# Réseaux - Cours 6 Datagrammes IP, ARP et ICMP

Cyril Pain-Barre

IUT Informatique Aix-en-Provence

version du 18/3/2013

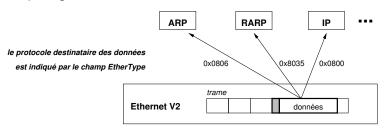
# Datagrammes IP (RFC 791)

# Datagramme IP sur Ethernet V2

• trame Ethernet v2 contenant un datagramme IP (EtherType en Hexa) :

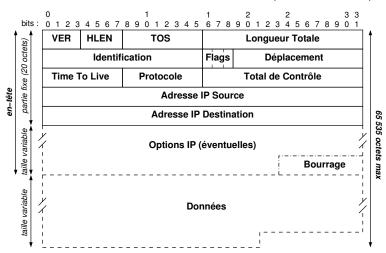
			type	données	
Préambule	Adresse Destination	Adresse Source	08 00	Datagramme IP	CRC

• (dé)multiplexage Ethernet v2 :



### Format du datagramme IP

- en-tête : nombre variable d'octets (multiple de 4)
- données : nombre quelconque d'octets (limité à 65 315)



# Champ Version



- o codé sur 4 bits
- identifie la version du (format du) datagramme
- actuellement, la version est 4 (codée 0100 en binaire)
- dans le datagramme IPv6, ce champ est maintenu et vaut 6
- permet de s'assurer que le datagramme sera correctement interprété

# Champs Adresses



- adresse IP Source : (32 bits)
   identifie l'hôte à l'origine du datagramme
- adresse IP Destination: (32 bits)
   identifie le destinataire final du datagramme



Ces adresses ne sont pas modifiées par les routeurs. Toutefois, en cas de NAT/NAPT (traduction d'adresse), nous verrons qu'une NATbox peut les modifier.

# Champ Longueur d'en-tête (HLEN)

| VED | NLEN | TOS | Longoure Totale | Deplement | | Deplem

- (internet) Header LENgth
- codée sur 4 bits
- indique le nombre de mots de 32 bits de l'en-tête (comprenant les options) :
  - en-tête de 20 à 60 octets
  - 5 < HLEN < 15
- si *HLEN* > 5 alors il y a des options

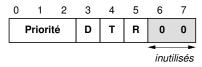
# Champ Longueur Totale



- o codée sur 16 bits
- indique le nombre de total d'octets du datagramme (en-tête + données)
- comprise entre 20 et 65 535

# Champ Type Of Service (TOS)

codé sur 8 bits :





- Priorité : de 0 à 7
  - distinction entre "normal" et "contrôle"
  - routeurs : infos trafic 6 et 7
- bits **D**, **T** et **R** : type d'acheminement **désiré** :
  - D(elay) : délai d'acheminement court
  - T(hroughput) : débit de transmission élevé
  - R(eliability) : grande fiabilité

Priorité		
val <sub>2</sub>	val <sub>2</sub> signification	
000	routine	
001	priority	
010	immediate	
011	flash	
100	flash override	
101	critic	
110	internetwork control	
111	network control	

• le *TOS* est un **souhait** que les routeurs peuvent ignorer

À la fin des années 1990, remplacé par le champ *Differentiated Services*, offrant plus de finesse pour exprimer une qualité de service (QoS) désirée.

