

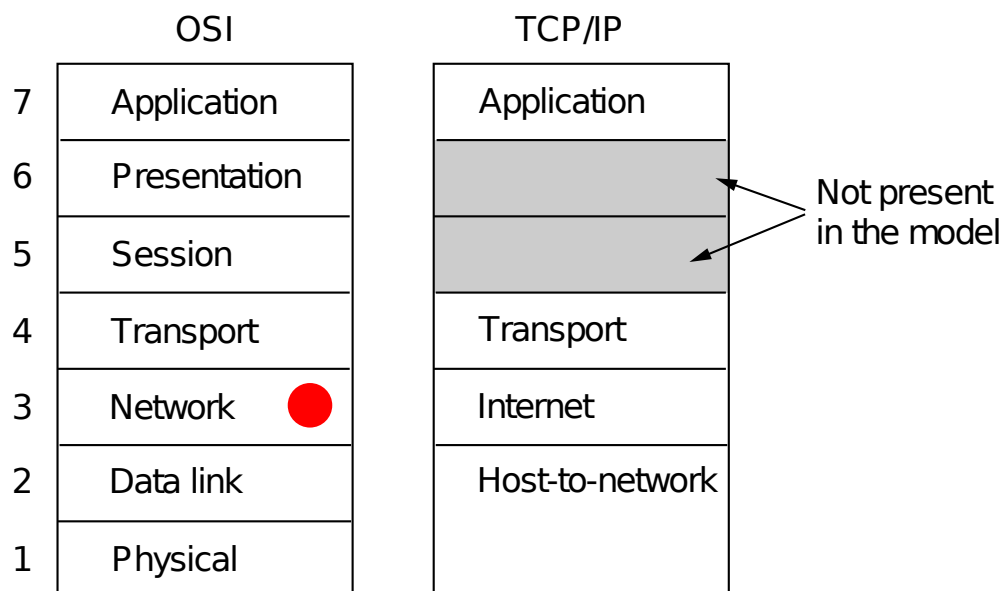
Séance 9 – Routage

Notes de Cours

LE ROUTAGE RASSEMBLE LES TECHNIQUES ISSUES DE L'ALGORITHMIQUE DISTRIBUÉE permettant de transférer les datagrammes de proche en proche, ainsi que de maintenir à jour, en fonction de l'évolution réelle et inévitable de la topologie du réseau, les informations permettant de réaliser ce transfert de manière optimale.

1 Routage

1.a "Vous êtes Ici"



1.b Définition du Routage

Définition : **algorithme distribué** ayant pour objectif d'acheminer des données depuis une source jusqu'à une destination.

Difficultés :

- le réseau évolue, son graphe est *dynamique*
- impossible de maintenir *localement* une connaissance **complète** et **en temps réel** de la topologie *globale*.
- => deux problématiques

 1. acheminement à l'aide d'une information **locale**, présente immédiatement sur le routeur.
 2. maintenance (en parallèle) de ces informations locales par **échange global** d'informations

1.c Aspect Algorithmique

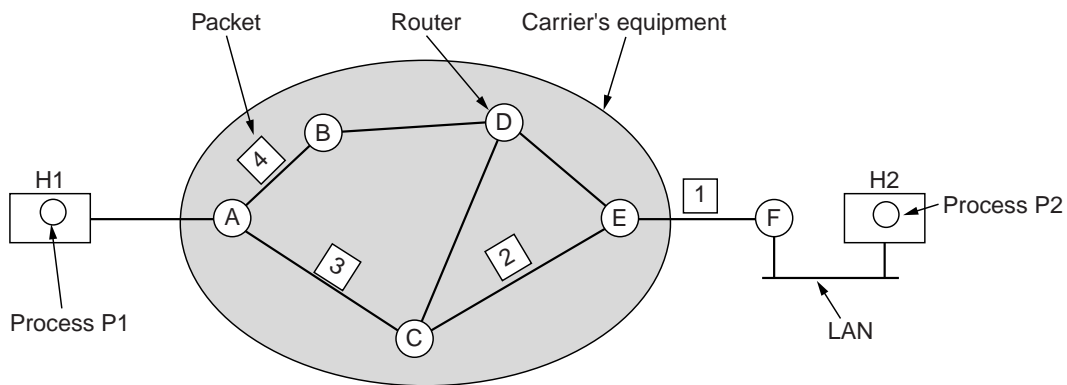
Acheminement — acheminement **de proche en proche**

- prend en compte l'information sur la destination ou l'origine (paquets vs circuits)
- rapidité de l'algorithme de traitement => **simplicité** de la *structure de données* => *tables de routage*
- nextHop [dest] vs nextHop (dest)

Maintenance — algorithmes distribués réagissant aux modifications topologiques

- la propagation des mises-à-jour se fait nécessairement avec du retard
- stabilité vs réactivité

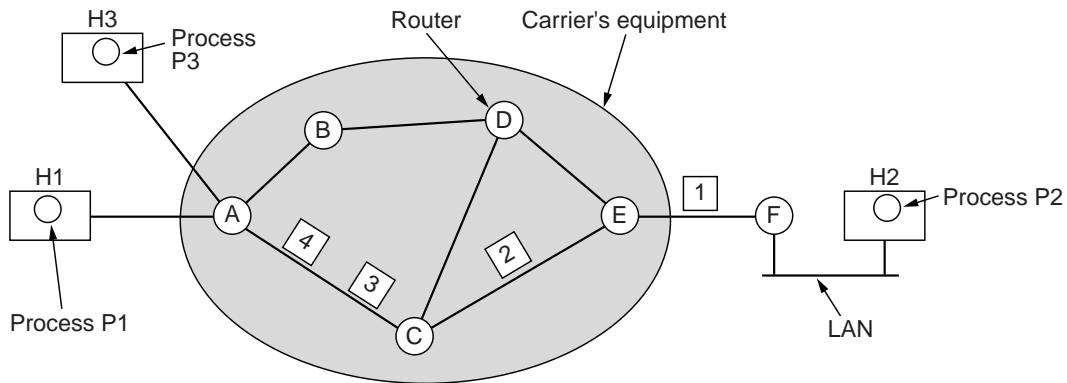
1.d Commutation de Paquets



| A's table | | C's table | E's table |
|---|-------|-----------|-----------|
| initially | later | | |
| A - | A - | A A | A C |
| B B | B B | B A | B D |
| C C | C C | C - | C C |
| D B | D B | D D | D D |
| E C | E B | E E | E - |
| F C | F B | F E | F F |
| <div style="border-top: 1px solid black; width: 50%; margin: 0 auto;"></div> Dest. Line | | | |

