Groupe de travail Réseau
Request for Comments : 4623
Catégorie : Sur la voie de la normalisation
Traduction Claude Brière de L'Isle

A. Malis, Tellabs
M. Townsley, Cisco Systems
août 2006

Fragmentation et réassemblage d'émulation de pseudo-filaire bord à bord (PWE3)

Résumé
Le présent document définit une méthode généralisée pour effectuer la fragmentation pour l'usage des protocoles et services d'émulation de pseudo-filaire bord à bord (PWE3, *Pseudowire Emulation Edge-to-Edge*).

Table des matières

1. Introduction

Le document d'architecture d'émulation de pseudo-filaire de bord à bord [RFC3985] définit un modèle de référence de réseau pour PWE3 :

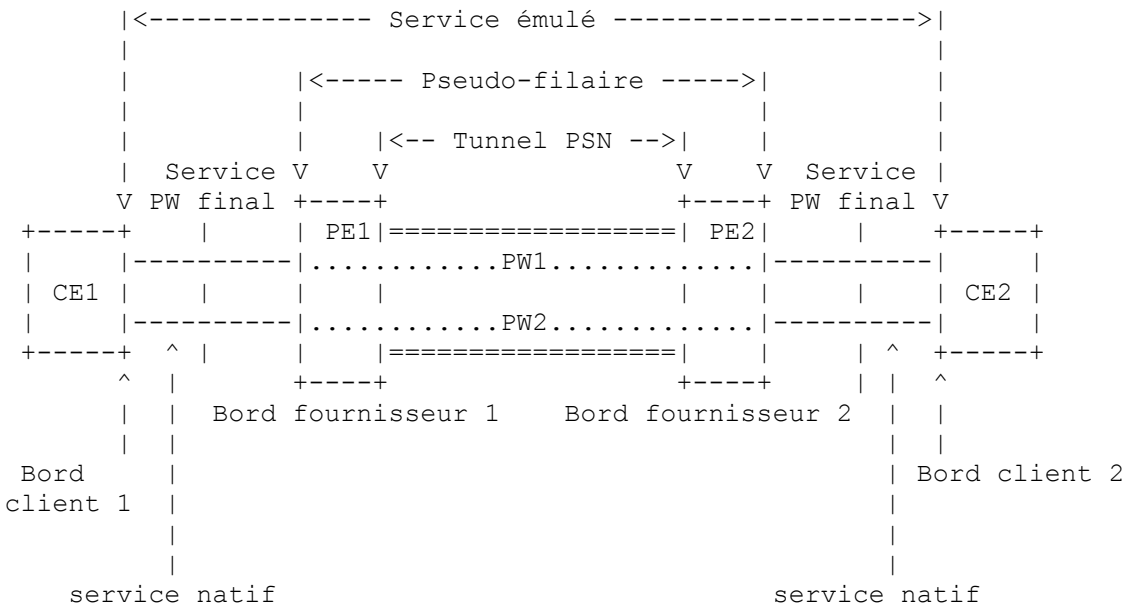

Figure 1 : Modèle de référence de réseau PWE3

Une charge utile de pseudo-filaire (PW, *Pseudowire*) est normalement relayée à travers le PW comme une seule unité de données de protocole (PDU, *Protocol Data Unit*) IP ou de réseau à commutation de paquets (PSN, *Packet Switched Network*) MPLS. Cependant, il y a des cas où la taille combinée de la charge utile et de ses en-têtes PWE3 et PSN associés peut excéder l'unité de transmission maximum (MTU, *Maximum Transmission Unit*) de chemin du PSN. Quand un paquet excède la MTU d'un certain réseau, la fragmentation et le réassemblage vont permettre au paquet de traverser le réseau et d'atteindre la destination prévue.

L'objet du présent document est de définir une méthode généralisée pour effectuer la fragmentation à utiliser avec tous les protocoles et services de PWE3. Cette méthode devrait être utilisée seulement dans les cas où les méthodes de gestion de MTU échouent. Du fait de l'augmentation des frais généraux de traitement, la fragmentation et le réassemblage dans les appareils de cœur de réseau devrait toujours être considéré comme quelque chose à éviter chaque fois que possible.

