœud de branchement

FIG. 5.7 – Protection multicast par tunnel P2P de secours

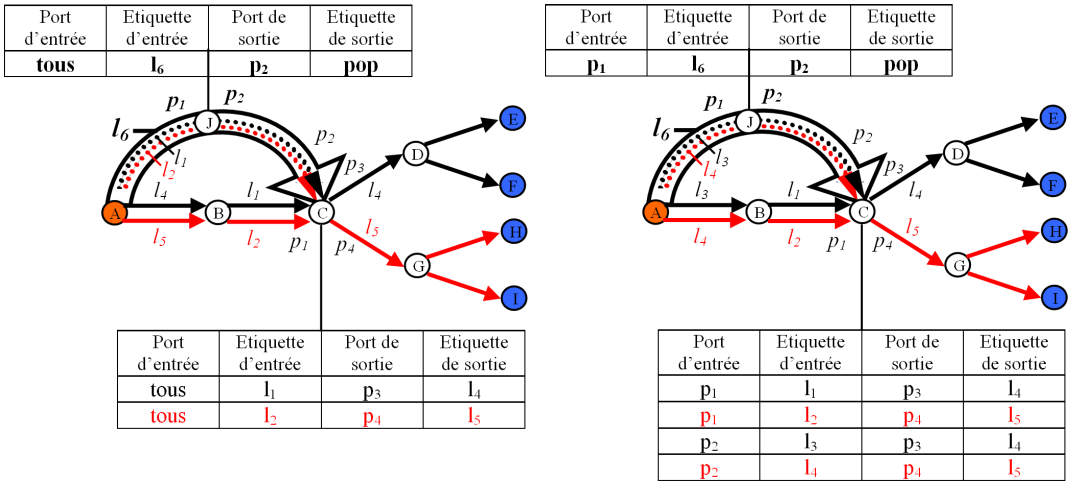
en l'absence de panne, un LSP P2MP primaire p_1 composé de trois LSP S2L : p_{1_1} = ($C \rightarrow E$), p_{1_2} = ($C \rightarrow E \rightarrow F \rightarrow D$), p_{1_3} = ($C \rightarrow E \rightarrow F \rightarrow H$).

Pour protéger contre la panne du lien C-E (resp. du nœud E), un tunnel de secours b_{1_1} = $C \rightarrow B \rightarrow D \rightarrow F \rightarrow E$ (resp. b_{1_2} = $C \rightarrow B \rightarrow D \rightarrow F$) reliant les nœuds C et E (resp. les nœuds C et F) est établi. Ce tunnel peut être utilisé pour protéger tous les LSP S2L (resp. les deux LSP S2L p_{1_2} et p_{1_3}) du LSP primaire p_1 puisque tous ces LSP S2L (ces deux LSP S2L) passent par les nœuds d'extrémité C et E (resp. C et F) du tunnel de secours et du composant protégé (cf. figure 5.7 (a)).

Pour protéger contre la panne du nœud F qui possède deux nœuds fils dans le LSP P2MP primaire p_1, deux tunnels de secours b_{1_3} = $E \rightarrow C \rightarrow B \rightarrow D$ et b_{1_4} = $E \rightarrow I \rightarrow H$ sont établis (figure 5.7 (b)). Le premier tunnel NNHOP a comme extrémités les nœuds E et D ; il ne peut donc protéger que le LSP S2L primaire p_{1_2} qui passe par le nœud protégé F. De même, le second tunnel NNHOP a comme extrémités les nœuds E et H ; ce tunnel ne peut donc protéger que le LSP S2L primaire p_{1_3} qui traverse le nœud protégé F. Nous notons que les deux tunnels P2P de secours (b_{1_3} et b_{1_4}) permettent une récupération complète de la communication multicast s lors d'une panne affectant le nœud F (ou une panne affectant le lien E-F). En effet, tous les LSP S2L permettant de supporter la communication multicast s et passant par le nœud F (le lien E-F) sont entièrement protégés contre la panne de ce dernier.

5.4.3.1 Signalisation des tunnels P2P de secours

Un tunnel de secours est un LSP interconnectant deux nœuds : un nœud PLR et un nœud MP. Il est établi pour protéger un ensemble de LSP primaires traversant ces deux nœuds. Il peut être identifié sur chaque nœud par une étiquette et éventuellement une



(a) Espace d'étiquettes global à chaque nœud (b) Espace d'étiquettes local à chaque interface

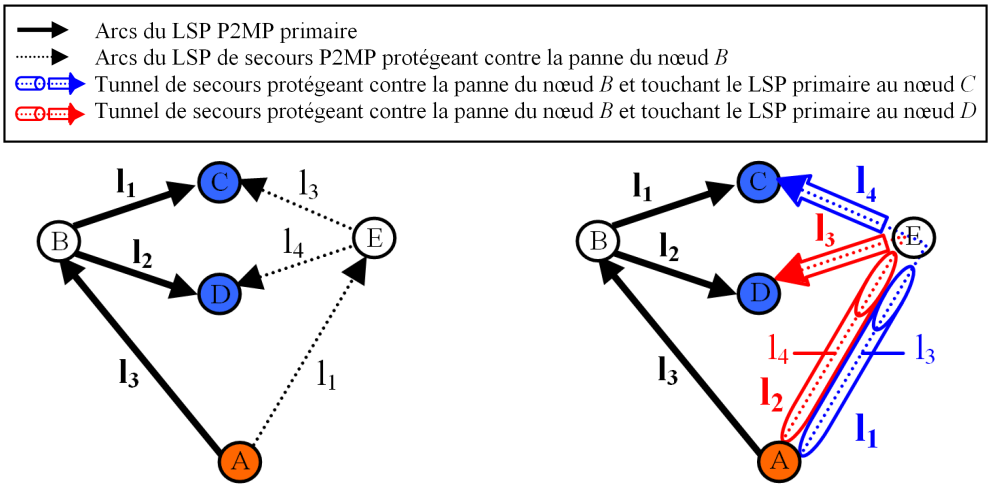
FIG. 5.8 – Allocation d'étiquettes pour un tunnel de secours

interface. A la détection d'une panne, le trafic de tous les LSP primaires protégés par un tunnel P2P de secours est redirigé vers ce dernier. Pour ce faire, le nœud PLR empile dans les paquets associés aux LSP primaires affectés une deuxième étiquette (située au sommet de la pile MPLS) identifiant le tunnel P2P de secours, en plus de la première étiquette permettant de séparer les flux à la sortie du tunnel (sur le nœud MP).

Sur la figure 5.8, une première session multicast s_1 (resp. seconde session multicast s_2) est établie afin d'alimenter les deux membres multicast E et F (resp. H et I) à partir de la source multicast A. La première session multicast emploie le LSP P2MP primaire p_1 constitué des arcs $A \rightarrow B$, $B \rightarrow C$, $C \rightarrow D$, $D \rightarrow E$, $D \rightarrow F$ et la seconde session utilise le LSP P2MP primaire p_2 composé des arcs $A \rightarrow B$, $B \rightarrow C$, $C \rightarrow G$, $G \rightarrow H$, $G \rightarrow I$. Pour protéger ces deux sessions multicast contre la panne du nœud B et du lien A-B, un tunnel de secours T_1 ($A \rightarrow J \rightarrow C$) reliant le PLR A au MP C est calculé et est configuré. Ce tunnel est destiné à supporter les flux des deux sessions multicast s_1 et s_2, entre les deux nœuds d'extrémité A et C, lors de la panne du nœud B ou du lien $A \rightarrow B$. Il est idendifié grâce à l'étiquette au sommet de la pile MPLS. Nous notons que ces flux sont opaques au nœud intermédiaire J situé entre le PLR A et le MP C.

Pour permettre au MP C d'aiguiller correctement les paquets à la sortie du tunnel T_1 (vers les LSP primaires p_1 et p_2), le PLR A insère dans les paquets émis une seconde étiquette (étiquette située au deuxième rang dans la pile MPLS des paquets). Cette dernière a pour rôle d'identifier les différents flux au sein d'un même tunnel de secours. Selon le type d'espace d'étiquettes utilisé par le MP (espace d'étiquettes global à toutes les interfaces ou espace d'étiquettes local à chaque interface), le PLR peut utiliser deux procédés pour déterminer ces étiquettes : la *déduction* ou la *découverte*.

Lorsque l'espace d'étiquettes utilisé par le MP est global à toutes ses interfaces, le PLR peut utiliser la même étiquette (en seconde position sur la pile MPLS) que celle retournée par le MP à son nœud en amont sur le LSP primaire. Cette étiquette peut



(a) Protection multicast un-à-un : une seule réservation de bande est effectuée sur l'arc $A \rightarrow E$

(b) Protection multicast par tunnels P2P de secours : deux réservations de bande sont nécessaires sur l'arc $A \rightarrow E$

FIG. 5.9 – Allocation de la bande passante avec la protection multicast un-à-un et la protection multicast par tunnels P2P de secours

être déduite en consultant l'objet RECORD_ROUTE du message *resv* qui configure le LSP primaire. Ainsi, lors de la panne du nœud B dans l'exemple de la figure 5.8 (a), le PLR A insère les étiquettes l_6 (sommet de la pile) et l_1 (seconde étiquette) dans la pile MPLS de tout message appartenant à la session s_1. Pour les messages de la session s_2, le PLR A ajoute les étiquettes l_6 (sommet de la pile) et l_2 (seconde étiquette) dans leur pile MPLS. Le nœud J (*penultimate hop*), recevant ces paquets, dépile l'étiquette l_6 avant de retransmettre les messages au MP C. Ce dernier aiguillera correctement les paquets reçus en utilisant les mêmes entrées multicast que celles allouées aux LSP P2MP primaires correspondants (p_1 et p_2).

Lorsque l'espace d'étiquettes utilisé par le MP est local à chacune de ses interfaces, le PLR doit découvrir la seconde étiquette permettant de séparer les flux à la sortie du tunnel de secours avant de basculer le trafic vers ce dernier. Pour ce faire, à chaque LSP P2MP primaire protégé par le tunnel de secours, le PLR envoie un message *path* directement au MP. Ce dernier alloue alors les étiquettes (une étiquette pour chaque LSP P2MP primaire protégé) et les transmet directement au PLR dans des messages *resv*. Par exemple, sur la figure 5.8 (b), le MP a réservé au LSP P2MP primaire p_1 l'étiquette l_3 sur le tunnel P2P de secours T_1 et il a alloué au LSP P2MP primaire p_2 l'étiquette l_4 sur le même tunnel P2P de secours (voir les deux dernières entrées de la table d'étiquettes du nœud C, illustrée sur la figure 5.8 (b)). Ainsi, lors de la panne du nœud B, le PLR A insère les étiquettes l_6 (sommet de la pile) et l_3 (seconde étiquette) dans la pile MPLS de tout paquet appartenant à s_1. De la même manière, le PLR A ajoute les étiquettes l_6 (sommet de la pile) et l_4 (seconde étiquette) dans la pile MPLS de tout paquet appartenant à s_2. Le nœud J (*penultimate hop*) recevant ces

paquets, dépile l'étiquette l_6 avant de les retransmettre au MP C. Ce dernier aiguillera correctement les paquets reçus vers les LSP primaires associés en utilisant les étiquettes allouées et transmises auparavant au PLR.

5.4.3.2 Partage et calcul de la bande passante de secours

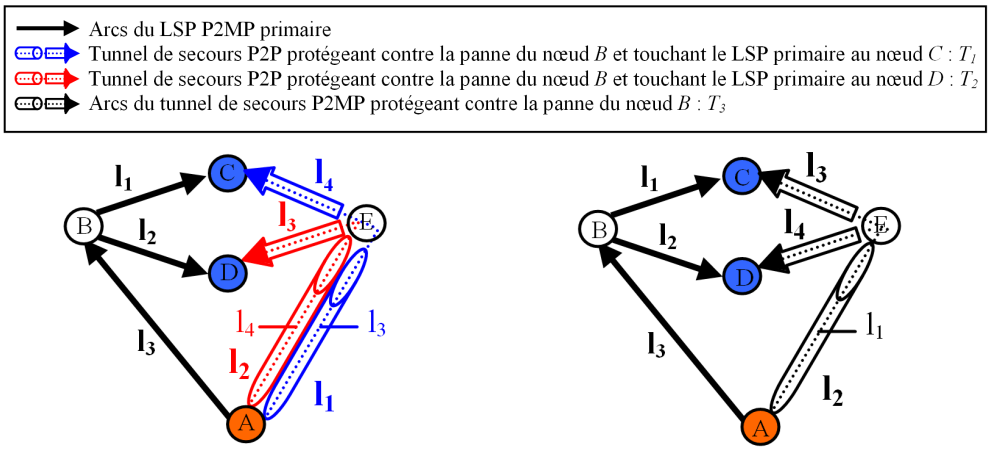
Bien que la protection multicast par tunnels P2P de secours interdise tout type de fusion entre les tunnels de secours, le partage de la bande passante reste possible mais légèrement diminué. En effet, avec ce type de protection, la fusion de flux est restreinte et n'est applicable qu'entre les LSP de secours[11] ayant les mêmes nœuds d'extrémité et protégeant des LSP S2L primaires d'une même session multicast. Le partage de la bande passante entre les tunnels de secours (partage inter-tunnels) reste possible et peut être effectué entre tous les tunnels P2P de secours protégeant des ensembles de risques de panne disjoints.

Sur la figure 5.9, une session multicast s, nécessitant une unité de bande passante et ayant comme source le nœud A et comme membres multicast les nœuds C et D, est établie. Cette session utilise le LSP P2MP primaire composé des arcs $A \rightarrow B$, $B \rightarrow C$ et $B \rightarrow D$ afin de router son trafic en l'absence de pannes.

En adoptant la protection multicast un-à-un (figure 5.9 (a)), un seul LSP P2MP de secours, constitué des arcs $A \rightarrow E$, $E \rightarrow C$ et $E \rightarrow D$, est suffisant et est utilisé pour protéger la session s contre la panne du nœud de branchement B. Ce LSP de secours ne réserve qu'une seule unité de bande passante sur l'arc $A \rightarrow E$ (et sur les arcs $E \rightarrow C$ et $E \rightarrow D$). Avec la protection par tunnels P2P de secours (figure 5.9 (b)) par contre, deux tunnels P2P de secours T_1 ($A \rightarrow E \rightarrow C$) et T_2 ($A \rightarrow E \rightarrow D$) sont établis et utilisés pour protéger contre la panne du nœud de branchement B. Puisque le partage de la bande passante n'est autorisé qu'entre des tunnels de secours protégeant contre des ensembles de risques de panne disjoints, les deux tunnels T_1 et T_2 ne peuvent pas partager leur allocation de la bande passante. En conséquence, deux unités de bande passante sont nécessaires sur l'arc $A \rightarrow E$ pour protéger contre la panne du nœud de branchement B (cependant, une seule unité est suffisante et est allouée sur les arcs $E \rightarrow C$ et $E \rightarrow D$).

Si nous calculons la quantité de bande passante d'un tunnel de secours comme la somme des quantités de bande passante allouées aux LSP primaires P2MP protégés par ce tunnel et si nous redéfinissons le prix de protection $\delta_r^\lambda\ (k)$ comme étant la quantité de bande passante de secours cumulée de tous les tunnels P2P de secours traversant l'arc λ et protégeant les k premières sessions contre le risque r, nous obtenons la quantité minimale de bande passante de secours à allouer sur tout arc λ comme suit :
Lorsque le partage de la bande passante est restreint : $G^\lambda\ (k) = Max_r\ (\delta_r^\lambda\ (k))$
Lorsque le partage de la bande passante est global : $G^\lambda\ (k) = Max_r\ (\delta_r^\lambda\ (k)$ - $L_r^\lambda\ (k), 0)$

[11] Un LSP de secours encapsulé dans un tunnel P2P de secours relie directement les nœuds d'extrémité de ce tunnel. Il est identifié par une étiquette MPLS située au dessous de l'étiquette identifiant le tunnel P2P de secours.



(a) Protection multicast par tunnels P2P de secours : deux réservations de bande sont nécessaires sur l'arc $A \rightarrow E$

(b) Protection multicast par tunnels P2MP de secours : une seule réservation de bande est effectuée sur l'arc $A \rightarrow E$

FIG. 5.10 – Allocation d'étiquettes et de bande passante pour les tunnels P2P et P2MP de secours

5.4.3.3 Information annoncée dans le réseau pour permettre le calcul des tunnels de secours

Comme pour la méthode de protection multicast un-à-un, le calcul des tunnels P2P de secours requiert la diffusion/transmission d'une information permettant de déterminer, sur chaque arc, la quantité de bande passante primaire cumulée et les prix de protection de tous les risques de panne. Cela peut être obtenu en adoptant les mêmes approches que celles décrites dans la section 5.4.2.5, à savoir :

- diffusion, pour chaque arc, de la quantité cumulé de la bande passante primaire et des prix de protection de tous les risques,
- diffusion des structures et propriétés des LSP primaires et des tunnels P2P de secours,
- association d'une entité de calcul à chaque risque de panne afin de calculer les tunnels P2P de secours protégeant contre la panne de ce risque.

5.4.4 Protection multicast par tunnel P2MP de secours

Afin d'accroître la disponibilité de la bande passante, [LRAVV07] propose d'utiliser des tunnels P2MP de secours afin de protéger les sessions multicast (SSM). Nous notons qu'un tunnel de secours P2MP est un LSP P2MP établi entre un nœud PLR et un ensemble de nœuds MP, afin de protéger un ou plusieurs LSP primaires traversant le PLR et les nœuds feuilles du tunnel. Il permet d'éviter la double allocation de la bande passante sur les arcs des structures de secours protégeant les mêmes ensembles de risques, ce qui évite la duplication du trafic multicast sur les arcs lors d'une panne d'un nœud (branchement ou pas) ou d'un lien.

Sur la figure 5.10, une session multicast s nécessitant une unité de bande passante et ayant comme source le nœud A et comme membres multicast les nœuds C et D est établie. Cette session utilise le LSP P2MP primaire composé des arcs $A \rightarrow B$, $B \rightarrow C$ et $B \rightarrow D$. Pour illustrer la différence entre la protection multicast par tunnels P2P de secours et la protection multicast par tunnels P2MP de secours, nous avons utilisé ces deux types de tunnels pour protéger la session s contre la panne du nœud de branchement B. Ainsi, avec la protection multicast par tunnels P2P (figure 5.10 (a)), deux tunnels P2P de secours T_1 ($A \rightarrow E \rightarrow C$) et T_2 ($A \rightarrow E \rightarrow D$) sont nécessaires pour protéger les deux LSP S2L primaires associés à la session s et passant par le nœud de branchement B. Vu que seuls des tunnels de secours indépendants peuvent partager leurs allocations de bande passante, deux unités de bande passante seront allouées aux deux tunnels de secours T_1 et T_2 (tunnels non indépendants) afin d'assurer la disponibilité de la bande passante lors de la récupération de la panne du nœud de branchement B. Avec la protection multicast par tunnels P2MP de secours (figure 5.10 (b)), un seul tunnel P2MP T_3 ($A \rightarrow E$, $E \rightarrow C$, $E \rightarrow D$), couvrant le PLR A et les deux nœuds C et D situés en aval du nœud protégé B, est nécessaire. Ce tunnel T_3 élimine la duplication du flux de la session multicast s sur l'arc $A \rightarrow E$ lors de la panne du nœud de branchement B. En conséquence, une seule allocation de la bande passante (i.e. une unité de bande passante) est suffisante et est allouée sur l'arc $A \rightarrow E$ afin de protéger contre la panne du nœud de branchement B.

