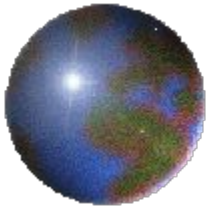
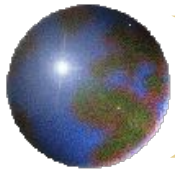
CENTRE NATIONAL
DE LA RECHERCHE
SCIENTIFIQUE



Réseaux, éléments de base : Modèle en couches de l'ISO TCP/IP

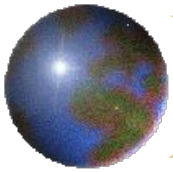
Formation SIARS



Auteur(s)

Cette présentation, élaborée dans le cadre de la formation SIARS, ne peut être utilisée ou modifiée qu'avec le consentement de ses auteur(s).

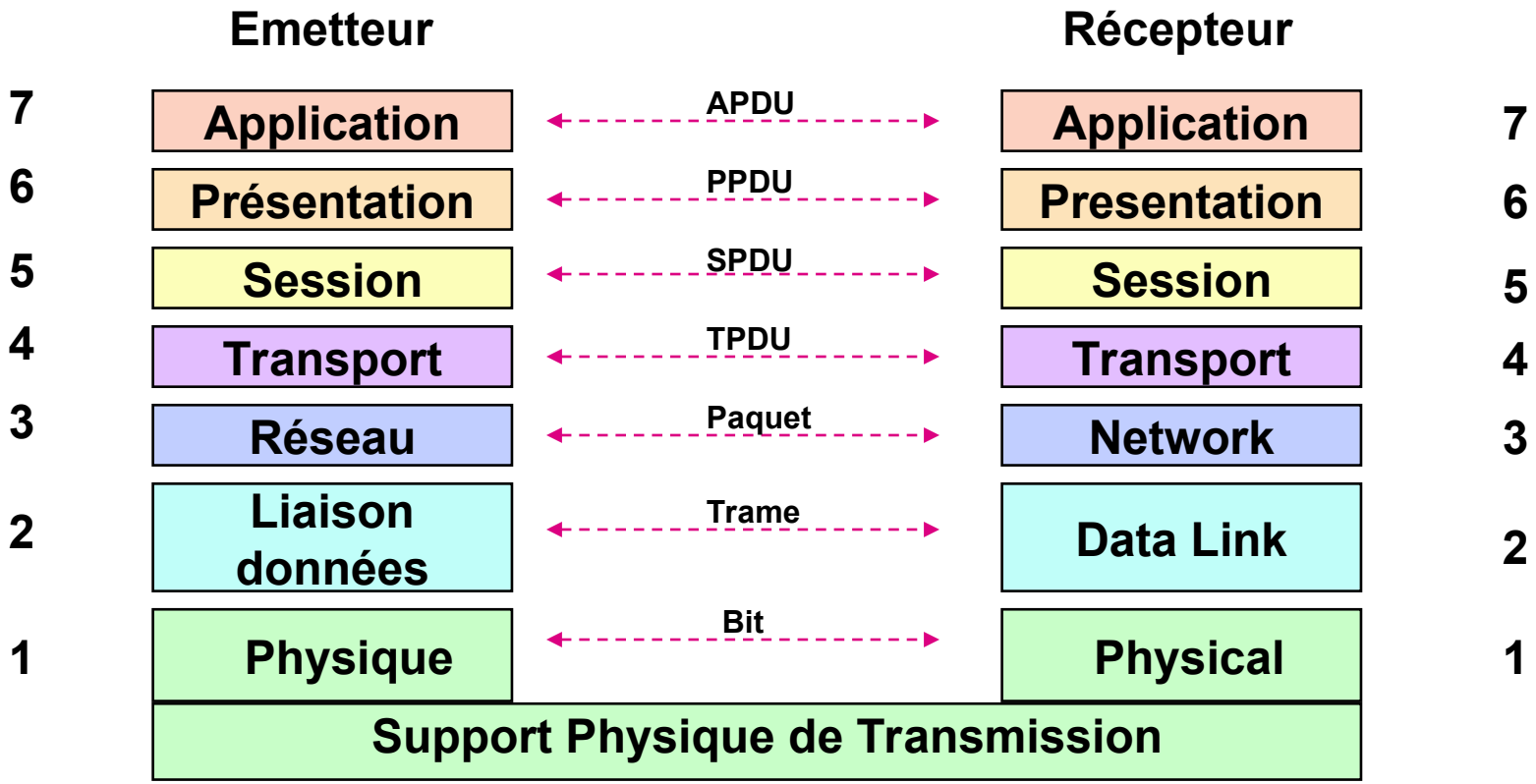
- Version 1.0 : 06/2002 : Denis.Pugnere@igh.cnrs.fr
-



