

Groupe de travail Réseau
Request for Comments : 3353
 Catégorie : Information
 août 2002

D. Ooms, Alcatel
 B. Sales, Alcatel
 W. Livens, Colt Telecom
 A. Acharya, IBM
 F. Griffoul, Ulticom
 F. Ansari, Bell Labs

Traduction Claude Brière de L'Isle

Vue d'ensemble de la diffusion groupée IP dans un environnement de commutation d'étiquettes multi-protocoles (MPLS)

Statut de ce mémoire

Le présent mémoire apporte des informations pour la communauté de l'Internet. Il ne spécifie aucune norme de l'Internet. La distribution du présent mémoire n'est soumise à aucune restriction.

Notice de copyright

Copyright (C) The Internet Society (2002). Tous droits réservés

Résumé

Le présent document offre un cadre pour le déploiement de la diffusion groupée IP dans un environnement MPLS. Les questions qui se posent lorsque les techniques MPLS sont appliquées à la diffusion groupée IP sont présentées. Les avantages et inconvénients des protocoles existants d'acheminement de diffusion groupée IP dans le contexte de MPLS sont décrits et les relations aux différentes méthodes de déclenchement et aux modes de distribution d'étiquettes sont discutées. Les conséquences des diverses technologies de couche 2 (L2) sont présentées. Les réseaux point à point et multi-accès sont considérés.

Table des Matières

1. Introduction.....	2
2. Caractéristiques de la couche 2.....	3
3. Taxonomie des protocoles d'acheminement de diffusion groupée IP dans le contexte MPLS.....	3
3.1 Agrégation.....	3
3.2 Arrosage et élagage.....	3
3.3 Arborescences de source/partagées.....	4
3.4 Coexistence d'arborescence de source et d'arborescence partagée.....	4
3.5 Arborescences partagées uni/bidirectionnelles.....	6
3.6 Données de diffusion groupée encapsulées.....	7
3.7 Suppression des boucles.....	7
3.8 Transposition des caractéristiques sur les protocoles existants.....	7
4. Transmission L2/L3 mixte dans un seul nœud.....	7
5. Taxonomie des déclencheurs de LSP de diffusion groupée IP.....	9
5.1 Piloté par la demande.....	9
5.2 Piloté par la topologie.....	10
5.3 Piloté par le trafic.....	10
5.4 Combinaisons des déclencheurs et des modes de distribution d'étiquettes.....	11
6. Portage.....	11
7. Acheminement explicite.....	12
8. QoS/CoS.....	12
8.1 DiffServ.....	12
8.2 IntServ et RSVP.....	12
9. Réseaux multi-accès.....	13
10. Autres questions.....	14
10.1 Champ TTL.....	14
10.2 Contrôle de distribution d'étiquettes indépendant ou ordonné.....	14
10.3 Mode de rétention d'étiquette prudent ou libéral.....	14
10.4 Allocation d'étiquettes vers l'amont ou vers l'aval.....	15
10.5 Distribution d'étiquettes explicite ou implicite.....	15
11. Considérations pour la sécurité.....	15
12. Remerciements.....	16
Références.....	16
Adresse des auteurs.....	17

Déclaration complète de droits de reproduction.....17

Tableau des abréviations

ATM (<i>Asynchronous Transfer Mode</i>)	technique temporelle asynchrone
CBT (<i>Core Based Tree</i>)	arborescence fondée sur le cœur
CoS (<i>class of service</i>)	classe de service
DLCI (<i>Data Link Connection Identifier</i>)	identifiant de connexion de liaison de données (RFC2427)
DRrecv (<i>Designated Router of the receiver</i>)	routeur désigné du receveur
DRsend (<i>Designated Router of the sender</i>)	routeur désigné de l'expéditeur
DVMRP (<i>Distance Vector Multicast Routing Protocol</i>)	protocole d'acheminement de diffusion groupée par vecteur de distance
FEC (<i>Forwarding Equivalence Class</i>)	classe d'équivalence d'acheminement (RFC3031)
FR (<i>Frame relay</i>)	relais de trame
IGMP (<i>Internet Group Management Protocol</i>)	protocole de gestion de groupe sur Internet
IP (<i>Internet Protocol</i>)	protocole Internet
L2 (layer 2)	couche 2 (par exemple ATM, relais de trame)
L3 (layer 3)	couche 3 (par exemple, IP)
LSP (<i>Label Switched Path</i>)	chemin commuté par étiquettes
LSR (<i>label switched router</i>)	routeur commuté par étiquettes
LSRd (<i>Downstream LSR</i>)	routeur commuté par étiquettes aval
LSRu (<i>Upstream LSR</i>)	routeur commuté par étiquettes amont
MOSPF (<i>Multicast OSPF</i>)	diffusion groupée en OSPF
mp2mp (<i>multipoint-to-multipoint</i>)	multipoint à multipoint
MRT (<i>Multicast Routing Table</i>)	tableau d'acheminement de diffusion groupée
p2mp (<i>point-to-multipoint</i>)	point à multipoint
PIM-DM (<i>Protocol Independent Multicast-Dense Mode</i>)	diffusion groupée indépendante du protocole – mode dense
PIM-SM (<i>Protocol Independent Multicast-Sparse Mode</i>)	diffusion groupée indépendante du protocole – mode épars
QoS (<i>Quality of Service</i>)	qualité de service
RP (<i>Rendezvous Point</i>)	point de rendez-vous
RPT-bit (<i>RP Tree bit</i>)	bit d'arborescence de point de rendez-vous
RSVP (<i>Resource ReserVation Protocol</i>)	= protocole de réservation de ressources
SPT-bit (<i>Shortest Path Tree bit</i>)	bit d'arborescence de plus court chemin
SSM (<i>Source-Specific Multicast</i>)	diffusion groupée spécifique de la source
TCP (<i>Transmission Control Protocol</i>)	protocole de contrôle de transmission (SCSSI)
UDP (<i>user datagram protocol</i>)	protocole de datagramme d'utilisateur
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

1. Introduction

Dans un nuage MPLS, les chemins sont déterminés par un protocole d'acheminement de couche 3. Ces chemins peuvent alors être transposés en chemins de couche 2 pour améliorer les performances du réseau. À côté de cela, MPLS offre un véhicule pour des services réseau améliorés tels que QoS/CoS, ingénierie du trafic, etc.

Les protocoles courants d'acheminement en envoi individuel génèrent un même chemin le plus court (optimal) en régime permanent pour une certaine paire source-destination. On remarque que les protocoles d'envoi individuel peuvent se comporter de façon légèrement différente par rapport à des chemins de coût égal.

Pour la diffusion groupée, la solution optimale (coût minimum d'interconnexion de N nœuds) imposerait le calcul d'une arborescence de Steiner. Malheureusement, aucun protocole d'acheminement de diffusion groupée n'est aujourd'hui capable d'entretenir une telle arborescence optimale. Différents protocoles de diffusion groupée vont donc en général, générer des arborescences différentes.

La discussion se concentre sur les protocoles d'acheminement de diffusion groupée intra-domaine. Les aspects de l'acheminement inter-domaine sortent du domaine d'application du présent document.

2. Caractéristiques de la couche 2

Bien que MPLS soit multi protocoles aux deux couches 3 et 2, en pratique IP est le seul protocole de couche 3 considéré. MPLS peut fonctionner par dessus plusieurs technologies de couche 2 (PPP/Sonet, Ethernet, ATM, FR, ...).

Lorsque la commutation d'étiquettes est transposée sur des capacités de commutation de couche 2 (par exemple, VPI/VCI est utilisé comme étiquette) l'attention se concentre principalement sur la transposition en ATM [RFC3035]. ATM offre de fortes capacités de commutation et de QS, mais dans le contexte de MPLS, il présente plusieurs limitations qui sont décrites dans la [RFC3035]. Des considérations similaires sont faites pour le relais de trame sur L2 dans la [RFC3034]. On peut résumer ainsi ces limitations :

- Espace limité d'étiquettes : le nombre de bits normalisé ou mis en œuvre disponible pour une étiquette peut être faible (par exemple, espace de VPI/VCI, espace DLCI) limitant le nombre de LSP qui peuvent être établis.
- Fusion : certaines technologies ou mises en œuvre de couche 2 de ces technologies ne prennent pas en charge les 'connexions' multipoint à point et/ou multipoint à multipoint, empêchant la fusion des LSP.
- TTL : les technologies de couche 2 ne prennent pas en charge la fonction de décrémentation du TTL.