# Champ Time To Live (TTL)

- codé sur 8 bits
- indiqué par l'émetteur pour limiter :
  - la "durée de vie" du datagramme en secondes (plus vraiment d'actualité)
  - le nombre de routeurs traversés par le datagramme (nombre de sauts)



- décrémenté par routeurs et stations traitant le datagramme :
  - de 1 à chaque traversée d'un routeur (saut)
  - du temps passé en file d'attente
- si atteint 0, le datagramme est détruit, et l'émetteur est informé par un message ICMP
- évite qu'un datagramme ne circule indéfiniment
- évite que des fragments d'un datagramme ne soient gardés inutilement
- Souvent fixé bien en deçà de la valeur maximale. Par défaut, sur Linux :
  - \$ cat /proc/sys/net/ipv4/ip\_default\_ttl
    64

# Champ Protocole



- o codé sur 8 bits
- joue le même rôle pour IP que le champ Type pour Ethernet
- indique le protocole devant recevoir les données du datagramme (démultiplexage)
- les valeurs de ce champ sont gérées par l'autorité centrale et sont accessibles sur www.iana.org

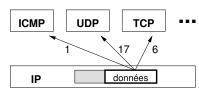
# Champ Protocole

• quelques valeurs officielles :

$val_{10}$	protocole
0	IP
1	ICMP
2	IGMP
6	TCP
17	UDP
89	OSPF

131	VER	HLEN	TOS		Longueur Totale	
		Identification		Flags	Déplacement	
	Time	To Live	Protocole	1	Total de Contrôle	
	Adresse IP Source					
S.		Adresse IP Destination				
ollo variable		Options IP (éventuelles)				
2	2				Bourrage	
	1		D	onnées		

• démultiplexage IP :



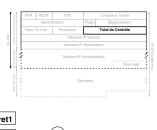
# Champ Total de Contrôle d'en-tête (checksum)

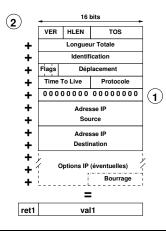
- codé sur 16 bits
- contrôle l'intégrité de l'en-tête uniquement

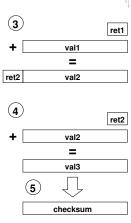
IP ne vérifie pas si les données ont subi des erreurs de transmission : cette vérification est laissée aux couches supérieures

- calculé par l'émetteur
- vérifié lors de la réception (routeurs et destinataire) :
  - stocker le checksum
  - calculer le checksum
  - si différents alors détruire le datagramme
- recalculé et modifié par les routeurs (car modifient au moins le TTL)

## Champ Total de Contrôle d'en-tête : calcul







- 1 mettre checksum à 0
- 2 calculer somme des mots de 16 bits de l'en-tête
- 3 ajouter ret1 à val1
- 4 ajouter ret2 à val2
- 5 checksum = complément à 1 de val3

# MTU et fragmentation

#### Maximum Transfer Unit (MTU) :

 taille maximale des données transportées sur un réseau physique, appelée aussi charge utile (payload)

Ethernet : 1 500 octetsToken Ring : 4 ou 16 Ko

• X.25: 128 octets recommandés (max 255)

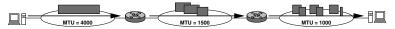
• SMDS: 9 188 octets

• Frame Relay: 1 600 octets

• ... et donc des datagrammes



• IP **fragmente** tout datagramme plus grand que le MTU du réseau qui doit le transporter :



 chaque fragment est un datagramme acheminé indépendamment (peut suivre une route différente des autres fragments) et peut être à son tour fragmenté

# Champs pour la fragmentation



- identification: valeur identifiant le datagramme d'origine.
   Les fragments d'un datagramme ont les mêmes identification, IP Source, IP Destination et Protocole
- bit *D(on't Fragment)* : le datagramme ne doit pas être fragmenté (détruit et message ICMP si impossible)
- bit M(ore): à 0 si datagramme non fragmenté ou si c'est le dernier fragment; sinon vaut 1
- déplacement (Offset) : déplacement × 8 est la position absolue (ou numéro) du premier octet de données de ce datagramme dans le datagramme d'origine.
  - Vaut 0 si pas de fragmentation

# Réassemblage



- réalisé par le destinataire final :
  - met en attente les fragments des datagrammes incomplets
  - les réordonne
  - détruit tous les fragments d'un datagramme si le TTL de l'un d'eux passe à 0 (et envoie un message ICMP à l'émetteur)

