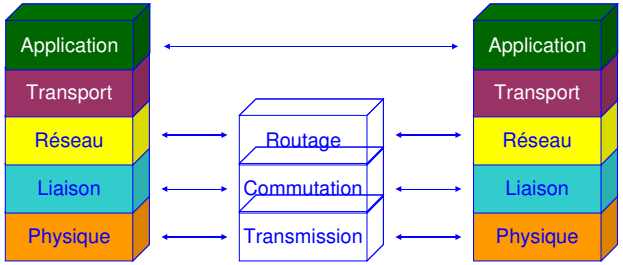
11/2008

Architecture tcp/ip améliorée :-)



1

couche liaison

- liaison:
  - permet la communiccation entre deux machines directement reliées (sur le même lien physique)
  - Adresse de couche liaison: adresse MAC
  -

2

En s'appuyant sur la couche liaison, deux machines directement connectées peuvent communiquer.

Les adresses MAC son les adresses de la couche liaison. Ce sont en général des adresses sur 48 bits (6 octets). Quand la liaison a lieu via un câble ethernet, on les appelle adresses ethernet.

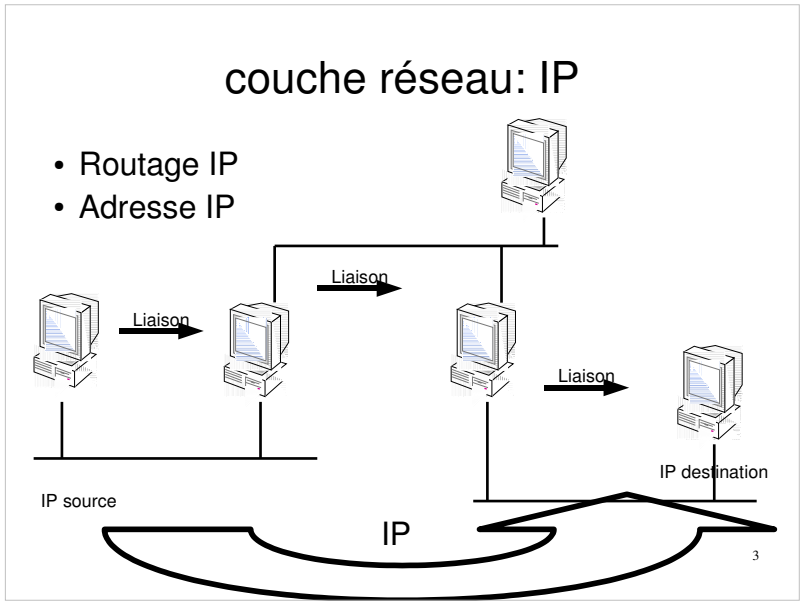
Exemple:

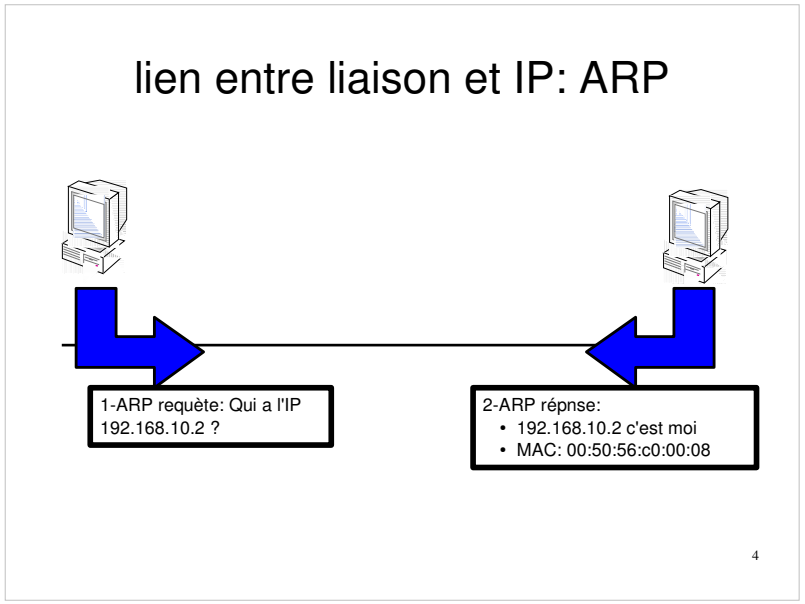
les machines de la salle et une partie des machines de l'étages sont directement reliées

elles peuvent directement communiquer

votre poste et www.google.fr ne sont pas directement reliés

11/2008





apport principal de la couche réseau: permettre la communication
- entre deux machines non directement reliées
  - via un chemin constitué de de machine directement reliées
  - s'appuie sur la couche liaison pour chaque « saut de puce »

