Notification des erreurs

(Z:\Polys\Internet_gestion_reseau\2.ICMP.fm- 26 septembre 2008 14:00)

PLAN

1. Introduction

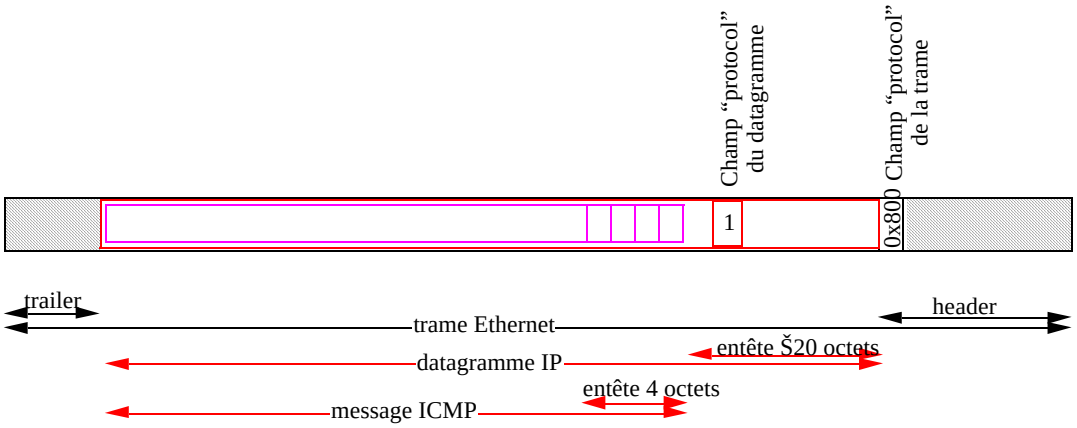
Internet Control Message Protocol
RFC 792, septembre 1981

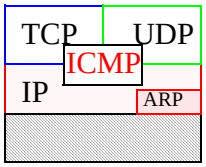
Notification des erreurs lors de la transmission de données (IP, TCP, UDP).
Messages d'administration : contrôle réseau.

. Découvertes de nouveaux mécanismes
. Rappels sur Internet

2. Généralités

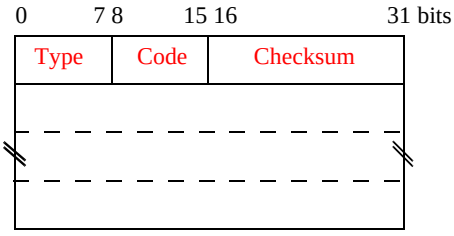
2.1. ICMP et IP et Ethernet





. ICMP gère les événements pour IP, TCP et UDP
. Les messages ICMP sont transportés par des datagrammes IP

2.2. Format général des messages ICMP



Type (8 bits) :
. type du message
. + de 20 types de messages ICMP différents
. 2 grandes catégories :
- message généré à la suite d'une erreur
- message d'administration

Code (8 bits) :
. sous-type du message ICMP

Checksum (16 bits) :
. protège la totalité du message
. procédé de calcul identique à celui de IP, TCP, UDP:
- somme de mots de 16 bits en complément à 1.
. obligatoire

La structure du reste du message dépend du type (1 mot minimum).

2.3. Les types des messages ICMP

Type	Code	Description	Erreur/Administration
0	0	echo reply	A
3		destination unreachable :	
	0	network unreachable	E
	1	host unreachable	E
	2	protocol unreachable	E
	3	port unreachable	E
	4	fragmentation needed and don't seg. bit set	E
	5	source route failed	E
	6	destination network unknown	E
	7	destination host unknown	E
	8	source host isolated	E
	9	destination network administratively prohibited	E
	10	destination host administratively prohibited	E
	11	network unreachable for TOS	E
	12	host unreachable for TOS	E
	13	communication administratively prohibited by filtering	E
	14	host precedence violation	E
	15	precedence cutoff in effect	E
4	0	source quench	E
5		redirect :	
	0	redirect for network	E
	1	redirect for host	E
	2	redirect for TOS and network	E
	3	redirect for TOS and host	E

Les types des messages ICMP (suite)

