

Groupe de travail Réseau  
**Request for Comments : 5004**  
 Catégorie : Sur la voie de la normalisation  
 Traduction Claude Brière de L'Isle

E. Chen, Cisco Systems  
 S. Sangli, Cisco Systems  
 01/09/07

# Éviter les transitions de meilleur chemin BGP d'un externe à un autre

## Statut du présent mémoire

Le présent document spécifie un protocole de l'Internet sur la voie de la 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 (2007).

## Résumé

Dans ce document, on propose une extension aux règles de choix de chemin de BGP qui éviterait des transitions inutiles de meilleur chemin entre des chemins externes dans certaines conditions. L'extension proposée aiderait à la stabilité du réseau global, et plus important, éliminerait certaines oscillations de chemin BGP dans lesquelles plus d'un chemin externe provenant d'un des locuteurs BGP contribue à l'oscillation.

## Table des Matières

1. Introduction.....	1
2. Spécification des exigences.....	1
3. L'algorithme.....	2
4. Les avantages.....	2
5. Remarques.....	3
6. Considérations sur la sécurité.....	3
7. Remerciements.....	3
8. Références normatives.....	3
9. Références pour information.....	3
Adresse des auteurs.....	4
Déclaration complète de droits de reproduction.....	4

## 1. Introduction

Les deux dernières étapes du choix de chemin BGP (au paragraphe 9.1.2.2 de la [RFC4271]) impliquent de comparer les identifiants BGP et les adresses d'échange de trafic. L'identifiant BGP (traité comme une adresse IP ou juste comme un entier [RFC6286]) pour un locuteur BGP est alloué par le système autonome (AS, *Autonomous System*) auquel le locuteur appartient. Par suite, pour un locuteur BGP local, l'identifiant BGP d'un chemin reçu d'un homologue externe est juste un nombre aléatoire. Quand les chemins considérés viennent d'homologues externes, le résultat des deux dernières étapes du choix de chemin est donc "aléatoire" pour ce qui concerne le locuteur BGP local.

C'est sur la base de cette observation qu'on propose une extension aux règles de choix de chemin BGP qui éviterait des transitions inutiles de meilleur chemin entre des chemins externes dans certaines conditions. L'extension proposée aiderait à la stabilité du réseau global, et plus important, éliminerait certaines oscillations de chemin BGP dans lesquelles plus d'un chemin externe provenant d'un locuteur BGP contribue à l'oscillation.

## 2. Spécification 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. L'algorithme

Considérons le cas dans lequel le meilleur chemin existant A vient d'un homologue externe, et où un autre chemin externe B est ensuite choisi comme nouveau meilleur chemin par l'algorithme de choix de chemin décrit dans la [RFC4271]. Quand on compare tous les chemins dans le choix de chemin, si ni le chemin A ni le chemin B ne sont éliminés par l'algorithme de choix de chemin avant l'étape f) -- la comparaison d'identifiants BGP (au paragraphe 9.1.2.2 de la [RFC4271]) -- on propose que le meilleur chemin existant (chemin A) soit conservé comme meilleur chemin (évitant donc de passer le meilleur chemin au chemin B). Cet algorithme NE DEVRAIT PAS être appliqué quand l'un ou l'autre des chemins vient d'un homologue de confédération BGP.

De plus, l'algorithme NE DEVRAIT PAS être appliqué quand les deux chemins proviennent d'homologues qui ont un identifiant BGP identique (c'est-à-dire qu'il existe des sessions BGP parallèles entre les deux locuteurs BGP). Comme les adresses d'échange de trafic pour les sessions parallèles sont normalement allouées par un AS (éventuellement avec des considérations de choix de chemin) l'algorithme (si il est appliqué) pourrait impacter l'établissement de l'acheminement existant. De plus, en n'appliquant pas l'algorithme, l'allocation des adresses d'échange de trafic resterait un outil simple et efficace pour influencer le choix de chemin quand il existe des sessions BGP parallèles.

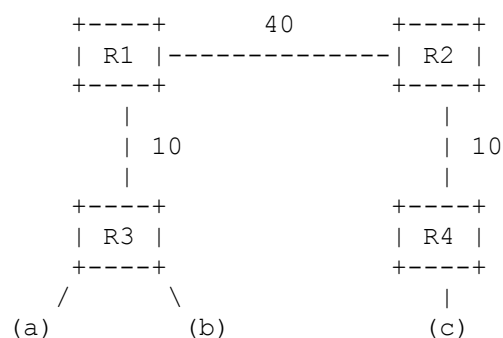
### 4. Les avantages

L'extension proposée aux règles de choix de chemin BGP évite des transitions inutiles de meilleur chemin entre des chemins externes sous certaines conditions. En clair, l'extension aiderait à réduire les changements d'acheminement et de transmission dans un réseau, aidant donc à la stabilité du réseau global.