En plus de l'avantage lié à la diminution de la consommation de la bande passante de secours, les tunnels de secours P2MP peuvent aussi réduire les délais de traitement des paquets sur certains nœuds. Typiquement, dans l'exemple de la figure 5.10, deux recherches d'entrées multicast sont effectuées sur le nœud E si nous utilisons la protection multicast par tunnel P2P de secours alors qu'une seule recherche suffit si nous optons pour la protection multicast par tunnel P2MP de secours. Nous reviendrons sur ce point, avec plus de détails, dans la section 5.4.4.2.

5.4.4.1 Signalisation des tunnels P2MP de secours

Un tunnel P2MP de secours est un LSP établi entre un nœud PLR et un ensemble de nœuds MP afin de protéger un ou plusieurs LSP P2MP primaires traversant le PLR et les nœuds feuilles du tunnel. Lorsqu'une panne est détectée, le trafic de tout LSP primaire affecté est basculé vers son tunnel P2MP de secours en remplaçant l'étiquettes associée au LSP primaire affecté par un couple d'étiquettes : une première étiquette, située au sommet de la pile MPLS, permettant de basculer et d'acheminer le trafic du LSP P2MP primaire affecté sur le tunnel de secours et une seconde étiquette (située au dessous du sommet de la pile MPLS) permettant d'identifier le trafic du LSP primaire à la sortie du tunnel de secours. Cette substitution d'étiquettes est effectuée de manière à assurer que chaque tunnel de secours activé ne reçoit que le trafic des LSP primaires qu'il protège.

Comme la seconde étiquette insérée par le PLR n'est consultée que par les MP, il en ressort que cette étiquette (identifiant le flux d'une seule session multicast) est identique sur tous les MP d'un tunnel de secours. Afin de faciliter le calcul et la distribution de

cette seconde étiquette, [ALR08] et [ARR08] suggèrent de fixer sa valeur par les nœuds PLR et de l'associer à un contexte commun qui est le tunnel de secours (contexte commun au PLR et à tous les nœuds feuilles du tunnel de secours). Évidemment, ce mécanisme d'association des paquets à leurs LSP primaires sur les nœuds MP, nécessite la désactivation du mécanisme de *penultimate hop popping* (i.e. il est nécessaire d'allouer une étiquette sur tout lien d'un tunnel P2MP de secours).

Avec RSVP-TE ([ALR08]), l'association de la seconde étiquette au contexte (tunnel de secours) est effectuée par le nœud PLR qui insert des objets IF_ID RSVP_HOP[12] et UPSTREAM_ASSIGNED_LABEL[13] dans les messages *path* des LSP P2MP de secours (encapsulés dans les tunnels P2MP de secours) protégeant le LSP P2MP primaire.

Sur la figure 5.10 (b), le nœud A (source multicast) insère deux étiquettes dans la pile MPLS de chaque paquet appartenant à la session s, après la détection d'une panne affectant le nœud de branchement B : Une première étiquette l_2 (sommet de la pile) permettant d'acheminer le paquet sur le tunnel de secours T_3 et une seconde étiquette l_1 identifiant le LSP de secours protégeant la session s au sein du tunnel T_3. Lorsqu'un paquet redirigé sur le tunnel T_3 est reçu par le nœud E, ce dernier le duplique et échange l'étiquette l_2 (située au sommet de la pile MPLS) du premier paquet (resp. du second paquet) par l'étiquette l_3 (resp. par l'étiquette l_4) avant de l'envoyer au prochain nœud C (resp. nœud D). A la réception d'un paquet envoyé par le nœud E et ayant emprunté le tunnel de secours T_3, le nœud MP C (resp. le nœud MP D) déduit les destinations du paquet en consultant les deux étiquettes MPLS l_3 et l_1 (resp . l_4 et l_l). Concrètement, en examinant l'étiquette l_3 (resp. l_4) située au sommet de la pile MPLS, le MP C (resp. le MP D) déduit le contexte[14] dans lequel est définie la seconde étiquette l_1, ce qui permet d'utiliser ensuite l'étiquette l_1 afin de déterminer le LSP P2MP primaire protégé associé au paquet.

5.4.4.2 Délais de traitement des paquets sur les nœuds : tunnels de secours P2MP vs. tunnel de secours P2P

En plus de la minimisation de la quantité de bande passante allouée pour fournir la protection, la diminution du délai de traitement des paquets sur les nœuds des tunnels de secours, après la récupération d'une panne, est fondamentale pour passer à l'échelle. Essentiellement, il existe trois[15] paramètres permettant de contrôler, sur un nœud donné u, la valeur du délai de traitement Dt_u des paquets redirigés vers des tunnels de secours après la récupération d'une panne : le temps moyen d'accès à une entrée de la table d'étiquettes t_u (déterminé généralement par le nombre d'entrées dans la table d'étiquettes du nœud u), le nombre moyen de paquets n_u acheminés sur des tunnels de secours et reçus par le nœud u et le nombre moyen de copies c_u effectuées

[12]L'objet IF_ID RSVP_HOP contient l'objet SESSION et éventuellement l'objet SENDER_TEMPLATE du tunnel P2MP de secours.

[13]L'objet UPSTREAM_ASSIGNED_LABEL contient une étiquette allouée par un PLR et utilisée pour associer un LSP de secours au sein d'un tunnel de secours au LSP primaire protégé.

[14]Le contexte peut être vu comme une sous-table d'étiquettes associée au tunnel de secours sur un nœud MP.

[15]Le délai de transmission des paquets est ignoré et n'est pas pris en compte.

sur les paquets acheminés sur les tunnels de secours et reçus par le nœud u. Ainsi : $Dt_u = n_u.t_u + c_u.tc_u$
où : tc_u est le temps moyen nécessaire au nœud u pour effectuer une copie d'un paquet.

Concernant le délai global Dg consacré au traitement de tous les paquets redirigés sur les tunnels de secours activés après une panne, il est déterminé comme la somme des délais de traitement sur tous les nœuds des tunnels actifs. Il est calculé comme suit : $Dg = \sum_{u \in V} Dt_u = \sum_{u \in V} (n_u.t_u + c_u.tc_u) = \sum_{u \in V} n_u.t_u + \sum_{u \in V} c_u.tc_u$

Pour évaluer l'impact du type de tunnel de secours choisi (P2P ou P2MP) sur le délai de traitement des paquets redirigés sur les tunnels de secours suite à une panne, il suffira d'analyser les paramètres intervenant dans le calcul du délai global Dg. Ainsi, si nous définissons pour chaque nœud u les paramètres : t_u^{P2MP} (resp. t_u^{P2P}) comme étant le temps moyen d'accès à une entrée de la table d'étiquettes du nœud u, n_u^{P2MP} (resp. n_u^{P2P}) comme le nombre moyen de paquets acheminés sur des tunnels de secours P2MP (resp. P2P) traversant le nœud u et c_u^{P2MP} (resp. c_u^{P2P}) comme étant le nombre moyen de copies effectuées par le nœud u sur les paquets reçus et acheminés sur des tunnels de secours P2MP (resp. P2P), nous aurons pour une panne affectant le risque r : $\sum_{u \in V} c_u^{P2MP}.tc_u = \sum_{u \in V} c_u^{P2P}.tc_u$ car $\sum_{u \in V} c_u^{P2MP}$ et $\sum_{u \in V} c_u^{P2MP}$ sont égaux et ne dépendent que du nombre de nœuds primaires situés en aval du risque en panne.

En conséquence, tout écart de délai de traitement des paquets sur les tunnels de secours P2P et P2MP sera essentiellement dû à la différence entre les valeurs de $\sum_{u \in V} n_u^{P2MP}.t_u^{P2MP}$ et $\sum_{u \in V} n_u^{P2P}.t_u^{P2P}$. Bien évidemment, les valeurs de ces deux derniers éléments dépendent fortement de la matrice de trafic, de la topologie du réseau et de l'algorithme de calcul des chemins adopté. Souvent, nous avons : $\sum_{u \in V} n_u^{P2MP} \leq \sum_{u \in V} n_u^{P2P}$ (car le coût moyen, en nombre de hops, d'un tunnel de secours P2MP protégeant un risque donné est souvent inférieur à la somme des coûts des tunnels de secours P2P protégeant le même risque) et $\sum_{u \in V} t_u^{P2MP} \geq \sum_{u \in V} t_u^{P2P}$ (pour un même taux de protection, le nombre moyen de tunnels de secours P2MP traversant un nœud donné est généralement supérieur au nombre de tunnels de secours P2P traversant le même nœud).

Nombre maximum de tunnels de secours protégeant contre la panne d'un nœud.

Pour approximer la différence entre les temps d'accès aux tables d'étiquettes obtenus avec les deux méthodes de protection multicast (par tunnel P2MP et par tunnel P2P), il serait intéressant dans un premier temps d'évaluer la différence entre le nombre de tunnels de secours P2MP et le nombre de tunnels de secours P2P nécessaire à la protection d'un composant du réseau.

Si nous considérons que la bande passante disponible sur les liens est suffisante (infinie), il suffira de construire $n_u^{P2MP} = d.(2^{(d-1)} - 1)$ tunnels P2MP de secours pour protéger tout LSP primaire contre la panne d'un nœud u de degré d. En effet, nous avons d possibilités de choix du PLR et, pour chaque PLR, nous disposons de $(\mathcal{C}_1^{d-1} + \mathcal{C}_2^{d-1} + .. + \mathcal{C}_{d-1}^{d-1} = 2^{(d-1)} - 1)$ choix d'ensembles de nœuds MP. Si nous utilisons la protection par tunnel P2P de secours, le nombre de tunnels P2P de secours nécessaires à la protection de tout nœud u de degré d sera égal à $n_u^{P2P} = d.(d-1)$ (d choix de PLR

et $d-1$ choix de MP pour chaque PLR).

En supposant que le temps de recherche d'une entrée dans la table d'étiquettes d'un nœud u est linéairement dépendant du logarithme de la taille de la table, nous obtenons :
$t_u^{P2MP}/t_u^{P2P} \approx [log_2\ (2^{(d-1)}-1) + log_2\ d] \ / \ [log_2\ d + log_2\ (d-1)]$
$\approx (d-1+log_2\ d) \ / \ [log_2\ d + log_2\ (d-1)]$ si le nombre de LSP primaires est quasi-nul.
$t_u^{P2MP}/t_u^{P2P} \approx 1$ si le nombre de LSP primaires est très élevé par rapport au nombre de tunnels de secours (ce qui est souvent le cas).

Notons que le nombre de tunnels P2MP de secours peut être réduit considérablement en adoptant la technique d'agrégation des LSP avec perte de la bande passante [FCGF01, Mou06].

Distribution de la charge de duplication des paquets redirigés suite à une panne.

Avec la protection par tunnels P2MP de secours, la duplication des paquets redirigés, suite à une panne, est effectuée par l'ensemble des nœuds de branchement des tunnels P2MP de secours. Cet ensemble peut contenir n'importe quel nœud du réseau (mis-à-part le nœud en panne s'il s'agit d'une panne d'un nœud ou les extrémités du lien en panne s'il s'agit d'une panne d'un lien). En conséquence, la charge de duplication des paquets est souvent partagée entre plusieurs nœuds, ce qui diminue des risques d'apparition de goulots d'étranglement.

Cependant, avec la protection multicast par tunnels P2P de secours, la duplication des paquets redirigés, suite à une panne d'un nœud u, n'est effectuée que par les PLR voisins au composant en panne (i.e. nœuds voisins du nœud en panne s'il s'agit d'une panne d'un nœud ou nœuds d'extrémité du lien en panne s'il s'agit d'une panne d'un lien). La charge de duplication des paquets est alors centralisée sur ces nœuds PLR, ce qui augmente leur charge et accroît les risques de congestion.

Constats.

Deux techniques de protection multicast locale par tunnel de secours existent : protection par tunnel P2P de secours et protection par tunnel P2MP de secours.

La première technique permet de réduire le nombre de LSP dans le réseau (et donc, elle diminue le nombre d'états RSVP à maintenir) au détriment d'une surconsommation de la bande passante due à la duplication du trafic multicast sur certains arcs, après la récupération d'une panne. La seconde technique, quant à elle, diminue la consommation de la bande passante en fusionnant les flux des LSP P2P de secours protégeant une même session multicast contre les mêmes ensembles de risques de panne. C'est une technique qui partage la tâche de duplication des paquets après une panne sur un ensemble de nœuds étendu (contrairement à la protection par tunnel P2P de secours où la duplication des paquets redirigés suite à une panne est centralisée sur l'ensemble des PLR), ce qui diminue le risque d'apparition de goulots d'étranglement autour des nœuds adjacents au composant défaillant. Son seul désavantage est qu'elle augmente le nombre de LSP dans le réseau, ce qui accroît le trafic de contrôle nécessaire à l'établissement et au maintien de ces LSP.

5.4.4.3 Allocation de la bande passante de secours et information transmise dans le réseau pour le calcul distribué des tunnels P2MP de secours

La protection multicast par tunnels P2MP de secours offre une meilleure disponibilité de la bande passante par rapport à la protection multicast par tunnels P2P de secours. Si nous redéfinissons le prix de protection $\delta_r^\lambda\ (k)$ comme étant la quantité de bande passante de secours cumulée de tous les tunnels P2MP de secours traversant l'arc λ et protégeant les k premières sessions contre le risque r, nous obtenons la quantité minimale de bande passante de secours à allouer sur tout arc λ comme suit :
Lorsque le partage de la bande passante est restreint : $G^\lambda\ (k) = Max_r\ (\delta_r^\lambda\ (k))$
Lorsque le partage de la bande passante est global : $G^\lambda\ (k) = Max_r\ (\delta_r^\lambda\ (k)$ - $L_r^\lambda\ (k), 0)$

Pour permettre le calcul distribué des tunnels P2MP de secours, les mêmes méthodes que celles décrites dans la section 5.4.3.3 peuvent être réutilisées. Nous notons que l'application de ces méthodes pour le calcul des tunnels P2MP de secours offre les mêmes avantages et inconvénients que ceux fournis par l'application de ces méthodes au calcul des tunnels P2P de secours.

5.4.5 Constats

La protection multicast locale sous MPLS permet de réduire les délais de récupération des pannes tout en augmentant le taux de protection des communications multicast. Hormis la protection par forêt duale qui nécessite une signalisation des chemins de secours après la panne, nous avons décrit trois méthodes de protection multicast locale : la protection multicast un-à-un, la protection multicast par tunnels P2P de secours et la protection multicast par tunnels P2MP de secours.

La première méthode de protection offre un taux élevé de disponibilité de la bande passante mais alloue un grand nombre d'étiquettes pour les LSP de secours (nombre élevé d'états RSVP à maintenir). Cette méthode privilégie donc l'optimisation de la bande passante à la diminution du trafic de contrôle et des délais de traitement des paquets sur les routeurs MPLS. Elle est surtout efficace lorsque le nombre de LSP établis est relativement faible (par exemple, les quantités de bande réclamées par les LSP sont tellement élevées que le nombre de LSP susceptibles d'être établis est petit).

La seconde méthode de protection réduit considérablement le nombre d'étiquettes allouées dans le réseau au détriment d'une diminution des possibilités de partage de la bande passante. Lorsqu'une panne d'un nœud (resp. d'un lien) est détectée, toute la charge de duplication et de redirection des paquets sera effectuée par les nœuds qui sont voisins directs du nœud en panne (resp. par les nœud d'extrémité du lien en panne). Ceci risque de générer des goulots d'étranglement autour du nœud ou lien en panne. Par ailleurs, la protection de sessions multicast denses avec des tunnels P2P de secours conduit à une augmentation de la probabilité de réception multiple des mêmes données multicast, ce qui diminue la disponibilité de la bande passante et augmente les délais de traitement des paquets.

Enfin, la troisième méthode de protection permet d'atteindre des taux de disponibilité de la bande passante élevés et légèrement inférieurs à ceux obtenus avec la

protection multicast un-à-un, et ceci tout en réduisant le nombre d'étiquettes réservées. Comparativement à la protection multicast par tunnels P2P, cette troisième méthode de protection réduit la probabilité de réception multiple des mêmes données multicast, ce qui améliore l'utilisation de la bande passante.

Pour permettre le calcul distribué des structures de secours utilisées par ces trois dernières méthodes de protection multicast, il serait nécessaire de transmettre et/ou la diffuser une certaine quantité d'informations sur les LSP établis. Deux alternatives peuvent être adoptées : soit nous adaptons au multicast les méthodes de transmission et/ou diffusion développées pour l'unicast (cf. section 5.4.2.5, section 5.4.3.3 et section 5.4.4.3), soit nous développons de nouvelles méthodes de transmission et/ou diffusions propres au multicast. Actuellement, seule la première alternative est exploitée car les méthodes de protection multicast locale sont elles-mêmes inspirées de l'unicast.

5.5 Conclusion

Dans ce chapitre, nous avons donné un large aperçu des méthodes de protection multicast sous MPLS. Pour chacune des méthodes décrites, nous avons expliqué ses principaux principes avant de nous intéresser de près aux mécanismes permettant le calcul distribué des structures de secours qu'elle emploie. Deux stratégies de partage de la bande passante ont été appliquées à chaque méthode de protection afin de calculer la quantité de bande passante minimale à allouer aux structures de secours : (1) le partage de la bande passante restreint aux LSP de secours où seuls les chemins de secours peuvent partager la bande passante et (2) le partage de la bande passante global où le partage de la bande passante est étendu et est appliqué entre les chemins primaires et les chemins de secours.

Deux grandes familles de méthodes de protection multicast sont distinguées : la protection globale et la protection locale. La première famille de méthodes de protection fournit des délais de restauration élevés et un taux de protection réduit en présence de risques de type SRLG (en effet, quelques risques de type SRLG ne peuvent pas être protégés avec ce type de protection). Avec la seconde famille de méthodes de protection, les délais de récupération et le taux de protection sont améliorés.

