Internet Engineering Task Force (IETF)
Request pour Comments : 8029
RFC rendues obsolètes : 4379, 6424, 6829, 7537
RFC mise à jour : 1122
Catégorie : Sur la voie de la normalisation
ISSN : 2070-1721
Traduction Claude Brière de L'Isle

K. Kompella, Juniper Networks, Inc.
G. Swallow, Cisco
C. Pignataro (éditeur), Cisco
N. Kumar, Cisco
S. Aldrin, Google
M. Chen, Huawei
mars 2017

Détection des défaillances de plan des données en commutation d'étiquettes multi protocole (MPLS)

Résumé
Le présent document décrit un mécanisme simple et efficace pour détecter les défaillances de plan des données dans les chemins à commutation d'étiquette (LSP, *Label Switched Path*) de commutation d'étiquette multi protocoles (MPLS, *Multi-Protocol Label Switching*). Il définit un message de sondage appelé une "demande d'écho MPLS" et un message de réponse appelé une "réponse d'écho MPLS" pour retourner le résultat du sondage. La demande d'écho MPLS est destinée à contenir des informations suffisantes pour vérifier le fonctionnement correct du plan de données et vérifier le plan des données par rapport au plan de contrôle, localisant ainsi les fautes.

Le présent document rend obsolètes les RFC 4379, 6424, 6829, et 7537, et met à jour la RFC 1122.

Statut de ce mémoire
Ceci est un document de l'Internet en cours de normalisation.

Le présent document a été produit par l'équipe d'ingénierie de l'Internet (IETF). Il représente le consensus de la communauté de l'IETF. Il a subi une révision publique et sa publication a été approuvée par le groupe de pilotage de l'ingénierie de l'Internet (IESG). Tous les documents approuvés par l'IESG ne sont pas candidats à devenir une norme de l'Internet ; voir la Section 2 de la RFC5741.

Les informations sur le statut actuel du présent document, tout errata, et comment fournir des réactions sur lui peuvent être obtenues à http://www.rfc-editor.org/info/rfc8029

Table des matières

1. Introduction

Le présent document décrit un mécanisme simple et efficace pour détecter les défaillances de plan des données dans les chemins à commutation d'étiquette (LSP, *Label Switched Path*) de commutation d'étiquette multi protocoles (MPLS, *Multi-Protocol Label Switching*). Il définit un message de sondage appelé une "demande d'écho MPLS" et un message de réponse appelé une "réponse d'écho MPLS" pour retourner le résultat du sondage. La demande d'écho MPLS est destinée à contenir des informations suffisantes pour vérifier le fonctionnement correct du plan de données, ainsi qu'un mécanisme pour vérifier le plan des données par rapport au plan de contrôle, localisant ainsi les fautes.

Une considération importante de la conception de cette demande d'écho MPLS est qu'elle suit le même chemin de données que les paquets MPLS normaux traversent. Les demandes d'écho MPLS sont principalement destinées à valider le plan des données et secondairement à vérifier le plan des données par rapport au plan de contrôle. Les mécanismes pour vérifier le plan de contrôle sont précieux mais ne sont pas couverts par le présent document.

Le présent document fait une utilisation particulière de la gamme d'adresses 127/8. C'est une exception au comportement défini dans la [RFC1122], et la présente spécification met à jour cette RFC. La motivation de ce changement et les détails de cette utilisation exceptionnelle sont discutées au paragraphe 2.1.

Le présent document rend obsolètes les [RFC4379], [RFC6424], [RFC6829], et [RFC7537].

1.1 Conventions

Les mots clés "DOIT", "NE DOIT PAS", "EXIGE", "DEVRA", "NE DEVRA PAS", "DEVRAIT", "NE DEVRAIT PAS",

"RECOMMANDE", "PEUT", et "FACULTATIF" en majuscules dans ce document sont à interpréter comme décrit dans le BCP 14, [RFC2119].

L'expression "doit être à zéro" (MBZ, *Must Be Zero*) est utilisée dans les descriptions d'objets pour les champs réservés. Ces champs DOIVENT être réglés à zéro à l'émission et ignorés à réception.

La terminologie relevant des couches 2 et 3 des réseaux privés virtuels (VPN, *Virtual Private Network*) est définie dans la [RFC4026].

Comme le présent document se réfère à la durée de vie (TTL, *Time to Live*) MPLS beaucoup plus fréquemment qu'au TTL IP, les auteurs ont choisi par convention d'utiliser "TTL" non qualifié pour signifier le "TTL MPLS" et d'utiliser "TTL IP" pour la valeur de TTL de l'en-tête IP.

1.2 Structure du document

Le corps de ce mémoire contient quatre parties principales : motivation, format du paquet de demande/réponse d'écho MPLS, fonctionnement du ping de LSP, et chemin de retour fiable. Il est suggéré que lors de la première lecture, on saute les formats de paquet et qu'on lise d'abord la "théorie du fonctionnement" (Section 4) ; le document est structuré ainsi pour éviter des références vers l'avant.

1.3 Domaine d'application de la spécification

Le principal but du présent document est de fournir une spécification nette et à jour du ping de LSP.

La [RFC4379] définit le mécanisme de base pour la validation de LSP MPLS qui peut être utilisé pour la détection et l'isolement des fautes. Le domaine d'application de ce document inclut aussi les diverses mises à jour du ping de LSP MPLS, incluant :

- La mise à jour de toutes les références et citations.
  - Les RFC 2434, 2030, et 3036 rendues obsolètes sont respectivement remplacées par les RFC 5226, 5905, et 5036.
  - De plus, certaines références pour informations ont été publiées comme RFC : les RFC 4761, 5085, 5885, et 8077.
- L'incorporation de tous les errata de RFC en cours.
  - Voir [Err108], [Err742], [Err1418], [Err1714], [Err1786], [Err2978], [Err3399].
- Remplacement de EXP par Classe de trafic (TC), sur la base de la mise à jour par la RFC 5462.
- Incorporation des mises à jour par la RFC 6829, en ajoutant les classes d'équivalence de transmission (FEC, *Forwarding Equivalence Classes*) de pseudo filaire (PW) annoncées sur IPv6 et rendant obsolète la RFC 6829.
- Incorporation des mises à jour de la RFC 7506, en ajoutant l'option d'alerte de routeur IPv6 (RAO, *Router Alert Option*) pour le fonctionnement, l'administration, et la maintenance (OAM) de MPLS.
- Incorporation des nouveaux bits définis sur le champ Fanions globaux à partir des RFC 6425 et 6426.
- Mise à jour des adresses IPv4 utilisées dans les exemples pour utiliser le préfixe de documentation. Ajout d'exemples avec des adresses IPv6.
- Incorporation des mises à jour de la RFC 6424, en déconseillant le TLV Transposition vers l'aval (DSMAP, *Transposition vers l'aval*) et ajout du TLV Transposition détaillée vers l'aval (DDMAP, *Downstream Detailed Mapping*) ; mise à jour de deux nouveaux codes de retour ; ajout des motivations des LSP tunnelés ou commutés ; mise à jour des procédures, des considérations relatives à l'IANA, et des considérations sur la sécurité ; et rendant obsolète la RFC 6424.
- Incorporation des mises à jour de la RFC 7537, en mettant à jour la Section Considérations relatives à l'IANA et rendant obsolète la RFC 7537.
- Finalement, en rendant obsolète la RFC 4379.

2. Motivation

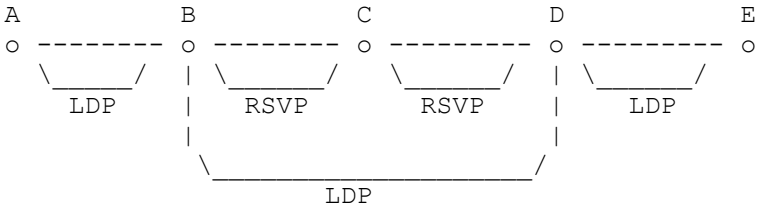
Quand un LSP échoue à livrer le trafic d'utilisateur, la défaillance ne peut pas toujours être détectée par le plan de contrôle MPLS. Il est nécessaire de fournir un outil qui permette aux utilisateurs de détecter un tel "trou noir" de trafic ou un défaut d'acheminement dans un délai raisonnable et un mécanisme pour isoler les fautes.

Dans le présent document, on décrit un mécanisme qui atteint ce but. Ce mécanisme est modélisé d'après le paradigme ping/traceroute : le ping (demande d'écho ICMP [RFC0792]) est utilisé pour des vérifications de connexité, et traceroute est utilisé pour la localisation de faute bond par bond ainsi que pour le retraçage du chemin. Le présent document spécifie un

mode "ping" et un mode "traceroute" pour tester les LSP MPLS.

L'idée de base est de vérifier que les paquets qui appartiennent à une certaine FEC terminent en fait leur chemin MPLS sur un routeur de commutation d'étiquettes (LSR, *Label Switching Router*) qui soit une sortie pour cette FEC. Le présent document propose que ce test soit effectué par l'envoi d'un paquet (appelé une "demande d'écho MPLS") le long du même chemin de données que les autres paquets appartenant à cette FEC. Une demande d'écho MPLS porte aussi des informations sur la FEC dont le chemin MPLS est vérifié. Cette demande d'écho est transmise comme tout autre paquet appartenant à cette FEC. En mode "ping" (vérification de connexité de base) le paquet devrait atteindre la fin du chemin, point auquel il est envoyé au plan de contrôle du LSR de sortie, qui vérifie alors si il est bien une sortie pour la FEC. En mode "traceroute" (isolement de faute) le paquet est envoyé au plan de contrôle de chaque LSR de transit, qui effectue diverses vérifications pour confirmer qu'il est bien un LSR de transit pour ce chemin ; ce LSR retourne aussi des informations qui aident à confronter le plan de contrôle au plan des données, c'est-à-dire, que la transmission correspond à ce que les protocoles de transmission ont déterminé comme le chemin.

Un LSP traceroute peut traverser un LSP tunnelé ou recousu en route pour la destination. Tout en effectuant la validation de LSP de bout en bout dans de tels scénarios, les informations de FEC incluses dans le paquet par l'initiateur peuvent être différentes de celles allouées par le nœud de transit dans un segment différent d'un LSP recousu ou tunnel. Considérons un cas simple :



Quand un LSP traceroute est initié du Routeur A au Routeur E, les informations de FEC incluses dans le paquet vont être LDP tandis que le Routeur C sur le chemin est un pur nœud RSVP et n'utilise pas LDP. Par conséquent, le nœud C va être incapable d'effectuer la validation de la FEC. La demande d'écho MPLS devrait contenir des informations suffisantes pour permettre à tout nœud de transit au sein d'un LSP recousu ou tunnelé d'effectuer la validation de la FEC pour détecter toute demande d'écho en fausse direction.

Une façon d'utiliser cet outil est de faire un ping périodique sur la FEC pour s'assurer de la connexité. Si le ping échoue, on peut alors initier un traceroute pour déterminer où se trouve la faute. On peut aussi faire périodiquement un traceroute sur les FEC pour vérifier que la transmission correspond au plan de contrôle ; cependant, cela fait peser un lourd fardeau sur les LSR de transit et donc devrait être utilisé avec modération.

2.1 Utilisation de la gamme d'adresses 127/8

Comme décrit ci-dessus, le ping de LSP est destiné à être un outil de diagnostic. Il est destiné à permettre aux fournisseurs de services fondés sur MPLS d'isoler les fautes du réseau. En particulier, le ping de LSP doit diagnostiquer des situations où les plans de contrôle et de données ne sont plus synchrones. Il effectue cela en acheminant un paquet de demande d'écho MPLS sur la seule base de sa pile d'étiquettes. C'est-à-dire que l'adresse de destination IP n'est jamais utilisée dans une décision de transmission. En fait, l'envoyeur d'un paquet de demande d'écho MPLS peut ne pas savoir, à priori, l'adresse du routeur à la fin du LSP.

Les fournisseurs de services fondés sur MPLS doivent aussi être capables de tracer tous les chemins possibles que peut prendre un LSP. Comme la plupart des services MPLS se fondent sur la transmission IP en envoi individuel, ces chemins sont soumis au partage de charge de multi chemins de coût égal (ECMP, *Equal-Cost Multipath*).

Cela conduit aux exigences suivantes :

1 Bien que le LSP en question puisse être cassé sans qu'on le sache, la probabilité qu'un paquet de diagnostic soit livré à un usager d'un service MPLS DOIT être tenue comme un minimum absolu.

2 Si un LSP est cassé de telle façon qu'il se termine prématurément, le paquet de diagnostic NE DOIT PAS être transmis par IP.

3 Un moyen de varier les paquets de diagnostic de façon à ce qu'ils éprouvent tous les chemins ECMP est donc EXIGÉ.

Il est clair que l'utilisation des adresses générales en envoi individuel ne satisfait pas aux deux premières exigences. Un certain nombre d'autres options pour les adresses ont été considérées, incluant une portion de l'espace d'adresses privé (comme déterminé par l'opérateur du réseau) et les adresses IPv4 de liaison locale. L'utilisation de l'espace d'adresses privé a été réputé inefficace car le service principal fondé sur MPLS est un VPN IPv4. Les VPN utilisent souvent des adresses privées.

Les adresses IPv4 de liaison locale sont plus attirantes en ce que la portée sur laquelle elles peuvent être transmises est limitée. Cependant, si on veut utiliser une adresse de cette gamme, il serait toujours possible pour le premier receveur d'un paquet de diagnostic qui s'est "échappé" d'un LSP cassé d'avoir cette adresse allouée à l'interface sur lequel il est arrivé et donc il pourrait par erreur recevoir un tel paquet. Les déploiement plus anciens de routeurs peuvent ne pas mettre (correctement) en œuvre les adresses IPv4 de liaison locale et transmettre un paquet avec une adresse dans cette gamme sur un chemin par défaut.

La gamme 127/8 pour IPv4 et cette même gamme incorporée dans une adresse IPv6 transposée en IPv4 pour IPv6 a été choisie pour un certain nombre de raisons.

La RFC 1122 alloue le 127/8 comme "adresse de bouclage arrière d'hôte interne" et déclare : "Les adresses de cette forme NE DOIVENT PAS apparaître en dehors d'un hôte". Donc, le comportement par défaut des hôtes est d'éliminer de tels paquets. Cela aide à s'assurer que si un paquet de diagnostic est mal dirigé sur un hôte, il va être éliminé en silence.

La [RFC1812] déclare : "Un routeur NE DEVRAIT PAS transmettre, sauf sur une interface de rebouclage, un paquet qui a une adresse de destination sur le réseau 127. Un routeur PEUT avoir un commutateur qui permet au gestionnaire de réseau de désactiver ces vérifications. Si un tel commutateur est fourni, il DOIT par défaut effectuer les vérifications".
Cela aide à s'assurer que les paquets de diagnostic ne sont jamais transmis sur IP.

La gamme d'adresses 127/8 fournit 16M adresses donnant une grande souplesse pour faire varier les adresses pour évaluer les chemins ECMP. Finalement, comme optimisation de mise en œuvre, la gamme 127/8 donne un moyen aisé d'identifier de possibles paquets LSP.

2.2 Option d'alerte de routeur

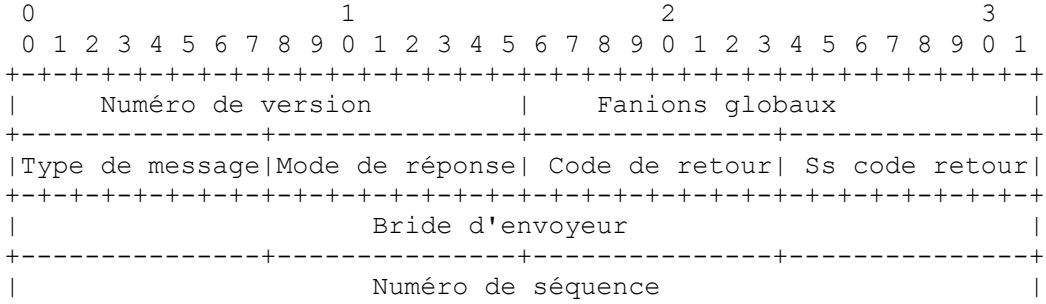
Le présent document exige l'utilisation de l'option Alerte de routeur (RAO, *Routeur Alert Option*) établie dans un en-tête IP afin que le nœud de transit traite la charge utile OAM MPLS.

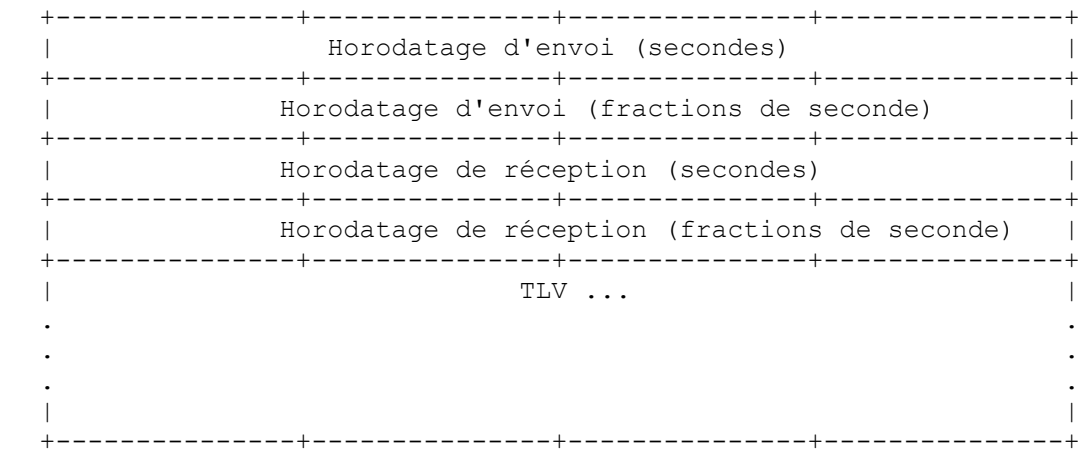
La [RFC2113] définit une valeur d'option générique 0x0 pour la RAO IPv4 qui alerte le routeur de transit afin d'examiner le paquet IPv4. La [RFC7506] définit la valeur d'option OAM MPLS de 69 pour la RAO IPv6 afin d'alerter les routeurs de transit pour qu'ils examinent le paquet IPv6 de plus près pour les besoins de l'OAM MPLS .

L'utilisation de l'option Alerte de routeur IP dans le présent document est comme suit : dans le cas d'un en-tête IPv4, la valeur générique de RAO IPv4 de 0x0 [RFC2113] DEVRAIT être utilisée. Dans le cas d'un en-tête IPv6, la valeur de RAO IPv6 de 69 pour OAM MPLS [RFC7506] DOIT être utilisée.

3. Format de paquet

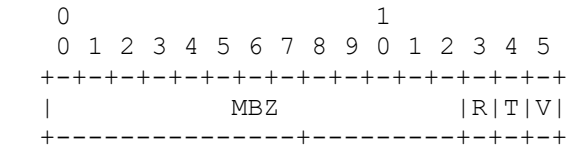
Une demande/réponse d'écho MPLS est un paquet UDP (éventuellement étiqueté) IPv4 ou IPv6 ; le contenu du paquet UDP a le format suivant :





Le numéro de version est actuellement 1. (Note : le numéro de version doit être incrémenté chaque fois qu'est fait un changement qui affecte la capacité d'une mise en œuvre d'analyser ou traiter correctement une demande/réponse d'écho MPLS. Ces changements incluent tout changement syntaxique ou sémantique à un des champs fixés, ou à toute allocation ou format de Type-Longueur-Valeur (TLV) ou sous-TLV défini à un certain numéro de version. Le numéro de version peut n'avoir pas besoin d'être changé si un TLV ou sous-TLV facultatif est ajouté.)

Le champ Fanions globaux est un vecteur de bits du format suivant :



Au moment de la rédaction, trois fanions sont définis : les bits R, T, et V ; le reste DOIT être réglé à zéro à l'émission et ignoré à réception.

Le fanion V (Valider la pile de FEC) est réglé à 1 si l'envoyeur veut que le receveur effectue la validation de pile de FEC ; si V est 0, le choix est laissé au receveur.

Le fanion T (Répondre seulement si le TTL est expiré) DOIT être établi seulement dans le paquet de demande d'écho par l'envoyeur. Si le fanion T est réglé à 1 dans une demande d'écho entrante, et si le TTL de l'étiquette MPLS entrante est supérieur à 1, le nœud receveur DOIT alors éliminer la demande d'écho entrante et NE DOIT PAS envoyer de réponse d'écho à l'envoyeur. Ce fanion NE DOIT PAS être établi dans le paquet de réponse d'écho. Si ce fanion est établi dans un paquet de réponse d'écho, il DOIT être ignoré. Le fanion T est défini au paragraphe 3.4 de la [RFC6425].

Le fanion R (Valider le chemin inverse) est défini dans la [RFC6426]. Quand ce fanion est établi dans la demande d'écho, celui qui répond DEVRAIT retourner les informations de FEC du chemin inverse, comme décrit au paragraphe 3.4.2 de la [RFC6426].

Le type de message est un des suivants :

Valeur	**Signification**
1	Demande d'écho MPLS
2	Réponse d'écho MPLS

Le mode de réponse peut prendre une des valeurs suivantes :

Valeur	**Signification**
1	Ne pas répondre
2	Répondre via un paquet UDP IPv4/IPv6
3	Répondre via un paquet UDP IPv4/IPv6 avec alerte de routeur
4	Répondre via un canal de contrôle de niveau application

Une demande d'écho MPLS avec 1 (Ne pas répondre) dans le champ Mode de réponse peut être utilisée pour des essais de

connexité unidirectionnelle ; le routeur receveur peut enregistrer les trous dans les numéros de séquence et/ou tenir des statistiques de délai/gigue. Une demande d'écho MPLS va normalement avoir 2 (Réponse via un paquet UDP IPv4/IPv6) dans le champ Mode de réponse. Si le chemin normal de retour IP est réputé non fiable, on peut utiliser 3 (Réponse via un paquet UDP IPv4/IPv6 avec alerte de routeur). Noter que cela exige que tous les routeurs intermédiaires comprennent et sachent comment transmettre les réponses d'écho MPLS. La réponse d'écho utilise le même numéro de version IP que la demande d'écho reçue, c'est-à-dire, une réponse d'écho IPv4 encapsulée est envoyée en réponse à une demande d'écho IPv4 encapsulée.

