

Groupe de travail Réseau
Request for Comments : 3031
 Catégorie : En cours de normalisation
 Traduction Claude Brière de L'Isle

E. Rosen, Cisco Systems, Inc.
 A. Viswanathan, Force10 Networks, Inc.
 R. Callon, Juniper Networks, Inc.
 janvier 2001

Architecture de commutation d'étiquettes multi protocoles

Statut de ce mémoire

Le présent document spécifie un protocole Internet en cours de normalisation pour la communauté de l'Internet, et appelle à des discussions et des suggestions pour son amélioration. Prière de se reporter à l'édition actuelle du STD 1 "Normes des protocoles officiels de l'Internet" pour connaître l'état de normalisation et le statut de ce protocole. La distribution du présent mémoire n'est soumise à aucune restriction.

Notice de copyright

Copyright (C) The Internet Society (2001). Tous droits réservés.

Résumé

Le présent document spécifie l'architecture pour la commutation d'étiquettes multi protocole (MPLS, *Multiprotocol Label Switching*).

Table des matières

1. Spécification des exigences.....	2
2. Introduction à MPLS.....	2
2.1 Généralités.....	2
2.2 Terminologie.....	3
2.3 Acronymes et abréviations.....	5
2.4 Remerciements.....	5
3. Bases de MPLS.....	5
3.1 Étiquettes.....	5
3.2 LSR vers l'amont et vers l'aval.....	6
3.3 Paquets étiquetés.....	6
3.4 Allocation et distribution des étiquettes.....	6
3.5 Attributs d'un lien d'étiquettes.....	6
3.6 Protocoles de distribution d'étiquettes.....	6
3.7 Vers l'aval non sollicité ou vers l'aval à la demande.....	7
3.8 Mode de rétention d'étiquette.....	7
3.9 Pile d'étiquettes.....	7
3.10 Entrée de transmission d'étiquette de prochain bond (NHLFE).....	8
3.11 Transposition d'étiquette entrante (ILM).....	8
3.12 Transposition de FEC à NHLFE (FTN).....	8
3.13 Échange d'étiquettes.....	8
3.14 Portée et unicité des étiquettes.....	9
3.15 Chemin commuté d'étiquette (LSP), LSP d'entrée, LSP de sortie.....	9
3.16 Saut de l'avant dernier bond.....	10
3.17 LSP de prochain bond.....	11
3.18 Étiquettes entrantes invalides.....	11
3.19 Contrôle de LSP : ordonné ou indépendant.....	11
3.20 Agrégation.....	12
3.21 Sélection du chemin.....	13
3.22 Manque d'étiquette sortante.....	13
3.23 Durée de vie (TTL).....	13
3.24 Contrôle de boucle.....	14
3.25 Codages des étiquettes.....	15
3.26 Fusion d'étiquettes.....	16
3.27 Tunnels et hiérarchie.....	18
3.28 Transport du protocole de distribution d'étiquettes.....	20
3.29 Pourquoi plus d'un protocole de distribution d'étiquettes ?.....	20
3.30 Diffusion groupée.....	21

4. Quelques applications de MPLS.....	21
4.1 MPLS et trafic acheminé bond par bond.....	21
4.2 MPLS et LSP à acheminement explicite.....	24
4.3 Piles d'étiquettes et homologie implicite.....	24
4.4 MPLS et acheminement multi chemins.....	25
4.5 Arborescences de LSP comme entités de multipoint à point.....	25
4.6 LSP tunnel entre routeurs frontière BGP.....	25
4.7 Autres utilisations des LSP tunnels acheminés bond par bond.....	26
4.8 MPLS et diffusion groupée.....	26
5. Procédures de distribution d'étiquettes (bond par bond).....	27
5.1 Procédures d'annonce et d'utilisation des étiquettes.....	27
5.2 Schémas MPLS : combinaisons de procédures prises en charge.....	30
6. Considérations pour la sécurité.....	32
7. Propriété intellectuelle.....	33
8. Adresse des auteurs.....	33
9. Références.....	33
10. Déclaration complète de droits de reproduction.....	33

1. Spécification des exigences

Les mots clés "DOIT", "NE DOIT PAS", "EXIGE", "DEVRA", "NE DEVRA PAS", "DEVRAIT", "NE DEVRAIT PAS", "RECOMMANDE", "PEUT", et "FACULTATIF" dans le présent document sont à interpréter comme décrit dans la RFC2119.

2. Introduction à MPLS

Le présent document spécifie l'architecture de la commutation d'étiquettes multi protocoles (MPLS, *Multiprotocol Label Switching*).

Noter que l'utilisation de MPLS pour la diffusion groupée est laissée à des travaux ultérieurs.

2.1 Généralités

Lorsque un paquet d'un protocole de couche réseau sans connexion voyage d'un routeur au suivant, chaque routeur prend une décision de transmission indépendante pour ce paquet. C'est à dire que chaque routeur analyse l'en-tête du paquet, et chaque routeur fait fonctionner un algorithme d'acheminement de couche réseau. Chaque routeur choisit indépendamment un prochain bond pour le paquet, sur la base de son analyse de l'en-tête du paquet et du résultat du fonctionnement de l'algorithme d'acheminement.

Les en-têtes de paquet contiennent considérablement plus d'informations qu'il n'est nécessaire pour simplement choisir le prochain bond. Le choix du prochain bond peut être vu comme la composante de deux fonctions. La première fonction partage la totalité de l'ensemble des paquets possibles en un ensemble de "classes d'équivalence de transmission" (FEC, *Forwarding Equivalence Classes*). La seconde transpose chaque FEC sur un prochain bond. Dans la mesure où la décision de transmission est concernée, différents paquets qui sont transposés sur la même FEC sont indistincts. Tous les paquets qui appartiennent à une FEC particulière et qui voyagent à partir un certain nœud vont suivre le même chemin (ou si on utilise certaines sortes d'acheminement multi-chemins, ils vont tous suivre un des ensembles de chemins associés à cette FEC).

En transmission IP conventionnelle, un routeur particulier va normalement considérer deux paquets comme étant dans la même FEC si il y a un certain préfixe d'adresse X dans les tableaux d'acheminement de ce routeur tel que X soit la "plus longue correspondance" pour l'adresse de destination de chaque paquet. Lorsque le paquet traverse le réseau, chaque bond réexamine à son tour le paquet et l'affecte à une FEC.

Dans MPLS, l'affectation d'un paquet particulier à une certaine FEC n'est faite qu'une seule fois, lorsque le paquet entre dans le réseau. La FEC à laquelle le paquet est affecté est codée comme une courte valeur de longueur fixe appelée une "étiquette". Lorsque un paquet est transmis à son prochain bond, l'étiquette est envoyée avec lui ; c'est-à-dire que les paquets sont "étiquetés" avant qu'ils soient transmis.

Aux bonds suivants, il n'y a pas d'autre analyse de l'en-tête de couche réseau du paquet. L'étiquette est utilisée comme un

indice dans un tableau qui spécifie le prochain bond, et une nouvelle étiquette. La vieille étiquette est remplacée par la nouvelle étiquette, et le paquet est transmis à son prochain bond.

Selon le paradigme de la transmission MPLS, une fois qu'un paquet est affecté à une FEC, aucune autre analyse d'en-tête n'est effectuée par les routeurs suivants ; toute la transmission est conduite par les étiquettes. Cela présente un certain nombre d'avantages sur la transmission de couche réseau conventionnelle.

- La transmission MPLS peut être faite par des commutateurs qui sont capables de faire la recherche et le remplacement d'étiquette, mais ne sont capables ni d'analyser les en-têtes de couche réseau, ni de le faire à la vitesse adéquate.
- Comme un paquet est affecté à une FEC lorsque il entre dans le réseau, le routeur d'entrée peut utiliser, pour déterminer l'affectation, toute information qu'il possède sur le paquet, même si cette information ne peut pas être tirée d'un en-tête de couche réseau. Par exemple, les paquets qui arrivent sur des accès différents peuvent être affectés à des FEC différentes. La transmission conventionnelle, d'un autre côté, peut seulement considérer les informations qui voyagent avec le paquet dans son en-tête.
- Un paquet qui entre dans le réseau sur un routeur particulier peut être étiqueté différemment du même paquet entrant dans le réseau sur un routeur différent, et il en résulte que les décisions de transmission qui dépendent du routeur d'entrée peuvent être prises facilement. Cela ne peut pas être fait avec la transmission conventionnelle, car l'identité du routeur d'entrée d'un paquet ne voyage pas avec le paquet.
- Les considérations qui déterminent comment un paquet est affecté à une FEC peuvent devenir de plus en plus compliquées, sans aucun impact du tout sur les routeurs qui transmettent simplement les paquets étiquetés.
- Il est parfois souhaitable de forcer un paquet à suivre un chemin particulier qui est explicitement choisi au moment, ou avant le moment, où le paquet entre dans le réseau, plutôt que d'être choisi par l'algorithme normal d'acheminement dynamique lorsque le paquet traverse le réseau. Cela peut relever de la politique, ou bien en soutien de l'ingénierie du trafic. Dans la transmission conventionnelle, cela exige que le paquet porte avec lui un codage du chemin qu'il doit suivre ("acheminement de source"). Dans MPLS, une étiquette peut être utilisée pour représenter le chemin, de sorte que l'identité du chemin explicite n'a pas besoin d'être portée avec le paquet.

Certains routeurs analysent l'en-tête de couche réseau d'un paquet non seulement pour choisir le prochain bond du paquet, mais aussi pour déterminer la "préséance" ou la "classe de service" d'un paquet. Ils peuvent alors appliquer des seuils d'élimination différents ou des disciplines de programmation différentes aux différents paquets. MPLS permet (mais n'exige pas) que la préséance ou la classe de service soit pleinement ou partiellement déduite de l'étiquette. Dans ce cas, on peut dire que l'étiquette représente la combinaison d'une FEC et d'une préséance ou classe de service.

MPLS veut dire commutation d'étiquette "multi protocole", parce que ses techniques sont applicables à TOUT protocole de couche réseau. Dans le présent document, cependant, on se concentre sur l'utilisation de IP comme protocole de couche réseau.

Un routeur qui prend en charge MPLS est appelé un "routeur de commutation d'étiquettes", LSR (*Label Switching Router*).

2.2 Terminologie

Ce paragraphe fait un survol conceptuel général des termes utilisés dans le présent document. Certains de ces termes sont plus précisément définis dans des sections ultérieures du document.

DLCI (*Data Link Circuit Identifier*) c'est une étiquette utilisée dans les réseaux en relais de trame pour identifier les circuits de relais de trame.

classe d'équivalence de transmission groupe de paquets IP qui sont transmis de la même manière (par exemple, sur le même chemin, avec le même traitement de transmission).

fusion de trame fusion d'étiquette, quand elle est appliquée au fonctionnement sur des supports à base de trames, de sorte que le problème potentiel d'entrelacement de cellules ne se pose pas.

étiquette identifiant court de longueur fixe physiquement contigu qui est utilisé pour identifier une FEC, normalement de signification locale.

fusion d'étiquette remplacement de plusieurs étiquettes entrantes pour une FEC particulière par une seule étiquette

sortante.

- échange d'étiquette opération de transmission de base consistant à examiner une étiquette entrante pour déterminer l'étiquette sortante, son encapsulation, son accès, et les autres informations sur le traitement des données.
- échange d'étiquetage paradigme de transmission qui permet une transmission réduite au minimum des données par l'utilisation d'étiquettes pour identifier les classes de paquets de données qui sont traitées de façon non distinguable lors de la transmission.
- bond d'étiquettes commutées bond entre deux nœuds MPLS, sur lesquels la transmission est effectuée en utilisant des étiquettes.
- chemin d'étiquettes commutées chemin à travers un ou plusieurs LSR à un niveau de la hiérarchie, suivi par un paquet dans une FEC particulière.
- routeur de commutation d'étiquettes nœud MPLS qui est capable de transmettre des paquets natifs de couche 3.
- couche 2 couche de protocole en dessous de la couche 3 (qui offre donc les services utilisés par la couche 3). La transmission, lorsque elle est faite par l'échange d'étiquettes courtes de longueur fixe, survient à la couche 2 sans considérer si l'étiquette examinée est un VPI/VCI ATM, un DLCI de relais de trame, ou une étiquette MPLS.
- couche 3 couche de protocole à laquelle IP et ses protocoles d'acheminement associés opèrent, couche de liaison synonyme de couche 2.
- détection de boucle méthode de traitement des boucles dans laquelle il est permis d'établir des boucles, et les données peuvent être transmises sur la boucle, mais où la boucle est détectée ultérieurement.
- prévention de boucle méthode de traitement de boucles dans laquelle les données ne sont jamais transmises sur une boucle.
- pile d'étiquettes ensemble ordonné d'étiquettes
- point de fusion nœud auquel est effectuée la fusion d'étiquette.
- domaine MPLS ensemble contigu de nœuds qui opère l'acheminement et la transmission MPLS et qui sont aussi dans un domaine d'acheminement ou administratif.
- nœud bordure MPLS nœud MPLS qui connecte un domaine MPLS avec un nœud qui est en dehors du domaine, soit parce qu'il n'utilise pas MPLS, soit parce qu'il est dans un domaine différent. Noter que si un LSR a un hôte voisin qui ne fonctionne pas avec MPLS, ce LSR est un nœud bordure MPLS.
- nœud MPLS de sortie nœud bordure MPLS dans son rôle de traitement du trafic lorsque il quitte un domaine MPLS.
- nœud MPLS d'entrée nœud bordure MPLS dans son rôle de traitement du trafic entrant dans un domaine MPLS.
- étiquette MPLS étiquette qui est portée dans un en-tête de paquet, et qui représente la FEC du paquet.
- nœud MPLS nœud qui fonctionne avec MPLS. Un nœud MPLS connaît les protocoles de contrôle MPLS, qui va fonctionner sur un ou plusieurs protocoles d'acheminement de couche 3, et qui sera capable de transmettre les paquets sur la base des étiquettes. Un nœud MPLS peut facultativement être aussi capable de transmettre des paquets natifs de L3.
- Commutation d'étiquettes multi protocoles groupe de travail de l'IETF et les travaux associés au groupe de travail.
- couche réseau synonyme de couche 3.
- pile synonyme de pile d'étiquettes.
- chemin commuté synonyme de chemin d'étiquettes commutées
- circuit virtuel circuit utilisé par une technologie de couche 2 orientée connexion telle que ATM ou le relais de trame qui exige la conservation des informations d'état dans les commutateurs de couche 2.

- fusion de VC fusion d'étiquette où l'étiquette MPLS est portée dans le champ VCI ATM (ou le champ VPI/VCI combiné) de façon à permettre que plusieurs VC fusionnent en un seul VC.
- fusion de VP fusion d'étiquette où l'étiquette MPLS est portée dans le champ VCI ATM, de façon à permettre que plusieurs VP fusionnent en un seul VP. Dans ce cas, deux cellules n'auront la même valeur de VCI que si elles proviennent du même nœud. Cela permet que des cellules provenant de différentes sources soient distinguées via le VCI.
- VPI/VCI étiquette utilisée dans les réseaux ATM pour identifier les circuits.

2.3 Acronymes et abréviations

ATM	(<i>Asynchronous Transfer Mode</i>) mode de transfert asynchrone
BGP	(<i>Border Gateway Protocol</i>) protocole de passerelle frontière
DLCI	(<i>Data Link Circuit Identifier</i>) identifiant de circuit de liaison de données
FEC	(<i>Forwarding Equivalence Class</i>) classe d'équivalence de transmission
FTN	(<i>FEC to NHLFE Map</i>) transposition de FEC en NHLFE
IGP	(<i>Interior Gateway Protocol</i>) protocole de passerelle intérieure
ILM	(<i>Incoming Label Map</i>) transposition d'étiquette entrante
IP	(<i>Internet Protocol</i>) protocole Internet
LDP	(<i>Label Distribution Protocol</i>) protocole de distribution d'étiquettes
L2	(<i>Layer 2</i>) couche 2
L3	(<i>Layer 3</i>) couche 3
LSP	(<i>Label Switched Path</i>) chemin d'étiquettes commutées
LSR	(<i>Label Switching Router</i>) routeur de commutation d'étiquettes
MPLS	(<i>MultiProtocol Label Switching</i>) commutation d'étiquettes multi protocoles
NHLFE	(<i>Next Hop Label Forwarding Entry</i>) entrée de transmission d'étiquette du prochain bond
SVC	(<i>Switched Virtual Circuit</i>) circuit virtuel commuté
SVP	(<i>Switched Virtual Path</i>) chemin virtuel commuté
TTL	(<i>Time-To-Live</i>) durée de vie
VC	(<i>Virtual Circuit</i>) circuit virtuel
VCI	(<i>Virtual Circuit Identifier</i>) identifiant de circuit virtuel
VP	(<i>Virtual Path</i>) chemin virtuel
VPI	(<i>Virtual Path Identifier</i>) identifiant de chemin virtuel

2.4 Remerciements

Les idées et le texte du présent document ont été collectés d'un grand nombre de sources et des commentaires ont été reçus. Nous tenons à remercier Rick Boivie, Paul Doolan, Nancy Feldman, Yakov Rekhter, Vijay Srinivasan, et George Swallow de leurs apports et de leurs idées.

3. Bases de MPLS

Dans cette section, on introduit les concepts de base de MPLS et on décrit l'approche générale à utiliser.

3.1 Étiquettes

Une étiquette est un identifiant court, de longueur fixe, de signification locale qui est utilisé pour identifier une FEC. L'étiquette mise sur un certain paquet représente la classe d'équivalence de transmission à laquelle ce paquet est affecté.

En bref, un paquet est affecté à une FEC sur la base (complète ou partielle) de son adresse de destination de couche réseau. Cependant, l'étiquette n'est jamais un codage de cette adresse.

Si Ru et Rd sont des LSR, ils peuvent s'accorder pour que lorsque Ru transmet un paquet à Rd, Ru ne va étiqueter le paquet avec la valeur d'étiquette L que si et seulement si le paquet est membre de la FEC F. C'est-à-dire qu'ils peuvent s'accorder sur un "lien" entre l'étiquette L et la FEC F pour les paquets qui se déplacent de Ru à Rd. Il résulte d'un tel accord que L

devient "l'étiquette sortante" de Ru qui représente la FEC F, et L devient "l'étiquette entrante" de Rd qui représente la FEC F.