Pour permettre le placement des structures de secours dans un environnement MPLS distribué, les nœuds supportant les entités de calcul doivent disposer de toute l'information permettant de déduire, sur chaque arc, la quantité de bande passante primaire cumulée et réservée, les prix de protection de tous les risques protégés et les quantités de bande passante primaire libérées suite à la panne de n'importe quel risque. Pour ce faire, trois différentes approches peuvent être adoptées :

- Transmission ciblée ou diffusion totale des quantités de bande passante primaire, des prix de protection et des quantités de bande passante libérées sur chaque arc ;
- Transmission ciblée ou diffusion totale d'une information synthétique (compressée) permettant de déduire les valeurs des quantités de bande passante primaire, des prix de protection et des quantités de bande passante libérées sur chaque arc (exemple : la diffusion des structures et propriétés des tunnels de secours) ;

– Transmission ciblée ou diffusion totale d'une information partielle permettant d'approximer les valeurs des quantités de bande passante primaire, des prix de protection et des quantités de bande passante libérées sur chaque arc.

Dans ce chapitre, seules les techniques de diffusion basées sur les deux premières approches ont été décrites. Nous notons que, souvent, ces deux approches requièrent la transmission d'une grande quantité d'informations dans le réseau, ce qui peut poser des problèmes lors du passage à l'échelle. Concernant la troisième approche, elle réduit la quantité d'informations diffusées dans le réseau aux dépens d'une diminution des possibilités de partage de la bande passante.

Pour compléter cette étude des méthodes de protection multicast, il serait judicieux de mesurer l'impact du choix de la stratégie de partage de la bande passante sur les performances des différentes méthodes de protection multicast, notamment la protection locale. Contrairement à l'unicast où la bande passante n'est libérée que sur un segment du LSP primaire, avec la protection multicast, la bande passante est libérée sur un sous-arbre du LSP primaire. L'adoption de la stratégie de partage de la bande passante global au lieu de la stratégie de partage restreint de la bande passante augmente donc la quantité moyenne de bande passante libérée dans le réseau suite à une panne, ce qui devrait diminuer le nombre moyen de requêtes de protection bloquées.

Chapitre 6

Protection locale point à multipoint : application, exigences et performances

6.1 Introduction

Dans le chapitre précédent, nous avons décrit différentes méthodes proactives de protection multicast. Nous les avons classées en deux catégories : méthodes de protection globale et méthodes de protection locale. En dépit de la taille élevée de l'ensemble des structures de secours nécessaires à la protection d'un arbre multicast avec les méthodes locales, ces dernières sont souvent préférées aux méthodes globales pour deux principales raisons :

1. Contrairement aux méthodes de protection multicast globale, les méthodes de protection multicast locale permettent de diminuer les délais de récupération puisqu'elles ne requièrent pas la notification de la panne au nœud source multicast avant de basculer le trafic vers les chemins de secours ;
2. Les méthodes de protection multicast locale permettent de protéger contre la panne de n'importe quel risque lorsque la topologie du réseau le permet, contrairement aux méthodes de protection multicast globale qui ne peuvent assurer une protection idéale dans un réseau disposant de SRLG.

En plus de la classification des méthodes de protection multicast en deux catégories, nous avons identifié deux stratégies de partage de la bande passante de secours : partage restreint de la bande passante (PRB) et partage global de la bande passante (PGB). Avec le premier type de stratégie, le partage de la bande passante est restreint aux LSP de secours, ce qui implique une dépendance complète de la quantité de bande passante de secours allouée des structures et propriétés des LSP de secours. Avec le second type de stratégie, le partage de la bande passante est étendu et est appliqué entre les LSP primaires et LSP de secours d'un côté, et entre les LSP de secours de l'autre côté.

La combinaison des deux catégories de méthodes de protection multicast avec les deux stratégies de partage de la bande passante permet d'obtenir différentes variantes

de mécanismes et techniques de placement de LSP P2MP de secours. Dans ce chapitre, nous ne nous intéressons qu'aux mécanismes permettant un placement efficace des LSP de secours dans un environnement distribué, où la protection locale est préférée et est employée pour protéger en ligne les différents LSP primaires chargés de router le trafic en l'absence des pannes.

Comme pour l'unicast, le placement de LSP P2MP de secours requiert la connaissance d'une quantité élevée d'informations parmi lesquelles nous citons (cf. chapitre précédent) : la topologie du réseau, les quantités cumulées de bande passante des LSP P2MP primaires, les quantités de bandes passante libérées par les LSP P2MP primaires suite à une panne, les prix de protection des risques, les structures et propriétés de certains LSP, etc. Afin de passer à l'échelle et pour éviter l'inondation du réseau, nous décrirons dans la deuxième section de ce chapitre une solution efficace permettant de distribuer l'information nécessaire au contrôle d'admission sur les arcs, lors de la configuration des LSP. Ensuite, nous nous focaliserons sur les algorithmes et/ou heuristiques agrégeant ou diminuant la taille de l'information nécessaire au calcul des LSP P2MP de secours. Nous verrons qu'il est possible de généraliser l'algorithme (TDRA) et les heuristiques (DBSH et PLRH) proposés initialement pour la protection unicast (cf. chapitre 3) afin de protéger efficacement les LSP P2MP primaires. Dans la troisième section de ce chapitre, nous mesurons l'impact du choix de la stratégie de partage de la bande passante sur les performances des mécanismes de placement de LSP de secours. Pour ce faire, nous utilisons plusieurs topologies de réseau et différentes matrices de trafic générées aléatoirement et nous appliquons les deux stratégies de partage PRB et PGB pour le placement des LSP de secours. Dans notre étude de performances, nous ne nous intéresserons qu'aux métriques indépendantes des implantations (des mécanismes de placement des LSP P2MP de secours) comme le taux de rejet des LSP de secours, les bandes passantes cumulées des LSP de secours établis sur chaque arc et les quantités de bande passante réservées sur les arcs pour fournir la protection (i.e. capacités de secours des arcs).

Dans la dernière section de ce chapitre, nous donnons quelques conclusions et énonçons quelques perspectives qui permettraient d'améliorer notre travail.

6.2 Annonces des paramètres requis pour le contrôle d'admission et le placement des LSP de secours

Bien que toute entité de calcul de LSP de secours vérifie les contraintes de la bande passante sur les arcs lors du calcul d'un nouveau LSP de secours, il est souvent nécessaire d'effectuer une seconde vérification des contraintes de la bande passante sur les arcs lors de la configuration de tout LSP de secours. La première vérification des contraintes de la bande passante est utile pour permettre un calcul distribué et efficace des LSP de secours par les routeurs qui sont proches des liens et nœuds à protéger. A cause des délais non nuls d'acheminement des messages de mise-à-jour de la bande passante et de la distribution des calculs de LSP de secours concurrents[1] sur plusieurs entités, cette

[1] Deux LSP de secours sont dits concurrents s'ils ne sont pas indépendants.

première vérification n'est souvent pas sûre. En conséquence, pour assurer le respect des contraintes de la bande passante à tout instant, une deuxième tâche de vérification, appelée aussi contrôle d'admission, est effectuée et est centralisée pour chaque arc sur une même entité.

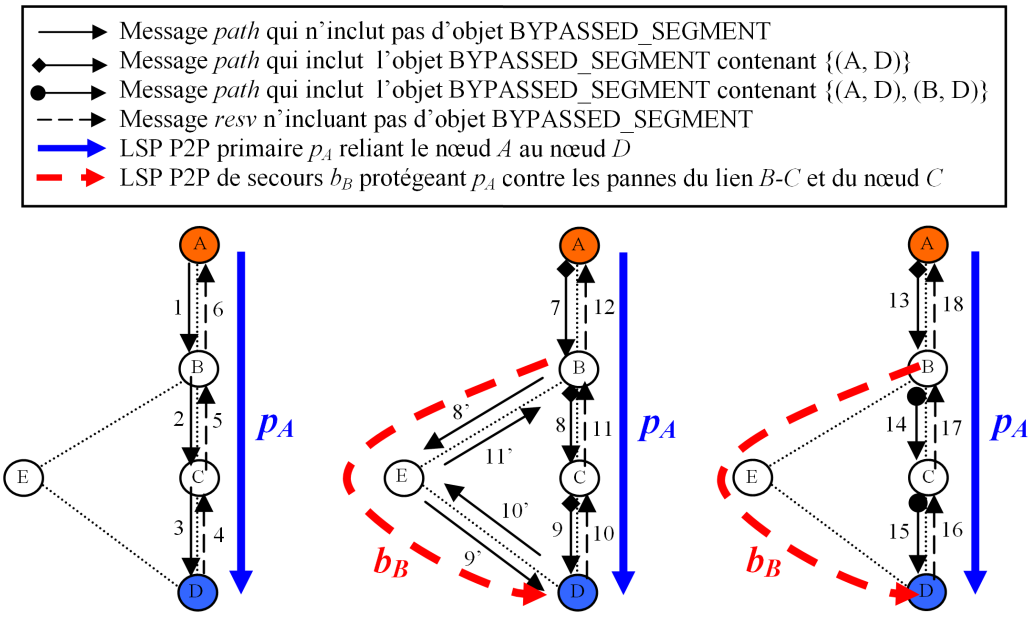
6.2.1 Information requise au contrôle d'admission

Dans un environnement distribué où la configuration de plusieurs LSP P2MP de secours concurrents et calculés séparément par plusieurs entités est autorisée, un contrôle d'admission sur tout arc doit être préconisé. Pour des considérations liées aux protocoles de signalisation et pour contrôler la quantité de trafic effectivement transmise sur un arc, le contrôle d'admission est souvent effectué pour tout arc par son nœud sortant (il peut aussi bien être effectué systématiquement par le nœud d'entrée de chaque arc). Ceci garantit le respect de contraintes de la bande passante par la sérialisation des requêtes de configuration des LSP concurrents sur chaque lien.

En fonction de la stratégie de partage de la bande passante adoptée, la tâche de contrôle d'admission peut nécessiter l'envoi d'une quantité d'informations plus ou moins élevée. Ainsi, pour un partage restreint de la bande passante (PRB), le contrôle d'admission peut être effectué sans aucun surcoût. En effet, l'information nécessaire à la vérification des contraintes de la bande passante sur un arc adjacent peut être obtenue à partir des messages des protocoles de signalisation ([ABG+01] et [PSA05]) et de routage ([KR05b] et [KR05a]). Pour rappel, cette information doit permettre le calcul des prix de protection de tous les risques, de la bande passante primaire cumulée et de l'ensemble des risques de pannes protégés. Elle peut consister, par exemple, en les structures des SRLG et propriétés des LSP (identifiant, quantité de bande passante réclamée, type de LSP, lien protégé, etc.) à configurer.

Lorsque c'est la seconde stratégie de partage de la bande passante qui est utilisée (i.e. PGB), le contrôle d'admission sur tout arc nécessitera, en plus des informations citées plus haut, la connaissance des quantités de bande passante primaire libérées sur les arcs suite à la panne de tout risque. Pour obtenir cette information, différentes solutions étendant les protocoles existants et/ou définissant de nouveaux protocoles pourraient être envisagées.

Pour des raisons de facilité d'implantation et de déploiement et afin d'atténuer le coût introduit par la transmission de l'information permettant de déduire les quantités de bande passante primaire libérées sur les arcs suite aux pannes, nous proposons ci-après une solution n'induisant que de très légères modifications des protocoles de signalisation (RSVP-TE). Dans notre approche, il suffira de définir et d'associer un nouvel objet BYPASSED_SEGMENT à tout LSP S2L appartenant au LSP P2MP primaire à protéger. Cet objet, qui doit avoir la même structure que l'objet DETOUR défini dans RSVP-TE ([PSA05]) et qui est transmis uniquement dans les messages *path* du LSP protégé, contiendra pour chaque LSP S2L primaire les identifiants de tous les nœuds d'extrémité des LSP P2P de secours qui le protègent. Ainsi, tout nœud PLR ayant fini de configurer son LSP de secours (resp. ayant échoué à déterminer un LSP de secours fournissant la protection de son lien et de son nœud en aval) ajoutera, à l'objet BY-



(a) Configuration du LSP P2P primaire p_A

(b) Configuration du LSP P2P de secours b_B et mise-à-jour de la bande passante primaire libérée suite à la panne du lien A-B et du nœud B sur les liens du LSP primaire p_A

(c) Mise-à-jour de la bande passante primaire libérée suite à la panne du lien B-C et du nœud C sur les liens du LSP primaire p_A

FIG. 6.1 – Fusion de LSP de secours et perte de la bande passante

PASSED_SEGMENT reçu, son identifiant et l'identifiant de l'autre nœud d'extrémité du LSP de secours qu'il vient de configurer (resp. l'identifiant du nœud de destination du LSP S2L primaire à protéger).

Pour faciliter la compréhension de notre proposition, nous expliquons ci-après ses principes à travers un exemple d'application au protocole RSVP-TE. Sur la figure 6.1 (a), le nœud A déclenche la configuration du LSP primaire p_A ($A \rightarrow B \rightarrow C \rightarrow D$) en envoyant un message *path* à son nœud voisin B. Ce message sera reçu par le nœud B qui le traitera et l'enverra à son prochain nœud C sur le LSP primaire p_A. A la réception du message *path*, le nœud C effectuera les mêmes traitements que ceux accomplis par le nœud B avant de le transmettre à son prochain nœud D. Lorsque le message *path* atteint le nœud de destination D, ce dernier répondra à son nœud en amont C, après avoir vérifié les contraintes de la bande passante sur son arc en amont. Ainsi, le nœud de destination D enverra un message *resv* au nœud C afin de réserver les ressources sur l'arc $C \rightarrow D$. A la réception du message *resv*, le nœud C (resp. le nœud B) le traite, effectue un contrôle d'admission sur l'arc situé en amont. Ensuite, le nœud C (resp. le nœud B) ré-achemine le message *resv* obtenu après traitement à son nœud en amont B (resp. A) afin de réserver les ressources sur l'arc $B \rightarrow C$ (resp. l'arc $A \rightarrow B$). La configuration du LSP primaire finit avec succès lorsque le nœud A reçoit un message *resv* du

nœud B. Nous notons que ces étapes de configuration du LSP primaire p_A (étapes 1 à 6 sur la figure 6.1 (a)) suivent les mêmes règles et procédures que celles décrites dans [ABG+01] et [APY07].

A la fin de la configuration du LSP primaire p_A, le nœud A (ayant conservé un état RSVP-TE pour le LSP p_A) déduit que p_A requiert une protection et initie le calcul d'un LSP de secours permettant de le protéger contre les pannes du nœud B et du lien A-B. Ne déterminant aucun LSP de secours capable de protéger p_A, le nœud A construira un objet BYPASSED_SEGMENT contenant son identifiant et l'identifiant du nœud de destination D du LSP primaire p_A (i.e. BYPASSED_SEGMENT = {entête identifiant l'objet, (A, D)}). Cet objet sera conservé et acheminé dans tous les messages *path* de rafraîchissement envoyés par le nœud A et les nœuds situés à son aval sur le LSP primaire p_A (étapes 7, 8 et 9 sur la figure 6.1 (b)). Il ne sera pas non plus modifié tant que les nœuds parcourus n'ont pas fini de configurer leur LSP de secours (ou ont déterminé qu'il n'y a aucun LSP de secours capable de fournir la protection désirée).

Lorsque le prochain nœud B finit de configurer le LSP de secours $b_B = B \rightarrow E \rightarrow D$ (i.e. les étapes 8', 9', 10' et 11' sur la figure 6.1 (b) ont été accomplies avec succès), il ajoutera (étape 14 sur la figure 6.1 (c)) à l'objet BYPASSED_SEGMENT reçu son identifiant et l'identifiant du nœud D qui est le nœud de fusion du LSP de secours b_B avec le LSP primaire p_A (i.e. BYPASSED_SEGMENT = {entête identifiant l'objet, (A, D), (B, D)}). Ce dernier traitement sera effectué par tous les nœuds intermédiaires, situés entre le nœud courant (i.e. le nœud B) et le nœud situé à l'amont du nœud de destination du LSP primaire (i.e. le nœud C).

Bien évidemment, pour que le contrôle d'admission puisse exploiter efficacement la bande passante primaire libérée, tout nœud recevant un message *path* contenant l'objet BYPASSED_SEGMENT doit le traiter en mettant à jour ses tables des quantités de bande passante primaire libérées sur les arcs dont il est le nœud sortant. Par exemple, lorsque le nœud D reçoit (pour la première fois) l'objet BYPASSED_SEGMENT contenant les identifiants suivants (A, D), (B, D), il incrémentera, de la quantité de bande passante bw (p_A) réclamée par le LSP primaire p_A, les quantités de bande passante primaire libérées sur l'arc $C \rightarrow D$ suite aux pannes des risques contenant les liens et nœuds en aval des PLR A et B (i.e. les risques B, A-B, C et $B \rightarrow C$).

Comme nous le constatons, notre approche ne nécessite que de très légères extensions aux protocoles de signalisation pour fonctionner. Elle a l'avantage de minimiser l'augmentation de la taille des messages de signalisation tout en conservant leur fréquence d'envoi. Cependant, elle introduit des délais supplémentaires pour prendre en charge les quantités de bande passante primaire libérées. Ceci n'est pas très gênant surtout si l'on considère des systèmes où le temps d'inter-arrivées des LSP est assez élevé par rapport à ces délais.

D'autres approches pourraient être appliquées pour le contrôle d'admission qui tient compte du partage global de la bande passante. Par exemple, [BML06] propose d'inclure les quantités de bande passante primaire libérées suite aux pannes simples directement dans les messages *resv*. Contrairement à notre approche, cette solution pourrait induire le rejet des configurations de certains LSP, lors de la présence des SRLG dans le réseau.

6.2.2 Annonces des paramètres nécessaires au calcul distribué des LSP de secours

Afin de diminuer le risque de rejet de la configuration des LSP lors du contrôle d'admission, il serait judicieux de doter le réseau de mécanismes permettant aux entités de calcul de sélectionner localement les arcs vérifiant les contraintes de la bande passante à chaque calcul d'un nouveau LSP de secours. De la même façon que pour l'unicast, ces mécanismes doivent annoncer dans le réseau une information permettant de déduire ou d'approximer, sur chaque arc, tous les paramètres de la bande passante nécessaires au contrôle d'admission (comme la bande passante primaire cumulée, les prix de protection de tous les risques et les quantités de bande passante primaire libérées suite aux différentes pannes possibles, etc.).

Comme nous l'avons indiqués dans les chapitres précédents, une implantation efficace des mécanismes distribuant l'information nécessaire au calcul des LSP de secours passe par une agrégation ou une diminution de la taille de cette information. Vu l'efficacité des mécanismes de distribution de l'information nécessaire au calcul des LSP de secours proposés pour l'unicast et étant donné la similarité des procédures de vérification des contraintes de la bande passante entre l'unicast et le multicast, il serait judicieux de réadapter les solutions unicast au multicast. En plus de l'information annoncée dans les protocoles de routage (bandes passantes primaires cumulées, capacités de tous les arcs, structures des risques, etc.) et les protocoles de signalisation (structures et propriétés des LSP traversant les liens adjacents à un nœud), de nouvelles modifications et/ou extensions des mécanismes proposés pour l'unicast (i.e. l'algorithme TDRA, l'heuristique DBSH et l'heuristique PLRH décrits dans le chapitre 3) pourraient être nécessaires afin de les étendre au multicast. Ces dernières sont explicitées dans les sous-sections suivantes.

