Réseaux : IP, Protocoles et Communications

HAI404I : Licence 2 Informatique

Anne-Elisabeth Baert – baert@lirmm.fr

2021-2022

Table des matières

1 Chapitre 6 – Gestion des Erreurs sur IP

Présentation du Problème

Constat : l'acheminement de datagrammes dans l'Internet se fait **au mieux**, sans garantie de livraison.

Action : Si un routeur ne peut acheminer un datagramme alors il tente d'en avertir l'hôte expéditeur.

ICMP (*Internet Control Message Protocol*)est le protocole d'annonce d'erreurs.

Il est utilisé par le logiciel de la couche réseau (IP), non seulement dans le sens *routeur* → *hôte*, mais aussi par des hôtes ou routeurs pour des utilisations *détournées* comme par exemple des tests d'accessibilité.

Remarque : noter qu'un routeur ne peut annoncer l'erreur qu'à l'hôte source (seule adresse figurant dans le paquet IP). C'est le logiciel de la couche réseau sur l'hôte source qui traite l'erreur ou la fait suivre à l'application correspondante.

Types d'Erreurs

Les exemples suivants permettent de voir l'étendue des dégâts et de constater qu'annoncer une erreur à la source n'est pas toujours la bonne solution.

Un routeur peut se trouver dans une situation désagréable comme :

— pas de chemin vers l'adresse destination dans sa table de routage,
— l'hôte de destination n'existe pas (détection par le dernier routeur),
— le réseau par lequel il veut acheminer est en panne ou congestionné,
— obligation de détruire le datagramme, par exemple, suite à une erreur du code de contrôle, ou à une durée de vie dépassée.

ICMP intègre aussi la possibilité d'obtenir diverses informations entre routeurs, entre hôtes ou les deux. Une des utilisations les plus connues est

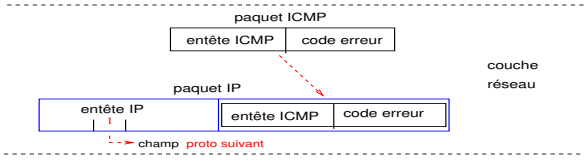
- la demande d'écho et
- la réponse associée à cette demande,

par le logiciel `ping`.

Où Traiter ?

ICMP fait partie de IP. C'est-à-dire que dans la couche *réseau* il y a du logiciel et des paquets ICMP au même titre que IP.

Les paquets ICMP sont acheminés dans des datagrammes IP. On en déduit l'encapsulation suivante :



Noter que le champ *protocole suivant* dans l'entête IP est utilisé pour désigner le *suivant*, soit dans la couche transport (tcp, udp, autre), soit ICMP, avec une valeur différente bien sûr.

Noter aussi qu'une erreur dans l'adresse source du datagramme va aboutir à la perte de l'annonce d'erreur.

Rappel du Paquet IP

Un rappel de la forme d'un paquet IP

octet 1	octet 2	octet 3	octet4
Vers. | lg. ent.	type service	lg. paquet	
Identification		drapeaux	place frag.
dur e vie	proto. suiv.	contr le ent te	
adresse IP source			
adresse IP destination			
options ...			
...			bourrage
Donn es ...			

Paquet ICMP

•L'entête de tout paquet ICMP est de la forme :

type	code	contrôle
8bits	8bits	16 bits

•Le champ *type* désigne le type d'erreur. **Exemples** :

8	demande d'écho
0	réponse écho

3	destination inaccessible
4	congestion
11	dépassement durée de vie

•Le champ *code* comporte une information complétant le type d'erreur. **Exemples** :

0	réseau inaccessible

1	hôte inaccessible
6	réseau inconnu

•Dans tous les cas d'erreur, ICMP ajoute dans la donnée les 64 premiers bits du datagramme ayant provoqué l'erreur. Plus généralement, la donnée permet de compléter plus explicitement les indications de l'entête.

