

Groupe de travail Réseau  
**Request for Comments : 4610**  
Catégorie : Sur la voie de la normalisation

D. Farinacci, Cisco Systems  
Y. Cai, Cisco Systems  
août 2006  
Traduction Claude Brière de L'Isle

## Point de rendez-vous d'envoi à la cantonnade utilisant la diffusion groupée indépendante du protocole (PIM)

### Statut du présent mémoire

Le présent document spécifie un protocole de l'Internet en cours de normalisation 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 "Protocoles officiels de l'Internet" (STD 1) pour voir l'état de normalisation et le statut de ce protocole. La distribution du présent mémoire n'est soumise à aucune restriction.

### Notice de Copyright

Copyright (C) The Internet Society (2006).

### Résumé

La présente spécification permet d'utiliser un point de rendez vous en envoi à la cantonnade (*Anycast Rendezvous Point*) à l'intérieur d'un domaine qui fonctionne seulement sous le régime de la diffusion groupée indépendante du protocole (PIM, *Protocol Independent Multicast*). D'autres protocoles de diffusion groupée (comme le protocole de découverte de source de diffusion groupée (MSDP, *Multicast Source Discovery Protocol*) qui a été traditionnellement utilisé pour résoudre ce problème) ne sont pas obligés de prendre en charge le point de rendez vous en envoi à la cantonnade.

## 1. Introduction

Le point de rendez vous en envoi à la cantonnade décrit dans la [RFC3446] est un mécanisme que les cœurs de réseau de fournisseurs d'accès Internet (FAI) ont utilisé pour obtenir une convergence rapide quand un routeur de point de rendez-vous (RP, *Rendezvous Point*) PIM a une défaillance. Pour permettre aux receveurs et sources de se donner rendez-vous au plus proche RP, les paquets provenant d'une source ont besoin d'obtenir tous les RP pour trouver les receveurs joints.

Cette notion de receveurs qui trouvent les sources est le problème fondamental de découverte de source que MSDP était destiné à résoudre. Cependant, si on veut garder le bénéfice du point de rendez vous en envoi à la cantonnade (*Anycast-RP*) de la [RFC3446] avec moins de complications de protocole, retirer MSDP de l'espace de solution est une option.

Le présent mémoire étend le mécanisme Register de PIM de façon que la fonction de point de rendez vous en envoi à la cantonnade puisse être conservé sans utiliser MSDP.

### 1.1 Terminologie

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].

## 2. Généralités

- o Une adresse IP d'envoi individuel est choisie pour être utilisée comme adresse de RP. Cette adresse est de configuration statique, ou distribuée en utilisant un protocole dynamique, à tous les routeurs PIM à travers le domaine.
- o Un ensemble de routeurs dans le domaine est choisi pour agir comme RP pour cette adresse de RP. Ces routeurs sont appelés l'ensemble de points de rendez vous en envoi à la cantonnade.
- o Chaque routeur dans l'ensemble de points de rendez vous en envoi à la cantonnade est configuré avec une interface de

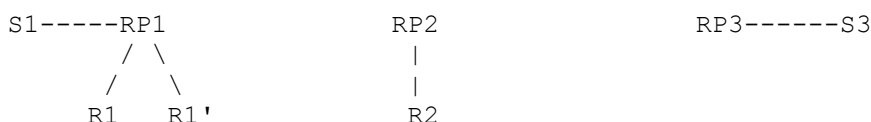
rebouclage utilisant l'adresse de RP.

- o Chaque routeur dans l'ensemble de points de rendez vous en envoi à la cantonnade a aussi besoin d'une adresse IP séparée, à utiliser pour la communication entre les RP.
- o L'adresse de RP, ou un préfixe qui couvre l'adresse de RP, est injectée dans le système d'acheminement en envoi individuel à l'intérieur du domaine.
- o Chaque routeur dans l'ensemble de points de rendez vous en envoi à la cantonnade est configuré avec les adresses de tous les autres routeurs dans l'ensemble de points de rendez vous en envoi à la cantonnade. Ceci doit être configuré de façon cohérente dans tous les RP de l'ensemble.

