Groupe de travail Réseau
Request for Comments : 1812
STD004
RFC rendues obsolètes : 1716, 1009
Catégorie : Norme

F. Baker, éditeur, Cisco Systems
juin 1995

Traduction Claude Brière de L'Isle
août 2007

Exigences pour les routeurs IP Version 4

Statut de ce mémoire
Le présent document spécifie un protocole Internet normalisé pour la communauté Internet, et appelle à des discussions et suggestions pour son amélioration. Prière de se référer à l'édition actuelle des "Normes officielles des protocoles Internet" (STD 1) pour l'état de la normalisation et le statut de ce protocole. La distribution du présent mémoire n'est soumise à aucune restriction.

Préface

Le présent document est une version mise à jour de la RFC 1716, le document historique des exigences pour les routeurs. Cette RFC a préservé le travail significatif effectué dans le groupe de travail, mais n'a pas réussi à décrire de façon adéquate la technologie actuelle pour que l'IESG le considère comme la norme actuelle.

Il a été demandé à l'éditeur de mettre le document à jour, de sorte qu'il soit utile comme spécification de transition et comme guide pour les développeurs. En cela, il suit fièrement les traces de ceux qui l'ont précédé, et dépend dans une large mesure des textes des experts qui y ont contribué. Tout le crédit leur revient ; les erreurs sont les siennes.

Le contenu et la forme de ce document sont dues, en grande partie, au président du groupe de travail, et à l'éditeur et auteur du document original : Philip Almquist. Ils sont aussi largement dus aux efforts de son éditeur précédent, Frank Kastenholz. Sans leurs efforts, le présent document n'aurait pas vu le jour.

Table des matières

1 Introduction

Le présent mémoire remplace la RFC 1716, "Exigences pour les passerelles Internet" ([INTRO:1]).

Le présent mémoire définit et discute des exigences concernant les appareils qui effectuent la fonction de transmission à la couche réseau de la suite de protocoles Internet. La communauté de l'Internet se réfère habituellement à de tels appareils sous le nom de routeurs IP ou simplement de routeurs ; la communauté OSI se réfère à de tels appareils sous le nom de systèmes intermédiaires. De nombreux documents plus anciens de l'Internet se réfèrent à de tels appareils sous le nom de passerelles, nom qui est tombé en disgrâce récemment pour éviter la confusion avec les passerelles d'application.

Un routeur IP peut être distingué des autres sortes d'appareils de commutation de paquets en ce que un routeur examine l'en-tête de protocole IP au titre du processus de commutation. Il retire généralement l'en-tête de couche de liaison avec lequel le message a été reçu, modifie l'en-tête IP, et remplace l'en-tête de couche de liaison pour la retransmission.

Les auteurs du présent mémoire reconnaissent, comme devraient le faire de nombreux lecteurs, que de nombreux routeurs prennent en charge plus d'un protocole. La prise en charge de plusieurs suites de protocoles sera à l'avenir exigée de parties croissantes de l'Internet. Le présent mémoire, n'essaye cependant pas de spécifier des exigences Internet pour d'autres suites de protocole que TCP/IP.

Le présent document énumère les protocoles standard que doit utiliser un routeur connecté à l'Internet, et y incorpore par référence les RFC et les autres documents qui décrivent les spécifications actuelles de ces protocoles. Il corrige des erreurs dans les documents référencés et ajoute des développement et conseils supplémentaires pour les développeurs.

Pour chaque protocole, le présent mémoire contient aussi un ensemble explicite d'exigences, recommandations, et options. Le lecteur doit comprendre que la liste des exigences du présent mémoire est par elle-même incomplète. L'ensemble d'exigences complet pour un routeur de protocole Internet est principalement défini dans les documents standard de spécification de protocoles, avec les corrections, amendements, et suppléments contenus dans le présent mémoire.

Le présent mémoire devrait être lu en conjonction avec les RFC sur les exigences pour les hôtes Internet ([INTRO:2] et [INTRO:3]). Les hôtes et routeurs Internet doivent être tous deux capables de se trouver à l'origine de datagrammes IP et de recevoir des datagrammes IP qui leurs sont destinés. La distinction majeure entre les hôtes et les routeurs Internet est que les routeurs mettent en œuvre des algorithmes de transmission, alors que les hôtes Internet n'exigent pas de capacités de transmission. Tout hôte Internet qui agit comme routeur doit adhérer aux exigences contenues dans le présent mémoire.

Le but de l'interconnexion des systèmes ouverts impose que les routeurs fonctionnent correctement comme hôtes Internet lorsque nécessaire. Pour y arriver, le présent mémoire donne des lignes directrices pour ces cas. Pour simplifier et faciliter la mise à jour du document, le présent mémoire essaye d'éviter le chevauchement des discussions sur les exigences pour les hôtes figurant dans [INTRO:2] et [INTRO:3] et incorpore les exigences pertinentes de ces documents par référence. Dans certains cas, les exigences établies dans [INTRO:2] et [INTRO:3] sont supplantées par celles du présent document.

Une mise en œuvre de bonne foi des protocoles produits après une lecture attentive des RFC ne devrait différer que de façon mineure des exigences du présent mémoire. La production d'une telle mise en œuvre exige souvent une certaine interaction avec la communauté technique de l'Internet, et doit respecter les bonnes pratiques d'ingénierie des logiciels de communications. Dans de nombreux cas, les exigences du présent document sont déjà établies ou impliquées par les documents de protocole standard, de sorte que leur inclusion ici est, en un sens, redondante. Elles ont été incluses parce que certaines mises en œuvre passées ont fait de mauvais choix, ce qui pause des problèmes d'interopérabilité, de performance, et/ou de robustesse.

Le présent mémoire inclut une discussion et l'explication de beaucoup des exigences et des recommandations. Une simple liste des exigences serait dangereuse, parce que :

- Certaines caractéristiques exigées sont plus importantes que d'autres, et certaines sont facultatives.
- Certaines caractéristiques sont critiques dans certaines applications de routeurs mais non pertinentes dans d'autres.
- Il peut y avoir des raisons valides pour que les produits d'un fabricant particulier conçus pour des contextes restreints puissent choisir d'utiliser des spécifications différentes.

Cependant, les spécifications du présent mémoire doivent être suivies pour satisfaire à l'objectif général d'interopérabilité d'un routeur quelconque à travers la diversité et la complexité de l'Internet. Bien que la plupart des mises en œuvre actuelles échouent de diverses façons à satisfaire ces exigences, certaines mineures et certaines majeures, la présente spécification est un idéal vers lequel il est nécessaire de progresser.

Ces exigences se fondent sur le niveau actuel de l'architecture Internet. Le présent mémoire sera mis à jour en tant que de besoin pour fournir des précisions supplémentaires ou pour inclure des informations supplémentaires dans les domaines dans lesquels les spécifications sont encore en évolution.

1.1 Lecture du présent document

1.1.1 Organisation

Le présent mémoire émule l'organisation en couches utilisée par [INTRO:2] et [INTRO:3]. Et donc, la section 2 décrit les couches qui se trouvent dans l'architecture Internet. La section 3 traite la couche de liaison des données. Les sections 4 et 5 s'occupent du protocole et des algorithmes de transmission de la couche Internet. La section 6 couvre la couche Transport. Les protocoles des souches supérieures sont divisés entre les sections 7, 8, et 9. La section 7 discute des protocoles qu'utilisent les routeurs pour échanger les informations d'acheminement les uns avec les autres. La section 8 discute de la gestion du réseau. La section 9 discute des autres protocoles de couche supérieure. La dernière section couvre les caractéristiques de fonctionnement et de maintenance. Cette organisation a été choisie pour sa simplicité, sa clarté, et sa cohérence avec les RFC sur les exigences pour les hôtes. Les appendices au présent mémoire incluent une bibliographie, un glossaire, et quelques conjectures sur les futures directions des normes sur les routeurs.

En décrivant les exigences, on suppose qu'une mise en œuvre reflète strictement la mise en couches des protocoles. Cependant, une stratification stricte est un modèle imparfait, à la fois pour la suite des protocoles et pour les approches de mises en œuvre recommandées. Les protocoles dans les différentes couches interagissent de façons complexes et parfois subtiles, et des fonctions particulières impliquent souvent plusieurs couches. Il y a de nombreux choix de conception dans une mise en œuvre, dont beaucoup impliquent une rupture créatrice de la stratification stricte. Chaque développeur devrait impérativement lire [INTRO:4] et [INTRO:5].

Chaque section majeure du présent mémoire est organisée selon les paragraphes suivants :

(1) Introduction
(2) Découverte du protocole – considère les documents de spécification du protocole paragraphe par paragraphe, en corrigeant les erreurs, établissant les exigences qui pourraient être ambiguës ou mal définies, et en fournissant d'autres éclaircissements ou explications.
(3) Questions spécifiques – discute les questions de conception et de mise en œuvre du protocole qui n'étaient pas incluse dans la découverte.

Dans beaucoup des sujets individuels du présent mémoire, se trouvent des passages entre parenthèses indiqués par Discussion ou MISE EN ŒUVRE. Ces passages sont destinés à donner une justification, des éclaircissements ou des explication sur le texte d'exigences précédent. Les passages de mise en œuvre contiennent les suggestions d'approches qu'un développeur pourrait vouloir prendre en considération. Les paragraphes Discussion et MISE EN ŒUVRE ne font pas partie de la norme.

1.1.2 Exigences

Dans le présent mémoire, les mots qui sont utilisés pour définir la signification de chaque exigence particulière sont en majuscules. Ces mots sont :

DOIT
Ce mot signifie que l'élément est une exigence absolue de la spécification. Les violations d'une telle exigence sont des erreurs fondamentales ; elles ne sont justifiées en aucun cas.

DOIT METTRE EN ŒUVRE
Cette phrase signifie que la présente spécification exige que l'élément soit mis en œuvre, mais n'exige pas qu'il soit activé par défaut.

NE DOIT PAS
Cette phrase signifie que l'élément est une interdiction absolue de la spécification.

DEVRAIT
Ce mot signifie qu'il peut exister des raisons valides dans des circonstances particulières pour ignorer cet élément, mais les implications complètes devraient en être comprises et le cas soigneusement évalué avant de choisir une voie différente.

DEVRAIT METTRE EN ŒUVRE
Cette phrase a une signification similaire à DEVRAIT, mais est utilisée lorsque on recommande qu'une caractéristique particulière soit fournie, mais pas nécessairement recommandé qu'elle soit activée par défaut.

NE DEVRAIT PAS
Cette phrase signifie qu'il peut exister des raisons valides dans des circonstances particulières où le comportement décrit est acceptable ou même utile. Même alors, les implications complètes devraient être comprises et le cas soigneusement évalué avant de mettre en œuvre tout comportement décrit avec cette mention.

PEUT
Ce mot signifie que l'élément est vraiment facultatif. Un fabricant peut choisir d'inclure l'élément parce qu'un marché particulier l'exige ou parce qu'il améliore le produit, par exemple ; un autre fabricant peut omettre le même élément.

1.1.3 Conformité

Certaines exigences sont applicables à tous les routeurs. D'autres exigences ne sont applicables qu'à ceux qui mettent en œuvre des caractéristiques ou protocoles particuliers. Dans les paragraphes qui suivent, pertinent se réfère à l'union des exigences applicables à tous les routeurs et à l'ensemble d'exigences applicables à un routeur particulier à cause de

l'ensemble de caractéristiques et protocoles qu'il a mis en œuvre.

Noter que toutes les exigences pertinentes ne sont pas établies directement dans le présent mémoire. Diverses parties du présent mémoire incorporent par référence des paragraphes de la spécification sur les exigences des hôtes, [INTRO:2] et [INTRO:3]. Pour déterminer la conformité au présent mémoire, il importe peu qu'une exigence pertinente soit établie directement dans le présent mémoire ou simplement incorporée par référence tirée d'un de ces documents.

Une mise en œuvre est dite être conditionnellement conforme si elle satisfait à toutes les exigences pertinentes DOIT, DOIT METTRE EN ŒUVRE, et NE DOIT PAS. Une mise en œuvre est dite être inconditionnellement conforme si elle est conditionnellement conforme et satisfait aussi à toutes les exigences pertinentes DEVRAIT, DEVRAIT METTRE EN ŒUVRE, et NE DEVRAIT PAS. Une mise en œuvre n'est pas conforme si elle n'est pas conditionnellement conforme (c'est-à-dire, échoue à satisfaire une ou plusieurs des exigences pertinentes DOIT, DOIT METTRE EN ŒUVRE, ou NE DOIT PAS).

La présente spécification indique occasionnellement qu'une mise en œuvre DEVRAIT mettre en œuvre une variable de gestion, et qu'elle DEVRAIT avoir une certaine valeur par défaut. Une mise en œuvre inconditionnellement conforme met en œuvre le comportement par défaut, et si il y a d'autres comportements mis en œuvre, met en œuvre la variable. Une mise en œuvre conditionnellement conforme documente clairement ce qu'est le réglage par défaut de la variable ou, en l'absence de la mise en œuvre d'une variable, peut le faire par construction. Une mise en œuvre qui échoue à la fois à mettre en œuvre la variable et choisit un comportement différent est non conforme.

Pour toute exigence DEVRAIT et NE DEVRAIT PAS, un routeur peut fournir une option de configuration qui amènera le routeur à agir autrement que ce qui est spécifié par l'exigence. Une telle option de configuration n'est pas un obstacle à la revendication de conformité inconditionnelle d'un routeur si l'option a un réglage par défaut, et si ce réglage amène le routeur à fonctionner de la façon requise.

De même, les routeurs peuvent fournir, excepté lorsque explicitement interdit par le présent mémoire, des options qui l'amènerait à violer des exigences DOIT ou NE DOIT PAS. Un routeur qui fournit de telles options est conforme (pleinement ou conditionnellement) si et seulement si chacune de ces options a un réglage par défaut qui amène le routeur à se conformer aux exigences du présent mémoire. Prière de noter que les auteurs du présent mémoire, bien que conscients des réalités du marché, recommandent fortement de ne pas fournir de telles options. Les exigences sont marquées DOIT ou NE DOIT PAS parce que les experts du domaine les ont jugées particulièrement importantes pour l'interopérabilité ou le fonctionnement approprié dans l'Internet. Les fabricants devraient peser soigneusement les coûts pour le consommateur de la fourniture d'options qui violent ces règles.

Bien sûr, le présent mémoire n'est pas une spécification complète d'un routeur IP, mais est plutôt proche de ce qu'on appelle un profil dans le monde de l'OSI. Par exemple, le présent mémoire exige la mise en œuvre d'un certain nombre de protocoles. Bien que la plus grande partie du contenu de leurs spécifications de protocole ne soit pas répétée dans le présent mémoire, les développeurs sont cependant invités à mettre en œuvre les protocoles conformément à ces spécifications.

1.2 Relations avec d'autres normes

Il est intéressant de vérifier le statut des spécifications et des normes de protocole dans les documents de référence :

- NORMES OFFICIELLES DE PROTOCOLE DE L'INTERNET

Le présent document décrit le processus de normalisation de l'Internet et fait la liste des protocoles du statut de norme. Au moment de la rédaction, la version actuelle du présent document est STD 1, RFC1780, [ARCH:7]. Le présent document est périodiquement republié. Il faut toujours consulter le répertoire des RFC et utiliser la dernière version du présent document.

- Numéros alloués

Le présent document fait la liste des valeurs allouées aux paramètres utilisés dans les divers protocoles. Par exemple, il fait la liste des codes de protocoles IP, des numéros d'accès TCP, des codes d'option Telnet, des types de matériel ARP, et des noms de types de terminaux. Au moment où elle a été rédigée, la version actuelle de ce document est le STD 2, RFC1700, [INTRO:7]. Ce document est périodiquement réédité. Il faut toujours consulter le répertoire des RFC et utiliser la dernière version de ce document.

- Exigences pour les hôtes

Cette paire de documents passe en revue les spécifications qui s'appliquent aux hôtes et donnent des directives et des éclaircissements sur toutes les ambiguïtés. Noter que ces exigences s'appliquent aussi aux routeurs, excepté lorsque le présent mémoire en dispose autrement. Au moment de la rédaction, les versions actuelles de ces documents sont la

RFC 1122 et la RFC1123 (STD 3), [INTRO:2] et [INTRO:3].

- Exigences pour les routeurs (anciennement Exigences pour les passerelles) : Le présent mémoire.

Noter que ces documents sont révisés et mis à jour à des moments différents ; en cas de différences entre ces documents, le plus récent doit avoir la préséance.

Le présent document et les autres protocoles de l'Internet peuvent être obtenus à :
The InterNIC
DS.INTERNIC.NET
InterNIC Directory and Database Service
info@internic.net
+1-908-668-6587
URL: http://ds.internic.net/

1.3 Considérations générales

Plusieurs importantes leçons ont été apprises par les vendeurs de logiciels de l'Internet, qui devraient être sérieusement prises en considération par les nouveaux entrants.

1.3.1 Poursuite de l'évolution de l'Internet

L'énorme croissance de l'Internet a révélé des problèmes de gestion et d'échelle dans les grands systèmes de communication par paquet fondés sur le datagramme. Ces problèmes sont en cours de résolution, et il en résulte que l'évolution des spécifications décrites dans le présent mémoire va se poursuivre. De nouveaux protocoles d'acheminement, des algorithmes, et des architectures sont constamment en cours de développement. Des nouveaux protocoles de couche internet, et des modifications aux protocoles existants sont constamment imaginés. Les routeurs jouent un rôle crucial dans l'Internet, et le nombre des routeurs déployés dans l'Internet est bien plus petit que le nombre des hôtes. Les vendeurs devraient donc s'attendre à ce que les normes de routeurs continuent d'évoluer beaucoup plus vite que les normes d'hôtes. Ces changements seront soigneusement planifiés et contrôlés dans la mesure où à lieu une large participation des vendeurs à cette planification ainsi que des organisations responsables du fonctionnement des réseaux.

Développement, évolution, et révision sont les caractéristiques des protocoles de réseaux d'ordinateurs d'aujourd'hui, et cette situation va persister plusieurs années. Un vendeur qui développe des logiciels de communications d'ordinateur pour la suite des protocoles de l'Internet (ou toute autre suite de protocoles !) et qui manque à assurer la maintenance et la mise à jour de ce logiciel en fonction du changement des spécifications va laisser derrière lui des hordes de consommateurs mécontents. L'Internet est un grand réseau de communications, et les utilisateurs sont en contact constant à travers lui. L'expérience a montré que la connaissance des déficiences du logiciel du vendeur se propage rapidement à travers la communauté technique de l'Internet.

1.3.2 Principe de robustesse

À chaque couche des protocoles, il y a une règle générale (tirée de [TRANS:2] par Jon Postel) dont l'application peut donner d'énormes bénéfices en termes de robustesse et d'interopérabilité :

Soyez conservateurs dans ce que vous faites, et libéral dans ce que vous acceptez des autres.

Le logiciel devrait être écrit pour traiter toute erreur imaginable, quelle que soit sa vraisemblance. Un jour un paquet arrivera dans cette combinaison particulière d'erreurs et d'attributs, et si le logiciel n'y est pas préparé, le chaos peut survenir ; Il est préférable de supposer que le réseau est rempli d'entités malveillantes qui sont prêtes à envoyer des paquets conçus pour avoir le pire effet possible. Cette hypothèse va conduire à des conceptions suffisamment protectrices. Les plus sérieux problèmes de l'Internet ont été causés par des mécanismes imprévus déclanchés par des événements à faible probabilité ; la simple malveillance humaine n'arriverait jamais à suivre des voies aussi tortueuses !

La capacité à s'adapter aux changements doit être conçue à tous les niveaux des logiciels de routeurs. Comme simple exemple, considérons une spécification de protocole qui contient une énumération de valeurs pour un champ d'en-tête particulier – par exemple, un champ de type, un numéro d'accès, ou un code d'erreur ; cette énumération doit être supposée incomplète. Si la spécification de protocole définit quatre codes d'erreur possibles, le logiciel ne doit pas tomber si un cinquième code est défini. Un code indéfini peut être noté, mais il ne doit pas causer une défaillance. La seconde partie du principe est presque aussi importante : Le logiciel sur les hôtes ou autres routeurs peut contenir des déficiences qui rendent peu sage d'exploiter des caractéristiques de protocole légales mais obscures. Il n'est pas raisonnable de s'éloigner de ce qui est évident et simple, de crainte que des effets malencontreux n'en résultent quelque

part ailleurs. Un corollaire de ce principe est de surveiller les hôtes qui se conduisent mal ; les logiciels de routeur devraient être prêts à survivre en présence d'hôtes au mauvais comportement. Une fonction importante des routeurs dans l'Internet est de limiter la quantité de perturbations que de tels hôtes peuvent infliger aux facilités de communications partagées.

1.3.3 Consignation des erreurs

L'Internet comporte une grande variété de systèmes, dont chacun met en œuvre de nombreux protocoles et couches de protocole, et certains d'entre eux contiennent des erreurs et des caractéristiques mal conçues dans leur logiciel de protocole Internet. Il en résulte de la complexité, de la diversité, et une répartition des fonctions, dont le diagnostic des problèmes est souvent très difficile.

Le diagnostic des problèmes sera plus facile si les routeurs comportent des facilités conçues soigneusement pour noter les événements erronés ou étranges. Il est important d'inclure autant d'informations de diagnostic que possible lorsqu'une erreur est notée. En particulier, il est souvent utile d'enregistrer le ou les en-têtes d'un paquet qui a causé une erreur. Cependant, il faut veiller à s'assurer que noter les erreurs ne consomme pas une quantité prohibitive de ressources ou n'interfère pas avec le fonctionnement du routeur.

Il y a une tendance à submerger les fichiers d'enregistrement d'erreurs avec des événements de protocole anormaux mais sans danger ; ceci peut être évité en utilisant un enregistrement circulaire, ou en n'activant l'enregistrement que lors du diagnostic d'une défaillance connue.
Il peut être utile de filtrer et compter les messages dupliqués successifs. Une stratégie qui semble bien fonctionner est de faire à la fois :

- toujours compter les anomalies et rendre un tel compte accessible à travers le protocole de gestion (voir Section 8) ;
- permettre l'activation sélective de la notation d'une grande variété d'événements. Par exemple, il peut être utile d'être capable de tout noter ou de tout noter pour l'hôte X.

Ce sujet sera approfondi dans [MGT:5].

1.3.4 Configuration

Dans un monde idéal, les routeurs seraient faciles à configurer, et peut-être même entièrement auto-configurables. Cependant, l'expérience pratique de la réalité suggère que c'est un but impossible, et que de nombreuses tentatives des fabricants de rendre la configuration facile causent en réalité aux consommateurs plus de maux qu'elles n'en préviennent. Un exemple extrême est celui d'un routeur qui, conçu pour démarrer et commencer à acheminer des paquets sans exiger aucune information de configuration, choisirait certainement des paramètres incorrects, causant éventuellement de sérieux problèmes sur tous les réseaux assez infortunés pour lui être connectés.

Le présent mémoire exige souvent qu'un paramètre soit une option configurable. Il y a plusieurs raisons à cela. Dans quelques cas, il y a actuellement un peu d'incertitude ou des désaccords sur la meilleure valeur et il peut être nécessaire de mettre à jour la valeur qui sera recommandée à l'avenir. Dans d'autres cas, la valeur dépend en fait de facteurs externes – par exemple, la distribution de sa charge de communication, ou la vitesse et la topologie des réseaux voisins – et les algorithmes d'autoréglage ne sont pas disponibles et peuvent être insuffisants. Dans certains cas, la possibilité de configurer est nécessaire à cause d'exigences administratives.

Finalement, certaines options de configuration sont nécessaires pour communiquer avec des mises en œuvre obsolètes ou incorrectes des protocoles, distribués sans source, qui persistent dans de nombreuses parties de l'Internet. Pour faire coexister les systèmes corrects avec ces systèmes déficients, les administrateurs doivent occasionnellement mal configurer les systèmes corrects. Ce problème se corrigera lui-même graduellement si les systèmes déficients sont retirés, mais ne peut être ignoré par les fabricants.

Lorsque nous disons qu'un paramètre doit être configurable, nous ne voulons pas exiger que sa valeur soit lue explicitement à partir d'un fichier de configuration à chaque amorçage. Pour de nombreux paramètres, il y a une valeur qui est appropriée pour toutes les situations sauf les plus inhabituelles. Dans de tels cas, il est assez raisonnable que le paramètre par défaut soit cette valeur si elle n'est pas établie explicitement.

Le présent mémoire exige dans certains cas une valeur particulière par défaut. Le choix d'une valeur par défaut est une question sensible lorsque l'élément de configuration contrôle le réglage des systèmes déficients existants. Si l'Internet doit réussir à converger vers une interopérabilité complète, les valeurs par défaut installées dans les mises en œuvre doivent appliquer le protocole officiel, et non les mauvaises configurations pour s'accommoder des mises en œuvres fautives. Bien que les considérations de commercialisation aient conduit certains vendeurs à choisir par défaut les mauvaises configurations, les fabricants sont instamment priés de choisir des configurations par défaut qui soient conformes aux normes.

Finalement, on note qu'un fabricant a besoin de fournir une documentation adéquate sur tous les paramètres de configuration, sur leurs limites et leurs effets.

1.4 Algorithmes

À différents endroits du présent mémoire sont spécifiés des algorithmes spécifiques qu'un routeur devrait suivre. Ces algorithmes ne sont pas, par eux-mêmes, exigés du routeur. Un routeur n'a pas besoins de mettre en œuvre chaque algorithme décrit dans le présent document. Une mise en œuvre doit plutôt présenter au monde extérieur un comportement qui soit le même que celui d'une mise en œuvre stricte et littérale de l'algorithme spécifié.

Les algorithmes sont décrits d'une manière qui diffère de la façon dont une bonne mise en œuvre les appliquerait. Pour les besoins de la présentation, il a été choisi un style qui privilégie la concision, la clarté et l'indépendance à l'égard des détails de mise en œuvre. Un bon développeur choisira les algorithmes et les méthodes de mise en œuvre qui produisent les mêmes résultats que ces algorithmes, mais peuvent être plus efficaces ou moins généraux.

On notera que l'art de la mise en œuvre du routeur efficace sort du domaine d'application du présent mémoire.

2 Architecture internet

Cette section ne contient aucune exigence. Cependant, elle contient des informations de base utilisées sur l'architecture générale de l'Internet et des routeurs.

On trouvera les fondements et l'exposé sur l'architecture Internet et les suites de protocole de prise en charge dans le Guide du protocole DDN *(DDN Protocol Handbook)* [ARCH:1] ; pour les fondements, voir par exemple [ARCH:2], [ARCH:3], et [ARCH:4]. L'architecture et les protocoles de l'Internet sont aussi traités dans un nombre toujours croissant de manuels, tels que [ARCH:5] et [ARCH:6].

2.1 Introduction

Le système Internet consiste en un certain nombre de réseaux de paquets interconnectés qui prennent en charge les communications entre les ordinateurs hôtes en utilisant les protocoles de l'Internet. Ces protocoles incluent le protocole Internet (IP), le protocole de messages de commande de l'Internet (ICMP), le protocole de gestion de groupe de l'Internet (IGMP), et divers protocoles de transport et d'application qui en dépendent. Comme exposé au paragraphe 1.2, le groupe de pilotage de l'ingénierie de l'Internet publie périodiquement un mémoire officiel des protocoles qui fait la liste de tous les protocoles de l'Internet.

Tous les protocoles Internet utilisent IP comme mécanisme de base de transport de données. IP est un service inter-réseaux de datagrammes, ou sans connexion, qui inclut des dispositions pour l'adressage, la spécification du type de service, la fragmentation et le réassemblage, et la sécurité. ICMP et IGMP sont considérés comme des parties intégrantes de IP, bien qu'ils appartiennent à une couche d'architecture au dessus de IP. ICMP fournit les rapports d'erreurs, les contrôles des flux, la redirection du routeur du premier bond, et d'autres fonctions de maintenance et de contrôle. IGMP fournit les mécanismes par lesquels les hôtes et les routeurs peuvent se joindre et quitter des groupes de diffusion groupée IP.

La livraison fiable des données est fournie dans la suite des protocoles de l'Internet par les protocoles de la couche Transport tels que le protocole de commande de transmission (TCP), qui fournit la retransmission de bout en bout, le re-séquençage et le contrôle de connexion. Le service de couche Transport sans connexion est fourni par le protocole de datagrammes d'utilisateur (UDP).

2.2 Éléments de l'architecture

2.2.1 Mise en couche des protocoles

Pour communiquer en utilisant le système Internet, un hôte doit mettre en œuvre l'ensemble en couches de protocoles qui comprend la suite des protocoles de l'Internet. Normalement, un hôte doit mettre en œuvre au moins un protocole de chaque couche.

Les couches de protocole utilisées dans l'architecture Internet sont comme suit [ARCH:7] :

- Couche Application

La couche Application est la couche supérieure de la suite des protocoles Internet. La suite des protocoles Internet ne fait pas d'autres subdivisions de la couche Application, bien que certains protocoles de couche d'application contiennent effectivement certaines sous-couches internes. La couche application de la suite Internet combine essentiellement des fonctions des deux couches supérieures - Présentation et Application – du modèle de référence OSI [ARCH:8]. La couche Application dans la suite de protocole Internet inclut aussi certaines des fonctions reléguées à la couche Session dans le modèle de référence OSI.

On distingue deux catégories de protocoles de couche d'application : les protocoles d'utilisateur qui fournissent le service directement à l'utilisateur, et les protocoles de soutien qui fournissent des fonctions de système communes. Les protocoles d'utilisateur Internet les plus courants sont :
- Telnet (connexion distante)
- FTP (transfert de fichiers)
- SMTP (messagerie électronique)

Il y a de nombreux autres protocoles d'utilisateur normalisés et de nombreux protocoles d'utilisateur privés.

Les protocoles de soutien, utilisés pour la transposition de nom d'hôte, l'amorçage et la gestion incluent SNMP, BOOTP, TFTP, le protocole du système des noms de domaine (DNS), et divers protocoles d'acheminement.

Les protocoles de la couche Application qui sont pertinents pour les routeurs sont exposés aux Sections 7, 8, et 9 du présent mémoire.

- Couche Transport

La couche Transport fournit des services de communication de bout en bout. Cette couche est en gros équivalente à la couche Transport dans le modèle de référence OSI, excepté qu'elle incorpore aussi des fonctions d'établissement et de destruction de la couche Session OSI.

Il y a présentement deux principaux protocoles de la couche Transport :
- Protocole de commande de transmission (TCP)
- Protocole de datagramme d'utilisateur (UDP)

TCP est un service de transport fiable orienté connexion qui fournit de la fiabilité, du reséquençage et du contrôle de flux de bout en bout. UDP est un service de transport sans connexion (datagrammes). D'autres protocoles de transport ont été développés par la communauté de la recherche, et l'ensemble des protocoles de transport officiels de l'Internet sera élargi à l'avenir.

Les protocoles de la couche Transport pertinents pour les routeurs sont exposés à la section 6.

- Couche Internet

Tous les protocoles de transport Internet utilisent le protocole Internet (IP) pour transporter les données de l'hôte source à l'hôte de destination. IP est un service inter-réseaux sans connexion ou de datagrammes, qui fournit des garanties de livraison, de bout en bout. Les datagrammes IP peuvent arriver endommagés à leur hôte de destination, ou dupliqués, ou déclassés, ou pas du tout. Les couches au-dessus d'IP sont responsables du service de livraison fiable lorsqu'il est exigé. Le protocole IP inclut des dispositions pour l'adressage, la spécification du type de service, la fragmentation et le réassemblage, et la sécurité.

La nature sans connexion ou par datagrammes de IP est une caractéristique fondamentale de l'architecture de l'Internet.

Le protocole de message de commande de l'Internet (ICMP, *Internet Control Message Protocol*) est un protocole de commande qui est considéré comme faisant partie intégrante de IP, bien qu'il soit architecturalement dans la couche au-dessus d'IP - il utilise IP pour transporter ses données de bout en bout. ICMP fournit les rapports d'erreurs, d'encombrement, et de redirection de routeur de premier bond.

Le protocole de gestion de groupe de l'Internet (IGMP, *Internet Group Management Protocol*) est un protocole de couche Internet utilisé pour établir des groupes d'hôtes dynamiques pour la diffusion groupée IP.

Les protocoles de couche Internet IP, ICMP, et IGMP sont exposés à la section 4.

- Couche de liaison des données

Pour communiquer sur un réseau directement connecté, un hôte doit mettre en œuvre le protocole de communication utilisé pour s'interfacer avec ce réseau. On appelle cela un protocole de couche de liaison.

Certains plus anciens documents de l'Internet se réfèrent à cette couche sous le nom de couche Réseau, mais ce n'est pas la même que la couche réseau du modèle de référence OSI.

Cette couche contient tout ce qui se trouve au-dessous de la couche Internet et au-dessus de la couche Physique (qui est la couche de la connectivité avec le support, normalement électrique ou optique, qui code et transporte les messages). Sa responsabilité est la livraison correcte des messages, entre lesquels elle ne fait pas de différence.

Les protocoles de cette couche sortent normalement du domaine d'application de la normalisation de l'Internet ; l'Internet utilise (intentionnellement) les normes existantes chaque fois que possible. Et donc, les normes de couche de liaison de l'Internet ne s'occupent habituellement que de résolution d'adresse et des règles de transmission des paquets IP sur des protocoles spécifiques de la couche de liaison. Les normes Internet de la couche de liaison sont exposées à la section 3.

2.2.2 Réseaux

Les réseaux constitutifs du système Internet sont tenus de ne fournir que du transport de paquet (sans connexion). Conformément à la spécification de service IP, les datagrammes peuvent être livrés dans le désordre, être perdus ou dupliqués, et/ou contenir des erreurs.

Pour une qualité de service raisonnable des protocoles qui utilisent IP (par exemple, TCP), le taux de perte du réseau devrait être très faible. Dans les réseaux qui fournissent des services orientés connexion, la fiabilité supplémentaire apportée par les circuits virtuels améliore la robustesse de bout en bout du système, mais n'est pas nécessaire pour le fonctionnement de l'Internet.

Les réseaux constitutifs peuvent généralement être divisés en deux classes :

- Réseaux de zone locale (LAN)
  Les LAN peuvent avoir des conceptions diverses. Ils couvrent normalement une petite zone géographique (par exemple, un seul bâtiment ou site) et fournissent une bande passante élevée avec de faibles délais. Les LAN peuvent être passifs (comme Ethernet) ou actifs (comme ATM).
- Réseaux de grande zone (WAN)
  Des hôtes géographiquement dispersés et des LAN sont interconnectés par des réseaux de grande zone, appelés aussi réseaux à longue portée. Ces réseaux peuvent avoir une structure interne complexe de lignes et de commutateurs de paquets, ou ils peuvent n'être que de simples lignes point à point.

2.2.3 Routeurs

Dans le modèle Internet, les réseaux constitutifs sont connectés ensemble par des transmetteurs de datagrammes IP qu'on appelle routeurs ou routeurs IP. Dans le présent document, chaque utilisation du terme routeur est équivalente à routeur IP. De nombreux documents plus anciens de l'Internet se réfèrent aux routeurs sous le nom de passerelles.

Historiquement, les routeurs ont été réalisés avec un logiciel de commutation de paquets s'exécutant sur un CPU d'usage général. Cependant, comme le développement de matériel personnalisé est devenu meilleur marché et qu'un débit plus élevé est nécessaire, le matériel spécialisé est devenu de plus en plus courant. La présente spécification s'applique aux routeurs indépendamment de la façon dont ils sont mis en œuvre.

Un routeur se connecte à deux interfaces logique ou plus, représentées par des sous-réseaux IP ou un nombre illimité de liaisons point à point (exposé au paragraphe 2.2.7). Et donc, il a au moins une interface physique. La transmission d'un datagramme IP exige généralement que le routeur choisisse l'adresse et l'interface pertinentes du routeur du prochain bond (ou du bond final) de l'hôte de destination. Ce choix, qu'on appelle relayage ou transmission, dépend d'une base de données d'acheminements au sein du routeur. La base de données d'acheminements est aussi appelée un tableau d'acheminements ou tableau de transmissions. Le terme "routeur" dérive du processus de construction de cette base de données de routage ; les protocoles d'acheminement et la configuration interagissent dans un processus appelé acheminement,

La base de données d'acheminement devrait faire l'objet d'une maintenance dynamique pour refléter la topologie actuelle du système Internet. Un routeur accomplit normalement cela en participant à un acheminement distribué et avec des algorithmes d'accessibilité avec les autres routeurs.

Les routeurs ne fournissent que le transport de datagrammes, et on cherche à minimiser les informations d'état nécessaires à ce service dans l'intérêt de la souplesse et de la robustesse de l'acheminement.

Les appareils de commutation de paquet peuvent aussi fonctionner à la couche de liaison ; de tels appareils sont normalement appelés des ponts. Les segments de réseau qui sont connectés par des ponts partagent le même préfixe de réseau IP qui forme un seul sous-réseau IP. Ces autres appareils sortent du domaine d'application du présent document.

2.2.4 Systèmes autonomes

Un système autonome (AS) est un segment connecté d'une topologie de réseau qui consiste en une collection de sous-réseaux (avec les hôtes rattachés) interconnectés par un ensemble de routes. Les sous-réseaux et les routeurs sont supposés être sous le contrôle d'une seule organisation d'opérations et maintenance (O&M). Au sein d'un AS, les routeurs peuvent utiliser un ou plusieurs protocoles d'acheminement intérieurs, et parfois, plusieurs ensembles de mesures. Un AS est supposé présenter aux autres AS l'apparence d'un plan d'acheminement intérieur cohérent, et une image cohérente des destinations atteignables à travers l'AS. Un AS est identifié par un numéro de système autonome.

Le concept d'AS jour un rôle important dans l'acheminement de l'Internet (voir au paragraphe 7.1).

2.2.5 Architecture d'adressage

Un datagramme IP porte des adresses de source et de destination de 32 bits, dont chacune est divisée en deux parties – un préfixe de réseau constitutif et un numéro d'hôte sur ce réseau. Symboliquement :

adresse-IP ::= { <préfixe-réseau>, <numéro-d'hôte> }

Pour livrer finalement le datagramme, le dernier routeur sur son chemin doit transposer la partie numéro d'hôte (ou reste) d'une adresse IP en adresse de couche liaison de l'hôte.

2.2.5.1 Architecture classique d'adressage IP

Bien qu'il soit bien documenté par ailleurs [INTER:2], il est utile de décrire l'histoire de l'utilisation du préfixe de réseau. Le langage développé pour le décrire est utilisé dans le présent document et dans d'autres et imprègne les idées sous-jacentes de nombreux protocoles.

Le préfixe de réseau classique le plus simple est le préfixe de réseau de classe A, B, C, D, ou E. Ces gammes d'adresses sont distinguées en observant les valeurs des bits de plus fort poids de l'adresse, et en cassant l'adresse en champs de simple préfixe et de numéro d'hôte. Ceci est décrit dans [INTER:18]. En bref, la classification est :

0xxx - Classe A – adresses en envoi individuel à usage général avec préfixe standard de 8 bits
10xx - Classe B – adresses en envoi individuel à usage général avec préfixe standard de 16 bits
110x - Classe C – adresses en envoi individuel à usage général avec préfixe standard de 24 bits
1110 - Classe D – adresses IP en envoi groupé – préfixe de 28 bits, non agrégeable
1111 - Classe E – réservée pour utilisation expérimentale

Cette notion simple a été étendue par le concept de sous-réseaux. Ils ont été introduits pour permettre une complexité arbitraire des structures de LAN interconnectées au sein d'une organisation, tout en isolant le système Internet de la croissance explosive des préfixes de réseau alloués et de la complexité de l'acheminement. Les sous-réseaux fournissent une structure d'acheminement hiérarchique à plusieurs niveaux pour le système Internet. L'extension de sous-réseau, décrite dans [INTER:2], est une partie nécessaire de l'architecture de l'Internet. L'idée de base est de partager le champ <numéro-d'hôte> en deux parties : un numéro de sous-réseau, et un vrai numéro d'hôte sur ce sous-réseau :

adresse-IP ::= { <numéro-de-réseau>, <numéro-de-sous-réseau>, <numéro-d'hôte> }

Les réseaux physiques interconnectés au sein d'une organisation utilisent le même préfixe de réseau mais des numéros de sous-réseau différents. La distinction entre les sous-réseaux d'un tel réseau n'est pas normalement visible à l'extérieur de ce réseau. Et donc, l'acheminement dans le reste de l'Internet n'utilise que la partie <préfixe-réseau> de l'adresse de destination IP. Les routeurs en-dehors du réseau traitent ensemble <préfixe-réseau> et <numéro-d'hôte> comme une partie restante non interprétée de l'adresse IP de 32 bits. Au sein du réseau subdivisé en sous-réseaux, les routeurs utilisent le préfixe de réseau étendu :

{ <numéro-de-réseau>, <numéro-de-sous-réseau> }

Les positions binaires qui contiennent ce numéro de réseau étendu ont dans le passé été indiquées par un gabarit de 32 bits appelé le gabarit de sous-réseau. Les bits du <numéro-de-sous-réseau> DEVRAIENT être contigus et tomber entre les champs de <numéro-de-réseau> et de <numéro-d'hôte>. Les protocoles plus à jour ne se réfèrent pas au gabarit de sous-réseau, mais à une longueur de préfixe ; la portion "préfixe" d'une adresse est ce qui serait choisi par un gabarit de sous-réseau dont les bits de plus fort poids sont tous des uns et le reste sont des zéros. La longueur du préfixe est égale au nombre de uns dans le gabarit de sous-réseau. Le présent document suppose que tout gabarit de sous-réseau peut s'exprimer par des longueurs de préfixes.

Les inventeurs du mécanisme de sous-réseau présumaient que chaque morceau du réseau d'une organisation aurait seulement un seul numéro de sous-réseau. En pratique, il s'est souvent révélé nécessaire ou utile que plusieurs sous-

réseaux partagent un seul câble physique. Pour cette raison, les routeurs devraient être capables de configurer plusieurs sous-réseaux sur les mêmes interfaces physiques, et de les traiter (du point de vue de l'acheminement ou de la transmission) comme si il y avait des interfaces physiques distinctes.

2.2.5.2 Routage inter domaine sans classe

La croissance explosive de l'Internet a forcé à réviser les politiques d'allocation d'adresses. L'utilisation traditionnelle des réseaux d'utilisation générale (Classes A, B, et C) a été modifiée pour obtenir une meilleure utilisation de l'espace d'adresse de 32 bits d'IP. L'acheminement inter domaine sans classe (CIDR, *Classless Inter Domain Routing*) [INTER:15] est une méthode actuellement déployée dans les dorsales de l'Internet pour réaliser cette efficacité accrue. Le CIDR dépend du déploiement et de l'acheminement sur des réseaux dimensionnés arbitrairement. Dans ce modèle, les hôtes et les routeurs ne font pas d'hypothèses sur l'utilisation de l'adressage dans l'internet. Les espaces d'adresse de Classe D (diffusion groupée IP) et de Classe E (expérimental) sont préservés, bien que ceci soit principalement une politique d'allocation.

Par définition, le CIDR comprend trois éléments :
- une allocation d'adresse topologiquement significative,
- des protocoles d'acheminement capables d'agréger les informations d'accessibilité de couche réseau,
- un algorithme de transmission cohérent ("à la plus longue correspondance").

L'utilisation des réseaux et des sous-réseaux est maintenant dépassée, bien que le langage utilisé pour les décrire reste dans l'usage courant. Ils ont été remplacés par le concept plus maniable de préfixe de réseau. Un préfixe de réseau est, par définition, un ensemble contigu de bits à l'extrémité de plus fort poids de l'adresse qui définit un ensemble de systèmes ; les numéros d'hôtes choisissent parmi ces systèmes. Il n'y a pas d'exigence que tout l'internet utilise uniformément des préfixes de réseau. Pour dégonfler de volume des informations d'acheminement, il est utile de diviser l'internet en domaines d'adressage. Au sein d'un tel domaine, des informations détaillées sont disponibles sur les réseaux constitutifs ; en-dehors de lui, seul le préfixe de réseau commun est publié.

L'architecture classique d'adressage IP utilisait les adresses et les gabarits de sous-réseau pour distinguer le numéro d'hôte du préfixe de réseau. Avec les préfixes de réseau, il est suffisant d'indiquer le nombre de bits dans le préfixe. Les deux représentations sont d'utilisation courante. Les gabarits de sous-réseau architecturalement corrects sont capables de se représenter à l'aide de la description de longueur de préfixe. Ils comprennent un sous-ensemble de tous les schémas binaires possibles qui ont :
- une chaîne contiguë de uns à l'extrémité de plus fort poids,
- une chaîne contiguë de zéros à l'extrémité de plus faible poids,
- pas de bits interposés.

Les routeurs DEVRAIENT toujours traiter un acheminement comme un préfixe de réseau, et DEVRAIENT rejeter les informations de configuration et d'acheminement incohérentes avec ce modèle.

adresse-IP ::= { <préfixe-réseau>, <numéro-d'hôte> }

Un effet de l'utilisation de CIDR est que l'ensemble de destinations associé aux préfixes d'adresse dans le tableau d'acheminement peut afficher des relations de sous-ensembles. Un acheminement décrivant un plus petit ensemble de destinations (un préfixe plus long) est dit être plus spécifique qu'un acheminement décrivant un plus grand ensemble de destinations (un préfixe plus court) ; de même, un acheminement décrivant un plus grand ensemble de destinations (un préfixe plus court) est dit être moins spécifique qu'un acheminement décrivant un plus petit ensemble de destinations (un plus long préfixe). Les routeurs doivent utiliser l'acheminement correspondant le plus spécifique (le préfixe de réseau correspondant le plus long) lorsqu'ils transmettent le trafic.

2.2.6 Diffusion groupée sur IP

La diffusion groupée sur IP est une extension de la diffusion groupée de couche de liaison aux internets IP. En utilisant les diffusions groupées IP, un seul datagramme peut être adressé à plusieurs hôtes sans l'envoyer à tous. Dans le cas étendu, ces hôtes peuvent résider dans des domaines d'adresse différents. Cette collection d'hôtes s'appelle un groupe de diffusion groupée *(multicast group)*. Chaque groupe de diffusion est représenté comme une adresse IP de classe D. Un datagramme IP envoyé au groupe sera délivré à chaque membre du groupe avec la même livraison au mieux que celle fournie pour le trafic IP en envoi individuel. L'envoyeur du datagramme n'a pas besoin d'être lui-même membre du groupe de destination.
La sémantique de l'adhésion à un groupe de diffusion IP est définie dans [INTER:4]. Ce document décrit comment les hôtes et les routeurs rejoignent et quittent les groupes de diffusion. Il définit aussi un protocole, le protocole de gestion de groupe Internet (IGMP, *Internet Group Management Protocol*), qui surveille l'adhésion au groupe de diffusion IP.

La transmission des datagrammes de diffusion groupée IP est accomplie soit par des informations d'acheminement statiques, soit via un protocole d'acheminement de diffusion groupée. Les appareils qui transmettent des datagrammes de diffusion groupée s'appellent des routeurs de diffusion groupée. Ils peuvent ou non transmettre aussi de l'envoi individuel IP. Les datagrammes de diffusion groupée sont transmis sur la base à la fois de leur adresse de source et de destination. La transmission des paquets de diffusion groupée IP est décrite plus en détails au paragraphe 5.2.1. L'appendice D discute des protocoles d'acheminement en diffusion groupée.

2.2.7 Lignes non numérotées et préfixes de réseau

Traditionnellement, chaque interface de réseau sur un hôte ou routeur IP a sa propre adresse IP. Ceci peut être la cause d'une utilisation inefficace de l'espace d'adresse IP, car cela force l'allocation d'un préfixe de réseau IP à chaque liaison point à point.

Pour résoudre ce problème, un certain nombre de gens ont proposé et mis en œuvre le concept de lignes point à point non numérotées. Une ligne en point à point non numérotée n'a pas de préfixe de réseau associé. Il en résulte que les interfaces de réseau connectées à une ligne en point à point non numérotée n'ont pas d'adresse IP.

Comme l'architecture IP a traditionnellement supposé que toutes les interfaces avaient des adresses IP, ces interfaces non numérotées causent quelques dilemmes intéressants. Par exemple, certaines options IP (par exemple, Record Route) spécifient qu'un routeur doit insérer l'adresse de l'interface dans l'option, mais une interface non numérotée n'a pas d'adresse IP. Encore plus fondamental (comme nous le verrons à la section 5) est que les routages contiennent l'adresse IP du routeur du prochain bond. Un routeur s'attend à ce que cette adresse IP soit sur un réseau ou sous-réseau IP auquel le routeur est connecté. Cette hypothèse est bien sûr violée si la seule connexion est une ligne point à point non numérotée.

Pour contourner ces difficultés, deux schémas ont été conçus. Le premier dit que deux routeurs connectés par une ligne point à point non numérotée ne sont pas réellement des routeurs, mais plutôt deux demi routeurs qui ensemble forment un seul routeur virtuel. La ligne point à point non numérotée est essentiellement considérée comme un bus interne dans le routeur virtuel. Les deux moitiés du routeur virtuel doivent coordonner leurs activités de telle sorte qu'elles agissent exactement comme un seul routeur.