1.1 Erreurs Liées au Routage

Destination Inaccessible

• Lorsqu'un routeur ne peut pas délivrer ou faire suivre un datagramme, il construit un message d'erreur ICMP, avec dans le champ *type* la valeur 3, dans le champ *code* une valeur de 0 à 12, calcule la somme de contrôle et ajoute au paquet ICMP les 64 premiers bits du datagramme, extrait l'adresse de l'hôte source Hs puis détruit ce datagramme non routable.

Ce paquet est encapsulé dans un datagramme IP, contenant en source le routeur expéditeur et en destinataire Hs, avec dans le champ *protocole suivant* le code 1, désignant ICMP.

L'hôte source peut ainsi analyser *plus sérieusement* la cause du rejet et faire suivre à l'application un retour d'erreur.

•Noter qu'un routeur peut faire suivre des datagrammes **sans se rendre compte** que la destination est inaccessible.

Exercice : Donner deux exemples démontrant ce phénomène, l'un concernant un hôte destinataire (penser à ethernet par exemple pour répondre), l'autre concernant un routeur destinataire.

Dépassement de Durée de Vie

Associer une durée de vie au datagramme IP permet de faire en sorte qu'un datagramme ne puisse circuler indéfiniment dans l'Internet sans arriver à destination.

Est-ce possible ? Oui, pour des erreurs de routage provoquant des aller-retours d'un datagramme entre deux routeurs, chacun ayant malheureusement une interprétation erronée des informations de routage, ou pire, une boucle de routage entre plusieurs routeurs (voir le chapitre sur le routage).

Solution : le champ *durée de vie* contient dans sa forme la plus simple (l'actuelle, dans IPV4), le nombre maximum de routeurs que le datagramme peut traverser. Chaque datagramme IP se voit appliquer le principe suivant :

Algorithme TTL

Appelons *TTL* le champ *durée de vie* du datagramme IP.

L'hôte source du datagramme initialise ce champ à une valeur déterminée, dans le logiciel de la couche réseau.

Chaque routeur applique ensuite l'algorithme suivant :

```
TTL - - ;
(TTL == 0) expédier message ICMP (dépassement
détruire datagramme ;
```

Les Échos

La demande d'écho dans ICMP permet aux routeurs de savoir si les routeurs voisins sont actifs ou non. Lorsqu'un routeur reçoit un message ICMP de *demande d'écho*, il doit répondre par un message ICMP de *réponse écho*.

Cette caractéristique est utilisée non seulement entre routeurs, mais aussi entre hôtes pour tester leurs présences, comme nous l'avons déjà vu pour le logiciel `ping`.

Noter que ping visualise la valeur du champ *Durée de Vie* et affiche aussi le temps d'aller-retour du datagramme.

Exercice : Pour quelles raisons est-ce que la durée d'aller-retour du premier datagramme dans `ping` est souvent supérieure aux suivants ?

Et si ICMP Provoquait une Erreur ?

Remarque Importante : Tout paquet ICMP est encapsulé puis routé dans un datagramme IP. Dès lors, ce datagramme peut subir les mêmes avatars que tout datagramme IP, perte, congestion, abandon.

Les pertes et erreurs engendrent des pertes et des erreurs (d'après Rez O.)...

Dans leur sagesse, les concepteurs ont décidé qu'on ne devait construire un message ICMP relatif à un datagramme contenant déjà un message ICMP...

Conséquence : voici encore une raison pour laquelle des protocoles comme `TCP` doivent inclure des garanties, ajouter des délais, tenir actifs les circuits virtuels, et alerter les applications avec des moyens complémentaires.

1.2 Utilisation Détournée

Détournemesnt de TTL

Le comportement des routeurs relativement au champ *Durée de Vie*, permet d'en faire une utilisation détournée, afin de déterminer le chemin d'accès à un hôte.

La commande `traceroute` applique un algorithme dont le principe est :

```
HdestNonAtteint = vrai ;
TTL=0 ;
(HdestNonAtteint) TTL + + ;
expédier (datagramme, Hdest) ;//de
 (expéditeur erreur ICMP) ;
(réponse de Hdest) HdesNonAtteint
```

Algorit

Question : Est-ce vraiment un chemin correct ?

