Groupe de travail Réseau
**Request for Comments : 4364**
RFC rendue obsolète : 2547
Catégorie : Sur la voie de la normalisation

E. Rosen, Cisco Systems
Y. Rekhter, Juniper Networks
février 2006
Traduction Claude Brière de L'Isle

# Réseaux privés virtuels IP BGP/MPLS

**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.



**Résumé**
Le présent document décrit une méthode par laquelle un fournisseur de services peut utiliser un cœur de réseau IP pour fournir des réseaux privés virtuels (VPN, *Virtual Private Network*) à ses clients. Cette méthode utilise un "modèle d'homologues", dans lequel les routeurs du côté consommateur (CE, *customers' edge*) envoient leurs chemins aux routeurs du côté fournisseur de service (PE, *Provider's edge*) ; il n'y a pas de "recouvrement" visible pour l'algorithme d'acheminement du consommateur, et les routeurs CE aux différents sites n'échangent pas de trafic les uns avec les autres. Les paquets de données sont tunnelés à travers le cœur de réseau, de sorte que les routeurs du cœur n'ont pas besoin de connaître les chemins du VPN.

Le présent document rend obsolète la RFC 2547.

## Table des matières

# 1. Introduction

Le présent document décrit une méthode par laquelle un fournisseur de services peut utiliser un cœur de réseau IP pour fournir des réseaux privés virtuels (VPN, *Virtual Private Network*) IP à ses consommateurs. Cette méthode utilise un "modèle d'homologues", dans lequel les routeurs du côté consommateur (routeurs CE) envoient leurs chemin aux routeurs côté fournisseur de services (routeurs PE). Le protocole de routeur frontière (BGP, *Border Gateway Protocol*) [RFC4271], [RFC2858] est alors utilisé par le fournisseur de services pour échanger les chemins d'un VPN particulier entre les routeurs PE qui sont rattachés à ce VPN. Ceci est fait d'une façon qui assure que les chemins provenant de différents VPN restent distincts et séparés, même si deux VPN ont un espace d'adresses qui se recouvre. Les routeurs PE distribuent, aux routeurs CE dans un VPN particulier, les chemins provenant des autres routeurs CE dans ce VPN. Les routeurs CE n'échangent pas de trafic les uns avec les autres, donc il n'y a pas de "recouvrement" visible pour l'algorithme d'acheminement du VPN. Le terme "IP" dans "VPN IP" est utilisé pour indiquer que le PE reçoit des datagrammes IP du CE, qu'il examine leurs en-têtes IP, et les achemine en conséquence.

Une étiquette de commutation d'étiquettes multi protocoles (MPLS, *Multiprotocol Label Switching*) [RFC3031], [RFC3107], [RFC3032] est allouée à chaque chemin au sein d'un VPN ; quand BGP distribue un chemin de VPN, il distribue aussi une étiquette MPLS pour ce chemin. Avant qu'un paquet de données de client traverse le cœur de réseau du fournisseur de services, il est encapsulé avec l'étiquette MPLS qui correspond, dans le VPN du client, au chemin qui est la meilleure correspondance avec l'adresse de destination du paquet. Ce paquet MPLS est encore encapsulé (par exemple, avec une autre étiquette MPLS ou avec un en-tête de tunnel IP ou d'encapsulation d'acheminement générique (GRE, *Generic Routing Encapsulation*) [RFC4023]) de façon à être tunnelé à travers le cœur de réseau au routeur PE approprié. Donc, les routeurs du cœur de réseau n'ont pas besoin de connaître les chemins de VPN.

Le principal but de cette méthode est de prendre en charge le cas où un client obtient des services IP de cœur de réseau d'un ou de fournisseurs de services avec lesquels il a des relations contractuelles. Le client peut être une entreprise, un groupe d'entreprises qui ont besoin d'un extranet, un fournisseur de services Internet, un fournisseur de services d'application, un autre fournisseur de services de VPN qui utilise cette même méthode pour offrir des VPN à ses propres clients, etc. La méthode rend très simple au client l'utilisation des services de cœur de réseau. Elle est aussi très adaptable et souple pour le fournisseur de services, et lui permet d'ajouter de la valeur.

## 1.1 Réseaux virtuels privés

Considérons un ensemble de "sites" qui sont rattachés à un réseau commun qu'on va appeler "le cœur de réseau". On applique ensuite une politique pour créer un certain nombre de sous ensembles de cet ensemble, et on impose la règle suivante : deux sites ne peuvent avoir l'inter connexité IP sur ce cœur de réseau que si au moins un de ces sous ensembles les contient tous deux.

Ces sous ensembles sont des réseaux privés virtuels (VPN, *Virtual Private Network*). Deux sites ont la connexité IP sur le cœur de réseau commun seulement si il y a un VPN qui les contient tous les deux. Deux sites qui n'ont pas de VPN en commun n'ont pas de connexité sur ce cœur de réseau.

Si tous les sites dans un VPN sont possédés par la même entreprise, le VPN peut être vu comme un "intranet" d'entreprise. Si les divers sites dans un VPN sont possédés par différentes entreprises, le VPN peut être vu comme un "extranet". Un site peut être dans plus d'un VPN ; par exemple, dans un intranet et dans plusieurs extranets. En général, quand on utilise le terme "VPN" on ne va pas distinguer entre intranets et extranets.

On se réfère aux possesseurs des sites comme à des "consommateurs". On se réfère aux possesseurs/opérateurs du cœur de réseau comme "fournisseurs de services" (SP, *Service Provider*). Les consommateurs obtiennent un "service de VPN" des SP.

Un client peut être une seule entreprise, un ensemble d'entreprises, un fournisseur de services Internet, un fournisseur de services d'application, un autre SP qui offre la même sorte de service de VPN à ses propres consommateurs, etc.

Les politiques qui déterminent si une collection de sites particulière est un VPN sont les politiques des consommateurs. Certains consommateurs vont vouloir que la mise en œuvre de ces politiques soit entièrement de la responsabilité du SP. D'autres consommateurs peuvent vouloir partager avec le SP la responsabilité de la mise en œuvre de ces politiques. Le présent document spécifie des mécanismes qui peuvent être utilisés pour mettre en œuvre ces politiques. Les mécanismes décrits sont assez généraux pour permettre que ces politiques soient mises en œuvre soit par le SP seul, soit par un client de VPN avec le SP. La plus grande partie de la discussion est cependant concentrée sur ce dernier cas.

Les mécanismes discutés dans le présent document permettent la mise en œuvre d'une large gamme de politiques. Par exemple, au sein d'un certain VPN, on peut permettre que chaque site ait un chemin direct à chaque autre site ("maillage complet"). Autrement, on peut forcer le trafic entre certaines paires de sites à être acheminé via un troisième site. Cela peut être utile, par exemple, si on désire que le trafic entre une paire de sites soit passé à travers un pare-feu, et que le pare-feu est dans ce troisième site.

Dans le présent document, on restreint la discussion au cas dans lequel le client recherche explicitement le service de VPN auprès d'un SP, ou d'un ensemble de SP qui se sont mis d'accord pour coopérer à fournir le service de VPN. C'est-à-dire, le client ne recherche pas simplement un accès Internet auprès d'un SP, et le trafic de VPN ne passe pas à travers une collection aléatoire de réseaux de SP interconnectés.

On restreint aussi la discussion au cas dans lequel le cœur de réseau fournit un service IP au client, plutôt que, par exemple, un service de couche 2 comme du relais de trame, du mode de transfert asynchrone (ATM, *Asynchronous Transfer Mode*), ethernet, du contrôle de liaison de données de haut niveau (HDLC, *High Level Data Link Control*), ou du protocole point à point (PPP). Le client peut se rattacher au cœur de réseau via un de ces services de couche 2 (ou autre) mais le service de couche 2 se termine au "bord" du cœur de réseau, où les datagrammes IP du client sont retirés de toute encapsulation de couche 2.

Dans le reste de cette introduction, on spécifie des propriétés que devraient avoir les VPN. Le reste de ce document spécifie un ensemble de mécanismes qui peuvent être déployés pour fournir un modèle de VPN qui a toutes ces propriétés. Cette section introduit aussi la terminologie technique utilisée dans le reste de ce document.

### 1.2 Côté consommateur et côté fournisseur

Les routeurs peuvent être rattachés l'un à l'autre, ou à des systèmes d'extrémité, de diverses façons : des connexions PPP, des circuits virtuels (VC) ATM, des VC en relais de trame, des interfaces ethernet, des réseaux de zone locale virtuels (VLAN, *Virtual Local Area Network*) sur des interfaces ethernet, des tunnels GRE, des tunnels du protocole de tunnelage de couche 2 (L2TP, *Layer 2 Tunneling Protocol*), des tunnels IPsec, etc. On utilise le terme de "circuit de rattachement" pour se référer généralement à un moyen de rattachement à un routeur. Un circuit de rattachement peut être la sorte de connexion qui est usuellement vue comme une "liaison de données", ou ce peut être un tunnel de quelque sorte ; ce qui importe est qu'il soit possible à deux appareils d'être des homologues de couche réseau sur le circuit de rattachement.

Chaque site de VPN doit contenir un ou plusieurs appareils de côté consommateur (CE, *Customer Edge*). Chaque appareil CE est rattaché, via un circuit de rattachement, à un ou plusieurs routeurs de côté fournisseur (PE, *Provider Edge*).

Les routeurs dans le réseau du SP qui ne se rattachent pas à des appareils CE sont appelés des "routeurs P".

Les appareils CE peuvent être des hôtes ou des routeurs. Dans un cas normal, un site contient un ou plusieurs routeurs, dont certains sont rattachés aux routeurs PE. Les routeurs du site qui se rattachent aux routeurs PE vont alors être les appareils CE, ou "routeurs CE". Cependant, il n'y a rien pour empêcher un hôte non acheminant de se rattacher directement à un routeur PE , dans ce cas l'hôte va être un appareil CE.

Parfois, ce qui est rattaché physiquement à un routeur PE est un commutateur de couche 2. Dans ce cas, on ne dit pas que le commutateur de couche 2 est un appareil CE. Les appareils CE sont plutôt les hôtes et routeurs qui communiquent avec le routeur PE à travers le commutateur de couche 2 ; l'infrastructure de couche 2 est transparente. Si l'infrastructure de couche 2 fournit un service multipoint, plusieurs appareils CE peuvent alors être rattachés au routeur PE sur le même circuit de rattachement.

Les appareils CE font logiquement partie d'un VPN de client. Les routeurs PE et P font logiquement partie du réseau du SP.

Le circuit de rattachement sur lequel voyage un paquet quand il va du CE au PE est appelé le "circuit de rattachement d'entrée" de ce paquet , et le PE est le "PE d'entrée" du paquet. Le circuit de rattachement sur lequel un paquet voyage en allant du PE au CE est appelé le "circuit de rattachement de sortie" de ce paquet , et le PE est le "PE de sortie" du paquet.

On dit qu'un routeur PE est rattaché à un VPN particulier si il est rattaché à un appareil CE qui est dans un site de ce VPN. De même, on dit qu'un routeur PE est rattaché à un site particulier si il est rattaché à un appareil CE qui est dans ce site.

Quand l'appareil CE est un routeur, il est un homologue d'acheminement du ou des PE auxquels il est rattaché, mais il n'est PAS un homologue d'acheminement de routeurs CE d'autres sites. Les routeurs de sites différents n'échangent pas directement d'informations d'acheminement les uns avec les autres ; en fait, ils n'ont même pas besoin de se connaître les uns les autres. Par conséquent, le client n'a pas de cœur de réseau ou de "cœur de réseau virtuel" à gérer, et n'a pas à traiter de questions d'acheminement inter sites. En d'autres termes, dans le schéma décrit dans le présent document, un VPN n'est PAS une "superstructure" par dessus le réseau du SP.

Par rapport à la gestion des appareils de bordure, de claires limites administratives sont maintenues entre le SP et ses clients. Les clients ne sont pas obligés d'accéder aux routeurs PE ou P pour les besoins de la gestion, et le SP n'est pas obligé d'accéder aux appareils CE pour les besoins de gestion.

### 1.3 VPN avec espaces d'adresses en chevauchement

Si deux VPN n'ont pas de site en commun, ils peuvent alors avoir des espaces d'adresses qui se chevauchent. C'est-à-dire, une certaine adresse pourrait être utilisée dans le VPN V1 comme adresse du système S1, mais dans le VPN V2 comme adresse d'un système S2 complètement différent. C'est une situation courante quand les VPN utilisent chacun un espace d'adresses privées de la RFC 1918. Bien sûr, au sein de chaque VPN, chaque adresse doit être non ambiguë.

Même deux VPN qui ont des sites en commun peuvent avoir des espaces d'adresses qui se chevauchent, pour autant qu'il ne soit pas besoin de communications entre les systèmes avec de telles adresses et systèmes dans les sites communs.

### 1.4 VPN avec des chemins différents sur le même système

Bien qu'un site puisse être dans plusieurs VPN, il n'est pas nécessairement le cas que le chemin pour un certain système à ce site devrait être le même dans tous les VPN. Supposons, par exemple, qu'on ait un intranet consistant en les sites A, B, et C, et un extranet consistant en A, B, C, et le site "étranger" D. Supposons qu'au site A il y ait un serveur, et qu'on veuille que les clients de B, C, ou D soient capables d'utiliser ce serveur. Supposons aussi qu'au site B il y ait un pare-feu. On veut que tout le trafic du site D au serveur passe à travers le pare-feu, afin que le trafic provenant de l'extranet puisse être d'accès contrôlé. Cependant, on ne veut pas que le trafic provenant de C passe par le pare-feu sur le chemin du serveur, car c'est du trafic intranet.

