Groupe de travail Réseau
Request for Comments: 5496
Catégorie : Sur la voie de la normalisation
Traduction Claude Brière de L'Isle

IJ. Wijnands, Cisco Systems, Inc.
A. Boers, Cisco Systems, Inc.
E. Rosen, Cisco Systems, Inc.
mars 2009

TLV de vecteur de transmission de chemin inverse (RPF)

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

Notice de droits de reproduction
Copyright (c) 2009 IETF Trust et les personnes identifiées comme auteurs du document. Tous droits réservés.

Le présent document est soumis au BCP 78 et aux dispositions légales de l'IETF Trust qui se rapportent aux documents de l'IETF (http://trustee.ietf.org/license-info) en vigueur à la date de publication de ce document. Prière de revoir ces documents avec attention, car ils décrivent vos droits et obligations par rapport à ce document.

Résumé
Le présent document décrit une utilisation de l'attribut Join de la diffusion groupée indépendante du protocole (PIM, *Protocol Independent Multicast*) comme défini dans la RFC 5384, qui permet à PIM de construire des arborescences de diffusion groupée à travers un réseau à capacité MPLS, même si l'IGP de ce réseau n'a pas de chemin pour la source de l'arborescence.

Table des matières

1. Introduction....1
2. Notation des exigences....2
3. Utilisation du TLV Vecteur RPF....2
3.1 Jonctions d'attribut et d'arborescence partagée....2
3.2 Messages d'attribut et d'amorçage....2
3.3 Attribut Vecteur....3
4. Format de TLV de l'attribut Vecteur....4
5. Considérations relatives à l'IANA....4
6. Considérations sur la sécurité....4
7. Remerciements....4
8. Références normatives....4
Adresse des auteurs....5

1. Introduction

Il est parfois pratique de distinguer les routeurs d'un réseau particulier selon deux catégories : les "routeurs de bordure" et les "routeurs de cœur". Les routeurs de bordure se rattachent directement aux utilisateurs ou à d'autres réseaux, mais les routeurs de cœur se rattachent seulement aux autres routeurs du même réseau. Si le réseau est à capacité MPLS, alors tout paquet en envoi individuel qui doit voyager en dehors du réseau peut être "tunnelé" via MPLS d'un routeur de bordure à un autre. Pour traiter un paquet en envoi individuel qui doit voyager en dehors du réseau, un routeur de bordure a besoin de savoir quels autres routeurs de bordure sont le meilleur point de sortie du réseau pour cette adresse de destination IP du paquet. Les routeurs de cœur, n'ont cependant pas besoin d'avoir connaissance des chemins qui conduisent hors du réseau ; comme ils traitent seulement les paquets tunnelés, ils ont seulement besoin de savoir comment atteindre les routeurs de bordure et les autres routeurs de cœur.

Considérons, par exemple, le cas où le réseau est un système autonome (AS, *Autonomous System*) les routeurs de bordure sont des locuteurs du protocole de passerelle frontière externe (EBGP, *External Border Gateway Protocol*) les routeurs de cœur peuvent être dits constituer un "cœur libre de BGP". Les routeurs de bordure se distribuent les chemins BGP les uns aux autres, mais pas aux routeurs de cœur.

Cependant, quand des paquets en diffusion groupée sont considérés, la stratégie de garder les routeurs de cœur libres des chemins "externes" est plus problématique. Quand on utilise le mode PIM épars (PIM-SIM, *PIM Sparse-Mode*) [RFC4601], le mode PIM spécifique de source (PIM-SSM, *PIM Source-Specific Mode*) [RFC4607], ou PIM bidirectionnel (BIDIR-PIM) [RFC5015] pour créer une arborescence de distribution de diffusion groupée pour un groupe de diffusion groupée particulier, on veut que les routeurs de cœur soient des participants complets du protocole PIM, afin que la diffusion groupée puisse être faite efficacement dans le cœur. Cela signifie que les routeurs de cœur doivent être capables de traiter correctement les messages PIM Join pour le groupe, ce qui à son tour signifie que les routeurs de cœur doivent être capables d'envoyer les messages Join vers la racine de l'arborescence de distribution. Si la racine de l'arborescence se tient en dehors des frontières du réseau (par exemple, est dans un AS différent) et si les routeurs de cœur ne tiennent pas de chemins pour les destinations externes, alors les messages PIM Join ne peuvent pas être traités, et l'arborescence de distribution de diffusion groupée ne peut pas être créée.

