Groupe de travail Réseau
Request for Comments : 4667
Catégorie : Sur la voie de la normalisation

W. Luo, Cisco Systems, Inc.
septembre 2006
Traduction Claude Brière de L'Isle

Extensions de couche 2 de réseau virtuel privé (L2VPN) pour le protocole de tunnelage de couche 2 (L2TP)

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 protocole de tunnelage de couche 2 (L2TP, *Layer 2 Tunneling Protocol*) donne une méthode standard pour établir et gérer des sessions L2TP à tunneler divers protocoles de couche 2. Un des modèles de référence pris en charge par L2TP décrit l'utilisation d'une session L2TP pour connecter deux circuits de couche 2 rattachés à une paire de concentrateurs d'accès L2TP (LAC, *L2TP Access Concentrator*) qui échangent du trafic, qui est une forme de base de réseau privé virtuel de couche 2 (L2VPN, *Layer 2 Virtual Private Network*). Le présent document définit les extensions de protocole pour que L2TP établisse différents types de L2VPN d'une façon unifiée.

Table des matières

1. Introduction

La [RFC3931] définit un mécanisme dynamique de tunnelage pour porter plusieurs protocoles de couche 2 à côté du protocole point à point (PPP), le seul protocole pris en charge dans la [RFC2661], sur un réseau à commutation de paquets. Le protocole de base prend en charge divers types d'applications, qui ont été mentionnées dans les différents modèles de référence de protocole de tunnelage de couche 2 (L2TP, *Layer 2 Tunneling Protocol*) dans la [RFC3931]. Un concentrateur d'accès L2TP (LAC, *L2TP Access Concentrator*) est un point d'extrémité de connexion de contrôle L2TP (LCCE, *L2TP Control Connection Endpoint*) qui interconnecte les circuits de rattachement et les sessions L2TP. Les applications de

réseau privé virtuel de couche 2 (L2VPN, *Layer 2 Virtual Private Network*) sont typiquement dans le domaine d'application du modèle de référence de LAC à LAC.

Le présent document discute des ressemblances et différences entre les applications de L2VPN par rapport à l'utilisation de L2TPv3 comme protocole de signalisation. Dans le présent document, l'acronyme "L2TP" se réfère à L2TPv3 ou à L2TP en général.

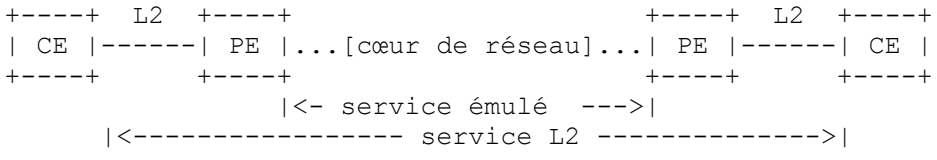
1.1 Spécification des exigences

Les mots clés "DOIT", "NE DOIT PAS", "EXIGE", "DEVRA", "NE DEVRA PAS", "DEVRAIT", "NE DEVRAIT PAS", "RECOMMANDE", "PEUT", et "FACULTATIF" en majuscules dans ce document sont à interpréter comme décrit dans le BCP 14, [RFC2119].

2. Modèle de référence réseau

Dans le modèle de référence de LAC à LAC, un LAC sert d'interconnexion entre les circuits de rattachement et les sessions L2TP. Chaque session L2TP agit comme un circuit émulé, aussi appelé un pseudo filaire. Un pseudo filaire est utilisé pour lier ensemble deux "transmetteurs". Pour des applications différentes de L2VPN, des types différents de transmetteurs sont définis.

Dans le cadre L2VPN [RFC4664], un LAC est un appareil de côté fournisseur (PE, *Provider Edge*). LAC et PE sont des termes interchangeables dans le contexte du présent document. Les systèmes distants dans le modèle de référence de LAC à LAC sont des appareils de côté utilisateur (CE, *Customer Edge*).