Ces trois limitations peuvent impacter la mise en œuvre de la diffusion groupée dans MPLS comme on va le décrire dans le présent document.

Lorsque MPLS natif est déployé, les limitations ci-dessus disparaissent. De plus, sur les liaisons PPP et Ethernet, la même étiquette peut être utilisée au même moment pour un LSP d'envoi individuel et un LSP de diffusion groupée parce que des EtherTypes différents sont définis pour MPLS en envoi individuel et en diffusion groupée [RFC3032].

3. Taxonomie des protocoles d'acheminement de diffusion groupée IP dans le contexte MPLS

Pour le moment, une abondance de protocoles d'acheminement de diffusion groupée IP est proposée et développée. Tous ces protocoles ont des caractéristiques différentes (adaptabilité, complexité de calcul, latence, redondance de message de contrôle, type d'arborescence, etc...). L'objet du présent document n'est pas de donner une taxonomie complète des protocoles d'acheminement de diffusion groupée IP, et on donnera seulement leurs caractéristiques pertinentes pour la technologie MPLS.

On examinera les caractéristiques suivantes :

- Agrégation
- Arrosage et élagage
- Arborescence de source/partagée
- Coexistence d'arborescences de source et partagées
- Arborescences partagées uni/bidirectionnelles
- Données de diffusion groupée encapsulées
- Absence de boucle

La discussion de ces caractéristiques ne va pas conduire à la sélection d'un protocole d'acheminement de diffusion groupée supérieur. Il n'est pas impossible que différents protocoles d'acheminement de diffusion groupée IP soient déployés dans l'Internet.

3.1 Agrégation

En envoi individuel, différentes adresses de destination sont agrégées à une entrée dans le tableau d'acheminement, donnant une FEC et un LSP.

La granularité des flux de diffusion groupée est (*, G) pour une arborescence partagée et (S, G) pour une arborescence de source, S étant l'adresse de source et G l'adresse du groupe de diffusion groupée. L'agrégation des arborescences de diffusion groupée avec des adresses de 'destination' de diffusion groupée différentes sur un LSP est le sujet d'études en cours.

3.2 Arrosage et élagage

Pour établir une arborescence de diffusion groupée, certains protocoles d'acheminement de diffusion groupée IP (par exemple, DVMRP, PIM-DM) arrosent le réseau avec des données en diffusion groupée. Les branches peuvent ensuite être élaguées par les nœuds qui ne veulent pas recevoir les données du groupe de diffusion groupée spécifique. Ce processus est répété périodiquement.

Les protocoles d'acheminement de diffusion groupée d'arrosage et élagage ont des caractéristiques qui diffèrent significativement de celles des protocoles d'acheminement d'envoi individuel :

- a) Volatile. Du fait de la nature d'arrosage et élagage du protocole, des structures d'arborescence très volatiles sont générées. Les solutions pour transposer une arborescence dynamique de point en multipoint de couche 3 en LSP point à multipoint de couche 2 doivent être efficaces en termes de redondance de signalisation et de temps d'établissement de LSP. Le LSP volatile de couche 2 va consommer une grande quantité d'étiquettes tout au long du réseau, ce qui est un inconvénient lorsque l'espace d'étiquettes est limité.
- b) Piloté par le trafic. Le routeur crée seulement l'état pour un certain groupe lorsque les données arrivent pour ce groupe. Les routeurs décident aussi indépendamment de retirer l'état lorsque un temporisateur d'inactivité arrive à expiration.
 - Donc, les LSP ne peuvent pas être pré établis comme on le fait habituellement en envoi individuel. Pour minimiser le délai entre l'arrivée du trafic et l'établissement du LSP, une méthode d'établissement rapide de LSP est favorable.
 - Comme la création et la suppression d'un chemin de couche 3 à chaque nœud est déclenchée par le trafic, cela suggère que le LSP associé au chemin soit aussi établi et supprimé d'une manière pilotée par le trafic.
 - Si un LSR ne prend pas en charge la couche 3, la transmission de nature pilotée par le trafic exige que le LSR amont prenne l'initiative de créer un LSP (annonce d'étiquette vers l'amont à la demande ou non sollicitée vers l'aval).

3.3 Arborescences de source/partagées

Les protocoles d'acheminement de diffusion groupée IP créent soit des arborescences de source (S, G), c'est-à-dire, une arborescence par source (S) et par groupe de diffusion groupée (G), soit des arborescences partagées (*, G), c'est-à-dire, par groupe de diffusion groupée (Figure 1).

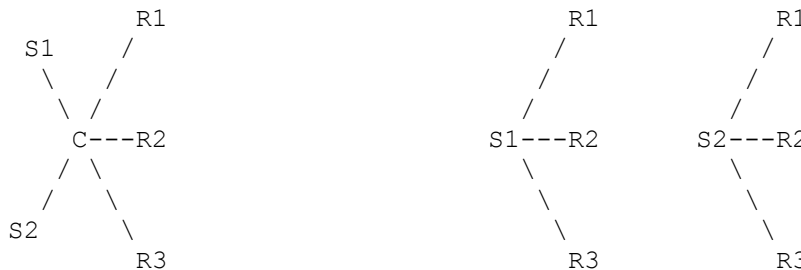


Figure 1 : arborescence partagée et arborescence de source

L'avantage d'utiliser les arborescences partagées, lorsque on applique la commutation d'étiquettes, est qu'elles consomment moins d'étiquettes que les arborescences de source (une étiquette par groupe contre une étiquette par source et par groupe).

Cependant, transposer une arborescence partagée de bout en bout sur la couche 2 implique d'établir des LSP de multipoint à multipoint (mp2mp). Le problème de la mise en œuvre de LSP mp2mp revient au problème de fusion discuté précédemment.

Noter qu'en pratique, les arborescences partagées ne sont souvent utilisées que pour découvrir les nouvelles sources du groupe et un passage à une arborescence de source est fait à un débit binaire très lent.

3.4 Coexistence d'arborescence de source et d'arborescence partagée

Certains protocoles prennent en charge les deux arborescences de source et partagées (par exemple, PIM-SM) et un routeur peut conserver les deux états (*, G) et (S, G) pour le même groupe G. Deux cas de coexistence d'état sont décrits ci-dessous. On suppose des topologies avec les envoyeurs S_i et les receveurs R_i . RP est le point de rendez-vous. N_i sont les LSR. Les numéros sont ceux des interfaces, "Reg" est l'interface d'enregistrement. Tous les messages IGMP et PIM Joindre/Élaguer sont montrés dans les figures. On indique aussi si le bit RPT est établi pour l'état (S, G).

- 1) La Figure 2 montre un passage d'une arborescence partagée à une arborescence de source. On suppose que le plus court chemin de R_1 à RP est via N_1 - N_2 - N_5 . N_1 , le routeur désigné du receveur R_1 (DRrecv), décide d'initier une arborescence de source pour la source S_1 . Après l'arrivée de données via l'arborescence de source dans N_2 , N_2 va envoyer un élagage à N_5 pour la source S_1 . La coexistence d'état survient dans le nœud lorsque commence le chevauchement de l'arborescence partagée et de source (N_2) et dans le nœud lorsque S_1 n'a plus besoin de transmettre sur l'arborescence partagée (N_5).

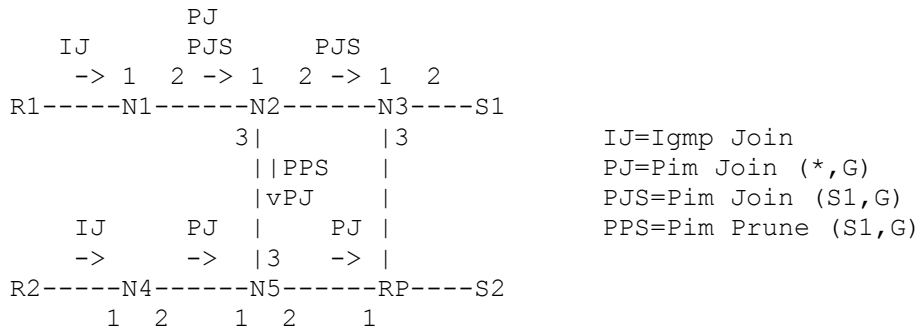


Figure 2

Les états d'acheminement en diffusion groupée créés dans le tableau d'acheminement de diffusion groupée (MRT, *Multicast Routing Table*) sont :

dans RP : (*,G):Reg->1 (c'est-à-dire, itf=Reg entrant ; itf=1 sortant)
 dans N1 : (*,G):2->1
 dans N2 : (*,G):3->1
 (S1,G):2->1
 dans N3 : (S1,G):2->Reg,1
 dans N4 : (*,G):2->1
 dans N5 : (*,G):2->1,3
 (S1,G)RPT-bit:2->1

2) La Figure 3 montre que même sans une commutation, la coexistence d'état peut se produire. Le trafic de diffusion groupée provenant d'un envoyeur va créer l'état (S, G) dans le routeur désigné de l'envoyeur (DRsend ; N3 dans la Figure 3 est le DRsend de S). Chaque nœud sur une arborescence partagée a l'état (*, G). Donc, un DRsend sur l'arborescence a les deux états (*, G) et (S, G). Si le DRsend est sur l'arborescence, il va aussi envoyer un élagage pour S vers le RP, créant l'état (S, G) dans tous les nœuds jusqu'au premier routeur qui a une branche (N1 et N2 dans la Figure 3).