6.2.2.1 Algorithme TDRA

Selon la stratégie de partage utilisée, les modifications et/ou extensions nécessaires aux mécanismes distribuant l'information nécessaire au calcul des LSP de secours avec l'algorithme TDRA peuvent être plus ou moins simples. Ainsi, lorsque le partage de la bande passante est restreint aux LSP de secours, il suffira de transmettre pour chaque lien u-v du LSP P2MP primaire, les structures et les propriétés des LSP de secours qui le protègent, à tous les nœuds d'extrémité des liens appartenant aux SRLG contenant le lien protégé u-v. De cette manière, tout nœud du réseau pourra déduire les prix de protection de tous les risques contenant le nœud lui-même ou un de ses liens adjacents. En se servant de ces prix de protection, des quantités de bande passante primaire annoncées par les protocoles de routage et des structures et propriétés des LSP S2L primaires à protéger (déduites de l'information transmise par les protocoles de signalisation), tout nœud pourra calculer les LSP de secours le protégeant lui-même et/ou protégeant un de ses liens adjacents (cf. section 3.2).

Lorsque le partage de la bande passante est global, l'information annoncée dans le réseau doit être étendue pour tenir compte des quantités de la bande passante primaire

libérées suite aux pannes des SRLG. Ainsi, tout nœud ayant calculé un nouveau LSP de secours protégeant contre la panne d'un lien u-v, doit ajouter à l'information envoyée aux nœuds d'extrémité des liens appartenant aux SRLG contenant le lien u-v, la structure de la partie du LSP primaire contournée par le LSP de secours déterminé.

6.2.2.2 Heuristiques DBSH et PLRH

L'heuristique DBSH (resp. PLRH) approxime les prix de protection des SRLG (resp. les prix de protection de tous les risques de panne) en diffusant dans le réseau, pour chaque arc λ, les x^λ plus grands prix de SRLG (resp. les y^λ plus grands prix de protection) et les risques associés.

Lorsque le partage de la bande passante est restreint aux LSP de secours, les heuristiques DBSH et PLRH ne nécessitent aucune nouvelle extension et/ou modification (hormis celles décrites dans les sections 3.3.2 et 3.4.2) pour permettre la protection de LSP P2MP primaires. En effet, la diffusion des vecteurs $\{x^\lambda\_\, vecteur\}_{\lambda \in E}$ (resp. vecteurs $\{y^\lambda\_\, vecteur\}_{\lambda \in E}$) permet de déterminer ou d'approximer tous les prix de protection des risques en utilisant l'heuristique DBSH (resp. l'heuristique PLRH). En combinant ces prix de protection avec l'information transmise par les protocoles de routage et les protocoles de signalisation étendus (cf. chapitre 3), les contraintes de la bande passante pourraient être vérifiées sur chaque arc susceptible de supporter un nouveau LSP de secours.

Lorsque le partage de la bande passante est global, une légère modification doit être apportée à l'information contenue dans les vecteurs $\{x^\lambda\_\, vecteur\}_{\lambda \in E}$ (resp. vecteurs $\{y^\lambda\_\, vecteur\}_{\lambda \in E}$) transmis avec l'heuristique DBSH (resp. avec l'heuristique PLRH). Concrètement, l'information incluse dans les vecteurs $\{x^\lambda\_\, vecteur\}_{\lambda \in E}$ (resp. vecteurs $\{y^\lambda\_\, vecteur\}_{\lambda \in E}$) diffusés dans le réseau avec DBSH (resp. PLRH) n'est pas appropriée pour tenir compte des quantités de la bande passante primaire libérées suite à une panne puisqu'elle n'inclut que les prix de protection (et risques associés). Pour un partage étendu, nous proposons de transmettre dans chaque vecteur $x^\lambda\_\, vecteur_{\lambda \in E}$ (resp. vecteur $y^\lambda\_\, vecteur_{\lambda \in E}$), les x^λ (resp. les y^λ) plus hautes valeurs (supérieures au seuil Sc^λ) des différences entre les prix de protection et les bandes passantes primaires libérées pour un même risque de panne.

Risques / Valeurs	u	u-v	$srlg_1$	$srlg_2$	$srlg_3$
Prix de protection	0	20	**80**	**70**	**65**
Bande primaire libérée suite à une panne	40	0	35	10	0
Différence	-40	20	45	**60**	**65**

TAB. 6.1 – Information requise au contrôle d'admission

Exemple.

Considérons la table 6.1 correspondant à l'information permettant au nœud sortant de l'arc λ d'effectuer le contrôle d'admission sur ce dernier.

Si nous utilisons l'heuristique PLRH pour le calcul des LSP de secours et si nous fixons la valeur de y^{λ} (resp. Sc^{λ}, C^{λ}) à 2 (resp. à 50 unités, 100 unités) sur l'arc λ, le nœud sortant à l'arc λ transmettra le vecteur suivant à la prochaine annonce :
$y^{\lambda}\_vecteur = [(srlg_1, 80), (risque\_générique, 70)]$ si le partage de la bande passante est restreint aux LSP de secours.
$y^{\lambda}\_vecteur = [(srlg_3, 65), (srlg_2, 60)]$ si le partage de la bande passante est global.

En appliquant l'algorithme 6, nous pourrons approximer tous les risques protégés par l'arc λ à partir du vecteur $y^{\lambda}\_vecteur$ diffusé. Ainsi, tout calcul d'un nouveau LSP de secours réclamant une quantité de bande passante inférieure ou égale à 30 unités (resp. 50 unités) inclura ou exclura l'arc λ sans erreur, si le partage de la bande passante est restreint (resp. si le partage de la bande passante est global).

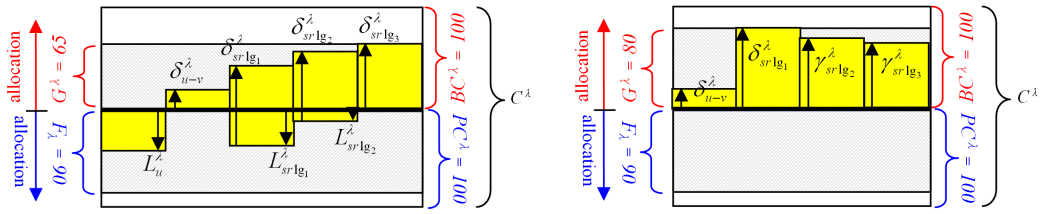
6.3 Impact du choix de la stratégie de partage sur les performances de la protection multicast un-à-un

Dans la section précédente, nous avons vu que l'application de la stratégie du partage global de la bande passante (PGB) au lieu du partage restreint de la bande passante (PRB) nécessiterait de nouvelles extensions aux protocoles de signalisation et/ou protocoles de routage. Ces extensions permettent d'augmenter le taux de partage au détriment d'une augmentation du trafic de contrôle dû à la croissance du volume de l'information transmise dans le réseau (pour calculer les LSP de secours et effectuer le contrôle d'admission qui tient compte du partage global de la bande passante).

Dans cette section, nous mesurons par simulations l'impact du choix de la stratégie de partage de la bande passante sur les performances de la protection multicast un-à-un. Concrètement, en évitant la perte de la bande passante (due à l'approximation de certains paramètres de la bande passante) et en choisissant des métriques indépendantes des implantations, nous allons quantifier le gain de performances obtenu par l'adoption de la stratégie PGB à la place de la stratégie PRB.

Dans nos simulations, nous avons opté pour la protection multicast locale un-à-un pour diverses raisons :

1. La protection multicast locale un-à-un permet de minimiser les allocations de la bande passante de secours et elle réduit davantage le taux de blocage des requêtes de protection en la comparant aux méthodes de protection par tunnel de secours ;
2. La protection multicast locale un-à-un augmente les quantités de bande passante libérées sur les arcs par rapport aux méthodes de protection locale par tunnel de secours. Par conséquent, le gain en performances obtenu en appliquant la stratégie PGB au lieu de la stratégie PRB est plus élevé avec la protection multicast locale un-à-un (par rapport à la protection multicast local par tunnel de secours).



(a) Allocation de la bande passante avec la stratégie de partage global (PGB)

(b) Allocation de la bande passante avec la stratégie de partage restreint (PRB)

FIG. 6.2 – Modèle d'allocation de la bande passante

6.3.1 Modèle d'allocation de la bande passante

Pour nous concentrer uniquement sur l'impact du choix de la stratégie de partage sélectionnée sur les performances de la protection multicast un-à-un, nous divisons la capacité C^λ de tout arc λ en deux pools (figure 6.2) : pool de bande passante primaire de capacité PC^λ et pool de bande passante de secours de capacité BC^λ. Le premier pool est utilisé pour allouer la bande passante aux LSP primaires alors que le second pool est exclusivement utilisé pour allouer la bande passante aux LSP de secours. En conséquence, quelle que soit la stratégie de partage de la bande passante adoptée, la bande passante cumulée de tous les LSP primaires $F_\lambda\ (n)_{\forall k}$ traversant l'arc λ doit être inférieure ou égale à la capacité du pool primaire PC^λ. De même, la bande passante allouée pour configurer tous les LSP de secours $G^\lambda\ (n)_{\forall k}$ traversant l'arc λ doit être inférieure ou égale à la capacité du pool de secours BC^λ.

Avec l'adoption de ce modèle pour le placement des LSP P2MP, nous déduisons qu'un arc λ peut être utilisé pour établir un LSP primaire p_{k+1} couvrant la source et les membres multicast de la session $k+1$ si et seulement si :
$F_\lambda\,(k) + F^{k+1} \leq PC^\lambda$

Lorsque c'est la stratégie PGB qui est appliquée, l'arc λ pourra être sélectionné pour établir un nouveau LSP P2MP de secours b qui protège le LSP primaire p_{k+1} si et seulement si :
$Max_{r\,\in\,PFRG(b)}\,(\delta^\lambda_r\,(k) - L^\lambda_r\,(k)) + F^{k+1} \leq BC^\lambda$

En appliquant la stratégie PGB à l'exemple de la figure 6.2 (a) (la figure 6.2 (a) reproduit les valeurs des paramètres de la bande passante de l'arc λ illustrées dans la table 6.1), nous déduisons que tout nouveau LSP de secours b protégeant contre les risques de panne u, u-v, et $srlg_2$ ne peut utiliser l'arc λ que s'il réclame une quantité de bande passante inférieure ou égale à 40 unités. En effet :
$Max_{r\,\in\,\{u,\,u\text{-}v,\,srlg_2\}}\,(\delta^\lambda_r\,(k) - L^\lambda_r\,(k)) + F^{k+1} \leq BC^\lambda \Leftrightarrow 60 + F^{k+1} \leq 100 \Leftrightarrow F^{k+1} \leq 40$

Lorsque c'est la stratégie PRB qui est choisie, l'arc λ pourra être sélectionné pour établir un nouveau LSP P2MP de secours b qui protège le LSP primaire p_{k+1} si et seulement si :
$Max_{r\,\in\,PFRG(b)}\,\delta^\lambda_r\,(k) + F^{k+1} \leq BC^\lambda$

En appliquant la stratégie PRB à l'exemple de la figure 6.2 (b) (la figure 6.2 (b) reproduit les valeurs des paramètres de la bande passante de l'arc λ illustrées dans la

table 6.1), nous déduisons que tout nouveau LSP de secours b protégeant contre les risques de panne u, u-v, et $srlg_2$ ne peut utiliser l'arc λ que s'il réclame une quantité de bande inférieure ou égale à 30 unités. En effet :
$Max_{r \in \{u,\, u\text{-}v,\, srlg_2\}}\, \delta_r^{\lambda}(k) + F^{k+1} \leq BC^{\lambda} \Leftrightarrow 70 + F^{k+1} \leq 100 \Leftrightarrow F^{k+1} \leq 30$

6.3.2 Métriques de comparaison

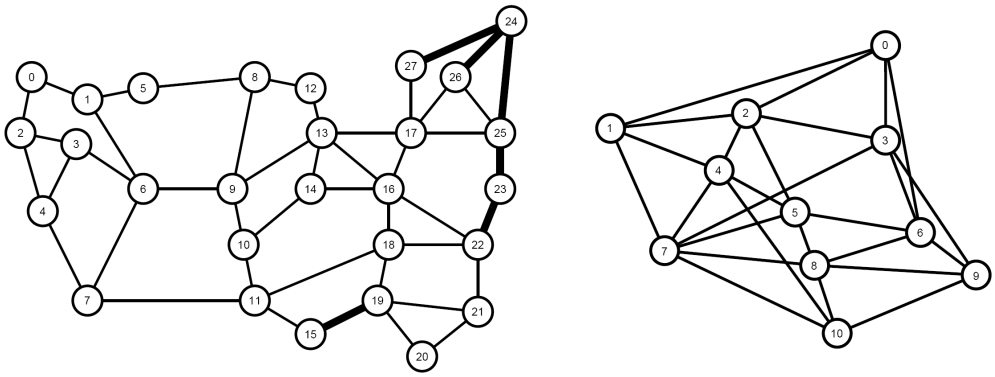
Afin de quantifier l'impact du choix de la stratégie de partage de la bande passante (PGB ou PRB) sur la protection multicast un-à-un, nous avons sélectionné les métriques suivantes : taux de rejet des LSP P2MP de secours (*RP2MP*), taux de rejet des LSP P2P de secours (*RP2P*), taux d'utilisation de la bande passante de secours (*TUBS*) et le taux d'occupation des pools de secours (*TOPS*).

La première métrique *RP2MP* correspond au taux moyen de rejet des requêtes de protection des nœuds et liens des LSP *P2MP* primaires. Elle est déterminée comme le rapport entre le nombre de requêtes de protection des nœuds et liens des LSP P2MP primaires qui sont partiellement rejetées sur le nombre total de requêtes de protection. Nous notons qu'une requête de protection d'une session multicast contre la panne d'un nœud (resp. d'un lien) est dite partiellement rejetée s'il existe au moins un nœud membre de la session qui n'est pas protégé contre la panne du nœud précédent (resp. lien précédent). Par conséquent, cette métrique permet de mesurer, suite à une panne simple, la sûreté des communications multicast où le départ d'un (deux, trois, etc.) membre peut provoquer l'arrêt des applications supportées. Ainsi, plus la valeur de cette métrique est basse pour une session multicast donnée, plus sont sûres les applications (coopératives) supportées.

La seconde métrique *RP2P* mesure le taux de connectivité des nœuds membres multicast au nœud source multicast, après une panne. Elle est calculée comme le rapport entre le nombre de requêtes d'établissement de LSP P2P de secours rejetées sur le nombre total de requêtes d'établissement de LSP P2P de secours. Plus cette métrique est basse, plus de membres sont desservis en moyenne après une panne.

La troisième métrique *TUBS* mesure l'efficacité du procédé d'allocation de la bande passante de secours. C'est une métrique permettant d'estimer efficacement le taux de partage de la bande passante de secours sur les différents arcs de la topologie du réseau. Elle est déterminée comme le rapport entre la moyenne des quantités de bande passante cumulée des LSP P2MP de secours sur les arcs et la capacité de secours moyenne. Formellement (u-v et λ ne doivent partager aucun SRLG) :
$TUBS = \sum_{(\lambda,\, u\text{-}v) \setminus (\lambda \in E) \wedge (u \to v \in E \setminus \{\lambda\} \wedge v \to u \in E \setminus \{\lambda\})} \delta_{u\text{-}v}^{\lambda}(k) \,/\, \sum_{\lambda \in E} BC^{\lambda}$
où k est le nombre de sessions multicast.

Enfin, la quatrième métrique *TOPS* détermine le taux moyen d'occupation des pools de secours. Elle permet de déduire la quantité moyenne de bande passante réellement allouée sur les arcs afin de rendre la fonctionnalité de protection. Elle est calculée comme suit (k est le nombre de sessions multicast) :
$TOPS = \sum_{\lambda \in E} G^{\lambda}(k) \,/\, \sum_{\lambda \in E} BC^{\lambda}$, avec :
$G^{\lambda}(k) = Max_{r \in R}\,(\delta_r^{\lambda}(k) - L_r^{\lambda}(k),\, 0)$ lorsque le partage de la bande passante est global,
$G^{\lambda}(k) = Max_{r \in R}\, \delta_r^{\lambda}(k)$ lorsque le partage de la bande passante est restreint.



(a) Allocation de la bande passante avec la stratégie de partage de bande passante global

(b) Allocation de la bande passante avec la stratégie de partage de bande passante restreint

FIG. 6.3 – Modèle d'allocation de la bande passante

6.3.3 Environnement de simulation

Pour mesurer l'impact du choix de la stratégie de partage sur le performances de la protection multicast locale un-à-un, nous avons utilisé deux topologies de réseau (figure 6.3), très couramment utilisées dans la littérature. La première topologie de réseau (topologie du réseau Long Haul), illustrée sur la figure 6.2 (a), est composée de 28 nœuds et 45 liens bidirectionnels (83 risques de panne). Les capacités des pools de protection sont égales à 600 unités sur les liens en gras et à 200 unités sur les autres liens (dans les deux sens). Cette topologie de réseau est d'une taille relativement élevée et a un degré de connectivité moyen et égal à 3.21. La seconde topologie de réseau (topologie du réseau Cost 239), illustrée sur la figure 6.2 (b), est constituée de 11 nœuds et 26 liens bidirectionnels (37 risques de panne). C'est une topologie de réseau de petite taille qui est fortement connectée. En effet, le degré moyen de ses nœuds est égal à 4.73. Sur tous les liens (dans les deux sens) de cette topologie de réseau, les pools de protection ont une capacité égale à 200 unités.

Différents scenarios de tests ont été considérés lors de la génération des matrices de trafic. Pour chaque scénario de test, nous avons créé 1000 demandes de connexions multicast arrivant séquentiellement l'une après l'autre, toutes associées à des couples formés d'une source multicast et d'un groupe multicast dont la taille varie de 1 (unicast) à 27 (tous les nœuds différents de la source sont alors des membres multicast) pour la topologie du réseau Long Haul et de 1 (unicast) à 10 (tous les nœuds différents de la source multicast sont alors des membres multicast) pour la topologie du réseau Cost 239 (les sources multicast ne sont pas membres multicast). Nous notons que tous les groupes multicast utilisés dans un même scénario de test sont de tailles égales. A l'arrivée de chaque demande de connexion multicast, un LSP P2MP primaire couvrant le groupe et la source multicast correspondants est calculé en utilisant l'algorithme des plus courts chemins de Dijkstra (afin d'optimiser le nombre de sauts entre chaque membre multicast et la source multicast). Nous avons choisi cet algorithme car il est très connu, performant et couramment utilisé.