Modèle de référence de réseau L2VPN

Dans une application simple d'interconnexion, un circuit de rattachement est un transmetteur directement lié à un pseudo filaire. C'est une transposition biunivoque. Le trafic reçu du circuit de rattachement sur un PE local est transmis au PE distant à travers le pseudo filaire. Quand le PE distant reçoit du trafic du pseudo filaire, il transmet le trafic au circuit de rattachement correspondant sur son extrémité. La décision de transmission est fondée sur l'identifiant de démultiplexage du circuit de rattachement ou du pseudo filaire.

Avec le service de LAN privé virtuel (VPLS, *Virtual Private LAN Service*) une instance de commutation virtuelle (VSI, *Virtual Switching Instance*) est un transmetteur connecté à un ou plusieurs circuits de rattachement et pseudo filaires. Un seul pseudo filaire est utilisé pour connecter une paire de VSI sur deux PE qui échangent du trafic. Le trafic reçu d'un circuit de rattachement ou d'un pseudo filaire est d'abord transmis à la VSI correspondante sur la base de l'identifiant de démultiplexage du circuit de rattachement ou pseudo filaire. La VSI effectue une recherche supplémentaire pour déterminer si il faut encore transmettre le trafic.

Avec le service filaire privé virtuel (VPWS, *Virtual Private Wire Service*) les circuits de rattachement sont groupés en "réservoirs colorés". Chaque réservoir est un transmetteur et est connecté par un réseau d'interconnexions en point à point. La perspective de transmission des données est identique à l'application d'interconnexion. Cependant, la construction de réservoir colorés implique des procédures de signalisation plus compliquées.

3. Identifiant d'émetteur

Un identifiant de transmetteur est alloué à chaque transmetteur sur un certain PE et est unique dans le contexte du PE. Il est défini comme l'enchaînement d'un identifiant de groupe de rattachement (AGI, *Attachment Group Identifier*) et d'un identifiant de rattachement individuel (AII, *Attachment Individual Identifier*) noté <AGI, AII>. L'AGI est utilisé pour grouper un ensemble de transmetteurs pour les besoins de la signalisation. Un AII est utilisé pour distinguer les transmetteurs au sein d'un groupe. L'AII peut être unique par plate-forme ou par groupe.

Pour autant que les procédures de signalisation sont concernées, un identifiant de transmetteur est une chaîne arbitraire d'octets. Il appartient aux mises en œuvre de décider des valeurs pour AGI et AII.

Quand deux transmetteurs se connectent, tous deux DOIVENT avoir le même AGI au titre de leurs identifiants de transmetteur. L'AII du transmetteur source est appelé AII de source (SAII), et l'AII du transmetteur cible est appelé AII cible (TAII, *Target AII*). Donc, le transmetteur source et le transmetteur cible peuvent être notés respectivement <AGI, SAII> et <AGI, TAII>.

4. Composants du protocole

4.1 Messages de contrôle

L2TP définit deux ensembles de procédures de gestion de session : appel entrant et appel sortant. Bien qu'il soit entièrement possible d'utiliser les procédures d'appel sortant pour la signalisation des L2VPN, les procédures d'appel entrant ont certains avantages en termes de pertinence de la sémantique. [PWE3L2TP] donne plus de détails sur la raison pour laquelle les procédures d'appel entrant sont plus appropriées pour l'établissement des pseudo filaires.