### 3. Mécanisme

Le diagramme suivant illustre un domaine qui utilise 3 RP où les receveurs se joignent au plus proche RP en accord avec le point où la métrique d'acheminement d'envoi individuel les prend et deux sources qui envoient des paquets à leurs RP respectifs.

Les règles décrites dans cette section n'outrepassent pas les règles de la [RFC4601]. Elles sont destinées à se mêler aux règles de la [RFC4601]. Si il y a des questions d'interprétation, la préséance est donnée à la [RFC4601].



On suppose que le scénario ci-dessus est complètement connecté où R1, R1', et R2 sont des receveurs pour un groupe, et S1 et S3 envoient à ce groupe. On suppose que RP1, RP2, et RP3 ont tous la même adresse IP allouée, qui est utilisée comme adresse de RP d'envoi à la cantonnade (disons que l'adresse IP est RPA).

Noter que l'adresse utilisée pour l'adresse de RP dans le domaine (l'adresse Anycast-RP) doit être différente des adresses utilisées par les routeurs de point de rendez vous à la cantonnade pour communiquer les uns avec les autres.

La procédure suivante est utilisée quand S1 commence à envoyer du trafic :

- o S1 envoie un paquet en diffusion groupée.
- o le routeur désigné (DR) directement rattaché à S1 va former un message PIM Register pour envoyer à l'adresse de point de rendez vous à la cantonnade (RPA). Le système d'acheminement d'envoi individuel va livrer le message PIM Register au plus proche RP, dans ce cas, RP1.
- o RP1 va recevoir le message PIM Register, le désencapsuler, et envoyer le paquet le long de l'arborescence partagée pour livrer le paquet aux receveurs R1 et R1'.
- o RP1 est configuré avec l'adresse IP de RP2 et RP3. Comme le message Register ne vient pas d'un des RP de l'ensemble anycast-RP, RP1 suppose que le paquet venait d'un DR. Si le Register n'est pas adressé à l'adresse Anycast-RP, une erreur a été faite et elle devrait être enregistrée en débit limité.
- o RP1 va alors envoyer une copie du message Register provenant du DR de S1 à RP2 et RP3. RP1 va utiliser sa propre adresse IP comme adresse de source pour le message PIM Register.
- o RP1 PEUT se joindre à l'arborescence de source en déclenchant un message (S1,G) Join vers S1. Cependant, RP1 DOIT créer l'état (S1,G).
- o RP1 renvoie un Register-Stop au DR. Si, pour une raison quelconque, les messages Register à RP2 et RP3 sont perdus, alors quand le temporisateur de suppression de Register arrive à expiration au DR, il va renvoyer des Register pour donner une autre chance à tous les RP dans l'ensemble Anycast-RP d'obtenir l'état (S,G).
- o RP2 reçoit le message Register de RP1, le désencapsule, et envoie aussi le paquet le long de l'arborescence partagée

pour donner le paquet au receveur R2.

- o RP2 renvoie un Register-Stop à RP1. RP2 PEUT attendre pour envoyer le Register-Stop si il décide de se joindre à l'arborescence de source. RP2 devrait attendre jusqu'à ce qu'il ait reçu des données de la source sur l'arborescence de source avant d'envoyer le Register-Stop. Si RP2 décide d'attendre, le Register-Stop va être envoyé quand le prochain Register sera reçu. Si RP2 décide de ne pas attendre, le Register-Stop est envoyé tout de suite.
- o RP2 PEUT se joindre à l'arborescence de source en déclenchant un message Join (S1,G) à S1. Cependant, RP2 DOIT créer l'état (S1,G).
- o RP3 reçoit le message Register de RP1, le désencapsule, mais comme il n'y a pas de receveur joint pour le groupe, il peut éliminer le paquet.
- o RP3 renvoie un Register-Stop à RP1.
- o RP3 crée l'état (S1,G) de sorte que quand un receveur se joint après que S1 a commencé d'envoyer, RP3 peut se joindre rapidement à l'arborescence de source pour S1.
- o RP1 traite le Register-Stop provenant de RP2 et de RP3. Il n'y a pas d'action spécifique lors du traitement des messages Register-Stop.

