Groupe de travail Réseau
Request for Comments : 2080
Catégorie : En cours de normalisation
Traduction Claude Brière de L'Isle

G. Malkin, Xylogics
R. Minnear, Ipsilon Networks
janvier 1997

RIPng pour IPv6

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 "Normes officielles des protocoles de l'Internet" (STD 1) pour connaître l'état de la normalisation et le statut de ce protocole. La distribution du présent mémoire n'est soumise à aucune restriction.

Résumé
Le présent document spécifie un protocole d'acheminement pour un Internet IPv6. Il se fonde sur les protocoles et algorithmes actuellement de large utilisation dans l'Internet IPv4.

Cette spécification représente les changements minimum au protocole des informations d'acheminement (RIP, *Routing Information Protocol*) tel que spécifié dans les RFC1058 [1] et RFC1723 [2], nécessaires pour le fonctionnement en IPv6 [3].

Remerciements
Le présent document est une version modifiée de la RFC1058, écrite par Chuck Hedrick [1]. Les modifications reflètent les améliorations de RIP-2 et de IPv6, mais la formulation d'origine est la sienne.

Nous souhaitons remercier Dennis Ferguson et Thomas Narten pour leurs apports.

Table des matières

1. Introduction

Le présent mémoire décrit un protocole parmi une série de protocoles d'acheminement qui se fondent sur l'algorithme de Bellman-Ford (ou du vecteur de distance). Cet algorithme a été utilisé pour les calculs d'acheminement dans les réseaux informatiques depuis les premiers jours de l'ARPANET. Les formats et protocoles de paquet particuliers décrits ici se fondent sur le programme "routed," qui est inclus dans la distribution Berkeley de Unix.

Dans un réseau international tel que l'Internet, il est très peu vraisemblable qu'un seul protocole d'acheminement soit utilisé pour la totalité du réseau. Le réseau va plutôt être organisé comme une collection de systèmes autonomes (AS, *Autonomous Systems*), dont chacun va, en général, être administré par une seule entité. Chaque AS va avoir sa propre technologie d'acheminement, qui peut varier selon l'AS. Le protocole d'acheminement utilisé au sein d'un AS est appelé un protocole de passerelle intérieure (IGP, *Interior Gateway Protocol*). Un protocole distinct, appelé un protocole de passerelle extérieure (EGP, *Exterior Gateway Protocol*) est utilisé pour transférer les informations d'acheminement entre les AS. RIPng a été

conçu pour fonctionner comme un IGP dans les AS de taille modérée. Il n'est pas destiné à être utilisé dans des environnements plus complexes. Pour des informations sur le contexte auquel on s'attend à ce que convienne RIP version 1 (RIP-1) voir Braden et Postel [6].

RIPng fait partie d'une classe d'algorithmes appelés algorithmes de vecteur de distance. La première description connue de l'auteur de cette classe d'algorithmes se trouve dans Ford et Fulkerson [8]. À cause de cela, ils sont parfois appelés algorithmes de Ford-Fulkerson. Le terme Bellman-Ford est aussi utilisé, et vient du fait que leur formulation se fonde sur l'équation de Bellman [4]. Sa présentation dans le présent document s'appuie directement sur [5]. Le présent document contient la spécification du protocole. Pour une introduction aux mathématiques des algorithmes d'acheminement, voir [1]. Les algorithmes de base utilisés dans ce protocole ont été utilisés dans l'acheminement informatique depuis 1969 dans l'ARPANET. Cependant, l'ancienneté spécifique de ce protocole remonte aux protocoles réseau de Xerox. Les protocoles PUP [7] utilisaient le protocole d'informations de passerelles pour échanger des informations d'acheminement. Une version quelque peu mise à jour de ce protocole a été adoptée pour l'architecture des systèmes réseau de Xerox (XNS, *Xerox Network Systems*) sous le nom de protocole d'informations d'acheminement [9]. L'acheminement Berkeley "routed" est sensiblement le même que le protocole d'informations d'acheminement, les adresses XNS étant remplacées par un format d'adresse plus général capable de traiter IPv4 et les autres types d'adresse, et les mises à jour de l'acheminement étant limitées à une toute les 30 secondes. À cause de cette similitude, le terme de protocole d'informations d'acheminement (ou simplement RIP) est utilisé pour se référer à la fois au protocole XNS et au protocole utilisé par routed.

1.1 Fondements théoriques

Une introduction à la théorie et aux mathématiques qui sous tendent les protocoles de vecteur de distance est fournie dans [1]. Elle n'a pas été incorporée dans le présent document par souci de concision.

1.2 Limitations du protocole

Ce protocole ne résout pas tous les problèmes d'acheminement possibles. Comme on l'a mentionné plus haut, il est principalement destiné à être utilisé comme un IGP dans les réseaux de taille modérée. De plus, les limitations spécifiques doivent être mentionnées :

- Le protocole se limite aux réseaux dont le chemin le plus long (le diamètre du réseau) est de 15 bonds. Les concepteurs pensent que le concept de base du protocole n'est pas approprié pour les plus grands réseaux. Noter que cette déclaration sur les limites suppose qu'un coût de 1 est utilisé pour chaque réseau. C'est la façon dont RIPng est normalement configuré. Si l'administrateur du système choisit d'utiliser des coûts plus élevés, la limite supérieure de 15 peut facilement devenir un problème.

- Le protocole dépend du "compte à l'infini" pour résoudre certaines situations inhabituelles (voir le paragraphe 2.2 dans [1]). Si le système de réseaux a plusieurs centaines de réseaux, et si une boucle d'acheminement s'est formée qui les implique tous, la résolution de la boucle va exiger beaucoup de temps (si la fréquence des mises à jour d'acheminement est limitée) ou de bande passante (si les mises à jour doivent être envoyées chaque fois que des changements sont détectés). Une telle boucle va consommer une grande quantité de bande passante du réseau avant qu'elle ne soit corrigée. Nous pensons que dans des cas réalistes, cela ne sera pas un problème sauf sur des lignes lentes. Même alors, le problème sera très inhabituel, car les diverses précautions qui sont prises vont empêcher cela dans la plupart des cas.