### Résolution d'adresse

#### Nécessité de la résolution d'adresse

**situation**: une station/routeur S a un datagramme à transmettre à une station/routeur D **du même réseau**. D est l'adresse IP de la destination finale du datagramme (remise directe) ou celle d'un routeur obtenue par consultation de la table de routage (remise indirecte)



- la transmission doit se faire en utilisant le service de la couche hôte-réseau (réseau physique)
- la couche hôte-réseau (niveau trame) n'utilise pas les adresses IP (niveau réseau) mais des adresses physiques (adresses MAC)
- la **résolution d'adresse** est le mécanisme permettant à une station d'obtenir l'adresse physique (de l'interface/carte réseau) d'une station possédant une certaine adresse IP dans le même réseau

# Méthodes possibles pour la résolution d'adresse

#### résolution directe

- l'adresse physique est déterminée comme une fonction de l'adresse IP
- méthode simple à mettre en œuvre si les adresses physiques sont configurables

#### interrogation d'un serveur

- un serveur est chargé de collecter les adresses physiques et IP des hôtes du réseau
- les stations interrogent le serveur pour résoudre les adresses
- méthode souvent utilisée lorsque le réseau ne permet pas la diffusion
- mais la résolution n'est plus possible si le serveur devient injoignable...

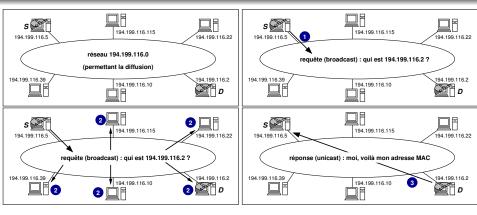
### Résolution dynamique par ARP

Pour les réseaux permettant la diffusion, la méthode privilégiée est ARP (Address Resolution Protocol), défini dans la RFC 826.

- ARP a l'avantage d'être à la fois dynamique et décentralisé :
  - les changements d'association adresse IP/adresse MAC sont automatiquement et rapidement pris en compte
  - aucun serveur n'est nécessaire et une panne d'une station n'a aucun impact global
- ARP a été originellement défini pour IP et Ethernet.
   Mais il est plus général et peut être utilisé sur tout type de réseau permettant la diffusion, pour le compte de différents protocoles réseau (dont IP)

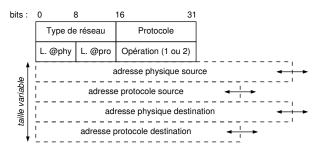
#### IP sur Ethernet utilise systématiquement ARP

# Principe de la résolution par ARP



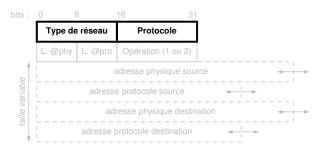
- S envoie en broadcast une requête ARP signifiant qu'il souhaite obtenir l'adresse physique correspondant à D
- 2 la requête est reçue et traitée par toutes les stations du réseau
- seule la station d'adresse D répond en envoyant en unicast à S une réponse ARP contenant l'adresse physique demandée

# Format du datagramme ARP



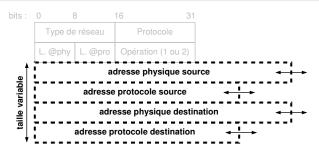
- la taille du datagramme ARP dépend des protocoles concernés :
  - la taille (en octets) des adresses physiques (comme Ethernet) est indiquée par le champ Longueur adresses physiques (L. @phy)
  - la taille (en octets) des adresses protocole (comme IP) est indiquée par le champ *Longueur adresses protocole* (L. @pro)
- les requêtes et les réponses ont le même format ; le champ *Opération* indique s'il s'agit d'une requête (*Opération* vaut 1) ou d'une réponse (*Opération* vaut 2)