Ce schéma va bien dans l'architecture IP, mais souffre de deux inconvénients importants. Le premier est que, bien qu'il traite le cas courant d'une seule ligne point à point non numérotée, il n'est pas extensible au traitement du cas d'un mélange de routeurs et de lignes point à point non numérotées. Le second inconvénient est que les interactions entre les demi routeurs sont nécessairement complexes et qu'elles ne sont pas normalisées, empêchant effectivement la connexion d'équipements de fabricants différents en utilisant des lignes point à point non numérotées.

A cause de ces inconvénients, le présent mémoire a adopté un schéma de remplacement, qui a été inventé de nombreuses fois mais qui peut probablement être attribuable à l'origine à Phil Karn. Dans ce schéma, un routeur qui a des lignes point à point non numérotées a aussi une adresse IP spéciale, appelée identifiant de routeur (routeur-id) dans le présent mémoire. Le routeur-id est une des adresses IP du routeur (un routeur est tenu d'avoir au moins une adresse IP). Ce routeur-id est utilisé comme si c'était l'adresse IP de toutes les interfaces non numérotées.

2.2.8 Particularités remarquables

2.2.8.1 Routeurs incorporés

Un routeur peut être un système informatique autonome, dédié à ses fonctions de routeur IP. Autrement, il est possible d'incorporer les fonctions de routeur au sein d'un système d'exploitation d'hôte qui prend en charge les connexions avec deux réseaux ou plus. L'exemple le plus connu d'un système d'exploitation avec un code de routeur incorporé est le système BSD de Berkeley. Le dispositif de routeur incorporé semble rendre facile la construction d'un réseau, mais il y a un certain nombre de pièges cachés :

(1) Si un hôte a seulement une interface constitutive de réseau, il ne devrait pas agir comme un routeur. Par exemple, les hôtes avec un code de routeur qui transmet gratuitement les paquets ou datagrammes en diffusion sur le même réseau causent souvent des avalanches de paquets.

(2) Si un hôte (à rattachements multiples) agit comme un routeur, il est soumis aux exigences pour les routeurs contenues dans le présent document. Par exemple, les questions de protocole d'acheminement et les problèmes de commande et de surveillance du routeur sont difficiles et importants pour les routeurs incorporés comme pour les routeurs autonomes. Les exigences et spécifications de routeur Internet peuvent changer indépendamment des changements du système d'exploitation. Il est fortement conseillé à une administration qui fait fonctionner un routeur incorporé dans l'Internet d'assurer la maintenance et la mise à jour du code de routeur. Ceci peut exiger le code source du routeur.

(3) Lorsque un hôte exécute un code de routeur incorporé, il s'intègre à l'infrastructure de l'Internet. Et donc, les

erreurs de logiciel ou de configuration peuvent gêner les communications entre d'autres hôtes. Par conséquent, l'administrateur d'hôte doit perdre un peu d'autonomie. Dans de nombreuses circonstances, un administrateur d'hôte aura besoin de désactiver le code de routeur incorporé dans le système d'exploitation. Pour cette raison, il devrait pouvoir désactiver directement la fonction de routeur incorporé.

(4) Lorsque un hôte qui fait fonctionner un code de routeur incorporé est utilisé concurremment pour d'autres services, les exigences d'opération et maintenance pour les deux modes d'utilisation peuvent entrer en conflit. Par exemple, l'O&M du routeur sera dans de nombreux cas effectuée à distance par un centre d'opérations ; cela peut exiger un accès système privilégié que l'administrateur de l'hôte voudrait normalement ne pas voir divulgué.

2.2.8.2 Routeurs transparents

Il y a deux modèles de base pour interconnecter les réseaux de zone locale et de zone large (ou à longue portée) dans l'Internet. Dans le premier, le réseau de zone locale se voit allouer un préfixe de réseau et tous les routeurs de l'Internet doivent savoir comment acheminer vers ce réseau. Dans le second, le réseau de zone locale partage (une petite partie de) l'espace d'adresse du réseau de large zone. Les routeurs qui prennent en charge ce second modèle sont appelés des routeurs à partage d'adresse ou des routeurs transparents. Le présent mémoire se concentre sur les routeurs qui prennent en charge le premier modèle, mais il n'est pas prévu d'exclure l'utilisation des routeurs transparents.

L'idée de base d'un routeur transparent est que les hôtes sur le réseau de zone locale derrière un tel routeur partagent l'espace d'adresse du réseau de zone large en front du routeur. Dans certaines situations, c'est une approche très utile et les limitations ne présentent pas d'inconvénients significatifs.

Les termes en front et derrière indiquent une des limitations de cette approche : ce modèle d'interconnexion ne convient que pour un environnement géographiquement (et topologiquement) limité. Il exige qu'il y ait une forme d'adressage logique dans l'adressage de niveau réseau du réseau de zone large. Les adresses IP dans l'environnement local se transposent en peu d'adresses (habituellement une) physiques dans le réseau de zone large. Cette transposition survient d'une façon cohérente avec la transposition { adresse IP <-> adresse réseau } utilisée dans le réseau de zone large.

L'origine multiple est possible sur un réseau de zone large, mais peut présenter des problèmes d'acheminement si les interfaces sont géographiquement ou topologiquement séparées. L'origine multiple sur deux (ou plus) réseaux de zone large pose un problème dû à la confusion des adresses.

Les comportements que les hôtes voient chez les autres hôtes dans ce qui est apparemment le même réseau peut être différent si le routeur transparent ne peut pas pleinement émuler le service normal de réseau de zone large. Par exemple, ARPANET utilisait un protocole de couche de liaison qui fournissait une indication Destination morte en réponse à une tentative d'envoi sur un hôte qui n'est pas sous tension. Cependant, si il y avait eu un routeur transparent entre ARPANET et un Ethernet, un hôte sur ARPANET ne recevrait pas d'indication Destination morte pour les hôtes Ethernet.

2.3 Caractéristiques de routeur

Un routeur Internet remplit les fonctions suivantes :

(1) Se conformer aux protocoles spécifiques de l'Internet spécifiés dans le présent document, y compris le protocole Internet (IP), le protocole de message de commande de l'Internet (ICMP), et d'autres, en tant que de besoin.

(2) Assurer l'interface avec deux réseaux de paquets ou plus. Pour chaque réseau connecté, le routeur doit assurer les fonctions exigées par ce réseau. Ces fonctions comportent normalement :
- l'encapsulation et la désencapsulation des datagrammes IP avec le tramage du réseau connecté (par exemple, un en-tête Ethernet et la somme de contrôle),
- L'envoi et la réception des datagrammes IP jusqu'à la taille maximale prise en charge par ce réseau, cette taille est l'unité maximale de transmission du réseau ou MTU,
- Traduire l'adresse de destination IP en une adresse de niveau réseau appropriée pour le réseau connecté (par exemple, une adresse de matériel Ethernet), si nécessaire, et
- Répondre aux indications de commande et d'erreur de flux du réseau, s'il en est.
  Voir à la section 3 (couche de liaison).

(3) Recevoir et transmettre les datagrammes Internet. Les questions importantes dans ce processus sont la gestion de mémoire tampon, le contrôle de l'encombrement, et la loyauté.
- Reconnaître les conditions d'erreur et générer des messages ICMP d'erreur et d'information en tant que de besoin,
- Détruire les datagrammes dont les champs de durée de vie ont atteint zéro.

- Fragmenter lorsque nécessaire les datagrammes pour coller à la MTU du prochain réseau.
  Voir à la section 4 (Couche Internet - Protocoles) et à la section 5 (Couche Internet - Transmission) pour des précisions.

(4) Choisir une destination de prochain bond pour chaque datagramme IP, sur la base des informations de sa base de données d'acheminement. Voir à la section 5 (Couche Internet - Transmission) pour des précisions.

(5) (Habituellement) prendre en charge un protocole de passerelle intérieure (IGP) pour résoudre les algorithmes d'acheminement et d'accessibilité distribués avec les autres routeurs dans le même système autonome. De plus, certains routeurs auront besoin de prendre en charge un protocole de passerelle extérieure (EGP) pour échanger des informations topologiques avec d'autres systèmes autonomes. Voir à la section 7 (Couche d'application – Protocoles d'acheminement) pour des précisions.

(6) Fournir des facilités de gestion de réseau et de prise en charge de système, incluant le chargement, le débogage, les rapports d'état, les rapports et le contrôle de situations d'exception. Voir à la section 8 (Couche d'application – Protocoles de gestion de réseau) et à la section 10 (Opérations et maintenance) pour des précisions.

Un fabricant de routeur aura de nombreux choix sur l'alimentation, la complexité, et les caractéristiques d'un produit de routeur particulier. Il peut être utile d'observer que le système Internet n'est ni homogène ni pleinement connecté. Pour des raisons technologiques et géographiques, il croît dans un système interconnecté mondialement plus une frange de LAN sur le pourtour. De plus en plus de ces LAN excentrés deviennent richement interconnectés, les rendant de moins en moins marginaux et de plus en plus exigeants envers les routeurs.

- Le système mondial d'interconnexion se compose d'un certain nombre de réseaux de zone large auxquels sont rattachés les routeurs de plusieurs systèmes autonomes (AS) ; il y a relativement peu d'hôtes connectés directement au système.

- La plupart des hôtes sont connectés à des LAN. De nombreuses organisations ont des grappes de LAN interconnectées par des routeurs locaux. Chacune de ces grappes est connectée par les routeurs à un ou plusieurs points dans le système mondial d'interconnexion. Si il est connecté en un seul point, un LAN est appelé un réseau d'amorce.

Dans le système mondial d'interconnexion, les routeurs ont besoin :
- d'algorithmes évolués d'acheminement et de transmission.
  Ces routeurs ont besoin d'algorithmes d'acheminement qui soient très dynamiques, imposent une charge minimale de traitement et de communication, et qui offrent un acheminement par type de service. L'encombrement est une question qui n'est toujours pas complètement résolue (voir au paragraphe 5.3.6). Des améliorations sont espérées dans ce domaine, car la recherche travaille activement sur ces questions.

- Forte disponibilité
  Ces routeurs doivent être très fiables, pour servir 24 heures sur 24, 7 jours sur 7. Les défaillances d'équipement et de logiciel peuvent avoir des conséquences très étendues (parfois mondiales). En cas de défaillance, la récupération doit être rapide. Dans tout environnement, un routeur doit être très robuste et capable de fonctionner, éventuellement en état dégradé, dans des conditions d'extrême encombrement ou de défaillance des ressources du réseau.

- Caractéristiques O&M évoluées
  Les routeurs Internet fonctionnent normalement en mode sans surveillance. Ils vont normalement fonctionner à distance à partir d'un centre de surveillance centralisé. Il est nécessaire de fournir des moyens sophistiqués de surveillance et de mesure du trafic et autres événements, et pour le diagnostic des fautes.

- Hautes performances
  Les lignes à longue portée dans l'Internet d'aujourd'hui sont le plus fréquemment du 56 kbit/s en duplex, du DS1 (1,544 Mbit/s), ou du DS3 (45 Mbit/s). Les LAN, qui sont un support multi accès semi duplex, sont normalement de l'Ethernet (10 Mbit/s) et, dans une moindre mesure, du FDDI (100 Mbit/s). Cependant, la technologie des supports de réseau est en progrès constant et des vitesses supérieures sont attendues à l'avenir.

Les exigences pour les routeurs utilisés dans la frange de LAN (par exemple, les réseaux de campus) dépendent largement des demandes sur les réseaux locaux. Ce peuvent être des appareils à hautes ou moyennes performance, probablement acquis en faisant jouer la concurrence entre plusieurs fabricants différents et mis en œuvre par une organisation interne (par exemple, un centre informatique de campus). La conception de ces routeurs devrait mettre l'accent sur un délai de latence moyen et de bonnes performances de crête, jointes à une gestion de ressources sensible au délai et au type de service. Dans cet environnement, il peut y avoir moins d'O&M formelle mais elle n'en sera pas moins importante. Le besoin d'un mécanisme d'acheminement très dynamique deviendra plus important lorsque les

réseaux deviennent plus complexes et interconnectés. Les utilisateurs demanderont plus à leurs connexions locales à cause de la vitesse de l'interconnexion mondiale.

Avec la croissance des réseaux, et comme plus de réseaux sont devenus assez anciens pour éliminer leurs plus vieux équipements, il est devenu impératif que les routeurs interopèrent avec les routeurs d'autres fabricants.

Bien que le système Internet ne soit pas complètement interconnecté, de nombreuses parties du système doivent avoir une connectivité redondante. Une connectivité riche permet un service fiable en dépit des défaillances des lignes de communication et des routeurs, et elle peut aussi améliorer le service en raccourcissant les chemins de l'Internet et en fournissant des capacités supplémentaires. Malheureusement, cette topologie plus riche peut rendre plus difficile le choix du meilleur chemin vers une destination particulière.

2.4 Hypothèses architecturales

L'architecture actuelle de l'Internet se fonde sur un ensemble d'hypothèses sur le système de communication. Les hypothèses les plus pertinentes sur les routeurs sont les suivantes :

- L'Internet est un réseau de réseaux.
  Chaque hôte est directement connecté à un ou des réseaux particuliers ; sa connexion à l'Internet est seulement conceptuelle. Deux hôtes sur le même réseau communiquent l'un avec l'autre en utilisant le même ensemble de protocoles qu'ils utiliseraient pour communiquer avec des hôtes sur des réseaux distants.
- Les routeurs ne conservent pas les informations d'état de connexion.

  Pour améliorer la robustesse du système de communication, les routeurs sont conçus sans état, et ils transmettent chaque paquet IP indépendamment des autres paquets. Il en résulte que des chemins redondants peuvent être exploités pour fournir un service robuste en dépit des défaillances des routeurs et réseaux intermédiaires.

  Toutes les informations d'état nécessaires pour la commande de flux de bout en bout et la fiabilité sont mises en œuvre dans les hôtes, dans la couche de transport ou dans les programmes d'application. Toutes les informations de commande de connexion sont donc co-localisées dans les points d'extrémité de la communication, de sorte qu'elle ne seront perdues que si le point d'extrémité est défaillant. Les routeurs ne contrôlent les flux de message qu'indirectement, en abandonnant des paquets ou en augmentant le délai du réseau.

  Noter que les développements futurs des protocoles pourraient bien finir par mettre plus d'états dans les routeurs. Ceci est particulièrement probable pour l'acheminement en diffusion groupée, pour la réservation de ressources, et pour la transmission fondée sur le flux.

- La complexité d'acheminement devrait être dans les routeurs.
  L'acheminement est un problème complexe et difficile, et il devrait être effectué par les routeurs, et non par les hôtes. Un objectif important est d'isoler le logiciel d'hôte des changements causés par l'évolution inévitable de l'architecture d'acheminement de l'Internet.

- Le système doit tolérer les variations des grands réseaux.
  Un objectif de base de la conception de l'Internet est de tolérer une large gamme de caractéristiques de réseau - par exemple, bande passante, délai, perte de paquet, réarrangement de paquets, et taille maximum de paquet. Un autre objectif est la robustesse aux défaillances des réseaux individuels, des routeurs, et des hôtes, en utilisant la bande passante restant disponible quelle qu'elle soit. Finalement, le but est l'interconnexion complète des systèmes ouverts : un routeur Internet doit être capable d'interopérer de façon robuste et effective avec tout autre routeur ou hôte Internet, à travers divers chemins de l'Internet.

Les développeurs ont parfois conçu des objectifs moins ambitieux. Par exemple, l'environnement de LAN est normalement beaucoup plus bénin que l'Internet dans son ensemble ; les LAN ont peu de perte de paquet et de délai et ne réordonnent pas les paquets. Certains fabricants ont mis sur le marché des mises en œuvre qui sont adéquates pour un environnement de LAN simple, mais fonctionnent mal pour un interfonctionnement général. Le fabricant justifie un tel produit par son côté économique sur le marché restreint des LAN. Cependant, les LAN isolés restent rarement isolés longtemps. Ils sont bientôt connectés les uns aux autres, en internet aux dimensions de l'organisation, et finalement au système Internet mondial. Finalement, ni le consommateur ni le fabricant ne sont satisfaits d'un routeur incomplet ou inférieur à la norme.

Les exigences du présent document sont conçues pour un routeur accomplissant la totalité de ses fonctions. On souhaite que les routeurs pleinement conformes puissent être utilisés dans presque toutes les parties de l'Internet.

3 Couche de liaison

Bien que [INTRO:1] traite des normes de la couche de liaison (IP sur diverses couches de liaison, ARP, etc.), le présent document anticipe que les matériaux de couche de liaison seront couverts par un document d'exigences de couche de liaison séparé. Un document d'exigences de couche Liaison serait applicable à la fois aux hôtes et aux routeurs. Et donc, le présent document ne rendra pas obsolètes les parties de [INTRO:1] qui traitent des questions de couche de liaison.

3.1 Introduction

Les routeurs ont essentiellement les mêmes exigences de protocole de couche de liaison que les autres sortes de systèmes Internet. Ces exigences sont données à la section 3 des Exigences pour les passerelles Internet [INTRO:1]. Un routeur DOIT se conformer à ses exigences et DEVRAIT se conformer à ses recommandations. Comme certains matériaux du présent document commencent à dater, certaines exigences et explications supplémentaires figurent ci-dessous.

Discussion

On s'attend à ce que la communauté de l'Internet produise des exigences pour une norme de la couche de liaison Internet qui remplacera à la fois ce chapitre et celui intitulé "PROTOCOLES DE COUCHE INTERNET" dans [INTRO:1].

3.2 Interface de couche Liaison/Internet

Le présent document n'essaye pas de spécifier l'interface entre la couche de liaison et les couches supérieures. Cependant, noter bien que les autres parties du présent document, en particulier la section 5, demandent que diverses sortes d'informations soient passées à travers cette frontière de couche.

La présente section utilise les définitions suivantes :

- Adresse physique de source
  C'est l'adresse de couche de liaison de l'hôte ou du routeur d'où le paquet a été reçu.
- Adresse physique de destination
  C'est d'adresse de couche de liaison à laquelle le paquet a été envoyé.

Les informations qui doivent passer de la couche Liaison à la couche Inter réseau pour chaque paquet reçu sont :

(1) Le paquet IP (5.2.2),

(2) La longueur de la portion de données (c'est-à-dire, non inclus le tramage de couche de liaison) de la trame de couche de liaison (5.2.2),

(3) L'identité de l'interface physique d'où le paquet IP a été reçu (5.2.3), et

(4) La classification de l'adresse physique de destination du paquet en envoi individuel, diffusion ou diffusion groupée de couche Liaison (4.3.2), (5.3.4).

De plus, la couche Liaison devrait aussi fournir :

(5) L'adresse physique de source.

Les informations qui doivent passer de la couche Internet à la couche Liaison pour chaque paquet transmis sont :

(1) Le paquet IP (5.2.1)
(2) La longueur du paquet IP (5.2.1)
(3) L'interface physique de destination (5.2.1)
(4) L'adresse IP du prochain bond (5.2.1)

De plus, la couche Internet devrait aussi fournir :

(5) la valeur de priorité de la couche Liaison (5.3.3.2)

La couche Liaison doit aussi notifier à la couche Internet si le paquet à transmettre cause une erreur de couche de liaison en rapport avec la préséance (5.3.3.3).

3.3 Questions spécifiques

3.3.1 Encapsulation d'en-queue

Les routeurs qui peuvent se connecter aux Ethernets à dix mégabits PEUVENT être capables de recevoir et transmettre

des paquets Ethernet encapsulés en utilisant l'encapsulation d'en-queue décrite dans [LINK:1]. Cependant, un routeur NE DEVRAIT PAS être à l'origine de paquets encapsulant un en-queue. Un routeur NE DOIT PAS être à l'origine de paquets encapsulant un en-queue sans vérifier d'abord, en utilisant le mécanisme décrit dans [INTRO:2], que la destination immédiate du paquet veut et est capable d'accepter les paquets qui encapsulent des en-queues. Un routeur NE DEVRAIT PAS accepter (en utilisant ces mécanismes) les paquets qui encapsulent des en-queues.

3.3.2 Protocole de résolution d'adresse - ARP

Les routeurs qui mettent en œuvre ARP DOIVENT être conformes et DEVRAIENT être inconditionnellement conformes aux exigences de [INTRO:2].

La couche de liaison NE DOIT PAS rapporter une erreur Destination Inaccessible à IP seulement parce qu'il n'y a pas d'entrée d'antémémoire ARP pour une destination ; il DEVRAIT mettre en file d'attente jusqu'à un petit nombre de datagrammes pendant une brève durée pendant qu'il effectue la séquence demande/réponse ARP, et ne répondre que la destination est inaccessible à un des datagrammes mis en file d'attente que lorsque ceci se révèle sans conséquences.

Un routeur NE DOIT croire aucune réponse ARP qui revendique l'adresse de couche de liaison d'un autre hôte ou routeur qui serait une adresse de diffusion ou de diffusion groupée.

3.3.3 Coexistence Ethernet et 802.3

Les routeurs qui peuvent se connecter aux Ethernets à dix mégabits DOIVENT être conformes et DEVRAIENT être inconditionnellement conformes aux exigences Ethernet de [INTRO:2].

3.3.4 Unité de transmission maximum - MTU

La MTU de chaque interface logique DOIT être configurable dans la gamme des MTU légales pour l'interface.

De nombreux protocoles de couche de liaison définissent une taille maximale de trame d'envoi. Dans de tels cas, un routeur NE DOIT PAS permettre d'établir une MTU qui permettrait d'envoyer des trames plus longues que celles permises par le protocole de couche de liaison. Cependant, un routeur DEVRAIT accepter de recevoir un paquet aussi long que la taille maximale de trame même si elle est plus longue que la MTU.

Discussion

Noter que ceci est une exigence plus stricte que celle imposée aux hôtes par [INTRO:2], qui exige que la MTU de chaque interface physique soit configurable.

Si un réseau utilise une MTU plus petite que la taille de trame maximum pour la couche de Liaison, un routeur peut recevoir des paquets plus longs que la MTU de la part d'hôtes mal configurés et incomplètement initialisés. Le principe de robustesse indique que le routeur devrait si possible réussir à recevoir ces paquets.

3.3.5 Protocole point à point - PPP

Au contraire de [INTRO:1], l'Internet a un protocole de ligne point à point normalisé : le protocole point à point (PPP), défini dans [LINK:2], [LINK:3], [LINK:4], et [LINK:5].

Une interface point à point est toute interface qui est conçue pour envoyer des données sur une ligne point à point. De telles interfaces incluent les lignes téléphoniques, les liaisons louées, dédiées ou directes (à deux ou quatre fils) et peuvent utiliser des canaux point à point ou des circuits virtuels d'interfaces multiplexées telles que le RNIS. Elles utilisent normalement un modem normalisé ou une interface binaire en série (comme RS-232, RS-449 ou V.35), avec des horloges synchrones ou asynchrones. Les interfaces multiplexées ont souvent des interfaces physiques particulières.

Une interface en série d'utilisation générale utilise le même support physique qu'une ligne point à point, mais prend en charge l'utilisation de réseaux de couche de liaison aussi bien que la connectivité point à point. Les réseaux de couche de liaison (tels que X.25 ou à relais de trame) utilisent une spécification de couche de liaison IP différente.

Les routeurs qui mettent en œuvre des interfaces point à point ou d'utilisation générale DOIVENT mettre en œuvre PPP.

PPP DOIT être pris en charge sur toutes les interfaces en série d'utilisation générale sur un routeur. Le routeur PEUT permettre que la ligne soit configurée pour utiliser les protocoles de ligne point à point autres que PPP. Les interfaces point à point DEVRAIENT utiliser PPP par défaut lorsqu'elles sont activées ou exiger la configuration du protocole de couche de liaison avant d'être activées. Les interfaces en série d'utilisation générale DEVRAIENT exiger la configuration du protocole de couche de liaison avant d'être activées.

3.3.5.1 Introduction

Ce paragraphe donne des lignes directrices aux développeurs de routeurs de sorte qu'ils puissent s'assurer de l'interopérabilité avec les autres routeurs qui utilisent PPP sur des liaisons synchrones ou asynchrones.

Il est particulièrement important qu'un développeur comprenne la sémantique du mécanisme de négociation d'option.

Les options sont un moyen pour un appareil local d'indiquer à un homologue distant ce que l'appareil local va accepter de la part de l'homologue distant, et non ce qu'il souhaite envoyer. Il appartient à l'homologue distant de décider de ce qu'il est convenable d'envoyer dans les limites de l'ensemble des options que l'appareil local a déclaré pouvoir accepter. Donc, il est parfaitement acceptable et normal qu'un homologue distant accuse réception de toutes les options indiquées dans une demande de configuration (CR) LCP même si l'homologue distant ne prend en charge aucune de ces options.

On répète que les options sont simplement un mécanisme pour que chaque appareil indique à son homologue ce qu'il va accepter, et pas nécessairement ce qu'il va envoyer.

3.3.5.2 Options du protocole de commande de liaison (LCP)

Le protocole de commande de liaison PPP (LCP) offre un certain nombre d'options qui peuvent être négociées. Ces options incluent (entre autres) la compression de champ d'en-tête et de contrôle, la transposition de caractère asynchrone, l'unité maximale de réception (MRU, *Maximum Receive Unit*), la surveillance de qualité de liaison (LQM, *Link Quality Monitoring*), le numéro magique (pour la détection de boucles), le protocole d'authentification par mot de passe (PAP, *Password Authentication Protocol*), le protocole d'authentification par dialogue à énigme (CHAP, *Challenge Handshake Authentication Protocol*), et la séquence de vérification de trame (FCS) à 32 bits.

Un routeur PEUT utiliser la compression de champ d'adresse/contrôle sur des liaisons aussi bien synchrones qu'asynchrones. Un routeur PEUT utiliser la compression de champ de protocole sur des liaisons aussi bien synchrones qu'asynchrones. Un routeur qui indique qu'il peut accepter ces compressions DOIT être capable d'accepter aussi les informations d'en-tête PPP non compressées.

Discussion

Ces options contrôlent l'apparition de l'en-tête PPP. Normalement l'en-tête PPP comporte l'adresse, le champ de contrôle, et le champ de protocole. L'adresse, sur une ligne point à point, est 0xFF, indiquant "diffusion". Le champ de contrôle est 0x03, indiquant "Informations non numérotées". L'identifiant de protocole est une valeur de deux octets qui indique le contenu de la zone de données de la trame. Si un système négocie la compression de champ d'adresse et de contrôle, il indique à son homologue qu'il acceptera les trames PPP qui ont ou n'ont pas ces champs au début de l'en-tête. Il n'indique pas qu'il enverra des trames dont ces champs auraient été retirés.

La compression de champ de protocole, lorsque elle est négociée, indique que le système est d'accord pour recevoir des champs de protocole compressés à un octet, lorsque c'est légal. Il n'est pas exigé que l'envoyeur le fasse.

L'utilisation de la compression de champ d'adresse/contrôle n'est pas cohérente avec l'utilisation de PPP en mode numéroté (fiable).

MISE EN ŒUVRE

Certains matériels ne traitent pas bien les informations d'en-tête de longueur de variable. Dans ces cas, il est plus significatif pour l'homologue distant d'envoyer l'en-tête PPP complet. Les mises en œuvre peuvent s'assurer de cela en n'envoyant pas les options de champ d'adresse/contrôle et compression de champ de protocole à l'homologue distant. Même si l'homologue distant a indiqué sa capacité à recevoir des en-têtes compressés, il n'est pas exigé que le routeur local envoie des en-têtes compressés.

Un routeur DOIT négocier la transposition de caractère de contrôle asynchrone (ACCM, *Asynchronous Control Character Map*) pour les liaisons PPP asynchrones, mais NE DEVRAIT PAS négocier l'ACCM pour des liaisons synchrones. Si un routeur reçoit une tentative de négociation de l'ACCM sur une liaison synchrone, il DOIT accuser réception *(ACKnowledge)* de l'option puis l'ignorer.

Discussion

Il y a des mises en œuvre qui offrent à la fois le mode de fonctionnement synchrone et asynchrone et peuvent utiliser le même code pour mettre en œuvre la négociation d'option. Dans cette situation il est possible qu'une extrémité ou l'autre puisse envoyer l'option ACCM sur une liaison synchrone.

Un routeur DEVRAIT négocier de façon appropriée l'unité maximum de réception (MRU). Même si un système négocie une MRU inférieure à 1 500 octets, il DOIT être capable de recevoir une trame de 1 500 octets.

Un routeur DEVRAIT négocier et activer l'option de surveillance de qualité de liaison (LQM, *link quality monitoring*).

Discussion

Le présent mémoire ne spécifie pas de politique pour décider si la qualité de la liaison est adéquate. Cependant, il est important (voir au paragraphe 3.3.6) qu'un routeur désactive les liaisons défaillantes.

Un routeur DEVRAIT mettre en œuvre et négocier l'option de numéro magique pour la détection de boucle.

Un routeur PEUT prendre en charge les options d'authentification (PAP – protocole d'authentification par mot de passe, et/ou CHAP – protocole d'authentification par dialogue à énigmes).

Un routeur DOIT prendre en charge la séquence de vérification de trame par CRC à 16 bits (FCS) et PEUT prendre en charge le CRC à 32 bits.

3.3.5.3 Options de protocole de commande IP (IPCP)

Un routeur PEUT offrir d'effectuer la négociation d'adresse IP. Un routeur DOIT accepter un refus (REJect) d'effectuer une négociation d'adresse IP de la part de l'homologue.

Les routeurs qui fonctionnent à des vitesses de liaison de 19 200 bit/s ou moins DEVRAIENT mettre en œuvre et offrir d'effectuer la compression d'en-tête de Van Jacobson (VJ). Les routeurs qui mettent en œuvre la compression VJ DEVRAIENT mettre en œuvre une commande administrative l'activant ou la désactivant.

3.3.6 Essai d'interface

Un routeur DOIT avoir un mécanisme pour permettre aux logiciels d'acheminement de déterminer si une interface physique est disponible pour envoyer ou non des paquets ; sur les interfaces multiplexées où des circuits virtuels permanents sont ouverts pour des ensembles limités de voisins, le routeur doit aussi être capable de déterminer si les circuits virtuels sont viables. Un routeur DEVRAIT avoir un mécanisme pour permettre au logiciel d'acheminement de juger de la qualité d'une interface physique. Un routeur DOIT avoir un mécanisme pour informer le logiciel d'acheminement du moment où une interface physique devient disponible ou indisponible pour envoyer des paquets à cause d'une action administrative. Un routeur DOIT avoir un mécanisme pour informer le logiciel d'acheminement du moment où il détecte qu'une interface de niveau liaison devient disponible ou indisponible, quelle qu'en soit la raison.

Discussion

Il est crucial que les routeurs aient des mécanismes exploitables pour déterminer que leurs connexions de réseau fonctionnent correctement. La défaillance à détecter une perte de liaison, ou la défaillance à effectuer les actions appropriées lorsqu'un problème est détecté, peut conduire à des situations sans issue.

Les mécanismes disponibles pour détecter les problèmes avec les connexions de réseau varient considérablement selon les protocoles de couche Liaison utilisés et le matériel d'interface. L'objectif est de maximiser la capacité de détection des défaillances dans les contraintes de la couche de liaison.

4 Couche INTERNET - Protocoles

4.1 Introduction

Cette section et la section 5 discutent des protocoles utilisés à la couche Internet : IP, ICMP, et IGMP. Comme la transmission est visiblement un sujet crucial dans un document qui discute des routeurs, la section 5 se limite aux aspects des protocoles qui se rapportent directement à la transmission. La présente section contient le reste de la discussion sur les protocoles de la couche Internet.

4.2 Protocole INTERNET - IP

4.2.1 Introduction

Les routeurs DOIVENT mettre en œuvre le protocole IP, comme défini dans [INTER:1]. Ils DOIVENT aussi mettre en œuvre ses extensions obligatoires : sous-réseaux (définis dans [INTER:2]), diffusion IP (définie dans [INTER:3]), et acheminement inter domaine sans classe (CIDR, *Classless Inter-Domain Routing*, défini dans [INTER:15]).

Les développeurs de routeurs n'ont pas besoin de prendre en considération la conformité au paragraphe de [INTRO:2] intitulé "Protocole Internet -- IP," car ce paragraphe est entièrement reproduit ou remplacé par le présent document. Un routeur DOIT être conforme, et DEVRAIT être inconditionnellement conforme aux exigences du paragraphe intitulé "Questions particulières" au sujet de IP dans [INTRO:2].

Dans ce qui suit, l'action spécifiée dans certains cas est d'éliminer en silence un datagramme reçu. Cela signifie que le datagramme sera éliminé sans autre traitement et que le routeur n'enverra aucun message d'erreur ICMP (voir au paragraphe [4.3]). Cependant, pour le diagnostic des problèmes, un routeur DEVRAIT fournir la capacité de localiser l'erreur (voir au paragraphe 1.3.3), y compris le contenu du datagramme éliminé en silence, et DEVRAIT compter les datagrammes éliminés.

4.2.2 Découverte du protocole

La RFC 791 [INTER:1] est la spécification pour le protocole Internet.

4.2.2.1 Options : RFC 791 paragraphe 3.2

Dans les datagrammes reçus par le routeur lui-même, la couche IP DOIT interpréter les options IP qu'elle comprend et préserver le reste inchangé pour qu'il soit utilisé par les protocoles de couches supérieures.

Les protocoles de couches supérieures peuvent avoir besoin de la capacité à établir des options IP dans les datagrammes qu'ils envoient ou d'examiner les options IP dans les datagrammes qu'ils reçoivent. Des paragraphes ultérieurs du présent document discutent de la prise en charge des options IP spécifiques exigées par les protocoles de couches supérieures.

Discussion

Ni le présent mémoire ni [INTRO:2] ne définissent l'ordre dans lequel un receveur doit traiter plusieurs options dans le même en-tête IP. Les hôtes et les routeurs qui sont à l'origine de datagrammes qui contiennent plusieurs options doivent savoir que cela introduit des ambiguïtés dans la signification de certaines options lorsqu'elles sont combinées avec une option source-route.

Les exigences pour les options spécifiques d'IP figurent ci-après :

(a) Option Security
Certains environnements exigent l'option Sécurité dans chaque paquet émis ou reçu. Les routeurs DEVRAIENT METTRE EN ŒUVRE l'option de sécurité révisée décrite dans [INTER:5].

Discussion

Noter que les options de sécurité décrites dans [INTER:1] et la RFC 1038 ([INTER:16]) sont obsolètes.

(b) Option Stream Identifier *(identifiant de flux)*
Cette option est obsolète ; les routeurs NE DEVRAIT PAS placer cette option dans un datagramme que le routeur émet. Cette option DOIT être ignorée dans les datagrammes reçus par le routeur.

(c) Options Source Route *(acheminement depuis la source)*
Un routeur DOIT être capable d'agir comme destination finale d'un acheminement de source. Si un routeur reçoit un paquet qui contient un acheminement de source terminé, le paquet a atteint sa destination finale. Dans une telle option, le pointeur est dirigé au delà du dernier champ et l'adresse de destination dans l'en-tête IP vise le routeur. L'option telle que reçue (l'acheminement indiqué) DOIT être passée jusqu'à la couche de transport (ou au traitement de message ICMP).

Dans le cas général, une réponse correcte à un datagramme à acheminement de source emprunte le même acheminement. Un routeur DOIT fournir un moyen par lequel les protocoles et applications de transport peuvent renverser l'acheminement de source dans un datagramme reçu. Cet acheminement de source inversé DOIT être inséré dans les datagrammes qu'ils émettent (voir les précisions dans [INTRO:2]) lorsque le routeur ne connaît pas les contraintes de politique. Cependant, si le routeur est au courant de la politique, il PEUT choisir un autre chemin.

Certaines applications dans le routeur PEUVENT exiger que l'utilisateur soit capable d'entrer un acheminement de source.

Un routeur NE DOIT PAS être à l'origine d'un datagramme contenant plusieurs options d'acheminement de source. Ce qu'un routeur devrait faire si on lui demande de transmettre un paquet contenant plusieurs options d'acheminement de source est décrit au paragraphe [5.2.4.1].

Lorsqu'une option d'acheminement de source est créée (ce qui peut arriver lorsque le routeur est à l'origine d'un datagramme acheminé depuis la source ou qu'il insère une option d'acheminement de source par suite d'un filtre spécial) il DOIT être correctement formé même si il est créé par inversion d'un acheminement enregistré qui inclut par erreur l'hôte de source (voir le cas (B) dans la discussion ci-dessous).

Discussion

Supposons qu'un datagramme acheminé depuis la source est à acheminer à partir de la source S à la destination D via les routeurs G1, G2, Gn. La source S construit un datagramme avec l'adresse IP de G1 comme adresse de destination, et une option d'acheminement de source pour amener le datagramme à sa destination sur le reste du chemin. Cependant, il y a une ambiguïté dans la spécification sur le point de savoir si l'option d'acheminement de source dans un datagramme envoyé par S devrait être (A) ou (B) :

(A): {>>G2, G3, ... Gn, D} <--- CORRECT

(B): {S, >>G2, G3, ... Gn, D} <---- FAUX

(où >> représente le pointeur). Si (A) est envoyé, le datagramme reçu à D contiendra l'option : {G1, G2, ... Gn >>}, avec S et D comme adresses IP de source et de destination. Si (B) était envoyé, le datagramme reçu à D contiendrait encore S et D comme mêmes adresses IP de source et de destination, mais l'option serait : {S, G1, .Gn >>} ; c'est-à-dire que l'hôte d'origine serait le premier bond sur le chemin.

(d) Option Record Route *(chemin enregistré)*
Les routeurs PEUVENT prendre en charge l'option Record Route dans des datagrammes dont l'origine est le routeur.

(e) Option Timestamp *(horodatage)*
Les routeurs PEUVENT prendre en charge l'option Timestamp dans des datagrammes dont l'origine est le routeur. Les règles suivantes s'appliquent :

- Lorsqu'il est à l'origine d'un datagramme qui contient une option Timestamp, un routeur DOIT enregistrer un horodatage dans l'option si
  - ses champs d'adresse Internet ne sont pas pré-spécifiés ou
  - sa première adresse pré-spécifiée est l'adresse IP de l'interface logique sur laquelle le datagramme est envoyé (ou l'identifiant du routeur si le datagramme est envoyé sur une interface non numérotée).
- Si le routeur reçoit lui-même un datagramme contenant une option Timestamp, le routeur DOIT insérer l'heure en cours dans l'option Timestamp (si il a de l'espace dans l'option pour le faire) avant de passer l'option à la couche transport ou à l'ICMP pour traitement. Si il n'y a pas d'espace, le routeur DOIT incrémenter le compteur Overflow dans l'option.
- Une valeur d'horodatage DOIT suivre les règles définies dans [INTRO:2].

MISE EN ŒUVRE
Pour maximiser l'utilité de l'horodatage contenu dans l'option timestamp, l'horodatage inséré devrait être, autant que possible, l'heure à laquelle le paquet est arrivé au routeur. Pour les datagrammes dont l'origine est le routeur, l'horodatage inséré devrait être, autant que possible, l'heure à laquelle le datagramme a été passé pour transmission à la couche de liaison.

L'option timestamp permet l'utilisation d'une horloge temporelle non standard, mais l'utilisation d'une horloge non synchronisée limite l'utilité de l'horodatage. Donc, les routeurs seront bien avisés de mettre en œuvre le protocole de l'heure du réseau pour les besoins de la synchronisation de leurs horloges.

4.2.2.2 Adresses dans les options : RFC 791 paragraphe 3.1

Les routeurs sont appelés à insérer leurs adresses dans les options Record Route, Strict Source et Record Route, Loose Source et Record Route, ou Timestamp. Lorsque un routeur insère son adresse dans une telle option, il DOIT utiliser l'adresse IP de l'interface logique à laquelle le paquet est envoyé. Lorsque cette règle ne peut pas être respectée parce que l'interface de sortie n'a pas d'adresse IP (c'est-à-dire que c'est une interface non numérotée), le routeur DOIT insérer à la place son identifiant de routeur (routeur-id). Le routeur-id du routeur est une des adresses IP du routeur. L'identifiant du routeur peut être spécifié selon le système ou selon la liaison. Quelle que soit l'adresse utilisée comme routeur-id, elle NE DOIT PAS changer (même après réamorçage) sauf si elle est changée par le gestionnaire du réseau. Les changements de gestion pertinents incluent la reconfiguration du routeur comme l'adresse IP utilisée comme routeur-id cesse d'être une des adresses IP du routeur. Les routeurs qui ont plusieurs interfaces non numérotées PEUVENT avoir plusieurs routeur-id. Chaque interface non numérotée DOIT être associée à un routeur-id particulier. Cette association NE DOIT PAS changer (même lors des réamorçages) sans une reconfiguration du routeur.

Discussion
La présente spécification ne permet pas que les routeurs aient moins d'une adresse IP. Ceci n'est pas perçu comme une limitation grave, car un routeur a besoin d'une adresse IP pour satisfaire aux exigences de gestion de la section 8 même si le routeur n'est connecté qu'à des liaisons point à point.

MISE EN ŒUVRE
Une méthode de choix possible pour que le routeur-id satisfasse cette exigence est d'utiliser la plus petite (ou la plus grande) adresse IP numérique (en traitant l'adresse comme un entier de 32 bits) qui soit allouée au routeur.

4.2.2.3 Bits d'en-tête IP non utilisés : RFC 791 paragraphe 3.1

L'en-tête IP contient deux bits réservés : un dans l'octet Type de service et l'autre dans le champ Fanions. Un routeur NE DOIT PAS régler un de ces bits à un dans les datagrammes dont l'origine est le routeur. Un routeur NE DOIT PAS abandonner (refuser de recevoir ou de transmettre) un paquet simplement parce que un ou plusieurs de ces bits réservés a une valeur différente de zéro ; c'est-à-dire que le routeur NE DOIT PAS vérifier les valeurs de ces bits.

Discussion

Des révisions ultérieures du protocole IP pourront faire usage de ces bits inutilisés. Ces règles sont destinées à s'assurer que ces révisions pourront intervenir sans qu'on soit obligé en même temps de mettre à jour tous les routeurs de l'Internet.

4.2.2.4 Type de service : RFC 791 paragraphe 3.1

L'octet Type de Service dans l'en-tête IP est divisé en trois sections : le champ Precedence *(préséance)* (3 bits de plus fort poids), un champ qui est habituellement appelé Type de Service ou TOS (4 bits suivants), et un bit réservé (le bit de moindre poids).

Les règles qui gouvernent le bit réservé sont décrites au paragraphe 4.2.2.3.

Un exposé plus développé du champ TOS et de son utilisation figure dans [ROUTE:11].

La description du champ Precedence IP est remplacée par celle du paragraphe 5.3.3. La RFC 795, Transpositions de service, est obsolète et NE DEVRAIT PAS être mise en œuvre.

4.2.2.5 Somme de contrôle d'en-tête : RFC 791 paragraphe 3.1

Comme décrit au paragraphe 5.2.2, un routeur DOIT vérifier la somme de contrôle IP de tout paquet reçu, et il DOIT éliminer les messages qui contiennent des sommes de contrôle invalides. Le routeur NE DOIT PAS fournir de moyens de désactiver cette vérification de somme de contrôle.

Un routeur PEUT utiliser une mise à jour de somme de contrôle d'en-tête IP par incrémentation lorsque le seul changement de l'en-tête IP est la durée de vie restante. Ceci réduira la possibilité de corruption non détectée de l'en-tête IP par le routeur. Voir dans [INTER:6] une discussion sur la mise à jour par incrémentation de la somme de contrôle.

MISE EN ŒUVRE

Une description plus complète de la somme de contrôle IP, comprenant des conseils très complets de mise en œuvre, se trouve dans [INTER:6] et [INTER:7].

4.2.2.6 Options d'en-tête non reconnue : RFC 791 paragraphe 3.1

Un routeur DOIT ignorer les options IP qu'il ne reconnaît pas. Un corollaire de cette exigence est qu'un routeur DOIT mettre en œuvre l'option End of Option List *(fin de liste d'options)* et l'option No Operation, car aucune d'elles ne contient de longueur explicite.

Discussion

Toutes les options IP futures contiendront une longueur explicite.

4.2.2.7 Fragmentation : RFC 791 paragraphe 3.2

La fragmentation, telle que décrite dans [INTER:1], DOIT être acceptée par un routeur.

Lorsqu'un routeur fragmente un datagramme IP, il DEVRAIT minimiser le nombre de fragments. Lorsqu'un routeur fragmente un datagramme IP, il DEVRAIT envoyer les fragments dans l'ordre. Une méthode de fragmentation qui peut générer un fragment IP significativement plus petit que l'autre PEUT être la cause de ce que le premier fragment IP est le plus petit.

Discussion

Plusieurs techniques de fragmentation sont d'usage courant dans l'Internet. Une d'elle implique le découpage du datagramme IP en fragments IP dont le premier a la taille de la MTU, et les autres sont approximativement de la même taille, plus petits que la MTU. La raison de ce comportement est double. Le premier fragment IP dans la séquence sera la MTU effective du chemin actuel entre les hôtes, et les fragments IP suivants sont dimensionnés de façon à minimiser la fragmentation ultérieure du datagramme IP. Une autre technique est de partager le datagramme IP en fragments IP de la taille de la MTU, le dernier fragment étant le seul qui soit plus petit, comme décrit dans [INTER:1].

Une astuce courante utilisée par certaines des mises en œuvre de TCP/IP est de fragmenter un datagramme IP en fragments IP qui ne dépassent pas 576 octets lorsque le datagramme IP doit traverser un routeur. Ceci est destiné à permettre que les fragments IP qui en résultent passent le reste du chemin sans autre fragmentation. Ceci pourrait cependant faire peser une charge plus lourde sur l'hôte de destination, car il aurait un plus grand nombre de fragments IP à rassembler en un datagramme IP. Cela ne serait pas non plus très efficace sur les réseaux où la MTU ne change qu'une seule fois et reste très supérieure à 576 octets. On a comme exemples les réseaux LAN comme un réseau IEEE 802.5 avec une MTU de 2048 ou un réseau Ethernet avec une MTU de 1500).

Une autre technique de fragmentation était de partager le datagramme IP en fragments IP de taille approximativement égale, avec des tailles inférieures ou égales à la MTU du réseau du prochain bond. Ceci est destiné à minimiser le nombre de fragments qui résulteraient de fragmentation supplémentaires plus loin sur le chemin, et assure un délai égal pour chaque fragment.

Les routeurs DEVRAIENT générer le plus faible nombre possible de fragments IP.

Le travail avec des machines lentes nous conduit à penser qu'il est nécessaire de fragmenter les messages, et l'envoi en

premier du plus petit fragment IP maximise les chance qu'un hôte qui a une interface lente reçoive tous les fragments.

4.2.2.8 Réassemblage : RFC 791 paragraphe 3.2

Comme spécifié dans le paragraphe correspondant de [INTRO:2], un routeur DOIT prendre en charge le réassemblage des datagrammes qu'il se livre à lui-même.

4.2.2.9 Durée de vie restante : RFC 791 paragraphe 3.2

Le traitement de la durée de vie restante (TTL, *Time to Live*) pour les paquets dont le routeur est l'origine ou qu'il reçoit est gouvernée par [INTRO:2] ; ce paragraphe ne change aucune de ses stipulations. Cependant, comme le reste de la section Protocole IP de [INTRO:2] est réécrit, ce paragraphe l'est aussi.

Noter en particulier qu'un routeur NE DOIT PAS vérifier la TTL d'un paquet excepté quand il le transmet.

Un routeur NE DOIT PAS envoyer ou transmettre un datagramme avec une durée de vie restante (TTL) de zéro.

Un routeur NE DOIT PAS éliminer un datagramme seulement parce qu'il l'a reçu avec une TTL égale à zéro ou un ; si il est pour le routeur et valide par ailleurs, le routeur DOIT essayer de le recevoir.

Sur les messages dont le routeur est à l'origine, la couche IP DOIT fournir un moyen pour que la couche transport règle le champ TTL de chaque datagramme envoyé. Lorsque une valeur de TTL fixe est utilisée, elle DOIT être configurable. Le nombre DEVRAIT excéder le diamètre internet normal, et le sens commun suggère qu'il devrait dépasser deux fois le diamètre internet pour permettre un accroissement. Les valeurs courantes suggérées sont normalement indiquées dans la RFC des numéros alloués. Le champ TTL a deux fonctions : limiter la durée de vie des segments TCP (voir la RFC 793 [TCP:1], p. 28), et de mettre un terme aux boucles d'acheminement Internet. Bien que la TTL soit une durée en secondes, elle a aussi certains attributs d'un compte de bonds, car chaque routeur est obligé de réduire le champ TTL d'au moins un.

L'expiration de la TTL est destinée à provoquer l'élimination des datagrammes par les routeurs, mais pas par l'hôte de destination. Les hôtes qui agissent comme des routeurs en transmettant les datagrammes doivent donc suivre les règles des routeurs pour la TTL.