Analyse de Traceroute

Exercice : prendre le schéma de réseau suivant et montrer que l'algorithme précédent peut afficher des chemins faux ou pire, inexistants. On suppose que S cherche un chemin vers D et que les Rx représentent des routeurs.

On peut définir
— *faux* par : le résultat donné ne sera pas un chemin suivi par un paquet,
— *inexistant* par : le chemin affiché contient au moins un arc (ou un sommet) inexistant.

1.3 Compléments Datagramme IP

Fragmentation

Un datagramme IP peut être *fragmenté*, c'est-à-dire découpé en morceaux, sur un ou même plusieurs routeurs, en fonction des caratéristiques des réseaux que le routeur interconnecte.

Chaque fragment circule comme un datagramme indépendant, donc peut suivre un chemin différent d'un autre fragment.

Conséquences :
— le réassemblage ne peut se faire que sur le hôte destinataire final,
— dans la couche IP qui doit attendre la réception de tous les fragments, tout en acceptant entre temps d'autres datagrammes,
— chaque fragment doit contenir les informations nécessaires à l'identification du datagramme d'origine et à l'insertion correcte du fragment dans ce datagramme.

2 Chapitre 7 – Routage

2.1 Introduction au Routage

Le Problème du Routage

Déterminer en fonction de l'adresse réseau du destinataire final d'un datagramme le prochain destinataire.

On peut compléter très légèrement le tableau déjà vu dans le cas d'un hôte quelconque sur un réseau local.

Destination	Contact	Masque	Interface
201.202.203.0	direct	255.255.255.192	eth0
autre	201.202.203.1	0.0.0.0	eth0

Contact indique soit le prochain routeur, soit une destination sur le même réseau.

Problèmes

— Cette table peut prendre des dimensions gigantesques dans le cas d'un routeur censé connaître l'ensemble des destinations de l'Internet ou d'un sous-ensemble.

— Comment construire cette table ?

Routage Statique

Le routage statique constitue une solution simple à la construction de la table : elle est figée et modifiée uniquement par une intervention d'un administrateur.

Cette solution est parfaitement bien adaptée à un réseau local avec un seul routeur assurant la connectivité vers le monde extérieur.

L'utilisation d'une route par défaut permet de passer rapidement le relai d'un hôte à un routeur, d'un routeur à un autre routeur. On comprend mieux pourquoi des incohérences sont possibles. Et il y a pire, par exemple des boucles...

Routage Dynamique

Dans le cas d'un routeur reliant plusieurs réseaux, un routage adaptatif ou dynamique permettra de tenir compte de :

— l'infrastructure des réseaux connectés,

— l'arrivée et la réparation de pannes, la création de nouveaux liens,

— la charge des réseaux (congestions, oscillations),

— de la qualité de service requise, etc.

Une distribution *intelligente* des adresses de réseaux permettrait de réduire la taille des tables de routage (voir routage hiérarchique). Hélas, ce n'est pas le cas, du moins ceci n'a pas été fait systématiquement, dans l'Internet.

Connaissance Partielle et Erreurs

Constat : Dans tous les cas, le résultat de l'algorithme de routage est l'adresse du *suivant*.

On espère que les routeurs sont cohérents entre eux, c'est-à-dire que le *suivant* peut continuer à acheminer correctement le paquet. Sinon, on aura des **erreurs de routage**.

Ceci reste vrai même si un routeur connaît le chemin complet, car on ne peut pas **forcer** une décision sur un **autre** routeur ; sauf cas spécifiques de tests de chemins, ce serait néfaste de le forcer.

Traitement des Erreurs - Un Début

On peut empêcher un paquet de vivre indéfiniment dans l'Internet :

Principe :

— Un champ *durée de vie (ttl)* est attaché à chaque paquet (cf. entête du paquet IP) ; il est initialisée par l'hôte source du paquet ;
— Chaque routeur décrémente la durée de vie ;
— Le routeur qui arrive à une durée de vie nulle ou négative détruit le paquet.