Pour chacune des deux stratégies de partage de bande (PGB et PRB) employée, un ensemble de LSP P2MP de secours, permettant de protéger le dernier LSP P2MP primaire déterminé contre les pannes de ses liens et nœuds internes, est calculé. Pour ce faire, deux types de contrôle d'admission sont effectués sur les liens de la topologie originale du réseau : le premier contrôle d'admission permet d'obtenir une première topologie de réseau réduite constituée des liens vérifiant les contraintes de la bande passante en appliquant la stratégie PGB alors que le second contrôle d'admission permet d'obtenir une seconde topologie de réseau réduite composée des liens respectant les contraintes de la bande en appliquant la stratégie PRB (cf. section 6.3.1). Ensuite, le même algorithme de calcul (i.e. l'algorithme CSPF qui est très répandu et qui est décrit à la fin de la section refsection1.3.1) de LSP P2MP de secours est lancé séparément sur les deux topologies de réseau réduites pour déterminer les deux LSP P2MP fournissant la protection avec les deux stratégies de partage de bande. Nous notons que tous les LSP P2MP de secours calculés relient un nœud PLR à l'ensemble des nœuds membres situés en aval du lien et nœud protégés sur le LSP P2MP primaire. De plus, ces LSP P2MP de secours sont constitués des plus courts chemins (en termes du nombre de sauts) dans les topologies de réseau réduites.

Pour la pertinence des résultats, nous avons choisi d'appliquer les deux stratégies de partage PGB et PRB pour la protection des mêmes ensembles de LSP P2MP primaires. Pour ce faire, nous avons adopté le modèle d'allocation de la bande passante, décrit dans la section 6.3.1, dans lequel la quantité de bande passante allouée aux LSP primaires sur chaque arc est indépendante de la stratégie de partage de bande passante employée. Pour éviter les effets de bord, nous avons aussi initialisé les capacités primaires des liens à des valeurs suffisamment élevées pour accepter toutes les requêtes d'établissement des LSP P2MP primaires. Nous notons que toutes les requêtes d'établissement des LSP primaires et des LSP de secours les protégeant arrivent en ligne et réclament des quantités de bande passante uniformément distribuées entre 1 et 10 unités.

Afin de tenir compte de la charge du réseau dans nos simulations, nous avons fait varier le nombre de LSP P2MP primaires. Ainsi, à chaque établissement de 20 LSP P2MP primaires, nous retournons les valeurs des quatre métriques décrites dans la section 6.3.2 (*RP2MP*, *RP2P*, *TUBS* et *TOPS*) pour toutes les tailles des groupes multicast.

6.3.4 Résultats et analyse

Nous notons que les valeurs des métriques représentées sur toutes les figures de cette section correspondent aux moyennes des résultats de 100 tests.

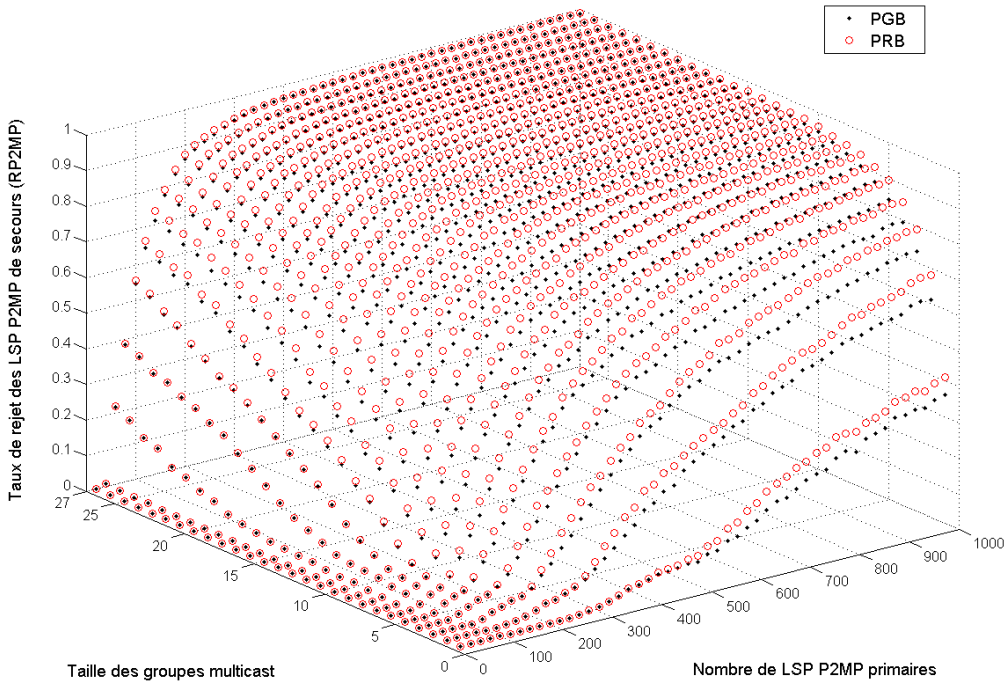


FIG. 6.4 – Taux de rejet des LSP P2MP de secours (RP2MP) dans le réseau Long Haul

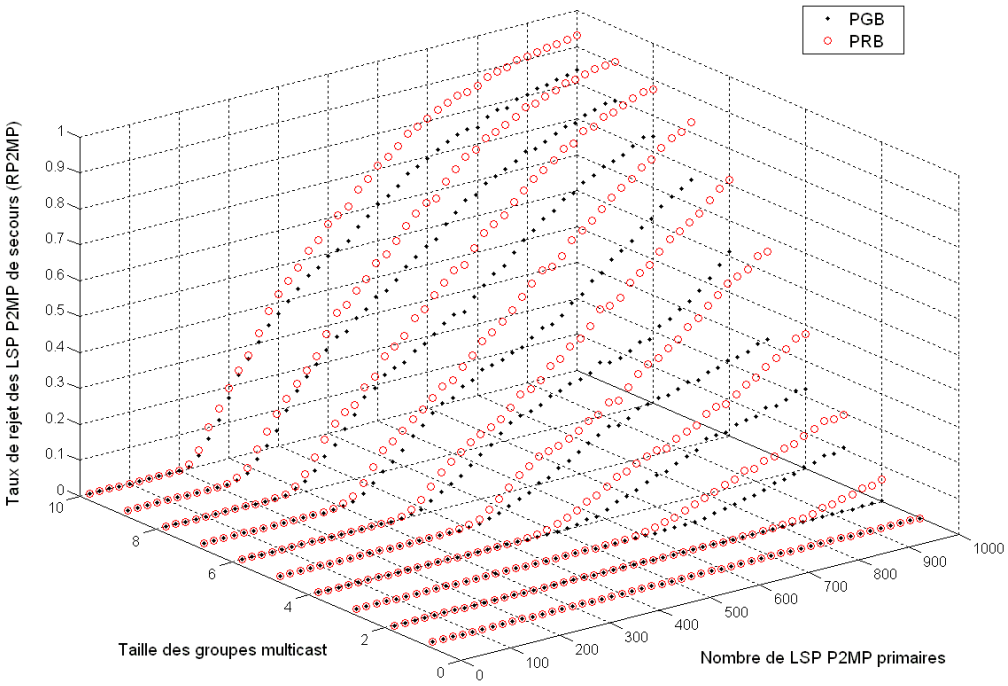


FIG. 6.5 – Taux de rejet des LSP P2MP de secours (RP2MP) dans le réseau Cost 239

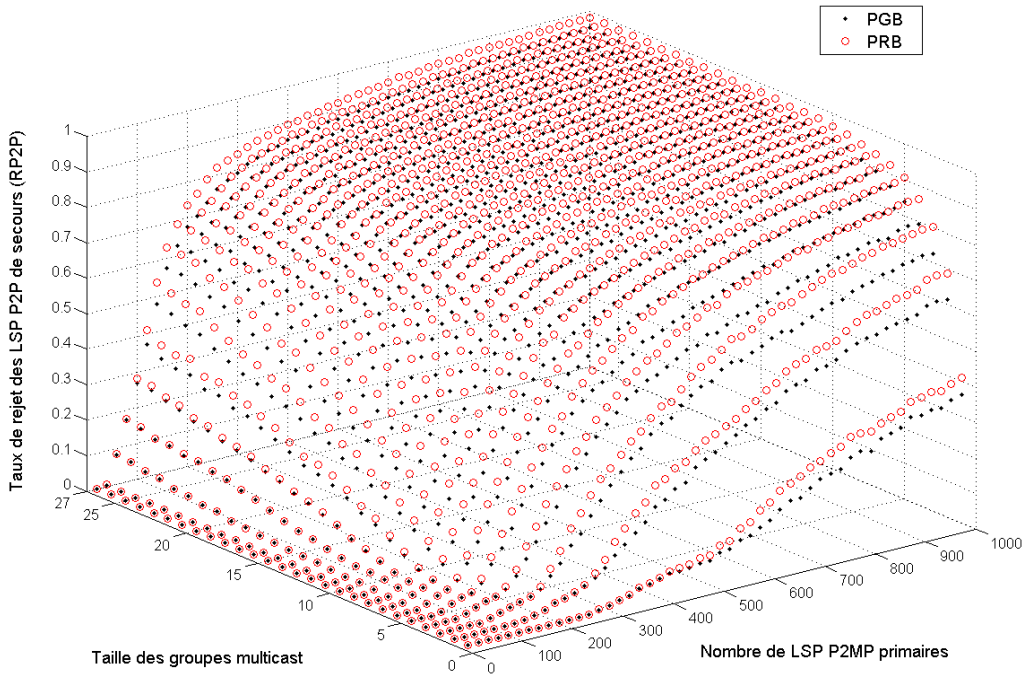


FIG. 6.6 – Taux de rejet des LSP P2P de secours (RP2P) dans le réseau Long Haul

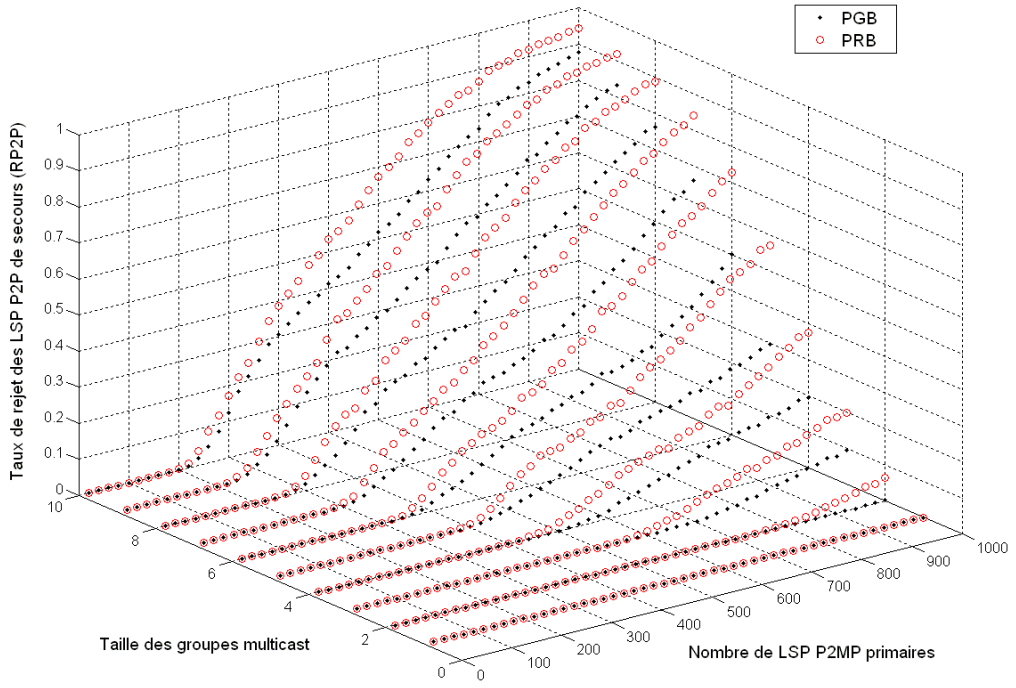


FIG. 6.7 – Taux de rejet des LSP P2P de secours (RP2P) dans le réseau Cost 239

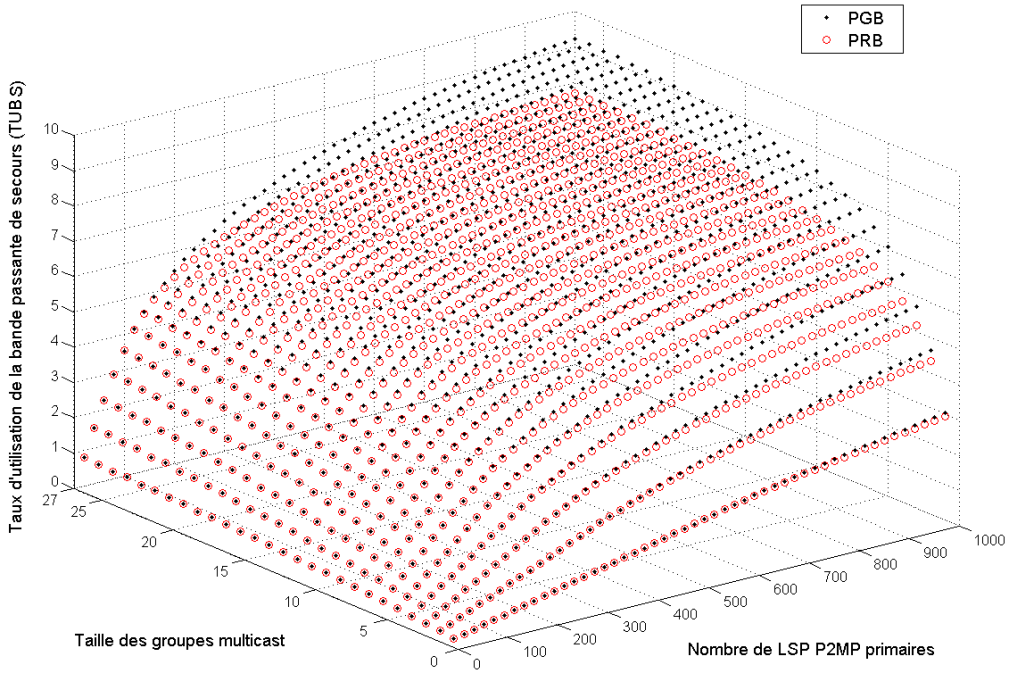


FIG. 6.8 – Taux d'utilisation de la bande passante de secours (*TUBS*) dans le réseau Long Haul

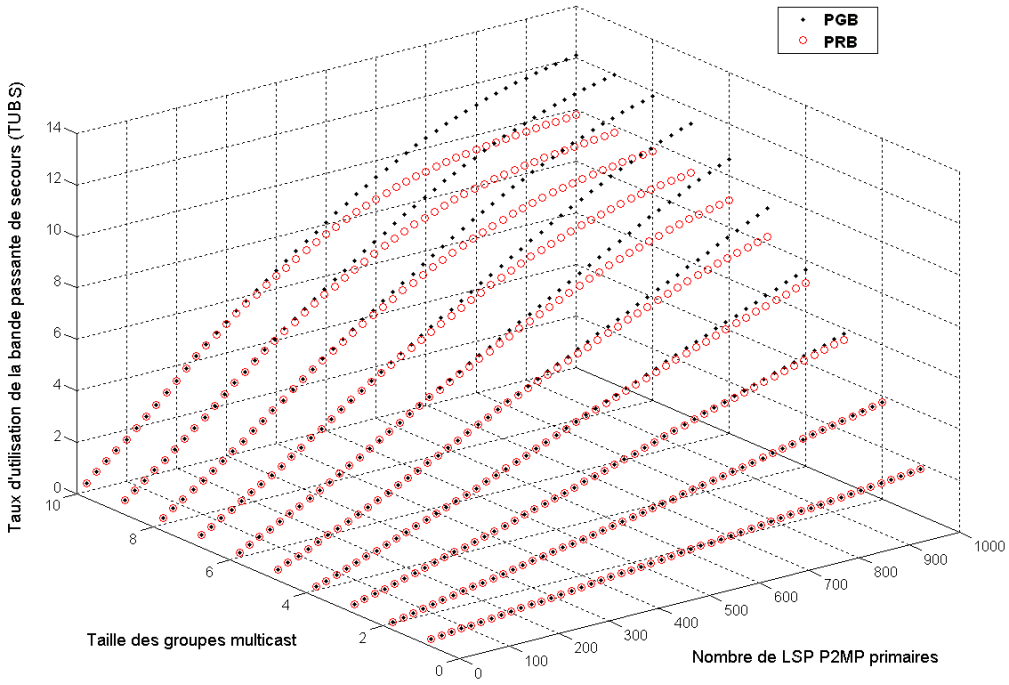


FIG. 6.9 – Taux d'utilisation de la bande passante de secours (*TUBS*) dans le réseau Cost 239

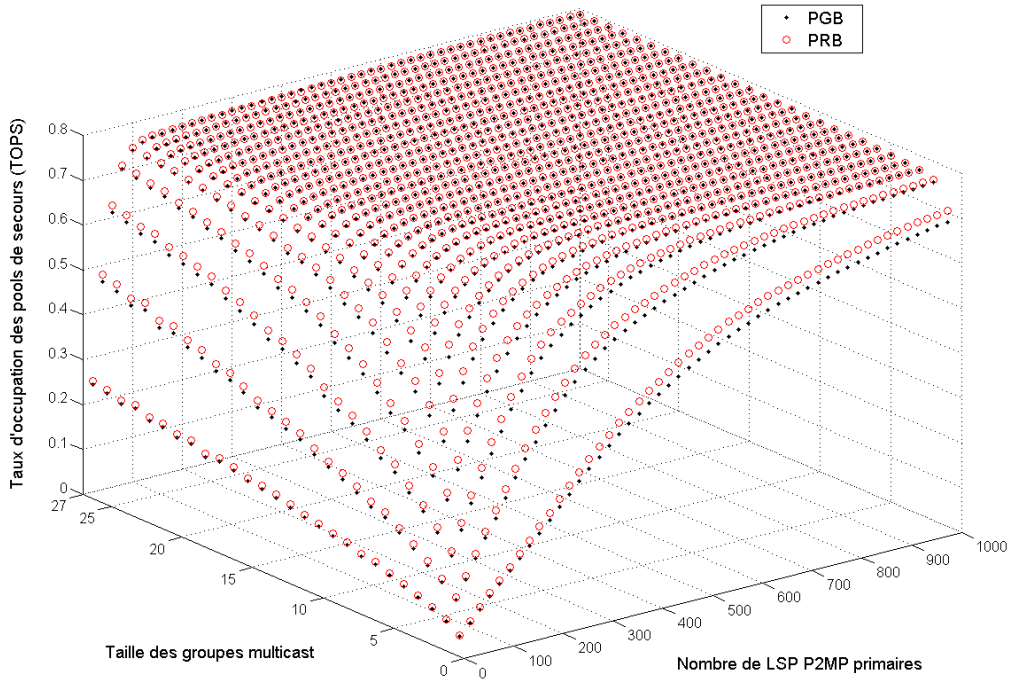


FIG. 6.10 – Taux d'occupation des pools de secours (*TOPS*) dans le réseau Long Haul