La procédure pour l'envoi par S3 suit les mêmes règles que ci-dessus mais c'est RP3 qui envoie une copie du Register généré par le DR de S3 à RP1 et RP2. Donc, cet exemple montre comment des sources n'importe où dans le domaine, associées à différents RP, peuvent atteindre tous les receveurs, aussi associés aux différents RP, dans le même domaine.

#### 4. Observations et lignes directrices sur cette proposition

- o Un RP ne va envoyer une copie d'un Register que si le Register est reçu d'une adresse IP qui n'est pas dans la liste Anycast-RP (c'est-à-dire, le Register vient d'un DR et non d'un autre RP). Une mise en œuvre DOIT se garder contre les incohérences de configuration des ensembles de Anycast-RP dans chaque RP en copiant la durée de vie (TTL, *Time to Live*) provenant d'un message Register dans les messages Register qu'il copie et envoie aux autres RP.
- o Chaque DR que PIM enregistre pour une source va envoyer le message à l'adresse Anycast-RP (ce qui a pour résultat que le paquet va au plus proche RP physique). Donc, il n'y a pas de changement de la logique de DR.
- o Les paquets s'écoulent à tous les receveurs quel que soit le RP auquel ils se sont joints.
- o La source va être enregistrée à un seul RP par le DR. Il est de la responsabilité du RP qui reçoit les messages PIM Register du DR (le plus proche RP du DR sur la base de la métrique d'acheminement) de faire passer le paquet à tous les autres RP dans l'ensemble Anycast-RP.
- o La logique n'est changée que dans les RP. Le changement de logique est pour l'envoi des copies des messages Register. Le traitement de Register-Stop est inchangé. Cependant, une mise en œuvre PEUT supprimer l'envoi des messages Register-Stop en réponse à un Register reçu d'un RP.
- o La limitation en débit des messages Register et Register-Stop est faite de bout en bout. C'est-à-dire de DR -> RP1 -> {RP2 et RP3}. Il n'y a pas besoin d'une logique de limitation en débit spécifique entre les RP.
- o Quand des changements de topologie surviennent, l'arborescence de source existante s'ajuste comme elle le fait aujourd'hui en accord avec la [RFC4601]. Les arborescences partagées existantes s'ajustent aussi comme elles le font aujourd'hui en accord avec la [RFC4601].
- o Les changements physiques de RP sont au rythme de la convergence de chemin en envoi individuel, en conservant le bénéfice de la [RFC3446].
- o Un RP qui ne prend pas en charge la présente spécification peut être mêlé à des RP qui le font. Cependant, le RP qui ne prend pas en charge ne devrait pas avoir de sources qui s'y enregistrent, mais peut avoir des receveurs qui s'y joignent.
- o Si des Register Nuls sont envoyés (des Register avec un en-tête IP et pas de charge utile IP) ils DOIVENT être

dupliqués à tous les RP dans l'ensemble Anycast-RP afin que l'état de source reste vivant pour les sources actives.

- o Le nombre de RP dans l'ensemble Anycast-RP devrait rester petit afin que la quantité de duplication non native reste à un minimum.
- o Comme le RP, qui reçoit un Register du DR, va envoyer des copies du Register aux autres RP au même moment qu'il envoie un Register-Stop au DR, il pourrait y avoir de la perte de paquet et d'état dans les autres RP jusqu'au moment où le DR envoie de nouveau des messages Register.

