Groupe de travail Réseau
Request for Comments : 5195
Catégorie : Sur la voie de la normalisation
Traduction Claude Brière de L'Isle

H. Ould-Brahim, Nortel
D. Fedyk, Nortel
Yakov Rekhter, Juniper Networks
juin 2008

Auto découverte fondée sur BGP pour les VPN de couche 1

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.

Résumé

L'objet du présent document est de définir un mécanisme d'auto-découverte fondé sur BGP pour les VPN de couche 1 (L1VPN, *Layer-1 VPN*). Le mécanisme d'auto-découverte pour les L1VPN permet aux appareils réseau du fournisseur de découvrir dynamiquement l'ensemble des appareils de côté fournisseur (PE, *Provider Edge*) qui ont des accès rattachés aux membres du côté consommateur (CE, *Customer Edge*) du même VPN. Cette information est nécessaire pour achever la phase de signalisation des connexions de L1VPN. Un des principaux objectifs d'un mécanisme d'auto-découverte de L1VPN est la prise en charge du modèle de "provisionnement sur une seule extrémité", où l'ajout d'un nouvel accès à un certain L1VPN n'impliquerait de changements de configuration qu'au PE qui a cet accès et sur le CE qui est connecté au PE via cet accès.

Table des Matières

1. Introduction

L'objet du présent document est de définir un mécanisme d'auto-découverte fondé sur BGP pour les VPN de couche 1 (L1VPN, *Layer-1 VPN*). Le mécanisme d'auto-découverte pour les L1VPN permet aux appareils réseau du fournisseur de découvrir dynamiquement l'ensemble des appareils de côté fournisseur (PE, *Provider Edge*) qui ont des accès rattachés aux membres du côté consommateur (CE, *Customer Edge*) du même VPN. Cette information est nécessaire pour achever la phase de signalisation des connexions de L1VPN. Un des principaux objectifs d'un mécanisme d'auto-découverte de L1VPN est la prise en charge du modèle de "provisionnement sur une seule extrémité", où l'ajout d'un nouvel accès à un certain L1VPN n'impliquerait de changements de configuration qu'au PE qui a cet accès et sur le CE qui est connecté au PE via cet accès.

Le mécanisme d'auto-découverte procède en ayant un PE qui annonce aux autres PE les informations suivantes, au minimum : sa propre adresse IP et la liste des couples de <adresse privée, adresse de fournisseur> locaux pour ce PE. Une fois ces informations reçues, les PE distants vont identifier la liste des membres du VPN qu'ils ont en commun avec le PE annonceur, et utiliser les informations portées dans le mécanisme de découverte pour effectuer la résolution d'adresse durant la phase de signalisation des connexions de VPN de couche 1.

La Figure 1 illustre la référence de réseau pour utiliser le mécanisme d'auto-découverte fondé sur BGP pour les VPN de couche 1. Pour les besoins du mécanisme d'auto-découverte, BGP fonctionne seulement sur le réseau de fournisseur. Les PE tiennent des tableaux d'informations par VPN appelées des tableaux d'informations d'accès (PIT, *Port Information Table*) relatifs aux informations de <adresse privée, adresse de fournisseur>. Il y a plus d'informations sur les PIT à la Section 2.