Le domaine de fragmentation et réassemblage PWE3 est montré à la Figure 2 :

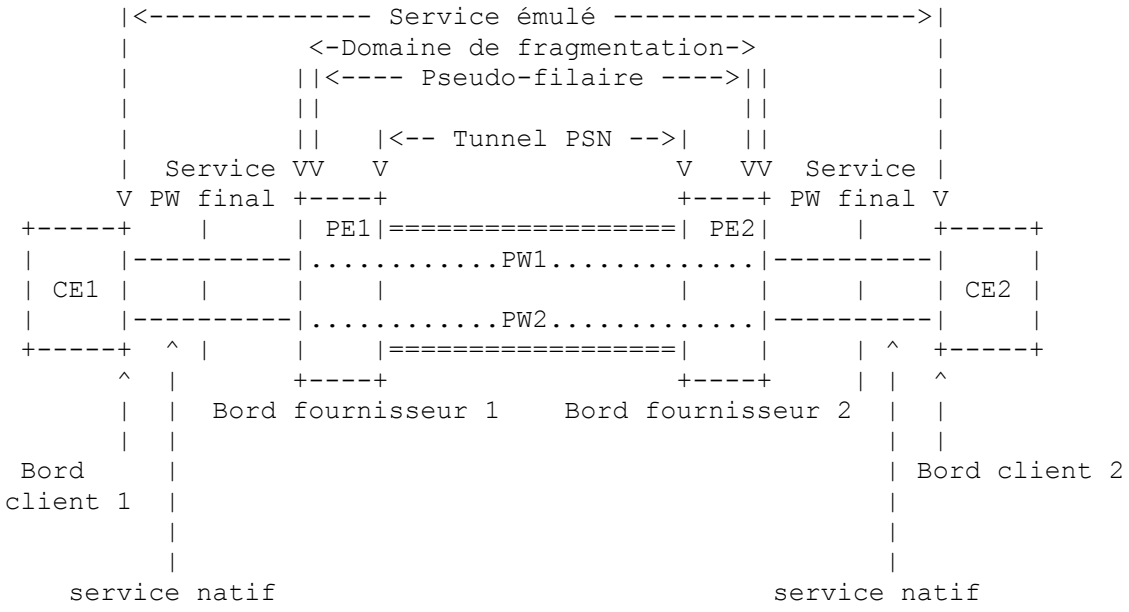

Figure 2 : Domaine de fragmentation/réassemblage PWE3

La fragmentation a lieu dans le PE émetteur immédiatement avant l'encapsulation de PW, et le réassemblage a lieu dans le PE receveur immédiatement après la désencapsulation de PW.

Comme un numéro de séquence est nécessaire pour les procédures de fragmentation et réassemblage, l'utilisation du champ Numéro de séquence sur les paquets fragmentés est EXIGÉE (voir aux paragraphes 4.1 et 5.5 la localisation des champs Numéro de séquence pour les encapsulations, respectivement, MPLS et L2TPv3). L'ordre des opérations est que d'abord la fragmentation est effectuée, et ensuite les fragments résultant ont des numéros de séquence alloués à la suite.

Selon l'encapsulation PWE3 spécifique utilisée, la valeur 0 peut ne pas faire partie de l'espace de numéros de séquence, et dans ce cas son utilisation pour la fragmentation doit suivre cette même règle : lorsque le numéro de séquence est incrémenté, il saute zéro et revient de 65535 à 1. À l'inverse, si la valeur 0 fait partie de l'espace de séquence, le même espace de séquence est aussi utilisé pour la fragmentation et le réassemblage.

2. Conventions utilisées dans ce document

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

3. Solutions de remplacement à la fragmentation/réassemblage PWE3

La fragmentation et le réassemblage dans les équipements de réseau exige généralement des resssources significatlivement plus grandes que d'envoyer un paquet comme une seule unité. À ce titre, la fragmentation et le réassemblage devraient être évités chaque fois que possible. Les solutions idéales pour éviter la fragmentation incluent une configuration appropriée et la gestion des tailles de MTU entre le routeur de bord client (CE, *Customer Edge*) et le routeur de bord fournisseur (PE, *Provider Edge*) et à travers le PSN, ainsi que des mesures adaptatives qui opèrent avec l'hôte d'origine (par exemple, [RFC1191], [RFC1981]) pour réduire les tailles de paquet à la source.

