Équipe d'ingénierie de l'Internet (IETF)
Request for Comments : 8201
STD : 87
RFC rendue obsolète : 1981
Catégorie : Norme
ISSN: 2070-1721

J. McCann, Digital Equipment Corporation
R. Hinden, Check Point Software
S. Deering, retraité
J. Mogul, Digital Equipment Corporation
juillet 2017
Traduction Claude Brière de L'Isle

Découverte de la MTU de chemin pour IP version 6

Résumé
Le présent document décrit la découverte de l'unité maximum de transmission de chemin (PMTUD, *Path MTU Dicovery*) pour IP version 6. Il est largement dérivé de la RFC 1191, qui décrit la découverte de la MTU de chemin pour IP version 4. Il rend obsolète la RFC 1981.

Statut du présent mémoire
Ce document est sur la voie de la normalisation.

Le présent document a été produit par l'équipe d'ingénierie de l'Internet (IETF, *Internet Engineering Task Force*). Il représente le consensus de la communauté de l'IETF. Il a subi une révision publique et sa publication a été approuvée par le groupe de pilotage de l'ingénierie de l'Internet (IESG, *Internet Engineering Steering Group*). D'autres informations sur les normes de l'Internet sont disponibles à la Section 2 de la RFC 7841.

Des informations sur le statut actuel de ce document, les errata éventuels, et comment y contribuer peuvent être obtenues à http://www.rfc-editor.org/info/rfc8201.

Table des matières

1. Introduction

Quand un nœud IPv6 a une grande quantité de données à envoyer à un autre nœud, les données sont transmises dans une série de paquets IPv6. Ces paquets peuvent avoir une taille inférieure ou égale à la MTU de chemin (PMTU). Autrement, il peut y avoir de plus gros paquets qui sont fragmentés en une série de fragments dont chacun a une taille inférieure ou égale à la PMTU.

Il est généralement préférable que ces paquets soient de la plus grande taille qui peut avec succès traverser le chemin du nœud de source au nœud de destination sans qu'il soit besoin de la fragmentation IPv6. Cette taille de paquet est appelée la MTU de chemin, et elle est égale à la MTU minimum de liaison de toutes les liaisons dans un chemin. Le présent document définit un mécanisme standard pour qu'un nœud découvre la PMTU d'un chemin arbitraire.

Les nœuds IPv6 devraient mettre en œuvre la découverte de la MTU de chemin afin de découvrir et tirer parti des chemins avec une PMTU supérieure à la MTU minimum de liaison IPv6 [RFC8200]. Une mise en œuvre IPv6 minimale (par exemple, dans une ROM d'amorçage) peut choisir d'omettre la mise en œuvre de la découverte de la MTU de chemin.

Les nœuds qui ne mettent pas en œuvre la découverte de la MTU de chemin doivent utiliser la MTU minimum de liaison IPv6 définie dans la [RFC8200] comme taille maximum de paquet. Dans la plupart des cas, il va en résulter une utilisation de plus petits paquets que nécessaire, parce que la plupart des chemins ont une PMTU supérieure à la MTU minimum de liaison IPv6. Un nœud qui envoie des paquets beaucoup plus petits que ce que la MTU de chemin permet gaspille les ressources du réseau et probablement obtient un débit sous optimal.

Les nœuds qui mettent en œuvre la découverte de la MTU de chemin et envoient des paquets plus grands que la MTU minimum de liaison IPv6 sont susceptibles de problèmes de connexité si des messages ICMPv6 [RFC4443] sont bloqués ou non transmis. Par exemple, il va en résulter que des connexions qui achèvent correctement la prise de contact TCP à trois étapes vont échouer à transférer des données. Cet état est appelé un trou noir de connexion [RFC2923]. La découverte de la MTU de chemin s'appuie sur le message ICMPv6 "paquet trop gros" (PIB, *Packet Too Big*) pour déterminer la MTU du chemin.

Une extension à la découverte de la MTU de chemin définie dans ce document se trouve dans la [RFC4821]. La RFC4821 définit une méthode pour la découverte de la MTU de chemin de couche de mise en paquet (PLPMTUD, *Packetization Layer Path MTU Discovery*) conçue pour être utilisée sur les chemins où la livraison des messages ICMPv6 à un hôte n'est pas assurée.

Note : Le présent document est une mise à jour de la [RFC1981] qui a été publiée avant celle de la [RFC2119]. Par conséquent, bien que la RFC 1981 utilise le langage de style "devrait/doit" en majuscules et minuscules, le présent document ne cite pas les définitions de la RFC 2119 et utilise seulement les minuscules pour ces mots.