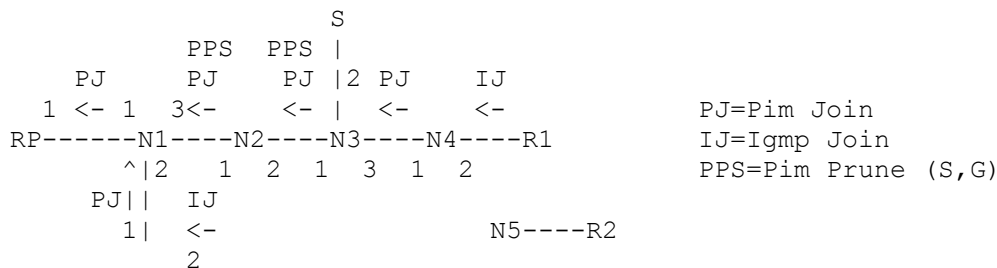


Figure 3

Les états d'acheminement de diffusion groupée créés dans le MRT sont :

dans RP : (*,G):Reg->1 (c'est-à-dire, itf=Reg entrant ; itf=1 sortant)
 dans N1 : (*,G):1->2,3
 (S,G)RPT-bit:1->2
 dans N2 : (*,G):1->2
 (S,G)RPT-bit:1->none
 dans N3 : (*,G):1->3
 (S,G):2->Reg,3
 dans N4 : (*,G):1->2
 dans N5 : (*,G):1->2

Dans les exemples, on peut observer que deux types de coexistence d'état surviennent :

1) (S, G) avec le bit RPT non établi (N2 dans la Figure 2, N3 dans la Figure 3). Les états (*, G) et (S, G) ont des interfaces entrantes différentes, mais des interfaces sortantes communes. Il est possible que le trafic de S arrive sur les deux interfaces (*, G) et (S, G). Dans la transmission normale de couche 3, l'entrée du bit SPT (S, G) interdit la transmission du trafic venant de S sur l'interface entrante (*, G). Le trafic de S peut seulement temporairement arriver sur les interfaces entrantes des deux entrées (*, G) et (S, G) (jusqu'à ce que N5 dans la Figure 2 et N1 dans la Figure 3 aient traité les messages d'élagage). Pour éviter la transmission temporaire de paquets dupliqués, la transmission de couche 3 peut être appliquée dans ce type de nœud. Si on ne se soucie pas de paquets temporairement dupliqués, la transmission de couche 2 peut être appliquée. Dans ce cas, les flux (*, G) et (S, G) doivent être fusionnés dans le LSP (*, G) sur leurs interfaces sortantes

communes.

- 2) (S, G) avec le bit RPT établi (N5 dans la Figure 2, N1 dans la Figure 3). Les états (*, G) et (S, G) ont la même interface entrante. Le trafic (S, G) doit être extrait du flux (*, G). Dans MPLS, cette coexistence d'état peut être traitée de plusieurs façons. On va décrire quatre approches de ce problème :
- Une première méthode pour traiter cette coexistence d'état est de terminer les LSP et transmettre tout le trafic de ce groupe à la L3. Cependant un retour à L3 peut être évité dans le cas où une entrée (S, G) sans interface sortante est ajoutée à la MRT (N2 dans la Figure 3). Cette entrée va seulement recevoir temporairement du trafic. Dans ce cas particulier, on peut ignorer l'état (S, G) et conserver le LSP existant (*, G), l'inconvénient étant le trafic dupliqué pour un très court instant.
 - Une seconde approche est d'allouer des étiquettes spécifiques de source sur les nœuds de l'arborescence partagée. Plusieurs étiquettes seront associées à une entrée (*, G), correspondant à une étiquette par source active. Comme les nœuds connaissent seulement les sources qui sont actives lorsque arrive le trafic provenant de ces sources, les LSP ne peuvent pas les pré établir et une méthode d'établissement rapide de LSP est favorable.
 - Une troisième façon est que seules les arborescences de source soient à commutation d'étiquettes et que le trafic sur les arborescences partagées soit toujours transmis à L3. Cela suppose que l'arborescence partagée ne soit utilisée que comme moyen pour les receveurs de découvrir qui sont les sources. En configurant un seuil de commutation de faible débit binaire, on peut s'assurer que les receveurs commutent très rapidement sur les arborescences de source.
 - Dans la quatrième approche, un LSR qui a l'état bit RPT (S, G) avec un oif non nul, annonce une étiquette pour (S, G) au LSR amont et cette annonce d'étiquette est alors propagée par chaque LSR amont vers le RP. De cette façon, un LSP dédié est créé pour le trafic (S, G) du RP au LSR avec l'état bit RPT (S, G). Dans le dernier LSR, le LSP (S, G) est fusionné avec le LSP (*, G) pour les interfaces sortantes appropriées. Cela assure que les paquets (S, G) qui voyagent sur l'arborescence partagée ne le font pas à travers un LSR qui a élagué S.

3.5 Arborecences partagées uni/bidirectionnelles

Les arborescences partagées bidirectionnelles (par exemple, CBT [RFC2201]) présentent l'inconvénient de créer beaucoup de points de fusion (M) dans les nœuds (N) de l'arborescence partagée. La Figure 4 montre ces points de fusion qui résultent des deux envoyeurs S1 et S2 sur une arborescence bidirectionnelle.

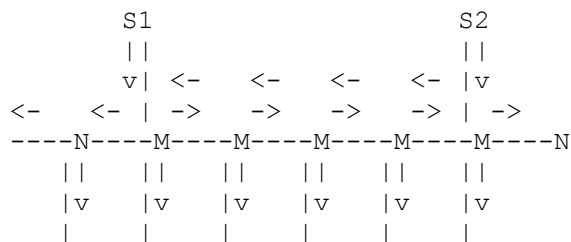


Figure 4 : Flux de trafic de diffusion groupée provenant de deux envoyeurs sur une arborescence bidirectionnelle

Dans la Figure 5, on décrit la même situation pour les arborescences partagées unidirectionnelles. Dans ce cas, les données des envoyeurs sont tunnelées vers le nœud racine R, donnant seulement un point de fusion, à savoir la racine de l'arborescence partagée elle-même.

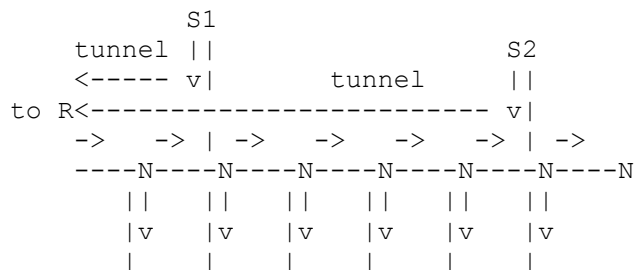


Figure 5 : Flux de trafic de diffusion groupée provenant de deux envoyeurs sur une arborescence unidirectionnelle

3.6 Données de diffusion groupée encapsulées

Les sources d'arborescences partagées unidirectionnelles et les sources d'arborescences partagées bidirectionnelles non membres encapsulent les données vers le nœud racine. Les données sont alors désencapsulées dans le nœud racine. L'encapsulation et la désencapsulation de données de diffusion groupée sont des processus de couche 3.

Donc, dans le cas d'encapsulation/désencapsulation, un chemin ne peut jamais être transposé en un LSP de bout en bout : le trafic ne peut pas être transmis sur la couche 2 sur l'interface Register du DRsend (encapsulation) ni traverser la racine (désencapsulation) à la couche 2.