Pour permettre à PIM de fonctionner correctement dans un environnement où les routeurs de cœur ne maintiennent pas les chemins externes, une extension à PIM est nécessaire. Quand un routeur de bordure envoie un message PIM Join dans le cœur, il DOIT inclure dans ce message un "vecteur" qui spécifie l'adresse IP du prochain routeur de bordure le long du chemin vers la racine de l'arborescence de distribution de diffusion groupée. Les routeurs de cœur peuvent alors traiter le message Join en l'envoyant vers le routeur de bordure spécifié (c'est-à-dire, vers le vecteur). En effet, le vecteur sert d'attribut, au sein d'un réseau particulier, pour la racine de l'arborescence.

Le présent document définit un nouveau TLV dans le message Attribut PIM Join [RFC5384]. Il consiste en un seul vecteur qui identifie le point de sortie du réseau.

2. Notation des exigences

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],.

3. Utilisation du TLV Vecteur RPF

Avant qu'un routeur puisse commencer à transmettre des paquets de diffusion groupée, il est nécessaire de construire une arborescence de transmission en envoyant des PIM Join bond par bond. Chaque routeur dans le chemin crée un état de transmission et propage le Join vers la racine de l'arborescence de transmission. La construction de cette arborescence est faite par le receveur. Voir la Figure 1.

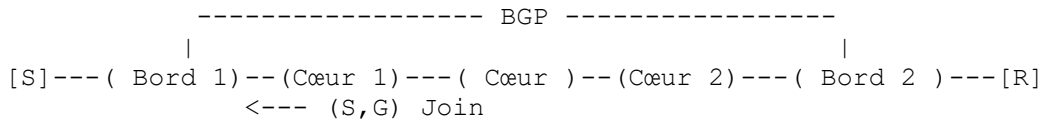


Figure 1

Dans cet exemple, les deux routeurs de bordure sont des locuteurs BGP. Les routeurs de cœur ne sont pas des locuteurs BGP et n'ont aucun chemin BGP distribué. Le chemin pour S est un chemin BGP distribué ; donc, il est connu de la bordure mais pas du cœur. Le routeur Bord 2 détermine l'interface conduisant à S, et envoie un PIM Join au routeur amont. Dans cet exemple cependant, le routeur amont est un routeur de cœur, sans chemin pour S. Sans les extensions à PIM spécifiées dans le présent document, le routeur de cœur ne peut pas déterminer où envoyer le Join, donc l'arborescence ne peut pas être construite.

Pour permettre au routeur de cœur de participer à la construction de l'arborescence, le routeur Bord 2 inclut un TLV "vecteur RPF" *(Reverse Path Forwarding)* dans l'attribut PIM Join [RFC5384] du PIM Join. Dans cet exemple, le TLV Vecteur RPF va contenir l'adresse IP de Bord 1. Bord 2 transmet le PIM Join à Bord 1. Chaque routeur de cœur intermédiaire fait sa vérification de RPF [RFC4601] sur l'adresse contenue dans le TLV Vecteur RPF (c'est-à-dire, sur l'adresse IP de Bord 1) au lieu de faire la vérification de RPFsur l'adresse S. Cela permet de construire l'arborescence.

3.1 Jonctions d'attribut et d'arborescence partagée

Dans l'exemple ci-dessus, on construit une arborescence de source pour illustrer le comportement d'attribut. L'utilisation de l'attribut n'est cependant pas restreinte à la construction des arborescences de source. Il peut aussi être utilisé pour construire une arborescence partagée. Dans ce cas, le TLV Vecteur RPF contient l'adresse IP d'un point de rendez-vous (RP). Les procédures définies dans le présent document pour les Join (S,G) sont également applicables aux Join (*,G) et (*,*,RP) sauf mention contraire.