Un protocole de couche supérieure peut vouloir régler la TTL afin de mettre en œuvre une recherche à "portée expansive" de certaines ressources Internet. Ceci est utilisé par certains outils de diagnostic, et est supposé être utile pour localiser le "plus proche" serveur d'une classe donnée en utilisant la diffusion groupée IP, par exemple. Un protocole de transport particulier peut aussi vouloir spécifier sa propre limite de TTL sur la durée de vie maximum de datagramme.

Une valeur par défaut fixe doit être au moins assez grande pour le "diamètre" Internet, c'est-à-dire, le chemin le plus long possible. Une valeur raisonnable est d'environ deux fois le diamètre, pour permettre la poursuite de la croissance de l'Internet. Au moment de cette rédaction, les messages qui traversent fréquemment les États Unis traversent 15 à 20 routeurs ; ceci milite en faveur d'une valeur de TTL par défaut dépassant 40, et 64 est une valeur courante.

4.2.2.10 Diffusions sur plusieurs sous-réseaux : RFC 922

Toutes les diffusions sur des sous-réseaux (appelées diffusions sur plusieurs sous réseaux dans [INTER:3]) ont été déconseillées. Voir au paragraphe [5.3.5.3].

4.2.2.11 Adressage : RFC 791 paragraphe 3.2

Comme noté au paragraphe 2.2.5.1, il y a maintenant cinq classes d'adresses IP : les classes de A à E. Les adresses de classe D sont utilisées pour la diffusion groupée IP [INTER:4], alors que les adresses de classe E sont réservées pour utilisation expérimentale. La distinction entre les adresses de classe A, B, et C n'a plus d'importance ; elles sont utilisées comme préfixes de réseau en diffusion individuelle généralisée avec seulement un intérêt historique pour leurs classes.

Une adresse IP en diffusion individuelle est une adresse logique de 28 bits qui désigne un groupe d'hôtes, et peut être permanente ou temporaire. Les adresses de diffusion groupée permanentes sont allouées par l'Autorité d'allocation des numéros de l'Internet (IANA) [INTRO:7], alors que les adresses temporaires peuvent être allouées dynamiquement aux groupes temporaires. L'adhésion aux groupes est déterminée de façon dynamique en utilisant IGMP [INTER:4].

Nous allons maintenant résumer les cas particuliers importants pour les adresses IP d'utilisation générale en diffusion individuelle, en utilisant la notation suivante pour une adresse IP :

{ <Préfixe-de-réseau>, <Numéro-d'hôte> }

et la notation -1 pour un champ qui contient tous les bits à 1 et la notation 0 pour un champ qui contient tous les bits à 0.

(a) { 0, 0 }
Cet hôte sur ce réseau. Il NE DOIT PAS être utilisé comme adresse de source par les routeurs, sauf que le routeur PEUT

utiliser ceci comme adresse de source au titre d'une procédure d'initialisation (par exemple, si le routeur utilise BOOTP pour charger ses informations de configuration).

Les datagrammes entrants avec une adresse de source de { 0, 0 } qui sont reçus pour une livraison locale (voir au paragraphe 5.2.3), DOIVENT être acceptés si le routeur met en œuvre le protocole associé et si ce protocole définit clairement l'action appropriée à entreprendre. Autrement, un routeur DOIT éliminer en silence tout datagramme de livraison locale dont l'adresse de source est { 0, 0 }.

Discussion
Certains protocoles définissent des actions spécifiques à prendre en réponse à un datagramme reçu dont l'adresse de source est { 0, 0 }. Les demandes de gabarit *(Mask Request)* BOOTP et ICMP sont deux exemples. Le fonctionnement approprié de ces protocoles dépend souvent de la capacité à recevoir des datagrammes dont l'adresse de source est { 0, 0 }. Cependant, pour la plupart des protocoles, il est préférable d'ignorer les datagrammes qui ont une adresse de source de { 0, 0 } car ils ont probablement été générés par un hôte ou routeur mal configuré. Et donc, si un routeur sait comment traiter un datagramme donné qui a une adresse de source de { 0, 0 }, le routeur DOIT l'accepter. Autrement, le routeur DOIT l'éliminer.

Voir aussi au paragraphe 4.2.3.1 une utilisation non normalisée de { 0, 0 }.

(b) { 0, <Numéro-d'hôte> }
Hôte spécifié sur ce réseau. Il NE DOIT PAS être envoyé par les routeurs, excepté que le routeur PEUT l'utiliser comme adresse de source au titre d'une procédure d'initialisation par laquelle il apprend sa propre adresse IP.

(c) { -1, -1 }
Diffusion limitée. Il NE DOIT PAS être utilisé comme adresse de source.
Un datagramme avec cette adresse de destination sera reçu par chaque hôte et routeur sur le réseau physique connecté, mais il ne sera pas transmis en-dehors de ce réseau.

(d) { <Préfixe-réseau>, -1 }
Diffusion dirigée – une diffusion dirigée sur le préfixe de réseau spécifié. Il NE DOIT PAS être utilisé comme adresse de source. Un routeur PEUT être à l'origine de paquets en diffusion dirigée sur le réseau *(Network Directed Broadcast)*. Un routeur DOIT recevoir les paquets en diffusion dirigée sur le réseau ; cependant un routeur PEUT avoir une option de configuration pour empêcher la réception de ces paquets. Une telle option DOIT par défaut permettre la réception.

(e) { 127, <any> }
Adresse de repli d'hôte interne. Les adresses de cette forme NE DOIVENT PAS apparaître en-dehors d'un hôte.

Le <Préfixe-réseau> est alloué administrativement de telles sorte que sa valeur soit unique dans le domaine d'acheminement auquel l'appareil est connecté.

Les valeurs de 0 ou – 1 ne sont pas permises aux adresses IP pour les champs <Numéro-d'hôte> ou <Préfixe-réseau> excepté dans les cas particuliers figurant ci-dessus. Ceci implique que chacun de ces champs sera au moins long de deux bits.

Discussion
Les précédentes versions du présent document notaient aussi que ce numéro de sous-réseau ne doit être ni 0 ni – 1, et doit avoir au moins une longueur de deux bits.
Dans l'univers CIDR, le numéro de sous-réseau est clairement une extension du préfixe de réseau et ne peut pas être interprété sans le reste du préfixe. Cette restriction aux numéros de sous-réseau est donc sans signification du point de vue de CIDR et peut être ignorée sans risque.

Pour des précisions sur les adresses de diffusion, voir au paragraphe 4.2.3.1.

Lorsque un routeur est à l'origine d'un datagramme, l'adresse IP de source DOIT être une de ses propres adresses IP (mais pas une adresse de diffusion ou de diffusion groupée). La seule exception est durant l'initialisation.

Dans la plupart des cas, un datagramme adressé à une destination de diffusion ou diffusion groupée est traité comme si il avait été adressé à une des adresses IP du routeur ; c'est-à-dire :
- Un routeur DOIT recevoir et traiter normalement tout paquet qui a une adresse de destination de diffusion.
- Un routeur DOIT recevoir et traiter normalement tout paquet envoyé à une adresse de destination de diffusion groupée que le routeur a demandé à recevoir.

Le terme adresse de destination spécifique signifie l'adresse IP locale équivalente de l'hôte. L'adresse de destination

spécifique est définie comme étant l'adresse de destination dans l'en-tête IP sauf si l'en-tête contient une adresse de diffusion ou de diffusion groupée, auquel cas la destination spécifique est une adresse IP allouée à l'interface physique sur laquelle le datagramme est arrivé.

Un routeur DOIT éliminer en silence tout datagramme reçu qui contient une adresse IP de source qui est invalide selon les règles de ce paragraphe. Cette validation pourrait être effectuée par la couche IP ou (lorsque c'est approprié) par chaque protocole dans la couche transport. Comme avec tout datagramme qu'élimine un routeur, l'élimination du datagramme DEVRAIT être comptée.

Discussion

Un datagramme mal adressé peut être causé par la diffusion à la couche Liaison d'un datagramme en envoi individuel ou par un autre routeur ou hôte qui s'est trompé ou est mal configuré.

4.2.3 Questions spécifiques

4.2.3.1 Adresses de diffusion IP

Pour des raisons historiques, il y a un certain nombre d'adresses IP (certaines standard et d'autres non) qui sont utilisées pour indiquer d'un paquet IP est une diffusion IP. Un routeur :

(1) DOIT traiter comme diffusions IP les paquets adressés à 255.255.255.255 ou { <Préfixe-réseau>, -1 }.

(2) DEVRAIT éliminer en silence à réception (c'est-à-dire, ne même pas le livrer aux applications dans le routeur) tout paquet adressé à 0.0.0.0 ou { <Préfixe-réseau>, 0 }. Si ces paquets ne sont pas éliminés en silence, ils DOIVENT être traités comme des diffusions IP (voir le paragraphe 5.3.5). Il PEUT y avoir une option de configuration pour permettre la réception de ces paquets. Cette option DEVRAIT par défaut être leur élimination.

(3) DEVRAIT (par défaut) utiliser l'adresse de diffusion limitée (255.255.255.255) lors de la production d'une diffusion IP destinée à un (sous-) réseau connecté (sauf lors de l'envoi d'une réponse de gabarit d'adresse ICMP, comme exposé au paragraphe 4.3.3.9). Un routeur DOIT recevoir des diffusions limitées.

(4) NE DEVRAIT PAS être à l'origine de datagrammes adressés à 0.0.0.0 ou { <Préfixe-réseau>, 0 }. Il peut y avoir une option de configuration pour permettre la génération de ces paquets (au lieu d'utiliser la diffusion au format 1s pertinente). Cette option DEVRAIT par défaut ne pas les générer.

Discussion

Au second point, le routeur ne peut visiblement pas reconnaître les adresses de la forme { <Préfixe-réseau>, 0 } si le routeur n'a pas d'interface pour ce préfixe de réseau. Dans ce cas, les règles du point (2) ne s'appliquent pas parce que, du point de vue du routeur, le paquet n'est pas un paquet de diffusion IP.

4.2.3.2 Diffusion groupée IP

Un routeur IP DEVRAIT satisfaire aux exigences d'hôte par rapport à la diffusion groupée IP, comme spécifié dans [INTRO:2]. Un routeur IP DEVRAIT prendre en charge la diffusion groupée IP locale sur tous les réseaux connectés. Lorsqu'une transposition d'adresses IP en diffusion groupée a été spécifiée pour des adresses de couche de liaison (voir les diverses spécifications IP-sur-xxx), il DEVRAIT utiliser cette transposition, et PEUT être configurable pour utiliser à la place la diffusion de couche de liaison. Sur les liaisons point à point et toutes les autres interfaces, les diffusions groupées sont encapsulées comme des diffusions de couche de liaison. La prise en charge de la diffusion groupée IP locale inclut d'être à l'origine de datagrammes de diffusion groupée, de se joindre à des groupes de diffusion et de recevoir des datagrammes de diffusion groupée, et de quitter des groupes de diffusion. Cela implique la prise en charge de tout [INTER:4] y compris IGMP (voir au paragraphe [4.4]).

Discussion

Bien que [INTER:4] soit intitulé Extensions d'hôtes pour la diffusion groupée IP, il s'applique à tous les systèmes IP, aussi bien hôtes que routeurs. En particulier, comme les routeurs peuvent se joindre à des groupes de diffusion, il est normal pour eux d'effectuer la partie hôte de IGMP, faisant rapport de leurs adhésions de groupe à tous les routeurs en diffusion groupée qui peuvent être présents sur leurs réseaux de rattachement (qu'ils soient eux-mêmes ou non des routeurs en diffusion groupée).

Certains protocoles de routeur peuvent exiger une prise en charge spécifique de la diffusion groupée IP (par exemple, OSPF [ROUTE:1]), ou peuvent la recommander (par exemple, ICMP Router Discovery [INTER:13]).

4.2.3.3 Découverte de la MTU d'un chemin

Pour éliminer la fragmentation ou la minimiser, il est souhaitable de savoir quelle est la MTU de chemin tout le long du chemin de la source à la destination. La MTU de chemin est le minimum des MTU de chaque bond du chemin. [INTER:14] décrit une technique pour la découverte dynamique de l'unité de transmission maximale (MTU) d'un

chemin internet arbitraire. Pour un chemin qui passe à travers un routeur qui ne prend pas en charge [INTER:14], cette technique pourrait ne pas découvrir la MTU correcte du chemin, mais elle choisira toujours une MTU de chemin aussi adaptée, et dans de nombreux cas, mieux adaptée, que la MTU de chemin qui aurait été choisie par de plus vieilles techniques ou par la pratique courante.

Lorsqu'un routeur est à l'origine d'un datagramme IP, il DEVRAIT utiliser le schéma décrit dans [INTER:14] pour limiter la taille du datagramme. Si l'acheminement du routeur vers la destination du datagramme a été apprise d'un protocole d'acheminement qui fournit des informations de MTU de chemin, le schéma décrit dans [INTER:14] est encore utilisé, mais les informations de MTU de chemin provenant du protocole d'acheminement DEVRAIENT être utilisées comme supposition initiale sur la MTU de chemin et aussi comme limite supérieure de la MTU de chemin.

4.2.3.4 Sous-réseautage

Dans certaines circonstances, il peut être désirable de prendre en charge des sous-réseaux d'un réseau particulier qui ne sont interconnectés qu'à travers un chemin qui ne fait pas partie du réseau subdivisé. Ceci est connu sous le nom de prise en charge de sous-réseau discontinu.

Les routeurs DOIVENT prendre en charge les sous-réseaux discontinus.

MISE EN ŒUVRE

Dans les réseaux IP classiques, ceci était très difficile à réaliser ; dans les réseaux CIDR, c'est un sous produit naturel. Donc, un routeur NE DEVRAIT PAS faire de suppositions sur l'architecture de sous-réseau, mais DEVRAIT traiter chaque route comme un préfixe de réseau généralisé.

Discussion

L'Internet a crû dernièrement à une vitesse incroyable. Cela a fait peser de sévères contraintes sur la technologie d'adressage IP. Un facteur majeur de ces contraintes est l'existence de frontières strictes de classes d'adresses IP. Cela rend difficile de dimensionner efficacement les préfixes de réseau à leurs réseaux et d'agréger plusieurs préfixes de réseau en une seule information de routage. En éliminant les frontières de classe strictes de l'adresse IP et en traitant chaque route comme un préfixe de réseau généralisé, ces contraintes peuvent être réduites.

La technologie actuelle pour faire cela est l'acheminement inter domaine sans classe (CIDR, *Classless Inter Domain Routing*) [INTER:15].

Pour des raisons similaires, un bloc d'adresse associé à un préfixe de réseau donné pourrait être subdivisé en sous blocs de différentes tailles, de sorte que les préfixes de réseau associés aux sous blocs aient des longueurs différentes. Par exemple, au sein d'un bloc dont le préfixe de réseau est long de 8 bits, un sous bloc peut avoir un préfixe de réseau de 16 bits, un autre peut avoir un préfixe de réseau de 18 bits, et un troisième un préfixe de réseau de 14 bits.

Les routeurs DOIVENT prendre en charge des préfixes de réseau de longueurs variables à la fois dans leurs configurations d'interface et dans leurs bases de données d'acheminement.

4.3 Protocole de message de commande de l'Internet - ICMP

4.3.1 Introduction

ICMP est un protocole auxiliaire, qui fournit des fonctions d'acheminement, de diagnostic et d'erreur pour IP. Il est décrit dans [INTER:8]. Un routeur DOIT prendre en charge ICMP.

Les messages ICMP sont groupés en deux classes qui sont exposées dans les paragraphes suivants :

Messages d'erreur ICMP :

Destination inaccessible	paragraphe 4.3.3.1
Redirection	paragraphe 4.3.3.2
Source disparue	paragraphe 4.3.3.3
Temps dépassé	paragraphe 4.3.3.4
Problème de paramètre	paragraphe 4.3.3.5

Messages d'interrogation ICMP :

Echo	paragraphe 4.3.3.6
Information	paragraphe 4.3.3.7
Horodatage	paragraphe 4.3.3.8
Gabarit d'adresse	paragraphe 4.3.3.9
Découverte du routeur	paragraphe 4.3.3.10

Les exigences générales et l'exposé d'ICMP sont dans la section suivante.

4.3.2 Questions générales

4.3.2.1 Types de message inconnus

Si un message ICMP de type inconnu est reçu, il DOIT être passé à l'interface d'utilisateur ICMP (si le routeur en a une) ou éliminé en silence (si le routeur n'en a pas).

4.3.2.2 TTL de message ICMP

Lorsqu'il est à l'origine d'un message ICMP, le routeur DOIT initialiser la TTL. La TTL pour les réponses ICMP ne doit pas être tirée du paquet qui a déclenché la réponse.

4.3.2.3 En-tête du message d'origine

Autrefois, chaque message d'erreur ICMP incluait l'en-tête Internet et au moins les huit premiers octets de données du datagramme qui avait déclenché l'erreur. Ceci n'a plus cours, du fait de l'utilisation du tunnelage IP dans IP et autres technologies. Donc, le datagramme ICMP DEVRAIT contenir autant qu'il est possible du datagramme original sans que la longueur du datagramme ICMP excède 576 octets. L'en-tête IP retourné (et les données d'utilisateur) DOIVENT être identiques à ce qui a été reçu, sauf que le routeur n'est pas obligé de défaire les modifications de l'en-tête IP qui sont normalement effectuées en transmettant ce qui a été effectué avant la détection de l'erreur (par exemple, décrémenter la TTL, ou mettre à jour les options). Noter que les exigences du paragraphe 4.3.3.5 supplantent cette exigence dans certains cas (c'est-à-dire, pour un message de problème de paramètre, si le problème est dans un champ modifié, le routeur doit défaire la modification). Voir au paragraphe 4.3.3.5).

4.3.2.4 Adresse de source de message ICMP

Excepté sur spécification contraire du présent document, l'adresse de source IP dans un message ICMP dont l'origine est le routeur DOIT être une des adresses IP associées à l'interface physique sur laquelle le message ICMP est transmis. Si l'interface n'a pas d'adresse IP associée, on utilise à la place l'identifiant du routeur (voir au paragraphe 5.2.5).

4.3.2.5 TOS et préséance

Les messages d'erreur ICMP DEVRAIT avoir leurs bits de TOS réglés à la même valeur que les bits de TOS dans le paquet qui a provoqué l'envoi du message d'erreur ICMP, sauf si le réglage à cette valeur causerait l'élimination immédiate du message d'erreur ICMP à cause de l'impossibilité de l'acheminer à sa destination. Autrement, les messages d'erreur ICMP DOIVENT être envoyés avec un TOS normal (c'est-à-dire, zéro). Un message de réponse ICMP DEVRAIT avoir ses bits de TOS réglés à la même valeur que les bits de TOS dans la demande ICMP qui a provoqué la réponse.

Les messages d'erreur ICMP Source disparue, s'ils sont envoyés, DOIVENT avoir leur champ Préséance IP réglé à la même valeur que le champ Préséance IP dans le paquet qui avait provoqué l'envoi du message ICMP Source disparue. Tous les autres messages d'erreur ICMP (Destination inaccessible, Redirection, Temps dépassé, et Problème de paramètre) DEVRAIENT avoir leur valeur de préséance réglée à 6 (INTERNETWORK CONTROL) ou 7 (NETWORK CONTROL). La valeur de Préséance IP pour ces messages d'erreur PEUT être réglable.

Un message de réponse ICMP DOIT avoir son champ Préséance IP réglé à la même valeur que le champ Préséance IP dans la demande ICMP qui avait provoqué la réponse.

4.3.2.6 Route de source

Si le paquet qui a provoqué l'envoi d'un message d'erreur ICMP contient une option de route de source, le message d'erreur ICMP DEVRAIT aussi contenir une option de route de source du même type (strict ou lâche), créée en inversant la portion se trouvant avant le pointeur du chemin enregistré dans l'option de route de source du paquet d'origine SAUF SI le message d'erreur ICMP est un Problème de paramètre ICMP qui se plaint d'une option de route de source dans le paquet d'origine, ou si le routeur est averti d'une politique qui empêcherait la livraison du message d'erreur ICMP.

Discussion

Dans des environnements qui utilisent l'option de sécurité du Ministère de la Défense des U.S.A (définie dans [INTER:5]), les messages ICMP peuvent avoir besoin d'inclure une option de sécurité. Des informations précises sur ce sujet devraient être disponibles auprès de l'Agence de communication de la Défense.

4.3.2.7 Quand ne pas envoyer d'erreurs ICMP

Au message d'erreur ICMP NE DOIT PAS être envoyé à la suite de la réception :

- d'un message d'erreur ICMP,
- d'un paquet qui échoue aux essais de validation d'en-tête IP décrits au paragraphe 5.2.2 (excepté lorsque ce paragraphe permet spécifiquement l'envoi d'un message d'erreur ICMP),
- d'un paquet destiné à une adresse de diffusion IP ou de diffusion de groupe IP,
- d'un paquet envoyé comme diffusion ou diffusion groupée de couche Liaison,
- d'un paquet dont l'adresse de source a un préfixe de réseau de zéro ou est une adresse de source invalide (comme défini au paragraphe 5.3.7),
- de tout fragment de datagramme autre le premier fragment (c'est-à-dire, un paquet pour lequel le fragment copié dans l'en-tête IP est différent de zéro).

De plus, un message d'erreur ICMP NE DOIT PAS être envoyé dans tous les cas où le présent mémoire établit qu'un paquet est à éliminer en silence.

NOTE : CES RESTRICTIONS PRENNENT LE PAS SUR TOUTE EXIGENCE FIGURANT AILLEURS DANS LE PRÉSENT DOCUMENT POUR L'ENVOI DES MESSAGES D'ERREUR ICMP.

Discussion

Ces règles visent à empêcher les tempêtes de diffusion qui résultent du retour par les routeurs ou hôtes des messages d'erreur ICMP en réponse à des paquets en diffusion. Par exemple, un paquet UDP en diffusion à un accès qui n'existe pas peut déclencher un flot de datagrammes ICMP Destination inaccessible de la part de tous les appareils qui n'ont pas de client pour ce port de destination. Sur un grand Ethernet, les collisions résultantes peuvent rendre le réseau inutilisable pendant une seconde ou plus.

Chaque paquet qui est diffusé sur le réseau connecté devrait avoir une adresse IP valide comme destination IP (voir au paragraphe 5.3.4 et [INTRO:2]). Cependant, certains appareils violent cette règle. Pour être certain de détecter les paquets en diffusion, il est donc exigé des routeurs qu'ils vérifient une adresse de diffusion de couche Liaison aussi bien que de couche IP.

MISE EN ŒUVRE+ Ceci exige que la couche Liaison informe la couche IP lorsque un paquet en diffusion de couche de liaison est reçu ; voir au paragraphe 3.1.

4.3.2.8 Limitation de débit

Un routeur qui envoie des messages ICMP Source éteinte DOIT être capable de limiter le débit auquel les messages peuvent être générés. Un routeur DEVRAIT aussi être capable de limiter le débit auquel il envoie d'autres sortes de messages d'erreur ICMP (Destination inaccessible, Redirection, Temps dépassé, Problème de paramètre). Les paramètres de limite de débit DEVRAIENT être réglables au titre de la configuration du routeur. La façon d'appliquer les limites (par exemple, par routeur ou par interface) est laissée à la discrétion de la mise en œuvre.

Discussion

Deux problèmes d'un routeur qui envoie un message d'erreur ICMP sont :

(1) La consommation de bande passante sur le chemin de retour, et
(2) L'utilisation des ressources du routeur (par exemple, mémoire, heure de CPU)

Pour aider à résoudre ces problèmes, un routeur peut limiter la fréquence à laquelle il génère les messages d'erreur ICMP. Pour des raisons similaires, un routeur peut limiter la fréquence à laquelle sont générées certaines autres sortes de messages, comme les réponses d'écho ICMP.

MISE EN ŒUVRE

Divers mécanismes on été utilisés ou proposés pour limiter le débit auquel les messages ICMP sont envoyés :

(1) Fondés sur le comptage - Par exemple, envoyer un message d'erreur ICMP pour chaque N paquets abandonnés globalement ou par hôte de source. Ce mécanisme pourrait être approprié pour un message ICMP Source éteinte, s'il est utilisé, mais probablement pas pour d'autres types de messages ICMP.

(2) Fondés sur le temps - Par exemple, envoi d'un message d'erreur ICMP à un hôte de source donnée ou globale au moins une fois toutes les T millisecondes.

(3) Fondés sur la bande passante - Par exemple, limiter le débit auquel les messages ICMP sont envoyés sur une interface particulière à une certaine fraction de la bande passante du réseau de rattachement.

4.3.3 Questions spécifiques

4.3.3.1 Destination inaccessible

Si un routeur ne peut pas transmettre un paquet parce qu'il n'a pas du tout de chemins (y compris de chemin par défaut) pour la destination spécifiée dans le paquet, le routeur DOIT alors générer un message ICMP Destination inaccessible, Code 0 (Réseau inaccessible). Si le routeur n'a pas de chemin pour le réseau de destination spécifié dans le paquet mais si le TOS spécifié pour les chemins n'est ni le TOS par défaut (0000) ni le TOS du paquet que le routeur essaye d'acheminer, le routeur DOIT alors générer un message ICMP Destination inaccessible, Code 11 (Réseau inaccessible pour TOS).

Si un paquet est à transmettre à un hôte sur un réseau qui est directement connecté au routeur (c'est-à-dire, le routeur est le routeur du dernier bond) et si le routeur a la certitude qu'il n'y a pas de chemin vers l'hôte de destination, le routeur DOIT alors générer un message ICMP Destination inaccessible, Code 1 (Hôte inaccessible). Si un paquet est à transmettre à un hôte qui est sur un réseau directement connecté au routeur et si le routeur ne peut pas transmettre le paquet parce qu'aucun chemin vers la destination n'a un TOS égal au TOS demandé dans le paquet ou au TOS par défaut (0000) le routeur DOIT alors générer un message ICMP Destination inaccessible, Code 12 (Hôte inaccessible pour TOS).

Discussion

L'objectif est qu'un routeur génère le hôte/réseau inaccessible "générique" si il n'a pas de chemin du tout (y compris de chemin par défaut) pour la destination. Si le routeur a un ou plusieurs chemins pour la destination, mais qu'aucun de ces chemins n'a un TOS acceptable, le routeur génère alors le message "inaccessible pour TOS".

4.3.3.2 Redirection

Le message ICMP Redirection est généré pour informer un hôte local qu'il devrait utiliser un routeur de prochain bond différent pour certain trafic.

A la différence de [INTRO:2], un routeur PEUT ignorer les messages ICMP Redirection lorsqu'il choisit un chemin pour un paquet dont l'origine est le routeur si le routeur utilise un protocole d'acheminement ou si la retransmission est activée sur le routeur et sur l'interface sur laquelle le paquet est envoyé.

4.3.3.3 Source éteinte

Un routeur NE DEVRAIT PAS être à l'origine de messages ICMP Source éteinte. Comme spécifié au paragraphe 4.3.2, un routeur qui est à l'origine de messages Source éteinte DOIT être capable de limiter le débit de leur émission.

Discussion

Les recherches semblent suggérer que Source éteinte consomme de la bande passante du réseau mais est un antidote inefficace (et déloyal) à l'encombrement. Voir, par exemple, [INTER:9] et [INTER:10]. Le paragraphe 5.3.6 discute des idées actuelles sur la façon dont les routeurs devraient traiter les surcharges et l'encombrement du réseau.

Un routeur PEUT ignorer tous les messages ICMP Source éteinte qu'il reçoit.

Discussion

Un routeur peut lui-même recevoir un message Source éteinte par suite de l'émission d'un paquet envoyé à un autre routeur ou hôte. De tels datagrammes pourraient être, par exemple, une mise à jour EGP envoyée à un autre routeur, ou un flux telnet envoyé à un hôte. Un mécanisme a été proposé ([INTER:11], [INTER:12]) pour que la couche IP réponde directement à Source éteinte en contrôlant le débit auquel les paquets sont envoyés, cependant, cette proposition est actuellement expérimentale et non recommandée.

4.3.3.4 Temps dépassé

Lorsqu'un routeur transmet un paquet et que le champ TTL est réduit à 0, les exigences du paragraphe 5.2.3.8 s'appliquent.

Lorsque le routeur réassemble un paquet qui est destiné au routeur, il agit comme un hôte Internet. Les exigences de réassemblage de [INTRO:2] s'appliquent donc.

Lorsque le routeur reçoit (c'est-à-dire, est destinataire comme routeur) d'un message Temps dépassé, il DOIT se conformer à [INTRO:2].

4.3.3.5 Problème de paramètre

Un routeur DOIT générer un message Problème de paramètre pour toute erreur non spécifiquement couverte par un autre message ICMP. Le champ d'en-tête IP ou option qui comporte l'octet indiqué par le champ de pointeur DOIT être inclus inchangé dans l'en-tête IP retourné avec ce message ICMP. Le paragraphe 4.3.2 définit une exception à cette exigence.

Une nouvelle variante du message Problème de paramètre a été définie dans [INTRO:2] :

Code 1 = l'option demandée manque.

Discussion

Cette variante est actuellement en usage dans la communauté militaire pour une option de sécurité manquante.

4.3.3.6 Demande/Réponse d'écho

Un routeur DOIT mettre en œuvre une fonction de serveur d'écho ICMP qui reçoive les demandes d'écho envoyées au routeur, et envoie les réponses d'écho correspondantes. Un routeur DOIT être prêt à recevoir, réassembler et faire écho à un datagramme de demande d'écho ICMP au moins comme le maximum de 576 et des MTU de tous les réseaux connectés.

La fonction de serveur d'écho PEUT choisir de ne pas répondre aux demandes d'écho ICMP adressées aux adresses IP en diffusion ou diffusion groupée.

Un routeur DEVRAIT avoir une option de configuration qui, si activée, amène le routeur à ignorer en silence toutes les demandes d'écho ICMP ; si elle est fournie, cette option DOIT permettre les réponses par défaut.

Discussion

La disposition neutre en matière de réponse d'écho en diffusion et diffusion groupée découle du paragraphe "Demande/Réponse d'écho" de [INTRO:2].

Comme expliqué au paragraphe 10.3.3, un routeur DOIT aussi mettre en œuvre une interface de couche utilisateur/application pour l'envoi d'une Demande d'écho et la réception d'une Réponse d'écho, pour les besoins du diagnostic. Tous les messages de réponse d'écho ICMP DOIVENT être passés à cette interface.

L'adresse IP de source dans une Réponse d'écho ICMP DOIT être la même que l'adresse de destination spécifique du message Demande d'écho ICMP correspondant.

Les données reçues dans une Demande d'écho ICMP DOIVENT être entièrement incluses dans la réponse d'écho résultante.

Si une option Record Route et/ou Horodatage est reçue dans une Demande d'écho ICMP, cette/ces options DEVRAIENT être mises à jour pour inclure le routeur actuel et incluses dans l'en-tête IP du message de réponse d'écho, sans coupure. Et donc, le chemin enregistré sera pour l'aller-retour total.

Si une option de route de source est reçue dans une Demande d'écho ICMP, le chemin de retour DOIT être inversé et utilisé comme option de route de source pour le message de réponse d'écho, sauf si le routeur est au courant d'une politique qui empêcherait la livraison du message.

4.3.3.7 Demande/réponse d'informations

Un routeur NE DEVRAIT PAS être à l'origine de ces messages ou leur répondre.

Discussion

La paire Demande/réponse d'informations était destinée à prendre en charge les systèmes auto configurables tels que les stations de travail sans disque, pour leur permettre de découvrir leurs préfixes de réseau IP au moment de l'amorçage. Cependant, ces messages sont maintenant obsolètes. Les protocoles RARP et BOOTP fournissent de meilleurs mécanismes à un hôte pour découvrir sa propre adresse IP.

4.3.3.8 Horodatage et réponse d'horodatage

Un routeur PEUT mettre en œuvre l'horodatage et la réponse d'horodatage. Si elles sont mises en œuvre, alors :

- La fonction de serveur d'horodatage ICMP DOIT retourner une réponse d'horodatage à chaque message d'horodatage reçue. Elle DEVRAIT être conçue pour une variabilité de délai minimum.

- Un message de demande d'horodatage ICMP à une adresse IP en diffusion ou en diffusion groupée PEUT être éliminée en silence.

- L'adresse IP de source dans une réponse d'horodatage ICMP DOIT être la même que l'adresse de destination

spécifique du message de demande d'horodatage correspondant.

- Si une option de route de source est reçue dans une demande d'horodatage ICMP, le chemin de retour DOIT être inversé et utilisé comme option de route de source pour le message de réponse d'horodatage, sauf si le routeur est au courant d'une politique qui empêcherait la délivrance du message.

- Si une option Record Route et/ou horodatage est reçue dans une demande d'horodatage, cette ou ces options DEVRAIENT être mises à jour pour inclure le routeur actuel et incluses dans l'en-tête IP du message de réponse d'horodatage.

- Si le routeur fournit une interface de couche application pour l'envoi de messages de demande d'horodatage, les messages de réponse d'horodatage entrants DOIVENT alors être passés à l'interface d'utilisateur ICMP.

La forme préférée pour une valeur d'horodatage (la valeur standard) est celle de millisecondes à partir de minuit, en temps universel. Cependant, il peut être difficile de fournir cette valeur avec une résolution à la milliseconde. Par exemple, de nombreux systèmes utilisent des horloges qui se mettent à jour seulement à la fréquence de la ligne, 50 ou 60 fois par seconde. Donc, une certaine latitude est permise dans une valeur standard :

(a) Une valeur standard DOIT être mise à jour au moins 16 fois par seconde (c'est-à-dire, au plus les six bits de moindre poids de la valeur peuvent être indéfinis).

(b) La précision d'une valeur standard DOIT approximer celle des horloges CPU gérées par les opérateurs, c'est-à-dire, correcte en quelques minutes.

MISE EN ŒUVRE

Pour satisfaire à la seconde condition, un routeur peut avoir besoin d'interroger un serveur horaire lors de l'amorçage ou du redémarrage du routeur. Il est recommandé que le protocole de serveur de temps UDP soit utilisé à cette fin. Une mise en œuvre plus évoluée devrait utiliser le Protocole de l'heure du réseau (NTP) pour réaliser une synchronisation d'horloge à la milliseconde près ; cependant, ceci n'est pas exigé.

4.3.3.9 Demande/réponse de gabarit d'adresse

Un routeur DOIT mettre en œuvre la prise en charge de la réception des messages ICMP de demande de gabarit d'adresse et leur réponse par des messages ICMP de réponse de gabarit d'adresse. Ces messages sont définis dans [INTER:2].

Un routeur DEVRAIT avoir une option de configuration pour chaque interface logique, spécifiant si le routeur est autorisé à répondre aux demandes de gabarit d'adresse pour cette interface ; cette option DOIT autoriser les réponses par défaut. Un routeur NE DOIT PAS répondre à une demande de gabarit d'adresse avant que le routeur ne connaisse le gabarit d'adresse correct.

Un routeur NE DOIT PAS répondre à une demande de gabarit d'adresse qui a une adresse de source de 0.0.0.0 et qui arrive sur une interface physique associée à plusieurs interfaces logiques dont les gabarits d'adresse ne sont pas tous les mêmes.

Un routeur DEVRAIT examiner toutes les réponses de gabarit d'adresse ICMP qu'il reçoit pour déterminer si les informations qu'il contient correspondent à la connaissance du routeur du gabarit d'adresse. Si la réponse de gabarit d'adresse ICMP paraît être erronée, le routeur DEVRAIT enregistrer le gabarit d'adresse et l'adresse IP de l'envoyeur. Un routeur NE DOIT PAS utiliser le contenu d'une réponse de gabarit d'adresse ICMP pour déterminer le gabarit d'adresse correct.

Comme les hôtes peuvent n'être pas capables d'acquérir le gabarit d'adresse si un routeur est en dérangement alors que l'hôte s'amorce, un routeur PEUT diffuser une réponse de gabarit d'adresse ICMP gratuite sur chacune de ses interfaces logiques après avoir configuré ses propres gabarits d'adresse. Cependant, cette caractéristique peut être dangereuse dans des environnements qui utilisent des gabarits d'adresse de longueurs variables. Donc, si cette caractéristique est mise en œuvre, des réponses de gabarit d'adresse gratuites NE DOIVENT PAS être diffusées sur une ou des interfaces logiques qui :

- ne sont pas configurées pour envoyer des réponses de gabarit d'adresse gratuites. Chaque interface logique DOIT avoir un paramètre de configuration qui contrôle cela, et ce paramètre DOIT par défaut de ne pas envoyer les réponses de gabarit d'adresse gratuites. Ou alors,
- partager des préfixes de réseau de même genre (mais pas identiques) et une interface physique.

La forme { <Préfixe-de-réseau>, -1 } de l'adresse IP de diffusion DOIT être utilisée pour les réponses de gabarit d'adresse de diffusion.

Discussion

La capacité à désactiver l'envoi des réponses de gabarit d'adresse par les routeurs est exigée par quelques sites qui mentent intentionnellement à leurs hôtes au sujet du gabarit d'adresse. On suppose que le besoin de cette caractéristique va disparaître avec la conformité croissante des hôtes aux normes sur les exigences pour les hôtes.

La raison à la fois du second tiret ci-dessus et de l'exigence sur les adresses IP en diffusion à utiliser est d'empêcher les problèmes lorsque plusieurs préfixes de réseau IP sont utilisés sur le même réseau physique.

4.3.3.10 Publicité et sollicitations du routeur

Un routeur IP DOIT prendre en charge la partie routeur du protocole de découverte de routeur ICMP [INTER:13] sur tous les réseaux connectés sur lesquels le routeur prend en charge l'adressage IP en diffusion groupée ou en diffusion. La mise en œuvre DOIT inclure toutes les variables de configuration spécifiées pour les routeurs, avec les valeurs par défaut spécifiées.

Discussion

Les routeurs ne sont pas obligés de mettre en œuvre la partie hôte du protocole de découverte de routeur ICMP, mais pourraient le trouver utile pour le fonctionnement lorsque la transmission IP est désactivée (c'est-à-dire, en fonctionnant comme hôte).

Discussion

On note qu'il est assez courant pour les hôtes d'utiliser RIP version 1 comme protocole de découverte de routeur. De tels hôtes écoutent le trafic RIP et utilisent les informations extraites de ce trafic pour découvrir les routeurs et prendre des décisions sur le routeur à utiliser comme routeur de premier bond pour une destination donnée. Bien que ce comportement soit déconseillé, il est toujours courant et les développeurs devraient en être conscients.

4.4 Protocole de gestion de groupe Internet - IGMP

IGMP [INTER:4] est un protocole utilisé entre les hôtes et les routeurs en diffusion groupée sur un seul réseau physique pour établir l'adhésion des hôtes dans des groupes de diffusion groupée particuliers. Les routeurs de diffusion groupée utilisent cette information, en conjonction avec un protocole d'acheminement de diffusion groupée, pour prendre en charge la transmission IP en diffusion groupée à travers l'Internet.

Un routeur DEVRAIT mettre en œuvre la partie hôte de IGMP.

5 Couche INTERNET - Transmission

5.1 Introduction

La présente section décrit le processus de transmission des paquets.

5.2 Généralités sur la transmission

Il n'y a pas de spécification séparée pour la fonction de transmission dans IP. Au lieu de cela, la transmission est couverte par les spécifications des protocoles de la couche internet ([INTER:1], [INTER:2], [INTER:3], [INTER:8], et [ROUTE:11]).

5.2.1 Algorithme de transmission

Comme aucun des principaux documents de protocole ne décrit en détail l'algorithme de transmission, il est présenté ici. Il s'agit simplement d'une présentation générale, qui omet des détails importants, tels que le traitement de l'encombrement, qui sera vu dans des sections ultérieures.

Il n'est pas exigé qu'une mise en œuvre suive exactement les algorithmes donnés dans les paragraphes 5.2.1.1, 5.2.1.2, et 5.2.1.3. Le plus gros défi de la rédaction d'un logiciel de routeur est de maximiser le débit auquel le routeur peut transmettre les paquets tout en réalisant le même effet de l'algorithme. Les détails sur la façon de le faire sortent du domaine d'application du présent document, en partie parce que ils sont très dépendants de l'architecture du routeur. On se contentera de souligner la dépendance à l'ordre des étapes :

(1) Un routeur DOIT vérifier l'en-tête IP, comme décrit au paragraphe 5.2.2, avant d'effectuer aucune action fondée

sur le contenu de l'en-tête. Ceci permet au routeur de détecter et éliminer les mauvais paquets avant la consommation d'autres ressources.

(2) Le traitement de certaines des options IP exige que le routeur insère son adresse IP dans l'option. Comme noté au paragraphe 5.2.4, l'adresse insérée DOIT être l'adresse de l'interface logique sur laquelle le paquet est envoyé, ou l'identifiant du routeur si le paquet est envoyé sur une interface non numérotée. Et donc, le traitement de ces options ne peut pas être terminé avant que l'interface de sortie ait été choisie.

(3) Le routeur ne peut pas vérifier et décrémenter la TTL avant d'avoir vérifié si le paquet devrait être livré au routeur lui-même, pour les raisons mentionnées au paragraphe 4.2.2.9.

(4) Plus généralement, lorsque un paquet est délivré localement au routeur, son en-tête IP NE DOIT être modifié en aucune façon (excepté qu'il peut être demandé à un routeur d'insérer un horodatage dans toutes les options Timestamp de l'en-tête IP). Et donc, avant que le routeur ne détermine si le paquet est à délivrer localement au routeur, il ne peut mettre à jour l'en-tête IP d'aucune façon qu'il ne soit prêt à défaire.

5.2.1.1 Généralités

Ce paragraphe traite de l'algorithme général de transmission. Cet algorithme s'applique à toutes les formes de paquets à transmettre : en envoi individuel, en diffusion groupée, et en diffusion.

(1) Le routeur reçoit le paquet IP (plus des informations supplémentaires à son sujet, comme décrit au paragraphe 3.1) de la couche de liaison.

(2) Le routeur valide l'en-tête IP, comme décrit au paragraphe 5.2.2. Noter que le réassemblage IP n'est pas fait, excepté sur les fragments IP à mettre en file d'attente pour une livraison locale à l'étape (4).

(3) Le routeur effectue la plupart du traitement de toutes les options IP. Comme décrit au paragraphe 5.2.4, certaines des options IP exigent un traitement additionnel après la prise de décision d'acheminement.

(4) Le routeur examine l'adresse IP de destination du datagramme IP, comme décrit au paragraphe 5.2.3, pour déterminer comment il devrait continuer le traitement du datagramme IP. Il y a trois possibilités :
- Le datagramme IP est destiné au routeur, et devrait être mis en file d'attente pour une livraison locale, en faisant le réassemblage si nécessaire.
- Le datagramme IP n'est pas destiné au routeur, et devrait être mis en file d'attente pour transmission.
- Le datagramme IP devrait être mis en file d'attente pour transmission, mais (une copie) doit aussi être mise en file d'attente pour livraison locale.

5.2.1.2 Envoi individuel

Comme le cas de la livraison locale est bien couvert par [INTRO:2], ce qui suit suppose que le datagramme IP a été mis en file d'attente pour transmission. Si la destination est une adresse IP en envoi individuel :

(5) Le transmetteur détermine l'adresse IP du prochain bond pour le paquet, normalement en regardant la destination du paquet dans le tableau d'acheminement du routeur. Cette procédure est décrite plus en détail au paragraphe 5.2.4. Cette procédure décide aussi quelle interface de réseau devrait être utilisée pour envoyer le paquet.

(6) Le transmetteur vérifie que la transmission du paquet est permise. Les adresses de source et de destination devraient être valides, comme décrit aux paragraphes 5.3.7 et 5.3.4. Si le routeur accepte les contraintes administratives sur la transmission, telles que décrites au paragraphe 5.3.9, ces contraintes doivent être satisfaites.

(7) Le transmetteur décrémente (au moins d'un) et vérifie la TTL du paquet, comme décrit au paragraphe 5.3.1.

(8) Le transmetteur effectue tous les traitements d'option IP qui n'auraient pu être achevés à l'étape 3.

(9) Le transmetteur effectue toute fragmentation IP nécessaire, comme décrit au paragraphe 4.2.2.7. Comme cette étape survient après le choix de l'interface de sortie (étape 5), tous les fragments du même datagramme seront transmis sur la même interface.

(10) Le transmetteur détermine l'adresse de couche Liaison du prochain bond du paquet. Les mécanismes pour ce faire dépendent de la couche Liaison (voir à la section 3).

(11) Le transmetteur encapsule le datagramme IP (ou chacun de ses fragments) dans une trame de couche Liaison appropriée et le place en file d'attente pour émission sur l'interface choisie à l'étape 5.

(12) Le transmetteur envoie une redirection ICMP si nécessaire, comme décrit au paragraphe 4.3.3.2.

5.2.1.3 Diffusion groupée

Si la destination est une diffusion groupée IP, on suivra les étapes ci-après.

Noter que les principales différences entre la transmission des envois individuels IP et la transmission IP en diffusion groupée sont :

- Les diffusions groupées IP sont normalement transmises sur la base des adresses IP à la fois de source et de destination du datagramme,
- la diffusion groupée IP utilise une recherche en anneau expansive,
- les diffusions groupées IP sont transmises comme diffusions groupées de niveau Liaison, et
- les erreurs ICMP ne sont jamais envoyées en réponse à des datagrammes IP en diffusion groupée.

Noter que la transmission de diffusions groupées IP est toujours un peu expérimentale. Il en résulte que l'algorithme présenté ci-dessous n'est pas obligatoire, et n'est fourni qu'à titre d'exemple.

(5a) Sur la base des adresses IP de source et de destination trouvées dans l'en-tête du datagramme, le routeur détermine si le datagramme a été reçu sur l'interface appropriée pour la transmission. Sinon, le datagramme est éliminé en silence. La méthode de détermination de l'interface de réception appropriée dépend du ou des algorithmes d'acheminement de diffusion groupée utilisés. Dans un des plus simples algorithmes, la transmission sur chemin inverse (RPF, *reverse path forwarding*), l'interface appropriée est celle qui serait utilisée pour transmettre des envois individuels en retour à la source du datagramme.

(6a) Sur la base des adresses IP de source et de destination trouvées dans l'en-tête du datagramme, le routeur détermine les interfaces de sortie du datagramme. Pour mettre en œuvre la recherche en anneau expansive de la diffusion groupée IP (voir [INTER:4]) une valeur minimum de TTL est spécifiée pour chaque interface de sortie. Une copie du datagramme en diffusion groupée est transmise sur chaque interface de sortie dont la valeur de TTL minimum est inférieure ou égale à la valeur de TTL figurant dans l'en-tête du datagramme, en appliquant séparément les étapes restantes à chacune de ces interfaces.

(7a) Le routeur décrémente de un la TTL du paquet.

(8a) Le transmetteur effectue tous les traitements d'option IP qui n'auraient pu être achevés à l'étape (3).

(9a) Le transmetteur effectue toute fragmentation IP nécessaire, comme décrit au paragraphe 4.2.2.7.

(10a) Le transmetteur détermine l'adresse de couche Liaison à utiliser dans l'encapsulation de niveau Liaison. Les mécanismes pour ce faire dépendent de la couche Liaison. Sur les LAN, une diffusion ou diffusion groupée de niveau Liaison est sélectionnée, comme une traduction algorithmique de l'adresse de diffusion groupée IP du datagramme. Voir les diverses spécifications IP-sur-xxx pour des précisions.

(11a) Le transmetteur encapsule le paquet (ou chacun de ses fragments) dans une trame de couche Liaison appropriée et la met en file d'attente de sortie sur l'interface appropriée.

5.2.2 Validation d'en-tête IP

Avant qu'un routeur ne puisse traiter aucun paquet IP, il DOIT effectuer les vérifications de validité de base suivantes sur l'en-tête IP du paquet pour s'assurer que l'en-tête est significatif. Si le paquet échoue à l'un des essais suivants, il DOIT être éliminé en silence, et l'erreur DEVRAIT être enregistrée.

(1) La longueur du paquet rapportée par la couche Liaison doit être suffisante pour contenir la longueur minimale légale de datagramme IP (20 octets).

(2) La somme de contrôle IP doit être correcte.

(3) Le numéro de version IP doit être 4. Si le numéro de version n'est pas 4, le paquet peut alors relever d'une autre version d'IP, telle que IPng ou ST-II.

(4) Le champ de longueur d'en-tête IP doit être suffisant pour contenir la longueur minimale légale de datagramme IP (20 octets = 5 mots).

(5) Le champ de longueur totale IP doit être suffisant pour contenir l'en-tête de datagramme IP, dont la longueur est spécifiée dans le champ de longueur d'en-tête IP.

Un routeur NE DOIT PAS avoir une option de configuration qui permette la désactivation de ces essais.

Si le paquet passe le second et le troisième essai, si le champ de longueur d'en-tête IP est d'au moins 4, et si le champ de longueur totale IP et la longueur du paquet rapportée par la couche Liaison sont toutes deux d'au moins 16, alors, malgré la règle précédente, le routeur PEUT répondre par un message Problème de paramètre ICMP, dont le pointeur pointe sur le champ de longueur d'en-tête IP (si il a échoué au quatrième essai) ou sur le champ de longueur totale IP (si il a échoué au cinquième essai). Cependant, il DOIT encore éliminer le paquet et DEVRAIT encore enregistrer l'erreur.

