

Groupe de travail Réseau

Request for Comments: 5640

Catégorie : Sur la voie de la normalisation

Traduction Claude Brière de L'Isle

C. Filsfils, Cisco Systems

P. Mohapatra, Cisco Systems

C. Pignataro, Cisco Systems

août 2009

Équilibrage de charge pour circuits logiciels maillés

Résumé

Les charges utiles transportées sur un service de circuits logiciels maillés (comme défini par l'échange d'informations BGP d'encapsulation de l'identifiant de famille d'adresse suivante (SAFI, *Subsequent Address Family Identifier*)) portent souvent un certain nombre de flux distincts, identifiables. Il peut, dans certaines circonstances, être souhaitable de distribuer ces flux selon des chemins multiples de coût égal (ECMP, *equal cost multiple paths*) qui existent dans le réseau de commutation de paquets. Actuellement, la charge utile d'un paquet qui entre sur le circuit logiciel peut seulement être interprété par les routeurs d'entrée et de sortie. Donc, la décision d'équilibrage de charge d'un routeur de cœur est seulement fondée sur l'en-tête d'encapsulation, présentant beaucoup moins d'entropie que ce qui est disponible dans la charge utile ou dans l'en-tête encapsulé car l'encapsulation de circuit logiciel agit à la façon d'un tunnelage. Le présent document décrit une méthode pour réaliser une efficacité d'équilibrage de charge comparable dans un réseau portant un service de circuits logiciels maillés sur le protocole de tunnelage de couche deux (L2TPv3, *Layer Two Tunneling Protocol - Version 3*) sur IP ou sur encapsulation d'acheminement générique (GRE, *Generic Routing Encapsulation*) à ce qui serait réalisé sans une telle encapsulation.

Statut du présent mémoire

Le présent document spécifie un protocole de l'Internet sur la voie de la 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.

Notice de droits de reproduction

Copyright (c) 2009 IETF Trust et les personnes identifiées comme auteurs du document. Tous droits réservés.

Le présent document est soumis au BCP 78 et aux dispositions légales de l'IETF Trust qui se rapportent aux documents de l'IETF (<http://trustee.ietf.org/license-info>) en vigueur à la date de publication de ce document. Prière de revoir ces documents avec attention, car ils décrivent vos droits et obligations par rapport à ce document.

Table des matières

1. Introduction.....	1
1.1 Langage des exigences.....	2
2. Sous TLV Bloc d'équilibrage de charge.....	2
2.1 Applicabilité aux types de tunnel.....	2
2.2 Considérations d'encapsulation.....	3
3. Considérations relatives à l'IANA.....	3
4. Considérations sur la sécurité.....	3
5. Remerciements.....	3
6. Références normatives.....	3
Adresse des auteurs.....	4

1. Introduction

Considérons le cas d'un routeur R1 qui encapsule un paquet P dans un circuit logiciel lié à un routeur R3. R2 est un routeur sur le plus court chemin de R1 à R3. Le plus court chemin de R2 à R3 implique plusieurs chemins de coût égal (ECMP, *Equal Cost Multiple Paths*). Le but pour R2 est d'être capable de choisir quel chemin utiliser sur la base de l'entropie complète du paquet P.

Ceci est réalisé en portant dans l'en-tête d'encapsulation une signature de l'en-tête interne, améliorant donc l'entropie des flux comme vue par les routeurs de cœur. La signature est portée au titre d'un des champs de l'en-tête d'encapsulation. Pour

aider à une meilleure description dans le document, on définit le terme générique de "champ d'équilibrage de charge" comme signifiant une valeur spécifique d'un type d'encapsulation. Par exemple, pour l'encapsulation L2TPv3 sur IP [RFC3931], le champ d'équilibrage de charge est l'identifiant de session. Pour l'encapsulation GRE [RFC2784], le champ Clé [RFC2890], si il est présent, représente le champ d'équilibrage de charge. Ce mécanisme suppose que les routeurs de cœur fondent leurs décisions d'équilibrage de charge sur une définition de flux qui inclut le champ d'équilibrage de charge. C'est une fonction évidente et générique comme, par exemple, pour les tunnels L2TPv3 sur IP, l'identifiant de session est au même décalage constant bien connu que les accès TCP/UDP dans l'en-tête encapsulant.