Il est possible d'établir deux chemins pour le serveur. Un chemin, utilisé par les sites B et C, prend le trafic directement au site A. Le second chemin, utilisé par le site D, prend le trafic pour le pare-feu au site B. Si le pare-feu permet le passage du trafic, il va alors apparaître comme étant du trafic qui vient du site B, et suit le chemin du site A.

### 1.5 Routeurs de cœur de réseau de fournisseur de services

Le cœur de réseau du SP consiste en routeurs PE, ainsi que d'autres routeurs ("routeurs P") qui ne se rattachent pas aux appareils CE.

Si chaque routeur dans le cœur de réseau d'un SP devait conserver les informations d'acheminement pour tous les VPN pris en charge par le SP, il y aurait de sévères problèmes d'adaptabilité ; le nombre de sites qui pourraient être pris en charge devrait être limité par la quantité d'informations d'acheminement qui pourrait être détenue pas un seul routeur. Il est donc important que les informations d'acheminement sur un VPN particulier aient seulement besoin d'être présentes dans les routeurs PE qui se rattachent à ce VPN. En particulier, les routeurs P n'ont pas besoin d'avoir d'informations d'acheminement sur quel que VPN que ce soit. (Cette condition peut devoir être un peu relâchée quand l'acheminement de diffusion groupée est en cause. Il n'est pas considéré plus avant dans ce travail, mais est examiné dans la [RFC6037].)

Ainsi, tout comme les possesseurs de VPN n'ont pas de cœur de réseau ou de "cœur de réseau virtuel" à administrer, les SP eux-mêmes n'ont pas un cœur de réseau ou "cœur de réseau virtuel" séparé à administrer pour chaque VPN. L'acheminement de site à site dans le cœur de réseau est optimal (dans les contraintes des politiques utilisées pour former les VPN) et n'est contraint en aucune façon par une "topologie virtuelle" artificielle de tunnels.

La Section 10 discute certaines des questions particulières qui se posent quand le cœur de réseau s'étend sur plusieurs fournisseurs de services.

### 1.6 Sécurité

Les VPN de la sorte discutée ici, même sans utiliser de mesures de sécurité cryptographiques, sont destinés à fournir un niveau de sécurité équivalent à celui obtenu avec un cœur de réseau de couche 2 (par exemple, relais de trame). C'est-à-dire, en l'absence de mauvaise configuration ou d'interconnexion délibérée de différents VPN, il n'est pas possible que les systèmes dans un VPN obtiennent l'accès aux systèmes d'un autre VPN. Bien sûr, les méthodes décrites ici ne chiffrent pas par elles-même les données pour la confidentialité, ni ne fournissent de moyen de déterminer si les données ont été altérées en chemin. Si c'est désiré, des mesures de chiffrement doivent être appliquées en plus. (Voir, par exemple, [VPN-IP].) La sécurité est discutée plus en détails à la Section 13.

## 2. Sites et CE

Du point de vue d'un cœur de réseau particulier, un ensemble de systèmes IP peut être regardé comme un "site" si ces systèmes ont une inter connexité IP mutuelle qui n'exige pas l'usage du cœur de réseau. En général, un site va consister en un ensemble de systèmes qui sont géographiquement proches. Cependant, ceci n'est pas universellement vrai. Si deux localisations géographiques sont connectées via une ligne louée, sur laquelle fonctionne le protocole de plus court chemin ouvert en premier (OSPF, *Open Shortest Path First*) [RFC2328], et si cette ligne est la façon préférée de communiquer entre les deux localisations, alors les deux localisations peuvent être considérées comme un seul site, même si chaque localisation a son propre routeur CE. (Cette notion de "site" est topologique, plus que géographique. Si la ligne louée est en panne, ou cesse par ailleurs d'être le chemin préféré, mais si les deux localisations géographiques peuvent continuer de communiquer en utilisant le VPN cœur de réseau, alors ce site en devient deux.)

Un appareil CE est toujours regardé comme étant dans un seul site (mais comme on le verra au paragraphe 3.2, un site peut consister en de multiples "sites virtuels"). Un site peut cependant appartenir à plusieurs VPN.

Un routeur PE peut se rattacher aux appareils CE provenant d'un nombre quelconque de sites différents, que ces appareils CE soient dans le même VPN ou dans des VPN différents. Un appareil CE peut, pour sa robustesse, se rattacher à plusieurs routeurs PE, du même ou de différents fournisseurs de service. Si l'appareil CE est un routeur, le routeur PE et le routeur CE vont apparaître comme des adjacences de routeur l'un à l'autre.

Bien qu'on parle le plus souvent de "sites" comme étant l'unité de base de l'interconnexion, rien ici n'empêche un degré de granularité plus fin dans le contrôle de l'inter connexité. Par exemple, certains systèmes d'un site peuvent être membres d'un intranet ainsi que d'un ou plusieurs extranets, tandis que d'autres systèmes du même site peuvent être restreints à être seulement membres de l'intranet. Cependant, cela peut exiger que le site ait deux circuits de rattachement au cœur de réseau, un pour l'intranet et un pour l'extranet ; cela peut de plus exiger que la fonction de pare-feu soit appliquée sur le circuit de rattachement de l'extranet.

## 3. VRF : tableaux de transmission multiples dans les PE

Chaque routeur PE tient un certain nombre de tableaux de transmission séparés. Un des tableaux de transmission est le "tableau de transmission par défaut". Les autres sont des "tableaux d'acheminement et de transmission de VPN" (VRF, *VPN Routing and Forwarding table*).

### 3.1 VRF et circuits de rattachement

Chaque circuit de rattachement PE/CE est associé, par configuration, à un ou plusieurs VRF. Un circuit de rattachement qui est associé à un VRF est appelé un "circuit de rattachement de VRF".

Dans le cas le plus simple et normal, un circuit de rattachement PE/CE est associé à exactement un VRF. Quand un paquet IP est reçu sur un circuit de rattachement particulier, son adresse de destination IP est cherchée dans le VRF associé. Le résultat de cette recherche détermine comment acheminer le paquet. Le VRF utilisé par le PE d'entrée d'un paquet pour acheminer un certain paquet est appelé le "VRF d'entrée" du paquet. (Il y a aussi la notion de "VRF de sortie" d'un paquet,

situé au PE de sortie du paquet ; ceci est discuté à la Section 5.)

Si un paquet IP arrive sur un circuit de rattachement qui n'est associé à aucun VRF, l'adresse de destination du paquet est cherchée dans le tableau de transmission par défaut, et le paquet est acheminé en conséquence. Les paquets transmis conformément au tableau de transmission par défaut incluent des paquets provenant de routeurs PE ou P du voisinage, ainsi que de paquets provenant de circuits de rattachement face au client qui n'ont pas été associés à des VRF.

Intuitivement, on peut voir le tableau de transmission par défaut comme contenant des "chemins publics", et les VRF comme contenant des "chemins privés". On peut de façon similaire voir les circuits de rattachement de VRF comme étant "privés", et les circuits de rattachement non VRF comme étant "publics".

Si un circuit de rattachement de VRF particulier connecte le site S à un routeur PE, la connexité avec S (via ce circuit de rattachement) peut alors être restreinte en contrôlant l'ensemble de routes qui sont entrées dans le VRF correspondant. L'ensemble de routes dans ce VRF devrait être limité à l'ensemble de routes conduisant aux sites qui ont au moins un VPN en commun avec S. Alors un paquet envoyé de S sur un circuit de rattachement de VRF peut seulement être acheminé par le PE à un autre site S' si S' est dans un des mêmes VPN que S. C'est-à-dire que la communication (via les routeurs PE) est empêchée entre toute paire de sites de VPN qui n'ont pas de VPN en commun. La communication entre les sites de VPN et les sites non VPN est empêchée en gardant les routes pour les sites de VPN hors du tableau de transmission par défaut.

Si il y a plusieurs circuits de rattachement conduisant de S à un ou plusieurs routeurs PE, il peut alors y avoir plusieurs VRF qui pourraient être utilisés pour acheminer le trafic provenant de S. Pour restreindre correctement la connectivité de S, le même ensemble de routes devrait exister dans tous les VRF. Autrement, on pourrait imposer différentes restrictions de connexité sur différents circuits de rattachement à partir de S. Dans ce cas, certains des VRF associés aux circuits de rattachement à partir de S contiendraient des ensembles de routes différents de certains des autres.

On permet le cas où un seul circuit de rattachement est associé à un ensemble de VRF, plutôt qu'à un seul VRF. Cela peut être utile si on désire diviser un seul VPN en plusieurs "sous VPN", chacun ayant des restrictions de connexité différentes, où des caractéristiques des paquets de client sont utilisées pour choisir parmi les sous VPN. Pour rester simple, on parlera cependant généralement d'un circuit de rattachement associé à un seul VRF.

### 3.2 Association de paquets IP à des VRF

Quand un routeur PE reçoit un paquet d'un appareil CE, il doit déterminer le circuit de rattachement sur lequel le paquet est arrivé, car cela détermine à son tour le VRF (ou ensemble de VRF) qui peut être utilisé pour transmettre ce paquet. En général, pour déterminer le circuit de rattachement sur lequel un paquet est arrivé, un routeur PE prend note de l'interface physique sur laquelle le paquet est arrivé, et éventuellement prend aussi note de certains aspects de l'en-tête de couche 2 du paquet. Par exemple, si le circuit de rattachement d'entrée d'un paquet est un VC en relais de trame, l'identité du circuit de rattachement peut être déterminée à partir de l'interface physique de relais de trame sur laquelle le paquet est arrivé, ainsi que du champ Identifiant de connexion de liaison de données (DLCI, *Data Link Connection Identifier*) dans l'en-tête de relais de trame du paquet.

Bien que la conclusion du PE qu'un paquet particulier est arrivé sur un circuit de rattachement particulier puisse être partiellement déterminée par l'en-tête de couche 2 du paquet, il doit être impossible à un client, en écrivant les champs d'en-tête, de tromper le SP en l'amenant à penser qu'un paquet reçu sur un circuit de rattachement est en réalité arrivé sur un autre. Dans l'exemple ci-dessus, bien que le circuit de rattachement soit déterminé partiellement par l'inspection du champ DLCI dans l'en-tête de relais de trame, ce champ ne peut pas être réglé librement par le client. Il doit plutôt être réglé à une valeur spécifiée par le SP, ou autrement le paquet ne peut pas arriver au routeur PE.

Dans certains cas, un site particulier peut être divisé par le client en plusieurs "sites virtuels". Le SP peut désigner un ensemble particulier de VRF à utiliser pour acheminer les paquets à partir de ce site et peut permettre au client de régler certaines caractéristiques du paquet, qui sont alors utilisées pour choisir un VPN particulier parmi l'ensemble.

Par exemple, chaque site virtuel peut être réalisé comme un VLAN. Le SP et le client pourraient s'accorder pour que sur les paquets arrivant d'un CE particulier, certaines valeurs de VLAN soient utilisées pour identifier certains VRF. Bien sûr, les paquets provenant de ce CE vont être éliminés par le PE si ils portent des valeurs d'étiquettes de VLAN qui ne sont pas dans l'ensemble objet de l'accord. Une autre façon de réaliser cela est d'utiliser les adresses IP de source. Dans ce cas, le PE utilise l'adresse IP de source dans un paquet reçu du CE, ainsi que l'interface sur laquelle le paquet est reçu, pour allouer le paquet à un VRF particulier. Là encore, le client va seulement être capable de choisir parmi l'ensemble particulier de VRF que ce client est autorisé à utiliser.

Si on désire qu'un hôte particulier soit dans plusieurs sites virtuels, cet hôte doit alors déterminer, pour chaque paquet, à quel site virtuel le paquet est associé. Il peut le faire, par exemple, en envoyant des paquets à partir de différents sites virtuels sur différents VLAN, ou à partir de différentes interfaces réseau.

### 3.3 Remplissage des VRF

Avec quel ensemble de routes les VRF sont ils remplis ?

Par exemple, soient PE1, PE2, et PE3 trois routeurs PE, et soient CE1, CE2, et CE3 trois routeurs CE. Supposons que PE1 apprenne, de CE1, les routes qui sont accessibles au site de CE1. Si PE2 et PE3 sont rattachés, respectivement, à CE2 et CE3, et si il y a un VPN V contenant CE1, CE2, et CE3, alors PE1 utilise BGP pour distribuer à PE2 et PE3 les routes qu'il a apprises de CE1. PE2 et PE3 utilisent ces routes pour remplir les VRF auxquels ils sont respectivement associés, avec les sites de CE2 et CE3. Les routes provenant de sites qui ne sont pas dans le VPN V n'apparaissent pas dans ces VRF, ce qui signifie que les paquets provenant de CE2 ou CE3 ne peuvent pas être envoyés aux sites qui ne sont pas dans le VPN V.

Quand on parle de PE "apprenant" les chemins d'un CE, on ne présuppose aucune technique d'apprentissage particulière. Le PE peut apprendre les chemins au moyen d'un algorithme d'acheminement dynamique, mais il peut aussi "apprendre" les chemins en les ayant configurés (c'est-à-dire, acheminement statique). (Dans ce cas, dire que le PE "a appris" les chemins du CE est peut-être un peu un exercice de licence poétique.)