3.2 Messages d'attribut et d'amorçage

Il n'y a pas de moyen de porter un TLV Vecteur RPF dans un message d'amorçage d'un routeur d'amorçage (BSR, *Bootstrap Router*). Les procédures du présent document ne définissent pas de moyen pour que les messages BSR soient transmis à travers un cœur dans lequel l'adresse IP du BSR n'est pas acheminable.

3.3 Attribut Vecteur

3.3.1 Insertion d'un attribut Vecteur dans un Join

Dans l'exemple de la Figure 1, quand le routeur Bord 2 cherche le chemin de la source de l'arborescence de distribution de diffusion groupée, il va trouver un chemin distribué par BGP dont le "prochain bond BGP" est Bord 1. Bord 2 cherche alors le chemin pour Bord 1 pour trouver le prochain bond pour la source, à savoir Cœur 2.

Quand Bord 2 envoie un Join PIM à Cœur 2, il inclut un attribut Vecteur qui spécifie l'adresse de Bord 1. Cœur 2, et les routeurs de cœur suivants, vont transmettre le Join le long du vecteur (c'est-à-dire, vers Bord 1) au lieu d'essayer de le transmettre vers S.

Si un attribut est réellement nécessaire dépend de si les routeurs de cœur ont un chemin pour la source de l'arborescence de diffusion groupée. Comment le routeur de bordure sait si c'est ou non le cas (et donc comment le routeur de bordure détermine si il insère ou non un champ d'attribut) sort du domaine d'application de ce document.

3.3.2 Traitement d'un attribut Vecteur reçu

Quand il traite un Join PIM reçu qui contient un attribut Vecteur, un routeur DOIT d'abord vérifier si l'adresse IP du vecteur est une de ses propres adresses IP. Si c'est le cas, l'attribut Vecteur est éliminé, et n'est pas passé plus loin en amont. Autrement, l'attribut Vecteur est utilisé pour trouver le chemin vers la source, et est passé le long de ce chemin quand un Join PIM est envoyé vers l'amont. Noter qu'un routeur qui reçoit un attribut Vecteur DOIT l'utiliser, même si ce routeur se trouve avoir un chemin pour la source. Un routeur qui élimine un attribut Vecteur PEUT bien sûr insérer un nouvel attribut Vecteur. Cela va normalement arriver si un Join PIM nécessaire pour passer à travers une séquence de routeurs de bordure, dont chaque paire est séparée par un cœur qui n'a pas de chemins externes. En l'absence de rafraîchissements périodiques, les vecteurs expirent avec l'état (S,G) correspondant.

3.3.3 Attribut Vecteur et messages Assert

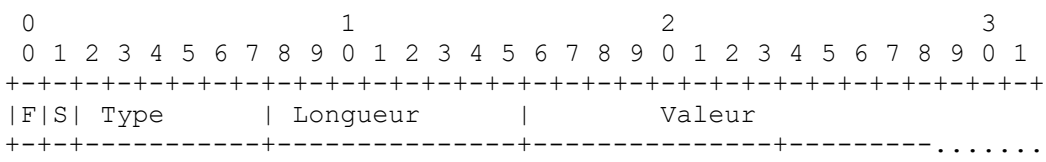
Un message PIM Assert inclut la "métrique" du protocole d'acheminement à la source de l'arborescence. Cette information est utilisée pour le choix du vainqueur de l'assertion. Si un PIM Join est envoyé sur un vecteur, plutôt que vers la source, le message Assert DOIT avoit la métrique du vecteur plutôt que la métrique de la source. Le message Assert n'a cependant pas de champ Attribut et ne mentionne pas le vecteur.

