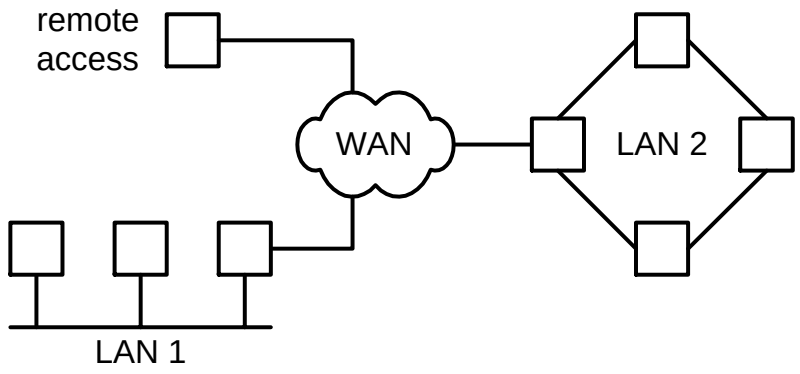
RESEAUX : *Internetworking*

1 Motivation

Pourquoi interconnecter des réseaux locaux ?

a) Relier des sites dispersés géographiquement



Etendre la portée du réseau local : **LAN *to* LAN**
intranet

Accès à distance : ***remote access***
domicile, portable, …
extranet

WAN *(Wide Area Network)* : Circuit spécialisé et modems
Réseau téléphonique
RNIS
internet
ATM, *frame relay*

b) Communiquer entre équipements différents

Assurer le transfert des données dans un **réseau hétérogène** → interopérabilité

Conversion de protocoles propriétaires : IBM-SNA, DECNET, …

Systèmes de messagerie : smtp, X.400, Lotus-Notes, …

c) Subdiviser un réseau en sous-réseaux

Répartition de la **charge**

Créer des **groupes logiques** d'utilisateurs *(Virtual LAN)* : développement, production, …

Sécurité

d) Communiquer facilement au niveau mondial

Rôle fédérateur du réseau *internet* (réseau public)
→ *Internet Service Provider*

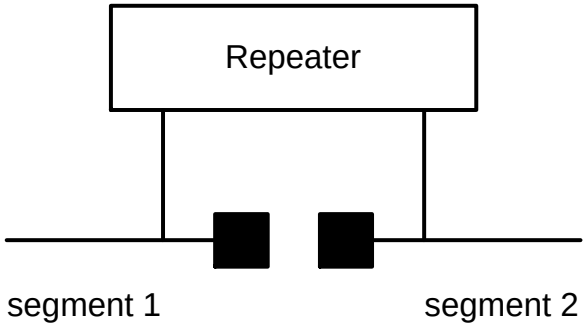
Mêmes technologies au niveau privé
→ *Service Provider*
Qualité de service, sécurité, …

2 Caractéristiques d'un répéteur

Répéteur = ***repeater***

Fonction de niveau 1 (couche physique)

Régénérer le signal 500 m avec 10Base5
200 m avec 10Base2
100 m avec 10BaseT



Tronçon en fibre optique pour étendre la portée du réseau local

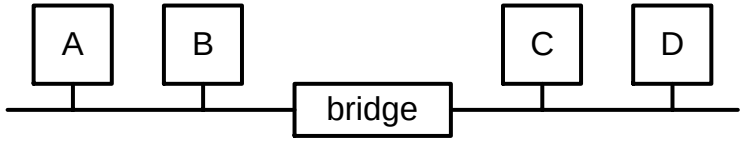
Interfaces électriques : connecteur AUI (15 pins)
connecteur BNC
connecteur RJ45

Certains affirment que 50% des anomalies d'un réseau provient de cette couche physique.

3 Fonctions d'un pont

Pont = ***bridge***

Subdiviser un réseau en sous-réseaux (segments) à partir des adresses de niveau 2 ou MAC (*Medium Access Control*)



Apprentissage des adresses MAC :

Mise sous tension du *bridge*
A émet une trame à B
Bridge learns : A se trouve sur le segment gauche
Bridge forwards la trame sur le segment de droite
B répond à A
Bridge learns : B se trouve sur le segment gauche
Bridge filters cette trame

Un segment *ethernet* à 10 Mbit/s peut compter jusqu'à 14880 trame/s.
→ caractéristiques essentielles :

Learning rate capacité de mémorisation
Max address
Filtering rate capacité de filtrage
Forwarding rate capacité de transfert

Bits de CRC : aucune règle !

Certains *bridges* transmettent ce CRC
→ protocole d'extrémité à extrémité
D'autres *bridges* gèrent ce CRC
Les trames avec erreurs de CRC sont ignorées
→ protocole en cascade

Un *bridge* est qualifié de :

Local bridge s'il relie directement 2 réseaux locaux
Remote bridge s'il utilise une liaison WAN.

Ex 1 : Comparer le degré de transparence entre 2 segments *ethernet* reliés par :

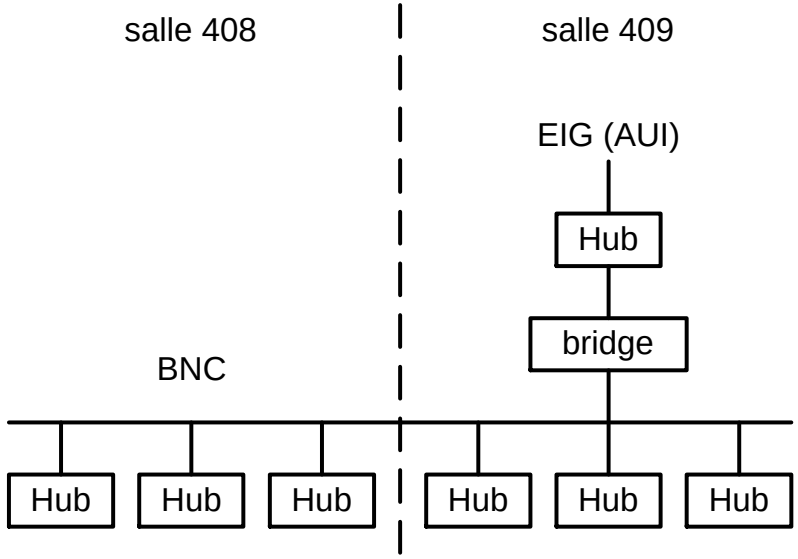
paramètre	répéteur	pont
charge (trame/s)		
collisions		
broadcast		
analyseur de protocole		
erreur de CRC		

4 Réseau *ethernet* du laboratoire

Les nœuds sont raccordés via l'interface 10 Base T (RJ 45).