Les PE doivent aussi apprendre, des autres PE, les chemins qui appartiennent à un certain VPN. Les procédures à utiliser pour remplir les VRF avec les ensembles de routes appropriées sont spécifiées à la Section 4.

Si il y a plusieurs circuits de rattachement qui conduisent d'un certain routeur PE à un site particulier, ils peuvent être tous transposés dans le même tableau de transmission. Mais si la politique l'impose, ils pourraient être transposés dans des tableaux de transmission différents. Par exemple, la politique pourrait être qu'un circuit de rattachement particulier provenant d'un site est utilisé seulement pour le trafic intranet, tandis qu'un autre circuit de rattachement provenant de ce site est utilisé seulement pour le trafic extranet. (Peut-être, par exemple, le CE rattaché au circuit de rattachement extranet est un pare-feu, tandis que le CE rattaché au circuit de rattachement intranet ne l'est pas.) Dans ce cas, les deux circuits de rattachement vont être associés à des VRF différents.

Noter que si deux circuits de rattachement sont associés au même VRF, les paquets que le PE reçoit sur l'un d'eux vont alors être capables d'accéder exactement au même ensemble de destinations que les paquets que le PE reçoit sur l'autre. Ainsi deux circuits de rattachement ne peuvent pas être associés au même VRF si chaque CE n'est pas exactement dans le même ensemble de VPN que l'autre.

Si un circuit de rattachement conduit à un site qui est dans plusieurs VPN, le circuit de rattachement peut quand même être associé à un seul VRF, et dans ce cas le VRF va contenir des chemins provenant de tout l'ensemble de VPN dont le site est membre.

## 4. Distribution de chemin de VPN via BGP

Les routeurs PE utilisent BGP pour distribuer les chemins de VPN à chaque autre routeur (de façon plus précise, pour faire que les chemins de VPN soient distribués à chacun des autres).

On permet que chaque VPN ait son propre espace d'adresses, ce qui signifie qu'une certaine adresse peut noter des systèmes différents dans des VPN différents. Si deux chemins pour le même préfixe d'adresse IP sont en fait des chemins pour des systèmes différents, il est important de s'assurer que BGP ne les traite pas comme comparables. Autrement, BGP pourrait choisir d'installer seulement l'un d'eux, rendant l'autre système inaccessible. De plus, on doit s'assurer que POLICY est utilisé pour déterminer quels paquets sont envoyés sur quelles chemins ; étant donné que plusieurs de ces chemins sont installés par BGP, seulement un d'eux doit apparaître dans tout VRF.

On satisfera ces buts par l'utilisation d'une nouvelle famille d'adresses, comme spécifié ci-dessous.

### 4.1 Famille d'adresses VPN-IPv4

Les extensions BGP multi protocoles [RFC2858] permettent à BGP de porter des chemins provenant de plusieurs "familles d'adresses". On introduit la notion de "famille d'adresse VPN-IPv4". Une adresse VPN-IPv4 est une quantité de 12 octets,

commençant par un différenciateur de chemin (RD, *Route Distinguisher*) de huit octets et se terminant par une adresse IPv4 de quatre octets. Si plusieurs VPN utilisent le même préfixe d'adresse IPv4, les PE le traduisent en préfixes d'adresse IPv4 uniques de VPN-IPv4. Cela assure que si la même adresse est utilisée dans plusieurs VPN différents, il est possible à BGP de porter plusieurs chemins complètement différents pour cette adresse, une pour chaque VPN.

Comme les adresses de VPN-IPv4 et les adresses IPv4 sont des familles d'adresses différentes, BGP ne les traite jamais comme des adresses comparables.

Un RD est simplement un numéro, et il ne contient aucune information inhérente ; il n'identifie pas l'origine du chemin ou l'ensemble de VPN auxquels le chemin est à distribuer. L'objet du RD est seulement de permettre de créer des chemins distincts pour un préfixe commun d'adresses IPv4. D'autres moyens sont utilisés pour déterminer où redistribuer le chemin (voir le paragraphe 4.3).

Le RD peut aussi être utilisé pour créer plusieurs chemins différents pour le même système. On a déjà discuté d'une situation où le chemin pour un serveur particulier devrait être différent pour le trafic intranet et pour le trafic extranet. Cela peut être réalisé en créant deux chemins de VPN-IPv4 différents qui ont la même partie IPv4, mais des RD différents. Cela permet à BGP d'installer plusieurs chemins différents pour le même système, et permet d'utiliser des politiques (voir au paragraphe 4.3.5) pour décider quels paquets utilisent quel chemin.

Les RD sont structurés de façon telle que chaque fournisseur de services peut administrer son propre "espace de numérotation" (c'est-à-dire, peut faire ses propres allocations de RD) sans conflit avec les allocations de RD faites par tout autre fournisseur de services. Un RD consiste en trois champs : un champ Type de deux octets, un champ Administrateur, et un champ Numéro alloué. La valeur du champ Type détermine la longueur des deux autres champs, ainsi que la sémantique du champ Administrateur. Le champ Administrateur identifie une autorité d'allocation de numéros, et le champ Numéro alloué contient le numéro qui a été alloué, par l'autorité identifiée, pour un objet particulier. Par exemple, on pourrait avoir un RD dont le champ Administrateur contiendrait un numéro de système autonome (ASN, *Autonomous System Number*), et dont le champ Numéro (4 octets) contiendrait un numéro alloué par le SP à qui appartient cet ASN (ayant été alloué à ce SP par l'autorité appropriée).

Les RD ont reçu cette structure afin d'assurer qu'un SP qui fournit un service de cœur de réseau de VPN peut toujours créer un RD unique quand il en a besoin. Cependant, la structure n'est pas significative dans BGP ; quand BGP compare deux de ces préfixes d'adresse, il ignore entièrement la structure.

Un PE a besoin d'être configuré de telle façon que les chemins qui conduisent à un certain CE soient associés à un certain RD. La configuration peut amener à ce que tous les chemins qui conduisent au même CE soient associés au même RD, ou que des chemins différents soient associés à des RD différents, même si ils conduisent au même CE.

### 4.2 Codage des différenciateurs de chemins

Comme on l'a déclaré, une adresse de VPN-IPv4 consiste en un différenciateur de chemin de 8 octets suivi par une adresse IPv4 de 4 octets. Les RD sont codés comme suit :
- champ Type : 2 octets
- champ Valeur : 6 octets

L'interprétation du champ Valeur dépend de la valeur du champ Type. Pour l'instant, trois valeurs de champ Type sont définies : 0, 1, et 2.

- Type 0 : le champ Valeur comporte deux sous champs :
* sous champ Administrateur : 2 octets
* sous champ Numéro alloué : 4 octets
Le sous champ Administrateur doit contenir un numéro de système autonome. Si cet ASN est de l'espace d'ASN public, il doit avoir été alloué par l'autorité appropriée (l'utilisation de valeurs d'ASN provenant de l'espace privé d'ASN est fortement déconseillée). Le sous champ Numéro alloué contient un numéro provenant d'un espace de numérotation qui est administré par l'entreprise à laquelle l'ASN a été alloué par une autorité appropriée.

- Type 1 : le champ Valeur comporte deux sous champs :
* sous champ Administrateur : 2 octets
* sous champ Numéro alloué : 4 octets
Le sous champ Administrateur doit contenir une adresse IP. Si cette adresse IP est de l'espace d'adresses IP public, elle doit avoir été allouée par une autorité appropriée (l'utilisation d'adresses provenant de l'espace d'adresses IP privé est fortement

déconseillée). Le sous champ Numéro alloué contient un numéro provenant d'un espace de numérotation qui est administré par l'entreprise à laquelle l'adresse IP a été allouée par une autorité appropriée.

- Type 2 : le champ Valeur comporte deux sous champs :
* sous champ Administrateur : 4 octets
* sous champ Numéro alloué : 2 octets
Le sous champ Administrateur doit contenir un numéro de système autonome de quatre octets [RFC4893]. Si cet ASN est de l'espace d'ASN public, il doit avoir été alloué par l'autorité appropriée (l'utilisation de valeurs d'ASN provenant de l'espace privé d'ASN est fortement déconseillée). Le sous champ Numéro alloué contient un numéro provenant d'un espace de numérotation qui est administré par l'entreprise à laquelle l'ASN a été alloué par une autorité appropriée.

### 4.3 Contrôle de la distribution des chemins

Ce paragraphe expose la façon dont est contrôlée la distribution des chemins de VPN-IPv4.

Si un routeur PE est rattaché à un VPN particulier (en étant rattaché à un CE particulier dans ce VPN) il apprend certains des chemins IP de ce VPN du routeur CE rattaché. Les chemins appris d'un homologue d'acheminement CE sur un circuit de rattachement particulier peuvent être installés dans le VRF associé à ce circuit de rattachement. Quels chemins sont exactement installés de cette manière est déterminé par la façon dont le PE apprend les chemins du CE. En particulier, quand le PE et le CE sont des homologues de protocole d'acheminement, ceci est déterminé par le processus de décision du protocole d'acheminement ; ceci est discuté à la Section 7.

Ces chemins sont alors convertis en chemins VPN-IP4, et "exportés" à BGP. Si il y a plus d'un chemin pour un préfixe d'adresse de VPN-IPv4 particulier, BGP choisit "le meilleur", en utilisant le processus de décision BGP. Ce chemin est alors distribué par BGP à l'ensemble des autres PE qui ont besoin d'en connaître. À ces autres PE, BGP va là encore choisir le meilleur chemin pour un préfixe d'adresse de VPN-IPv4 particulier. Alors les chemins de VPN-IPv4 choisis sont reconvertis en chemins IP, et "importés" dans un ou plusieurs VRF. Si ils sont réellement installés dans les VRF dépend du processus de décision de la méthode d'acheminement utilisée entre le PE et les CE qui sont associés au VRF en question. Finalement, tout chemin installé dans un VRF peut être distribué aux routeurs CE associés.

#### 4.3.1 Attribut Cible de chemin

Chaque VRF est associé à un ou plusieurs attributs Cible de chemin (RT, *Route Target*).

Quand un chemin VPN-IPv4 est créé (à partir d'un chemin IPv4 que le PE a appris d'un CE) par un routeur PE, il est associé à un ou plusieurs attributs Cible de chemin. Ceux-ci sont portés dans BGP comme attributs du chemin.

Tout chemin associé à la cible de chemin T doit être distribué à tout routeur PE qui a un VRF associé à la cible de chemin T. Quand un tel chemin est reçu par un routeur PE, il est éligible à être installé dans les VRF de ce PE qui sont associés à la cible de chemin T. (Si il est réellement installé dépend du résultat du processus de décision BGP, et du résultat du processus de décision de l'IGP (c'est-à-dire, le protocole d'acheminement intra domaine) qui fonctionne sur l'interface PE/CE.)

Un attribut Cible de chemin peut être vu comme identifiant un ensemble de sites. (Bien qu'il serait plus précis de le voir comme identifiant un ensemble de VRF.) Associer un attribut Cible de chemin particulier à un chemin permet que ce chemin soit placé dans les VRF qui sont utilisés pour acheminer le trafic reçu des sites correspondants.

Il y a un ensemble de cibles de chemin qu'un routeur PE rattache à un chemin reçu du site S ; elles peuvent être appelées les "cibles exportées". Et il y a un ensemble de cibles de chemin qu'un routeur PE utilise pour déterminer si un chemin reçu d'un autre routeur PE pourrait être placé dans le VRF associé au site S ; elles peuvent être appelées les "cibles importées". Les deux ensembles sont distincts, et n'ont pas besoin d'être les mêmes. Noter qu'un chemin VPN-IPv4 particulier est seulement éligible à l'installation dans un VRF particulier si il y a une cible de chemin qui est à la fois une des cibles de chemin du chemin et une des cibles importées du VRF.

La fonction exercée par l'attribut Cible de chemin est similaire à celle exercée par l'attribut BGP Communautés. Cependant, le format de ce dernier est inadéquat pour le présent objet, car il permet seulement un espace de numérotation de deux octets. Il est souhaitable de structurer le format, de façon similaire à ce qui a été décrit pour les RD (voir au paragraphe 4.2) afin qu'un champ Type définisse la longueur d'un champ Administrateur, et que le reste de l'attribut soit un numéro provenant de l'espace de numérotation spécifié par l'administrateur. Cela peut être fait en utilisant les Communautés

étendues de BGP. Les cibles de chemin discutées ici sont codées comme des cibles de chemin de communauté étendu BGP [RFC4360]. Leur structure est similaire à celle des RD.

Quand un locuteur BGP a reçu plus d'un chemin sur le même préfixe VPN-IPv4, les règles de BGP pour la préférence de chemin sont utilisées pour choisir quel chemin VPN-IPv4 est installé par BGP.

Noter qu'un chemin peut seulement avoir un RD, mais il peut avoir plusieurs cibles de chemin. Dans BGP, l'adaptabilité est améliorée si on a un seul chemin avec plusieurs attributs, par opposition à plusieurs chemins. On pourrait éliminer l'attribut Cible de chemin en créant plus de chemins (c'est-à-dire, en utilisant plus de RD) mais les propriétés d'adaptabilité seraient moins favorables.

Comment un PE détermine t-il quels attributs Cible de chemin associer à un certain chemin ? Il y a un certain nombre de différentes façons possibles. Le PE pourrait être configuré à associer tous les chemins qui conduisent à un site spécifié avec une cible de chemin spécifiée. Ou le PE pourrait être configuré à associer certains chemins conduisant à un site spécifié avec une cible de chemin, et certains avec une autre.