Type	Code	Description	Erreur/Administration
8	0	echo request	A
9	0	router advertisement	A
10	0	router solicitation	A
11		time to live exceeded :	
	0	during transit	E
	1	during fragment reassembly	E
12		parameter problem:	
	0	bad IP header	E
	1	required option missing	E
13	0	timestamp request	A
14	0	timestamp reply	A
15	0	information request	A
16	0	information reply	A
17	0	address mask request	A
18	0	address mask reply	A
30	0	traceroute	A
31	0	datagram conversion error	A
32	0	mobile station redirection	A
33	0	IPv6 station localisation request	A
34	0	IPv6 station localisation response	A
35	0	mobile station recording request	A
36	0	mobile station recording response	A

2.4. Génération conditionnelle d'un message ICMP

Un message ICMP de la catégorie erreur n'est jamais émis en réponse à :
. un message ICMP de la catégorie erreur
. un datagramme contenant une adresse IP broadcast ou multicast
. un datagramme utilisant une trame contenant une adresse de broadcast ou de groupe
. un fragment de datagramme autre que le premier
. un datagramme dont l'origine n'est pas une vraie station (zero address, loopback address)

=> Pour limiter les risques d'avalanche (“broadcast storm”)
ou lorsque l’on ne connait pas la station émettrice.

Par contre un message ICMP de la catégorie administrative de code “response” est généralement émis après un message ICMP de la catégorie administrative de code “request”.

3. Inaccessibilité

3.1. Présentation

Lorsqu’un noeud reçoit un datagramme qu’il ne sait pas acheminer !

Message d’erreur émis par les noeuds (routeurs) d’extrémité :

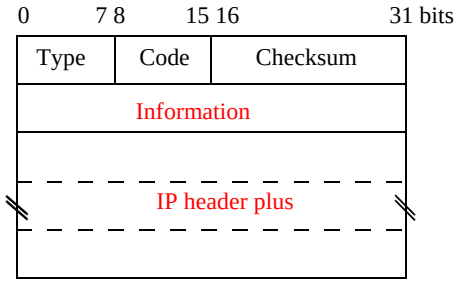
- le noeud ne connaît aucune station ayant l’adresse de destination du datagramme
- le noeud ne gère pas le sous-réseau associé à cette adresse

Le message d’erreur est retourné à l’expéditeur. Le champ de données du message d’erreur contient l’entête du datagramme erroné.

Quelques causes :

- l’adresse du datagramme (ou le n° de port du message) a été corrompue
- la station destinatrice a disparue (en panne, éteinte, déplacée, etc.)
- la station destinatrice n’est pas prête à recevoir un tel datagramme :
  - . le message contenu dans le datagramme ne lui convient pas
  - . par exemple : aucun processus n’est affecté au n° port spécifié
    utilisé par le processus de recherche de route : “traceroute”

3.2. Format des messages d'inaccessibilité



ICMP destination unreachable message

Type
- 3 : le destinataire est inaccessible

Code : la cause de l'inaccessibilité :
- hôte ou réseau inaccessible : émis par un routeur ne pouvant retransmettre le datagramme
- port inaccessible : le numéro de port n'a pas de processus affecté (⇒ traceroute)
- fragmentation nécessaire (⇒ MTU discovery)

Information (32 bits) :
- information générale sur le traitement de l'erreur
- champ parfois inutilisé (=0)

IP header :
permet de reconnaître le paquet ayant généré l'erreur.
- l'entête du datagramme IP ayant provoqué la l'erreur.
- inclusivement, les options de l'entête IP.
- plus (au moins) les 8 premiers octets du champ de données du datagramme IP (c'est-à-dire l'entête du protocole de niveau supérieur).