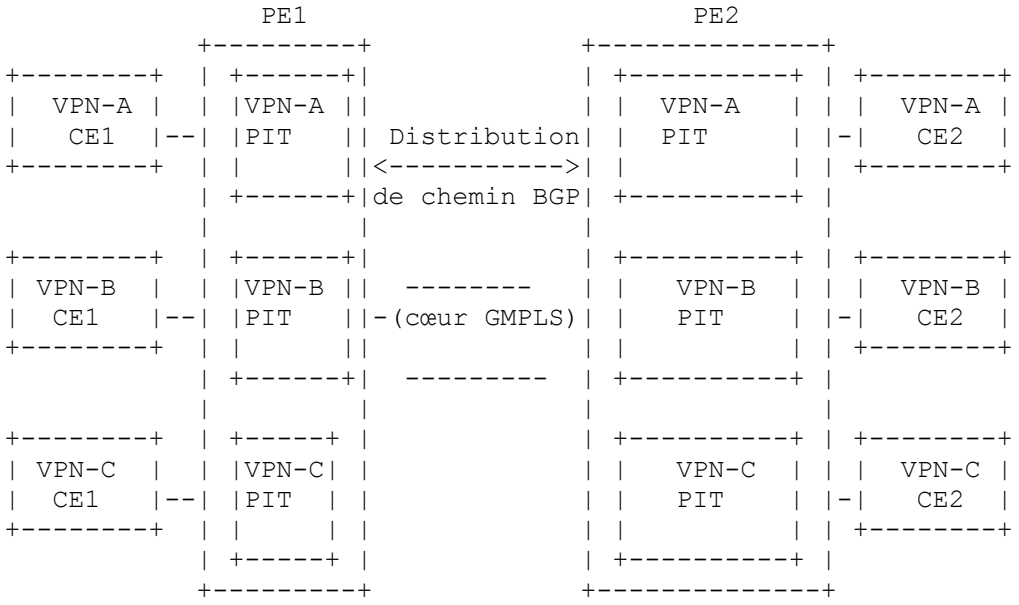

Figure 1 : Auto découverte BGP pour L1VPN

La [RFC4847] décrit deux modes de fonctionnement pour un L1VPN : le mode de base et le mode amélioré. Le présent document décrit un mécanisme d'auto-découverte seulement pour le mode de base.

1.1 Langage 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].

2. Procédures

Dans le contexte des L1VPN, un CE est connecté à un PE via un ou plusieurs accès, où chaque accès peut consister en un ou plusieurs canaux ou sous canaux. Chaque accès sur un CE qui connecte le CE à un PE a un identifiant qui est unique dans ce L1VPN (mais n'a pas besoin d'être unique sur plusieurs L1VPN). On appelle cet identifiant l'identifiant d'accès consommateur (CPI, *Customer Port Identifier*). Chaque accès sur un PE a aussi un identifiant qui est unique au sein du réseau de fournisseur. On appelle cet identifiant l'identifiant d'accès fournisseur (PPI, *Provider Port Identifier*). Noter que les adresses IP utilisées pour les CPI ou PPI peuvent être des adresses IPv4 ou IPv6.

Pour chaque L1VPN qui a au moins un acccès configuré sur un PE, le PE tient un tableau d'informations d'accès (PIT, *Port Information Table*). Un PIT contient une liste des couples <CPI, PPI> pour tous les accès dans son L1VPN. Noter qu'un PIT peut aussi contenir des informations d'acheminement (par exemple, quand les CPI sont appris en utilisant un protocole d'acheminement).

Un PIT sur un certain PE est rempli avec deux types d'informations :
- des informations relatives aux accès de CE rattachés aux accès sur le PE. Ces informations peuvent être configurées en local au PE ou être reçues des CE ;
- des informations reçues d'autres PE par le mécanisme d'auto découverte.

On se réfère aux premières comme à des informations locales, et aux dernières comme à des informations distantes. La propagation des informations locales aux autres PE est accomplie en utilisant les extensions multi protocoles à BGP

[RFC4760]. Pour restreindre le flux de ces informations aux seuls PIT dans un L1VPN donné, on utilise le filtrage de chemin BGP fondé sur la communauté étendue de cible de chemin [RFC4360], comme suit :

Chaque PIT sur un PE est configuré avec une ou plusieurs communautés de cible de chemin, appelées "cibles de chemin exportées", qui sont utilisées pour étiqueter les informations locales quand elles sont exportées dans le BGP du fournisseur. La granularité de cet étiquetage pourrait être aussi fine qu'une seule paire de <CPI, PPI>. De plus, chaque PIT sur un PE est configuré (au moment du provisionnement) avec une ou plusieurs communautés de cible de chemin, appelées "cibles de chemin importées", qui restreignent l'ensemble de chemins qui pourraient être importés du BGP du fournisseur dans le PIT pour seulement les chemins qui ont au moins une de ces communautés.