Ex : TCP/IP

1.e Commutation de Circuits



| A's table | | C's table | | E's table | |
|-----------|---|-----------|---|-----------|---|
| H1 | 1 | C | 1 | C | 1 |
| H3 | 1 | C | 2 | C | 2 |
| In | | Out | | | |

Ex : ATM, MPLS

2 Principes du Routage

2.a Tables de Routage pour la Commutation de Paquets

Principe général : faire correspondre adresse et *segment local*

— interface physique ↔ segment local

Algorithme simplifié

- le destinataire appartient au même réseau physique => envoi direct (avec ARP)
- sinon transmettre le paquet à un hôte (*dans le même réseau physique*) plus proche de la destination => "récursion"

Rappel : Comment déterminer si une adresse IP appartient au *même* réseau physique?

=> *masque de sous-réseau*

2.b Optimalité du Routage

Rôles différents :

- routeur
- hôte

Compacité des tables =>

- adresses regroupées en sous-réseaux => *plage d'adresses*
- *route par défaut*.

- Indication de route efficace :

- plus court chemin
- limite la congestion
- Temps de stabilisation après changement topologie.

2.c Tables vs Plan d'adressage

Une table de routage est toujours (schématiquement)

| Destination | sortie |
|----------------|-----------------------|
| plage1 | sortie1 |
| plage2 | sortie2 |
| ... | ... |
| <i>default</i> | passerelle par défaut |

Une plage d'adresse est un intervalle d'adresses, organisé avec CIDR :

- préfixe /pp
- machine : sur 32 – pp bits

L'espace des adresses peut être vu comme étant décomposé de manière arborescente.

Le plan d'adressage doit viser à limiter la taille des tables.

2.d Exemples : Tables de Routage Réelles

```

ras> ip route status
Dest          FF Len Device  Gateway      Metric stat Timer  Use
14.4.232.1    00 32  mpoa00     14.4.232.1   1      03a9 0     0
192.168.1.0   00 24  enet0      192.168.1.1  1      041b 0
16279
default     00 0   mpoa00     ChangeMe     1      00ab 0
11209

$ /sbin/route
Table de routage IP du noyau
Destination Passerelle  Genmask          Indic Metric Ref Use Iface
route1.fai  *           255.255.255.255 UH    0     0   0 ppp0
127.0.0.0   *           255.0.0.0       U     0     0   0 lo
default    route1.fai  0.0.0.0         UG    0     0   0 ppp0

$ netstat -nr
Table de routage IP du noyau
Destination Passerelle  Genmask          Indic MSS Fenêtre irtt Iface
192.168.1.0 0.0.0.0    255.255.255.0   U     0 0         0 eth0
0.0.0.0     192.168.1.1 0.0.0.0         UG    0 0         0 eth0

```

2.e Interprétation

| Intervalle de destination | | routeur le plus proche |
|----------------------------|----------------|---------------------------------|
| <i>Adresse</i> | <i>Masque</i> | <i>Passerelle</i> |
| IP locale 1 | masque local 1 | interface locale 1 |
| IP locale 2 | masque local 2 | interface locale 2 |
| ... | ... | ... |
| IP 1 | masque 1 | IP voisin 1 |
| IP 2 | masque 2 | IP voisin 2 |
| ... | ... | ... |
| toutes les autres adresses | | <i>IP passerelle par défaut</i> |

2.f Commutations de Paquets : Algorithmes

Algorithmes de Routage :

- Acheminement d'un paquet
- Maintenance de la structure de données distribuée

Contraintes

- Décentralisation et expansion Internet : => algorithmes efficaces
- Multiples constructeurs : => standardisation et conseils d'implémentations (RFC)
- Deux techniques principales (basé sur le *coût* associé à un lien) : Vecteur de distance vs Etat des liens

2.g Vecteurs de Distance

- Un nœud N connaît le coût pour joindre chacun de ses voisins
- Il transmet cette information à tous ses voisins
- Chaque voisin additionne cette information au coût pour joindre N et retient le total minimum pour déterminer le meilleur intermédiaire pour chaque destination.
- On répète jusqu'à ce que les tables convergent (ie ne soient plus modifiées)

2.h État des Liens

Pour chaque interface

- adresse IP
- masque de sous-réseau
- type de réseaux auxquels celle-ci est connecté
- adresses des routeurs connectés
- débit
- ...

2.i Algorithme

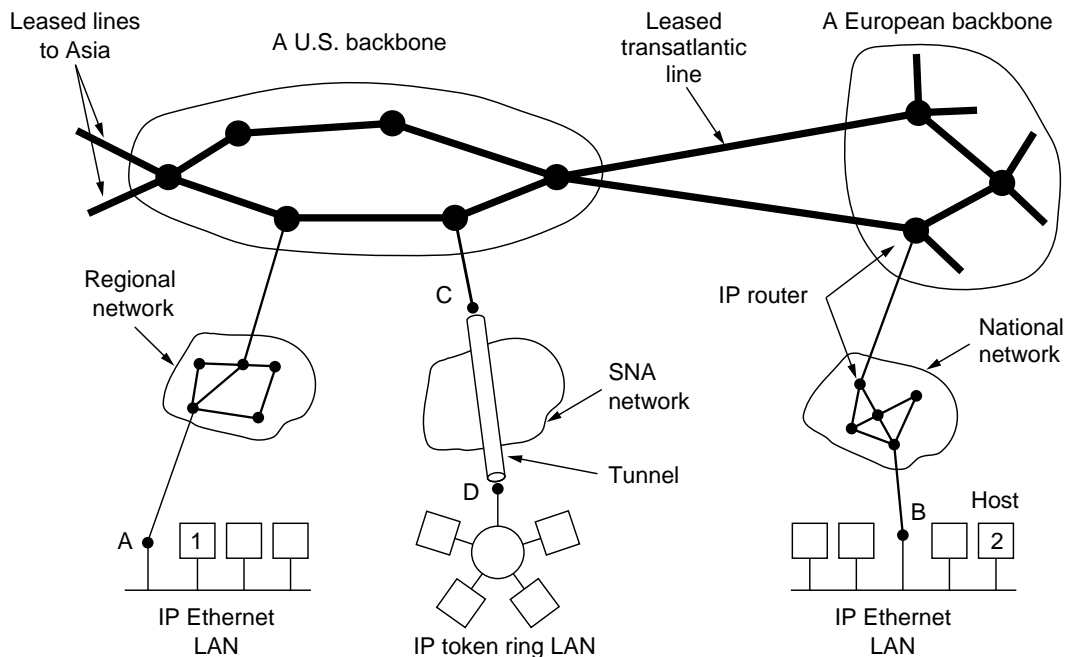
Le but est de maintenir *le plus court chemin* pour chaque destination connue :