- Ce protocole utilise des "métriques" fixes pour comparer les différents chemins. Il n'est pas approprié pour les situations où les chemins doivent être choisis sur la base de paramètres en temps réel comme le délai mesuré, la fiabilité, ou la charge. Les extensions évidentes pour permettre des métriques de ce type vont vraisemblablement introduire des instabilités que le protocole n'est pas conçu pour traiter.

2. Spécification du protocole

RIPng est destiné à permettre aux routeurs d'échanger des informations pour calculer des chemins à travers un réseau fondé sur IPv6. RIPng est un protocole de vecteur de distance, tel que décrit dans [1]. RIPng ne devrait être mis en œuvre que dans les routeurs ; IPv6 procure d'autres mécanismes de découverte des routeurs [10]. Tout routeur qui utilise RIPng est supposé avoir des interfaces avec un ou plusieurs réseaux, autrement ce n'est pas vraiment un routeur. On les appelle les réseaux directement connectés. Le protocole s'appuie sur l'accès à certaines informations sur chacun de ces réseaux, dont la plus importante est sa métrique. La métrique RIPng d'un réseau est un entier entre 1 et 15, inclus. La façon de la régler n'est pas spécifiée dans ce protocole ; cependant, étant donnée la limite maximum de chemin de 15, une valeur de 1 est utilisée habituellement. Les mises en œuvre devraient permettre à l'administrateur de système de régler la métrique de chaque réseau. En plus de la métrique, chaque réseau aura un préfixe d'adresse de destination IPv6 et une longueur de préfixe qui

lui est associée. La façon dont ils sont réglés par l'administrateur de système n'est pas spécifiée dans le présent protocole.

Chaque routeur qui met en œuvre RIPng est supposé avoir un tableau d'acheminement. Ce tableau a une entrée pour chaque destination qui est accessible par le système qui fait fonctionner RIPng. Chaque entrée contient au moins les informations suivantes :

- le préfixe IPv6 de la destination ;
- une métrique, qui représente le coût total de l'obtention d'un datagramme de la part du routeur pour cette destination. Cette métrique est la somme des coûts associés aux réseaux qui seraient traversés pour arriver à la destination ;
- L'adresse IPv6 du prochain routeur le long du chemin vers la destination (c'est-à-dire, le prochain bond). Si la destination est sur un des réseaux directement connectés, cet élément n'est pas nécessaire ;
- un fanion pour indiquer que les informations sur ce chemin ont changé récemment. On appellera cela le "fanion de changement de chemin" ;
- divers temporisateurs associés au chemin. Voir au paragraphe 2.3 des détails sur les temporisateurs.

Les entrées pour les réseaux directement connectés sont établies par le routeur en utilisant les informations rassemblées par des moyens qui ne sont pas spécifiés dans le présent protocole. La métrique pour un réseau directement connecté est réglée au coût de ce réseau. Comme mentionné précédemment, 1 est le coût usuel. Dans ce cas, la métrique RIPng est réduite à un simple compte de bonds. Des métriques plus complexes peuvent être utilisées quand il est souhaitable de montrer une préférence pour certains réseaux par rapport à d'autres (par exemple, pour indiquer des différences de bande passante ou de fiabilité).

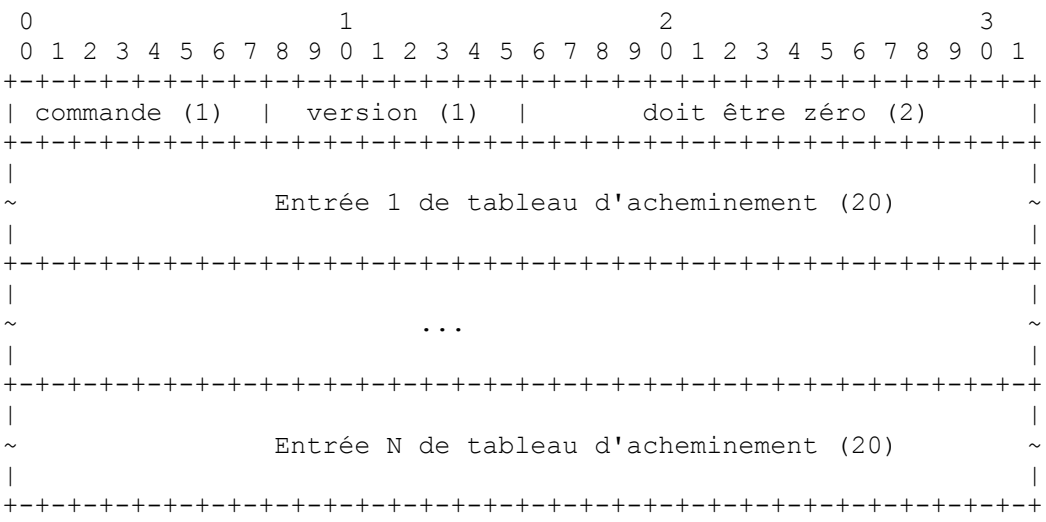
Une mise en œuvre peut aussi choisir de permettre à l'administrateur de système d'entrer des chemins supplémentaires. Ils seront vraisemblablement des chemins vers des hôtes ou réseaux en dehors de la portée du système d'acheminement. On les appelle des chemins "statiques". Les entrées pour les destinations autres que ces destinations initiales sont ajoutées et mises à jour par les algorithmes décrits dans les paragraphes suivants.

Pour que le protocole fournisse des informations complètes sur l'acheminement, chaque routeur de l'AS doit participer au protocole. Dans le cas où plusieurs IGP sont utilisés, il doit y avoir au moins un routeur qui puisse écouler les informations d'acheminement entre les protocoles.