Adresse IP : associée à l'interface réseau d'une machine
- un adresse ip correspond à une seule machine
- par contre, une machine peut avoir plusieurs adresses IP

adresse resolution protocol
- obtenir l'adresse ethernet (MAC) connaissant l'adresse IP
- la machine demandeuse diffuse un message sur le réseau auquel la machine concernée répond

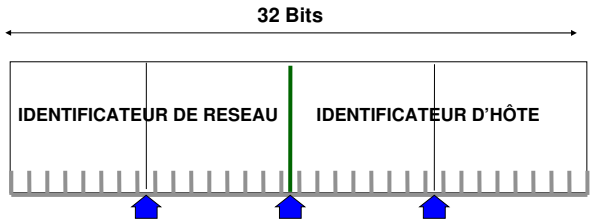
11/2008

Adresse IP

- identifie l'inface réseau d'une machine
- constituée de deux parties :
  - une partie qui identifie le réseau où se trouve la machine
  - une partie qui identifie la machine sur ce réseau
- toutes les machines situées sur le même réseau ont la même partie réseau
- deux machines différentes ne doivent pas avoir la même adresse
- une machine peut avoir plusieurs adresses

5

Adresse IP



W . X . Y . Z

132.109.4.20

6

Adresse IP/Adresse Postale

une adresse postale identifie une boîte aux lettres

deux maisons différentes ne doivent pas avoir la même adresse

deux boîtes aux lettres différentes ne doivent pas avoir la même adresse

une boîtes aux lettre peut avoir plusieurs noms

une maison peur avoir plusieurs boîtes aux lettres

adresse IP: adresse postale

maison : machine

une adresse IP identifie une carte réseau

deux machines différentes ne doivent pas avoir la même adresse

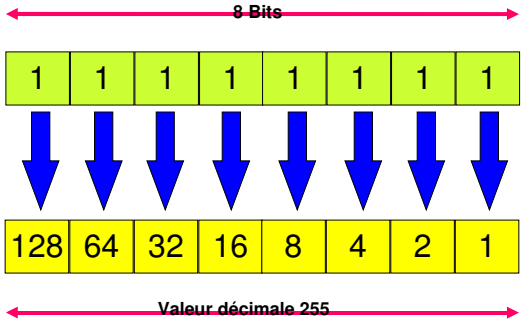
deux cartes différentes ne doivent pas avoir la même adresse

une machine peut avoir plusieurs adresses

une machine peur avoir plusieurs cartes

- L'adresse est sur 32 bits
  - Répartis en 4 séries de 8 bits appelés octets
  - Les octets sont séparés par des points (notation décimale pointée.
  - Format décimal
    - 131.107.3.24
  - Format binaire
    - 10000011 01101011 00000011 00011000
- Les bits de gauches identifient le réseau où est l'hôte
- Les bits de droite identifient l'hôte sur son réseau
- A l'origine, on autorisait une partie réseau :
  - De 8 bits (ou 1 octets) : classe A
  - De 16 bits (ou 2 octets) : classe B
  - De 24 bits (ou 3 octets) : classe C
- De nos jours, on autorise des parties réseaux au bit près.
- Dans la suite, nous allons présenter d'abord le système des classes A, B, ... puis généraliser.