- A chaque changement, un routeur "publie" sa table d'état de tous ses liens

- La diffusion est faite par inondation
- A partir des données obtenues, l'arbre des plus "courts" chemins est calculé *localement* par algorithme de Dijkstra (ou variante)
- La table de routage indique pour chaque adresse le routeur suivant (correspondant au plus court chemin)

3 Exemples de Protocoles de Routage : IPv4

3.a Topologie et Routage IP



Internet utilise la *commutation de paquets*. Les algorithmes de routage sont hiérarchiques.

3.b Communication Hors-Bande : ICMP

Internet Control Message Protocol permet d'envoyer des informations "hors-bande" à l'expéditeur d'un datagramme.

- Destination unreachable : problème de routage
 - Time exceed : le TTL atteint 0
 - Parameter problem : entête incorrect
 - Redirect : rerouter => apprendre la géographie à un routeur
 - Echo : demander à une machine si elle est en ligne => *ping*
 - Echo reply : oui je suis en ligne
 - Timestamp request : Echo avec horodatage
 - Timestamp reply : Echo reply avec horodatage
- => Commandes `ping`, `tracert`, ainsi que `ping6` et `tracert6`...

3.c Mise à Jour des Informations de Routage I

— Statiquement :

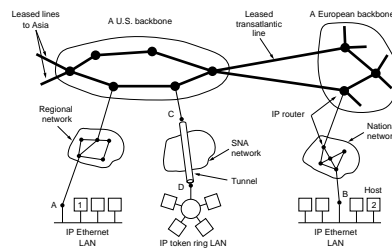
Gestion “manuelle” des routes, en salle TP (debian/ubuntu) :

- `ip addr ifconfig`
- `ip route add, route add`
- `ifup, ifdown, configuration via /etc/network/interfaces`
- pour *activer le routage* :

```
$ sysctl -w net.ipv4.ip_forward=1
```

pour que cela soit **permanent**, dans `/etc/sysctl.conf`, ajouter
`net/ipv4/ip_forward=1`

3.d Mise à Jour des Informations de Routage II

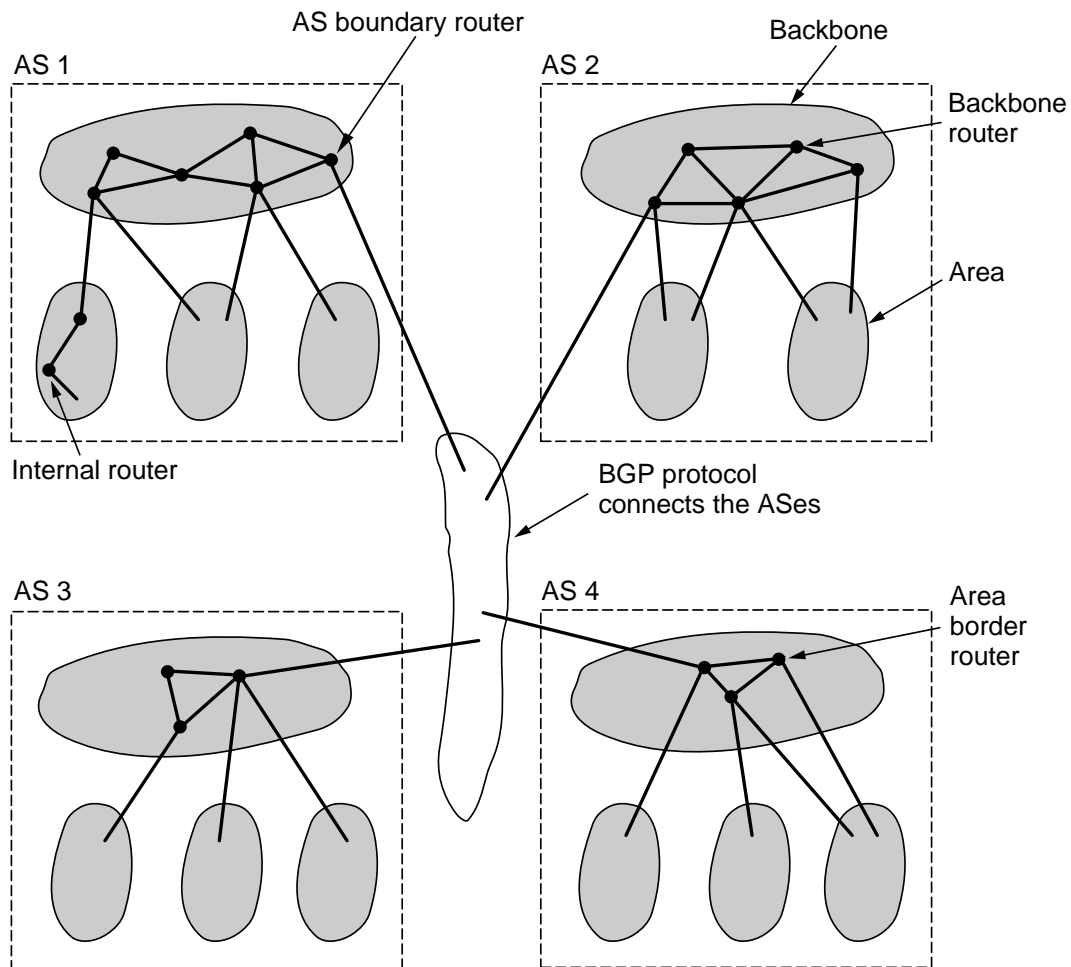


— Dynamiquement :

- *impossibilité d'échanger toutes les routes*
- protocoles pour le routage hiérarchique
- parties de réseau organisées de manière autonome
 - en *intérieur* : protocoles IGP Ex : RIP, OSPF, IGRP
 - en *extérieur* : protocoles EGP Ex : EGP, BGP

3.e Architecture Globale

Les réseaux sont organisées en *systèmes autonomes (AS)* **géré par une même institution**. La *politique de routage* est y cohérente.