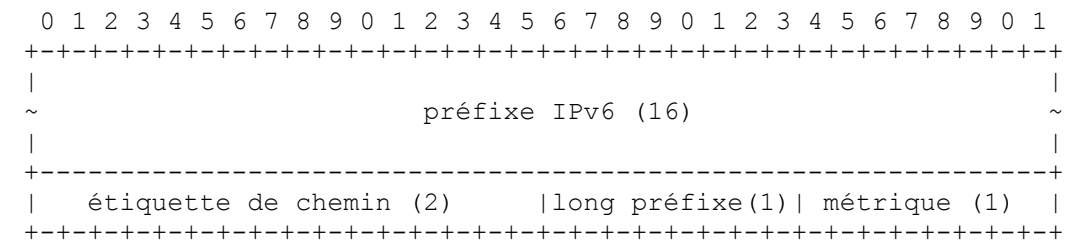
2.1 Format de message

RIPng est un protocole fondé sur UDP. Chaque routeur qui utilise RIPng a un processus d'acheminement qui envoie et reçoit les datagrammes sur le numéro d'accès UDP 521, qui est l'accès RIPng. Toutes les communications destinées au processus RIPng d'un autre routeur sont envoyées à l'accès RIPng. Tous les messages de mise à jour d'acheminement sont envoyés à partir de l'accès RIPng. Les messages de mise à jour non sollicitée d'acheminement ont l'accès aussi bien de source que de destination égaux à l'accès RIPng. Ceux qui sont envoyés en réponse à une demande sont envoyés à l'accès d'où est venue la demande. Des interrogations spécifiques peuvent être envoyées depuis des accès autres que l'accès RIPng, mais ils doivent être dirigés sur l'accès RIPng de la machine cible.

Le format du paquet lRIPng est :



où chaque entrée de tableau d'acheminement (RTE, *Route Table Entry*) a le format suivant :



Le nombre maximum de RTE est défini ci-dessous.

Les tailles de champ sont données en octets. Sauf spécification contraire, les champs contiennent des entiers binaires, dans l'ordre des octets du réseau, avec l'octet de poids forte en premier (gros boutien). Chaque marque de pas représente un bit.

Chaque message contient un en-tête RIPng qui consiste en une commande et un numéro de version. Le présent document décrit la version 1 du protocole (voir au paragraphe 2.4). Le champ commande est utilisé pour spécifier l'objet de ce message. Les commandes mises en œuvre dans la version 1 sont :

1 – demande Une demande que le système qui répond envoie tout ou partie de son tableau d'acheminement.

2 – réponse Un message contenant tout ou partie du tableau d'acheminement de l'envoyeur. Ce message peut être envoyé en réponse à une demande, ou peut être une mise à jour non sollicitée d'acheminement générée par l'envoyeur.

Pour chacun de ces types de message, le reste du datagramme contient une liste des RTE. Chaque RTE de cette liste contient un préfixe de destination, le nombre de bits significatifs du préfixe, et le comput pour atteindre cette destination (métrique).

Le préfixe de destination est le préfixe usuel d'adresse IPv6 de 128 bits, mémorisé par 16 octets dans l'ordre des octets du réseau.

Le champ d'étiquette de chemin est un attribut alloué à un chemin qui doit être préservé et réannoncé avec un chemin. L'utilisation prévue de l'étiquette de chemin est de donner une méthode de séparation des chemins RIPng "internes" (chemins pour les réseaux au sein du domaine d'acheminement RIPng) des chemins RIPng "externes", qui peuvent avoir été importés d'un EGP ou d'un autre IGP.

Les routeurs qui prennent en charge des protocoles autres que RIPng devraient être configurables pour permettre que l'étiquette de chemin soit configurée pour des chemins importés de différentes sources. Par exemple, des chemins importés d'un EGP devraient pouvoir avoir leur étiquette de chemin réglée à une valeur arbitraire, ou au moins au numéro du système autonome d'où le chemin a été appris.

D'autres utilisations des étiquettes de chemin sont valides, pour autant que tous les routeurs dans le domaine RIPng les utilisent de façon cohérente.

Le champ Longueur de préfixe donne la longueur en bits de la partie significative du préfixe (une valeur comprise entre 0 et 128 inclus) commençant par la gauche du préfixe.

Le champ Métrique contient une valeur entre 1 et 15 inclus, qui spécifie la métrique actuelle pour la destination; ou la valeur 16 (infini) qui indique que la destination n'est pas accessible.

La taille maximum de datagramme est limitée par la MTU du support sur lequel le protocole est utilisé. Comme une mise à jour non sollicitée de RIPng n'est jamais propagée à travers un routeur, il n'y a pas de risque de discordance de MTU. La détermination du nombre des RTE qui peuvent être mises dans un message donné est fonction de la MTU du support, du nombre d'octets d'informations d'en-tête précédant le message RIPng, de la taille de l'en-tête RIPng, et de la taille d'une RTE. La formule est :

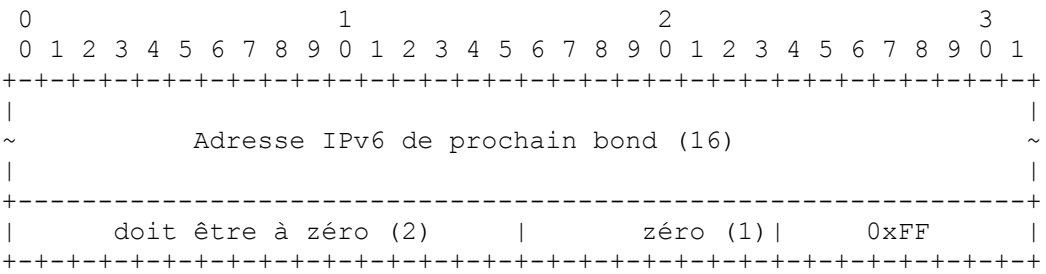
$$\text{n° RTE} = \text{INT}\left(\frac{\text{MTU} - \text{taille de (en-tête\_IPv6)} - \text{longueur d'en-tête\_UDP} - \text{longueur d'en-tête\_RIPng}}{\text{Taille\_de\_RTE}}\right)$$

2.1.1 Prochain bond

RIPng fournit la capacité de spécifier l'adresse IPv6 du prochain bond immédiat auquel devraient être transmis les paquets pour une destination spécifiée par une entrée de tableau d'acheminement (RTE) tout à fait de la même façon qu'avec RIP-2 [2]. Dans RIP-2, chaque entrée de tableau d'acheminement a un champ Prochain bond. Inclure un champ Prochain bond pour chaque RTE dans RIPng doublerait presque la taille de la RTE. Donc, dans RIPng, le prochain bond est spécifié par une RTE spéciale et s'applique à toutes les adresses des RTE qui suivent la RTE de prochain bond jusqu'à la fin du message ou jusqu'à ce qu'une antre RTE de prochain bond soit rencontrée.