Ces règles (et le présent document tout entier) n'appliquent que la version 4 du protocole Internet. Ces règles ne devraient pas être perçues comme interdisant aux routeurs de prendre en charge les autres versions d'IP. De plus, si un routeur peut vraiment classer un paquet comme relevant d'une autre version d'IP, il ne devrait alors pas traiter ce paquet

comme paquet erroné dans le contexte du présent mémoire.

MISE EN ŒUVRE

Il est souhaitable pou ce qui concerne le rapport des erreurs, bien que pas toujours entièrement possible, de déterminer pourquoi un en-tête était invalide. Il y a quatre raisons possibles :

- la couche Liaison a tronqué l'en-tête IP
- le datagramme utilise une version d'IP autre que la version standard (version 4).
- l'en-tête IP a été corrompu dans le transit.
- l'envoyeur a généré un en-tête IP illégal.

Il est probablement souhaitable d'effectuer les vérifications dans l'ordre de la liste, car nous pensons que cet ordre a le plus de chances de catégoriser correctement la cause de l'erreur. Pour les besoins du rapport d'erreur, il peut aussi être souhaitable de vérifier si les paquets qui ont échoué à ces essais ont un numéro de version IP qui indique IPng ou ST-II ; ils devraient alors être traités conformément à leurs spécifications respectives.

De plus, le routeur DEVRAIT vérifier que la longueur de paquet rapportée â la couche de liaison est au moins aussi grande que la longueur IP totale enregistrée dans l'en-tête IP du paquet. Si il apparaît que le paquet a été tronqué, le paquet DOIT être éliminé, l'erreur DEVRAIT être enregistrée, et le routeur DEVRAIT répondre par un message Problème de paramètre ICMP dont le pointeur pointe sur le champ longueur totale IP.

Discussion

Comme tout protocole de couche supérieure qui est concerné par la corruption de données va détecter la troncature des données du paquet lorsqu'il atteindra sa destination finale, il n'est pas absolument nécessaire que les routeurs effectuent la vérification évoquée ci-dessus pour le maintien d'un protocole correct. Cependant, en effectuant cette vérification, un routeur peut simplifier considérablement la tâche de détermination du bond qui tronque les paquets sur le chemin. Cela réduira aussi la dépense en ressources vers l'aval du routeur qui dans ces systèmes aval n'auront pas besoin de traiter le paquet.

Finalement, si l'adresse de destination dans l'en-tête IP n'est pas une des adresses du routeur, le routeur DEVRAIT vérifier que le paquet ne contient pas une option Strict Source et Record Route. Si un paquet échoue à cet essai (si il contient une option de route de source stricte), le routeur DEVRAIT enregistrer l'erreur et DEVRAIT répondre par une erreur ICMP Problème de paramètre avec le pointeur pointant sur l'adresse de destination IP du paquet en défaut.

Discussion

Certains pourraient suggérer que le routeur devrait répondre par un message Mauvais chemin de source au lieu du message Problème de paramètre. Cependant, lorsque un paquet échoue à cet essai, il indique habituellement une erreur de protocole du routeur du bond précédent, alors que Mauvais chemin de source suggèrerait que l'hôte de source avait demandé un chemin non existant ou endommagé à travers le réseau.

5.2.3 Décision de livraison locale

Lorsque un routeur reçoit un paquet IP, il doit décider si le paquet est adressé au routeur (et devrait être livré localement) ou si le paquet est adressé à un autre système (et devrait être traité par le transmetteur). Il y a aussi un cas hybride, lorsque certaines diffusions IP et diffusions groupées IP sont toutes deux délivrées localement et transmises. Un routeur DOIT déterminer lequel de ces trois cas s'applique en utilisant les règles suivantes.

- Une option de route de source non expirée est celle dont la valeur de pointeur ne pointe pas au delà de la dernière entrée dans le chemin de source. Si le paquet contient une option de route de source non expirée, le pointeur dans l'option est avancé jusqu'à ce que le pointeur ne pointe pas au delà de la dernière adresse dans l'option ou alors que la prochaine adresse ne soit plus une des propres adresses du routeur. Dans ce dernier cas (normal), le paquet est transmis (et non délivré localement) indépendamment des règles ci-dessous.

- Le paquet est délivré localement et non destiné à la transmission dans les cas suivants :
  - L'adresse de destination du paquet correspond exactement à une des adresses IP du routeur,
  - L'adresse de destination du paquet est une adresse de diffusion limitée ({-1, -1}), ou
  - La destination du paquet est une adresse IP en diffusion groupée qui n'est jamais transmise (comme 224.0.0.1 ou 224.0.0.2) et (au moins) une des interfaces logiques associées à l'interface physique sur laquelle est arrivé le paquet est un membre du groupe de diffusion de destination.

- Le paquet est passé au transmetteur ET délivré localement dans les cas suivants :
  - L'adresse de destination du paquet est une adresse de diffusion IP qui s'adresse à au moins une des interfaces logiques du routeur mais ne s'adresse à aucune des interfaces logiques associées à l'interface physique sur

laquelle est arrivé le paquet
- La destination du paquet est une adresse IP en diffusion groupée qui a la permission d'être transmise (à la différence de 224.0.0.1 et 224.0.0.2) et une (au moins) des interfaces logiques associées à l'interface physique sur laquelle est arrivé le paquet est un membre du groupe de diffusion groupée de destination.

- Le paquet est délivré localement si l'adresse de destination du paquet est une adresse de diffusion IP (autre qu'une adresse de diffusion limitée) qui s'adresse à au moins une des interfaces logiques associées à l'interface physique sur laquelle est arrivé le paquet. Le paquet est AUSSI passé au transmetteur sauf si la liaison sur laquelle est arrivé le paquet utilise une encapsulation IP qui n'encapsule pas les diffusions différemment des envois individuels (par exemple, en utilisant des adresses de destination de couche Liaison différentes).

- Le paquet est passé au transmetteur dans tous les autres cas.

Discussion

L'objet de cette exigence de la dernière phrase du quatrième tiret est de traiter le cas d'une diffusion dirigée vers un autre préfixe de réseau sur le même câble physique. Normalement, ceci fonctionne comme prévu : l'envoyeur envoie la diffusion au routeur comme un envoi individuel de couche Liaison. Le routeur note qu'elle est arrivée comme un envoi individuel, et doit donc être destinée à un préfixe de réseau différent de celui sur lequel l'envoyeur l'a expédiée. Donc, le routeur peut l'envoyer en toute sécurité comme une diffusion de couche Liaison sur la même interface (physique) que celle sur laquelle elle est arrivée. Cependant, si le routeur ne peut pas dire si le paquet a été reçu comme envoi individuel de couche Liaison, la phrase propose que le routeur fasse la chose sûre mais fausse plutôt que la pas sûre mais juste.

MISE EN ŒUVRE

Comme décrit au paragraphe 5.3.4, les paquets reçus comme des diffusions de couche Liaison ne sont généralement pas transmis. Il peut être avantageux d'éviter de passer au transmetteur des paquets qu'il devrait ensuite éliminer à cause des règles de ce paragraphe.

Certaines couches Liaison (à cause du matériel ou à cause d'un code particulier dans les pilotes) peuvent délivrer au routeur des copies de toutes les diffusions et diffusions groupées de couche Liaison qu'elles transmettent. L'utilisation de ce dispositif peut simplifier la mise en œuvre des cas où un paquet doit à la fois être passé au transmetteur et délivré localement, car la transmission du paquet causera automatiquement la réception par le routeur d'une copie du paquet qu'il peut alors délivrer localement. On doit dans ces circonstances veiller à empêcher le traitement d'un paquet reçu en boucle comme un paquet reçu normal (et donc soumis aux règles de la transmission, etc.).

Même sans une telle couche de liaison, il est bien sûr très nécessaire de faire une copie d'un paquet entier à mettre en file d'attente à la fois pour transmission et pour livraison locale, bien qu'il faille faire attention avec les fragments, car le réassemblage est effectué sur les paquets livrés localement mais pas sur les paquets transmis. Un schéma simple est d'associer un fanion à chaque paquet sur la file d'attente de sortie du routeur, qui indique si il devrait être mis en file d'attente pour livraison locale après son envoi.

5.2.4 Détermination de l'adresse du prochain bond

Lorsque un routeur s'apprête à transmettre un paquet, il doit déterminer si il peut l'envoyer directement à sa destination, ou si il a besoin de le passer à un autre routeur. Dans ce dernier cas, il a besoin de déterminer quel routeur utiliser. Ce paragraphe explique comment sont faites ces déterminations.

Ce paragraphe utilise les définitions suivantes :
- LSRR – option IP Loose Source and Record Route *(source vague et route enregistrée)*
- SSRR – option IP Strict Source and Record Route *(source stricte et route enregistrée)*
- Option de route de source - LSRR ou SSRR
- Adresse ultime de destination – où le paquet va être envoyé :
  la dernière adresse dans le chemin de source d'un paquet à acheminement de source, ou
  l'adresse de destination dans l'en-tête IP d'un paquet qui n'est pas à acheminement de source
- Adjacent – accessible sans passer à travers aucun routeur IP
- Adresse de prochain bond - adresse IP de l'hôte ou routeur adjacent auquel le paquet devrait ensuite être envoyé
- Adresse IP de destination - adresse de destination ultime, excepté dans les paquets à acheminement de source, chez lesquels c'est la prochaine adresse spécifiée dans le chemin de source
- Destination immédiate - nœud, système, routeur, système final, ou ce qui est visé par l'adresse de destination IP.

5.2.4.1 Adresse de destination IP

Si :

- l'adresse de destination dans l'en-tête IP est une des adresses du routeur,
- le paquet contient une option de route de source, et
- le pointer dans l'option de route de source ne pointe pas après la fin de l'option,

la prochaine adresse de destination IP est l'adresse pointée par le pointeur dans cette option. Si :

- l'adresse de destination dans l'en-tête IP est une des adresses du routeur,
- le paquet contient une option de route de source, et
- le pointeur dans l'option de route de source pointe après la fin de l'option,

le message est alors adressé au système qui analyse le message.

Un routeur DOIT utiliser l'adresse de destination IP, et non l'adresse de destination ultime (la dernière adresse dans l'option de route de source), lorsque il détermine comment traiter un paquet.

L'apparition de plus d'une option de route de source dans un datagramme est une erreur. Si il reçoit un tel datagramme, il DEVRAIT éliminer le paquet et répondre par un message Problème de paramètre ICMP dont le pointeur pointe sur le début de la seconde option de route de source.

5.2.4.2 Décision locale/distante

Après avoir déterminé que le paquet IP doit être transmis conformément aux règles spécifiées au paragraphe 5.2.3, l'algorithme suivant DOIT être utilisé pour déterminer si la destination immédiate est directement accessible (voir [INTER:2]).

(1) Pour chaque interface réseau à laquelle n'a été allouée aucune adresse IP (les lignes non numérotées comme décrit au paragraphe 2.2.7), comparer le routeur-id de l'autre extrémité de la ligne à l'adresse de destination IP. Si ils sont exactement égaux, le paquet peut être transmis à travers cette interface.

Discussion

En d'autres termes, le routeur ou hôte à l'extrémité distante de la ligne est la destination du paquet ou est la prochaine étape dans la route de source d'un paquet à acheminement de source.

(2) Si aucune interface réseau n'a été choisie à la première étape, pour chaque adresse IP allouée au routeur :
   (a) Isoler le préfixe de réseau utilisé par l'interface.

MISE EN ŒUVRE

Le résultat de cette opération aura normalement été calculé et sauvegardé durant l'initialisation.

   (b) Isoler l'ensemble de bits correspondant de l'adresse de destination IP du paquet.
   (c) Comparer les préfixes de réseau résultants. Si ils sont égaux l'un à l'autre, le paquet peut être transmis à travers l'interface de réseau correspondante.

(3) Si la destination n'était ni le routeur-id d'un voisin sur une interface non numérotée ni un membre d'un préfixe de réseau directement connecté, la destination IP n'est accessible qu'à travers un autre routeur. Le choix du routeur et l'adresse IP du prochain bond sont décrits au paragraphe 5.2.4.3. Dans le cas d'un hôte qui n'est pas aussi un routeur, ce peut être le routeur par défaut configuré.

Des travaux en cours à l'IETF [ARCH:9], [NRHP] examinent des cas tels que celui où plusieurs (sous) réseaux IP se recouvrent sur le même réseau de couche de liaison. Sauf restrictions de politique, les hôtes et les routeurs qui utilisent un réseau commun de couche de liaison peuvent communiquer directement même si ils ne sont pas dans le même (sous)réseau IP, si les informations adéquates sont présentes. Le Protocole d'acheminement du prochain bond (NHRP) permet aux entités IP de déterminer l'adresse de couche de liaison "optimale" à utiliser pour traverser un tel réseau de couche de liaison vers une destination distante.

(4) Si le "prochain bond" choisi est accessible à travers une interface configurée pour utiliser NHRP, les étapes suivantes sont alors appliquées :
   (a) Comparer l'adresse de destination IP aux adresses de destination dans l'antémémoire de NHRP. Si l'adresse est dans l'antémémoire, envoyer alors le datagramme à l'adresse de couche de liaison correspondante de l'antémémoire.
   (b) Si l'adresse n'est pas dans l'antémémoire, construire alors un paquet de demande NHRP contenant l'adresse de destination IP. Ce message est envoyé au serveur NHRP configuré pour cette interface. Ce peut être un processus séparé logiquement ou une entité dans le routeur lui-même.
   (c) Le serveur NHRP répondra par l'adresse de couche de liaison appropriée à utiliser pour transmettre le datagramme et les datagrammes suivants à la même destination. Le système PEUT transmettre le ou les datagrammes au routeur traditionnel du "prochain bond" tout en attendant la réponse NHRP.

5.2.4.3 Adresse du prochain bond

COMMENTAIRES DE L'ÉDITEUR

Le routeur applique l'algorithme du paragraphe précédent pour déterminer si l'adresse de destination IP est adjacente. Si elle l'est, l'adresse de prochain bond est la même que l'adresse de destination IP. Autrement, le paquet doit être transmis à travers un autre routeur pour atteindre sa destination immédiate. Le choix de ce routeur est l'objet de ce paragraphe.

Si le paquet contient un SSRR, le routeur DOIT éliminer le paquet et répondre par une erreur ICMP Mauvais chemin de source. Autrement, le routeur recherchera l'adresse de destination IP dans son tableau d'acheminement pour déterminer une adresse de prochain bond appropriée.

Discussion

Selon la spécification IP, une route de source stricte doit spécifier une séquence de nœuds au travers desquels le paquet doit passer ; le paquet doit aller d'un nœud du chemin de source au suivant, en traversant seulement les réseaux intermédiaires. Et donc, si le routeur n'est pas adjacent à la prochaine étape du chemin de source, le chemin de source ne peut être suivi. Donc, le routeur rejette avec une erreur ICMP Mauvais chemin de source.

L'objectif du processus de choix du prochain bond est d'examiner les entrées dans la base d'informations de transmission (FIB, *Forwarding Information Base*) du routeur et de choisir le meilleur chemin (s'il en est un) pour le paquet parmi ceux disponibles dans le FIB.

Conceptuellement, tout algorithme de recherche de chemin commence par un ensemble de chemins candidats qui comporte la totalité du contenu du FIB. L'algorithme comporte une série d'étapes qui élimine des chemins de l'ensemble. Ces étapes sont désignées sous le nom de règles d'élagage. Normalement, lorsque l'algorithme s'achève, il reste exactement un chemin dans l'ensemble. Si l'ensemble devient vide, le paquet est éliminé parce que la destination est inaccessible. Il est aussi possible que l'algorithme se termine alors que plus d'un chemin reste dans l'ensemble. Dans ce cas, le routeur peut arbitrairement éliminer tous les chemins sauf un, ou peut effectuer un "partage de charge" en choisissant n'importe lequel des chemins récemment utilisés.

À l'exception de la règle 3 (TOS faible), un routeur DOIT utiliser les règles d'élagage suivantes lors du choix du prochain bond d'un paquet. Si un routeur ne prend pas en considération le TOS en prenant les décisions de choix du prochain bond, la règle 3 doit être appliquée dans l'ordre indiqué ci-dessous. Ces règles DOIVENT être (conceptuellement) appliquées au FIB dans l'ordre de leur présentation. (Pour une perspective historique, des règles d'élagage supplémentaires, et les autres algorithmes courants utilisés, voir l'Appendice E.)

Discussion

La règle 3 est facultative en ce que le paragraphe 5.3.2 dit qu'un routeur ne DEVRAIT considérer le TOS que lors de la prise de décisions de transmission.

(1) Correspondance de base

Cette règle élimine tous les chemins vers des destinations autres que l'adresse de destination IP du paquet. Par exemple, si l'adresse de destination IP d'un paquet est 10.144.2.5, cette étape éliminerait un chemin pour 128.12.0.0/16 mais garderait tout chemin pour les préfixes de réseau 10.0.0.0/8 et 10.144.0.0/16, et tous les chemins par défaut.

Plus précisément, on suppose que chaque chemin a un attribut de destination, appelé route.dest et une longueur de préfixe correspondante, appelée route.length, pour spécifier quels bits de route.dest sont significatifs. L'adresse de destination IP du paquet à transmettre est ip.dest. Cette règle élimine tous les chemins de l'ensemble de candidats excepté ceux pour lesquels les bits de plus fort poids de route.length, de route.dest et de ip.dest sont égaux.

Par exemple, si l'adresse de destination IP d'un paquet est 10.144.2.5 et qu'il y a des préfixes de réseau de 10.144.1.0/24, 10.144.2.0/24, et 10.144.3.0/24, cette règle ne garderait que 10.144.2.0/24 ; c'est le seul chemin dont le préfixe ait la même valeur que les bits correspondants dans l'adresse de destination IP du paquet.

(2) Plus longue correspondance

Plus longue correspondance est un raffinement de Correspondance de base, décrite ci-dessus. Après avoir effectué l'élagage par Correspondance de base, l'algorithme examine les chemins restants pour déterminer parmi eux ceux qui ont les plus grandes valeurs de route.length. Tous sont éliminés sauf celui-là.

Par exemple, si une adresse de destination IP d'un paquet est 10.144.2.5 et qu'il y a des préfixes de réseau de 10.144.2.0/24, 10.144.0.0/16, et 10.0.0.0/8, cette règle ne garderait que le premier (10.144.2.0/24) parce que son préfixe est le plus long.

(3) TOS faible

Chaque chemin a un attribut de type de service, appelé route.tos, dont les valeurs possibles sont supposées être identiques à celles utilisées dans le champ TOS de l'en-tête IP. Les protocoles d'acheminement qui distribuent les informations de TOS remplissent route.tos de façon appropriée avec les chemins qu'ils ajoutent au FIB ; les chemins provenant d'autres protocoles d'acheminement sont traités comme si ils avaient le TOS par défaut (0000). Le champ TOS dans l'en-tête IP du paquet à acheminer s'appelle ip.tos.

L'ensemble des chemins candidats est examiné pour déterminer si il contient des chemins pour lesquels route.tos = ip.tos. S'il en est, tous les chemins excepté celui pour lequel route.tos = ip.tos sont éliminées. Sinon, tous les chemins excepté celui pour lequel route.tos = 0000 sont éliminés de l'ensemble des chemins candidats.

On trouvera un exposé complémentaire sur l'acheminement fondé sur le TOS faible dans [ROUTE:11].

Discussion

L'effet de cette règle est de ne choisir que les chemins qui ont un TOS qui corresponde au TOS demandé dans le paquet. Si il n'existe pas de tels chemins, on prend alors en compte ceux qui ont le TOS par défaut. Les chemins qui ont un TOS explicite (pas par défaut) qui n'est pas le TOS demandé dans le paquet ne sont jamais utilisés, même si de tels chemins sont les seuls disponibles pour la destination du paquet.

(4) Meilleure métrique

Chaque chemin a un attribut de métrique, appelé route.metric, et un identifiant de domaine d'acheminement, appelé route.domain. Chaque membre de l'ensemble des chemins candidats est comparé à chaque autre membre de l'ensemble. Si route.domain est égal pour les deux chemins et si route.metric est strictement inférieur pour l'un lors de la comparaison avec l'autre, alors celui qui a la mesure inférieure est éliminé de l'ensemble. La détermination de l'inférieur est habituellement effectuée par simple comparaison arithmétique, bien que certains protocoles puissent avoir des métriques structurées qui exigent des comparaisons plus complexes.

(5) Politique du fabricant

Politique du fabricant est une sorte de fourre-tout pour tenir compte du fait que les règles énoncées précédemment sont souvent inadéquates pour le choix des chemins possibles. Les règles d'élagage de Politique du fabricant sont extrêmement spécifiques du fabricant. Voir au paragraphe 5.2.4.4.

Cet algorithme a deux désavantages distincts. Vraisemblablement, une mise en œuvre de routeur va développer des techniques pour compenser ces désavantages et les intégrer à la règle d'élagage Politique du fabricant.

(1) Les classes d'acheminement IS-IS et OSPF ne sont pas traitées directement.

(2) Les propriétés de chemin autres que le type de service (par exemple, MTU) sont ignorées.

Il vaut aussi de noter une déficience dans la façon dont le TOS est pris en charge : les protocoles d'acheminement qui prennent en charge le TOS sont implicitement préférés lors de la transmission de paquets qui ont des valeurs de TOS différentes de zéro.

Les règles d'élagage Correspondance de base et Plus longue correspondance généralisent le traitement d'un certain nombre de types de chemins particuliers. Ces chemins sont choisis dans l'ordre de préférence décroissant suivant :

(1) Chemin d'hôte : C'est un chemin pour un système terminal spécifique.

(2) Chemins à préfixe de réseau hiérarchique : C'est un chemin pour un préfixe de réseau particulier. Noter que le FIB peut contenir plusieurs chemins vers des préfixes de réseau qui s'englobent les uns les autres (un préfixe est l'autre préfixe avec des bits additionnels). Ils sont choisis en ordre de longueur de préfixe décroissante.

(5) Chemin par défaut : C'est un chemin vers tous les réseaux pour lesquels il n'y a pas de chemin explicite. Il est par définition le chemin dont la longueur de préfixe est zéro.

Si, après application des règles d'élagage, l'ensemble des chemins est vide (c'est-à-dire, aucun chemin n'a été trouvé) le paquet DOIT être éliminé et une erreur ICMP appropriée est générée (Mauvaise route de source ICMP si l'adresse de destination IP venait d'une option de route de source ; autrement, le message ICMP approprié Hôte de destination inaccessible ou Destination réseau inaccessible, comme décrit au paragraphe 4.3.3.1).

5.2.4.4 Préférence administrative

Un mécanisme suggéré pour la règle d'élagage Politique du fabricant est d'utiliser la préférence administrative, qui est un simple algorithme de priorité. L'idée est de prioriser manuellement les routes dont on peut avoir besoin pour faire le choix parmi elles.

Chaque chemin est associé à une valeur de préférence, sur la base de divers attributs du chemin (des mécanismes

spécifiques pour l'allocation des valeurs de préférence sont suggérés ci-dessous). Cette valeur de préférence est un entier dans la gamme [0..255], zéro étant la préférée et 254 étant le plus mauvais choix. 255 est une valeur spéciale qui signifie que le chemin ne devrait jamais être utilisé. La première étape de la règle d'élagage Politique du fabricant élimine tous les chemins sauf les préférés (et élimine toujours les chemins dont la valeur de préférence est 255).

Cette politique n'est pas sûre en ce qu'on peut facilement en mésuser et créer des acheminements en boucle. Comme aucun protocole n'assure que les préférences configurées pour un routeur sont cohérentes avec les préférences configurées dans ses voisins, les gestionnaires de réseau doivent faire attention en configurant les préférences.

- Correspondance d'adresse

Il est utile d'être capable d'allouer une seule valeur de préférence à tous les chemins (acquis auprès du même domaine d'acheminement) pour un ensemble de destinations spécifié, où l'ensemble de destinations est toutes les destinations qui correspondent à un préfixe de réseau spécifié.

- Classe de chemin

Pour les protocoles d'acheminement qui conservent la distinction, il est utile d'être capable d'allouer une seule valeur de préférence à tous les chemins (acquis auprès du même domaine d'acheminement) qui ont une classe de chemin particulière (intra-zone, inter-zone, externe avec métrique interne, ou externe avec métrique externe).

- Interface

Il est utile d'être capable d'allouer une seule valeur de préférence à toutes les routes (acquises auprès du même domaine d'acheminement) qui amèneraient les paquets à être acheminés sur une interface logique particulière sur le routeur (les interfaces logiques se transposent généralement d'une à une sur les interfaces réseau du routeur, sauf que toute interface réseau qui a plusieurs adresses IP aura plusieurs interfaces logiques associées).

- Routeur de source

Il est utile d'être capable d'allouer une seule valeur de préférence à tous les chemins (acquis auprès du même domaine d'acheminement) qui ont été acquis auprès de tout ensemble de routeurs, où l'ensemble des routeurs est ceux dont les mises à jour ont une adresse de source qui correspond à un préfixe de réseau spécifié.

- AS d'origine

Pour les protocoles d'acheminement qui fournissent l'information, il est utile d'être capable d'allouer une seule valeur de préférence à tous les chemins (acquis auprès d'un domaine d'acheminement particulier) qui ont leur origine dans un autre domaine d'acheminement particulier. Pour les chemins BGP, l'AS d'origine est le premier AS figurant sur la liste des attributs AS_PATH du chemin. Pour les chemins externes OSPF, l'AS d'origine peut être considéré comme les 16 bits de moindre poids de l'étiquette de chemin externe du chemin si le bit Automatique de l'étiquette est établi est si la longueur de chemin de l'étiquette n'est pas égale à 3.

- Étiquette de chemin externe

Il est utile d'être capable d'allouer une seule valeur de préférence à tous les chemins OSPF externes (acquis auprès du même domaine d'acheminement) dont les étiquettes de chemin externe correspondent à une de celles d'une liste de valeurs spécifiées. Parce que l'étiquette de chemin externe peut contenir une valeur structurée, il peut être utile de donner la capacité de correspondre à des sous champs particuliers de l'étiquette.

- Chemin d'AS

Il est utile d'être capable d'allouer une seule valeur de préférence à tous les chemins BGP (acquis auprès du même domaine d'acheminement) dont le chemin d'AS "correspond" à l'un d'un ensemble de valeurs spécifiées. On ne sait pas encore exactement quelle sorte de correspondance est la plus utile. Une simple option serait de permettre la correspondance de tous les chemins pour lesquels un nombre d'AS particulier apparaît (ou n'apparaît pas) n'importe où dans l'attribut AS_PATH du chemin. Une solution de remplacement plus générale mais un peu plus difficile serait de permettre la correspondance de tous les chemins pour lesquels le chemin d'AS correspond à une expression régulière spécifiée.

5.2.4.5 Partage de charge

À la fin du processus de choix du prochain bond, il peut rester encore plusieurs chemins. Un routeur a plusieurs options lorsque cela arrive. Il peut arbitrairement éliminer certains chemins. Il peut réduire le nombre de chemins candidats en comparant la métrique des chemins à partir de domaines d'acheminement qui ne sont pas considérés comme équivalents. Il peut retenir plus d'un chemin et employer un mécanisme de partage de charge pour diviser le trafic entre eux. Peut-être la seule chose à dire sur les mérites relatifs des options est que le partage de charge est utile dans des situations mais pas dans d'autres, aussi un gestionnaire avisé qui met en œuvre le partage de charge prévoira un moyen pour que le gestionnaire de réseau puisse le désactiver.

5.2.5 Bits d'en-tête IP inutilisés : RFC-791 paragraphe 3.1

L'en-tête IP contient plusieurs bits réservés, dans le champ Type de service et dans le champ Etiquettes. Les routeurs NE DOIVENT PAS abandonner des paquets simplement parce que un ou plusieurs de ces bits réservés a une valeur différente de zéro.

Les routeurs DOIVENT ignorer et DOIVENT passer inchangées les valeurs de ces bits réservés. Si un routeur fragmente un paquet, il DOIT copier ces bits dans chaque fragment.

Discussion

Des révisions futures du protocole IP pourraient faire usage de ces bits inutilisés. Ces règles sont destinées à s'assurer que ces révisions pourront être déployées sans qu'il soit nécessaire de mettre à jour tous les routeurs de l'Internet.

5.2.6 Fragmentation et réassemblage : RFC-791 paragraphe 3.2

Comme exposé au paragraphe 4.2.2.7, un routeur DOIT prendre en charge la fragmentation IP.

Un routeur NE DOIT PAS réassembler un datagramme avant de le transmettre.

Discussion

Certains ont suggéré qu'il pourrait exister des topologies dans lesquelles le réassemblage des datagrammes en transit par les routeurs pourrait améliorer les performances. Le fait que les fragments puissent prendre des chemins différents pour une même destination empêche une utilisation sûre d'un tel dispositif.

Dans ce paragraphe, rien n'est conçu pour contrôler ou limiter la fragmentation ou le réassemblage effectué au titre de la fonction de couche de liaison par le routeur.

De même, si un datagramme IP est encapsulé dans un autre datagramme IP (par exemple, il est tunnelé), ce datagramme est fragmenté à son tour, et les fragments doivent être réassemblés dans l'ordre pour la transmission du datagramme d'origine. Ce paragraphe ne l'interdit pas.

5.2.7 Protocole de message de commande Internet - ICMP

Les exigences générales pour ICMP sont exposées au paragraphe 4.3. Ce paragraphe discute des messages ICMP qui sont seulement envoyés par les routeurs.

5.2.7.1 Destination inaccessible

Le message ICMP Destination inaccessible est envoyé par un routeur en réponse à un paquet qu'il ne peut pas transmettre parce que la destination (ou le prochain bond) est inaccessible ou qu'un service est indisponible. Des exemples de tels cas sont le message adressé à un hôte qui n'est pas là et donc ne répond pas aux demandes ARP, et les messages adressés à des préfixes de réseau pour lesquels le routeur n'a pas de route valide.

Un routeur DOIT être capable de générer des messages ICMP Destination inaccessible et DEVRAIT choisir un code de réponse qui corresponde très étroitement à la raison pour laquelle le message a été généré.

Les codes suivants sont définis dans [INTER:8] et [INTRO:2] :

0 = Réseau inaccessible – généré par un routeur si un chemin de transmission vers le réseau de destination n'est pas disponible ;
1 = Hôte inaccessible – généré par un routeur si un chemin de transmission vers l'hôte de destination sur un réseau directement connecté n'est pas disponible (ne répond pas à l'ARP) ;
2 = Protocole inaccessible – généré si le protocole de transport désigné dans un datagramme n'est pas accepté dans la couche transport de la destination finale ;
3 = Port inaccessible – généré si le protocole de transport désigné (par exemple, UDP) est incapable de démultiplexer le datagramme dans la couche transport de la destination finale mais n'a pas de mécanisme de protocole pour en informer l'envoyeur ;
4 = Fragmentation nécessaire et DF établi – généré si un routeur a besoin de fragmenter un datagramme mais ne peut pas le faire car le fanion DF est mis ;
5 = Échec de route de source – généré si un routeur ne peut pas transmettre un paquet au prochain bond dans une option de route de source ;
6 = Réseau de destination inconnu – Ce code NE DEVRAIT PAS être généré car il implique de la part du routeur que le réseau de destination n'existe pas (le code 0, Réseau inaccessible DEVRAIT être utilisé plutôt que le code 6) ;

7 = Hôte de destination inconnu – généré seulement lorsque un routeur peut déterminer (sur avis de la couche de liaison) que l'hôte de destination n'existe pas ;
11 = Réseau inaccessible pour ce type de service – généré par un routeur si un chemin de transmission pour le réseau de destination avec le TOS demandé ou par défaut n'est pas disponible ;
12 = Hôte inaccessible pour ce type de service – généré si un routeur ne peut pas transmettre un paquet parce que son ou ses chemins pour la destination ne correspondent ni au TOS demandé dans le datagramme ni au TOS par défaut (0).

Les codes supplémentaires suivants sont maintenant définis :

13 = Interdiction administrative de communication – généré si un routeur ne peut pas transmettre un paquet à cause d'un filtrage administratif ;
14 = Violation de préséance d'hôte – envoyé par le routeur du premier bond à un hôte pour indiquer qu'une préséance demandée n'est pas permise pour une combinaison particulière de source/hôte de destination ou de réseau, de protocole de couche supérieure, et d'accès de source/destination ;
15 = Seuil de préséance en effet – les opérateurs de réseau ont imposé un niveau minimum de préséance exigé pour l'opération ; le datagramme a été envoyé avec une préséance inférieure à ce niveau ;

NOTE : [INTRO:2] a défini le code 8 pour un hôte de source isolé. Les routeurs NE DEVRAIENT PAS générer le code 8 ; ce sont les codes 0 (Réseau inaccessible) et 1 (Hôte inaccessible) qui DEVRAIENT être utilisés à la place, selon celui qui est appropriés. [INTRO:2] a aussi défini le code 9 pour les communications avec un réseau de destination interdit administrativement et le code 10 pour les communications avec un hôte de destination interdit administrativement. Ces codes étaient destinés à être utilisés par des appareils de codage de bout en bout des organes de l'armée des USA. Les routeurs DEVRAIENT utiliser le code 13 nouvellement défini (Interdiction administrative de communication) s'il y a un filtrage administratif des paquets.

Les routeurs PEUVENT avoir une option de configuration qui empêche la génération des messages de code 13 (Interdiction administrative de communication). Lorsque cette option est activée, aucun message d'erreur ICMP n'est envoyé en réponse à un paquet abandonné parce que sa transmission est administrativement interdite.

De même, les routeurs PEUVENT avoir une option de configuration option qui empêche la génération des messages de code 14 (Violation de préséance d'hôte) et de code 15 (Seuil de préséance en effet). Lorsque cette option est activée, aucun message d'erreur ICMP n'est envoyé en réponse à un paquet abandonné à cause d'une violation de préséance.

Les routeurs DOIVENT utiliser les codes Hôte inaccessible ou Hôte de destination inconnu chaque fois que d'autres hôtes pourraient être accessibles sur le même réseau de destination ; autrement, l'hôte de source peut conclure par erreur que tous les hôtes sur ce réseau sont inaccessibles, alors que ce n'est pas forcément le cas.

[INTER:14] décrit une légère modification de la forme des messages Destination inaccessible qui contiennent le code 4 (Fragmentation nécessaire et DF établi). Un routeur DOIT utiliser cette forme modifiée lorsqu'il est à l'origine de messages de code 4 Destination inaccessible.

5.2.7.2 Redirection

Le message ICMP Redirection est généré pour informer un hôte local qu'il devrait utiliser un routeur de prochain bond différent pour une certaine classe de trafic.

Les routeurs NE DOIVENT PAS générer les messages Redirection pour le réseau et Redirection pour le réseau et le type de service (codes 0 et 2) spécifiés dans [INTER:8]. Les routeurs DOIVENT être capables de générer le message Redirection pour l'hôte (code 1) et DEVRAIENT être capables de générer le message Redirection pour le type de service et l'hôte (code 3) spécifié dans [INTER:8].

Discussion

Si le réseau directement connecté n'est pas subdivisé en sous-réseau (dans le sens classique), un routeur peut normalement générer un Redirection de réseau qui s'applique à tous les hôtes sur un réseau distant spécifié. Utiliser un message Redirection réseau plutôt que Redirection d'hôte peut économiser un peu sur le trafic du réseau et sur la taille mémoire du tableau d'acheminement de l'hôte. Cependant, ces économies ne sont pas significatives, et les sous-réseaux créent une ambiguïté sur le gabarit de sous-réseau à utiliser pour interpréter un message Redirection réseau. Dans un environnement CIDR, il est difficile de spécifier précisément les cas dans lesquels les messages Redirection réseau peuvent être utilisés. Donc, les routeurs ne doivent envoyer que les messages Redirection d'hôte (ou hôte et type de service).

Un message de code 3 (Redirection pour l'hôte et le type de Service) est généré lorsque le paquet qui provoque la redirection a une destination pour laquelle le chemin choisi par le routeur dépendrait (en partie) du TOS demandé.

Les routeurs qui peuvent générer des redirections de code 3 (Hôte et type de service) DOIVENT avoir une option de configuration (activée par défaut) pour activer la substitution de la redirection de code 1 (Hôte) à la redirection de code 3. Un routeur DOIT envoyer une redirection de code 1 à la place d'une redirection de code 3 si il a été configuré pour le faire.

Si un routeur n'est pas capable de générer des redirections de code 3, il DOIT alors générer des redirections de code 1 dans les situations qui appellent une redirection de code 3.

Les routeurs NE DOIVENT PAS générer de message Redirection si toutes les conditions suivantes ne sont pas réunies :
- Le paquet est transmis sur la même interface physique que celle sur laquelle il a été reçu,
- L'adresse IP de source dans le paquet est sur le même (sous) réseau IP logique que l'adresse IP du prochain bond, et
- Le paquet ne contient aucune option IP de route de source.

L'adresse de source utilisée dans le message Redirection ICMP DOIT appartenir au même (sous) réseau logique que l'adresse de destination.

Un routeur qui utilise un protocole d'acheminement (autre que les chemins statiques) NE DOIT PAS considérer les chemins acquis à partir des redirections ICMP lors de la transmission d'un paquet. Si un routeur n'utilise pas de protocole d'acheminement, il PEUT avoir une configuration qui, si elle est établie, permette au routeur de prendre en considération des chemins acquis à travers des redirections ICMP lors de la transmission de paquets.

Discussion

La redirection ICMP est un mécanisme pour que les routeurs convoient les informations d'acheminement aux hôtes. Les routeurs utilisent d'autres mécanismes pour acquérir ces informations d'acheminement, et n'ont donc pas de raisons d'obéir aux redirections. Croire une redirection en contradiction avec les autres informations du routeur créerait vraisemblablement des boucles dans l'acheminement.

D'un autre côté, lorsqu'un routeur n'agit pas comme un routeur, il DOIT se conformer au comportement exigé d'un hôte.

5.2.7.3 Délai dépassé

Un routeur DOIT générer un message Délai dépassé de code 0 (En transit) lorsque il élimine un paquet du fait d'un champ de durée de vie expirée. Un routeur PEUT avoir une option interface par interface pour désactiver la génération de ces messages sur cette interface, mais par défaut, cette option DOIT permettre la génération des messages.

5.2.8 Protocole des messages de gestion de l'Internet - IGMP

IGMP [INTER:4] est un protocole utilisé entre les hôtes et les routeurs de diffusion groupée sur un seul réseau physique pour établir l'adhésion des hôtes à des groupes de diffusion groupée particuliers. Les routeurs de diffusion groupée utilisent ces informations, en conjonction avec un protocole d'acheminement de diffusion groupée, pour prendre en charge la diffusion groupée IP à travers l'Internet.

Un routeur DEVRAIT mettre en œuvre la partie routeur de diffusion groupée du protocole IGMP.

5.3 Questions spécifiques

5.3.1 Durée de vie restante (TTL)

Le champ Durée de vie restante (TTL, *Time-to-Live*) de l'en-tête IP est défini comme un temporisateur qui limite la durée de vie d'un datagramme. C'est un champ de 8 bits et les unités sont des secondes. Chaque routeur (ou autre module) qui traite un paquet DOIT décrémenter la TTL d'au moins un, même si le temps écoulé était de bien moins d'une seconde. Comme c'est très souvent le cas, la TTL est effectivement une limite du compte des bonds sur la distance sur laquelle un datagramme peut se propager à travers l'Internet.

Lorsque un routeur transmet un paquet, il DOIT réduire la TTL d'au moins un. Si il détient un paquet pour plus d'une seconde, il PEUT décrémenter la TTL d'un pour chaque seconde.

Si la TTL est réduite à zéro (ou moins), le paquet DOIT être éliminé, et si la destination n'est pas une adresse de diffusion groupée, le routeur DOIT envoyer un message ICMP Délai dépassé, code 0 (TTL dépassée en transit) à la source. Noter qu'un routeur NE DOIT PAS éliminer un paquet IP en envoi individuel ou en diffusion avec une TTL

différente de zéro simplement parce qu'il peut prédire qu'un autre routeur sur le chemin de la destination finale du paquet va décrémenter la TTL jusqu'à zéro. Cependant, un routeur PEUT le faire pour les diffusions groupées IP, afin de mettre en œuvre de façon plus efficace l'algorithme de recherche en anneau expansif de la diffusion groupée IP (voir [INTER:4]).

Discussion

La TTL IP est utilisée, de façon un peu schizophrénique, à la fois comme une limite de nombre et comme une limite de temps. Sa fonction de compte numérique est critique pour s'assurer que des problèmes d'acheminement ne risquent pas de couler le réseau en causant la ùise en boucle infinie de paquets dans le réseau. La fonction de limite de temps est utilisée par les protocoles de transport tels que TCP pour assurer un transfert fiable des données. De nombreuses mises en œuvre courantes traitent la TTL comme un pur compte de bonds, et dans des pans entiers de la communauté de l'Internet il y a un fort sentiment que la fonction de limite de temps devrait à la place être effectuée par les protocoles de transport qui en ont besoin.

Dans la présente spécification, nous avons décidé à contre cœur de suivre la forte tendance des fabricants de routeurs à penser que la fonction de limite de temps devrait être facultative. Ils avancent que la mise en œuvre de la fonction de limite de temps est assez difficile pour n'être pas effectuée habituellement. Ils avancent de plus le manque de cas documentés où ce raccourci aurait causé la lésion de données par TCP (bien sûr, on supposera que les problèmes créés seraient rares et difficiles à reproduire, de sorte que l'absence de cas documentés ne donne que peu d'assurance qu'il n'y aurait pas de nombreux cas non documentés).

Des notions de diffusion groupée IP comme la recherche par anneau d'expansion peuvent ne pas fonctionner comme prévu sauf si la TTL est traitée comme un pur compte de bonds. C'est à peu près la même chose pour traceroute.

Les messages ICMP Délai dépassé sont nécessaires parce que l'outil de diagnostic traceroute en dépend.

Et donc, le compromis est entre estropier sévèrement, sinon éliminer, deux outils très utiles et éviter un problème très rare et transitoire de transport de données qui peut ne pas survenir du tout. Nous avons choisi de préserver les outils.

5.3.2 Type de service (TOS)

L'octet Type-de-Service dans l'en-tête IP est divisé en trois sections : le champ Préséance (3 bits de plus fort poids), un champ qui est traditionnellement appelé Type de service ou "TOS" (4 bits suivants), et un bit réservé (le bit de moindre poids). Les règles qui gouvernent le bit réservé ont été décrites au paragraphe 4.2.2.3. Le champ Préséance sera exposé au paragraphe 5.3.3. Un exposé plus complet du champ TOS et de son utilisation figure dans [ROUTE:11].

Un routeur DEVRAIT considérer le champ TOS dans l'en-tête IP d'un paquet pour décider comment le transmettre. Le reste de ce paragraphe décrit les règles applicables aux routeurs qui se conforment à cette exigence.

Un routeur DOIT maintenir une valeur de TOS pour chaque chemin de son tableau d'acheminement. Les chemins acquis à travers un protocole d'acheminement qui n'accepte pas le TOS DOIVENT recevoir un TOS de zéro (le TOS par défaut).

Pour choisir un chemin vers une destination, un routeur DOIT utiliser un algorithme équivalent au suivant :

(1) Le routeur localise dans son tableau d'acheminement toutes les routes disponibles vers la destination (paragraphe 5.2.4).
(2) Si il n'y en a aucune, le routeur abandonne le paquet parce que la destination est inaccessible (paragraphe 5.2.4).
(3) Si un ou plusieurs de ces chemins ont un TOS qui correspond exactement au TOS spécifié dans le paquet, le routeur choisit le chemin qui a la meilleure métrique.
(4) Autrement, le routeur répète l'étape ci-dessus, mais en cherchant les chemins dont le TOS est zéro.
(5) Si aucun chemin n'a été choisi, le routeur abandonne le paquet parce que la destination est inaccessible. Le routeur retourne une erreur ICMP Destination inaccessible qui spécifie le code approprié : Réseau inaccessible avec le type de service (code 11) ou Hôte inaccessible pour ce type de service (code 12).

Discussion

Bien que le TOS ait été peu utilisé dans le passé, son utilisation par les hôtes est maintenant rendue obligatoire par les RFC sur les exigences pour les hôtes Internet ([INTRO:2] et [INTRO:3]). La prise en charge du TOS dans les routeurs peut devenir un DOIT à l'avenir, mais est un DEVRAIT actuellement, jusqu'à ce qu'on ait un peu plus d'expérience et qu'on puisse mieux juger de ses avantages et de ses inconvénients.

Plusieurs personnes ont proposé que le TOS puisse affecter d'autres aspects de la fonction de transmission.
Par exemple :

(1) Un routeur pourrait placer les paquets qui ont le bit Faible délai mis avant les autres paquets dans sa file d'attente de sortie.
(2) Si un routeur est forcé d'éliminer des paquets, il pourrait essayer d'éviter ceux qui ont le bit Haute fiabilité mis.

Ces idées ont été explorées plus en détail dans [INTER:17] mais nous n'avons pas encore assez d'expérience de tels schémas pour édicter des exigences dans ce domaine.

5.3.3 Préséance IP

Ce paragraphe spécifie les exigences et les lignes directrices pour un traitement approprié du champ Préséance IP dans les routeurs. Préséance est un schéma d'allocation des ressources dans le réseau fondé sur l'importance relative des différents flux de trafic. La spécification IP définit des valeurs spécifiques à utiliser dans ce champ pour divers types de trafic.

Les mécanismes de base pour le traitement de la préséance dans un routeur sont l'allocation préférentielle de ressource, y compris à la fois le service de mise en file d'attente ordonnée selon la préséance et le contrôle d'encombrement fondé sur la préséance, et la sélection des caractéristiques de priorité de couche Liaison. Le routeur choisit aussi la préséance IP pour l'acheminement, la gestion et le contrôle du trafic qu'il génère. Pour un exposé détaillé de la préséance IP et sa mise en œuvre, voir [FORW:6].

Le service de mise en file d'attente ordonnée selon la préséance, comme exposé dans ce paragraphe, inclut, sans s'y limiter, la file d'attente pour le processus de transmission et les files d'attente pour les liaisons sortantes. Il est destiné à ce qu'un routeur qui prend en charge la préséance utilise aussi l'indication de préséance à tout point du traitement qui serait concerné par l'allocation de ressources finies, telles que les mémoires tampon de paquet ou des connexions de couche Liaison. L'ensemble de tels points est fonction de la mise en œuvre.

Discussion

Bien que le champ Préséance ait été à l'origine fourni pour être utilisé dans les systèmes du DOD où de grosses pointes de trafic ou des dommages majeur au réseau sont vues comme des menaces inhérentes, il a des applications utiles pour de nombreux réseaux IP non militaires. Bien que la capacité d'absorption du trafic des réseaux ait considérablement crû ces dernières années, la capacité à générer du trafic des utilisateurs s'est aussi accrue, et les conditions de surcharge des réseaux surviennent parfois. Comme les protocoles de gestion de d'acheminement fondés sur IP sont devenus plus essentiels pour le bon fonctionnement de l'Internet, les surcharges présentent deux risques supplémentaires pour le réseau :

(1) De forts retards peuvent résulter en perte de paquets de protocole d'acheminement. Cela peut amener le protocole d'acheminement à conclure faussement à un changement de topologie et à propager cette fausse information aux autres routeurs. Cela peut non seulement causer l'oscillation des routes, mais aussi à faire peser une charge de traitement supplémentaire sur les autres routeurs.

(2) De forts retards peuvent interférer avec l'utilisation des outils de gestion du réseau pour analyser et peut-être corriger ou soulager le problème qui a causé dans le réseau la survenance de la condition de surcharge.

La mise en œuvre et l'utilisation appropriée du mécanisme Préséance pallie ces deux problèmes.

5.3.3.1 Service de mise en file d'attente ordonnée selon la préséance

Les routeurs DEVRAIENT mettre en œuvre le service de mise en file d'attente ordonnée selon la préséance. "Service de mise en file d'attente ordonnée selon la préséance" signifie que lorsque un paquet est retenu pour la sortie sur une liaison (logique), le paquet de plus forte préséance qui est en file d'attente pour cette liaison est envoyé. Les routeurs qui mettent en œuvre le service de mise en file d'attente ordonnée selon la préséance DOIVENT aussi avoir une option de configuration pour supprimer le service de mise en file d'attente ordonnée selon la préséance dans la couche Internet.

Tout routeur PEUT mettre en œuvre d'autres procédures de gestion du débit fondées sur une politique qui résultera en un ordre de préséance autre que strict, mais il DOIT être possible de les supprimer par configuration (c'est-à-dire, utiliser l'ordre strict).

Comme précisé au paragraphe 5.3.6, les routeurs qui mettent en œuvre le service de mise en file d'attente ordonnée selon la préséance éliminent les paquets de faible préséance avant d'éliminer les paquets de forte préséance pour les besoins du contrôle d'encombrement.