Les routeurs d'un même système autonome partagent la même table BGP.

Les zones sont organisées en général avec

- un *backbone* : réseau central à *haut débit*
- des *aires* où l'on utilise un algorithme de routage interne

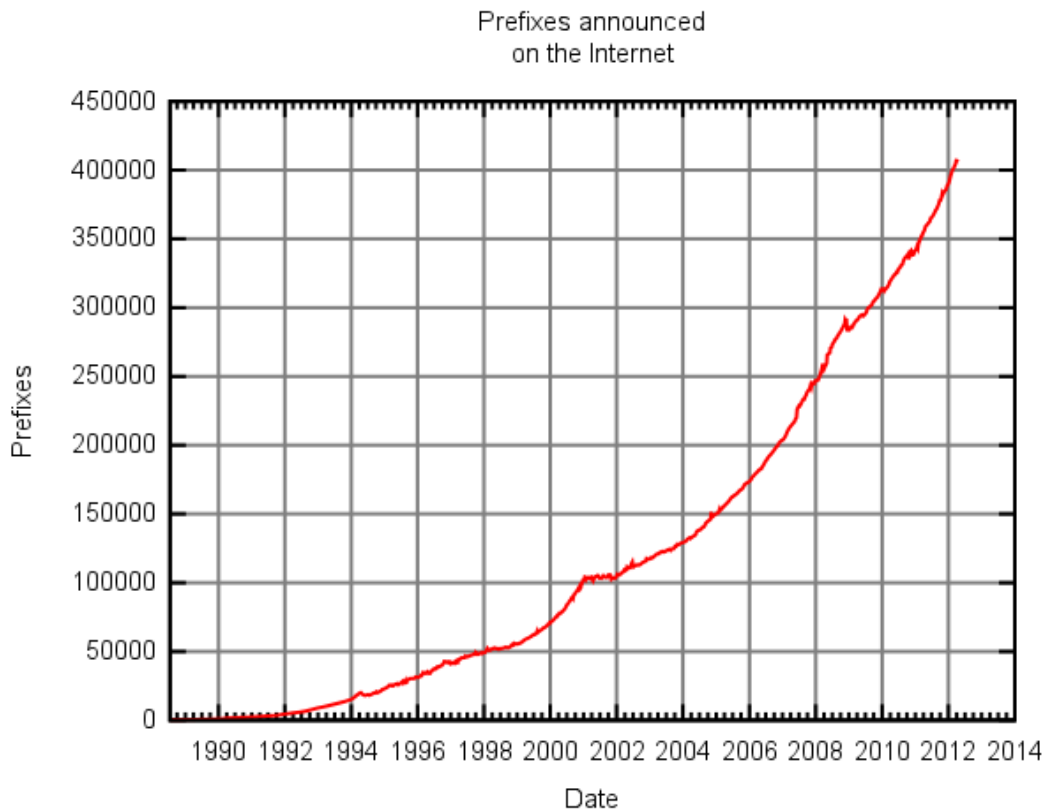
Une AS possède un *ASN* (sur 32 bits) délivré par un RIR (comme les adresses IP).

Une AS possède une politique de routage cohérente.

3.f Préfixe BGP : Le nombre de préfixes a augmenté

Un *préfixe BGP* désigne un ensemble d'adresses.

depuis les débuts d'Internet :



3.g Routes BGP

Une *route BGP* est une séquence de système autonome (désignés par leur ASN) permettant de rejoindre un ensemble d'adresses donné.

Ceci permet de savoir quel est le routeur suivant mais également l'ensemble des réseaux traversés pour parvenir à destination.

3.h BGP et Politiques de Routage

Le *Border Gateway Protocol* permet d'annoncer et de choisir des routes de manière sélective :

- sécurité (rappel : l'information TCP/IP circule en clair)
- géopolitique
- économique (facturation, échange de trafic)

3.i Organisation Tarifaires

Lorsque deux entités économiques s'échangent du trafic

- si facturation : *transit*
- si (presque) pas facturation : *appairement (peering)* ou échange de trafic

Tier 1 Peering uniquement

Tier 2 Mixte (transit et peering)

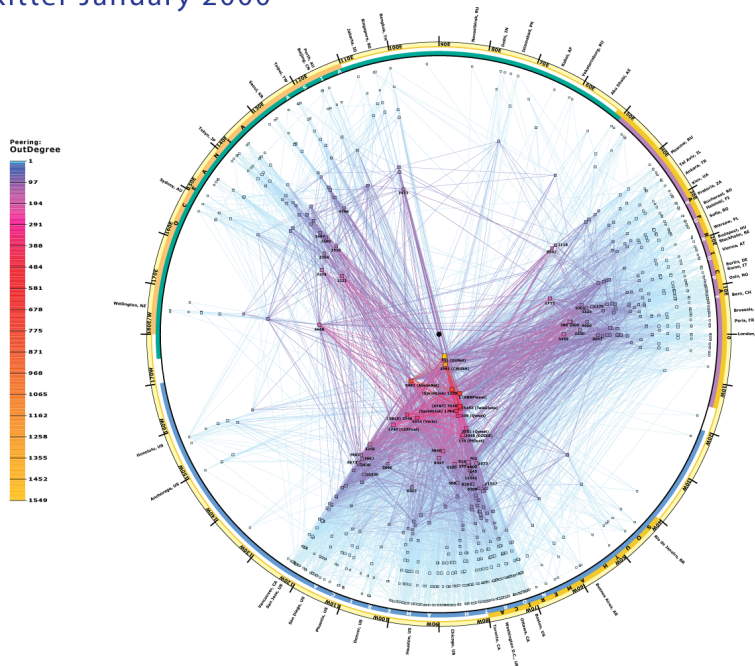
Tier 3 Transit (=> facturation) => votre FAI

3.j Le coeur d'Internet

Les AS dont les routeurs ont l'ensemble des routes d'Internet
(source CAIDA.org)