Dans certains cas, un PE peut être capable de fragmenter un paquet IPv4 [RFC0791] avant qu'il entre dans un PW. Par exemple, si le PE peut fragmenter et transmettre des paquets IPv4 avec le bit DF à zéro d'une manière identique à un routeur IPv4, il peut fragmenter les paquets arrivant d'un CE, transmettre les fragments IPv4 avec le tramage associé pour ce circuit de rattachement (AC, *attachment circuit*)sur le PW. Architecturalement, la fragmentation IPv4 se produit avant d'atteindre le PW, présentant plusieurs trames au PW à transmettre de la manière normale pour ce type de PW. Donc, cette méthode est entièrement transparente à l'encapsulation de PW et à l'extrémité distante du PW lui-même. Les fragments de paquet sont finalement réassemblés chez l'hôte de destination IPv4 de la façon normale. Les paquets IPv6 ne sont pas fragmentés de cette manière.

4. Fragmentation de PWE3 avec MPLS

Quand on utilise les procédures de signalisation de la [RFC4447], il y a un type de sous TLV Paramètre d'interface de pseudo-filaire qui est utilisé pour signaler l'utilisation de la fragmentation quand on annonce une étiquette de VC [RFC4446] :

Paramètre	Longueur	Description
0x09	4	Indicateur de fragmentation

La présence de ce paramètre dans l'élément FEC de VC indique que le receveur est capable de réassembler les fragments quand le mot de contrôle est utilisé pour l'étiquette de VC annoncée. Cela n'oblige pas l'envoyeur à utiliser la fragmentation ; c'est simplement une indication que l'envoyeur PEUT utiliser la fragmentation. L'envoyeur NE DOIT PAS utiliser la fragmentation si ce paramètre n'est pas présent dans l'élément FEC de VC.

Si la signalisation de la [RFC4447] n'est pas utilisée, il DOIT être configuré par l'envoyeur que la fragmentation doit ou non être utilisée.

4.1 Localisations de bit de fragment pour MPLS

PWE3 fondé sur MPLS utilise le format de mot de contrôle suivant [RFC4385], avec les bits de fragmentation B et E identifiés dans les positions 8 et 9 :

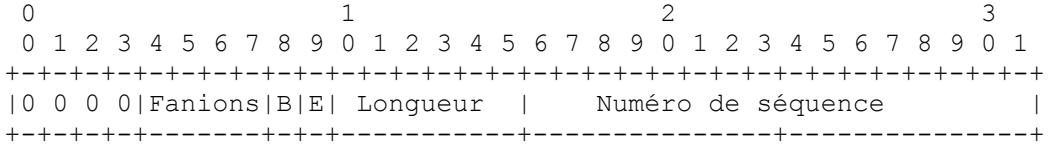

Figure 3 : Mot de contrôle MPLS de PW préféré

Les bits B et E sont définis comme suit :

BE
00 indique que la charge utile entière (non fragmentée) est portée dans un seul paquet
01 indique que le paquet porte le premier fragment
10 indique que le paquet porte le dernier fragment
11 indique un paquet portant un fragment intermédiaire

Voir à l'Appendice A la discussion de la déduction de ces valeurs pour les bits B et E.

Voir à la Section 1 la description de l'utilisation du champ Numéro de séquence.

4.2 Autres considérations

La MTU du chemin [RFC1191], [RFC1981], peut être utilisée pour déterminer dynamiquement la taille maximum des fragments. L'application de la MTU de chemin à MPLS est discutée dans la [RFC3032]. La taille maximum des fragments peut aussi être configurée. Le paramètre de MTU d'interface signalé dans la [RFC4447] DEVRAIT être utilisé pour régler la taille maximum de la mémoire tampon de réassemblage pour les paquets reçus pour faire un usage optimal des ressources de mémoire tampon de réassemblage.

5. Fragmentation PWE3 avec L2TP

Cette Section définit la localisation des bits B et E pour les en-têtes L2TPv3 [RFC3931] et L2TPv2 [RFC2661], ainsi que le mécanisme de signalisation pour annoncer les valeurs d'unité maximum de réception (MRU, *Maximum Receive Unit*) et la prise en charge de la fragmentation sur un certain PW. Comme IP est le PSN le plus couramment utilisé avec L2TP, la fragmentation et le réassemblage de PSN IP est aussi discutée.

5.1 Fragmentation spécifique de PW contre fragmentation IP

Lorsque une gestion appropriée de MTU échoue à travers un réseau, la fragmentation et réassemblage de PSN IP peut être utilisée pour s'accommoder des discordances de MTU entre les points d'extrémité de tunnel. Si le trafic global qui exige la fragmentation et réassemblage est très léger, ou lorsque il y a des mécanismes optimisés suffisants disponibles pour la fragmentation et réassemblage de PSN IP, la fragmentation et réassemblage de PSN IP peut être suffisante.