Modèle OSI (sommaire)

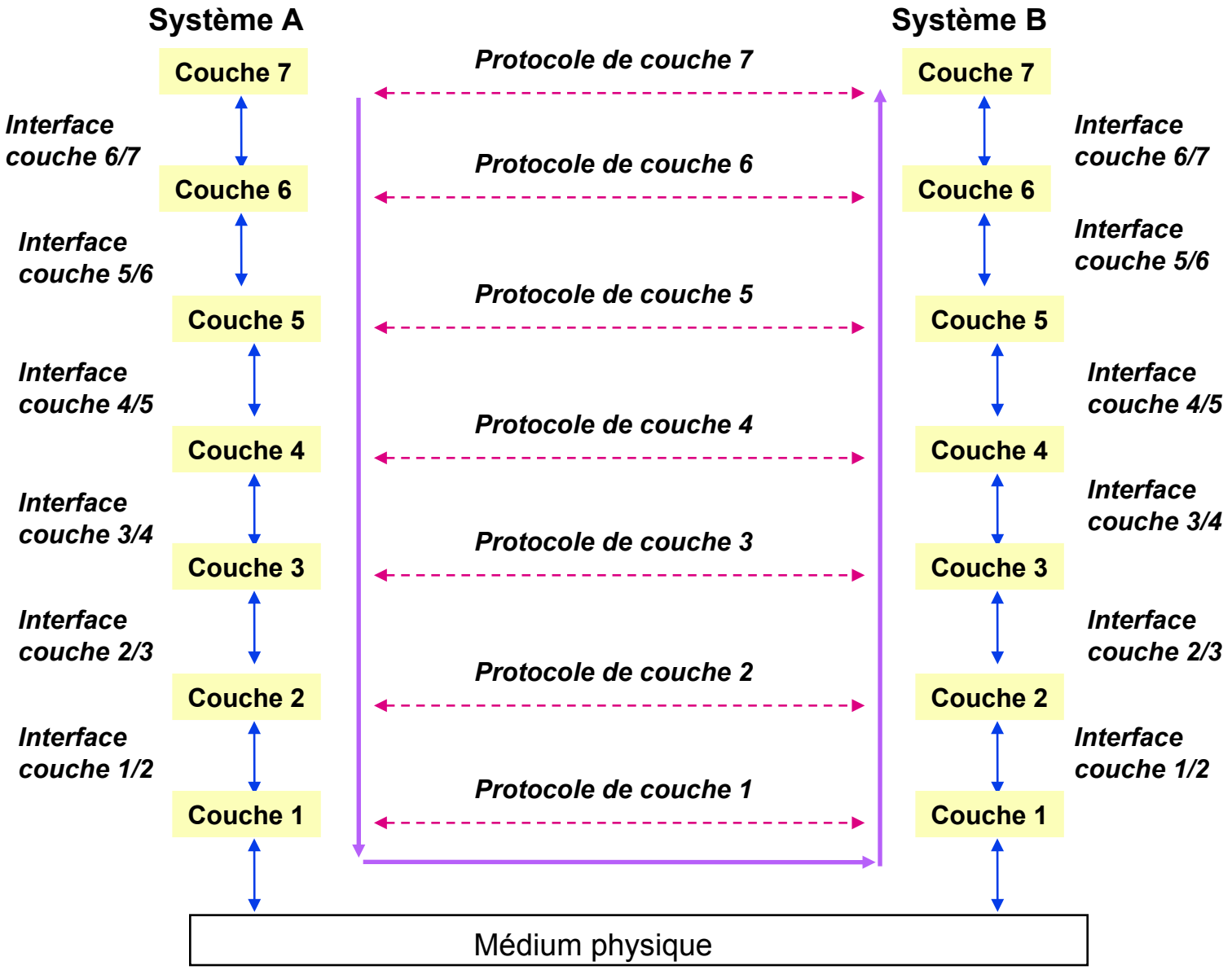
- Le modèle ISO (ISO 7498:1994)
- La couche physique
- La couche liaison
- Les couches réseau et transport
- Les couches supérieures