# Datagramme ARP : type de réseau et protocole



- Type de réseau précise le réseau physique utilisé et donc le type d'adresse recherchée
- *Protocole* précise la couche réseau utilisée et donc le type d'adresse à partir duquel la résolution doit être opérée
- les valeurs que peuvent prendre ces champs sont définies par l'IANA (www.iana.org):
  - Type de réseau vaut 0x0001 pour Ethernet
  - Protocole vaut 0x0800 pour IP

# Datagramme ARP: adresses



- qu'il s'agisse d'une requête ou d'une réponse :
  - adresse physique source contient l'adresse physique (MAC) de l'émetteur du datagramme
  - adresse protocole source contient son adresse réseau (IP)
- adresse physique destination :
  - dans requête : inconnue (à 0, soit 00:00:00:00:00:00 pour Ethernet)
  - dans réponse : adresse MAC du destinataire
- adresse protocole destination contient l'adresse réseau (IP) du destinataire (dans une requête, c'est l'adresse à résoudre)

# Requête ARP pour IP sur Ethernet V2

#### adresses de l'émetteur

IP: 194.199.116.5

ethernet: 08:00:05:0e:ab:51

#### adresses de la cible

IP: 194.199.116.2

ethernet (recherchée) : 08:00:07:5c:10:0a

Trame Ethernet V2 (en hexadécimal) :



• Requête ARP (en binaire) :

## Réponse ARP pour IP sur Ethernet v2

#### émetteur de la réponse

IP: 194.199.116.2

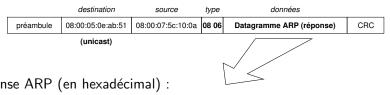
ethernet: 08:00:07:5c:10:0a

#### destinataire de la réponse

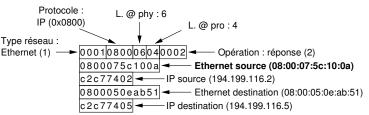
IP: 194.199.116.5

ethernet: 08:00:05:0e:ab:51

Trame Ethernet V2 (en hexadécimal) :



Réponse ARP (en hexadécimal) :



# Optimisations d'ARP

- cache (mémoire temporaire) ARP obligatoire stocké sur les hôtes :
  - ullet contient une liste d'associations  $\prec$  adresse MAC, adresse IP  $\succ$
  - évite d'émettre une nouvelle requête lorsque l'association a déjà été obtenue
  - une association a une durée de vie (TTL) limitée, variable selon les systèmes (1 min, 20 min, . . . )
  - chaque fois qu'une association est confirmée, sa durée de vie est remise à son max
  - les associations dont la durée de vie expire sont supprimées
- traitement de la requête :
  - les requêtes étant envoyées en broadcast, toutes les stations les traitent
  - or elles incluent l'adresse MAC et l'adresse IP de l'émetteur
  - en recevant une requête, les stations peuvent mettre à jour leur cache avec les infos sur l'émetteur
- possibilité d'émission d'une requête ARP fictive si changement de carte (et donc d'adresse MAC) :
  - en plaçant sa propre adresse IP comme celle recherchée
  - personne ne répondra, mais en recevant la requête les stations peuvent mettre à jour leur cache

#### ICMP:

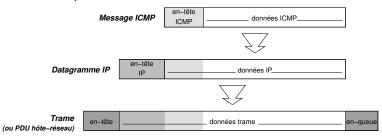
Internet Control and Error Message Protocol (RFC 792, 1122, 950, 1256)

#### Statut et rôle d'ICMP

- ICMP est un module obligatoire d'IP
- qui assure deux fonctions principales :
  - rendre compte d'un problème réseau
  - tester l'accessibilité d'une machine
- les messages ICMP sont de deux natures :
  - les messages d'erreur : suite à une erreur constatée sur un datagramme (qui entraîne le plus souvent sa destruction)
  - les messages d'interrogation/information : messages divers contribuant au (ou informant sur le) bon fonctionnement des équipements