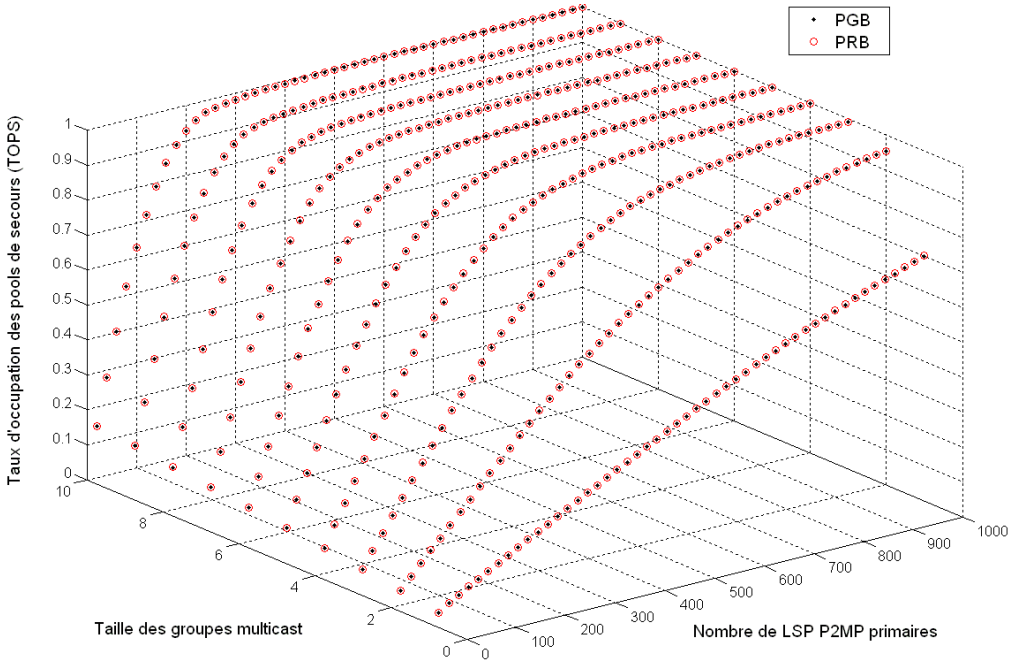


FIG. 6.11 – Taux d'occupation des pools de secours (*TOPS*) dans le réseau Cost 239

Les figures 6.4 et 6.5 montrent les taux de rejet des LSP P2MP de secours obtenus respectivement dans les réseaux Long Haul et Cost 239, en fonction du nombre de LSP P2MP primaires établis et de la taille des groupes multicast. Comme il apparaît sur ces deux figures, les taux de rejet des LSP P2MP de secours obtenus, pour toutes les tailles des groupes multicast avec les deux stratégies de partage (PGB et PRB), passent successivement par trois phases lorsque la charge du réseau (nombre de LSP P2MP primaire) augmente : phase de satisfaction de toutes les requêtes de protection, phase de disparité et phase de saturation.

Dans la première phase, caractérisée par une charge du réseau faible, les taux de rejet des LSP P2MP de secours des stratégies PGB et PRB sont similaires et très proches de zéro. Dans cette phase, les prix de protection de tous les risques sont faibles, ce qui permet de protéger entièrement (presque tous) les LSP P2MP primaires établis contre les pannes de leurs nœuds et liens internes.

Après la première phase, les valeurs des taux de rejet des LSP P2MP de secours des deux stratégies PGB et PRB commencent à se distinguer rapidement les unes des autres pour diverger ensuite : on parle alors de la phase de disparité. Dans cette phase, le taux de rejet des LSP P2MP de secours de la stratégie PGB est statistiquement inférieur à celui obtenu avec la stratégie PRB. Ceci s'explique par la qualité de partage employée par PGB qui est meilleure et plus efficace que celle utilisée par PRB. En effet, l'utilisation dans PGB de la bande passante primaire libérée suite à une panne (en plus du partage de bande entre les LSP P2MP de secours) permet de réduire les coûts de protection (*effectifs*) sur certains liens (surtout, les liens adjacents aux PLR), ce qui donne plus de flexibilité dans le choix des LSP de secours.

Lorsque la charge du réseau est très élevée, la phase de saturation, dans laquelle le taux de rejet est très proche de 1 (i.e. presqu'aucun risque de panne ne peut être protégé), sera entamée. Nous notons que le passage par cette phase n'est pas obligatoire lorsque les capacités primaires ne sont pas saturées (i.e. lorsque les capacité primaires sont infinies). En effet, la combinaison de la stratégie PGB avec certains algorithmes de calcul des LSP P2MP primaires pourrait entraîner, suite à une panne d'un risque r, la libération de la bande passante sur un ensemble de liens capables de router un LSP de secours protégeant contre la panne du même risque r. Par exemple, sur la figure 6.4, la phase de saturation est franchie pour les groupes multicast de taille moyenne ou élevée (taille des groupes multicast > 15) lorsque le nombre de LSP P2MP primaires est élevé (500 LSP P2MP primaires) ; elle ne l'est pas par contre sur la figure 6.5 où le nombre de LSP P2MP primaires n'est pas suffisamment élevé pour provoquer la saturation des capacités de secours. Nous notons et disons que les capacités de secours sont saturées lorsque les valeurs des coûts de protection des risques de panne sur les liens sont distribuées de telle sorte que presqu'aucun nouveau LSP de secours ne peut être établi.

Enfin, en comparant les résultats illustrés sur la figure 6.4 avec ceux de la figure 6.5, nous apercevons que la différence entre les taux de rejet des LSP P2MP de secours des stratégies PGB et PRB obtenue sur la topologie du réseau Long Haul est plus petite que celle obtenue sur la topologie du réseau Cost 239. Ceci est dû à la faible connectivité et à l'asymétrie de la topologie du réseau Long Haul où la portion des LSP P2MP de

secours rejetés pour manque de la bande passante sur les liens adjacents aux PLR est plus faible par rapport à celle obtenue sur la topologie du réseau Cost 239. Par exemple, sur la topologie du réseau Long Haul, tout rejet d'une demande d'établissement d'un LSP P2MP de secours protégeant contre la panne du lien *6-9* sera souvent dû au manque de la bande passante sur les liens *7-11* et *5-8* (et *5-1*) et non pas sur les liens adjacents aux PLR *9* et *6*. Comme aucune unité de bande passante ne peut être libérée sur les liens *7-11* et *5-8* suite à la panne du lien *6-9* (car aucun plus court chemin ne peut traverser successivement les liens *6-9* et *7-11* ou *6-9* et *5-8*), nous déduisons que l'exploitation de la bande passante primaire libérée suite à une panne ne diminue pas (significativement) le taux de rejet des LSP P2MP de secours protégeant contre la panne du lien *6-9*. Sur la topologie du réseau Cost 239 par contre, le taux de LSP de secours rejetés pour manque de la bande passante sur les liens adjacents aux PLR est élevé. Comme la quantité de la bande passante primaire libérée suite à la panne d'un risque est plus élevée sur les liens adjacents aux PLR, l'application de la stratégie PGB au lieu de la stratégie PRB permettra de diminuer (sensiblement) le taux de protection.

Concernant la seconde métrique *RP2P*, les résultats illustrés sur les figures 6.6 et 6.7 montrent que les valeurs de ces taux suivent un comportement assez semblable à celui de la métrique *RP2MP*. Pour les mêmes raisons que précédemment (métrique *RP2MP*), les taux de rejet des LSP P2P de secours obtenus sur les deux figures 6.6 et 6.7 passent successivement par les deux premières phases qui sont la phase de satisfaction de toutes les requêtes de protection et la phase de disparité. Après ces deux phases, les taux de rejet des LSP P2P de secours des deux stratégies PGB et PRB atteignent la dernière phase, non illustrée sur les figures 6.6 et 6.7, qui est la phase de saturation (la charge du réseau n'est pas suffisamment élevée pour atteindre cette phase). Dans cette dernière phase, aucune nouvelle demande d'établissement de LSP P2P de secours n'est satisfaite par manque de la bande passante dans le pool de secours. Nous notons que cette dernière phase n'est franchissable que pour certaines familles d'algorithmes de calcul des LSP primaires comme l'algorithme des plus courts chemins utilisé dans nos simulations (cf. théorème 6.5.1 de l'annexe du chapitre).

Au regard de la troisième métrique *TUBS*, les figures 6.8 et 6.9 montrent que les taux d'utilisation de la bande passante de secours obtenus avec PGB sont proches de ceux obtenus avec PRB lorsque la charge du réseau est faible et ils leurs sont supérieurs pour les autres charges du réseau (charges moyennes et élevées). Concrètement, lorsque la charge du réseau est faible, (presque) aucune requête de protection n'est rejetée. Comme les LSP primaires à protéger et les algorithmes de calcul des LSP de secours employés dans nos simulations sont identiques, il est évident que les LSP de secours déterminés soient identiques. En conséquence, les valeurs des quantités de bande passante cumulée des LSP de secours correspondants et construits avec les stratégies PGB et PRB sont similaires. Lorsque la charge du réseau augmente et atteint un niveau assez élevé (pour que certaines requêtes de demande de protection soient rejetées avec PRB), la différence dans les valeurs de la métrique *TUBS* obtenues avec les stratégies PGB et PRB croit (en moyenne) jusqu'à la saturation complète du réseau, c'est-à-dire jusqu'à ce qu'il ne soit plus possible d'établir de nouveaux LSP de secours avec les deux stratégies PGB et PRB. En effet, avec l'augmentation de la charge, le taux de rejet des LSP de secours

augmente, ce qui accroît davantage les quantités de bande passante primaire libérées sur les liens (spécifiquement sur les liens adjacents aux PLR) suite à une panne. Ceci réduit les (l'augmentation des) coûts de protection (*effectifs*) des risques sur certains liens avec PGB, ce qui permet de satisfaire de nouvelles requêtes (contrairement à PRB ou aucune réduction de coût de protection n'est possible). Ceci explique l'augmentation de la différence entre les taux d'utilisation de la bande passante de secours obtenus avec les deux stratégies PGB et PRB. Lorsque la charge du réseau atteint une valeur très élevée, la phase de saturation sera franchie. Dans ce cas, aucun LSP de secours ne sera établi et donc, la différence entre les valeurs de la métrique *TUBS*, obtenus avec les deux stratégie PGB et PRB, se stabilise.

Concernant la dernière métrique *TOPS* et selon la topologie du réseau utilisée, la différence entre les taux d'occupation du pool de secours obtenus avec les stratégies PGB et PRB peut être apparente ou quasi-nulle.

Sur la topologie du réseau Long Haul, la figure 6.10 montre que le taux d'occupation du pool de secours obtenu avec PGB est inférieur à celui qui correspond à PRB, lorsque la charge du réseau est faible. Pour les fortes charges du réseau, les taux d'occupation du pool de secours des deux stratégie PGB et PRB sont très semblables. Ceci peut être expliqué par l'exploitation dans PGB (contrairement à PRB) de la bande passante primaire libérée suite à une panne lors du calcul de la quantité de bande passante de secours. Ainsi, avec l'augmentation du nombre de LSP P2MP primaires à protéger, le taux d'occupation du pool de secours obtenu avec PRB croît (en moyenne) plus vite que celui de PGB, jusqu'à atteindre sa valeur maximale égale à 0.8. A ce moment, la quantité de bande passante réellement utilisée pour fournir la protection avec PRB se stabilise (aux alentours de 0.8) alors que le taux d'utilisation de la bande passante de secours correspondant à PGB continue de grimper pour atteindre lui aussi la valeur maximale 0.8.

Sur la figure 6.11 par contre, nous constatons que les taux d'occupation du pool de secours obtenus avec les deux stratégies PGB et PRB sont quasi-identiques (la différence est de l'ordre de quelques millièmes et elle n'est pas perceptible sur la figure 6.11). Ceci s'explique par la grande connectivité de la topologie du réseau Cost 239 et par l'algorithme de calcul des chemins de secours qui, lors du calcul d'un nouveau LSP protégeant contre la panne d'un risque donné, favorisent l'utilisation des liens ne libérant pas la bande passante suite à la panne du risque précédent. Typiquement, pour récupérer la bande passante primaire sur un lien $B \rightarrow C$ suite à la panne du lien A-B, les deux liens $A \rightarrow B$ et $B \rightarrow C$ doivent appartenir à un même segment d'un LSP primaire (donc, le nœud C est à une distance de deux sauts du nœud A). Maintenant, pour utiliser la bande passante libérée sur le lien $B \rightarrow C$ suite à la panne du lien A-B, le LSP de secours protégeant contre la panne du lien A-B doit traverser le lien $B \rightarrow C$ avant d'atteindre le nœud A. En conséquence, le LSP de secours protégeant contre la panne du lien A-B doit être d'une longueur minimale égale à 3, ce qui très rare, pour les premiers LSP de secours établis, dans une topologie de réseau fortement connectée.

Contrairement aux taux d'utilisation obtenus sur la topologie du réseau Long Haul (qui sont très inférieurs à 1), ceux correspondant à la topologie Cost 239 atteignent des valeurs très proches de 1. Ceci s'explique par les symétries de la matrice de trafic et de

la topologie du réseau Cost 239, ce qui permet de bien répandre les allocations de la bande passante sur tous les liens du réseau.

6.3.5 Constats

Deux stratégies d'allocation de la bande passante ont été comparées dans cette section : le partage de la bande passante restreint aux LSP de secours et le partage global de la bande passante. Avec la première stratégie, seuls les LSP P2P de secours partagent la bande passante sur leurs liens communs en utilisant la fusion de flux et la fusion de LSP. Aucun partage, par contre, n'est appliqué entre les LSP primaires et les LSP de secours (même pas la fusion de LSP). Avec la seconde stratégie, tous les types de partage (partage de la bande passante entre les LSP primaires et les LSP de secours d'un côté et partage de la bande passante entre les LSP de secours d'un autre côté) sont appliqués aux LSP afin de diminuer la consommation de la bande passante sur les liens.

Comme la différence entre les deux stratégies PGB et PRB se situe dans la capacité de la première stratégie (contrairement à la seconde stratégie) à exploiter la bande passante primaire libérée sur les liens suite aux pannes, il est évident que le gain en performances (métriques *RP2MP*, *RP2P*, *TUBS* et *TOPS*) résultant de l'utilisation de PGB au lieu de PRB, lors du calcul des nouveaux LSP de secours, dépendra essentiellement des quantités de bande passante primaire libérées ainsi que de la distribution de ces dernières sur les liens du réseau.

Divers paramètres contrôlent les quantités et la qualité de distribution de la bande passante primaire libérée sur les liens suite à une panne, parmi lesquels nous citons : la topologie du réseau (connectivité et symétrie), l'algorithme de calcul des LSP primaires employé, la matrice de trafic et le taux de rejet des LSP de secours autorisé.

Dans notre étude de performances, nous n'avons considéré que le cas d'algorithmes de calcul de LSP primaires basés sur les plus courts chemins où la métrique à optimiser est statique. Nous avons ainsi remarqué qu'indépendamment de l'algorithme de calcul adopté pour calculer les LSP de secours, le nombre maximum de ces LSP (de secours) pouvant être établis avec PRB (la preuve est triviale) et PGB (la preuve est apportée par le théorème 6.5.1 de l'annexe du chapitre) dans un réseau est borné, lorsque les capacités de secours sont limitées. Nous avons aussi constaté qu'avec un choix des membres et sources multicast uniformément distribués sur les nœuds du réseau, l'utilisation de topologies de réseau présentant une symétrie et une connectivité élevée pourrait permettre d'augmenter le gain en performances avec la substitution de PGB à PRB. Enfin, nous avons fait observer que la différence dans les performances des stratégies PGB et PRB n'est apparente et perceptible que lorsque le taux de rejet des LSP de secours est assez élevé. Cela réduit l'intérêt pour la stratégie PGB lorsque le blocage des requêtes de protection n'est pas autorisé d'autant plus que cette stratégie induit de nouveaux traitements (liés à la gestion de la bande passante primaire libérée dans le réseau suite à une panne) et pourrait augmenter le trafic de contrôle traversant le réseau.

6.4 Conclusion et perspectives

Dans ce chapitre, nous nous sommes intéressés aux mécanismes permettant de protéger localement et en ligne les communications point à multipoint sous MPLS. Nous nous sommes d'abord focalisés sur la structure de l'information nécessaire au contrôle d'admission et au calcul efficace des LSP P2MP de secours dans un environnement distribué en utilisant deux stratégies de partage de la bande passante : le partage restreint de la bande passante (PRB) et le partage global de la bande passante (PGB). Ensuite, nous avons mesuré par simulations l'impact du choix de la stratégie de partage de la bande passante employée sur les performances des mécanismes de placement des LSP P2MP de secours.

Avec la première stratégie de partage de la bande passante (PRB), nous avons vu que le contrôle d'admission distribué peut être effectué pour chaque lien unidirectionnel sur son nœud sortant sans aucun surcoût, grâce à l'exploitation de l'information transmise par les protocoles de routage et les protocoles de signalisation étendus. De même, nous avons montré que les algorithmes et heuristiques distribuant, agrégeant et approximant l'information nécessaire au calcul des LSP de secours protégeant des communications unicast peuvent être facilement étendus pour protéger les communications point à multipoint.

Avec la seconde stratégie de partage (PGB) par contre, l'utilisation des quantités de la bande passante primaire libérées sur les liens suite à une panne, lors du contrôle d'admission et du calcul distribués des LSP P2MP de secours, requiert sa distribution dans le réseau. Pour ce faire, différentes approches définissant de nouveaux protocoles et/ou étendant les protocoles existants peuvent être adoptées. Dans ce chapitre, nous avons proposé différentes modifications et/ou extensions des protocoles de signalisation et des protocoles de routage pour tenir compte des quantités de bande passante primaire libérées suite à une panne lors du contrôle d'admission et du calcul des LSP P2MP de secours. Nos propositions ont l'avantage de la facilité d'implantation et de déploiement puisqu'elles ne requièrent que de très légères extensions aux algorithmes et heuristiques proposés pour protéger les communications unicast.