Si le PE et le CE sont eux-mêmes des homologues BGP (voir la Section 7) alors le SP peut permettre au client, dans certaines limites, de spécifier comment ses chemins sont à distribuer. Le SP et le client vont avoir besoin de se mettre d'accord à l'avance sur l'ensemble de RT qui sont autorisés à se rattacher aux chemins de VPN du client. Le CE pourrait alors rattacher un ou plusieurs de ces RT à chaque chemin IP qu'il distribue au PE. Cela donne au client la liberté de spécifier en temps réel, dans les limites de l'accord, ses politiques de distribution de chemins. Si il est permis au CE de rattacher les RT à ses chemins, le PE DOIT filtrer tous les chemins qui contiennent des RT que le client n'est pas autorisé à utiliser. Si le CE n'est pas autorisé à rattacher des RT à ses chemins, mais le fait de toutes façons, le PE DOIT retirer le RT avant de convertir le chemin du client en chemin VPN-IPv4.

### 4.3.2 Distribution de chemins aux PE par BGP

Si deux sites d'un VPN se rattachent aux PE qui sont dans le même système autonome, les PE peuvent distribuer les chemins VPN-IPv4 à chacun des autres au moyen d'une connexion IBGP entre eux. (Le terme "IBGP" se réfère à l'ensemble de protocoles et procédures utilisé quand il y a une connexion BGP entre deux locuteurs BGP dans le même système autonome. Ceci se distingue de "EBGP", l'ensemble de procédures utilisées entre deux locuteurs BGP dans des systèmes autonomes différents.) Autrement, chacun peut avoir une connexion IBGP à un réflecteur de chemin [RFC2796].

Quand un routeur PE distribue un chemin VPN-IPv4 via BGP, il utilise sa propre adresse comme "prochain bond BGP". Cette adresse est codée comme une adresse VPN-IPv4 avec un RD de 0. (La [RFC2858] exige que l'adresse de prochain bond soit dans la même famille d'adresses que les informations d'accessibilité de la couche réseau (NLRI, *Network Layer Reachability Information*).) Elle alloue aussi et distribue une étiquette MPLS. (Essentiellement, les routeurs PE ne distribuent pas de chemins VPN-IPv4, mais des chemins VPN-IPv4 étiquetés. Voir la [RFC3107].) Quand le PE traite un paquet reçu qui a cette étiquette au sommet de la pile, le PE va éclater la pile, et traiter le paquet de façon appropriée.

Le PE peut distribuer l'ensemble exact de chemins qui apparaît dans le VRF, ou il peut effectuer un résumé et distribuer des agrégats de ces chemins, ou il peut faire un peu de l'un et un peu de l'autre.

Supposons qu'un PE ait alloué l'étiquette L au chemin R, et ait distribué cette transposition d'étiquette via BGP. Si R est un agrégat d'un ensemble de chemins dans le VRF, le PE va savoir que les paquets provenant du cœur de réseau qui arrivent avec cette étiquette doivent avoir leurs adresses de destination recherchées dans un VRF. Quand le PE cherche l'étiquette dans sa base de données d'informations d'étiquettes, il apprend quel VRF doit être utilisé. Par ailleurs, si R n'est pas un agrégat, alors quand le PE cherche l'étiquette, il apprend le circuit de rattachement de sortie, ainsi que l'en-tête d'encapsulation pour le paquet. Dans ce cas, aucune recherche n'est faite dans le VRF.

On s'attendrait à ce que le cas le plus courant soit celui où le chemin N'est PAS un agrégat. Le cas où il est un agrégat peut être très utile même si le VRF contient un grand nombre de chemins d'hôtes (par exemple, comme dans un appel entrant) ou si le VRF a une interface associée de réseau de zone locale (LAN, *Local Area Network*) (où il y a un en-tête de couche 2 sortant différent pour chaque système sur le LAN, mais où un chemin n'est pas distribué pour chacun de ces systèmes).

Si chaque chemin a ou non une étiquette distincte est une affaire de mise en œuvre. Il y a un certain nombre d'algorithmes possibles qu'on pourrait utiliser pour déterminer si deux chemins ont la même étiquette allouée :

- On peut choisir d'avoir une seule étiquette pour tout un VRF, de sorte qu'une seule étiquette soit partagée par tous les chemins venant de ce VRF. Alors quand le PE de sortie reçoit un paquet qui a cette étiquette, il doit chercher l'adresse

de destination IP du paquet dans ce VRF (le "VRF de sortie" du paquet) afin de déterminer le circuit de rattachement de sortie du paquet et l'encapsulation de liaison des données correspondante.

- On peut choisir d'avoir une seule étiquette pour chaque circuit de rattachement, de sorte qu'une seule étiquette soit partagée par tous les chemins avec le même "circuit de rattachement sortant". Cela permet d'éviter de faire une recherche dans le VRF de sortie, bien qu'une puisse devoir être faite afin de déterminer l'encapsulation de liaison des données, par exemple, une recherche du protocole de résolution d'adresse (ARP, *Address Resolution Protocol*).

- On peut choisir d'avoir une étiquette distincte pour chaque chemin. Alors, si un chemin est potentiellement accessible sur plus d'un circuit de rattachement, l'acheminement PE/CE peut passer au chemin préféré pour un chemin d'un circuit de rattachement à un autre, sans qu'il soit besoin de distribuer une nouvelle étiquette pour ce chemin.

Il peut aussi y avoir d'autres algorithmes possibles. Le choix de l'algorithme est entièrement à la discrétion du PE de sortie, et est par ailleurs transparent.

En utilisant de cette manière les étiquettes MPLS distribuée par BGP, on présuppose qu'un paquet MPLS portant une telle étiquette peut être tunnelé à partir du routeur qui installe le chemin correspondant distribué par BGP au routeur qui est le prochain bond BGP de ce chemin. Cela exige qu'un chemin à commutation d'étiquette existe entre ces deux routeurs ou que quelque autre technologie de tunnelage (par exemple, [RFC4023]) puisse être utilisée entre eux.

Ce tunnel peut suivre un chemin "au mieux", ou il peut suivre un chemin à ingénierie du trafic. Entre une certaine paire de routeurs, il peut y avoir un tel tunnel, ou il il peut y en avoir plusieurs, peut-être avec des caractéristiques de qualité de service (QS) différentes. Tout ce qui importe pour l'architecture de VPN est que ces tunnels existent. Pour assurer l'interopérabilité entre les systèmes qui mettent en œuvre cette architecture de VPN en utilisant des chemins à commutation d'étiquette MPLS comme technologie de tunnelage, tous les systèmes de cette sorte DOIVENT prendre en charge le protocole de distribution d'étiquettes (LDP, *Label Distribution Protocol*) [RFC3036]. En particulier, le mode non sollicité vers l'aval DOIT être pris en charge sur les interfaces qui ne sont ni ATM contrôlé par étiquettes (LC-ATM, *Label Controlled ATM*) [RFC3035] ni relais de trame contrôlé par étiquettes (LC-FR, *Label Controlled Frame Relay*) [RFC3034], et le mode à la demande vers l'aval DOIT être pris en charge sur les interfaces LC-ATM et LC-FR.

Si le tunnel suit un chemin au mieux, alors le PE trouve le chemin vers le point d'extrémité distant en cherchant son adresse IP dans le tableau de transmission par défaut.

Un routeur PE, SAUF si il est un réflecteur de chemin (voir le paragraphe 4.3.3) ou un routeur frontière de système autonome (ASBR, *Autonomous System Border Router*) pour un VPN inter fournisseurs (voir la Section 10) ne devrait pas installer un chemin de VPN-IPv4 sauf si il a au moins un VRF avec une cible d'importation identique à un des attributs de cible de chemin du chemin. Le filtrage entrant devrait être utilisé pour causer l'élimination de tels chemins. Si une nouvelle cible d'importation est ultérieurement ajoutée à un des VRF du PE (une opération"VPN Join") elle doit alors acquérir les chemins qu'elle peut avoir éliminé précédemment. Cela peut être fait en utilisant le mécanisme de rafraîchissement décrit dans la [RFC2918]. Le mécanisme de filtrage de chemin sortant de la [RFC5291] peut aussi être utilisé pour rendre le filtrage plus dynamique.

De même, si une cible d'importation particulière n'est plus présente dans aucun des VRF d'un PE (par suite d'une ou plusieurs opérations "d'élagage de VPN") le PE peut éliminer tous les chemins qui, par suite, n'ont plus aucune cible d'importation de VRF de PE comme un de leurs attributs de cible de chemin.

Un routeur qui n'est rattaché à aucun VPN et qui n'est pas un réflecteur de chemin (c'est-à-dire, un routeur P) n'installe jamais de chemin de VPN-IPv4.

Noter que les opérations VPN Join et Prune sont non interruptives et n'exigent qu'aucune connexion BGP soit arrêtée, pour autant 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 VPN ; ceci est une importante considération d'adaptabilité.

### 4.3.3 Utilisation des réflecteurs de chemins

Plutôt que d'avoir un maillage IBGP complet parmi les PE, il est avantageux d'utiliser les réflecteurs de chemin BGP [RFC2796] pour améliorer l'adaptabilité. Toutes les techniques usuelles d'utilisation des réflecteurs de chemin pour améliorer l'adaptabilité (par exemple, des hiérarchies de réflecteurs de chemin) sont disponibles.

Les réflecteurs de chemin sont les seuls systèmes qui ont besoin d'avoir des informations d'acheminement pour les VPN auxquels ils ne sont pas directement rattachés. Cependant, il n'est pas nécessaire qu'un réflecteur de chemin connaisse tous les chemins VPN-IPv4 pour tous les VPN pris en charge par le cœur de réseau.

On présente ci-dessous deux façons différentes de partager l'ensemble de chemins de VPN-IPv4 parmi un ensemble de réflecteurs de chemin.

1. Chaque réflecteur de chemin est pré configuré avec une liste de cibles de chemin. Pour avoir une redondance, plus d'un réflecteur de chemin peut être pré configuré avec la même liste. Un réflecteur de chemin utilise la liste pré configurée de cibles de chemin pour construire son filtrage de chemin entrant. Le réflecteur de chemin peut utiliser les techniques de la [RFC5291] pour installer sur chacun de ses homologues (sans considérer si l'homologue est un autre réflecteur de chemin ou un PE) l'ensemble de filtres de chemins sortants (ORF, *Outbound Route Filter*) qui contient la liste de ses cibles de chemin pré configurées. Noter que les réflecteurs de chemin devraient accepter les ORF provenant d'autres réflecteurs de chemin, ce qui signifie que les réflecteurs de chemin devraient annoncer la capacité ORF aux autres réflecteurs de chemin.

   Un fournisseur de services peut modifier la liste des cibles de chemin pré configurées sur un réflecteur de chemin. Quand c'est fait, le réflecteur de chemin modifie les ORF qu'il installe sur tous ses homologues IBGP. Pour réduire la fréquence des changements de configuration sur les réflecteurs de chemin, chaque réflecteur de chemin peut être pré configuré avec un bloc de cibles de chemin. De cette façon, quand une nouvelle cible de chemin est nécessaire pour un nouveau VPN, il y a déjà un ou plusieurs réflecteurs de chemin qui sont (pré)configurés avec cette cible de chemin.

   Sauf si un PE est un client de tous les réflecteurs de chemin, quand un nouveau VPN est ajouté au PE ("jonction de VPN") il va devoir devenir un client du ou des réflecteurs de chemin qui tiennent les routes pour ce VPN. De même, supprimer un VPN existant du PE ("élagage de VPN") peut résulter en une situation où le PE n'a plus besoin d'être un client d'un réflecteur de chemin. Dans l'un et l'autre cas, les opérations Join ou Prune ne sont pas interruptives (pour autant que la [RFC2918] soit utilisée, et n’exigent jamais qu'une connexion BGP soit arrêtée, mais seulement qu'elle soit rétablie.

   (Par "ajouter un nouveau VPN à un PE", on veut dire en réalité ajouter une nouvelle cible de chemin importée de ses VRF, ou d'ajouter un nouveau VRF avec une cible de chemin importée que n'avait aucun autre VRF du PE.)

2. Une autre méthode est de faire que chaque PE soit un client d'un sous ensemble des réflecteurs de chemin. Un réflecteur de chemin n'est pas pré configuré avec la liste des cibles de chemin, et n'effectue pas de filtrage de chemin entrant des routes reçues de ses clients (les PE) ; il accepte plutôt tous les chemins reçus de tous ses clients (PE). Le réflecteur de chemin garde trace de l'ensemble des cibles de chemin portées par tous les chemins qu'il reçoit. Quand le réflecteur de chemin reçoit de son client un chemin avec une cible de chemin qui n'est pas dans cet ensemble, cette cible de chemin est immédiatement ajoutée à l'ensemble. Par ailleurs, quand le réflecteur de chemin n'a plus de chemin avec une cible de chemin particulière qui est dans l'ensemble, le réflecteur de chemin devraient retarder (de quelques heures) la suppression de cette cible de chemin de l'ensemble.

Le réflecteur de chemin utilise cet ensemble pour former les filtres de chemins entrants qu'il applique aux routes reçues des autres réflecteurs de chemin. Le réflecteur de chemin peut aussi utiliser des ORF pour installer le filtrage approprié de chemin sortant sur les autres réflecteurs de chemin. Tout comme avec la première approche, un réflecteur de chemin devrait accepter les ORF provenant des autres réflecteurs de chemin. Pour faire cela, un réflecteur de chemin annonce la capacité d'ORF aux autres réflecteurs de chemin.