## Transport des messages ICMP

• Les messages ICMP sont encapsulés dans des datagrammes IP (champ *Protocole* vaut 1) :



- mais IP utilisant ICMP pour envoyer ses messages d'erreur, les deux se situent au même niveau!
- les protocoles de transport (TCP et UDP) utilisent aussi ICMP pour certaines erreurs (sur la station destinataire)

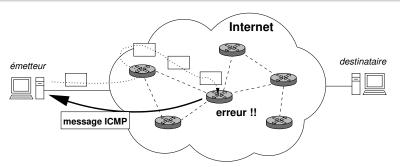


### IP: un service non fiable

- les causes rendant impossible la remise d'un datagramme peuvent être nombreuses :
  - panne de ligne de transmission, ou d'un processeur
  - destinataire déconnecté
  - TTL insuffisant
  - congestion des routeurs intermédiaires
  - checksum erroné
  - mauvaises tables de routage
  - ...
- néanmoins, les erreurs ne sont pas toutes détectables!!

lorsqu'une erreur (dans la remise d'un datagramme) est détectée par un routeur ou la station destinataire, un rapport d'erreur (message ICMP) est envoyé à l'émetteur (d'origine) du datagramme

# Les messages d'erreur ICMP



- le message inclut au moins 28 octets (l'en-tête et les 64 premiers bits) du datagramme ayant causé l'erreur, contenant une partie de l'en-tête du protocole de niveau supérieur, notamment afin de déterminer le processus émetteur pour l'informer
- il n'est envoyé qu'à l'émetteur du datagramme qui, parfois, n'y peut rien (exemple d'une mauvaise table de routage d'un routeur intermédiaire)
- aucun message d'erreur n'est envoyé si le datagramme en cause contient un message d'erreur ICMP!

# Messages ICMP

- destinés au module ICMP du logiciel IP
- pas de précaution particulière pour acheminement : possibilité de perte, duplication, etc.
- certains messages d'erreur sont envoyés avec une priorité 6 ou 7
- tous les messages ICMP commencent par les 3 champs :

Туре	Code	Total de Contrôle
(8 bits)	(8 bits)	(16 bits)

# Messages ICMP: Type et Code

le champ Type
 (codé sur 8 bits)
 indique la nature du
 message :



valeur (décimal)	signification
0	réponse à une demande d'echo
3	destination inacessible
4	limitation de production à la source
5	redirection (changement de route)
8	demande d'echo
9	annonce de routeur
10	sollicitation de routeur
11	TTL de datagramme expiré
12	problème de paramètre d'un datagramme
13	demande d'horodatage
14	réponse à une demande d'horodatage
17	demande de masque de sous-réseau
18	réponse à une demande de masque de sous-réseau

• le champ *Code* (8 bits) est utilisé pour préciser le message pour le *Type* donné



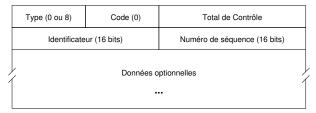
# Messages ICMP : Total de Contrôle

- le champ Total de Contrôle (ou checksum) :
  - est codé sur 16 bits
  - porte sur (contrôle) la totalité du message ICMP
  - calculé de la même manière que le checksum IP



# Test d'accessibilité (PING)

- permet de s'assurer qu'un hôte est joignable et en état de communiquer
- exploité par la commande ping
- échange de messages "ECHO" différenciés par le Type :



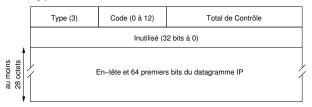
- Type = 8 : demande d'echo (ECHO REQUEST)
- Type = 0: réponse à une demande d'echo (ECHO REPLY)
- le Code est toujours à 0
- Identitificateur permet d'associer la réponse à la demande
- Numéro de Séquence incrémenté à chaque demande