L'identifiant de famille d'adresse d'encapsulation suivante [RFC5512] est étendu de façon qu'un bloc contigu du champ d'équilibrage de charge soit lié au circuit logiciel annoncé par un prochain bond BGP. Sur la base du flux interne, le côté fournisseur (PE, *Provider Edge*) d'entrée choisit une valeur du champ d'équilibrage de charge provenant du bloc pour préserver l'ordre par flux et, en même temps, pour améliorer l'entropie à travers les flux.

1.1 Langage 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. Sous TLV Bloc d'équilibrage de charge

Le présent document définit un nouveau sous TLV à utiliser avec l'attribut Encapsulation de tunnel défini dans la [RFC5512]. Le nouveau sous TLV est appelé le "sous TLV Bloc d'équilibrage de charge" et PEUT être inclus dans tout message UPDATE de SAFI d'encapsulation où l'équilibrage de charge est désiré.

Le type du sous TLV Bloc d'équilibrage de charge est 5. La longueur du sous TLV est 2 octets. La valeur représente la longueur du bloc en bits et NE DOIT PAS excéder la taille du champ Équilibrage de charge. Ce format est très similaire à celui du gabarit de sous réseau de longueur variable (VLSM, *Variable-Length Subnet Masking*) utilisé dans les adresses IP pour permettre des préfixes de longueur arbitraire. Le bloc est déterminé en extrayant la séquence initiale de bits de "taille de bloc" du champ Équilibrage de charge.

Si un champ Équilibrage de charge n'est pas signalé (par exemple, si le sous TLV Encapsulation n'est pas inclus dans une annonce comme dans le cas de GRE sans une Clé) alors le sous TLV Bloc d'équilibrage de charge NE DOIT PAS être inclus.

Plus la valeur du champ de sous TLV Bloc d'équilibrage de charge est petite, plus grand est l'espace pour l'identification par flux, et donc meilleure l'entropie pour l'équilibrage de charge potentiel dans le cœur, et également plus petite sera la polarisation quand la transposition s'écoule sur les chemins à ECMP. Cependant, réduire la taille de bloc d'équilibrage de charge consomme plus d'identifiants de session L2TPv3 ou de clés GRE, résultant en potentiellement moins de services pris en charge. Un déploiement typique exigerait un arbitrage entre ces compromis.

Par exemple, supposons qu'il y ait un circuit logiciel entre R1 et R3 avec un type de tunnel L2TPv3 sur IP. Supposons que R3 code l'identifiant de session avec la valeur 0x1234ABCD dans le sous TLV Encapsulation. Cela inclut aussi le sous TLV Bloc d'équilibrage de charge et code la valeur 24. Cela devrait être interprété comme suit :

- o Si un routeur d'entrée ne comprend pas le sous TLV Bloc d'équilibrage de charge, il continue d'utiliser l'identifiant de session 0x1234ABCD et encapsule tous les paquets avec cet identifiant de session.
- o Si un routeur d'entrée comprend le sous TLV Bloc d'équilibrage de charge, il prend les 24 premiers bits de l'identifiant de session (0x1234AB) à utiliser comme bloc et remplit les 8 bits de moindre poids avec un identifiant par flux (par exemple, il peut être déterminé sur la base de la source du paquet interne, des adresses de destination, et des accès TCP/UDP). Ce choix préserve l'ordre par flux des paquets.

Cette exigence et cette solution s'appliquent également à GRE où la Clé joue le même rôle que l'identifiant de session dans L2TPv3.

Inutile de dire que si un routeur de sortie ne prend pas en charge le sous TLV Bloc d'équilibrage de charge, le circuit logiciel continue de fonctionner avec un seul champ d'équilibrage de charge avec lequel tous les routeurs d'entrée

s'encapsulent.