Une RTE de prochain bond est identifiée par une valeur de 0xFF dans le champ Métrique d'une RTE. Le champ Préfixe spécifie l'adresse IPv6 du prochain bond. L'étiquette de chemin et la longueur du préfixe dans la RTE de prochain bond doivent être réglées à zéro à l'envoi et ignorées à réception.

Le format de l'entrée de tableau d'acheminement (RTE) du prochain bond est le suivant :



Spécifier une valeur de 0:0:0:0:0:0:0:0 dans le champ Préfixe d'une RTE de prochain bond indique que l'adresse du prochain bond devrait être le générateur de l'annonce RIPng. Une adresse spécifiée comme étant celle d'un prochain bond doit être une adresse de liaison locale.

L'objet d'une RTE de prochain bond est d'éliminer les paquets qui sont acheminés par des bonds supplémentaires dans le système. C'est particulièrement utile quand RIPng ne fonctionne pas sur tous les routeurs d'un réseau. Noter qu'une RTE de prochain bond est "consultative". C'est à dire que si les informations fournies sont ignorées, un chemin éventuellement sous optimal, mais absolument valide, peut être pris. Si l'adresse de prochain bond reçue n'est pas une adresse de liaison locale, elle devrait être traitée comme étant 0:0:0:0:0:0:0:0.

2.2 Considérations d'adressage

La distinction entre chemin de réseau, de sous-réseau et d'hôte n'a pas besoin d'être faite pour RIPng parce qu'un préfixe d'adresse IPv6 n'est pas ambigu.

Tout préfixe d'une longueur de préfixe zéro est utilisé pour désigner un chemin par défaut. Il est suggéré que le préfixe 0:0:0:0:0:0:0:0 soit utilisé quand on spécifie le chemin par défaut, bien que le préfixe soit essentiellement ignoré. Un chemin par défaut est utilisé lorsque il n'est pas pratique d'énumérer tous les réseaux possibles dans les mises à jour RIPng, et quand un ou plusieurs routeurs dans le système sont prêts à traiter le trafic vers les réseaux qui ne sont pas explicitement énumérés. Ces "routeurs par défaut" utilisent le chemin par défaut comme chemin pour tous les datagrammes pour lesquels ils n'ont pas de chemin explicite. La décision sur la façon dont un routeur devient un routeur par défaut (c'est-à-dire, comment une entrée de chemin par défaut est créée) est à la charge de la mise en œuvre. En général, l'administrateur de système va disposer d'un moyen pour spécifier quels routeurs devraient créer et annoncer des entrées de chemin par défaut. Si ce mécanisme est utilisé, la mise en œuvre devrait permettre à l'administrateur de réseau de choisir la métrique associée à l'annonce de chemin par défaut. Cela va rendre possible l'établissement d'une préséance parmi plusieurs routeurs par défaut. Les entrées de chemin par défaut sont traitées par RIPng exactement de la même manière que tout autre préfixe de destination. Les administrateurs de système devraient veiller à s'assurer que les chemins par défaut ne se propagent pas plus loin que prévu. Généralement, chaque AS a son propre routeur par défaut préféré. Donc, les chemins par défaut ne devraient généralement pas franchir les frontières d'un AS. Les mécanismes pour mettre en application cette restriction ne sont pas spécifiés dans le présent document.

2.3 Temporisateurs

Ce paragraphe décrit tous les événements qui sont déclenchés par des temporisateurs.

Toutes les 30 secondes, le processus RIPng est réveillé pour envoyer un message de réponse non sollicitée, contenant le tableau d'acheminement complet (voir au paragraphe 2.6 "Horizon partagé") à tous les routeurs du voisinage. Lorsque il y a de nombreux routeurs sur un seul réseau, il y a une tendance à la synchronisation qui fait qu'ils vont tous produire leur mise

à jour en même temps. Cela peut arriver chaque fois que le temporisateur de 30 secondes est affecté par la charge de traitement sur le système. Il n'est pas souhaitable que les messages de mise à jour soient synchronisés, parce que cela peut conduire à des collisions inutiles sur les réseaux en diffusion (voir les détails dans [13]). Il est donc demandé aux mises en œuvre de prendre une des deux précautions suivantes :

- Déclencher les mises à jour de 30 secondes avec une horloge dont le débit n'est pas affecté par la charge du système ou par le moment du service du précédent temporisateur de mise à jour.

- Le temporisateur de 30 secondes est décalé d'une petite durée aléatoire (+/- 0 à 15 secondes) chaque fois qu'il est établi. Le décalage est déduit de 0,5 * la période de mise à jour (c'est-à-dire, 30).

Deux temporisateurs sont associés à chaque chemin, une "fin de temporisation" et un "délai de collecte des déchets." À l'expiration de la fin de temporisation, le chemin n'est plus valide ; cependant, il est conservé dans le tableau d'acheminement pendant un bref délai afin que les voisins puissent être notifiés de l'abandon du chemin. À expiration du délai de collecte des déchets, le chemin est finalement retiré du tableau d'acheminement.

La fin de temporisation est initialisée lors de l'établissement d'un chemin, et à chaque fois qu'un message de mise à jour est reçu pour le chemin. Si 180 secondes s'écoulent depuis la dernière fois que la temporisation a été initialisée, le chemin est considéré comme ayant expiré, et le processus de suppression décrit ci-dessous commence pour ce chemin.

