Groupe de travail Réseau
Request for Comments : 4576
Catégorie : Sur la voie de la normalisation
Traduction Claude Brière de L'Isle

E. Rosen, Cisco Systems, Inc.
P. Psenak, Cisco Systems, Inc.
P. Pillay-Esnault, Cisco Systems, Inc.
juin 2006

Utilisation d'un bit d'option de LSA pour empêcher les boucles dans les VPN IP BGP/MPLS

Résumé
Le présent document spécifie une procédure qui traite d'un problème particulier pouvant survenir quand un fournisseur de service (SP, *Service Provider*) offre un réseau virtuel privé (VPN, *Virtual Private Network*) IP BGP/MPLS à un consommateur qui utilise OSPFv2 pour annoncer ses chemins au SP. Dans cette situation, un routeur côté consommateur (CE, *Customer Edge*) et un routeur côté fournisseur (PE, *Provider Edge*) sont des homologues OSPF, et les chemins du consommateur sont envoyés via OSPFv2 du CE au PE. Les chemins du consommateur sont convertis en chemins BGP, et BGP les porte à travers le cœur de réseau aux autres routeurs PE. Les chemins sont alors reconvertis en chemins OSPF envoyés via OSPF aux autres routeurs CE. Par suite de cette conversion, certaines des informations nécessaires pour empêcher les boucles peuvent être perdues. Une procédure est nécessaire pour s'assurer qu'une fois qu'un chemin est envoyé d'un PE à un CE, le chemin sera ignoré par tout PE qui le reçoit en retour d'un CE. Le présent document spécifie la procédure nécessaire, en utilisant un des bits d'options dans l'annonce d'état de liaison (LSA, *Link State Advertisement*) pour indiquer qu'une LSA a déjà été transmise par un PE et devrait être ignorée par tous les autres PE qui la voient.

Table des matières

1. Introduction....1
2. Spécification des exigences....2
3. Perte d'informations et boucles....2
4. Utilisation des options de LSA pour empêcher les boucles....3
5. Considérations sur la sécurité....3
6. Remerciements....3
7. Références normatives....4
Adresse des auteurs....4
Déclaration complète de droits de reproduction....4

1. Introduction

La [RFC4364] décrit une méthode par laquelle un fournisseur de service (SP) peut utiliser son cœur de réseau IP pour fournir un service de "VPN IP" aux consommateurs. Dans cette sorte de service, les appareils côté consommateur (appareils CE) sont connectés aux routeurs côté fournisseur (routeurs PE). Chaque appareil CE est dans un seul réseau virtuel privé (VPN, *Virtual Private Network*). Chaque appareil PE peut se rattacher à plusieurs CE du même VPN ou de VPN différents. Un VPN consiste donc en un ensemble de "segments de réseau" connectés par le cœur de réseau du SP.

Un CE échange des routes avec un PE, en utilisant un protocole d'acheminement sur lequel le consommateur et le SP se mettent d'accord. Le PE fait fonctionner le processus de décision de ce protocole d'acheminement (c'est-à-dire, il effectue les calculs d'acheminement) pour déterminer l'ensemble des préfixes d'adresses IP pour lesquels tiennent les deux conditions suivantes :

- Chaque préfixe d'adresse de l'ensemble peut être joint via ce CE.
- Le chemin de ce CE à chacun de ces préfixes d'adresse N'inclut PAS le cœur de réseau du SP (c'est-à-dire, il n'inclut

aucun routeur PE).

Les routeurs PE qui se rattachent à un VPN particulier redistribuent les routes à ces préfixes d'adresses dans BGP, de sorte qu'ils peuvent utiliser BGP pour distribuer les routes du VPN à chacun des autres. BGP porte ces routes dans la "famille d'adresses VPN-IPv4", afin qu'elles soient distinctes des routes ordinaires de l'Internet. La famille d'adresses VPN-IPv4 étend aussi les adresses IP sur la gauche de sorte que les préfixes d'adresses provenant de deux VPN différents soient toujours distinct pour BGP, même si les deux VPN utilisent le même élément de l'espace d'adresse privée de la RFC 1918. Donc, les routes pour les différents VPN peuvent être portées par une seule instance BGP et peuvent être mémorisés dans un tableau BGP commun sans crainte de conflit.