2.1 Applicabilité aux types de tunnel

Le sous TLV Bloc d'équilibrage de charge est applicable aux types de tunnel qui définissent un champ d'équilibrage de charge. Le présent document définit les champ d'équilibrage de charge pour les types de tunnel 1 (L2TPv3 sur IP) et 2 (GRE) comme suit :

- o L2TPv3 sur IP - identifiant de session. Une attention particulière doit être apportée à toujours créer un identifiant de session non zéro. Quand un routeur de sortie inclut un sous TLV Bloc d'équilibrage de charge, il DOIT coder le champ Identifiant de session du sous TLV Encapsulation d'une façon qui assure que les bits de poids fort de l'identifiant de session, après extraction du bloc, sont différents de zéro.
- o GRE - Clé GRE.

Le présent document ne définit pas de champ d'équilibrage de charge pour le type de tunnel IP dans IP (types de tunnel 7). De futurs types de tunnel qui désireraient utiliser le sous TLV Bloc d'équilibrage de charge DOIVENT définir un champ Équilibrage de charge qui fait partie de l'en-tête encapsulant.

2.2 Considérations d'encapsulation

Les champs inclus dans l'en-tête d'encapsulation à côté du champ Équilibrage de charge ne sont pas affectés par le sous TLV Bloc d'équilibrage de charge. Tous les autres champs d'encapsulation sont partagés entre les variations du champ Équilibrage de charge. Par exemple, pour le type de tunnel L2TPv3 sur IP, si le mouchard facultatif est inclus dans le sous TLV encapsulation par le routeur de sortie durant la signalisation de circuit logiciel, il s'applique à toutes les valeurs de "identifiant de session" déduites au routeur d'entrée après l'application du bloc d'équilibrage de charge comme décrit dans le présent document.

3. Considérations relatives à l'IANA

L'IANA a alloué la valeur 5 pour le sous TLV Bloc d'équilibrage de charge, dans le registre des sous TLV Attributs d'encapsulation de tunnel BGP (espace de numéros créé au titre de la publication de la [RFC5512]):

Nom de sous TLV	Valeur
Bloc d'équilibrage de charge	5

4. Considérations sur la sécurité

Le présent document définit un nouveau sous TLV pour l'attribut Encapsulation de tunnel BGP. Les considérations sur la sécurité pour la SAFI Encapsulation BGP et l'attribut Encapsulation de tunnel BGP sont couvertes dans la [RFC5512]. Il n'y a pas de risque supplémentaire pour la sécurité introduit par ce concept.

5. Remerciements

Les auteurs tiennent à remercier Stewart Bryant, Mark Townsley, Rajiv Asati, Kireeti Kompella, et Robert Raszuk de leur relecture et de leurs commentaires.

6. Références normatives

[RFC2119] S. Bradner, "[Mots clés à utiliser](#) dans les RFC pour indiquer les niveaux d'exigence", BCP 14, mars 1997. DOI 10.17487/RFC2119, (MàJ par [RFC8174](#))

- [RFC2784] D. Farinacci, T. Li, S. Hanks, D. Meyer et P. Traina, "[Encapsulation d'acheminement générique \(GRE\)](#)", DOI 10.17487/RFC2784, mars 2000.
- [RFC2890] G. Dommety, "[Extensions de clé et de numéro de séquence](#) à GRE", DOI 10.17487/RFC2890, septembre 2000. (*P.S.*)
- [RFC3931] J. Lau et autres, "[Protocole de tunnelage de couche deux](#) - version 3 (L2TPv3)", DOI 10.17487/RFC3931, mars 2005. (*P.S.*)
- [RFC5512] P. Mohapatra, E. Rosen, "[Identifiant de famille d'adresse suivante](#) (SAFI) d'encapsulation BGP et attribut d'encapsulation de tunnel de BGP", DOI 10.17487/RFC5512, avril 2009. (*P. S.*)

Adresse des auteurs

Clarence Filsfil
Cisco Systems
Brussels,
Belgium
mél : cfilsfil@cisco.com

Pradosh Mohapatra
Cisco Systems
170 W. Tasman Drive
San Jose, CA 95134
USA
mél : pmohapat@cisco.com

Carlos Pignataro
Cisco Systems
7200 Kit Creek Road, PO Box 14987
Research Triangle Park, NC 27709
USA
mél : cpignata@cisco.com