Remarques :

- 1) Si le LSR prend en charge la transmission mixte de couche 2/3 (section 4), le trafic (S, G) dans DRsend peut quand même être transmis à la couche L2 sur toutes les interfaces sortantes autres que l'interface Register.
- 2) Le trafic encapsulé peut aussi bénéficier de MPLS en faisant la commutation d'étiquette sur les tunnels.
- 3) Si le nœud racine décide de se joindre à la source (pour éviter l'encapsulation/désencapsulation) un LSP de bout en bout (S, G) peut être construit.

3.7 Suppression des boucles

Les protocoles d'acheminement de diffusion groupée qui dépendent d'un protocole d'acheminement d'envoi individuel souffrent des mêmes boucles transitoires que les protocoles d'envoi individuel, cependant, l'effet de boucles sera bien pire dans le cas de la diffusion groupée. La raison en est que chaque fois qu'un paquet en diffusion groupée passe dans une boucle, des copies du paquet peuvent être émises à partir de la boucle si des branches existent dans la boucle.

La détection de boucle est actuellement un option configurable dans LDP et une décision sur le mécanisme de prévention de boucle est remise à plus tard.

3.8 Transposition des caractéristiques sur les protocoles existants

Les caractéristiques ci-dessus sont résumées dans le Tableau 1 pour donner une liste non exhaustive des protocoles d'acheminement de diffusion groupée IP existants : DVMRP [PUSA], MOSPF [RFC1584], CBT [RFC2201], PIM-DM [RFC3973], PIM-SM [RFC2362], SSM [RFC4607], SM [PERL].

	DVMRP	MOSPF	CBT	PIM-DM	PIM-SM	SSM	SM	
Agrégation	non	non	non	non	non	non	non	
Arrosage& élagage	oui	non	non	oui	non	non	option	
Type d'arborescence	source	source	partagé	source	les deux	source	partagé	
Coexistence d'état	non	non	non	non	oui	non	non	
Uni/bidirectionnelle	N/A	N/A	bi	N/A	uni	uni	bi	
Encapsulation	non	non	oui	non	oui	non	oui	Sans boucle non
	non	non	non	non	non	non		

Table 1 : Taxonomie des protocoles d'acheminement de diffusion groupée IP

Du Tableau 1 on peut déduire par exemple, que DVMRP va consommer beaucoup d'étiquettes lorsque l'arborescence L3 d'arrosage et élagage est transposée en une arborescence L2. De plus, comme DVMRP utilise des arborescences de source, il ne subit aucun problème de fusion lorsque on applique la commutation d'étiquettes. Le tableau peut être interprété de la même façon pour les autres protocoles.

4. Transmission L2/L3 mixte dans un seul nœud

Comme le trafic en envoi individuel a une interface entrante et une interface sortante, le trafic est soit transmis à la couche 2 SOIT à la couche 3 (Figure 6). Comme le trafic en diffusion groupée peut être transmis à plus d'une interface sortante on peut considérer le cas où le trafic pour certaines branches est transmis sur L2 et pour d'autres branches sur L3 (Figure 7).

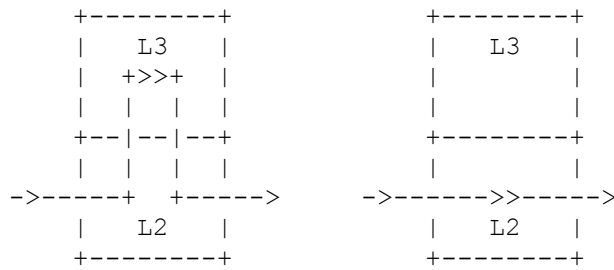


Figure 6 : Transmission en envoi individuel respectivement sur L3 ou L2

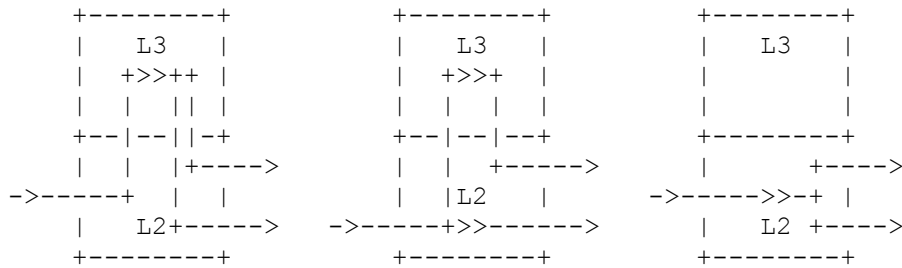


Figure 7 : Transmission en diffusion groupée sur respectivement L3, L2/L3 mixtes, ou L2

Les nœuds qui prennent en charge cette caractéristique de 'transmission mixte L2/L3' permettent le partage d'une arborescence de diffusion groupée entre des branches dans lesquelles certaines sont transmises à la couche L3 tandis que d'autres sont commutées à L2.

La transmission à la couche L3 doit veiller à ce que le trafic ne soit pas transmis à des branches qui ont déjà leur trafic sur la couche L2. Cela peut se réaliser, par exemple, en fournissant un bit supplémentaire dans le tableau d'acheminement de diffusion groupée.

Bien que la transmission mixte L2/L3 exige le traitement du trafic à la couche L3, la charge du moteur de transmission à la couche L3 est généralement moindre que dans un pur nœud L3.

La prise en charge de cette caractéristique de 'transmission mixte L2/L3' présente les avantages suivants :

- Supposons que le LSR A (Figure 8) soit un nœud bordure MPLS pour la branche vers le LSR B et un nœud cœur MPLS pour la branche vers le LSR C. La transmission mixte L2/L3 permet que la branche vers C ne soit pas dérangée par un retour à L3 dans le LSR A.

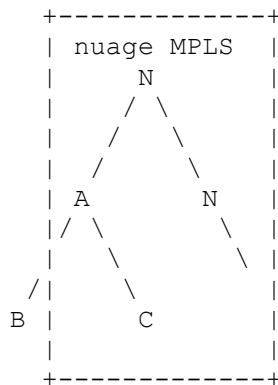


Figure 8 : Transmission mixte L2/L3 dans le nœud A

- Permet l'utilisation du déclenchement piloté par le trafic avec le mode de distribution d'étiquette vers l'aval non sollicité ou vers l'amont à la demande, comme expliqué au paragraphe 5.3.1.

5. Taxonomie des déclencheurs de LSP de diffusion groupée IP

La création d'un LSP pour les flux de diffusion groupée peut être déclenchée par différents événements, qui peuvent être transposés sur les catégories bien connues de 'pilote par la demande', 'pilote par la topologie' et 'pilote par le trafic'.

- Piloté par la demande : intercepte l'envoi ou la réception des messages de contrôle (par exemple, les messages d'acheminement de diffusion groupée, les messages de réservation de ressources).
- Piloté par la topologie : transpose l'arborescence L3, qui est disponible dans le Tableau d'acheminement de diffusion groupée, en une arborescence L2. La transposition est faite même si il n'y a pas de trafic.
- Piloté par le trafic : l'arborescence L3 est transposée en arborescence L2 lorsque les données arrivent sur l'arborescence.

5.1 Piloté par la demande

5.1.1 Généralités

L'établissement des LSP peut être déclenché par l'interception de messages de contrôle sortants (ce qui exige que l'étiquette soit demandée par le LSR aval) ou entrants (ce qui exige que l'étiquette soit demandée par le LSR amont). La Figure 9 illustre ces deux cas.