# Exemple de ping (sous Linux)

```
$ ping www.free.fr
PING www.free.fr(212.27.48.10) 56(84) bytes of data.
64 bytes from www.free.fr(212.27.48.10): icmp_seq=1 ttl=107 time=17.6 ms
64 bytes from www.free.fr(212.27.48.10): icmp_seq=2 ttl=107 time=16.6 ms
64 bytes from www.free.fr(212.27.48.10): icmp_seq=3 ttl=107 time=17.2 ms
64 bytes from www.free.fr(212.27.48.10): icmp_seq=3 ttl=107 time=17.2 ms
64 bytes from www.free.fr(212.27.48.10): icmp_seq=4 ttl=107 time=16.7 ms
---www.free.fr ping statistics---
4 packets transmitted, 4 received, 0% packet loss, time 3012ms
rttmin/avg/max/mdev= 16.619/17.094/17.662/0.441 ms
```

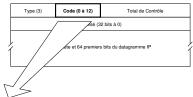
#### Compte rendu de destination inaccessible

- envoyé à l'Adresse Source d'un datagramme IP lorsqu'une station/routeur se rend compte qu'il ne peut pas atteindre sa destination (ordinateur ou processus)
- messages avec Type = 3



- le *Code* précise la nature du problème
- comme les autres messages d'erreur, contient le début du datagramme ayant provoqué l'erreur

# Codes de compte rendu de destination inaccessible

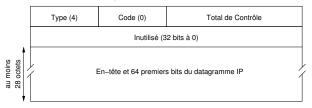


#### plusieurs codes ont été définis :

valeur (décimal)	signification			
0	réseau inaccessible			
1	ordinateur inaccessible			
2	protocole inaccessible			
3	port inacessible			
4	fragmentation nécessaire mais bit Don't Fragment positionné			
5	échec de routage à la source			
6	réseau de destination inconnu			
7	ordinateur de destination inconnu			
8	ordinateur source isolé			
9	communication avec le réseau de destination interdite par l'administrateur réseau			
10	communication avec l'ordinateur destinataire interdite par l'administrateur réseau			
11	réseau inacessible pour le service demandé			
12	ordinateur inaccessible pour le service demandé			

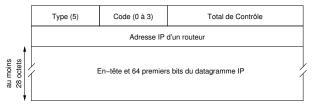
#### Message de contrôle de la congestion

- appelé aussi message de limitation de production de la source (Source Quench)
- envoyé par un routeur à l'émetteur d'un datagramme détruit pour cause de saturation (congestion)
- un message par datagramme détruit
- le message contient le début du datagramme
- le récepteur doit alors réduire ses envois pour cette destination
- format (Type = 4, Code = 0):



#### Message de redirection

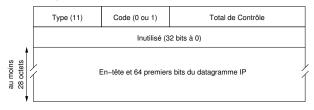
- appelé aussi message de modification de la route
- envoyé par un routeur à l'émetteur (source) d'un datagramme, situé sur le même réseau, si ce routeur n'était pas le meilleur choix
- format (*Type* = 5) :



- le *Code* précise pourquoi il faut changer de route :
  - 0 : changer de route pour atteindre le réseau
  - 1 : changer de route pour atteindre l'ordinateur
  - 2 : changer de route pour le type de service et le réseau
  - 3 : changer de route pour le type de service et l'ordinateur
- Adresse IP d'un routeur est l'adresse du routeur qui aurait dû être utilisé
- l'émetteur devrait en tenir compte et modifier sa table de routage

# Message de durée de vie expirée

• format (Type = 11):



- envoyé à l'émetteur (source) d'un datagramme dans l'un des deux cas suivants :
  - Code = 0 : le TTL a expiré sur un routeur
  - ullet Code =1: le délai de réassemblage des fragments a expiré sur l'hôte destinataire

