Groupe de travail Réseau
Request for Comments : 1853
Catégorie : Information

W. Simpson, Daydreamer
octobre 1995
Traduction Claude Brière de L'Isle

Tunnelage IP dans IP

Statut du présent mémoire
Le présent mémoire apporte des informations pour la communauté de l'Internet. Il ne spécifie aucune norme de l'Internet. La distribution du présent mémoire n'est soumise à aucune restriction.

Note de l'IESG :
Noter que le présent mémoire est un effort individuel de l'auteur. Ce document reflète des pratiques informelles actuelles dans l'Internet. Des travaux sont en cours au sein du groupe de travail IP Mobile de l'IETF pour produire une proposition de norme appropriée pour traiter cette question.

Résumé
Le présent document expose les techniques mises en œuvre pour utiliser l'encapsulation numéro 4 de protocole/charge utile IP pour le tunnelage avec la sécurité IP et d'autres protocoles.

Table des matières

1. Introduction

L'encapsulation de protocole/charge utile de IP dans IP numéro 4 de la [RFC1700] a longtemps été utilisé pour faire un pont entre des portions de l'Internet qui ont des capacités ou politiques disjointes. Le présent document décrit les techniques de mise en œuvre utilisées depuis de nombreuses années par le réseau par paquets des radioamateurs pour constituer un grand réseau mobile, et aussi par les premières mises en œuvre des protocoles de sécurité IP.

L'utilisation de l'encapsulation IP dans IP diffère des techniques plus récentes de tunnelage (par exemple, les numéros de protocole 98 [RFC1241], 94 [IDM91a], 53 [swIPe], et 47 [RFC1701]) en ce qu'elle n'insère pas son propre en-tête particulier entre les en-têtes IP. Au lieu de cela, l'en-tête IP original non modifié est conservé, et simplement enveloppé dans un autre en-tête IP standard.

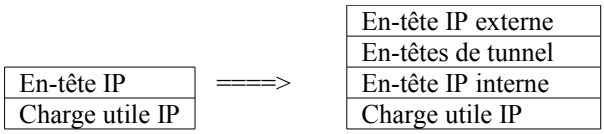
Ces informations s'appliquent principalement à l'encapsulation de IP version 4. D'autres versions IP seront décrites dans des documents ultérieurs.

2. Encapsulation

La technique d'encapsulation est très simple. Un en-tête IP externe est ajouté avant l'en-tête IP original. Entre eux se trouvent tous les autres en-têtes pour le chemin, tels que les en-têtes de sécurité spécifiques de la configuration de tunnel.

La source et la destination de l'en-tête IP externe identifient les "points d'extrémité" du tunnel. La source et la destination de l'en-tête IP interne identifient l'envoyeur et le receveur original du datagramme.

Chaque en-tête s'enchaîne au suivant en utilisant les valeurs du protocole IP [RFC1700].



Le format des en-têtes IP est décrit dans la [RFC0791].

Type de service : copié de l'en-tête IP interne. Facultativement, un autre TOS peut être utilisé entre homologues coopérants. Ceci est conforme au principe de transparence qui veut que si l'utilisateur s'attend à un certain niveau de service, le tunnel devrait fournir ce niveau de service. Cependant, certains tunnels peuvent être construits spécifiquement pour fournir un niveau de service différent selon un choix de politique.

Identification : un nouveau numéro est généré pour chaque en-tête IP externe.
Le datagramme encapsulé peut avoir été déjà fragmenté, et un autre niveau de fragmentation peut survenir du fait de l'encapsulation dans le tunnel. Ces fragments de tunnel seront rassemblés par le désencapsuleur, plutôt que par la destination finale.

Réservé : ignoré (mis à zéro).
Ce fanion non officiel a eu une utilisation expérimentale, et bien qu'il reste dans l'en-tête IP interne, n'affecte pas le tunnel.

Ne pas fragmenter *(DF)* : copié de l'en-tête IP interne. Cela permet à l'origine de contrôler le niveau des compromis sur les performances. Voir à "Découverte de la MTU de tunnel".

Autres fragments *(MF)* : réglé comme il convient en cas de fragmentation.
Le fanion n'est pas copié parce qu'une identification distincte est utilisée.