Rappels: Calcul en base 2



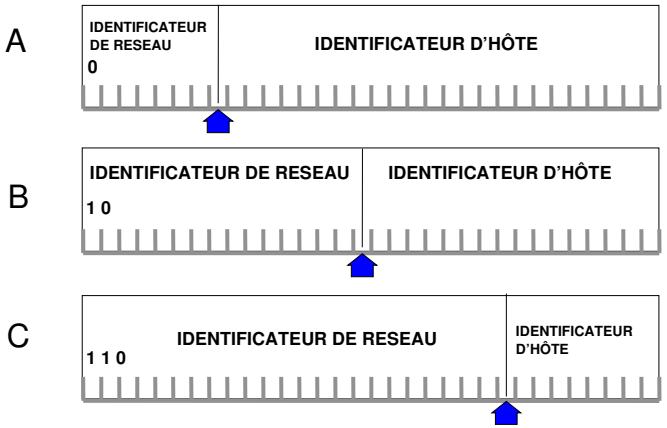
7

Base 10: $n=n_2*10^2+n_1*10^1+n_0*10^0$

base 2: $n=b7*2^7+b6*2^6++b5*2^5+b4*2^4+b3*2^3+b2*2^2+b1*2^1+b0*2^0$

Exemple: 135=ràf

Classes d'adresses



8

❑ **Adresses de classe A**

- Réseaux comportant de très nombreux hôtes, le bit de poids fort est toujours défini à 0 en classe A, les 7 bits suivants définissent l'identificateur de réseau, les 24 autres (3 derniers octets) définissent les hôtes.
  - **126 réseaux et 17 millions d'hôtes par réseau**

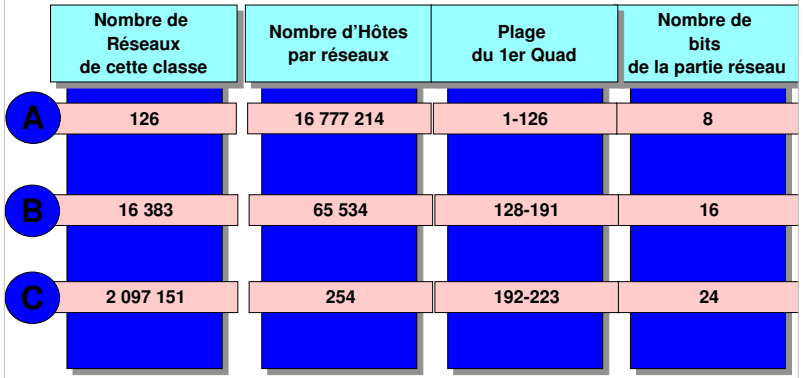
❑ **Adresses de classe B**

- Réseaux de taille moyenne ou grande, les deux bits de poids fort sont toujours définis à 1 0 en classe B, les 14 bits suivants définissent l'identificateur de réseau, les 16 autres (2 derniers octets) définissent les hôtes.
  - **16384 réseaux et 65000 hôtes par réseau**

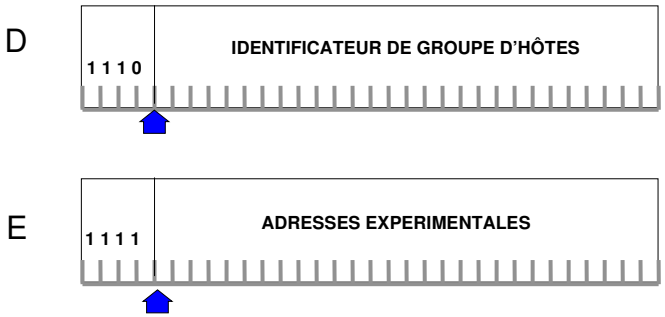
❑ **Adresses de classe C**

- Réseaux comportant très peu d'hôtes, les trois bits de poids fort sont toujours définis à 1 1 0 en classe C, les 21 bits suivants définissent l'identificateur de réseau, les 8 autres (dernier octet) définissent les hôtes.
  - **2 millions de réseaux et 254 hôtes par réseau**

Classes d'adresses



Classes d'adresses



- Adresses de classe D
  - Réseaux utilisés par des groupes multicast.
    - Un groupe multicast peut contenir un ou plusieurs hôtes. Le bits de poids fort sont toujours définis à 1 1 1 0 en classe D, les bits suivants définissent l'identificateur du groupe auquel le client participe. Les opérations en multicast ne comportent pas de bits de réseau ou d'hôtes. Les paquets sont transmis à un sous ensemble d'hôtes sélectionné du sous-réseau.
    - Les hôtes enregistrés pour l'adresse de groupe recevront seuls le paquet. WINS et NetShow utilisent des adresses multicast.
- Adresses de classe E
  - Réseaux expérimentaux
    - (réservé à des usages ultérieurs), les quatre bits de poids fort sont toujours définis à 1 1 1 1.