La préemption (interruption du traitement ou de la transmission d'un paquet) n'est pas envisagée comme fonction de la couche Internet. Certains protocoles à d'autres couches peuvent fournir des dispositifs de préemption.

5.3.3.2 Transpositions de préséance de couche inférieure

Les routeurs qui mettent en œuvre la mise en file d'attente ordonnée selon la préséance DOIVENT METTRE EN ŒUVRE, et les autres routeurs DEVRAIENT METTRE EN ŒUVRE, la transposition de préséance de couche inférieure.

Un routeur qui met en œuvre la transposition de préséance de couche inférieure :

- DOIT être capable de transposer la préséance IP en mécanismes de priorité de couche de liaison pour les couches de liaison qui ont un tel dispositif défini.
- DOIT avoir une option de configuration pour choisir le traitement de priorité par défaut de couche de liaison pour tout le trafic IP.
- DEVRAIT être capable de configurer des transpositions non standard spécifiques des valeurs de préséance IP en valeurs de priorité de couche de liaison pour chaque interface.

Discussion

Certaines recherches mettent en question la faisabilité des dispositifs de priorité de certains protocoles de couche de liaison, et certains réseaux peuvent avoir des mises en œuvre déficientes du mécanisme de priorité de couche de liaison. Il semble prudent de fournie un mécanisme d'échappement au cas où de tels problèmes apparaîtraient dans un réseau.

D'un autre côté, il y a des propositions pour utiliser de nouvelles stratégies de mise en file d'attente pour mettre en œuvre des services spéciaux tels qu'une réservation de bande passante multimédia ou un service à faible délai. Les services spéciaux et les stratégies de mise en file d'attente pour les prendre en charge sont des sujets de recherche en cours et sont en cours de normalisation.

Les développeurs peuvent souhaiter considérer qu'une transposition de couche de liaison correcte de la préséance IP est nécessaire à la politique du DOD pour les systèmes TCP/IP utilisés sur les réseaux du DOD. Comme ces exigences sont destinées à encourager (mais non à forcer) l'utilisation des dispositifs de préséance dans l'espoir de fournir un meilleur service Internet à tous les utilisateurs, les routeurs qui prennent en charge le service de file d'attente ordonnée selon la préséance devrait par défaut maintenir un ordre de préséance strict sans considération du type de service demandé.

5.3.3.3 Traitement de la préséance pour tous les routeurs

Un routeur (qu'il utilise ou non le service de file d'attente ordonnée selon la préséance) :

(1) DOIT accepter et traiter normalement le trafic entrant de tous les niveaux de préséance, sauf s'il a été configuré administrativement pour faire autrement.

(2) PEUT mettre en œuvre un filtre de validation pour restreindre administrativement l'utilisation des niveaux de préséance par des sources de trafic particulières. S'il est fourni, ce filtre NE DOIT PAS filtrer ou couper les types de messages d'erreur ICMP suivants : Destination inaccessible, Redirection, Délai dépassé, et Problème de paramètre. Si ce filtre est fourni, les procédures requises pour le filtrage de paquet par les adresses sont aussi requises pour ce filtre.

Discussion

Le filtrage de préséance devrait être applicable à des paires spécifiques d'adresses IP de source/destination, des protocoles spécifiques, des accès spécifiques, et ainsi de suite.

Un message ICMP Destination inaccessible avec le code 14 DEVRAIT être envoyé lorsque un paquet est abandonné par le filtre de validation, sauf si cela a été supprimé par un choix de configuration.

(3) PEUT mettre en œuvre une fonction de coupure qui permet au routeur d'être réglé pour refuser ou abandonner du trafic avec une préséance inférieure à un niveau spécifié. Cette fonction peut être activée par des actions de gestion ou par une heuristique dépendante de la mise en œuvre, mais il DOIT y avoir une option de configuration pour désactiver tout mécanisme heuristique qui fonctionne sans intervention humaine. Un message ICMP Destination inaccessible avec le code 15 DEVRAIT être envoyé lorsque un paquet est abandonné par la fonction de coupure, sauf si cela a été supprimé par un choix de configuration.

Un routeur NE DOIT PAS refuser de transmettre des datagrammes avec une préséance IP de 6 (Contrôle inter réseau) ou 7 (Contrôle réseau) seulement à cause de la coupure de préséance. Cependant, d'autres critères peuvent être utilisés en conjonction avec la coupure de préséance pour filtrer le trafic à haute préséance.

Discussion

La coupure de préséance sans restriction a pour résultat une coupure non intentionnelle du trafic d'acheminement et de contrôle. Dans le cas général, le trafic d'hôte devrait être restreint à une valeur de 5 (CRITIC/ECP) ou inférieure, mais ceci n'est pas exigé et peut n'être pas correct dans certains systèmes.

(4) NE DOIT PAS changer le réglage de la préséance sur les paquets dont il n'est pas à l'origine.

(5) DEVRAIT être capable de configurer des valeurs de préséance distinctes à utiliser pour chaque acheminement ou protocole de gestion pris en charge (excepté pour ces protocoles, tels que OSPF, qui spécifie quelle valeur de préséance doit être utilisée).

(6) PEUT être capable de configurer les valeurs de préséance de trafic d'acheminement ou de gestion indépendamment pour chaque adresse d'homologue.

(7) DOIT répondre de façon appropriée aux indications d'erreur qui se rapportent à la préséance de couche Liaison lorsque elles sont fournies. Un message ICMP Destination inaccessible avec le code 15 DEVRAIT être envoyé lorsque un paquet est abandonné parce qu'une liaison ne peut pas l'accepter à cause d'une condition en rapport avec la préséance, à moins que cela n'ait été supprimé par un choix de configuration.

Discussion

Le mécanisme de coupure de la préséance décrit en (3) est assez controversé. Selon la localisation topologique de la zone affectée par la coupure, le trafic de transit peut être dirigé par les protocoles d'acheminement dans la zone de coupure, où il sera abandonné. Ce n'est un problème que si un autre chemin qui n'est pas affecté par la coupure existe entre les points en communication. Les façons qui ont été proposées pour éviter ce problème sont de fournir une bande passante minimale à tous les niveaux de préséance même dans des conditions de surcharge, ou de propager les informations de coupure dans les protocoles d'acheminement. En l'absence d'une solution largement acceptée (et mise en œuvre) à ce problème, une grande prudence est recommandée pour l'activation de mécanismes de coupure dans les réseaux de transit.

Un relais de couche transport pourrait légitimement fournir la fonction interdite en (4) ci-dessus. Changer les niveaux de préséance peut causer des interactions subtiles avec TCP et peut-être d'autres protocoles ; une conception correcte n'est pas une tâche triviale.

L'intention de (5) et (6) (et de l'exposé sur la préséance IP dans les messages ICMP au paragraphe 4.3.2) est que les bits de préséance IP devraient être réglés de façon appropriée, que ce routeur agisse ou non sur ces bits d'une autre façon. On s'attend à ce que de futures spécifications des protocoles d'acheminement et de protocoles de gestion de réseau spécifient comment la préséance IP devrait être réglée pour les messages envoyés par ces protocoles.

La réponse appropriée pour (7) dépend du protocole de couche de liaison utilisé. Normalement, le routeur devrait arrêter d'essayer d'envoyer du trafic présentant un risque de surcharge pendant un certain temps pour cette destination, et devrait retourner un message ICMP Destination inaccessible avec le code 15 (service non disponible pour la préséance demandée) à la source du trafic. Il ne devrait pas non plus essayer de rétablir une connexion de couche Liaison préemptée pendant un certain temps.

5.3.4 Transmission de diffusions de couche Liaison

L'encapsulation de paquets IP dans la plupart des protocoles de couche Liaison (excepté PPP) permet à un receveur de distinguer les diffusions et diffusions groupées des envois individuels simplement en examinant les en-têtes de protocole de couche Liaison (et plus couramment, l'adresse de destination de couche Liaison). Les règles de ce paragraphe qui se réfèrent aux diffusions de couche Liaison ne s'appliquent qu'aux protocoles de couche Liaison qui permettent de distinguer les diffusions ; de même, les règles qui se réfèrent à la diffusion groupée de couche Liaison ne s'appliquent qu'aux protocoles de couche Liaison qui permettent de distinguer les diffusions groupées.

Un routeur NE DOIT PAS transmettre de paquet qu'il a reçu comme diffusion de couche Liaison, sauf si il est dirigé sur une adresse IP en diffusion groupée. Dans ce dernier cas, on supposera que la diffusion de couche de liaison a été utilisée à cause de l'absence d'un service effectif de diffusion groupée.

Un routeur NE DOIT PAS transmettre de paquets qu'il a reçus comme diffusion groupée de couche Liaison, sauf si l'adresse de destination du paquet est une adresse IP de diffusion groupée.

Un routeur DEVRAIT éliminer en silence un paquet reçu via une diffusion de couche Liaison mais ne spécifie pas d'adresse de destination IP de diffusion ou de diffusion groupée.

Lorsque un routeur envoie un paquet comme diffusion de couche de liaison, l'adresse de destination IP DOIT être une adresse IP légale de diffusion ou de diffusion groupée.

5.3.5 Transmission de diffusions de couche Internet

Il y a deux types majeurs d'adresses de diffusion IP ; les diffusions limitées et les diffusions dirigées. De plus, il y a trois sous-types de diffusion dirigée : une diffusion dirigée vers un préfixe de réseau spécifié, une diffusion dirigée vers un sous-réseau spécifié, et une diffusion dirigée vers tous les sous-réseaux d'un réseau spécifié. La classification par un

routeur d'une diffusion dans une de ces catégories dépend de l'adresse de diffusion et de la compréhension du routeur (s'il en a une) de la structure de sous-réseaux du réseau de destination. La même diffusion sera classée différemment par les différents routeurs.

Une adresse de diffusion IP limitée est définie comme étant toute de uns : { -1, -1 } ou 255.255.255.255.

Une diffusion dirigée sur un préfixe de réseau se compose du préfixe de réseau de l'adresse IP avec une partie locale toute de uns ou de { <Préfixe-réseau>, -1 }. Par exemple, une adresse de diffusion de réseau de classe A est net.255.255.255, une adresse de diffusion de réseau de classe B est net.net.255.255 et une adresse de diffusion de réseau de classe C est net.net.net.255 où net est un octet de l'adresse réseau.

La diffusion dirigée sur tous les sous-réseaux n'est pas très bien définie dans un environnement CIDR, et a été déconseillée dans la version 1 du présent mémoire.

Comme décrit au paragraphe 4.2.3.1, un routeur peut rencontrer certaines adresses de diffusion IP non standard :
- 0.0.0.0 est une forme obsolète de l'adresse de diffusion limitée
- { <Préfixe-réseau>, 0 } est une forme obsolète d'adresse de diffusion dirigée par préfixe de réseau.

Comme décrit dans ce paragraphe, les paquets adressés à l'une de ces adresses DEVRAIENT être éliminés en silence, mais s'il apparaît qu'ils ne le sont pas, ils DOIVENT être traités selon les mêmes règles qui s'appliquent aux paquets adressés aux formes non obsolètes des adresses de diffusion décrites ci-dessus. Ces règles sont décrites dans les paragraphes suivants.

5.3.5.1 Diffusions limitées

Les diffusions limitées NE DOIVENT PAS être retransmises. Les diffusions limitées NE DOIVENT PAS être éliminées. Les diffusions limitées PEUVENT être envoyées et DEVRAIENT être envoyées à la place de diffusions dirigées lorsque des diffusions limitées suffiront.

Discussion

Certains routeurs contiennent des serveurs UDP qui fonctionnent en renvoyant les demandes (comme envois individuels ou comme diffusions dirigées) à d'autres serveurs. Cette exigence ne devrait pas être interprétée comme interdisant de tels serveurs. Noter, cependant, que de tels serveurs peuvent aisément causer la mise en boucle du paquet si ils sont mal configurés. Et donc, les fournisseurs de tels serveurs seraient probablement bien avisés de documenter soigneusement leurs réglages et de considérer avec attention la TTL sur les paquets envoyés.

5.3.5.2 Diffusions dirigées

Un routeur DOIT classer en diffusions dirigées par préfixe de réseau toutes les diffusions dirigées valides destinées à un réseau distant ou un réseau rattaché sans sous-réseau. Noter que du point de vue de CIDR, c'est comme cela qu'apparaissent les adresses d'hôte au sein du préfixe de réseau ; l'inspection de la partie hôte de tels préfixes de réseau est empêchée. Pour un chemin donné et sans politique surajoutée, un routeur DOIT alors transmettre les diffusions dirigées par préfixe de réseau. Les diffusions dirigées par préfixe de réseau PEUVENT être envoyées.

Un routeur PEUT avoir une option pour désactiver la réception des diffusions dirigées par préfixe de réseau sur une interface et DOIT avoir une option pour désactiver la transmission des diffusions dirigées par préfixe de réseau. Ces options DOIVENT par défaut permettre de recevoir et transmettre des diffusions dirigées par préfixe de réseau.

Discussion

Il y a eu quelques débats sur la question de la transmission ou non des diffusions dirigées. Dans le présent mémoire nous avons fait dépendre la décision de transmission de la connaissance du préfixe de réseau de destination par le routeur. Les routeurs ne peuvent pas déterminer qu'un message est en envoi individuel ou en diffusion dirigée sans cette connaissance. La décision de transmettre ou non le message n'est par définition possible qu'au routeur du dernier bond.

5.3.5.3 Diffusions dirigées sur tous les sous-réseaux

La première version du présent mémoire décrivait un algorithme pour la distribution d'une diffusion dirigée sur tous les sous-réseaux d'un numéro de réseau classique. Cet algorithme a été déclaré "cassé," et certains cas de défaillances étaient spécifiés.

Dans un domaine d'acheminement CIDR, dans lequel les numéros de réseau IP classiques sont sans signification, le concept de diffusion dirigée sur tous les sous-réseaux n'a pas non plus de signification. À la connaissance du groupe de travail, la facilité n'a jamais été mise en œuvre ou développée, et est maintenant reléguée dans les poubelles de l'histoire.

5.3.5.4 Diffusions dirigées sur un sous-réseau

La première version du présent mémoire énonçait des procédures pour traiter les diffusions dirigées sur les sous-réseaux. Dans un domaine d'acheminement CIDR, ce sont des diffusions non distinguables des diffusions dirigées sur

un réseau. Les deux sont donc traitées ensemble au paragraphe 5.3.5.2 "Diffusions dirigées", et devraient être vues comme des diffusions dirigées sur préfixe de réseau.

5.3.6 Contrôle d'encombrement

L'encombrement dans un réseau est en gros défini comme une condition dans laquelle la demande de ressources (normalement de la bande passante ou du temps CPU) excède les capacités. Éviter l'encombrement revient à empêcher la demande d'excéder les capacités, alors que la récupération sur encombrement essaye de restaurer un état fonctionnel. Il est possible à un routeur de contribuer à ces deux mécanismes. De nombreux efforts ont été fournis pour étudier le problème. Le lecteur est invité à lire [FORW:2] qui récapitule ces travaux. Des articles importants sur ce sujet sont entre autres [FORW:3], [FORW:4], [FORW:5], [FORW:10], [FORW:11], [FORW:12], [FORW:13], [FORW:14], et [INTER:10].

La quantité de mémoire qu'un routeur devrait avoir disponible pour traiter la demande instantanée de pointe lorsque les hôtes utilisent des politiques d'encombrement raisonnables, comme celles décrites dans [FORW:5], est fonction du produit de la bande passante de la liaison par le délai de chemin des flux qui utilisent la liaison, et donc la mémoire devrait augmenter avec la croissance de ce produit (bande passante*délai). La fonction exacte reliant la capacité de mémoire à la probabilité d'élimination n'est pas connue.

Lorsqu'un routeur reçoit un paquet au-delà de sa capacité de mémoire, il doit (par définition, pas par décret) l'éliminer, lui ou un ou des autres paquets. Quel paquet éliminer est le sujet de nombreuses études, mais malheureusement, de peu d'accord jusqu'à présent. La suggestion la plus sage à ce jour est d'éliminer un paquet du flux de données qui utilise le plus lourdement la liaison. Cependant, un certain nombre de facteurs additionnels peuvent être pertinents, y compris la préséance du trafic, la réservation active de bande passante, et la complexité associée au choix de ce paquet.

Un routeur PEUT éliminer le paquet qu'il vient de recevoir ; c'est la politique la plus simple, mais pas la meilleure. Dans l'idéal, le routeur devrait choisir un paquet dans une des sessions qui abusent le plus lourdement de la liaison, selon ce que permet la politique de qualité de service applicable. Une politique recommandée dans les environnements de datagrammes qui utilisent des files d'attentes FIFO, est d'éliminer un paquet choisi de façon aléatoire dans la file d'attente (voir [FORW:5]). Un algorithme équivalent dans les routeurs qui utilisent la file d'attente normale est l'éliminer le paquet dans la plus longue file d'attente ou celle qui utilise le plus long temps virtuel (voir [FORW:13]). Un routeur PEUT utiliser ces algorithmes pour déterminer quel paquet éliminer.

Si un routeur met en œuvre une politique d'élimination (comme l'abandon au hasard) selon laquelle il choisit un paquet à éliminer d'un groupe de paquets éligibles :

- Si le service de file d'attente à ordre de préséance (décrit au paragraphe 5.3.3.1) est mis en œuvre et activé, le routeur NE DOIT PAS éliminer un paquet dont la préséance IP est plus forte que celle d'un paquet qui n'est pas éliminé.
- Un routeur PEUT protéger des paquets dont l'en-tête IP demande le TOS de fiabilité maximisée, sauf lorsque le faire serait en violation de la règle précédente.
- Un routeur PEUT protéger des paquets IP fragmentés, selon la théorie où abandonner un fragment d'un datagramme peut accroître l'encombrement en provoquant la retransmission de tous les fragments du datagramme par la source.
- Pour aider à empêcher les perturbations d'acheminement ou les interruptions des fonctions de gestion, le routeur PEUT protéger les paquets utilisés pour le contrôle de l'acheminement, le contrôle de liaison, ou la gestion du réseau contre l'élimination. Les routeurs dédiés (c'est-à-dire, les routeurs qui ne sont pas aussi des hôtes d'utilisation générale, les serveurs de terminaux, etc.) peuvent effectuer une approximation de cette règle en protégeant les paquets dont la source ou la destination est le routeur lui-même.

Des méthodes évoluées de contrôle de l'encombrement incluent une notion de traitement équitable, de sorte que l'utilisateur qui est pénalisé par la perte d'un paquet soit celui qui a le plus contribué à l'encombrement. Peu importe le mécanisme mis en œuvre pour traiter le contrôle d'encombrement de la bande passante, il est important que l'effort de CPU effectué soit suffisamment petit pour que le routeur ne soit pas conduit aussi à un encombrement de CPU.

Comme décrit au paragraphe 4.3.3.3, le présent document recommande qu'un routeur NE DEVRAIT PAS envoyer un Source Quench à l'envoyeur du paquet qu'il élimine. Le Source Quench ICMP est un mécanisme très faible, de sorte qu'il n'est pas nécessaire qu'un routeur l'envoie, et le logiciel hôte ne devrait pas l'utiliser exclusivement comme un indicateur d'encombrement.

5.3.7 Filtrage d'adresse martienne

Une adresse IP de source est invalide si elle est une adresse IP spéciale, comme défini dans 4.2.2.11 ou 5.3.7, ou si ce n'est pas une adresse en diffusion individuelle.

Une adresse de destination IP est invalide si elle fait partie de celles définies comme destinations illégales en 4.2.3.1, ou si elle est une adresse de classe E (excepté 255.255.255.255).

Un routeur NE DEVRAIT PAS transmettre de paquet ayant une adresse IP de source invalide ou une adresse de source sur le réseau 0. Un routeur NE DEVRAIT PAS transmettre, excepté sur une interface de bouclage, de paquet qui ait une adresse de source sur le réseau 127. Un routeur PEUT avoir un commutateur qui permette au gestionnaire de réseau de désactiver ces vérifications. Si un tel commutateur est fourni, il DOIT par défaut effectuer les vérifications.

Un routeur NE DEVRAIT PAS transmettre de paquet avec une adresse de destination IP invalide ou une adresse de destination sur le réseau 0. Un routeur NE DEVRAIT PAS transmettre, excepté sur une interface de bouclage, de paquet avec une adresse de destination sur le réseau 127. Un routeur PEUT avoir un commutateur qui permette au gestionnaire de réseau de désactiver ces vérifications. Si un tel commutateur est fourni, il DOIT par défaut effectuer ces vérifications.

Si un routeur élimine un paquet à cause de ces règles, il DEVRAIT enregistrer au moins l'adresse IP de source, l'adresse de destination IP, et, si le problème est avec l'adresse de source, l'interface physique sur laquelle le paquet a été reçu et l'adresse de couche Liaison de l'hôte ou routeur d'où le paquet a été reçu.

5.3.8 Validation d'adresse de source

Un routeur DEVRAIT METTRE EN ŒUVRE la capacité à filtrer le trafic sur la base d'une comparaison de l'adresse de source d'un paquet et du tableau de transmission pour une interface logique sur laquelle le paquet a été reçu. Si ce filtre est activé, le routeur DOIT éliminer en silence un paquet si l'interface sur laquelle il est reçu n'est pas l'interface sur laquelle le paquet aurait dû être transmis pour atteindre l'adresse contenue dans l'adresse de source. En termes plus simples, si un routeur ne devait pas acheminer un paquet contenant cette adresse par une interface particulière, il ne devrait pas croire l'adresse si elle apparaît comme une adresse de source dans un paquet lu à partir de cette interface.

Si ce dispositif est mis en œuvre, il DOIT être désactivé par défaut.

Discussion

Ce dispositif peut fournir des améliorations utiles de la sécurité dans certaines situations, mais peut par erreur éliminer des paquets valides dans des situations où les chemins sont asymétriques.

5.3.9 Filtrage de paquet et listes d'accès

Un routeur DEVRAIT fournir la capacité de transmettre de façon sélective (ou de filtrer) les paquets, au titre de la fourniture de sécurité et/ou de la limitation du trafic sur des portions d'un réseau. Si cette capacité est fournie, le filtrage de paquets DEVRAIT être configurable soit pour transmettre tous les paquets soit pour les transmettre de façon sélective selon les préfixes de source et de destination, et PEUT filtrer sur d'autres attributs de message. Chaque adresse de source et de destination DEVRAIT permettre la spécification d'une longueur de préfixe arbitraire.

Discussion

Ce dispositif peut fournir une mesure de la confidentialité, lorsque les systèmes hors frontière ne sont pas autorisés à échanger certains protocoles avec les systèmes à l'intérieur des frontières, ou sont limités quant aux systèmes avec lesquels ils peuvent communiquer. Il peut aussi aider à prévenir certaines classes d'infractions à la sécurité, par lesquelles un système hors frontière se fait passer pour un système interne et contrefait une session avec lui.

S'il est pris en charge, un routeur DEVRAIT être configurable pour permettre soit :

- une liste d'inclusion - spécification d'une liste de définitions de message à transmettre, ou
- une liste d'exclusion - spécification d'une liste de définitions de message à NE PAS transmettre.

Une "définition de message", dans ce contexte, spécifie le préfixe de réseau de source et de destination, et peut inclure d'autres informations d'identification telles que le type de protocole IP ou le numéro d'accès TCP.

Un routeur PEUT fournir une commutation de configuration permettant un choix entre spécifier une liste d'inclusion ou d'exclusion, ou d'autres contrôles équivalents.

Une valeur correspondant à n'importe quelle adresse (par exemple, un mot clé "any", une adresse avec un gabarit de tout à zéro, ou un préfixe de réseau dont la longueur est zéro) DOIT être admis comme adresse de source et/ou destination.

En plus des paires d'adresses, le routeur PEUT permettre de spécifier toute combinaison de protocole de transport et/ou application et d'accès de source et de destination.

Le routeur DOIT permettre que les paquets soit éliminés en silence (c'est-à-dire, éliminés sans envoi d'un message d'erreur ICMP).

Le routeur DEVRAIT permettre l'envoi d'un message ICMP inaccessible approprié lorsque un paquet est éliminé. Le message ICMP DEVRAIT spécifier Communication administrativement interdite (code 13) comme raison de la destination inaccessible.

Le routeur DEVRAIT permettre l'envoi de messages ICMP de destination inaccessible (code 13) configurés pour

chaque combinaison de paires d'adresses, de types de protocole, et d'accès qu'il permet de spécifier.

Le routeur DEVRAIT compter et DEVRAIT permettre un enregistrement sélectif des paquets non transmis.

5.3.10 Acheminement en diffusion groupée

Un routeur IP DEVRAIT prendre en charge la transmission de paquets en diffusion groupée IP, sur la base de chemins de diffusion groupée statiques ou de chemins déterminés de façon dynamique par un protocole d'acheminement de diffusion groupée (par exemple, DVMRP [ROUTE:9]). Un routeur qui transmet des paquets IP en diffusion groupée est appelé un routeur de diffusion groupée.

5.3.11 Contrôles sur la transmission

Pour chaque interface physique, un routeur DEVRAIT avoir une option de configuration qui spécifie si la transmission est activée sur cette interface. Lorsqu'il transmet sur une interface désactivée, le routeur :

- DOIT éliminer en silence tous les paquets qui sont reçus sur cette interface mais non adressés au routeur
- NE DOIT PAS envoyer de paquets par cette interface, sauf pour les datagrammes dont il est lui-même l'origine
- NE DOIT PAS annoncer via un protocole d'acheminement la disponibilité de chemins par cette interface.

Discussion

Ce dispositif permet au gestionnaire de chemin réseau de couper au trafic une interface mais de la laisser accessible pour la gestion de réseau.

Dans l'idéal, cette commande devrait s'appliquer plutôt au logiciel qu'à des interfaces physiques. Cela ne peut être, parce que il n'y a pas de moyen connu pour qu'un routeur détermine sur quelle interface logique est arrivé un paquet en l'absence d'une correspondance bijective entre les interfaces logiques et physiques.

5.3.12 Changements d'état

Durant le fonctionnement du routeur, des interfaces peuvent tomber en panne ou être désactivées manuellement, ou peuvent devenir disponibles pour l'utilisation du routeur. De même, la transmission peut être désactivée sur une interface particulière ou pour le routeur tout entier, ou peut être (ré)activée. Bien que de telles transitions soient (habituellement) peu communes, il est important que les routeurs les traitent correctement.

5.3.12.1 Lorsqu'un routeur cesse de transmettre

Lorsqu'un routeur cesse de transmettre, il DOIT arrêter de publier tous les chemins, sauf pour les chemins de tiers. Il PEUT continuer à recevoir et utiliser les chemins provenant des autres routeurs dans ses domaines d'acheminement. Si la base de données de transmission est conservée, le routeur NE DOIT PAS cesser la temporisation des chemins dans la base de données de transmission. Si les chemins qui ont été reçus des autres routeurs sont mémorisés, le routeur NE DOIT PAS cesser la temporisation des chemins qu'il a mémorisés. Il DOIT éliminer tous les chemins dont la temporisation arrive à expiration alors que la transmission est désactivée, tout comme il le ferait si la transmission était activée.

Discussion

Lorsque un routeur cesse de transmettre, il cesse essentiellement d'être un routeur. Il est toujours un hôte, et doit suivre toutes les exigences des hôtes [INTRO:2]. Le routeur peut cependant toujours être un membre passif d'un ou plusieurs domaines d'acheminement. Comme tel, il lui est permis d'entretenir sa base de données de transmission en écoutant les autres routeurs de son domaine d'acheminement. Il ne peut cependant pas publier les chemins de sa base de données de transmission, car il n'effectue aucune transmission lui-même. La seule exception à cette règle est lorsque le routeur publie un chemin qui n'utilise qu'un autre routeur, mais que c'est ce routeur qui lui a demandé de le publier.

Un routeur PEUT envoyer des messages ICMP Destination inaccessible (hôte inaccessible) à l'expéditeur des paquets qu'il est incapable de transmettre. Il NE DEVRAIT PAS envoyer de messages ICMP redirect.

Discussion

Noter que l'envoi d'un message ICMP Destination inaccessible (hôte inaccessible) est une action de routeur. Ce message ne devrait pas être envoyé par des hôtes. Cette exception aux règles des hôtes est admise de façon que les paquets puissent être réacheminés dans les plus brefs délais possibles, afin qu'il n'y ait pas de trous noirs.

5.3.12.2 Lorsqu'un routeur commence à transmettre

Lorsque un routeur commence à transmettre, il DEVRAIT expédier l'envoi des nouvelles informations d'acheminement à tous les routeurs avec lesquels il échange normalement des informations d'acheminement.

5.3.12.3 Lorsqu'une interface échoue ou est désactivée

Si une interface est en panne ou est désactivée, un routeur DOIT le retirer de sa base de données de transmission et cesser de publier toutes les chemins qui font usage de cette interface. Il DOIT désactiver tous les chemins statiques qui utilisent cette interface. Si d'autres chemins pour la même destination et TOS sont acquis ou remémorés par le routeur, celui-ci DOIT choisir la meilleure solution de remplacement, et l'ajouter à sa base de données de transmission. Le routeur DEVRAIT envoyer des messages ICMP Destination inaccessible ou Redirect, selon le cas, en réponse à tous les paquets qu'il n'est pas capable de transmettre du fait que l'interface est devenue indisponible.

5.3.12.4 Lorsqu'une interface est activée

Si une interface qui était indisponible devient disponible, un routeur DOIT réactiver tous les chemins statiques qui utilisent cette interface. Si des chemins qui pourraient utiliser cette interface sont appris par le routeur, ces chemins DOIVENT alors être évalués par rapport à tous les autres chemins appris, et le routeur DOIT prendre une décision quant au chemin qui devrait être entré dans la base de données de transmission. Les développeurs sont invités à se reporter à la section 7, Couche d'application – Protocoles d'acheminement, pour des précisions sur la façon dont cette décision est prise.

Un routeur DEVRAIT expédier l'envoi de nouvelles informations d'acheminement à tous les routeurs avec lesquels il échange normalement des informations d'acheminement.

5.3.13 Options IP

Plusieurs options, telles que Record Route et Timestamp, contiennent des espaces dans lesquels un routeur insère son adresse lors de la transmission du paquet. Cependant, chacune de ces options a un nombre d'espaces fini, et donc un routeur peut trouver qu'il n'y a plus d'espace libre dans lequel il puisse insérer son adresse. Aucune des exigences figurant ci-dessous ne devrait être construite comme exigeant d'un routeur qu'il insère son adresse dans une option qui n'a pas d'espaces restants libres pour qu'il s'y insère. Le paragraphe 5.2.5 expose comment un routeur doit choisir laquelle de ses adresses insérer dans une option.

5.3.13.1 Options non reconnues

Les options IP non reconnues dans les paquets transmis DOIVENT être passées inchangées.

5.3.13.2 Option de sécurité

Certains environnements exigent l'option Sécurité dans chaque paquet ; une telle exigence sort du domaine d'application du présent document et de la spécification IP standard. Noter cependant, que les options de sécurité décrites dans [INTER:1] et [INTER:16] sont obsolètes. Les routeurs DEVRAIENT METTRE EN ŒUVRE l'option de sécurité révisée décrite dans [INTER:5].

Discussion

Les routeurs destinés à être utilisés dans des réseaux qui ont plusieurs niveaux de sécurité devraient prendre en charge le filtrage des paquets sur la base des étiquettes IPSO (RFC1108). Pour mettre en œuvre cette prise en charge, le routeur devra permettre à l'administrateur de routeur de configurer à la fois une limite de sensibilité inférieure (par exemple Unclassified) et une limite de sensibilité supérieure (par exemple Secret) sur chaque interface. C'est habituellement le cas, mais pas toujours, que les deux limites soient les mêmes (par exemple, une interface à un seul niveau). Les paquets qu'un filtre IPSO rejette comme hors gamme devraient être éliminés en silence et un compteur devrait noter le nombre de paquets abandonnés à cause d'étiquette IPSO les classant hors gamme.

5.3.13.3 Option d'identifiant de flux

Cette option est obsolète. Si l'option Identifiant de flux est présente dans un paquet transmis par le routeur, l'option DOIT être ignorée et passée inchangée.

5.3.13.4 Options de route de source

Un routeur DOIT mettre en œuvre la prise en charge de l'option de route de source dans les paquets transmis. Un routeur PEUT mettre en œuvre une option de configuration qui, lorsqu'elle est activée, provoque l'élimination de tous les paquets acheminés par la source. Cependant, une telle option NE DOIT PAS être activée par défaut.

Discussion

La capacité à générer des datagrammes d'acheminement à travers l'Internet est importante pour divers outils de diagnostic du réseau. Cependant, l'acheminement à partir de la source peut être utilisé pour contourner des contrôles administratifs et de sécurité au sein d'un réseau. En particulier, dans les cas où la manipulation de tableaux d'acheminement est utilisée pour faire une séparation administrative au lieu d'utiliser d'autres méthodes comme le filtrage de paquet, les paquets acheminés à partir de la source peuvent constituer une faiblesse.

COMMENTAIRES SUPPLEMENTAIRES DE L'EDITEUR

Le filtrage de paquet peut aussi bien être mis en échec par l'acheminement de source, si il est appliqué dans tous les routeurs excepté celui de la branche finale du chemin conduit par la source. Aucun filtre de route ni de paquet ne constitue une solution parfaite pour la sécurité.

5.3.13.5 Option d'enregistrement de route

Les routeurs DOIVENT prendre en charge l'option Record Route dans les paquets transmis.

Un routeur PEUT fournir une option de configuration qui, si elle est activée, va amener le routeur à ignorer (c'est-à-dire, passer inchangées) les options Record Route dans les paquets transmis. Si elle est fournie, une telle option DOIT activer l'enregistrement de chemin par défaut. Cette option ne devrait pas affecter le traitement des options Record Route dans les datagrammes reçus par le routeur lui-même (en particulier, les options Record Route dans les demandes d'écho ICMP seront toujours traitées conformément au paragraphe 4.3.3.6).

Discussion

Certaines personnes croient que l'option Record Route pose un problème de sécurité parce qu'elle divulgue des informations sur la topologie du réseau. Et donc, le présent document permet qu'elle soit désactivée.

5.3.13.6 Option d'horodatage

Les routeurs DOIVENT prendre en charge l'option d'horodatage dans les paquets transmis. Une valeur d'horodatage DOIT suivre les règles données dans [INTRO:2].

Si le champ des fanions = 3 (horodatage et adresse pré spécifiée), le routeur DOIT ajouter son horodatage si la prochaine adresse pré spécifiée correspond à une des adresses IP du routeur. Il n'est pas nécessaire que l'adresse pré spécifiée soit l'adresse de l'interface sur laquelle le paquet est arrivé ou l'adresse de l'interface sur laquelle il sera envoyé.

MISE EN ŒUVRE

Pour maximiser l'utilité de l'horodatage contenu dans l'option horodatage, il est suggéré que l'horodatage inséré soit, d'aussi près que possible, l'heure à laquelle le paquet est arrivé au routeur. Pour les datagrammes générés par le routeur, l'horodatage inséré devrait être, d'aussi près que possible, l'heure à laquelle le datagramme a été passé à la couche réseau pour transmission.

Un routeur PEUT fournir une option de configuration qui, si elle est activée, amènera le routeur à ignorer (c'est-à-dire, à les passer inchangées) les options d'horodatage dans les datagrammes transmis lorsque le mot fanion est réglé à zéro (seulement des horodatages) ou un (horodatage et adresse IP d'enregistrement). Si elle est fournie, une telle option DOIT être désactivée par défaut (c'est-à-dire que le routeur n'ignore pas l'horodatage). Cette option ne devrait pas affecter le traitement des options d'horodatage dans les datagrammes reçus par le routeur lui-même (en particulier, un routeur insèrera des horodatages dans les options horodatage dans les datagrammes reçus par le routeur, et les options d'horodatage dans les demandes d'écho ICMP seront toujours traitées conformément au paragraphe 4.3.3.6).

Discussion

Comme l'option Record Route, l'option horodatage peut révéler des informations sur la topologie du réseau. Certains considèrent que cela pose un problème pour la sécurité.

6. Couche Transport

Un routeur n'est pas nécessaire pour mettre en œuvre un protocole de couche Transport, sauf ceux nécessaires pour prendre en charge les protocoles de couche Application acceptés par le routeur. En pratique, cela signifie que la plupart des routeurs mettent en œuvre à la fois le protocole de commande de transmission (TCP) et le protocole de datagramme d'utilisateur (UDP).

6.1 Protocole de datagramme d'utilisateur - UDP

Le Protocole de datagramme d'utilisateur (UDP) est spécifié dans [TRANS:1].

Un routeur qui met en œuvre UDP DOIT être conforme, et DEVRAIT être inconditionnellement conforme, aux exigences de [INTRO:2], à l'exception de ce qui suit :

- La présente spécification ne spécifie pas les interfaces entre les diverses couches de protocole. Et donc, les interfaces d'un routeur ne doivent pas nécessairement être conformes à [INTRO:2], excepté lorsque la conformité est exigée pour le bon fonctionnement des protocoles de couche Application pris en charge par le routeur.
- Contrairement à [INTRO:2], une application NE DEVRAIT PAS désactiver la génération des sommes de contrôle UDP.

Discussion

Bien qu'un protocole d'application particulier puisse exiger que les datagrammes UDP qu'il reçoit contiennent une somme de contrôle UDP, il n'y a pas d'exigence générale que les datagrammes UDP reçus contiennent des sommes de contrôle UDP. Bien sûr, si une somme de contrôle UDP est présente dans un datagramme reçu, la somme doit être vérifiée et le datagramme doit être éliminé si la somme est incorrecte.

6.2 Protocole de commande de transmission - TCP

Le protocole de commande de transmission (TCP) est spécifié dans [TRANS:2].

Un routeur qui met en œuvre TCP DOIT être conforme, et DEVRAIT être inconditionnellement conforme, aux exigences de [INTRO:2], avec les exceptions suivantes :

La présente spécification ne spécifie pas les interfaces entre les diverses couches de protocole. Et donc, un routeur ne doit pas nécessairement se conformer aux exigences de [INTRO:2] suivantes (excepté bien sûr lorsque la conformité est exigée pour le bon fonctionnement des protocoles de couche Application acceptés par le routeur) :

Utilisation de Push (RFC-793 paragraphe 2.8) :

Passer un fanion PSH reçu à la couche Application est maintenant FACULTATIF.

Pointeur de données urgentes (RFC-793 paragraphe 3.1) :

Un TCP DOIT informer la couche application de façon asynchrone chaque fois qu'il reçoit un pointeur Urgent et qu'il n'y avait pas précédemment de données urgentes en cours, ou chaque fois que le pointeur Urgent avance dans le flux de données. Il DOIT être une façon pour l'application d'apprendre combien de données urgentes restent à lire depuis la connexion, ou au moins de déterminer s'il reste ou non des données urgentes à lire.

Échecs de connexion TCP :

Une application DOIT être capable de régler la valeur pour R2 sur une connexion particulière. Par exemple, une application interactive peut régler R2 à "infini", donnant à l'utilisateur le contrôle sur le moment de déconnexion.

Multi localisation TCP :

Si une application sur un hôte multi localisé ne spécifie pas l'adresse IP locale lorsqu'il ouvre de façon active une connexion TCP, le TCP DOIT alors demander à la couche IP de choisir une adresse IP locale avant d'envoyer le (premier) SYN. Voir la fonction GET_SRCADDR() au paragraphe 3.4.

Options IP :

Une application DOIT être capable de spécifier un chemin de source lorsqu'elle ouvre de façon active une connexion TCP, et cela DOIT prendre le pas sur un chemin de source reçu dans un datagramme.

Pour des raisons similaires, un routeur n'a pas besoin de se conformer aux exigences de [INTRO:2].

Les exigences concernant l'option de taille de segment maximum dans [INTRO:2] sont amendées comme suit : un routeur qui met en œuvre la portion hôte de découverte de MTU (exposée au paragraphe [4.2.3.3] du présent mémoire) n'utilise 536 comme valeur par défaut de SendMSS que si le MTU de chemin est inconnu ; si le MTU de chemin est connu, la valeur par défaut pour SendMSS est le MTU du chemin - 40.

Les exigences concernant l'option de taille de segment maximum dans [INTRO:2] sont amendées comme suit : les codes ICMP Destination inaccessible 11 et 12 sont des conditions d'erreur douce supplémentaires. Donc, ces messages NE DOIVENT PAS causer l'interruption d'une connexion par TCP.

Discussion

Il faut particulièrement noter qu'une mise en œuvre de TCP dans un routeur doit se conformer aux exigences suivantes de [INTRO:2] :

- Fournir un TTL configurable. (Durée de vie : RFC-793 paragraphe 3.9)
- Fournir une interface pour configurer le comportement garder en vie, si garder en vie est utilisé. [TCP Keep-Alive]
- Fournir un mécanisme de rapport d'erreur, et la capacité à le gérer. (Rapports asynchrones)
- Spécifier le type de service. (Type-de-Service)

Le paradigme général appliqué est que si une interface particulière est visible à l'extérieur du routeur, toutes les exigences pour l'interface doivent alors être suivies. Par exemple, si un routeur fournit une fonction telnet, il générera alors du trafic, vraisemblablement à acheminer dans les réseaux externes. Donc, il doit être capable de régler le type de service correctement ou alors le trafic telnet pourrait ne pas passer.

7 Couche Application - Protocoles d'acheminement

7.1 Introduction

Pour des raisons techniques, de gestion, et parfois de politique, le système d'acheminement de l'Internet comporte deux composantes – l'acheminement intérieur et l'acheminement extérieur. Le concept de système autonome (AS), comme défini au paragraphe 2.2.4 du présent document, joue un rôle clé dans la séparation de l'acheminement intérieur de l'acheminement extérieur, car ce concept permet de délimiter l'ensemble des routeurs où survient un changement de l'acheminement intérieur vers l'acheminement extérieur. Un datagramme IP peut avoir à traverser les routeurs de deux systèmes autonomes ou plus pour atteindre sa destination, et le système autonome doit fournir à chacun des autres les informations de topologie pour permettre une telle transmission. Les protocoles de passerelle intérieure (IGP) sont utilisés pour distribuer les informations d'acheminement au sein d'un AS (c'est-à-dire, l'acheminement intra-AS). Les protocoles de passerelle extérieure sont utilisés pour échanger les informations d'acheminement entre les AS (c'est-à-dire, l'acheminement inter-AS).

7.1.1 Considérations sur la sécurité de l'acheminement

L'acheminement est un des rares endroits où le principe de robustesse (être libéral dans ce que l'on accepte) ne s'applique pas. Les routeurs devraient être relativement soupçonneux pour accepter de données d'acheminement provenant d'autres systèmes d'acheminement.

Un routeur DEVRAIT fournir la capacité de hiérarchiser les sources d'informations d'acheminement de la plus fiable à la moins fiable et d'accepter en premier les informations d'acheminement sur toute destination particulière de la part des sources les plus fiables. Ceci était implicite dans le modèle d'origine d'acheminement de cœur/tronçon de système autonome utilisant EGP et divers protocoles d'acheminement intérieurs. Il est encore plus important avec la disparition d'un noyau central de confiance.

Un routeur DEVRAIT fournir un mécanisme pour éconduire les chemins évidement invalides (telles que ceux pour un réseau 127).

Les routeurs NE DOIVENT PAS redistribuer par défaut les données d'acheminement qu'ils n'utilisent pas eux-mêmes, leur faire confiance ou les considérer comme valides. Dans de rares cas, il peut être nécessaire de redistribuer des informations suspectes, mais cela ne devrait arriver que sous la supervision directe de quelque agent humain.

Les routeurs doivent être au moins un peu paranoïaques quant à l'acceptation de données d'acheminement provenant de quiconque, et doivent être particulièrement attentifs lorsqu'ils distribuent des informations d'acheminement qui leur ont été fournies par des tiers. Voir plus loin les lignes directrices spécifiques.

7.1.2 Préséance

Excepté lorsque la spécification d'un protocole d'acheminement particulier en décide autrement, un routeur DEVRAIT régler la valeur de préséance IP pour les datagrammes IP portant du trafic d'acheminement qu'il a généré à 6 (CONTROLE INTER RESEAU).

Discussion

Le trafic d'acheminement devrait avec TRES PEU d'exceptions être le trafic de plus forte préséance sur tous les réseaux. Si le trafic d'acheminement d'un système ne peut pas passer, il y a des chances que rien d'autre ne passe.

7.1.3 Validation de message

L'authentification d'homologue à homologue implique plusieurs vérifications. L'application de mots de passe de message et de listes explicites de voisins acceptables a dans le passé amélioré la robustesse de la base de données d'acheminement. Les routeurs DEVRAIENT METTRE EN ŒUVRE des contrôles de gestion qui activent la constitution de listes explicites de voisins d'acheminement valides. Les routeurs DEVRAIENT METTRE EN ŒUVRE l'authentification d'homologue à homologue pour les protocoles d'acheminement qui les prennent en charge.

Les routeurs DEVRAIENT valider les voisins d'acheminement sur la base de leur adresse de source et de l'interface sur laquelle est reçu un message ; les voisins, dans un sous-réseau directement relié, DEVRAIENT être limités aux communications avec le routeur via l'interface sur laquelle ce sous-réseau est positionné ou via des interfaces non numérotées. Les messages reçus sur d'autres interfaces DEVRAIENT être éliminés en silence.

Discussion

Les atteintes à la sécurité et de nombreux problèmes d'acheminement sont évités par cette vérification de base.

7.2 Protocoles de passerelles intérieures

7.2.1 Introduction

Un protocole de passerelle intérieure (IGP) est utilisé pour distribuer les informations d'acheminement entre les divers routeurs dans un AS particulier. Indépendamment de l'algorithme utilisé pour mettre en œuvre un IGP particulier, il devrait effectuer les fonctions suivantes :

(1) Répondre rapidement aux changements de la topologie interne d'un AS
(2) Fournir un mécanisme tel que la défaillance d'un circuit ne cause pas de mises à jour continuelles de l'acheminement
(3) Fournir une convergence rapide pour un acheminement sans bouclage
(4) Utiliser la bande passante minimale
(5) Fournir des routes de coût égal pour activer l'étalement de la charge
(6) Fournir un moyen d'authentification des mises à jour de l'acheminement.

Les IGP actuels utilisés dans l'Internet se caractérisent par leur vecteur distance ou par un algorithme d'état de liaison.

Plusieurs IGP sont exposés dans cette section, y compris ceux le plus couramment utilisés et certains protocoles récemment développés qui pourraient être largement utilisés à l'avenir. De nombreux autres protocoles destinés à être utilisés dans l'acheminement intra-AS existent dans la communauté de l'Internet.

Un routeur qui met en œuvre un protocole d'acheminement (autre que de chemins statiques) DOIT METTRE EN ŒUVRE OSPF (voir au paragraphe 7.2.2). Un routeur PEUT mettre en œuvre des IGP supplémentaires.

7.2.2 Plus court chemin ouvert d'abord - OSPF

Les protocoles d'acheminement fondés sur le plus court chemin d'abord (SPF, *Shortest Path First*) sont une classe d'algorithmes d'état de liaison qui se fondent sur l'algorithme du plus court chemin de Dijkstra. Bien que les algorithmes fondés sur SPF soient utilisés depuis l'invention d'ARPANET, c'est seulement récemment qu'ils sont réellement devenus populaires à la fois dans la communauté IP et OSI. Dans les systèmes fondés sur SPF, chaque routeur obtient la totalité de la base de données de topologie par un processus connu sous le nom de "inondation". L'inondation assure un transfert fiable des informations. Chaque routeur fait alors tourner l'algorithme SPF dans sa base de données pour construire le tableau des acheminements IP. Le protocole d'acheminement OSPF est une mise en œuvre d'un algorithme SPF. La version actuelle, OSPF version 2, est spécifiée dans [ROUTE:1]. Noter que la RFC 1131, qui décrit OSPF version 1, est obsolète.

Noter que pour se conformer au paragraphe 8.3 du présent mémoire, un routeur qui met en œuvre OSPF DOIT mettre en œuvre la MIB OSPF [MGT:14].

7.2.3 Système intermédiaire à système intermédiaire - DUAL IS-IS

Le comité X3S3.3 de l'Institut national américain de normalisation (ANSI) a défini un protocole d'acheminement intra domaine. Ce protocole est intitulé Protocole d'échange d'acheminement de système intermédiaire à système intermédiaire.

Son application à un réseau IP a été défini dans [ROUTE:2], et on s'y réfère sous le nom de Dual IS-IS (ou parfois sous le nom de IS-IS intégré). IS-IS se fonde sur un algorithme d'acheminement à état des liaisons (SPF) et bénéficie de tous les avantages de cette classe de protocoles.

7.3 Protocoles de passerelles extérieures

7.3.1 Introduction

Les protocoles de passerelles extérieures sont utilisés par les systèmes d'acheminement inter-autonomes pour échanger des informations d'accessibilité sur un ensemble de réseaux internes à un système autonome particulier vers un système autonome voisin.