Les suppressions peuvent survenir pour une de deux raisons : l'expiration de la fin de temporisation, ou le réglage de la métrique à 16 à cause d'une mise à jour reçue du routeur actuel (voir au paragraphe 2.4.2 une discussion sur le traitement des mises à jour à partir d'autres routeurs). Dans l'un et l'autre cas, les événements suivants se produisent :
- Le temporisateur de collecte des déchets est réglé à 120 secondes.
- La métrique pour le chemin est réglée à 16 (infini). Cela cause le retrait du service de ce chemin.
- Le fanion de changement de chemin pour indiquer que cette entrée a été changée.
- Le processus de sortie est signalé pour déclencher une réponse.

Jusqu'à l'expiration du délai de collecte des déchets, le chemin est inclus dans toutes les mises à jour envoyées par ce routeur. Lorsque le délai de collecte des déchets arrive à expiration, le chemin est supprimé du tableau d'acheminement.

Si un nouveau chemin pour ce réseau devait être établi alors que le délai de collecte des déchets court encore, le nouveau chemin remplacerait celui qui est sur le point d'être supprimé. Dans ce cas, le temporisateur de délai de collecte des déchets doit être supprimé.

Les mises à jour déclenchées utilisent aussi un petit temporisateur ; ceci est cependant mieux décrit au paragraphe 2.5.1.

2.4 Traitement des entrées

Ce paragraphe décrit le traitement des datagrammes reçus sur l'accès RIPng. Le traitement va dépendre de la valeur qui est dans le champ Commande. La version 1 ne prend en charge que deux commandes : Demande et Réponse.

2.4.1 Messages de demande

Une demande est utilisée pour réclamer une réponse contenant tout ou partie du tableau d'acheminement d'un routeur. Normalement, les demandes sont envoyées comme diffusion groupée, à partir de l'accès RIPng, par les routeurs qui viennent juste de s'activer et qui cherchent à remplir leur tableau d'acheminement aussi rapidement que possible. Cependant, il peut y avoir des situations (par exemple, la surveillance du routeur) où le tableau d'acheminement d'un seul routeur est nécessaire. Dans ce cas, la Demande devrait être envoyée directement à ce routeur à partir d'un accès UDP autre que l'accès RIPng. Si une telle Demande est reçue, le routeur répond directement à l'adresse et accès du demandeur avec une adresse de source valide mondialement car le demandeur peut ne pas résider sur le réseau directement rattaché.

La Demande est traitée entrée par entrée. Si il n'y a pas d'entrées, aucune réponse n'est donnée. Il y a un cas particulier. Si il y a exactement une entrée dans la demande, et si elle a un préfixe de destination de zéro, une longueur de préfixe de zéro, et une métrique de infini (c'est-à-dire, 16) c'est alors une demande d'envoyer la totalité du tableau d'acheminement. Dans ce cas, un appel est fait au processus de sortie pour envoyer le tableau d'acheminement à l'adresse/accès demandeur. Excepté ce cas particulier, le traitement est assez simple. Examiner une par une la liste des RTE dans la Demande. Pour chaque entrée, chercher la destination dans la base de données d'acheminement du routeur, et si il y a un chemin, mettre la métrique de ce chemin dans le champ Métrique de la RTE. Si il n'y a pas de chemin explicite pour la destination spécifiée, mettre infini dans le champ de métrique. Une fois que toutes les entrées ont été remplies, changer la commande de Demande à Réponse et renvoyer le datagramme au demandeur.

Noter qu'il y a une différence dans le traitement des métriques pour les demandes spécifiques et celles qui portent sur le tableau entier. Si la demande est pour un tableau d'acheminement complet, le traitement normal de sortie est fait, y compris l'horizon partagé (voir au paragraphe 2.6 "Horizon partagé"). Si la demande est pour des entrées spécifiques, il y a une recherche dans le tableau d'acheminement et les informations sont retournées telles quelles ; aucun traitement d'horizon partagé n'est effectué. La raison de cette distinction est qu'on s'attend à ce que ces demandes vont probablement être utilisés pour différents objets. Lorsque un routeur devient actif, il envoie une Demande en diffusion groupée sur tous les réseaux connectés et demande un tableau d'acheminement complet. On suppose que ces tableaux d'acheminement complets seront utilisés pour mettre à jour le tableau d'acheminement du demandeur. Pour cette raison, l'horizon partagé doit être effectué. On suppose de plus qu'une Demande pour un réseau spécifique n'est faite que par des logiciels de diagnostic, et qu'elle n'est pas utilisée pour l'acheminement. Dans ce cas, le demandeur voudra savoir le contenu exact du tableau d'acheminement et ne voudra pas d'informations mises en antémémoire ou modifiées.

2.4.2 Messages de réponse

Une Réponse peut être reçue pour une raison parmi plusieurs :
- réponse à une interrogation spécifique
- mise à jour régulière (réponse non sollicitée)
- mise à jour déclenchée causée par un changement de chemin

Le traitement est le même quelle que soit la cause qui a généré la Réponse.

Parce que le traitement d'une Réponse peut mettre à jour le tableau d'acheminement du routeur, la validité de la Réponse doit être vérifiée avec soin. La Réponse doit être ignorée si elle ne vient pas de l'accès RIPng. L'adresse IPv6 de source du datagramme devrait être vérifiée pour voir si le datagramme vient d'un voisin valide ; la source du datagramme doit être une adresse de liaison locale. Cela vaut aussi la peine de vérifier si la réponse vient d'une des propres adresses du routeur. Les interfaces sur les réseaux en diffusion peuvent recevoir immédiatement des copies de leurs propres diffusions groupées. Si un routeur traite ses propres sorties comme de nouvelles entrées, la confusion est probable, et de tels datagrammes doivent être ignorés. Comme vérification supplémentaire, les annonces périodiques doivent avoir leur compte de bonds réglé à 255, et les paquets entrants de diffusion groupée envoyés de l'accès RIPng (c'est-à-dire, les paquets d'annonce périodique ou les paquets de mise à jour déclenchée) doivent être examinés pour s'assurer que leur compte de bonds est 255. Cela garantit absolument qu'un paquet provient d'un voisin, parce que tout nœud intermédiaire aurait décrémenté le compte de bonds. Les interrogations et leurs réponses peuvent encore traverser des nœuds intermédiaires et n'exigent donc pas que soit effectuée la vérification du compte de bonds.