4. Echo

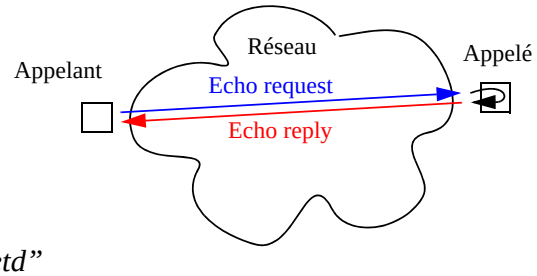
4.1. Principe

Permet de tester la présence (l'accessibilité) d'une station.
- attention : les mécanismes de contrôle d'accès peuvent rendre ce test instable ("firewall gateway")

Mesure du temps de propagation aller-retour (RTT : Round Trip Time)

=> commande *ping*

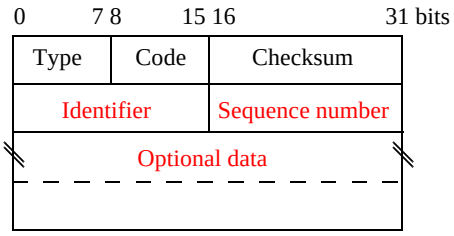
Exemple : `ping athena`
athena is alive
`ping xthena`
no answer from xthena
`ping -s -i 10`



Fonction interne au programme *"inetd"*

Note : les "firewalls" interceptent de tels messages pour lutter contre les "DoS" !

4.2. Format des messages d'écho



ICMP echo message

2 types de message d'écho :
- 0 : réponse
- 8 : demande

Code = 0.

Identifier :
- identifie le client (≈ n° de port)
- le PID du processus sous Unix

Sequence number :
- incrémenté à chaque envoi de messages
- "Ping" permet l'envoi périodique (1s)

⇒ Min/Moy/Max/taux d'erreur.

Optional data :
- "Ping" permet l'envoi de messages de taille quelconque.
- les mêmes données à aller et au retour.

⇒ le RTT est composé du délai de propagation et de la durée d'émission.

5. Durée de résidence dépassée

5.1. Principe

Message généré lorsque le champ TTL d'un datagramme IP arrive à 0, et que le datagramme est détruit.

Utilisé pour connaître la route entre 2 stations :
- l'accessibilité
- la liste des routeurs intermédiaires sur cette route
- la durée de transit entre chaque routeur de cette route

⇒ traceroute

Traceroute :
- utilisation conjointe du champ TTL des datagrammes IP
- et des messages ICMP de type "Time exceeded"

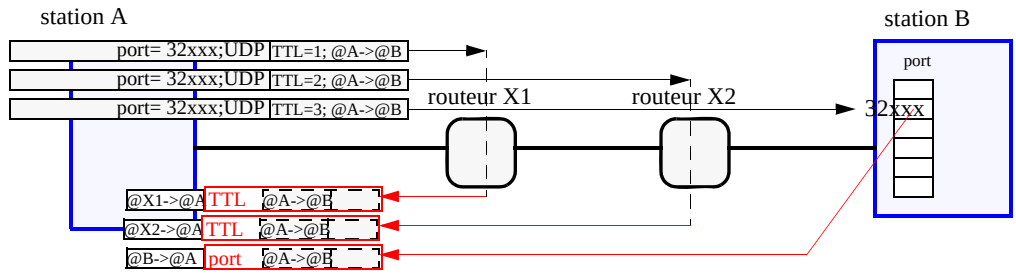
5.2. Traceroute

Principe :

Pour toutes les valeurs de TTL possibles (entre 1 et 255 (64!)) :
. émettre un message UDP avec un TTL donné
. mémoriser l'heure d'émission
Attendre la réception des messages ICMP "time exceeded" :
. mémoriser l'heure de réception
. imprimer le nom du routeur qui a répondu et calculer la valeur du RTT

Le destinataire est atteint lorsque l'on reçoit un message ICMP "port unreachable"
. on a pris soin d'émettre un message UDP sur un port improbablement lié à un processus (n° port >32000 : ou-logique entre 0x8000 et le pid du P[us] traceroute)

Exemple de fonctionnement du "traceroute" :

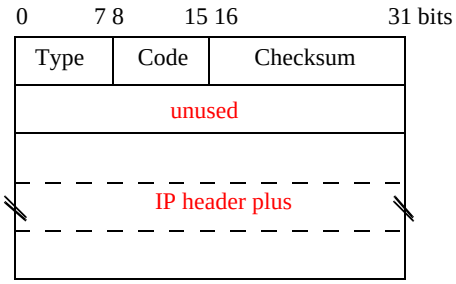


Exemple de résultat :

- `traceroute to B`
  - `X1 5 ms 4 ms 3 ms`
  - `X2 138 ms _`[2] `ms 141 ms`[1] <--- congestion
  - `B 141 ms 150ms 120`[3] `ms`

Pb : fragmentation ("don't fragment"), modification de la route ("source routing")

5.3. Format des messages de durée excessive



ICMP time exceeded message

Type du message :
- 11 : durée limite de vie du datagramme atteinte.