Chaque salle dispose de *hubs* reliés par l'interface 10 Base 2 (BNC).

Un pont filtre le trafic du laboratoire de celui de l'école.



5 *Ethernet* à 100 Mbit/s

La norme 100 Base T *(Fast ethernet)*, apparue en 1994, reste compatible avec les réseaux *ethernet* à 10 Mbit/s.

Elle conserve l'accès non déterministe, défini dans la norme IEEE 802.3 (CSMA/CD) et permet une distance maximale de 100 m sur paire torsadée entre nœud et *hub*.

Le temps de propagation aller et retour *(RTD : Round Trip Delay)* est cette fois de 5,1 μs (51 μs à 10 Mbit/s) alors que l'*interframe gap* = 0,96 μs (9,6 μs à 10 Mbit/s) correspond toujours à 12 octets

Variantes :
- 100 Base TX 2 paires torsadées
- 100 Base T4 4 paires torsadées
- 100 Base FX 2 fibres multimode

Détection automatique *(auto negotiation)*
Certains *hubs auto sensing* sont capables de détecter le type de nœud : 10 Base T, 100 Base TX, 100 Base T4

Réseau mixte 10/100 Mbit/s ?
L'introduction de nœuds 100 Base T se fait de façon assez naturelle car elle ne remet pas en cause le câblage existant.
La migration s'effectue progressivement avec des *hubs* hybrides 10/100 Base T.

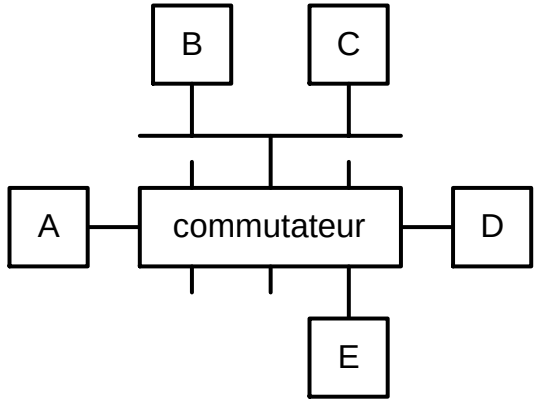
Ex 2 Peut-on assurer l'interfonctionnement 10 Base T - 100 Base T avec un *hub* ?

6 Commutateur *ethernet*

Ce commutateur ***(ethernet switch)***, à ne pas confondre avec un *hub*, fonctionne sur le principe du *multiport bridge*.

Il teste chaque trame (adresse de destination) et ouvre au besoin un canal entre ces 2 ports; plusieurs canaux pouvant être actifs simultanément (4 au maximum dans l'exemple ci-dessous).

Illustration avec un commutateur 8 ports :



Dans cet exemple, les éventuelles collisions sur le segment des nœuds B et C ne sont pas visibles sur les autres ports du commutateur.
Chaque port constitue donc son **propre domaine de collisions**.

L'adresse de diffusion est propagée sur tous les autres ports.
→ **Un seul domaine de diffusion** *(broadcast domain)*

Voir **annexe 1 (réseau EIG) :** on parle d'un réseau "bridgé"

a) ***Shared versus switched LAN***

Un réseau 10 Base T composé de *hubs* partage une bande passante de 10 Mbit/s → ***shared LAN***

Un commutateur 100 Base T permet l'émission de plusieurs trames simultanées → ***switched LAN***
Exemple : A → C et D → E

b) ***Full duplex***

La majorité des équipements 100 Base T autorisent l'émission et la réception simultanées.

Ex 3 Peut-on imaginer de travailler en *full duplex* avec un répéteur ?

Ex 4 Peut-on imaginer de travailler en *full duplex* avec un *hub* ?

Ex 5 Peut-on s'attendre, dans tous les cas, à une augmentation de 100 % des performances en passant de *half* à *full duplex* ?

c) **Principe de fonctionnement**

Le commutateur mémorise l'adresse source MAC de chaque trame dans sa table de filtrage.

Si l'adresse de destination de la trame est inconnue; la trame est diffusée sur tous les ports.

Chaque commutateur dispose d'une certaine capacité de mémorisation.

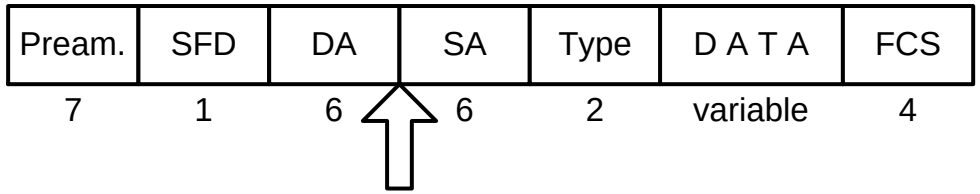
La capacité de transmission à l'intérieur du commutateur doit être suffisante pour écouler plusieurs trames simultanées.
→ Transmission en parallèle sur un bus.

d) **Sécurité**

Possibilité, comme avec le *hub*, d'isoler un nœud jugé défaillant.

e) Commutateur du type *cut-through*

Il commute la trame dès qu'il a reçu l'adresse de destination



Temps de latence minimum (15 μs) indépendant de la trame.

Retransmission des erreurs (CRC, trame trop courte)

Même débit 10 - 100 Mbit/s pas possible

Commutation au niveau matériel *(switch fabric, ASIC = Application Specific Integrated Circuit)*

f) Commutateur du type *store & forward*

La trame n'est commutée que lorsqu'elle a été complètement reçue

Temps de latence dépendant de la longueur de la trame.

Erreurs sont filtrées (CRC, trame trop courte)

Adaptation de débit 10 - 100 Mbit/s possible

Commutation au niveau logiciel (CPU, RISC)
Mises à jour du logiciel possibles

g) Critères de choix

Bien que tous les commutateurs *ethernet* fonctionnent sur le même principe, leur fonctionnement dépend :

- Nombre d'adresses MAC gérées par port (table de filtrage)
- Variantes de ports : 10Mbit/s, 100Mbit/s, 1Gbit/s
- Outils d'administration (web, snmp, telnet, RS232, RMON…)
- *Monitoring Port*
  Possibilité de copier le trafic d'un port sur un autre port relié à un analyseur
- *Automatic Broadcast Control*
  Limitation des trames de diffusion.
  Par exemple 5% de la bande passante
- Contrôle de congestion
  Que se passe-t-il si A et B émettent simultanément une trame destinée à D ?
  FIFO plein → perte de trame
  Retransmission au niveau MAC *(back pressure* par génération de collisions)
  Retransmission au niveau TCP, …
- Mode adaptatif
  Démarrage en mode *cut-through*
  Passage en *store & forward* si seuil d'erreurs dépassé
  Retour en mode *cut-through* en dessous du seuil.

h) *Gigabit ethernet*

Cette norme, apparue en 1999, conserve le format de trames *ethernet*.