Une fois que le datagramme a été validé dans sa globalité, on traite les RTE de la Réponse une par une. Là encore, on commence par faire une validation. Les métriques incorrectes et autres erreurs de format indiquent habituellement des voisins au mauvais comportement et devraient probablement être portées à l'attention de l'administrateur. Par exemple, si la métrique est supérieure à l'infini, ignorer l'entrée mais inscrire l'événement sur le journal. Les essais de validation de base sont :
- le préfixe de destination est-il valide ? (par exemple, ce n'est pas un préfixe de diffusion groupée ni une adresse de liaison locale.) Une adresse de liaison locale ne devrait jamais être présente dans une RTE.
- la longueur du préfixe est-elle valide ? (c'est-à-dire, entre 0 et 128, inclus)
- la métrique est-elle valide ? (c'est-à-dire, entre 1 et 16, inclus)

Si une de ces vérifications échoue, ignorer cette entrée et traiter la suivante. Là encore, l'enregistrement de l'erreur dans le journal est probablement une bonne idée.

Une fois que l'entrée a été validée, mettre à jour la métrique en ajoutant le coût du réseau sur lequel le message est arrivé. Si le résultat est supérieur à l'infini, utiliser l'infini. C'est à dire ,

métrique = MIN (métrique + coût, infini)

Ensuite, on vérifie qu'il y a déjà un chemin explicite pour le préfixe de la destination. Si il n"y a pas un tel chemin, l'ajouter au tableau d'acheminement, sauf si la métrique est l'infini (il n'y a pas de raison d'ajouter un chemin qui est inutilisable). Ajouter un chemin au tableau d'acheminement consiste à :
- Régler le préfixe et la longueur de destination à ce qui est dans la RTE.
- Régler la métrique à celle qui vient d'être calculée (comme décrit ci-dessus).
- Régler l'adresse de prochain bond à l'adresse du routeur d'où est venu le datagramme ou à l'adresse de prochain bond spécifiée par une RTE de prochain bond.
- Initialiser la temporisation pour ce chemin. Si le temporisateur de délai de collecte des déchets est en cours pour ce chemin, l'arrêter (voir au paragraphe 2.3 un exposé sur les temporisateurs).
- Établir le fanion de changement de chemin.

- Signaler au processus de sortie de déclencher une mise à jour (voir au paragraphe 2.5).

Si il y a un chemin existant, comparer l'adresse du prochain bond à l'adresse du routeur d'où est venu le datagramme. Si ce datagramme vient du même routeur que le chemin existant, réinitialiser le temporisateur. Ensuite, comparer les métriques. Si le datagramme vient du même routeur que le chemin existant, et si la nouvelle métrique est différente de l'ancienne; ou, si la nouvelle métrique est inférieure à l'ancienne, effectuer les actions suivantes :
- Adopter le chemin du datagramme. C'est-à-dire, y mettre la nouvelle métrique, et ajuster l'adresse de prochain bond (si nécessaire).
- Établir le fanion de changement de chemin et signaler au processus de sortie de déclencher une mise à jour.
- Si la nouvelle métrique est l'infini, commencer le processus de suppression (décrit ci-dessus) ; autrement, réinitialiser le temporisateur.

Si la nouvelle métrique est l'infini, le processus de suppression commence pour le chemin, qui n'est plus utilisé pour acheminer les paquets. Noter que le processus de suppression n'est commencé que lorsque la métrique est d'abord réglée à l'infini. Si la métrique était déjà à l'infini, on ne commence pas un nouveau processus de suppression.

Si la nouvelle métrique est la même que l'ancienne, il est plus simple de ne rien faire de plus (en dehors de réinitialiser la temporisation, comme spécifié ci-dessus) mais il y a une heuristique qui pourrait être appliquée. Normalement, cela n'a pas de sens de remplacer un chemin si le nouveau chemin a la même métrique que le chemin existant ; cela va provoquer un mouvement de va et vient du chemin et générer un nombre intolérable de déclenchements de mises à jour. Cependant, si le chemin existant montre des signes de fin de temporisation, il peut être préférable de passer immédiatement à un chemin de remplacement également bon, plutôt que d'attendre l'arrivée de la fin de temporisation. Si la nouvelle métrique est la même que celle de l'ancienne, on examine donc la temporisation du chemin existant. Si elle est au moins à la moitié du point d'expiration, on passe au nouveau chemin. Cette heuristique est facultative, mais vivement recommandée.

Toute entrée qui échoue à ces vérifications est ignorée, car elle n'est pas meilleure que le chemin actuel.

2.5 Traitement des résultats

Ce paragraphe décrit le traitement utilisé pour créer des messages de réponse qui contiennent tout ou partie du tableau d'acheminement. Ce traitement peut être déclenché d'une des façons suivantes :
- Par le traitement d'entrée, lors de la réception d'une Demande. Dans ce cas, la Réponse n'est envoyée qu'à une seule destination (c'est-à-dire, à l'adresse d'envoi individuel du demandeur).
- Par la mise à jour régulière d'acheminement. Toutes les 30 secondes, une Réponse contenant la totalité du tableau d'acheminement est envoyée à chaque routeur du voisinage.
- Par une mise à jour déclenchée. Chaque fois que la métrique d'un chemin est changée, une mise à jour est déclenchée.

Le traitement particulier nécessaire pour une Demande est décrit au paragraphe 2.4.1.

Lorsque une Réponse doit être envoyée à tous les voisins (c'est-à-dire, une mise à jour régulière ou déclenchée) un message Réponse est envoyé en diffusion groupée au groupe de diffusion groupée FF02::9, qui est le groupe de diffusion groupée de tous les routeurs rip, sur tous les réseaux connectés qui prennent en charge la diffusion ou sont des liaisons point à point. RIPng traite les liaisons point à point exactement comme des liaisons de diffusion groupée car la fourniture de la diffusion groupée est triviale sur de telles liaisons. Donc, une Réponse est préparée pour chaque réseau directement connecté, et envoyée au groupe de diffusion groupée tous-routeurs-rip. Dans la plupart des cas, elle atteint tous les routeurs voisins. Cependant, il y a des cas où cela peut n'être pas suffisant. Cela peut impliquer un réseau qui n'est pas un réseau de diffusion (par exemple, l'ARPANET) ou une situation impliquant des routeurs muets. Dans de tels cas, il peut être nécessaire de spécifier une liste des routeurs effectifs du voisinage et d'envoyer explicitement un datagramme à chacun d'eux. On laisse aux mises en œuvre le soin de déterminer si un tel mécanisme est nécessaire, et de définir comment est spécifiée la liste.