Quand on fait face à un grand nombre de paquets de PW qui exigent la fragmentation et réassemblage, une méthode spécifique du PW a des propriétés qui peuvent permettre des mises en œuvre plus économes en ressources. Précisément, la capacité d'allouer une utilisation de mémoire tampon sur la base du PW et le séquençage de PW peuvent être utilisés pour prendre l'avantage sur un mécanisme général s'appliquant à tous les paquets IP sur tous les PW. De plus, la fragmentation de PW peut être plus facilement activée d'une manière sélective pour certains PW ou tous, plutôt que d'activer le réassemblage pour tout le trafic IP qui arrive à un certain nœud.

Les déploiements DEVRAIENT éviter une situation qui utiliserait une combinaison de fragmentation et réassemblage de PSN IP et de PW sur le même nœud. Un tel fonctionnement irait clairement à l'encontre de l'objet du mécanisme défini dans le présent document. Ceci est particulièrement important pour les pseudo-filaires L2TPv3, car la fragmentation peut avoir lieu dans trois endroits différents (le PSN IP, le PW, et la charge utile encapsulée). On doit veiller à s'assurer que les

valeurs de MTU/MRU sont correctement réglées et annoncées à chaque point d'extrémité de tunnel pour éviter cela. Quand la fragmentation est activée au sein d'un certain PW, le bit DF DOIT être établi sur tous les L2TP sur les paquets IP pour ce PW.

Les nœuds L2TPv3 DEVRAIENT participer à la MTU de chemin ([RFC1191], [RFC1981]) pour l'ajustement automatique de la MTU du PSN. Quand la charge utile est IP, la MTU de chemin devrait être utilisée aussi au niveau de leur charge utile.

5.2 Annonce de la prise en charge du réassemblage dans L2TP

Les constructions définies dans ce paragraphe pour annoncer la prise en charge de la fragmentation dans L2TP sont applicables à L2TPv3 [RFC3931] et L2TPv2 [RFC2661].

Le présent document définit deux nouvelles AVP pour annoncer les valeurs d'unité de réception maximum et la prise en charge du réassemblage. Ces AVP PEUVENT être présentes dans les messages Demande d'appel entrant (ICRQ, *Incoming-Call-Request*), Réponse d'appel entrant (ICRP, *Incoming-Call-Reply*), Appel entrant connecté (ICCN, *Incoming-Call-Connected*), Demande d'appel sortant (OCRQ, *Outgoing-Call-Request*), Réponse d'appel sortant (OCRP, *Outgoing-Call-Reply*), Appel sortant connecté (OCCN, *Outgoing-Call-Connected*), ou Informations d'établissement de liaison (SLI, *Set-Link-Info*). La plus récente valeur reçue prend toujours la préséance sur une valeur antérieure et DOIT être dynamique sur la vie de la session si elle est reçue via le message SLI. Une des deux nouvelles AVP (MRRU) est utilisée pour annoncer que le réassemblage PWE3 est pris en charge par l'envoyeur de l'AVP. La prise en charge du réassemblage PEUT être unidirectionnelle.

5.3 AVP L2TP Unité maximum de réception (MRU)

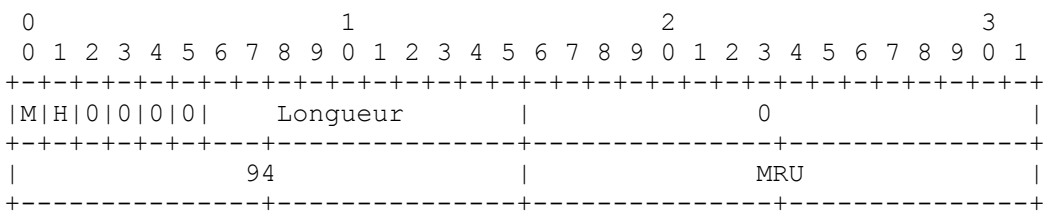

Figure 4 : AVP L2TP Unité maximum de réception (MRU)

L'AVP Unité maximum de réception (MRU, *Maximum Receive Unit*) numéro d'attribut 94, est la taille maximum, en octets, d'une trame PW fragmentée ou complète, incluant l'encapsulation L2TP, recevable par le côté du PW qui annonce cette valeur. La MRU annoncée N'inclut PAS l'en-tête de PSN (c'est-à-dire, l'en-tête IP et/ou UDP). Cette AVP n'implique pas que la fragmentation ou réassemblage PWE3 est pris en charge. Si le réassemblage n'est pas activé ou est indisponible, cette AVP peut être utilisée seule pour annoncer la MRU pour une trame complète.