## 5. Interaction avec MSDP dans un routeur PIM en envoi à la cantonnade

L'objectif de cette proposition de diffusion groupée indépendante du protocole en envoi à la cantonnade (Anycast-PIM) est de s'affranchir de la dépendance à l'utilisation de MSDP. Ceci peut se faire en supprimant l'échange de trafic MSDP entre des points de rendez-vous en envoi à la cantonnade (*Anycast-RP*). Cependant, pour annoncer les sources internes aux routeurs en dehors d'un domaine d'acheminement PIM et pour apprendre les sources externes provenant des autres domaines d'acheminement, MSDP peut être quand même nécessaire.

### 5.1 Fonction PIM en envoi à la cantonnade dans un domaine de bout

Dans cette capacité, quand il y a des sources internes qui doivent être annoncées en externe, un RP en envoi à la cantonnade qui reçoit un message Register, soit d'un DR, soit d'un RP en envoi à la cantonnade, devrait le traiter comme décrit dans la présente spécification ainsi que de la façon de traiter un message Register décrite dans la [RFC4601]. Cela signifie qu'un Source-Active (SA) pour la même source interne pourrait être généré par plusieurs RP en envoi à la cantonnade qui font l'échange de trafic MSDP. Il n'y a rien de fondamentalement faux dans cela à part que la source est annoncée dans l'infrastructure MSDP à partir de plusieurs endroits du domaine de source. Cependant, si ce n'est pas souhaitable, la configuration d'un ou plusieurs (plutôt que de tous) routeurs MSDP de RP en envoi à la cantonnade permettrait que seuls ces routeurs génèrent des SA pour la source interne. Et dans certaines situations, il y a de bonnes possibilités que tous les RP d'envoi à la cantonnade dans l'ensemble n'aient pas de sessions d'échange de trafic MSDP de sorte que ce problème peut être dans une certaine mesure atténué.

Du point de vue d'un RP en envoi à la cantonnade, une source devrait être considérée comme interne à un domaine quand elle est découverte par un Anycast-RP par la réception d'un message Register, sans considérer si le message Register a été envoyé par un DR, un autre membre Anycast-RP, ou le routeur lui-même.

Pour apprendre les sources externes à un domaine, les messages SA de MSDP pourraient arriver à plusieurs Anycast-RP d'échange de trafic MSDP. Les règles pour le traitement d'un SA, selon la [RFC3446], devraient être suivies. C'est-à-dire, si G se joint au domaine, un Join (S,G) est envoyé à la source. Et si des données accompagnent le SA, chaque RP Anycast-PIM qui fait de l'échange de trafic MSDP va transmettre les données le long de son arborescence partagée respective.

Ceci suppose que chaque Anycast-RP a des connexions d'échange de trafic MSDP externes. Si ce n'est pas le cas, les routeurs Anycast-PIM avec les connexions d'échange de trafic MSDP vont suivre la même procédure que si un Data-Register ou un Null-Register avait été reçu d'un DR ou d'un autre Anycast-RP. C'est-à-dire qu'ils vont envoyer des Register aux autres membres de l'ensemble Anycast-RP.

Si il y a un mélange de RP d'envoi à la cantonnade qui ont et qui n'ont pas de connexion d'échange de trafic MSDP externe, alors ceux qui en ont doivent être configurés avec l'ensemble de ceux qui n'en ont pas. Ainsi les messages Register sont envoyés seulement aux membres de l'ensemble de Anycast-RP qui n'a pas de connexion d'échange de trafic MSDP externe.

La quantité de trafic Register généré par ce RP d'échange de trafic MSDP serait égale au nombre de sources externes actives sur le domaine. L'état Source-Active devrait être transporté à tous les autres RP dans l'ensemble Anycast-RP car le RP d'échange de trafic MSDP ne saura rien des membres du groupe associés aux autres RP. Pour éviter ce trafic de contrôle périodique, il est recommandé que tous les Anycast-RP soient configurés avec des sessions d'échange de trafic MSDP externes afin qu'aucun RP dans l'ensemble Anycast-RP n'ait à générer de message Register au nom de source externes.