CAIDA's IPv4 AS Core AS-level INTERNET GRAPH

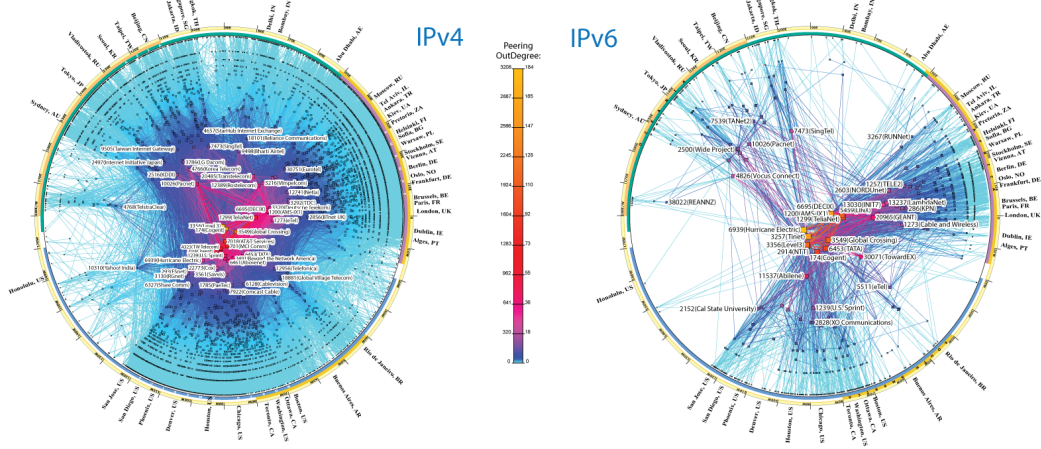
Skitter January 2000



copyright © 2000 UC Regents. all rights reserved.

**CAIDA'S IPv4 & IPv6 AS Core
AS-level INTERNET GRAPH**

Archipelago April 2011



Copyright © 2012 UC Regents. All rights reserved.

3.k OSPF vs RIP

- RIP (Routing Information Protocol)
 - vecteur de distance
 - diamètre limité (15)
 - pas de routage sans classe
 - diffusion régulière de toute la table de routage
 - convergence en minutes
 - métrique rudimentaire : nombre de *hop*
 - réseaux non structurés
- RIP2
 - améliorations
 - toujours diamètre de 15 et lente convergence

3.1 OSPF vs RIP

- OSPF (Open Shortest Path First)
 - état des liens
 - diamètre non borné
 - gestion fine des adresses
 - propagation des modifications par diffusion IP
 - meilleure convergence
 - métrique permettant l'équilibrage de charge
 - agrégation des réseaux en aires
 - authentification des routeurs
 - gestion des routes externes (injectées par BGP)

3.m Historique OSPF

- début des discussions : 1988
- début de la formalisation : 1991
- RFC 2328 : 1998
- Synopsis :
 - protocole non propriétaire pour remplacer RIP
 - *mais* plus complexe à configurer/maintenir

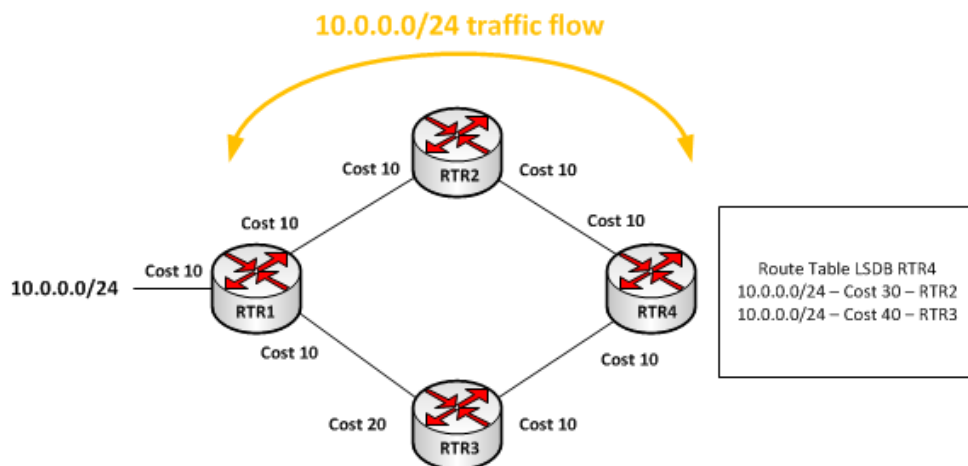
3.n Plus Courts Chemins et Métrique OSPF

La *métrique* reflète le coût de traversée d'un lien, pour que le plus court chemin soit vraiment *le plus court* :

$$\delta(\text{interface}) = \frac{10^8}{\text{débit}_{\text{bits/s}}}$$

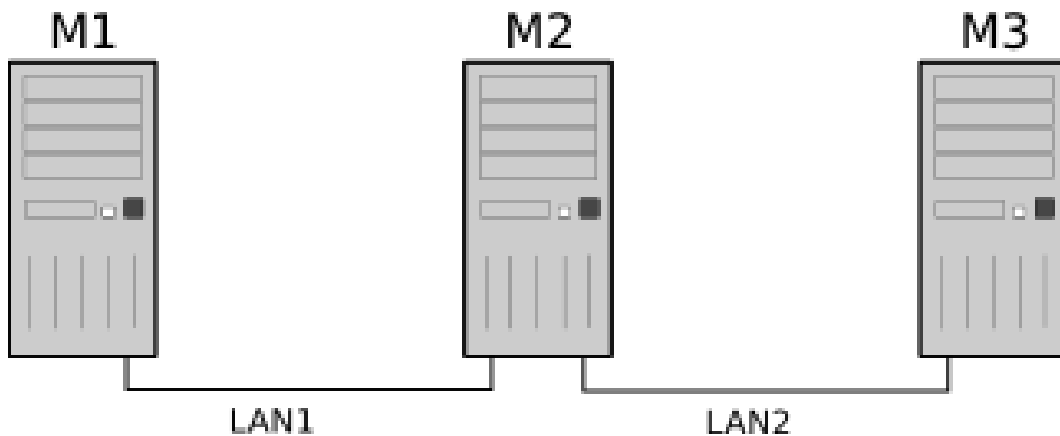
La métrique peut aussi être définie manuellement.

3.o Exemple



3.p Problème du Compte-à-l'infini

En utilisant uniquement les vecteurs de distance, il n'est pas possible de savoir exactement par où passe la route la plus courte. En particulier, il est impossible de savoir si elle passe par le nœud même.