Code :
- 0 = pendant le transit (aux routeurs).
- 1 = pendant le réassemblage (au récepteur)

Unused :
- champ à zéro

IP header :
- l'entête du datagramme IP ayant provoqué l'erreur.
- inclusivement les options de l'entête IP.
- plus (au moins) les 8 premiers octets du champ de données du datagramme IP.

⇒ au minimum : 28 octets

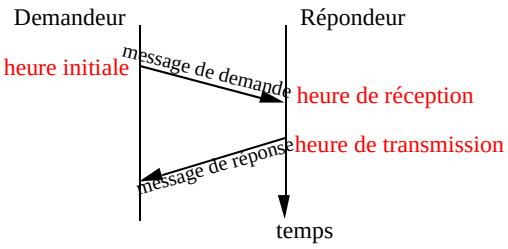
6. Horodatage

6.1. Principe

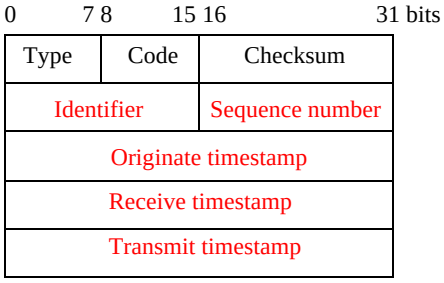
Permet à un noeud de demander et d'obtenir l'heure d'une autre station.
- (obsolète)
- donne l'heure universelle
- en millisecondes, écoulées depuis minuit
- trois horodatages (initial, de réception et de transmission)
- l'heure mais pas la date
- procédé de mise à jour de l'horloge : *timed*

⇒ commande *rdate <host>*

⇒ protocole NTP (Network Time Protocol) [RFC 1305]



6.2. Format des messages d'horodatage



ICMP timestamp message

Type du message :
- 13 : demande d'horodatage
- 14 : réponse d'horodatage

Code :
- 0.

Identifier et Sequence number :
- identifie l'échange

Originate timestamp :
- l'heure d'émission du message de demande.
- nombre de millisecondes depuis minuit UTC ("Coordinated Universal Time"), si le bit de poids fort est nul.

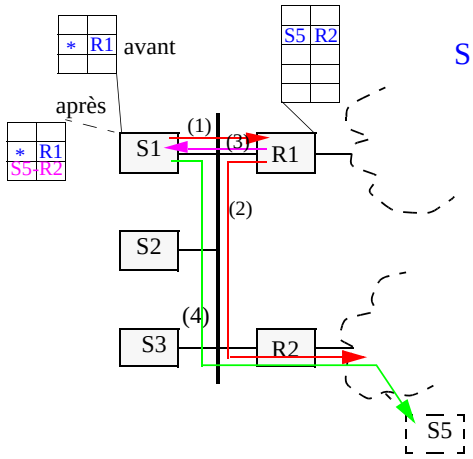
Receive timestamp :
- l'heure de réception du message de demande.

Transmit timestamp :
- l'heure d'émission du message de réponse.
- différent du "receive timestamp" si le temps de traitement est important (en général ils sont égaux).

7. La redirection

7.1. Principe

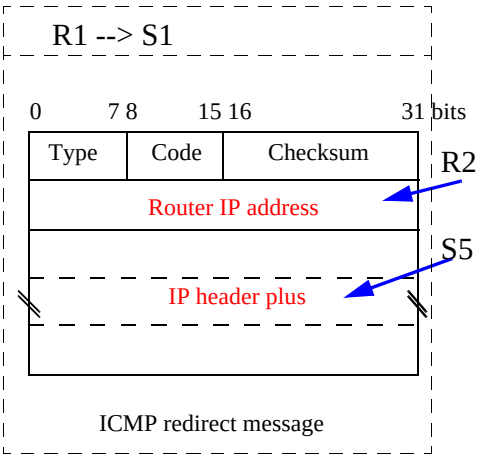
Message échangé entre un routeur et l'émetteur d'un datagramme lorsque le datagramme aurait du être envoyé à un autre routeur.