Certaines applications prennent en charge un canal de contrôle IP. Un exemple en est le canal de contrôle associé défini dans la vérification de connexité de circuit virtuel (VCCV, *Virtual Circuit Connectivity Verification*) [RFC5085], [RFC5885]. Toute application qui prend en charge un canal de contrôle IP entre ses entités de contrôle peut régler le mode de réponse à 4 (Réponse via canal de contrôle de niveau application) pour s'assurer que les réponses utilisent ce même canal. Une définition plus précise de ce codet est spécifique de l'application et sort donc du domaine d'application du présent document.

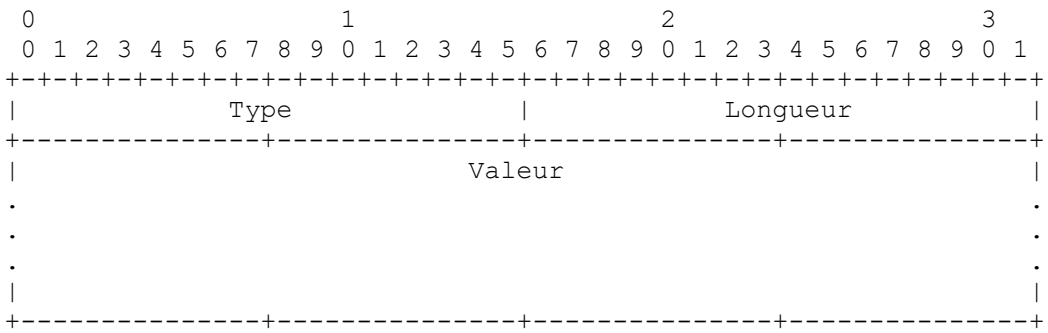
Les codes et sous codes de réponse sont décrits au paragraphe 3.1.

La bride de l'envoyeur est remplie par l'envoyeur et retournée inchangée par le receveur dans la réponse d'écho (si il y en a une). Il n'y a pas de sémantique associée à cette bride, bien qu'un envoyeur puisse trouver cela utile pour faire correspondre les demandes avec les réponses.

Le numéro de séquence est alloué par l'envoyeur de la demande d'écho MPLS et peut être (par exemple) utilisé pour détecter des réponses manquées.

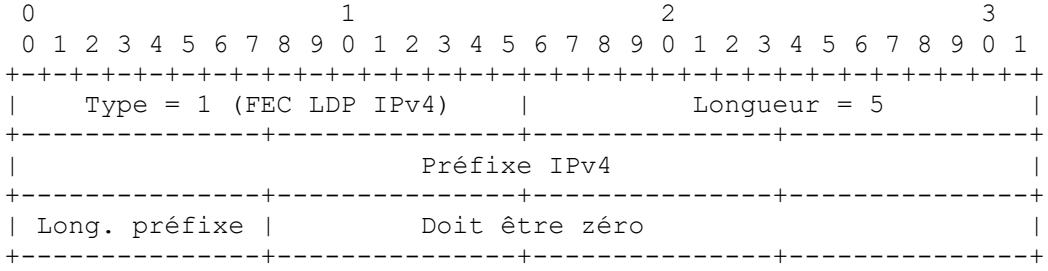
L'horodatage d'envoi est l'heure du jour (selon l'horloge de l'envoyeur) en format d'horodatage NTP de 64 bits [RFC5905] à laquelle la demande d'écho MPLS est envoyée. L'horodatage reçu dans une réponse d'écho est l'heure du jour (selon l'horloge du receveur) en format d'horodatage NTP de 64 bits à laquelle la demande d'écho correspondante a été reçue.

Les TLV (triplets de Type-Longueur-Valeur) ont le format suivant :



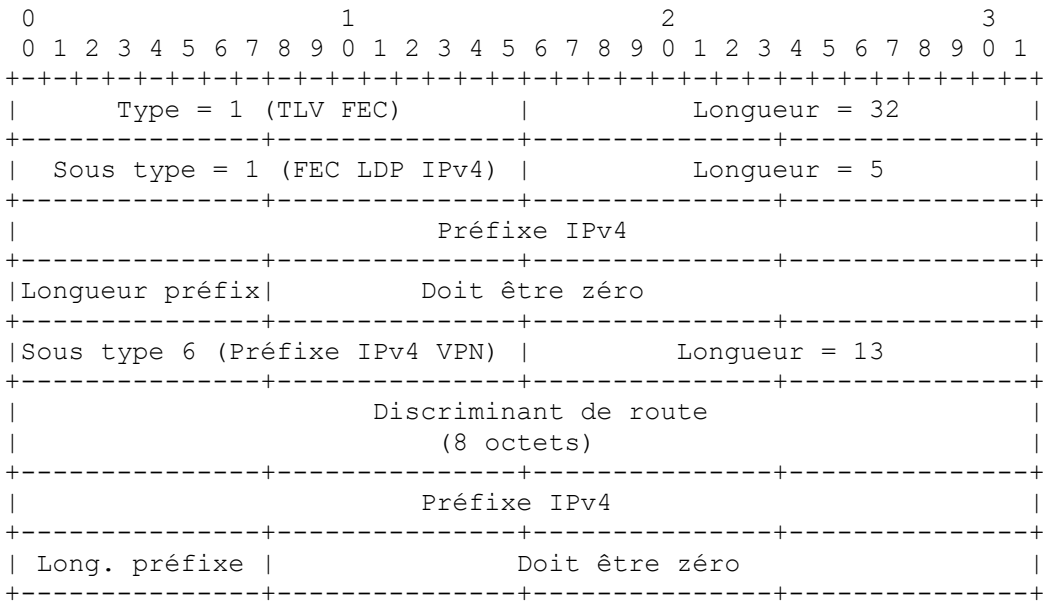
Les types sont définis ci-dessous ; Longueur est la longueur du champ Valeur en octets. Le champ Valeur dépend du type ; il est bourré de zéros pour l'aligner sur une limite de 4 octets. Les TLV peuvent être incorporés au sein d'autres TLV, auquel cas les TLV incorporés sont appelés des sous TLV. Les sous TLV ont des types indépendants et DOIVENT aussi être alignés sur quatre octets.

Deux exemples de la façon dont les longueurs de sous TLV sont calculées, et comment les sous TLV sont bourrés pour être alignés sur quatre octets, sont les suivants :



La longueur pour ce TLV est 5. Un TLV Pile de FEC cible qui contient un sous TLV FEC LDP IPv4 et un sous TLV

Préfixe de VPN IPv4 a le format suivant :



Une description des types et valeurs des TLV de niveau supérieur pour le ping de LSP est donnée ci-dessous :

Numéro de type	Champ de valeur
1	Pile de FEC cible
2	Transposition vers l'aval (déconseillé)
3	Bourrage
4	Non alloué
5	Numéro d'entreprise de fabricant
6	Non alloué
7	Interface et pile d'étiquette
8	Non alloué
9	TLV erroné
10	Octet TOS de réponse
20	Transposition détaillée vers l'aval

Les types de moins de 32 768 (c'est-à-dire, avec le bit de poids fort égal à 0) sont des TLV obligatoires qui DOIVENT soit être pris en charge par une mise en œuvre, soit résulter en l'envoi du code de retour de 2 ("Un ou plusieurs des TLV n'a pas été compris") dans la réponse d'écho.

Les types supérieurs ou égaux à 32 768 (c'est-à-dire, avec le bit de poids fort égal à 1) sont des TLV facultatifs qui DEVRAIENT être ignorés si la mise en œuvre ne les comprend pas ou ne les prend pas en charge.

Dans les paragraphes 3.2 à 3.9 et leurs divers sous paragraphes, seul le champ Valeur du TLV est inclus.

3.1 Codes de retour

Le code de retour est réglé à zéro par l'envoyeur d'une demande d'écho. Le receveur de ladite demande d'écho peut la régler à une des valeurs de la liste ci-dessous dans la réponse d'écho correspondante qu'il génère. La notation <RSC> se réfère au sous code de retour. Ce champ est rempli dans les codes qui en spécifient. Pour tous les autres codes, le sous code de retour DOIT être réglé à zéro.

Valeur	Signification
0	Pas de code de retour
1	La demande d'écho reçue est mal formée
2	Un ou plusieurs des TLV ne sont pas compris
3	Le routeur qui répond est une sortie pour la FEC à la profondeur de pile <RSC>
4	Le routeur qui répond n'a pas de transposition pour la FEC à la profondeur de pile <RSC>

5 Discordance de transposition vers l'aval (voir la Note 1)
6 Indice d'interface amont inconnu (voir Note 1)
7 Réservé
8 Étiquette commutée à la profondeur de pile <RSC>
9 Étiquette commutée mais pas de transmission MPLS à la profondeur de pile <RSC>
10 La transposition pour cette FEC n'est pas l'étiquette données à la profondeur de pile <RSC>
11 Pas d'entrée d'étiquette à la profondeur de pile <RSC>
12 Protocole non associé à l'interface à la profondeur de pile de FEC <RSC>
13 Terminaison prématurée de ping due à la réduction de la pile d'étiquettes à une seule étiquette
14 Voir le TLV DDMAP pour la signification du code et sous code de retour (voir la Note 2)
15 Étiquette commutée avec changement de FEC

Note 1 : Le sous code de retour (RSC) contient le point dans la pile d'étiquettes où le traitement s'est terminé. Si le RSC est 0, aucune étiquette n'a été traitée. Autrement, le paquet a été commuté par étiquette à la profondeur du RSC.

Note 2 : Le code de retour est selon le "TLV Transposition détaillée vers l'aval" (paragraphe 3.4). Ce code de retour DOIT être utilisé seulement dans l'en-tête de message et DOIT être établi seulement dans le message réponse d'écho MPLS. Si le code de retour est établi dans le message demande d'écho MPLS, il DOIT alors être ignoré. Quand ce code de retour est établi, chaque TLV Transposition détaillée vers l'aval DOIT avoir un code et sous code de retour approprié. Ce code de retour DOIT être utilisé quand il y a plusieurs aval pour un certain nœud (comme un point à multipoint (P2MP) ou un ECMP) et que le nœud a besoin d'un code/sous code de retour pour chaque aval. Ce code de retour PEUT être utilisé même quand il y a seulement un aval pour le nœud.

3.2 Pile de FEC cible

Une pile de FEC cible est une liste de sous TLV. Le nombre des éléments est déterminé en regardant dans le champ Longueur du sous TLV.

Sous type	Longueur	Champ Valeur
1	5	Préfixe LDP IPv4
2	17	Préfixe LDP IPv6
3	20	LSP RSVP IPv4
4	56	LSP RSVP IPv6
5		Non alloué
6	13	Préfixe VPN IPv4
7	25	Préfixe VPN IPv6
8	14	Point d'extrémité de couche 2 de VPN
9	10	"FEC 128" de pseudo filaire - IPv4 (déconseillé)
10	14	"FEC 128" de pseudo filaire - IPv4
11	16+	"FEC 129" de pseudo filaire - IPv4
12	5	Préfixe IPv4 étiqueté BGP
13	17	Préfixe IPv6 étiqueté BGP
14	5	Préfixe IPv4 générique
15	17	Préfixe IPv6 générique
16	4	FEC nulle
24	38	"FEC 128" de pseudo filaire - IPv6
25	40+	"FEC 129" de pseudo filaire - IPv6

D'autres types de FEC ont été définis et seront définis en tant que de besoin.

Noter que ces TLV définissent une pile de FEC, le premier élément de FEC correspondant au sommet de la pile d'étiquettes, etc.

Une demande d'écho MPLS DOIT avoir une pile de FEC cible qui décrit la pile de FEC soumise à l'essai. Par exemple, si un LSR X a une transposition de LDP [RFC5036] pour 192.0.2.1 (disons, l'étiquette 1001) alors pour vérifier que l'étiquette 1001 atteint bien un LSR de sortie qui annonçait ce préfixe via LDP, X peut envoyer une demande d'écho MPLS avec un TLV Pile de FEC avec une FEC dedans, à savoir, de type Préfixe LDP IPv4, avec le préfixe 192.0.2.1/32, et envoyer la demande d'écho avec une étiquette de 1001.

Disons que le LSR X voulait vérifier qu'une pile d'étiquettes de <1001, 23456> est la bonne pile d'étiquettes à utiliser pour

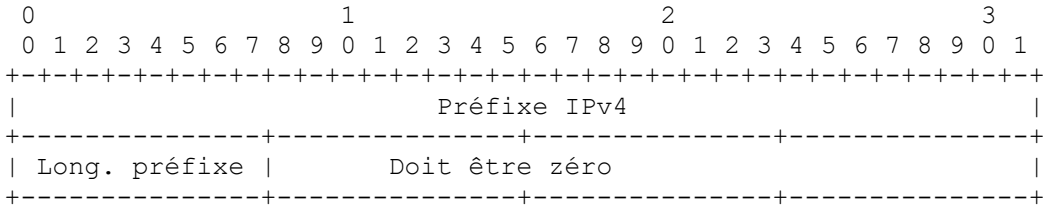
atteindre un préfixe de VPN IPv4 (voir au paragraphe 3.2.5) de 203.0.113.0/24 dans le VPN foo. Disons de plus que le LSR Y avec l'adresse de rebouclage arrière 192.0.2.1 a annoncé le préfixe 203.0.113.0/24 avec le discriminateur de route (RD, *Route Distinguisher*) RD-foo-Y (qui peut en général être différent du RD que le LSR X utilise dans ses propres annonces pour le VPN foo), l'étiquette 23456, et le prochain bond BGP 192.0.2.1 [RFC4271]. Finalement, supposons que le LSR X reçoive un lien d'étiquette de 1001 pour 192.0.2.1 via LDP. X a le choix entre deux solutions : envoyer une demande d'écho MPLS ; X peut envoyer une demande d'écho MPLS avec un TLV Pile de FEC avec une seule FEC de type Préfixe VPN IPv4 avec un préfixe de 203.0.113.0/24 et un RD de RD-foo-Y. Autrement, X peut envoyer un TLV Pile de FEC avec deux FEC, la première de type LDP IPv4 avec un préfixe de 192.0.2.1/32 et la seconde de type VPN IP avec un préfixe 203.0.113.0/24 et un RD de RD-foo-Y. Dans l'un et l'autre cas, la demande d'écho MPLS va avoir une pile d'étiquettes de <1001, 23456>. (Note : dans cet exemple, 1001 est l'étiquette "externe" et 23456 est l'étiquette "interne".)

Si, par exemple, un LSR Y a une transposition de LDP pour l'adresse IPv6 2001:db8::1 (disons, l'étiquette 2001), alors pour vérifier que l'étiquette 2001 atteint bien un LSR de sortie qui a annoncé ce préfixe via LDP, le LSR Y peut envoyer une demande d'écho MPLS avec un TLV Pile de FEC avec une FEC Préfixe de LDP IPv6, avec le préfixe 2001:db8::1/128, et une étiquette de 2001.

Si un chemin de bout en bout se compose d'un ou plusieurs LSP tunnelés ou recousus, chaque nœud de transit qui est le point d'origine d'un nouveau tunnel ou segment DEVRAIT répondre en notifiant le changement de pile de FEC ainsi que les détails de la nouvelle FEC, par exemple, si le LSR X a une transposition de LDP pour le préfixe IPv4 192.0.2.10 sur le LSR Z (disons, l'étiquette 3001). Disons de plus que le LSR A et le LSR B sont des nœuds de transit le long du chemin, qui ont aussi un tunnel RSVP sur lequel LDP est activé. Tout en répondant, A DEVRAIT notifier que la FEC change de LDP à <RSVP, LDP>. Si le nouveau tunnel est un tuyau transparent, c'est-à-dire, si la trace du plan de données ne va pas arriver à expiration au milieu du tunnel, alors le nœud de transit NE DEVRAIT PAS répondre en notifiant le changement de pile de FEC ou les détails de la nouvelle FEC. Si le nœud de transit souhaite cacher la nature du tunnel à partir de l'entrée de la demande d'écho, le nœud de transit PEUT alors notifier le changement de pile de FEC et inclure la FEC nulle comme nouvelle FEC.

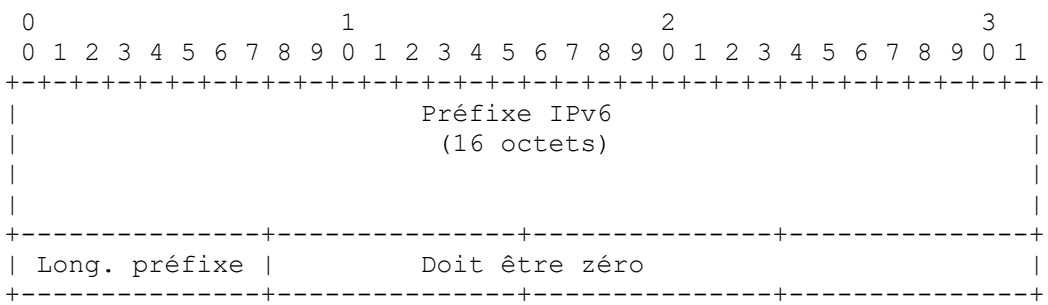
3.2.1 Préfixe LDP IPv4

La FEC Préfixe IPv4 est définie dans la [RFC5036]. Quand un préfixe LDP IPv4 est codé dans une pile d'étiquettes, le format suivant est utilisé. La valeur consiste en 4 octets d'un préfixe IPv4 suivi par 1 octet de longueur de préfixe en bits ; le format est donné ci-dessous. Le préfixe IPv4 est dans l'ordre des octets du réseau ; si le préfixe fait moins de 32 bits, les bits de queue DEVRAIENT être réglés à zéro. Voir dans la [RFC5036] un exemple de transposition f pour une FEC IPv4.



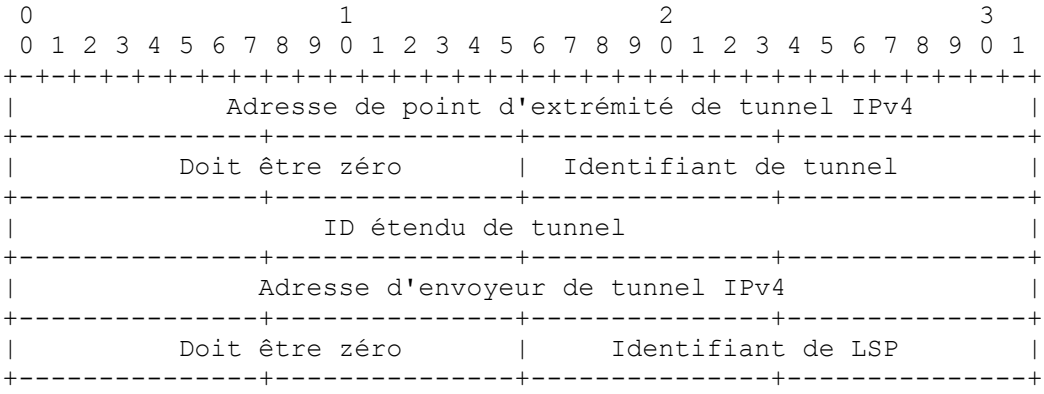
3.2.2 Préfixe LDP IPv6

La FEC Préfixe IPv6 est définie dans la [RFC5036]. Quand un préfixe IPv6 de LDP est codé dans une pile d'étiquettes, le format qui suit est utilisé. La valeur consiste en 16 octets d'un préfixe IPv6 suivi par un octet de longueur de préfixe en bits ; le format est donné ci-dessous. Le préfixe IPv6 est dans l'ordre des octets du réseau ; si le préfixe fait moins de 128 bits, les bits de queue DEVRAIENT être réglés à zéro. Voir dans la [RFC5036] un exemple de transposition pour une FEC IPv6.



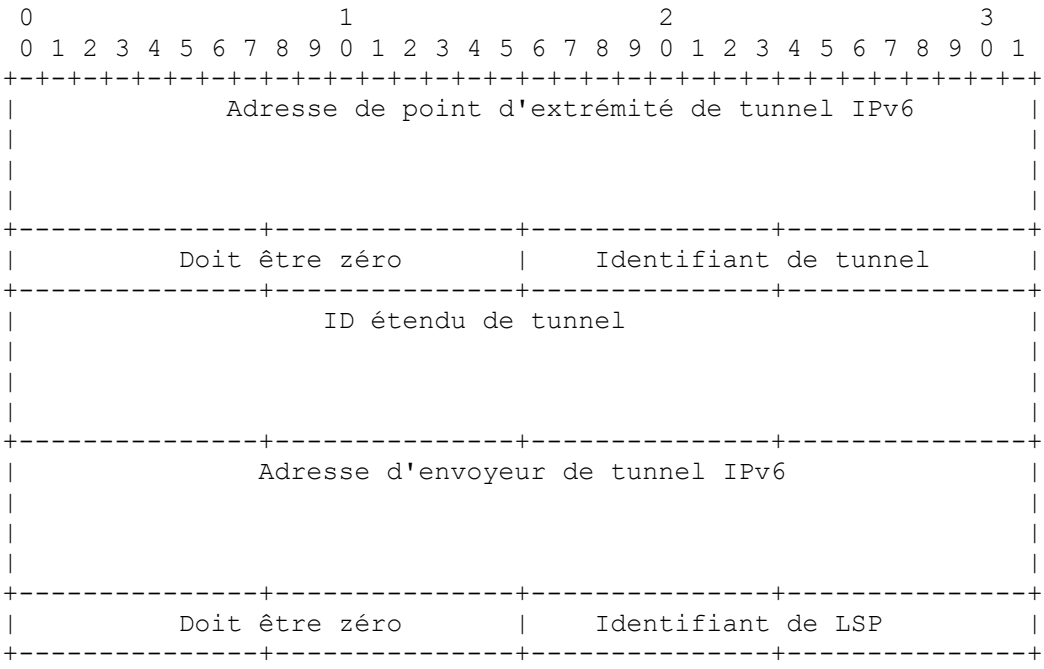
3.2.3 LSP RSVP IPv4

La valeur a le format ci-dessous. Les champs Valeur sont tirés de la [RFC3209], paragraphes 4.6.1.1 et 4.6.2.1.