Elle est principalement utilisée en mode ***full duplex*** où la méthode d'accès CSMA/CD est désactivée.
La charge théorique max est alors de 1,488 Mio trame/s

La méthode d'accès CSMA/CD, active qu'en *half duplex*, a été modifiée afin de permettre une distance suffisante.
La durée minimale d'émission de la trame est fixée à 512 octets permettant ainsi une domaine de collision de 200 m.
Le contrôleur *ethernet*, qui doit émettre une trame de longueur inférieure, génère un signal particulier tout en détectant les colisions *(carrier extension)*.
Le débit utile, dans le cas le plus défavorable (émission de trames de 64 octets), atteint ainsi 120 Mbit/s !

Principales caractéristiques de 802.3z :

- débit = 1 Gbit/s
- 1000 Base CX 25 m *shielded balanced copper*
- 1000 Base SX 550 m fibre optique multimode
- 1000 Base LX 3 km fibre optique monomode

Variante 802.3ab

- 1000 Base T 100 m UTP cat. 5 (4 paires)

i) 10 Gbit/s

Ratification du standard prévue en 2002

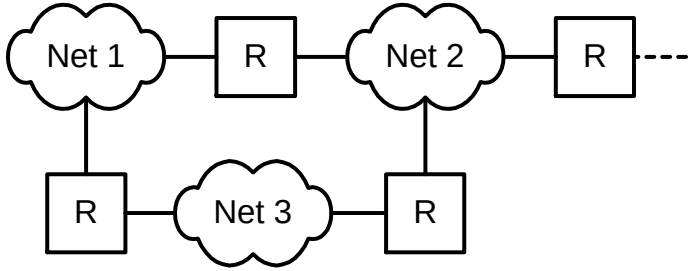
Extension *full duplex*

100 m à 40 km selon la fibre (monomode, multimode, noire)

7 *Internet*

a) Définition

Internet est constitué d'un ensemble de réseaux ***(Network)*** reliés par des routeurs ***(Routers)*** :

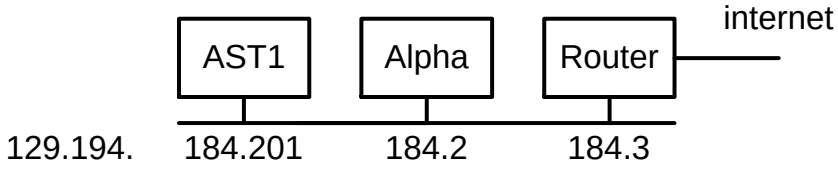


Rappel : adresse IP = partie ***Network*** + partie ***Host***

Exemple : serveur Alpha de l'EIG = 129.194.184.2
→ adresse de classe B
→ ***Network*** = 129.194
→ ***Host*** = 184.2

b) Configuration du poste de travail **AST1 :**

Adresse IP	*IP address*	129.194.184.201
Masque de sous-réseau	*subnet mask*	255.255.0.0
Passerelle par défaut	*router*	129.194.184.3



c) *Subnet Mask*

Valeur par défaut dans notre cas (adresse de classe B) :

255.255.0.0

255	255	0	0
11111111	11111111	00000000	00000000

La partie à 1 correspond à la partie ***Network***
La partie à 0 correspond à la partie ***Host***

Ce masque permet de distinguer, parmi toutes les destinations possibles, entre destination **directe** ou **indirecte** :

Direct destination

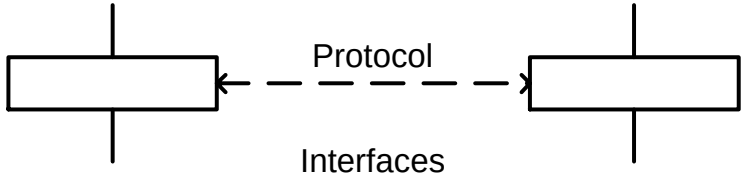
ping	129.194.184.2	
subnet mask	255.255.0.0	→ ***same network***

Indirect destination

ping	130.59.1.40	
subnet mask	255.255.0.0	→ ***other network***
		router = 129.194.184.3

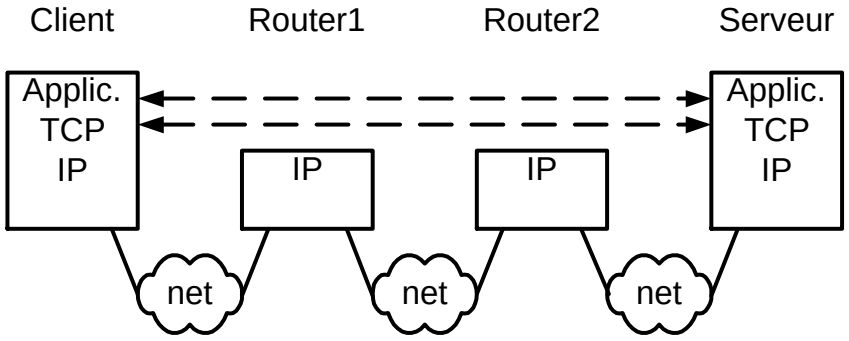
d) Modèle en couches

Chaque couche (*ethernet*, IP, TCP, …) communique avec la couche opposée selon des règles précises définies par le protocole



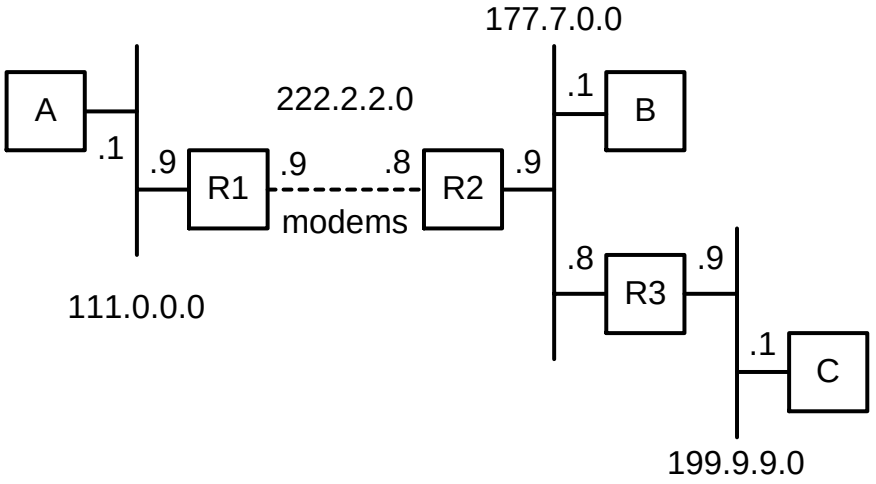
Certains protocoles, comme TCP, ont une signification d'extrémité à extrémité *(end to end protocol)*