S1 ⇒ S5

(1) l'émetteur émet son datagramme vers un premier routeur (le routeur par défaut).
(2) le premier routeur route le datagramme vers le 2ème routeur (après consultation de sa table de routage).
(3) le premier routeur, constatant que le datagramme emprunte l'interface d'où il provient, retourne un message ICMP vers l'émetteur.
(4) la prochaine fois, l'émetteur émettra son datagramme directement vers le bon routeur !

7.2. Format des messages de redirection



Type du message :

- 5 :message de redirection d'une route.

Code :

- 0 = redirection pour accéder à une station
- 1 = redirection pour accéder à un sous-réseau
- 2 = redirection pour accéder à une station et pour un TOS
- 3 = redirection pour accéder à un sous-réseau et pour un TOS

Router IP address :

- adresse du routeur qui aurait dû servir.

IP header :

- l'entête du datagramme IP ayant provoqué la redirection.
- inclusivement les options de l'entête IP.
- plus les 8 premiers octets du champ de données du datagramme IP.

En général, les messages de redirection sont générés par les routeurs pour des stations. Les routeurs n'envoient des redirections que pour accéder à des stations.

8. Découverte des routeurs

8.1. Introduction

Message ICMP d'annonce ou de découverte de routeur (au sein d'un (sous-) réseau IP.

=> RFC 1256

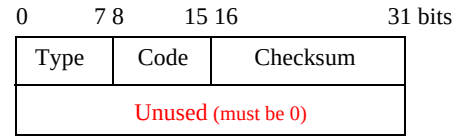
Utilisation :

. lors du démarrage d'une station,
elle diffuse (ou multicaste) un message ICMP de découverte de routeurs
(3 fois max. à 3 secondes d'intervalle)
- un ou plusieurs routeurs peuvent répondre par un message d'annonce.

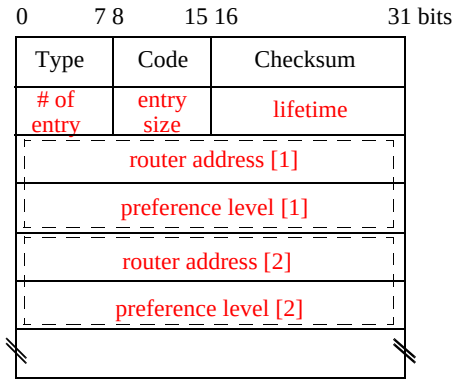
. de plus, les routeurs diffusent périodiquement
(toutes les 10 mn, aléatoirement entre [450-600 s])
des messages publics d'annonce (la durée de vie des infos obtenues est, par défaut, 30 mn).

- surveillance de la présence des routeurs
une station invalide son entrée si elle n'a pas été rafraichie à temps par un message ICMP d'annonce

8.2. Format des messages de découverte des routeurs



ICMP router solicitation message



ICMP router advertisement message

Type du message :
- 10 : message de découverte des routeurs
- 9 : message d'annonce de présence de routeurs (ou de réponse).

Code : 0 !

# of entry :
- nombre d'entrées dans la liste des routeurs

Entry size :
- longueur en mots de 32 bits de chaque entrée
- 2 (toujours)

Lifetime :
- durée de vie (en secondes) des entrées (par défaut : 1800)

Router address :
- l'adresse IP des routeurs connus par le routeur qui émet ce message.

Preference level :
- indique le niveau de préférence du routeur vis-à-vis des autres routeurs pour servir de routeur par défaut.
- la plus forte valeur à la préférence
- 0x80000000 : ce routeur ne doit pas être utilisé comme routeur par défaut.
- 0 : normalement.

9. Masque d'adresse

9.1. Principe

L'adressage IP est adaptée pour permettre d'introduire la notion de subnetID.

. Chaque station est munie d'un mot de 32 bits (appelé "subnet mask"), indiquant la position du champ subnetID dans le champ hostID de l'adresse de la station.
. Toute station a donc besoin de connaître cette information (au même titre que sa propre adresse).
. Par exemple, les stations sans disque utilise les messages ICMP de demande et de réponse de masque d'adressage lors de leur phase d'initialisation (au même titre qu'elles utilisent les messages RARP).