3.2.4 LSP RSVP IPv6

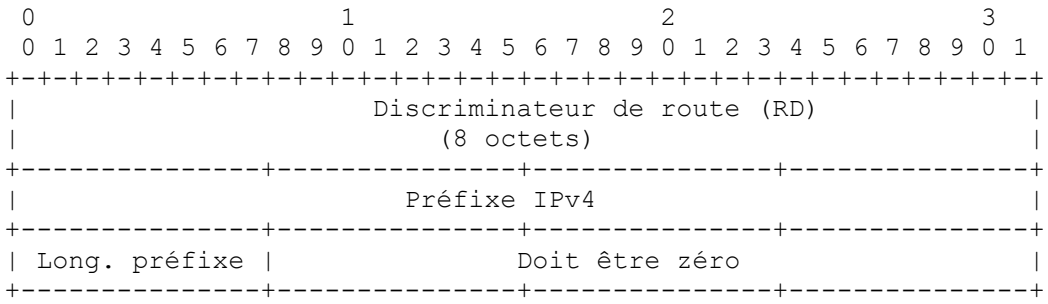
La valeur a le format ci-dessous. Les champs Valeur sont tirés de la [RFC3209], paragraphes 4.6.1.2 et 4.6.2.2.



3.2.5 Préfixe de VPN IPv4

Les informations d'acheminement de couche réseau (NLRI, *Network Layer Routing Information*) de VPN IPv4 sont définies dans la [RFC4365]. Le présent document utilise le terme de Préfixe de VPN IPv4 pour des NLRI de VPN IPv4 qui ont été annoncées avec une étiquette MPLS dans BGP. Voir la [RFC3107].

Quand un préfixe de VPN IPv4 est codé dans une pile d'étiquettes, le format suivant est utilisé. Le champ Valeur consiste en le RD annoncé avec le préfixe de VPN IPv4, le préfixe IPv4 (avec des bits à zéro en queue pour faire 32 bits en tout) et une longueur de préfixe, comme suit :

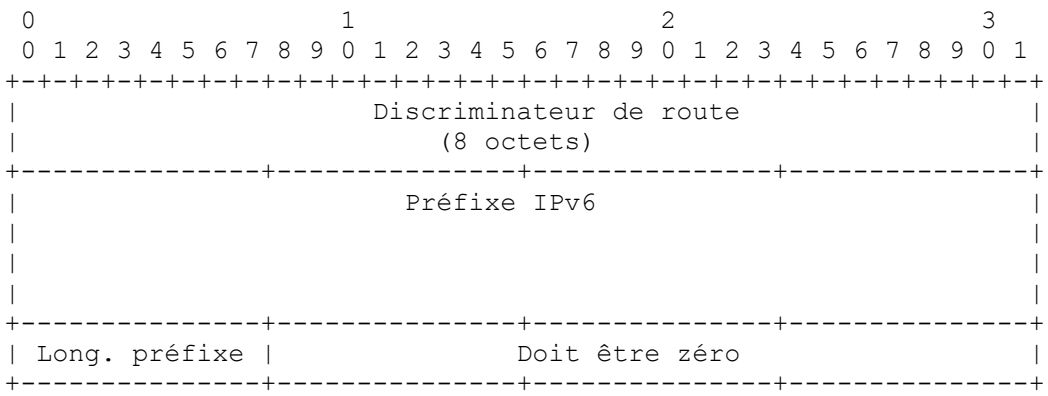


Le RD est un identifiant de 8 octets ; il ne contient aucune information inhérente. L'objet du RD est seulement de permettre de créer des routes distinctes à un préfixe d'adresse IPv4 commun. Le codage du RD n'a pas d'importance ici. Quand on confronte ce champ aux informations locales de FEC, il est traité comme une valeur opaque.

3.2.6 Préfixe de VPN IPv6

Les NLRI de VPN IPv6 sont définies dans la [RFC4365]. Le présent document utilise le terme "préfixe de VPN IPv6" pour des NLRI de VPN IPv6 qui ont été annoncées avec une étiquette MPLS dans BGP. Voir la [RFC3107].

Quand un préfixe de VPN IPv6 est codé dans une pile d'étiquettes, on utilise le format suivant. Le champ Valeur consiste en le RD annoncé avec le préfixe de VPN IPv6, le préfixe IPv6 (avec des bits 0 en queue pour faire 128 bits en tout) et une longueur de préfixe, comme suit :

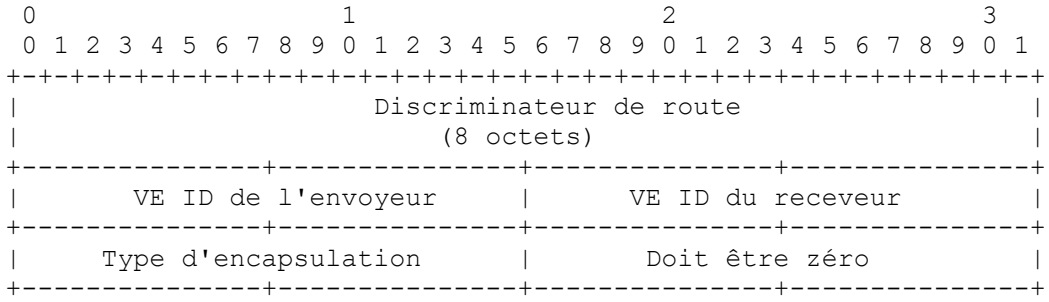


Le RD est identique au RD de préfixe de VPN IPv4, sauf qu'il fonctionne ici pour permettre la création de routes distinctes aux préfixes IPv6. Voir au paragraphe 3.2.5. Quand on confronte ce champ aux informations de FEC locales, il est traité comme une valeur opaque.

3.2.7 Point d'extrémité de VPN de couche 2

On utilise l'abréviation VPLS pour "service de LAN privé virtuel". Les termes NLRI BGP de VPLS et identifiant de bordure VPLS (VE ID, *VPLS Edge Identifier*) sont définis dans la [RFC4761]. Le présent document utilise le terme plus simple de point d'extrémité de VPN de couche 2 pour se référer à des NLRI BGP de VPLS. Le RD est un identifiant de 8 octets utilisé pour distinguer les informations sur les divers VPN de couche 2 annoncés par un nœud. Le VE ID est un identifiant de 2 octets utilisé pour identifier un certain nœud qui sert comme point de rattachement de service au sein d'un VPLS. La structure de ces deux identifiants est sans importance ici ; lorsque on confronte ces champs aux informations locales de FEC, elles sont traitées comme des valeurs opaques. Le type d'encapsulation est identique au type pseudo filaire (PW, *Pseudowire*) du paragraphe 3.2.9.

Quand un point d'extrémité de VPN de couche 2 est codé dans une pile d'étiquettes, on utilise le format suivant. Le champ Valeur consiste en un RD (8 octets), l'identifiant de bordure VPLS (VE ID) de l'envoyeur du ping (2 octets), le VE ID du receveur (2 octets), et un type d'encapsulation (2 octets), formaté comme suit :



3.2.8 FEC 128 pseudo filaire - IPv4 (déconseillé)

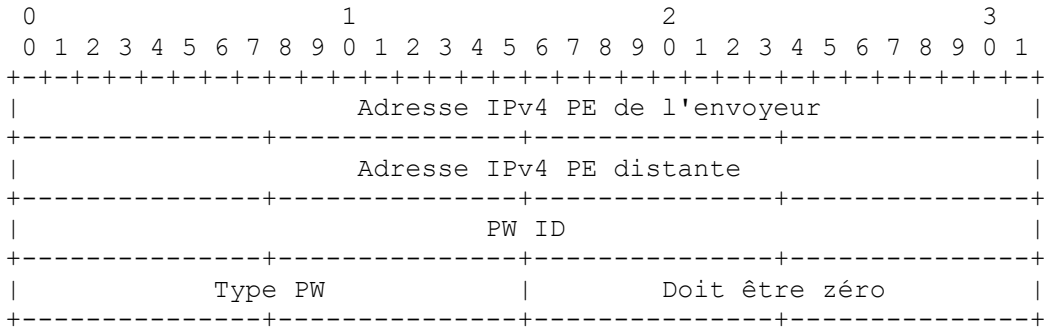
Voir les détails à l'Appendice A.1.1.

3.2.9 FEC 128 pseudo filaire - IPv4 (courant)

La FEC 128 (0x80) est définie dans la [RFC8077], comme le sont les termes "identifiant de pseudo filaire" (PW ID, *Pseudowire ID*) et "type de pseudo filaire (PW Type, *Pseudowire Type*). Un PW ID est un identifiant de connexion non zéro de 32 bits. Le PW Type est un nombre de 15 bits qui indique le type d'encapsulation. Il est porté justifié à droite dans le champ appelé ci-dessous "Type d'encapsulation" avec le bit de poids fort à zéro.

Ces deux champs sont traités comme des valeurs opaques dans ce protocole. Quand on confronte ces champs aux informations locales de FEC, la correspondance DOIT être exacte.

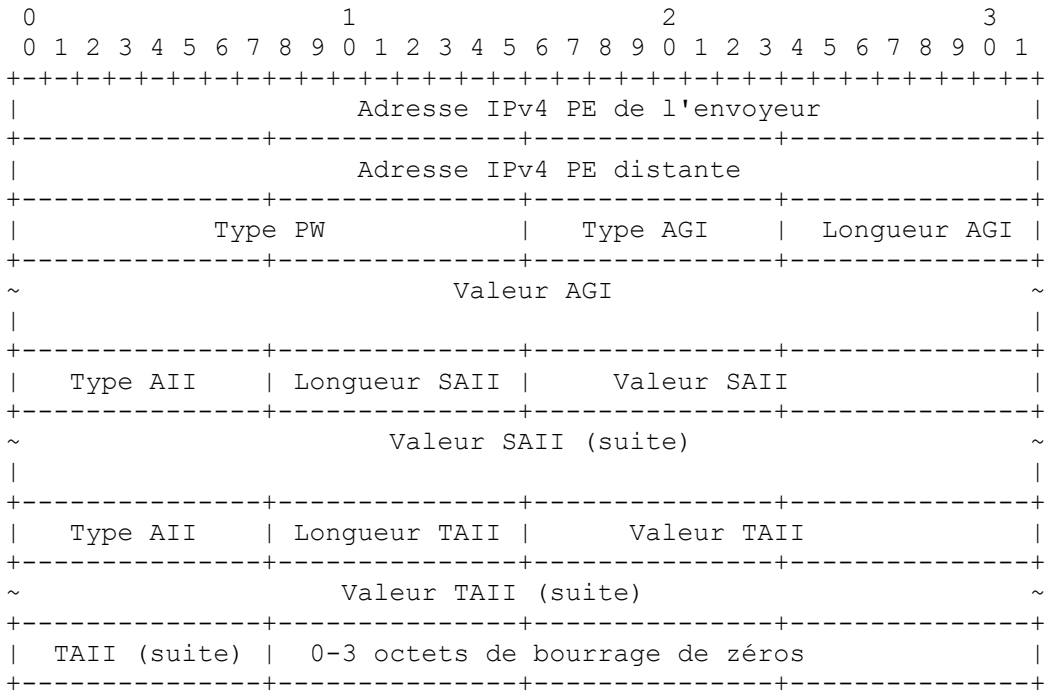
Quand une FEC 128 est codée dans une pile d'étiquettes, on utilise le format suivant. Le champ Valeur consiste en l'adresse IPv4 côté fournisseur (PE, *Provider Edge*) de l'envoyeur (l'adresse de source de la session LDP ciblée), en l'adresse IPv4 PE distante (l'adresse de destination de la session LDP ciblée), le PW ID, et le type d'encapsulation comme suit :



3.2.10 FEC 129 pseudo filaire - IPv4

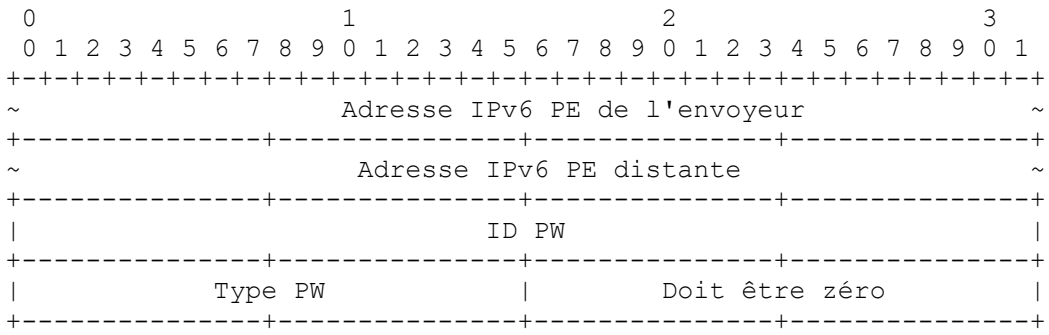
La FEC 129 (0x81) et les termes "type PW", "identifiant de groupe de rattachement" (AGI, *Attachment Group Identifier*), "type d'identifiant de groupe de rattachement" (AGI Type, *Attachment Group Identifier Type*), "type d'identifiant individuel de rattachement" (AII Type, *Attachment Individual Identifier Type*), "identifiant individuel de rattachement de source" (SAII, *Source Attachment Individual Identifier*), et "identifiant individuel de rattachement cible" (TAII, *Target Attachment Individual Identifier*) sont définis dans la [RFC8077]. Le type de PW est un nombre de 15 bits qui indique le type d'encapsulation. Il est porté justifié à droite dans le champ en dessous du type PW avec le bit de poids fort réglé à zéro. Tous les autres champs sont traités comme des valeurs opaques et copiés directement du format de FEC 129. Toutes ces valeurs ensemble définissent de façon univoque la FEC dans la portée de la session LDP identifiée par les adresses IPv4 de PE de source et distant.

Quand une FEC 129 est codée dans une pile d'étiquettes, on utilise le format suivant. Le champ Longueur de ce TLV est 16 + longueur de AGI + longueur de SAII + longueur de TAII. Un bourrage est utilisé pour faire de la longueur totale un multiple de 4 ; la longueur du bourrage n'est pas incluse dans le champ Longueur.



3.2.11 FEC 128 pseudo filaire - IPv6

Le sous TLV FEC 128 pseudo filaire IPv6 a une structure cohérente avec le sous TLV FEC 128 pseudo filaire IPv4 décrit au paragraphe 3.2.9. Le champ Valeur consiste en l'adresse IPv6 PE de l'envoyeur (l'adresse de source de la session LDP ciblée) l'adresse IPv6 PE distante (l'adresse de destination de la session LDP ciblée) le PW ID, et le type d'encapsulation comme suit :



Adresse IPv6 PE de l'envoyeur : adresse IP de source de la session LDP IPv6 cible ; 16 octets.

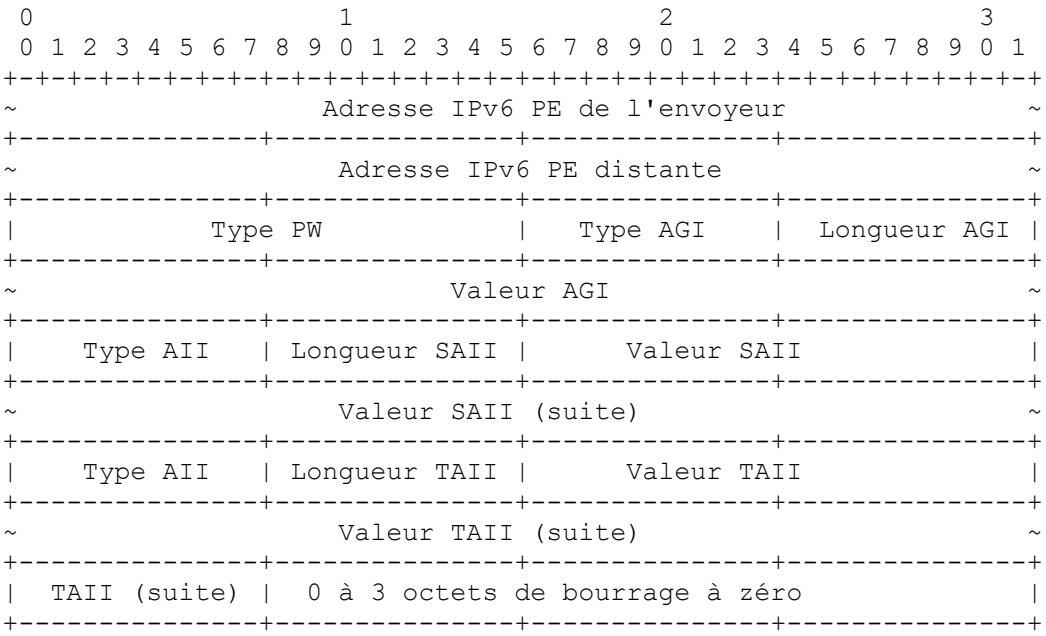
Adresse IPv6 PE distante : adresse IP de destination de la session LDP IPv6 cible ; 16 octets.

ID PW : le même que celui de la FEC 128 pseudo filaire IPv4 du paragraphe 3.2.9.

Type PW : le même que celui de la FEC 128 pseudo filaire IPv4 du paragraphe 3.2.9.

3.2.12 FEC 129 pseudo filaire - IPv6

Le sous TLV FEC 129 pseudo filaire IPv6 a une structure cohérente avec celle de la sous TLV de la FEC 129 pseudo filaire IPv4 décrite au paragraphe 3.2.10. Quand une FEC 129 est codée dans une pile d'étiquettes, on utilise le format suivant. La longueur de cette TLV est 40 + longueur d'AGI (Identifiant de groupe de rattachement) + longueur de SAII (identifiant individuel de rattachement de source) + longueur de TAII (identifiant individuel de rattachement de cible). Le bourrage est utilisé pour faire de la longueur totale un multiple de 4 ; la longueur du bourrage n'est pas incluse dans le champ Longueur.



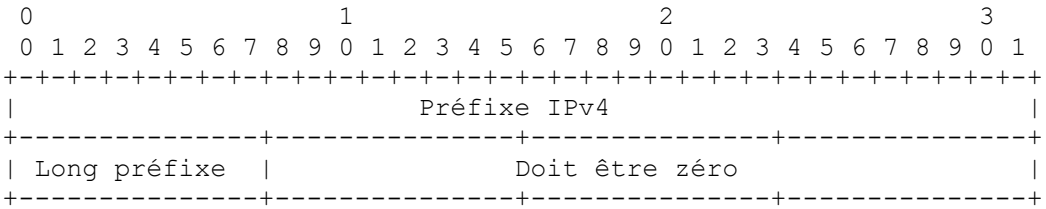
Adresse IPv6 PE de l'envoyeur : adresse IP de source de la session LDP IPv6 cible ; 16 octets.

Adresse IPv6 PE distante : adresse IP de destination de la session LDP IPv6 cible ; 16 octets.

Les autres champs sont les mêmes que ceux de la FEC 129 de pseudo filaire IPv4 du paragraphe 3.2.10.

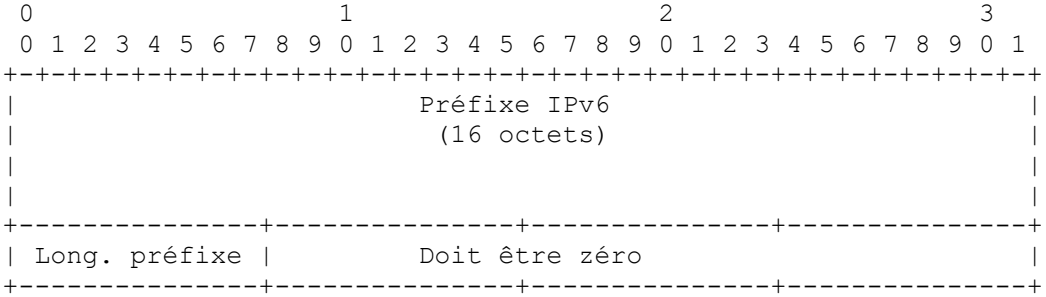
3.2.13 Préfixe IPv4 étiqueté BGP

Les préfixes IPv4 étiquetés BGP sont définis dans la [RFC3107]. Quand un préfixe IPv4 étiqueté BGP est codé dans une pile d'étiquettes, on utilise le format suivant. Le champ Valeur consiste en le préfixe IPv4 (avec des bits 0 en queue pour faire 32 bits en tout) et la longueur de préfixe, comme suit :



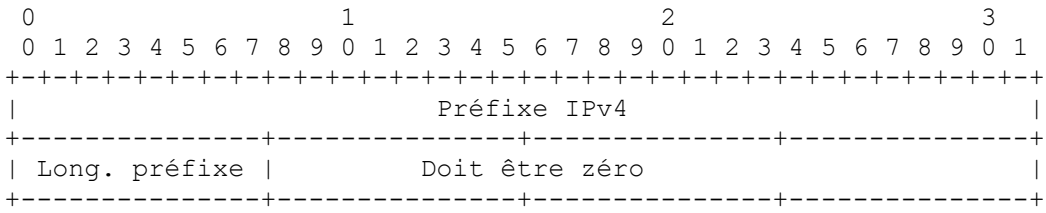
3.2.14 Préfixe IPv6 étiqueté BGP

Les préfixes IPv6 étiquetés BGP sont définis dans la [RFC3107]. Quand un préfixe IPv6 étiqueté BGP est codé dans une pile d'étiquettes, on utilise le format suivant. La valeur consiste en 16 octets d'un préfixe IPv6 suivi par 1 octet de longueur de préfixe en bits ; le format est donné ci-dessous. Le préfixe IPv6 est dans l'ordre des octets du réseau ; si le préfixe fait moins de 128 bits, les bits de queue DEVRAIENT être réglés à zéro.



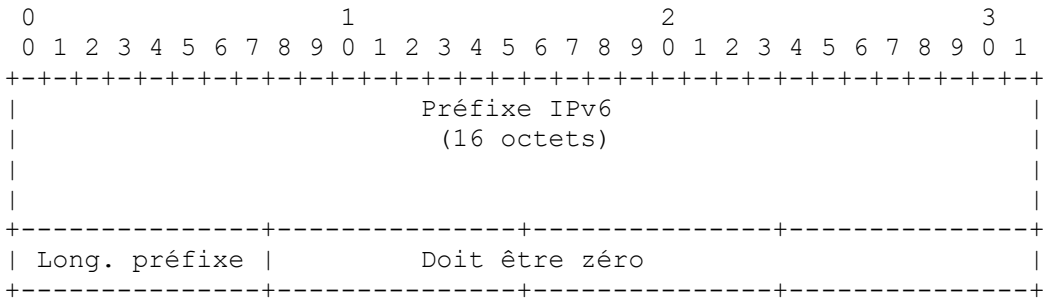
3.2.15 Préfixe IPv4 générique

La valeur consiste en 4 octets de préfixe IPv4 suivis par 1 octet de longueur de préfixe en bits ; le format est donné ci-dessous. Le préfixe IPv4 est dans l'ordre des octets du réseau ; si le préfixe fait moins de 32 bits, les bits de queue DEVRAIENT être réglés à zéro. Cette FEC est utilisée si le protocole qui annonce l'étiquette est inconnu ou peut changer durant le cours du LSP. Un exemple est un LSP inter AS qui peut être signalé par LDP dans un système autonome (AS), par RSVP-TE [RFC3209] dans un autre AS, et par BGP entre les AS, comme cela est courant pour les VPN inter AS.