2. Terminologie

nœud : appareil qui met en œuvre IPv6.

routeur : nœud qui transmet des paquets IPv6 non explicitement adressés à lui-même.

hôte : tout nœud qui n'est pas un routeur. (Voir la note ci-dessous.)

couche supérieure : couche de protocole immédiatement au dessus de IPv6. Des exemples sont des protocoles de transport comme TCP et UDP, des protocoles de contrôle comme ICMP, des protocoles d'acheminement comme OSPF, et des protocoles de couche internet ou en dessous qui sont "tunnelés" sur (c'est-à-dire, encapsulés dans) IPv6 comme l'échange

de paquets inter réseaux (IPX, *Internetwork Packet Exchange*), AppleTalk, ou IPv6 lui-même.

liaison : facilité ou support de communication sur lequel les nœuds peuvent communiquer à la couche de liaison, c'est-à-dire, à la couche immédiatement en-dessous de IPv6. Des exemples sont les Ethernets (simples ou pontés) les liaisons PPP, X.25, le relais de trame, ou les réseaux ATM, et les "tunnels" de couche internet ou supérieure, comme les tunnels sur IPv4 ou IPv6 lui-même.

interface : rattachement d'un nœud à une liaison.

adresse : identifiant de couche IPv6 pour une interface ou un ensemble d'interfaces.

paquet : un en-tête IPv6 plus une charge utile.

MTU de liaison : unité maximum de transmission, c'est-à-dire, taille maximum de paquet en octets, qui peut être convoyée sur une liaison.

MTU de chemin : MTU minimum de liaison de toutes les liaisons dans un chemin entre un nœud de source et un nœud de destination.

paquet : en-tête IPv6 plus charge utile. Le paquet peut avoir une taille inférieure ou égale à la PMTU. Autrement, ce peut être un paquet plus grand qui est fragmenté en une série de fragments dont chacun aune taille inférieure ou égale à la PMTU.

MTU de liaison : unité maximum de transmission, c'est-à-dire, la taille de paquet maximum en octets, qui peut être convoyée en une seule fois sur une liaison.

chemin : ensemble des liaisons traversées par un paquet entre un nœud de source et un nœud de destination.

MTU de chemin : MTU minimum de liaison de toutes les liaisons dans un chemin entre un nœud de source et un nœud de destination.

PMTU : MTU de chemin.

découverte de la MTU de chemin : processus par lequel un nœud apprend la PMTU d'un chemin.

EMTU_S : MTU effective d'envoi : utilisée par les protocoles de couche supérieure pour limiter la taille des paquets IP qu'ils mettent en file d'attente pour l'envoi [RFC6691], [RFC1122].

EMTU_R : MTU effective de réception : le plus grand paquet qui peut être réassemblé chez le receveur [RFC1122].

flux : séquence de paquets envoyés d'une source particulière à une destination particulière (envoi individuel ou en diffusion groupée) pour laquelle la source désire un traitement spécial de la part des routeurs intermédiaires.

identifiant de flux : combinaison d'une adresse de source et d'une étiquette de flux non zéro.

3. Vue d'ensemble du protocole

Le présent mémoire décrit une technique pour la découverte dynamique de la PMTU d'un chemin. L'idée de base est qu'un nœud source suppose initialement que la PMTU d'un chemin est la MTU (connue) du premier bond du chemin. Si un des paquets envoyés sur ce chemin est trop gros pour être transmis à un nœud le long du chemin, ce nœud va l'éliminer et retourner un message ICMPv6 "Paquet trop gros". À réception d'un tel message, le nœud source réduit sa PMTU supposée pour le chemin sur la base de la MTU du bond qui fait l'étranglement comme rapporté dans le message "Paquet trop gros". La PMTU diminuée fait que la source va envoyer de plus petits paquets ou changer EMTU_S pour faire que la couche supérieure réduise la taille des paquets IP qu'elle envoie.

Le processus de découverte de la MTU de chemin se termine quand l'estimation de la PMTU du nœud de source est inférieure ou égale à la PMTU réelle. Noter que plusieurs itérations de cycles de paquet envoyé/message "Paquet trop gros" reçus peuvent se produire avant que le processus de découverte de la MTU de chemin se termine, car il peut y avoir des liaisons avec de plus petites MTU plus loin le long du chemin.

Autrement, le nœud peut choisir de terminer le processus de découverte en cessant d'envoyer des paquets plus grands que la MTU minimum de liaison IPv6.

La PMTU d'un chemin peut changer avec le temps, à cause de changements dans la topologie d'acheminement. Les réductions de la PMTU sont détectées par les messages "Paquet trop gros". Pour détecter les augmentations de la PMTU d'un chemin, un nœud augmente périodiquement sa PMTU supposée. Il en résulte presque toujours l'élimination de paquets et la génération de messages "Paquet trop gros", parce que dans la plupart des cas, la PMTU du chemin n'aura pas changé. Donc, les tentatives de détection d'augmentation de la MTU d'un chemin devraient être faites peu fréquemment.

La découverte de la MTU de chemin accepte les destinations en diffusion groupée aussi bien qu'en envoi individuel. Dans le cas d'une destination de diffusion groupée, des copies d'un paquet peuvent traverser de nombreux chemins différents vers de nombreux nœuds différents. Chaque chemin peut avoir une PMTU différente, et un seul paquet en diffusion groupée peut résulter en multiples messages "Paquet trop gros", chacun rapportant une MTU de prochain bond différente. La valeur de PMTU minimum à travers l'ensemble des chemins utilisés détermine la taille des paquets envoyés ensuite à la destination de diffusion groupée.