Fonctionnement :
. Diffusion d'un message de demande
. Réception d'un ou plusieurs messages de réponse

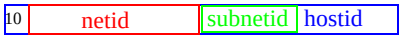
Spécification :
⇒ RFC 950

Note : il existe d'autres techniques pour qu'une station acquière ces informations (par ex. : protocoles BOOTP ou DHCP).

9.2. Subnetting

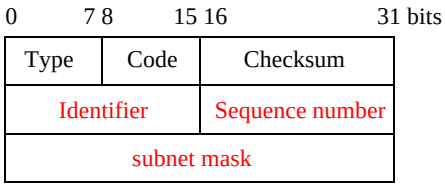
⇒ RFC 1009
Les administrations utilisant des adresses de classe A (ou B) peuvent avoir plusieurs sites ou veulent séparer différents trafics.
Les administrateurs d'un domaine d'adressage voudraient l'organiser en sous-domaines d'adressage.
Cette structuration peut faciliter le routage interne :
- chaque routeur s'occupant que des paquets de son sous-domaine.
- les tables de routage sont plus courtes.

Chaque station est munie d'un "subnet mask".
. indique la frontière entre subnetid et hostid.
. ex. 0x 11111111 11111111 111111**0** 00000000



=> commande *(/etc/)ifconfig -a*
fichier */etc/netmasks*

9.3. Format du message de masque d'adresse



ICMP subnet mask resquest/reply message

Type, code :
- 17/18, 0: demande/réponse du masque d'adresse

Identifier (16 bits) :
. identifie le message
. valeur choisie par le demandeur (par défaut 0)
. même valeur retournée par le répondeur

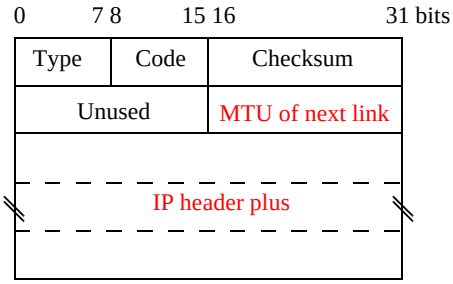
Sequence number (16 bits) :
. idem Identifier (par défaut 0)
. identifie les différentes occurrences du même message

Subnet mask (32 bits) :
. une suite de 1 suivie par une suite de 0.
. identifie la position de la frontière du champ subnet
. 0 dans la requête

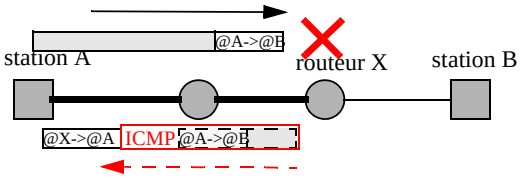
10. Fragmentation

10.1. Format des messages d'erreur de fragmentation

Type, code :
- 3, 4 : erreur lors d'une fragmentation



ICMP fragmentation error message



MTU of next link :
- donne la longueur maximum d'un datagramme pour qu'il puisse franchir la prochaine liaison sans fragmentation.

IP header plus :
- l'entête + options + 8 1er octets du champ data du datagramme IP ayant provoqué cette erreur.

Un message est généré lorsque la fragmentation est nécessaire :
- le MTU de la prochaine liaison est plus petit que la taille du datagramme

mais interdite :
- le datagramme a son bit "do not fragment" positionné.

Le datagramme est détruit par le routeur
Un message ICMP, indiquant que le destinataire est inaccessible à cause de la fragmentation, est émis par le routeur vers l'émetteur.