En supposant une distance de 1 par hop, on a que M3 annonce une distance de 2 pour M1. Si M1 est hors ligne, alors M2 recevra toujours une annonce de M3 pour une distance de 2 pour M1. **Par conséquent**, M2 mettra son vecteur de distance (vis-à-vis de M1) à 3. **Par conséquent**, M3 mettra son vecteur de distance à 4. **Par conséquent**, M2 mettra son vecteur de distance à 5. Etc...

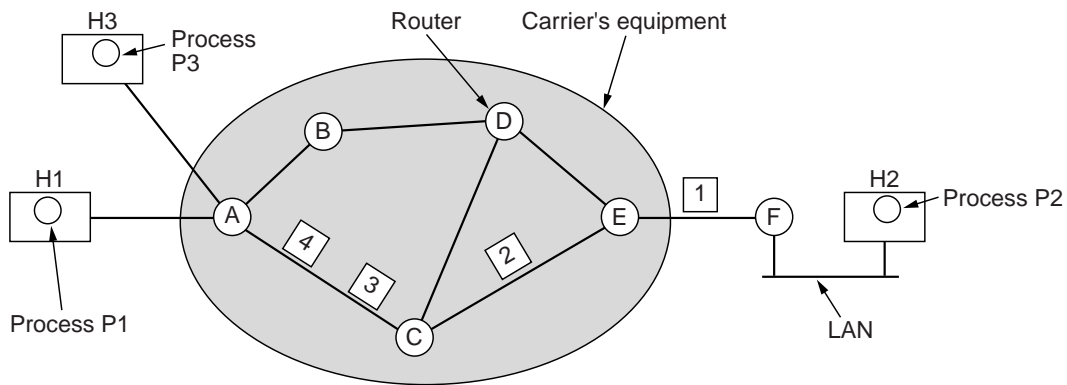
3.q Avancées Récentes

Des améliorations récentes de la technique de vecteur de distances permettent d'éviter le problème du *compte à l'infini*.

- EIGRP (Enhanced Interior Gateway Routing Protocol)
 - protocole de routage développé par Cisco
 - propriétaire
 - basé sur les vecteurs de distance,
 - **condition d'évitement des boucles**
- DSDV,AODV
- Babel

4 Routage en Commutation de Circuits

4.a Rappel du Principe



| A's table | | C's table | | E's table | |
|-----------|---|-----------|---|-----------|---|
| H1 | 1 | A | 1 | C | 1 |
| H3 | 1 | A | 2 | C | 2 |
| In | | Out | | | |

Ex : ATM, MPLS

ATM= Asynchronous Transfer Mode
 MPLS= Multi-Protocol Label Switching

4.b Le Protocole ATM

Historique :

- créé au CNET (Lannion, France) à partir de 1982
- adopté par l'ITU (International Telecommunication Union)
- répond initialement à des besoins spécifiques des opérateurs télécoms
- standards internationaux au sein de l'ATM Forum
 - UNI, LANE
 - MPOA
- utilisation
 - IPoA IP over ATM : transport de paquet IP par un opérateur télécom au sein de son réseau très haut débit
 - offre *triple play* les différentes qualités de service nécessaires à une offre télévision + téléphone + internet peuvent être satisfaites en ATM

4.c Réseau de cellules

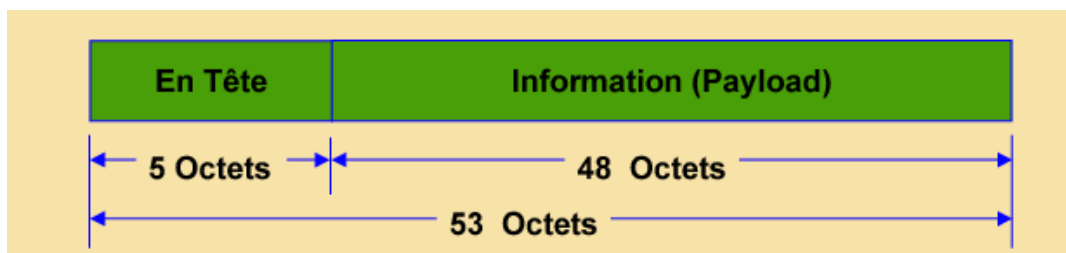
- découpage en cellules de **taille fixe**
- multiplexage temporel
- un circuit = un sous-intervalle dédié aux cellules du circuit

4.d Avantage des cellules

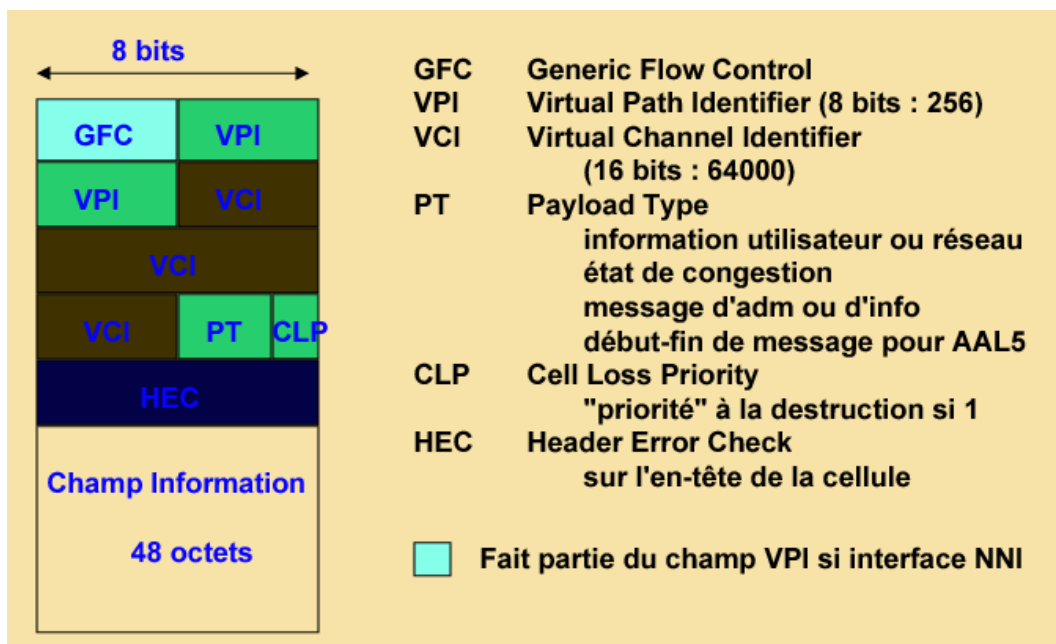
- la taille fixe simplifie la gestion du multiplexage
=> implémentation matérielle
- adapté aux réseaux homogènes (typiquement de télécoms)