Durée de vie *(TTL)* : c'est la valeur par défaut spécifiée dans la plus récente version des "Numéros alloués" [RFC1700]. Cela assure que de longs tunnels imprévus n'interrompent pas le flux des datagrammes entre les points d'extrémité. Le TTL interne est décrémenté une fois avant l'encapsulation, et n'est pas affecté par la désencapsulation.

Protocole : c'est celui du prochain e,-tête ; 4 pour l'en-tête IP interne, lorsque aucun autre en-tête de tunnel n'est utilisé.

Source ; c'est une adresse IP associée à l'interface utilisée pour envoyer le datagramme.

Destination : c'est une adresse IP du désencapsuleur du tunnel.

Options : elles ne sont pas copiées de l'en-tête IP interne. Cependant, de nouvelles options particulières au chemin PEUVENT être ajoutées.

Horodatage, Route de source lâche, Route de source stricte, et Route enregistrée sont délibérément cachés au sein du tunnel. Les tunnels sont souvent construits pour surmonter l'inadéquation de ces options.

La prise en charge d'options de sécurité de l'en-tête IP interne PEUT affecter le choix des options de sécurité pour le tunnel. Il n'est pas prévu qu'il y ait une correspondance biunivoque de telles options avec les en-têtes d'options ou de sécurité choisis pour le tunnel.

3. Gestion de tunnel

Il est possible que l'un des routeurs le long de l'intérieur du tunnel rencontre une erreur lors du traitement du datagramme, ce qui l'amènera à retourner un message d'erreur ICMP [RFC0792] à l'encapsuleur à la source IP du tunnel. Malheureusement, ICMP exige seulement des routeurs IP qu'ils retournent 8 octets (64 bits) du datagramme au delà de l'en-tête IP.

Ce n'est pas suffisant pour inclure l'en-tête encapsulé entier. Donc, il n'est généralement pas possible à un routeur encapsulant de refléter immédiatement un message ICMP provenant de l'intérieur d'un tunnel à l'hôte d'origine.

Cependant, en entretenant avec soin un "état conditionnel" sur ses tunnels, l'encapsuleur peut retourner des messages ICMP précis dans la plupart des cas. Le routeur DEVRAIT entretenir au moins les informations d'état conditionnel suivantes sur chaque tunnel :

- accessibilité de l'extrémité du tunnel,

- encombrement du tunnel,
- MTU du tunnel.

Le routeur utilise les messages ICMP qu'il reçoit de l'intérieur d'un tunnel pour mettre à jour les informations d'état conditionnel pour ce tunnel. Lorsque arrivent ultérieurement des datagrammes qui devraient transiter par le tunnel, le routeur vérifie l'état conditionnel pour le tunnel. Si le datagramme devrait violer l'état du tunnel (comme avec une MTU supérieure à la MTU du tunnel lorsque le fanion Ne pas fragmenter est mis) le routeur renvoie un message d'erreur ICMP approprié à l'origine, mais transmet aussi le datagramme dans le tunnel. La transmission du datagramme en dépit de l'envoi du message d'erreur assure que les changements de l'état du tunnel seront découverts.

En utilisant cette technique, les messages d'erreur ICMP provenant de routeurs encapsulants ne vont pas toujours correspondre de façon univoque avec les erreurs rencontrées au sein du tunnel, mais ils vont refléter précisément l'état du réseau.

3.1 Découverte de la MTU de tunnel

Lorsque le bit Ne pas fragmenter est mis par l'origine et copié dans l'en-tête IP externe, la MTU appropriée du tunnel sera apprise par les erreurs ICMP (Type 3 Code 4) "Datagramme trop gros" rapportées à l'encapsuleur. Pour prendre en charge les hôtes générateurs qui utilisent cette capacité, toutes les mises en œuvre DOIVENT prendre en charge la découverte de la MTU du chemin [RFC1191], [RFC1435] au sein de leurs tunnels.

Au bénéfice de la découverte de la MTU de tunnel, toute fragmentation qui survient à cause de la taille de l'en-tête d'encapsulation est faite une seule fois après l'encapsulation. Cela empêche d'avoir plus d'une fragmentation d'un seul datagramme, afin d'améliorer l'efficacité du traitement des routeurs du chemin et du désencapsuleur du tunnel.