Adresses réservées

- Réseau: 127.0.0.0/8
- Adresse de bouclage: 127.0.0.1
- Adresse du réseau: partie hôte à 0
- Adresses de diffusion:
  - 255.255.255.255
  - Partie hôte à 255: ce réseau (destination). ex.: 194.199.90.255 (classe C)
- Prévoir une adresse pour routeur par défaut

- Adresse de boucle
  - 127 est réservé aux fonctions de Loopback (127.0.0.1)
- Identificateur 255.255.255.255
  - Si les bits de l'identificateur de réseau et d'hôte sont à 1, l'adresse est interprétée comme adresse de diffusion.
- 255.255.255.255: ce réseau (adresse destination)
- A.255.255.255, B.B.255.255, C.255.255.255:
  - Adresses de diffusion vers tous les hôtes des réseaux en question (dangereux donc souvent désactivé de nos jours)
- Prévoir une adresse pour le routeur par défaut des postes du réseau. Cette adresse doti respecter une convention de site. Par ex.:
  - Partie hôte à 1 (1921.68.2.1)
  - Partie hôte à 249 (194.199.90.249) : université d'Evry
  - Partie hôte à 254 (192.168.0.254) : Freebox
  - ...

Adresses réservées

- Adresse de réseau à zéro (adresse source) :
  - 0.0.0.0: ce réseau (source)
  - 0.x.y.z : l'hôte x.y.z sur ce réseau
- Réseaux privés rfc 1918

- Quand le numéro de réseau est inconnu, on peut le remplacer par 0. Ca peut être le cas d'un hôte en cours d'initialisation.
  - Soit totalement: 0.0.0.0
  - Soit partiellement : 0.x.y.z sur une classe A
- Adresses non routables sur internet (adresses privées, RFC 1918)
  - classe A : 10.x.y.z, classe B : 172.16.y.z à 172.31.y.z
  - classe C : 192.168.y.z
    - où x,y,z représentent un nombre compris entre 0 et 254
  - Ces classes sont conçues pour un usage interne et ne doivent pas circuler sur internet.
  - Ces adresses sont routables au sein d'un site mais pas vers internet
  - Elles sont très utilisées :
    - Pour des machines internes sans accès internet
    - En liaison avec le mécanisme de traduction d'adresses

Masque

- Permet
  - De distinguer la partie réseau de la partie hôte d'une adresse
  - De déterminer si deux hôtes sont sur le même réseau

```
194.199. 90.1          194.199. 90.20
255.255.255.0          255.255.255.0
-------------          -------------
194.199. 90.0          194.199. 90.0
```

13

- Le Masque de Sous-Réseau
  - Une adresse spéciale sur 32 bits, utilisée pour :
    - Cacher une partie de l'adresse IP, ce qui distinguera l'adresse du réseau de l'adresse de l'hôte.
    - Permettre de distinguer si l'adresse de l'hôte destinataire est ou non sur le réseau physique de l'émetteur.
- Chaque hôte doit disposer d'un masque
- Pour trouver l'adresse du réseau, on fait un ET binaire entre l'adresse et le masque.
- Table du ET:
  - 1 ET 1 = 1, 0 ET 1 = 1 ET 0 = 0 ET 0 = 0
  - Traduit en base 10, cela donne :
    - X ET 0 = 0
    - X ET 255 = 255

CIDR/VLSM

- VLSM (Variable Length Subnet Mask) : le masque est défini au bit près
- Permet :
  - Un découpage précis des sous-réseaux d'un site
  - Permet de regrouper des réseaux contigus de classe C en un seul « sur-réseau » : CIDR (Classless Inter Domain Routing)
    - Diminue le nombre d'entrée dans les tables de routage
- Notation /nn avec nn: nombres de bits de la partie réseau du masque

14

VLSM: on s'autorise à découper au milieu des octets.
Correspondance:
- Classe A : 255..0.0.0 ou /8
- Classe B: 255.255.0.0 ou /16
- Classe C: 255.255.255.0 ou /24

Faute d'avoir assez de classe B, on a distribué plusieurs classes C aux entités qui en besoin de quelques milliers d'adresses.
Problème: chacune de ces classes C devait avoir une entrée dans les tables de routages des routeurs. La taille de ces tables de routage menaçaient les performances des routeurs.
Solution: regrouper des réseaux de classe dans des réseaux plus gros. Exemple:
- 4 classes C = un /22 et une entrée dans les tables de routage au lieu de 4.
- 16 classes C = un /20 et une entrée dans les tables de routage au lieu de 16.