Plus important, comme le montre l'exemple suivant, l'extension proposée peut être utilisée pour éliminer certaines oscillations de chemin BGP dans lesquelles plus d'un chemin externe provenant d'un locuteur BGP contribue à l'oscillation. Noter cependant qu'il y a des scénarios permanents d'oscillation de chemins BGP [RFC3345] que le mécanisme décrit dans le présent document n'élimine pas.

Considérons l'exemple de la Figure 1 où :

- o R1, R2, R3, et R4 appartiennent à un AS.
- o R1 est un réflecteur de chemin avec R3 comme client.
- o R2 est un réflecteur de chemin avec R4 comme client.
- o Les métriques IGP sont celles indiquées.
- o Les chemins externes (a), (b), et (c) sont comme décrit à la Figure 2.



**Figure 1**

Chemin	AS	MED	Identifiant
a	1	0	2
b	2	20	1
c	2	10	5

**Figure 2**

Du fait de l'interaction de la réflexion de chemin [RFC4456] et de l'attribut MULTI\_EXIT\_DISC (MED) (découverte multi sorties) le meilleur chemin sur R1 continue d'osciller entre (a) et (c), et le meilleur chemin sur R3 oscille entre (a) et (b).

Avec l'algorithme proposé, R3 ne changerait pas de meilleur chemin de (a) à (b) même après que R1 retire (c) à ses clients, cela suffit pour arrêter l'oscillation de chemin.

Bien que ce type d'oscillation de chemin puisse aussi être éliminé par d'autres améliorations de réflexion de chemin en cours de développement, l'algorithme proposé est extrêmement simple et peut être mis en œuvre et déployé immédiatement sans introduire de problème de rétro compatibilité.

## 5. Remarques

L'algorithme proposé est rétro compatible, et peut être déployé sur la base du locuteur BGP. Le déploiement de l'algorithme est fortement recommandé sur un locuteur BGP avec plusieurs homologues BGP externes (en particulier ceux qui connectent un point d'échange).

Comparé au comportement existant, l'algorithme proposé peut introduire un certain "non déterminisme" dans le choix de chemin BGP -- bien qu'on puisse objecter que la comparaison d'identifiant BGP dans le choix de chemin existant ait déjà introduit un certain "aléa" comme décrit dans l'introduction. Un tel "non déterminisme" ne s'est pas révélé désavantageux en pratique et peut être complètement éliminé en utilisant les mécanismes existants (comme d'établir LOCAL\_PREF ou MED) si on le désire.

## 6. Considérations sur la sécurité

Cette extension n'introduit aucun problème de sécurité.

## 7. Remerciements

L'idée présentée a été inspirée par un cas d'oscillation de chemin observé dans le réseau BBN/Genuity en 1998. L'algorithme a aussi été mis en œuvre et déployé à ce moment.

Les auteurs tiennent à remercier Yakov Rekhter et Ravi Chandra de leurs commentaires sur l'idée initiale.

## 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](#))
- [RFC4271] Y. Rekhter, T. Li et S. Hares, "[Protocole de routeur frontière](#) version 4 (BGP-4)", janvier 2006. (D.S.) (MàJ par [RFC6608](#), [RFC8212](#), [RFC9072](#))
- [RFC4456] T. Bates, E. Chen, R. Chandra, "[Réflexion de chemin BGP](#) : une solution de remplacement au BGP interne à maillage complet (IBGP)", avril 2006. (Remplace [RFC2796](#), [RFC1966](#)) (D.S.)

## 9. Références pour information

- [RFC3345] D. McPherson et autres, "Condition d'oscillation de chemin persistante du protocole de routeur frontière (BGP)", août 2002. (Information)
- [RFC6286] E. Chen, J. Yuan, "Identifiant BGP unique au niveau d'un système autonome pour BGP-4", juin 2011. (MàJ la RFC4271) (P.S.)

## Adresse des auteurs

Enke Chen  
Cisco Systems, Inc.  
170 W. Tasman Dr.  
San Jose, CA 95134  
mail : [enkechen@cisco.com](mailto:enkechen@cisco.com)

Srihari R. Sangli  
Cisco Systems, Inc.  
170 W. Tasman Dr.  
San Jose, Ca. 95134  
mail : [rsrihari@cisco.com](mailto:rsrihari@cisco.com)

## Déclaration complète de droits de reproduction

Copyright (C) The Internet Society (2007)

Le présent document est soumis aux droits, licences et restrictions contenus dans le BCP 78, 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, le IETF TRUST 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 pourraient 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 le 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 de l'IETF (IASA).