Chaque événement suivant se produit au moment du provisionnement : si un fournisseur de services ajoute un nouvel accès L1VPN à un PE particulier, cet accès est associé à un PIT sur ce PE, et ce PIT est associé à ce L1VPN.

Noter que comme le protocole utilisé pour remplir un PIT avec les informations distantes est BGP, et comme BGP fonctionne sur plusieurs systèmes autonomes (AS) il s'ensuit que le mécanisme décrit dans le présent document pourrait prendre en charge des L1VPN qui s'étendent sur plusieurs systèmes autonomes.

Bien que les L1VPN multi AS sortent actuellement du domaine d'application du mode de base, les mécanismes définis dans le présent document paraissent facilement applicables à un scénario multi AS, si un tel besoin devait apparaître à l'avenir. Pour l'instant, des travaux supplémentaires peuvent être nécessaires pour examiner divers aspects incluant la sécurité.

3. Portage des informations L1VPN dans BGP

La transposition de <CPI, PPI> est portée en utilisant les extensions multi protocoles à BGP [RFC4760]. La [RFC4760] définit le format de deux attributs BGP, MP_REACH_NLRI et MP_UNREACH_NLRI, qui peuvent être utilisés pour annoncer et retirer l'annonce des informations d'accessibilité. On introduit un nouvel identifiant de famille d'adresses suivante, appelé informations d'auto découverte de VPN de couche 1 (valeur 69) et aussi un nouveau format d'informations d'accessibilité de couche réseau (NLRI, *Network Layer Reachability Information*) pour porter les informations de CPI et PPI.

Un ou plusieurs couples <PPI, CPI> pourraient être portés dans les attributs BGP susmentionnés.

Le format des NLRI est décrit dans la Figure 2.

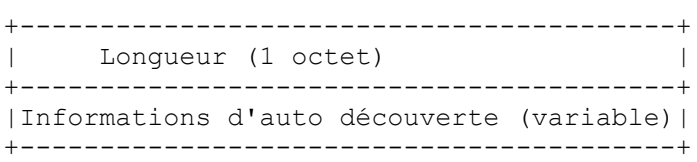

Figure 2 : Codage des NLRI

Noter que le codage des informations d'auto découverte est décrit dans la [RFC5251], et noter aussi que si la valeur de la longueur du champ Prochain bond (de l'attribut MP_REACH_NLRI) est 4, alors le prochain bond contient une adresse IPv4. Si cette valeur est 16, alors le prochain bond contient une adresse IPv6.

4. Portage des informations d'ingénierie du trafic L1VPN dans BGP

En plus des informations d'accessibilité, le mécanisme d'auto-découverte PEUT porter des informations d'ingénierie du trafic utilisées pour le choix du chemin de sortie. Par exemple, un PE peut apprendre la capacité de commutation et la bande passante maximum de LSP des interfaces L1VPN distantes des PE distants. Le présent document utilise l'attribut BGP Ingénierie du trafic [RFC5543] pour porter ces informations.

5. Adaptabilité

On rappelle que le réseau du fournisseur de services consiste en (a) des PE, (b) des réflecteurs de chemin BGP, (c) des nœuds P (qui ne sont ni des PE ni des réflecteurs de chemin) et dans le cas de VPN multi fournisseurs, (d) des routeurs de bordure de système autonome (ASBR, *Autonomous System Border Router*).

Un routeur PE, sauf si il est un réflecteur de chemin, ne conserve pas les informations relatives au L1VPN, sauf si il a au moins un VPN avec une cible de chemin importé identique à un des attributs de cible de chemin d'informations relatives au VPN. Si un PE n'a pas un VPN avec une cible de chemin importé correspondante, il DOIT alors éliminer les informations de L1VPN reçues. Le filtrage d'entrée DOIT être utilisé pour causer l'élimination de ces informations. Si une nouvelle cible de chemin importée est ensuite ajoutée à un des VPN du PE (opération "Jonction de VPN") il DOIT alors acquérir les informations relatives au VPN qu'il a précédemment éliminées.