Le modèle de référence OSI : 7 couches



PDU : Protocol Data Unit

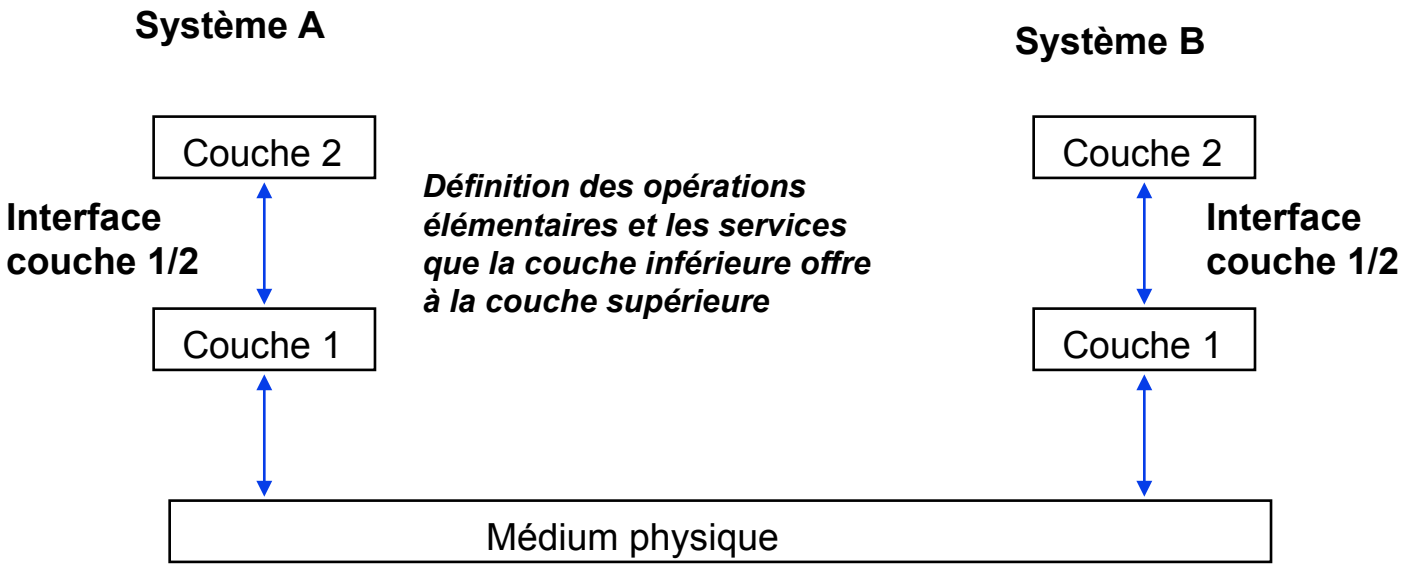
Communication entre couches



Animation...