Noter que la découverte de la MTU de chemin doit être effectuée même dans les cas où un nœud "pense" qu'une destination est rattachée à la même liaison que lui, car il se peut qu'elle ait une PMTU inférieure à la MTU de la liaison. Dans une situation où un routeur du voisinage agit comme mandataire [RFC4861] pour une certaine destination, la destination peut paraître être directement connectée, mais elle est en fait plus d'un bond plus loin.

4. Exigences du protocole

Comme expliqué à la Section 1, les nœuds IPv6 ne sont pas obligés de mettre en œuvre la découverte de la MTU de chemin. Les exigences de cette section ne s'appliquent qu'aux mises en œuvre qui incluent la découverte de la MTU de chemin.

Les nœuds devraient valider de façon appropriée la charge utile des messages PTB ICMPv6 pour s'assurer qu'ils sont reçus en réponse au trafic transmis (c'est-à-dire, une condition d'erreur rapportée qui correspond à un paquet IPv6 envoyé en fait par l'application) conformément à la [RFC4443].

Si un nœud reçoit a message "Paquet trop gros" rapportant une MTU de prochain bond qui est inférieure à la MTU minimum de liaison, il doit l'éliminer. Un nœud ne doit pas réduire son estimation de la MTU de chemin en dessous de la MTU minimum de liaison IPv6 à réception d'un message "Paquet trop gros".

Quand un nœud reçoit un message "Paquet trop gros", il doit réduire son estimation de la PMTU pour le chemin pertinent, sur la base de la valeur du champ MTU dans le message. Le comportement précis d'un nœud dans ces circonstances n'est pas spécifié, car des applications différentes peuvent avoir des exigences différentes, et des architectures de mise en œuvre différentes peuvent favoriser des stratégies différentes.

Après avoir reçu un message "Paquet trop gros", un nœud doit tenter d'éviter de provoquer de tels messages dans le proche avenir. Le nœud doit réduire la taille des paquets qu'il envoie le long du chemin. Utiliser une estimation de PMTU plus grande que la MTU minimum de liaison IPv6 peut continuer de provoquer des messages "Paquet trop gros". Parce que chacun de ces messages (et les paquets abandonnés auxquels ils répondent) consomme des ressources du réseau, les nœuds qui utilisent la découverte de la MTU de chemin doivent détecter les diminutions de la PMTU aussi rapidement que possible.

Les nœuds peuvent détecter les augmentations de PMTU, mais comme le faire exige d'envoyer des paquets plus gros que la PMTU estimée actuelle, et parce que la probabilité est que la PMTU n'aura pas augmenté, ceci doit n'être fait qu'à des intervalles non fréquents. Une tentative de détecter une augmentation (par l'envoi d'un paquet plus gros que l'estimation actuelle) ne doit pas être faite moins de 5 minutes après la réception d'un message "Paquet trop gros" pour le chemin concerné. Le réglage recommandé pour ce temporisateur est deux fois sa valeur minimum (10 minutes).

Un nœud ne doit pas augmenter son estimation de la MTU de chemin en réponse au contenu d'un message "Paquet trop gros". Un message prétendant annoncer une augmentation de la MTU de chemin pourrait être un paquet périmé qui a traîné dans le réseau, un paquet falsifié injecté au titre d'une attaque de déni de service, ou le résultat de l'existence de plusieurs chemins pour la destination, chacun avec une PMTU différente.

5. Questions de mise en œuvre

Cette section discute certains problèmes relatifs à la mise en œuvre de la découverte de la MTU de chemin. Ceci n'est pas une spécification, mais plutôt un ensemble de notes pour aider à la mise en œuvre.

Les questions incluent :
- Quelles couches mettent en œuvre la découverte de la MTU de chemin ?
- Comment sont mises en antémémoire les informations de PMTU ?
- Comment sont supprimées les informations de PMTU périmées ?
- Que doivent faire les couches de transport et supérieures ?

5.1 Mise en couches

Dans l'architecture IP, le choix de la taille du paquet à envoyer est fait par un protocole à une couche au dessus de IP. Le présent mémoire se réfère à un tel protocole comme "protocole de mise en paquets". Les protocoles de mise en paquets sont généralement des protocoles de transport (par exemple, TCP) mais peuvent aussi être des protocoles de couche supérieure (par exemple, des protocoles construits par dessus UDP).

La mise en œuvre de la découverte de la MTU de chemin dans les couches de mise en paquet simplifie certains des problèmes inter couches mais a plusieurs inconvénients : la mise en œuvre peut devoir être refaite pour chacun des protocoles de mise en paquet, il devient difficile de partager les informations de PMTU entre les différentes couches de mise en paquet, et l'état en mode connexion conservé par certaines couches de mise en paquets peut ne pas s'étendre facilement pour sauvegarder les informations de PMTU pour de longues périodes.