La zone d'acheminement inter-AS est un sujet de recherches actuel au sein de l'IETF. Le protocole de passerelle extérieure (EGP) décrit à l'Appendice F.1 a traditionnellement été le protocole inter-AS choisi, mais il est maintenant dépassé. Le protocole de passerelle frontière (BGP) élimine la plupart des restrictions et limitations d'EGP, et sa popularité est en croissance rapide. Il n'est pas exigé d'un routeur qu'il mette en œuvre un protocole d'acheminement inter-AS. Cependant, si un routeur met en œuvre EGP, il DOIT METTRE EN ŒUVRE aussi BGP. Bien qu'il ne soit pas conçu comme un protocole de passerelle extérieure, RIP (décrit au paragraphe 7.2.4) est parfois utilisé pour l'acheminement inter-AS.

7.3.2 Protocole de passerelle frontière - BGP

7.3.2.1 Introduction

Le protocole de passerelle frontière (BGP-4) est un protocole d'acheminement inter-AS qui échange des informations d'accessibilité de réseau avec d'autres interlocuteurs BGP. Les informations pour un réseau incluent la liste complète des AS par lesquelles le trafic doit transiter pour atteindre ce réseau. Ces informations peuvent alors être utilisées pour s'assurer de chemins sans mise en boucle. Ces informations sont suffisantes pour construire un graphe de la connexité d'AS d'où on puisse éliminer les boucles et où on puisse mettre en application certaines décisions de politique au niveau de l'AS.

BGP est défini par [ROUTE:4]. [ROUTE:5] spécifie le bon usage de BGP dans l'Internet, et donne quelques utiles conseils et lignes directrices de mise en œuvre. [ROUTE:12] et [ROUTE:13] fournissent des informations supplémentaires utiles.

Pour se conformer aux exigences du paragraphe 8.3 du présent mémoire, un routeur qui met en œuvre BGP doit mettre en œuvre la MIB BGP [MGT:15].

Pour caractériser l'ensemble de décisions de politique qui peuvent être mises en application en utilisant BGP, on doit se focaliser sur la règle selon laquelle un AS ne prévient ses AS voisins que des chemins qu'il utilise lui-même. Cette règle reflète le paradigme d'acheminement bond par bond généralement utilisé partout sur l'Internet actuel. Noter que certaines politiques ne peuvent pas être prises en charge par le paradigme d'acheminement bond par bond et exigent donc des techniques telles que l'acheminement de source pour être appliquées. Par exemple, BGP n'active pas un AS pour envoyer du trafic à l'AS voisin sachant que ce trafic prend un chemin différent de celui emprunté par le trafic qui prend son origine dans l'AS voisin. D'un autre côté, BGP peut prendre en charge toute politique conforme au paradigme d'acheminement bond par bond.

Les développeurs de BGP sont vivement encouragés à suivre les recommandations formulées à la section 6 de [ROUTE:5].

7.3.2.2 Tour d'horizon des protocoles

Alors que BGP fournit la prise en charge de politiques d'acheminement assez complexes (voir à titre d'exemple le paragraphe 4.2 de [ROUTE:5]) il n'est pas exigé que toutes les mises en œuvre de BGP prennent en charge de telles politiques. Au minimum, cependant, une mise en œuvre BGP :

(1) DEVRAIT permettre qu'un AS contrôle les annonces des chemins BGP appris dans les AS adjacents. Des mises en œuvre DEVRAIENT prendre en charge de tels contrôles avec au moins une granularité d'un seul réseau. Des mises en œuvre DEVRAIENT aussi prendre en charge de tels contrôles avec la granularité d'un système autonome, où le système autonome peut être le système autonome à l'origine du chemin, ou le système autonome qui a publié le chemin auprès du système local (système autonome adjacent).

(2) DEVRAIT permettre à un AS de préférer un chemin particulier vers une destination (lorsque plus d'un chemin est disponible). Une telle fonction DEVRAIT être mise en œuvre en permettant à un administrateur de système d'allouer des pondérations aux systèmes autonomes, et en faisant que le processus de sélection choisisse le chemin du moindre poids (où le poids d'un chemin est défini comme une somme de poids de tous les AS dans l'attribut de chemin AS_PATH associé à ce chemin).

(3) DEVRAIT permettre à un AS d'ignorer les chemins avec certains AS dans l'attribut de chemin AS_PATH. Une telle fonction peut être mise en œuvre en utilisant la technique exposée en (2), et en allouant l'infini comme pondération pour de tels AS. Le processus de choix de chemin doit ignorer ceux qui ont une pondération égale à l'infini.

7.3.3 Acheminement inter-AS sans protocole extérieur

Il est possible d'échanger des informations d'acheminement entre deux systèmes autonomes ou domaines d'acheminement sans utiliser de protocole d'acheminement extérieur standard entre deux protocoles d'acheminement intérieurs standard. La façon la plus courante de la faire est de faire fonctionner les deux protocoles intérieurs indépendamment dans un des routeurs frontière avec un échange des informations d'acheminement entre les deux traitements.

Comme avec l'échange d'information d'un EGP à un IGP, sans contrôles appropriés, ces échanges d'informations d'acheminement entre deux IGP dans un seul routeur sont sujets à la création de boucles.

7.4 Acheminement statique

L'acheminement statique donne un moyen de définition explicite du prochain bond à partir d'un routeur pour une destination particulière. Un routeur DEVRAIT fournir un moyen pour définir un chemin statique vers une destination, où la destination est définie par un préfixe de réseau. Le mécanisme DEVRAIT aussi permettre de spécifier une

métrique pour chaque chemin statique.

Un routeur qui prend en charge un protocole d'acheminement dynamique DOIT permettre la définition de chemins statiques par toute métrique valide pour le protocole d'acheminement utilisé. Le routeur DOIT fournir à l'utilisateur la capacité de spécifier une liste de chemins statiques qui peut être ou non propagée à travers le protocole d'acheminement. De plus, un routeur DEVRAIT prendre en charge les informations supplémentaires si il prend en charge un protocole d'acheminement qui pourrait utiliser les informations. Ce sont :

- TOS,
- Gabarit de sous-réseau,
- Longueur de préfixe,
- Une métrique spécifique d'un protocole d'acheminement donné qui peut importer le chemin.

Discussion

Notre intention est qu'on ne prenne en charge que les choses utiles au protocole d'acheminement donné. Le besoin du TOS ne devrait pas exiger du fabricant qu'il mettre en œuvre les autres parties si elles ne sont pas utilisées.

Si un routeur préfère un chemin statique à un chemin dynamique (ou vice versa) ou si la métrique associée est utilisée pour faire un choix entre les chemins statiques et dynamiques en compétition DEVRAIT être configurable pour chaque chemin statique.

Un routeur DOIT permettre qu'une métrique soit assignée à un chemin statique pour chaque domaine d'acheminement qu'il prend en charge. Chacune de ces métriques DOIT être explicitement assignée à un domaine d'acheminement spécifique. Par exemple :

chemin 10.0.0.0/8 via 192.0.2.3 rip métrique 3
chemin 10.21.0.0/16 via 192.0.2.4 ospf inter-area métrique 27
chemin 10.22.0.0/16 via 192.0.2.5 egp 123 métrique 99

Discussion

Il a été suggéré que, dans l'idéal, les chemins statiques devraient avoir des valeurs leur donnant la préférence sur les métriques (dans la mesure où les métriques ne peuvent être comparées qu'avec les métriques des autres chemins dans le même domaine d'acheminement, la métrique d'un chemin statique ne pourrait être comparée qu'avec la métrique d'autres chemins statiques). Ceci est contraire à certaines mises en œuvre courantes, où les chemins statiques ont réellement une métrique, et cette métrique est utilisée pour déterminer si un chemin dynamique particulier supplante le chemin statique pour la même destination. Et donc, le présent document utilise le terme de métrique plutôt que celui de préférence.

Cette technique fait essentiellement le chemin statique dans un chemin RIP, ou dans un chemin OSPF (ou autre, selon le domaine de la métrique). Et donc, l'algorithme de recherche de chemin de ce domaine s'applique. Cependant, ce N'EST PAS une fuite de chemin, en ce que contraindre un chemin statique à passer dans un domaine d'acheminement dynamique n'autorise pas le routeur à redistribuer le chemin au sein du domaine d'acheminement dynamique.

Pour les chemins statiques qui ne sont pas mis dans un domaine d'acheminement spécifique, l'algorithme de recherche de chemin est :

(1) correspondance de base
(2) plus longue correspondance
(3) TOS faible (si le TOS est pris en charge)
(4) meilleure métrique (où la métrique est définie par la mise en œuvre)

La dernière peut n'être pas nécessaire, mais elle est utile dans le cas où on veut avoir un chemin statique principal sur une interface et un chemin statique secondaire sur une interface de remplacement, avec solution de secours sur le chemin de remplacement en cas de défaillance de l'interface pour le chemin principal.

7.5 Filtrage des informations d'acheminement

Chaque routeur au sein d'un réseau prend des décisions de transmission sur la base d'informations contenues dans sa base de données de transmission. Dans un réseau simple, le contenu de la base de données peut être configuré de façon statique. Comme la complexité du réseau augmente, le besoin d'une mise à jour dynamique de la base de données de transmission devient critique pour l'efficacité du fonctionnement du réseau.

Si le flux des données au travers d'un réseau doit être aussi efficace que possible, il est nécessaire de fournir un mécanisme de contrôle de la propagation des informations qu'un routeur utilise pour construire sa base de données de transmission. Ce contrôle prend la forme du choix des sources d'informations d'acheminement qui devraient être

considérées comme de confiance et du choix des éléments d'information qu'on peut croire. La base de données de transmission résultante est une version filtrée des informations d'acheminement disponibles.

En plus de l'efficacité, le contrôle de la propagation des informations d'acheminement peut réduire l'instabilité en prévenant l'étalage d'informations d'acheminement incorrectes ou fausses.

Dans certains cas, la politique locale peut exiger que la totalité des informations d'acheminement ne soit pas largement divulguée.

Ces exigences de filtrage ne s'appliquent qu'aux protocoles non fondés sur SPF (et donc pas à tous les routeurs qui ne mettent pas en œuvre de protocoles à vecteur de distance).

7.5.1 Validation de route

Un routeur DEVRAIT enregistrer comme une erreur toute mise à jour d'acheminement qui publie un chemin violant les spécifications du présent mémoire, sauf si le protocole d'acheminement d'où provient la mise à jour reçue utilise ces valeurs pour coder des chemins particuliers (comme des chemins par défaut).

7.5.2 Filtrage de route de base

Le filtrage des informations d'acheminement permet le contrôle des chemins utilisés par un routeur pour transmettre les paquets qu'il reçoit. Un routeur devrait faire un choix parmi les sources d'informations d'acheminement qu'il écoute et les chemins qu'il croit. Donc, un routeur DOIT avoir la capacité de spécifier :

- sur quelles interfaces logiques les informations d'acheminement seront acceptées et quels chemins seront acceptés de la part de chaque interface logique,
- si tous les chemins ou seulement un chemin par défaut est publié sur une interface logique.

Certains protocoles d'acheminement ne reconnaissent pas les interfaces logiques comme source d'informations d'acheminement. Dans de tels cas, le routeur DOIT fournir la capacité de spécifier à partir de quels autres routeurs les informations d'acheminement seront acceptées.

Par exemple, supposons un routeur connectant un ou plusieurs réseaux terminaux à la portion principale ou au cœur de réseau d'un plus grand réseau. Comme chacun des réseaux terminaux n'a qu'un seul chemin d'entrée et de sortie, le routeur peut seulement avoir un seul chemin par défaut avec eux. Il publie les réseaux terminaux sur le réseau principal.

7.5.3 Filtrage d'acheminement évolué

Lorsque la topologie d'un réseau devient plus complexe, le besoin d'un filtrage d'acheminement plus complexe se fait sentir. Donc, un routeur DEVRAIT avoir la capacité de spécifier de façon indépendante pour chaque protocole d'acheminement :

- Quelles interfaces logiques ou quelles informations d'acheminement des routeurs (quels chemins) seront acceptées et quels chemins seront crus par chaque autre routeur ou interface logique,
- Quels chemins seront envoyés via quelles interfaces logiques, et
- Quelles informations d'acheminement des routeurs seront envoyées, si cela est accepté par le protocole d'acheminement utilisé ?

Dans de nombreuses situations, il est souhaitable d'allouer une hiérarchie de fiabilité aux informations d'acheminement reçues d'un autre routeur au lieu du simple choix entre croire ou ne pas croire figurant dans la première solution ci-dessus. Un routeur PEUT donner la capacité de spécifier :

- Une fiabilité ou une préférence allouée à chaque chemin reçu. Un chemin avec une plus forte fiabilité sera choisi face à un de moindre fiabilité sans considération de la métrique d'acheminement associée à chaque chemin.

Si un routeur accepte l'allocation de préférences, le routeur NE DOIT PAS propager de chemins qu'il ne préfère pas comme informations de premier rang. Si le protocole d'acheminement utilisé pour propager les chemins n'accepte pas de distinguer entre informations de premier et de troisième rang, le routeur NE DOIT PAS propager de chemins qu'il ne préfère pas.

Discussion

Par exemple, supposons qu'un routeur reçoive un chemin pour le réseau C du routeur R et un chemin pour le même réseau du routeur S. Si le routeur R est considéré comme plus fiable que le routeur S, le trafic destiné au réseau C sera transmis au routeur R sans considération du chemin reçu du routeur S.

Les informations d'acheminement pour des chemins que n'utilise pas le routeur (le routeur S dans l'exemple ci-dessus) NE DOIVENT être passées à aucun autre routeur.

7.6 Échange d'informations entre protocoles d'acheminement

Les routeurs DOIVENT être capables d'échanger des informations d'acheminement entre des protocoles

d'acheminement intérieur IP séparés, si des processus d'acheminement IP indépendants peuvent tourner dans le même routeur. Les routeurs DOIVENT fournir un mécanisme évitant les boucles d'acheminement lorsque les routeurs sont configurés pour l'échange bidirectionnel d'informations d'acheminement entre deux processus d'acheminement intérieur séparés. Les routeurs DOIVENT fournir un mécanisme de priorité pour choisir des chemins provenant de processus d'acheminement indépendants. Les routeurs DEVRAIENT fournir un contrôle administratif des échanges IGP-IGP lorsqu'ils sont utilisés à travers des frontières administratives.

Les routeurs DEVRAIENT fournir un mécanisme de traduction ou de transformation des métriques réseau par réseau. Les routeurs (ou protocoles d'acheminement) PEUVENT permettre la préférence globale de chemins extérieurs importés dans un IGP.

Discussion

Différents IGP utilisent différentes métriques, exigeant une technique de traduction lors de l'introduction d'informations provenant d'un protocole dans un autre protocole ayant une forme de métrique différente. Certains IGP peuvent faire tourner plusieurs instances dans le même routeur ou ensemble de routeurs. Dans ce cas, les informations de métrique peuvent être préservée exactement ou traduites.

Il y a au moins deux techniques de traduction entre différents processus d'acheminement. L'approche statique (ou accessibilité) utilise l'existence de la publication du chemin dans un IGP pour générer une publication de chemin dans l'autre IGP avec une métrique donnée. La traduction ou approche tabulaire utilise la métrique d'un IGP pour créer une métrique dans l'autre IGP par l'utilisation d'une fonction (comme l'ajout d'une constante) ou d'une recherche dans un tableau.

L'échange bidirectionnel d'informations d'acheminement est dangereux en l'absence de mécanismes de contrôle limitant les rétroactions. C'est le même problème que doivent affronter les protocoles d'acheminement à vecteur de distance avec le partage de l'horizon technique et que traite EGP avec la règle du tiers. Les acheminements en boucle peuvent être évités explicitement par l'utilisation de tableaux ou listes de chemins permis/refusés, ou implicitement par l'utilisation d'une règle d'horizon partagé, d'une règle d'exclusion de tiers, ou d'un mécanisme de marquage de chemin. Les fabricants sont invités à utiliser des techniques implicites lorsque c'est possible, pour faciliter l'administration par les opérateurs de réseau.

8 Protocoles de gestion de réseau de couche application

Noter que cette section subroge toutes les exigences établies dans "Gestion à distance" dans [INTRO:3].

8.1 Protocole simple de gestion de réseau - SNMP

8.1.1 Éléments du protocole SNMP

Les routeurs DOIVENT être gérables par SNMP [MGT:3]. SNMP DOIT opérer en utilisant UDP/IP comme protocoles de transport et de réseau. D'autres PEUVENT être pris en charge (par exemple, voir [MGT:25], [MGT:26], [MGT:27], et [MGT:28]). Les opérations de gestion SNMP DOIVENT s'effectuer comme si SNMP était mis en œuvre sur le routeur lui-même. Précisément, les opérations de gestion DOIVENT être effectuées par l'envoi de demandes de gestion SNMP à toutes les adresses IP allouées à toutes les interfaces du routeur. Le fonctionnement de la gestion réelle peut être effectué soit par le routeur soit par un mandataire pour le compte du routeur.

Discussion

Cette formulation est destinée à permettre la gestion soit par un mandataire, où l'appareil mandataire répond aux paquets SNMP qui ont une des adresses IP du routeur dans le champ d'adresse de destination des paquets, soit SNMP est directement mis en œuvre dans le routeur lui-même et reçoit les paquets et leur répond de la façon appropriée.

Il est important que les opérations de gestion puissent être envoyées à une des adresses IP du routeur. Lors du diagnostic des problèmes de réseau, le seul élément disponible pour identifier le routeur peut être une des adresses IP du routeur ; peut-être obtenue en cherchant dans le tableau d'acheminement d'un autre routeur.

Toutes les opérations SNMP (get, get-next, get-response, set, et trap) DOIVENT être mises en œuvre.

Les routeurs DOIVENT fournir un mécanisme de limitation du débit de génération des messages trap SNMP. Les routeurs PEUVENT fournir ce mécanisme par les algorithmes de gestion d'alerte asynchrone décrits dans [MGT:5].

Discussion

Bien qu'il y ait un accord général sur la nécessité de limiter le débit des traps, il n'y a pas encore de consensus sur la meilleure façon de le faire. La référence citée est considérée comme expérimentale.

8.2 Tableau de communauté

Pour les besoins de la présente spécification, nous supposons qu'il existe un "tableau de communauté" abstrait dans le routeur. Ce tableau contient plusieurs entrées, une pour chaque communauté spécifique, qui contient les paramètres nécessaires à une définition complète des attributs de cette communauté. La méthode de mise en œuvre réelle est, bien sûr, spécifique de la mise en œuvre.

Un tableau de communauté d'un routeur DOIT permettre au moins une entrée et DEVRAIT permettre au moins deux entrées.

Discussion

Un tableau de communauté de capacité zéro est inutile. Il signifierait que le routeur ne reconnaîtra aucune communauté et donc, toutes les opérations SNMP seront rejetées. Donc une entrée est la taille minimale utile du tableau. Avoir deux entrées permet à une entrée d'être limitée à un accès en lecture seule tandis que l'autre aurait des capacités d'écriture.

Les routeurs DOIVENT permettre à l'utilisateur d'examiner, ajouter, supprimer et changer manuellement (c'est-à-dire, sans utiliser SNMP) les entrées dans le tableau de communauté SNMP. L'utilisateur DOIT être capable de régler le nom de la communauté ou de construire une vue de MIB. L'utilisateur DOIT être capable de configurer les communautés en lecture seule (c'est-à-dire, de ne pas admettre SET) ou en lecture écriture (c'est-à-dire, d'admettre SET).

L'utilisateur DOIT être capable de définir au moins une adresse IP à laquelle sont envoyées les notifications pour chaque communauté ou vue de MIB, si des traps sont utilisés. Ces adresses DEVRAIENT être définissables par communauté ou vue de MIB. Il DEVRAIT être possible d'activer ou désactiver les notifications par communauté ou vue de MIB.

Un routeur DEVRAIT fournir la capacité de spécifier une liste des gestionnaires de réseau valides pour toute communauté particulière. Si elle est activée, un routeur DOIT valider l'adresse de source du datagramme SNMP par rapport à la liste et DOIT éliminer le datagramme si son adresse n'apparaît pas. Si le datagramme est éliminé, le routeur DOIT prendre toutes les actions appropriées pour un échec d'authentification SNMP.

Discussion

Ceci est un système d'authentification assez limité, mais couplé avec diverses formes de filtrage de paquet, il peut fournir une petite augmentation de la sécurité.

Le tableau de communauté DOIT être sauvegardé dans une mémoire non volatile.

L'état initial du tableau de communauté DEVRAIT contenir une entrée, avec l'accès public et en lecture seule de la chaîne du nom de communauté. L'état par défaut de cette entrée NE DOIT PAS envoyer de traps. Si elle est mise en œuvre, cette entrée DOIT alors rester dans le tableau de communauté jusqu'à ce que l'administrateur la change ou la supprime.

Discussion

Par défaut, les traps ne sont pas envoyés à cette communauté. Les PDU de traps sont envoyés aux adresses IP en diffusion individuelle. Cette adresse doit être configurée de quelque manière dans le routeur. Avant que ne survienne cette configuration, il n'y a pas de telle adresse, donc, à qui envoyer le trap ? Par défaut, l'envoi de traps à la communauté publique doit donc être désactivé. Ceci peut bien sûr être changé par une opération administrative une fois que le routeur est opérationnel.

8.3 MIB standard

Toutes les MIB appartenant à la configuration d'un routeur doivent être mises en œuvre ; à savoir :

- Les groupes système, interface, IP, ICMP, et UDP du MIB-II [MGT:2] DOIVENT être mis en œuvre.
- Les MIB d'extension d'interface [MGT:18] DOIVENT être mises en œuvre.
- La MIB de tableau de transmission IP [MGT:20] DOIT être mise en œuvre.
- Si le routeur met en œuvre TCP (par exemple, pour Telnet) le groupe TCP de MIB-II [MGT:2] DOIT être mis en œuvre.
- Si le routeur met en œuvre EGP, le groupe EGP de MIB-II [MGT:2] DOIT alors être mis en œuvre.
- Si le routeur accepte OSPF, la MIB OSPF [MGT:14] DOIT alors être mise en œuvre.
- Si le routeur accepte BGP, la MIB BGP [MGT:15] DOIT alors être mise en œuvre.
- Si le routeur a des interfaces Ethernet, 802.3, ou StarLan, la MIB genre Ethernet [MGT:6] DOIT être mise en œuvre.
- Si le routeur a des interfaces 802.4, la MIB 802.4 [MGT:7] DOIT être mise en œuvre.
- Si le routeur a des interfaces 802.5, la MIB 802.5 [MGT:8] DOIT être mise en œuvre.
- Si le routeur a des interfaces FDDI qui mettent en œuvre ANSI SMT 7.3, la MIB FDDI [MGT:9] DOIT être mise en œuvre.

- Si le routeur a des interfaces FDDI qui mettent en œuvre ANSI SMT 6.2, la MIB FDDI [MGT:29] DOIT être mise en œuvre.
- Si le routeur a des interfaces qui utilisent la signalisation V.24, telles que RS-232, V.10, V.11, V.35, V.36, ou RS-422/423/449, la MIB RS-232 [MGT:10] DOIT alors être mise en œuvre.
- Si le routeur a des interfaces T1/DS1, la MIB T1/DS1 [MGT:16] DOIT alors être mise en œuvre.
- Si le routeur a des interfaces T3/DS3, la MIB T3/DS3 [MGT:17] DOIT alors être mise en œuvre.
- Si le routeur a des interfaces SMDS, la MIB de protocole d'interface SMDS [MGT:19] DOIT être mise en œuvre.
- Si le routeur accepte PPP sur une de ses interfaces les MIB PPP [MGT:11], [MGT:12], et [MGT:13] DOIVENT être mises en œuvre.
- Si le routeur accepte RIP Version 2, la MIB RIP Version 2 [MGT:21] DOIT alors être mise en œuvre.
- Si le routeur accepte X.25 sur une des ses interfaces, les MIB X.25 [MGT:22, MGT:23 et MGT:24] DOIVENT alors être mises en œuvre.

8.4 MIB spécifiques de fabricants

Les MIB Internet standard et expérimentaux ne couvrent pas la totalité de la gamme des informations statistiques, d'état, de configuration et de contrôle qui peuvent être disponibles dans un élément de réseau. Ces informations sont néanmoins extrêmement utiles. Les fabricants de routeurs (et d'autres appareils de réseau) ont généralement développé des extensions de MIB qui traitent ces informations. Ces extensions de MIB sont appelées MIB spécifiques de fabricant.

La MIB spécifique de fabricant pour le routeur DOIT fournir l'accès à toutes les informations statistiques, d'état, de configuration, et de contrôle qui ne sont pas disponibles par les MIB standard et expérimentales qui ont été mises en œuvre. Ces informations DOIVENT être disponibles pour les opérations à la fois de surveillance et de commande.

Discussion

L'intention de cette exigence est de donner la possibilité de faire sur le routeur par SNMP tout ce qui peut être fait sur une console, et vice versa. Une quantité de configuration minimale est nécessaire avant que SNMP puisse fonctionner (par exemple, le routeur doit avoir une adresse IP). Cette configuration initiale ne peut pas être réalisée par SNMP. Cependant, une fois que la configuration initiale est faite, les pleines capacités devraient être disponibles à travers la gestion du réseau.

Le fabricant DEVRAIT rendre disponibles les spécifications de toutes les variables de MIB spécifique de fabricant. Ces spécifications DOIVENT être conformes au SMI [MGT:1] et les descriptions DOIVENT être sous la forme spécifiée dans [MGT:4].

Discussion

La mise à disposition de l'utilisateur de la MIB spécifique de fabricant est nécessaire. Sans ces informations, l'usager ne sera pas capable de configurer son système de gestion de réseau pour pouvoir accéder aux paramètres spécifiques du fabricant. Ces paramètres seraient alors dépourvus d'utilité.

Le format de la spécification de la MIB est aussi spécifié. Les analyseurs qui lisent les spécifications de MIB et génèrent les tableaux nécessaires pour la station de gestion du réseau sont disponibles. Ces analyseurs ne comprennent généralement que le format de spécification de MIB standard.

8.5 Sauvegarde des changements

Les paramètres altérés par SNMP PEUVENT être sauvegardés dans des mémoires non volatiles.

Discussion

Les raisons pour lesquelles cette exigence est un PEUT sont :

- La nature physique exacte d'une mémoire non volatile n'est pas spécifiée dans le présent document. Et donc, les paramètres peuvent être sauvegardés sur NVRAM/EEPROM, sur un disque local ou un disque dur, ou dans un serveur de fichiers TFTP, ou un serveur BOOTP, etc. Supposons que ces informations soient dans un fichier qui est restitué par TFTP. Dans ce cas, un changement apporté à un paramètre de configuration sur le routeur aurait besoin d'être retransmis au serveur de fichier qui détient le fichier de configuration. Autrement, l'opération SNMP aurait besoin d'être dirigée sur le serveur de fichier, et ensuite le changement serait répercuté d'une façon ou d'une autre jusqu'au routeur. La réponse à ce problème ne paraît pas évidente. Cela fait aussi peser plus d'exigences sur l'hôte qui détient les informations de configuration que s'il y avait juste un serveur TFTP disponible, tellement plus qu'il n'est probablement pas sûr pour un fabricant de supposer qu'un acheteur potentiel aura un hôte convenable disponible.
- Le moment où il faut confier les paramètres à une mémoire non volatile est une question toujours débattue. Certains préfèrent envoyer immédiatement tous les changements. D'autres préfèrent ne confier les changements à une mémoire non volatile que sur une commande explicite.

9 Couche d'application – protocoles divers

Pour tous les protocoles d'application supplémentaires qu'un routeur met en œuvre, le routeur DOIT être conforme et DEVRAIT être inconditionnellement conforme aux exigences pertinentes de [INTRO:3].

9.1 BOOTP

9.1.1 Introduction

Le protocole Bootstrap (BOOTP) est un protocole fondé sur UDP/IP qui permet à un hôte qui s'amorce de se configurer lui-même de façon dynamique et sans supervision de l'utilisateur. BOOTP fournit le moyen de notifier à un hôte son adresse IP allouée, l'adresse IP d'un hôte serveur d'amorce, et le nom d'un fichier à charger en mémoire et à exécuter [APPL:1]. D'autres informations de configuration telles que la longueur du préfixe local ou le gabarit de sous-réseau, le décalage horaire local, les adresses des routeurs par défaut, et les adresses de divers serveurs Internet peuvent aussi être communiquées à un hôte utilisant BOOTP [APPL:2].

9.1.2 Agents relais de BOOTP

Dans de nombreux cas, les clients BOOTP et leurs serveurs BOOTP associés ne résident pas sur le même (sous)-réseau IP. Dans de tels cas, un agent tiers est nécessaire pour transférer les messages BOOTP entre clients et serveurs. Un tel agent était à l'origine appelé un agent de transmission BOOTP. Cependant, pour éviter des confusions avec la fonction de transmission IP d'un routeur, le nom d'agent relais de BOOTP a été adopté à la place.

Discussion

Un agent relais BOOTP effectue une tâche qui est distincte de la fonction de transmission IP normale d'un routeur. Alors qu'un routeur commute normalement les datagrammes IP entre les réseaux de façon plus ou moins transparente, un agent relais de BOOTP peut plus précisément être vu comme recevant des messages BOOTP comme destination finale puis comme générant de nouveaux messages BOOTP par la suite. On devrait résister à la tentation de voir une simple transmission directe de message BOOTP comme un paquet normal.

Cette fonctionnalité d'agent relais localisée de façon très opportune dans les routeurs qui interconnectent les clients et les serveurs (bien qu'elle puisse aussi être localisée dans un hôte qui est directement connecté au (sous)-réseau du client.

Un routeur PEUT fournir la capacité d'agent relais de BOOTP. S'il le fait, il DOIT se conformer aux spécifications de [APPL:3].

Le paragraphe 5.2.3 expose les circonstances dans lesquelles un paquet est délivré localement (au routeur). Tous les messages UDP livrés localement dont le numéro de port de destination UDP est BOOTPS (67) sont pris en considération pour un traitement spécial par l'agent relais logique BOOTP du routeur.

Les paragraphes 4.2.2.11 et 5.3.7 discutent des adresses IP de source invalides. Conformément à ces règles, un routeur ne doit pas transmettre un datagramme reçu dont l'adresse IP de source est 0.0.0.0. Cependant, les routeurs qui acceptent un agent relais de BOOTP DOIVENT accepter pour livraison locale à l'agent relais les messages BOOTREQUEST dont l'adresse IP de source est 0.0.0.0.

10 Fonctionnement et maintenance

Cette section subroge toutes les exigences de [INTRO:3] se rapportant aux "Extensions au module IP."

Les facilités de prise en charge des activités de fonctionnement et maintenance (O&M) forment une part essentielle de toute mise en œuvre de routeur. Bien que ces fonctions ne semblent pas se rapporter directement à l'interopérabilité, elles sont essentielles pour le gestionnaire de réseau qui doit rendre le routeur interopérable et doit chercher où est le problème lorsqu'il ne l'est pas. Cette section comporte aussi un exposé sur l'initialisation du routeur et sur les facilités pour assister les gestionnaires de réseau dans leurs tâches de sécurisation et de comptabilité de leurs réseaux.

10.1 Introduction

Les types d'activités suivants sont inclus dans l'O&M d'un routeur :

- Diagnostic des problèmes de matériel dans le processeur du routeur, dans ses interfaces réseau, ou dans ses réseaux connectés, ses modems, ou ses lignes de communication.
- Installation de nouveau matériel
- Installation de nouveau logiciel

- Redémarrage ou réamorçage du routeur après une défaillance
- Configuration (ou reconfiguration) du routeur.
- Détection et diagnostic des problèmes d'Internet tels que l'encombrement, acheminements en boucle, mauvaises adresses IP, trous noirs, avalanches de paquets, et hôtes douteux.
- Changement de la topologie de réseau, temporaire (par exemple, pour contourner un problème de ligne de communication) ou permanente.
- Surveillance de l'état et des performances des routeurs et des réseaux connectés.
- Collecte des statistiques de trafic à utiliser pour la planification (inter-)réseau.
- Coordination des activités ci-dessus avec les fabricants et spécialistes de télécommunications appropriés.

Les routeurs et leurs lignes de communication connectées fonctionnent souvent comme un système commandé par une organisation O&M centralisée. Cette organisation peut comporter un centre d'opérations (inter-)réseau, ou NOC, pour mener à bien ses fonctions O&M. Il est essentiel que les routeurs prennent en charge la commande à distance et la surveillance à partir d'un tel NOC à travers un chemin Internet, car les routeurs pourraient n'être pas connectés au même réseau que leur NOC. Comme une défaillance de réseau peut temporairement empêcher l'accès au réseau, de nombreux NOC insistent pour que les routeurs soient accessibles pour la gestion de réseau à travers des moyens de remplacement, souvent des modems numérotables attachés aux accès par console sur les routeurs.

Comme un paquet IP qui traverse un internet va souvent utiliser des routeurs sous le contrôle de plus d'un NOC, le diagnostic de problèmes Internet va souvent impliquer la coopération de personnels de plus d'un NOC. Dans certains cas, le même routeur peut avoir besoin d'être surveillé par plus d'un NOC, mais seulement si nécessaire, car une surveillance excessive pourrait avoir un impact sur les performances du routeur.

Les outils disponibles pour la surveillance au niveau d'un NOC peuvent couvrir une large gamme de sophistication. Les mises en œuvre courantes incluent l'affichage multifenêtre dynamique du système routeur tout entier. L'utilisation de techniques d'indicateur d'action pour le diagnostic automatique de problème est proposée à l'avenir.

Les facilités d'O&M pour les routeurs exposées ici sont seulement une partie du vaste et difficile problème de la gestion de l'Internet. Ces problèmes n'englobent pas seulement plusieurs organisations de gestion, mais aussi plusieurs couches de protocole. Par exemple, au stade actuel de l'évolution de l'architecture de l'Internet, il y a un fort couplage entre les mises en œuvre d'hôte TCP et l'encombrement éventuel de niveau IP du système routeur [OPER:1]. Donc, le diagnostic des problèmes d'encombrement va parfois exiger la surveillance des statistiques TCP dans les hôtes. Il y a actuellement un certain nombre d'efforts de R&D en cours dans le domaine de la gestion de l'Internet et plus spécifiquement de l'O&M de routeur. Ces efforts de R&D ont déjà produit des normes pour l'O&M de routeur. C'est aussi un domaine dans lequel la créativité des fabricants peut apporter une contribution significative.

10.2 Initialisation du routeur

10.2.1 Configuration minimale du routeur

Il existe un ensemble minimum de conditions qui doivent être satisfaites avant qu'un routeur puisse transmettre des paquets. Un routeur NE DOIT PAS activer la transmission sur une interface physique tant que :

(1) le routeur ne connaît pas l'adresse IP et le gabarit de sous-réseau ou la longueur du préfixe de réseau associé d'au moins une interface logique associée à cette interface physique, ou

(2) le routeur ne sait pas si cette interface est une interface non numérotée et ne connaît pas son ID de routeur.

Ces paramètres DOIVENT être explicitement configurés :

- Un routeur NE DOIT PAS utiliser de valeurs par défaut configurées par fabrication pour ses adresses IP, ses longueurs de préfixe ou ses ID de routeur, et
- Un routeur NE DOIT PAS supposer qu'une interface non configurée est une interface sans numéro.

Discussion

Il y a eu des instances dans lesquelles les routeurs ont été installés avec des adresses par défaut fixées par le fabricant pour les interfaces. Dans quelques cas, il en est résulté que les routeurs ont publié ces adresses par défaut dans des réseaux actifs.

10.2.2 Initialisation d'adresse et de préfixe

Un routeur DOIT permettre que ses adresses IP et leurs gabarits d'adresse ou longueurs de préfixes soient configurées de façon statique et sauvegardées dans des mémoires non volatiles.

Un routeur PEUT obtenir ses adresses IP et leurs gabarits d'adresse correspondants de façon dynamique comme effet secondaire du processus d'initialisation du système (voir au paragraphe 10.2.3).

Si la méthode dynamique est fournie, le choix de la méthode à utiliser dans un routeur particulier DOIT être configurable.

Comme décrit au paragraphe 4.2.2.11, il n'est pas permis aux adresses IP d'avoir la valeur 0 ou -1 dans les champs

<Host-number> ou <Network-prefix>. Donc, un routeur NE DEVRAIT PAS permettre qu'une adresse IP ou gabarit d'adresse soit réglé à une valeur qui ferait qu'un des champs ci-dessus ait la valeur zéro ou -1.

Discussion
Il est possible qu'en utilisant des gabarits d'adresse arbitraires on crée des situations dans lesquelles l'acheminement est ambigu (c'est-à-dire que deux chemins avec des gabarits de sous-réseau différents mais également spécifiques correspondent à une adresse de destination particulière). C'est un des plus forts arguments en faveur de l'utilisation de préfixes de réseau, et la raison pour laquelle l'utilisation de gabarits de sous-réseau discontigus n'est pas permise.

Un routeur DEVRAIT faire les vérifications suivantes sur tout gabarit d'adresse qu'il installe :
- Le gabarit n'est pas tout en uns ou tout en zéros (la longueur du préfixe n'est ni zéro ni 32).
- Les bits qui correspondent à la partie préfixe de réseau de l'adresse sont tous mis à 1.
- Les bits qui correspondent au préfixe de réseau sont contigus.

Discussion
Les gabarits associés aux chemins sont aussi parfois appelés gabarits de sous-réseau, cette vérification ne devrait pas leur être appliquée.

10.2.3 Amorçage de réseau avec BOOTP et TFTP

Il y a eu de nombreuses discussions sur la façon dont les routeurs peuvent et devraient s'amorcer à partir du réseau. Ces discussions ont tourné autour de BOOTP et de TFTP. Actuellement, il y a les routeurs qui s'amorcent avec TFTP à partir du réseau. Il n'y a pas de raison que BOOTP ne puisse pas être utilisé pour localiser le serveur à partir duquel l'image d'amorce devrait être chargée.

BOOTP est un protocole utilisé par les systèmes d'amorçage d'extrémité, et exige un peu de gymnastique pour accommoder son utilisation avec les routeurs. Si un routeur utilise BOOTP pour localiser l'hôte d'amorce actuel, il devrait envoyer une demande BOOTP avec l'adresse de son matériel pour sa première interface, ou, si il a été précédemment configuré autrement, avec une autre adresse de matériel d'interface, ou autre numéro à mettre dans le champ d'adresse de matériel du paquet BOOTP. Ceci est destiné à permettre aux routeurs sans adresse de matériel (comme les routeurs avec seulement une ligne synchrone) d'utiliser BOOTP pour la découverte de chargement d'amorce. TFTP peut alors être utilisé pour restituer l'image trouvée dans la réponse BOOTP. S'il n'y a pas d'interface configurée ou de numéro à utiliser, un routeur PEUT circuler entre les adresses de matériel d'interface qu'il a, jusqu'à ce qu'une correspondance soit trouvée par le serveur BOOTP.

Un routeur DEVRAIT METTRE EN ŒUVRE la capacité à mémoriser les paramètres appris par l'intermédiaire de BOOTP dans une mémoire locale non volatile. Un routeur PEUT mettre en œuvre la capacité à mémoriser une image système chargée sur le réseau dans une mémoire locale stable.

Un routeur PEUT avoir une facilité permettant à un utilisateur distant de demander que le routeur obtienne une nouvelle image d'amorce. Une différentiation devrait être faite entre l'obtention d'une nouvelle image d'amorce à partir d'une des trois localisations : celle incluse dans la demande, à partir du serveur de la dernière image d'amorce, et utiliser BOOTP pour localiser un serveur.

10.3 Opération et maintenance

10.3.0 Introduction

Il y a toute une gamme de modèles possibles pour effectuer les fonctions O&M sur un routeur. À un extrême se trouve le modèle local seul, dans lequel les fonctions O&M ne peuvent être exécutées qu'en local (par exemple, à partir d'un terminal branché sur la machine routeur). À l'autre extrême, le modèle entièrement à distance ne permet qu'à un minimum absolu de fonctions d'être effectuées localement (par exemple, forcer une amorce) la plus grande partie de l'O&M étant faite à distance à partir du NOC. Il y a des modèles intermédiaires, comme celui dans lequel le personnel du NOC peut enregistrer dans le routeur comme si c'était un hôte, en utilisant le protocole Telnet, pour effectuer des fonctions qui peuvent aussi être invoquées en local. Le modèle local seul peut être adéquat dans quelques installations de routeurs, mais le fonctionnement à distance à partir d'un NOC est normalement exigé, et donc les dispositions d'O&M à distance sont exigées pour la plupart des routeurs.

Les fonctions O&M distantes peuvent être exercées à travers un agent de contrôle (programme). Dans l'approche directe, le routeur prendrait directement en charge les fonctions d'O&M distantes à partir du NOC en utilisant les protocoles Internet standard (par exemple, SNMP, UDP ou TCP) ; dans l'approche indirecte, l'agent de contrôle prendrait en charge ces protocoles et contrôleraient le routeur lui-même en utilisant des protocoles privés. L'approche

directe est préférée, bien que les deux approches soient acceptables. L'utilisation de matériel et/ou logiciel d'hôte spécialisé exigeant un investissement supplémentaire significatif est déconseillée ; néanmoins, certains fabricants peuvent choisir de fournir l'agent de contrôle comme partie intégrante du réseau dans lequel sont les routeurs. Si c'est le cas, il est exigé qu'il y ait un moyen disponible pour faire fonctionner l'agent de contrôle à partir d'un site distant utilisant les protocoles et chemins de l'Internet et avec des fonctionnalités équivalentes par rapport à un terminal d'agent local.

Il est souhaitable qu'un agent de contrôle et tous les autres outils logiciels de NOC que fournit un fabricant fonctionnent comme des programmes d'utilisateur dans un système d'exploitation standard. L'utilisation des protocoles standard Internet UDP et TCP pour communiquer avec les routeurs devrait le faciliter.

La surveillance à distance du routeur et (en particulier) la commande de routeur à distance présente d'importants problèmes de contrôle d'accès qui doivent être examinés. Il faut veiller aussi à s'assurer du contrôle de l'utilisation des ressources du routeur pour ces fonctions. Il n'est par exemple pas souhaitable de laisser la surveillance du routeur prendre plus qu'une fraction limitée du temps CPU du routeur. D'un autre côté, les fonctions O&M doivent recevoir la priorité, de sorte qu'elles puissent être exercées lorsque le routeur est encombré, car souvent c'est alors que l'O&M est la plus nécessaire.

10.3.1 Accès hors bande

Les routeurs DOIVENT prendre en charge l'accès hors bande. L'accès hors bande DEVRAIT fournir la même fonctionnalité que l'accès dans la bande. Cet accès DEVRAIT mettre en œuvre des contrôles d'accès, pour empêcher l'accès non autorisé.

Discussion

Cet accès hors bande va permettre au NOC un moyen d'isoler l'accès des routeurs durant des périodes où l'accès au réseau n'est pas disponible. L'accès hors bande est un important outil de gestion pour l'administrateur de réseau. Il permet l'accès aux équipements indépendamment des connexions réseau. Cet accès peut être réalisé de nombreuses façons. Quelle que soit celle utilisée, il est important que l'accès soit indépendant de la connexion réseau. Un exemple d'accès hors bande serait celui d'un accès en série connecté à un modem qui fournit un accès par numérotage au routeur.

Il est important que l'accès hors bande fournisse la même fonctionnalité que dans l'accès dans la bande. L'accès dans la bande, ou l'accès aux équipements à travers la connexion réseau existante, est limitant, parce que la plupart du temps, les administrateurs ont besoin d'atteindre l'équipement pour comprendre pourquoi il est inaccessible. L'accès dans la bande est aussi très important pour configurer un routeur, et pour dépanner des problèmes plus subtils.

10.3.2 Fonctions O&M de routeur

10.3.2.1 Maintenance – Diagnostic de matériels

Chaque routeur DEVRAIT fonctionner comme un appareil autonome pour les besoins de la maintenance matérielle locale. Des moyens DEVRAIENT être disponibles pour faire tourner des programmes de diagnostic sur le site du routeur en utilisant uniquement des outils du site. Un routeur DEVRAIT être capable de faire des diagnostics en cas de faute. Voir au paragraphe 10.3.3 les diagnostics suggérés de matériel et logiciel.

10.3.2.2 Contrôle – Décharge et réamorçage

Un routeur DOIT inclure des mécanismes à la fois dans la bande et hors bande pour permettre au gestionnaire de réseau de recharger, arrêter, et redémarrer le routeur. Un routeur DEVRAIT aussi contenir un mécanisme (comme un temporisateur de garde) qui va réamorcer automatiquement le routeur si il s'arrête suite à une faute logicielle ou matérielle.

Un routeur DEVRAIT METTRE EN ŒUVRE un mécanisme pour décharger le contenu de la mémoire d'un routeur (et/ou d'autres états utiles pour nettoyer des défauts de fabrication après une défaillance), et le sauvegarder sur un appareil local de mémorisation stable sur le routeur ou le sauvegarder sur un autre hôte via un mécanisme de déchargement en ligne tel que TFTP (voir [OPER:2], [INTRO:3]).

10.3.2.3 Contrôle – Configurer le routeur

Chaque routeur a des paramètres de configuration qui peuvent devoir être réglés. Il DEVRAIT être possible de mettre à jour les paramètres sans réamorcer le routeur ; au pire, un redémarrage PEUT être nécessaire. Il peut y avoir des cas où il n'est pas possible de changer les paramètres sans réamorcer le routeur (par exemple, changer l'adresse IP d'une

interface). Dans ces cas, il faut veiller à minimiser l'interruption entre le routeur et le réseau environnant.

Il DEVRAIT y avoir un moyen de configurer le routeur sur le réseau, soit manuel, soit automatique. Un routeur DEVRAIT être capable de télécharger ses paramètres à partir d'un hôte ou d'un autre routeur. Un moyen DEVRAIT être fourni, comme un programme d'application ou une fonction de routeur, pour faire la conversion entre le format du paramètre et un format lisible par l'homme. Un routeur DEVRAIT avoir une ressource de mémoire stable pour sa configuration. Un routeur NE DEVRAIT PAS croire des protocoles tels que RARP, Réponse de gabarit d'adresse ICMP, et PEUT ne pas croire BOOTP.

Discussion

Il est nécessaire de noter ici qu'à l'avenir, RARP, Réponse de gabarit d'adresse ICMP, BOOTP et d'autres mécanismes pourront être nécessaires pour permettre à un routeur de s'auto-configurer. Bien que les routeurs puissent à l'avenir être capables de configuration automatique, l'intention ici est de déconseiller cette pratique dans un environnement de production jusqu'à ce que l'auto configuration ait été testée plus en détail. L'intention N'EST PAS de déconseiller l'auto configuration dans l'absolu. Si un routeur est prévu pour une configuration automatique, il peut être avisé de lui permettre de croire ces programmes lors du démarrage puis de les ignorer une fois qu'il a sa configuration.

10.3.2.4 Amorçage réseau de logiciel système

Un routeur DEVRAIT conserver son image système dans une mémoire locale non volatile comme un PROM, NVRAM, ou un disque. Il PEUT aussi être capable de charger son logiciel système sur le réseau à partir d'un hôte ou autre routeur.

Un routeur qui conserve son image système dans une mémoire locale non volatile PEUT être configurable pour amorcer son image système sur le réseau. Un routeur qui offre cette option DEVRAIT être configurable pour amorcer l'image système dans sa mémoire locale non volatile s'il n'est pas capable d'amorcer son image système sur le réseau.

Discussion

Il est important que le routeur soit capable de démarrer et fonctionner de façon autonome. NVRAM peut être une solution particulière pour les routeurs utilisés dans de grands réseaux, car changer les PROM peut prendre du temps à un gestionnaire de réseau responsable de nombreux routeurs peut-être géographiquement dispersés. Il est important d'être capable d'amorcer à partir du réseau l'image système parce qu'il devrait être facile à un routeur de réparer une erreur de programme ou installer un nouveau dispositif plus rapidement que d'installer les PROM. Aussi si le routeur a NVRAM à la place des PROM, il va amorcer l'image à partir du réseau et la mettra ensuite en NVRAM.

Les routeurs DEVRAIENT effectuer des vérifications de cohérence de base sut toute image chargée, pour détecter et peut-être prévenir des images incorrectes.

Un routeur PEUT aussi être capable de distinguer entre différentes configurations sur la base du logiciel qui fonctionne. Si les commandes de configuration changent d'une version de logiciel à l'autre, il serait utile que le routeur puisse utiliser la configuration qui est compatible avec le logiciel.

10.3.2.5 Détection et réplique aux mauvaises configurations

Il DOIT y avoir un mécanisme de détection et de réplique aux mauvaises configurations. Si une commande est exécutée de façon incorrecte, le routeur DEVRAIT donner un message d'erreur. Le routeur NE DEVRAIT PAS accepter une commande mal tournée comme si elle était correcte.

Discussion

Il y a des cas où il n'est pas possible de détecter les erreurs : la commande est correctement formée, mais incorrecte par rapport au réseau. Ceci peut être détecté par le routeur, mais peut n'être pas possible.

Une autre forme de mauvaise configuration est celle du réseau auquel le routeur est rattaché. Un routeur PEUT détecter les mauvaises configurations dans le réseau. Le routeur PEUT enregistrer ces découvertes dans un fichier, sur le routeur ou sur un hôte, de sorte que le gestionnaire de réseau puisse voir qu'il pourrait y avoir des problèmes sur le réseau.