Dans ce cas, le mécanisme de rafraîchissement décrit dans la [RFC2918] DOIT être utilisé. Les mécanismes de filtrage de chemin sortant des [RFC5291] et [RFC4684] peuvent aussi être utilisés pour rendre le filtrage plus dynamique.

De même, si une cible de chemin importée particulière n'est plus présente dans les VPN d'un PE (par suite d'une ou plusieurs "opérations d'élagage de VPN") le PE PEUT éliminer tous les chemins BGP de L1VPN qui, par suite, n'ont plus de cibles de chemin importé du PIT du PE comme un de leurs attributs Cible de chemin.

Noter que les opérations "Jonction de VPN" et "Élagage de VPN" ne sont pas interruptives, et n'exigent pas que des connexions BGP soient fermées, tant que le mécanisme de rafraîchissement de la [RFC2918] est utilisé.

Par suite de ces règles de distribution, aucun PE n'a jamais besoin de conserver tous les chemins pour tous les L1VPN ; c'est une importante considération d'adaptabilité.

Les réflecteurs de chemin peuvent être partagés entre les VPN de telle sorte que chaque partition porte des chemins pour seulement un sous ensemble des L1VPN pris en charge par le fournisseur de services. Donc, aucun réflecteur de chemin unique n'est exigé pour conserver les informations relatives au VPN pour tous les VPN.

Pour les VPN inter fournisseurs, si le BGP multi bonds externe (EBGP) est utilisé, alors les ASBR n'ont pas du tout besoin de conserver et distribuer les informations relatives au VPN. Les routeurs P ne conservent aucune information relative au VPN.

Par suite, aucun composant unique au sein du réseau du fournisseur de service n'a à conserver toutes les informations relatives au VPN pour tous les VPN. Donc la capacité totale du réseau pour prendre en charge la croissance du nombre de VPN n'est pas limitée par la capacité d'un composant individuel.

Une importante considération à rappeler est qu'on peut avoir un nombre quelconque de systèmes BGP INDÉPENDANTS qui portent des informations relatives au VPN. Ceci est différent du cas de l'Internet, où le système BGP de l'Internet DOIT porter tous les chemins de l'Internet. Donc, une distinction significative (mais peut-être subtile) entre l'utilisation de BGP pour l'acheminement Internet et pour la distribution des informations relatives au VPN, comme décrite dans le présent document, est que la première n'est pas susceptible de partition, tandis que la seconde l'est.

6. Considérations sur la sécurité

Le présent document décrit un mécanisme d'auto-découverte fondé sur BGP qui permet à un PE qui se rattache à un L1VPN particulier de découvrir l'ensemble des autres routeurs PE qui se rattachent au même VPN. Chaque routeur PE qui est rattaché à un certain VPN utilise BGP pour l'annoncer. Les autres routeurs PE qui se rattachent au même VPN reçoivent ces annonces BGP. Cela permet que des ensembles de PE se découvrent les uns les autres. Noter qu'un PE ne va pas toujours recevoir ces annonces directement des PE distants ; les annonces peuvent être reçues de locuteurs BGP "intermédiaires".

Il est d'une importance critique qu'un PE particulier NE DOIT PAS être "découvert" comme étant rattaché à un VPN particulier si il n'est pas réellement rattaché à ce VPN, et est bien sûr proprement autorisé à être rattaché à ce VPN. Si un nœud arbitraire de l'Internet pouvait commencer d'envoyer ces annonces BGP, et si ces annonces étaient capables d'atteindre les nœuds de PE, et si les nœuds de PE acceptaient ces annonces, alors n'importe qui pourrait ajouter n'importe quel site à n'importe quel L1VPN. Donc, les procédures d'auto-découverte décrites ici présupposent qu'un PE particulier fasse confiance à ses homologues BGP pour être qui ils paraissent être, et de plus, qu'il peut faire confiance à ces

homologues pour sécuriser de façon appropriée leurs rattachements locaux. (C'est-à-dire, un PE DOIT faire confiance à ses homologues qui sont rattachés, et sont autorisés à être rattachés aux L1VPN auxquels ils prétendent être rattachés.)