Il est donc suggéré que la couche IP mémorise les informations de PMTU et que la couche ICMPv6 traite les messages "Paquet trop gros" reçus. Les couches de mise en paquet peuvent répondre aux changements de la PMTU en changeant la taille des messages qu'elles envoient. Pour prendre en charge cette mise en couches, les couches de mise en paquets exigent un moyen pour apprendre les changements de valeur de MMS_S, la"taille maximum de message transport envoyée" [RFC1122].

MMS_S est un message de taille de transport calculé en soustrayant la taille de l'en-tête IPv6 (incluant les en-têtes d'extension IPv6) du plus gros paquet IP qui peut être envoyé, EMTU_S. MMS_S est limité par une combinaison de facteurs, incluant la PMTU, la prise en charge de la fragmentation et du réassemblage de paquet, et la limite de réassemblage de paquet (voir le paragraphe 4.5 "En-tête de fragment", de la [RFC8200]). Quand la fragmentation de source est disponible, EMTU_S est réglé à EMTU_R, comme indiqué par le receveur en utilisant un protocole couche supérieure ou sur la base des exigences du protocole (1500 octets pour IPv6). Quand un message plus grand que la PMTU est à transmettre, la source crée des fragments, chacun limité par la PMTU. Quand la fragmentation de source n'est pas désirée, EMTU_S est réglé à PMTU, et le protocole de couche supérieure est supposé effectuer sa propre fragmentation et réassemblage ou autrement limiter la taille de ses messages en conséquence.

Cependant, les couches de mise en paquets sont encouragées à éviter d'envoyer des messages qui vont exiger la fragmentation de source (pour les cas contre la fragmentation, voir [FRAG]).

5.2 Mémorisation des informations de PMTU

Idéalement, une valeur de PMTU devrait être associée à un chemin spécifique traversé par les paquets échangés entre les nœuds de source et de destination. Cependant, dans la plupart des cas, un nœud n'aura pas assez d'informations pour identifier complètement et précisément un tel chemin. Un nœud doit plutôt associer une valeur de PMTU à une représentation locale d'un chemin. Le choix de la représentation locale d'un chemin appartient à la mise en œuvre. Pour les nœuds avec plusieurs interfaces, les informations de MTU de chemin devraient être conservées pour chaque liaison IPv6.

Dans le cas d'une adresse de destination de diffusion groupée, des copies d'un paquet peuvent traverser de nombreux chemins différents pour atteindre de nombreux nœuds différents. La représentation locale du "chemin" pour une destination de diffusion groupée doit représenter un ensemble potentiellement grand de chemins.

Au minimum, une mise en œuvre pourrait conserver une seule valeur de PMTU à utiliser pour tous les paquets générés à partir du nœud. Cette valeur de PMTU serait la PMTU minimum apprise à travers l'ensemble de tous les chemins utilisés par le nœud. Cette approche va probablement résulter en l'utilisation de plus petits paquets que nécessaire pour beaucoup des chemins. Dans le cas d'acheminement multi chemins (par exemple, acheminement multi chemin de coût égal (ECMP,

Equal-Cost Multipath Routing)) un ensemble des chemins peut exister même pour une seule paire de source et destination.

Une mise en œuvre pourrait utiliser l'adresse de destination comme représentation locale d'un chemin. La valeur de PMTU associée à une destination serait la PMTU minimum apprise sur l'ensemble de tous les chemins utilisés pour cette destination. Cette approche va résulter en l'utilisation de paquets de taille optimale sur la base de la destination. Cette approche s'intègre bien avec le modèle conceptuel d'un hôte comme décrit dans la [RFC4861] : une valeur de PMTU pourrait être mémorisée avec l'entrée correspondante dans l'antémémoire de destination.

Si des flux [RFC8200] sont utilisés, une mise en œuvre pourrait utiliser les identifiants de flux comme représentation locale d'un chemin. Les paquets envoyés à une destination particulière mais appartenant à des flux différents peuvent utiliser des chemins différents, comme avec ECMP, dans lequel le choix des chemins peut dépendre de l'identifiant de flux. Cette approche peut résulter en l'utilisation de paquets de taille optimale sur la base du flux, donnant une plus fine granularité que les valeurs de PMTU tenues sur la base de la destination.

Pour les paquets en acheminement de source (c'est-à-dire les paquets contenant un en-tête Acheminement IPv6 [RFC8200]) la route de source peut qualifier la représentation locale d'un chemin.

Initialement, la valeur de PMTU pour un chemin est supposée être la MTU (connue) de la liaison de premier bond.

Quand un message "Paquet trop gros" est reçu, le nœud détermine à quel chemin le message s'applique sur la base du contenu du message "Paquet trop gros". Par exemple, si l'adresse de destination est utilisée comme représentation locale d'un chemin, l'adresse de destination provenant du paquet d'origine va être utilisée pour déterminer à quel chemin le message s'applique.

Note : si le paquet d'origine contenait un en-tête Acheminement, il devrait être utilisé pour déterminer la localisation de l'adresse de destination au sein du paquet d'origine. Si Segments restants est égal à zéro, l'adresse de destination est dans le champ Adresse de destination dans l'en-tête IPv6. Si Segments restants est supérieur à zéro, l'adresse de destination est la dernière adresse (Adresse[n]) dans l'en-tête Acheminement.