Cette AVP PEUT être cachée (le bit H PEUT être 0 ou 1). Le bit M (mandatory, *obligatoire*) pour cette AVP DEVRAIT être réglé à 0. Le champ Longueur (avant de cacher) est 8. L'identifiant de fabricant est identique à celui de l'IETF qui est 0.

5.4 AVP L2TP Unité maximum réassemblée de réception (MRRU)

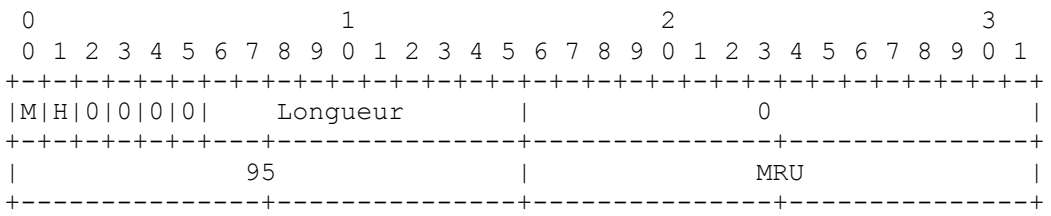

Figure 5 : AVP L2TP Unité maximum réassemblée de réception (MRRU)

L'AVP Unité maximum réassemblée de réception (MRRU, *Maximum Reassembled Receive Unit*) numéro d'attribut 95, est la taille maximum, en octets, d'une trame réassemblée, incluant tout tramage de PW, mais n'incluant pas l'encapsulation L2TP ou la sous couche spécifique de couche 2. La présence de cette AVP signifie la capacité de recevoir les fragments de

PW et de les réassembler. Les fragments de paquet NE DOIVENT PAS être envoyés à un homologue qui n'a pas reçu cette AVP dans un message de contrôle. Si la MRRU est présente dans un message, l'AVP MRU DOIT être aussi présente.

La MRRU DEVRAIT être utilisée pour régler la taille maximum de la mémoire tampon de réassemblage pour les paquets reçus pour faire un usage optimal des ressources de mémoire tampon de réassemblage.

Cette AVP PEUT être cachée (le bit H PEUT être 0 ou 1). Le bit M pour cette AVP DEVRAIT être réglé à 0. Le champ Longueur (avant de le cacher) est 8. L'identifiant de fabricant est identique à celui de l'IETF qui est 0.

5.5 Localisations de bit Fragment pour encapsulation L2TPv3

L'usage des bits B et E est décrit au paragraphe 4.1. Pour l'encapsulation L2TPv3, les bits B et E sont définis comme bits 2 et 3 dans les bits de tête de la sous couche spécifique de couche 2 par défaut (voir la Section 7).

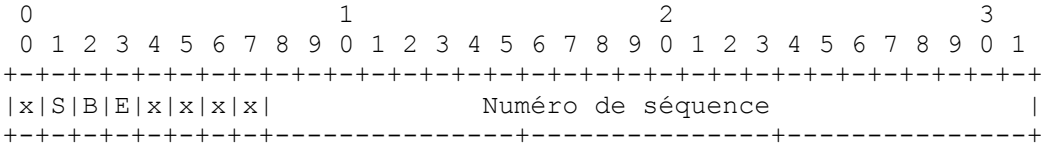

Figure 6 : Localisation des bits B et E dans la sous couche spécifique de couche 2 par défaut

Le bit S (Séquence) est comme défini dans la [RFC3931]. La localisation des bits B et E pour les types de PW qui utilisent une variante de sous couche spécifique de couche 2 sort du domaine d'application de ce document.

Quand la fragmentation est utilisée, une sous couche spécifique de couche 2 avec les bits B et E établis DOIT être présente dans tous les paquets de données pour une session donnée. La présence et le format de la sous couche spécifique de couche 2 est annoncée via l'AVP Sous couche spécifique de couche 2, type d'attribut 69, définie au paragraphe 5.4.4 de la [RFC3931].

Voir à la Section 1 la description de l'utilisation du champ Numéro de séquence.

5.6 Localisations des bits Fragment pour encapsulation L2TPv2

L'usage des bits B et E est décrit au paragraphe 4.1. Pour l'encapsulation L2TPv2, les bits B et E sont définis comme les bits 8 et 9 dans les bits de tête de l'en-tête L2TPv2 décrit ci-dessous (voir la Section 7).