Quand le réflecteur de chemin change l'ensemble, il devrait immédiatement changer son filtrage de chemin entrant. De plus, si le réflecteur de chemin utilise des ORF, alors les ORF doivent être immédiatement changés pour refléter les changements de l'ensemble. Si le réflecteur de chemin n'utilise pas d'ORF, et si une nouvelle cible de chemin est ajoutée à l'ensemble, le réflecteur de chemin, après avoir changé son filtrage de chemins entrants, doit produire un Refresh BGP aux autres réflecteurs de chemin.

Le délai de "quelques heures" mentionné plus haut permet à un réflecteur de chemin de conserver les routes avec un certain RT, même après avoir perdu le dernier de ses clients intéressés à ces routes. Cela protège contre le besoin de ré acquérir tous ces chemins si la "disparition" des clients est seulement temporaire.

Avec cette procédure, les opérations VPN Join et Prune sont aussi non interruptives.

Noter que cette technique ne va pas fonctionner correctement si un PE client a un VRF avec une cible de chemin importée

qui n'est pas une de ses cibles de chemin exportées.

Dans ces procédures, un routeur PE qui se rattache à un VPN particulier "auto découvre" les autres PE qui se rattachent au même VPN. Quand un nouveau routeur PE est ajouté, ou quand un routeur PE existant se rattache à un nouveau VPN, aucune reconfiguration des autres routeurs PE n'est nécessaire.

Tout comme aucun routeur PE n'a besoin de connaître tous les chemins de VPN-IPv4 pris en charge sur le cœur de réseau, ces règles de distribution assurent qu'il n'y a pas un seul réflecteur de chemin (RR, *Route Reflector*) qui ait besoin de connaître tous les chemins de VPN-IPv4 pris en charge sur le cœur de réseau. Par suite, le nombre total de ces chemins qui peuvent être pris en charge sur le cœur de réseau n'est pas limité par la capacité d'un seul appareil, et donc peut augmenter virtuellement sans borne.

### 4.3.4 Portage des NLRI de VPN-IPv4 dans BGP

Les extensions multi protocoles BGP [RFC2858] sont utilisées pour coder les NLRI. Si le champ Identifiant de famille d'adresses (AFI, *Address Family Identifier*) est réglé à 1, et si le champ Identifiant de famille d'adresse suivante (SAFI, *Subsequent Address Family Identifier*) est réglé à 128, les NLRI sont une adresse VPN-IPv4 étiquetée MPLS. L'AFI 1 est utilisé car le protocole de couche réseau associé aux NLRI est toujours IP. Noter que cette architecture de VPN n'exige pas la capacité de distribuer des adresses VPN-IPv4 non étiquetées.

Afin que deux locuteurs BGP échangent des NLRI VPN-IPv4 étiquetées, ils doivent utiliser des annonces de capacité BGP pour s'assurer qu'ils sont tous deux capables de traiter correctement de telles NLRI. Ceci est fait comme spécifié dans la [RFC2858], en utilisant le code de capacité 1 (BGP multi protocoles) avec un AFI de 1 et un SAFI de 128.

Les NLRI de VPN-IPv4 étiquetées elles mêmes sont codées comme spécifié dans la [RFC3107], où le préfixe consiste en un RD de 8 octets suivi par un préfixe IPv4.

### 4.3.5 Construction de VPN en utilisant les cibles de chemins

On peut construire différentes sortes de VPN en réglant de façon appropriée les cibles d'importation et d'exportation.

Supposons qu'on désire créer un groupe clos d'utilisateurs pleinement maillé, c'est-à-dire, un ensemble de sites où chacun peut envoyer du trafic directement aux autres, mais le trafic ne peut pas être envoyé ou reçu d'autres sites. Chaque site est alors associé à un VRF, un seul attribut de cible de chemin est choisi, cette cible de chemin est allouée à chaque VRF à la fois comme cible d'importation et d'exportation, et cette cible de chemin n'est pas allouée à d'autres VRF, ni comme cible d'importation ni comme cible d'exportation.

Autrement, supposons qu'on désire, quelle qu'en soit la raison, créer une sorte de VPN "concentrateur radial" *(hub and spoke)*. Ce pourrait être fait par l'utilisation de deux valeurs de cible de chemin, l'une signifiant "concentrateur" et l'autre signifiant "rayons". Aux VRF rattachés au sites de concentrateur, "concentrateur" est la cible d'export et "rayons" est la cible d'export. Aux VRF rattachés aux sites de rayons, "concentrateur" est la cible d'import et "rayons" est la cible d'export.

Donc, les méthodes pour contrôler la distribution des informations d'acheminement parmi les divers ensembles de sites sont très souples, ce qui à son tour fournit une grande souplesse pour construire les VPN.

### 4.3.6 Distribution de chemins aux VRF dans un seul PE

Il est possible de distribuer les chemins d'un VRF à un autre, même si les deux VRF sont dans le même PE, et même si dans ce cas on ne peut pas dire que le chemin a été distribué par BGP. Néanmoins, la décision de distribuer un chemin particulier provenant d'un VRF à un autre au sein d'un seul PE est la même décision que celle qui serait prise si les VRF étaient sur des PE différents. C'est-à-dire, elle dépend de l'attribut de cible de chemin alloué au chemin (ou lui serait alloué si le chemin était distribué par BGP) et de la cible d'importation du second VRF.

## 5. Transmission

Si les routeurs intermédiaires dans le cœur de réseau n'ont aucune information sur les chemins pour les VPN, comment les paquets sont ils transmis d'un site de VPN à un autre ?

Quand un PE reçoit un paquet IP provenant d'un appareil CE, il choisit un VPN particulier dans lequel chercher l'adresse de destination du paquet. Ce choix se fonde sur le circuit de rattachement d'entrée du paquet.

En supposant qu'une correspondance soit trouvée, on apprend le "prochain bond" du paquet.

Si le "prochain bond" du paquet est atteint directement sur un circuit de rattachement de VRF provenant de ce PE (c'est-à-dire, si le circuit de rattachement de sortie du paquet est sur le même PE que son circuit de rattachement d'entrée) alors le paquet est envoyé sur le circuit de rattachement de sortie, et aucune étiquette MPLS n'est poussée sur la pile d'étiquette du paquet.

Si les circuits de rattachement d'entrée et de sortie sont sur le même PE, mais sont associés à des VRF différents, et si le chemin qui correspond le mieux à l'adresse de destination dans le VRF du circuit de rattachement d'entrée est un agrégat de plusieurs chemins dans le VRF du circuit de rattachement de sortie, il peut être nécessaire de chercher aussi l'adresse de destination du paquet dans le VRF de sortie.

Si le prochain bond du paquet N'est PAS atteint par un circuit de rattachement du VRF, le paquet doit alors voyager au moins un bond à travers le cœur de réseau. Le paquet a donc un "prochain bond BGP", et le prochain bond BGP va avoir une étiquette MPLS allouée pour le chemin qui correspond le mieux à l'adresse de destination du paquet. On appelle cette étiquette "étiquette de chemin de VPN". Le paquet IP est transformé en paquet MPLS avec l'étiquette de chemin de VPN comme seule étiquette de la pile d'étiquettes.

Le paquet doit alors être tunnelé au prochain bond BGP.

Si le cœur de réseau prend en charge MPLS, cela se fait comme suit :

- Les routeurs PE (et tout routeur de bordure de système autonome) qui redistribuent les adresses de VPN-IPv4 doivent insérer des préfixes d'adresses /32 pour eux-mêmes dans les tableaux d'acheminement IGP du cœur de réseau. Cela permet à MPLS, à chaque nœud dans le cœur de réseau, d'allouer une étiquette correspondant au chemin à chaque routeur PE. Pour assurer l'interopérabilité parmi différentes mises en œuvre, il est exigé de prendre en charge LDP pour établir les chemins commutés par étiquette à travers le cœur de réseau. Cependant, d'autres méthodes d'établissement de ces chemins commutés par étiquette sont aussi possibles. (Certaines de ces autres méthodes peuvent ne pas exiger la présence des préfixes d'adresses /32 dans l'IGP.)

- Si il y a des tunnels à ingénierie du trafic pour le prochain bond BGP, et si un ou plusieurs d'entre eux est disponible pour être utilisé par le paquet en question, un de ces tunnels est choisi. Ce tunnel va être associé à une étiquette MPLS, "l'étiquette de tunnel". L'étiquette de tunnel est poussée sur la pile d'étiquettes MPLS, et le paquet est transmis au prochain bond du tunnel.

- Autrement,
* Le paquet va avoir un "prochain bond IGP", qui est le prochain bond le long du chemin IGP au prochain bond BGP.
* Si le prochain bond BGP et le prochain bond IGP sont les mêmes, et si le saut de l'avant dernier bond est utilisé, le paquet est alors envoyé au prochain bond IGP, portant seulement l'étiquette de chemin de VPN.
* Autrement, le prochain bond IGP va avoir une étiquette allouée pour le chemin qui correspond le mieux à l'adresse du prochain bond BGP. On appelle cela "l'étiquette de tunnel". L'étiquette de tunnel est poussée sur l'étiquette de sommet du paquet. Le paquet est alors transmis au prochain bond IGP.

- MPLS va alors porter le paquet à travers le cœur de réseau au prochain bond BGP, où l'étiquette de VPN va être examinée.

Si le cœur de réseau ne prend pas en charge MPLS, le paquet MPLS portant seulement l'étiquette de chemin de VPN peut être tunnelé au prochain bond BGP en utilisant les techniques de la [RFC4023]. Quand le paquet émerge du tunnel, il va être au prochain bond BGP, où l'étiquette de chemin de VPN va être examinée.

Au prochain bond BGP, le traitement du paquet dépend de l'étiquette de chemin de VPN (voir au paragraphe 4.3.2). Dans de nombreux cas, le PE va être capable de déterminer, à partir de l'étiquette, le circuit de rattachement sur lequel le paquet devrait être transmis (à un appareil CE) ainsi que l'en-tête de couche de liaison des données approprié pour cette interface. Dans d'autres cas, le PE peut seulement être capable de déterminer que l'adresse de destination du paquet doit être recherchée dans un VRF particulier avant d'être transmis à un appareil CE. Il y a aussi des cas intermédiaires où l'étiquette de chemin de VPN peut déterminer le circuit de rattachement de sortie du paquet, mais une recherche (par exemple, ARP) doit quand même être faite afin de déterminer l'en-tête de couche de liaison du paquet sur ce circuit de rattachement.

Les informations dans l'en-tête MPLS lui-même, et/ou les informations associées à l'étiquette, peuvent aussi être utilisées pour fournir la qualité de service sur l'interface vers le CE.

En tous cas, si le paquet était un paquet IP non étiqueté quand il est arrivé à son PE d'entrée, il va à nouveau être un paquet non étiqueté quand il quittera son PE de sortie.

Le fait que les paquets avec des étiquettes de chemin de VPN sont tunnelés à travers le cœur de réseau est ce qui rend possible de garder tous les chemins de VPN hors des routeurs P. Ceci est crucial pour assurer l'adaptabilité du schéma. Le cœur de réseau n'a même pas besoin d'avoir les chemins pour les CE, seulement pour les PE.

Par rapport aux tunnels, on notera que la présente spécification :
- N'exige PAS que les tunnels soient en point à point ; le multipoint à point peut être utilisé ;
- N'exige PAS qu'il y ait un établissement explicite des tunnels, via la signalisation ou via configuration manuelle ;
- N'exige PAS qu'il y ait de signalisation spécifique du tunnel ;
- N'exige PAS qu'il y ait d'état spécifique de tunnel dans les routeurs P ou PE, au delà de ce qui est nécessaire pour conserver les informations d'acheminement et (si utilisé) les informations d'étiquette MPLS.

Bien sûr, la présente spécification est compatible avec l'utilisation de tunnels point à point qui doivent être explicitement configurés et/ou signalés, et dans certaines situations il peut y avoir des raisons d'utiliser de tels tunnels.

Les considérations pertinentes pour le choix d'une technologie de tunnelage particulière sortent du domaine de la présente spécification.

## 6. Maintenir un isolement approprié des VPN

Pour maintenir un isolement approprié d'un VPN par rapport à un autre, il est important qu'aucun routeur dans le cœur de réseau n'accepte un paquet tunnelé provenant de l'extérieur du cœur de réseau, sauf si il est sûr que les deux points d'extrémité de ce tunnel sont extérieurs au cœur de réseau.

Si MPLS est utilisé comme technologie de tunnelisation, cela signifie qu'un routeur dans le cœur de réseau NE DOIT PAS accepter un paquet étiqueté provenant de tout appareil adjacent non de cœur de réseau sauf si les deux conditions suivantes sont satisfaites :
1. l'étiquette au sommet de la pile d'étiquettes a réellement été distribuée par ce routeur de cœur de réseau à cet appareil de non cœur de réseau, et
2. le routeur de cœur de réseau peut déterminer que l'utilisation de cette étiquette va causer le départ du paquet du cœur de réseau avant que toute étiquette plus bas dans la pile soit inspectée, et avant que l'en-tête IP soit inspecté.