4.e Cellule ATM

- cellule de petite taille
 - compromis débit latence

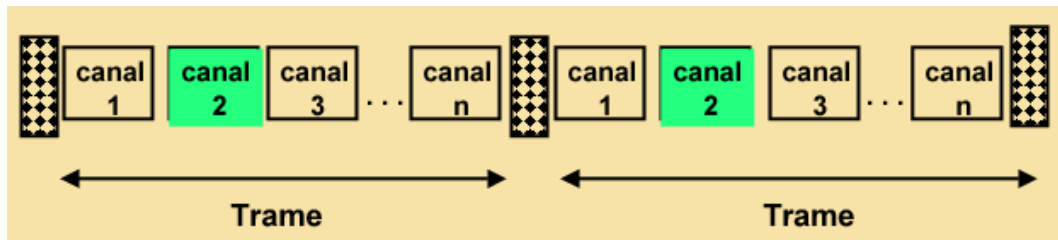


4.f Entête

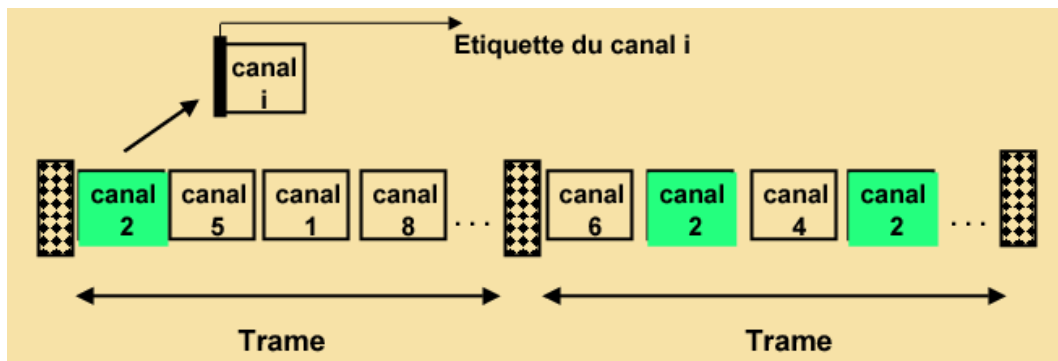


4.g Multiplexage

synchrone



asynchrone



Avantage : tous les sous-intervalles sont bien utilisés

4.h Principes d'ATM

- sur une connexion les cellules suivent toutes le même chemin
- la qualité de service peut être spécifique
- 2 couches principales (OSI 2-3)
 - ATM
 - AAL (ATM Adaptation Layer)

La couche AAL permet d'adpter le trafic à une qualité de service spécifique.

AAL1 voix

AAL2 audio-vidéo (débit variable)

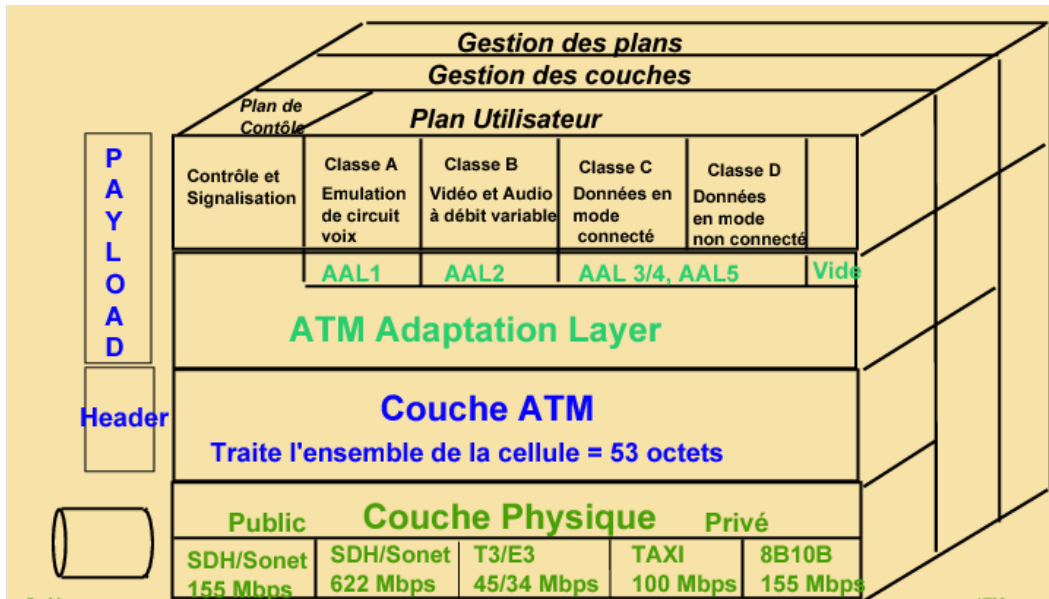
AAL3/4 données en mode connecté

peu utilisé

AAL5 données en mode non-connecté

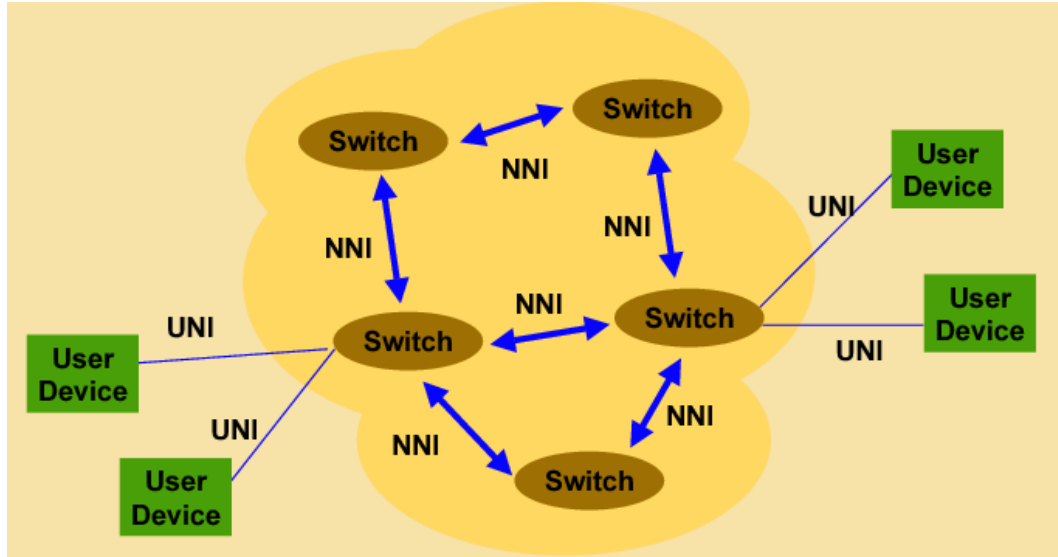
haut débit stupide informatique : utilisé pour IPoA