Les procédures de signalisation pour les L2VPN décrites dans les paragraphes qui suivent se fondent sur les procédures de gestion de connexion de contrôle et les procédures d'appel entrant définies respectivement aux paragraphes 3.3 et 3.4.1 de la [RFC3931]. Les types de messages de contrôle L2TP sont définies au paragraphe 3.1 de la [RFC3931]. Le présent document fait référence aux messages de contrôle L2TP suivants :
Demande de début de connexion de contrôle (SCCRQ, *Start-Control-Connection-Request*)
Réponse de début de connexion de contrôle (SCCRP, *Start-Control-Connection-Reply*)
Demande d'appel entrant (ICRQ, *Incoming-Call-Request*)
Appel entrant connecté (ICCN, *Incoming-Call-Connected*)
Informations d'établissement de liaison (SLI, *Set-Link-Info*)

4.2 AVP existantes pour L2VPN

Les paires d'attribut/valeur (AVP, *Attribute Value Pair*) suivantes, définies aux paragraphes 5.4.3, 5.4.4, et 5.4.5 de la [RFC3931], sont utilisées pour la signalisation des L2VPN.

Identifiant de routeur : c'est l'identifiant de routeur envoyé dans les SCCRQ et SCCRP durant l'établissement de connexion de contrôle qui établit l'identité unique de chaque PE.

Liste de capacités de pseudo filaire : la liste de capacités de pseudo filaire envoyée dans les SCCRQ et SCCRP indique les types de pseudo filaire pris en charge par le PE envoyeur. Elle sert simplement d'annonce au PE receveur. Son contenu ne devrait pas affecter l'établissement de la connexion de contrôle. Avant qu'un PE local initie une session d'un type particulier de pseudo filaire avec un PE distant, il DOIT examiner si le PE distant a annoncé ce type de pseudo filaire dans cet AVP et NE DEVRAIT PAS tenter d'initier la session si le type de pseudo filaire prévu n'est pas pris en charge par le PE distant.

Type de pseudo filaire : le type de pseudo filaire envoyé dans ICRQ signale le type de pseudo filaire prévu au PE receveur. Le PE receveur le vérifie par rapport à sa liste locale de capacités de pseudo filaire. Si il trouve une correspondance, il répond avec une ICRP sans AVP Type de pseudo filaire, qui accuse implicitement son acceptation du pseudo filaire prévu. Si il ne trouve pas de correspondance, il DOIT répondre avec une notification d'appel déconnecté (CDN, *Call-Disconnect-Notify*) avec un code de résultat "Type de pseudo filaire non pris en charge".

Sous couche spécifique de couche 2 : la sous couche spécifique de couche 2 peut être envoyée dans ICRQ, ICRP, et ICCN. Si le PE receveur prend en charge la sous couche spécifique de couche 2 spécifiée, il DOIT inclure la sous couche spécifique de couche 2 identifiée dans ses paquets de données envoyés au PE envoyeur. Autrement, il DOIT rejeter la connexion en envoyant une CDN au PE envoyeur.

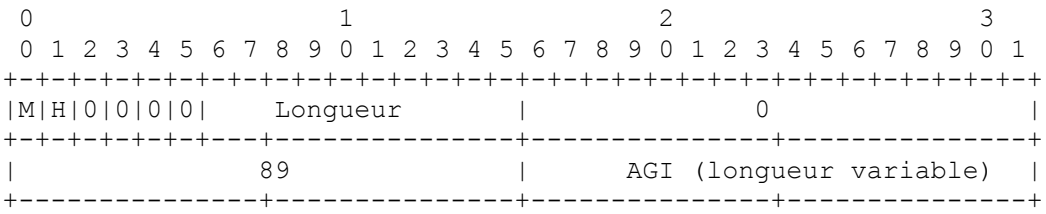
État de circuit : l'état de circuit est envoyé dans les deux ICRQ et ICRP pour informer le PE receveur de l'état de circuit sur le PE envoyeur. Il peut aussi être envoyé dans ICCN et SLI pour mettre à jour l'état.

Identifiant d'extrémité distante : la valeur de TAII *(identifiant de rattachement individuel cible)* est codée dans l'AVP Identifiant d'extrémité distante et envoyée dans la ICRQ avec le AGI facultatif pour demander au PE receveur de lier le pseudo filaire proposé au transmetteur qui correspond à l'identifiant de transmetteur spécifié.