### 5.2 Fonction PIM en envoi à la cantonnade dans un domaine de transit

Au sein d'un domaine d'acheminement, il est recommandé qu'un ensemble de Anycast-RP défini dans la présente spécification ne devrait pas être mêlé avec des membres qui font de l'échange de trafic MSDP. Dans certains cas, la

découverte de source va fonctionner mais il peut n'être pas évident pour une mise en œuvre quelles sources sont locales au domaine et quelles ne le sont pas. Ceci peut affecter l'annonce MSDP externe des sources internes.

Ceci étant dit, le présent document n'essaye pas de connecter des domaines d'échange de trafic MSDP entre eux en utilisant Anycast-PIM à l'intérieur d'un domaine de transit.

## 6 Considérations sur la sécurité

Cette Section décrit les considérations sur la sécurité pour les messages Register et Register-Stop entre les RP en envoi à la cantonnade. Pour les messages PIM entre DR et RP, se reporter à la [RFC4601].

### 6.1 Attaques fondées sur des messages falsifiés

Un attaquant peut falsifier un message Register en utilisant une des adresses de la liste des Anycast-RP afin de réaliser un ou plusieurs des effets suivants :

1. Submerger le RP cible dans une attaque de déni de service (DoS)
2. Injecter des données non autorisées chez les receveurs desservis pas le RP
3. Injecter des données non autorisées et créer des entrées de SA bogués dans d'autres domaines PIM si le RP cible a des échanges de trafic MSDP externes.

Un attaquant peut aussi falsifier un message Register-Stop en utilisant une des adresses de la liste de Anycast-RP. Cependant, à côté du déni de service, l'effet d'une telle attaque est limité parce que un RP ignore généralement les messages Register-Stop.

### 6.2 Protection des messages Register et Register-Stop

L'attaque de DoS en utilisant des messages Register ou Register-Stop falsifiés ne peut pas être empêchée. Mais le RP peut quand même être protégé. Par exemple, le RP peut limiter le débit d'entrée des messages. Il peut aussi choisir de refuser de traiter les messages Register-Stop. Le mécanisme réel de protection est spécifique de la mise en œuvre.

La distribution de données non autorisées et de messages Register bogués peut être empêchée en utilisant la méthode décrite au paragraphe 6.3.2 de la [RFC4601]. Quand RP1 envoie une copie d'un Register à RP2, RP1 agit comme la [RFC4601] décrit le DR et RP2 agit comme la [RFC4601] décrit le RP.

Comme décrit dans la [RFC4601], un RP peut être configuré en utilisant un unique SA et indice de paramètre de sécurité (SPI) pour le trafic (Register ou Register-Stop) pour chaque membre des Anycast-RP de la liste, mais il en résulte un problème de gestion de clés ; donc, il peut être préférable dans les domaines PIM où tous les points de rendez-vous sont sous un seul contrôle administratif, d'utiliser les mêmes paramètres d'algorithme d'authentification (y compris la clé) pour tous les paquets enregistrés dans un domaine.

## 7. Remerciements

Les auteurs ont réalisé un prototype de ce document dans les mises en œuvre cisco IOS et Procket, respectivement.

Les auteurs remercient John Zwiebel des ses essais d'interopérabilité des deux prototypes de mise en œuvre.

Les auteurs remercient Thomas Morin de France Telecom pour sa discussion extensive sur la diffusion groupée des Register dans un maillage complet fondé sur SSM parmi l'ensemble de Anycast-RP. Cette idée pourra faire l'objet d'un futur document.

Et finalement, les auteurs remercient les personnes suivantes de leurs commentaires sur les projets antérieurs :

Greg Shepherd (Procket Networks (maintenant Cisco Systems))

Lenny Giuliano (Juniper Networks)

Prashant Jhingran (Huawei Technologies)

Pekka Savola (CSC/FUNET)

Bill Fenner (AT&T)

James Lingard (Data Connection)

Amit Shukla (Juniper Networks)  
Tom Pusateri (Juniper Networks)

## 8. Références

### 8.1 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.)

### 8.2 Références pour information

[RFC3446] D. Kim et autres, "[Mécanisme de point de rendez-vous \(RP\)](#) en envoi à la cantonnade utilisant la diffusion groupée indépendante du protocole (PIM) et le protocole de découverte de source de diffusion groupée (MSDP)", janvier 2003. (Info.)