Conséquence
- sur le routage: les routeurs et les algorithmes de routage doivent tenir compte (et donc transmettre) de la taille du masque et pas seulement du premier octet de l'adresse.
- Les classes A, B et C deviennent des notions obsolètes

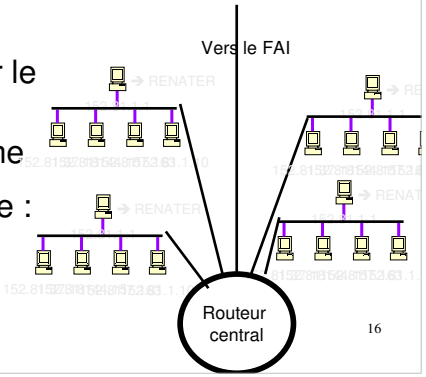
CIDR/VLSM: masque

Classe d'adresse	Bits utilisés pour le masque de Sous-Réseau				Notation Décimale
Classe A /8	11111111	00000000	00000000	00000000	255.0.0.0
Classe B /16	11111111	11111111	00000000	00000000	255.255.0.0
Classe C /24	11111111	11111111	11111111	00000000	255.255.255.0
/23	11111111	11111111	11111110	00000000	255.255.254.0
/18	11111111	11111111	11000000	00000000	255.255.192.0
/15	11111111	11111110	00000000	00000000	255.254.0.0
/9	11111111	10000000	00000000	00000000	255.128.0.0
	Bits 1 à 8	bits 9 à 16	bits 17 à 24	bits 25 à 32	

15

sous-réseaux

- Motivations :
  - Découper un gros réseau en réseaux plus petits
- Solution : partitionner le réseau en plusieurs entités à usage interne
- Architecture classique :



16

Les RFC CIDR:

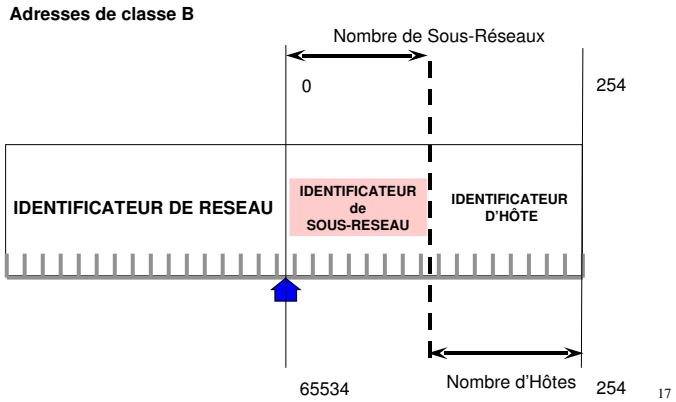
- **1517** Applicability Statement for the Implementation of Classless Inter-Domain Routing (CIDR). Internet Engineering Steering Group, R. Hinden. September 1993. (Status: HISTORIC)
- **1518** An Architecture for IP Address Allocation with CIDR. Y. Rekhter, T. Li. September 1993. (Status: HISTORIC)
- 1519 Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy. V. Fuller, T. Li, J. Yu, K. Varadhan. September 1993. (Obsoletes RFC1338) (Obsoleted by RFC4632) (Status: PROPOSED STANDARD)
- **1520** Exchanging Routing Information Across Provider Boundaries in the CIDR Environment. Y. Rekhter, C. Topolcic. September 1993. (Status: HISTORIC)
- **4632** Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan. V. Fuller, T. Li. August 2006. (Obsoletes RFC1519) (Also BCP0122) (Status: BEST CURRENT PRACTICE)

Dans l'ancien monde classfull, les classes se sont révélées mal adaptées : une classe A, c'est très rare et gros, une classe C, c'est trop petit. Les sites ont dont été amenés à demander des classes B. 65000 postes sur un réseau unique, c'est trop :

- Pour ethernet et son algothme CSMA-CD
- D'un point de vue taille, raccorder tous les services peut amener à dépasser la taille maxi (longueur, maxi 4 répéteurs, ...)