Discussion

Des exemples de telles mauvaises configurations pourraient être un routeur avec la même adresse que celui en question ou un routeur avec le mauvais gabarit d'adresse. Si un routeur détecte de tels problèmes, il n'est probablement pas le mieux placé pour essayer de régler la situation. Cela pourrait causer plus de mal que de bien.

10.3.2.6 Minimiser l'interruption

Changer la configuration d'un routeur DEVRAIT avoir un effet minimal sur le réseau. Les tableaux d'acheminement NE DEVRAIENT PAS bouger sans nécessité lorsqu'un simple changement est fait au routeur. Si un routeur fait tourner

plusieurs protocoles d'acheminement, arrêter un protocole d'acheminement NE DEVRAIT PAS interrompre les autres protocoles d'acheminement, excepté dans le cas où un réseau est en acquisition par plus d'un protocole d'acheminement.

Discussion

C'est l'objectif d'un gestionnaire de réseau de faire fonctionner un réseau de telle sorte que les usagers du réseau obtiennent la meilleure connectivité possible. Recharger un routeur pour de simples changements de configuration peut causer des interruptions dans l'acheminement et finalement causer des interruptions au réseau et à ses usagers. Si les tableaux d'acheminement sont modifiés sans nécessité, par exemple, le chemin par défaut sera perdu aussi bien que les chemins spécifiques vers les sites au sein du réseau. Cette sorte d'interruption causera des pertes de temps significatives aux usagers. L'objectif de ce paragraphe est de souligner que chaque fois que possible, ces interruptions devraient être évitées.

10.3.2.7 Contrôle – Résolution des problèmes

(1) Un routeur DOIT fournir l'accès réseau dans la bande, mais (excepté quand c'est exigé par le paragraphe 8.2) pour des considérations de sécurité, cet accès DEVRAIT être désactivé par défaut. Les fabricants DOIVENT documenter l'état par défaut dans tout accès dans la bande. Cet accès DEVRAIT mettre en œuvre des contrôles d'accès, pour empêcher les accès non autorisés.

Discussion

L'accès dans la bande se réfère principalement à l'accès par des protocoles réseau normaux qui peuvent affecter ou non l'état de fonctionnement permanent du routeur. Cela inclut, sans s'y limiter, l'accès par console Telnet/RLOGIN et le fonctionnement de SNMP.

Ceci était un point de dispute entre les partisans de la "sortie de boîte opérationnelle" et ceux de la "sortie de boîte sécurisée". Tout accès automatique au routeur peut introduire de l'insécurité, mais il peut être plus important pour le consommateur d'avoir un routeur qui soit accessible sur le réseau aussitôt qu'il est banché. Au moins un fabricant fournit des routeurs sans accès par console externe qui dépendent de la capacité à accéder au routeur à travers le réseau pour achever sa configuration.

Il appartient au fabricant de dire si l'accès dans la bande est activé par défaut ; mais il est aussi de la responsabilité du fabricant d'avertir le consommateur d'une possible insécurité.

(2) Un routeur DOIT fournir la capacité à initier un écho ICMP.
Les options suivantes DEVRAIENT être mises en œuvre :
Choix des gabarits de données
Choix de la taille de paquet
Les options Record route et les options supplémentaires suivantes PEUVENT être mises en œuvre :
Route de source lâche
Route de source stricte
Horodatage

(3) Un routeur DEVRAIT fournir la capacité à initier un traceroute. Si traceroute est fourni, le traceroute de tiers DEVRAIT être mis en œuvre.

Chacune des trois facilités ci-dessus (si elles sont mises en œuvre) DEVRAIT avoir des restrictions d'accès pour empêcher leur abus par des personnes non autorisées.

10.4 Considérations sur la sécurité

10.4.1 Audit et suites d'audit

L'audit et la facturation sont le lot de l'opérateur du réseau, mais sont deux des dispositifs les plus réclamés par ceux qui sont chargés de la sécurité du réseau et ceux qui sont responsables du payement des factures. Dans le contexte de la sécurité, l'audit est souhaitable si il vous aide à conserver votre réseau en état de fonctionnement et protège vos ressources contre une utilisation abusive, sans vous coûter plus que ce que ne valent ces ressources.

(1) Changements de configuration
Un routeur DEVRAIT avoir une méthode pour l'audit d'un changement de configuration d'un routeur, même si c'est quelque chose d'aussi simple que l'enregistrement des initiales de l'opérateur et l'heure du changement.

Discussion

L'enregistrement du changement de configuration (qui a fait un changement de configuration, qu'est ce qui a été changé, et quand ?) est très utile, en particulier lorsque le trafic est soudain acheminé par l'Alaska pour traverser la ville. Il en est de même de la capacité à inverser une configuration antérieure.

(2) Comptabilité des paquets
Les fabricants devraient fortement envisager la fourniture d'un système de suivi des niveaux de trafic entre les paires d'hôtes ou de réseaux. Un mécanisme de limitation de la collecte de ces informations à des paires d'hôtes ou réseaux spécifiques est aussi vivement conseillé.

Discussion
Une matrice de trafic d'hôte telle que décrite ci-dessus peut donner à l'opérateur de réseau une vue des tendances du trafic qui n'apparaît pas dans les autres statistiques. Elle peut aussi identifier les hôtes ou réseaux qui mettent à l'épreuve la structure des réseaux rattachés – par exemple, un seul hôte externe qui essaye d'envoyer des paquets à toutes les adresses IP dans la gamme des adresses réseau pour un réseau connecté.

(3) Audit de sécurité
Les routeurs DOIVENT avoir une méthode d'audit des défaillances ou des violations en rapport avec la sécurité incluant :
Les échecs d'autorisation : mauvais mots de passe, communautés SNMP invalides, jetons d'autorisation invalides,
Violations des contrôles de politique : routes de source interdites, destinations filtrées, et
Approbations d'autorisation : bons mots de passe – accès Telnet dans la bande, accès à la console.

Les routeurs DOIVENT avoir une méthode pour limiter ou désactiver un tel audit, mais l'audit DEVRAIT être activé par défaut. Des méthodes possibles pour l'audit incluent de faire la liste des violations à la console si il y en a une, de les enregistrer ou de les compter en interne, ou de les enregistrer sur un serveur de sécurité distant à travers le mécanisme trap de SNMP ou le mécanisme d'enregistrement Unix selon ce qui est approprié. Un routeur DOIT mettre en œuvre au moins un de ces mécanismes d'analyse – il PEUT en mettre en œuvre plus d'un.

10.4.2 Contrôle de configuration

Un fabricant a la responsabilité d'utiliser de bonnes pratiques de contrôle de configuration dans la création de l'ensemble de logiciels chargés dans les routeurs. En particulier, si un fabricant rend disponible sur l'Internet les mises à jour et chargements, il devrait aussi donner le moyen au consommateur de confirmer que le chargement est valide, peut-être par la vérification d'une somme de contrôle sur le chargement.

Discussion
De nombreux fabricants fournissent des avis de mise à jour de leurs produits logiciels sur l'Internet. C'est une bonne tendance qui devrait être encouragée, mais cela crée un point de vulnérabilité dans le processus de contrôle de configuration.

Si un fabricant donne au consommateur la capacité de changer à distance les paramètres de configuration d'un routeur, par exemple dans une session Telnet, cette capacité DEVRAIT être configurable et DEVRAIT être désactivée par défaut. Le routeur DEVRAIT exiger une authentification valide avant de permettre la reconfiguration à distance. Cette procédure d'authentification NE DEVRAIT PAS transmette de secret d'authentification sur le réseau. Par exemple, si Telnet est mis en œuvre, le fabricant DEVRAIT METTRE EN ŒUVRE Kerberos, S-Key, ou une procédure d'authentification similaire.

Discussion
Permettre à votre opérateur de réseau correctement identifié de tripoter vos routeurs est nécessaire ; permettre à n'importe qui d'en faire autant serait téméraire.

Un routeur NE DOIT PAS avoir d'accès de secours et de mots de passe maîtres non documentés. Un fabricant DOIT s'assurer que tout accès de cette sorte ajouté dans le but d'éliminer des défauts de fabrication ou de développement du produit est supprimé avant que le produit ne soit distribué aux consommateurs.

Discussion
Un fabricant a la responsabilité envers ses consommateurs de s'assurer qu'ils sont conscients des faiblesses présentes intentionnellement dans son code – par exemple, l'accès dans la bande. Les portes dérobées, les contournements et les mots de passe maîtres intentionnels ou involontaires peuvent transformer un routeur relativement sûr en un problème majeur pour un réseau opérationnel. Les bénéfices opérationnels supposés ne sont pas à la hauteur des problèmes potentiels.

11 Références

Les développeurs devraient être avertis que les normes de protocole de l'Internet sont occasionnellement mises à jour. Les références ci-après sont à jour au moment de la rédaction, mais un développeur vigilant vérifiera toujours sur une version récente de l'index des RFC pour s'assurer qu'une RFC n'a pas été mise à jour ou subrogée par une autre, plus récente. La référence [INTRO:6] explique diverses façons d'obtenir l'index des RFC en cours.

[APPL:1] B. Croft et J. Gilmore, "Protocole BOOTSTRAP (BOOTP)", RFC0951, septembre 1985.

[APPL:2] E. Gavron, "Problème de sécurité et proposition de correction avec le logiciel courant du DNS", RFC1535, octobre 1993. *(Information)*

[APPL:3] W. Wimer, "Précisions et extensions au protocole Boostrap", décembre 1993, RFC1542, octobre 1993. *(DS.)*

[ARCH:1] DDN Protocol Handbook, NIC-50004, NIC-50005, NIC-50006 (trois volumes), DDN Network Information Center, SRI International, Menlo Park, California, USA, décembre 1985.

[ARCH:2] V. Cerf et R. Kahn, "A Protocol for Packet Network Intercommunication", IEEE Transactions on Communication, mai 1974. Aussi inclus dans [ARCH:1].

[ARCH:3] J. Postel, C. Sunshine, et D. Cohen, "The ARPA Internet Protocol", Computer Networks, volume 5, n° 4, juillet 1981. Aussi inclus dans [ARCH:1].

[ARCH:4] B. Leiner, J. Postel, R. Cole, et D. Mills, "The DARPA Internet Protocol Suite", Proceedings of INFOCOM '85, IEEE, Washington, DC, mars 1985. Aussi dans IEEE Communications Magazine, mars 1985. Aussi disponible auprès de Information Sciences Institute, University of Southern California comme rapport technique ISI-RS-85-153.

[ARCH:5] D. Comer, "Internetworking With TCP/IP Volume 1: Principles, Protocols, et Architecture", Prentice Hall, Englewood Cliffs, NJ, 1991.

[ARCH:6] W. Stallings, "Handbook of Computer-Communications Standards Volume 3: The TCP/IP Protocol Suite", Macmillan, New York, NY, 1990.

[ARCH:7] J. Postel, éd., "Normes officielles de protocole de l'Internet", RFC1780, mars 1995. *(Obsolète, voir* RFC1800) *(Historique)*

[ARCH:8] Organisation Mondiale de Normalisation, Norme ISO 7489, "Systèmes de traitement de l'information - Interconnexion des systèmes ouverts - Modèle de référence de base", 1984.

[ARCH:9] R. Braden, J. Postel, Y. Rekhter, "Extensions à l'architecture de l'Internet pour les supports partagés", 20/05/1994

[FORW:1] C. Topolcic, éditeur, "Protocole expérimental de flux Internet, version 2 (ST-II)", RFC1190, octobre 1990. *(obsolète, voir la RFC 1819)*

[FORW:2] A. Mankin et K. Ramakrishnan, "Enquête de contrôle sur l'encombrement des routeurs", RFC1254, août 1991. *(Info)*

[FORW:3] J. Nagle, "On Paquet Switches with Infinite Storage", IEEE Transactions on Communications, volume COM-35, number 4, April 1987.

[FORW:4] R. Jain, K. Ramakrishnan, et D. Chiu, "Congestion Avoidance in Computer Networks With a Connectionless Network Layer", Technical Report DEC-TR-506, Digital Equipment Corporation.

[FORW:5] V. Jacobson, "Congestion Avoidance et Control", Proceedings of SIGCOMM '88, Association for Computing Machinery, August 1988.

[FORW:6] W. Barns, "Precedence et Priority Access Implementation for Department of Defense Data Networks", Technical Report MTR- 91W00029, The Mitre Corporation, McLean, Virginia, USA, July 1991.

[FORW:7] Fang, Chen, Hutchins, "Simulation Results of TCP Performance over ATM with et without Flow Control", presentation to the ATM Forum, November 15, 1993.

[FORW:8] V. Paxson, S. Floyd "Wide Area Traffic: the Failure of Poisson Modeling", version abrégée dans SIGCOMM '94.

[FORW:9] Leland, Taqqu, Willinger et Wilson, "On the Self-Similar Nature of Ethernet Traffic", Proceedings of SIGCOMM '93, septembre 1993.

[FORW:10] S. Keshav "A Control Theoretic Approach to Flow Control", SIGCOMM 91, pages 3-16

[FORW:11] K.K. Ramakrishnan et R. Jain, "A Binary Feedback Scheme for Congestion Avoidance in Computer Networks", ACM Transactions of Computer Systems, volume 8, n° 2, 1980.

[FORW:12] H. Kanakia, P. Mishara, et A. Reibman]. "An adaptive encombrement control scheme for real-time paquet video transport", dans Proceedings of ACM SIGCOMM 1994, pages 20-31, San Francisco, California, septembre 1993.

[FORW:13] A. Demers, S. Keshav, S. Shenker, "Analysis et Simulation of a Fair Queuing Algorithm", 1993, pp 1-12

[FORW:14] Clark, D., Shenker, S., et L. Zhang, "Supporting Real-Time Applications in an Integrated Services Packet Network: Architecture et Mechanism", 92 pages 14-26

[INTER:1] J. Postel, éd., "Protocole Internet - Spécification du protocole du programme Internet", RFC0791, STD 5,

septembre 1981.

[INTER:2] J. Mogul et J. Postel, "Procédure standard de sous-réseautage Internet", RFC0950, (STD 5) août 1985.

[INTER:3] J. Mogul, "Diffusion des datagrammes Internet en présence de sous-réseaux", RFC0922, oct. 1984.

[INTER:4] S. Deering, "Extensions d'hôte pour diffusion groupée sur IP", RFC1112, STD 5, août 1989.

[INTER:5] S. Kent, "Options de sécurité du Ministère US de la défense pour le protocole Internet", RFC1108, novembre 1991. *(Historique)*

[INTER:6] R. Braden, D. Borman et C. Partridge, "Calcul de la somme de contrôle Internet", RFC1071, septembre 1988.

[INTER:7] T. Mallory et A. Kullberg, "Mise à jour incrémentaire de la somme de contrôle Internet", RFC1141, janvier 1990.

[INTER:8] J. Postel, "Protocole du message de contrôle Internet – Spécification du protocole du programme Internet DARPA", RFC0792, STD 5, septembre 1981.

[INTER:9] A. Mankin, G. Hollingsworth, G. Reichlen, K. Thompson, R. Wilder, et R. Zahavi, "Evaluation of Internet Performance - FY89", Technical Report MTR-89W00216, MITRE Corporation, février 1990.

[INTER:10] G. Finn, A "Connectionless Congestion Control Algorithm", Computer Communications Review, volume 19, number 5, Association for Computing Machinery, octobre 1989.

[INTER:11] W. Prue et J. Postel, "Ce qu'un hôte peut faire avec l'extinction de source : retard introduit par l'extinction de source (SQUID)", RFC1016, juillet 1987.

[INTER:12] A. McKenzie, "Quelques commentaires sur SQUID", RFC1018, août 1987.

[INTER:13] S. Deering, éditeur, "Messages ICMP de découverte de routeur", RFC1256, septembre 1991.

[INTER:14] J. Mogul et S. Deering, "Découverte de la MTU de chemin", RFC1191, novembre 1990.

[INTER:15] V. Fuller, T. Li, J. Yu et K. Varadhan, "Acheminement inter domaine sans classe (CIDR : stratégie d'allocation et d'agrégation d'adresses", RFC1519, septembre 1993. *(D.S., rendue obsolète par la RFC4632)*

[INTER:16] M, St. Johns, "Projet révisé d'options de sécurité IP", RFC 1038, IETF, janvier 1988.

[INTER:17] W. Prue et J. Postel, "Algorithme de mise en file d'attente pour fournir le type de service pour les liaisons IP", RFC1046, février 1988.

[INTER:18] J. Postel, "Transpositions d'adresse", RFC0796, septembre 1981.

[INTRO:1] R. Braden et J. Postel, "Exigences pour les routeurs de l'Internet", RFC1009, juin 1987. *(obsolète voir RFC 1812) (Historique)*

[INTRO:2] R. Braden, "Exigences pour les hôtes Internet – couches de communication", RFC1122, STD 3, octobre 1989.

[INTRO:3] R. Braden, éditeur, "Exigences pour les hôtes Internet – Application et prise en charge", RFC1123, STD 3, octobre 1989.

[INTRO:4] D. Clark, "Modularité et efficacité dans une mise en œuvre de protocole", RFC0817, juillet 1982.

[INTRO:5] Clark, D., "The Structuring of Systems Using Upcalls", Proceedings of 10th ACM SOSP, décembre 1985.

[INTRO:6] O. Jacobsen et J. Postel, "Comment se procurer les documents de protocole (les RFC)", RFC0980, mars 1986.

[INTRO:7] J. Reynolds et J. Postel, "Numéros alloués", RFC1700, STD 2, octobre 1994. *(Historique)* Le document est maintenant mis en ligne sur le site www.iana.org

[INTRO:8] DoD Trusted Computer System Evaluation Criteria, DoD publication 5200.28-STD, U.S. Department of Defense, décembre 1985.

[INTRO:9] G. Malkin et La Quey Parker, "Glossaire des utilisateurs de l'Internet", RFC1392, janvier 1993. *(Remplacé par RFC1983)*

[LINK:1] S. Lekkler et M. Karels, "Encapsulations d'en-queues", RFC0893, avril 1984.

[LINK:2] W. Simpson, éd., "Protocole point à point (PPP)", RFC1661, STD 51, juillet 1994. *(MàJ par RFC 2153)*

[LINK:3] G. McGregor, "Protocole de contrôle de protocole Internet point à point (IPCP)", RFC1332, mai 1992. *(MàJ par RFC3241)*

[LINK:4] B. Lloyd et W. Simpson, "Protocoles d'authentification PPP", RFC1334, octobre 1992. *(Obs., voir 1994)*

[LINK:5] W. Simpson, "Surveillance de qualité de liaison PPP", RFC 1333, mai 1992.

[MGT:1] M. Rose et K. McCloghrie, "Structure et identification des informations de gestion pour les internets fondés sur TCP/IP", RFC1155, STD 16, mai 1990.

[MGT:2] K. McCloghrie et M. Rose, "Base de données d'informations de gestion pour la gestion de réseau des internets fondés sur TCP/IP : MIB-II", RFC1213, STD 17, mars 1991.

[MGT:3] J. Case, M. Fedor, M. Schoffstall et J. Davin, "Protocole simple de gestion de réseau", RFC1157, STD 15, mai 1990. *(Historique)*

[MGT:4] M. Rose et K. McCloghrie, "Définitions concises de MIB", RFC1212, STD 16, février 1991.

[MGT:5] L. Steinberg, "Techniques de gestion des alertes générées de façon asynchrone", RFC1224, mai 1991. *(Exp.)*

[MGT:6] F. Kasteholzn, "Définitions des objets gérés pour les types d'interface de style Ethernet", RFC1398, janvier 1993. *(D.S., Obsolète, voir 1623)*

[MGT:7] K. McCloghrie et R. Fox, "MIB de bus à jetons IEEE 802.4", RFC1230, mai 1991. *(Historique)*

[MGT:8] K. McCloghrie, R. Fox et E. Decker, "MIB d'anneau à jetons IEEE 802.5", RFC1231, mai 1991. *(Obsolète, voir les RFC 1743, 1748)*

[MGT:9] J. Case et A. Rijsinghani, "Base de données d'informations de gestion de FDDI", RFC1512, septembre 1993. *(Hist.)*

[MGT:10] B. Stewart, éditeur "Définitions des objets gérés pour appareils de type RS-232", RFC1317, avril 1992.

[MGT:11] F. Kastenholz, "Définitions des objets gérés pour le protocole de contrôle de liaison PPP", RFC1471, juin 1993. *(P.S.)*

[MGT:12] F. Kastenholz, "Définitions des objets gérés pour les protocoles de sécurité PPP", RFC1472, juin 1993.

[MGT:13] F. Kastenholz, "Définitions des objets gérés pour le protocole de contrôle de réseau IP PPP", RFC1473, juin 1993. *(P.S.)*

[MGT:14] F. Baker et R. Coltun, "Base de données d'informations de gestion OSPF version 2", RFC1253, août 1991. *(remplacée par la RFC 1850)*

[MGT:15] S. Willis et J. Burruss, "Définitions des objets gérés pour le protocole BGP version 3", RFC1269, octobre 1991. *(Obsolète, voir 4273)*

[MGT:16] F. Baker et J. Watt, éditeurs, "Définitions des objets gérés pour les types d'interface DS1 et E1", RFC1406, janvier 1993. *(P.S., remplacée par la RFC 2495)*

[MGT:17] T. Cox et K. Tesink, "Définitions des objets gérés pour le type d'interface DS3/E3", RFC1407, janvier 1993. *(Obsolète, voir RFC2496)*

[MGT:18] K. McCloghrie, éd,, "Extensions à la MIB d'interface générique", RFC1229, mai 1991. *(Obs. RFC1573)*

[MGT:19] T. Cox et K. Tesnik, "Définitions des objets gérés pour le type d'interface SIP", RFC1304, février 1992. *(Obsolète, voir RFC 1694)*

[MGT:20] F. Baker, "MIB de tableau de transmission IP", RFC1354, juillet 1992. *(Obsolète, voir RFC2096)*

[MGT:21] G. Malkin et F. Baker, "Extension de MIB de RIP v2", RFC1724, novembre 1994. *(D.S.)*

[MGT:22] D. Throop, "Extension de MIB SNMP pour la couche paquet X.25", RFC1382, novembre 1992. *(P.S.)*

MGT:23] D. Throop et F. Baker, "Extension de MIB SNMP pour LAPB X.25", RFC1381, novembre 1992. *(P.S.)*

[MGT:24] D. Throop et F. Baker, "Extension de MIB SNMP pour interconnexion multiprotocole sur X.25", RFC1461, mai 1993. *(Hist.)*

[MGT:25] M. Rose, "SNMP sur OSI", RFC1418, janvier 1993. *(Historique)*

[MGT:26] G. Minshall et M. Ritter, "SNMP sur AppleTalk", RFC1419, janvier 1993. *(Historique)*

[MGT:27] S. Bostock, "SNMP sur IPX", RFC1420, janvier 1993. *(P.S.)*

[MGT:28] M. Schoffstall, J. Davin, M. Fedor et J. Case "SNMP sur Ethernet", RFC1089, février 1989. *(Obsolète, voir la RFC 4789)*

[MGT:29] J. Case, "Base de données d'informations de gestion FDDI", RFC1285, janvier 1992. *(Historique)*

[OPER:1] M. Rose, "Proposition de norme pour la transposition d'en-tête de message", RFC0886, décembre 1983.

[OPER:2] K. Sollins, "Protocole TFTP (révision 2)", RFC1350, STD 33, juin 1992. *(MàJ par 1782-85, 2347_49)*

[ROUTE:1] Moy, J., "Spécification d'OSPF", RFC1583, mars 1994. *(D.S., remplacée par la RFC 2178)*

[ROUTE:2] U. Warrier et L. Besaw, "Services et protocole communs d'informations de gestion sur TCP/IP (CMOT)", RFC1095, avril 1989. *(rendue obsolète par la RFC1189, elle-même historique)*

[ROUTE:3] C. Hedrick, "Protocole d'informations d'acheminement", RFC1058, juin 1988. *(Historique)*

[ROUTE:4] K. Lougheed et Y Rekhter, "Protocole de routeur frontière 3 (BGP 3)", RFC1267, octobre 1991. *(Hist,)*

[ROUTE:5] Y. Rekhter, P. Gross, "Application du protocole de routeur frontière dans l'Internet", RFC1772, mars 1995. *(D.S.)*

[ROUTE:6] D. Mills, "Spécification formelle du protocole de passerelle extérieure", RFC0904, avril 1984. *(Hist,)*

[ROUTE:7] E. Rosen, "Protocole de routeur extérieur (EGP)", RFC0827, octobre 1982.

[ROUTE:8] L. Seamonson et E. Rosen, "Protocole STUB de routeur extérieur", RFC0888, janvier 1984.

[ROUTE:9] D. Waitzman et autres, "Protocole d'acheminement en diffusion groupée par vecteur de distance",

RFC1075, novembre 1988.

[ROUTE:10] Deering, S., "Multicast Routing in Internetworks et Extended LANs", Proceedings of '88, Association for Computing Machinery, août 1988.

[ROUTE:11] P. Almquist, "Type de service dans la suite de protocole Internet", RFC1349, juillet 1992. *(Remplacée par RFC2474)*

[ROUTE:12] Y. Rekhter, "Expérience sur le protocole BGP", RFC1266, octobre 1991. *(Information)*

[ROUTE:13] Y. Rekhter, "Analyse du protocole BGP", RFC1265, octobre 1991. *(Information)*

[ROUTE:14] G. Malkin, "RIP version 2", RFC2453, STD 56, novembre 1998. *(Mise à jour par la RFC 4822)*

[TRANS:1] J. Postel, "Protocole de datagramme d'utilisateur (UDP)", RFC0768, (STD 6), 28 août 1980.

[TRANS:2] J. Postel (éd.), "Protocole de commande de transmission – Spécification du protocole du programme Internet DARPA, "RFC0793, (STD 7), septembre 1981.

Appendice A Exigences pour les hôtes d'acheminement de source

Sous réserve des restrictions figurant ci-dessous, un hôte PEUT être capable d'agir comme bond intermédiaire dans une route de source, en transmettant un datagramme généré par la source au prochain bond spécifié.

Cependant, en effectuant cette fonction de routeur, l'hôte DOIT respecter toutes les règles pertinentes pour un routeur qui transmet des datagrammes générés par la source [INTRO:2]. Cela inclut les dispositions spécifiques suivantes :

(A) TTL
Le champ TTL DOIT être décrémenté et le datagramme peut-être éliminé comme spécifié pour un routeur dans [INTRO:2].

(B) Destination ICMP inaccessible
Un hôte DOIT être capable de générer des messages Destination inaccessible avec les codes suivants :

4 (Fragmentation requise mais DF mis) lorsqu'un datagramme généré par la source ne peut pas être fragmenté pour entrer dans le réseau cible ;

5 (Échec de route de source) lorsqu'un datagramme généré par la source ne peut pas être retransmis, par exemple, à cause d'un problème d'acheminement ou parce que le prochain bond d'une route de source stricte n'est pas sur un réseau connecté.

(C) Adresse IP de source
Un datagramme généré par la source étant transmis PEUT (et normalement doit) avoir une adresse de source qui n'est pas une des adresses IP de l'hôte qui transmet.

(D) Option Record Route
Un hôte qui transmet un datagramme généré par la source qui contient une option Record Route DOIT mettre à jour cette option, si il en a la place.

(E) Option horodatage
Un hôte qui transmet un datagramme généré par la source qui contient une option Horodatage DOIT ajouter l'horodatage en cours à cette option, conformément aux règles de cette option.

Pour définir les règles qui restreignent la retransmission par l'hôte de datagrammes générés par la source, on utilise le terme d'acheminement de source locale si le prochain bond sera à travers la même interface physique que celle à travers laquelle est arrivé le datagramme, autrement, c'est un acheminement de source non locale.

Il est permis à un hôte d'effectuer de l'acheminement de source locale sans restriction.

Un hôte qui prend en charge l'acheminement de source non locale DOIT avoir un commutateur configurable pour désactiver la transmission, et ce commutateur DOIT être sur désactivé par défaut.

L'hôte DOIT satisfaire à toutes les exigences des routeurs pour les filtres de politique configurables [INTRO:2] restreignant la transmission non locale.

Si un hôte reçoit un datagramme avec une route de source incomplète mais ne le transmet pas pour une raison quelconque, l'hôte DEVRAIT retourner un message ICMP Destination inaccessible (code 5, Échec de route de source),

sauf si le datagramme était lui-même un message d'erreur ICMP.

Appendice B Glossaire

Cet Appendice définit des termes spécifiques utilisés dans le présent mémoire. Il définit aussi quelques termes d'usage général qui paraissent intéressants. Voir aussi dans [INTRO:9] un ensemble plus général de définitions.

Système autonome (AS, *Autonomous System*)
Un système autonome est un segment connecté d'une topologie de réseau qui comporte une collection de sous-réseaux (avec des hôtes rattachés) interconnectés par un ensemble de chemins. Les sous-réseaux et les routeurs sont censés être sous le contrôle d'une seule organisation d'opérations et maintenance (O&M). Au sein d'un AS, les routeurs peuvent utiliser un ou plusieurs protocoles d'acheminement intérieurs, et parfois plusieurs ensembles de métriques. Un AS est censé présenter aux autres AS l'apparence d'un plan d'acheminement intérieur cohérent, et un tableau cohérent des destinations accessibles à travers l'AS. Un AS est identifié par un numéro de système autonome.

Réseau connecté *(Connected Network)*
Un préfixe de réseau auquel un routeur est interfacé est souvent connu comme réseau local ou sous-réseau de ce routeur. Cependant, ces termes peuvent prêter à confusion, et donc on utilise le terme de réseau connecté dans le présent mémoire.

Sous-réseau connecté *(Connected (Sub)Network)*
Un sous-réseau connecté est un sous-réseau IP auquel un routeur est interfacé, ou un réseau connecté si le réseau connecté n'a pas de sous-réseau. Voir aussi à Réseau connecté.

Datagramme
Unité transmise entre une paire de modules internet. Les données, appelées des datagrammes, vont des sources aux destinations. Le protocole Internet ne fournit pas de facilité de communication fiable. Il n'y a pas d'accusé de réception ni de bout en bout ni bond par bond. Il n'y a pas de signalisation d'erreur ni de retransmission. Il n'y a pas de contrôle de flux. Voir IP.

Route par défaut *(Default Route)*
C'est une entrée de tableau d'acheminement qui est utilisée pour diriger toutes données adressées à tous préfixes de réseau non explicitement énumérés dans le tableau d'acheminement.

Mode dense *(Dense Mode)*
Dans la transmission en distribution *(multicast)*, deux paradigmes sont possibles : en transmission en mode dense, une distribution sur le réseau est transmise comme une distribution de couche de liaison de données à toutes les interfaces sauf celle sur laquelle elle a été reçue, sauf et jusqu'à ce que le routeur ait reçu l'instruction de ne pas le faire de la part d'un voisin d'acheminement de distribution. Voir à Mode clairsemé.

EGP
Protocole de passerelle extérieure *(Exterior Gateway Protocol)* C'est un protocole qui distribue les informations d'acheminement aux passerelles (les routeurs) qui connectent les systèmes autonomes. Voir IGP.

EGP-2
Protocole de passerelle extérieure version 2. C'est un protocole d'acheminement EGP développé pour traiter le trafic entre les systèmes autonomes sur l'Internet.

Transmetteur *(Forwarder)*
Entité logique au sein d'un routeur qui est chargée de commuter les paquets entre les interfaces du routeur. Le transmetteur prend aussi les décisions de mise en file d'attente d'un paquet pour livraison locale, de mise en file d'attente d'un paquet pour transmission à une autre interface, ou les deux.

Transmission *(Forwarding)*
La transmission est le processus qu'un routeur effectue pour chaque paquet qu'il reçoit par le routeur. Le paquet peut être consommé par le routeur, il peut être sorti sur une ou plusieurs interfaces du routeur, ou les deux. La transmission inclut le processus de décision de ce qu'il convient de faire du paquet, aussi bien le mettre en file d'attente pour une (possible) sortie ou une consommation interne.

Base d'information de transmission (FIB, *Forwarding Information Base*)
Le tableau qui contient les informations nécessaires à la transmission des datagrammes IP, dans le présent document, est

appelé la base d'informations de transmission. Au minimum, il contient l'identifiant d'interface et les informations du prochain bond pour chaque préfixe de réseau de destination accessible.

Fragment
Un datagramme IP qui représente une portion d'un paquet de couche supérieure qui était trop grand pour être envoyé en entier sur le réseau de sortie.

Interface de série d'utilisation générale *(General Purpose Serial Interface)*
Support physique capable de connecter exactement deux systèmes, et donc configurable comme une ligne point à point, mais aussi configurable pour prendre en charge le réseautage de couche de liaison en utilisant des protocoles comme X.25 ou le relais de trame. Un réseau de couche de liaison connecte un autre système à un commutateur, et une couche supérieure de communication multiplexe des circuits virtuels sur la connexion. Voir Ligne point à point.

IGP
Protocole de passerelle intérieure *(Interior Gateway Protocol)* : Protocole qui distribue des informations d'acheminement avec un système autonome (AS). Voir EGP.

Adresse d'interface IP *(Interface IP Address)*
L'adresse IP et la longueur de préfixe de réseau qui sont allouées à une interface spécifique d'un routeur.

Adresse Internet *(Internet Address)*
Numéro alloué qui identifie un hôte dans un internet. Il a deux parties : une adresse IP et une longueur de préfixe. La longueur de préfixe indique combien des bits les plus spécifiques de l'adresse constituent le préfixe de réseau.

IP
Protocole Internet : Protocole de couche réseau pour l'Internet. C'est un protocole de commutation de paquets, en datagrammes défini dans la RFC 791. IP ne fournit pas de facilités de communications fiables ; c'est-à-dire qu'il n'y a pas d'accusé de réception de bout en bout des bonds.

Datagramme IP
Un datagramme IP est l'unité de transmission de bout en bout du protocole Internet. Un datagramme IP comporte un en-tête IP suivi par toutes les données de couche supérieure (telles que TCP, UDP, ICMP, et les autres). Un datagramme IP est un en-tête IP suivi par un message. Un datagramme IP est une unité de transmission complète de bout en bout. Un datagramme IP est composé d'un ou plusieurs fragments. Dans le présent mémoire, le terme datagramme non qualifié devrait être compris comme se référant à un datagramme IP.

Fragment IP
Un fragment IP est un composant d'un datagramme IP. Un fragment IP comporte un en-tête IP suivi de tout ou partie de la couche supérieure du datagramme IP d'origine. Un ou plusieurs fragments IP comprennent un seul datagramme IP. Dans le présent mémoire, le terme de fragment non qualifié devrait être compris comme se référant à un fragment IP.

Paquet IP
Un datagramme IP ou un fragment IP. Dans le présent mémoire, le terme non qualifié de paquet devrait généralement être compris comme se référant à un paquet IP.

Interface logique [de réseau]
On définit une interface logique [de réseau] comme un chemin logique, distingué par une adresse IP unique, à un réseau connecté.

Filtrage martien
Un paquet qui contient une adresse de source ou de destination invalide est considéré comme martien et éliminé.

Unité de transmission maximum (MTU, *Maximum Transmission Unit*)
C'est la taille du plus grand paquet qui peut être transmis ou reçu à travers une interface logique. Cette taille inclut l'en-tête IP mais n'inclut pas la taille de tout en-tête de couche de liaison ou tramage.

Diffusion groupée *(Multicast)*
Paquet qui est destiné à plusieurs hôtes. Voir diffusion.

Adresse de diffusion groupée *(Multicast address)*
Type spécial d'adresse reconnaissable par plusieurs hôtes. Une adresse de diffusion groupée est parfois appelée une adresse fonctionnelle ou adresse de groupe.

Préfixe de réseau *(Network Prefix)*
Portion d'une adresse IP qui signifie un ensemble de systèmes. Elle est sélectionnée à partir de l'adresse IP par l'ajout logique d'un gabarit de sous-réseau avec l'adresse, ou (de façon équivalente) en ne mettant pas les bits de l'adresse parmi les bits de plus fort poids de <prefix-length> de l'adresse à zéro.

Générés *(Originate)*
Les paquets peuvent être transmis par un routeur pour une de deux raisons : 1) le paquet a été reçu et va être transmis, ou 2) le routeur a lui-même créé le paquet pour le transmettre (comme des publications d'acheminements). Les paquets que le routeur crée pour les transmettre vont être générés au routeur.

Paquet
Un paquet est l'unité de données qui est passée à travers l'interface entre la couche Internet et la couche de liaison. Il inclut un en-tête IP et des données. Un paquet peut être un datagramme IP complet ou un fragment d'un datagramme IP.

Chemin, ou route *(Path)*
C'est la séquence des routeurs et (sous-)réseaux qu'un paquet traverse d'un routeur particulier à un hôte de destination particulier. Noter qu'un chemin est unidirectionnel ; il n'est pas inhabituel d'avoir différents chemins dans les deux directions entre une paire d'hôtes donnée.

Réseau physique *(Physical Network)*
Un réseau physique est un réseau (ou un morceau d'un internet) qui est contigu à la couche de liaison. Sa structure interne (s'il en est une) est transparente pour la couche Internet. Dans le présent mémoire, plusieurs composants de supports qui sont connectés en utilisant des appareils comme des ponts ou des répéteurs sont considérés comme un seul réseau physique car de tels appareils sont transparents pour IP.

Interface de réseau physique *(Physical Network Interface)*
C'est une interface physique à un réseau connecté et qui a une adresse de couche de liaison (qui peut être unique). Plusieurs interfaces de réseau physique sur un seul routeur peuvent partager la même adresse de couche de liaison, mais l'adresse doit être unique pour les différents routeurs sur le même réseau physique.

Ligne point à point *(Point to Point Line)*
Support physique capable de connecter exactement deux systèmes. Dans le présent document, c'est seulement utilisé pour se référer à de telles lignes lorsqu'elles sont utilisées pour connecter des entités IP. Voir à Interface en série d'utilisation générale.

Routeur
Ordinateur dédié à un usage particulier qui connecte plusieurs réseaux. Les routeurs commutent les paquets entre ces réseaux dans un processus appelé transmission. Ce processus peut être répété plusieurs fois sur un seul paquet par plusieurs routeurs jusqu'à ce que le paquet puisse être livré à la destination finale – commutant le paquet de routeur en routeur, jusqu'à ce que le paquet arrive à sa destination.

RPF
Transmission de chemin inverse *(Reverse Path Forwarding)* - Méthode utilisée pour déduire les prochains bonds pour les paquets en diffusion et diffusion groupée.

Éliminer en silence *(Silently Discard)*
Le présent mémoire spécifie plusieurs cas où un routeur va éliminer en silence un paquet (ou datagramme) reçu. Cela signifie que le routeur devrait éliminer le paquet sans autre traitement, et que le routeur n'enverra aucun message d'erreur ICMP (voir au paragraphe 4.3.2) en résultant. Cependant, pour le diagnostic des problèmes, le routeur devrait avoir la capacité d'enregistrer l'erreur (voir au paragraphe 1.3.3) y compris le contenu du paquet éliminé en silence, et devrait enregistrer l'événement dans un compteur statistique.

Ignorer en silence *(Silently Ignore)*
Un routeur est dit ignorer en silence une erreur ou condition si il n'entreprend aucune action autre que d'éventuellement générer un rapport d'erreur dans un enregistrement d'erreur ou au moyen de quelque protocole de gestion de réseau, et d'éliminer, ou ignorer, la source de l'erreur. En particulier, le routeur NE génère PAS de message d'erreur ICMP.

Mode clairsemé *(Sparse Mode)*
Dans le mode diffusion groupée, deux paradigmes sont possibles : dans la transmission en mode clairsemé, un datagramme de diffusion groupée de couche réseau est transmis comme une trame de diffusion groupée de couche de liaison des données aux routeurs et hôtes qui l'ont demandé. L'état de transmission initial est l'inverse du mode dense en ce qu'il suppose qu'aucune partie du réseau ne veut les données. Voir Mode dense.

Adresse de destination spécifique *(Specific-destination address)*
C'est l'adresse de destination dans l'en-tête IP sauf si l'en-tête contient une adresse IP en diffusion ou en diffusion groupée, auquel cas, la destination spécifique est une adresse IP allouée à l'interface physique sur laquelle le paquet est arrivé.

Sous-réseau *(subnet)*
Portion d'un réseau, qui peut être un réseau physiquement indépendant, qui partage une adresse réseau avec d'autres portions du réseau et se distingue par un numéro de sous-réseau. Un sous-réseau est à un réseau ce qu'un réseau est à un internet.

Numéro de sous-réseau
Partie de l'adresse internet qui désigne un sous-réseau. Il est ignoré pour les besoins de l'acheminement internet, mais utilisé pour le routage intranet.

TOS
Type de Service : Champ dans l'en-tête IP qui représente le degré de fiabilité attendu de la couche réseau par la couche transport ou application.

TTL
Durée de vie *(Time To Live)* : Champ dans l'en-tête IP qui représente la durée pendant laquelle un paquet est considéré comme valide. C'est une combinaison du compte de bonds et de la valeur d'un temporisateur.

Appendice C Directions futures

Le présent appendice fait la liste des travaux que les révisions à venir du présent document pourraient souhaiter réaliser.

Dans la préparation des exigences pour les routeurs, on butte sur plusieurs autres problèmes d'architecture. Chacun d'eux est un peu traité dans le document, mais mériterait d'être classé comme question ouverte dans l'architecture IP.

La plupart des sujets présentés ici indiquent généralement des domaines où la technologie est encore relativement jeune et il n'est pas approprié de développer des exigences spécifiques car la communauté en est toujours à acquérir l'expérience du fonctionnement.

D'autres sujets représentent de domaines de recherches en cours et concernent des domaines que le développeur prudent devrait surveiller attentivement.

(1) SNMP Version 2
(2) MIB SNMP supplémentaires
(7) Des exigences plus précises pour les routes de fuite entre les protocoles d'acheminement
(8) Sécurité des systèmes de routeur
(9) Sécurité du protocole d'acheminement
(10) Sécurité de la couche Protocole inter-réseaux. Il y a eu un travail considérable pour préciser la sécurité de IP depuis le travail original de rédaction du présent document. Ce travail sur la sécurité devrait être inclus ici.
(12) Partage de charge
(13) Envoi des fragments sur des chemins différents
(15) Plusieurs (sous-)réseaux logiques sur le même câble. Les exigences pour les routeurs ne portent pas sur cela. Nous avons faits quelques tentatives pour identifier les pièces de l'architecture (par exemple, transmission de diffusions dirigées et émission des redirections) où la formulation des règles était faite très soigneusement pour que les bons événements surviennent, et essai de distinction claire des interfaces logiques et des interfaces physiques. Cependant, nous n'avons pas étudié cette question en détail, et nous ne sommes pas certains que toutes les règles du document soient correctes en présence de plusieurs sous-réseaux logiques sur le même câble.
(15) Contrôle de l'encombrement et gestion des ressources. Suivant l'avis des experts de l'IETF (Mankin et Ramakrishnan) nous déconseillons (NE DEVRAIT PAS) l'extinction de source *(Source Quench)* et ne disons pas grand chose de concret (paragraphe 5.3.6).
(16) Développer un document d'exigences pour la couche de liaison qui seraient communes pour les routeurs et les hôtes.
(17) Développer un algorithme commun PPP LQM.
(18) Investiguer sur les autres informations (au dessus et au-delà du paragraphe [3.2]) qui passent entre les couches, comme les MTU de réseau physique, les transpositions de préséance IP en valeurs de priorité de couche de liaison, etc.
(19) La couche de liaison devrait elle notifier IP si la résolution d'adresse a échoué (juste comme elle notifie IP

lorsqu'il y a un problème de valeur de priorité de couche de liaison) ?
(20) Tous les routeurs devraient ils être obligé de mettre en œuvre un résolveur DNS ?
(21) Un utilisateur humain devrait-il être capable d'utiliser un nom d'hôte partout où on peut utiliser une adresse IP quand on configurer le routeur ? Même dans ping et traceroute ?
(22) Le projet de réflexions d'Almquist sur le prochain bond et ses réflexions sur les pertes de route doivent être revues, mises à jour et publiées.
(23) Des investigations sont nécessaires pour déterminer si un message redirect pour la préséance est nécessaire ou non. Si non, les redirections de type de service sont elles acceptables ?
(24) RIPv2 et RIP+CIDR et longueur variable des préfixes de réseau.
(25) BGP-4 CIDR va devenir important, et chacun parie sur BGP-4. On ne peut éviter de le mentionner. Il est probablement nécessaire de décrire les différences entre BGP-3 et BGP-4, et d'explorer les questions de mise à niveau...
(26) IP mobile à acheminement de source lâche *(Loose Source Route Mobile IP)* et certaines diffusions groupées pourraient avoir besoin de cela. Peut-être cela devrait être élevé à DEVRAIT (selon la suggestion de Fred Baker).

Appendice D Protocoles d'acheminement en diffusion groupée

La diffusion groupée est une technologie relativement nouvelle dans la famille du protocole Internet. Elle n'est pas encore largement déployée ou communément utilisée. Son importance, cependant, est destinée à croître dans les années à venir.

Le présent Appendice décrit certaines des technologies qui sont explorées pour l'acheminement des diffusions groupées à travers l'Internet.

Un développeur attentif se tiendra au courant des développements dans ce domaine pour développer de façon appropriée les facilités de diffusion groupée.

Le présent Appendice ne spécifie aucune norme ni exigence.

D.1 Introduction

Les protocoles d'acheminement de diffusion groupée permettent la transmission de datagramme IP en diffusion groupée à travers un internet TCP/IP. Généralement ces algorithmes transmettent le datagramme sur la base de ses adresses de source et de destination. De plus, le datagramme peut avoir besoin d'être transmis à plusieurs membres d'un groupe de diffusion groupée, ce qui exige à certains moments que le datagramme soit dupliqué et envoyé sur plusieurs interfaces.

L'état des protocoles d'acheminement de diffusion groupée est moins développé que celui des protocoles disponibles pour la transmission IP en envoi individuel. Trois protocoles d'acheminement de diffusion groupée expérimentaux ont été publiés pour TCP/IP. Chacun utilise le protocole IGMP (exposé au paragraphe 4.4) pour surveiller l'adhésion au groupe de diffusion groupée.

D.2 Protocole d'acheminement de diffusion groupée à vecteur de distance - DVMRP

DVMRP, documenté dans [ROUTE:9], se fonde sur le vecteur de distance ou la technologie Bellman-Ford. Il n'achemine que les datagrammes en diffusion groupée, et le fait au sein d'un seul système autonome. DVMRP est une mise en œuvre de l'algorithme de diffusion sur le chemin inverse tronqué *(Truncated Reverse Path Broadcasting)* décrit dans [ROUTE:10]. De plus, il spécifie le tunnelage des diffusions groupées IP à travers des domaines IP qui ne disposent pas de la capacité d'acheminement de la diffusion groupée.

D.3 Extensions de diffusion groupée à OSPF - MOSPF

MOSPF, actuellement en cours de développement, est un ajout rétro compatible à OSPF qui permet la transmission à la fois de diffusions groupées et d'envois individuels IP au sein d'un système autonome. Les routeurs MOSPF peuvent être mélés à des routeurs OSPF au sein d'un domaine d'acheminement, et ils vont interopérer dans la transmission des envois individuels. OSPF est un protocole d'état de liaison ou fondé sur SPF.

En ajoutant des annonces d'état de liaison qui précisent l'adhésion au groupe, les routeurs MOSPF peuvent calculer le chemin d'un datagramme en diffusion groupée comme un arbre dont la racine est la source du datagramme. Les branches qui ne contiennent pas de membre du groupe peuvent alors être éliminées, ce qui supprime les bonds de transmission du datagramme qui ne sont pas nécessaires.

D.4 Diffusion groupée indépendante du protocole - PIM

PIM, actuellement en cours de développement, est un protocole d'acheminement de diffusion groupée qui fonctionne sur une infrastructure d'envoi individuel existante. PIM est adapté à la fois pour les groupes à adhésion dense et clairsemée. Il est différent des autres protocoles, car il utilise un modèle conjoint explicite pour les groupes clairsemés. La jonction survient sur un arbre de partage et peut passer à un arbre par source. Lorsque la bande passante est abondante et que l'adhésion au groupe est dense, la redondance peut être réduite en écoulant des données sur toutes les liaisons et en élaguant ultérieurement les exceptions où il n'y a pas de membre du groupe.

Appendice E Algorithmes supplémentaires de choix du prochain bond

Le paragraphe 5.2.4.3 spécifie un algorithme que les routeurs devraient utiliser pour choisir le prochain bond d'un paquet.

Cet appendice fournit une perspective historique du problème du choix du prochain bond. Il présente aussi plusieurs règles supplémentaires d'élagage et d'algorithmes de choix du prochain bond qu'on pourra trouver dans l'Internet.

Cet appendice présente des matériaux tirés d'un travail antérieur, non publié de Philip Almquist : Réflexions sur le prochain bond.

Cet appendice ne spécifie aucune norme ni exigence.