3.2.16 Préfixe IPv6 générique

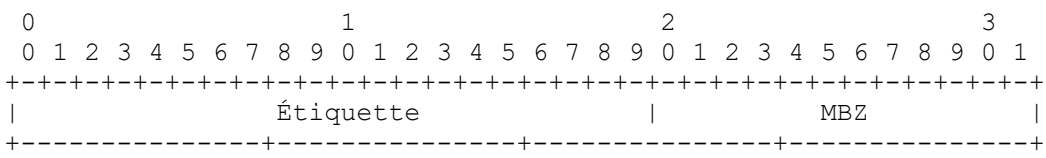
La valeur consiste en 16 octets d'un préfixe IPv6 suivi par un octet de longueur de préfixe en bits ; le format est donné ci-dessous. Le préfixe IPv6 est dans l'ordre des octets du réseau ; si le préfixe fait moins de 128 bits, les bits de queue DEVRAIENT être réglés à zéro.



3.2.17 FEC nulle

Parfois, des étiquettes provenant de la gamme réservée, par exemple, Alerte de routeur et Nulle explicite, peuvent être ajoutées à la pile d'étiquettes pour divers diagnostics comme sur l'équilibrage de charge. Ces étiquettes peuvent n'avoir pas une FEC explicite associée. La pile de FEC nulle est définie pour permettre qu'un sous TLV de pile de FEC cible soit ajouté à la pile de FEC cible pour prendre en compte de telles étiquettes afin qu'une validation appropriée puisse quand même être effectuée.

La longueur est 4. Les étiquettes sont des valeurs de 20 bits traitées comme des nombres.



Étiquette est la valeur réelle d'étiquette insérée dans la pile d'étiquettes ; le champ MBZ DOIT être à zéro à l'émission et ignoré à réception.

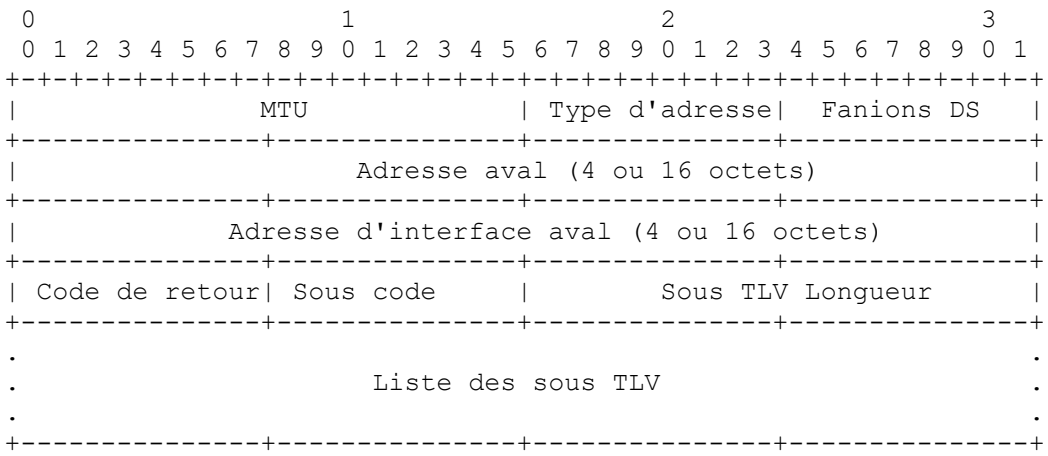
3.3 Transposition vers l'aval (déconseillée)

Voir les détails à l'Appendice A.2.

3.4 TLV de transposition vers l'aval détaillée

L'objet Transposition détaillée vers l'aval *(Downstream Detailed Mapping)* est un TLV qui PEUT être inclus dans un

message Demande d'écho MPLS. Un seul objet Transposition détaillée vers l'aval peut apparaître dans une demande d'écho. La présence d'un objet Transposition détaillée vers l'aval est une demande que les objets Transposition détaillée vers l'aval soient inclus dans la réponse d'écho MPLS. Si le routeur qui répond est la destination (Routeur de bordure d'étiquettes) de la FEC, une TLV Transposition détaillée vers l'aval NE DEVRAIT PAS être incluse dans la réponse d'écho MPLS. Autrement, le routeur qui répond DEVRAIT inclure un objet Transposition détaillée vers l'aval pour chaque interface sur laquelle cette FEC pourrait être transmise. Pour une définition plus précise de la notion de "vers l'aval", voir au paragraphe 3.4.2, "Routeur et interface aval".



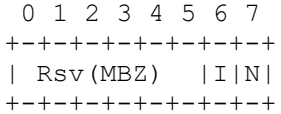
Le format du TLV Transposition détaillée vers l'aval est déduit du format du TLV déconseillé Transposition vers l'aval (voir l'Appendice A.2). Le changement clé est que les champs de longueur variable et facultatifs ont été convertis en sous TLV.

Unité maximum de transmission (MTU, *Maximum Transmission Unit*) : la MTU est la taille en octets de la plus grande trame MPLS (incluant la pile d'étiquettes) qui tient sur l'interface vers le LSR aval.

Type d'adresse : il indique si l'interface est numérotée ou non numérotée. Il détermine aussi la longueur des champs Adresse IP aval et Interface aval. Le type d'adresse est réglé à une des valeurs suivantes :

Numéro de type	**Type d'adresse**
1	IPv4 numérotée
2	IPv4 non numérotée
3	IPv6 numérotée
4	IPv6 non numérotée

Fanions DS : le champ Fanions DS est un vecteur binaire de divers fanions avec le format suivant :



Deux fanions sont actuellement définis , I et N. Les fanions restants DOIVENT être réglés à zéro à l'émission et ignorés à réception.

Fanion	**Nom et signification**
I	Demande d'objet Interface et pile d'étiquettes. Quand ce fanion est établi, il indique que le routeur qui répond DEVRAIT inclure un objet Interface et pile d'étiquettes dans le message de réponse d'écho.
N	Traiter comme paquet non IP. Les messages de demande d'écho vont être utilisés pour diagnostiquer les flux non IP. Cependant, ces messages sont portés dans des paquets IP. Pour un routeur qui altère son algorithme ECMP sur la base de la FEC ou d'un examen de paquet en profondeur, ce fanion demande que le routeur traite cela comme si la détermination d'une charge utile IP avait échoué.

Adresse aval et Adresse d'interface aval : les adresses IPv4 et les indices d'interface sont codés sur 4 octets ; les adresses IPv6 sont codées sur 16 octets.

Si l'interface vers le LSR aval est numérotée, le type d'adresse DOIT alors être réglé à IPv4 ou IPv6, l'adresse aval DOIT être réglée soit à l'identifiant de routeur du LSR aval, soit à l'adresse d'interface du LSR aval, et l'adresse de l'interface aval DOIT être réglée à l'adresse d'interface du LSR aval.
Si l'interface vers le LSR aval n'est pas numérotée, le type d'adresse DOIT être IPv4 non numéroté ou IPv6 non numéroté, l'adresse aval DOIT être l'identifiant de routeur du LSR aval, et l'adresse d'interface aval DOIT être réglée à l'indice alloué à l'interface par le LSR amont.
Si un LSR ne connaît pas l'adresse IP de son voisin, il DOIT alors régler le type d'adresse à IPv4 non numéroté ou à IPv6 non numéroté. Pour IPv4, il doit régler l'adresse aval à 127.0.0.1 ; pour IPv6, l'adresse est réglée à 0::1. Dans les deux cas, l'indice d'interface DOIT être réglé à 0. Si un LSR reçoit un paquet de demande d'écho avec une de ces adresses dans le champ Adresse aval, cela indique qu'il DOIT outrepasser la vérification d'interface mais continuer la validation d'étiquette.
Si le générateur d'un paquet de demande d'écho souhaite obtenir les informations de transposition détaillée vers l'aval mais ne connaît pas la pile d'étiquettes attendue, il DEVRAIT alors régler le type d'adresse à IPv4 non numéroté ou à IPv6 non numéroté. Pour IPv4, il DOIT régler l'adresse aval à 224.0.0.2 ; pour IPv6, l'adresse DOIT être réglée à FF02::2. Dans les deux cas, l'indice d'interface DOIT être réglé à 0. Si un LSR reçoit un paquet de demande d'écho avec l'adresse de diffusion groupée TOUS ROUTEURS, cela indique alors qu'il DOIT outrepasser la validation de l'interface et de la pile d'étiquettes mais retourner les TLV Transposition vers l'aval en utilisant les informations fournies.

Code de retour : le code de retour est réglé à zéro par l'envoyeur d'une demande d'écho. Le receveur de ladite demande d'écho peut le régler dans la réponse d'écho correspondante qu'il génère à une des valeurs spécifiées au paragraphe 3.1, autre que 14.
Si le receveur règle à une valeur non zéro le champ Code de retour dans le TLV Transposition détaillée vers l'aval, le receveur DOIT alors aussi régler le champ Code de retour dans l'en-tête de la réponse d'écho à "voir le TLV DDMAP pour le code de retour et le sous code de retour" (paragraphe 3.1). Une exception à cela est si le receveur est un nœud de bout [RFC4461] et répond à la fois comme nœud de transit et comme nœud de sortie avec le code de retour de 3 ("le routeur qui répond est une sortie pour la FEC à la profondeur de pile <RSC>") dans l'en-tête de réponse d'écho.
Si le code de retour du message de réponse d'écho n'est pas réglé à "voir le TLV DDMAP pour le code de retour et le sous code de retour" (paragraphe 3.1) ou à "le routeur qui répond est une sortie pour la FEC à la profondeur de pile <RSC>", le code de retour spécifié dans le TLV Transposition détaillée vers l'aval DOIT alors être ignoré.

Sous code de retour : le sous code de retour est réglé à zéro par l'envoyeur. Le receveur peut régler ce champ à une valeur appropriée comme spécifié au paragraphe 3.1 : le sous code de retour est rempli avec la profondeur de pile pour les codes qui spécifient la profondeur de pile. Pour tous les autres codes, le sous code de retour DOIT être réglé à zéro.
Si le code de retour du message de réponse d'écho n'est réglé ni à "voir le TLV DDMAP pour le code de retour et le sous code de retour" (paragraphe 3.1) ni à "le routeur qui répond est une sortie pour la FEC à la profondeur de pile <RSC>", le sous code de retour spécifié dans le TLV Transposition détaillée vers l'aval DOIT alors être ignoré.

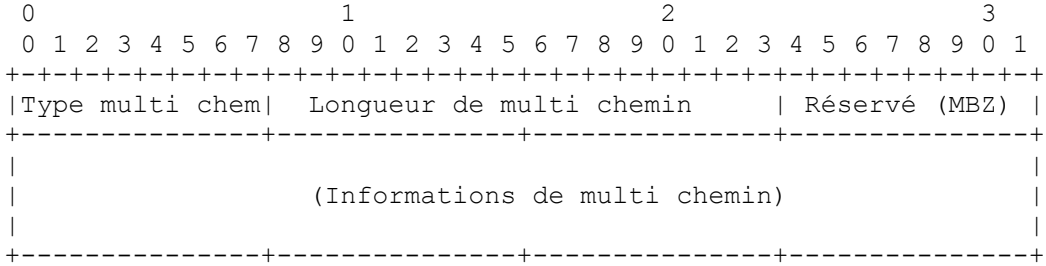
Longueur de sous TLV : longueur totale en octets des sous TLV associés à ce TLV.

3.4.1 Sous TLV

Ce paragraphe définit les sous TLV qui PEUVENT être inclus au titre du TLV Transposition détaillée vers l'aval.

Sous type	Champ Valeur
1	données de multi chemins
2	pile d'étiquettes
3	changement de pile de FEC

3.4.1.1 Sous TLV Données de multi chemins



Le sous TLB Données de multi chemins inclut des informations de multi chemins.

Type de multi chemins : Type du codage pour les informations de multi chemins.

Les types de multi chemins suivants sont définis dans le présent document :

Clé	Type	Informations de multi chemins
0	pas de multi chemins	vide (Longueur de multi chemin = 0)
2	adresse IP	adresses IP
4	gamme d'adresses IP	paires d'adresses haute/basse
8	gabarit binaire IP	préfixe d'adresse IP et gabarit binaire d'ensemble d'adresses
9	gabarit binaire d'ensemble d'étiquettes	préfixe d'étiquette et gabarit binaire

Le type 0 indique que tous les paquets seront transmis à partir de cette seule interface.
Les types 2, 4, 8, et 9 spécifient que les informations de multi chemins fournies vont servir à appliquer ce chemin.

Longueur de multi chemins : longueur en octets des informations de multi chemins.

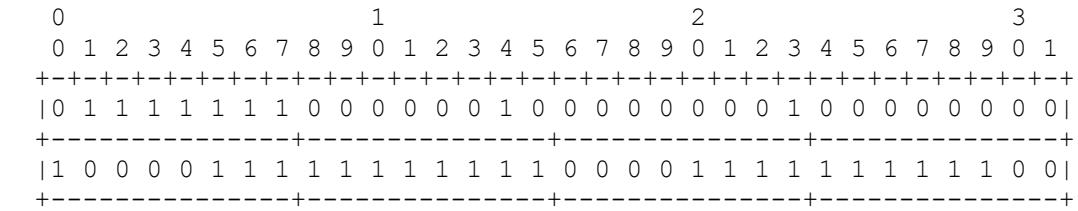
MBZ : DOIT être réglé à zéro en émission ; DOIT être ignoré à réception.

Informations de multi chemins : données codées de multi chemins (par exemple, adresse ou valeurs d'étiquette codées) conformément au type de multi chemins. Voir au paragraphe suivant les détails du codage.

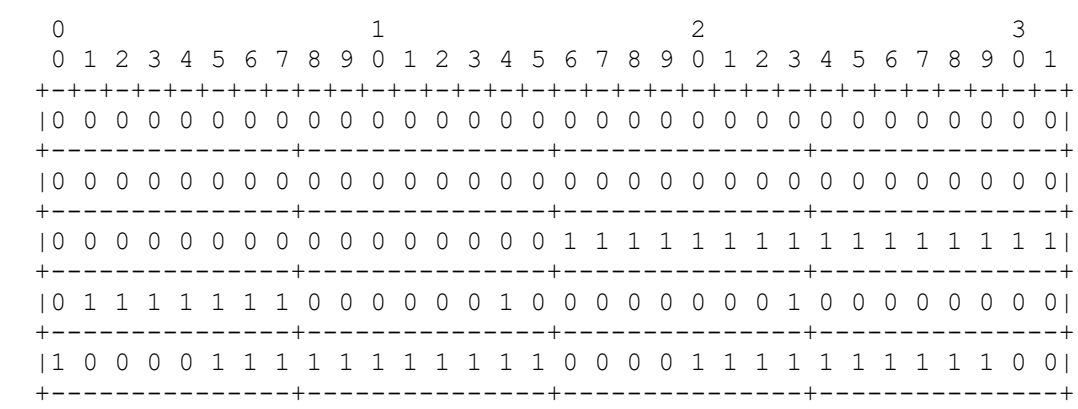
3.4.1.1.1 Codage des informations de multi chemins

Les informations de multi chemins codent les étiquettes ou adresses qui vont s'appliquer à ce chemin. Les informations de multi chemins dépendent du type de multi chemins. Le contenu du champ est montré dans le tableau ci-dessous. Les adresses IPv4 sont tirées de la gamme 127/8 ; les adresses IPv6 sont tirées de la gamme 0:0:0:0:0:FFFF:7F00:0/104. Les étiquettes sont traitées comme des nombres, c'est-à-dire, elles sont justifiées à droite dans le champ. Pour le type 4, les gammes indiquées par des paires d'adresses NE DOIVENT PAS se chevaucher et DOIVENT être en suite ascendante.

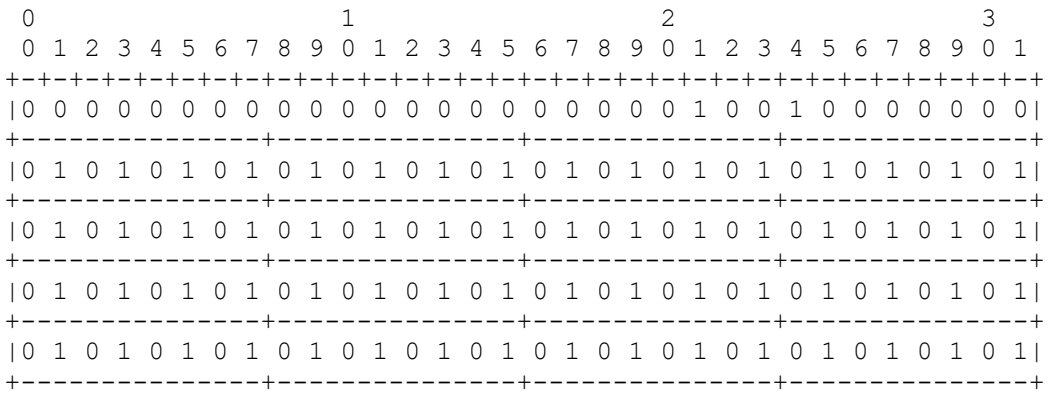
Le type 8 permet un codage plus dense des adresses IP. Le préfixe IP est formaté comme une adresse IP de base avec les bits de moindre poids du non préfixe réglés à zéro. La longueur maximum du préfixe est de 27. Suivant le préfixe se trouve un gabarit de longueur 2^(32 - longueur de préfixe) bits pour IPv4 et de 2^(128 - longueur de préfixe) bits pour IPv6. Chaque bit réglé à 1 représente une adresse valide. L'adresse est l'adresse IPv4 de base plus la position du bit dans le gabarit où les bits sont numérotés de gauche à droite en commençant par zéro. Par exemple, les adresses IPv4 127.2.1.0, 127.2.1.5-127.2.1.15, et 127.2.1.20-127.2.1.29 vont être codées comme suit :



Les mêmes adresses incorporées dans IPv6 vont être codées comme suit :



Le type 9 permet un codage plus dense des étiquettes. Le préfixe d'étiquette est formaté comme une valeur d'étiquette de base avec les bits de moindre poids qui ne sont pas du préfixe réglés à zéro. La longueur de préfixe maximum (incluant les zéros de tête dus au codage) est 27. À la suite du préfixe se trouve un gabarit de longueur 2^(32 - longueur de préfixe) bits. Chaque bit à un représente une étiquette valide. L'étiquette est celle de base plus la position du bit dans le gabarit où les bits sont numérotés de gauche à droite en commençant par zéro. Les valeurs d'étiquettes de tous les nombres impairs entre 1152 et 1279 vont être codés comme suit :



Si les informations de multi chemins reçues sont non nulles, les étiquettes et adresses IP DOIVENT être prises dans l'ensemble fourni. Si aucune de ces étiquettes ou adresses ne se transpose en une interface aval particulière, alors pour cette interface, le type DOIT être réglé à 0. Si les informations de multi chemins sont nulles (c'est-à-dire, Longueur de multi chemins = 0, ou pour les types 8 et 9, un gabarit tout de zéros) le type DOIT être réglé à 0.

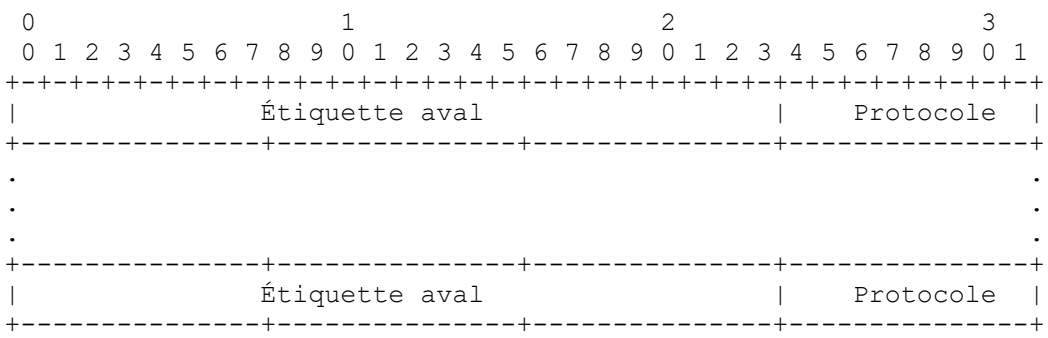
Par exemple, supposons que le LSR X au bond 10 a deux LSR aval, Y et Z, pour la FEC en question. Le X reçu pourrait retourner le type de multi chemins 4, avec les adresses IP haute/basse de 127.1.1.1->127.1.1.255 pour le LSR aval Y et 127.2.1.1->127.2.1.255 pour le LSR aval Z. L'extrémité de tête reflète ces informations au LSR Y. Y, qui a trois LSR aval, U, V, et W, calcule que 127.1.1.1->127.1.1.127 va aller à U et 127.1.1.128-> 127.1.1.255 à V. Y va alors répondre par trois TLV Transposition détaillée vers l'aval : à U, avec le type de multi chemins 4 (127.1.1.1->127.1.1.127) ; à V, avec le type de multi chemins 4 (127.1.1.127->127.1.1.255) ; et à W, avec le type de multi chemins 0.

Noter que le calcul des informations de multi chemins peut imposer une charge de traitement significative au receveur. Un receveur PEUT donc choisir de traiter un sous ensemble des préfixes reçus. L'envoyeur, à réception d'une réponse à une transposition détaillée vers l'aval avec des informations partielles, DEVRAIT supposer que les préfixes manquants dans la réponse ont été sautés par le receveur et PEUT redemander des informations sur eux dans une nouvelle demande d'écho.

Le codage des informations de multi chemins dans les scénarios où quelques LSR appliquent l'équilibrage de charge fondé sur l'entropie des étiquettes tandis que d'autres LSR ne le font pas (fondés sur IP) sera défini dans un autre document.