Si un routeur PE reçoit un chemin VPN-IPv4 particulier via BGP, et si ce PE est rattaché à un CE dans le VPN auquel appartient le chemin, le processus de décision de BGP peut alors installer ce chemin dans le tableau d'acheminement BGP. Si il en est ainsi, le PE retraduit le chemin en route IP et la redistribue au protocole d'acheminement qui fonctionne sur la liaison qui va à ce CE.

Cette méthodologie donne un "modèle d'homologue". Les routeurs CE échangent du trafic avec les routeurs PE, mais les routeurs CE sur des sites différents n'échangent pas de trafic les uns avec les autres.

Si un VPN utilise OSPFv2 comme protocole d'acheminement interne, les routeurs CE de ce VPN n'utilisent pas nécessairement OSPFv2 pour échanger du trafic avec les routeurs PE. Chaque site dans un VPN peut utiliser OSPFv2 comme protocole d'acheminement intra site tout en utilisant BGP ou RIP (par exemple) pour distribuer les chemins à un routeur PE. Cependant, il est certainement pratique quand OSPFv2 est utilisé en intra site de l'utiliser aussi sur les liaisons PE-CE, et la [RFC4364] le permet explicitement. Dans ce cas, un PE va faire fonctionner une instance séparée de OSPFv2 pour chaque VPN rattaché au PE ; le PE va en général avoir un tableau d'acheminement OSPFv2 spécifique pour chaque VPN.

Quand OSPFv2 est utilisé sur une liaison PE-CE qui appartient à un VPN particulier, le routeur PE doit redistribuer à l'instance OSPFv2 de ce VPN certaines routes qui ont été installées dans le tableau d'acheminement de BGP. De même, un routeur PE doit redistribuer à BGP les routes qui ont été installées dans les tableaux d'acheminement spécifiques de VPN de OSPF. Les procédures pour cela sont spécifiées dans la [RFC4577].

Les routes qui sont redistribuées de BGP à OSPFv2 sont annoncées dans des LSA générées par le PE. Le PE agit comme routeur frontière OSPF, annonçant certaines de ces routes dans des LSA d'AS externes, et certaines dans des LSA résumées, comme spécifié dans la [RFC4577].

De même, quand un routeur PE reçoit une LSA d'un routeur CE, il fait le calcul d'acheminement OSPF. Tout chemin qui se trouve installé dans le tableau d'acheminement OSPF doit être traduit en chemin de VPN-IPv4 et ensuite redistribué dans BGP. BGP va alors distribuer ces routes aux autres routeurs PE.

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. Perte d'informations et boucles

Un PE, disons PE1, peut apprendre un chemin pour un préfixe particulier d'adresse de VPN IPv4 via BGP. Cela peut l'amener à générer une LSA résumée ou une LSA d'AS externe dans laquelle il rapporte ce préfixe d'adresse. Cette LSA peut ensuite être distribuée à un CE particulier, disons CE1. La LSA peut ensuite être distribuée dans toute une zone OSPF particulière, atteignant un autre CE, disons CE2. CE2 peut ensuite distribuer la LSA à un autre PE, disons PE2.

Comme déclaré dans la section précédante, PE2 doit faire le calcul d'acheminement OSPF pour déterminer si un préfixe d'adresse particulier, rapporté dans une LSA venant de CE2, est accessible depuis CE2 via un chemin qui n'inclut aucun routeur PE. Malheureusement, il n'y a pas une façon standard pour faire cela. Les LSA OSPFv2 ne portent pas nécessairement les informations nécessaires pour permettre à PE2 de déterminer qu'un chemin pour le préfixe d'adresse X dans une certaine LSA venant de CE2 est en fait un chemin qui inclut, par exemple, PE1. Si PE2 diffuse dans BGP X

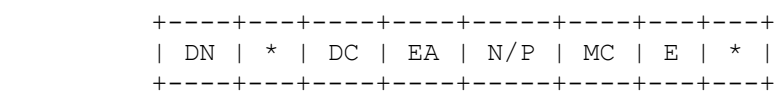
comme chemin de VPN IPv4, alors PE2 viole une des contraintes pour l'exclusion de boucle dans BGP ; à savoir que les routes apprises d'un certain domaine BGP ne sont pas redistribuées dans ce domaine BGP. Cela pourrait causer la création d'une boucle d'acheminement.