10.2. La recherche de MTU path

"MTU path discovery"

La fragmentation est coûteuse :
. on peut l'éviter aux routeurs si l'émetteur connaît le MTU path

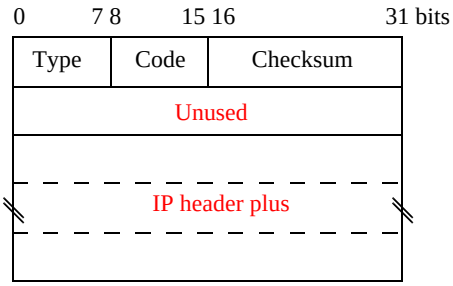
Cette valeur est variable :
. elle dépend du chemin emprunté
. de la topologie du réseau

=> Procédé de recherche du MTU_P (RFC 1191) :
. on émet les paquets avec le bit "don't fragment" positionné.
. le premier MTU_P choisi est le MTU_L de la source.
. tant qu'on reçoit un message d'erreur de fragmentation, on diminue la taille du paquet (en fonction des longueurs proposées).
. une liste ordonnée de longueurs privilégiées est proposée pour augmenter la vitesse de convergence du processus de recherche.
(min: 68, X25: 508, Ethernet: 1492, 802.5: 2002, FDDI: 4352, 802.4: 8166, TK: 17914, 32000, max: 65535)

Régulièrement chaque station tente d'émettre un paquet avec un MTU_P plus grand (par défaut, toutes les 10 mn, ou 2 mn après avoir augmenté le MTU_P).

11. Congestion

11.1. Format des messages d'information de congestion



ICMP source quench message

Type, code :
- 4, 0 : Message généré par une station ou un routeur qui reçoit des datagrammes plus vite qu'il ne peut les traiter.

. un "quench message" à chaque datagramme détruit par congestion.

. optionnel :
- consomme de la bande passante supplémentaire
- inefficace (délai de réaction)
- inéquitable (certaines stations les ignorent)

. utilisation par TCP :
- la fenêtre de congestion (*cwnd*) est réinitilisée à un segment ("slow start").

. pas de message d'augmentation du débit (implicite) !

IP header plus :
- l'entête + options + 8 premiers octets du champ data du datagramme IP ayant provoqué cette congestion.

12. Erreur des paramètres

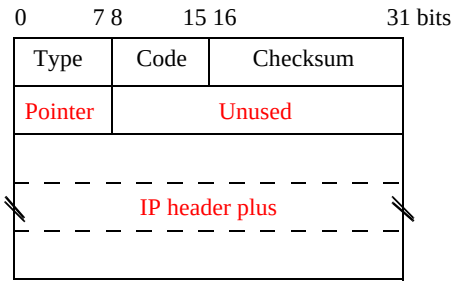
12.1. Introduction

Le format des datagrammes peut être incorrect :
- soit le datagramme a été corrompu lors de la transmission :
  - . la non-détection de la corruption est improbable mais pas impossible,
  - . donc on effectue un contrôle supplémentaire (structurel et sémantique).
- soit le processus émetteur fonctionne mal :
  - . bugs, version non-compatible, etc.,
  - . champ devant contenir une valeur précise ou absence d'une option obligatoire.

Tous les protocoles d'Internet contrôlent les messages qu'ils reçoivent (notamment IP). Mais à part l'action de contrôle/détection, ils entreprennent peu d'action de correction.

L'émetteur sait mieux que celui qui a détecté l'erreur ce qu'il convient de faire :
- retour à l'envoyeur !

12.2. Format des messages d'erreur des paramètres



ICMP parameter error message

Type :

12 : Message généré par une station qui reçoit un datagramme qu'elle ne peut décoder.

Le format du datagramme reçu est incorrect, la valeur de certains champs (de certaines options) est incorrecte.

Code :

- 0 : erreur de paramètre
- -1 : option obligatoire absente

Pointer :

- référence l'octet dans le datagramme qui provoque l'erreur.

Unused :

- 0 à l'émission, ignoré à la réception.

IP header plus :

- l'entête + options + 8 premiers octets du champ data du datagramme IP ayant provoqué cette erreur.

13. Conclusion

Le protocole ICMP est indispensable au bon fonctionnement d'Internet :

. informe des cas d'erreur survenant sur IP, TCP et UDP :
- destination inaccessible
- corruption de messages

. permet d'administrer le réseau :
- redirection de route
- annonce de masque d'adressage, de routeurs, etc.
- vérification de l'accessibilité : commandes *traceroute*, *ping*
- gestion de la fragmentation, calcul du "MTU path"
- mise à l'heure (obsolète !)
- aide au contrôle de congestion (inutilisée !)
- station mobile et IPv6 !

Ils existent d'autres fonctions d'administration rendues par d'autres protocoles tout aussi indispensables au fonctionnement d'Internet.