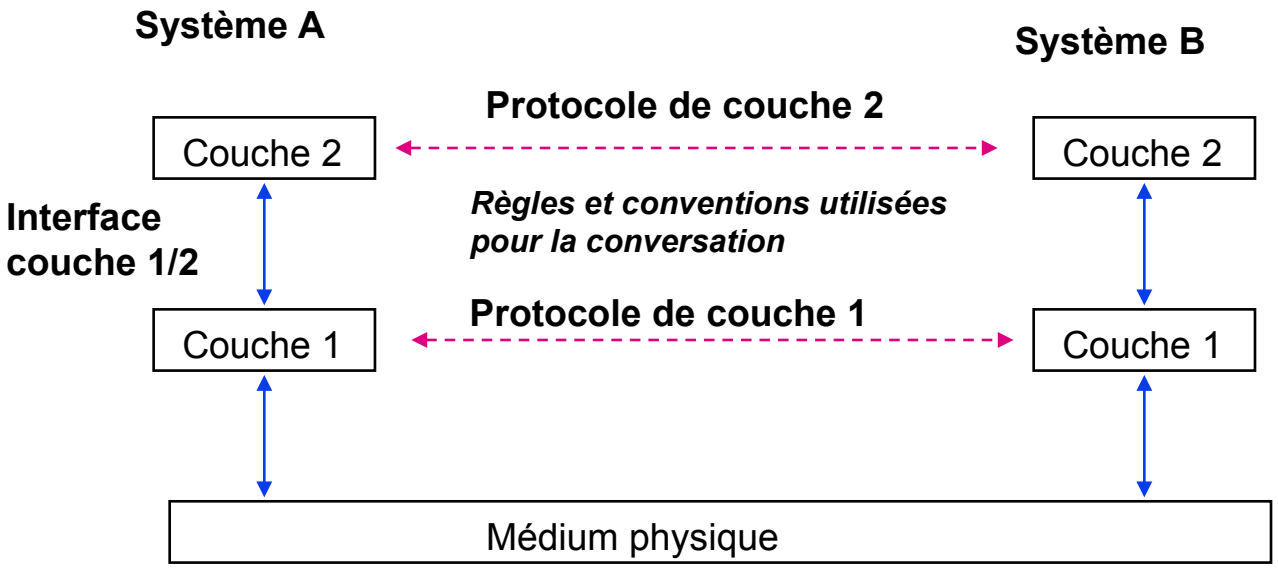
Communication entre couches : interface

- l'objet de chaque couche est d'offrir certains services aux couches plus hautes
- ces dernières ne connaissant pas la mise en œuvre de ces services