# Message d'obtention du masque de sous-réseau

- permet à une station de demander son masque de sous-réseau. La demande est adressée (par ordre de préférence) en unicast à un routeur, en multicast 224.0.0.2 (All Routers) ou en broadcast 255.255.255.255
- un routeur du réseau répondra en unicast en indiquant le masque
- format :

Type (17 ou 18)	Code (0)	Total de Contrôle			
Identifi	cateur	Numéro de séquence			
Masque de sous-réseau					

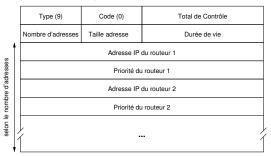
#### où:

- Type indique s'il s'agit d'une demande (17) ou d'une réponse (18)
- Code est à 0
- Identificateur et Numéro de séquence sont utilisés pour associer les réponses aux demandes
- Masque contient le masque dans une réponse (à 0 dans une demande)

Peu utilisé car généralement des protocoles de configuration automatique sont employés, tels que BOOTP et DHCP.

#### Message d'annonce d'un routeur

- envoyé par les routeurs toutes les 10 minutes et destiné aux stations de leurs réseaux (224.0.0.1 ou à défaut en broadcast)
- format (Type = 9, Code = 0):



- Nombre d'adresses indique le nombre d'adresses de routeurs
- Taille adresse ×4 donne la taille en octets des adresses
- Durée de vie indique le temps en secondes de validité des infos (généralement 30 min)
- à chaque adresse est associée une priorité; à utiliser par ordre de priorité croissante

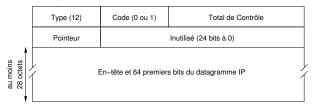
#### Message de sollicitation d'un routeur

- permet à un ordinateur venant de démarrer de demander aux routeurs d'envoyer immédiatement un message d'annonce de routeur
- ce message est envoyé en multicast 224.0.0.2 (All Routers) ou, à défaut, en broadcast
- les routeurs recevant ce message sont censés y répondre dès que possible
- format (Type = 10, Code = 0):

Type (10)	Code (0)	Total de Contrôle				
Réservé (32 bits à 0)						

# Message de problème de paramètre

- envoyé par un routeur ou l'ordinateur destinataire à l'émetteur (source)
   d'un datagramme, lorsque celui-ci comporte une erreur dans un champ
- format (Type = 12):



- le Code précise s'il faut tenir compte du champ Pointeur :
  - 0 : tenir compte du *Pointeur*
  - 1 : ignorer *Pointeur* (qui devrait être à 0). Dans ce cas, il s'agit d'un paramètre manquant (comme une option qui aurait dû être présente)
- Pointeur, si utilisé, indique le numéro de l'octet du datagramme reçu qui comporte l'erreur

# Synchronisation horloges/Estimation temps de transit

- messages échangés entre 2 ordinateurs pour synchroniser leur horloge ou estimer le temps de transit
- format :

Type (13 ou 14)	Code (0)	Total de Contrôle			
Identifi	cateur	Numéro de séquence			
Horodatage de l'émission					
Horodatage de la réception					
Horodatage de la transmission					

- où *Type* indique la nature du message : 13 pour une demande d'horodatage, et 14 pour une réponse à une demande d'horodatage
- le *Code* est toujours à 0
- Identificateur et Numéro de séquence sont utilisés pour associer les réponses aux demandes

# Synchronisation horloges/Estimation temps de transit

- les 3 champs horodatage contiennent les dates, exprimées en millisecondes écoulées depuis le 1<sup>er</sup> janvier 1970 à 0h00 :
  - d'émission de la demande
  - de réception de la demande
  - d'envoi de la réponse



• plusieurs échanges sont nécessaires pour une synchronisation correcte

Le protocole NTP (*Network Time Protocol*) a spécifiquement été développé pour synchroniser son horloge avec des serveurs de référence (et des horloges atomiques).

Sa version actuelle (version 3) est définie par la RFC 1305.