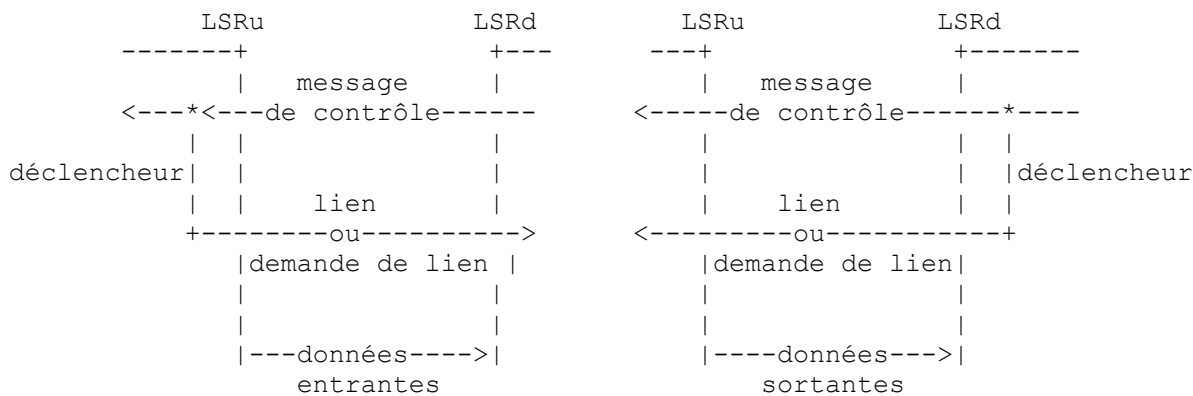


Figure 9 : Déclenchement piloté par la demande (interception des messages de contrôle respectivement entrants et sortants)

Le LSR aval (LSRd) envoie un message de contrôle au LSR amont (LSRu). Dans le cas où les messages de contrôle entrants sont interceptés et où le module MPLS dans le LSRu décide d'établir un LSP, il va envoyer une demande de lien LDP (mode non sollicité vers l'amont) ou une demande de lien LDP (mode à la demande vers l'amont) au LSRd.

Actuellement, pour la diffusion groupée, on peut identifier des types importants de messages de contrôle : les messages d'acheminement de diffusion groupée et les messages de réservation de ressources.

5.1.2 Messages d'acheminement de diffusion groupée

En principe, ce mécanisme ne peut être utilisé que par les protocoles d'acheminement de diffusion groupée IP qui utilisent la signalisation explicite : par exemple, les messages Join dans PIM-SM ou CBT. On remarque que DVMRP et PIM-DM peuvent être adaptés pour prendre en charge ce type de déclenchement [FARI], cependant, cela modifie le protocole d'acheminement de diffusion groupée IP lui-même !

Les messages d'acheminement de diffusion groupée IP peuvent créer à la fois des états fixes (par exemple, CBT Join + CBT Join-Ack) et des états conditionnels (par exemple, les Join PIM-SM sont envoyés périodiquement). Ces derniers génèrent plus de trafic de contrôle pour la maintenance de l'arborescence et donc exigent plus de traitement dans le module MPLS.

Les déclencheurs fondés sur les messages du protocole d'acheminement de diffusion groupée présentent l'inconvénient que les calculs d'acheminement effectués par le robot d'acheminement de diffusion groupée pour déterminer le tableau d'acheminement de diffusion groupée (MRT) sont répétés par le module MPLS. Le premier détermine l'arborescence qui va être utilisée à la couche L3, le second calcule une arborescence identique qui sera utilisée par la couche L2. Comme la même tâche est effectuée deux fois, il est préférable de créer le LSP de diffusion groupée sur la base des informations extraites du MRT lui-même (voir les paragraphes 5.2 et 5.3). Les calculs d'acheminement deviennent plus complexes pour les protocoles

qui prennent en charge la commutation d'une arborescence (*, G) en une arborescence (S, G) parce que plus de messages doivent être interprétés.

Lorsque un hôte a une connexion point à point avec le premier routeur, on peut créer des LSP jusqu'à l'utilisateur final en interceptant non seulement les messages d'acheminement de diffusion groupée mais aussi les messages Join/Prune IGMP [RFC2236].

5.1.3 Messages de réservation de ressources

Comme dans le cas de l'envoi individuel, le message Resv de RSVP peut être utilisé comme déclencheur pour établir les LSP. Une source d'un groupe de diffusion groupée va envoyer un message Path RSVP descendant l'arborescence, et les receveurs peuvent alors répondre par un message Resv RSVP. RSVP s'adapte aussi bien pour la diffusion groupée que pour l'envoi individuel parce que :

- les messages RSVP Resv peuvent fusionner,
- les messages RSVP Resv ne sont envoyés qu'à la première branche qui a fait la demande de réservation.

5.2 Piloté par la topologie

Le tableau d'acheminement de diffusion groupée (MRT, *Multicast Routing Table*) est entretenu par le robot du protocole d'acheminement de diffusion groupée IP. Le module MPLS transpose ces informations de topologie d'arborescence de couche L3 en LSP de point à multipoint de couche L2. Le module MPLS peut interroger le MRT pour extraire les topologies d'arborescence. Autrement, le robot de diffusion groupée peut être modifié pour notifier directement au module MPLS tout changement du MRT. L'inconvénient de cette méthode est que les étiquettes sont consommées même lorsque il n'y a pas de trafic.

5.3 Piloté par le trafic

5.3.1 Généralités

Une méthode de déclenchement pilotée par le trafic ne va construire de LSP que pour les arborescences qui portent du trafic. Elle consomme moins d'étiquettes que la méthode pilotée par la topologie, car les étiquettes ne sont allouées que lorsque il y a du trafic sur l'arborescence de diffusion groupée.

Si la capacité de transmission mixte L2/L3 (voir la section 4) n'est pas prise en charge, le déclenchement piloté par le trafic exige un mode de distribution d'étiquettes dans lequel l'étiquette est demandée par le LSRu (mode vers l'aval à la demande ou mode non sollicité vers l'amont). Dans la Figure 10, on suppose qu'un LSP pour un certain groupe existe vers LSRd1 et qu'un autre LSRd2 veut se joindre à l'arborescence. Afin que le LSRd2 initie un déclenchement, il doit déjà recevoir le trafic provenant de l'arborescence. Cela peut être aussi bien à L2 qu'à L3. Le premier cas est le problème de la poule et de l'œuf. Le second cas exige une capacité de transmission mixte L2/L3 dans le LSRu pour ajouter la branche L3.

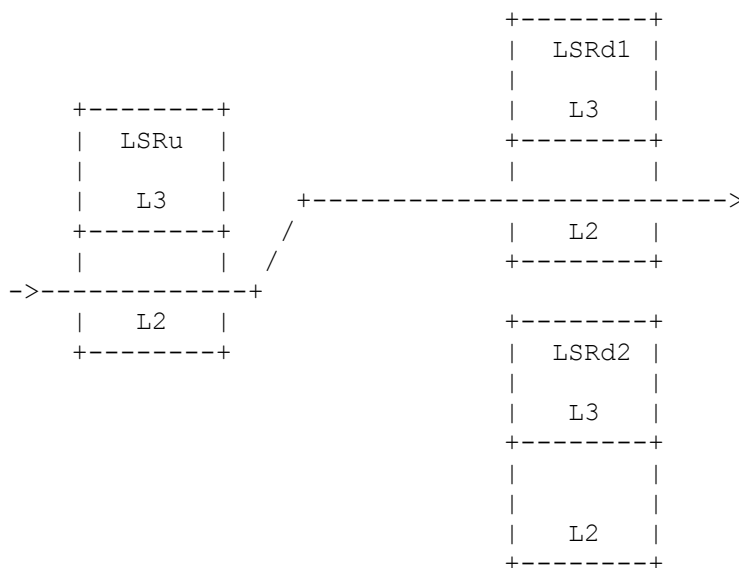


Figure 10 : Le LSRu doit demander l'étiquette

5.3.2 Exemple de mise en œuvre

Pour illustrer qu'en choisissant un déclencheur approprié, on peut conclure que la diffusion groupée MPLS est indépendante du protocole d'acheminement de diffusion groupée déployé, on donne les exemples de mise en œuvre suivants.

Les mises en œuvre actuelles sur des plates-formes Unix des protocoles d'acheminement de diffusion groupée IP (DVMRP, PIM) ont une antémémoire de transmission de diffusion groupée (MFC, *Multicast Forwarding Cache*). La MFC est une copie en antémémoire du tableau d'acheminement de diffusion groupée. La MFC demande une entrée pour un certain groupe de diffusion groupée lorsque il rencontre un 'vide d'antémémoire' pour un paquet de diffusion groupée entrant. Les informations d'acheminement manquantes sont fournies par le robot de diffusion groupée. Si ultérieurement quelque chose change dans l'acheminement (des interfaces sortantes ajoutées ou retirées) le robot de diffusion groupée va mettre à jour la MFC.

La MFC est mise en œuvre comme un composant commun (une partie du noyau) qui rend ce déclenchement très intéressant parce que il peut être utilisé de façon transparente pour tout protocole d'acheminement de diffusion groupée IP.

Les entrées dans la MFC sont retirées lorsque il n'y a plus de trafic reçu pour cette entrée pendant un certain temps. Lorsque la commutation d'étiquettes est appliquée à une certaine entrée de la MFC, la couche L3 ne verra plus arriver aucun paquet. Pour garder le comportement normal de la MFC, les compteurs de L3 de la MFC ont besoin d'être mis à jour par des mesures de L2.