2.5.1 Mises à jour déclenchées

Les mises à jour déclenchées exigent un traitement particulier pour deux raisons. D'abord, l'expérience montre que les mises à jour déclenchées causent une charge excessive sur les réseaux qui ont des capacités limitées ou sur les réseaux qui portent de nombreux routeurs. Donc, le protocole exige que les mises en œuvre comportent des dispositions pour limiter la fréquence des mises à jour déclenchées. Après l'envoi d'une mise à jour déclenchée, un temporisateur devrait être établi pour un intervalle aléatoire entre 1 et 5 secondes. Si d'autres changements qui déclencheraient des mises à jour surviennent avant l'expiration de la temporisation, une seule mise à jour est déclenchée lorsque la temporisation arrive à expiration. Le temporisateur est alors réinitialisé à une autre valeur aléatoire entre une et cinq secondes. Les mises à jour déclenchées peuvent être supprimées si une mise à jour régulière est prévue au moment où la mise à jour déclenchée devrait être envoyée.

Ensuite, les mises à jour déclenchées n'ont pas besoin d'inclure le tableau d'acheminement entier. En principe, seuls les chemins qui ont changé doivent être inclus. Donc les messages générés au titre d'une mise à jour déclenchée doivent inclure au moins les chemins qui ont leur fanion Changement de chemin établi. Ils peuvent inclure des chemins supplémentaires, à la discrétion de la mise en œuvre ; cependant, l'envoi de mises à jour d'acheminement complètes est fortement déconseillé. Lorsque une mise à jour déclenchée est traitée, les messages devraient être générés pour chaque réseau directement connecté. Le traitement d'horizon partagé est effectué lors de la génération de mises à jour déclenchées aussi bien que pour les mises à jour normales (voir au paragraphe 2.6). Si, après le traitement d'horizon partagé pour un réseau donné, un chemin modifié apparaît inchangé sur ce réseau (par exemple, il apparaît avec une métrique de infini) le chemin n'a pas besoin d'être envoyé. Si aucun chemin n'a besoin d'être envoyé sur ce réseau, la mise à jour peut être omise. Une fois que toutes les mises à jour déclenchées ont été générées, les fanions de changement de chemin devraient être supprimés.

Si le traitement d'entrées est permis pendant que les résultats sont générés, les verrouillages appropriés doivent être effectués. Les fanions de changement de chemin ne devraient pas être changés par suite d'un traitement d'entrée alors qu'un message de mise à jour déclenchée est généré.

La seule différence entre une mise à jour déclenchée et d'autres messages de mise à jour est l'omission possible de chemins qui n'ont pas changé. Les mécanismes restants, décrits dans le paragraphe suivant, doivent être appliqués à toutes les mises à jour.

2.5.2 Générer des messages de réponse

Ce paragraphe décrit comment un message Réponse est généré pour un réseau directement connecté particulier :

L'adresse de source IPv6 doit être une adresse de liaison locale des adresses possibles de l'interface du routeur envoyeur, sauf lors d'une réponse à un message Demande en envoi individuel provenant d'un accès autre que l'accès RIPng. Dans ce dernier cas, l'adresse de source doit être une adresse valide mondialement. Dans le premier cas, il est important d'utiliser une adresse de liaison locale parce que l'adresse de source est mise dans les tableaux d'acheminement (comme prochain bond) chez les routeurs qui reçoivent cette Réponse. Si une adresse de source incorrecte est utilisée, les autres routeurs pourraient être incapables d'acheminer les datagrammes. Parfois, les routeurs sont établis avec plusieurs adresses IPv6 sur une seule interface physique. Normalement, cela signifie que plusieurs réseaux IPv6 logiques sont portés sur un seul support physique. Il est possible qu'un routeur ait plusieurs adresses de liaison locale pour une seule interface. Dans ce cas, le routeur doit générer seulement un message Réponse avec une adresse de source de l'adresse de liaison locale désignée pour une interface donnée. Le choix de l'adresse de liaison locale à utiliser ne devrait changer que lorsque le choix actuel n'est plus valide. Ceci est nécessaire parce que les nœuds qui reçoivent les messages Réponse utilisent l'adresse de source pour identifier l'envoyeur. Si plusieurs paquets du même routeur contiennent des adresses de source différentes, les nœuds vont supposer qu'ils viennent de routeurs différents, ce qui conduirait à des comportements indésirables.

Régler le numéro de version à la version actuelle de RIPng. La version décrite dans le présent document est la version 1. Régler la commande à Réponse. Régler les octets étiquetés "doit être à zéro" à zéro. Commencer à remplir les RTE. Se rappeler que la taille maximum de datagramme est limitée par la MTU du réseau. Quand il n'y a plus d'espace dans le datagramme, envoyer la Réponse actuelle et en commencer une nouvelle.

Pour remplir les RTE, examiner chaque chemin du tableau d'acheminement. Les chemins pour les adresses de liaison locale ne doivent jamais être inclus dans une RTE. Si une mise à jour déclenchée est générée, seules les entrées dont le fanion Changement de chemin est établi doivent être incluses. Si, après le traitement d'horizon partagé, le chemin ne devrait pas être inclus, le sauter. Si le chemin est à inclure, le préfixe de destination, la longueur de préfixe, et la métrique sont alors mis dans la RTE. L'étiquette de chemin est remplie comme défini au paragraphe 2.1. Les chemins doivent être inclus dans le datagramme même si leur métrique est l'infini.