La première condition assure que tous les paquets étiquetés reçus de routeurs non de cœur de réseau ont une étiquette légitime et correctement allouée au sommet de la pile d'étiquettes. La seconde condition assure que les routeurs de cœur de réseau ne vont jamais chercher en dessous de cette étiquette sommitale. Bien sûr, la façon la plus simple de satisfaire ces deux conditions est juste de faire que les appareils de cœur de réseau refusent d'accepter les paquets étiquetés provenant d'appareils non de cœur de réseau.

Si MPLS n'est pas utilisé comme technologie de tunnelage, alors le filtrage doit être fait pour assurer qu'un paquet MPLS-dans-IP ou MPLS-dans-GRE peut être accepté dans le cœur de réseau seulement si l'adresse de destination IP du paquet va causer son envoi en dehors du cœur de réseau.

## 7. Comment les PE apprennent les chemins des CE

Les routeurs PE qui se rattachent à un VPN particulier ont besoin de savoir, pour chaque circuit de rattachement conduisant à ce VPN, quelles adresses du VPN devraient être atteintes sur ce circuit de rattachement.

Le PE traduit ces adresses en adresses de VPN-IPv4, en utilisant un RD configuré. Le PE traite alors ces chemins VPN-IPv4 comme des entrées à BGP. Les chemins provenant d'un site de VPN NE sont PAS communiqués à l'IGP du cœur de réseau.

Exactement quelles techniques de distribution de chemin PE/CE sont possibles dépend de si un CE particulier CE est ou

non dans un "VPN de transit". Un "VPN de transit" est celui qui contient un routeur qui reçoit des routes d'un "tiers" (c'est-à-dire, d'un routeur qui n'est pas dans le VPN, mais n'est pas un routeur PE) et qui redistribue ces routes à un routeur PE. Un VPN qui n'est pas un VPN de transit est un "VPN d'extrémité". La vaste majorité des VPN, incluant presque tous les réseaux d'entreprises, vont être supposés être "d'extrémité" dans ce sens.

Les techniques de distribution de PE/CE possibles sont :

1. L'acheminement statique (c'est-à-dire, par configuration) peut être utilisé. (Cela va probablement n'être utile que dans les VPN d'extrémité.)

2. Les routeurs PE et CE peuvent être des homologues du protocole d'informations d'acheminement (RIP) [RFC2453], et le CE peut utiliser RIP pour dire au routeur PE l'ensemble de préfixes d'adresses accessibles au site du routeur CE. Quand RIP est configuré dans le CE, il faut veiller à s'assurer que les préfixes d'adresses provenant d'autres sites (c'est-à-dire, les préfixes d'adresses appris par le routeur CE du routeur PE) ne sont jamais annoncés au PE. Plus précisément : si un routeur PE, disons, PE1, reçoit un chemin VPN-IPv4 R1, et si par suite il distribue un chemin IPv4 R2 à un CE, R2 ne doit alors pas être redistribué de ce site de CE à un routeur PE, disons, PE2, (où PE1 et PE2 peuvent être le même routeur ou des routeurs différents) sauf si PE2 transpose R2 en un chemin de VPN-IPv4 qui est différent de R1 (c'est-à-dire, contient un RD différent).

3. Les routeurs PE et CE peuvent être des homologues OSPF. Un routeur PE qui est un homologue OSPF d'un routeur CE apparaît, au routeur CE, comme étant un routeur de zone 0. Si un routeur PE est un homologue OSPF de routeurs CE qui sont dans des VPN distincts, le PE doit bien sûr faire fonctionner plusieurs instances de OSPF.

   Les chemins IPv4 que le PE apprend du CE via OSPF sont redistribués dans BGP comme chemins de VPN-IPv4. Les attributs de communauté étendue sont utilisés pour porter, avec le chemin, toutes les informations nécessaires pour permettre que le chemin soit distribué aux autres routeurs CE dans le VPN dans le type approprié d'annonce d'état de liaison (LSA, *Link State Advertisement*). L'étiquetage de chemin OSPF est utilisé pour s'assurer que les chemins reçus du cœur de réseau MPLS/BGP ne sont pas renvoyés dans le cœur de réseau.

   La spécification de l'ensemble complet de procédures pour l'utilisation de OSPF entre PE et CE se trouve dans les [RFC4577] et [RFC4576].

4. Les routeurs PE et CE peuvent être des homologues BGP, et le routeur CE peut utiliser BGP (en particulier, EBGP) pour dire au routeur PE l'ensemble de préfixes d'adresses qui sont au site du routeur CE. (Cette technique peut être utilisée dans les VPN d'extrémité ou de transit.)

   Cette technique présente un certain nombre d'avantages sur les autres :
   a) À la différence de IGP, cela n'exige pas que le PE fasse fonctionner plusieurs instances d'algorithme d'acheminement afin de parler à plusieurs CE.
   b) BGP est explicitement conçu pour cette seule fonction : passer les informations d'acheminement entre les systèmes de différentes administrations.
   c) Si le site contient des "portes dérobées BGP", c'est-à-dire, des routeurs avec des connexions BGP avec des routeurs autres que des routeurs PE, cette procédure va fonctionner correctement dans toutes les circonstances. Les autres procédures peuvent ou non fonctionner, selon les circonstances.
   d) l'utilisation de BGP rend facile au CE de passer les attributs des chemins au PE. Il sort du domaine d'application du présent document de faire une spécification complète de l'ensemble des attributs et de leur utilisation. Cependant, des exemples de la façon de les utiliser sont :
      – Le CE peut suggérer une cible de chemin particulière pour chaque chemin, parmi les cibles de chemin que le PE est autorisé à rattacher au chemin. Le PE va alors rattacher seulement la cible de chemin suggérée, plutôt que l'ensemble complet. Cela donne à l'administrateur de CE un contrôle dynamique de la distribution des chemins à partir du CE.
      - Des types supplémentaires d'attributs Communauté étendue peuvent être définis, où l'intention est que ces attributs soient passés de façon transparente (c'est-à-dire, sans être changés par les routeurs PE) de CE à CE. Cela permettrait aux administrateurs de CE de mettre en œuvre un filtrage de chemin supplémentaire, au delà de ce qui est fait par les PE. Ce filtrage supplémentaire n'exigerait pas de coordination avec le SP.

   Par ailleurs, l'utilisation de BGP peut être quelque chose de nouveau pour les administrateurs de CE.

   Si un site n'est pas dans un VPN de transit, on notera qu'il n'est pas besoin d'avoir un numéro unique de système autonome (ASN). Chaque CE dont le site n'est pas dans un VPN de transit peut utiliser le même ASN. Il peut être choisi dans l'espace d'ASN privé, et sera éliminé par le PE. Les boucles d'acheminement sont évitées par l'utilisation de

l'attribut Site d'origine (voir plus loin).

Qu'en est-il si un ensemble de sites constitue un VPN de transit ? Ce sera généralement le cas seulement si le VPN est lui-même un réseau de fournisseur d'accès Internet (FAI) où le FAI est lui-même acheteur de services de cœur de réseau d'un autre SP. Ce dernier SP peut être appelé un "transporteur de transporteur". Dans ce cas, la meilleure façon de fournir le VPN est de faire que les routeurs CE prennent en charge MPLS, et d'utiliser la technique décrite à la Section 9.

Quand on n'a pas besoin de distinguer entre les différentes façons dont un PE peut être informé des préfixes d'adresses qui existent à un certain site, on dit simplement que le PE a "appris" les chemins de ce site. Cela inclut le cas où le PE a été configuré manuellement avec les chemins.

Avant qu'un PE puisse redistribuer un chemin de VPN-IPv4 appris d'un site, il doit allouer un attribut Cible de chemin (voir au paragraphe 4.3.1) au chemin, et il peut allouer un attribut Site d'origine au chemin.

L'attribut Site d'origine, si il est utilisé, est codé comme Communauté étendue d'origine de chemin [RFC4360]. L'objet de cet attribut est d'identifier de façon univoque l'ensemble des chemins appris d'un site particulier. Cet attribut est nécessaire dans certains cas pour s'assurer qu'un chemin appris d'un site particulier via une certaine connexion PE/CE n'est pas redistribué au site par une connexion PE/CE différente. Il est particulièrement utile si BGP est utilisé comme protocole PE/CE, mais si les différents sites n'ont pas eu d'allocation d'ASN distincts.

## 8. Comment les CE apprennent les chemins des PE

Dans cette Section, on suppose que l'appareil CE est un routeur.

Si le PE place un chemin particulier dans le VRF qu'il utilise pour acheminer les paquets reçus d'un CE particulier, alors en général, le PE peut distribuer ce chemin au CE. Bien sûr, le PE ne peut distribuer ce chemin au CE que si c'est permis par les règles du protocole PE/CE. (Par exemple, si un protocole PE/CE particulier a un "horizon partagé", certains chemins dans le VRF ne peuvent pas être redistribués au CE.) On ajoute une restriction de plus à la distribution des chemins du PE au CE : si un attribut de chemin Site d'origine identifie un site particulier, ce chemin ne doit jamais être redistribué à un CE à ce site.

Dans la plupart des cas, cependant, il va être suffisant que le PE distribue simplement le chemin par défaut au CE. (Dans certains cas, il peut même être suffisant au CE d'être configuré avec un chemin par défaut pointant sur le PE.) Cela va généralement fonctionner sur tout site qui n'a pas lui-même besoin de distribuer le chemin par défaut aux autres sites. (Par exemple, si un site dans un VPN d'entreprise a l'accès d'entreprise à l'Internet, ce site peut avoir besoin d'avoir le chemin par défaut distribué à l'autre site, mais on ne pourrait pas distribuer le chemin par défaut à ce site lui-même.)

Quelle que soit la procédure utilisée pour distribuer les chemins de CE à PE, elle sera aussi utilisée pour distribuer les chemins de PE à CE.

## 9. Transporteur de transporteur

Parfois un VPN peut être en fait le réseau d'un FAI, avec ses propres politiques d'échange de trafic et d'acheminement. Parfois un VPN peut être le réseau d'un fournisseur de services qui offre des services de VPN à ses propres consommateurs. Des VPN comme ceux-là peuvent aussi obtenir des services de cœur de réseau d'un autre SP, le "transporteur de transporteur", en utilisant essentiellement la même méthode que décrit dans le présent document. Cependant, il est nécessaire dans ces cas que les routeurs CE prennent en charge MPLS. En particulier :

- Les routeurs CE devraient distribuer aux routeurs PE SEULEMENT les chemins internes au VPN. Cela permet au VPN d'être traité comme VPN d'extrémité.
- Les routeurs CE devraient prendre en charge MPLS, en ce qu'ils devraient être capables de recevoir des étiquettes provenant des routeurs PE, et envoyer des paquets étiquetés aux routeurs PE. Ils n'ont cependant pas besoin de distribuer d'eux-mêmes des étiquettes.
- Les routeurs PE devraient distribuer aux routeurs CE les étiquettes pour les chemins qu'ils distribuent aux routeurs CE.

Le PE ne doit pas distribuer la même étiquette à deux CE différents sauf si une des conditions suivantes est satisfaite :

* les deux CE sont associés à exactement le même ensemble de VRF ;
* le PE conserve une transposition d'étiquette entrante différente ([RFC3031]) pour chaque CE.

De plus, quand le PE reçoit un paquet étiqueté d'un CE, il doit vérifier que l'étiquette du sommet est une de celles qui ont été distribuées à ce CE.
- Les routeurs aux différents sites devraient établir des connexions BGP entre eux pour l'échange des chemins externes (c'est-à-dire, les chemins qui conduisent à l'extérieur du VPN).
- Tous les chemins externes doivent être connus des routeurs CE.

Ensuite, quand un routeur CE cherche l'adresse de destination d'un paquet, la recherche d'acheminement va donner une adresse interne, généralement l'adresse du prochain bond BGP du paquet. Le CE étiquette le paquet de façon appropriée et l'envoie au PE. Le PE, plutôt que de chercher l'adresse IP de destination du paquet dans un VRF, utilise l'étiquette MPLS sommitale du paquet pour choisir le prochain bond BGP. Par suite, si le prochain bond BGP est plus d'un bond plus loin, l'étiquette sommitale va être remplacée par deux étiquettes, une étiquette de tunnel et une étiquette de VPN de chemin. Si le prochain bond BGP est un bond plus loin, l'étiquette sommitale peut être remplacée par juste l'étiquette de VPN de chemin. Si le PE d'entrée est aussi le PE de sortie, l'étiquette sommitale va juste être sautée. Quand le paquet est envoyé de son PE de sortie à un CE, le paquet va avoir une étiquette MPLS de moins que ce qu'il avait quand il a été reçu par son PE d'entrée.

Dans la procédure ci-dessus, les routeurs CE sont les seuls routeurs dans le VPN qui doivent prendre en charge MPLS. Si, par ailleurs, tous les routeurs à un site de VPN particulier prennent en charge MPLS, il n'est alors plus exigé que les routeurs CE connaissent tous les chemins externes. Tout ce qui est requis est que les chemins externes soient connus de tout routeur chargé de mettre la pile d'étiquettes sur un paquet jusqu'alors non étiqueté et qu'il y ait un chemin à commutation d'étiquette qui conduise de ces routeurs à leurs homologues BGP dans les autres sites. Dans ce cas, pour chaque chemin interne qu'un routeur CE distribue à un routeur PE, il doit aussi distribuer une étiquette.

## 10. Cœur de réseau multi AS

Qu'en est-il si deux sites d'un VPN sont connectés à différents systèmes autonomes (par exemple, parce que les sites sont connectés à des SP différents) ? Les routeurs PE rattachés à ce VPN ne vont alors pas être capables de conserver des connexions IBGP de l'un avec chaque autre, ou avec un réflecteur de chemin commun. Il doit plutôt y avoir quelque moyen d'utiliser EBGP pour distribuer les adresses de VPN-IPv4.