Noter que L ne représente pas nécessairement la FEC F pour tous les paquets autres que ceux qui sont envoyés de Ru à Rd. L est une valeur arbitraire dont le lien à F est local pour Ru et Rd.

Quand on parle ci-dessus de paquets "envoyés" de Ru à Rd, cela n'implique ni que le paquet soit généré par Ru ni que sa destination soit Rd. On veut plutôt dire que cela inclut les paquets qui sont des "paquets en transit" à l'un des LSR ou aux deux.

Il peut parfois être difficile ou même impossible à Rd de dire, à l'arrivée d'un paquet qui porte l'étiquette L, que l'étiquette L a été placée dans le paquet par Ru, plutôt que par quelque autre LSR. (Cela sera normalement le cas lorsque Ru et Rd ne sont pas des voisins directs.) Dans de tels cas, Rd doit s'assurer que le lien de l'étiquette à la FEC est biunivoque. C'est-à-dire que Rd NE DOIT PAS s'accorder avec Ru1 pour lier L à la FEC F1, tout en se mettant aussi d'accord avec un autre LSR Ru2 pour lier L à une différente FEC F2, SAUF SI Rd peut toujours dire, lorsque il reçoit un paquet avec l'étiquette entrante L, si l'étiquette a été mise sur le paquet par Ru1 ou si elle a été mise par Ru2.

Il est de la responsabilité de chaque LSR de s'assurer qu'il peut interpréter de façon univoque ses étiquettes entrantes.

3.2. LSR vers l'amont et vers l'aval

Supposons que Ru et Rd se sont accordés pour lier l'étiquette L à la FEC F, pour les paquets envoyés de Ru à Rd. Par rapport à cette liaison, Ru est ensuite "le LSR amont", et Rd est "le LSR aval".

Dire qu'un nœud est vers l'amont et que l'autre est vers l'aval par rapport à un certain lien signifie seulement qu'une certaine étiquette représente une FEC particulière dans les paquets qui voyagent du nœud amont vers le nœud aval. Cela N'EST PAS destiné à impliquer que les paquets de cette FEC vont en fait être acheminés du nœud amont vers le nœud aval.

3.3 Paquets étiquetés

Un "paquet étiqueté" est un paquet dans lequel une étiquette a été codée. Dans certains cas, l'étiquette réside dans un en-tête d'encapsulation qui existe spécifiquement à cette fin. Dans d'autres cas, l'étiquette peut résider dans un en-tête existant de couche liaison des données ou réseau, pour autant qu'il y ait un champ disponible pour cela. La technique particulière de codage à utiliser doit être acceptée par les deux entités, celle qui code l'étiquette et celle qui la décode.

3.4 Allocation et distribution des étiquettes

Dans l'architecture MPLS, la décision de lier une étiquette L particulière à une FEC particulière F est prise par le LSR qui est AVANT par rapport à ce lien. Le LSR aval informe alors le LSR amont du lien. Les étiquettes sont donc "affectées par l'aval", et les liens d'étiquette sont distribués dans la direction "de l'aval vers l'amont".

Si un LSR a été conçu de telle sorte qu'il ne puisse examiner que les étiquettes qui entrent dans une certaine gamme numérique, il a alors seulement besoin de s'assurer qu'il ne se lie qu'à des étiquettes qui sont dans cette gamme.

3.5 Attributs d'un lien d'étiquettes

Un lien particulier de l'étiquette L à la FEC F, distribué par Rd à Ru, peut avoir des "attributs" associés. Si Ru, agissant comme un LSR aval, distribue aussi un lien d'une étiquette à la FEC F, alors, sous certaines conditions, il peut être aussi nécessaire de distribuer l'attribut correspondant qu'il a reçu de Rd.

3.6 Protocoles de distribution d'étiquettes

Un protocole de distribution d'étiquettes est un ensemble de procédures par lesquelles un LSR en informe un autre des liens d'étiquette/FEC qu'il a fait. Deux LSR qui utilisent un protocole de distribution d'étiquettes pour échanger des informations de lien d'étiquette/FEC sont appelés des "homologues de distribution d'étiquettes" par rapport aux

informations de liens qu'ils échangent. Si deux LSR sont des homologues de distribution d'étiquettes, on parlera d'eux comme étant dans une relation "d'adjacence de distribution d'étiquettes".

(Note : Deux LSR peuvent être des homologues de distribution d'étiquettes par rapport à un ensemble de liens, mais pas par rapport à un autre ensemble de liens.)

Le protocole de distribution d'étiquettes comporte aussi toutes les négociations dans lesquelles deux homologues de distribution d'étiquettes ont besoin de s'engager afin de d'apprendre leurs capacités MPLS réciproques.

L'architecture ne suppose pas qu'il y a un seul protocole de distribution d'étiquettes. En fait, un certain nombre de différents protocoles de distribution d'étiquettes sont normalisés. Les protocoles existants ont été étendus de telle sorte que la distribution d'étiquettes puisse être portée par eux (voir, par exemple, [RFC3107], [RFC3209]). De nouveaux protocoles ont aussi été définis dans le but explicite de distribuer les étiquettes (voir, par exemple, les [RFC3036], [RFC3212]).

Dans le présent document, on essaye d'utiliser l'acronyme "LDP" pour se référer spécifiquement au protocole défini dans la [RFC3036] ; quand on parle de protocoles de distribution d'étiquettes en général, on essaye d'éviter l'acronyme.

3.7 Vers l'aval non sollicité ou vers l'aval à la demande

L'architecture MPLS permet à un LSR de demander explicitement, de son prochain bond pour une FEC particulière, un lien d'étiquette pour cette FEC. C'est ce qu'on appelle une distribution d'étiquettes "vers l'aval à la demande".

L'architecture MPLS permet aussi qu'un LSR distribue des liens avec des LSR qui ne les ont pas explicitement demandés. C'est ce qu'on appelle la distribution d'étiquettes "non sollicitée vers l'aval".

Il est prévu que certaines mises en œuvre de MPLS ne fournissent que la distribution d'étiquettes vers l'aval à la demande, et que certaines ne fournissent que la distribution d'étiquettes non sollicitée vers l'aval, et que certaines fournissent les deux. Ce qui est fourni peut dépendre des caractéristiques des interfaces qui sont prises en charge par une mise en œuvre particulière. Cependant, ces deux techniques de distribution d'étiquettes peuvent être utilisées dans le même réseau en même temps. Sur toute adjacence de distribution d'étiquettes, le LSR amont et le LSR aval doivent se mettre d'accord sur la technique à utiliser.

3.8 Mode de rétention d'étiquette

Un LSR Ru peut recevoir (ou avoir reçu) un lien d'étiquette pour une FEC particulière d'un LSR Rd, même si Rd n'est pas le prochain bond de Ru (ou n'est plus le prochain bond de Ru) pour cette FEC.

Ru a alors le choix de garder trace de tels liens, ou d'éliminer de tels liens. Si Ru garde trace de ces liens, il peut alors commencer immédiatement à utiliser le lien à nouveau si Rd devient finalement son prochain bond pour la FEC en question. Si Ru élimine ces liens, alors si Rd devient plus tard le prochain bond, le lien devra être acquis à nouveau.

Si un LSR prend en charge le "mode libéral de rétention d'étiquette", il conserve les liens entre une étiquette et une FEC qui sont reçus des LSR qui ne sont pas son prochain bond pour cette FEC. Si un LSR prend en charge le "mode prudent de rétention d'étiquette", il élimine ces liens.

Le mode libéral de rétention d'étiquette permet une adaptation plus rapide aux changements d'acheminement, mais le mode prudent de rétention d'étiquette exige d'un LSR qu'il conserve beaucoup moins d'étiquettes.

3.9 Pile d'étiquettes

Jusqu'à présent, on a présenté les choses comme si un paquet étiqueté portait seulement une étiquette. Comme on va le voir, il est utile d'avoir un modèle plus général dans lequel un paquet étiqueté porte un certain nombre d'étiquettes, organisées comme une pile au dernier entré, premier sorti. On appelle cela une "pile d'étiquettes".

Bien que, comme on va le voir, MPLS prenne en charge une hiérarchie, le traitement d'un paquet étiqueté est complètement indépendant du niveau de la hiérarchie. Le traitement est toujours fondé sur l'étiquette supérieure, sans considération pour la possibilité qu'un certain nombre d'autres étiquettes puissent avoir été "au dessus d'elle" dans le passé, ou qu'un certain nombre d'autres étiquettes puissent être en dessous d'elle à présent.

Un paquet non étiqueté peut être vu comme un paquet dont la pile d'étiquette est vide (c'est-à-dire, dont la pile d'étiquettes

a une profondeur de 0).

Si la pile d'étiquettes d'un paquet est de profondeur m , on se réfère à l'étiquette du bas de la pile comme étiquette de niveau 1, à l'étiquette au dessus d'elle (si elle existe) comme étiquette de niveau 2, et à l'étiquette au sommet de la pile comme étiquette de niveau m .

L'utilité de la pile d'étiquettes va devenir clair quand on aura introduit la notion de tunnel LSP et de hiérarchie MPLS (paragraphe 3.27).

3.10 Entrée de transmission d'étiquette de prochain bond (NHLFE)

L'entrée de transmission d'étiquette de prochain bond (NHLFE, *Next Hop Label Forwarding Entry*) est utilisée lors de la transmission d'un paquet étiqueté. Elle contient les informations suivantes :

1. le prochain bond du paquet
2. l'opération à effectuer sur la pile d'étiquettes du paquet ; c'est une des opérations suivantes :
 - a) remplacer l'étiquette du sommet de la pile d'étiquettes par une nouvelle étiquette spécifiée
 - b) sauter la pile d'étiquettes
 - c) remplacer l'étiquette du sommet de la pile d'étiquettes par une nouvelle étiquette spécifiée, et puis pousser une ou plusieurs nouvelles étiquettes spécifiées dans la pile d'étiquettes.Elle peut aussi contenir :
 - d) l'encapsulation de liaison de données à utiliser lors de la transmission du paquet,
 - e) la façon de coder la pile d'étiquettes lors de la transmission du paquet,
 - f) toutes autres informations nécessaires afin de traiter correctement le paquet.

Noter qu'à un LSR donné, le "prochain bond" du paquet pourrait être ce LSR lui-même. Dans ce cas, le LSR aura besoin de sauter l'étiquette du sommet, et puis de "transmettre" le paquet résultant à lui-même. Il devra alors prendre une autre décision de transmission, en se fondant sur ce qui reste après avoir sauté la pile d'étiquettes. Cela peut être encore un paquet étiqueté, ou cela peut être le paquet IP natif.

Cela implique que dans certains cas, le LSR peut avoir besoin de travailler sur l'en-tête IP afin de transmettre le paquet.

Si le "prochain bond" du paquet est le LSR en cours, l'opération sur la pile d'étiquettes DOIT être de "sauter la pile".

3.11 Transposition d'étiquette entrante (ILM)

La transposition d'étiquette entrante (ILM, *Incoming Label Map*) transpose chaque étiquette entrante en un ensemble de NHLFE. Elle est utilisée lors de la transmission de paquets qui arrivent comme paquets étiquetés.

Si la ILM transpose une étiquette particulière en un ensemble de NHLFE qui contiennent plus d'un élément, exactement un élément de l'ensemble doit être choisi avant de transmettre le paquet. Les procédures pour choisir un élément dans l'ensemble sortent du domaine d'application du présent document. Avoir l'ILM qui transpose une étiquette en un ensemble contenant plus d'une NHLFE peut être utile si, par exemple, on souhaite faire un équilibrage de charge sur plusieurs chemins de coût égal.

3.12 Transposition de FEC à NHLFE (FTN)

La transposition de FEC à NHLFE (FTN) transpose chaque FEC en un ensemble de NHLFE. Elle est utilisée lors de la transmission de paquets qui arrivent sans étiquette, mais qui sont à étiqueter avant transmission.

Si la FTN transpose une étiquette particulière en un ensemble de NHLFE qui contiennent plus d'un élément, exactement un élément de l'ensemble doit être choisi avant la transmission du paquet. Les procédures pour choisir un élément dans l'ensemble sortent du domaine d'application du présent document. Avoir la FTN qui transpose une étiquette en un ensemble contenant plus d'une NHLFE peut être utile si, par exemple, on souhaite faire un équilibrage de charge sur plusieurs chemins de coût égal.

3.13 Échange d'étiquettes

L'échange d'étiquettes est l'utilisation des procédures suivantes pour transmettre un paquet.

Afin de transmettre un paquet étiqueté, un LSR examine l'étiquette au sommet de la pile d'étiquettes. Il utilise la ILM pour transposer cette étiquette en une NHLFE. En utilisant les informations de la NHLFE, il détermine où transmettre le paquet,

et effectue une opération sur la pile d'étiquette du paquet. Il code ensuite la nouvelle pile d'étiquettes dans le paquet, et transmet le résultat.

Afin de transmettre un paquet non étiqueté, un LSR analyse l'en-tête de couche réseau, pour déterminer la FEC du paquet. Il utilise ensuite la FTN pour transposer cela en une NHLFE. En utilisant les informations qui sont dans la NHLFE, il détermine où transmettre le paquet, et effectue une opération sur la pile d'étiquettes du paquet. (Sauter la pile d'étiquettes serait, bien sûr, illégal dans ce cas.) Il code ensuite la nouvelle pile d'étiquettes dans le paquet, et transmet le résultat.

Il est important de noter que lorsque l'échange d'étiquettes est utilisé, le prochain bond est toujours tiré de la NHLFE ; cela peut dans certains cas être différent de ce que le prochain bond aurait été si on n'utilisait pas MPLS.

3.14 Portée et unicité des étiquettes

Un certain LSR Rd peut lier l'étiquette L1 à la FEC F, et distribuer ce lien à l'homologue de distribution d'étiquettes Ru1. Rd peut aussi lier l'étiquette L2 à la FEC F, et distribuer ce lien à l'homologue de distribution d'étiquettes Ru2. Savoir si $L1 = L2$ ou non n'est pas déterminé par l'architecture ; c'est une affaire locale.

Un certain LSR Rd peut lier l'étiquette L à la FEC F1, et distribuer ce lien à l'homologue de distribution d'étiquettes Ru1. Rd peut aussi lier l'étiquette L à la FEC F2, et distribuer ce lien à l'homologue de distribution d'étiquettes Ru2. Si (et seulement si) Rd peut dire, quand il reçoit un paquet dont l'étiquette de sommet est L, si l'étiquette a été mise là par Ru1 ou par Ru2, alors l'architecture n'exige pas que $F1 = F2$. Dans de tels cas, on peut dire que Rd utilise un "espace d'étiquettes" différent pour les étiquettes qu'il distribue à Ru1 que pour les étiquettes qu'il distribue à Ru2.

En général, Rd peut seulement dire si c'est Ru1 ou Ru2 qui a mis une certaine valeur d'étiquette L au sommet de la pile d'étiquettes si les conditions suivantes sont réunies :

- Ru1 et Ru2 sont les seuls homologues de distribution d'étiquette auxquels Rd a distribué un lien de la valeur d'étiquette L, et
- Ru1 et Ru2 sont chacun directement connecté à Rd via une interface point à point.

Lorsque ces conditions sont réunies, un LSR peut utiliser des étiquettes qui ont une portée "pour l'interface", c'est-à-dire, qui ne sont univoques que pour l'interface. On peut dire que le LSR utilise un "espace d'étiquette par interface". Lorsque ces conditions ne sont pas réunies, les étiquettes doivent être univoques sur le LSR qui les a affectées, et on peut dire que le LSR utilise un "espace d'étiquette par plateforme".

Si un certain LSR Rd est rattaché à un LSR Ru particulier sur deux interfaces point à point, Rd peut alors distribuer à Ru un lien de l'étiquette L à la FEC F1, ainsi qu'un lien de l'étiquette L à la FEC F2, $F1 \neq F2$, si et seulement si chaque lien n'est valide que pour les paquets que Ru envoie à Rd sur une des deux interfaces particulières. Dans tous les autres cas, Rd NE DOIT PAS distribuer à Ru de liens de la même valeur d'étiquette de deux FEC différentes.

Cette interdiction tient même si les liens sont considérés comme étant à différents "niveaux de la hiérarchie". Dans MPLS, il n'y a pas la notion d'un espace d'étiquettes différent pour les différents niveaux de la hiérarchie ; lors de l'interprétation d'une étiquette, le niveau de l'étiquette est sans conséquence.

La question se pose de savoir si il est possible à un LSR d'utiliser plusieurs espaces d'étiquettes par plateforme, ou d'utiliser plusieurs espaces d'étiquette par interface pour la même interface. Ceci n'est pas interdit par l'architecture. Cependant, dans de tels cas, le LSR doit avoir des moyens, non spécifiés par l'architecture, de déterminer, pour une étiquette entrante particulière, à quel espace d'étiquette cette étiquette appartient. Par exemple, la [RFC3032] spécifie qu'un espace d'étiquette différent est utilisé pour les paquets en envoi individuel et pour les paquets en diffusion groupée, et utilise un codet de couche de liaison des données différent pour distinguer les deux espaces d'étiquette.

3.15 Chemin commuté d'étiquette (LSP), LSP d'entrée, LSP de sortie

Un chemin d'étiquettes commutées de niveau m pour un paquet P particulier est une séquence de routeurs, $\langle R1, \dots, Rn \rangle$ avec les propriétés suivantes :

1. R1, le "LSP d'entrée", est un LSR qui pousse une étiquette sur la pile d'étiquettes de P, résultant en une pile d'étiquettes de profondeur m ;
2. Pour tout i, $1 < i < n$, P a une pile d'étiquettes de profondeur m quand il est reçu par le LSR Ri ;
3. À aucun moment durant le transit de P de R1 à R[n-1] sa pile d'étiquettes n'a une profondeur inférieure à m ;

4. Pour tout $i, 1 < i < n$: R_i transmet P à R_{i+1} au moyen de MPLS, c'est-à-dire, en utilisant l'étiquette au sommet de la pile d'étiquettes (l'étiquette de niveau m) comme indice dans une ILM ;
5. Pour tout $i, 1 < i < n$: si un système S reçoit et transmet P après que P est transmis par R_i mais avant que P soit reçu par R_{i+1} (par exemple, R_i et R_{i+1} peuvent être connectés via un sous-réseau de liaison de données commuté, et S peut être un des commutateurs de la liaison de données) alors la décision de transmission de S ne se fonde pas sur le niveau m de l'étiquette, ni sur l'en-tête de couche réseau. Cela peut être parce que :
 - a) la décision ne se fonde pas du tout sur la pile d'étiquettes ni sur l'en-tête de couche réseau ;
 - b) la décision se fonde sur une pile d'étiquettes sur laquelle des étiquettes supplémentaires ont été poussées (c'est-à-dire, sur une étiquette de niveau $m+k$, où $k > 0$).