4.3 Nouvelles AVP pour L2VPN

Identifiant de groupe de rattachement (AGI, *Attachment Group Identifier*) : l'AVP AGI, type d'attribut 89, est un identifiant utilisé pour associer un transmetteur à un groupe logique. L'AVP AGI est utilisé en conjonction avec l'AVP Identifiant d'extrémité locale et l'AVP Identifiant d'extrémité distante, qui codent respectivement le SAII et le TAII, pour identifier un transmetteur spécifique. Quand l'AVP AGI est omis dans les messages de contrôle ou contient une valeur de longueur zéro, les transmetteurs sont considérés utiliser l'AGI par défaut. Noter qu'il y a seulement une valeur d'AGI par défaut désignée pour tous les transmetteurs.

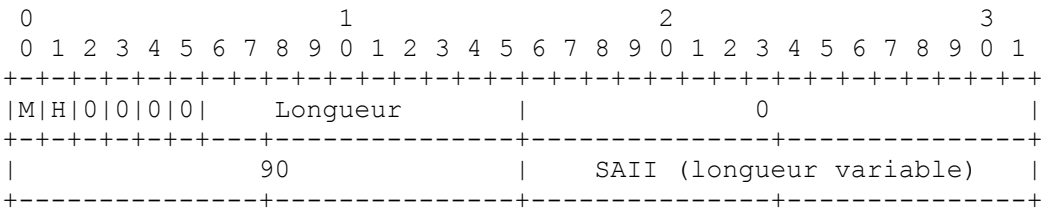
Le champ Valeur d'attribut pour cette AVP a le format suivant :



Le champ AGI est de longueur variable. Cette AVP PEUT être présente dans ICRQ.
Cette AVP PEUT être cachée (le bit H PEUT être 0 ou 1). La dissimulation des valeurs d'attribut d'AVP est définie au paragraphe 5.3 de la [RFC3931]. Le bit M pour cette AVP DEVRAIT être réglé à 0. La longueur (avant de la cacher) de cette AVP est 6 octets plus la longueur du champ AGI.

Identifiant d'extrémité locale : AVP Identifiant d'extrémité locale, type d'attribut 90, code la valeur de SAII. Le SAII peut aussi être utilisé en conjonction avec le TAII pour détecter une concurrence de pseudo filaires. Quand il est omis dans les messages de contrôle, il est supposé avoir la même valeur que le TAII.

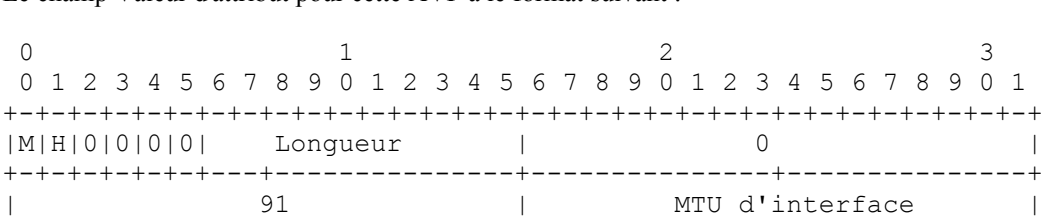
Le champ Valeur d'attribut pour cette AVP a le format suivant :



Le champ SAII est de longueur variable. Cette AVP PEUT être présente dans ICRQ.
Cette AVP PEUT être cachée (le bit H PEUT être 0 ou 1). Le bit M pour cette AVP DEVRAIT être réglé à 0. La longueur (avant d'être cachée) de cette AVP est 6 octets plus la longueur du champ SAII.

Unité maximum de transmission d'interface : l'AVP MTU d'interface, type d'attribut 91, indique la MTU en octets d'un paquet qui peut être envoyé de l'interface qui fait face au CE. Les valeurs de MTU d'un certain pseudo filaire, si elles sont annoncées dans les deux directions, DOIVENT être identiques. Si elles ne correspondent pas, le pseudo filaire NE DEVRAIT PAS être établi. Quand cette AVP est omise dans les messages de contrôle dans l'une ou l'autre direction, il est supposé que le PE distant a la même MTU d'interface que le PE local pour le pseudo filaire signalé.