5.4 Combinaisons des déclencheurs et des modes de distribution d'étiquettes

Le Tableau 2 montre les combinaisons valides des modes de distribution d'étiquettes et de types de déclencheurs qui ont été discutées dans les paragraphes précédents. Le (X) signifie que la combinaison est valide lorsque la caractéristique de transmission mixte L2/L3 est prise en charge dans le LSR.

	Étiquette demandée par			
	LSRu		LSRd	
	amont non sollicitée	aval à la demande	aval non sollicitée	amont à la demande
Piloté demande (msg entrant)	X	X		
Piloté demande (msg sortant)	X	X		
Piloté topologie	X	X	X	X
Piloté trafic	X	X	(X)	(X)

Tableau 2 : Combinaisons valides de déclencheurs et de mode de distribution d'étiquettes

6. Portage

Dans la Figure 9 (cas sortant) on peut observer qu'au lieu d'envoyer deux messages séparés l'annonce d'étiquette peut être portée sur les messages de contrôle existants. Pour la diffusion groupée, deux candidats au portage existent :

- Messages d'acheminement de diffusion groupée : des protocoles comme PIM-SM et CBT ont des messages Join explicites qui pourraient porter les transpositions d'étiquettes. Cette approche est décrite dans [FARI]. Lorsque différents protocoles d'acheminement de diffusion groupée sont déployés, une extension à chacun de ces protocoles a été définie.
- Messages Resv RSVP : un objet d'extension de transposition d'étiquette pour RSVP, aussi applicable à la diffusion groupée, est proposé dans la [RFC3209].

On va maintenant décrire les avantages et inconvénients du portage sur les messages d'acheminement de diffusion groupée.

Le portage a les avantages suivants :

- Si l'annonce d'étiquette est portée sur des messages d'acheminement de diffusion groupée, la distribution des chemins et la distribution des étiquettes sont alors étroitement synchronisées. Cela élimine les cas difficiles comme "que faire avec une étiquette si on n'a pas (encore) une entrée de tableau d'acheminement pour la rattacher ?". Cela minimise aussi l'intervalle entre l'établissement du chemin de diffusion groupée et la transposition d'une étiquette en chemin.
- Le nombre de messages de contrôle nécessaire pour prendre en charge l'annonce d'étiquette au delà de ceux nécessaires pour prendre en charge l'acheminement de diffusion groupée lui-même est de zéro.

Les inconvénients du portage suivants peuvent être identifiés :

- Dans les protocoles en mode dense, il n'y a pas de message sur lequel l'annonce d'étiquette puisse être portée. [FARI] propose d'ajouter des messages périodiques aux protocoles en mode dense pour les besoins d'annonce d'étiquette, ce qui a un impact lourd sur le protocole d'acheminement de diffusion groupée et il élimine le message en conservant le bénéfice du

portage.

- b) La seconde solution au problème de la coexistence d'état (paragraphe 3.4) ne peut pas être appliquée dans la combinaison avec le portage.
- c) Le portage exige l'extension du protocole d'acheminement de diffusion groupée, et donc devient moins attractif si l'annonce d'étiquette a besoin d'être prise en charge par plusieurs protocoles d'acheminement de diffusion groupée. En particulier lorsque non seulement l'annonce d'étiquette mais aussi les deux autres fonctions de LDP (découverte et adjacence) sont portées.
- d) Le portage suppose le mode de distribution d'étiquette non sollicité vers l'aval, ce qui exclut un certain nombre de méthodes de déclenchement (voir le Tableau 2).
- e) LDP fonctionne normalement par dessus TCP, assurant une communication fiable entre les nœuds homologues. L'annonce d'étiquette portée remplace souvent la communication fiable par des annonces périodiques d'étiquette d'état conditionnel. À cause de cette annonce périodique d'étiquette, le trafic de contrôle va augmenter (en nombre d'octets).
- f) Si un mécanisme de notification de VCID [RFC3038] est nécessaire, la notification (dans la bande) peut être faite normalement par l'envoi du lien LDP à travers le VC nouvellement établi. De cette façon, un seul message est nécessaire. Cette méthode ne peut pas être combinée avec le portage parce que le message d'acheminement est envoyé avant que le VC puisse être établi. Un message de prise de contact supplémentaire est donc nécessaire, diminuant le bénéfice du portage.

Que le portage ait un sens ou non dépend fortement du protocole d'acheminement de diffusion groupée déployé, de si LDP est déjà utilisé pour l'envoi individuel, et du mécanisme de déclenchement utilisé, ... le portage est juste un composant possible de la solution de diffusion groupée MPLS.

7. Acheminement explicite

L'acheminement explicite pour l'envoi individuel se réfère à la substitution des LSP au tableau d'acheminement en envoi individuel.

Une première manière d'interpréter "acheminement explicite de diffusion groupée" est de remplacer l'arborescence établie par le protocole d'acheminement de diffusion groupée par une autre arborescence de LSP (par exemple, une arborescence de Steiner calculée par un outil hors ligne). Dans cette interprétation, le protocole d'acheminement de diffusion groupée actuel de 'plus court chemin' devient obsolète et peut être remplacé par des messages d'annonce d'étiquette qui suivent un chemin explicite (par exemple, une branche de l'arborescence de Steiner).

Une seconde manière d'interpréter "acheminement explicite de diffusion groupée" est que les protocoles d'acheminement de diffusion groupée connus fonctionnent, mais que les messages générés par ces protocoles utilisent des chemins explicites d'envoi individuel (au lieu des plus courts chemins de l'IGP) pour construire les arborescences.

8. QoS/CoS

8.1 DiffServ

L'approche des services différenciés peut être aussi appliquée à la diffusion groupée. Elle introduit des granularités de flux plus fines (le codet DiffServ (DSCP) comme différenciateur supplémentaire). Un expéditeur peut construire une ou plusieurs arborescences avec des DSCP différents.

Ces arborescences (S, G, DSCP) ou (*, G, DSCP) peuvent être transposées très facilement en LSP lorsque on utilise le déclenchement piloté par le trafic. Dans ce cas, on peut créer des LSP avec des attributs différents pour les divers DSCP. Noter cependant que ces LSP utilisent encore le même chemin tant que le mécanisme de construction d'arborescence lui-même ne prend pas le DSCP en entrée.

8.2 IntServ et RSVP

RSVP peut être utilisé pour établir des arborescences de diffusion groupée avec qualité de service. Une question importante de la diffusion groupée est le problème de comment transposer le paradigme des 'receveurs hétérogènes' dans la couche L2 (remarquer qu'il n'est pas résolu non plus dans IP). Ce sujet est abordé dans la [RFC2382]. Des approches pragmatiques sont le 'modèle d'hétérogénéité limitée' qui permet un service au mieux et une seule qualité de service de remplacement (par exemple, une QS proposée par l'expéditeur dans un message RSVP Path) et le 'modèle homogène' qui ne permet qu'une seule QS.

La première approche va construire une arborescence complète pour chaque classe de service. L'expéditeur doit envoyer son trafic deux fois à travers le réseau (par exemple, une arborescence au mieux et une avec QS). Les deux arborescences peuvent être à commutation d'étiquette.

La seconde approche construit une arborescence et les utilisateurs au mieux sont connectés à l'arborescence à QS. Si les branches créées pour les utilisateurs au mieux ne sont pas à commutation d'étiquettes, (donc portées par un LSP par défaut bon par bond) le trafic de diffusion groupée à QS doit être fusionné sur ces LSP par défaut. Cette fonction peut être fournie par la caractéristique de 'transmission mixte L2/L3' décrite à la Section 4. Si ceci n'est pas disponible, la fusion est nécessaire pour éviter un retour à la L3 dans le LSP à QS.

La transposition des catégories de service IntServ sur la couche L2 en catégories de service ATM est étudiée dans la [RFC2381].