En d'autres termes, on peut parler du LSP de niveau m pour le paquet P comme une séquence des routeurs :

1. qui commencent par un LSR (un "LSP d'entrée") qui pousse une étiquette de niveau m ,
2. tous les LSR intermédiaires qui prennent leur décision de transmission en commutant l'étiquette sur une étiquette de niveau m ,
3. qui terminent (à un "LSP de sortie") lorsque une décision de transmission est prise en commutant l'étiquette sur une étiquette de niveau $m-k$, où $k > 0$, ou lorsque une décision de transmission est prise par des procédures de transmission "ordinaires", non MPLS.

Cela a pour conséquence (ou peut-être un présupposé) que chaque fois qu'un LSR pousse une étiquette sur un paquet déjà étiqueté, il doit s'assurer que la nouvelle étiquette correspond à une FEC dont le LSP de sortie est le LSR qui a affecté l'étiquette qui est maintenant la seconde dans la pile.

On appelle une séquence de LSR le "LSP pour une FEC F particulière" si c'est un LSP de niveau m pour un paquet P particulier lorsque l'étiquette de niveau m de P est une étiquette qui correspond à la FEC F .

Considérons l'ensemble de nœuds qui peuvent être les nœuds LSP d'entrée pour la FEC F . Il y a alors un LSP pour la FEC F qui commence avec chacun de ces nœuds. Si un certain nombre de ces LSP ont le même LSP de sortie, on peut alors considérer que l'ensemble de ces LSP est une arborescence, dont la racine est le LSP de sortie. (Comme les données voyagent le long de cette arborescence vers la racine, cela peut être appelé une arborescence de multipoint à point.) On peut donc parler de "l'arbre des LSP" pour une FEC F particulière.

3.16 Saut de l'avant dernier bond

Noter que selon les définitions du paragraphe 3.15, si $\langle R_1, \dots, R_n \rangle$ est un LSP de niveau m pour le paquet P , P peut être transmis de R_{n-1} à R_n avec une pile d'étiquettes de profondeur $m-1$. C'est-à-dire, la pile d'étiquettes peut être sautée à l'avant dernier LSR du LSP, plutôt qu'au LSP de sortie.

Du point de vue architectural, ceci est parfaitement approprié. L'objet de l'étiquette de niveau m est d'amener le paquet à R_n . Une fois que R_{n-1} a décidé d'envoyer le paquet à R_n , l'étiquette n'a plus aucune fonction, et n'a plus besoin d'être portée.

Il y a aussi un avantage pratique à faire le saut de l'avant dernier bond. Si on ne le fait pas, lorsque le LSP de sortie reçoit un paquet, il regarde d'abord l'étiquette du sommet de la pile, et détermine par suite de cet examen que c'est bien le LSP de sortie. Puis il doit sauter la pile, et examiner ce qui reste du paquet. Si il y a une autre étiquette sur la pile, le LSP de sortie va la regarder et transmettre le paquet selon ce qu'il aura vu. (Dans ce cas, le LSP de sortie pour le LSP de niveau m du paquet est aussi un nœud intermédiaire pour son LSP de niveau $m-1$.) Si il n'y a pas d'autre étiquette sur la pile, le paquet est alors transmis conformément à son adresse de destination de couche réseau. Noter que cela exigerait que le LSP de sortie fasse DEUX examens, soit deux examens d'étiquette, soit un examen d'étiquette suivi par un examen d'adresse.

Si, d'un autre côté, on utilise le saut de l'avant dernier bond, alors lorsque l'avant dernier bond regarde l'étiquette, il détermine :

- que c'est l'avant dernier bond, et
- qui est le prochain bond.

L'avant dernier nœud saute alors la pile, et transmet le paquet sur la base des informations obtenues en regardant l'étiquette qui était précédemment au sommet de la pile. Lorsque le LSP de sortie reçoit le paquet, l'étiquette qui est maintenant au sommet de la pile sera l'étiquette qu'il doit regarder afin de prendre sa propre décision de transmission. Ou, si le paquet ne portait qu'une seule étiquette, le LSP de sortie va simplement voir le paquet de couche réseau, ce qui est juste ce qu'il a besoin de voir afin de prendre sa décision de transmission.

Cette technique permet au LSP de sortie de faire un seul examen, et n'exige aussi qu'un seul examen par l'avant dernier nœud.

La création de ce "raccourci" de transmission d'un un produit de commutation d'étiquette peut être grandement favorisée si on sait qu'un seul examen est toujours nécessaire :

- le code peut être simplifié si on peut supposer qu'un seul examen est toujours nécessaire,
- le code peut se fonder sur un "budget temps" qui suppose qu'un seul examen est toujours nécessaire.

En fait, lorsque le saut de l'avant dernier bond est fait, le LSP de sortie n'a même pas besoin d'être un LSR.

Cependant, certains moteurs de commutation matériels peuvent n'être pas capables de sauter la pile d'étiquettes, de sorte que cette exigence ne peut être universelle. Il peut aussi y avoir certaines situations dans lesquelles le saut de l'avant dernier bond n'est pas souhaitable. Donc, l'avant dernier nœud ne saute la pile d'étiquettes que si c'est spécifiquement demandé par le nœud de sortie, OU si le prochain nœud dans le LSP ne prend pas MPLS en charge. (Si le prochain nœud dans le LSP prend bien MPLS en charge, mais ne fait pas une telle demande, l'avant dernier nœud n'a aucun moyen de savoir qu'il est en fait l'avant dernier nœud.)

Un LSR qui est capable de sauter la pile d'étiquettes DOIT faire le saut de l'avant dernier bond lorsque cela lui est demandé par l'homologue aval de distribution d'étiquettes.

Les négociations initiales de protocole de distribution d'étiquettes DOIVENT permettre à chaque LSR de déterminer si ses LSR voisins sont capables de sauter la pile d'étiquettes. Un LSR NE DOIT PAS demander à un homologue de distribution d'étiquettes de sauter la pile d'étiquettes si il n'est pas capable de le faire.

On peut se demander si le nœud de sortie peut toujours interpréter l'étiquette du sommet de la pile d'un paquet correctement reçu si le saut de l'avant dernier bond est utilisé. Pour autant que soient respectées les règles d'unicité et de portée du paragraphe 3.14, il est toujours possible d'interpréter sans ambiguïté l'étiquette du sommet de la pile d'un paquet reçu.

3.17 LSP de prochain bond

Le LSP de prochain bond pour un paquet étiqueté particulier dans un LSR particulier est le LSR qui est le prochain bond, comme choisi par l'entrée NHLFE utilisée pour transmettre ce paquet.

Le LSP de prochain bond pour une FEC particulière est le prochain bond tel que sélectionné par l'entrée de NHLFE indexé par une étiquette qui correspond à cette FEC.

Noter que le LSP de prochain bond peut différer du prochain bond qui aurait été choisi par l'algorithme de couche réseau. On utilisera le terme "prochain bond de L3" quand on se réfèrera à ce dernier.

3.18 Étiquettes entrantes invalides

Que devrait faire un LSR si il reçoit un paquet étiqueté avec une étiquette entrante particulière, mais si il n'a pas de lien pour cette étiquette ? Il est tentant de penser que l'étiquettes peut juste être retirée, et le paquet transmis comme un paquet IP non étiqueté. Cependant, dans certains cas, faire ainsi pourrait causer une boucle. Si le LSR amont pense que l'étiquette est liée à un chemin explicite, et si le LSR aval pense que l'étiquette n'est liée à rien, et si l'acheminement bond par bond du paquet IP non étiqueté ramène le paquet au LSR amont, une boucle est alors formée.

Il est aussi possible que l'étiquette ait été destinée à représenter un chemin qui ne peut pas être déduit de l'en-tête IP.

Donc, lorsque un paquet étiqueté est reçu avec une étiquette entrante invalide, il DOIT être éliminé, SAUF SI il est déterminé par quelque moyen (qui n'entre pas dans le domaine d'application du présent document) que le transmettre non étiqueté ne peut pas causer de dommage.

3.19 Contrôle de LSP : ordonné ou indépendant

Certaines FEC correspondent à des préfixes d'adresses qui sont distribués via un algorithme d'acheminement dynamique. L'établissement des LSP pour ces FEC peut être fait d'une de deux façons : le contrôle de LSP indépendant ou le contrôle de LSP ordonné.

Dans le contrôle de LSP indépendant, chaque LSR, en notant qu'il reconnaît une FEC particulière, prend une décision indépendante de lier une étiquette à cette FEC et de distribuer ce lien à ses homologues de distribution d'étiquettes. Cela correspond à la façon dont fonctionne l'acheminement conventionnel des datagrammes IP ; chaque nœud prend une décision indépendante sur la façon de traiter chaque paquet, et s'appuie sur l'algorithme d'acheminement pour converger rapidement de façon à assurer que chaque datagramme est livré correctement.

Dans le contrôle de LSP ordonné, un LSR ne lie une étiquette à une FEC particulière que si il est le LSR de sortie pour cette FEC, ou si il a déjà reçu un lien d'étiquette pour cette FEC de la part de son prochain bond pour cette FEC.

Si on veut s'assurer que le trafic dans une FEC particulière suit un chemin avec un ensemble spécifié de propriétés (par exemple, que le trafic ne traverse aucun nœud deux fois, qu'une quantité spécifiée de ressources est disponible pour le trafic, que le trafic suit un chemin explicitement spécifié, etc.) le contrôle ordonné doit être utilisé. Avec le contrôle indépendant, certains LSR peuvent commencer la commutation d'étiquette sur un trafic dans la FEC avant que le LSP soit complètement établi, et donc une partie du trafic dans la FEC peut suivre un chemin qui n'a pas l'ensemble de propriétés spécifié. Le contrôle ordonné doit aussi être utilisé si la reconnaissance de la FEC est une conséquence de l'établissement du LSP correspondant.

L'établissement du LSP ordonné peut être à l'initiative de l'entrée ou de la sortie.

Contrôle ordonné et contrôle indépendant sont parfaitement inter opérables. Cependant, sauf si tous les LSR dans un LSP utilisent le contrôle ordonné, l'effet global sur le comportement de réseau est largement celui du contrôle indépendant, car on ne peut pas être sûr qu'un LSP n'est pas utilisé tant qu'il n'est pas pleinement établi.

Cette architecture permet que le choix entre contrôle indépendant et contrôle ordonné soit une affaire locale. Comme les deux méthodes interopèrent, un certain LSR n'a besoin de prendre en charge que l'un ou l'autre. Généralement parlant, le choix du contrôle indépendant ou du contrôle ordonné ne paraît pas avoir d'effet sur les mécanismes de distribution d'étiquette qui doivent être définis.

3.20 Agrégation

Une façon de partager le trafic entre les FEC est de créer une FEC distincte pour chaque préfixe d'adresses qui apparaît dans le tableau d'acheminement. Cependant, au sein d'un domaine MPLS particulier, cela peut résulter en un ensemble de FEC tel que tout le trafic dans toutes ces FEC suive le même chemin. Par exemple, un ensemble de préfixes d'adresses distincts peuvent tous avoir le même nœud de sortie, et l'échange d'étiquetage ne pourrait être utilisé que pour obtenir le trafic pour le nœud de sortie. Dans ce cas, au sein du domaine MPLS, l'union de ces FEC est lui-même une FEC. Cela crée un choix : une étiquette distincte devrait elle être liée à chaque FEC composante, ou une seule étiquette devrait elle être liée à l'union, et cette étiquette appliquée à tout le trafic dans l'union ?

La procédure de liaison d'une seule étiquette à une union de FEC qui est elle-même une FEC (au sein d'un certain domaine) et d'appliquer cette étiquette à tout le trafic dans l'union, est appelée "agrégation". L'architecture MPLS permet l'agrégation. L'agrégation peut réduire le nombre d'étiquettes nécessaires pour traiter un ensemble particulier de paquets, et peut aussi réduire la quantité de trafic de contrôle de distribution d'étiquettes nécessaire.

Étant donnée un ensemble de FEC qui sont "agrégables" en une seule FEC, il est possible de (a) les agréger en une seule FEC, (b) les agréger en un ensemble de FEC, ou (c) ne pas les agréger du tout. Donc, on peut parler de "granularité" de l'agrégation, avec (a) qui est la "plus grosse granularité", et (c) qui est la "plus fine granularité".

Quand on utilise le contrôle d'ordre, chaque LSR devrait adopter, pour un ensemble de FEC donné, la granularité utilisée par son prochain bond pour ces FEC.

Quand on utilise le contrôle indépendant, il est possible qu'il y ait deux LSR adjacents, Ru et Rd, qui agrègent différemment un ensemble de FEC.

Si Ru a une plus fine granularité que Rd, cela ne pose pas de problème. Ru distribue plus d'étiquettes pour cet ensemble de FEC que ne le fait Rd. Cela signifie que quand Ru a besoin de transmettre des paquets étiquetés dans ces FEC à Rd, il peut avoir besoin de transposer n étiquettes en m étiquettes, lorsque $n > m$. En option, Ru peut retirer l'ensemble de n étiquettes qu'il a distribué, puis distribuer un ensemble de m étiquettes, correspondant au niveau de granularité de Rd. Cela n'est pas nécessaire pour assurer un fonctionnement correct, mais il en résulte une réduction du nombre d'étiquettes distribuées par Ru, et Ru ne tire aucun avantage particulier de distribuer le plus grand nombre d'étiquettes. La décision de faire ou non cela est une affaire locale.

Si Ru a une plus grosse granularité que Rd (c'est-à-dire, si Rd a distribué n étiquettes pour l'ensemble des FEC, alors que Ru en a distribué m, avec $n > m$) il a le choix de :

- adopter le plus fin niveau de granularité de Rd. Cela exigerait qu'il retire les m étiquettes qu'il a distribué, et qu'il distribue n étiquettes. C'est l'option préférée.
- transposer simplement ses m étiquettes en un sous-ensemble de n étiquettes de Rd, si il peut déterminer que cela va produire le même acheminement. Par exemple, supposons que Ru applique une seule étiquette à tout le trafic qui a

besoin de passer à travers un certain LSR de sortie, tandis que Rd lie un certain nombre de différentes étiquettes à un tel trafic, en fonction des adresses individuelles de destination des paquets. Si Ru connaît l'adresse du routeur de sortie, et si Rd a lié une étiquette à la FEC qui est identifiée par cette adresse, alors Ru peut simplement appliquer cette étiquette.

En tout état de cause, chaque LSR a besoin de savoir (par configuration) quelle granularité utiliser pour les étiquettes qu'il alloue. Lorsque le contrôle ordonné est utilisé, cela exige que chaque nœud sache la granularité pour les seules FEC qui quittent le réseau MPLS à ce nœud. Pour le contrôle indépendant, les meilleurs résultats peuvent être obtenus en s'assurant que tous les LSR sont configurés de façon cohérente pour connaître la granularité pour chaque FEC. Cependant, dans de nombreux cas, cela peut être fait en utilisant un seul niveau de granularité qui s'applique à toutes les FEC (comme "une étiquette par préfixe IP dans le tableau de transmission", ou "une étiquette par nœud de sortie").

3.21 Sélection du chemin

La sélection de chemin se réfère à la méthode utilisée pour choisir le LSP pour une FEC particulière. L'architecture du protocole MPLS proposé prend en charge deux options pour la sélection de chemin : (1) l'acheminement bond par bond, et (2) l'acheminement explicite.

L'acheminement bond par bond permet à chaque nœud de choisir indépendamment le prochain bond pour chaque FEC. C'est aujourd'hui le mode usuel dans les réseaux IP existants. Un "LSP acheminé bond par bond" est un LSP dont le chemin est choisi en utilisant l'acheminement bond par bond.

Dans un LSP à acheminement explicite, chaque LSR ne choisit pas le prochain bond de façon indépendante ; un seul LSR, généralement le LSP d'entrée ou le LSP de sortie, spécifie plutôt plusieurs (ou la totalité) des LSR dans le LSP. Si un seul LSR spécifie le LSP entier, le LSP est "strictement" à acheminement explicite. Si un seul LSR ne spécifie que quelques uns des LSP, le LSP est "vaguement" à acheminement explicite.

La séquence de LSR suivie par un LSP à acheminement explicite peut être choisie par configuration, ou peut être choisie de façon dynamique par un seul nœud (par exemple, le nœud de sortie peut faire usage des informations topologiques apprises de la base de données d'état de liaison afin de calculer le chemin entier pour l'arborescence qui se termine au nœud de sortie).

L'acheminement explicite peut être utile pour un certain nombre d'objets, comme l'acheminement d'après la politique, ou l'ingénierie du trafic. Dans MPLS, le chemin explicite doit être spécifié au moment où les étiquettes sont affectées, mais le chemin explicite n'a pas à être spécifié avec chaque paquet IP. Cela rend l'acheminement explicite MPLS beaucoup plus efficace que la solution de l'acheminement de source IP.

Les procédures pour faire usage des chemins explicites, strict ou vague, sortent du domaine d'application du présent document.

3.22 Manque d'étiquette sortante

Lorsque un paquet étiqueté voyage le long d'un LSP, il peut arriver à l'occasion qu'il atteigne un LSR auquel la ILM ne transpose pas l'étiquette entrante du paquet en une NHLFE, même si l'étiquette entrante est par elle-même valide. Cela peut arriver à cause de conditions transitoires, ou à cause d'une erreur au LSR qui devrait être le prochain bond du paquet.

Il est tentant dans de tels cas de retirer la pile d'étiquettes et de tenter de transmettre le paquet plus loin via la transmission conventionnelle, sur la base de son en-tête de couche réseau. Cependant, en général, ce n'est pas une procédure sûre :

- si le paquet a suivi un LSP à acheminement explicite, il peut en résulter une boucle ;
- l'en-tête réseau du paquet peut ne pas contenir assez d'informations pour donner à ce LSR particulier le moyen d'une transmission correcte.