4.i En couches



4.j Structure du Réseau ATM

coeur du réseau (NNI) vs utilisateur (UNI)



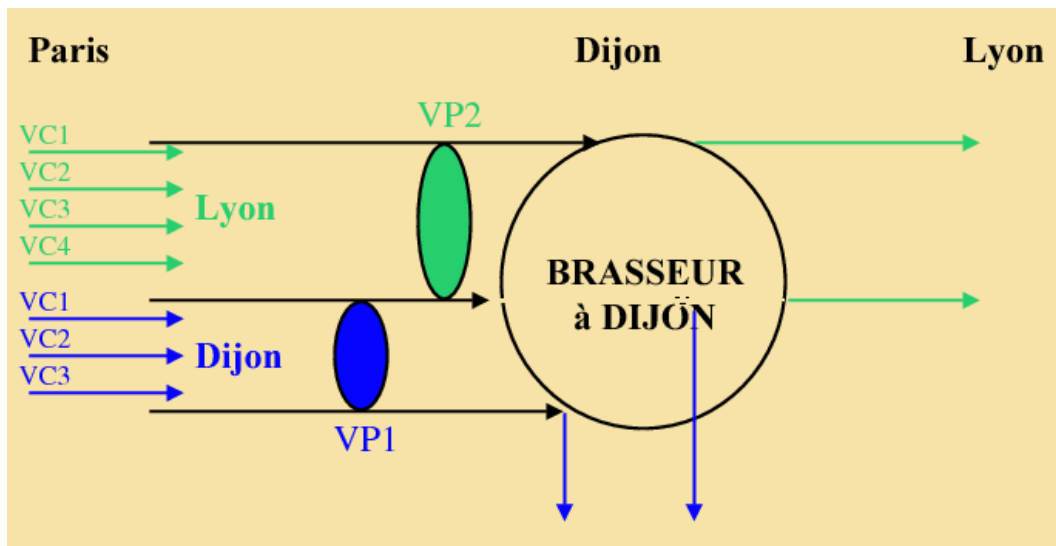
NNI Network Node Interface

UNI User Network Interface

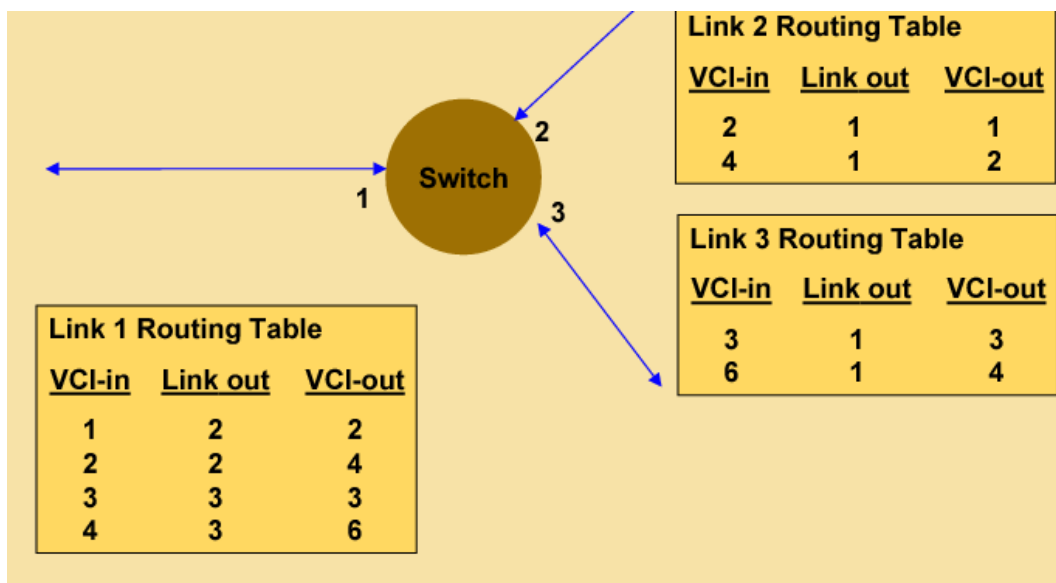
4.k Routage ATM

Les circuits sont structurés à 2 niveaux
Chemin virtuel (permanent) => VPI

Canal virtuel (à la demande) => VCI

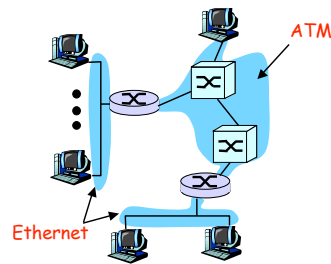


4.1 Commutateur ATM



4.m Utiliser ATM pour Internet

Le but de IP sur ATM est d'utiliser le réseau ATM de la même manière qu'IP peut être déployé au dessus d'un LAN ethernet 802.3.



Limitation Le réseau ATM n'a pas de mécanisme de diffusion, on ne peut donc configurer le réseau que *statiquement*.

4.n Le Protocole MPLS

On recherche le meilleur des deux mondes : MPLS

| | | | |
|------------|--------------------|-----------|----------------------|
| entête ATM | entête MPLS | entête IP | reste de la trame... |
|------------|--------------------|-----------|----------------------|

=> couche 2,5!

5 Routage : Conclusion

Le routage c'est

- un algorithme et une structure de données **distribués**
- organisation hiérarchique sur IP :
 - en interne : algorithmes basés sur un calcul distribué des meilleures routes
 - en externe : moins automatique, politiques de routage explicites car peu expressibles en terme de métrique.
- voir aussi *commutation de circuits*

6 Crédits

- Figures A. Tanenbaum. Libre d'utilisation pour l'enseignement
- Cours ATM par Unité Réseaux CNRS (UREC, 1997)
- Wikimedia CC-BY-SA