Il y a un certain nombre de façons différentes de traiter ce cas, qu'on présente en ordre décroissant d'adaptabilité.

a) Connexions RF-à-VRF aux routeurs de bordure de l'AS (système autonome).
   Dans cette procédure, un routeur PE dans un AS rattache directement un routeur PE dans un autre. Les deux routeurs PE vont être rattachés par plusieurs sous-interfaces, au moins une pour chaque VPN dont les chemins doivent être passés d'un AS à un AS. Chaque PE va traiter l'autre comme si c'était un routeur CE. C'est-à-dire, les PE associent chacune de ces sous-interfaces à un VRF, et utilisent EBGP pour distribuer les adresses IPv4 non étiquetées à chaque autre.
   C'est une procédure qui "fonctionne", et qui n'exige pas MPLS à la frontière entre les AS. Cependant, elle ne s'adapte pas aussi bien que les autres procédures présentées ci-dessous.

b) Redistribution par EBGP des chemins de VPN-IPv4 étiquetées de l'AS à l'AS voisin.
   Dans cette procédure, les routeurs PE utilisent IBGP pour redistribuer les chemins de VPN-IPv4 étiquetés soit à un routeur bordure de système autonome (ASBR, *Autonomous System Border Router*) soit à un réflecteur de chemin dont un ASBR est un client. L'ASBR utilise alors EBGP pour redistribuer ces chemins de VPN-IPv4 étiquetés à un ASBR dans un autre AS, qui à son tour les distribue aux routeurs PE dans cet AS, ou peut-être à un autre ASBR qui à son tour les distribue, et ainsi de suite.
   Quand on utilise cette procédure, les chemins de VPN-IPv4 devraient seulement être acceptés sur les connexions EBGP à des points d'échange de trafic privés, au titre d'accords de confiance entre les SP. Les chemins VPN-IPv4 ne devraient ni être distribués dans, ni acceptés de l'Internet public, ou de tout homologue BGP qui ne soit pas de confiance. Un ASBR ne devrait jamais accepter un paquet étiqueté provenant d'un homologue EBGP si il n'a pas en fait distribué l'étiquette sommitale à cet homologue.
   Si il y a de nombreux VPN qui ont des sites rattachés à des systèmes autonomes différents, il n'est pas nécessaire qu'il y ait un seul ASBR entre ces deux AS qui détienne tous les chemins pour tous les VPN ; il peut y avoir plusieurs ASBR, dont chacun détient seulement les chemins pour un sous ensemble particulier de VPN.
   Cette procédure exige qu'il y ait un chemin de commutation d'étiquettes conduisant du PE d'entrée de paquet au PE de sortie. Donc, des relations de confiance appropriées doivent exister entre et parmi l'ensemble des AS le long du chemin. Aussi, il doit y avoir un accord entre l'ensemble des SP sur quels routeurs de bordure ont besoin de recevoir les chemins avec quelles cibles de chemin.

c) Redistribution multi bonds EBGP de chemins de VPN-IPv4 étiquetés entre AS de source et de destination, avec redistribution EBGP des chemins IPv4 étiquetés de l'AS à l'AS voisin.
Dans cette procédure, les chemins de VPN-IPv4 ne sont ni conservés ni distribués par les ASBR. Un ASBR doit conserver les chemins /32 IPv4 étiquetés pour les routeurs PE au sein de son AS. Il utilise EBGP pour distribuer ces chemins aux autres AS. Les ASBR dans tout AS de transit vont aussi devoir utiliser EBGP pour passer les chemins /32 étiquetés. Il en résulte la création d'un chemin de commutation d'étiquettes du routeur PE d'entrée au routeur PE de sortie. Maintenant, les routeurs PE dans des AS différents peuvent établir des connexions EBGP multi bonds avec chaque autre, et peuvent échanger les chemins de VPN-IPv4 sur ces connexions.
Si les chemins /32 pour les routeurs PE sont portés à la connaissance des routeurs P de chaque AS, tout fonctionne normalement. Si les chemins /32 pour les routeurs PE NE sont PAS connus des routeurs P (autres que les ASBR) alors cette procédure exige que le PE d'entrée d'un paquet mette une pile de trois étiquette sur lui. L'étiquette du bas est allouée par le PE de sortie, correspondant à l'adresse de destination du paquet dans un VRF particulier. L'étiquette du milieu est allouée par l'ASBR, correspondant au chemin /32 pour le PE de sortie. L'étiquette du sommet est allouée par le prochain bond IGP du PE de sortie, correspondant au chemin /32 pour l'ASBR.
Pour améliorer l'adaptabilité, on peut faire que les connexions multi-bond EBGP n'existent qu'entre un réflecteur de chemin dans un AS et un réflecteur de chemin dans un autre. (Cependant, quand les réflecteurs de chemin distribuent les chemins sur cette connexion, ils ne modifient pas l'attribut de prochain bond BGP des chemins.) Les routeurs PE actuels vont seulement avoir les connexions IBGP avec les réflecteurs de chemin dans leur propre AS.
Cette procédure est très similaire à celle du "transporteur de transporteur" décrit à la Section 9. Comme la procédure précédente, elle exige qu'il y ait un chemin de commutation d'étiquettes conduisant du PE d'entrée du paquet à son PE de sortie.

## 11. Accès à l'Internet à partir d'un VPN

De nombreux sites de VPN ont besoin d'être capables d'accéder à l'Internet public, ainsi que d'accéder aux autres sites de VPN. On décrit ci-dessous certaines des différentes façons de le faire.

1. Dans certains VPN, un ou plusieurs des sites vont obtenir l'accès Internet au moyen d'une "passerelle Internet" (peut-être un pare-feu) rattachée à une non interface de VRF à un FAI. Le FAI peut ou non être la même organisation que le SP qui fournit le service de VPN. Le trafic de/vers la passerelle Internet va alors être acheminé en accord avec le tableau de transmission par défaut du routeur PE.
   Dans ce cas, les sites qui ont l'accès Internet peuvent distribuer un chemin par défaut à leurs PE, qui à leur tour le redistribuent aux autres PE et donc dans d'autres sites du VPN. Cela fournit l'accès Internet pour tous les sites du VPN.
   Afin de traiter correctement le trafic provenant de l'Internet, le FAI doit distribuer, à l'Internet, les chemins conduisant aux adresses qui sont dans le VPN. Ceci est complètement indépendant de toutes les procédures de distribution de chemin décrites dans ce document. La structure interne du VPN ne va en général pas être visible de l'Internet ; de tels chemins conduiraient simplement à la non interface de VRF qui s'attache à la passerelle Internet du VPN.
   Dans ce modèle, il n'y a pas d'échange de routes entre le tableau de transmission par défaut d'un routeur PE et ses VRF. Les procédures de distribution de chemin de VPN et les procédures de distribution de chemin de l'Internet sont complètement indépendantes.
   Noter que bien que certains sites de VPN utilisent une interface de VRF pour communiquer avec l'Internet, en fin de compte tous les paquets de/vers l'Internet traversent une non interface de VRF avant de quitter/entrer dans le VPN, de sorte qu'on se réfère à cela comme à un "accès Internet non VRF".
   Noter que le routeur PE auquel la non interface de VRF se rattache n'a pas nécessairement besoin de conserver tous les chemins Internet dans son tableau de transmission par défaut. Le tableau de transmission par défaut pourrait avoir jusqu'à un seul chemin, "défaut", qui conduise à un autre routeur (probablement adjacent) qui a les chemins Internet. Une variante de ce schéma est de tunneler les paquets reçus sur la non interface de VRF du routeur PE à un autre routeur, où cet autre routeur conserve l'ensemble complet de chemins Internet.

2. Certains VPN peuvent obtenir l'accès Internet via une interface de VRF ("accès Internet VRF"). Si un paquet est reçu par un PE sur une interface de VRF, et si l'adresse de destination du paquet ne correspond à aucun chemin dans le VRF, il peut alors être confronté au tableau de transmission par défaut du PE. Si une correspondance est trouvée, le paquet peut être transmis nativement à travers le cœur de réseau à l'Internet, au lieu d'être transmis par MPLS.
   Afin que le trafic s'écoule nativement dans la direction opposée (de l'Internet à l'interface de VRF) certains des chemins venant du VRF doivent être exportés au tableau de transmission de l'Internet. Inutile de dire que tout chemin de cette sorte doit correspondre à des adresses uniques au monde.
   Dans ce schéma, le tableau de transmission par défaut peut avoir le jeu complet des chemins de l'Internet, ou il peut avoir un seul chemin par défaut conduisant à un autre routeur qui lui a le jeu complet des chemins Internet dans son tableau de transmission par défaut.

3. Supposons que le PE ait la capacité de mémoriser les "non chemins de VPN" dans un VRF. Si l'adresse de destination d'un paquet correspond à un "non chemin de VPN", alors le paquet est transmis nativement, plutôt que via MPLS. Si le VRF contient un chemin par défaut non VPN, tous les paquets pour l'Internet public vont correspondre, et être transmis nativement au prochain bond du chemin par défaut. À ce prochain bond, les adresses de destination du paquet vont être recherchées dans le tableau de transmission par défaut, et peuvent correspondre à des chemins plus spécifiques.
   Cette technique ne va être disponible que si aucun des routeurs CE ne distribue de chemin par défaut.

4. Il est aussi possible d'obtenir l'accès Internet via une interface de VRF en faisant que le VRF contienne les chemins Internet. Comparé au modèle 2, ceci élimine la seconde recherche, mais cela a l'inconvénient d'exiger que les chemins Internet soient dupliqués dans chacun de ces VRF.
   Si cette technique est utilisée, le SP peut vouloir que son interface à l'Internet soit une interface de VRF, et utiliser les techniques de la Section 4 pour distribuer les chemins Internet, comme des chemins de VPN-IPv4, aux autres VRF.

Il devrait être clairement compris que par défaut, il n'y a pas d'échange de chemins entre un VRF et le tableau de transmission par défaut. Ceci n'est fait QUE par accord entre un client et un SP, et seulement si cela convient à la politique du client.

## 12. Gestion des VPN

La présente spécification n'exige pas que la sous interface connectant un routeur PE et un routeur CE soit une interface "numérotée". Si c'est une interface numérotée, la présente spécification permet que les adresses allouées à l'interface viennent de l'espace d'adresses du VPN ou de l'espace d'adresses du SP.

Si un routeur CE est géré par le fournisseur de services, celui-ci va probablement avoir un système de gestion de réseau qui doit être capable de communiquer avec le routeur CE. Dans ce cas, les adresses allouées à la sous interface qui connecte le CE et les routeurs PE devraient venir de l'espace d'adresses du SP, et devraient être uniques dans cet espace. Le système de gestion de réseau devrait lui-même se connecter à un routeur PE (plus précisément, être à un site qui se connecte à un routeur PE) via une interface de VRF. L'adresse du système de gestion de réseau va être exportée à tous les VRF qui sont associés aux interfaces avec les routeurs CE qui sont gérés par le SP. Les adresses des routeurs CE vont être exportées au VRF associé au système de gestion de réseau, mais à aucun autre VRF.

Cela permet la communication entre le CE et le système de gestion de réseau, mais ne permet pas de communication non désirée ou entre les routeurs CE.

Une façon d'assurer que les importations/exportations de chemin sont faites de façon appropriée est d'utiliser deux cibles de chemin ; appelons les T1 et T2. Si une interface de VRF particulière se rattache à un routeur CE qui est géré par le SP, alors ce VRF est configuré pour :
- importer les chemins qui ont T1 rattaché à eux, et
- rattacher T2 aux adresses allouées à chaque extrémité de ses interfaces de VRF.

Si une interface de VRF particulière se rattache au système de gestion de réseau du SP, alors ce VRF est configuré pour rattacher T1 à l'adresse de ce système, et pour importer les chemins qui ont T2 rattaché à eux.

## 13. Considérations sur la sécurité

### 13.1 Plan des données

Par sécurité dans le "plan des données", on entend la protection contre les possibilités suivantes :
- Les paquets provenant de l'intérieur d'un VPN voyagent vers un site extérieur au VPN, autrement que d'une manière cohérente avec les politiques de ce VPN.
- Les paquets provenant de l'extérieur d'un VPN entrent dans les sites du VPN, autrement que d'une manière cohérente avec les politiques de ce VPN.

Dans les conditions suivantes :
1. un routeur de cœur de réseau n'accepte pas les paquets étiquetés sur une certaine liaison de données, sauf si il est connu que cette liaison de données se rattache seulement à des systèmes de confiance, ou sauf si il est connu que de tels paquets vont quitter le cœur de réseau avant que l'en-tête IP ou que toute étiquette inférieure de la pile soient inspectés,

et
2. les chemins de VPN-IPv4 étiquetés ne sont pas acceptés d'homologues d'acheminement qui ne sont pas de confiance ou ne sont pas fiables,
3. aucune attaque réussie n'a été montée sur le plan de contrôle,

la sécurité du plan des données fournie par cette architecture est virtuellement identique à celle fournie au VPN par les cœurs de réseau de relais de trame ou ATM. Si les appareils sous le contrôle du SP sont proprement configurés, les données ne vont entrer ou quitter le VPN qu'autorisées à le faire.