Sauf si on peut déterminer (par des moyens qui sortent du domaine d'application du présent document) qu'aucune de ces situations ne se présente, la seule procédure sûre est d'éliminer le paquet.

3.23 Durée de vie (TTL)

Dans la transmission IP conventionnelle, chaque paquet porte une valeur "Durée de vie" (TTL, *Time To Live*) dans l'en-tête. Chaque fois qu'un paquet passe par un routeur, son TTL est décrémenté de 1 ; si le TTL atteint 0 avant que le paquet ait atteint sa destination, il est éliminé.

Cela donne un certain niveau de protection contre les boucles de transmission qui peuvent exister à cause de mauvaises configurations, ou à cause de défaillances ou mauvaise convergence de l'algorithme d'acheminement. Le TTL est parfois

aussi utilisé pour d'autres fonctions, comme la portée des diffusions groupées, et la prise en charge de la commande "traceroute". Cela implique qu'il y a deux problèmes en rapport avec le TTL auxquels MPLS est confronté : (i) le TTL comme façon de supprimer les boucles ; (ii) le TTL comme moyen d'accomplir d'autres fonctions, telles que de limiter la portée d'un paquet.

Lorsque un paquet voyage le long d'un LSP, il DEVRAIT émerger avec la même valeur de TTL qu'il aurait eu si il avait traversé la même séquence de routeurs sans avoir été en commutation d'étiquette. Si le paquet voyage le long d'une hiérarchie de LSP, le nombre total de bonds de LSR traversés DEVRAIT être reflété dans sa valeur de TTL lorsque il émerge de la hiérarchie de LSP.

La façon dont le TTL est traité peut varier selon que les valeurs des étiquette MPLS sont portées dans un en-tête "de calage" spécifique de MPLS [RFC3032], ou si les étiquettes MPLS sont portées dans un en-tête de couche 2, comme un en-tête ATM [RFC2935] ou un en-tête de relais de trame [RFC3034].

Si les valeurs d'étiquette sont codées dans un "calage de compensation" qui se tient entre l'en-tête de couche de liaison et l'en-tête de couche réseau, cette cale DOIT avoir un champ TTL qui DEVRAIT être initialement chargé à partie du champ TTL de l'en-tête de couche réseau, DEVRAIT être décrétementée à chaque bond de LSR, et DEVRAIT être copiée dans le champ TTL de l'en-tête de couche réseau lorsque le paquet émerge de son LSP.

Si les valeurs d'étiquette sont codées dans un en-tête de couche de liaison des données (par exemple, le champ VPI/VCI dans l'en-tête AAL5 d'ATM) et si les paquets étiquetés sont transmis par un commutateur de couche 2 (par exemple, un commutateur ATM) et si la couche de liaison des données (comme ATM) n'a pas elle-même un champ TTL, il ne sera alors pas possible de décrétement le TTL d'un paquet à chaque bond de LSR. Un segment LSP qui consiste en une séquence de LSR qui ne peuvent pas décrétement le TTL d'un paquet sera appelé un "segment LSP sans TTL".

Lorsque un paquet émerge d'un segment LSP sans TTL, il DEVRAIT cependant recevoir un TTL qui reflète le nombre de bonds de LSR qu'il a traversé. Dans le cas de l'envoi individuel, cela peut être réalisé en propageant une longueur de LSP significative pour les nœuds d'entrée, permettant à l'entrée de décrétement la valeur de TTL avant de transmettre les paquets dans un segment LSP sans TTL.

Il peut parfois être déterminé, à l'entrée dans un segment LSP sans TTL que le TTL d'un paquet particulier va arriver à expiration avant que le paquet atteigne la sortie de ce segment LSP sans TTL. Dans ce cas, le LSR à l'entrée du segment LSP sans TTL ne doit pas mettre le paquet en commutation d'étiquette. Cela signifie que des procédures spéciales doivent être développées pour prendre en charge la fonctionnalité traceroute, par exemple, les paquets traceroute peuvent être transmis en utilisant la transmission bond par bond conventionnelle.

3.24 Contrôle de boucle

Sur un segment LSP sans TTL, par définition, le TTL ne peut pas être utilisé pour protéger contre les boucles de transmission. L'importance du contrôle de boucle peut dépendre du matériel particulier utilisé pour fournir les fonctions de LSR le long du segment de LSP sans TTL.

Supposons, par exemple, qu'un matériel de commutation ATM soit utilisé pour fournir les fonctions de commutation MPLS, les étiquettes étant portées dans le champ VPI/VCI. Comme le matériel de commutation ATM ne peut pas décrétement le TTL, il n'y a pas de protection contre les boucles. Si le matériel ATM est capable de fournir un bon accès au réservoir de mémoires tampon pour les cellules entrantes qui portent les différentes valeurs de VPI/VCI, ces bouclages peuvent n'avoir pas d'effet délétère sur les autres trafics. Si le matériel ATM ne peut pas fournir d'accès de mémoire tampon de cette sorte, même des boucles transitoires peuvent cependant causer une sévère dégradation des performances totales du LSR.

Même si un bon accès de mémoire tampon peut être fourni, cela vaut tout de même la peine d'avoir quelque moyen de détection des boucles qui durent "plus longtemps que possible".

De plus, même lorsque le TTL et/ou la mise en file d'attente équitable par VC donne le moyen de survivre aux boucles, il peut tout de même être souhaitable lorsque c'est praticable d'éviter d'établir des LSP avec des boucles. Tout LSR qui peut se rattacher à des segments de LSP sans TTL va donc devoir prendre en charge une technique commune de détection de boucle; cependant, l'utilisation de la technique de détection de boucle est facultative. La technique de détection de boucle est spécifiée dans la [RFC2935] et la [RFC3036].

3.25 Codages des étiquettes

Afin de transmettre une pile d'étiquettes avec le paquet dont c'est la pile d'étiquettes, il est nécessaire de définir un codage concret de la pile d'étiquettes. L'architecture prend en charge plusieurs techniques de codage différentes ; le choix de la

technique de codage dépend du type particulier d'appareil qui est utilisé pour transmettre les paquets étiquetés.

3.25.1 Matériel et/ou logiciel spécifique de MPLS

Si on utilise un matériel et/ou logiciel spécifique de MPLS pour transmettre des paquets étiquetés, la façon la plus évidente de coder la pile d'étiquettes est de définir un nouveau protocole à utiliser comme une "cale" entre les en-têtes de la couche de liaison des données et celui de la couche réseau. Cette cale va en réalité être juste une encapsulation du paquet de couche réseau ; il sera "indépendant du protocole" de sorte qu'il puisse être utilisé pour encapsuler toute couche réseau. Donc nous appellerons cela "l'encapsulation générique MPLS".

L'encapsulation générique MPLS sera à son tour encapsulée dans un protocole de couche de liaison des données.

L'encapsulation générique MPLS est spécifiée dans la [RFC3032].

3.25.2 Commutateurs ATM comme LSR

On notera que les procédures de transmission MPLS sont similaires à celles des commutateurs traditionnels "d'échange d'étiquetage" comme les commutateurs ATM. Les commutateurs ATM utilisent l'accès d'entrée et la valeur du VPI/VCI entrant comme indice dans un tableau "d'interconnexion", duquel ils obtiennent un accès de sortie et une valeur de VPI/VCI de sortie. Donc, si une ou plusieurs étiquettes peuvent être codées directement dans les champs auxquels accèdent ces commutateurs traditionnels, ceux-ci peuvent alors, avec une mise à niveau convenable du logiciel, être utilisés comme des LSR. On appellera de tels appareils des "ATM-LSR".

Il y a trois façons évidentes de coder les étiquettes dans l'en-tête de cellule ATM (en supposant l'utilisation de AAL5) :

1. Codage SVC

Utilise le champ VPI/VCI pour coder l'étiquette qui est au sommet de la pile d'étiquettes. Cette technique peut être utilisée dans tout réseau. Avec cette technique de codage, chaque LSP est réalisé comme un SVC ATM, et le protocole de distribution d'étiquettes devient le protocole de "signalisation" ATM. Avec cette technique de codage, les LSR ATM ne peuvent pas effectuer d'opérations de "poussée" ou de "saut" sur la pile d'étiquettes.

2. Codage SVP

Utilise le champ VPI pour coder l'étiquette qui est au sommet de la pile d'étiquettes, et le champ VCI pour coder la seconde étiquette de la pile, s'il en est une présente. Cette technique présente certains avantages sur la précédente, en ce qu'elle permet d'utiliser la "commutation de VP" d'ATM. C'est-à-dire que les LSP sont réalisés comme des SVP ATM, avec le protocole de distribution d'étiquettes qui sert de protocole de signalisation ATM.

Cependant, cette technique ne peut pas toujours être utilisée. Si le réseau comporte un chemin virtuel ATM à travers un réseau ATM non MPLS, le champ VPI n'est alors pas nécessairement disponible pour son utilisation par MPLS.

Lorsque cette technique de codage est utilisée, le LSR ATM à la sortie du VP fait effectivement une opération de "saut".

3. Codage SVP multipoint

Utilise le champ VPI pour coder l'étiquette qui est au sommet de la pile d'étiquettes, utilise une partie du champ VCI pour coder la seconde étiquette de la pile, s'il en est une présente, et utilise le reste du champ VCI pour identifier l'entrée de LSP. Si cette technique est utilisée, les capacités conventionnelles de commutation de VP ATM peuvent être utilisées pour fournir des VP en multipoint à point. Les cellules provenant de différents paquets vont alors porter des valeurs de VCI différentes. Comme on le verra au paragraphe 3.26, cela nous permet de faire la fusion d'étiquette, sans entrer dans les problèmes d'entrelaçage de cellules, sur les commutateurs ATM qui peuvent fournir des VP en multipoint à point, mais qui n'ont pas la capacité de fusion de VC.

Cette technique dépend de l'existence d'une capacité d'allouer des valeurs de VCI de 16 bits à chaque commutateur ATM de telle sorte qu'aucune valeur d'un seul VCI ne soit allouée à deux commutateurs différents. (Si un nombre adéquat de telles valeurs pouvait être alloué à chaque commutateur, il serait possible de traiter aussi la valeur de VCI comme la seconde étiquette de la pile.)

Si il y a plus d'étiquettes sur la pile que ce qui peut être codé dans l'en-tête ATM, les codages ATM doivent être combinés avec l'encapsulation générique.

3.25.3 Interopérabilité entre les techniques de codage

Si $\langle R1, R2, R3 \rangle$ est un segment d'un LSP, il est possible que R1 utilise un codage de la pile d'étiquettes lorsque il transmet le paquet P à R2, mais R2 va utiliser un codage différent lorsque il transmettra un paquet P à R3. En général, l'architecture MPLS prend en charge les LSP avec des codages de pile d'étiquettes différents utilisés sur des bonds différents. Donc, quand on discute des procédures de traitement d'un paquet étiqueté, on parle en termes abstraits du fonctionnement de la pile d'étiquette du paquet. Lorsque un paquet étiqueté est reçu, le LSR doit le décoder pour déterminer la valeur courante de la pile d'étiquettes, puis il doit opérer sur la pile d'étiquettes pour déterminer la nouvelle valeur de la pile, et ensuite coder la nouvelle valeur de façon appropriée avant de transmettre le paquet étiqueté à son prochain bond.

Malheureusement, les commutateurs ATM n'ont pas de capacité de traduction d'une technique de codage à une autre. L'architecture MPLS exige donc que chaque fois qu'il est possible à deux commutateurs ATM d'être les LSR successifs le long d'un LSP de niveau m pour un certain paquet, ces deux commutateurs ATM utilisent la même technique de codage.

Naturellement il y aura des réseaux MPLS qui contiendront une combinaison de commutateurs ATM qui fonctionnent comme des LSR, et d'autres LSR qui fonctionnent en utilisant un en-tête MPLS de calage. Dans de tels réseaux, il peut y avoir des LSR qui ont des interfaces ATM aussi bien que des interfaces "de calage MPLS". C'est un exemple d'un LSR avec un codage de pile d'étiquettes différent sur les différents bonds. Un tel LSR peut changer une pile d'étiquettes codée en ATM sur une interface entrante et la remplacer par une pile d'étiquettes codée en en-tête MPLS de calage sur l'interface sortante.

3.26 Fusion d'étiquettes

Supposons qu'un LSR ait lié plusieurs étiquettes entrantes à une certaine FEC. Lors de la transmission de paquets de cette FEC, on aimerait avoir une seule étiquette sortante appliquée à tous ces paquets. Le fait que deux différents paquets dans la FEC soient arrivés avec des étiquettes entrantes différentes n'a pas d'importance ; on voudrait les transmettre avec la même étiquette sortante. La capacité de faire cela est appelée "fusion d'étiquette".

Disons qu'un LSR est capable de fusion d'étiquette si il peut recevoir deux paquets de deux interfaces entrantes différentes, et/ou avec des étiquettes différentes, et envoyer les deux paquets sur la même interface sortante avec la même étiquette. Une fois que les paquets sont transmis, les informations qui sont arrivées des différentes interfaces et/ou avec les différentes étiquettes entrantes sont perdues.

Disons qu'un LSR n'est pas capable de fusion d'étiquette si, pour deux paquets quelconques qui arrivent d'interfaces différentes, ou avec des étiquettes différentes, les paquets doivent être transmis sur des interfaces différentes, ou doivent avoir des étiquettes différentes. Les LSR ATM qui utilisent le codage SVC ou SVP ne peuvent pas effectuer la fusion d'étiquette. Ceci est exposé plus en détail au paragraphe suivant.

Si un LSR particulier ne peut pas effectuer la fusion d'étiquette, alors si deux paquets dans la même FEC arrivent avec des étiquettes entrantes différentes, ils doivent être transmis avec des étiquettes sortantes différentes. Avec la fusion d'étiquette, le nombre d'étiquettes sortantes par FEC a seulement besoin d'être 1 ; sans fusion d'étiquette, le nombre d'étiquettes sortantes par FEC pourrait être aussi grand que le nombre de nœuds dans le réseau.

Avec la fusion d'étiquette, le nombre d'étiquettes entrantes par FEC dont a besoin un certain LSR n'est jamais supérieur au nombre d'adjacences de distribution d'étiquettes. Sans fusion d'étiquette, le nombre d'étiquettes entrantes par FEC dont a besoin un certain LSR est égal au nombre de nœuds amonts qui transmettent du trafic dans la FEC au LSR en question. En fait, il est difficile à un LSR de déterminer même combien de telles étiquettes entrantes il doit prendre en charge pour une certaine FEC.

L'architecture MPLS s'accommode de la fusion et de la non fusion des LSR, mais permet qu'il puisse y avoir des LSR qui ne prennent pas en charge la fusion d'étiquettes. Cela conduit au problème de s'assurer de l'inter opération correcte entre les LSR qui font la fusion et ceux qui ne la font pas. Le problème est quelque peu différent dans le cas de supports de datagrammes par opposition au cas de l'ATM. Les différents types de supports vont donc être exposés séparément.

3.26.1 LSR sans fusion

Les procédures de transmission MPLS sont très similaires aux procédures de transmission utilisées par des technologies telles que ATM et le relais de trame. C'est-à-dire que lorsque arrive une unité de données, une étiquette (VPI/VCI ou DLCI) est recherchée dans un "tableau d'interconnexion", sur la base de cette recherche, un accès de sortie est choisi, et la valeur de l'étiquette est réécrite. En fait, il est possible d'utiliser de telles technologies pour la transmission MPLS ; un protocole de distribution d'étiquettes peut être utilisé comme "protocole de signalisation" pour établir les tableaux d'interconnexion.

Malheureusement, ces technologies ne prennent pas nécessairement en charge la capacité de fusion d'étiquette. Dans ATM, si on tente d'effectuer la fusion d'étiquette, le résultat peut être l'entrelacement de cellules provenant de divers paquets. Si des cellules provenant de différents paquets se trouvent entrelacées, il est impossible de ré-assembler les paquets. Certains commutateurs de relais de trame utilisent la commutation de cellules en arrière plan. Ces commutateurs peuvent aussi être incapables de prendre en charge la fusion d'étiquette, pour la même raison – les cellules de différents paquets peuvent se trouver entrelacées, et il n'y a alors aucun moyen de ré-assembler les paquets.

On propose de prendre en charge deux solutions à ce problème. D'abord, MPLS va contenir des procédures qui permettent d'utiliser des LSR sans fusion. Ensuite, MPLS va mettre en œuvre des procédures qui permettent à certains commutateurs ATM de fonctionner comme des LSR à fusion.

Comme MPLS prend en charge les LSR aussi bien à fusion que sans fusion, il contient aussi des procédures pour assurer l'inter fonctionnement correct entre eux.

3.26.2 Étiquettes pour LSR à fusion et LSR sans fusion

Un LSR amont qui prend en charge la fusion d'étiquettes n'a besoin de recevoir qu'une seule étiquette par FEC. Un voisin amont qui ne prend pas en charge la fusion d'étiquette a besoin de recevoir plusieurs étiquettes par FEC. Cependant, il n'y a aucun moyen de savoir a priori de combien d'étiquettes il a besoin. Cela va dépendre du nombre de LSR qui sont en amont de lui par rapport à la FEC en question.

Dans l'architecture MPLS, si un certain voisin amont ne prend pas en charge la fusion d'étiquettes, il ne lui est pas envoyé d'étiquette pour une FEC particulière à moins qu'il ne demande explicitement une étiquette pour cette FEC. Le voisin amont peut faire plusieurs demandes comme cela, et il lui est donné une nouvelle étiquette à chaque fois. Lorsque un voisin aval reçoit une telle demande de la part de l'amont, et que le voisin aval ne prend pas lui-même en charge la fusion d'étiquettes, il doit alors à son tour demander à son voisin aval une autre étiquette pour la FEC en question.

Il est possible qu'il y ait des nœuds qui prennent en charge la fusion d'étiquette, mais ne puissent fusionner qu'un nombre limité d'étiquettes entrantes en une seule étiquette sortante. Supposons par exemple qu'à cause de certaines limitations du matériel, un nœud soit capable de fusionner quatre étiquettes entrantes en une seule étiquette sortante. Supposons cependant que ce nœud ait six étiquettes entrantes qui lui arrivent pour une certaine FEC. Dans ce cas, ce nœud peut les fusionner en deux étiquettes sortantes.