Après avoir défini les outils permettant d'implanter les deux stratégies de partage de la bande passante PRB et PGB, nous nous sommes tournés vers l'étude de l'impact du type de la stratégie de partage employée sur les performances des mécanismes de placement des LSP P2MP de secours. Nous avons vu, qu'indépendamment de l'algorithme de calcul des LSP de secours, le nombre de LSP de secours pouvant être établis dans un réseau est borné lorsque les capacités de secours sont limitées et lorsque l'algorithme de calcul des LSP primaires est basé sur l'algorithme des plus courts chemins en termes d'une métrique statique (exemple : nombre de sauts). Par ailleurs, le gain apporté par la substitution de PGB à PRB lors du calcul des LSP de secours n'est perceptible et apparent que pour les grandes charges du réseau où le taux de rejet est relativement élevé. Typiquement, lorsque le blocage des demandes de protection n'est pas autorisée (ou est faible), PRB devrait être préférée à PGB car elle diminue la quantité d'informations distribuées dans le réseau et elle facilite le déploiement des mécanismes de placement des LSP de secours. Par contre, lorsque le taux de blocage autorisé est élevé et lorsque

les ressources (la bande passante) sont limitées sur les liens, PGB pourrait être préférée à PRB pour reporter au plus loin les investissements (augmentation des capacités des liens par exemple) et mieux satisfaire les requêtes de protection.

6.5 Annexe

Nous présentons dans cette section un théorème (et sa preuve) lié à l'exploitation de la bande passante primaire libérée sur les liens. Ce théorème garantit l'inexistence d'un LSP reliant les nœuds d'extrémité d'un lien *u-v* en panne (resp. reliant deux nœuds voisins à un nœud *u* en panne) via un chemin constitué exclusivement de liens sur lesquels est libérée de la bande passante après la panne du lien *u-v* (resp. du nœud *u*).

Théorème 6.5.1 *Soit un réseau représenté par un graphe G = (V, E) où V est l'ensemble des nœuds du réseau et E est l'ensemble de ses liens. Chaque lien u-v supporte deux arcs de sens opposés : u→v et v→u.*

Pour tout routage de LSP P2P primaires basé sur les plus courts chemins où la valeur de la métrique à optimiser est strictement positive et est constante sur chaque arc du graphe (métrique statique), l'ensemble des arcs sur lesquels est libérée de la bande passante primaire suite à une panne d'un nœud (resp. d'un lien) ne peut contenir un chemin permettant de protéger localement contre la panne du nœud en panne (resp. un chemin permettant de protéger contre le lien en panne).

Preuve :
Cas d'une panne de nœud u
Nous allons effectuer une preuve par contradiction. Supposons qu'il existe un plus court chemin primaire P contenant le segment $plr \rightarrow u \rightarrow p1 \rightarrow p2 \rightarrow .. \rightarrow p_x \rightarrow mp$ et un LSP de secours $B = plr \rightarrow b_1 \rightarrow b_2 \rightarrow .. \rightarrow b_y \rightarrow mp$ (avec b_1, b_2, .., $b_y \neq u$) permettant de le protéger contre la panne du nœud u en n'utilisant que des arcs sur lesquels est libérée de la bande passante suite à la panne du nœud u. Comme il n'est possible de libérer de la bande passante suite à la panne d'un nœud u que sur des arcs appartenant aux différents arbres des plus courts chemins de racine u (car les LSP P2P primaires sont des plus courts chemins), nous déduisons que tous les arcs $plr \rightarrow b_1$, $b_1 \rightarrow b_2$, $b_{y-1} \rightarrow b_{y-1}$, $b_{y-1} \rightarrow mp$ du LSP de secours B appartiennent à des arbres des plus courts chemins de racine u. Cela veut dire qu'il existe au moins un plus court chemin $u \rightarrow .. \rightarrow plr$ permettant de relier le nœud u au nœud plr. Comme l'arc $plr \rightarrow b_1$ appartient à au moins un arbre des plus courts chemins de racine u, nous déduisons qu'il existe au moins un plus court chemin $u \rightarrow .. \rightarrow plr \rightarrow b_1$ reliant le nœud u au nœud b_1, en passant par le nœud plr. De même, nous déduisons qu'il existe un plus court chemin $u \rightarrow .. \rightarrow plr \rightarrow b_1 \rightarrow b_2$ permettant de relier le nœud u au nœud b_2. En suivant le même raisonnement, nous concluons qu'il existe un plus court chemin $u \rightarrow .. \rightarrow plr \rightarrow b_1 \rightarrow b_2 \rightarrow .. \rightarrow b_y \rightarrow mp$ permettant de relier le nœud u au nœud mp, ce qui n'est pas possible. En effet, l'existence de deux plus courts chemins $u \rightarrow p_1 \rightarrow p_2 \rightarrow .. \rightarrow p_x \rightarrow mp$ (segment du chemin primaire) et $u \rightarrow .. \rightarrow plr \rightarrow b_1 \rightarrow b_2 \rightarrow .. \rightarrow b_y \rightarrow mp$ reliant le nœud u au nœud mp permet de déduire que le LSP primaire P ne peut correspondre à un plus court chemin (à moins que les coûts

primaires de l'arc $plr \rightarrow u$ et du segment $u \rightarrow .. \rightarrow plr$ soient nuls ou négatifs), ce qui contredit nos hypothèses.□

Cas d'une panne d'un lien $u \rightarrow v$
Nous allons effectuer une preuve par contradiction. Supposons qu'il existe un plus court chemin primaire P contenant le segment $u \rightarrow v \rightarrow p_1 \rightarrow p_2 \rightarrow .. \rightarrow p_x \rightarrow mp$ et un LSP de secours $B = u \rightarrow b_1 \rightarrow b_2 \rightarrow .. \rightarrow b_y \rightarrow mp$ (avec b_1, b_2, .., $b_y \neq v$) permettant de le protéger contre la panne du lien u-v en n'utilisant que des arcs sur lesquels est libérée de la bande passante suite à la panne du lien u-v.

Il est évident que, lors de la panne d'un lien u-v, de la bande passante primaire ne peut être libérée que sur les arcs partagés entre les arbres des plus courts chemins de racine u et les arbres des plus courts chemins de racine v. En conséquence, nous déduisons que tous les arcs du LSP de secours B appartiennent à au moins un arbre des plus courts chemins de racine v. En adoptant le même raisonnement que pour la panne d'un nœud, nous déduisons qu'il existe au moins un plus court chemin $v \rightarrow .. \rightarrow u \rightarrow b_1 \rightarrow b_2 \rightarrow .. \rightarrow b_y \rightarrow mp$ reliant le nœud v au nœud mp. Cela n'est pas possible car l'existence d'un autre chemin différent $v \rightarrow p_1 \rightarrow p_2 \rightarrow .. \rightarrow p_x \rightarrow mp$ (segment du LSP primaire) reliant le nœud v au nœud mp implique que le coût primaire de l'arc $u \rightarrow v$ et le coût primaire du segment $v \rightarrow .. \rightarrow u$ sont nuls ou négatifs, ce qui contredit nos hypothèses. □

Chapitre 7

Conclusion génarale et perspectives

7.1 Conclusion générale

Dans cette thèse, nous nous sommes intéressés de près à la protection proactive locale des communications point à point et point à multipoint sous MPLS. Ce type de protection permet de réduire les délais de récupération en dotant les routeurs de fonctions permettant le contournement rapide des pannes, par un basculement local du trafic des communications affectées vers les chemins de secours, dès qu'une panne est détectée. Notre objectif était de proposer de nouveaux mécanismes de placement distribué et en ligne des chemins de secours afin d'assurer le passage à l'échelle de la protection proactive locale et pour éviter l'explosion des délais d'établissement des communications après une panne.

Pour des raisons d'efficacité (c'est-à-dire, pour éviter le gaspillage des ressources du réseau), de structures des réseaux MPLS (contenant trois types de risques de panne), de facilité de mise en œuvre et de passage à l'échelle, nous avons tenu compte, au cours de notre étude, de trois enjeux principaux qui sont :

1. L'optimisation de l'utilisation des ressources : pour repousser au plus tard les investissements réseau (ajout ou augmentation des capacités de liens, installation de nouveaux routeurs, etc.), les mécanismes de placement des chemins de secours doivent profiter des gains offerts par le partage de ressources (particulièrement la bande passante) pour mieux exploiter ces ressources. Ils doivent aussi garantir la disponibilité des ressources après une panne et réduire les délais de récupération en diminuant autant que possible la longueur des chemins de secours ;
2. La protection contre tous les types de risque de panne : les mécanismes de placement des chemins de secours doivent permettre de protéger non seulement contre les risques de panne simples (nœud ou lien) mais aussi contre les risques de panne de type SRLG. Un risque SRLG est constitué d'un ensemble de liens logiques (liens MPLS, par exemple) partageant un ou plusieurs composants physiques (par exemple switch optique ou switch ATM) qui ne sont pas visibles par la couche MPLS. La défaillance du composant physique entraîne les pannes simultanées de tous les liens logiques l'utilisant. Actuellement, beaucoup de réseaux optiques et

ATM disposent de risques de type SRLG, d'où la nécessité de protéger contre ce type de risque de panne. Les SRLG peuvent être aussi utilisés pour protéger contre les pannes multiples de liens ;

3. La facilité d'implantation et le passage à l'échelle : les mécanismes de placement des chemins de secours doivent être faciles à déployer. Typiquement, de très légères extensions aux protocoles existants et/ou la définition de nouveaux protocoles légers devraient suffire à leur déploiement. Pour passer à l'échelle, les mécanismes de placement des chemins de secours ne doivent pas inonder le réseau avec des messages d'annonce des paramètres requis au calcul distribué des chemins de secours.

Pour répondre aux enjeux précédents, nous avons proposé différents algorithmes et heuristiques efficaces pour le placement des chemins de secours et nous avons amélioré le procédé d'allocation de la bande passante de secours. Nous avons aussi étudié différentes stratégies de partage de la bande passante et mesuré leur impact sur les performances des mécanismes de placement des chemins de secours.

Dans la première partie de cette thèse, nous nous sommes focalisés uniquement sur la protection locale des communications point à point. Afin d'éviter les fuites de la bande passante, nous avons proposé dans un premier temps l'algorithme TDRA (Targeted Distribution of Resource Allocation) qui cible les routeurs auxquels il envoie l'information requise pour le calcul des chemins de secours. Dans TDRA, les entités de calcul des chemins de secours sont distribuées sur les nœuds du réseau de manière à diminuer les échanges de messages. Cet algorithme est facile à déployer et ne nécessite que de légères extensions aux protocoles de signalisation et aux protocoles IGP-TE. Il est surtout efficace pour placer les chemins de secours dans un réseau large, contenant un nombre limité de SRLG.

Pour réduire la taille et la fréquence d'envoi de l'information nécessaire au placement des chemins de secours, nous avons proposé dans un deuxième temps d'agréger cette information (i.e. information nécessaire au placement des chemins de secours) avant de la diffuser dans le réseau.

Ainsi, dans une première heuristique DBSH (Distributed Bandwidth Sharing Heuristic), nous centralisons les calculs de tous les chemins protégeant contre les pannes d'un nœud et de ses liens adjacents sur le nœud lui-même. De cette manière, aucun échange d'information n'est nécessaire pour protéger contre les risques de panne simples (risques de type nœud ou lien). Par contre et pour permettre la protection contre les risques de panne de type SRLG, nous avons proposé de diffuser pour chaque lien (unidirectionnel) du réseau, une information partielle (notée $x^{\lambda}\_vecteur$) sur les SRLG protégés en utilisant le lien. En se servant de cette information partielle diffusée dans le réseau, nous avons montré (par simulation) qu'il est possible d'approximer efficacement tous les paramètres nécessaires au calcul distribué des chemins de secours. Cette heuristique DBSH est facile à déployer et ne nécessite que de légères extensions aux protocoles de signalisation et protocoles IGP-TE. Elle réduit aussi le volume de l'information diffusée dans le réseau pour le placement des chemins de secours au détriment d'une légère diminution du taux de partage et d'une distribution asymétrique de l'information requise

pour le calcul des chemins de secours (dans ce cas, l'information requise pour le calcul des chemins de secours est structurée en plusieurs partitions localisées sur plusieurs entités de calcul des chemins de secours).

Afin d'offrir une meilleure flexibilité de choix des emplacements des entités de calcul et pour une distribution symétrique de l'information requise pour le calcul des chemins de secours (en d'autres termes, afin que toutes les entités de calcul disposent de la même information requise pour le calcul des chemins de secours), nous avons proposé une deuxième heuristique PLRH (PLR-based Heuristic). Cette heuristique est une extension à DBSH dans laquelle nous envoyons pour chaque lien du réseau une information partielle ($y^{\lambda}\_vecteur$) sur tous les risques protégés (au lieu de l'envoyer pour les SRLG protégés uniquement) en utilisant ce même lien. Comme dans DBSH, PLRH permet de déterminer et d'approximer efficacement toute l'information requise au placement des chemins de secours. En plus de l'avantage de la réduction du volume d'informations diffusées dans le réseau, PLRH permet d'éliminer la communication entre les nœuds supportant les entités de calcul et les nœuds chargés de configurer les chemins de secours. L'heuristique PLRH permet aussi de coordonner le calcul des chemins de secours avec leur chemin primaire (pour améliorer la protection et faciliter le calcul des chemins de secours multi-domaines). Cette heuristique ne nécessite que de très légères extensions aux protocoles IGP-TE pour son implantation.

Après la validation de nos algorithmes et heuristiques par simulations, nous nous sommes intéressés à l'amélioration du procédé d'allocation de la bande passante de secours. En remarquant que certains chemins de secours activés suite à une panne d'un SRLG ne reçoivent aucun flux (chemins dits *inopérationnels*), nous avons proposé de tenir compte des structures des SRLG afin de limiter la concurrence pour l'allocation de la bande passante de secours aux chemins recevant effectivement des flux après une panne (chemins dits *opérationnels*). De plus, afin d'augmenter le taux de protection, nous avons réduit l'ensemble des risques de panne devant être protégés par chaque chemin de secours à ceux provoquant son opération. Nos simulations ont montré que l'exploitation des structures des SRLG réduit le taux moyen de rejet des chemins de secours et diminue le nombre moyen de messages diffusés dans le réseau.

Dans la seconde partie de cette thèse, nous nous sommes concentrés sur la protection des communications de groupes et particulièrement les communications point à multipoint. Après avoir montré que nos algorithmes et heuristiques (TDRA, DBSH et PLRH) sont facilement extensibles pour protéger les communications point à multipoint, nous nous sommes focalisés sur l'étude de l'impact du choix des stratégies de partage de la bande passante sur les performances des mécanismes de placement des chemins point à multipoint de secours. Deux stratégies de partage ont été ainsi comparées : (1) le partage restreint de la bande passante (PRB) qui limite le partage de la bande passante aux chemins de secours et (2) le partage global de la bande passante (PGB) qui applique le partage de la bande passante entre les chemins de secours d'un côté et entre les chemins primaires et les chemins de secours d'un autre côté. Dans notre étude, nous avons adoptée la protection multicast locale un-à-un où la distinction entre les deux stratégies PRB et PGB est très perceptible et est apparente.

Deux principaux résultats ressortent de notre étude. Premièrement, nous avons mon-

tré qu'indépendamment de la stratégie de partage employée (PRB ou PGB) et de l'algorithme de calcul des chemins de secours utilisé, le nombre de chemins de secours pouvant être établis dans un réseau est borné lorsque les capacités de secours sont limitées et lorsque l'algorithme de calcul des chemins primaires est basé sur l'algorithme des plus courts chemins en termes d'une métrique statique (par exemple, le nombre de sauts). Deuxièmement, nous avons remarqué (dans nos simulations) que la différence de performances entre PGB et PRB n'est perceptible et apparente (statistiquement PGB a toujours de meilleures performances) que lorsque les taux de rejet des requêtes de protection sont relativement élevés. Cela réduit sensiblement l'intérêt pour la stratégie PGB dans les réseaux destinés à éviter le blocage des requêtes de protection.

7.2 Perspectives

De nombreuses pistes de recherche peuvent être explorées afin de compléter notre étude du problème de placement des chemins de secours et pour améliorer les différentes approches proposées.

Ainsi, pour tenir compte des délais de propagation des messages d'annonce de l'information requise au calcul des chemins de secours, il serait intéressant d'étendre l'étude de performances des différents algorithmes et heuristiques proposés à d'autres métriques, comme le taux d'échec de configuration des chemins de secours, la longueur moyenne des chemins de secours ou le délai moyen/maximum de récupération, etc.

Pour optimiser le nombre moyen de messages diffusés dans le réseau avec les heuristiques DBSH et PLRH et pour bien calibrer les tailles des vecteurs qui transmettent l'information nécessaire au calcul des chemins de secours, il serait judicieux de mesurer, analytiquement et par simulation, la corrélation entre les topologies des réseaux, les matrices de trafic et les algorithmes de calcul des chemins primaires et de secours avec les valeurs du x^{eme} plus grand prix de protection sur chaque lien du réseau. De plus, afin de mieux exploiter la bande passante disponible sur les liens, il pourrait être intéressant d'ajuster les capacités des pools primaires et de secours en fonction des structures et propriétés des chemins primaires et de secours traversant chaque lien. Par exemple, les valeurs des capacités primaires et de secours des liens pourraient être fixées à l'initialisation du réseau selon un modèle de prévisions des futures requêtes (en utilisant une enveloppe primaire et une enveloppe de secours comme dans [Gro04]). Au fur et à mesure de l'arrivée de nouvelles requêtes (d'établissement ou de suppression des chemins), le modèle peut être réajusté et corrigé en augmentant ou diminuant les valeurs des capacités de certains pools.

Parallèlement aux études et améliorations ci-dessus, nous pouvons envisager de prendre en compte les niveaux de priorité et d'utiliser la préemption pour satisfaire les requêtes de protection de plus haute priorité. Différentes précautions doivent être prises afin d'implanter les priorités. Primo, selon le type de partage de la bande passante, la suppression d'un chemin primaire ou de secours peut ou pas permettre la libération de la bande passante. Secondo, la connaissance partielle de l'information requise pour le placement des chemins de secours peut s'avérer insuffisante pour choisir les chemins

à préempter. En conséquence, de légères extensions et/ou adaptations des heuristiques DBSH et PLRH pourraient être nécessaires pour prendre en compte les niveaux de priorités lors du placement des chemins de secours.

Pour compléter notre étude de l'impact du choix de la stratégie de partage de la bande passante employée sur les performances des mécanismes de placement des chemins de secours, il serait envisageable de l'étendre à des métriques dynamiques et à des algorithmes de calcul des chemins qui ne sont pas basés sur les plus courts. Une telle étude augmenterait l'efficacité des mécanismes de placement des chemins de secours en permettant d'opter pour la stratégie de partage de la bande passante la plus adéquate.