9. Réseaux multi-accès

La diffusion groupée MPLS sur des réseaux multi accès pose un problème particulier. Un LSR qui veut se joindre à un groupe doit toujours être prêt à accepter l'étiquette qui est déjà allouée au LSP du groupe (pour un autre LSR à l'aval sur la liaison). Ceci peut être réalisé de trois façons :

- 1) Chaque LSR sur la liaison multi-accès mémorise toutes les étiquettes annoncées sur la liaison, même si elle n'a pas reçu un message de jonction pour le groupe associé. Si un LSR est ajouté à la liaison multi-accès, il doit récupérer cette information d'un autre LSR sur la liaison ou dans le cas d'une annonce d'étiquette d'état conditionnel il peut attendre un certain temps avant qu'il puisse allouer lui-même des étiquettes. Si les LSR allouent une étiquette 'au même moment', le LSR qui a la plus forte adresse IP pourrait la garder, tandis que les autres LSR retirent leur étiquette.
- 2) Chaque LSR obtient sa propre gamme d'étiquettes d'où il les alloue. Un mécanisme de partition d'étiquettes est décrit dans [FARI]. Si un LSR est ajouté à la liaison multi-accès, les gammes d'étiquettes doivent être négociées à nouveau et éventuellement des LSP existants sont supprimés et sont reconstruits avec d'autres étiquettes.
- 3) Sur chaque liaison multi-accès, un LSR peut être choisi pour être responsable de l'allocation d'étiquettes. Lorsque un LSR a besoin d'une étiquette, il peut la demander à ce LSR d'allocation d'étiquettes.

À la différence du cas d'envoi individuel, un flux de diffusion groupée peut avoir plus d'un LSR aval qui doivent tous utiliser la même étiquette. Deux solutions peuvent être envisagées pour les annonces d'étiquettes :

- 1) [FARI] propose d'envoyer en diffusion groupée les annonces d'étiquettes à tous les LSR sur la liaison partagée. Comme la diffusion groupée n'est pas fiable, cela exige des annonces d'étiquettes périodiques, donnant des duplications d'annonces d'étiquettes dans le temps.
- 2) Une autre approche est qu'un LSR envoie en individuel ses annonces d'étiquettes d'une façon fiable (TCP) à tous les autres LSR (ou tous les intéressés) sur la liaison partagée. Dans cette approche le caractère d'état stable de LDP peut être conservé mais l'annonce d'étiquette est dupliquée dans l'espace.

Comme les LSP ne sont intéressants que si ils ont une longue durée de vie et comme le nombre de LSR sur une liaison partagée est limité, la seconde approche semble avantageuse.

Un autre problème de la diffusion groupée dans les réseaux multi-accès est la question d'utiliser l'allocation d'étiquettes vers l'amont ou vers l'aval. Pour le trafic de diffusion groupée, l'allocation d'étiquettes vers l'amont est plus simple car il ne peut y avoir qu'un seul nœud vers l'amont par liaison qui appartient à une arborescence de diffusion groupée. Ce nœud (amont) peut allouer une étiquette unique pour la FEC. Avec l'allocation vers l'aval, il peut y avoir plusieurs nœuds vers l'aval pour une certaine arborescence sur une liaison multi-accès ; chaque nœud peut proposer une allocation d'étiquette différente pour une FEC, ce qui exigerait un processus de résolution afin d'arriver à une seule étiquette par FEC de diffusion groupée sur la liaison.

Une fois qu'une étiquette a été allouée, il est possible que l'allocataire de l'étiquette quitte l'arborescence. Avec l'allocation d'étiquette vers l'aval, cela pourrait arriver lorsque l'allocataire de l'étiquette quitte le groupe. Avec l'allocation vers l'amont cela pourrait arriver lorsque le LSR amont change à cause d'un changement de la topologie d'envoi individuel.

10. Autres questions

10.1 Champ TTL

Le champ TTL dans l'en-tête IP est normalement utilisé pour la détection de boucles. Dans la diffusion groupée IP, il est aussi utilisé pour limiter la portée des paquets de diffusion groupée en réglant une valeur de TTL appropriée.

Donc, dans les LSR qui ne prennent pas en charge une fonction de décrémentation du TTL (par exemple, un LSR ATM) la fonction de restriction de portée est affectée. Supposons qu'on calcule à l'avance le nombre de bonds que traverse un LSP. dans un LSP en envoi individuel, la valeur du TTL pourrait alors être décrémentée au nœud d'entrée ou de sortie. Pour la diffusion groupée, toutes les branches de l'arborescence peuvent avoir des longueurs différentes de sorte que le TTL ne peut être décrémenté qu'au nœud de sortie, gaspillant potentiellement la bande passante si le TTL se retrouve à zéro ou négatif.

10.2 Contrôle de distribution d'étiquettes indépendant ou ordonné

La terminologie actuelle de distribution d'étiquettes n'est définie que pour l'envoi individuel. Les paragraphes qui suivent explorent ce que cette terminologie peut signifier dans un contexte de diffusion groupée.

En contrôle indépendant ([RFC3036]) chaque LSR peut prendre l'initiative de faire une transposition d'étiquette. Dans le contrôle ordonné ([RFC3036]) un LSR ne transpose une étiquette que lorsque il a déjà reçu une étiquette de son prochain bond.

Toutes les méthodes de déclenchement précédemment décrites (section 5) combinent le contrôle indépendant. Noter que si l'approche pilotée par la demande est utilisée avec le contrôle indépendant, la distribution d'étiquettes se comporte quand même comme dans le contrôle ordonné : le flux des messages de contrôle s'écoule du nœud de sortie amont, imposant la même séquence aux annonces d'étiquettes.

Le contrôle ordonné n'est pas applicable pour un déclenchement piloté par le trafic au cas où le nœud ne prendrait pas en charge la transmission mixte L2/L3. Conformément au Tableau 2, ce cas implique que les étiquettes soient demandées par le LSR amont. Supposons dans la Figure 11 qu'il existe un LSP de S à R1 et qu'une nouvelle branche doive être ajoutée à R2. B va seulement accepter une étiquette sur la liaison A-B si une étiquette est déjà allouée sur la liaison B-C. Cependant, pour établir une étiquette sur la liaison B-C, B doit déjà recevoir du trafic sur la liaison A-B.

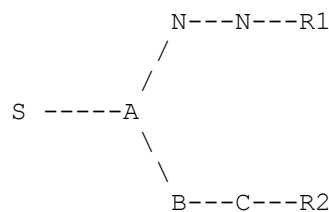


Figure 11.

10.3 Mode de rétention d'étiquette prudent ou libéral

Dans le mode prudent ([RFC3036]) seules les étiquettes qui sont utilisées pour transmettre des données (si le prochain bond pour la FEC est le LSR qui a annoncé l'étiquette) sont allouées et entretenues. Dans le mode libéral, les étiquettes sont annoncées et entretenues vers tous les voisins. Le mode libéral n'a pas de sens en diffusion groupée. Deux raisons peuvent être avancées pour cela :

- 1) Tous les LSR ont un chemin pour chaque FEC en envoi individuel. Ce n'est pas vrai pour les FEC de diffusion groupée.
- 2) Pour la diffusion groupée, un LSR sait toujours à quel voisin envoyer la demande d'étiquette ou les messages de transposition d'étiquette. Dans, par exemple, le mode d'envoi individuel non sollicité vers l'amont (voir ci-dessous) le LSR ne sait pas où envoyer les transpositions d'étiquettes et a donc à envoyer la transposition à tous ses voisins. Dans ce cas, prendre en charge le mode libéral ne génère pas de message supplémentaire (il exige quand même des informations d'état et d'espace d'étiquettes supplémentaires) et donc le seuil pour la prise en charge du mode libéral pourrait être établi plus bas.

Le Tableau 3 montre les cas où le LSR sait où envoyer ses demandes d'étiquettes.

	étiquette demandée par	
	LSRu	LSRd
envoi individuel	oui	non
diffusion groupée	oui	oui

Tableau 3 : un LSR sait-il où envoyer ses demandes d'étiquettes ?

Pour un flux en envoi individuel, un LSR peut déterminer le LSR de prochain bond, qui est celui à qui envoyer la demande en cas de mode non sollicité vers l'amont ou à la demande vers l'aval. Le LSR n'est cependant pas capable de trouver le bond précédent. Le bond précédent n'est pas nécessairement le prochain bond vers la source, parce que le chemin de A à B n'est pas nécessairement le même que le chemin de B à A. Une telle situation peut se produire par suite de mesures de liaison asymétriques ou lorsque plusieurs chemins de coût égal existent [PAXS].

Dans le cas de la diffusion groupée, un LSR connaît à la fois le ou les prochains bonds et le bond précédent. Parce que les arborescences de diffusion groupée sont construites en utilisant la méthode du plus court chemin inverse, le bond précédent est toujours le prochain bond vers la source ou vers la racine de l'arborescence.