Le codage des informations de multi chemins dans les scénarios où les LSR ont ECMP de couche 2 sur des interfaces de groupe d'agrégation de liaisons (LAG, *Link Aggregation Group*) sera défini dans un autre document.

3.4.1.2 Sous TLV Pile d'étiquettes



Le sous TLV Pile d'étiquettes contient l'ensemble d'étiquettes dans la pile d'étiquettes comme il apparaîtrait si ce routeur

transmettait le paquet à travers cette interface. Toutes les étiquettes nulles implicites sont explicitement incluses. Le nombre de paires étiquettes/protocole présentes dans le sous TLV est déterminé sur la base de la longueur des données du sous TLV. Quand le TLV Transposition détaillée vers l'aval est envoyé dans la réponse d'écho, ce sous TLV DOIT être inclus.

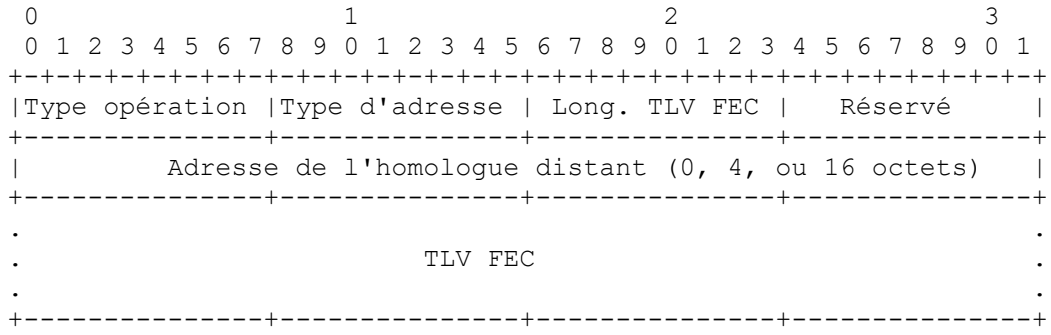
Étiquette aval : une étiquette aval fait 24 bits, dans le même format qu'une étiquette MPLS moins le champ TTL, c'est-à-dire que le bit de poids fort de l'étiquette est le bit 0, le bit de moindre poids est le bit 19, le champ TC [RFC5462] sont les bits 20-22, et S est le bit 23. Le routeur qui répond DEVRAIT remplir le champ TC et le bit S ; le LSR qui reçoit la réponse d'écho PEUT choisir de les ignorer.

Protocole : cela spécifie le protocole de distribution d'étiquettes pour l'étiquette aval. Les valeurs de Protocole sont prises dans le tableau suivant :

N° de protocole	Protocole de signalisation
0	inconnu
1	statique
2	BGP
3	LDP
4	RSVP-TE

3.4.1.3 Sous TLV Changement de pile de FEC

Un routeur DOIT inclure le sous TLV Changement de pile de FEC quand le nœud aval a dans la réponse d'écho une pile de FEC différente de la pile de FEC reçue dans la demande d'écho. Un ou plusieurs sous TLV Changement de pile de FEC PEUVENT être présents dans le TLV Transposition détaillée vers l'aval. Le format est montré ci-dessous.



Type opération : spécifie l'action associée au changement de pile de FEC. Les types d'opération suivants sont définis :

Type	Opération
1	Pousse *(PUSH)*
2	Saute *(POP)*

Type d'adresse : le type d'adresse indique le type d'adresse de l'homologue distant. Le type d'adresse est réglé à une des valeurs du tableau ci-dessous. La longueur de l'adresse de l'homologue est déterminée sur la base du type d'adresse. Le type d'adresse PEUT être différent du type d'adresse inclus dans le TLV Transposition détaillée vers l'aval. Cela peut arriver quand le LSP passe dans un tunnel d'une famille d'adresses différente. Le type d'adresse PEUT être réglé à Non spécifié si l'adresse de l'homologue est indisponible ou si le routeur de transit ne souhaite pas le fournir pour des raisons de sécurité ou des raisons administratives.

Type	Type d'adresse	Longueur d'adresse
0	non spécifié	0
1	IPv4	4
2	IPv6	16

Longueur TLV FEC : Longueur en octets du TLV FEC.

Réservé : ce champ est réservé pour une utilisation future et DOIT être réglé tout à zéro.

Adresse de l'homologue distant : l'adresse de l'homologue distant spécifie l'homologue distant qui est le prochain bond pour

la FEC actuellement tracée. Si le type d'opération est PUSH, l'adresse de l'homologue distant est l'adresse de l'homologue à partir duquel la FEC qui est poussée a été apprise. Si le type d'opération est POP, l'adresse de l'homologue distant PEUT être réglée à Non spécifié.

Pour les étiquettes allouées en amont [RFC5331], un type d'opération de POP va avoir une adresse d'homologue distant (le nœud amont qui a alloué l'étiquette) et elle DEVRAIT être incluse dans le sous TLV Changement de pile de FEC. L'adresse de l'homologue distant PEUT être réglée à Non spécifié si l'adresse doit être cachée.

TLV FEC : Le TLV FEC n'est présent que quand le champ Longueur de TLV FEC est non zéro. Le TLV FEC spécifie la FEC associée à l'opération de changement de la pile de FEC. Ce TLV PEUT être inclus quand le type d'opération est POP. Il DOIT être inclus quand le type d'opération est PUSH. Le TLV FEC contient exactement une FEC de la liste des FEC spécifiées au paragraphe 3.2. Une FEC Nulle PEUT être associée à une opération PUSH si le routeur qui répond souhaite cacher les détails de la FEC poussée.

Les règles de fonctionnement du sous TLV Changement de pile de FEC sont les suivantes :

a. Un sous TLV Changement de pile de FEC contenant une opération PUSH NE DOIT PAS être suivi par un sous TLV Changement de pile de FEC contenant une opération POP.
b. Une ou plusieurs opérations POP PEUVENT être suivies par une ou plusieurs opérations PUSH.
c. Un sous TLV Changement de pile de FEC DOIT être inclus par changement de pile de FEC. Par exemple, si deux étiquettes vont être poussées, un sous TLV Changement de pile de FEC DOIT alors être inclus pour chaque FEC.
d. Une opération de ligature de FEC (opération où une FEC se termine et une autre FEC commence) DOIT être effectuée en incluant un sous TLV Changement de pile de FEC de type POP suivi par un sous TLV Changement de pile de FEC de type PUSH.
e. Un TLV Transposition détaillée vers l'aval contenant seulement un sous TLV Changement de pile de FEC avec l'opération POP est équivalent à IS_EGRESS (code de retour 3, paragraphe 3.1) pour la FEC la plus externe dans la pile de FEC. Le routeur d'entrée qui effectue le LSP traceroute DOIT traiter ce cas comme un IS_EGRESS pour la FEC la plus externe.

3.4.2 Routeur et interface aval

La notion de "routeur aval" et "interface aval" devrait être expliquée. Considérons un LSR X. Si un paquet qui a été généré avec un TTL n > 1 est arrivé avec l'étiquette la plus externe L et le TTL = 1 au LSR X, X doit être capable de calculer quels LSR pourraient recevoir le paquet si il avait été généré avec le TTL = n + 1, sur quelle interface la demande arriverait et quelle pile d'étiquettes ces LSR verraient. (Il sort du domaine d'application du présent document de spécifier comment ce calcul est fait.) L'ensemble de ces LSR/interfaces constitue les routeurs/interfaces aval (et leurs étiquettes correspondantes) pour X par rapport à L. Chaque paire de routeur et interface aval exige qu'une transposition détaillée vers l'aval séparée soit ajoutée à la réponse.

Le cas où X est le LSR qui génère la demande d'écho est un cas particulier. X a besoin de se représenter quels LSR vont recevoir la demande d'écho MPLS pour une certaine pile de FEC que X génère avec le TTL = 1.

L'ensemble des routeurs aval à X peut être des chemins de remplacement (voir la discussion ci-dessous sur ECMP) ou des chemins simultanés (par exemple, pour la diffusion groupée MPLS). Dans le premier cas, les informations de multi chemins sont utilisées comme indication à l'envoyeur sur la façon dont il peut influencer le choix de ces remplaçants.

3.5 TLV Bourrage

La partie Valeur du TLV Bourrage *(Pad)* contient un nombre variable (≥ 1) d'octets. Le premier octet prend des valeurs du tableau ci-dessous ; tous les autres octets (si il en est) sont ignorés. Le receveur DEVRAIT vérifier que le TLV est reçu dans son intégralité, mais ignore par ailleurs le contenu de ce TLV, à part le premier octet.

Valeur	Signification
0	Réservé
1	Sortir le TLV Bourrage de la réponse
2	Copier le TLV Bourrage dans la réponse
3-250	Non alloué
251-254	Réservé pour utilisation expérimentale
255	Réservé

Le TLV Bourrage peut être ajouté à une demande d'écho pour créer un message d'une longueur spécifique dans les cas où des messages de diverses tailles sont nécessaires pour localiser des pannes. Le premier octet permet de contrôler l'inclusion de ce bourrage supplémentaire dans les réponses d'écho respectives.

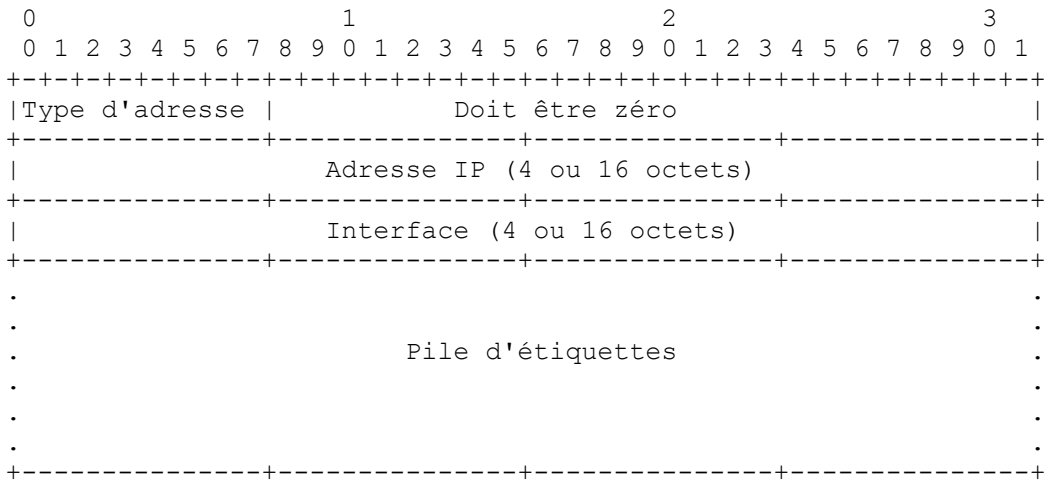
3.6 Numéro d'entreprise du fabricant

Les "numéros d'entreprise privée" [IANA-ENT] sont tenus par l'IANA. La longueur de ce TLV est toujours 4 ; la valeur est le code d'entreprise privée de la structure des informations de gestion (SMI, *Structure of Management Information*) dans l'ordre des octets du réseau, du fabricant avec une extension "Vendeur privé" à tous les champs de la partie fixe du message, auquel cas ce TLV DOIT être present. Si aucun des champs de la partie fixe du message n'a d'extension Vendeur privé, l'inclusion de ce TLV est FACULTATIVE. Les gammes de Vendeur privé pour les types de message, les modes de réponse, et les codes de retour ont été définis. Quand un de ceux-ci est utilisé, le TLV Numéro d'entreprise du fabricant DOIT être inclus dans le message.

3.7 Interface et pile d'étiquettes

Le TLV Interface et pile d'étiquette PEUT être inclus dans un message de réponse pour rapporter l'interface sur laquelle le message de demande a été reçu et la pile d'étiquettes qui était sur le paquet quand il a été reçu. Un seul de ces objets doit apparaître. Le but de cet objet est de permettre au routeur amont d'obtenir les informations exactes d'interface et de pile d'étiquettes comme elles apparaissent au LSR qui répond.

La longueur est K + 4*N octets ; N est le nombre d'étiquettes dans la pile d'étiquettes. Les valeurs pour K se trouvent dans la description du type d'adresse ci-dessous. Le champ Valeur de ce TLV a le format suivant :



Type d'adresse : il indique si l'interface est numérotée ou non numérotée. Il détermine aussi la longueur des champs Adresse IP et Interface. Le résultat total pour la partie initiale du TLV figure dans le tableau ci-dessous. Le type d'adresse est réglé à une des valeurs suivantes :

Type	Type d'adresse	Total (octets)
0	Réservé	4
1	IPv4 numéroté	12
2	IPv4 non numéroté	12
3	IPv6 numéroté	36
4	IPv6 non numéroté	24
5 à 250	Non alloué	
251 à 254	Réservé pour usage expérimental	
255	Réservé	

Adresse IP et interface : les adresses IPv4 et les indices d'interface sont codés sur 4 octets ; les adresses IPv6 sont codées sur 16 octets. Si l'interface sur laquelle le message de demande d'écho a été reçu est numérotée, le type d'adresse DOIT alors être réglé à IPv4 ou IPv6, l'adresse IP DOIT être réglée à l'identifiant de routeur du LSR ou à l'adresse de l'interface, et Interface DOIT être réglé à l'adresse de l'interface. Si l'interface n'est pas numérotée, le type d'adresse DOIT être IPv4 non numéroté ou IPv6 non numéroté, l'adresse IP DOIT être l'identifiant de routeur du LSR, et Interface DOIT être réglé à

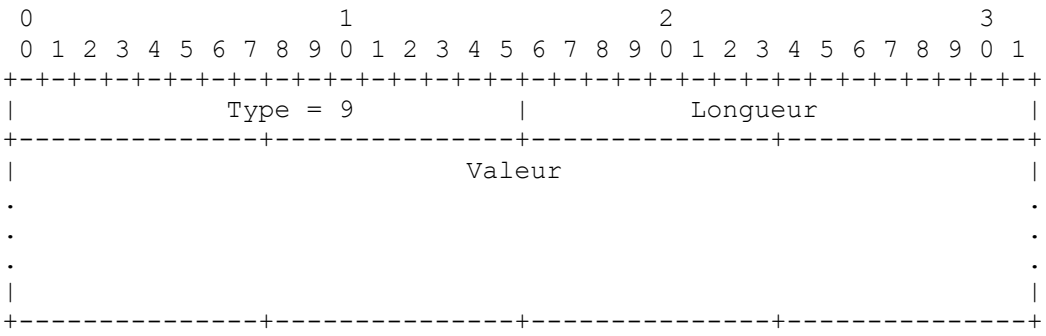
l'indice alloué à l'interface.

Pile d'étiquettes : c'est la pile d'étiquettes du message de demande d'écho reçu. Si des valeurs de TTL ont été changées par ce routeur, elles DEVRAIENT être restaurées.

3.8 TLV erronés

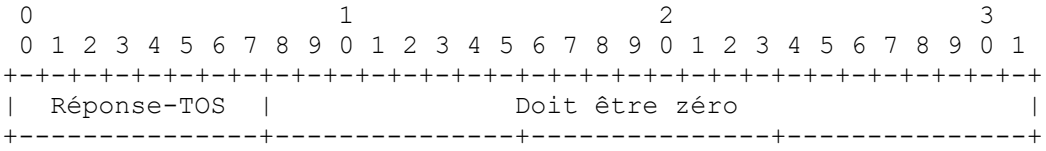
Le TLV suivant PEUT être inclus dans une réponse d'écho pour informer l'envoyeur d'une demande d'écho de TLV obligatoires non pris en charge par une mise en œuvre, ou trouvés erronés à l'analyse.

Le champ Valeur contient les TLV qui n'ont pas été compris, codés comme des sous TLV.



3.9 TLV de réponse à l'octet TOS

Ce TLV PEUT être utilisé par le générateur de la demande d'écho pour demander qu'une réponse d'écho soit envoyée avec l'octet Type de Service (TOS) de l'en-tête IP réglé à la valeur spécifiée dans le TLV. Ce TLV a une longueur de 4 avec le champ Valeur suivant.



4. Théorie du fonctionnement

Une demande d'écho MPLS est utilisée pour vérifier un certain LSP. Le LSP à vérifier est identifié par la "pile de FEC"; par exemple, si le LSP a été établi via LDP, et si une étiquette est transposée sur une adresse IP de sortie de 198.51.100.1, la pile de FEC contient un seul élément, à savoir une sous TLV Préfixe IPv4 de LDP d'une valeur de 198.51.100.1/32. Si le LSP vérifié est un LSP RSVP, la pile de FEC consiste en un seul élément qui capture la session RSVP et le gabarit d'envoyeur qui identifient le LSP de façon univoque.

Les piles de FEC peuvent être plus complexes. Par exemple, on peut souhaiter vérifier un préfixe IPv4 de VPN de 203.0.113.0/24 qui est tunnelé sur un LSP LDP avec la sortie 192.0.2.1. La pile de FEC va alors contenir deux sous TLV, celui du bas étant un préfixe IPv4, de VPN et celui du haut étant un préfixe IPv4 LDP. Si le tunnel (LDP) sous-jacent n'est pas connu, ou a été considéré comme non pertinent, la pile de FEC pourrait être un seul élément avec juste le sous TLV VPN IPv4.

Quand une demande d'écho MPLS est reçue, le receveur est supposé vérifier que le plan de contrôle et le plan des données sont tous deux en bonne santé (pour la pile de FEC qui fait l'objet du ping) et que les deux plans sont synchrones. Les procédures pour cela sont au paragraphe 4.4.

4.1 Traitement de multi chemins de coût égal (ECMP)

Les LSP n'ont pas besoin d'être de simples tunnels point à point. Fréquemment, un seul LSP peut avoir son origine à

plusieurs entrées et se terminer à plusieurs sorties ; ceci est très courant avec les LSP LDP. Les LSP pour une certaine FEC peuvent aussi avoir plusieurs "prochains bonds" aux LSR de transit. À une entrée, il peut aussi y avoir plusieurs LSP différents entre lesquels choisir pour arriver au point d'extrémité désiré. Finalement, les LSP peuvent avoir des chemins de secours, des chemins de détour, et autres chemins de remplacement à prendre si le LSP principal était défaillant.

Concernant ces deux derniers points, on suppose que le LSR à la source des demandes d'écho MPLS peut forcer la demande d'écho à passer dans le LSP désiré, de sorte que choisir parmi plusieurs LSP à l'entrée n'est pas un problème. Le problème de vérifier les diverses variantes de chemins de secours qui ne vont normalement pas être utilisés pour transmettre des données, sauf si le LSP principal est défaillant, ne sera pas traité ici.

Comme le LSP réel et le chemin qu'un certain paquet peut prendre peuvent n'être pas connus à priori, il est utile que les demandes d'écho MPLS puissent s'appliquer sur tous les chemins possibles. Cela, bien que désirable, peut n'être pas praticable parce que les algorithmes qu'un LSR donné utilise pour distribuer les paquets sur les chemins de remplacement peuvent être propriétaires.

Pour réaliser un certain degré de couverture des chemins de remplacement, on a une certaine latitude dans le choix de l'adresse IP de destination et de l'accès UDP de source pour une demande d'écho MPLS. Ceci n'est évidemment pas suffisant ; dans le cas de traceroute, plus de latitude est offerte au moyen des informations de multi chemins du TLV Transposition détaillée vers l'aval. C'est utilisé comme suit. Un LSR d'entrée envoie périodiquement un message LSP traceroute pour déterminer si il y a plusieurs chemins pour un certain LSP. Si oui, chaque bond va fournir des informations sur la façon dont chacun des chemins vers l'aval peut être utilisé. L'entrée peut alors envoyer des demandes d'écho MPLS qui utilisent ces chemins. Si plusieurs LSR de transit ont ECMP, l'entrée peut tenter de les composer pour appliquer tous les chemins possibles. Cependant, une couverture complète peut n'être pas possible.

4.2 Vérification des LSP utilisés pour porter des charges utiles MPLS

Pour détecter certaines coupures de LSP, il peut être nécessaire d'encapsuler un paquet de demande d'écho MPLS avec au moins une étiquette supplémentaire quand on vérifie les LSP qui sont utilisés pour porter des charges utiles MPLS (comme des LSP utilisés pour porter du trafic L2VPN et L3VPN. Par exemple, quand on vérifie des LSP LDP ou RSVP-TE, juste envoyer un paquet de demande d'écho MPLS peut ne pas détecter des instances où le routeur immédiatement en amont de la destination du ping de LSP peut transmettre avec succès la demande d'écho MPLS sur une interface non configurée pour porter des charges utiles MPLS à cause de l'utilisation du saut de l'avant dernier bond. Comme le routeur receveur n'a aucun moyen de s'assurer si le paquet IP a été envoyé non étiqueté ou étiqueté implicitement, l'ajout d'étiquettes calées par dessus la demande d'écho MPLS (en utilisant la FEC nulle) va empêcher un routeur de transmettre un tel paquet sur des interfaces non étiquetées.

4.3 Envoi d'une demande d'écho MPLS

Une demande d'écho MPLS est un paquet UDP. L'en-tête IP est réglé comme suit : l'adresse IP de source est une adresse acheminable de l'envoyeur ; l'adresse IP de destination est une adresse IPv4 (choisie au hasard) dans la gamme 127/8 ou une adresse IPv6 dans la gamme 0:0:0:0:0:FFFF:7F00:0/104. Le TTL IP est réglé à 1. L'accès UDP de source est choisi par l'envoyeur ; l'accès UDP de destination est réglé à 3503 (alloué par l'IANA pour les demandes d'écho MPLS). L'option IP Alerte de routeur de valeur 0x0 [RFC2113] pour IPv4 ou 69 [RFC7506] pour IPv6 DOIT être établie dans l'en-tête IP.

Une demande d'écho MPLS est envoyée avec une pile d'étiquettes correspondant à la pile de FEC qu'on vérifie. Noter que plus d'étiquettes pourraient être appliquées si, par exemple, le chemin normal pour la FEC du sommet de la pile est via un tunnel à ingénierie du trafic [RFC3209]. Si toutes les FEC de la pile correspondent à des étiquettes nulles implicites, la demande d'écho MPLS est considérée comme non étiquetée même si d'autres étiquettes vont être appliquées lors de l'envoi du paquet.