Comme obtenir de nouvelles adresse n'est pas possible, la solution consiste à découper la classe B en plusieurs sous-réseau. Cela impose de s'appuyer sur le masque de sous-réseau pour connaître la taille de la partie réseau et non plus sur la forme du premier octet de l'adresse

sous-réseaux



17

Table de sous-réseaux

Nombre de Sous-Réseau	Nombre de Bits requis	Masque de Sous-Réseau	Nombre d'Hôtes par Sous-Réseau
0	1	Invalide	Invalide
2	2	255.192.0.0	4.194.302
6	3	255.224.0.0	2.097.150
14	4	255.240.0.0	1.048.574
30	5	255.248.0.0	524.286
62	6	255.252.0.0	262.142
126	7	255.254.0.0	131.070
254	8	255.255.0.0	65.534

Nombre de masques de sous-réseaux déjà convertis en utilisant un octet

pour les réseaux de classe A

18

- Bits du masque de Sous-Réseau
  - Préalablement à la définition du masque
    - Le nombre de Sous-Réseaux doit être défini
  - Plus le nombre de bits affectés au masque de sous-réseau augmente
    - Plus le nombre de sous-réseaux possibles augmente
    - Plus le nombre d'hôtes par sous réseau diminue
- Planification
  - Elle est indispensable afin de déterminer les besoins et permet un cadrage précis du nombre de bits de masque réellement nécessaire.

La RFC1878 (Variable Length Subnet Table For Ipv4) qui remplace la rfc 950 donne les bonnes pratiques et des exemples concrets de création de sous réseau.
Le tableau ci-dessus et les suivants s'inspirent de cette rfc.

Le masque de sous-réseau doit avoir ses bits à 1 contigus et à gauche.

Table de sous-réseaux

Nombre de Sous-Réseau	Nombre de Bits requis	Masque de Sous-Réseau	Nombre d'Hôtes par Sous-Réseau
0	1	Invalide	Invalide
2	2	255.255.192.0	16.382
6	3	255.255.224.0	8190
14	4	255.255.240.0	4094
30	5	255.255.248.0	2046
62	6	255.255.252.0	1022
126	7	255.255.254.0	510
254	8	255.255.255.0	254

Nombre de masques de sous-réseaux déjà convertis en utilisant un octet

pour les réseaux de classe B

Table de sous-réseaux

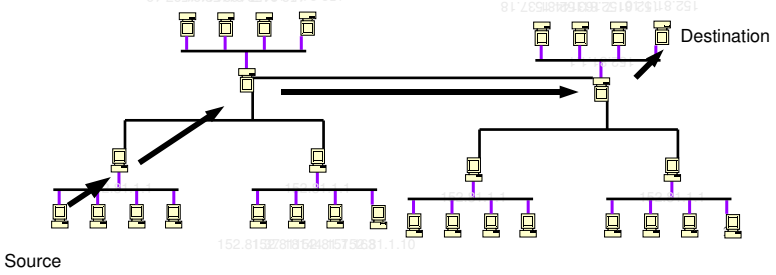
Nombre de Sous-Réseau	Nombre de Bits requis	Masque de Sous-Réseau	Nombre d'Hôtes par Sous-Réseau
0	1	Invalide	Invalide
2	2	255.255.255.192	62
6	3	255.255.255.224	30
14	4	255.255.255.240	14
30	5	255.255.255.248	6
62	6	255.255.255.252	2
126	7	255.255.255.254	Invalide
254	8	255.255.255.255	Invalide

Nombre de masques de sous-réseaux déjà convertis en utilisant un octet

pour les réseaux de classe C

Routage IP: problématique

- Une machine sait transmettre les paquets sur les sous-réseaux de ses interfaces (réseaux locaux)
- Les autres paquets sont envoyés à un routeur directement joignable (situé sur un réseau local)



Routeur (ou, par abus de langage, passerelle) : Une machine qui sait gérer et retransmettre des paquets qui ne lui sont pas destinés.

Le trajet complet est une suite de saut de puce entre machines directement connectées.
La machine source ne détermine pas le chemin. Son rôle est de déterminer la prochaine étape (**next hop**).
C'est aussi le rôle de chaque machine intermédiaire. A chaque étape, il y a détermination de la prochaine machine intermédiaire.
C'est l'**algorithme de routage** qui détermine le « next hop ».
Il le détermine en fonction:

- D'informations locales présentes sur la machine: la **table de routage**;
- Des informations (adresse ip destination) présente dans le paquet à transmettre.