Le champ Valeur d'attribut pour cette AVP a le format suivant :



Le champ MTU d'interface est une valeur d'entier de 2 octets. Cette AVP PEUT être présente dans ICRQ et ICRP. Quand

un PE reçoit une AVP MTU d'interface avec une valeur de MTU différente de la sienne, il PEUT répondre avec une CDN avec un nouveau code de résultat indiquant la cause de déconnexion de :

23 – MTU d'interface discordante

Cette AVP PEUT être cachée (le bit H PEUT être 0 ou 1). Le bit M pour cette AVP DEVRAIT être réglé à 0. La longueur (avant de la cacher) de cette AVP est 8 octets.

4.4 Interopérabilité d'AVP

Pour assurer l'interopérabilité, les réglages du bit obligatoire (M, *mandatory*) des AVP existantes utilisées dans les applications L2VPN devraient être les mêmes que ceux spécifiés dans la [RFC3931]. Le traitement générique du bit M est décrit au paragraphe 5.2 de la [RFC3931]. Établir à 1 le bit M des nouvelles AVP va impacter l'interopérabilité.

5. Procédures de signalisation

5.1 Généralités

Supposons qu'un PE alloue un identifiant de transmetteur à un de ses transmetteurs locaux et qu'il sache qu'il a besoin d'établir un pseudo filaire pour un transmetteur distant sur un PE distant qui a un certain identifiant de transmetteur. Cette connaissance peut être obtenue par configuration manuelle ou par une procédure d'auto-découverte.

Avant d'établir le pseudo filaire prévu, chaque paire de PE homologues échange des messages de connexion de contrôle pour établir une connexion de contrôle. Chacun annonce les types de pseudo filaire qu'il prend en charge, comme défini dans la [RFC4446], dans l'AVP Liste de capacités de pseudo filaire.

Après que la connexion de contrôle est établie, le PE local examine si le PE distant prend en charge le type de pseudo filaire qu'il a l'intention d'établir. C'est seulement si le PE distant prend en charge le type de pseudo filaire prévu qu'il devrait initier une demande de connexion de pseudo filaire.

Quand le PE local reçoit une ICRQ pour une connexion de pseudo filaire, il examine les identifiants de transmetteur codés dans le AGI, SAII, et TAII afin de déterminer ce qui suit :
- si il a un transmetteur local avec la valeur d'identifiant de transmetteur spécifiée dans la ICRQ,
- si il est permis au transmetteur distant avec l'identifiant de transmetteur spécifié dans la ICRQ de se connecter à ce transmetteur local.

Si les deux conditions sont satisfaites, il envoie une ICRP au PE distant pour accepter la demande de connexion. Si l'une ou l'autre des deux conditions échoue, il envoie une CDN au PE distant pour rejeter la demande de connexion.

Le PE local peut facultativement inclure un code de résultat dans la CDN pour indiquer la cause de la déconnexion. Les codes de résultat possibles sont :

24 – Tentative de connexion à un transmetteur non existant
25 - Tentative de connexion à un transmetteur non autorisé