2.6 Horizon partagé

L'horizon partagé est un algorithme pour éviter les problèmes causés par l'inclusion de mises à jour de chemins envoyés à la passerelle d'où ils ont été appris. L'algorithme de base d'horizon partagé omet les chemins appris d'un voisin dans les mises à jour envoyées à ce voisin. Dans le cas d'un réseau de diffusion, tous les chemins appris d'un voisin quelconque sur ce réseau sont omis des mises à jour envoyées sur ce réseau.

L'horizon partagé avec inversion empoisonnée (dite plus simplement inversion empoisonnée) n'inclut pas de tels chemins dans les mises à jour, mais règle leur métrique à l'infini. En effet, cela annonce que leurs chemins ne sont pas accessibles. C'est la méthode préférée de fonctionnement ; cependant, les mises en œuvre devraient fournir un contrôle par interface permettant de choisir entre pas d'horizon partagé, l'horizon partagé, et l'inversion empoisonnée.

Pour une discussion théorique sur l'horizon partagé et l'inversion empoisonnée et pourquoi ils sont nécessaires, voir au

paragraphe 2.1.1 de [1].

3. Fonctions de contrôle

Cette section décrit les contrôles administratifs. Ils ne font pas par eux-mêmes partie du protocole, cependant l'expérience des réseaux existants suggère qu'ils sont importants. Comme ils ne constituent pas une partie nécessaire du protocole, ils sont considérés comme facultatifs. Cependant, il est vivement recommandé qu'au moins certains d'entre aux soient inclus dans toute mise en œuvre. Ces contrôles sont destinés principalement à permettre à RIPng d'être connecté aux réseaux dont l'acheminement pourrait être instable ou sujet à l'erreur. Voici quelques exemples :

- Il est parfois souhaitable de restreindre les routeurs d'où les mises à jour seront acceptées, ou auxquels les mises à jour seront envoyées. Cela est généralement fait pour des raisons administratives, de politique d'acheminement.
- Un certain nombre de sites limitent l'ensemble de réseaux qu'ils permettent dans les messages Réponse. L'organisation A peut avoir une connexion avec l'organisation B qu'elles utilisent pour des communications directes. Pour des raisons de sécurité ou de performances, A peut n'être pas d'accord pour donner à d'autres organisations l'accès à cette connexion. Dans un tel cas, A ne devrait pas inclure les réseaux de B dans les mises à jour que A envoie à des tiers.

Voici quelques contrôles typiques. Noter cependant que le protocole RIPng ne les exige pas, ni eux, ni d'autres.

- Une liste des voisins qui permet à l'administrateur de réseau de définir une liste de voisins pour chaque routeur. Un routeur n'accepterait de messages de réponse que des routeurs qui sont sur sa liste de voisins. Une liste similaire pour les routeurs cibles devrait aussi être disponible pour l' administrateur. Par défaut, aucune restriction n'est définie.

- Un filtre pour des destinations spécifiques permettrait à l'administrateur de réseau de spécifier une liste de préfixes de destinations à permettre ou interdire. La liste serait associée à une interface particulière de la direction entrante et/ou sortante. Seuls les réseaux admis seraient mentionnés dans les messages Réponse sortants ou traités dans les messages Réponse entrants. Si une liste de préfixes permis est spécifiée, tous les autres préfixes sont interdits. Si une liste de préfixes interdits est spécifiée, tous les autres préfixes sont permis. Par défaut, aucun filtre n'est appliqué.

4. Considérations pour la sécurité

Depuis que RIPng fonctionne sur IPv6, RIPng s'appuie sur l'en-tête d'authentification IP (voir [11]) et la charge utile de sécurité encapsulée dans IP (voir [12]) pour assurer l'intégrité et l'authentification/confidentialité des échanges d'acheminement.

Références

[1] C. Hedrick, "Protocole d'informations d'acheminement", RFC1058, juin 1988. *(Historique)*

[2] G. Malkin, "RIP v2, portage d'informations supplémentaires", RFC1723, novembre 1994. *(remplacée par RFC2453)*

[3] R. Hinden, "Nouveau schéma d'acheminement et adressage Internet (ENCAPS) pour IPng", RFC1955, juin 1996.

[4] Bellman, R., "Dynamic Programming", Princeton University Press, Princeton, N.J., 1957.

[5] Bertsekas, D. P., and Gallaher, R. G., "Data Networks", Prentice- Hall, Englewood Cliffs, N.J., 1987.

[6] R. Braden et J. Postel, "Exigences pour les routeurs de l'Internet"[RFC1009], juin 1987. *(Obsolète voir RFC 1812) (Historique)*

[7] Boggs, D. R., Shoch, J. F., Taft, E. A., and Metcalfe, R. M., "Pup: An Internetwork Architecture", IEEE Transactions on Communications, avril 1980.

[8] Ford, L. R. Jr., and Fulkerson, D. R., "Flows in Networks", Princeton University Press, Princeton, N.J., 1962.

[9] Xerox Corp., "Internet Transport Protocols", Xerox System Integration Standard XSIS 028112, décembre 1981.

[10] T. Narten, E. Nordmark, W. Simpson, "Découverte du voisinage pour IP version 6 (IPv6)"RFC1970, août 1996. (*Obsolète, voir* RFC4861) *(P.S.)*

[11] R. Atkinson, "En-tête d'authentification IP", RFC1826, août 1995. *(Rendue obsolète par la RFC4302)*

[12] R. Atkinson, "Encapsulation dans IP de charge utile de sécurité (ESP)", RFC1827, août 1995. *(obsolète, voir RFC4303)*

[13] Floyd, S., and Jacobson, V., "The Synchronization of Periodic Routing Messages", Proceedings of ACM SIGCOMM '93, septembre 1993.

Adresse des auteurs

Gary Scott Malkin
Xylogics, Inc.
53 Third Avenue
Burlington, MA 01803
téléphone : (617) 272-8140
mél : gmalkin@Xylogics.com

Robert E. Minnear
Ipsilon Networks, Inc.
2191 E. Bayshore Road, Suite 100
Palo Alto, CA 94303
téléphone : (415) 846-4614
mél : minnear@ipsilon.com