La question de savoir si la fusion d'étiquettes est applicable aux LSP à acheminement explicite fera l'objet d'un complément d'étude.

3.26.3 Fusion sur ATM

3.26.3.1 Méthodes d'élimination de l'entrelacement de cellules

Plusieurs méthodes peuvent être utilisées pour éliminer le problème de l'entrelacement de cellules dans ATM, ce qui permet aux commutateurs ATM de prendre en charge la fusion de flux :

1. Fusion de chemin virtuel (VP) en utilisant le codage SVP multipoint
Lorsque la fusion de VP est utilisée, plusieurs chemins virtuels sont fusionnées en un chemin virtuel, mais les paquets provenant de sources différentes sont distingués en utilisant des VCI différents au sein du VP.
2. Fusion de circuit virtuel (VC)
Lorsque on utilise la fusion de VC, il est exigé des commutateurs qu'ils mettent en mémoire tampon les cellules provenant d'un paquet jusqu'à ce que la totalité du paquet ait été reçue (cela peut être déterminé en regardant l'indicateur de fin de trame AAL5).

La fusion de VP présente l'avantage d'être compatible avec un plus fort pourcentage des mises en œuvre existantes de commutateurs ATM. Cela rend plus probable que la fusion de VP puisse être utilisée dans les réseaux existants. À la différence de la fusion de VC, la fusion de VP ne souffre d'aucun délai aux points de fusion et n'impose pas non plus d'exigences de mémoire tampon particulières. Cependant, elle a l'inconvénient d'exiger la coordination de l'espace de VCI au sein de chaque VP. Cela peut être réalisé d'un certain nombre de façons. Le choix d'une ou plusieurs de ces méthodes fera l'objet d'un complément d'étude.

Ce compromis entre la compatibilité avec les équipements existants et la complexité et l'adaptabilité du protocole implique qu'il est souhaitable que le protocole MPLS prenne en charge aussi bien la fusion de VP que la fusion de VC. Pour ce faire,

chaque commutateur ATM qui participe à MPLS a besoin de savoir si son voisin ATM immédiat effectue la fusion de VP, la fusion de VC, ou pas de fusion du tout.

3.26.3.2 Inter opération : fusion de VC, fusion de VP, et non fusion

L'inter opération des diverses formes de fusion sur ATM est le plus facilement décrite en commençant par décrire l'inter opération de VC à fusion avec VC sans fusion.

Dans le cas où des nœuds avec des VC à fusion et sans fusion sont interconnectés, la transmission des cellules se fonde dans tous les cas sur un VC (c'est-à-dire, la concaténation du VPI et du VCI). Pour chaque nœud, si un voisin amont fait la fusion de VC, alors ce voisin amont exige seulement un seul VPI/VCI pour un flux particulier (ceci est analogue à l'exigence d'une seule étiquette dans le cas de fonctionnement sur un support de trame). Si le voisin amont ne fait pas la fusion, le voisin va alors exiger un seul VPI/VCI par flux pour lui-même, plus assez de VPI/VCI pour le passer à ses voisins en amont. Le nombre exigé sera déterminé en permettant aux nœuds en amont de demander des VPI/VCI supplémentaires à leurs voisins vers l'aval (c'est là aussi analogue à la méthode utilisée dans la fusion de trames).

Une méthode similaire est possible pour prendre en charge les nœuds qui effectuent la fusion de VP. Dans ce cas, le nœud qui fusionne les VP, plutôt que de demander un seul VPI/VCI ou un certain nombre de VPI/VCI à son voisin en aval, peut plutôt demander un seul VP (identifié par un VPI) mais plusieurs VCI au sein du VP. De plus, supposons qu'un nœud sans fusion soit en aval de deux nœuds différents de fusion de VP. Ce nœud peut avoir besoin de demander un VPI/VCI (pour le trafic généré par lui-même) plus deux VP (un pour chaque nœud amont) chacun associé à un ensemble spécifié de VCI (comme demandé au nœud amont).

Afin de prendre en charge toutes les fusions de VP, les fusions de VC et la non fusion, il est donc nécessaire de permettre aux nœuds amont de demander une combinaison de zéro, un, ou plusieurs identifiants de VC (consistant en un VPI/VCI) plus zéro, un, ou plusieurs VP (identifiés par des VPI) contenant chacun un nombre spécifié de VC (identifiés par un ensemble de VCI qui sont significatifs au sein d'un VP). Les nœuds à fusion de VP demanderont donc un VP, avec un VCI contenu pour le trafic qu'ils génèrent (si c'est approprié) plus un VCI pour chaque VC demandé comme ci-dessus (sans considérer si le VC fait partie ou non d'un VP contenant). Un nœud à fusion de VC va demander seulement un seul VPI/VCI (car il peut fusionner tout le trafic amont en un seul VC). Les nœuds sans fusion vont passer toutes les demandes qu'ils ont obtenues comme ci-dessus, plus une demande de VPI/VCI pour le trafic qu'ils génèrent (si c'est approprié).

3.27 Tunnels et hiérarchie

Parfois un routeur Ru entreprend une action explicite pour faire qu'un paquet particulier soit livré à un autre routeur Rd, même si Ru et Rd ne sont pas des routeurs consécutifs sur le chemin bond par bond pour ce paquet, et si Rd n'est pas la destination ultime du paquet. Par exemple, cela peut être fait en encapsulant le paquet à l'intérieur d'un paquet de couche réseau dont l'adresse de destination est l'adresse de Rd lui-même. Cela crée un "tunnel" de Ru à Rd. On appelle tout paquet ainsi traité un "paquet tunnelé".

3.27.1 Tunnel acheminé bond par bond

Si un paquet tunnelé suit le chemin bond par bond de Ru à Rd, on dit que c'est un "tunnel acheminé bond par bond" dont le "point d'extrémité de transmission" est Ru et dont le "point d'extrémité de réception" est Rd.

3.27.2 Tunnel à acheminement explicite

Si un paquet tunnelé voyage de Ru à Rd sur un chemin autre que le chemin bond par bond, on dit que c'est un "tunnel à acheminement explicite" dont le "point d'extrémité de transmission" est Ru et dont le "point d'extrémité de réception" est Rd. Par exemple, on peut envoyer un paquet à travers un tunnel à acheminement explicite en l'encapsulant dans un paquet qui est acheminé de source.

3.27.3 LSP tunnels

Il est possible de mettre en œuvre un tunnel MPLS comme LSP, et d'utiliser la commutation d'étiquette plutôt que l'encapsulation de couche réseau pour faire que le paquet voyage à travers le tunnel. Le tunnel serait un LSP <R1, ..., Rn>, où R1 est le point d'extrémité de transmission du tunnel, et Rn est le point d'extrémité de réception du tunnel. C'est ce qu'on appelle un "LSP tunnel".

L'ensemble des paquets qui sont à envoyer à travers le LSP tunnel constitue une FEC, et chaque LSR dans le tunnel doit allouer une étiquette à cette FEC (c'est-à-dire, doit allouer une étiquette au tunnel). Les critères pour allouer un certain

paquet à un LSP tunnel sont une affaire locale au point d'extrémité de transmission du tunnel. Pour mettre un paquet dans un LSP tunnel, le point d'extrémité de transmission pousse une étiquette pour le tunnel sur la pile d'étiquettes et envoie le paquet étiqueté au prochain bond dans le tunnel.

Si il n'est pas nécessaire pour le point d'extrémité de réception du tunnel d'être capable de déterminer quels paquets il reçoit à travers le tunnel, comme exposé plus haut, la pile d'étiquettes peut être sautée à l'avant dernier LSR dans le tunnel.

Un "LSP tunnel acheminé bond par bond" est un tunnel qui est mis en œuvre comme un LSP acheminé bond par bond entre le point d'extrémité de transmission et le point d'extrémité de réception.

Un "LSP tunnel à acheminement explicite" est un LSP tunnel qui est aussi un LSP à acheminement explicite.

3.27.4 Hiérarchie : LSP tunnels au sein des LSP

Considérons un LSP <R1, R2, R3, R4>. Supposons que R1 reçoive un paquet étiqueté P, et pousse sur sa pile d'étiquettes l'étiquette qui va l'amener à suivre ce chemin, et que c'est en fait le chemin bond par bond. Cependant, supposons de plus que R2 et R3 ne sont pas directement connectés, mais sont "voisins" par le fait qu'ils sont les points d'extrémité d'un LSP tunnel. Ainsi, la séquence réelle des LSR traversés par P est <R1, R2, R21, R22, R23, R3, R4>.

Lorsque P voyage de R1 à R2, il va avoir une pile d'étiquettes de profondeur 1. R2, en activant l'étiquette, détermine que P doit entrer dans le tunnel. R2 remplace d'abord l'étiquette entrante par une étiquette qui aura une signification pour R3. Puis il pousse une nouvelle étiquette. Cette étiquette de niveau 2 a une valeur qui a une signification pour R21. La commutation est faite sur l'étiquette de niveau 2 par R21, R22, R23. R23, qui est l'avant dernier bond dans le tunnel R2-R3, saute la pile d'étiquettes avant de transmettre le paquet à R3. Lorsque R3 voit le paquet P, P a seulement une étiquette de niveau 1, étant maintenant sorti du tunnel. Comme R3 est l'avant dernier bond dans le LSP de niveau 1 de P, il saute la pile d'étiquettes, et R4 reçoit P non étiqueté.

Le mécanisme de pile d'étiquettes permet au tunnelage de LSP d'aller à toutes profondeurs.

3.27.5 Homologie et hiérarchie de distribution d'étiquettes

Supposons que le paquet P voyage le long d'un LSP de niveau 1 <R1, R2, R3, R4>, et que lorsque il va de R2 à R3, il voyage le long d'un LSP de niveau 2 <R2, R21, R22, R3>. Dans la perspective du LSP de niveau 2, l'homologue de distribution d'étiquette de R2 est R21. Du point de vue du LSP de niveau 1, les homologues de distribution d'étiquettes de R2 sont R1 et R3. On peut avoir des homologues de distribution d'étiquettes à chaque couche de la hiérarchie. On verra aux paragraphes 4.6 et 4.7 des moyens d'utiliser cette hiérarchie. Noter que dans cet exemple, R2 et R21 doivent être des voisins IGP, mais que R2 et R3 n'ont pas besoin de l'être.

Lorsque deux LSR sont des voisins IGP, on les appelle des "homologues locaux de distribution d'étiquettes". Lorsque deux LSR peuvent être des homologues de distribution d'étiquettes, mais ne sont pas des voisins IGP, on les appelle des "homologues distants de distribution d'étiquettes". Dans l'exemple ci-dessus, R2 et R21 sont des homologues locaux de distribution d'étiquettes, mais R2 et R3 sont des homologues distants de distribution d'étiquettes.

L'architecture MPLS accepte deux façons de distribuer les étiquettes aux différentes couches de la hiérarchie : l'homologie explicite et l'homologie implicite.

L'une effectue la distribution d'étiquettes avec son homologue local de distribution d'étiquette en envoyant des messages de protocole de distribution d'étiquettes qui sont adressés à l'homologue. On peut effectuer la distribution d'étiquettes avec son homologue distant de distribution d'étiquettes d'une des deux façons suivantes :

1. Homologie explicite

Dans l'homologie explicite, on distribue les étiquettes à un homologue en envoyant des messages de protocole de distribution d'étiquettes qui sont adressés à l'homologue, exactement comme on le ferait pour des homologues locaux de distribution d'étiquettes. Cette technique est des plus utiles lorsque le nombre d'homologues distants de distribution d'étiquettes est faible, ou que le nombre de liens d'étiquettes de niveau supérieur est grand, ou que les homologues distants de distribution d'étiquettes sont dans les zones d'acheminement ou domaines distincts. Bien sûr, on a besoin de savoir quelles étiquettes distribuer à quels homologues ; c'est ce qui est traité au paragraphe 4.1.2.

On trouvera des exemples d'utilisation de l'homologie explicite aux paragraphes 4.2.1 et 4.6.

2. Homologie implicite

Dans l'homologie implicite, on n'envoie pas de message de protocole de distribution d'étiquettes adressé à son homologue. Pour distribuer des étiquettes de niveau supérieur à ses homologues distants de distribution d'étiquettes, on code plutôt une étiquette de niveau supérieur comme un attribut d'une étiquette de niveau inférieur, et on distribue ensuite l'étiquette de niveau inférieur, avec cet attribut, à ses homologues locaux de distribution d'étiquettes. Les homologues locaux de distribution d'étiquettes propagent alors les informations à leurs homologues locaux de distribution d'étiquettes. Ce processus continue jusqu'à ce que les informations atteignent l'homologue distant.

Cette technique est des plus utiles lorsque le nombre d'homologues distants de distribution d'étiquettes est grand. L'homologie implicite n'exige pas un maillage de n au carré relations d'échange de trafic pour distribuer les étiquettes aux homologues distants de distribution d'étiquettes parce que les informations sont portées à travers les homologues locaux de distribution d'étiquettes. Cependant, l'homologie implicite exige que les nœuds intermédiaires mémorisent des informations qui peuvent ne pas les intéresser directement.

Un exemple de l'utilisation de l'homologie implicite se trouve au paragraphe 4.3.

3.28 Transport du protocole de distribution d'étiquettes

Un protocole de distribution d'étiquettes est utilisé entre les nœuds dans un réseau MPLS pour établir et entretenir les liens d'étiquettes. Afin que MPLS fonctionne correctement, les informations de distribution d'étiquettes doivent être transmises de façon fiable, et les messages de distribution d'étiquettes qui relèvent d'une certaine FEC ont besoin d'être transmis en séquence. Le contrôle des flux est aussi souhaitable, comme l'est la capacité à porter plusieurs messages d'étiquette dans un seul datagramme.

Une façon d'atteindre ces objectifs est d'utiliser TCP comme transport sous-jacent, comme on le fait dans les [RFC3036] et [RFC3107].

3.29 Pourquoi plus d'un protocole de distribution d'étiquettes ?

La présente architecture n'établit pas de règles fermes et rapides pour choisir quel protocole de distribution d'étiquettes utiliser dans telles ou telles circonstances. Cependant, il est possible de souligner certaines considérations.

3.29.1 BGP et LDP

Dans de nombreux scénarios, il est souhaitable de lier les étiquettes à des FEC qui peuvent être identifiées avec des chemins vers des préfixes d'adresses (voir au paragraphe 4.1). Si il y a un algorithme d'acheminement standard, largement déployé, qui distribue ces chemins, il peut être avancé que la distribution d'étiquettes est le mieux réalisée par le portage de la distribution d'étiquette par la distribution des chemins eux-mêmes.

Par exemple, BGP distribue de tels chemins, et si un locuteur BGP a besoin aussi de distribuer les étiquettes à ses homologues BGP, l'utilisation de BGP pour faire la distribution d'étiquettes (voir la [RFC3107]) présente un certain nombre d'avantages. En particulier, il permet aux réflecteurs de chemins BGP de distribuer les étiquettes, fournissant ainsi un avantage d'adaptabilité significatif par rapport à l'utilisation de LDP pour distribuer les étiquettes entre les homologues BGP.

3.29.2 Étiquettes pour les spécifications de flux RSVP

Lorsque RSVP est utilisé pour établir des réservations de ressource pour certains flux, il peut être souhaitable d'étiqueter les paquets dans ces flux, afin que la spécification de filtre RSVP n'ait pas besoin d'être appliquée à chaque bond. On peut rétorquer que d'avoir RSVP pour distribuer les étiquettes au titre de son processus d'établissement de réservation de chemin est la méthode la plus efficace de distribuer les étiquettes pour cet objet.

3.29.3 Étiquette pour les LSP à acheminement explicite

Dans certaines applications de MPLS, en particulier celles qui se rapportent à l'ingénierie du trafic, il est souhaitable d'établir un chemin à acheminement explicite, de l'entrée à la sortie. Il est aussi souhaitable d'appliquer les réservations de ressources le long de ce chemin.

On peut imaginer deux approches pour cela :

- Commencer par un protocole existant qui est utilisé pour établir les réservations de ressources, et l'étendre pour la prise

en charge de l'acheminement explicite et la distribution d'étiquette.

- Commencer par un protocole existant qui est utilisé pour la distribution d'étiquette, et l'étendre pour accepter l'acheminement explicite et les réservations de ressources.

La première approche a donné naissance au protocole spécifié dans la [RFC3209], la seconde, à l'approche spécifiée dans la [RFC3212].

3.30 Diffusion groupée

Cette section fera l'objet de travaux ultérieurs.

4. Quelques applications de MPLS

4.1 MPLS et trafic acheminé bond par bond

Un certain nombre d'utilisations de MPLS exigent que les paquets avec une certaine étiquette soient transmis le long du même chemin d'acheminement bond par bond que celui qui serait utilisé pour transmettre un paquet avec une adresse spécifiée dans son champ d'adresse de destination de couche réseau.

4.1.1 Étiquettes pour les préfixes d'adresses

En général, le routeur R détermine le prochain bond pour le paquet P en trouvant le préfixe d'adresse X dans son tableau d'acheminement qui est la plus longue correspondance pour l'adresse de destination de P. C'est-à-dire, les paquets dans une FEC donnée sont juste les paquets qui correspondent à un préfixe d'adresse donné dans le tableau d'acheminement de R. Dans ce cas, une FEC peut être identifiée avec un préfixe d'adresse.

Noter qu'un paquet P peut être affecté à une FEC F, et la FEC F peut être identifiée avec le préfixe d'adresse X, même si l'adresse de destination de P ne correspond pas à X.