La table de routage peut être construite :

- Par l'administrateur système: on parle de **routage statique**
- Par des programmes qui échangent automatiquement des informations avec les autres routeurs : on parle de **routage dynamique**.

Routage: routeur par défaut, routes statique

- Table de routage :
  - Une entrée pour chaque réseau directement connecté
  - Routeur par défaut: pour les destinations non traités par les autres entrées
  - Routes statiques: pour les destinations pour lesquelles le routeur par défaut ne convient pas
- Le parcours de la table est récursif
  - Les cas d'arrêt sont les réseaux directement connectés à l'hôte.

Chaque machine s'appuie sur sa table de routage pour savoir que faire d'un paquet à émettre. La décision se fait en fonction de l'adresse destination du paquet.
La table de routage contient une entrée pour chaque réseau auquel la machine est directement connecté. En principe, les paquets à destination de ces réseaux sont directement transmis à la destination
Elle contient une entrée par défaut qui fait office de règle balai : les paquets dont la destination n'a pas été traitée par les autres règles sont transmises au routeur par défaut.
Elle contient d'autres régles pour les cas où le routeur par défaut n'est pas la bonne solution.
Les règles/entrées ont grosso-modo la forme suivante :

- Destination (adresse, masque), adresse du routeur à utiliser

si la destination est un hôte, le masque est 255.255.255.255 sinon, c'est le masque du réseau destination.
Les réseaux directement connecté à la machine ont 0.0.0.0 comme routeur destination
Exemple d'entrée:
destination: 192.168.10.0, 255.255.255.0, routeur: 0.0.0.0
destination: 192.168.20.0, 255.255.255.0, routeur: 192.168.10.1

Routage : algo de routage (faux)

- quand une machine M a un paquet à transmettre, elle applique l'algorithme suivant :
  - si le paquet est pour une machine située sur l'un des sous-réseaux d'une de ses cartes réseau, il est envoyé directement à la destination
  - si le paquet est pour un hôte pour lequel M a une route définie, il est envoyé au routeur défini dans la route
  - si le paquet est pour un réseau pour lelque M a une route définie, => envoyé au routeur défini dans la route
  - sinon, le paquet est envoyé au routeur par défaut de M

L'algorithme ci-dessus correspond globalement au comportement de la majeure partie des tables de routage.

Il peut cependant être factuellement faux car les entrées sont traitées dans l'ordre des maques croissants. Ainsi, une route vers un hôtes (masque 255.255.255.255) sera prioritaire par rapport à une route vers un réseau directement connecté.

Un poste de travail a en général une table de routage très simple :

- Une entrée vers le sous réseau auquel est relié le poste
- Un routeur par défaut

Il est déconseillé que des postes de travail aient des tables de routage complexes. Ca rend le réseau difficile à gérer et son fonctionnement difficile à prédire.

En général, ce sont les routeurs centraux qui ont des tables de routage plus complexes.

Routage: algo de routage (juste)

- Tri des lignes par taille de masque décroissante
- On prend la première entrée qui convient
- Fonctionnement récursif

Table de routage IP du noyau

Destination	Passerelle	Genmask	Indic	Iface
195.221.162.0	0.0.0.0	255.255.255.0	U	eth0
172.18.0.0	192.168.120.102	255.255.0.0	UG	eth1
172.17.0.0	0.0.0.0	255.255.0.0	U	eth2
192.168.0.0	0.0.0.0	255.255.0.0	U	eth1
172.20.0.0	0.0.0.0	255.255.0.0	U	eth3
0.0.0.0	195.221.162.249	0.0.0.0	UG	eth0

On considère l'adresse de destination du paquet à traiter. Parmi les règles correspondant à cette destination, on utilise celle qui le masque le plus long. Ainsi, si on a une règle avec un masque /20, une avec un masque /24 et une avec un masque /0 (routeur par défaut), on utilise celle avec le masque /24.

Cette méthode fait que les routes vers des hôtes sont prioritaires par rapport aux routes vers les réseaux, y compris les réseaux directement connectés.