Il est donc nécessaire d'avoir un moyen pour qu'une LSA puisse porter l'information qu'un préfixe d'adresse particulier a été appris d'un routeur PE. Tout routeur PE qui reçoit une LSA avec cette information va omettre l'information de cette LSA de son calcul d'acheminement OSPF, et donc ne va pas répandre en retour cette information dans BGP.

Quand un PE génère une LSA d'AS externe, il pourrait utiliser une valeur d'étiquette distincte pour indiquer que la LSA porte des informations sur un préfixe d'adresse pour lequel le chemin comporte un routeur PE. Cependant, cette méthode n'est pas disponible dans le cas où le PE génère une LSA résumée. Selon la [RFC4577], chaque routeur PE doit fonctionner comme un routeur OSPF de zone 0. Si la liaison PE-CE est une liaison de zone 0, il est alors possible au PE de recevoir, sur cette liaison, une LSA résumée qui a pour origine un autre routeur PE. Donc, on a besoin d'un moyen pour marquer une LSA résumée comme indiquant qu'elle porte des informations sur un chemin via un routeur PE.

4. Utilisation des options de LSA pour empêcher les boucles

Le bit de poids fort du champ Options de LSA (bit jusqu'à présent non utilisé) est utilisé pour résoudre le problème décrit à la Section 3. On se réfère à ce bit comme bit DN. Quand une LSA de type 3, 5, ou 7 est envoyée d'un PE à un CE, le bit DN DOIT être établi. Le bit DN DOIT être à zéro dans tous les autres types de LSA.



Champ Options avec le bit DN (RFC 2328, paragraphe A.2)

Quand le PE reçoit, d'un routeur CE, une LSA de type 3, 5, ou 7 avec le bit DN établi, les informations de cette LSA NE DOIVENT PAS être utiliséees durant le calcul de chemin OSPF. Par suite, la LSA n'est pas traduite en un chemin BGP. Le bit DN DOIT être ignoré dans tous les autres types de LSA.

Cela empêche les routes apprises via BGP d'être redistribuées à BGP. (Cette restriction est analogue à la restriction OSPF usuelle que les chemins inter zones qui sont appris d'une zone 0 ne sont pas repassés à la zone 0.)

Noter que le bit DN n'a pas d'autre effet sur le traitement de LSA. En particulier, une LSA avec le bit DN établi sera mise dans la base de données topologique, vieillie, arrosée, etc., juste comme si le DN n'était pas établi.

5. Considérations sur la sécurité

Un attaquant peut causer le non établissement du bit DN, dans une LSA voyageant du CE au PE, quand le bit DN devrait en fait être établi. Cela peut être cause que les préfixes d'adresses mentionnés dans cette LSA soient inaccessibles à partir d'autres sites du VPN. De même, un attaquant peut causer le non établissement du bit DN, dans une LSA voyageant dans l'une ou l'autre direction, quand le bit DN devrait en fait être établi. Cela peut causer des boucles d'acheminement pour le trafic destiné aux préfixes d'adresses mentionnés dans cette LSA.

Ces possibilités peuvent être éliminées en utilisant l'authentification cryptographique spécifiée à la Section D de la [RFC2328].

6. Remerciements

L'idée d'utiliser le bit d'option de poids fort à cette fin est due à Derek Yeung. Merci à Yakov Rekhter de sa contribution au présent travail. Nous remercions aussi Acee Lindem de ses utiles commentaires.

7. 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)

[RFC2328] J. Moy, "OSPF version 2", STD 54, avril 1998. *(MàJ par la* RFC6549, RFC8042)

[RFC4364] E. Rosen et Y. Rekhter, "Réseaux privés virtuels IP BGP/MPLS", février 2006. (*P.S., MàJ par* RFC4577, RFC4684)

[RFC4577] E. Rosen et autres, "OSPF comme protocole de bord fournisseur/consommateur pour les réseaux privés virtuels (VPN) IP BGP/MPLS", juin 2006. (*MàJ* RFC4364) *(P.S.)*

Adresse des auteurs

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

Peter Psenak
Cisco Systems, Inc.
BA Business Center, 9th Floor
Bratislava 82109
Slovakia
mél : ppsenak@cisco.com

Padma Pillay-Esnault
Cisco Systems, Inc.
3750 Cisco Way
San Jose, CA 95134
USA
mél : ppe@cisco.com

Déclaration complète de droits de reproduction

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