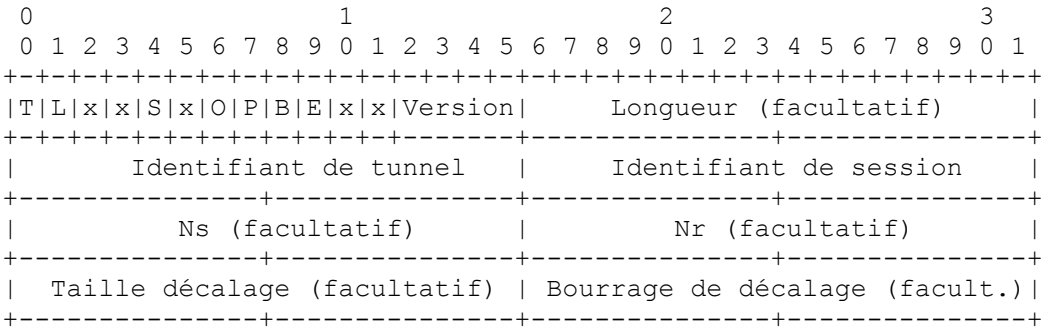

Figure 7 : Localisation des bits B et E dans l'en-tête de message L2TPv2

6. Considérations sur la sécurité

Comme avec toute construction de protocole supplémentaire, chaque niveau de complexité ajoute un potentiel d'exploitation des erreurs de protocole et de mise en œuvre. Les mises en œuvre devraient être particulièrement attentives à ne pas lier une abondance de ressources, même pour les plus pathologiques combinaisons de fragments de paquet qui pourraient être reçues. Au delà de ces problèmes de qualité générale de mise en œuvre, il n'y a pas de problème de sécurité notable connu d'utilisation du mécanisme défini dans ce document. On devrait noter que la RFC 1990, sur laquelle se fonde

le présent document, et ses dérivés, a été largement mise en œuvre et utilisée de façon extensive dans l'Internet et ailleurs.

Les [RFC1858] et [RFC3128] décrivent les attaques potentielles de réseau associées à la fragmentation et réassemblage IP. Les problèmes décrits dans ces documents tentent de contourner les contrôles d'accès IP par l'envoi de "petits fragments" formés avec soin, ou en exploitant le champ Décalage IP de façon à causer le chevauchement des fragments et de reécrire les portions intéressantes d'un paquet IP après que les vérifications d'accès ont été effectuées. Ce dernier n'est pas un problème avec la méthode de fragmentation spécifique de PW décrite dans ce document, car il n'y a pas de champ de décalage. Cependant, les mises en œuvre DOIVENT être sûres de ne pas permettre que plus d'un fragment complet en écrase un autre dans une trame reconstruite. Le premier peut être un souci si le filtrage de paquet et les contrôles d'accès sont placés sur des trames tunnelées au sein de l'encapsulation de PW. Pour circonvenir toute attaque possible dans l'un et l'autre cas, tous les filtrages et contrôles d'accès devraient être appliqués à la trame reconstruite résultante plutôt qu'à tout fragment de PW.

7. Considérations relatives à l'IANA

Le présent document ne définit aucun nouveau registre à tenir par l'IANA.

Noter que la [RFC4446] a déjà alloué le paramètre d'interface d'indicateur de fragmentation, de sorte qu'aucune autre action de l'IANA n'est requise.

Le présent document demande à l'IANA d'allouer de nouvelles valeurs pour les registres déjà gérés par l'IANA (paragraphes 7.1 et 7.2) et deux bits réservés dans un en-tête existant (paragaphe 7.3).

7.1 Paires Attribut Valeur (AVP) Message de contrôle

Deux attributs d'AVP supplémentaires sont spécifiés aux paragraphes 5.3 et 5.4. Il est demandé qu'ils soient définis par l'IANA comme décrit au paragraphe 2.2 de la [RFC3438].

Paires Attribut Valeur Message de contrôle
94 : AVP Unité de réception maximum (MRU)
95 : AVP Unité de réception maximum réassemblée (MRRU)

7.2 Bits de sous couche spécifique de couche 2 par défaut

Ce registre a été créé au titre de la publication de la [RFC3931]. Le présent document définit deux bits réservés dans la sous couche spécifique de couche 2 par défaut au paragraphe 5.5, qui peuvent être alloués par consensus de l'IETF [RFC2434]. Il est demandé qu'ils soient alloués par l'IANA.