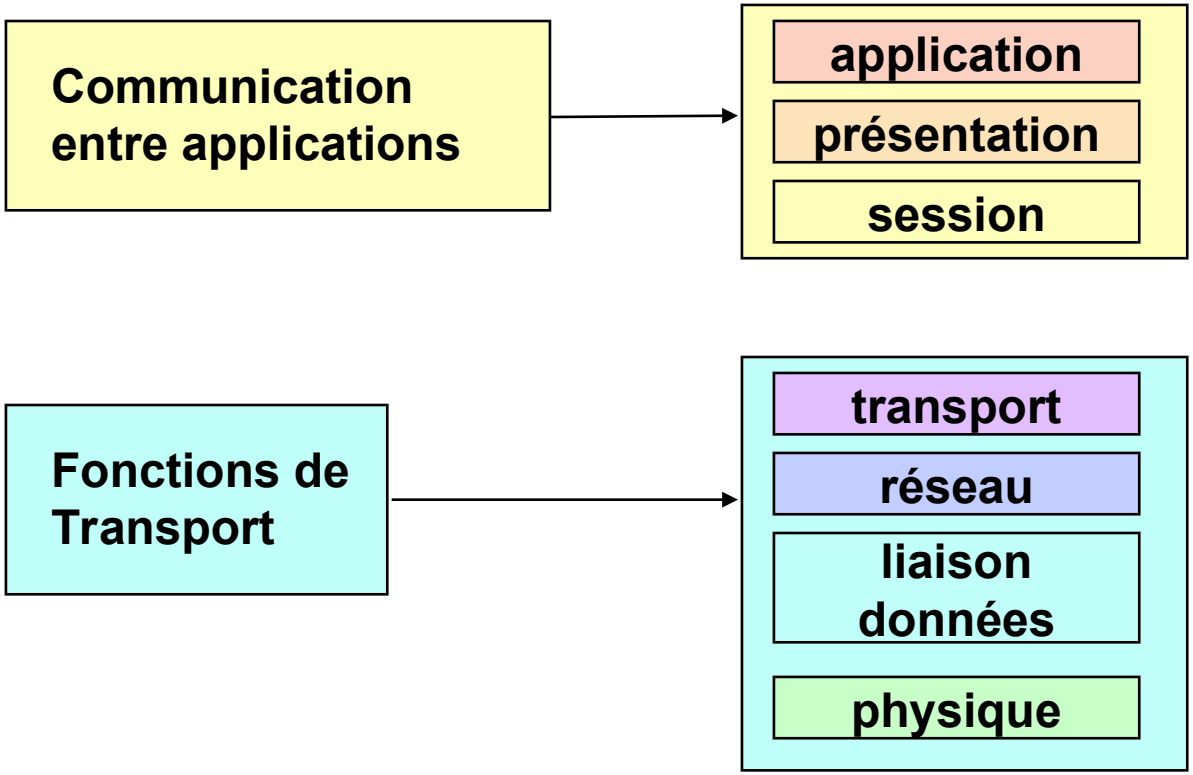


Communication entre couches : protocole

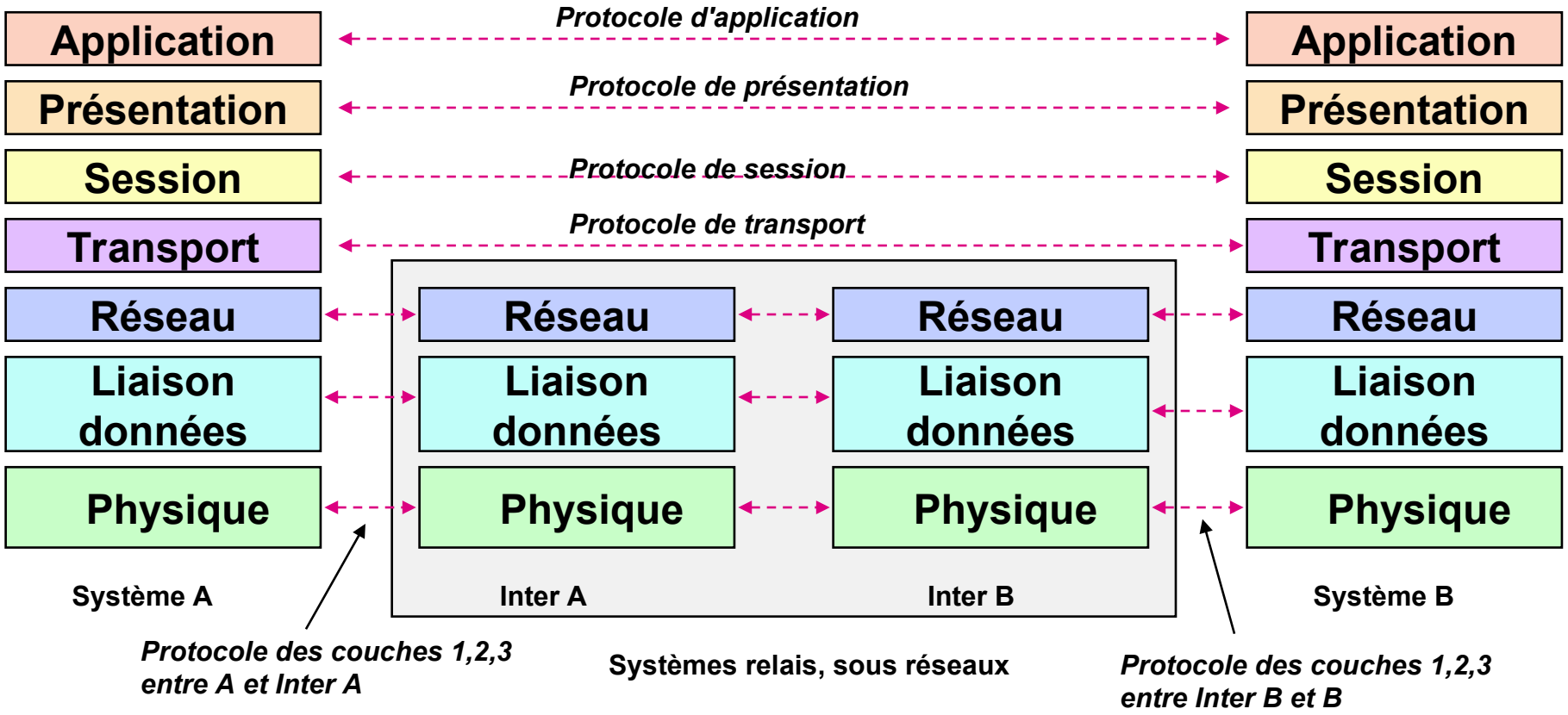
- Les applications sur les machines A et B se comprennent grâce à l ’utilisation d ’un protocole commun d ’une même couche.