Si un PE distant particulier est un homologue BGP du PE local, alors les procédures d'authentification de BGP de la [RFC2385] DEVRAIENT être utilisées pour s'assurer que le PE distant est qui il prétend être, c'est-à-dire, qu'il est un PE de confiance.

Si un PE distant particulier n'est pas un homologue BGP du PE local, alors les informations qu'il annonce sont distribuées au PE local à travers une chaîne de locuteurs BGP. Le PE local DOIT avoir confiance que ses homologues acceptent seulement des informations provenant d'homologues en qui ils ont à leur tour confiance, et cette relation de confiance DOIT être transitive. BGP ne fournit pas de moyen de déterminer qu'un élément d'information particulier reçu provient d'un locuteur BGP qui était autorisé à annoncer cet élément d'information particulier. Donc, les procédures du présent document DOIVENT être seulement utilisées dans des environnements où des relations de confiance adéquates existent entre les locuteurs BGP (comme dans le cas de l'utilisation du mécanisme d'auto découverte au sein d'un seul réseau de fournisseur).

7. Considérations relatives à l'IANA

Le présent document alloue un nouveau SAFI, appelé informations d'auto découverte de VPN de couche 1 (Section 3). Cette allocation a été faite dans le registre des identifiants de famille d'adresse suivante (SAFI, *Subsequent Address Family Identifier*) en utilisant la procédure d'allocation d'action de normalisation. La valeur est 69.

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)

[RFC2918] E. Chen, "Capacité de rafraîchissement de chemin pour BGP-4", septembre 2000. *(P.S., MàJ par* RFC7313)

[RFC4760] T. Bates, R. Chandra, D. Katz et Y. Rekhter, "Extensions multi protocoles pour BGP-4", janvier 2007.

8.2 Références pour information

[RFC2385] A. Heffernan, "Protection des sessions de BGP via l'option de signature MD5 de TCP", août 1998. *(P.S. ; MàJ par la* RFC6691 *; remplacée par* RFC5925)

[RFC4360] S. Sangli et autres, "Attribut BGP-4 Communauté étendue", février 2006. *(P.S.)*

[RFC4684] P. Marques et autres, "Distribution de chemin contraint pour réseaux privés virtuels (VPN) au protocole Internet selon le protocole de routeur frontière/commutation d'étiquettes multiprotocoles (BGP/MPLS)", novembre 2006. *(P.S.)*

[RFC4847] T. Takeda, éd., "Cadre et exigences pour la couche 1 des réseaux privés virtuels", avril 2007. *(Information)*

[RFC5251] D. Fedyk et autres, "Mode de base de VPN de couche 1", juillet 2008. *(P.S.)*

[RFC5291] E. Chen, Y. Rekhter, "Capacité de filtrage de chemin sortant pour BGP-4", août 20080. *(P.S.)*

[RFC5543] H. Ould-Brahim, D. Fedyk, Y. Rekhter, "Attribut BGP d'ingénierie du trafic", mai 2009. *(P. S., MàJ par* RFC7606*)*

9. Remerciements

Merci à Adrian Farrel de ses utiles commentaires.

Adresse des auteurs

Hamid Ould-Brahim
Nortel
PO Box 3511 Station C
Ottawa ON K1Y 4H7
Canada
téléphone : +1 (613) 763 4730
mél : hbrahim@nortel.com

Yakov Rekhter
Juniper Networks
1194 N. Mathilda Avenue
Sunnyvale, CA 94089
USA
mél : yakov@juniper.net

Don Fedyk
Nortel
600 Technology Park
Billerica, MA 01821
USA
téléphone : +1 (978) 288 3041
mél : dwfedyk@nortel.com

Déclaration complète de droits de reproduction

Copyright (C) The IETF Trust (2008).

Le présent document est soumis aux droits, licences et restrictions contenus dans le BCP 78, et à 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 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.