Enfin, concernant les communications multicast, notre travail dans cette thèse pourrait être amélioré de différentes manières. Premièrement, il est intéressant d'étendre l'étude de performances effectuée sur la protection locale un-à-un des communications P2MP à la protection par tunnels P2P et P2MP de secours. Deuxièmement, adapter les solutions proposées pour la protection des communications point à multipoint à la protection locale multipoint à multipoint serait un nouveau challenge et une autre étape pour le déploiement de la protection multicast. Troisièmement, l'adoption de l'agrégation des chemins de secours pour réduire le nombre d'états (états RSVP) à gérer et à maintenir pourrait être une piste très prometteuse pour le passage à l'échelle de la protection multicast.

Glossaire

ASM : Any Source Multicast
BFD : Bidirectional Forwarding Detection
BPCE : Backup Path Computation Element
COS : Class Of Service
CR-LDP : Constraint-Based LSP Setup using LDP
CSPF : Constrained Shortest Path First
DBSH : Distributed Bandwidth Sharing Heuristic
FEC : Forwarding Class Equivalence
IETF : Internet Engineering Task Force
IGP : Interior Gateway Protocol
ILP : Integer Linear Programming
IP : Internet Protocol
ISIS : Intermediate System to Intermediate System
ISO : International Organization for Standardization
LDP : Label Distribution Protocol
LER : Label Edge Router
LFIB : Label Forwarding Information Base
LSA : Link State Advertisement
LSP : Label Switched Path
LSPDU : Link State Protocol Data Unit
LSR : Label Switched Routers
MP : Merge Point
MPLS : MultiProtocol Label Switching
MST : Minimal Spanning Tree
NHOP : Next HOP
NNHOP : Next Next HOP
OSPF : Open Shortest Path First
P2P : Point to Point
P2MP : Point to MultiPoint
PGB : Partage Global de la Bande passante
PHP : Penultimate Hop Popping
PLR : Point of Local Repair
PLRH : PLR-based Heuristic
PRB : Partage Restreint de la Bande passante

QoS : Quality of Service
RSVP : Resource reSerVation Protocol
S : Bottom of Stack
S2L : Source to Leaf
SDH : Synchronous Digital Hierarchy
SPF : Shortest Path First
SRLG : Shared Risk Link Group
SSM : Source-Specific multicast
TDRA : Targeted Distribution of Resource Allocation
TE : Traffic Engineering
TOS : Type Of Service
TV : TeleVision
VoIP : Voice over IP
VPN : Virtual Private Network

Bibliographie

[ABG+01] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, and G. Swallow. RSVP-TE : Extensions to RSVP for LSP Tunnels. RFC 3209, December 2001.

[ADF+01] L. Andersson, P. Doolan, N. Feldman, A. Fredette, and B. Thomas. LDP Specification. RFC 3036, January 2001.

[AKNS08] R. Aggarwal, K. Kompella, T. Nadeau, and G. Swallow. BFD For MPLS LSPs. Internet Draft draft-ietf-bfd-mpls-07.txt, IETF, June 2008.

[ALR08] R. Aggarwal and J. L. Le Roux. MPLS Upstream Label Assignement for RSVP-TE. Internet Draft draft-ietf-mpls-rsvp-upstream-03.txt, IETF, July 2008.

[AMA+99] D. Awduche, J. Malcolm, J. Agogbua, M. O'Dell, and J. McManus. Requirements for Traffic Engineering Over MPLS. RFC 2702, September 1999.

[APY07] R. Aggarwal, D. Papadimitriou, and S. Yasukawa. Extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for Point-to-Multipoint TE Label Switched Paths (LSPs). RFC 4875, May 2007.

[ARR08] R. Aggarwal, Y. Rekhter, and E. Rosen. MPLS Upstream Label Assignment and Context-Specific Label Space. RFC 5331, August 2008.

[ARUK06] F. Aslam, S. Raza, Z. Afzal Uzmi, and Y. C. Kim. Bandwidth Sharing with Primary Paths for Protection Routing in an MPLS Network. In *Proceedings of 25th IEEE International Conference on Computer Communications (INFOCOM 2006)*, pages 1–6, April 2006.

[BCG+07] Wojtek Bigos, Bernard Cousin, Stéphane Gosselin, Morgane Le Foll, and Hisao Nakajima. Survivable MPLS Over Optical Transport Networks : Cost and Resource Usage Analysis. *IEEE Journal on Selected Areas in Communications*, 25(5) :949–962, 2007.

[Ber03] L. Berger. Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions. RFC 3473, January 2003.

[Bha99] R. Bhandari. *Survivable Networks : Algorithms for Diverse Routing*. Kluwer Academic Publishers, 1999.

[BML06] S. Balon, L. Mélon, and G. Leduc. A Scalable and Decentralized Fast-Rerouting Scheme with Efficient Bandwidth Sharing. *Computer Networks*, 50(16) :3043–3063, November 2006.

[BZB+97] R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin. Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification. RFC 2205, September 1997.

[Dee91] S. E. Deering. *Multicast routing in a datagram internetwork*. PhD thesis, Stanford University (USA), 1991.

[Dij59] E. W. Dijkstra. A note on two problems in connection with graphs. *Numerische Mathematik*, 1 :269–271, 1959.

[FAV08] A. Farrel, A. Ayyangar, and J. P. Vasseur. Inter-Domain MPLS and GMPLS Traffic Engineering – Resource Reservation Protocol-Traffic Engineering (RSVP-TE) Extensions. RFC 5151, February 2008.

[FCGF01] A. Fei, J. Cui, M. Gerla, and M. Faloutsos. Aggregated multicast : an approach to reduce multicast state. In *IEEE Global Telecommunications Conference, 2001 (GLOBECOM '01)*, volume 3, pages 1595–1599, San Antonio, Texas (USA), 2001.

[FPVA06] A. Farrel, D. Papadimitriou, J. P. Vasseur, and A. Ayyangar. Encoding of Attributes for Multiprotocol Label Switching (MPLS) Label Switched Path (LSP) Establishment Using Resource ReserVation Protocol-Traffic Engineering (RSVP-TE). RFC 4420, February 2006.

[FVA06] A. Farrel, J. P. Vasseur, and J. Ash. A Path Computation Element (PCE)-Based Architecture. RFC 4655 (Informational), 2006.

[Gro04] W. D. Grover. The Protected Working Capacity Envelope Concept : An Alternative Paradigm for Automated Service Provisioning. *IEEE Communications Magazine*, pages 62–69, 2004.

[GS98] W. Grover and D. Stamatelakis. Cycle-Oriented Distributed Preconfiguration : Ring-like Speed with Mesh-like Capacity for Self-planning Network Restoration. In *Proceedings International Conference on Communications*, pages 537–543, 1998.

[Hu03] Jian Qiang Hu. Diverse Routing in Optical Mesh Networks. *IEEE Transactions on Communications*, 51(3) :489– 494, 2003.

[JAC+02] B. Jamoussi, L. Andersson, R. Callon, R. Dantu, L. Wu, P. Doolan, T. Worster, N. Feldman, A. Fredette, M. Girish, E. Gray, J. Heinanen, T. Kilty, and A. Malis. Constraint-Based LSP Setup using LDP. RFC 3212, January 2002.

[KKL+01] S. Kini, K. Kodialam, T. V. Lakshman, S. Sengupta, and C. Villamizar. Shared Backup Label Switched Path Restoration. Internet Draft draft-kini-restoration-shared-backup-01.txt, IETF, May 2001.

[KL01] M. S. Kodialam and T. V. Lakshman. Dynamic Routing of Locally Restorable Bandwidth Guaranteed Tunnels using Aggregated Link Usage Information. In *Proceedings of 20th IEEE International Conference on Computer Communications (INFOCOM 2001)*, pages 376–385, 2001.

[KL03] M. S. Kodialam and T. V. Lakshman. Dynamic Routing of Restorable Bandwidth-Guaranteed Tunnels using Aggregated Network Resource Usage Information. *IEEE/ACM Transactions On Networking*, 11(3) :399–410, June 2003.

[KL04] K. Kompella and J. Lang. Procedures for Modifying the Resource reSerVation Protocol (RSVP). RFC 3936, October 2004.

[KMB81] L. Kou, G. Markowsky, and L. Berman. A Fast Algorithm for Steiner Trees. *Acta Informatica*, 15 :141–145, June 1981.

[KR05a] K. Kompella and Y. Rekhter. Intermediate System to Intermediate System (IS-IS) Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS). RFC 4205, October 2005.

[KR05b] K. Kompella and Y. Rekhter. OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS). RFC 4203, October 2005.

[KR05c] K. Kompella and Y. Rekhter. Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS). RFC 4202, October 2005.

[Kru56] J. B. Kruskal. On the Shortest Spanning Subtree and the Traveling Salesman Problem. In *Proceedings of the American Mathematical Society*, volume 7, pages 48–50, 1956.

[LCC] H. Lee, J. Chung, and S. J. Chung. A State-dependent Preplanned ATM VP Restoration Scheme. In *Global Telecommunications Conference, 1998 (IEEE GLOBECOM '98)*, volume 3, pages 1766–1771.

[LFDC07] C. Y. Lee, A. Farrel, and S. De Cnodder. Exclude Routes - Extension to Resource ReserVation Protocol-Traffic Engineering (RSVP-TE). RFC 4874, April 2007.

[LR06] J. L. Le Roux. MPLS : applications à l'ingénierie de trafic et à la sécurisation. *Techniques de l'Ingénieur*, Mai 2006.

[LRAVV07] J. L. Le Roux, R. Aggarwal, P. Vasseur, J., and M. Vigoureux. P2MP MPLS-TE Fast Reroute with P2MP Bypass Tunnels. Internet Draft draft-ietf-mpls-p2mp-te-bypass-02.txt, IETF, March 2007.

[LRC02] J. L. Le Roux and G. Calvignac. A Method for an Optimized Online Placement of MPLS Bypass Tunnels. Internet Draft draft-leroux-mpls-bypass-placement-00.txt, IETF, February 2002.

[MBL03] L. Mélon, F. Blanchy, and G. Leduc. Decentralized Local Backup LSP Calculation with Efficient Bandwidth Sharing. In *Proceedings of 10th International Conference on Telecommunications*, Papeete (Tahiti), February 2003.

[MFB99] M. Medard, S. G. Finn, and R. A. Barry. Redundant Trees for Preplanned Recovery in Arbitrary Vertex-Redundant or Edge-Redundant Graphs. *IEEE/ACM Transactions on Networking*, 7(5) :641–652, 1999.

[MK98] K. Murakami and H. S. Kim. Optimal Capacity and Flow Assignment for Self-Healing ATM Networks based on Line and End-to-end Restoration. *IEEE/ACM Transactions on Networking*, 6(2) :207–221, 1998.

[Mou06] J. Moulierac. *Agrégation des communications multicast.* PhD thesis, 2006.

[Moy98a] J. Moy. *OSPF Anatomy of an Iternet Routing Protocol.* Addison-Wesley, 1998.

[Moy98b] J. Moy. OSPF Version 2. RFC 2328, April 1998.

[Moy04] J. Moy. *Network Recovery : Protection and Restoration of Optical, SONET-SDH, IP, and MPLS.* Morgan Kaufmann Publishers, 2004.

[MP06] E. Mannie and D. Papadimitriou. Recovery (Protection and Restoration) Terminology for Generalized Multi-Protocol Label Switching (GMPLS). RFC 4427 (Informational), 2006.

[MVDBD04] P. Meyer, S. Van Den Bosch, and N. Degrande. High Availability in MPLS-based Networks. Alcatel telecommunication review, Alcatel, 4th Quarter 2004.

[Ora90] D. Oran. OSI IS-IS Intra-domain Routing Protocol. RFC 1142, February 1990.

[Pri57] R. C. Prim. Shortest Connection Networks and some Generalisations. *Numerische Mathematik*, 36 :1389–1401, 1957.

[PSA05] P. Pan, G. Swallow, and A. Atlas. Fast Reroute Extensions to RSVP-TE for LSP Tunnels. RFC 4090, May 2005.

[RM99a] S. Ramamurthy and B. Mukherjee. Survivable WDM Mesh Networks (Part I - Protection). In *Proceedings of 18th IEEE International Conference on Computer Communications (INFOCOM 2001)*, volume 2, pages 744–751, 1999.

[RM99b] S. Ramamurthy and B. Mukherjee. Survivable WDM Mesh Networks (Part II - Restoration). In *IEEE International Conference on Communications*, pages 2023–2030, Vancouver (Canada), June 1999.

[RVC01] E. Rosen, A. Viswanathan, and R. Callon. Multiprotocol Label Switching Architecture. RFC 3031, January 2001.

[SC07] M. Y. Saidi and B. Cousin. Evaluation par simulation des performances des mécanismes de placement des LSP de secours unicast. Rapport R2+R3 du projet de recherche *Méthode de contrôle distribué du placement de LSP de secours* entre l'université de Rennes 1 et France Télécom, June 2007.

[SC08] M. Y. Saidi and B. Cousin. Evaluation par simulation des performances des mécanismes de placement des LSP de secours multicast. Rapport R5 du projet de recherche *Méthode de contrôle distribué du placement de LSP de secours* entre l'université de Rennes 1 et France Télécom, June 2008.

[SCLRa] M. Y. Saidi, B. Cousin, and J. L. Le Roux. Backup Path Classification based on Failure Risks for Efficient Backup Path Computation. In *IFIP Networking 2009*.

[SCLRb] M. Y. Saidi, B. Cousin, and J. L. Le Roux. Distributed PLR-Based Backup Path Computation in MPLS Networks. In *IFIP Networking 2008*.

[SCLR07] M. Y. Saidi, B. Cousin, and J. L. Le Roux. A Distributed Bandwidth Sharing Heuristic for Backup LSP Computation. In *Global Telecommunications Conference, 2007 (IEEE GLOBECOM '07)*, pages 2477–2482, Washington (USA), November 2007.

[SCLR08] M. Y. Saidi, B. Cousin, and J. L. Le Roux. Targeted Distribution of Resource Allocation for Backup LSP Computation. In *Seventh European Dependable Computing Conference (EDCC-7)*, Kaunas (Lithuania), May 2008.

[SCLR09a] M. Y. Saidi, B. Cousin, and J. L. Le Roux. PLR-based Backup Path Computation in MPLS Network. *Computer Networks*, 2009.

[SCLR09b] M. Y. Saidi, B. Cousin, and J. L. Le Roux. Using Shared Risk Link Group Structures for an Efficient Protection. *Computer Networks*, 2009.

[SCM06a] M. Y. Saidi, B. Cousin, and M. Molnar. An Efficient Multicast Protection Scheme based on a Dual-Forest. Irisa Research Report number 1786, March 2006.

[SCM06b] M. Y. Saidi, B. Cousin, and M. Molnar. Improved Dual-Forest for Multicast Protection. In *Second Conference on Next Generation Internet Design and Engineering Conference*, Valence (Spain), April 2006.

[SCM06c] M. Y. Saidi, B. Cousin, and M. Molnar. Protection Remontante des Communications Multicast. In *Conférence MajecSTIC 2006*, Lorient (France), November 2006.

[SM03a] N. K. Singhal and B. Mukherjee. Algorithms for Provisioning Survivable Multicast Sessions against Link Failures in Mesh Networks. In *Proceedings of the international workshop on distributed computing*, Kolkata (India), December 2003.

[SM03b] N. K. Singhal and B. Mukherjee. Protecting Multicast Sessions in WDM Optical Mesh Networks. *Journal of Lightwave Technology*, 21(4) :884–892, 2003.

[SOM06] N. K. Singhal, C. Ou, and B. Mukherjee. Cross-sharing vs. Self-sharing Trees for Protecting Multicast Sessions in Mesh Networks. *Computer Networks*, 50(2) :200–206, 2006.

[SSM03] N. K. Singhal, H. Sahasrabuddhe, and B. Mukherjee. Provisioning of Survivable Multicast Sessions Against Single Link Failures in Optical WDM Mesh Networks. *Journal of Lightwave Technology*, 21(11) :2587–2594, 2003.

[TM80] H. Takahashi and A. Matsuyama. An Approximate Solution for the Steiner Problem in Graphs. *Mathematica Japonica*, 24(6), 1980.

[VC02] J. P. Vasseur and A. Charny. Distinguish a Link from a Node failure using RSVP Hello Extensions. Internet Draft draft-vasseur-mpls-linknode-failure-00.txt, IETF, November 2002.

[VCLF+04] J. P. Vasseur, A. Charny, F. Le Faucheur, J. Achirica, and J. L. Le Roux. Framework for PCE-based MPLS-TE Fast Reroute Backup Path Computation. Internet Draft draft-leroux-pce-backup-comp-frwk-00.txt, IETF, July 2004.

[XCX+06] D. Xu, Y. Chen, Y. Xiong, C. Qiao, and X. He. On the complexity of and algorithms for finding the shortest path with a disjoint counterpart. *IEEE/ACM Transactions on Networking*, 14(1) :147–158, 2006.

[YJ05] S. Yuan and J. P. Jue. Dynamic Lightpath Protection in WDM Mesh Networks Under Wavelength-Continuity and Risk-Disjoint Constraints. *Computer Networks and ISDN Systems*, 48(2) :91–112, 2005.

Table des figures

Liste des algorithmes

Résumé

Dans cette thèse, nous proposons plusieurs algorithmes et heuristiques efficaces pour le placement distribué des chemins de secours. Ces derniers sont calculés en ligne et permettent de protéger localement des communications unicast et multicast contre les pannes simples dans un réseau MPLS.

En plus de leur rapidité de récupération après une panne, nos algorithmes et heuristiques sont faciles à déployer (ils ne nécessitent que de très légères extensions aux protocoles de signalisation et/ou protocoles de routage) et ils augmentent la disponibilité de la bande passante en partageant efficacement la bande passante entre les chemins.

Deux stratégies de partage de la bande passante sont ainsi employées, étudiées et comparées : le partage restreint de la bande passante et le partage global de la bande passante. Dans la première stratégie, seuls les chemins de secours peuvent partager leur bande passante alors que dans la seconde stratégie, le partage de la bande passante est étendu et est appliqué entre les chemins primaires et les chemins de secours d'un côté, et entre les chemins de secours d'un autre côté.

Mots clés : Réseaux informatiques ; protection locale ; partage de la bande passante ; SRLG ; placement des chemins de secours ; routage ; MPLS ; unicast ; multicast.

Abstract

In this thesis, we propose efficient algorithms and heuristics for the distributed and on-line computation of backup paths. These paths provide local protection against the simple failures to unicast and multicast communications in a MPLS network.

Our algorithms and heuristics recover quickly from failures, easy to be deployed (they require only slight extensions to the signaling protocols and/or routing protocols) and they increase the bandwidth availability by applying the bandwidth sharing between the paths.

Two bandwidth sharing strategies are studied and compared : restricted bandwidth sharing and global bandwidth sharing. In the first strategy, only the backup paths can share their bandwidth. In the second strategy, the bandwidth sharing is extended and applied between the primary and backup paths and between the backup paths themselves.

Key words: Networks ; local protection ; bandwidth sharing ; SRLG ; backup path computation ; routing ; MPLS ; unicast ; multicast.