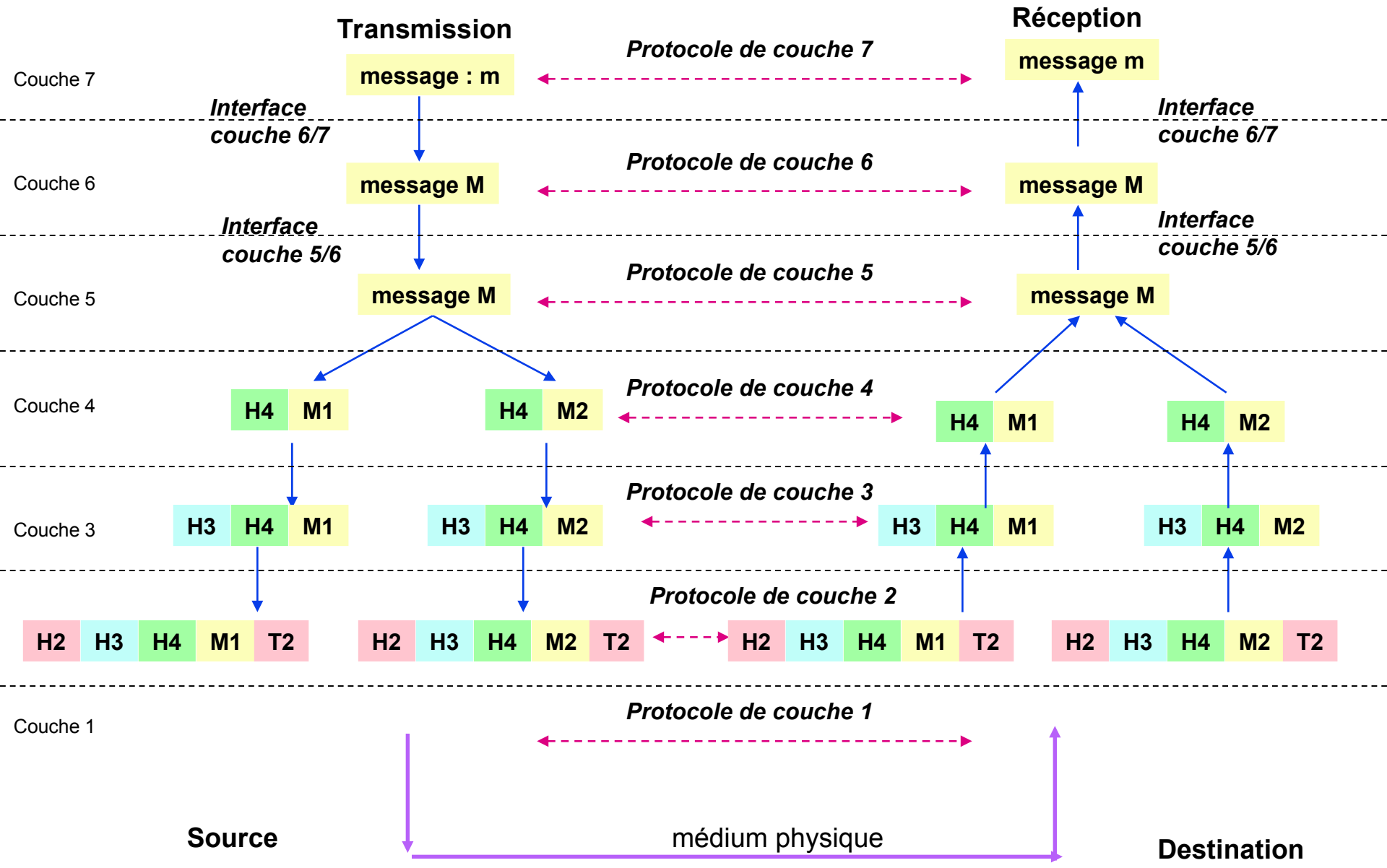
Les couches hautes et basses



Le modèle de référence OSI

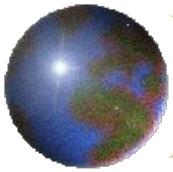


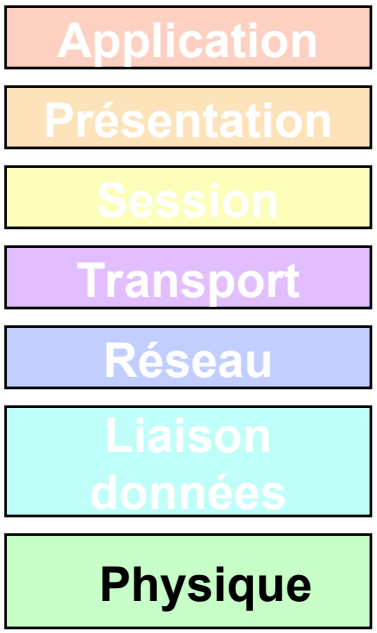
Cheminement des messages entre couches



Animation...

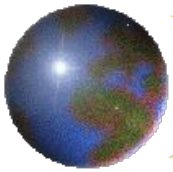
Couche physique





- Support Physique + Couche Physique
- Elle fournit les moyens mécaniques, électriques, fonctionnels, au maintien et à la désactivation des connexions physiques destinées à la transmission des éléments binaires entre entités de liaisons
- Transmission des bits sur un circuit de communication
- Éléments de la couche physique
  - Support physique
  - Codeurs, Modulateurs,
  - Multiplexeurs
- La conception de la couche physique peut-être réellement considérée comme faisant partie du domaine de l'ingénieur électronicien.
- Exemples de supports physiques : câble cuivre paire torsadée, câble cuivre coaxial, fibre optique monomode ou multimode, faisceau hertzien...