Bits de sous couche spécifique de couche 2 par défaut – selon la [RFC3931]
Bit 2 – bit B (Fragmentation)
Bit 3 – bit E (Fragmentation)

7.3 Bits de tête de l'en-tête de message L2TPv2

Le présent document demande la définition de deux bits réservés dans l'en-tête L2TPv2 [RFC2661]. Les localisations sont notées par les bits "B" et "E" au paragraphe 5.6.

Bits de tête de l'en-tête de message L2TPv2 – selon les [RFC2661], [RFC3931]
Bit 8 – bit B (Fragmentation)
Bit 9 – bit E (Fragmentation)

8. Remerciements

Les auteurs tiennent à remercier Eric Rosen et Carlos Pignataro, de Cisco Systems, de leur relecture de ce document.

9. Références normatives

[RFC1191] J. Mogul et S. Deering, "Découverte de la MTU de chemin", novembre 1990.

[RFC1981] J. McCann, S. Deering, J. Mogul, "Découverte de la MTU de chemin pour IP version 6", août 1996. *(D.S. ; Remplacé par* [RFC8201], STD87)

[RFC1990] K. Sklower et autres, "Protocole multi liaisons en PPP (MP)", août 1996. (*Remplace* RFC1717) *(D.S.)*

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

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

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

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

[RFC4385] S. Bryant et autres, "Mot de contrôle d'émulation bord à bord pseudo filaire (PWE3) à utiliser sur un PSN MPLS", février 2006. *(P.S.)*

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

[RFC4447] L. Martini et autres, "Établissement et maintenance de pseudo filaires avec le protocole de distribution d'étiquettes", avril 2006. (*MàJ par la* RFC6723) *(P.S. ; Remplacé par* RFC8077 STD 84)

10. Références pour information

[FAST] ATM Forum, "Frame Based ATM over SONET/SDH Transport (FAST)", af-fbatm-0151.000, juillet 2000.

[FRF.12] Frame Relay Forum, "Frame Relay Fragmentation Implementation Agreement", FRF.12, décembre 1997.

[RFC0791] J. Postel, éd., "Protocole Internet - Spécification du protocole du programme Internet", STD 5, septembre 1981.

[RFC1858] G. Ziemba, D. Reed, P. Traina, "Considérations sur la sécurité pour le filtrage de fragment IP", octobre 1995. *(Mise à jour par la* RFC3128*) (Information)*

[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)

[RFC3128] I. Miller, "Protection contre une variante de l'attaque de petit fragment (RFC1858)", juin 2001. *(Info.)*

[RFC3438] W. Townsley, "Mise à jour des considérations de l'IANA sur le protocole de tunnelage de couche deux (L2TP)", décembre 2002. (BCP0068)

[RFC3985] S. Bryant et autres, "Architecture d'émulation bord à bord pseudo-filaire (PWE3)", mars 2005. *(Information)*

Appendice A. Relations entre ce document et la RFC 1990

La fragmentation de grands paquets en plus petites unités pour la transmission n'est pas nouvelle. Une méthode de fragmentation et réassemblage a été définie dans la [RFC1990] pour PPP multi liaison. Cette méthode a aussi été adoptée pour les technologies de réseau de relais de trame [FRF.12] et ATM [FAST]. Le présent document adopte aussi les procédures de fragmentation et réassemblage de la RFC 1990, avec quelques modifications décrites dans cet appendice. On

suppose une certaine familiarité avec la RFC 1990.

La RFC 1990 a été conçue pour être utilisée dans des environnements où des fragments de paquet peuvent arriver dans le désordre à cause de leur transmission sur plusieurs liaisons en parallèle, en spécifiant que la mise en mémoire tampon soit utilisée pour placer les fragments dans l'ordre correct. Pour PWE3, la capacité de réordonner les fragments avant le réassemblage est FACULTATIVE ; les receveurs PEUVENT choisir d'éliminer les trames quand un fragment perdu est détecté. Donc, quand le numéro de séquence sur les fragments reçus montre qu'un fragment a été sauté, le paquet partiellement réassemblé PEUT être éliminé, ou le receveur PEUT souhaiter attendre que le fragment arrive décalé. Dans ce cas, un temporisateur de réassemblage DOIT être utilisé pour éviter de bloquer des ressources de mémoire tampon pendant trop longtemps.