10.4 Allocation d'étiquettes vers l'amont ou vers l'aval

L'étiquette peut être allouée soit par le LSR aval (aval à la demande, aval non sollicité) soit par le LSR amont (amont à la demande, amont non sollicité, implicite). Les avantages de l'allocation d'étiquettes vers l'aval sont :

- C'est le même mode que pour LDP en envoi individuel, éliminant donc le besoin de développer les procédures de distribution d'étiquettes vers l'amont.
- La même étiquette peut être conservée lorsque le LSR amont change à cause d'un changement de chemin, ce qui est un avantage sur les réseaux multi-accès (section 9).
- Compatible avec le portage (en particulier le mode de distribution vers l'aval).

Les avantages de l'allocation d'étiquettes vers l'amont sont :

- Une allocation d'étiquettes plus facile dans les réseaux multi-accès (section 9).
- La même étiquette peut être conservée lorsque le LSR amont (qui aurait été l'allocataire d'étiquettes dans le mode vers l'aval dans un réseau multi-accès) quitte le groupe (section 9).
- Le mode de distribution vers l'amont et implicite permet un établissement de LSP plus rapide lorsque le LSP est déclenché par le trafic.

La question de savoir si il faut utiliser la distribution d'étiquettes vers l'amont ou vers l'aval sort du domaine d'application du présent cadre. La complexité relative entre les extensions de protocole nécessaires et les mécanismes de résolution nécessaires, ainsi que la complexité opérationnelle relative vont influencer la voie à suivre.

10.5 Distribution d'étiquettes explicite ou implicite

À côté des modes de distribution explicites (qui utilisent un protocole de signalisation) [ACHA] propose une méthode de distribution implicite d'étiquettes en utilisant des étiquettes inconnues. Cette méthode a tous les avantages de la méthode d'allocation d'étiquettes vers l'amont et est probablement la méthode la plus rapide d'annonce d'étiquettes pour le trafic déclenché par des LSP.

La distribution implicite d'étiquettes n'est pas applicable si le lien entre la FEC et l'étiquette n'a pas été annoncée avant l'arrivée du trafic, par exemple, l'acheminement explicite (c'est-à-dire, si toutes les informations nécessaires pour identifier la FEC ne sont pas présentes dans le paquet).

La distribution explicite permet le pré établissement (avant l'arrivée des données) des LSP avec des déclencheurs pilotés par la topologie ou par la demande.

11. Considérations pour la sécurité

En général, l'utilisation de la diffusion groupée dans un environnement MPLS ne pose pas de problèmes de sécurité supplémentaires au delà de ceux qui existent déjà dans les protocoles de diffusion groupée et MPLS en tant que tels.

Les protocoles décrits dans ce document ne conviennent cependant pas pour franchir des frontières administratives.

Lorsque l'arborescence de diffusion groupée est déterminée par un protocole d'acheminement de diffusion groupée existant (c'est l'hypothèse qui est faite dans le présent document, sauf pour la section Acheminement explicite) il est clair qu'aucun problème de sécurité supplémentaire n'est introduit par rapport à la forme de l'arborescence (par exemple, jonction non autorisée, écoute, création de trou noir, injection de trafic, ...). Ces problèmes de sécurité devraient avoir été traités dans les spécifications des protocoles d'acheminement de diffusion groupée.

Dans le contexte MPLS, il est possible que les messages de contrôle déclenchent des allocations de ressources de couche 2 (par exemple, des LSP). Si des fautes de sécurité étaient encore présentes dans les protocoles existants (qui pourraient n'être pas trop dommageables dans leur contexte d'origine) l'envoi abusif de tels messages de contrôle peut donner de plus sévères attaques de déni de service.

Dans le cas d'une arborescence "à acheminement explicite" qui est calculée centralement, une authentification suffisante doit être faite sur les messages de contrôle qui établissent l'arborescence dans les nœuds du réseau.

12. Remerciements

Les auteurs tiennent à remercier Eric Rosen, Piet Van Mieghem, Philip Dumortier, Hans De Neve, Jan Vanhoutte, Alex Mondrus et Gérard Gastaud pour les fructueuses discussions et/ou leur revision attentive du présent document.

Références

- [ACHA] A. Acharya, R. Dighe and F. Ansari, "IP Switching Over Fast ATM Cell Transport (IPSOFACTO) : Switching Multicast flows", IEEE Globecom '97.
- [FARI] D. Farinacci, Y. Rekhter, E. Rosen and T. Qian, "Using PIM to Distribute MPLS Labels for Multicast Routes", Non publiée.
- [PERL] R. Perlman, C-Y. Lee, A. Ballardie, J. Crowcroft, Z. Wang, T. Maufer, "Simple Multicast", Non publiée.
- [PUSA] T. Pusateri, "Distance Vector Multicast Routing Protocol", Non publiée.
- [PAXS] V. Paxson, "End-to-End Routing Behavior in the Internet", IEEE/ACM Transactions on Networking 5(5), pp. 601-615.
- [RFC1584] J. Moy, "Extensions de diffusion groupée à OSPF", mars 1994. (*Historique*)
- [RFC2201] A. Ballardie, "Architecture d'acheminement de diffusion groupée d'arborescences fondées sur le cœur (CBT)", septembre 1997. (*Historique*)
- [RFC2236] W. Fenner, "Protocole de gestion de groupe Internet, version 2", novembre 1997. (*Remplacée par RFC3376*)
- [RFC2362] D. Estrin et autres, "Mode éparé de diffusion groupée indépendante du protocole (PIM-SM) : Spécification du protocole", juin 1998. (*Obsolète, voir RFC4601, RFC5059*)
- [RFC2381] M. Garrett, M. Borden, "[Interopération du service à charge contrôlée](#) et du service garanti avec ATM", août 1998.
- [RFC2382] E. Crawley et autres, "Cadre des services intégrés et de RSVP sur ATM", août 1998. (*Information*)
- [RFC3209] D. Awduche, et autres, "[RSVP-TE : Extensions à RSVP pour les tunnels LSP](#)", décembre 2001. (*Mise à jour par RFC3936, RFC4420, RFC4874, RFC5151, RFC5420, RFC6790*)
- [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.*)
- [RFC3035] B. Davie et autres, "[Utilisation de MPLS](#) dans la commutation de circuit virtuel LDP et ATM", janvier 2001. (*P.S.*)

- [RFC3036] L. Andersson et autres, "Spécification de LDP", janvier 2001. (*Rendue obsolète par la RFC5036*)
- [RFC3038] K. Nagami et autres, "[Notification de VCID sur liaison ATM](#) pour LDP", janvier 2001. (*P.S.*)
- [RFC3973] A. Adams et autres, "Diffusion groupée indépendante du protocole - Mode dense (PIM-DM) : Spécification du protocole (révisée)", janvier 2005. (*Expérimentale*)
- [RFC4607] H. Holbrook, B. Cain, "[Diffusion groupée spécifique de source pour IP](#)", août 2006. (*P.S.*)

Adresse des auteurs

Dirk Ooms
Alcatel Corporate Research Center
Fr. Wellesplein 1, 2018 Antwerp,
Belgique
téléphone : 32 3 2404732
Fax : 32 3 2409879
mél : Dirk.Ooms@alcatel.be

Bernard Sales
Alcatel Corporate Research Center
Fr. Wellesplein 1, 2018 Antwerp,
Belgique
téléphone : 32 3 2409574
mél : Bernard.Sales@alcatel.be

Wim Livens
Colt Telecom
Zweefvliegtuigstraat 10, 1130 Brussels,
Belgique
téléphone : 32 2 7901705
Fax : 32 2 7901711
mél : WLivens@colt-telecom.be

Arup Acharya
IBM TJ Watson Research Center
30 Saw Mill River Road, Hawthorne
NY 10532
téléphone : 1 914 784 7481
mél : arup@us.ibm.com

Frederic Griffoul
Ulticom, Inc.
Les Algorithmes,
BP29
06901 Sophia-Antipolis, FRANCE
mél : griffoul@ulticom.com

Furquan Ansari
Bell Labs, Lucent Tech.
101 Crawfords Corner Rd., Holmdel, NJ 07733
téléphone : 1 732 949 5249
Fax : 1 732 949 4556
mél : furquan@dnrc.bell-labs.com

Déclaration complète de droits de reproduction

Copyright (C) The Internet Society (2002). 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 droits de reproduction ci-dessus et le présent 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 ou ses successeurs ou ayant droits.

Le présent document et les informations contenues sont fournis 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 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.