Le nœud utilise alors la valeur dans le champ MTU dans le message "Paquet trop gros" comme valeur d'essai de PMTU ou la MTU minimum de liaison IPv6 si elle est plus grande, et compare la PMTU d'essai à la PMTU existante. Si la PMTU d'essai est inférieure à l'estimation de MPTU existante, la PMTU d'essai remplace la PMTU existante comme valeur de PMTU pour le chemin.

Les couches de mise en paquet doivent être notifiées des diminutions de la PMTU. Toute instance de couche de mise en paquets (par exemple, une connexion TCP) qui utilise activement le chemin doit être notifiée si l'estimation de PMTU est diminuée.

Note : même si le message "Paquet trop gros" contient un en-tête du paquet d'origine qui se réfère à un paquet UDP, la couche TCP doit être notifiée si une de ses connexions utilise le chemin en question.

Aussi, l'instance qui envoie le paquet qui a provoqué le message "Paquet trop gros" devrait être notifiée que son paquet a été éliminé, même si l'estimation de PMTU n'a pas changé, afin qu'elle puisse retransmettre les données éliminées.

Note : une mise en œuvre peut éviter d'utiliser un mécanisme de notification asynchrone pour les diminutions de la PMTU en retardant la notification jusqu'à la prochaine tentative d'envoi d'un paquet plus grand que l'estimation de PMTU. Dans cette approche, quand une tentative est faite d'envoyer un paquet plus grand que l'estimation de PMTU, la fonction SEND devrait échouer et retourner une indication d'erreur convenable. Cette approche peut être plus convenable à une couche de mise en paquet sans connexion (comme celles qui utilisent UDP) qui (dans certaines mises en œuvre) peuvent être difficiles à "notifier" à partir de la couche ICMPv6. Dans ce cas, les mécanismes normaux fondés sur la fin de temporisation vont être utilisés pour récupérer de la perte des paquets éliminés.

Il est important de comprendre que la notification des instances de couche de mise en paquets en utilisant le chemin sur le changement de la PMTU est distincte de la notification d'une instance spécifique qu'un paquet a été éliminé. Cette dernière devrait être faite aussitôt que possible (c'est-à-dire, en asynchrone du point de vue de l'instance de couche de mise en paquets) tandis que la première peut être retardée jusqu'à ce que l'instance de couche de mise en paquets veuille créer un paquet.

5.3 Purge des informations de PMTU périmées

La topologie de l'inter réseaux est dynamique ; les routes changent avec le temps. Alors que la représentation locale d'un chemin peut rester constante, les chemins réels utilisés peuvent changer. Donc, les informations de PMTU mises en antémémoire par un nœud peut être périmées.

Si la valeur de PMTU périmée est trop grande, cela va être découvert presque immédiatement une fois qu'un paquet assez gros est envoyé sur le chemin. Aucun mécanisme n'existe pour réaliser qu'une valeur de PMTU périmée est trop petite, de sorte qu'une mise en œuvre devrait "faire vieillir" les valeurs en antémémoire. Quand une valeur de PMTU n'a pas été diminuée pendant un certain temps (de l'ordre de 10 minutes) on devrait sonder pour voir si une plus grosse PMTU est supportée.

Note : une mise en œuvre devrait fournir un moyen pour changer la durée de temporisation, incluant de la régler à "infini". Par exemple, les nœuds rattachés à une liaison qui a une grosse MTU qui est ensuite rattachée au reste de l'Internet via une liaison qui a une petite MTU ne vont jamais découvrir une nouvelle PMTU non locale, de sorte qu'ils ne devraient pas commencer à éliminer des paquets toutes les 10 minutes.

5.4 Actions de la couche de mise en paquets

Une couche de mise en paquets (par exemple, TCP) doit utiliser la PMTU pour le ou les chemins utilisés par une connexion ; elle ne devraient pas envoyer de segments qui résulteraient en paquets plus grands que la PMTU, sauf pour sonder durant la découverte de PMTU (ce paquet sonde ne doit pas être fragmenté à la PMTU). Une simple mise en œuvre pourrait demander à la couche IP cette valeur chaque fois qu'elle a créé un nouveau segment, mais ceci pourrait être inefficace. Une mise en œuvre met normalement en antémémoire les autres valeurs déduites de la PMTU. Il peut être plus simple de recevoir des notifications asynchrones quand la PMTU change, de sorte que ces variables puissent aussi être mises à jour.

Une mise en œuvre de TCP doit aussi mémoriser la valeur de taille maximum de segment (MSS, *Maximum Segment Size*) reçue de son homologue, qui représente la EMTU_R, le plus gros paquet qui peut être réassemblé par le receveur, et ne doit pas envoyer de segment plus grand que cette MSS, sans considération de la PMTU.