Actuellement, la durée de vie est mesurée en *nombre de routeurs traversés*, dit aussi *nombre de sauts*.

ttl- - ;
(ttl $>$ 0)router paquet ;
expédier (erreur routage) à hôte source ;

Classes d'Algorithmes

Plusieurs classes d'algorithmes dynamiques existent en fonction de l'étendue des réseaux reliés.

Globalement, on a besoin de trouver des chemins dans un graphe dynamique. Il faut se poser les questions de

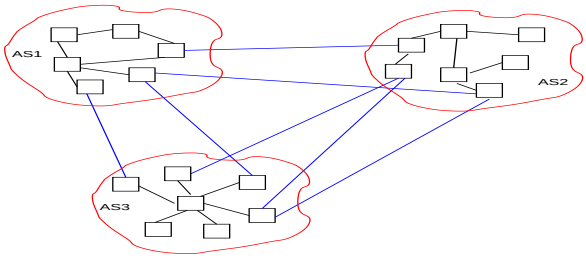
— l'efficacités des algorithmes distribués ou centralisés ;
— convergence, stabilité et cohérence des algorithmes.

Pourquoi plusieurs classes d'algorithmes ? Parce qu'on ne peut plus se contenter d'une organisation centralisée.

Organisation en Systèmes Autonomes de l'Internet

Aujourd'hui, l'Internet est organisé en *Systèmes Autonomes* (AS ci-après), vastes réseaux (en général), administrés chacun par une entité unique.

Chaque AS possède des routeurs *intérieurs* reliant les sous-réseaux entre eux, et des routeurs *extérieurs* reliés à des routeurs extérieurs d'autres AS.



Organisation du Routage dans l'Internet

Chaque AS organise son propre routage interne librement. Plusieurs algorithmes sont connus : RIP, OSPF.

Pour la partie externe, il faut un protocole commun.

Les AS communiquent entre eux par un seul algorithme lié à un seul protocole : aujourd'hui, BGP.

Les routeurs *extérieurs* d'un AS doivent connaître **toutes** les adresses des autres AS afin de constituer un routage cohérent. On insiste sur **toutes**, pas seulement celles des AS et routeurs adjacents.

On pourra voir qu'il est difficile de concilier le fonctionnement extérieur, visible, avec les choix politiques internes. C'est un sujet de recherches actuellement.

2.2 Algorithmes à Vecteurs de Distances

Algorithme à Vecteur de Distance RIP

Principes :

- Diffusion d'informations sur le routage à base de la **distance** ; quelle métrique pour exprimer une distance ? le plus fréquent : *nombre de sauts* ;
- Chaque routeur dispose d'une table contenant des triplets $(destination, numéro\_de\_liaison, coût)$ c'est-à-dire pour telle destination, envoyer sur telle liaison pour tel coût ;
- La table est mise à jour dynamiquement ; il y a diffusion périodique d'informations $(destination, coût)$
- Chaque routeur qui reçoit une information la compare au contenu courant de sa table ;
- Si l'information est *meilleure* il la prend ;
- Sinon, il y a des cas où l'on est obligé d'accepter une information fut-elle *moins bonne*, d'autres où on la rejettera.

Algorithme RIP

table de routage ; des doublets $(dest, cout)$ reçus sur une liaison l_{recue} table de routage toutes les données arrivant $dest_{recue}$ trouvée dans table $l_{recue} == l_{table} c_{table} = c_{recue} + 1$;
$c_{table} > (c_{recue} + 1) l_{table} = l_{recue}$;
$c_{table} = c_{recue} + 1$;
ajouter $(d_{recue}, l_{recue}, c_{recue} + 1)$ dans table

Vecteurs de Distances - Exemple

Initialisation ;
A diffuse $(A, 0)$ sur les liaisons 1 et 3 ;
B diffuse $(A, 1), (B, 0)$ sur 1, 2 et 4 ;
information perdue sur 4 ;
C diffuse $(C, 0), (A, 2), (B, 1)$ sur 2 et 5 ;
D diffuse $(D, 0), (A, 1)$ sur 3 et 6.