Représenter le *protocol stack* suivant :



8 Routage statique

Dans cet exemple, chaque équipement est configuré manuellement :

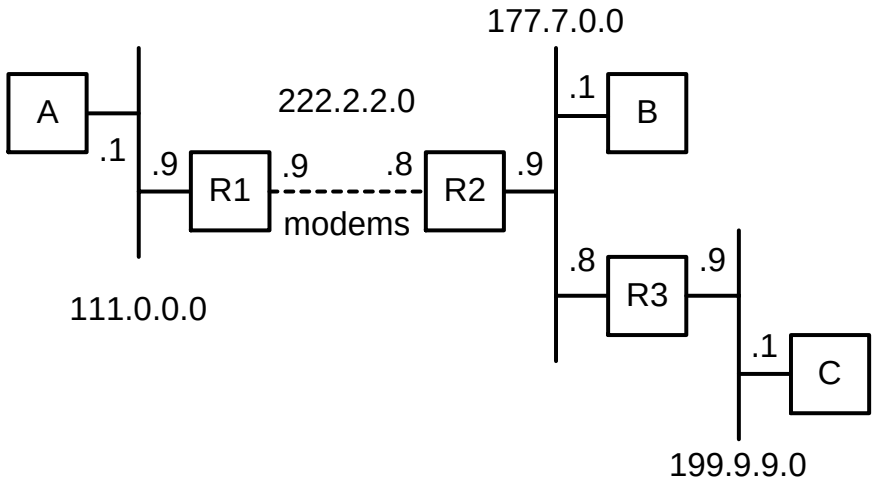


- Sur la machine **B** ifconfig eth0 177.7.0.1 mask 255.255.0.0
  route add 111.0.0.0 177.7.0.9
  route add 199.9.9.0 177.7.0.8
- Sur la machine **A** ifconfig eth0 111.0.0.1 mask 255.0.0.0
  route add default 111.0.0.9
- Sur la routeur **R1** ifconfig eth0 111.0.0.9 mask 255.255.0.0
  ifconfig le0 222.2.2.9 mask 255.255.255.0
  route add default 222.2.2.8
- Sur la routeur **R2** ifconfig eth0 177.7.0.9 mask 255.255.0.0
  ifconfig le0 222.2.2.8 mask 255.255.255.0
  route add 111.0.0.0 222.2.2.9
  route add 199.9.9.0 177.7.0.8

9 Routage dynamique

Le routage dynamique met en œuvre un protocole de communication inter-routeurs.

Chaque routeur informe son voisin des réseaux auxquels il est connecté.



Configuration du routeur **R1** :

Interface LAN	*IP address*	111.0.0.9	classe A
	subnet mask	255.0.0.0	
	router		
Interface WAN	*IP address*	222.2.2.9	classe C
	subnet mask	255.255.255.0	
	router	222.2.2.8	

Illustration avec le protocole RIP *(Routing Information Protocol)* qui utilise la notion de distance.

A intervalles réguliers (30 s), un routeur transmet à ses voisins une copie de sa table de routage :

R2 envoie à ses voisins :	222.2.2.0	d=0
	177.7.0.0	d=0
R1 envoie à ses voisins :	111.0.0.0	d=0
	222.2.2.0	d=0
R3 envoie à ses voisins :	177.7.0.0	d=0
	199.9.9.0	d=0

Le paramètre d *(metric)* indique le nombre de routeurs intermédiaires pour atteindre le réseau de destination.

Chaque routeur met à jour sa table de routage :

R2	222.2.2.0	d=0
	177.7.0.0	d=0
	111.0.0.0	d=1
	222.2.2.0	*d=1*
	177.7.0.0	*d=1*
	199.9.9.0	d=1

Chaque routeur gère une **table de routage IP** (*IP routing table*) qu'il consulte à chaque fois qu'il reçoit un datagramme.

Caractéristiques :
- Protocole de routage très simple
- Distance est une information sommaire
- Distance max = 15
- Pas de garantie sur l'origine des informations
- Convient pour de petits réseaux

Temps de convergence
- Un algorithme comme RIP, de type *distance vector*, exprime la distance en nombre de sauts *(hops count)*.
- Un saut est représenté par un routeur reliant 2 segments.
- Chaque routeur construit sa table de routage à partir des informations diffusées par les autres.
- Ce mécanisme d'apprentissage progressif n'est pas immédiat et n'aboutit à une description complète du réseau qu'après un temps de convergence de l'algorithme.

Ex 6 : Déterminer les champs (ethernet, ip) qu'un routeur doit modifier.

Ex 7 : Quelle est la portée d'une table de routage ?

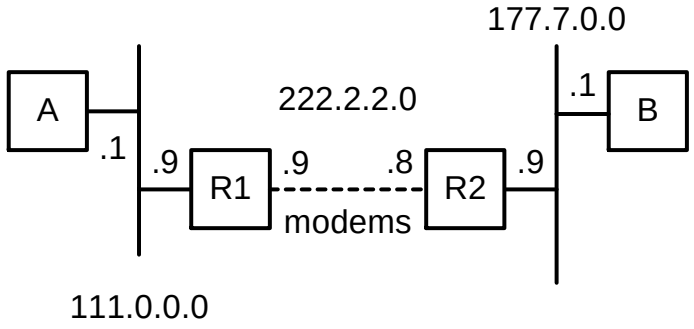
Ex 8 : Comment un réseau inconnu (par la table) est-il atteint ?

Ex 9 : Un paquet peut-il se perdre ?