La valeur envoyée dans l'option TCP MSS est indépendante de la PMTU ; elle est déterminée par la limite de réassemblage du receveur EMTU_R. Cette valeur d'option MSS est utilisée par l'autre extrémité de la connexion, qui peut utiliser une valeur de PMTU sans rapport avec elle. Voir la Section 5, "Questions de taille de paquet", et le paragraphe 8.3, "Taille maximum de charge utile de couche supérieure", de la [RFC8200] pour des informations sur le choix d'une valeur pour l'option TCP MSS.

La réception d'un message "Paquet trop gros" implique qu'un paquet a été éliminé par le nœud qui a envoyé le message ICMPv6. Un protocole fiable de couche supérieure va détecter cette perte par ses propres moyens, et le récupérer par les méthodes normales de retransmission. La retransmission pourrait résulter en retards, selon la méthode de détection de perte utilisée par le protocole de couche supérieure. Si le processus de découverte de la MTU de chemin exige plusieurs étapes pour trouver la PMTU du chemin complet, cela pourrait finalement retarder la retransmission de nombreux temps d'aller-retour.

Autrement, la retransmission pourrait être faite dans une réponse immédiate à une notification que la MTU de chemin a été diminuée, mais seulement pour la connexion spécifiée dans le message "Paquet trop gros". La taille de paquet utilisée dans la retransmission devrait être pas plus que la nouvelle PMTU.

Note : une couche de mise en paquet qui détermine qu'un paquet sonde est perdu a besoin d'adapter la taille de segment de la retransmission. Utiliser la taille rapportée dans le dernier message "Paquet trop gros" peut cependant conduire à d'autres perte car il pourrait y avoir de plus faibles limites de PMTU chez des routeurs plus loin sur le chemin. Cela conduirait à la perte de tous les segments retransmis et causerait donc un encombrement inutile ainsi que l’envoi de paquets supplémentaires chaque fois qu'un nouveau routeur annonce une plus petite MTU. Toute couche de mise en paquet qui utilise la retransmission est donc aussi responsable du contrôle d'encombrement de ses retransmissions [RFC8085].

Une perte causée par une sonde de PMTU indiquée par la réception d'un message "Paquet trop gros" ne doit pas être considérée comme une notification d’encombrement, et donc la fenêtre d'encombrement peut ne pas changer.

5.5 Problèmes pour les autres protocoles de transport

Certains protocoles de transport n'ont pas la permission de remettre en paquet lors d'une retransmission. C'est-à-dire, une fois qu'est faite une tentative de transmission d'un segment d'une certain taille, le transport ne peut pas partager le contenu du segment en plus petits segments pour une retransmission. Dans de tels cas, le segment d'origine peut être fragmenté par la couche IP durant la retransmission. Les segments suivants, lorsque transmis pour la première fois, devraient n'être pas plus grands que ce qui est permis par la MTU de chemin.

La découverte de la MTU de chemin pour IPv4 [RFC1191] utilisait le système de fichier réseau (NFS, *Network File System*) comme exemple d'application fondée sur UDP qui bénéficie de la découverte de la PMTU. Depuis, la [RFC7530] déclare que la couche de transport prise en charge entre NFS et IP doit être un protocole de transport normalisé par l'IETF qui est spécifié comme évitant l'encombrement du réseau ; de tels transports incluent TCP, le protocole de transmission de contrôle de flux (SCTP, *Stream Control Transmission Protocol*) [RFC4960], et le protocole de contrôle d'encombrement de datagrammes (DCCP, *Datagram Congestion Control Protocol*) [RFC4340]. Dans ce cas, le transport est responsable de la vérification que les segments transmis (sauf les sondes) se conforment à la MTU de chemin, incluant de prendre en charge les transmissions des sondes de découverte de la PMTU comme nécessaire.

5.6 Interface de gestion

Il est suggéré qu'une mise en œuvre fournisse un moyen pour qu'un programme d'utilitaire système :
- spécifie que la découverte de la MTU de chemin ne soit pas faite sur un certain chemin,
- change la valeur de la PMTU associée à un certain chemin.

Le premier peut être réalisé en associant un fanion au chemin ; quand un paquet est envoyé sur un chemin avec ce fanion établi, la couche IP n'envoie pas de paquet plus grand que la MTU minimum de liaison IPv6.

Ces caractéristiques pourraient être utilisées pour traiter une situation anormale ou par une mise en œuvre de protocole d'acheminement qui est capable d'obtenir les valeurs de MTU de chemin.

La mise en œuvre devrait aussi fournir un moyen de changer la période de temporisation pour le vieillissement des informations de PMTU jusqu'à ce qu'elles deviennent périmées.

6. Considérations sur la sécurité

Ce mécanisme de découverte de la MTU de chemin rend possible deux attaques de déni de service, toutes deux fondées sur un acteur malveillant qui envoie des messages "Paquet trop gros" falsifiés à un nœud.

Dans la première attaque, le message falsifié indique une PMTU beaucoup plus petite qu'en réalité. En réponse, le nœud victime ne devrait jamais régler son estimation de PMTU plus bas que la MTU minimum de liaison IPv6. Un envoyeur qui réduit sa taille de paquet à cette fausse MTU va observer une chute de ses performances.