La condition 1 ci-dessus peut être formulée plus précisément. On devrait éliminer un paquet étiqueté reçu d'un certain voisin sauf si une des deux conditions suivantes est satisfaite :
- l'étiquette sommitale du paquet a une valeur que le système receveur a distribuée à ce voisin, ou
- l'étiquette sommitale du paquet a une valeur que le système receveur a distribuée à un système au delà de ce voisin (c'est-à-dire, quand on sait que le chemin provenant du système auquel l'étiquette a été distribuée au système receveur peut être via ce voisin).

La condition 2 ci-dessus est de grand intérêt dans le cas de VPN inter fournisseurs (voir la Section 10). Pour les VPN inter fournisseurs construits en accord avec le schéma b) de la Section 10, la condition 2 est facile à vérifier. (La question de la sécurité quand le schéma (c) de la Section 10 est utilisé fera l'objet d'études ultérieures.)

On notera que l'utilisation de MPLS rend beaucoup plus simple de fournir la sécurité du plan des données qu'il ne serait possible si on tentait d'utiliser une forme de tunnelage IP à la place de l'étiquette extérieure MPLS. Il est simple de faire que les routeurs bordures refusent d'accepter un paquet étiqueté si la première des conditions ci-dessus ne s'applique pas. Il est plus difficile de configurer un routeur à refuser d'accepter un paquet IP si ce paquet est un paquet IP tunnelé dont l'adresse de destination est celle d'un routeur PE ; certainement, ce n'est pas impossible, mais cela a des implications de gestion et de performances.

Les tunnelages MPLS-dans-IP et MPLS-dans-GRE sont spécifiés dans la [RFC4023]. Si on désire utiliser de tels tunnels pour porter les paquets de VPN, alors les considérations de sécurité décrites à la Section 8 de ce document doivent être pleinement comprises. Toute mise en œuvre de VPN IP BGP/MPLS qui permet aux paquets de VPN d'être tunnelés comme décrit dans le présent document DOIT contenir une mise en œuvre de IPsec qui puisse être utilisée comme décrit ici. Si le tunnel n'est pas sécurisé par IPsec, alors la technique de filtrage d'adresse IP aux routeurs de bordure, décrite au paragraphe 8.2 de ce document, est le seul moyen de s'assurer qu'un paquet qui sort du tunnel à un PE de sortie particulier a réellement été placé dans le tunnel par le nœud de tête de tunnel approprié (c'est-à-dire, que le paquet n'a pas une adresse de source usurpée). Comme les routeurs bordures filtrent fréquemment seulement les adresses de source, le filtrage de paquets peut n'être pas efficace sauf si le PE de sortie peut vérifier l'adresse IP de source de tout paquet tunnelé qu'il reçoit, et la comparer à une liste d'adresses IP qui sont des adresses de tête de tunnel valides. Toute mise en œuvre qui permet l'utilisation du tunnelage MPLS-dans-IP et/ou MPLS-dans-GRE sans IPsec DOIT permettre au PE de sortie de valider de cette manière l'adresse IP de source de tout paquet tunnelé qu'elle reçoit.

Dans le cas où un certain nombre de routeurs CE se rattachent à un routeur PE via une interface de LAN, pour assurer une sécurité appropriée, une des conditions suivantes doit être satisfaite :
1. Tous les routeurs CE du LAN appartiennent au même VPN, ou
2. Un commutateur de LAN de confiance et sécurisé divise le LAN en plusieurs VLAN, chaque VLAN contenant seulement des systèmes d'un seul VPN ; dans ce cas, le commutateur va rattacher l'étiquette de VLAN appropriée à tout paquet avant de le transmettre au routeur PE.

La confidentialité cryptographique n'est pas fournie par cette architecture, ni par les VPN en relais de trame ni ATM. Ces architectures sont toutes compatibles avec l'utilisation de cryptographie CE par CE, si c'est désiré.

L'utilisation de la cryptographie sur la base du PE fera l'objet d'une étude ultérieure.

### 13.2 Plan de contrôle

La sécurité du plan des données du paragraphe précédent dépend de la sécurité du plan de contrôle. Pour assurer la sécurité, ni les connexions BGP ni les connexions LDP ne devraient être faites avec des homologues qui ne sont pas de confiance. L'option d'authentification MD5 TCP/IP [RFC2385] devrait être utilisée avec ces deux protocoles. Le protocole d'acheminement dans le réseau du SP devrait aussi être sécurisé d'une manière similaire.

### 13.3 Sécurité des appareils P et PE

Si la sécurité physique de ces appareils est compromise, la sécurité du plan des données peut aussi être compromise.

Les étapes usuelles devraient être suivies pour s'assurer que le trafic IP provenant de l'Internet public ne peut pas être utilisé pour modifier la configuration de ces appareils, ou pour monter contre eux des attaques de déni de service.

## 14. Qualité de service

Bien que ce ne soit pas le point principal de ce document, la qualité de service est un composant clé de tout service de VPN. Dans les VPN MPLS/BGP, les capacités existantes de qualité de service de couche 3 peuvent être appliquées pour étiqueter les paquets par l'utilisation des bits "expérimentaux" dans l'en tête d'ajustement *(shim header)* [RFC3032], ou, lorsque ATM est utilisé comme cœur de réseau, par l'utilisation des capacités de qualité de service d'ATM. L'ingénierie de trafic discutée dans la [RFC3209] est aussi directement applicable aux VPN MPLS/BGP. L'ingénierie de trafic pourrait même être utilisée pour établir des chemins à commutation d'étiquettes avec des caractéristiques de qualité de service particulières entre des paires particulières de sites, si c'est désirable. Lorsque un VPN MPLS/BGP s'étend sur plusieurs SP, l'architecture décrite dans la [RFC2430] peut être utile. Un SP peut appliquer les capacités intserv (services intégrés) ou diffserv (services différenciés) à un VPN particulier, comme approprié.

## 15. Adaptabilité

On a discuté des questions d'adaptabilité tout au long de ce papier. Dans cette section, on résume brièvement les principales caractéristiques de notre modèle par rapport à l'adaptabilité.

Le cœur de réseau du fournisseur de services consiste en (a) des routeurs PE, (b) des réflecteurs de chemin BGP, (c) des routeurs P (qui ne sont ni des routeurs PE ni des réflecteurs de chemin) et, dans le cas de VPN multi fournisseurs, (d) des ASBR.

Les routeurs P ne conservent aucun chemin de VPN. Afin de transmettre correctement le trafic de VPN, les routeurs P ont seulement à conserver les chemins pour les routeurs PE et les ASBR. L'utilisation de deux niveaux d'étiquetage est ce qui rend possible de garder les chemins de VPN en dehors des routeurs P.

Un routeur PE conserve les chemins de VPN, mais seulement pour les VPN auxquels il est directement rattaché.

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

Pour les VPN inter fournisseurs, si les ASBR conservent et distribuent les chemins VPN-IPv4, alors les ASBR peuvent être partagés entre les VPN d'une manière similaire, avec pour résultat qu'aucun ASBR seul n'est exigé pour conserver les chemins pour tous les VPN inter fournisseurs. Si EBGP multi bonds est utilisé, alors les ASBR n'ont pas besoin du tout de conserver et distribuer les chemins VPN-IPv4.

Par suite, aucun composant seul dans le réseau du fournisseur de services n'a à conserver tous les chemins pour tous les VPN. Donc la capacité totale du réseau pour prendre en charge un nombre croissant de VPN n'est pas limitée par la capacité d'un composant individuel.

## 16. Considérations relatives à l'IANA

L'Autorité d'allocation des numéros de l'Internet (IANA) a créé un nouveau registre pour le "Champ Type de discriminateur de chemin" (voir le paragraphe 4.2). C'est un champ de deux octets. Les types 0, 1, et 2 sont définis par le présent document. Des valeurs supplémentaires de champ Type de discriminateur de chemin avec un bit de poids fort de 0 peuvent être allouées par l'IANA sur la base du "premier arrivé, premier servi" [RFC2434]. Des valeurs avec un bit de poids fort de 1 peuvent être allouées par l'IANA sur la base du "consensus de l'IETF" [RFC2434].

Le présent document spécifie (paragraphe 4.3.4) l'utilisation de la valeur 1 d'identifiant de famille d'adresse BGP (AFI, *Address Family Identifier*) ainsi que la valeur 128 d'identifiant suivant de famille d'adresse BGP (SAFI, *Subsequent Address Family Identifier*) pour représenter la famille d'adresses "Adresses étiquetées VPN-IPv4", qui est définie dans ce document.

L'utilisation de la valeur d'AFI 1 pour IP est comme actuellement spécifié dans le registre IANA "Identifiant de famille d'adresses", de sorte que l'IANA n' a pas besoin d'une action à son égard.

La valeur de SAFI 128 était à l'origine spécifiée comme "utilisation privée" dans le registre IANA "Identifiant suivant de famille d'adresse". L'IANA a changé la valeur de SAFI 128 de "utilisation privée" à "Adresse de VPN étiquetée MPLS".

## 17. Remerciements

La liste complète des contributeurs se trouve à la Section 18.
Des contributions significatives au présent travail ont aussi été faites par Ravi Chandra, Dan Tappan, et Bob Thomas.
Merci aussi à Shantam Biswas de sa relecture et ses contributions.

## 18. Contributeurs

Tony Bogovic, Telcordia Technologies ; Stephen John Brannon, Swisscom AG ; Marco Carugi, Nortel Networks S.A. ; Christopher J. Chase, AT&T ; Ting Wo Chung, Bell Nexxia ; Eric Dean ; Jeremy De Clercq, Luyuan Fang, AT&T ; Paul Hitchen, BT ; Manoj Leelanivas, Juniper Networks, Inc. ; Dave Marshall, Worldcom ; Luca Martini, Cisco Systems ; Monique Jeanne Morrow, Cisco Systems ; Ravichander Vaidyanathan, Telcordia Technologies ; Adrian Smith, BT ; Vijay Srinivasan, cosinecom ; Alain Vedrenne, Equant.

## 19. Références normatives

[RFC2858] T. Bates et autres, "Extensions multiprotocoles pour BGP-4", juin 2000. (*Obsolète, voir* RFC4760) *(P.S.)*

[RFC3031] E. Rosen, A. Viswanathan, R. Callon, "Architecture de commutation d'étiquettes multi protocoles", janvier 2001. *(P.S.)* (*MàJ par la* RFC6790)

[RFC3032] E. Rosen et autres, "Codage de pile d'étiquettes MPLS", janvier 2001.

[RFC3107] Y. Rekhter et E. Rosen, "Portage des informations d'étiquette dans BGP-4", mai 2001. (*MàJ par* RFC6790, RFC8277)

[RFC4271] Y. Rekhter, T. Li et S. Hares, "Protocole de routeur frontière version 4 (BGP-4)", janvier 2006. *(D.S.) (MàJ par* RFC6608, RFC8212)

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

## 20. Références pour information

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

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

[RFC2430] T. Li, Y. Rekhter, "Architecture fournisseur pour les services différenciés et l'ingénierie du trafic", octobre 1998. *(Info.)*

[RFC2434] T. Narten et H. Alvestrand, "Lignes directrices pour la rédaction d'une section Considérations relatives à l'IANA dans les RFC", BCP 26, octobre 1998. *(Rendue obsolète par la* RFC5226)

[RFC2453] G. Malkin, "RIP version 2", STD 56, novembre 1998. *(Mise à jour par la RFC 4822)*

[RFC2796] T. Bates, R. Chandra, E. Chen, "Réflexion de chemin BGP - une alternative à IBGP à maillage complet", avril 2000. (*Obsolète, voir* RFC4456) *(P.S.)*

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

[RFC3034] A. Conta, P. Doolan, A. Malis, "Spécification de l'utilisation de la commutation d'étiquettes sur les réseaux en relais de trame", janvier 2001. *(P.S.)*

[RFC3035] B. Davie et autres, "Utilisation de MPLS dans la commutation de circuit virtuel LDP et ATM", janvier 2001. *(P.S.)*

[RFC3036] L. Andersson et autres, "Spécification de LDP", janvier 2001. *(Obsolète, voir la* RFC5036)

[RFC3209] D. Awduche, et autres, "RSVP-TE : Extensions à RSVP pour les tunnels LSP", décembre 2001. (*Mise à jour par* RFC3936, RFC4420, RFC4874, RFC5151, RFC5420, RFC6790)

[RFC4023] T. Worster et autres, "Encapsulation de MPLS dans IP ou encapsulation d'acheminement générique (GRE)", mars 2005. (*MàJ par* RFC5332) *(P.S.)*

[RFC4576] E. Rosen et autres, "Utilisation d'un bit d'option d'annonce d'état de liaison (LSA) pour empêcher les boucles dans les réseaux privés virtuels (VPN) IP BGP/MPLS", juin 2006. *(P.S.)*

[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.)*

[RFC4893] Q. Vohra, E. Chen, "Prise en charge par BGP de l'espace de numéros d'AS à quatre octets", mai 2007. *(P.S.)*

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

[RFC6037] E. Rosen, Y. Cai, IJ. Wijnands, "Solution de Cisco Systems pour la diffusion groupée dans les VPN IP BGP/MPLS", octobre 2010. *(Historique)*

[VPN-IP] Rosen, E., De Clercq, J., Paridaens, O., T'Joens, Y., and C. Sargor, "Architecture for the Use of PE-PE IPsec Tunnels in BGP/MPLS IP VPNs", travail en cours, mars 2004.

## Adresse des auteurs

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

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

## Déclaration complète de droits de reproduction

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.

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