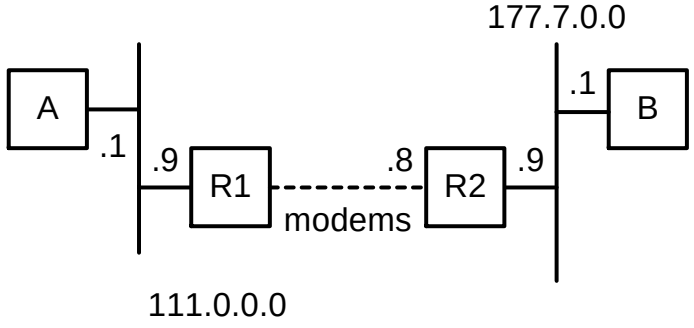
Ex 10 : Déterminer l'architecture de la couche IP d'un routeur composé des blocs fonctionnels suivants (FIFOs, filtrages, interfaces réseaux, table de routage, RIP, ICMP).

10 *Proxy ARP*

Dans l'exemple précédent, il est possible de ne plus monopoliser des adresses IP pour une simple liaison série.



Le routeur **R1**doit être configuré en mode *proxy ARP* (RFC 1027).



Lorsque **A** effectue une commande **ping 111.0.0.8 :**
- A émet une requête ARP car le masque = 255.0.0.0
- R1 répond (se fait passer) pour R2

A atteint B et réciproquement à partir des tables de routage.

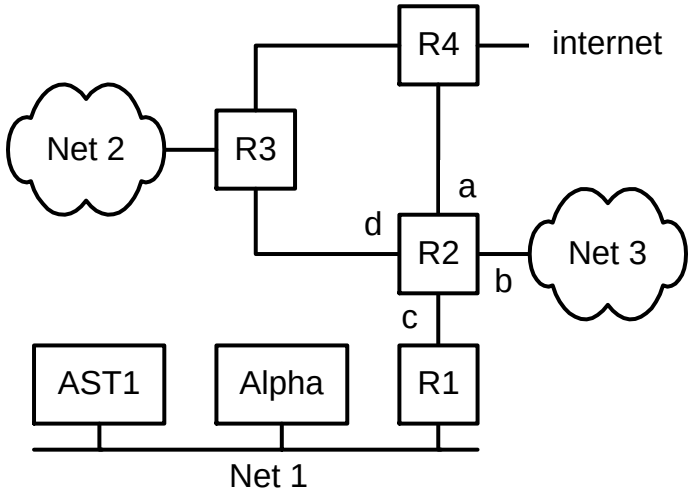
Cette fonction est également utilisée par les équipements *dial-up router*

11 Hiérarchie

Un réseau de la taille d'*internet* dispose de domaines de routage appelés, *top-level routing domains*, qui en constituent l'épine dorsale *(backbone)* :

NSFNET	*National Science Foundation Network*
CIX	*Commercial Internet Exchange*
EBONE	*European IP Backbone*

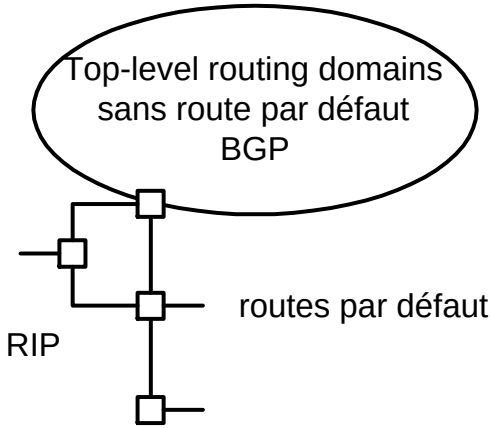
Les routeurs périphériques utilisent la route par défaut :



Dans cet exemple tous les routeurs, à l'exceptions de R1, effectuent l'opération de **choix du chemin** des paquets IP sur la base des informations contenues dans la table de routage.

Table de routage R2 :	Net 3	d=0	interface b
	Net 1	d=1	interface c
	Net 2	d=1	interface d
	Défaut		interface a

Vue d'ensemble



Epine dorsale *(top-level routing domains)* avec protocole de routage BGP *(Border Gateway Protocol)*

Parmi l'information transmise figure le chemin complet permettant d'atteindre ces destinations.
Cette information peut servir à établir un graphe de connectivité; les boucles de routage pouvant ensuite être éliminées.

Ce protocole, qui exige une connexion TCP, peut facilement détecter la défaillance d'une liaison.

Le protocole RIP plus simple, se contente d'échanges UDP.

La route par défaut simplifie les tables de routages périphériques.

12 *Traceroute*

La commande Traceroute permet de localiser chaque routeur situé sur le chemin en envoyant des datagrammes successifs avec le champ TTL égal à 1, 2, ...

Un message ICMP "*time exceeded"* est retourné par chaque routeur rejetant le datagramme et un message ICMP "*port unreachable"* est émis par la destination finale.

```
TraceRoute to host www.luth.se

#    Address         Host Name                 Response Time
1    129.194.184.3   Unavailable                  4 ms
2    129.194.12.3    unige-gw.unige.ch            2 ms
3    192.33.214.3    swige1.unige.ch              2 ms
4    130.59.33.45    swiCE1-A4-0-0-2.SWITCH.ch    3 ms
5    212.1.192.169   switch.ch.ten-155.NET        4 ms
6    212.1.192.46    ch-se.se.ten-155.NET        58 ms
7    212.1.192.154   sw-gw.nordu.NET             53 ms
8    193.10.252.178  STK-BB-1.SUNET.SE           54 ms
9    130.242.200.10  SVL-BB-1.SUNET.SE           58 ms
10   130.242.200.126 lulea-pos.SUNET.SE          66 ms
11   130.242.202.116 Unavailable                 68 ms
12   130.240.42.42   www.luth.se                 69 ms
```


13 Adressage de sous-réseaux

Nous avons considéré qu'une adresse IP était composée de 2 identificateurs : ***Network + Host***.

En fait, l'identificateur ***Host*** peut identifier des **sous-réseaux.**

L'Université de Genève, qui dispose d'une adresse IP de classe B (*Network* = 129.194) a décidé de subdiviser son réseau en 64 sous-réseaux (*subnet*).

6 bits étant nécessaires, chaque nœud IP du réseau de l'Université doit être configuré avec le masque de sous-réseau suivant :

255.255.252.0

255	255	252	0
11111111	11111111	11111100	00000000

Chaque *subnet* ne peut plus contenir que 1024 ordinateurs.

Voir **annexe 2 : réseau UNIGE**

On parle dans ce cas d'un réseau "routé" par opposition à un réseau "bridgé".