Dans la seconde attaque, le faux message indique une PMTU plus grande qu'en réalité. Si il est cru, cela va causer un blocage temporaire lorsque la victime envoie des paquets qui vont être éliminés par un routeur. En un temps d'aller-retour, le nœud va découvrir sa faute (en recevant des messages "Paquet trop gros" de ce routeur) mais la fréquente répétition de cette attaque pourrait causer l'élimination de nombreux paquets. Cependant, un nœud ne doit pas relever son estimation de la PMTU sur la base d'un message "Paquet trop gros", de sorte qu'il ne devrait pas être vulnérable à cette attaque.

Ces deux attaques peuvent causer une connexion en trou noir, c'est-à-dire que la prise de contact TCP en trois phases s'achève correctement mais la connexion reste inactive quand des données sont transférées.

Un acteur malveillant pourrait aussi causer des problèmes si il pouvait empêcher une victime de recevoir des messages "Paquet trop gros" légitimes, mais dans ce cas, il y a des attaques de déni de service plus simples qui sont disponibles.

Si le filtrage ICMPv6 empêche la réception des messages ICMPv6 "Paquet trop gros", la source ne va pas apprendre la MTU de chemin réelle. La [RFC4821] "Découverte de la MTU de chemin de couche de mise en paquet" ne s'appuie pas sur la prise en charge du réseau pour les messages ICMPv6 et est donc considérée comme plus robuste que la PMTUD standard. Elle n'est pas susceptible de connexions "en trou noir" causées par le filtrage des messages ICMPv6. Voir dans la [RFC4890] des recommandations sur le filtrage des messages ICMPv6.

7. Considérations relatives à l'IANA

Le présent document n'exige aucune action de la part de l'IANA.

8. Références

8.1 Références normatives

[RFC4443] A. Conta et autres, "Spécification du protocole de message de contrôle Internet (ICMPv6) pour la version 6 du protocole Internet (IPv6)", mars 2006. STD 89, DOI 10.17487/RFC4443. (*Remplace* RFC2463 ; *MàJ* RFC2780 ; *MàJ par* RFC4884 ; *D.S.)*

[RFC8200] S. Deering, R. Hinden, "Spécification du protocole Internet version 6 (IPv6)", juillet 2017. STD 86, DOI 10.17487/RFC8200. *(Remplace* 2460)

8.2 Références pour information

[FRAG] Kent, C. et J. Mogul, "Fragmentation Considered Harmful", dans Proc. SIGCOMM '87 Workshop on Frontiers in Computer Communications Technology, DOI 10.1145/55483.55524, août 1987.

[RFC1122] R. Braden, "Exigences pour les hôtes Internet – couches de communication", STD 3, octobre 1989, DOI 10.17487/RFC1122. *(MàJ par* RFC6633, 8029)

[RFC1191] J. Mogul et S. Deering, "Découverte de la MTU de chemin", novembre 1990, DOI 10.17487/RFC1191.

[RFC1981] J. McCann, S. Deering, J. Mogul, "Découverte de la MTU de chemin pour IP version 6", août 1996, DOI 10.17487/RFC1981. *(D.S. ; Remplacé par* [RFC8398], STD87)

[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)

[RFC2923] K. Lahey, "Problèmes de TCP avec la découverte de MTU de chemin", septembre 2000, DOI 10.17487/RFC2923. *(Information)*

[RFC4340] E. Kohler et autres, "Protocole de contrôle d'encombrement de datagrammes (DCCP)", mars 2006, DOI 10.17487/RFC4340. *(P.S.) (MàJ par* 6773)

[RFC4821] M. Mathis, J. Heffner, "Découverte de la MTU de chemin de couche de mise en paquet", mars 2007, DOI 10.17487/RFC4821. *(P.S.)*

[RFC4861] T. Narten et autres, "Découverte du voisin pour IP version 6 (IPv6)", septembre 2007, DOI 10.17487/RFC4861. (*Remplace* RFC2461) *(D.S. ; MàJ par* RFC8028, RFC8319, RFC8425)

[RFC4890] E. Davies, J. Mohacsi, "Recommandations pour le filtrage des messages ICMPv6 dans les pare-feu", mai 2007, DOI 10.17487/RFC4890. *(Information)*

[RFC4960] R. Stewart, éd., "Protocole de transmission de commandes de flux (SCTP)", septembre 2007, DOI 10.17487/RFC4960. (*Remplace* RFC2960, RFC3309 ; *P.S. ; voir errata dans* RFC8540 *; MàJ par* RFC8899)

[RFC6691] D. Borman, "Options TCP et taille maximum de segment (MSS)", juillet 2012, DOI 10.17487/RFC6691. (*MàJ les* RFC0879, RFC2385) *(Information)*

[RFC7530] T. Haynes, D. Noveck, "Protocole du système de fichier réseau (NFS) version 4", mars 2015, DOI 10.17487/RFC7530. *(P.S. ; Remplace* 3580, *MàJ par* RFC7931, RFC8587)

[RFC8085] L. Eggert, et autres, "Lignes directrices pour l'utilisation de UDP", mars 2017. BCP 145, DOI

10.17487/RFC8085. *(MàJ* 5405 *; MàJ par* RFC8899)

Appendice A. Comparaison avec la RFC1191

La RFC1981 (rendue obsolète par le présent document) était fondée en grande partie sur la RFC1191, qui décrit la découverte de la MTU de chemin pour IPv4. Certaines portions de la RFC 1191 n'étaient pas nécessaires dans la RFC1981 :

Spécification de routeur : les messages "Paquet trop gros" et le comportement de routeur correspondant sont définis dans la [RFC4443].