Si la demande d'écho est étiquetée, on PEUT (selon ce qui fait l'objet du ping) régler à 1 le TTL de l'étiquette la plus interne, pour empêcher la demande de ping d'aller plus loin qu'elle ne devrait. Des exemples de où cela DEVRAIT être fait incluent de faire un ping sur un préfixe de VPN IPv4 ou IPv6, sur un point d'extrémité de VPN L2, ou sur un pseudo filaire. Empêcher la demande de ping d'aller trop loin peut aussi être accompli en insérant une étiquette Alerte de routeur au dessus de cette étiquette ; cependant, cela peut conduire à l'effet collatéral non désiré que les demandes d'écho MPLS prennent un chemin de données différent de celui des données réelles. Pour plus d'informations sur la façon dont ces mécanismes peuvent être utilisés pour la vérification de connexité de pseudo filaire, voir les [RFC5085] et [RFC5885].

En mode "ping" (vérification de connexité de bout en bout) le TTL dans l'étiquette la plus externe est réglé à 255. En mode

"traceroute" (mode d'isolation de faute) le TTL est réglé successivement à 1, 2, et ainsi de suite.

L'envoyeur choisit une Bride d'envoyeur et un Numéro de séquence. Lors de l'envoi des demandes d'écho MPLS suivantes, l'envoyeur DEVRAIT incrémenter le numéro de séquence de 1. Cependant, un envoyeur PEUT choisir d'envoyer un groupe de demandes d'écho avec le même numéro de séquence pour améliorer les chances d'arrivée d'au moins un paquet avec ce numéro de séquence.

L'horodatage d'envoi est réglé à l'heure en format NTP à laquelle est envoyée la demande d'écho. L'horodatage de réception est réglé à zéro.

Une demande d'écho MPLS DOIT avoir un TLV Pile de FEC. Aussi, le mode de réponse doit être réglé au mode de réponse désiré ; le code et sous code de retour sont réglés à zéro. Dans le mode "traceroute", la demande d'écho DEVRAIT inclure un TLV Transposition détaillée vers l'aval.

4.4 Réception d'une demande d'écho MPLS

L'envoi d'une demande d'écho MPLS au plan de contrôle est déclenchée par une des exceptions de traitement de paquet suivantes : option d'alerte de routeur, expiration du TTL IP, expiration du TTL MPLS, étiquette d'alerte de routeur MPLS, ou l'adresse de destination dans la gamme d'adresses 127/8. Le plan de contrôle l'identifie de plus par l'accès de destination UDP 3503.

Pour les besoins de rapports, le bas de la pile est considéré comme étant à la profondeur de pile de 1. C'est pour établir une référence absolue pour le cas où la pile réelle aurait plus d'étiquettes qu'il n'y a de FEC dans la pile de FEC cible.

De plus, dans tous les codes de retour mentionnés dans le présent document, une profondeur de pile de 0 signifie "pas de valeur spécifiée". Cela permet la compatibilité avec les mises en œuvre existantes qui n'utilisent pas le champ Sous code de retour.

Un LSR X qui reçoit une demande d'écho MPLS la traite comme suit.

1. La santé générale du paquet est vérifiée. Si le paquet n'est pas bien formé, le LSR X DEVRAIT envoyer une réponse d'écho MPLS avec le code de retour réglé à "demande d'écho reçue mal formée" et le sous code réglé à zéro. Si il y a des TLV non marqués comme "Ignore" (c'est-à-dire, si le type de TLV est inférieur à 32768, voir la Section 3) que le LSR X ne comprend pas, le LSR X DEVRAIT envoyer un message MPLS "TLV non compris" (comme approprié) et régler le sous code à zéro. Dans le dernier cas, les TLV incompris (et eux seuls) sont inclus comme des sous TLV dans un TLV "TLV Erronés" dans la réponse. Les champs d'en-tête Bride d'envoyeur, Numéro de séquence, et Horodatage d'envoi ne sont pas examinés mais sont inclus dans le message réponse d'écho MPLS.

L'algorithme utilise les variables et identifiants suivants :

Interface-I : l'interface sur laquelle la demande d'écho MPLS a été reçue.

Pile-R : la pile d'étiquettes sur le paquet comme elle a été reçue.

Pile-D : la pile d'étiquettes portée dans le sous TLV "Pile d'étiquettes" dans le TLV Transposition détaillée vers l'aval (pas toujours présent).

Étiquette-L : l'étiquette provenant de la pile réelle examinée. N'exige pas d'initialisation.

Profondeur de pile d'étiquettes : profondeur de l'étiquette vérifiée. Initialisée au nombre d'étiquettes dans la pile d'étiquettes S reçue.

Profondeur de pile de FEC : profondeur de la FEC dans la pile de FEC cible qui devrait être utilisée pour vérifier l'étiquette réelle actuelle. N'exige pas d'initialisation.

Meilleur code de retour : contient le code de retour pour le paquet de réponse d'écho comme il est le mieux connu actuellement. Avec la progression de l'algorithme, ce code peut changer selon les résultats des autres vérifications effectuées.

Meilleur sous code de retour : similaire au Meilleur code de retour, mais pour le sous code de réponse d'écho.

État de FEC : valeur du résultat retourné par l'algorithme de vérification de FEC décrit au paragraphe 4.4.1.

/* Sauvegarder les informations de contexte reçues */

2. Si la demande d'écho est bonne, le LSR X mémorise l'interface sur laquelle l'écho a été reçu dans Interface-I, et la pile d'étiquettes avec laquelle il est venu dans Pile-R.

/* Le reste de l'algorithme fait une itération sur les étiquettes de Pile-R, vérifie la validité des valeurs d'étiquettes, rapporte les opérations de commutation d'étiquette associées (pour traceroute) vérifie la correspondance entre la Pile-R et la description de la pile de FEC cible dans le corps de la demande d'écho, et rapporte toutes les erreurs. */

/* L'algorithme itère comme suit. */

3. Validation d'étiquette :

Si Profondeur de pile d'étiquettes est 0 {

/* Le LSR doit rapporter qu'il est une extrémité de queue pour le LSP */

Régler Profondeur de pile d'étiquettes à 1, régler Étiquette-L à 3 (Nulle implicite).
Régler Meilleur code de retour à 3 ("Le routeur qui répond est une sortie pour la FEC à la profondeur de pile") régler Meilleur sous code de retour à la valeur de Profondeur de pile de FEC (1), et passer à l'étape 5 (Traitement de sortie).
}

/* Cette étape suppose qu'il y a toujours une entrée pour les valeurs d'étiquette bien connues */

Régler Étiquette-L à la valeur extraite de Pile-R à la profondeur de Profondeur de pile d'étiquettes. Chercher Étiquette-L dans la transposition d'étiquette entrante (ILM, *Incoming Label Map*) pour déterminer si l'étiquette a été allouée et l'opération qui lui est associée.

Si il n'y a pas d'entrée pour Étiquette-L {
/* Indique un problème temporaire ou permanent de synchronisation d'étiquettes, et le LSR doit rapporter une erreur */

Régler Meilleur code de retour à 11 ("Pas d'entrée d'étiquette à cette profondeur de pile") et Meilleur sous code de retour à Profondeur de pile d'étiquettes. Passer à l'étape 7 (Envoyer le paquet de réponse).
}

Autrement {
Reprendre l'opération d'étiquette associée de l'entrée de transmission d'étiquette de prochain bond (NHLFE, *Next Hop Label Forwarding Entry*) correspondante, et procéder à l'étape 4 (Vérification d'opération d'étiquette).
}

4. Vérification d'opération d'étiquette

Si l'opération d'étiquette est "Sauter et continuer le traitement" {
/* Inclut les cas d'étiquettes explicites nulles et d'alerte de routeur */
Itérer à la prochaine étiquette en décrémentant la profondeur de pile d'étiquettes, et revenir à l'étape 3 (Validation d'étiquette).
}

Si l'opération d'étiquette est "Échanger ou sauter et commuter sur la base de l'étiquette sautée" {
Régler Meilleur code de retour à 8 ("Étiquette commutée à la profondeur de pile") et Meilleur sous code de retour à Profondeur de pile d'étiquette pour rapporter le changement de transit.
Si un TLV Transposition détaillée vers l'aval est présent dans la demande d'écho reçue {
Si l'adresse IP dans le TLV est 127.0.0.1 ou 0::1 {
Régler Meilleur code de retour à 6 ("Indice d'interface amont inconnu"). Un TLV Interface et pile d'étiquettes DEVRAIT être inclus dans la réponse et rempli avec Interface-I et Pile-R.
}
Autrement {
Vérifier que l'adresse IP, l'adresse d'interface, et la pile d'étiquettes dans le TLV Transposition détaillée vers l'aval

correspondent à Interface-I et Pile-R. Si il y a une discordance, régler Meilleur code de retour à 5, "Discordance de transposition vers l'aval". Un TLV Interface et pile d'étiquette DEVRAIT être inclus dans la réponse et rempli sur la base de Interface-I et Pile-R. Passer à l'étape 7 (Envoi du paquet de réponse).
}
}

Pour chaque chemin ECMP aval disponible {
Restituer l'interface de sortie à partir de l'entrée de NHLFE.

/* Note : ce code de retour est établi même si Profondeur de pile d'étiquettes est un */

Si l'interface de sortie n'a pas la capacité MPLS {
Régler Meilleur code de retour au code de retour 9, "Étiquette commutée mais pas de transmission MPLS à la profondeur de pile" et régler Meilleur sous code de retour à Profondeur de pile d'étiquettes et passer à l'étape 7 (Envoi du paquet de réponse).
}

Si un TLV Transposition détaillée vers l'aval est présent {
Un TLV Transposition détaillée vers l'aval DEVRAIT être inclus dans la réponse d'écho (voir au paragraphe 3.4) rempli avec des informations sur le chemin ECMP actuel.
}
}

Si aucun TLV Transposition détaillée vers l'aval n'est présent, ou si l'adresse IP aval est réglée à l'adresse de diffusion groupée TOUS LES ROUTEURS, passer à l'étape 7 (Envoi du paquet de réponse).

Si le fanion "Valider la pile de FEC" n'est pas établi et si le LSR n'est pas configuré à effectuer la vérification de FEC par défaut, passer à l'étape 7 (Envoi du paquet de réponse).

/* Valider la pile de FEC cible dans la demande d'écho reçue. Déterminer d'abord la profondeur de pile de FEC à partir du TLV Transposition détaillée vers l'aval. Ceci se fait en parcourant la Pile-D (les étiquettes vers l'aval) à partir du bas, en décrémentant le numéro d'étiquette pour chaque étiquette nulle non implicite, tout en incrémentant la profondeur de pile de FEC pour chaque étiquette. Si le TLV Transposition détaillée vers l'aval contient une ou plusieurs étiquettes nulles implicites, la profondeur de pile de FEC peut être supérieure à la profondeur de pile d'étiquettes. Pour être cohérent avec les profondeurs de pile ci-dessus, le bas est considéré être l'entrée 1. */

Régler profondeur de pile de FEC à 0. Régler i à Profondeur de pile d'étiquettes.

Si (i > 0) faire {
++profondeur de pile de FEC.
Si Pile-D [profondeur de pile de FEC] != 3 (nulle implicite)
--i.
}

Si le nombre de FEC dans la pile de FEC est supérieur ou égal à la profondeur de pile de FEC {
Effectuer la procédure de vérification de FEC (voir au paragraphe 4.4.1).
Si État de FEC est 2, régler Meilleur code de retour à 10 ("La transposition pour cette FEC n'est pas l'étiquette donnée à la profondeur de pile").
Si le code de retour est 1, régler Meilleur code de retour à Code de retour de FEC et Meilleur sous code de retour à Profondeur de pile de FEC.
}
Passer à l'étape 7 (Envoi du paquet de réponse).
}

5. Traitement de sortie

/* Ces étapes sont effectuées par le LSR qui s'est identifié comme LSR d'extrémité de queue pour un LSP. */

Si la demande d'écho reçue ne contient pas de TLV Transposition détaillée vers l'aval, ou si l'adresse IP aval est réglée à 127.0.0.1 ou 0::1, passer à l'étape 6 (Validation de FEC de sortie).

Vérifier que l'adresse IP, l'adresse d'interface, et la pile d'étiquettes dans le TLV Transposition détaillée vers l'aval correspondent à Interface-I et Pile-R.
Sinon, régler Meilleur code de retour à 5, "Discordance de transposition vers l'aval".
Un TLV Interface reçue et Pile d'étiquettes DEVRAIT être créé pour le paquet réponse d'écho. Passer à l'étape 7 (Envoi du paquet de réponse).

6. Validation de la FEC de sortie :

/* C'est une boucle pour toutes les entrées dans la pile de FEC cible commençant par Profondeur de pile de FEC. */

Effectuer la vérification de FEC en suivant l'algorithme décrit au paragraphe 4.4.1 pour Étiquette-L et la FEC à la profondeur de pile de FEC.

Régler Meilleur code de retour à Code de FEC et Meilleur sous code de retour à la valeur dans Profondeur de pile de FEC.

Si État de FEC (résultat de la vérification) est 1, passer à l'étape 7 (Envoi du paquet de réponse).

/* Itérer la prochaine entrée de FEC */
++profondeur de pile de FEC.
Si profondeur de pile de FEC > nombre de FEC dans la pile de FEC,
passer à l'étape 7 (Envoi du paquet de réponse).

Si État de FEC est 0 {
++Profondeur de pile d'étiquettes.
Si Profondeur de pile d'étiquettes > nombre d'étiquettes dans Pile-R,
passer à l'étape 7 (Envoi du paquet de réponse).

Étiquette-L = étiquette extraite de Pile-R à la profondeur Profondeur de pile d'étiquettes.
Revenir à l'étape 6 (Validation de la FEC de sortie).
}

7. Envoi du paquet de réponse :

Envoi d'une réponse d'écho MPLS avec un code de retour de Meilleur code de retour et un sous code de retour de Meilleur sous code de retour. Inclure tous les TLV créés durant le processus ci-dessus. Les procédures pour l'envoi de la réponse d'écho sont au paragraphe 4.5.

4.4.1 Validation de FEC

/* Ce paragraphe décrit la validation d'une entrée de FEC dans la pile de FEC cible et accepte une FEC, Étiquette-L, et Interface-I. Si la FEC la plus externe de la pile de FEC cible est la FEC nulle, le nœud DOIT alors sauter complètement la validation de la FEC cible. C'est pour prendre en charge la dissimulation de la FEC, dans laquelle la FEC externe cachée peut être la FEC nulle. Autrement, l'algorithme effectue les étapes suivantes. */

1. Deux valeurs de retour, État de FEC et Code de retour de FEC, sont initialisées à 0.

2. Si la FEC est la FEC nulle {
Si Étiquette-L est Explicit_Null ou Router_Alert, retour.
Autrement {
Régler FEC-return-code à 10 ("La transposition pour cette FEC n'est pas l'étiquette donnée à la profondeur de pile").
Régler État de FEC à 1
Retour.
}
}

3. Vérifier la transposition d'étiquette de FEC qui décrit comment le trafic reçu sur le LSP est ensuite commuté ou quelle application lui est associée. Si il n'existe pas de transposition, régler Code de retour de FEC à 4, "Le routeur qui répond n'a pas de transposition pour la FEC à cette profondeur de pile". Régler État de FEC à 1. Retour.

4. Si la transposition d'étiquette pour la FEC est Nul implicite, régler État de FEC à 2 et passer à l'étape 5. Autrement, si la

transposition d'étiquette pour la FEC est Étiquette-L, passer à l'étape 5. Autrement, régler Code de retour de FEC à 10 ("La transposition pour cette FEC n'est pas l'étiquette donnée à cette profondeur de pile") régler État de FEC à 1, et retour.

5. C'est une vérification de protocole. Vérifier quel protocole va être utilisé pour annoncer la FEC. Si il peut être déterminé qu'aucun protocole associé à Interface-I ne va annoncer une FEC de ce type de FEC, régler Code de retour de FEC à 12 ("Protocole non associé à l'interface à cette profondeur de pile de FEC"). Régler État de FEC à 1.

6. Retour.

4.5 Envoi d'une réponse d'écho MPLS

Une réponse d'écho MPLS est un paquet UDP. Elle DOIT SEULEMENT être envoyée en réponse à une demande d'écho MPLS. L'adresse IP de source est une adresse acheminable de celui qui répond ; l'accès de source est l'accès UDP bien connu pour les ping de LSP. L'adresse IP de destination et l'accès UDP sont copiés de l'adresse IP de source et de l'accès UDP de la demande d'écho. Le TTL IP est réglé à 255. Si le mode de réponse dans la demande d'écho est "Réponse via un paquet UDP IPv4 avec Alerte de routeur", l'en-tête IP DOIT alors contenir l'option IP Alerte de routeur de valeur 0x0 [RFC2113] pour IPv4 ou 69 [RFC7506] pour IPv6. Si la réponse est envoyée sur un LSP, l'étiquette du sommet DOIT dans ce cas être l'étiquette Alerte de routeur (1) (voir la [RFC3032]).

Le format de la réponse d'écho est le même que celui de la demande d'écho. La bride d'envoyeur, le numéro de séquence, et l'horodatage sont copié de la demande d'écho ; l'horodatage de réception est réglé à l'heure à laquelle la demande d'écho a été reçue (noter que ces informations sont surtout utilsées si les horloges du demandeur et de celui qui répond sont synchronisées). Le TLV Pile de FEC provenant de la demande d'écho PEUT être copié dans la réponse.

Celui qui répond DOIT remplir le code et sous code de retour, comme déterminé au paragraphe précédent.

Si la demande d'écho contient un TLV Bourrage, celui qui répond DOIT interpréter le premier octet comme des instructions sur la façon de répondre.

Si le routeur qui répond est la destination de la FEC, des TLV Transposition détaillée vers l'aval NE DEVRAIENT PAS être alors inclus dans la réponse d'écho.

Si la demande d'écho contient un TLV Transposition détaillée vers l'aval, et si le routeur qui répond n'est pas la destination de la FEC, celui qui répond DEVRAIT calculer ses routeurs aval et les étiquettes correspondantes pour l'étiquette entrante et ajouter des TLV Transposition détaillée vers l'aval pour chacune des réponses d'écho qu'il renvoie. Un nœud qui répond devrait suivre les procédures définies au paragraphe 4.5.1 si il y a un changement de pile de FEC dû au LSP tunnelé. Si la pile de FEC change à cause d'une couture de LSP, il devrait suivre les procédures définies au paragraphe 4.5.2.

Si le TLV Transposition détaillée vers l'aval contient des informations de multi chemins qui exigent plus de traitement que ce que veut faire le routeur receveur, le routeur qui répond PEUT choisir de répondre avec seulement un sous ensemble des multi chemins contenus dans la transposition détaillée vers l'aval de demande d'écho. (Note : le générateur de la demande d'écho PEUT envoyer une autre demande d'écho avec les informations de multi chemins qui n'étaient pas incluses dans la réponse.)

Sauf dans le cas de réponse du mode 4, "Réponse via canal de contrôle de niveau application", les réponses d'écho sont toujours envoyées dans le contexte du réseau IP/MPLS.

4.5.1 Ajout d'un nouveau tunnel

Un nœud de transit sait quand la FEC tracée est sur le point d'entrer dans un tunnel à ce nœud. Donc, il sait quelle sera la nouvelle FEC externe. Tous les nœuds de transit qui sont le point d'origine d'un nouveau tunnel DEVRAIENT ajouter le sous TLV Changement de pile de FEC (paragraphe 3.4.1.3) au TLV Transposition détaillée vers l'aval dans la réponse d'écho. Le nœud de transit DEVRAIT ajouter un sous TLV Changement de pile de FEC de type d'opération PUSH, par nouveau tunnel généré au nœud de transit.

Un nœud de transit qui envoie un sous TLV Changement de pile de FEC vers l'aval dans la réponse d'écho DEVRAIT la remplir avec l'adresse de l'homologue distant, qui est l'homologue du LSP en cours qui est tracé. Si le nœud de transit ne sait pas l'adresse de l'homologue distant, il DOIT régler le type d'adresse à Non spécifié.

Le sous TLV Pile d'étiquettes DOIT contenir une étiquette supplémentaire par FEC poussée. L'étiquette DOIT être codée comme défini au paragraphe 3.4.1.2. La valeur de l'étiquette DOIT être celle utilisée pour commuter le trafic de données. Si le tunnel est un tuyau transparent pour le nœud, c'est-à-dire, si la trace de plan de données ne va pas expirer au milieu du nouveau tunnel, un sous TLV Changement de pile de FEC NE DEVRAIT PAS être alors ajouté, et le sous TLV Pile d'étiquettes NE DEVRAIT PAS contenir d'étiquette correspondant au tunnel caché.

Si le nœud de transit souhaite cacher la nature du tunnel à l'entrée de la demande d'écho, il PEUT alors vouloir envoyer des détails sur la FEC du nouveau tunnel à l'entrée. Dans ce cas, le nœud de transit DEVRAIT utiliser la FEC nulle. La réponse d'écho va alors contenir un sous TLV Changement de pile de FEC avec un type d'opération PUSH et une FEC nulle. La valeur de l'étiquette dans la FEC nulle DOIT être réglée à zéro. Le type d'adresse de l'homologue distant DOIT être réglé à Non spécifié. Le nœud de transit DEVRAIT ajouter un sous TLV Changement de pile de FEC de type d'opération PUSH, par nouveau tunnel généré au nœud de transit. Le sous TLV Pile d'étiquettes DOIT contenir une étiquette supplémentaire par FEC poussée. La valeur de l'étiquette DOIT être la valeur utilisée pour commuter le trafic de données.

4.5.2 Transition entre tunnels

Un nœud de transit qui rattache deux LSP DEVRAIT inclure deux sous TLV Changement de pile de FEC. Un avec une opération POP pour la vieille FEC (entrée) et un avec l'opération PUSH pour la nouvelle FEC (sortie). Le nœud qui répond DEVRAIT régler le code de retour à "Étiquette commutée avec changement de FEC" pour indiquer le changement dans la FEC tracée.

Si le nœud qui répond souhaite effectuer la dissimulation de la FEC, il DEVRAIT répondre avec deux sous TLV Changement de pile de FEC, un POP suivi par un PUSH. L'opération POP PEUT exclure le TLV FEC (en réglant la longueur du TLV FEC à 0) ou régler le TLV FEC à contenir la FEC LDP. L'opération PUSH DEVRAIT avoir le TLV FEC contenant la FEC nulle. Le code de retour DEVRAIT être réglé à "Étiquette commutée avec changement de FEC".