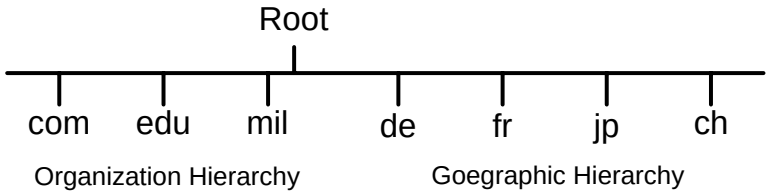
- Le pont utilise les adresses MAC de niveau 2
- Le routeur utilise les adresse IP de niveau 3

14 *Domain Name System* (DNS)

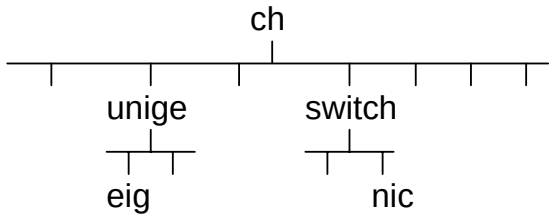
L'utilisateur du réseau *internet* préfère souvent utiliser une adresse facilement mémorisable comme http://**www.cern.ch** ou ping **nic.switch.ch** plutôt qu'une adresse IP comme 130.59.1.40

On parle alors de ***Full Qualified Domain Name*** *(FQDN)*
A l'origine, un seul fichier HOSTS.TXT gérait l'équivalence entre une adresse FQDN et une adresse IP.

Le *domain name system* est une **base de données distribuée.**



Structure arborescente : eig.unige.ch
(host=eig domain=unige.ch)

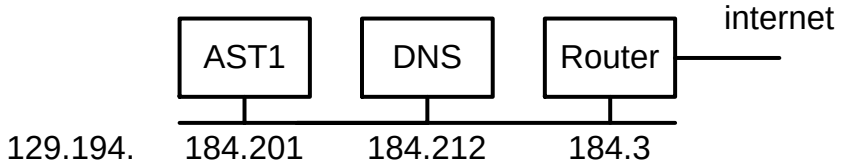


Possibilité de déléguer l'autorité : Université gère unige.ch
labo de TD gère td.unige.ch

L'échange est du type **client** *(resolver)* - **serveur** *(name server).*
Lien entre père et fils par leur adresse ip et leur domaine.
Mémoire cache dans chaque serveur.

a) Configuration du poste de travail **AST1 :**

Adresse IP	*IP address*	129.194.184.201
Masque de sous-réseau	*subnet mask*	255.255.0.0
Passerelle par défaut	*router*	129.194.184.3
Serveur de nom	***domain name server***	**129.194.184.212**



Ex 11 : Déterminer le trafic DNS si le client AST1 exécute **ping nic.switch.ch**

Ex 12 : Client et serveur DNS doivent-ils être sur le même réseau (physique, IP) ?

Ex 13 : Que se passe-t-il si le serveur DNS est hors service ?

Ex 14 : Que contient le serveur DNS du domaine **td.unige.ch** géré par le laboratoire

b) Fichier de configuration du serveur DNS du labo

fichier td.unige.ch.dns

```
@       IN  SOA     netserver2.td.unige.ch.
@       IN  NS      netserver2

ast1    IN  A       129.194.184.200
hp3     IN  A       129.194.184.99

www     IN  CNAME   hp3
ftp     IN  CNAME   hp3
```

Types définis de la classe Internet (IN) selon RFC 1700 :

- SOA *marks the start of zone of authority*
- NS *an authoritative name server*
- A *a host address*
- CNAME *the canonical name for an alias*
- PTR *a domain name pointer*
- MX *mail exchange*

c) Requête inverse

Il est possible d'activer la fonction inverse sur un serveur DNS afin de lui demander le DNFQ correspondant à une adresse IP.

La commande Traceroute utilise cette fonctionnalité.

Ex 15 : Pourquoi une requête sur www.altavista.com retourne plusieurs adresses IP

15 Configuration automatique

a) Motivation

La configuration d'un nœud IP demande des connaissances spécifiques que certains utilisateurs n'ont pas.
L'essor d'équipements grand public comme portable, *mobile IP*, *phone IP*, *thin client*, … exige une procédure de configuration automatique.
Certaines entreprises changent d'ISP *(Internet Service Provider)* tous les ans.

b) RARP *(Reverse Address Resolution Protocol)* - RFC 903

Ce protocole de résolution d'adresse inverse permet à un système minimum sans disque de demander l'adresse IP correspondant à l'adresse physique transmise.
Le serveur RARP contient le fichier d'équivalence entre adresse physique (MAC) et adresse logique (IP).

c) BOOTP *(Bootstrap Protocol)* - RFC 951

Il offre une alternative à RARP en offrant des informations supplémentaires comme le masque de sous réseau, l'adresse IP du routeur, l'adresse IP du serveur DNS, …

d) DHCP (*Dynamic Host Configuration Protocol*) - RFC 2131

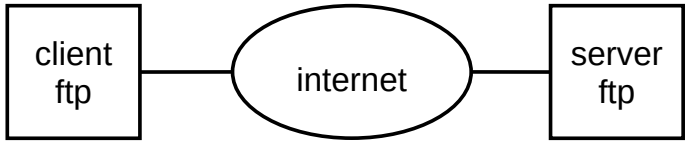
Un serveur DHCP va allouer dynamiquement les adresses IP dans l'espace qu'il gère sans connaître l'adresse physique du nœud.
Il transmet également les informations supplémentaires décrites précédemment.
Le client doit être dans le mode (par défaut de Win2000 : *Obtain an IP address from a DHCP server.*
La configuration présente une certaine durée de vie *(lease time).*

16 Performances

L'utilisateur souhaite naturellement les meilleures performances au prix le plus bas !

Améliorer sans cesse le **débit utile** est une opération complexe !

Considérons une application courante comme **FTP :**



Plusieurs **paramètres** vont influencer ce débit utile :

- Performances des ordinateurs (CPU, DMA, débit du disque dur, taille de la mémoire cache, optimisation des couches de protocole, carte réseau, ...)
- Dissymétrie (client rapide - serveur lent, ...)
- Caractéristique du réseau (nombres de routeurs traversés, temps aller-retour, débit binaire le plus lent,...)

L'ingénieur en télécommunications, dans un travail d'optimisation (débit utile, coût, temps de réponse, ...) doit être capable de décomposer et de caractériser les différents composants de cette chaîne.

a) Longueur optimum

A tout bloc d'information correspond une longueur maximale, appelée **MTU** *(Maximum Transmission Unit)*, qui dépend du type de réseau (ethernet, ATM, ppp, ...).

La spécification *ethernet DIX*, par exemple, fixe cette taille à 1500 octets.