Bit "Ne pas fragmenter" : il n'y a pas de bit DF dans les paquets IPv6.

Discussion de la MSS TCP : le choix d'une valeur à envoyer dans l'option MSS de TCP est discuté dans la [RFC8200].

Messages d'ancien style : tous les messages "Paquet trop gros" rapportent la MTU de la liaison qui restreint le débit.

Tableaux de plateau de MTU : ils ne sont pas nécessaire parce que il n'y a pas de messages d'ancien style.

Appendice B. Changements depuis la RFC 1981

Le présent document se fonde sur la RFC 1981 et a introduit les changements suivants par rapport à elle :

- Précision dans la Section 1, "Introduction", que l'objet de la PMTUD est de réduire le besoin de fragmentation IPv6.
- Ajout de texte à la Section 1, "Introduction", sur les effets sur la PMTUD du blocage de messages ICMPv6.
- Ajout d'une "Note" à l'introduction du document que cette spécification ne cite pas la RFC 2119 et utilise seulement le langage "devrait/doit" en minuscules. Changement de tous les "devrait/doit" majuscules en minuscules.
- Ajout d'un bref résumé à la Section 1, "Introduction", sur la PLPMTUD et une référence à la RFC 4821 qui la définit.
- Alignement du texte de la Section 2, "Terminologie", sur la terminologie actuelle de la couche de mise en paquet.
- Ajout d'une précision à la Section 4, "Exigences de protocole", que les nœuds devraient valider la charge utile des messages PIB ICMP selon la RFC 4443, et que les nœuds devraient détecter aussitôt que possible les diminutions de la PMTU.
- Suppression d'une "Note" à la Section 4, "Exigences du protocole", sur un message "Paquet trop gros" rapportant une MTU de prochain bond inférieure à la MTU minimum de liaison IPv6 parce que cela a été supprimé de la [RFC8200].
- Ajout d'une précision au paragraphe 5.2, "Mémorisation des informations de PMTU", pour éliminer un message ICMPv6 "Paquet trop gros" si il contient une MTU inférieure à la MTU minimum de liaison IPv6.
- Ajout d'une précision au paragraphe 5.2, "Mémorisation des informations de PMTU", que pour les nœuds avec plusieurs interfaces, les informations de MTU de chemin devraient être mémorisées pour chaque liaison.
- Suppression de texte au paragraphe 5.2, "Mémorisation des informations de PMTU", sur l'en-tête d'acheminement de type 0 (RH0) parce qu'il est déconseillé par la RFC 5095.
- Suppression de texte sur la classification de sécurité obsolète au paragraphe 5.2, "Mémorisation des informations de PMTU".
- Changement du titre du paragraphe 5.4 en "Actions de la couche de mise en paquet" et changement du texte du premier alinéa pour généraliser la couverture de ce paragraphe à toutes les couches de mise en paquet, en plus de TCP.
- Précision au texte du paragraphe 5.4, "Actions de la couche de mise en paquet", pour utiliser les méthodes normales de retransmission de couche de mise en paquets.
- Suppression de texte au paragraphe 5.4, "Actions de la couche de mise en paquet", qui décrivait BSD 4.2 parce que il est obsolète, et suppression de la référence à TP4.
- Mise à jour du texte du paragraphe 5.5, "Problèmes pour les autres protocoles de transport", sur le NFS, incluant d'ajouter la référence actuelle à NFS et suppression du texte obsolète.
- Ajout d'un paragraphe à la Section 6, "Considérations sur la sécurité", sur les connexions en trou noir si des messages PTB ne sont pas reçus et comparaison à la PLPMTUD.
- Mise à jour des "Remerciements".
- Changements rédactionnels.

Remerciements

Nous tenons à remercier les auteurs et contributeurs de la [RFC1191], de laquelle la majorité de ce document est dérivée. Merci aussi aux membres du groupe de travail IPng de leur relecture attentive et de leurs critiques constructives.
Merci aussi aux contributeurs de cette mise à jour de la "découverte de la MTU de chemin pour IP Version 6". Cela inclut les membres du groupe de travail 6MAN, les réviseurs de la zone de direction, l'IESG, et en particukier Joe Touch et Gorry Fairhurst.

Adresse des auteurs

Robert M. Hinden
Check Point Software
959 Skyway Road
San Carlos, CA 94070
USA
mèl : bob.hinden@gmail.com

Stephen E. Deering
Vancouver, British Columbia
Canada

Jeffrey Mogul
Digital Equipment Corporation

Jack McCann
DEC