Si le nœud qui répond souhaite effectuer la dissimulation de la FEC, il PEUT choisir de n'envoyer aucun sous TLV Changement de pile de FEC dans la réponse d'écho si le nombre d'étiquettes ne change pas pour le nœud aval et si le type de FEC ne change pas non plus (FEC nulle). Dans ce cas, le nœud qui répond NE DOIT PAS régler le code de retour à "Étiquette commutée avec changement de FEC".

4.6 Réception d'une réponse d'écho MPLS

Un LSR X devrait recevoir une réponse d'écho MPLS seulement en réponse à une demande d'écho MPLS qui est envoyée. Donc, à réception d'une réponse d'écho MPLS, X devrait analyser le paquet pour s'assurer qu'il est bien formé, puis tenter de faire correspondre la réponse d'écho à une demande d'écho qu'il a envoyée précédemment, en utilisant l'accès UDP de destination et la bride d'envoyeur. Si aucune correspondance n'est trouvée, alors X élimine la réponse d'écho ; autrement, il vérifie le numéro de séquence pour voir si il correspond.

Si la réponse d'écho contient des transpositions détaillées vers l'aval, et si X souhaite de plus faire un traceroute, il DEVRAIT copier la ou les transpositions détaillées vers l'aval dans la ou les prochaines demande d'écho (avec le TTL incrémenté de un).

Si un ou plusieurs sous TLV Changement de pile de FEC sont reçus dans la réponse d'écho MPLS, le nœud d'entrée DEVRAIT les traiter et effectuer les validations.

Les changements de pile de FEC sont associés à un voisin vers l'aval et le long d'un chemin particulier du LSP. Par conséquent, l'entrée va devoir tenir une pile de FEC par chemin tracé (en cas de multi chemins). Tous les changements à la pile de FEC résultant du traitement de sous TLV Changement de pile de FEC devraient être appliqués seulement pour le chemin vers un certain voisin aval. L'algorithme suivant devrait être utilisé pour traiter les sous TLV Changement de pile de FEC :

```
PUSH vu = FAUX
profondeur de pile de FEC = profondeur actuelle de la pile de FEC tracée
pile de FEC sauvegardée = pile de FEC actuelle

quand (sous TLV = obtenir le prochain sous TLV(TLV transposition détaillée vers l'aval))
   si (sous TLV == NUL) rompre
```

```
   si (type de sous TLV == Changement de pile de FEC) {
      si (opération de sous TLV == POP) {
         si (PUSH vu) {
            Éliminer la réponse d'écho
            pile de FEC actuelle = pile de FEC sauvegardée
            retour
         }
         si (profondeur de pile de FEC == 0) {
            Éliminer la réponse d'écho
            pile de FEC actuelle = pile de FEC sauvegardée
            retour
         }
         Sauter la FEC de la pile de FEC tracée
         profondeur de pile de FEC -- ;
      }
      si (opération de sous TLV == PUSH) {
         PUSH vu = 1
         Pousser la FEC sur la pile de FEC tracée
         profondeur de pile de FEC++ ;
      }
   }
}

 Si (profondeur de pile de FEC == 0) {
    Éliminer la réponse d'écho
    pile de FEC actuelle = pile de FEC sauvegardée
    retour
}
```

La prochaine demande d'écho MPLS le long du même chemin devrait utiliser la pile de FEC modifiée obtenue après le traitement des sous TLV Changement de pile de FEC. Une FEC non nulle garantit que la prochaine demande d'écho le long du même chemin aura le TLV Transposition détaillée vers l'aval validé à l'égard des discordances d'adresse IP, d'adresse d'interface, et de pile d'étiquettes.

Si le sommet de la pile de FEC est une FEC nulle et si la réponse d'écho MPLS ne contient aucun sous TLV Changement de pile de FEC, cela ne signifie alors pas nécessairement que le LSP n'a pas commencé à traverser un tunnel différent. Il se pourrait que le LSP associé à la FEC nulle se soit terminée à un nœud de transit, et qu'en même temps, un nouveau LSP ait commencé au même nœud de transit. La FEC nulle serait alors associée au nouveau LSP (et l'entrée n'a aucun moyen de le savoir). Donc, il n'est pas possible de construire une topologie hiérarchique de LSP précise si un traceroute contient des FEC nulles.

Une réponse d'un nœud aval avec le code de retour 3 peut ne pas nécessairement être pour la FEC en cours de traçage. Elle pourrait être pour une des nouvelles FEC qui ont été ajoutées. À réception d'une réponse IS_EGRESS, le LSP d'entrée devrait vérifier si la profondeur de la FEC cible envoyée au nœud qui a juste répondu était la même que la profondeur de la FEC qui est en cours de traçage. Si cela n'est pas, il devrait alors sauter une entrée de la pile de FEC cible et renvoyer la demande avec le même TTL (que celui envoyé précédemment). Le processus de saut d'une FEC est à répéter jusqu'à ce que le LSP d'entrée reçoive une réponse non -IS_EGRESS ou jusqu'à ce que toutes les FEC supplémentaires ajoutées à la pile de FEC aient déjà été sautées. En utilisant une réponse IS_EGRESS, une entrée peut construire une carte de la structure hiérarchique de LSP traversée par une FEC donnée.

Quand le code de retour de la réponse d'écho MPLS est "Étiquette commuté avec changement de FEC", le nœud d'entrée DEVRAIT manipuler la pile de FEC conformément aux sous TLV Changement de pile de FEC contenus dans le TLV Transposition détaillée vers l'aval. Un nœud de transit peut utiliser ce code de retour pour coller des LSP et pour des LSP hiérarchiques. Dans le cas de ECMP ou P2MP, il pourrait y avoir plusieurs chemins et TLV Transposition détaillée vers l'aval avec des codes de retour différents (voir au paragraphe 3.1, Note 2). Le nœud d'entrée devrait construire la topologie sur la base du code de retour par chemin ECMP/branche P2MP.

4.7 Problème des préfixes IPv4 et IPv6 de VPN

Normalement, un ping de LSP pour un Préfixe de VPN IPv4 ou un préfixe de VPN IPv6 est envoyé avec une pile

d'étiquettes de profondeur supérieure à 1, avec l'étiquette la plus interne d'un TTL de 1. C'est pour terminer le ping au PE *(côté fournisseur)* de sortie, avant qu'il soit envoyé à l'appareil du consommateur. Cependant, dans certaines circonstances, la pile d'étiquettes peut se réduire à une seule étiquette avant que le ping touche le PE de sortie ; il va en résulter que le ping se termine de façon prématurée. Un tel scénario est celui d'un VPN de transporteur d'un transporteur multi AS.

Pour contourner ce problème, une approche est que le LSR qui reçoit un tel ping réalise que le ping s'est terminé de façon prématurée et qu'il renvoie le code de retour 13. Dans ce cas, le LSR initiateur peut réessayer le ping après avoir incrémenté le TTL sur l'étiquette de VPN. De cette façon, le LSR d'entrée va essayer à la suite les valeurs de TTL jusqu'à ce qu'il en trouve une qui permette au ping de VPN d'atteindre le PE de sortie.

4.8 Routeurs non conformes

Si la sortie pour la pile de FEC qui fait l'objet du ping ne prend pas en charge le ping de LSP, aucune réponse ne va être envoyée, résultant en un possible "faux négatif". En mode "traceroute", si un LSR de transit ne prend pas en charge le ping de LSP, aucune réponse ne suivra de la part de ce LSR pour un TTL, disons de n. Le LSR qui génère la demande d'écho DEVRAIT essayer d'envoyer la demande d'écho avec un TTL = n+1, n+2, ..., n+k pour sonder les LSR plus loin sur le chemin. Dans ce cas, la demande d'écho pour un TTL > n DEVRAIT être envoyée avec le TLV Transposition détaillée vers l'aval et le champ "Adresse IP vers l'aval" réglée à l'adresse de diffusion groupée TOUS ROUTEURs jusqu'à ce qu'une réponse soit reçue avec un TLV Transposition détaillée vers l'aval. Le TLV Pile d'étiquettes PEUT être omis du TLV Transposition détaillée vers l'aval. De plus, le fanion "Valider la pile de FEC" NE DEVRAIT PAS être établi jusqu'à la réception d'un paquet de réponse d'écho avec un TLV Transposition détaillée vers l'aval.

5. Considérations sur la sécurité

Globalement, les besoins de sécurité pour le ping de LSP sont similaires à ceux du ping ICMP.

Il y a au moins trois approches pour attaquer des LSR en utilisant les mécanismes définis ici. L'une d'elles est l'attaque de déni de service (DoS) en envoyant des demandes/réponses d'écho MPLS aux LSR et augmentant par là leur charge de travail. La seconde est de troubler l'état du plan de données MPLS en usurpant, capturant, répétant ou altérant par ailleurs les demandes et réponses d'écho MPLS. La troisième est une source non autorisée qui utilise un ping de LSP pour obtenir des informations sur le réseau.

Pour éviter de potentielles attaques de DoS, il est RECOMMANDÉ que les mises en œuvre régulent le trafic de ping de LSP qui arrive au plan de contrôle. Un limiteur de taux DEVRAIT être appliqué à l'accès UDP bien connu défini au paragraphe 6.1.

Les attaques non sophistiquées de répétition et d'usurpation d'identité qui impliquent de contrefaire ou répéter des messages de réponse d'écho MPLS ont peu de chances d'être efficaces. Ces réponses devraient correspondre à la bride d'envoyeur et au numéro de séquence d'un message en cours de demande d'écho MPLS. Une réponse non correspondante va être éliminée lorsque la séquence est passée, et donc une usurpation a seulement une petite fenêtre d'opportunité. Cependant, pour fournir une plus forte défense, une mise en œuvre PEUT aussi valider l'horodatage d'envoi en exigeant une correspondance exacte sur ce champ.

Pour protéger contre les sources non autorisées qui utilisent le message de demande d'écho MPLS pour obtenir des informations sur le réseau, il est RECOMMANDÉ que les mises en œuvre fournissent un moyen de vérifier les adresses de source des messages de demande d'écho MPLS par rapport à une liste d'accès avant d'accepter le message.

On ne sait pas clairement comment empêcher la capture (non livraison) des demandes ou réponses d'écho ; cependant, si ces messages sont effectivement capturés, le ping de LSP va rapporter que le plan des données ne fonctionne pas comme il devrait.

Il ne semble pas vital (pour l'instant) de sécuriser les données portées dans les demandes et réponses d'écho MPLS, bien que la connaissance de l'état du plan des données MPLS puisse être considérée comme confidentielle par certains. Les mises en œuvre DEVRAIENT cependant fournir un moyen de filtrer les adresses auxquelles les messages de réponse d'écho peuvent être envoyées.

La partie valeur du TLV Bourrage contient un nombre variable d'octets. À l'exception du premier octet, ce contenu, si il en est, est ignoré à réception, et peut donc servir de canal clandestin.

Quand le ping de LSP MPLS est utilisé au sein d'un domaine administratif, un déploiement peut augmenter la sécurité en utilisant le filtrage en bordure des paquets entrants de ping de LSP ainsi que des paquets sortants de ping de LSP.

Bien que le présent document fasse une utilisation particulière des adresses 127/8, elles ne sont utilisées qu'en conjonction avec l'accès UDP 3503. De plus, ces paquets ne sont traités que par les routeurs. Tous les autres hôtes DOIVENT traiter tous les paquets avec une adresse de destination dans la gamme 127/8 conformément à la RFC 1122. Tout paquet reçu par un routeur avec une adresse de destination dans la gamme 127/8 sans un accès de destination UDP de 3503 DOIT être traité conformément à la RFC 1812. En particulier, le comportement par défaut est de traiter les paquets destinés à une adresse 127/8 comme des "martiens" *(voir ce terme dans la RFC4949)*.

Si un opérateur réseau veut empêcher le traçage à l'intérieur d'un tunnel, il peut utiliser le modèle "Pipe" [RFC3443], c'est-à-dire, cacher le tunnel MPLS externe en ne propageant pas le TTL MPLS dans le tunnel externe (au début du tunnel externe). En faisant ainsi, les paquets LSP traceroute ne vont pas expirer dans le tunnel externe, et celui-ci ne sera pas "tracé".

Si on ne veut pas exposer les détails du nouveau LSP externe, la FEC nulle peut alors être utilisée pour cacher ces détails. Utiliser la FEC nulle assure que la trace progresse sans faux négatifs et que tous les nœuds de transit (du nouveau tunnel externe) effectuent des validations minimales sur les demandes d'écho MPLS reçues.

6. Considérations relatives à l'IANA

6.1 Numéro d'accès TCP et UDP

Le numéro d'accès TCP et UDP 3503 a été alloué par l'IANA pour les demandes et réponses d'écho de LSP.

6.2 Paramètres de ping de LSP MPLS

L'IANA tient le registre des "Paramètres de ping des chemins d'étiquettes commutées (LSP) de commutation d'étiquettes multi protocoles (MPLS)" à < http://www.iana.org/assignments/mpls-lsp-ping-parameters >.

Les paragraphes qui suivent détaillent les espaces de noms gérés par l'IANA. Pour certains de ces espaces de noms, l'espace est divisé en gammes d'allocation ; les termes suivants sont utilisés pour décrire les procédures par lesquelles l'IANA alloue les valeurs : "Action de normalisation" (comme défini dans la [RFC5226]), "Spécification exigée", et "Usage privé de fabricant".

Les valeurs des gammes "Spécification exigée" DOIVENT être enregistrées auprès de l'IANA. La demande DOIT être faite via une RFC qui décrit le format et les procédures pour utiliser le codet ; l'allocation réelle est faite durant les actions de l'IANA pour la RFC.

Les valeurs des gammes "Usage privé de fabricant" NE DOIVENT PAS être enregistrées auprès de l'IANA ; cependant, le message DOIT contenir un code d'entreprise enregistré auprès de l'IANA dans les SMI de numéros d'entreprise privées de gestion de réseau . Pour chaque espace de noms qui a une gamme de fabricant privé, il doit être spécifié où se trouve exactement le SMI de numéros d'entreprise privée ; voir ci-dessous des exemples. De cette façon, plusieurs entreprises (fabricants) peuvent utiliser le même codet sans crainte de collision.

6.2.1 Types de message, modes de réponse, codes de retour

L'IANA a créé et va tenir des registres pour les types de message, les modes de réponse, et les codes de retour. Chacun d'eux peut prendre des valeurs dans la gamme de 0 à 255. Les allocations dans la gamme de 0 à 191 sont via "action de normalisation" ; les allocations dans la gamme de 192 à 251 sont faites via "Spécification exigée" ; les valeurs dans la gamme de 252 à 255 sont pour utilisation de fabricant privé et NE DOIVENT PAS être allouées.

Si un de ces champs tombe dans la gamme de fabricant privé, un TLV Numéro d'entreprise de fabricant de niveau supérieur DOIT être présent dans le message.

Les types de message définis dans le présent document sont :

Valeur	Signification
1	demande d'écho MPLS
2	Réponse d'écho MPLS

Les modes de réponse définis dans le présent document sont :

Valeur	Signification
1	Ne pas répondre
2	Réponse via un paquet UDP IPv4/IPv6
3	Réponse via un paquet UDP IPv4/IPv6 UDP avec Alerte de routeur
4	Réponse via un canal de contrôle de niveau application

La liste des codes de retour définis dans le présent document figure au paragraphe 3.1.
L'IANA a mis à jour la référence au présent document pour chacune de ces valeurs.

6.2.2 TLV

L'IANA a créé et tient un registre pour le champ Type des TLV de niveau supérieur ainsi que pour tous les sous TLV associés. Noter que la signification d'un sous TLV est dans la portée du TLV. Les espaces de nombres pour les sous TLV des divers TLV sont indépendants.

La gamme valide pour les TLV et les sous TLV est de 0 à 65 535. Les allocations dans les gammes de 0 à 16 383 et de 32 768 à 49 161 sont faites via "Action de normalisation" comme défini dans la [RFC5226]; les allocations dans les gammes de 16 384 à 31 743 et 49 162 à 64 511 sont faites via "Spécification exigée" ; les valeurs dans les gammes de 31 744 à 32 767 et de 64 512 à 65 535 sont pour utilisation privée de fabricant et NE DOIVENT PAS être allouées.

Si un TLV ou sous TLV a un type qui tombe dans la gamme pour utilisation privée de fabricant, la longueur DOIT être d'au moins 4, et les quatre premiers octets DOIVENT être le numéro d'entreprise privée de SMI de ce fabricant, dans l'ordre des octets du réseau. Le reste du champ Valeur est l'affaire privée du fabricant.

Les TLV et sous TLV définis dans le présent document sont :

Type	Sous type	Champ Valeur
1		Pile de FEC cible
	1	Préfixe LDP IPv4
	2	Préfixe LDP IPv6
	3	LSP RSVP IPv4
	4	LSP RSVP IPv6
	5	Non alloué
	6	Préfixe VPN IPv4
	7	Préfixe VPN IPv6
	8	Point d'extrémité de VPN de couche 2
	9	"FEC 128" Pseudo filaire - IPv4 (déconseillé)
	10	"FEC 128" Pseudo filaire - IPv4
	11	"FEC 129" Pseudo filaire - IPv4
	12	Préfixe IPv4 étiqueté BGP
	13	Préfixe IPv6 étiqueté BGP
	14	Préfixe IPv4 générique
	15	Préfixe IPv6 générique
	16	FEC nulle
	24	"FEC 128" Pseudo filaire - IPv6
	25	"FEC 129" Pseudo filaire - IPv6
2		Transposition vers l'aval (déconseillé)
3		Bourrage
4		Non alloué
5		Numéro d'entreprise de fabricant
6		Non alloué
7		Interface et pile d'étiquette
8		Non alloué
9		TLV erronés

toute valeur	TLV non compris
10	Octet TOS de réponse
20	Transposition détaillée vers l'aval

L'IANA a mis à jour la référence au présent document pour chacune de ces valeurs.

6.2.3 Fanions globaux

L'IANA a créé un sous registre "Global Flags" *(fanions globaux)* du registre des "Paramètres de ping des chemins d'étiquettes commutées (LSP) de la commutation d'étiquettes multi protocoles (MPLS)".
Ce registre retrace les allocations de 16 fanions dans le champ Fanions globaux du message demande d'écho de ping de LSP MPLS. Les fanions sont numérotés de 0 (bit de poids fort, transmis en premier) à 15.
Les nouvelles entrées seront allouées par action de normalisation.
Les entrées initiales du registre sont :

N° de bit	Nom	Référence
15	Fanion V	[RFC8029]
14	Fanion T	[RFC6425]
13	Fanion R	[RFC6426]
12 à 0	non alloué	[RFC8029]

6.2.4 Type d'adresse Transposition détaillée vers l'aval

Le présent document étend la RFC 4379 en définissant un nouveau type d'adresse à utiliser avec les TLV Transposition vers l'aval et transposition détaillée vers l'aval. L'IANA a établi un registre pour les types d'adresses à utiliser avec les TLV Transposition vers l'aval et Transposition détaillée vers l'aval, qui alloue initialement les valeurs suivantes :

N° de type	Type d'adresse	K octets	Référence
1	IPv4 numéroté	16	[RFC8029]
2	IPv4 non numéroté	16	[RFC8029]
3	IPv6 numéroté	40	[RFC8029]
4	IPv6 non numéroté	28	[RFC8029]
5	Non IP	12	[RFC6426]

Registre de type d'adresse Transposition détaillée vers l'aval : Comme le champ est dans ce cas de 8 bits, la politique d'allocation pour ce registre est "Action de normalisation".

6.2.5 Fanions DS

Le présent document définit les TLV Transposition vers l'aval (DSMAP) et Transposition détaillée vers l'aval (DDMAP) qui ont respectivement les types 2 et 20, alloués dans le sous registre "TLV" du registre "Paramètres de ping des chemins d'étiquettes commutées (LSP) de la commutation d'étiquettes multi protocoles (MPLS)".

DSMAP a été déconseillé au profit de DDMAP, mais les deux TLV partagent le champ Fanions DS.
L'IANA a créé et tient maintenant un registre intitulé "Fanions DS".
La politique d'enregistrement pour ce registre est Action de normalisation [RFC5226].
L'IANA a fait les allocations suivantes :

Numéro de bit	Nom	Référence
7	N : Traiter comme paquet non IP	RFC8029]
6	I : Demande d'objet Interface et Pile d'étiquettes	[RFC8029]
5	E : indicateur de poussée ELI/EL	[RFC8012]
4	L : indicateur d'équilibrage de charge fondé sur l'étiquette	[RFC8012]
3-0	Non alloué	

6.2.6 Types de multi chemins

L'IANA a créé et tient maintenant un registre intitulé "Types de multi chemins". La politique d'enregistrement [RFC5226] pour ce registre est Action de normalisation. L'IANA a fait les allocations suivantes :

Valeur	Signification	Référence
0	pas de multi chemins	[RFC8029]
1	Non alloué	
2	adresse IP	[RFC8029]
3	Non alloué	
4	gamme d'adresses IP	[RFC8029]
5-7	Non alloué	
8	ensemble d'adresses IP à gabarit binaire	RFC8029]
9	ensemble d'étiquettes à gabarit binaire	[RFC8029]
10	IP et ensemble d'étiquettes	[RFC8012]
11-250	Non alloué	
251-254	Réservé pour utilisation expérimentale	[RFC8029]
255	Réservé	[RFC8029]

6.2.7 Type de bourrage

L'IANA a créé et tient maintenant un registre intitulé "Types de bourrage".
La politique d'enregistrement [RFC5226] pour ce registre est Action de normalisation.
L'IANA a fait les allocations initiales suivantes :

Nom de registre : Types de bourrage

Valeur	Signification	Référence
0	Réservé	[RFC8029]
1	Éliminer le TLV Bourrage de la réponse	[RFC8029]
2	Copier le TLV Bourrage dans la réponse	[RFC8029]
3-250	Non alloué	
251-254	Utilisation expérimentale	[RFC8029]
255	Réservé	[RFC8029]

6.2.8 Type d'adresse Interface et pile d'étiquettes

L'IANA a créé et tient maintenant un registre intitulé "Type d'adresse Interface et pile d'étiquettes".
La politique d'enregistrement [RFC5226] pour ce registre est Action de normalisation.
L'IANA a fait les allocations initiales suivantes :
Nom de registre : Type d'adresse Interface et pile d'étiquettes

Valeur	Signification	Référence
0	Réservé	[RFC8029]
1	IPv4 numéroté	[RFC8029]
2	IPv4 non numéroté	[RFC8029]
3	IPv6 numéroté	[RFC8029]
4	IPv6 non numéroté	[RFC8029]
5-250	Non alloué	
251-254	Utilisation expérimentale	[RFC8029]
255	Réservé	[RFC8029]

6.3 Registre d'adresses IPv4 d'utilisation particulière

L'IANA a mis à jour la référence dans la Note 1 du registre IANA "Adresses IPv4 d'utilisation particulière" [IANA-SPECIAL-IPv4] pour pointer sur le présent document.