Extrait de *Le routage dans l'Internet.*

A			B			C			D			E		
→	l	c	→	l	c	→	l	c	→	l	c	→	l	c
A	loc	0	B	loc	0	C	loc	0	D	loc	0	E	loc	0
			A	1	1				A	3	1			
B	1	1				A	2	2						
						B	2	1						
			C	2	1							C	5	1
												A	5	3
												A	6	2

Justifications

Traitement global : Chaque routeur diffuse périodiquement sa table, ensemble de couples $(destination, coût)$ vers ses voisins adjacents.

Important : Si une nouvelle information sur une destination arrive par la **même** liaison que celle par laquelle on route, alors il faut la prendre, qu'elle soit bonne ou mauvaise.

Pourquoi ? Considérer le point de vue d'un routeur.

> Si A annonce à R qu'une destination X est atteinte pour un coût 10 et que le même, A annonce ensuite un coût différent (envisager 7 puis 12) pour la même destination, R doit accepter ce nouveau coût, sauf si entre temps il a un meilleur chemin.

Remarques

Attention : Nous avons vu **une** simulation possible de la diffusion. L'ordre de diffusion peut être totalement différent, avec des résultats différents sur les tables.

On peut démontrer la convergence de cet algorithme si aucune modification (panne, apparition de nouvelles liaisons) n'arrive. Mais les pannes et modifications sont **fréquentes**. On peut montrer que pour toute panne ou apparition de liaison, l'algorithme converge si un nouvel incident n'a pas lieu avant l'aboutissement de la convergence.

Le défaut de fonctionnement reproché à RIP est résumé ainsi :
Les bonnes nouvelles se propagent vite, les mauvaises se propagent doucement.

Traitement d'une Panne dans RIP

Principe du traitement d'une panne : annoncer un coût ∞ pour chaque voisin en panne. Cette détection peut se faire par un outil comme `ping` permettant de tester l'existence d'un hôte.

Exemple : Supposons que la liaison 1 tombe en panne. B corrige sa table avec le triplet $(A, 1, \infty)$ et la diffuse.

Selon l'ordre de propagation de l'information, on peut constater une convergence rapide, ou des phénomènes connus sous le nom de *comtage à l'infini* :

Supposons que C diffuse sa table, **avant** que B ne diffuse la sienne. C diffuse entre autres informations $(A, 2)$, qui lors du traitement sur B va engendrer le triplet $(A, 2, 3)$!!!

D'autres défauts de fonctionnement de cet algorithme ont été répertoriés (voir bibliographie), accompagnés de solutions plus ou moins heureuses. Elles font l'objet de cours ultérieurs.

2.3 Algorithmes à états de liens-OSPF

Algorithmes à États de Liens - Principes

- Chaque routeur possède la topologie complète du réseau ;
- Deux tâches sont accomplies pour arriver à connaître cette topologie :
  - test d'activité de tous les routeurs adjacents : échanges courts (type `ping`)
  - diffusion périodique de l'état des liens : c'est un compte-rendu des communications possibles. L'état est rediffusée *autant que nécessaire* à tous les routeurs participants, avec horodatage (numéro de séquence) des messages.
- Chaque routeur calcule un plus court chemin vers chacun des autres routeurs.

Exemple : *Open Shortest Path First* (OSPF), est actuellement un algorithme à état de liaisons très utilisé à l'intérieur des systèmes autonomes de l'Internet.

Algorithmes à états de liens - Algo

ensembles $(voisins, couts, séquence)$; N : ensemble de nœuds dont le pcc est connu ; $D(v)$: coût parcours vers v nouvelle topologie et arbre des plus courts chemins //Initialisation
N={A} ; D(v) infini (tous) ;
tous les nœuds v v adjacentD(v)=coût(A,v) ;
//mise à jour
tous les nœuds v explorés trouver w extérieur à N tel que D(w) min ;
ajouter w à N ;
tout v adjacent à w D(v)=min(D(v),D(w)+c(W,v)) ;

3 Chapitre 9 Grandes Applications – Serveurs de Noms

Serveurs de Noms

L'application *Serveurs de Noms* est aujourd'hui une base fondamentale de l'Internet.