Un routeur peut changer son voisin amont sur une arborescence de diffusion groupée particulière par suite de la réception de messages Assert. Cependant, un attribut Vecteur NE DOIT PAS être envoyé dans un PIM Join à un voisin en amont choisi par suite du traitement de Assert, si ce voisin est différent du voisin amont d'origine. L'accessibilité du vecteur est seulement garantie par le routeur qui annonce l'accessibilité au vecteur dans son IGP. Si le vainqueur de l'assertion en amont n'est pas le prochain bond préféré réel, il est possible que le vainqueur de l'assertion ne connaisse pas le chemin pour le vecteur. Dans le pire des cas, le vainqueur de l'assertion a un chemin pour le vecteur qui est sur l'interface même où l'assertion a été gagnée. Cela va pointer l'interface RPF sur cette interface et va avoir pour résultat que la O-liste sera NULLE. L'attribut Vecteur NE DOIT donc PAS être inséré si le voisin RPF a été choisi via un processus Assert et si le voisin RPF est différent du voisin RPF qui aurait été choisi via le tableau d'acheminement local. Dans tous les autres cas, le vecteur DOIT être inclus dans le message Join.

3.3.4 Attribut Vecteur et suppression de Join

Si un routeur reçoit un PIM Join sur l'interface de LAN amont pour un état de diffusion groupée particulier, la suppression de Join peut être appliquée si ce PIM Join est ciblé sur le même voisin amont. Quel routeur va supprimer son PIM Join dépend du moment et n'est pas prévisible. Les routeurs en aval sur un LAN PEUVENT inclure des vecteurs RPF différents dans les PIM Join. Donc, un routeur en amont sur ce LAN peut recevoir et utiliser différents vecteurs RPF au fil du temps pour atteindre la destination (selon les routeurs en aval qui ont supprimé leur Join). Pour rendre le comportement du routeur amont plus prévisible, l'adresse du vecteur RPF DOIT être utilisée comme condition supplémentaire de la logique de suppression de Join. La logique de suppression n'est appliquée que si le vecteur RPF dans le PIM Join correspond au vecteur RPF dans l'état de diffusion groupée. Il est aussi possible de désactiver la suppression de Join sur ce LAN.

4. Format de TLV de l'attribut Vecteur



Bit F : Transmettre le TLV inconnu. Si ce bit est établi, le TLV est transmis sans considération de si le routeur comprend le type. Si le TLV est connu, le bit F est ignoré.

Bit S : Bas de pile. Si ce bit est établi, c'est alors le dernier TLV de la pile.

Type : le type d'Attribut Vecteur est 0.

Longueur : elle dépend de la famille d'adresses de l'adresse codée en envoi individuel.

Valeur : adresse codée en envoi individuel.

5. Considérations relatives à l'IANA

L'IANA a alloué la valeur 0 au vecteur RPF dans le registre "Types d'attribut de jonction PIM".

6. Considérations sur la sécurité

La sécurité de l'attribut Vecteur RPF est seulement garantie par la sécurité du paquet PIM, de sorte que les considérations sur la sécurité pour les paquets Join de PIM comme décrites dans PIM-SM [RFC4601] s'appliquent ici.

7. Remerciements

Les auteurs tiennent à remercier Yakov Rekhter et Dino Farinacci de leurs idées initiales sur ce sujet et Su Haiyang pour ses commentaires sur le document.

8. Références normatives

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

[RFC4601] B. Fenner et autres, "Diffusion groupée indépendante du protocole - Mode épars (PIM-SM) : spécification du protocole (révisée)", août 2006. (*Remplace* RFC2362) (*MàJ par* RFC5059 ; *Remplacée par* RFC7761, STD83) *(P.S.)*

[RFC4607] H. Holbrook, B. Cain, "Diffusion groupée spécifique de source pour IP", août 2006. *(P.S.)*

[RFC5015] M. Handley et autres, "Diffusion groupée bidirectionnelle indépendante du protocole (BIDIR-PIM)", octobre 2007. *(P.S.)*

[RFC5384] A. Boers et autres, "Format de l'attribut Join dans la diffusion groupée indépendante du protocole (PIM)", novembre 2008. *(P.S. ; MàJ par* RFC7887)

Adresse des auteurs

IJsbrand Wijnands
Cisco Systems, Inc.
De kleetlaan 6a
Diegem 1831
Belgium
mél : ice@cisco.com

Arjen Boers
Cisco Systems, Inc.
Avda. Diagonal, 682
Barcelona 08034
Spain
mél : aboers@cisco.com

Eric Rosen
Cisco Systems, Inc.
1414 Massachusetts Avenue
Boxborough, Ma 01719
USA
mél : erosen@cisco.com