5.2 Détection de concurrence de pseudo-filaires

On peut concevoir que dans le modèle de référence réseau, comme l'un et l'autre PE peut initier un pseudo filaire à un autre PE à tout moment, les PE pourraient finir par initier un pseudo filaire simultanément l'un à l'autre. Afin d'éviter d'établir des pseudo filaires dupliqués entre deux transmetteurs, chaque PE doit être capable de détecter indépendamment une telle concurrence de pseudo filaires. Les procédures suivantes doivent être suivies pour détecter une concurrence :
- Si le TAII et le SAII sont tous deux présents dans la ICRQ, le PE receveur compare le TAII et le SAII aux SAII et TAII précédemment envoyés au PE envoyeur. Si le TAII reçu correspond au SAII envoyé et si le SAII reçu correspond au TAII envoyé, une concurrence est détectée.
- Si seul le TAII est présent dans la ICRQ, le SAII est supposé avoir la même valeur que le TAII. Le PE receveur compare le TAII reçu avec le SAII envoyé précédemment au PE envoyeur. Si le SAII dans cette ICRQ est aussi omis, alors la valeur codée dans le TAII envoyé est utilisée pour la comparaison. Si ils correspondent, une concurrence est détectée.

- Si l'AGI est présent, il est d'abord ajouté aux valeurs de TAII et de SAII avant que se produise la détection de concurrence.

Une fois qu'une concurrence est découverte, le PE utilise la procédure standard de départage de L2TP, décrite au paragraphe 5.4.4 de la [RFC3931], pour déconnecter le pseudo filaire dupliqué.

5.3 Algorithme générique

On utilise ci-dessous un algorithme générique pour illustrer les interactions de protocole quand on construit un L2VPN en utilisant la signalisation L2TP.

Chaque PE forme d'abord une liste, SOURCE_FORWARDERS, consistant en tous les transmetteurs locaux d'un certain VPN. Ensuite il met tous les transmetteurs locaux qui doivent être interconnectés et tous les transmetteurs distants du même VPN dans une autre liste, TARGET_FORWARDERS. La formation de la topologie du réseau dépend du contenu des listes SOURCE_FORWARDERS et TARGET_FORWARDERS. Ces deux listes peuvent être construites par configuration manuelle ou par une procédure d'auto découverte.

L'algorithme est utilisé pour établir le maillage complet des interconnexions entre SOURCE_FORWARDERS et TARGET_FORWARDERS. Un L2VPN est formé quand l'algorithme est fini dans chaque PE participant de ce L2VPN.

1. Prendre le prochain transmetteur, dans SOURCE_FORWARDERS. Si aucun transmetteur n'est disponible pour le traitement, le processus est achevé.

2. Prendre le prochain transmetteur, dans TARGET_FORWARDERS. Si aucun transmetteur n'est disponible pour le traitement, retourner à l'étape 1.

3. Si les deux transmetteurs sont associés à des identifiants de routeur différents, un pseudo filaire doit être établi entre eux. Passer à l'étape 6.

4. Comparer les valeurs <AGI, AII> des deux transmetteurs. Si elles coïncident, les transmetteurs de source et cible sont les mêmes, de sorte qu'aucune autre action n'est nécessaire. Revenir à l'étape 2.

5. Lorsque les transmetteurs de source et de cible résident tous deux sur le PE local, aucun pseudo filaire n'est nécessaire. Le PE crée simplement une inter connexion locale entre les deux transmetteurs. Revenir à l'étape 2.

6. Lorsque les transmetteurs de source et de cible résident sur des PE différents, un pseudo filaire doit être établi entre eux. Le PE examine d'abord si le transmetteur de source a déjà établi un pseudo filaire pour le transmetteur cible. Si oui, revenir à l'étape 2.

7. Si aucun pseudo filaire n'est déjà établi entre les transmetteurs de source et de cible, le PE local obtient l'adresse du PE distant et établit une connexion de contrôle au PE distant si il n'en existe pas déjà une.

8. Le PE local envoie une ICRQ au PE distant. Les valeurs de AGI, TAII, et SAII sont codées, respectivement, dans l'AVP AGI, l'AVP Identifiant d'extrémité distante, et l'AVP Identifiant d'extrémité locale.

9. Si le PE local reçoit une réponse correspondant à la ICRQ qu'il a juste envoyé, passer à l'étape 10. Autrement, si le PE local reçoit une ICRQ provenant du même PE distant, passer à l'étape 11.