La couche IP, pour émettre un datagramme dont la longueur est supérieure au MTU, devra employer la **fragmentation**, qui consiste à casser ce datagramme en "morceaux" dont la taille ne dépasse pas ce MTU.

La norme RFC1191 précise quelques valeurs courantes de **MTU de réseau :**

réseau	MTU (octets)
ethernet DIX	1500
IEEE 802.3	1492
X.25	576
PPP	296
Token Ring - 16	17814

b) MTU de chemin

Quel MTU faut-il considérer lorsque client et serveur communiquent à travers plusieurs réseaux ?
Logiquement le plus petit appelé **MTU de chemin.**

Remarquons que ce MTU peut varier en fonction du chemin pris par les datagrammes IP à travers les routeurs d'*internet*.

c) Taille maximum de segment (TCP)

La taille maximum de segment (MSS : *Maximum Segment Size*) est le plus grand "morceau" de données que TCP enverra à l'autre extrémité.

Lors de l'établissement (SYN), chaque extrémité peut annoncer son MSS qu'elle s'attend à recevoir ; sinon une valeur par défaut de 536 est utilisée.

En général, un MSS le plus grand possible est souhaitable jusqu'à ce que la fragmentation apparaisse.

17 Performances de TCP

a) Débit utile maximum

Commençons par déterminer le débit utile maximum théorique sur un réseau ethernet de 10 Mbit/s :

Champ	Données	ACK
préambule ethernet	8	8
en-tête ethernet	14	14
en-tête IP	20	20
en-tête TCP	20	20
données utilisateur	1460	0
bourrage	0	6
CRC ethernet	4	4
interframe gap (9,6 μs)	12	12
TOTAL	**1538**	**84**

Si la taille de fenêtre TCP exige un *ACK* par segment :
débit utile = (1460 / 1538 + 84) x 10 Mbit/s
= 9,00 Mbit/s

Par contre 44 segments de 1460 octets peuvent être envoyés si la fenêtre est ouverte à sa taille maximale (65535) :
débit utile = (44 x 1460 / ((44 x 1538) + 84)) x 10
= 9,48 Mbit/s

Il s'agit d'une limite théorique qui fait l'hypothèse que les accusés de réception ne provoquent aucune collision.

b) Produit débit - délai

Ce produit définit le nombre d'octets émis dans l'intervalle nécessaire à son acquittement (contrôle de flux).

Ainsi 3750 octets sont transmis sur un réseau ethernet de 10 Mbit/s pendant un temps aller - retour égal à 3 ms.

Une taille de fenêtre inférieure réduirait d'autant le débit utile entre producteur et consommateur.

Quelques produits débit - délai :

Réseau	Débit binaire [bit/s]	Temps aller retour [ms]	Produit [octets]
ethernet	10.000.000	3	3.750
tél. transcontinent. T1	1.544.000	60	11.580
tél. satellite T1	1.544.000	500	95.500
tél. transconti. T3	45.000.000	60	337.500
gigabit transcontinent.	1.000.000.000	60	7.500.000

Le débit binaire remplace le débit utile calculé précédemment par soucis de simplicité.

Le temps nécessaire à l'envoi d'un paquet dépend de 2 facteurs : un **délai de propagation** (vitesse de la lumière, latences dans les équipements de transmission, ...) et un délai de transmission fonction du débit binaire.

Le **délai de transmission** domine à faible débit binaire alors que le délai de propagation domine à des débits de l'ordre du Gbit/s.

c) Limites

Des études sérieuses montrent que vous ne pouvez pas aller plus vite que :

- le maillon (la liaison) le plus lent,
- la taille de la fenêtre offerte par le récepteur, divisée par le temps aller-retour.

L'utilisateur dispose pour cela des paramètres suivants :

- *buffer size* taille de la mémoire cache
- *window size* taille de la fenêtre
- *MTU* *maximum transmission unit*

18 TCP de l'intérieur

a) Machine d'états

La figure suivante résume le diagramme d'états de cette couche TCP orientée connexion :

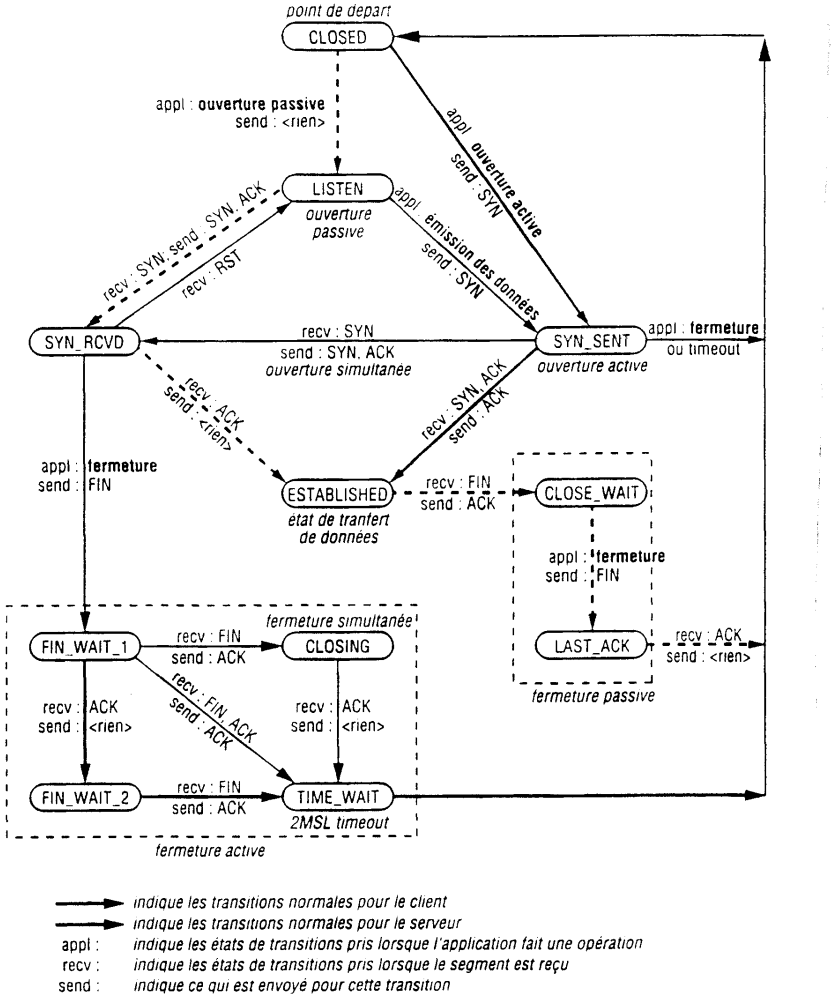


Figure 18.12 Diagramme des états de transitions de TCP.

b) Temporisateurs

Ce protocole, chargé de sécuriser l'échange, présente une gestion complexe des temporisateurs : 4 temporisateurs différents étant nécessaires par connexion.