4.1.2 Distribution des étiquettes pour les préfixes d'adresses

4.1.2.1 Homologues de distribution d'étiquettes pour un préfixe d'adresse

Les LSR R1 et R2 sont considérés comme étant des homologues de distribution d'étiquettes pour le préfixe d'adresse X si et seulement si une des conditions suivantes est réalisée :

1. Le chemin de R1 pour X est un chemin qui est appris via une instance particulière d'un IGP particulier, et R2 est un voisin de R1 dans cette instance de cet IGP
2. Le chemin de R1 vers X est un chemin qui est appris par une instance de l'algorithme d'acheminement A1, et ce chemin est redistribué dans une instance de l'algorithme d'acheminement A2, et R2 est un voisin de R1 dans cette instance de A2
3. R1 est le point d'extrémité de réception d'un LSP tunnel qui est au sein d'un autre LSP, et R2 est un point d'extrémité de transmission de ce tunnel, et R1 et R2 sont des participants à une instance commune d'un IGP, et sont dans la même zone d'IGP (si l'IGP en question a des zones) et le chemin de R1 pour X a été appris via cette instance d'IGP, ou est redistribué par R1 dans cette instance d'IGP.
4. Le chemin de R1 pour X est un chemin qui a été appris via BGP, et R2 est un homologue BGP de R1

En général, ces règles assurent que si le chemin vers un préfixe d'adresses particulier est distribué via un IGP, les homologues de distribution d'étiquette pour ce préfixe d'adresse sont les voisins IGP. Si le chemin pour un préfixe d'adresse particulier est distribué via BGP, les homologues de distribution d'étiquette pour ce préfixe d'adresse sont les homologues BGP. Dans les autres cas de LSP de tunnelage, les points d'extrémité du tunnel sont les homologues de distribution d'étiquettes.

4.1.2.2 Distribution des étiquettes

Afin d'utiliser MPLS pour la transmission des paquets conformément au chemin bond par bond correspondant à tout préfixe d'adresse, chaque LSR DOIT :

1. lier une ou plusieurs étiquettes à chaque préfixe d'adresse qui apparaît dans son tableau d'acheminement;

- pour chacun de ces préfixe d'adresse X, utiliser un protocole de distribution d'étiquettes pour distribuer le lien d'une étiquette avec X à chacun de ses homologues de distribution d'étiquette pour X.

Il y a aussi une circonstance dans laquelle un LSR doit distribuer un lien d'étiquette pour un préfixe d'adresse, même si il n'est pas le LSR qui a lié cette étiquette à ce préfixe d'adresse :

- Si R1 utilise BGP pour distribuer un chemin pour X, en désignant un autre LSR R2 comme prochain bond BGP pour X, et si R1 sait que R2 a alloué l'étiquette L à X, alors R1 doit distribuer le lien entre L et X à tout homologue BGP auquel il distribue ce chemin.

Ces règles assurent que les étiquettes correspondant aux préfixes d'adresse qui correspondent aux chemins BGP sont distribuées aux voisins IGP si et seulement si les chemins BGP sont distribués dans l'IGP. Autrement, les étiquettes liées aux chemins BGP ne sont distribuées qu'aux autres locuteurs BGP.

Ces règles ne sont destinées qu'à indiquer quel lien d'étiquette doit être distribué par un certain LSR à quels autres LSR.

4.1.3 Utilisation du chemin bond par bond comme le LSP

Si le chemin bond par bond que le paquet P a besoin de suivre est $\langle R1, \dots, Rn \rangle$, alors $\langle R1, \dots, Rn \rangle$ peut être un LSP dans la mesure où :

- il y a un seul préfixe d'adresse X, tel que, pour tout i , $1 \leq i < n$, X soit la plus longue correspondance dans le tableau d'acheminement de R_i pour l'adresse de destination de P ;
- pour tout i , $1 < i < n$, R_i ait affecté une étiquette à X et distribué cette étiquette aux $R[i-1]$.

Noter que le LSP d'un paquet ne peut s'étendre que jusqu'à ce qu'il rencontre un routeur dont les tableaux de transmission ont une plus longue meilleure correspondance de préfixe d'adresse pour l'adresse de destination du paquet. À ce point, le LSP doit se terminer et l'algorithme de meilleure correspondance doit être effectué à nouveau.

Supposons, par exemple, que le paquet P, avec l'adresse de destination 10.2.153.178 ait besoin d'aller de R1 à R2 puis à R3. Supposons aussi que R2 annonce le préfixe d'adresse 10.2/16 à R1, mais que R3 annonce 10.2.153/23, 10.2.154/23, et 10.2/16 à R2. C'est-à-dire, R2 annonce un "chemin agrégé" à R1. Dans cette situation, le paquet P peut avoir la commutation d'étiquette jusqu'à ce qu'il atteigne R2, mais comme R2 a effectué l'agrégation de chemin, il doit exécuter l'algorithme de meilleure correspondance pour trouver la FEC de P.

4.1.4 Sortie de LSP et mandataire de sortie de LSP

Un LSR R est considéré comme étant un LSR "de sortie de LSP" pour le préfixe d'adresse X si et seulement si une des conditions suivantes est vérifiée :

- R a une adresse Y, telle que X soit le préfixe d'adresse dans le tableau d'acheminement de R qui a la plus longue correspondance pour Y, ou
- R contient dans ses tableaux d'acheminement un ou plusieurs préfixes d'adresse Y tels que X soit une sous-chaîne initiale appropriée de Y, mais que les "précédents bonds de LSP" de R pour X ne contiennent aucun de ces préfixes d'adresse Y ; c'est-à-dire que R soit un "point de désagrégation" pour le préfixe d'adresse X.

Un LSR R1 est considéré comme LSR "mandataire de sortie de LSP" pour le préfixe d'adresse X si et seulement si :

- le prochain bond de R1 pour X est R2, et si R1 et R2 ne sont pas des homologues de distribution d'étiquettes par rapport à X (peut-être parce que R2 ne prend pas en charge MPLS) ou
- R1 a été configuré pour agir comme un mandataire de sortie de LSP pour X

La définition d'un LSP permet que la sortie de LSP soit un nœud qui ne prend pas en charge MPLS ; dans ce cas, l'avant dernier nœud dans le LSP est le mandataire de sortie.

4.1.5 Étiquette NUL implicite

L'étiquette NUL implicite est une étiquette avec une sémantique particulière qu'un LSR peut lier à un préfixe d'adresse. Si le LSR R_u , en consultant son ILM, voit que le paquet étiqueté P doit être transmis ensuite à R_d , mais que R_d a distribué un lien de NUL implicite au préfixe d'adresse correspondant, au lieu de remplacer alors la valeur de l'étiquette du sommet de la pile d'étiquettes, R_u saute la pile d'étiquettes, puis transmet le paquet résultant à R_d .

Le LSR R_d distribue un lien entre NUL implicite et un préfixe d'adresse X au LSR R_u si et seulement si :

1. les règles du paragraphe 4.1.2 indiquent que Rd distribue à Ru un lien d'étiquette pour X, et
2. Rd sait que Ru peut prendre en charge l'étiquette NUL implicite (c'est-à-dire, il peut sauter la pile d'étiquettes) et
3. Rd est une sortie de LSP (pas un mandataire de sortie) pour X.

Cela est cause que l'avant dernier LSR sur un LSP saute la pile d'étiquettes. Cela est assez approprié ; si la sortie de LSP est une sortie MPLS pour X, alors si l'avant dernier LSR ne saute pas la pile d'étiquettes, la sortie de LSP aura besoin de regarder l'étiquette, de sauter la pile d'étiquettes, puis de regarder la prochaine étiquette (ou de regarder l'adresse de L3, si il n'y a plus d'étiquette présente). En faisant que l'avant dernier LSR saute la pile d'étiquettes, la sortie de LSP épargne le travail d'examen de deux étiquettes afin de prendre sa décision de transmission.

Cependant, si l'avant dernier LSR est un commutateur ATM, il peut n'avoir pas la capacité de sauter la pile d'étiquettes. Donc, un lien de NUL implicite peut n'être distribué qu'aux LSR qui peuvent prendre en charge cette fonction.

Si l'avant dernier LSR sur un LSP pour le préfixe d'adresse X est un mandataire de sortie de LSP, il agit juste comme si la sortie de LSP avait distribué un lien de NUL implicite pour X.

4.1.6 Option d'allocation d'étiquette ciblée sur la sortie

Il y a des situations dans lesquelles une sortie de LSP, Ri, sait que les paquets de plusieurs FEC différentes doivent tous suivre le même LSP, se terminant, par exemple, à la sortie de LSP Re. Dans ce cas, l'acheminement approprié peut être réalisé en utilisant une seule étiquette pour toutes ces FEC ; il n'est pas nécessaire d'avoir une étiquette distincte pour chaque FEC. Si (et seulement si) les conditions suivantes sont réunies :

1. l'adresse du LSR Re est elle-même dans le tableau d'acheminement comme "chemin d'hôte", et
2. il y a un moyen pour Ri de déterminer que Re est le LSP de sortie pour tous les paquets dans un ensemble particulier de FEC.

Ensuite, Ri peut lier une seule étiquette à toutes les FEC de l'ensemble. C'est ce qu'on appelle "allocation d'étiquette ciblée sur la sortie."

Comment le LSR Ri détermine-t-il qu'un LSR Re est la sortie de LSP pour tous les paquets dans une certaine FEC ? Il y a un certain nombre de façons possibles :

- Si le réseau utilise un algorithme d'acheminement à état de liaison, et si tous les nœuds de la zone prennent en charge MPLS, alors l'algorithme d'acheminement fournit Ri avec suffisamment d'informations pour déterminer les routeurs à travers les paquets de cette FEC qui doivent quitter le domaine ou la zone d'acheminement.
- Si le réseau utilise BGP, Ri peut être capable de déterminer que les paquets d'une certaine FEC doivent quitter le réseau via un routeur particulier qui est le "prochain bond BGP" pour cette FEC.
- Il est possible d'utiliser le protocole de distribution d'étiquettes pour passer les informations sur les LSR de sortie auxquels sont "rattachés" les préfixes d'adresse. Cette méthode présente l'avantage de ne pas dépendre de la présence d'un acheminement par état de liaison.

Si l'allocation d'étiquette ciblée sur la sortie est utilisée, le nombre d'étiquettes qui doivent être prises en charge dans l'ensemble du réseau peut être considérablement diminué. Cela peut être significatif si on utilise le matériel de commutation traditionnel pour faire MPLS, et le matériel de commutation ne peut prendre en charge qu'un nombre limité d'étiquettes.

Une approche possible serait de configurer le réseau à utiliser l'allocation d'étiquette ciblée sur la sortie par défaut, mais de configurer certains LSR à NE PAS utiliser l'allocation d'étiquette ciblée sur la sortie pour un ou plusieurs des préfixes d'adresse pour lesquelles ils sont LSP de sortie. Cela impose la règle suivante :

- Si un certain LSR N'EST PAS LSP de sortie pour un ensemble de préfixes d'adresse, il devrait alors allouer des étiquettes aux préfixes d'adresse de la même façon que ce qui est fait par son prochain bond de LSP pour ces préfixes d'adresse. C'est-à-dire, supposons que Rd soit le prochain bond de LSP pour Ru pour les préfixes d'adresse X1 et X2. Si Rd alloue la même étiquette à X1 et à X2, Ru devrait le faire aussi. Si Rd alloue des étiquettes différentes à X1 et à X2, Ru devrait le faire aussi.

Par exemple, supposons qu'on veuille faire de l'allocation d'étiquette ciblée sur la sortie la solution par défaut, mais allouer des étiquettes distinctes à ceux des préfixes d'adresse pour lesquels il y a plusieurs LSP de sortie possibles (c'est-à-dire, pour les préfixes d'adresse qui sont multi rattachement). On peut configurer tous les LSR à utiliser l'allocation d'étiquette ciblée sur la sortie, et configurer ensuite une poignée de LSR à allouer des étiquettes distinctes aux préfixes d'adresse qui sont multi rattachement. Pour un certain préfixe d'adresse multi rattachement X, on aura seulement besoin de configurer cela dans les LSR qui sont soit des sorties de LSP, soit des mandataires de sortie de LSP pour X.

Il est important de noter que si Ru et Rd sont des LSR adjacents dans un LSP pour X1 et X2, la transmission est quand

même faite correctement si Ru alloue des étiquettes distinctes à X1 et à X2 alors que Rd alloue juste une étiquette aux deux. Cela signifie juste que R1 va transposer différentes étiquettes entrantes en la même étiquette sortante, ce qui est un cas ordinaire.

De même, si Rd alloue des étiquettes distinctes à X1 et à X2, mais si Ru leur alloue à tous deux l'étiquette correspondant à l'adresse de leur sortie de LSP ou mandataire de sortie de LSP, la transmission sera quand même faite correctement. Ru va juste transposer l'étiquette entrante en l'étiquette que Rd a alloué à l'adresse de cette sortie de LSP.

4.2 MPLS et LSP à acheminement explicite

Il peut être souhaitable pour un certain nombre de raisons d'utiliser l'acheminement explicite au lieu de l'acheminement bond par bond. Par exemple, cela permet que les chemins soient fondés sur des politiques administratives, et cela permet que les chemins que prennent les LSP soient dessinés avec soin pour permettre l'ingénierie du trafic [RFC2702].

4.2.1 LSP tunnels à acheminement explicite

Dans certaines situations, les administrateurs de réseau peuvent désirer transmettre certaines classes de trafic le long de certains chemins pré-spécifiés, lorsque ces chemins diffèrent du chemin bond par bond que le trafic suivrait ordinairement. Cela peut être fait à l'appui d'une politique d'acheminement, ou en soutien de l'ingénierie du trafic. Le chemin explicite peut être un chemin configuré, ou il peut être déterminé de façon dynamique par certains moyens, par exemple, par un acheminement fondé sur la contrainte.

MPLS permet de faire cela facilement au moyen des LSP tunnel à acheminement explicite. Tout ce qui est nécessaire est :

1. un moyen de choisir les paquets qui sont à envoyer dans le LSP tunnel à acheminement explicite ;
2. un moyen d'établir le LSP tunnel à acheminement explicite ;
3. un moyen de s'assurer que les paquets envoyés dans le tunnel ne vont pas être en boucle entre le point d'extrémité de réception et le point d'extrémité de transmission.

Si le point d'extrémité de transmission du tunnel souhaite mettre un paquet étiqueté dans le tunnel, il doit d'abord remplacer la valeur d'étiquette au sommet de la pile par une valeur d'étiquette qui lui a été distribuée par le point d'extrémité de réception du tunnel. Puis il doit pousser l'étiquette qui correspond au tunnel lui-même, comme elle lui a été distribuée par le prochain bond le long du tunnel. Pour permettre cela, les points d'extrémité du tunnel devraient être des homologues explicites de distribution d'étiquettes. Les liens d'étiquettes qu'ils ont besoin d'échanger ne présentent pas d'intérêt pour les LSR le long du tunnel.

4.3 Piles d'étiquettes et homologie implicite

Supposons qu'un certain LSR Re soit un mandataire de sortie de LSP pour dix préfixes d'adresses, et qu'il atteigne chaque préfixe d'adresse à travers une interface distincte.

On peut affecter une seule étiquette à tous les dix préfixes d'adresses. Ensuite, Re est un LSP de sortie pour tous les dix préfixes d'adresses. Cela assure que les paquets pour tous les dix préfixes d'adresses sont livrés à Re. Cependant, Re aurait alors à examiner l'adresse de couche réseau de chacun de ces paquets afin de choisir la bonne interface sur laquelle envoyer le paquet.

Autrement, on pourrait affecter une étiquette distincte à chaque interface. Alors Re est un mandataire de sortie de LSP pour les dix préfixes d'adresses. Cela élimine le besoin que Re examine les adresses de couche réseau afin de transmettre les paquets. Cependant, il peut en résulter l'utilisation d'un grand nombre d'étiquettes.

Une solution de remplacement serait de lier tous les dix préfixes d'adresses à la même étiquette de niveau 1 (qui est aussi liée à l'adresse du LSR lui-même) puis de lier chaque préfixe d'adresse à une étiquette de niveau 2 distincte. L'étiquette de niveau 2 serait traitée comme un attribut du lien d'étiquette de niveau 1, que nous appelons "attribut de pile". Les règles suivantes s'imposent :

- Lorsque le LSR Ru étiquette initialement un paquet jusqu'alors non étiqueté, si la plus longue correspondance pour l'adresse de destination du paquet est X, et si le prochain bond du LSP de Ru pour X est Rd, et si Rd a distribué à Ru un lien d'étiquette L1 à X, ainsi qu'un attribut de pile de L2, alors
 1. Ru doit pousser L2 et ensuite L1 sur la pile d'étiquette du paquet, et ensuite transmettre le paquet à Rd ;
 2. Lorsque Ru distribue les liens d'étiquettes pour X à ses homologues de distribution d'étiquettes, il doit inclure L2 comme attribut de pile ;
 3. Chaque fois que change l'attribut de pile (éventuellement par suite d'un changement du prochain bond de LSP de Ru pour X) Ru doit distribuer le nouvel attribut de pile.

Noter que bien que la valeur d'étiquette liée à X puisse être différente à chaque bond le long du LSP, la valeur de l'attribut de pile est passée inchangée, et est établie par le mandataire de sortie de LSP.

Donc, le mandataire de sortie de LSP pour X devient un "homologue implicite" avec chaque autre LSR dans la zone ou domaine d'acheminement. Dans ce cas, l'homologie explicite serait trop peu maniable, parce que le nombre d'homologues deviendrait trop important.

4.4 MPLS et acheminement multi chemins

Si un LSR prend en charge plusieurs chemins pour un certain flux, il peut alors affecter plusieurs étiquettes au flux, un pour chaque chemin. Donc la réception d'un second lien d'étiquette de la part d'un voisin particulier pour un certain préfixe d'adresse devrait être comprise comme signifiant que l'une ou l'autre étiquette peut être utilisée pour représenter ce préfixe d'adresse.

Si plusieurs liens d'étiquettes sont spécifiés pour un certain préfixe d'adresse, ils peuvent avoir des attributs distincts.

4.5 Arborescences de LSP comme entités de multipoint à point

Considérons le cas des paquets P1 et P2, dont chacun a une adresse de destination dont la plus longue correspondance, à travers un certain domaine d'acheminement, est le préfixe d'adresse X. Supposons que le chemin bond par bond pour P1 soit <R1, R2, R3>, et que le chemin bond par bond pour P2 soit <R4, R2, R3>. Supposons que R3 lie l'étiquette L3 à X, et distribue ce lien à R2. R2 lie l'étiquette L2 à X, et distribue ce lien aux deux R1 et R4. Lorsque R2 reçoit le paquet P1, son étiquette entrante sera L2. R2 va remplacer L2 par L3, et envoyer P1 à R3. Lorsque R2 reçoit le paquet P2, son étiquette entrante sera aussi L2. R2 remplace à nouveau L2 par L3, et envoie P2 à R3.