7. Références

7.1 Références normatives

[RFC1122] R. Braden, "Exigences pour les hôtes Internet – couches de communication", STD 3, octobre 1989. *(MàJ par* RFC6633, 8029)

[RFC1812] F. Baker, "Exigences pour les routeurs IP version 4", juin 1995. *(MàJ par les* RFC2644, RFC6633)

[RFC2113] D. Katz, "Option d'alerte de routeur IP", février 1997. (*MàJ par* RFC5350, RFC6398) *(P.S.)*

[RFC2119] S. Bradner, "Mots clés à utiliser dans les RFC pour indiquer les niveaux d'exigence", BCP 14, mars 1997. *(MàJ par* RFC8174)

[RFC3032] E. Rosen et autres, "Codage de pile d'étiquettes MPLS", janvier 2001.

[RFC4271] Y. Rekhter, T. Li et S. Hares, "Protocole de routeur frontière version 4 (BGP-4)", janvier 2006. *(D.S.) (MàJ par* RFC6608, RFC8212)

[RFC4379] K. Kompella et G. Swallow, "Détection des défaillances de plan des données en commutation d'étiquettes multi protocole (MPLS)", février 2006. (*MàJ par la* RFC6424 *; Rendue obsolète par* RFC8029) *(P.S.)*

[RFC5226] T. Narten et H. Alvestrand, "Lignes directrices pour la rédaction d'une section Considérations relatives à l'IANA dans les RFC", BCP 26, mai 2008. (*Remplace* RFC2434 *; remplacée par* RFC8126)

[RFC5905] D. Mills, J. Martin, J. Burbank, W. Kasch, "Protocole de l'heure du réseau version 4 (NTPv4) : Spécification du protocole et des algorithmes ", juin 2010. (*Remplace* RFC1305, RFC4330) *(P. S ; MàJ par* RFC7822, RFC8573)

[RFC6424] N. Bahadur, K. Kompella, G. Swallow, "Mécanisme pour effectuer un Ping de chemin de commutation d'étiquette (LSP Ping) sur des tunnels MPLS", DOI 10.17487/RFC6424, novembre 2011. (*MàJ la* RFC4379 *; Rendue obsolète par* RFC8029) *(P.S.)*

[RFC7506] K. Raza, N. Akiya, C. Pignataro, "Option d'alerte de routeur IPv6 pour les opérations, l'administration et la maintenance de MPLS", DOI 10.17487/RFC7506, avril 2015. *(P.S.)*

7.2 Références pour information

[Err108] RFC Errata, Erratum ID 108, RFC 4379.

[Err742] RFC Errata, Erratum ID 742, RFC 4379.

[Err1418] RFC Errata, Erratum ID 1418, RFC 4379.

[Err1714] RFC Errata, Erratum ID 1714, RFC 4379.

[Err1786] RFC Errata, Erratum ID 1786, RFC 4379.

[Err2978] RFC Errata, Erratum ID 2978, RFC 4379.

[Err3399] RFC Errata, Erratum ID 3399, RFC 4379.

[IANA-ENT] IANA, "PRIVATE ENTERPRISE NUMBERS", < http://www.iana.org/assignments/enterprise-numbers >.

[IANA-MPLS-LSP-PING] IANA, "Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters", < http://www.iana.org/assignments/mpls-lsp-ping-parameters >.

[IANA-SPECIAL-IPv4] IANA, "IANA IPv4 Special-Purpose Address Registry", < http://www.iana.org/assignments/iana-ipv4-special-registry >.

[RFC0792] J. Postel, "Protocole du message de contrôle Internet – Spécification du protocole du programme Internet DARPA", STD 5, septembre 1981. *(MàJ par la* RFC6633)

[RFC3107] Y. Rekhter et E. Rosen, "Portage des informations d'étiquette dans BGP-4", mai 2001. (*MàJ par* RFC6790, RFC8277)

[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)

[RFC3443] P. Agarwal, B. Akyol, "Traitement de la durée de vie (TTL) dans les réseaux à commutation d'étiquettes multi-protocoles (MPLS)", janvier 2003. *(P.S.)*

[RFC4026] L Andersson et T. Madsen, "Terminologie des réseaux privés virtuels (VPN) approvisionnés par le fournisseur", mars 2005.

[RFC4365] E. Rosen, "Déclaration d'applicabilité pour les réseaux privés virtuels (VPN) IP BGP/MPLS", février 2006. *(Info.)*

[RFC4461] S. Yasukawa, éd., "Exigences de signalisation pour les chemins à commutation d'étiquettes (LPS) de MPLS à ingénierie de trafic en point à multipoint", avril 2006. *(Information)*

[RFC4761] K. Kompella et Y. Rekhter, éditeurs "Service de LAN privé virtuel (VPLS) utilisant BGP pour l'auto découverte et la signalisation", janvier 2007. *(P.S. ; MàJ par* RFC8395)

[RFC5036] L. Andersson, I. Minei et B. Thomas, éditeurs, "Spécification de LDP", janvier 2001. (*Remplace* RFC3036) (*MàJ par les* RFC6720 , RFC6790, RFC7552,) *(D.S)*

[RFC5085] T. Nadeau et C. Pignataro, éditeurs, "Vérification de connexité de circuit virtuel pseudo filaire (VCCV) : un canal de contrôle pour les pseudo filaires", décembre 2007. (*MàJ par* RFC5586)

[RFC5331] R. Aggarwal et autres, "Allocation d'étiquettes MPLS vers l'amont et espace d'étiquettes spécifiques du contexte", août 2008. *(P.S.)*

[RFC5462] L. Andersson, R. Asati, "Entrée de pile d'étiquettes de commutation d'étiquettes multiprotocoles (MPLS) : le champ "EXP" est renommé champ "Traffic Class"", février 2009. (*MàJ* RFC3032, RFC3270, RFC3272, RFC3443, RFC3469, RFC3564, RFC3985, RFC4182, RFC4364, RFC4379, RFC4448, RFC4761, RFC5129) *(P.S.)*

[RFC5885] T. Nadeau, C. Pignataro, "Détection de transmission bidirectionnelle (BFD) pour la vérification de connexité de pseudo circuit virtuel (VCCV)", juin 2010. *(P. S. ; MàJ par* RFC6478, RFC7885)

[RFC6425] S. Saxena et autres, "Détection des défaillances du plan des données dans MPLS en point à multipoint - Extensions au Ping LSP", novembre 2011. (*MàJ la* RFC4379) *(P.S.)*

[RFC6426] E. Gray, N. Bahadur, S. Boutros, R. Aggarwal, "Vérification de la connexité à la demande et du suivi de chemin MPLS", novembre 2011. (*MàJ la* RFC4379) *(P.S.)*

[RFC6829] M. Chen, P. Pan, C. Pignataro et R. Asati, "Ping de chemin de commutation d'étiquettes (LSP) pour les classes d'équivalence de transmission (FEC) de pseudofil annoncées sur IPv6", DOI 10.17487/RFC6829, janvier 2013. *(Rendue obsolète par* RFC8029)

[RFC7537] B. Decraene, et autres, "Registres IANA pour codets ping de LSP", DOI 10.17487/RFC7537, mai 2015. *(P.S. ; MàJ* 4379, 6424 *; Rendue obsolète par* RFC8029)

[RFC8012] N. Akiya, et autres, "Ping/Trace de LSP et pseudo filaire sur réseaux MPLS utilisant des étiquettes d'entropie", DOI 10.17487/RFC8012, novembre 2016. *(P.S. ; MàJ* RFC6790*)*

[RFC8077] L. Martini, G. Heron, "Établissement et maintenance de pseudo filaire avec le protocole de distribution d'étiquettes (LDP)", DOI 10.17487/RFC8077, février 2017. STD 84. *(Remplace* RFC4447 et 6723)

Appendice A. TLV et sous TLV déconseillés (non-normatif)

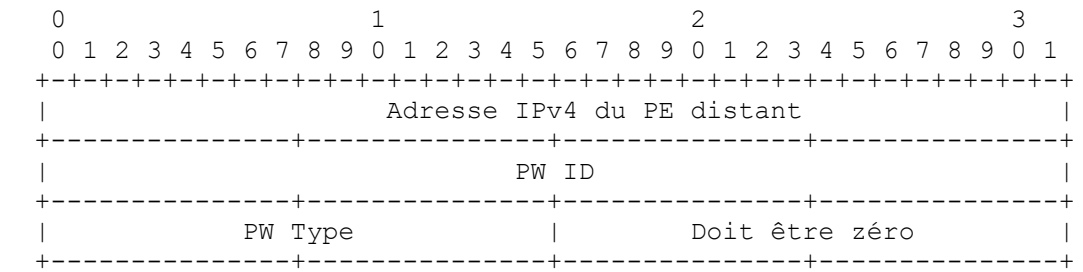
Le présent appendice décrit les éléments déconseillés, qui sont non normatifs pour une mise en œuvre. Ils sont inclus dans le présent document pour des raisons d'historique et d'information.

A.1 Pile de FEC cible

A.1.1 FEC 128 Pseudo filaire - IPv4 (déconseillé)

La FEC 128 (0x80) est définie dans la [RFC4447], comme le sont les termes PW ID (Identifiant de pseudo filaire) et PW Type (type de pseudo filaire). Un PW ID est un identifiant de connexion non zéro de 32 bits. Le PW Type est un nombre de 15 bits qui indique le type d'encapsulation. Il est porté justifié à droite dans le champ en dessous du champ appelé type d'encapsulation avec le bit de poids fort réglé à zéro. Ces deux champs sont traités dans ce protocole comme des valeurs opaques.

Quand une FEC 128 est codée dans une pile d'étiquettes, on utilise le format suivant. Le champ Valeur consiste en l'adresse IPv4 du PE distant (l'adresse de destination de la session LDP ciblée) le PW ID, et le type d'encapsulation comme suit :



Cette FEC est déconseillée et n'est conservée que pour la rétro compatibilité. Les mises en œuvre de ping de LSP DEVRAIENT accepter et traiter ce TLV, mais DEVRAIENT envoyer des demandes d'écho de ping de LSP avec le nouveau TLV (voir au paragraphe 3.2.9) sauf configuration explicite de l'utilisation du vieux TLV.

Un LSR qui reçoit ce TLV DEVRAIT utiliser l'adresse IP de source de la demande d'écho de LSP pour déduire l'adresse côté fournisseur de l'envoyeur.

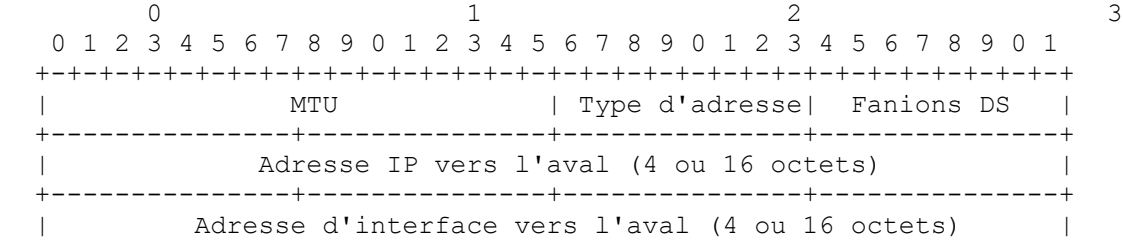
A.2 Transposition vers l'aval (déconseillée)

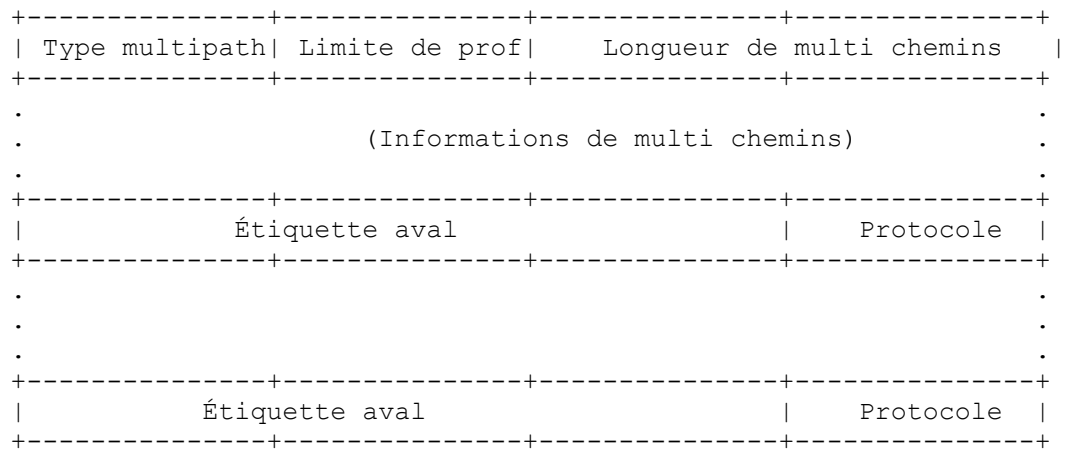
L'objet Transposition vers l'aval est un TLV qui PEUT être inclus dans un message de demande d'écho. Un seul objet Transposition vers l'aval peut apparaître dans une demande d'écho. La présence d'un objet Transposition vers l'aval est une demande que des objets Transposition vers l'aval soient inclus dans la réponse d'écho. Si le routeur qui répond est la destination de la FEC, un TLV Transposition vers l'aval NE DEVRAIT alors PAS être inclus dans la réponse d'écho.

Autrement, le routeur qui répond DEVRAIT inclure un objet Transposition vers l'aval pour chaque interface sur laquelle cette FEC pourrait être transmise. Pour une définition plus précise de la notion de "vers l'aval", voir au paragraphe 3.4.2, "Routeur et interface aval".

La longueur est K + M + 4*N octets, où M est la longueur de multi chemins, et N est le nombre d'étiquettes aval. Les valeurs pour K se trouvent dans la description du type d'adresse ci-dessous.

Le champ Valeur d'une transposition vers l'aval a le format suivant :



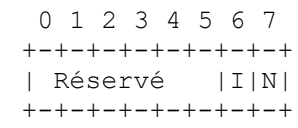


Unité de transmission maximum (MTU, *Maximum Transmission Unit*) : la MTU est la taille en octets de la plus grande trame MPLS (incluant la pile d'étiquettes) qui tient sur l'interface vers le LSR aval.

Type d'adresse : il indique si l'interface est numérotée ou non. Il détermine aussi la longueur des champs Adresse IP aval et Interface aval. Le total résultant pour la partie initiale du TLV figure dans le tableau ci-dessous. Le type d'adresse est réglé à une des valeurs suivantes :

Type	Type d'adresse	Longueur totale (octets)
1	IPv4 numéroté	16
2	IPv4 non numéroté	16
3	IPv6 numéroté	40
4	IPv6 non numéroté	28
5	Non IP	12

Fanions DS : le champ Fanions DS est un vecteur binaire du format suivant :



Deux fanions sont actuellement définis, I et N. Les fanions restants DOIVENT être réglés à zéro à l'émission et ignorés à réception.

Fanion	Nom et signification
I	Demande d'objet Interface et pile d'étiquettes. Quand ce fanion est établi, il indique que le routeur qui répond DEVRAIT inclure un objet Interface et pile d'étiquettes dans le message de réponse d'écho.
N	Traiter comme paquet non IP. Les messages de demande d'écho vont être utilisés pour diagnostiquer les flux non IP. Cependant, ces messages sont portés dans des paquets IP. Pour un routeur qui altère son algorithme ECMP sur la base de la FEC ou d'un examen de paquet en profondeur, ce fanion demande que le routeur traite cela comme il ferait si la détermination d'une charge utile IP avait échoué.

Adresse IP aval et Adresse d'interface aval : les adresses IP et les indices d'interface sont codés sur 4 octets ; les adresses IPv6 sont codées sur 16 octets.

Si l'interface au LSR aval est numérotée, le type d'adresse DOIT alors être réglé à IPv4 ou IPv6, l'adresse IP aval DOIT être réglée à l'identifiant de routeur du LSR aval ou à l'adresse d'interface du LSR aval, et l'adresse d'interface aval DOIT être réglée à l'adresse d'interface du LSR aval.

Si l'interface au LSR aval est non numérotée, le type d'adresse DOIT être IPv4 non numéroté ou IPv6 non numéroté, l'adresse IP aval DOIT être l'identifiant de routeur du LSR aval, et l'adresse de l'interface aval DOIT être réglée à l'indice alloué par le LSR amont à l'interface.

Si un LSR ne sait pas l'adresse IP de son voisin, il DOIT alors régler le type d'adresse à IPv4 non numéroté ou à IPv6 non

numéroté. Pour IPv4, il doit régler l'adresse IP aval à 127.0.0.1 ; pour IPv6, l'adresse est réglée à 0::1. Dans les deux cas, l'indice de l'interface DOIT être réglé à 0. Si un LSR reçoit un paquet de demande d'écho avec une de ces adresses dans le champ Adresse IP aval, cela indique qu'il DOIT sauter la vérification d'interface mais continuer la validation d'étiquettes.

Si le générateur d'un paquet de demande d'écho souhaite obtenir des informations de transposition vers l'aval mais ne connaît pas la pile d'étiquettes attendue, il DEVRAIT alors régler le type d'adresse à IPv4 non numéroté ou à IPv6 non numéroté. Pour IPv4, il DOIT régler l'adresse IP aval à 224.0.0.2 ; pour IPv6, l'adresse DOIT être réglée à FF02::2. Dans les deux cas, l'indice de l'interface DOIT être réglé à 0. Si un LSR reçoit un paquet de demande d'écho avec l'adresse de diffusion groupée TOUS LES ROUTEURS, cela indique qu'il DOIT sauter la validation de l'interface et de la pile d'étiquettes, mais retourner les TLV Transposition vers l'aval en utilisant les informations fournies.

Type de multi chemins : Les types de multi chemins suivants sont définis :

Clé	Type	Informations de multi chemins
0	pas de multi chemins	vide (Longueur = 0)
2	adresse IP	adresses IP
4	gamme d'adresses IP	paires d'adresse haute/basse
8	ensemble d'adresses IP à gabarit binaire	préfixe d'adresse IP et gabarit binaire
9	ensemble d'étiquettes à gabarit binaire	préfixe d'étiquette et gabarit binaire

Le type 0 indique que tous les paquets seront transmis à partir de cette interface.
Les types 2, 4, 8, et 9 spécifient que les Informations de multi chemins fournies vont servir à exploiter ce chemin.

Limite de profondeur : elle n'est applicable qu'à une pile d'étiquettes et est le nombre maximum d'étiquettes considérées dans le hachage ; ce DEVRAIT être réglé à zéro si elle n'est pas spécifiée ou non limitée.

Longueur de multi chemins : longueur en octets des informations de multi chemins.

Informations de multi chemins : Adresse ou valeurs d'étiquettes codées en accord avec le type de multi chemins. Voir au paragraphe 3.4.1.1.1 les détails du codage.

Étiquettes aval : ensemble d'étiquettes dans la pile d'étiquettes comme il apparaîtrait si ce routeur transmettait le paquet à travers cette interface. Toutes les étiquettes nulles implicites sont explicitement incluses. Les étiquettes sont traitées comme des nombres, c'est-à-dire, elles sont justifiées à droite dans le champ.

Une étiquette aval fait 24 bits, dans le même format qu'une étiquette MPLS moins le champ TTL, c'est-à-dire, le bit de poids fort de l'étiquette est le bit 0, le bit de moindre poids est le bit 19, les bits TC sont les bits 20 à 22, et le bit 23 est le bit S. Le routeur qui répond DEVRAIT remplir les bits TC et S ; le LSR qui reçoit la réponse d'écho PEUT choisir d'ignorer ces bits.

Protocole : le protocole est pris dans le tableau suivant :

N° de protocole	Protocole de signalisation
0	inconnu
1	statique
2	BGP
3	LDP
4	RSVP-TE

Remerciements

Les remerciements d'origine de la RFC 4379 déclarent : "Le présent document est le résultat de nombreuses discussions entre de nombreuses personnes, parmi lesquelles Manoj Leelanivas, Paul Traina, Yakov Rekhter, Der-Hwa Gan, Brook Bailey, Eric Rosen, Ina Minei, Shivani Aggarwal, et Vanson Lim. La description des sous champs Informations de multi chemins du TLV Transposition vers l'aval a été adaptée du texte suggéré par Curtis Villamizar.

Nous tenons à remercier Loa Andersson pour sa motivation à l'avancement de la présente spécification. Merci aussi à Alexander Vainshtein, Yimin Shen, Curtis Villamizar, David Allan, Vincent Roca, Mirja Kuhlewind, et Elwyn Davies pour leur relecture et leurs utiles commentaires.

Contributeurs

Un mécanisme utilisé pour détecter les défaillances du plan des données dans les LSP MPLS a été publié à l'origine dans la RFC 4379 en février 2006. Il a été produit par le groupe de travail MPLS de l'IETF et ses auteurs étaient Kireeti Kompella et George Swallow.
Les personnes suivantes ont fait des contributions vitales sur tous les aspects de la RFC 4379 d'origine, et beaucoup des matériaux du présent document viennent des débats et discussions au sein de ce groupe :
Ronald P. Bonica, Juniper Networks, Inc.
Dave Cooper, Global Crossing
Ping Pan, Hammerhead Systems
Nischal Sheth, Juniper Networks, Inc.
Sanjay Wadhwa, Juniper Networks, Inc.

Adresse des auteurs

Kireeti Kompella
Juniper Networks, Inc.
mél : kireeti.kompella@gmail.com

George Swallow
Cisco Systems, Inc.
mél : swallow.ietf@gmail.com

Carlos Pignataro
Cisco Systems, Inc.
mél : cpignata@cisco.com

Nagendra Kumar
Cisco Systems, Inc.
mél : naikumar@cisco.com

Sam Aldrin
Google
mél : aldrin.ietf@gmail.com

Mach(Guoyi) Chen
Huawei
mél : mach.chen@huawei.com