Le protocole TCP surveille la réception des accusés de réception grâce à son temporisateur de **retransmission** qui doit s'adapter au temps aller-retour de l'échange.

Un temporisateur **persistant** est activé par le producteur dont le flux est bloqué (contrôle de flux) par le récepteur opposé (fenêtre fermée).

Un temporisateur *keepalive* optionnel permet à une extrémité de sonder l'autre côté alors qu'aucune donnée ne circule depuis 2 h.
En effet, aucune donnée ne circulera si la couche application reste muette. Certains estiment que l'application doit posséder ses propres temporisateurs ; d'autres préfèrent utiliser ce temporisateur.

Un temporisateur 2MSL mesure le temps pendant lequel une connexion a été dans l'état TIME_WAIT.
Chaque implémentation doit choisir une durée de vie maximum du segment *(Maximum Segment Lifetime).*

19 *Network Address Translation* (NAT)

a) *Private internets*

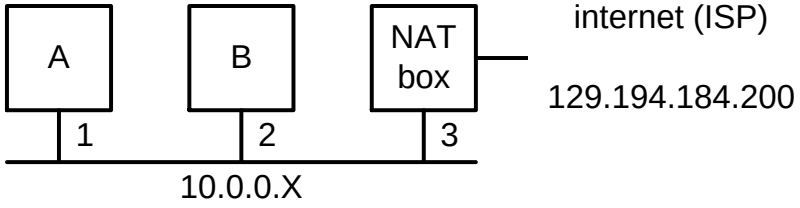
Certaines plages d'adresse ip sont réservées à un usage privé *(intranet)* et définies dans la RFC 1918 *(Address Allocation for Private Internets)* :

10.0.0.0	-	10.255.255.255	classe A
172.16.0.0	-	172.31.255.255	classe B
192.168.0.0	-	192.168.255.255	classe C

Les routeurs d'*internet* doivent être configurés pour ignorer ces adresses (adresses non routables).

b) Principe de NAT *(Network Address Translation)*

Il est possible, grâce à la fonction NAT *(Network Address Translation)*, d'accéder à *internet* avec plusieurs ordinateurs et **une seule adresse ip routable**.



Ce module NAT cache (mascarade) un réseau interne, utilisant une plage d'adresses privées, derrière une seule adresse ip publique.

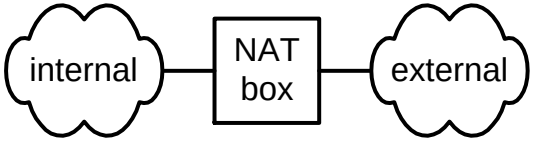
Il peut offrir un service DHCP du côté interne

Un client du réseau *internet* ne peut pas accéder aux serveurs du réseau privé (adresses privées non routables).

c) Terminologie selon RFC 1918

La RFC 1918 *(Address Allocation for Private Internets)* définit :

- un réseau privé interne *(internal)*
- un réseau externe *(external)*



d) Mode de fonctionnement statique

- Relation fixe entre adresses IP interne et externe
- *One to one mapping*
- Les 2 réseaux (interne et externe) sont de même taille
- Linux : *static address translation*

e) Mode de fonctionnement dynamique

- Un groupe d'adresses externes *(pool)* est partagé
- Nb adr IP internes > nb adr IP externes *(pool)*
- Linux : *dynamic address translation*

Cas particulier :

- Une seule adresse externe
- Linux : *masquerading*
- Cisco : *Port Address Translation*

Ex 16 : Déterminer le principe de fonctionnement de ce module utilisant le principe de la translation de port
- Quels champs doivent être modifiés ?
- Quelles informations doivent être mémorisées ?

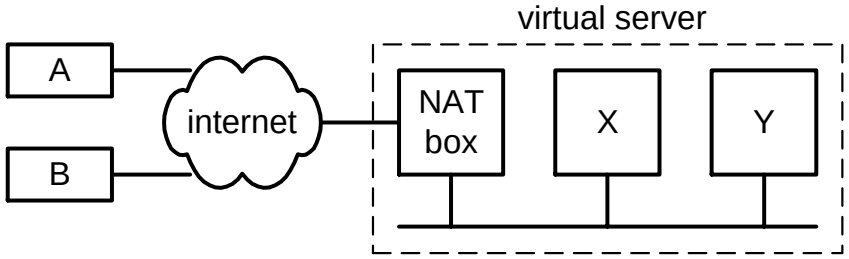
Ex 17 : Un client du réseau externe peut-il communiquer avec un serveur du réseau interne ?

f) Sécurité

La fonction NAT est disponible dans les routeurs, *dial-up routers* et *firewalls* (mur coupe-feu).

g) *Load balancing*

Une configuration particulière du module NAT permet de répartir la charge entre plusieurs serveurs.

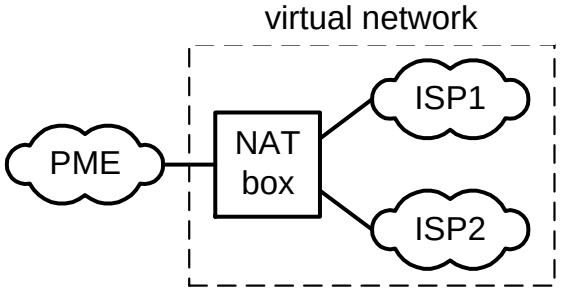


Les serveurs X et Y associés au module NAT forment un serveur virtuel *(virtual server)*
Ainsi le premier client utilisera le serveur X, le second client Y,…

Comment mesurer la charge des serveurs ?
- Le module NAT peut compter les paquets
- Le module NAT peut mesurer les temps de réponse
- …

h) NAT *router*

Imaginons une PME qui dispose de 2 accès à *internet* pour des raisons de sécurité :



Le module NAT peut décider d'utiliser ISP1 *(internet service provider)* ou ISP2 selon différents critères :
- disponibilité
- taux de charge
- tarification
- temps de réponse
- …

Chaque poste de travail ne voit qu'un seul routeur *(default router)* ; donc qu'un seul réseau *(virtual Network).*

i) Difficultés

Certains protocoles comme FTP, DNS, … échangent des informations relatives à l'adresse IP que le module NAT doit modifier.