Elle n'est pas seulement utilisée pour la correspondance nom↔adresse des hôtes, mais plus généralement, pour enregistrer des informations d'administration.

RFC1034 décrit les concepts de base. Suivent un tas de compléments et mises à jour.

Exemple : La messagerie électronique utilise les serveurs de noms pour trouver l'hôte à contacter pour l'acheminement des courriels sur un site.

En effet, que l'adresse électronique d'une personne comporte ou non un nom d'hôte, il y a souvent un (ou quelques) hôte(s) fixé(s) pour l'acheminement dans le domaine de destination.

L'application *serveur de noms* ou DNS, permet de déterminer en fonction d'un nom d'hôte ou du nom de domaine de l'adresse électronique, le serveur à contacter pour la messagerie.

Principes de Fonctionnement

- Les conventions de nommage.
- Le fonctionnement de base, consistant à avoir un serveur de noms dans le domaine local, contenant la correspondance noms↔adresse des hôtes locaux.
- La requête est acheminée vers un autre serveur (souvent un serveur racine), qui est capable de la refaire suivre.
- Le serveur qui connaît la correspondance répond directement au demandeur.

La partie du système d'exploitation prenant en charge la résolution s'appelle *resolver*. Cet outil existe forcément sur chaque hôte, consulte un fichier de base ($/etc/resolv.conf$), contenant au moins :

- le domaine dans lequel il faut chercher les noms simples (ceux donnés sans le caractère point) ;
- l'adresse IP d'au moins un serveur de noms, en général le serveur de noms local.

Exemple

`search info.rmatique.fr nameserver 123.231.111.12`

•Veut dire que tout nom simple, par exemple *hotel* sera recherché (complété) en tant que *hotel.info.rmatique.fr*. Ensuite, ayant le nom complet, la requête sera expédiée vers 123.231.111.12.

•Le serveur de noms reçoit alors un nom interne ou externe et réagit en fonction de ce qu'il trouve dans sa base.

•S'il ne trouve rien, il doit posséder l'adresse d'au moins un autre serveur de noms (souvent un serveur racine) afin de faire suivre la requête.

•Noter que les requêtes aux serveurs de noms utilisent udp.

•Noter aussi qu'il est indispensable d'avoir l'adresse IP du serveur de noms !

Remarques

- Il est possible de gérer des sous-domaines séparémment, c'est-à-dire en fait avoir dans un même domaine, géré par une même autorité administrative, plusieurs serveurs de noms relatifs à plusieurs domaines. Plus généralement, on parlera de *zone d'autorité* pour l'unité gérée par un serveur donné. Mais, un serveur de noms peut gérer plusieurs zones...
- Il est possible aussi d'avoir plusieurs serveurs pour une même zone : penser aux pannes en particulier. Dans ce cas, il y a un serveur dit *primaire* disposant de l'information **origine** et des serveurs *secondaires* disposant d'une **copie**.
- Les requêtes sont vulnérables et DNSSEC décrit dans la RFC2535 permet de construire une *chaine de confiance*.
- La mise à jour dynamique (RFC2136) devient indispensable avec la distribution dynamique des adresses avec des logiciels comme DHCP par exemple.

Plus loin

•Chaque serveur gère des informations comme la correspondance *inverse* (obtention d'un nom à partir d'une adresse), le contact messagerie, la durée de vie des informations, les noms multiples d'un même hôte, etc.

•L'application *serveur* s'appelle *named*, qui lit un fichier de configuration de démarrage (*named.conf*).

Plusieurs bases contiennent ensuite toutes les autres informations. La localisation de ces bases est donnée dans le fichier de démarrage.

De même, les serveurs secondaires ou primaires sont nommés dans ce fichier de démarrage.

Conclusion : le point d'entrée pour aller plus loin est : *named* en plus de DNS.

pgfexternal@did@a@shipout