10. Le PE local reçoit une réponse du PE distant. Si c'est une CDN, revenir à l'étape 2. Si c'est une ICRP, le PE local lie le transmetteur de source au pseudo filaire et envoie une ICCN au PE distant. Revenir à l'étape 2.

11. Si le PE local reçoit une ICRQ provenant du même PE distant, il doit effectuer une détection de concurrence de sessions, comme décrit au paragraphe 5.2. Si une concurrence de sessions est détectée, le PE effectue un départage.

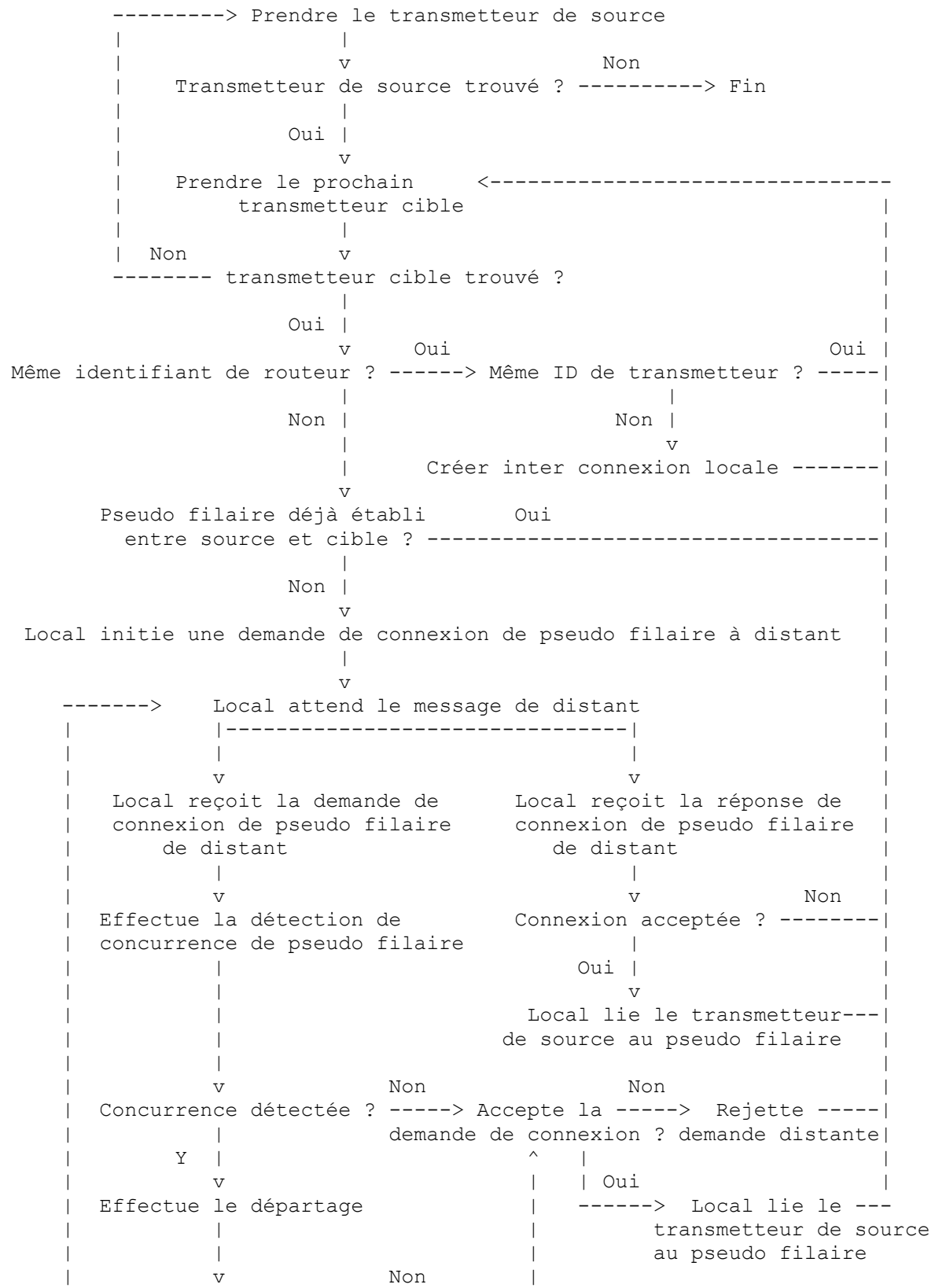
12. Si le PE local perd le départage, il envoie une CDN avec le code de résultat qui indique que la déconnexion est due à la perte du départage. Passer à l'étape 14.