Éliminer les fragments en désordre sur un certain PW peut fournir un avantage d'adaptabilité considérable pour les équipements de réseau qui effectuent le réassemblage. Si des fragments décalés sont un événement relativement rare sur un certain PW, le débit ne devrait pas être contrarié par cela. Noter cependant, que si il y a des cas où les fragments d'une certaine trame sont reçus en désordre de façon répétée (par exemple, un bref fragment est toujours passé devant un plus gros) l'élimination des fragments déclassés va être cause que la trame fragmentée ne sera jamais reçue. Cette condition peut résulter en un efficace déni de service à une application de niveau supérieur. À ce titre, les mises en œuvre qui fragmentent une trame PW DOIVENT au moins s'assurer que tous les fragments sont envoyés dans l'ordre à partir de leur propre point de sortie.

Une mise en œuvre peut aussi choisir de permettre le réassemblage d'un nombre limité de trames fragmentées sur un certain PW, ou à travers un ensemble de PW avec le réassemblage activé. Cela permet une distribution plus équitable des ressources de réassemblage, réduisant les chances qu'un seul ou un petit ensemble de PW puisse épuiser touts les ressources de réassemblage d'un nœud. Comme avec l'élimination des fragments en désordre, on peut percevoir des cas où cela peut aussi entraîner un déni de service effectif. Par exemple, si des fragments de multiples trames sont reçus de façon persistente avant que chaque trame puisse être reconstruite dans un ensemble limité de mémoires tampon de réassemblage de PW, une partie de ces trames fragmentées ne sera jamais livrée.

Les en-têtes de la RFC 1990 utilisent deux bits qui indiquent le premier et le dernier fragment d'une trame, et un numéro de séquence. Le numéro de séquence peut être long de 12 ou 24 bits (d'après la [RFC1990]) :

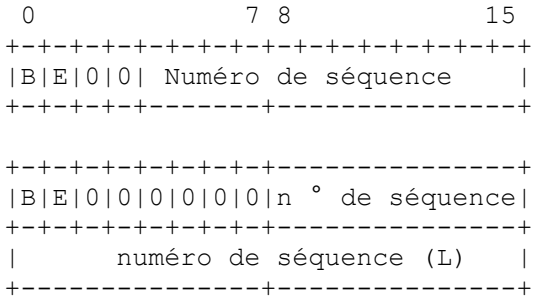

Figure 6 : Formats d'en-tête de la RFC 1990

La fragmentation PWE3 tire parti des numéros de séquence existants de PW et des champs de bits de contrôle chaque fois que possible, plutôt que de définir un en-tête séparé pour l'usage exclusif de la fragmentation. Donc, elle n'utilise aucun des formats de numéro de séquence de la RFC 1990 décrits ci-dessus, s'appuyant plutôt sur le numéro de séquence qui existe déjà dans l'en-tête PWE3.

La RFC 1990 définit deux champs d'un bit : un bit de début de fragment (B) et un bit de fin de fragment (E). Le bit B est réglé à 1 sur le premier fragment dérivé d'un paquet PPP et réglé à 0 pour tous les autres fragments du même paquet PPP. Le bit E est réglé à 1 sur le dernier fragment et réglé à 0 pour tous les autres fragments. Une trame complète non fragmentée a les deux bits B et E réglés à 1.

La fragmentation PWE3 inverse la valeur des bits B et E, tout en conservant le concept opérationnel de marquer le début et la fin d'une trame fragmentée. Donc, pour PW le bit B est réglé à 0 sur le premier fragment dérivé d'une trame PW et établi à 1 pour tous les autres fragments dérivés de la même trame. Le bit E est réglé à 0 sur le dernier fragment et établi à 1 pour tous les autres fragments. Une trame complète non fragmentée a les deux bits B et E réglés à 0. Le motif de cette inversion de valeur pour les bits B et E est de permettre que les trames complètes (et en particulier, les mises en œuvre qui ne prennent en charge que les trames complètes) laissent simplement les bits B et E réglés à 0 dans l'en-tête.

Pour prendre en charge la fragmentation, les bits B et E DOIVENT être définis ou identifiés pour tous les protocoles de tunnelage PWE3. Les Sections 4 et 5 définissent ces localisations pour les protocoles de tunnelage PWE3 MPLS [RFC4385], L2TPv2 [RFC2661], et L2TPv3 [RFC3931].

Adresse des auteurs

Andrew G. Malis
Tellabs
1415 West Diehl Road
Naperville, IL 60563
mél : Andy.Malis@tellabs.com

W. Mark Townsley
Cisco Systems
7025 Kit Creek Road
PO Box 14987
Research Triangle Park, NC 27709
mél : mark@townsley.net

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.