## Appendice A : Possible langage de configuration

Un ensemble possible de commandes à utiliser pourrait être :

```
ip pim anycast-rp <anycast-rp-addr> <rp-addr>
```

Où :

<anycast-rp-addr> décrit l'ensemble Anycast-RP pour le RP qui est alloué à la gamme du groupe. Cette adresse IP est l'adresse qu'utilisent les routeurs PIM de premier et dernier bonds pour s'enregistrer et se joindre.

<rp-addr> décrit l'adresse IP où les copies de messages Register sont envoyées. Cette adresse IP est toute adresse allouée au routeur RP, non incluse <anycast-rp-addr>.

Exemple :

À partir de l'illustration ci-dessus, les commandes de configuration seraient :

```
ip pim anycast-rp RPA RP1  
ip pim anycast-rp RPA RP2  
ip pim anycast-rp RPA RP3
```

Commentaire :

Il peut être utile d'inclure l'adresse IP du routeur local dans l'ensemble de commandes de sorte que les lignes ci-dessus peuvent être copiées collées ou écrites dans tous les RP de l'ensemble de Anycast-RP.

Mais la mise en œuvre devrait connaître sa propre adresse et ne pas envoyer par inadvertance un Register à elle même.

## Adresse des auteurs

Dino Farinacci  
Cisco Systems  
mél : [dino@cisco.com](mailto:dino@cisco.com)

Yiqun Cai  
Cisco Systems  
mél : [ycai@cisco.com](mailto:ycai@cisco.com)

## Déclaration complète de droits de reproduction

Copyright (C) The IETF Trust (2006).

Le présent document est soumis aux droits, licences et restrictions contenus dans le BCP 78, et à [www.rfc-editor.org](http://www.rfc-editor.org), et sauf pour ce qui est mentionné ci-après, les auteurs conservent tous leurs droits.

Le présent document et les informations contenues sont fournis sur une base "EN L'ÉTAT" et le contributeur, l'organisation qu'il ou elle représente ou qui le/la finance (s'il en est), la INTERNET SOCIETY et la INTERNET ENGINEERING TASK FORCE déclinent toutes garanties, exprimées ou implicites, y compris mais non limitées à toute garantie que l'utilisation des informations encloses ne viole aucun droit ou aucune garantie implicite de commercialisation ou d'aptitude à un objet particulier.

### Propriété intellectuelle

L'IETF ne prend pas position sur la validité et la portée de tout droit de propriété intellectuelle ou autres droits qui pourrait être revendiqués au titre de la mise en œuvre ou l'utilisation de la technologie décrite dans le présent document ou sur la mesure dans laquelle toute licence sur de tels droits pourrait être ou n'être pas disponible ; pas plus qu'elle ne prétend avoir accompli aucun effort pour identifier de tels droits. Les informations sur les procédures de l'ISOC au sujet des droits dans les documents de l'ISOC figurent dans les BCP 78 et BCP 79.

Des copies des dépôts d'IPR faites au secrétariat de l'IETF et toutes assurances de disponibilité de licences, ou le résultat de tentatives faites pour obtenir une licence ou permission générale d'utilisation de tels droits de propriété par ceux qui mettent en œuvre ou utilisent la présente spécification peuvent être obtenues sur répertoire en ligne des IPR de l'IETF à <http://www.ietf.org/ipr>.

L'IETF invite toute partie intéressée à porter son attention sur tous copyrights, licences ou applications de licence, ou autres droits de propriété qui pourraient couvrir les technologies qui peuvent être nécessaires pour mettre en œuvre la présente norme. Prière d'adresser les informations à l'IETF à [ietf-ipr@ietf.org](mailto:ietf-ipr@ietf.org).

### Remerciement

Le financement de la fonction d'édition des RFC est fourni par l'activité de soutien administratif (IASA) de l'IETF.