13. Si le PE local gagne le départage, il ignore la ICRQ du PE distant, mais accuse réception du message de contrôle et continue d'attendre la réponse du PE distant. Passer à l'étape 10.

14. Le PE local détermine si il devrait accepter la demande de connexion, comme décrit au paragraphe 5.1. Si il accepte la ICRQ, il envoie une ICRP au PE distant.

15. Le PE local reçoit une réponse du PE distant. Si c'est une CDN, revenir à l'étape 2. Si c'est une ICCN, le PE local lie le transmetteur de source au pseudo filaire, revenir à l'étape 2.

Le diagramme suivant illustre la procédure ci-dessus :



```
|   Gagne le départage ?  ------>   Déconnecte la connexion locale
|           |
|       Oui |
|           v
------ Ignore la demande de connexion distante
```

6. Considérations relatives à l'IANA

La procédure de registre de l'IANA suit dans le présent document celle de la Section 10 de la [RFC3931]. L'IANA a alloué les nouvelles valeurs suivantes aux registres existants gérés par l'IANA.

Le présent document définit trois nouvelles paires de valeur/attribut de message de contrôle L2TP qui ont été allouées par l'IANA. Elles sont décrites au paragraphe 4.3 et sont résumées ci-dessous :

89 – Identifiant de groupe de rattachement
90 – Identifiant d'extrémité locale
91 – Unité maximum de transmission d'interface

Les paragraphes 4.3 et 5.1 définissent trois nouveaux codes de résultat pour le message CDN, qui ont été alloués par l'IANA :

23 – MTU d'interface discordante
24 – Tentative de connexion à un transmetteur non existant
25 – Tentative de connexion à un transmetteur non autorisé

7. Considérations sur la sécurité

La présente spécification n'introduit aucune considérations de sécurité supplémentaires à celles discutées dans les [RFC3931] et [RFC4664].

8. Remerciements

L'auteur tient à remercier Mark Townsley et Carlos Pignataro de leurs précieux apports.

9. Références

9.1 Références normatives

[RFC2119] S. Bradner, "Mots clés à utiliser dans les RFC pour indiquer les niveaux d'exigence", BCP 14, mars 1997. *(MàJ par* RFC8174)

[RFC3931] J. Lau et autres, "Protocole de tunnelage de couche deux - version 3 (L2TPv3)", mars 2005. *(P.S.)*

9.2 Références pour information

[RFC2661] W. Townsley, A. Valencia, A. Rubens, G. Pall, G. Zorn et B. Palter, "Protocole de tunnelage de couche 2 "L2TP"", *(P.S.)*

[RFC4446] L. Martini, "Allocations de l'IANA pour l'émulation de bord à bord pseudo filaire (PWE3)", avril 2006. (BCP0116)

[RFC4664] L. Andersson et E. Rosen, éd., "Cadre pour les réseaux virtuels privés de couche 2 (L2VPN)", septembre

2006. *(Info.)*

[PWE3L2TP] W. Townsley, "Pseudowires and L2TPv3", *Travail en cours*.

Adresse de l'auteur

Wei Luo
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134

mél : luo@cisco.com

Déclaration complète de droits de reproduction

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