3.2 Encombrement

L'état conditionnel du tunnel va collecter les indications d'encombrement, tels qu'un Source éteinte ICMP (Type 4) dans les datagrammes provenant du désencapsuleur (homologue du tunnel). Lorsque on transmet un autre datagramme dans le tunnel, il est approprié d'envoyer des messages Source éteinte à l'origine.

3.3 Échecs d'acheminement

Comme le TTL est remis à zéro chaque fois qu'un datagramme est encapsulé, les boucles d'acheminement au sein d'un tunnel sont particulièrement dangereuses lorsque elles reviennent à l'encapsuleur. Si la source IP correspond à une de ses interfaces, un mise en œuvre NE DOIT PAS réencapsuler. Le datagramme doit plutôt être transmis normalement.

Les messages ICMP (Type 11) Temps dépassé rapportent des boucles d'acheminement au sein du tunnel lui-même. Les messages ICMP (Type 3) Destination injoignable rapportent des échecs de livraison au désencapsuleur. Cet état conditionnel DOIT être rapporté à l'origine comme (Type 3 Code 0) Réseau injoignable.

3.4 Autres messages ICMP

La plupart des messages d'erreur ICMP ne sont pas pertinents pour l'usage du tunnel. En particulier, les problèmes de paramètres sont vraisemblablement le résultat d'une mauvaise configuration de l'encapsuleur, et NE DOIVENT PAS être rapportés à l'origine.

4. Considérations pour la sécurité

Les questions de sécurité sont brièvement évoquées dans le présent mémoire. L'utilisation du tunnelage peut pallier certaines anciennes options IP de sécurité (étiquetage) mais va permettre une meilleure prise en charge des nouveaux en-têtes de sécurité IP.

Références

[IDM91a] Ioannidis, J., Duchamp, D., Maguire, G., "IP-based protocols for mobile internetworking", Proceedings of SIGCOMM '91, ACM, septembre 1991.

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

[RFC0792] J. Postel, "Protocole du message de contrôle Internet – Spécification du protocole du programme Internet DARPA", STD 5, septembre 1981.

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

[RFC1241] W. Woodburn et D. Mills, "Schéma pour un protocole d'encapsulation Internet v.1", juillet 1991. *(Exp)*

[RFC1435] S. Knowles, "Avis de l'IESG d'une expérience avec découverte de la MTU de chemin", mars 1993. *(Info)*

[RFC1700] J. Reynolds et J. Postel, "Numéros alloués", STD 2, octobre 1994. *(Historique)*

[RFC1701] S. Hanks, T. Li, D. Farinacci et P. Traina, "Encapsulation générique d'acheminement (GRE)", octobre 1994. *(Information)*

[swIPe] Ioannidis, J., and Blaze, M., "The Architecture and Implementation of Network-Layer Security Under Unix", Fourth Usenix Security Symposium Proceedings, octobre 1993.

Remerciements

Ces précisions sur le tunnelage IP sont en grande partie tirés d'un travail indépendant effectué en 1990 par Phil Karn et le groupe TCP en utilisant KA9Q NOS.

Des remerciements particuliers à John Ioannidis (alors à l'Université Columbia) pour son inspiration et l'expérimentation qui a commencé la dernière session de la mobilité IP et du développement de la sécurité d'IP. Une partie de ce texte a été tirée de [IDM91a] et [swIPe].

L'enchaînement des en-têtes aussi été décrit dans "Simple Internet Protocol", par Steve Deering (Xerox PARC).

L'organisation globale et certains passages sont tirés de la [RFC1241], par David Mills (U Delaware) et Robert Woodburn (SAIC).

Une partie du texte sur l'état conditionnel de tunnel a été tiré de "IP Address Encapsulation (IPAE)", par Robert E. Gilligan, Erik Nordmark, et Bob Hinden (tous de Sun Microsystems).

Adresse de l'auteur

Les questions au sujet du présent mémoire peuvent aussi être envoyées à :

William Allen Simpson
Daydreamer
Computer Systems Consulting Services
1384 Fontaine
Madison Heights, Michigan 48071
USA
mél : Bill.Simpson@um.cc.umich.edu bsimpson@MorningStar.com