Noter que lorsque P1 et P2 voyagent de R2 à R3, ils portent la même étiquette, et pour autant que cela concerne MPLS, ils ne peuvent pas être différenciés. Donc, au lieu de parler des deux LSP distincts, <R1, R2, R3> et <R4, R2, R3>, on devrait parler d'une seule "arborescence de LSP en multipoint à point", qu'on pourrait noter <{R1, R4}, R2, R3>.

Cela crée une difficulté lorsque on essaye d'utiliser les commutateurs ATM conventionnels comme LSR. Dans la mesure où les commutateurs ATM conventionnels ne prennent pas en charge les connexions en multipoint à point, il doit y avoir des procédures pour assurer que chaque LSP est réalisé comme un VC en point à point. Cependant, si les commutateurs ATM qui prennent en charge les VC en multipoint à point sont utilisés, alors les LSP peuvent être très efficacement réalisés comme VC en multipoint à point. Autrement, si le codage multipoint SVP (paragraphe 3.25.2) peut être utilisé, les LSP peuvent être réalisés comme des SVP en multipoint à point.

4.6 LSP tunnel entre routeurs frontière BGP

Considérons le cas d'un système autonome A, qui porte du trafic de transit entre d'autres systèmes autonomes. Le système autonome A aura un certain nombre de routeurs frontières BGP, et un maillage de connexions BGP entre eux, sur lesquelles sont distribués les chemins BGP. Dans nombre de ces cas, il est souhaitable d'éviter de distribuer le chemin BGP aux routeurs qui ne sont pas des routeurs frontières BGP. Si cela peut être évité, la "charge de distribution de chemin" à ces routeurs est significativement réduite. Cependant, il doit y avoir un moyen de s'assurer que le trafic de transit sera livré de routeur frontière à routeur frontière par les routeurs de l'intérieur.

Cela peut être fait facilement au moyen des LSP tunnels. Supposons que les chemins BGP soient distribués seulement aux routeurs frontières BGP, et non aux routeurs intérieurs qui sont le long du chemin bond pas bond entre les deux routeurs frontières. Les LSP tunnels peuvent alors être utilisés comme suit :

1. Chaque routeur frontière BGP distribue, à chaque autre routeur frontière BGP dans le même système autonome, une étiquette pour chaque préfixe d'adresse qu'il distribue à ce routeur via BGP.
2. Le IGP pour le système autonome entretient un chemin d'hôte pour chaque routeur frontière BGP. Chaque routeur intérieur distribue ses étiquettes pour ces chemins d'hôte à chacun de ses voisins IGP.
3. Supposons que :
 - a) le routeur frontière BGP B1 reçoive un paquet non étiqueté P,
 - b) le préfixe d'adresse X dans le tableau d'acheminement de B1 soit la plus longue correspondance pour l'adresse de destination de P,
 - c) le chemin vers X soit un chemin BGP,
 - d) le prochain bond BGP pour X soit B2,
 - e) B2 ait lié l'étiquette L1 à X, et ait distribué ce lien à B1,
 - f) le prochain bond IGP pour l'adresse de B2 soit I1,
 - g) l'adresse de B2 soit dans les tableaux d'acheminement IGP de B1 et de I1 comme un chemin d'hôte,

- h) Il ait lié l'étiquette L2 à l'adresse de B2, et ait distribué ce lien à B1. Alors, avant d'envoyer le paquet P à I1, B1 doit créer une pile d'étiquettes pour P, puis pousser l'étiquette L1, et ensuite pousser l'étiquette L2.
4. Supposons que le routeur frontière BGP B1 reçoive un paquet étiqueté P, où l'étiquette du sommet de la pile d'étiquettes correspond à un préfixe d'adresse, X, dont le chemin est un chemin BGP, et que les conditions 3b, 3c, 3d, et 3e soient toutes respectées. Alors, avant d'envoyer le paquet P à I1, B1 doit remplacer l'étiquette du sommet de la pile d'étiquettes par L1, puis pousser l'étiquette L2.

Avec ces procédures, un certain paquet P suit un LSP de niveau 1 dont tous les membres sont des routeurs frontières BGP, et entre chaque paire de routeurs frontières BGP dans le LSP de niveau 1, il suit un LSP de niveau 2.

Ces procédures créent effectivement un LSP tunnel acheminé bond par bond entre les routeurs frontières BGP.

Comme les routeurs frontières BGP échangent les liens d'étiquettes pour les préfixes d'adresses qui ne sont pas encore connus de l'acheminement IGP, les routeurs BGP devraient devenir des homologues explicites de distribution d'étiquettes les uns les autres.

Il est parfois possible de créer des LSP tunnels acheminés bond par bond entre deux routeurs frontières BGP, même si ils ne sont pas dans le même système autonome. Supposons, par exemple, que B1 et B2 soient dans AS1. Supposons que B3 soit un voisin EBGP de B2, et soit dans AS2. Finalement, supposons que B2 et B3 soient sur un réseau qui est commun aux deux systèmes autonomes (une "zone démilitarisée"). Dans ce cas, un LSP tunnel peut être établi directement entre B1 et B3 comme suit :

- B3 distribue les chemins à B2 (en utilisant EBGP) affectant facultativement des étiquettes aux préfixes d'adresses ;
- B2 redistribue ces chemins à B1 (en utilisant IBGP) en indiquant que le prochain bond BGP pour chacun de ces chemins est B3. Si B3 a affecté des étiquettes aux préfixes d'adresse, B2 passe ces étiquettes inchangées à B1.
- L'IGP de AS1 a un chemin d'hôte pour B3.

4.7 Autres utilisations des LSP tunnels acheminés bond par bond

L'utilisation des LSP tunnels acheminés bond par bond ne se restreint pas aux tunnels entre prochains bonds BGP. Toute situation dans laquelle aurait pu autrement avoir été utilisée une encapsulation de tunnel est appropriée à l'utilisation d'un LSP tunnel à acheminement bond par bond. Au lieu d'encapsuler le paquet avec un nouvel en-tête dont l'adresse de destination est l'adresse du point d'extrémité de réception du tunnel, l'étiquette correspondant au préfixe d'adresse qui est la plus longue correspondance pour l'adresse du point d'extrémité de réception du tunnel est poussée sur la pile d'étiquette du paquet. Le paquet qui est envoyé dans le tunnel peut être ou non déjà étiqueté.

Si le point d'extrémité de transmission du tunnel souhaite mettre un paquet étiqueté dans le tunnel, il doit d'abord remplacer la valeur de l'étiquette au sommet de la pile par la valeur d'étiquette qui lui a été distribuée par le point d'extrémité de réception du tunnel. Il doit ensuite pousser l'étiquette qui correspond au tunnel lui-même, comme elle lui a été distribuée par le prochain bond le long du tunnel. Pour permettre cela, les points d'extrémité du tunnel devraient être des homologues explicites de distribution d'étiquettes. Les liens d'étiquettes qu'ils ont besoin d'échanger ne présentent pas d'intérêt pour les LSR le long du tunnel.

4.8 MPLS et diffusion groupée

L'acheminement en diffusion groupée se fait par la construction d'arborescences de diffusion groupée. Les arborescences le long desquelles un certain paquet en diffusion groupée doit être transmis dépendent en général de l'adresse de source d'un paquet et de son adresse de destination. Chaque fois qu'un LSR particulier est un nœud dans une arborescence de diffusion groupée particulière, il lie une étiquette à cette arborescence. Il distribue alors ce lien à ses parents sur l'arborescence de diffusion groupée. (Si le nœud en question est sur un LAN, et a des frères ou sœurs sur ce LAN, il doit aussi distribuer le lien à ses frères et sœurs. Cela permet au parent d'utiliser une seule valeur d'étiquette lors d'une diffusion groupée à tous les enfants sur le LAN.)

Lorsque un paquet étiqueté arrive en diffusion groupée, la NHLFE correspondant à l'étiquette indique l'ensemble des interfaces de sortie pour ce paquet, ainsi que l'étiquette sortante. Si la même technique de codage d'étiquette est utilisée sur toutes les interfaces sortantes, c'est très exactement le même paquet qui peut être envoyés à tous les enfants.

5. Procédures de distribution d'étiquettes (bond par bond)

Dans la présente section, on ne considère que les liens d'étiquettes qui sont utilisés pour le trafic destiné à la commutation

d'étiquette le long de son chemin bond par bond. Dans ces cas, l'étiquette en question va correspondre à un préfixe d'adresse dans le tableau d'acheminement.

5.1 Procédures d'annonce et d'utilisation des étiquettes

Un certain nombre de procédures différentes peuvent être utilisées pour distribuer les liens d'étiquettes. Certaines sont exécutées par le LSR aval, et certaines autres par le LSR amont.

Le LSR aval doit effectuer :

- la procédure de distribution, et
- la procédure de retrait.

Le LSR amont doit effectuer :

- la procédure de demande,
- la procédure d'indisponibilité,
- la procédure de libération, et
- la procédure d'utilisation d'étiquette.

L'architecture MPLS prend en charge plusieurs variantes de chaque procédure.

Cependant, l'architecture MPLS ne prend pas en charge toutes les combinaisons possibles de toutes les variantes possibles. L'ensemble des combinaisons prises en charge sera décrit au paragraphe 5.2, où l'interopérabilité entre les différentes combinaisons sera aussi exposée.

5.1.1 LSR aval : procédure de distribution

La procédure de distribution est utilisée par un LSR aval pour déterminer quand il devrait distribuer un lien d'étiquette pour un certain préfixe d'adresse à ses homologues de distribution d'étiquettes. L'architecture prend en charge quatre différentes procédures de distribution.

Sans considération de la procédure particulière qui est utilisée, si un lien d'étiquette pour un certain préfixe d'adresse a été distribué par un LSR aval Rd à un LSR amont Ru, et si à tout moment les attributs (comme défini ci-dessus) de ce lien changent, alors Rd doit informer Ru des nouveaux attributs.

Si un LSR entretient plusieurs chemins vers un certain préfixe d'adresse, c'est une affaire locale de déterminer si ce LSR lie plusieurs étiquettes au préfixe d'adresse (une par chemin) et donc, distribue plusieurs liens.

5.1.1.1 PousséeInconditionnelle

Soit Rd un LSR. On suppose que :

1. X est un préfixe d'adresse dans le tableau d'acheminement de Rd,
2. Ru est un homologue de distribution d'étiquette de Rd par rapport à X.

Chaque fois que ces conditions sont remplies, Rd doit lier une étiquette à X et distribuer ce lien à Ru. Il est de la responsabilité de Rd de garder trace des liens qu'il a distribué à Ru, et de s'assurer que Ru a toujours ces liens.

Cette procédure sera utilisée par les LSR qui effectuent l'affectation d'étiquette non sollicitée vers l'aval dans le mode indépendant de contrôle de LSP.

5.1.1.2 PousséeConditionnelle

Soit Rd un LSR. On suppose que :

1. X est un préfixe d'adresse dans le tableau d'acheminement de Rd,
2. Ru est un homologue de distribution d'étiquette de Rd par rapport à X,
3. Rd est soit une sortie de LSP, soit un mandataire de sortie de LSP pour X, ou le prochain bond de L3 de Rd pour X est Rn, où Rn est distinct de Ru, et Rn a lié une étiquette à X et distribué ce lien à Rd.

Alors, aussitôt que toutes ces conditions sont remplies, Rd devrait lier une étiquette à X et distribuer ce lien à Ru.

Tandis que la poussée inconditionnelle cause la distribution des liens d'étiquettes pour tous les préfixes d'adresses dans le tableau d'acheminement, la poussée conditionnelle cause la distribution des liens d'étiquettes pour les seuls préfixes d'adresses pour lesquels on a reçu des liens d'étiquettes de son prochain bond de LSP, ou pour lesquels on n'a pas de prochain bond de L3 à capacité MPLS.

Cette procédure sera utilisée par des LSR qui effectuent une affectation non sollicitée d'étiquette vers l'aval en mode ordonné de contrôle de LSP.

5.1.1.3 TirageInconditionnel

Soit Rd un LSR. On suppose que :

1. X est un préfixe d'adresse dans le tableau d'acheminement de Rd,
2. Ru est un homologue de distribution d'étiquette de Rd par rapport à X,
3. Ru a explicitement demandé que Rd lie une étiquette à X et distribue le lien à Ru.

Alors, Rd devrait lier une étiquette à X et distribuer ce lien à Ru. Noter que si X n'est pas dans le tableau d'acheminement de Rd, ou si Rd n'est pas un homologue de distribution d'étiquette de Ru par rapport à X, alors Rd doit informer Ru qu'il ne peut pas fournir un lien pour le moment.

Si Rd a déjà distribué un lien pour le préfixe d'adresse X à Ru, et si il reçoit une nouvelle demande de Ru pour un lien pour le préfixe d'adresse X, il va lier une seconde étiquette, et distribuer le nouveau lien à Ru. Le premier lien d'étiquette reste en vigueur.

Cette procédure sera utilisée par les LSR qui effectuent la distribution d'étiquette vers l'aval à la demande en utilisant le mode de contrôle de LSP indépendant.

5.1.1.4 TirageConditionnel

Soit Rd un LSR. On suppose que :

1. X est un préfixe d'adresse dans le tableau d'acheminement de Rd,
2. Ru est un homologue de distribution d'étiquette de Rd par rapport à X,
3. Ru a explicitement demandé que Rd lie une étiquette à X et distribue le lien à Ru,
4. Rd est soit une sortie de LSP, soit un mandataire de sortie de LSP pour X, ou le prochain bond de couche 3 de Rd pour X est Rn, où Rn est distinct de Ru, et Rn a lié une étiquette à X et a distribué ce lien à Rd.

Alors, aussitôt que ces conditions sont toutes remplies, Rd devrait lier une étiquette à X et distribuer ce lien à Ru. Noter que si X n'est pas dans le tableau d'acheminement de Rd et si un lien pour X ne peut pas être obtenu via le prochain bond de Rd pour X, ou si Rd n'est pas un homologue de distribution d'étiquette de Ru par rapport à X, alors Rd doit informer Ru qu'il ne peut pas fournir un lien pour le moment.

Cependant, si la seule condition non remplie est que Rn n'a pas encore fourni une étiquette à Rd, alors Rd doit différer toute réponse à Ru jusqu'au moment où il aura reçu un lien de la part de Rn.

Si Rd a distribué un lien d'étiquette pour le préfixe d'adresse X à Ru, et qu'après un moment, un attribut quelconque du lien d'étiquette change, alors Rd doit redistribuer le lien d'étiquette à Ru, avec le nouvel attribut. Il doit faire cela même si Ru ne formule pas une nouvelle demande.

Cette procédure sera utilisée par les LSR qui effectuent l'allocation d'étiquette vers l'aval à la demande dans le mode ordonné de contrôle de LSP.

Au paragraphe 5.2, on expose comment choisir la procédure particulière à utiliser à un moment donné, et comment s'assurer de l'interopérabilité parmi les LSR qui choisissent des procédures différentes.

5.1.2 LSR amont : procédure de demande

La procédure de demande est utilisée par le LSR amont pour un préfixe d'adresse pour déterminer quand demander explicitement que le LSR aval lie une étiquette à ce préfixe et distribue le lien. Il y a trois procédures possibles qui peuvent être utilisées.

5.1.2.1 JamaisDeDemande

Ne jamais formuler de demande. C'est utile si le LSR aval utilise la procédure de poussée conditionnelle ou de poussée inconditionnelle, mais elle n'est pas utile si le LSR aval utilise la procédure Tirage inconditionnel ou Tirage conditionnel.

Cette procédure sera utilisée par un LSR lorsque sont utilisés la distribution non sollicitée d'étiquette vers l'aval et le mode libéral de rétention d'étiquette.

5.1.2.2 DemandeQuandNécessaire

Faire une demande chaque fois que change le prochain bond de couche 3 pour le préfixe d'adresse, ou quand un nouveau préfixe d'adresse est appris, et qu'on n'a pas déjà un lien d'étiquette provenant de ce prochain bond pour ce préfixe d'adresse.

Cette procédure sera utilisée par un LSR chaque fois que le mode prudent de rétention d'étiquette est utilisé.

5.1.2.3 DemandeSurDemande

Faire une demande chaque fois qu'une demande est reçue, en plus de produire une demande lorsque nécessaire (comme décrit au paragraphe 5.1.2.2). Si Ru n'est pas capable d'être une sortie de LSP, il ne peut faire une demande que lorsque il reçoit une demande de l'amont.

Si Rd reçoit une telle demande de Ru, pour un préfixe d'adresse pour lequel Rd a déjà distribué une étiquette à Ru, Rd devra affecter une nouvelle étiquette (distincte), la lier à X, et distribuer ce lien. (Savoir si Rd peut distribuer ce lien à Ru immédiatement ou non dépend de la procédure de distribution utilisée.)

Cette procédure sera utilisée par un LSR qui effectue la distribution d'étiquette vers l'aval à la demande, mais ne fait pas la fusion d'étiquette, par exemple, un LSR ATM qui n'est pas capable de fusion de VC.

5.1.3 LSR amont : procédure NonDisponible

Si Ru et Rd sont respectivement les homologues de distribution d'étiquettes amont et aval pour le préfixe d'adresse X, et si Rd est le prochain bond de couche 3 de Ru pour X, et si Ru demande un lien pour X de la part de Rd, mais si Rd répond qu'il ne peut pas fournir de lien pour le moment, parce qu'il n'a pas de prochain bond pour X, alors la procédure NonDisponible détermine comment Ru répond. Deux procédures possibles gouvernent le comportement de Ru.

5.1.3.1 Redemander

Ru devrait produire à nouveau la demande ultérieurement. C'est-à-dire que le demandeur est chargé de réessayer plus tard d'obtenir le lien nécessaire. Cette procédure sera utilisée avec la distribution d'étiquette vers l'aval à la demande.

5.1.3.2 NePasRedemander

Ru ne devrait jamais réessayer la demande, en supposant plutôt que Rd va fournir automatiquement le lien lorsque il sera disponible. Ceci est utile si Rd utilise la procédure de poussée inconditionnelle ou celle de poussée conditionnelle, c'est-à-dire, si on utilise la distribution d'étiquette non sollicitée vers l'aval.