E.1 Quelques perspectives historiques

Il est utile de revoir brièvement l'historique de ce sujet, qui commence avec ce qu'on appelle parfois le "modèle classique" de la façon dont un routeur prend ses décisions d'acheminement. Ce modèle est antérieur à IP. Dans ce modèle, un routeur ne parle qu'un seul protocole d'acheminement comme RIP. Le protocole détermine complètement le contenu de la base d'informations de transmission (FIB) du routeur. L'algorithme de recherche de route est trivial : le routeur cherche dans la FIB une route dont l'attribut de destination correspond exactement à la portion du préfixe de réseau de l'adresse de destination dans le paquet. S'il en trouve un, il l'utilise ; si aucun n'est trouvé, la destination est inaccessible. Comme le protocole d'acheminement conserve au plus une route pour chaque destination, le problème de ce qu'il convient de faire lorsqu'il y a plusieurs routes correspondant à la même destination ne peut pas se présenter.

Au fil des ans, ce modèle classique s'est un peu enrichi. Avec le déploiement de routes par défaut, de sous-réseaux, et routes d'hôte, il est devenu possible d'avoir plus d'une entrée du tableau d'acheminement qui corresponde d'une façon ou d'une autre à la destination. Ceci était facilement résolu par un consensus sur une hiérarchie des routes : les routes hôtes devraient être préférées aux routes de sous-réseau, les routes de sous-réseau aux routes de réseau, et les routes de réseau aux routes par défaut.

Avec le développement de technologies prenant en charge des gabarits de sous-réseau de longueur variable (préfixes de réseau à longueur variable), l'approche générale restait la même bien que sa description devienne un peu plus compliquée ; les préfixes de réseau ont été introduits comme une simplification consciente et une régularisation de l'architecture. Nous disons maintenant que pour chaque route un préfixe de route de réseau a une longueur de préfixe associée. Cette longueur de préfixe indique le nombre de bits du préfixe. Cela peut aussi être représenté en utilisant le gabarit de sous-réseau classique. Une route ne peut pas être utilisée pour acheminer un paquet si chaque bit de poids fort dans le préfixe de réseau de la route ne correspond pas au bit équivalent dans l'adresse de destination du paquet. Les routes avec plus de bits établis dans leur gabarit sont préférées aux routes qui ont moins de bits établis dans leur gabarit. C'est simplement une généralisation de la hiérarchie des routes décrite ci-dessus, et on l'appellera dans le reste du présent mémoire le choix d'une route par la préférence de la plus longue correspondance.

Une autre façon d'enrichir le modèle classique est par quelques assouplissements de la notion qu'un protocole d'acheminement a un complet contrôle sur le contenu du tableau d'acheminement. D'abord, les routes statiques ont été introduites. Pour la première fois, il a été possible d'avoir simultanément deux routes (une dynamique et une statique) pour la même destination. Lorsque cela arrivait, un routeur devait avoir une politique (configurable dans certains cas, et dans d'autres, choisie par l'auteur du logiciel du routeur) qui déterminait si la route statique ou la dynamique était préférée. Cependant, cette politique n'était utilisée que pour résoudre un problème lorsque la plus longue correspondance ne déterminait pas de façon univoque quelle route utiliser. Et donc, par exemple, une route statique par défaut ne serait jamais préférée à une route de réseau dynamique même si la politique était de préférer les routes statiques aux routes dynamiques.

Le modèle classique devait être encore amélioré lorsque furent inventés les protocoles d'acheminement inter domaines. Les protocoles d'acheminement traditionnels ont alors été appelés "protocole de passerelle intérieure" (IGP), et à chaque site Internet il y avait un nouvel animal étrange appelé une "passerelle extérieure", un routeur qui parlait EGP à plusieurs "passerelles centrales BBN" (les routeurs qui formaient à l'époque le cœur de réseau de l'Internet) en même

temps qu'il parlait son IGP aux autres routeurs sur son site. Les deux protocoles voulaient déterminer le contenu du tableau d'acheminement du routeur. Théoriquement, il pouvait en résulter qu'un routeur ait trois routes (EGP, IGP, et statique) pour la même destination. À cause de la topologie de l'Internet de l'époque, il fut résolu avec peu de débats que les routeurs seraient mieux servis par une politique de préférence des routes IGP sur les routes EGP. Cependant, le caractère sacré de la plus longue correspondance restait indiscuté : une route par défaut apprise de IGP ne serait jamais préférée à une route réseau apprise d'EGP.

Bien que la topologie de l'Internet, et par conséquent l'acheminement dans l'Internet, ait considérablement évolué depuis, cette version légèrement augmentée du modèle classique a survécu intacte jusqu'à ce jour dans l'Internet (sauf que BGP a remplacé EGP). Conceptuellement (et souvent dans les mises en œuvre) chaque routeur a un tableau d'acheminement et un ou plusieurs processus de protocole d'acheminement. Chacun de ces processus peut ajouter toute entrée qu'il lui plait, et peut supprimer ou modifier toute entrée qu'il a créée. Lorsqu'il achemine un paquet, le routeur prend la meilleure route en utilisant la plus longue correspondance, augmentée d'un mécanisme de politique pour résoudre les conflits. Bien que ce modèle classique augmenté nous ait bien servi, il a un certain nombre d'insuffisances :

- Il ignore (bien qu'il puisse être augmenté pour le prendre en compte) des caractéristiques de chemin telles que la qualité de service et la MTU.
- Il ne prend pas en charge les protocoles d'acheminement (comme OSPF et IS-IS intégré) qui exigent des algorithmes de recherche de route différents de la plus longue correspondance pure.
- Il n'y a pas eu de consensus fort sur ce que devrait être le mécanisme de résolution de conflit. Les mécanismes de résolution de conflit ont souvent été trouvés difficiles sinon impossibles à configurer d'une façon telle que le routeur prenne toujours ce que le gestionnaire de réseau considère comme le chemin "correct".

E.2 Règles de base supplémentaires

Le paragraphe 5.2.4.3 définit plusieurs règles d'élagage à utiliser pour choisir les chemins à partir du FIB. D'autres règles pourraient aussi être utilisées :

- Classe de chemin OSPF

Les protocoles d'acheminement qui ont des domaines ou font une distinction entre chemin interne et chemin externe divisent leurs chemins en classes selon le type d'informations utilisées pour les calculer. Un chemin est toujours choisi à partir de la classe préférée sauf si aucun n'est disponible, auquel cas il est choisi dans la seconde classe préférée, et ainsi de suite. En OSPF, les classes (dans l'ordre de préférence) sont intra zone, inter zone, type 1 externe (chemins externes avec métrique interne), et type 2 externe. Comme subtilité supplémentaire, un routeur est configuré pour savoir quelles adresses devraient être accessibles en utilisant des chemins intra zone, et pour ne pas utiliser de chemins inter zone ou externes pour atteindre ces destinations même si aucun chemin intra zone n'est disponible.

Plus précisément, on suppose que chaque chemin a un attribut de classe, appelé route.class, qui est alloué par le protocole d'acheminement. L'ensemble des chemins candidats est examiné pour déterminer si il en contient pour qui route.class = intra zone S'il en est ainsi, tous les chemins excepté ceux pour lesquels route.class = intra-zône sont éliminés. Autrement, le routeur vérifie si la destination du paquet entre dans les gammes d'adresses configurées pour la zone locale. S'il en est ainsi, tout l'ensemble des chemins candidats est supprimé. Autrement, l'ensemble des chemins candidats est examiné pour déterminer si il en contient pour qui route.class = inter zone. S'il en est ainsi, tous les chemins excepté ceux pour lesquels route.class = interzone sont éliminés. Autrement, l"ensemble des chemins candidats est examiné pour déterminer si il en contient pour lesquels route.class = type 1 externe. S'il en est ainsi, tous les chemins excepté ceux pour lesquels route.class = type 1 externe sont éliminés.

- Classe de chemin IS-IS

Les classes de chemin IS-IS fonctionnent de la même façon que celles d'OSPF. Cependant, l'ensemble des classes définies par IS-IS intégré est différent, de sorte qu'il n'y a pas de transposition biunivoque entre les classes de chemin IS-IS et les classes de chemin OSPF. Les classes de chemin utilisés par IS-IS intégré sont (dans l'ordre de préférence décroissante) intra zone, interzone, et externe.

La classe interne de IS-IS intégré est équivalente à la classe interne OSPF. De même, la classe externe d'IS-IS intégré est équivalente à la classe externe de type 2 d'OSPF. Cependant, IS-IS intégré ne fait pas de distinction entre les chemins interzone et les r chemins externes avec métrique interne – tous deux sont considérés comme des chemins interzone Et donc, OSPF préfère un vraie chemin interzone à un chemin externe avec métrique interne, tandis que IS-IS intégré donne aux deux types de chemins une préférence égale.

- Politique IDPR

Un cas spécifique de politique. Le groupe de travail Politique inter domaine de l'IETF travaille à un protocole d'acheminement appelé Acheminement à politique inter domaine (IDPR, *Inter-Domain Policy Routing*) pour prendre en

charge un acheminement fondé sur une vraie politique dans l'Internet. Les paquets avec certaines combinaisons d'attributs d'en-tête (comme une combinaison spécifique d'adresses de source et de destination ou une option de route de source IDPR spéciale) sont obligés d'utiliser les chemins fournis par le protocole IDPR. Et donc, à la différence des autres règles d'élagage de politique, la politique IDPR devrait être appliquée avant toute autre règle d'élagage excepté Basic Match.

Précisément, la politique IDPR examine le paquet transmis pour s'assurer si ses attributs exigent qu'il soit transmis en utilisant des chemins fondés sur la politique. S'il en est ainsi, la politique IDPR supprime tous les chemins qui ne sont pas fournis par le protocole IDPR.

E.3 Quelques algorithmes de recherche de chemin

Cette section examine plusieurs algorithmes de recherche de chemin qui sont utilisés ou ont été proposés. Chacun d'eux est décrit en donnant la séquence des règles d'élagage qu'il utilise. Les forces et faiblesses de chaque algorithme sont présentées.

E.3.1 Algorithme classique révisé

L'algorithme classique révisé est la forme de l'algorithme traditionnel qui a été exposé au paragraphe E.1. Les étapes de cet algorithme sont :

1. Correspondance de base
2. Plus longue correspondance
3. Meilleure métrique
4. Politique

Certaines des mises en œuvre omettent l'étape de politique, car elle n'est nécessaire que lorsque des chemins pourraient avoir des métriques qui ne sont pas comparables (parce qu'ils ont été acquis auprès de domaines d'acheminement différents).

Les avantages de cet algorithme sont :

(1) Il est largement mis en œuvre
(2) Excepté pour l'étape de politique (qu'une mise en œuvre peut choisir de rendre arbitrairement complexe) l'algorithme est simple à comprendre et à mettre en œuvre.

Ses inconvénients sont :

(1) Il ne traite pas les classes de chemin IS-IS ou OSPF, et donc ne peut pas être utilisé pour IS-IS intégré ou OSPF.
(2) Il ne traite pas le TOS ou autres attributs de chemin.
(3) Les mécanismes de politique ne sont pas du tout normalisés, et ils sont donc souvent spécifiques de la mise en œuvre. Cela cause un travail supplémentaire aux développeurs (qui doivent inventer des mécanismes de politique appropriés) et aux utilisateurs (qui doivent apprendre à utiliser le mécanisme. Cette absence d'un mécanisme normalisé rend aussi difficile la construction de configurations cohérentes pour les routeurs provenant de différents fabricants. Cela présente un obstacle pratique significatif à l'interopérabilité.
(4) Le mécanisme de politique propriétaire actuellement fourni par les fabricants est souvent inadéquat dans des parties complexes de l'Internet.
(5) L'algorithme n'a pas été rédigé dans un document ou norme disponible au public. Il fait parie du folklore de l'Internet.

E.3.2 Variante de l'algorithme des exigences de routeur

Certains membres du groupe de travail Exigences pour les routeurs ont proposé une légère variante de l'algorithme décrit au paragraphe 5.2.4.3. Dans cette variante, la correspondance avec le type de service demandé n'est pas considérée comme plus importante, et plutôt moins importante, que de correspondre autant que possible à l'adresse de destination. Par exemple, cet algorithme préfèrerait un chemin par défaut qui a le type de service correct à un chemin de réseau qui aurait le type de service par défaut, alors que l'algorithme du paragraphe 5.2.4.3 ferait le choix opposé.

Les étapes de l'algorithme sont :

1 Correspondance de base
2 TOS faible
3 Plus longue correspondance
4 Meilleure métrique
5 Politique

Le débat entre ceux qui proposent cet algorithme et les partisans de l'algorithme des exigences de routeur suggère que chaque camp peut montrer des cas où son algorithme conduit à un acheminement plus simple, plus intuitif que ne le fait l'autre algorithme. Cette variante a le même ensemble d'avantages et d'inconvénients que l'algorithme spécifié en 5.2.4.3, excepté que l'élagage sur le TOS faible avant l'élagage sur la plus longue correspondance rend cet algorithme

moins compatible avec OSPF et IS-IS intégré que l'algorithme des exigences de routeur standard.

E.3.3 Algorithme OSPF

OSPF utilise un algorithme qui est virtuellement identique à l'algorithme des exigences de routeur excepté sur un point crucial : OSPF prend en compte les classes d'acheminement OSPF.

L'algorithme est :
1 Correspondance de base
2 Classe d'acheminement OSPF
3 Plus longue correspondance
4 TOS faible
5 Meilleure métrique
6 Politique

La prise en charge du type de service n'est pas toujours présente. Si elle ne l'est pas, bien sûr, la quatrième étape sera omise.

Cet algorithme a quelques avantages sur l'algorithme classique révisé :
(1) Il prend en charge l'acheminement par type de service.
(2) Ses règles sont écrites, plutôt que de faire simplement partie du folklore Internet.
(3) Il travaille (évidemment) avec OSPF.

Cependant, cet algorithme conserve certains des inconvénients de l'algorithme classique révisé :
(1) Les propriétés de chemin autres que le type de service (par exemple, la MTU) sont ignorées.
(2) Comme dans l'algorithme classique révisé, les détails (ou même l'existence) de l'étape Politique sont laissés à la discrétion de la mise en œuvre.

L'algorithme OSPF a encore un autre inconvénient (qui n'est pas partagé par l'algorithme classique révisé). Les chemins internes OSPF (intra zone ou interzone) sont toujours considérés comme supérieurs aux chemins appris d'autres protocoles d'acheminement, même dans les cas où le chemin OSPF correspond par moins de bits à l'adresse de destination. C'est une décision de politique qui est inappropriée dans certains réseaux.

Finalement, il vaut de noter que la prise en charge du TOS par l'algorithme OSPF souffre d'une faiblesse en ce que les protocoles d'acheminement qui prennent en charge le TOS sont implicitement préférés lors de la transmission de paquets qui ont des valeurs de TOS différentes de zéro. Cela peut n'être pas approprié dans certains cas.

E.3.4 Algorithme IS-IS intégré

IS-IS intégré utilise un algorithme qui est similaire mais pas tout à fait identique à l'algorithme OSPF. IS-IS intégré utilise un ensemble de classes d'acheminement différent, et diffère légèrement dans le traitement du type de service. L'algorithme est :
1 Correspondance de base
2 Classes d'acheminement IS-IS
3 Plus longue correspondance
4 TOS faible
5 Meilleure métrique
6 Politique

Bien que IS-IS intégré utilise le TOS faible, le protocole n'est capable que de porter des chemins pour un petit sous-ensemble spécifique des valeurs possible pour le champ TOS dans l'en-tête IP. Les paquets qui contiennent d'autres valeurs dans le champ TOS sont acheminés en utilisant le TOS par défaut.

La prise en charge du type de service est facultative ; si elle est désactivée, la quatrième étape sera omise. Comme dans OSPF, la spécification n'inclut pas l'étape de politique.

Cet algorithme présente quelques avantages sur l'algorithme classique révisé :
(1) Il prend en charge l'acheminement par le type de service.
(2) Ses règles sont écrites, plutôt que de simplement faire partie du folklore de l'Internet.
(3) Il travaille (évidemment) avec IS-IS intégré.

Cependant, cet algorithme conserve aussi certains des inconvénients de l'algorithme classique révisé :

(1) Les propriétés de chemin autres que le type de service (par exemple, MTU) sont ignorées.
(2) Comme dans l'algorithme classique révisé, les détails (ou même l'existence) de l'étape de politique sont laissés à la discrétion de la mise en œuvre.
(3) Il ne travaille pas avec OSPF à cause des différences entre les classes d'acheminement IS-IS et les classes d'acheminement OSPF. Et aussi, parce que IS-IS ne prend en charge qu'un sous ensemble des valeurs de TOS possibles, certaines mises en œuvre évidentes de l'algorithme IS-IS intégré n'accepteraient pas l'interprétation du TOS d'OSPF.

L'algorithme IS-IS intégré a aussi d'autres inconvénients (qui ne sont pas partagés par l'algorithme classique révisé) : les chemins internes IS-IS (intra zone ou interzone) sont toujours considérés comme supérieurs aux chemins appris d'autres protocoles d'acheminement, même dans les cas où le chemins IS-IS correspond à moins de bits de l'adresse de destination et ne fournit pas le type de service demandé. C'est une décision de politique qui peut n'être pas appropriée dans tous les cas.

Finalement, il vaut de noter que la prise en charge du TOS par l'algorithme IS-IS intégré souffre de la même déficience que celle notée pour l'algorithme OSPF.

Considérations sur la sécurité

Bien que l'objet du présent document soit l'interopérabilité plutôt que la sécurité, de nombreux paragraphes du présent document ont des rapports évidents avec la sécurité du réseau.

Sécurité signifie des choses différentes selon les individus. Du point de vue d'un routeur, la sécurité est tout ce qui aide à garder opérationnel son propre réseau et en plus aide à garder l'Internet en bonne santé dans son ensemble. Pour les besoins du présent document, les services de sécurité qui nous importent sont le déni de service, l'intégrité, et l'authentification lorsqu'elle s'applique aux deux premiers. La confidentialité est importante comme service de sécurité, mais elle n'est qu'une préoccupation périphérique pour un routeur – au moins au jour de la rédaction du présent document.

Divers endroits du présent document possèdent des paragraphes intitulés Considérations sur la sécurité. Ces paragraphes exposent les considérations spécifiques qui s'appliquent au sujet général exposé.

Le présent document indique rarement "Faites ceci et votre routeur/réseau sera en sécurité". Plus vraisemblablement, il dit ceci est une bonne idée et si vous le faites cela *peut* améliorer la sécurité de l'Internet et de votre système local en général.

Malheureusement, ceci est l'état de l'art AU MOMENT PRÉSENT. Peu, s'il en est, de protocoles de routeur du réseau se soucient d'avoir des dispositifs de sécurité raisonnables incorporés. L'industrie et les concepteurs de protocoles ont été et continuent d'être en dispute sur ces questions. Il y a des progrès, mais très petits, comme l'authentification d'homologue à homologue disponible dans les protocoles d'acheminement BGP et OSPF.

En particulier, le présent document note les recherches actuelles pour développer et améliorer la sécurité du réseau. Des domaines de recherche spécifiques, des développements, et de l'ingénierie sont en bonne voie au moment de la rédaction (décembre 1993) dans la sécurité d'IP, la sécurité SNMP, et les technologies communes d'authentification .

Malgré ce qui précède, il y a des choses que les fabricants et les utilisateurs peuvent faire pour améliorer la sécurité de leur routeur. Les fabricants devraient avoir une copie de Trusted Computer System Interpretation [INTRO:8]. Même si un fabricant décide de ne pas soumettre son appareil à une vérification formelle selon ces lignes directrices, la publication donne d'excellents conseils sur la conception générale sur les pratiques de la sécurité pour le matériel informatique.

Appendice F Protocoles d'acheminement historiques

Certains protocoles d'acheminement sont courants sur l'Internet, mais les auteurs du présent document ne peuvent en conscience recommander leur utilisation. Ce n'est pas parce qu'ils ne fonctionnent pas correctement, mais parce que les caractéristiques de l'Internet supposées dans leur conception (acheminement simple, pas de politique, un réseau à un seul "routeur central" sous administration commune, complexité limitée, ou diamètre de réseau limité) ne sont pas des attributs de l'Internet d'aujourd'hui. Les parties de l'Internet qui les utilisent encore sont généralement des domaines "marginaux" limités avec une faible complexité.

En témoignage de bonne foi, un recueil de conseils sur leur mise en œuvre est exposé dans cette section.

F.1 Protocole de passerelle extérieure - EGP

F.1.1 Introduction

Le protocole de passerelle extérieure (EGP, *Exterior Gateway Protocol*) spécifie un EGP qui est utilisé pour échanger des informations d'accessibilité entre les routeurs du même système autonome ou de systèmes différents. EGP n'est pas considéré comme un protocole d'acheminement car il n'y a pas d'interprétation standard (c'est-à-dire métrique) des champs de distance dans le message EGP update, de sorte que les distances ne sont comparables que parmi les routeurs du même AS. Il est cependant conçu pour fournir des informations d'accessibilité de haute qualité, à la fois sur les routeurs voisins et sur les chemins vers les routeurs non voisins.

EGP est défini par [ROUTE:6]. Celui qui veut le mettre en œuvre devra très certainement vouloir lire [ROUTE:7] et [ROUTE:8], car ils contiennent des explications utiles et les notions de base.

Discussion
La spécification EGP actuelle a de sérieuses limites, dont la plus importante est une restriction qui limite les routeurs à ne publier que les réseaux qui sont accessibles de l'intérieur du système autonome du routeur. Cette restriction à l'encontre de la propagation des informations EGP à des tiers est destinée à empêcher des boucles d'acheminement durables. Cela limite en fait EGP à une hiérarchie à deux niveaux.

La RFC-975 ne fait pas partie de la spécification EGP, et devrait être ignorée.

F.1.2 Description rapide du protocole

Voisins indirects : RFC-888, page 26
Une mise en œuvre de EGP DOIT inclure la prise en charge du voisin indirect.

Intervalles d'interrogation : RFC-904, page 10
L'intervalle entre les retransmissions de commande Hello et l'intervalle entre les retransmissions de Poll DEVRAIT être configurable mais il DOIT y avoir une valeur minimum définie.

L'intervalle auquel une mise en œuvre va répondre aux commandes Hello et aux commandes Poll DEVRAIT être configurable mais il DOIT y avoir une valeur minimum définie.

Accessibilité du réseau : RFC-904, page 15
Une mise en œuvre DOIT par défaut ne pas fournir la liste externe des routeurs dans d'autres systèmes autonomes ; seule la liste interne des routeurs avec les réseaux qui sont accessibles à travers ces routeurs devrait être incluse dans un paquet Réponse/Indication de mise à jour. Cependant, une mise en œuvre PEUT choisir de fournir une option de configuration activant la fourniture de la liste externe. Une mise en œuvre NE DOIT PAS inclure dans la liste externe les routeurs qui ont été appris à travers la liste externe fournie par un routeur dans un autre système autonome. Une mise en œuvre NE DOIT PAS renvoyer un réseau au système autonome dont il a été appris, c'est-à-dire qu'il DOIT faire une coupure au niveau d'un système autonome.

Si plus de 255 routeurs internes ou externes doivent être spécifiés dans une mise à jour d'accessibilité de réseau, les réseaux accessibles à partir des routeurs qui ne peuvent pas figurer sur la liste DOIVENT être fusionnés en un seul sur la liste des routeurs. Le choix des routeurs retenus pour cela DEVRAIT être configurable par l'utilisateur, mais DEVRAIT par défaut revenir à l'adresse de source de la mise à jour EGP générée.

Une mise à jour EGP contient une série de blocs de numéros de réseau, et chaque bloc contient une liste de numéros de réseau accessibles à une distance particulière à travers u n routeur particulier. Si plus de 255 réseaux sont accessibles à une distance particulière à travers un routeur particulier, ils sont séparés en plusieurs blocs (qui ont tous la même distance). De même, si plus de 255 blocs sont nécessaires pour faire la liste des réseaux accessibles à travers un routeur particulier, l'adresse du routeur est inscrite autant de fois que nécessaire sur la liste pour inclure tous les blocs dans la mise à jour.

Mises à jour non sollicitées : RFC-904, page 16
Si un réseau est partagé avec son homologue, une mise en œuvre DOIT envoyer une mise à jour non sollicitée à l'entrée dans l'état Up si le réseau de source est le réseau partagé.

Accessibilité de voisinage : RFC-904, page 6, 13-15
Le tableau de la page 6 qui décrit les valeurs de j et k (les seuils de voisinage amont et aval) est incorrect. Il est reproduit correctement ici :

Nom	Actif	Passif	Description

j	3	1	seuil de voisinage amont
k	1	0	seuil de voisinage aval

La valeur de k en mode passif est aussi incorrecte à la page 14 de la RFC-904. Les valeurs entre parenthèses devraient se lire :

(j = 1, k = 0, et T3/T1 = 4)

Pour une optimisation, une mise en œuvre peut se retenir d'envoyer une commande Hello quand il faut une commande Poll. Si une mise en œuvre le fait, elle DEVRAIT fournir une option configurable par l'utilisateur pour désactiver cette optimisation.

Temporisateur d'interruption : RFC-904, pages 6, 12, 13
Une mise en œuvre EGP DOIT inclure la prise en charge du temporisateur d'interruption (tel qu'exposé au paragraphe 4.1.4 de la RFC-904). Une mise en œuvre DEVRAIT utiliser le temporisateur d'interruption dans l'état Idle *(repos)* pour produire automatiquement un événement Start *(démarrage)* pour relancer la machine du protocole. Les valeurs recommandées sont P4 pour une erreur critique (Interdiction administrative, Violation de protocole et Problème de paramètre) et P5 pour toutes les autres. Le temporisateur d'interruption NE DEVRAIT PAS être démarré lorsqu'un événement Stop a été initié à la main (comme à travers un protocole de gestion de réseau).

Commande Cease *(cesser)* reçue dans l'état Idle : RFC-904, page 13
Lorsque la machine d'état EGP est dans l'état Idle, elle DOIT répondre aux commandes Cease par Cease-ack.

Mode d'interrogation Hello : RFC-904, page 11
Une mise en œuvre EGP DOIT inclure la prise en charge à la fois des modes d'interrogation actif et passif.

Messages d'acquisition de voisins : RFC-904, page 18
On a noté que Hello et Poll Intervals ne devraient être présents que dans les messages Request et Confirm. Donc la longueur d'un message d'acquisition de voisin EGP est 14 octets pour un message Request ou Confirm et 10 octets pour un message Refuse, Cease ou Cease-ack. Des mises en œuvre NE DOIVENT PAS envoyer 14 octets pour des messages Refuse, Cease ou Cease-ack mais DOIVENT permettre des mises en œuvre qui envoient 14 octets pour ces messages.

Numéro de séquence : RFC-904, page 10
Les paquets de réponse ou d'indication reçus avec un numéro de séquence qui n'est pas égal à S DOIVENT être éliminés. Le numéro de séquence envoyé S DOIT être incrémenté juste avant le moment où une commande Poll est envoyée et à aucun autre moment.

F.2 Protocole d'informations d'acheminement - RIP

F.2.1 Introduction

RIP est spécifié dans [ROUTE:3]. Bien que RIP soit toujours assez important dans l'Internet, il est remplacé dans les applications sophistiquées par des IGP plus modernes tels que ceux décrit ci-dessus. Un routeur qui met en œuvre RIP DEVRAIT mettre en œuvre RIP Version 2 [ROUTE:14], car il prend en charge les chemins CIDR. Si un réseautage d'accès occasionnel est utilisé, un routeur qui met en œuvre RIP DEVRAIT mettre en œuvre Demand RIP [ROUTE:14].

Une autre utilisation commune de RIP est le protocole de découverte d'un routeur. Le paragraphe 4.3.3.10 évoque brièvement ce sujet.

F.2.2 Description rapide du protocole

Traitement des changements de topologie : [ROUTE:3], page 11
Une mise en œuvre de RIP DOIT fournir les moyens de fixer une durée d'expiration pour un chemin. Comme les messages sont parfois perdus, des mises en œuvre NE DOIVENT PAS invalider un chemin sur la base d'une seule mise à jour manquée.

Les mises en œuvre DOIVENT par défaut attendre six fois l'intervalle de mise à jour avant d'invalider un chemin. Un routeur PEUT avoir des options de configuration pour modifier cette valeur.

Discussion
Il est important pour la stabilité de l'acheminement que les routeurs dans un système RIP autonome utilisent une valeur de temporisation similaire pour l'invalidation des routes, et donc, il est important qu'une mise en œuvre revienne par défaut à la valeur de temporisation spécifiée dans la spécification de RIP.

Cependant, cette valeur de temporisation est trop conservatrice dans les environnements où la perte de paquet est raisonnablement rare. Dans de tels environnements, un gestionnaire de réseau peut souhaiter être capable de diminuer la période de temporisation au profit d'une accélération de la récupération sur défaillance.

MISE EN ŒUVRE

Il existe un mécanisme très simple qu'un routeur peut utiliser pour satisfaire aux exigences d'invalidation rapide des routes après la fin de la temporisation. Chaque fois qu'un routeur explore le tableau d'acheminement pour voir si une des routes est arrivée à expiration, il note aussi l'age des routes les plus récemment mises à jour qui ne sont pas encore arrivées à expiration. En soustrayant cet age de la durée de la temporisation, on a le délai dans lequel le routeur doit à nouveau explorer le tableau à la recherche de routes arrivées à expiration.

Horizon partagé : [ROUTE:3], page 14-15

Une mise en œuvre de RIP DOIT mettre en œuvre l'horizon partagé, schéma utilisé pour éviter les problèmes causés par l'inclusion de routes dans les mises à jour envoyées au routeur duquel elles ont été apprises.

Une mise en œuvre de RIP DEVRAIT mettre en œuvre l'horizon partagé *(Split horizon)* avec inversion empoisonnée *(poisoned reverse)*, une variante de l'horizon partagé qui inclut des routes apprises d'un routeur envoyées à ce routeur, mais avec leur métrique réglée à infini. À cause des redondances d'acheminement qui peuvent découler de la mise en œuvre de l'horizon partagé avec inversion empoisonnée, des mises en œuvre PEUVENT inclure l'option de choisir si l'inversion empoisonnée est effective. Une mise en œuvre DEVRAIT limiter le temps pendant lequel elle envoie des routes inverses avec une métrique infinie.

MISE EN ŒUVRE

Chacun des algorithmes suivants peut être utilisé pour limiter la durée pendant laquelle l'inversion empoisonnée est appliquée à une route. Le premier algorithme est plus complexe mais fait un travail précis de limitation de l'inversion empoisonnée aux seuls cas où elle est nécessaire.

Le but des deux algorithmes est de s'assurer que l'inversion empoisonnée est faite pour toute destination dont le chemin a changé dans la dernière durée de vie de chemin (normalement 180 s) sauf si on peut être sûr que le chemin précédent utilisait la même interface de sortie. La durée de vie de chemin est utilisée parce qu'elle est la durée pendant laquelle RIP va rester sur un vieux chemin avant de le déclarer périmé.

Les intervalles de temps (et les variables dérivées) utilisées dans les algorithmes suivants sont comme suit :

Tu : Temporisation de mise à jour ; nombre de secondes entre les mises à jour RIP.
Par défaut, normalement de 30 secondes.

Rl : Durée de vie du chemin, en secondes. C'est la quantité de temps pendant laquelle un chemin est présumé être bon, sans requérir de mise à jour. Par défaut, normalement de 180 s.

Ul : Perte de mise à jour ; nombre de mises à jour consécutives qui ont été perdues ou qui sont défaillantes à mentionner un chemin avant que RIP ne le supprime. Ul est calculé comme étant (Rl/Tu)+1. Le +1 est pour tenir compte du fait que la première fois que le ifcounter est décrémenté sera moins de Tu secondes après qu'il ait été initialisé. Normalement, Ul sera 7: (180/30)+1.

In : valeur à laquelle régler ifcounter lorsqu'une destination vient d'être apprise. Cette valeur est Ul-4, où le 4 est la temporisation /30 de collecte de déchets de RIP.

Le premier algorithme est :

- Un compteur est associé à chaque destination ; on l'appelle ifcounter. L'inversion empoisonnée est faite pour tout chemin dont l'ifcounter de destination est supérieur à zéro.
- Après l'envoi d'une mise à jour régulière (non déclenchée ou en réponse à une demande) tous les ifcounters qui ne sont pas à zéro sont décrémentés de un.
- Lorsque est créée un chemin pour une destination, son ifcounter est réglé comme suit :
  - Si le nouveau chemin se substitue à un chemin valide, et si le vieux chemin utilisait une interface de sortie différente (logique), le ifcounter est alors réglé à Ul.
  - Si le nouveau chemin se substitue à un chemin périmé, et si le vieux chemin utilisait une interface de sortie différente (logique), le ifcounter est alors réglé à MAX(0, Ul - INT(secondes que le chemin est périmé/Ut).
  - Si il n'y avait pas de chemin précédent vers cette destination, le ifcounter est réglé à In.
  - Autrement, le ifcounter est réglé à zéro
- RIP entretient aussi un temporisateur, appelé ci-dessous le resettimer. L'inversion empoisonnée est faite sur tous les chemins chaque fois que resettimer n'est pas arrivé à expiration (sans considération des valeurs de l'ifcounter).
- Lorsque RIP démarre, redémarre, se réinitialise, ou a autrement un tableau d'acheminement supprimé, il règle le resettimer pour s'arrêter dans Rl secondes.

Le second algorithme est identique au premier, excepté que :

- La règle qui établit le ifcounter à des valeurs différentes de zéro est changée pour toujours la régler à Rl/Tu, et

- Le resettimer est éliminé.

Mises à jour déclenchées : [ROUTE:3], page 15-16; page 29
Les mises à jour déclenchées (appelées aussi mises à jour flash) sont un mécanisme pour notifier immédiatement les voisins d'un routeur lorsque le routeur ajoute ou supprime des chemins ou change leur métrique. Un routeur DOIT envoyer une mise à jour déclenchée lorsque des chemins sont supprimés ou que leur métrique est augmentée. Un routeur PEUT envoyer une mise à jour déclenchée lorsque des chemins sont ajoutés ou que leur métrique est diminuée.

Comme les mises à jour déclenchées peuvent causer une redondance d'acheminement excessive, les mises en œuvre DOIVENT utiliser le mécanisme suivant pour limiter la fréquence des mises à jour déclenchées :
(1) Lorsque un routeur envoie une mise à jour déclenchée, il règle un temporisateur à une durée aléatoire entre une et cinq secondes dans le futur. Le routeur ne doit pas générer de mises à jour déclenchées supplémentaires avant l'expiration de ce temporisateur.
(2) Si le routeur devait générer une mise à jour déclenchée durant cet intervalle, il établirait un fanion indiquant qu'il souhaite une mise à jour déclenchée. Le routeur enregistre aussi la mise à jour déclenchée désirée.
(3) Lorsque le temporisateur de mise à jour déclenché arrive à expiration, le routeur vérifie le fanion de mise à jour déclenchée. Si le fanion est mis, le routeur envoie alors une seule mise à jour déclenchée qui inclut tous les changement qui ont été enregistrés. Le routeur supprime alors le fanion, et comme une mise à jour déclenchée a été envoyée, recommence cet algorithme.
(4) Le fanion est aussi supprimé chaque fois qu'une mise à jour régulière est envoyée.

Les mises à jour déclenchées DEVRAIENT inclure tous les chemins qui ont changé depuis la mise à jour régulière la plus récente (non déclenchée). Les mises à jour déclenchées NE DOIVENT PAS inclure des chemins qui n'ont pas changé depuis la plus récente mise à jour régulière.

Discussion
L'envoi de toutes les routes, qu'elles aient changé récemment ou non est inacceptable dans les mises à jour déclenchées parce que la taille énorme de nombreux tableaux d'acheminement Internet pourrait autrement résulter en un considérable gâchis de bande passante lors des mises à jour déclenchées.

Utilisation d'UDP : [ROUTE:3], page 18-19.
Les paquets RIP envoyés à une adresse de diffusion IP DEVRAIENT avoir leur TTL initial réglé à un.

Noter que pour se conformer au paragraphe 6.1 du présent mémoire, un routeur DEVRAIT utiliser la somme de contrôle UDP dans les paquets RIP qu'il génère, DOIT éliminer les paquets RIP reçus avec une somme de contrôle UDP invalide, mais NE DOIT PAS éliminer les paquets RIP reçus simplement parce qu'ils ne contiennent pas de somme de contrôle UDP.

Considérations sur l'adressage : [ROUTE:3], page 22
Une mise en œuvre RIP DEVRAIT prendre en charge les chemins d'hôte. Si elle ne le fait pas, elle DOIT (comme décrit à la page 27 de [ROUTE:3]) ignorer les chemins d'hôtes dans les mises à jour reçues. Un routeur PEUT enregistrer les chemins d'hôtes ignorés.

L'adresse spéciale 0.0.0.0 est utilisée pour décrire un chemin par défaut. Un chemin par défaut sert de chemin de dernier recours (c'est-à-dire, quand il n'existe pas de chemin pour le réseau spécifié dans le tableau d'acheminement). Le routeur DOIT être capable de créer une entrée RIP pour l'adresse 0.0.0.0.

Processus d'entrée – Réponse : [ROUTE:3], page 26
Lors du traitement d'une mise à jour, les vérifications de validité suivantes DOIVENT être effectuées :
- La réponse DOIT provenir de l'accès UDP 520.
- L'adresse de source DOIT être sur un sous-réseau directement connecté (ou sur un réseau sans sous-réseau directement connecté) pour être considérée comme valide.
- L'adresse de source NE DOIT PAS être une des adresses du routeur.

Discussion
Certains réseaux, supports, et interfaces permettent à un nœud d'envoi de recevoir des paquets qu'il diffuse. Un routeur ne doit pas accepter ses propres paquets comme mises à jour d'acheminement valides et les traiter. La dernière exigence empêche un routeur d'accepter ses propres mises à jour d'acheminement et de les traiter (en supposant qu'ils ont été envoyés par un autre routeur sur le réseau).

Une mise en œuvre NE DOIT PAS remplacer un chemin existant si la métrique reçue est égale à la métrique existante sauf conformément à l'heuristique suivante.

Une mise en œuvre PEUT choisir de mettre en œuvre l'heuristique suivante pour répondre à la situation ci-dessus. Normalement, il ne sert à rien de changer le chemin vers un réseau d'un routeur à un autre si tous deux ont publié la

même métrique. Cependant, le chemin publié par un des routeurs peut être dans un processus de fin de temporisation. Au lieu d'attendre l'expiration de la durée de vie du chemin, on peut utiliser le nouveau chemin après l'écoulement d'une durée spécifiée. Si cette heuristique est mise en œuvre, il DOIT attendre au moins la moitié de la durée d'expiration avant d'installer le nouveau chemin.

F.2.3 Questions spécifiques

Fermeture de RIP

Une mise en œuvre de RIP DEVRAIT fournir une fermeture gracieuse en utilisant les étapes suivantes :

(1) Le processus d'entrée est terminé,

(2) Quatre mises à jour sont générées à des intervalles aléatoires entre deux et quatre secondes. Ces mises à jour contiennent tous les chemins qui ont été précédemment annoncés, mais avec des changements de métrique. Les chemins qui avaient été annoncés avec une métrique infinie devraient continuer d'utiliser cette métrique. Les chemins qui avaient été annoncées avec une métrique non infinie devaient être annoncés avec une métrique de 15 (infini - 1).

Discussion

La métrique utilisée pour ce qui précède devrait bien être de 16 (infini) ; la régler à 15 est un bricolage pour éviter de froisser certains vieux hôtes qui enregistrent le protocole RIP. Un tel hôte va (à tort) interrompre une connexion TCP si il essaye d'envoyer un datagramme sur la connexion alors que l'hôte n'a pas de chemin pour la destination (même si la période pendant laquelle l'hôte n'a pas de chemin ne dure que quelques secondes alors que RIP choisit un chemin de remplacement pour la destination).

Horizon partagé RIP et chemins statiques

L'horizon partagé DEVRAIT être appliqué aux chemins statiques par défaut. Une mise en œuvre DEVRAIT fournir un moyen de spécifier, pour chaque chemin statique, que l'horizon partagé ne devrait pas s'appliquer à ce chemin.

F.3 Protocole de passerelle à passerelle - GGP

Le protocole de passerelle à passerelle est considéré comme obsolète et NE DEVRAIT PAS être mis en œuvre.

Remerciements

O that we now had here
But one ten thousand of those men in England
That do no work to-day!

What's he that wishes so?
My cousin Westmoreland? No, my fair cousin:
If we are mark'd to die, we are enow

To do our country loss; and if to live,
The fewer men, the greater share of honour.
God's will! I pray thee, wish not one man more.
By Jove, I am not covetous for gold,
Nor care I who doth feed upon my cost;
It yearns me not if men my garments wear;
Such outward things dwell not in my desires:
But if it be a sin to covet honour,
I am the most offending soul alive.
No, faith, my coz, wish not a man from England:
God's peace! I would not lose so great an honour
As one man more, methinks, would share from me
For the best hope I have. O, do not wish one more!
Rather proclaim it, Westmoreland, through my host,
That he which hath no stomach to this fight,
Let him depart; his passport shall be made
And crowns for convoy put into his purse:
We would not die in that man's company
That fears his fellowship to die with us.
This day is called the feast of Crispian:

Oh ! si nous avions tout à l'heure ici
Seulement dix mille de ces anglais
Qui chôment aujourd'hui !

Qui donc exprime ce vœu ?
Mon cousin Westmoreland ? Non, mon beau cousin :
Si nous sommes marqués pour mourir, notre pays n'a pas besoin
de perdre plus d'hommes que nous ne sommes ;
Et si nous devons vivre, moins nous serons,
Plus grande sera pour chacun notre part d'honneur.
C'est la volonté de Dieu ! Ne souhaite pas un homme de plus,
Je t'en prie. Par Jupiter, je ne suis pas affamé d'or,
Et je m'inquiète peu qu'on vive à mes dépens ; je regrette peu
Que d'autres usent mes vêtements ; ces choses extérieures
Ne font pas partie de mes désirs ; mais si convoiter l'honneur
Est un péché, je suis l'âme la plus pécheresse qui existe.
Non ma foi, mon cousin, ne souhaitez pas un anglais de plus.
Paix de Dieu ! Je ne voudrais pas m'exposer à perdre
Un si grand honneur, qu'un homme de plus pourrait partager
Avec moi, pour la plus grande de mes espérances.
Oh ! Ne souhaite pas un homme de plus ! Proclame plutôt
A travers mon armée, Westmoreland, que celui-là peut partir,
Qui ne porte pas de cœur au combat ; on lui donnera passeport
Et des écus dans sa bourse pour le voyage :
Nous ne voudrions pas mourir dans la compagnie d'un homme
Qui craindrait de mourir en notre compagnie.
Ce jour est appelé la fête de Saint Crépin :

He that outlives this day, and comes safe home,	Celui qui survivra à ce jour, et retournera sauf chez lui
Will stand a tip-toe when the day is named,	Se dressera sur ses orteils lorsque ce jour sera nommé
And rouse him at the name of Crispian.	Et se sentira transporté à ce nom de Crépin.
He that shall live this day, and see old age,	Celui qui survivra à ce jour et arrivera à la vieillesse
Will yearly on the vigil feast his neighbours,	Chaque année à la veille de cette fête invitera ses amis
And say 'To-morrow is Saint Crispian:'	Et leur dira “Demain est la Saint Crépin”.
Then will he strip his sleeve and show his scars.	Il retroussera alors ses manches et montrera ses cicatrices
And say 'These wounds I had on Crispin's day.'	Et dira “J’ai reçu ces blessures à la Saint Crépin”
Old men forget: yet all shall be forgot,	Les vieux oublient, et tout doit tomber dans l’oubli
But he'll remember with advantages	Mais lui se souviendra encore avec bonheur
What feats he did that day: then shall our names.	Des exploits de ce jour là. Et nos noms seront familiers
Familiar in his mouth as household words	A leurs bouches, comme ceux de leurs proches
Harry the king, Bedford et Exeter,	Les mémoires du roi Henry, Bedford et Exeter
Warwick and Talbot, Salisbury and Gloucester,	Warwick et Talbot, Salisbury et Gloucester,
Be in their flowing cups freshly remember'd.	Seront saluées à coupes débordantes.
This story shall the good man teach his son;	Le brave homme dira cette histoire à son fils.
And Crispin Crispian shall ne'er go by,	Et saint Crépin et Crépinien ne s'évoqueront jamais,
From this day to the ending of the world,	Depuis ce jour jusqu’à la fin du monde
But we in it shall be remember'd;	Sans qu’avec lui notre souvenir ne soit rappelé ;
We few, we happy few, we band of brothers;	Nous si peu, notre heureuse petite armée, notre bande de frères,
For he to-day that sheds his blood with me	Car celui qui ce jour a répandu son sang avec moi
Shall be my brother; be he ne'er so vile,	Sera mon frère, si vil qu’il soit.
This day shall gentle his condition:	Cette journée le fera noble
And gentlemen in England now a-bed	Et les gentilshommes d’Angleterre qui sont au lit en ce moment
Shall think themselves accursed they were not here,	Se considèreront comme maudits de n’avoir pas été ici
And hold their manhoods cheap whiles any speaks	et tiendront leur noblesse pour peu de chose en entendant
That fought with us upon Saint Crispin's day.	Les récits de ceux qui ont combattu à la Saint Crépin.

-- William Shakespeare, Henry V, acte 4 scène III (traduction Emile Montaigut)

Le présent mémoire a été produit par le groupe de travail Exigences des routeurs de l’IETF. Un mémoire comme celui-ci est nécessairement le travail de plus de gens que ce qu’on pourrait énumérer ici. Un grand nombre de fabricants, de gestionnaires de réseau, et autres experts provenant de la communauté de l’Internet ont graciéusement contribué de leur temps et de leur expérience pour améliorer la qualité du présent mémoire. L’éditeur souhaité etendre ses sincèree remerciements à eux tous.

L’éditeur actuel souhaite étendre sa sincère gratitude et ses remerciements à l’éditeur d’origine du présent document ; Philip Almquist. Sans le travail de Philip, à la fois comme premier éditeur et comme président du groupe de travail, le présent document n’aurait pu être produit. Il souhaite aussi exprimer sa profonde et sincère gratitude à l’éditeur précédent, Frank Kastenholz. Frank a changé le document d’origine d’une collection d’informations en une utile description de la technologie IP - selon ses propres mots, une "photographie" de la technology en 1991. On peut seulement espérer que cette photographie, celle de la technologie de 1994, est aussi claire.

Philip Almquist, Jeffrey Burgan, Frank Kastenholz, et Cathy Wittbrodt ont chacun écrit les chapitres majeurs du présent mémoire. D’autres ont fait des contributions majeures au document. Ce sont Bill Barns, Steve Deering, Kent England, Jim Forster, Martin Gross, Jeff Honig, Steve Knowles, Yoni Malachi, Michael Reilly, et Walt Wimer.

Des textes supplémentaires ont été fournis par Andy Malis, Paul Traina, Art Berggreen, John Cavanaugh, Ross Callon, John Lekashman, Brian Lloyd, Gary Malkin, Milo Medin, John Moy, Craig Partridge, Stephanie Price, Yakov Rekhter, Steve Senum, Richard Smith, Frank Solensky, Rich Woundy, et d’autres qui sont restés anonymes.

Une partie du texte du présent mémoire a été (honteusement) copié dans des documents antérieurs, en particulier de la RFC-1122 de Bob Braden et du groupe de travail sur les exigences pour les routeurs, et de la RFC-1009 de Bob Braden et Jon Postel. Le travail de ces auteurs précédents doit être salué.

Jim Forster était coprésident du groupe de travail sur les exigences pour les routeurs durant ses dernières réunions, et il a largement contribué en permettant un bon début des tavaux du groupe. Jon Postel, Bob Braden, et Walt Prue ont aussi contribué au succès en fournissant de bons conseils avant la première réunion du groupe. Plus tard, Phill Gross, Vint Cerf, et Noel Chiappa ont tous fourni de bons avis et leur soutien.

Mike St. Johns a coordonné les interactions du groupe de travail avec la communauté de la sécurité, et Frank Kastenholz a coordonné les interactions du groupe de travail avec le domaine de la gestion de réseau. Allison Mankin et K.K. Ramakrishnan ont fourni leur expertise sur les questions de contrôle d’encombrement et d’allocation de ressource.

Bien plus de personne qu’il n’est possible d’en faire la liste ont participé aux déliberations du groupe de travail des exigences pour les routeurs, soit par messagerie électronique soit en assistant aux réunions. Cependant, les efforts de Ross Callon et Vince Fuller pour résoudre la difficile question du choix de route et de la fuite de route doivent être

particulièrement salués.

L’éditeur remercie son employeur, Cisco Systems, pour lui avoir permis de passer le temps nécessaire pour produire l'état de la situation de 1994.

Adresse de l’éditeur

L’adresse de l’éditeur actuel du présent document est
Fred Baker
Cisco Systems
519 Lado Drive
Santa Barbara, California 93111
USA
téléphone :+1 805-681-0115
mél : fred@cisco.com