Noter que si Rd répond qu'il ne peut pas fournir un lien à Ru, à cause d'une condition d'erreur, plutôt que parce que Rd n'a pas de prochain bond, le comportement de Ru sera gouverné par les conditions de récupération d'erreur du protocole de distribution de l'étiquette, plutôt que par la procédure NePasRedemander.

5.1.4 LSR amont : procédure de libération

Supposons que Rd soit un LSR qui a lié une étiquette au préfixe d'adresse X, et ait distribué ce lien au LSR Ru. Si Rd ne se trouve pas être le prochain bond de couche 3 de Ru pour le préfixe d'adresse X, ou a cessé d'être le prochain bond de L3 de Ru pour le préfixe d'adresse X, alors Ru ne va pas utiliser l'étiquette. La procédure Libération détermine comment agit Ru dans ce cas. Il y a deux procédures possibles qui gouvernent le comportement de Ru.

5.1.4.1 LibérationSurChangement

Ru devrait libérer le lien, et informer Rd qu'il l'a fait. Cette procédure sera utilisée pour mettre en œuvre le mode prudent de rétention d'étiquette.

5.1.4.2 PasDeLibérationSurChangement

Ru devrait conserver le lien, afin qu'il puisse le réutiliser immédiatement si Rd devient ultérieurement le prochain bond de couche 3 de Ru pour X. Cette procédure sera utilisée pour mettre en œuvre le mode libéral de rétention d'étiquette.

5.1.5 LSR amont : procédure UsageD'Étiquette

Supposons que Ru soit un LSR qui a reçu un lien d'étiquette L pour le préfixe d'adresse X du LSR Rd, et que Ru soit en amont de Rd par rapport à X, et qu'en fait Rd soit le prochain bond L3 de Ru pour X.

Ru va utiliser le lien si Rd est le prochain bond L3 de Ru pour X. Si, au moment où le lien est reçu par Ru, Rd N'EST PAS le prochain bond L3 de Ru pour X, Ru n'utilise pas le lien à ce moment là. Ru peut cependant commencer à utiliser le lien ultérieurement, si Rd devient le prochain bond L3 de Ru pour X.

La procédure UsageD'Étiquette détermine seulement comment Ru utilise le lien de Rd.

Ru peut utiliser deux procédures.

5.1.5.1 UsageImmédiat

Ru peut mettre immédiatement le lien en service. À tout moment lorsque Ru a un lien pour X provenant de Rd, et quand Rd est le prochain bond de couche 3 de Ru pour X, Rd va aussi être le prochain bond de LSP de Ru pour X. Cette procédure est utilisée lorsque la détection de boucle n'est pas en service

5.1.5.2 UsageSiBoucleNonDétectée

Cette procédure est la même que UsageImmédiat, sauf que Ru a détecté une boucle dans le LSP. Si une boucle a été détectée, Ru va arrêter d'utiliser l'étiquette L pour transmettre les paquets à Rd.

Cette procédure est utilisée lorsque la détection de boucle est en service.

Cela va continuer jusqu'à ce que change le prochain bond pour X, ou jusqu'à ce que la boucle ne soit plus détectée.

5.1.6 LSR aval : Procédure de retrait

Dans ce cas, il n'y a qu'une seule procédure.

Lorsque le LSR Rd décide de rompre le lien entre l'étiquette L et le préfixe d'adresse X, cette rupture doit alors être distribuée à tous les LSR auxquels le lien avait été distribué.

Il est exigé que la rupture entre L et X soit distribuée par Rd à un LSR Ru avant que Rd ne distribue à Ru un nouveau lien de L à tout autre préfixe d'adresse Y, où $X \neq Y$. Si Ru devait apprendre le nouveau lien de L à Y avant d'apprendre la rupture entre L et X, et si des paquets correspondant à la fois à X et à Y étaient transmis par Ru à Rd, alors pendant un certain temps, Ru étiquetterait à la fois les paquets correspondants à X et ceux correspondant à Y avec l'étiquette L.

La distribution et le retrait des liens d'étiquettes sont faits via un protocole de distribution d'étiquettes. Tous les protocoles de distribution d'étiquettes exigent qu'une adjacence de distribution d'étiquette soit établie entre deux homologues de distribution d'étiquettes (sauf les homologues implicites). Si le LSR R1 a une adjacence de distribution d'étiquette au LSR R2, et a reçu des liens d'étiquettes du LSR R2 via cette adjacence, alors si l'adjacence est rompue par l'un ou l'autre homologue (par suite d'une défaillance ou au titre du fonctionnement normal) tous les liens reçus sur cette adjacence doivent être considérés comme ayant été retirés.

Tant que l'adjacence de distribution d'étiquette pertinente reste en place, les liens d'étiquettes qui sont retirés doivent toujours être retirés de façon explicite. Si une seconde étiquette est liée à un préfixe d'adresse, le résultat n'est pas de retirer implicitement la première étiquette, mais de lier les deux étiquettes ; ceci est nécessaire pour la prise en charge de l'acheminement multi-chemins. Si un second préfixe d'adresse est lié à une étiquette, le résultat n'est pas de retirer implicitement le lien de cette étiquette au premier préfixe d'adresse, mais d'utiliser cette étiquette pour les deux préfixes d'adresses.

5.2 Schémas MPLS : combinaisons de procédures prises en charge

Considérons deux LSR, Ru et Rd, qui sont des homologues de distribution d'étiquettes par rapport à un certain ensemble de préfixes d'adresses, où Ru est l'homologue amont et Rd est l'homologue aval.

Le schéma MPLS qui gouverne l'interaction de Ru et de Rd peut être décrit comme un quintuplet de procédures : <procédure de distribution, procédure de demande, procédure NonDisponible, procédure de libération, procédure UsageD'Étiquette>. (Comme il n'y a qu'une seule procédure de retrait, il n'est pas besoin de la mentionner.) Une "*" qui apparaît dans une des positions est un caractère générique, ce qui signifie que toute procédure dans cette catégorie peut être présente ; un "N/A" qui apparaît dans une certaine position indique qu'aucune procédure n'est nécessaire dans cette catégorie.

Seuls les schémas MPLS qui sont spécifiés ci-dessous sont acceptés par l'architecture MPLS. D'autres schémas pourront être ajoutés à l'avenir, si le besoin s'en faisait sentir.

5.2.1 Schémas pour les LSR qui acceptent la fusion d'étiquette

Si Ru et Rd sont des homologues de distribution d'étiquettes, et si tous deux prennent en charge la fusion d'étiquette, un des schémas suivants doit être utilisé :

1. <PousséeInconditionnelle, JamaisDeDemande, N/A, PasDeLibérationSurChangement, UsageImmédiat>

C'est la distribution non sollicitée d'étiquette vers l'aval avec le mode de rétention libéral d'étiquette, en contrôle indépendant, et pas de détection de boucle.

2. <PousséeInconditionnelle, JamaisDeDemande, N/A, PasDeLibérationSurChangement, UsageSiBoucleNonDétectée>

C'est la distribution non sollicitée d'étiquette vers l'aval avec mode de contrôle indépendant, rétention libérale d'étiquette, et détection de boucle.

3. <PousséeConditionnelle, DemandeQuandNécessaire, NePasRedemander, LibérationSurChangement, *>

C'est la distribution non sollicitée d'étiquette vers l'aval avec contrôle ordonné (à partir de la sortie) et mode prudent de rétention d'étiquette. La détection de boucle est facultative.

4. <PousséeConditionnelle, JamaisDeDemande, N/A, PasDeLibérationSurChangement, *>

C'est la distribution non sollicitée d'étiquette vers l'aval avec contrôle ordonné (à partir de la sortie) et mode libéral de rétention d'étiquette. La détection de boucle est facultative.

5. <TirageConditionnel, DemandeQuandNécessaire, Redemander, LibérationSurChangement, *>

C'est la distribution d'étiquette à la demande vers l'aval avec contrôle ordonné (initié par l'entrée), en mode prudent de rétention d'étiquette, et détection de boucle facultative.

6. <TirageInconditionnel, DemandeQuandNécessaire, N/A, LibérationSurChangement, UsageImmédiat>

C'est la distribution d'étiquette à la demande vers l'aval avec contrôle indépendant et mode prudent de rétention d'étiquette, sans détection de boucle.

7. <TirageInconditionnel, DemandeQuandNécessaire, N/A, LibérationSurChangement, UsageSiBoucleNonDétectée>

C'est la distribution d'étiquette à la demande vers l'aval avec contrôle indépendant et mode prudent de rétention d'étiquette, avec détection de boucle.

5.2.2 Schéma pour LSR qui n'acceptent pas la fusion d'étiquette

Supposons que R1, R2, R3, et R4 soient des commutateurs ATM qui ne prennent pas en charge la fusion d'étiquette, mais sont utilisés comme des LSR. Supposons de plus que le chemin bond par bond de couche 3 pour le préfixe d'adresse X soit <R1, R2, R3, R4>, et que les paquets destinés à X puissent entrer dans le réseau à l'un de ces LSR. Comme il n'y a pas de capacité multipoint à point, les LSP doivent être réalisés comme des circuits virtuels en point à point, ce qui signifie qu'il doit y avoir trois de ces circuits virtuels pour le préfixe d'adresse X : <R1, R2, R3, R4>, <R2, R3, R4>, et <R3, R4>.

Donc, si R1 et R2 sont des homologues MPLS, et soit sont des LSR qui sont mis en œuvre en utilisant un matériel conventionnel de commutation ATM (c'est-à-dire, sans suppression d'entrelacement de cellule) soit sont par ailleurs incapables d'effectuer la fusion d'étiquette, le schéma MPLS utilisé entre R1 et R2 doit être un des suivants :

1. <TirageConditionnel, DemandeSurDemande, Redemander, LibérationSurChangement, *>

C'est la distribution d'étiquette à la demande vers l'aval avec contrôle ordonné (initié par l'entrée), mode prudent de rétention d'étiquette, et détection de boucle facultative.

L'utilisation de la procédure DemandeSurDemande va être cause que R4 distribue trois étiquettes pour X à R3 ; R3 va distribuer deux étiquettes pour X à R2, et R2 va distribuer une étiquette pour X à R1.

2. <TirageInconditionnel, DemandeSurDemande, N/A, LibérationSurChangement, UsageImmédiat>

C'est la distribution d'étiquette à la demande vers l'aval avec contrôle indépendant et mode prudent de rétention d'étiquette, sans détection de boucle.

3. <TirageInconditionnel, DemandeSurDemande, N/A, LibérationSurChangement, UsageSiBoucleNonDétectée>

C'est la distribution d'étiquette à la demande vers l'aval avec contrôle indépendant et mode prudent de rétention d'étiquette, avec détection de boucle.

5.2.3 Considérations d'interopérabilité

Il est aisé de voir que certains quintuplets NE DONNENT PAS de schémas MPLS viables. Par exemple :

- <TirageInconditionnel, JamaisDeDemande, *, *, *>
- <TirageConditionnel, JamaisDeDemande, *, *, *>

Dans ces schémas MPLS, le LSR aval Rd ne distribue les liens d'étiquettes au LSR amont Ru que sur demande de Ru, mais Ru ne fera jamais de telles demandes. Il est évident que ces schémas ne sont pas viables, car ils ne résultent pas en une distribution correcte des liens d'étiquettes.

- <*, JamaisDeDemande, *, *, LibérationSurChangement>

Dans ces schémas MPLS, Rd libère les liens lorsque il ne les utilise pas, mais il ne les redemande jamais, même si il en a besoin ultérieurement. Donc ces schémas n'assurent pas que les liens d'étiquettes sont correctement distribués.

Dans ce paragraphe, on spécifie les règles pour empêcher une paire d'homologues de distribution d'étiquettes d'adopter des procédures qui conduisent à des schémas MPLS impraticables. Ces règles exigent l'échange d'informations entre les homologues de distribution d'étiquettes durant l'initialisation de l'adjacence de distribution d'étiquettes, ou une connaissance à priori des informations (obtenues par des moyens qui sortent du domaine d'application de ce document).

1. Chacun doit déclarer si il prend en charge la fusion d'étiquette.
2. Si Rd ne prend pas en charge la fusion d'étiquette, Rd doit choisir la procédure TirageInconditionnel ou la procédure TirageConditionnel. Si Rd choisit le TirageConditionnel, Ru est forcé d'utiliser la procédure Redemander.

C'est-à-dire que si le LSR aval ne prend pas en charge la fusion d'étiquette, ses préférences prennent le pas lors du choix du schéma MPLS.

3. Si Ru ne prend pas en charge la fusion d'étiquette, mais si Rd le fait, Ru doit choisir la procédure Redemander ou la procédure NePasRedemander. Cela force Rd à utiliser respectivement la procédure TirageConditionnel ou la procédure TirageInConditionnel.

C'est-à-dire que si un seul des LSR ne prend pas en charge la fusion d'étiquette, ses préférences ont la priorité lors du choix du schéma MPLS.

4. Si Ru et Rd prennent tous deux en charge la fusion d'étiquette, le choix entre mode libéral et prudent de rétention d'étiquette appartient alors à Ru. C'est-à-dire que Ru a à choisir entre utiliser DemandeQuandNécessaire/LibérationSurChangement (prudent) ou utiliser JamaisDeDemande/PasDeLibérationSurChangement (libéral). Cependant, le choix de "pousser" ou de "tirer" et de "conditionnel" ou de "inconditionnel" appartient à Rd. Si Ru choisit le mode libéral de rétention d'étiquette, Rd peut choisir PousséeInconditionnelle ou PousséeConditionnelle. Si Ru choisit le mode prudent de rétention d'étiquette, Rd peut choisir PousséeConditionnelle, TirageConditionnel, ou TirageInconditionnel.

Ces choix déterminent tous ensemble le schéma MPLS utilisé.

6. Considérations pour la sécurité

Certains routeurs peuvent mettre en œuvre des procédures de sécurité qui dépendent de ce que l'en-tête de couche réseau se trouve à un endroit fixé par rapport à l'en-tête de couche liaison des données. L'encapsulation générique MPLS insère une cale d'ajustement entre l'en-tête de couche liaison des données et l'en-tête de couche réseau. Cela peut causer l'échec de telles procédures de sécurité.

Une étiquette MPLS a sa signification par la vertu d'un accord entre le LSR qui met l'étiquette dans la pile d'étiquettes (le "rédacteur d'étiquettes") et le LSR qui interprète cette étiquette (le "lecteur d'étiquette"). Si les paquets étiquetés sont acceptés de sources qui ne sont pas de confiance, ou si une étiquette entrante particulière est acceptée d'un LSR auquel

cette étiquette n'a pas été distribuée, les paquets peuvent alors être acheminés d'une façon illégitime.

7. Propriété intellectuelle

Des droits de propriété intellectuelle revendiqués à l'égard de tout ou partie de la spécification contenue dans le présent document ont été notifiés à l'IETF. Pour plus d'informations, consulter la liste en ligne des revendications de droits.

8. Adresse des auteurs

Eric C. Rosen
Cisco Systems, Inc.
250 Apollo Drive
Chelmsford, MA, 01824
USA
mél : erosen@cisco.com

Arun Viswanathan
Force10 Networks, Inc.
1440 McCarthy Blvd.
Milpitas, CA 95035-7438
USA
mél : arun@force10networks.com

Ross Callon
Juniper Networks, Inc.
1194 North Mathilda Avenue
Sunnyvale, CA 94089
USA
mél : rcallon@juniper.net

9. Références

- [RFC2702] D. Awduche et autres, "Exigences d'[ingénierie du trafic sur MPLS](#)", septembre 1999. (*Information*)
- [RFC2935] D. Eastlake 3rd, C. Smith, "Supplément HTTP au protocole Internet de commerce ouvert (IOTP)", septembre 2000. (*P.S.*)
- [RFC3032] E. Rosen et autres, "[Codage de pile d'étiquettes MPLS](#)", janvier 2001.
- [RFC3034] A. Conta, P. Doolan, A. Malis, "Spécification de l'[utilisation de la commutation d'étiquettes](#) sur les réseaux en relais de trame", janvier 2001. (*P.S.*)
- [RFC3036] L. Andersson et autres, "[Spécification de LDP](#)", janvier 2001. (*Rendue obsolète par la RFC5036*)
- [RFC3107] Y. Rekhter et E. Rosen, "Portage des informations d'étiquette dans BGP-4", mai 2001.
- [RFC3209] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan et G. Swallow, "RSVP-TE : Extensions à RSVP pour les tunnels LSP", décembre 2001. (Mise à jour par [RFC3936](#), [RFC4420](#), [RFC4874](#), [RFC5151](#), [RFC5420](#))
- [RFC3212] B. Jamoussi et autres, "Établissement de [LSP fondé sur la contrainte](#) avec LDP", janvier 2002. (*MàJ par RFC3468*) (*P.S.*)

10. Déclaration complète de droits de reproduction

Copyright (c) The Internet Society (2001). Tous droits réservés.

Le présent document et ses traductions peuvent être copiés et fournis aux tiers, et les travaux dérivés qui les commentent ou les expliquent ou aident à leur mise en œuvre peuvent être préparés, copiés, publiés et distribués, en tout ou partie, sans restriction d'aucune sorte, pourvu que la déclaration de copyright ci-dessus et le présent et paragraphe soient inclus dans toutes telles copies et travaux dérivés. Cependant, le présent document lui-même ne peut être modifié d'aucune façon, en particulier en retirant la notice de droits de reproduction ou les références à la Internet Society ou aux autres organisations Internet, excepté autant qu'il est nécessaire pour le besoin du développement des normes Internet, auquel cas les procédures de droits de reproduction définies dans les procédures des normes Internet doivent être suivies, ou pour les besoins de la traduction dans d'autres langues que l'anglais.

Les permissions limitées accordées ci-dessus sont perpétuelles et ne seront pas révoquées par la Internet Society, ses successeurs ou ayant droits.

Le présent document et les informations y contenues sont fournies sur une base "EN L'ÉTAT" et le contributeur, l'organisation qu'il ou elle représente ou qui le/la finance (s'il en est), la INTERNET SOCIETY et la INTERNET ENGINEERING TASK FORCE déclinent toutes garanties, exprimées ou implicites, y compris mais non limitées à toute

garantie que l'utilisation des informations ci encloses ne violent aucun droit ou aucune garantie implicite de commercialisation ou d'aptitude à un objet particulier.

Remerciement

Le financement de la fonction d'édition des RFC est actuellement fourni par l'Internet Society.