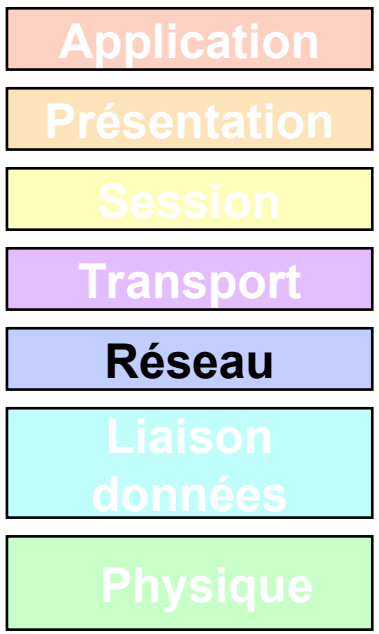
Couche liaison



Application
Présentation
Session
Transport
Réseau
Liaison données
Physique

- Utilise la couche physique
- Gestion de la liaison de données
  - données de l'émetteur en *trame de données,*
  - transmission des trames en séquence,
  - gestion des trames d'acquittement,
  - reconnaissance des frontières de trames envoyées par la couche physique.
- Détection et reprise sur erreur
  - régulation du trafic,
  - détection de collisions (conflit d'accès au média)
  - gestion des erreurs.
- Procédure de transmission (HDLC,LLC, DSC, ..)
- exemple de protocoles de couche 2 : 10BaseT, 100BaseT, 1000BaseX (IEEE 802.3z), 1000BaseTX (IEEE 802.3ab), ATM LLC

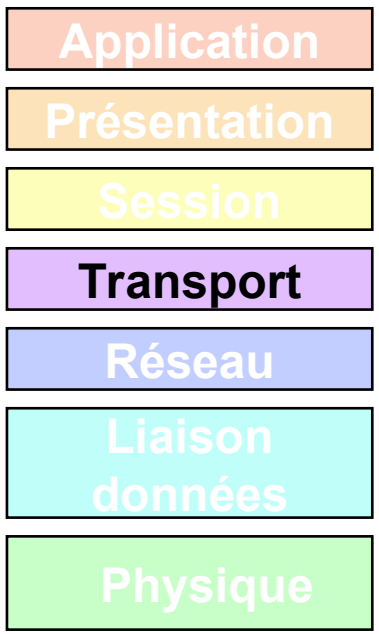
Couche réseau



- Fournit les moyens à la couche supérieure d'établir, de maintenir et de libérer des connexions de réseau entre des systèmes
  - gestion du sous-réseau,
  - acheminement des paquets de source vers la destination.
- Fonctionnalités
  - Adressage
  - Routage (statique, dynamique)
  - Contrôle de flux (congestion)
  - Comptabilité (comptage des octets transmis/reçus)
  - Fragmentation / ré assemblage des paquets
- Modes connecté/non connecté

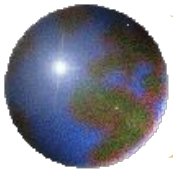
La couche réseau doit permettre l'interconnexion de réseaux hétérogènes

Couche transport

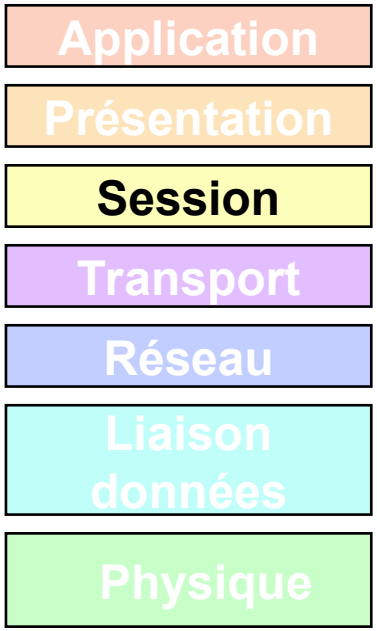


- Indépendance des réseaux sous-jacents
- Accepte les données de la couche session
  - les découpe éventuellement,
  - s'assure de l'ordonnancement
- Optimiser les ressources réseaux
- Fonctionnalités de bout en bout
  - multiplexage de plusieurs messages sur un canal
    - nécessité d'indiquer quel message appartient à quelle connexion.
  - Contrôle de flux (fenêtre d ’anticipation…)
  - gestion priorités
- Protocoles de Transport
  - TP0, 1, 2, 3 ou 4
  - TCP, UDP

Authentique couche de bout en bout

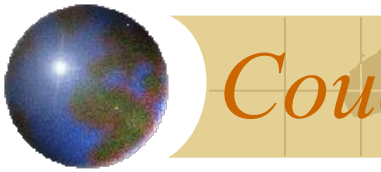


Couche session



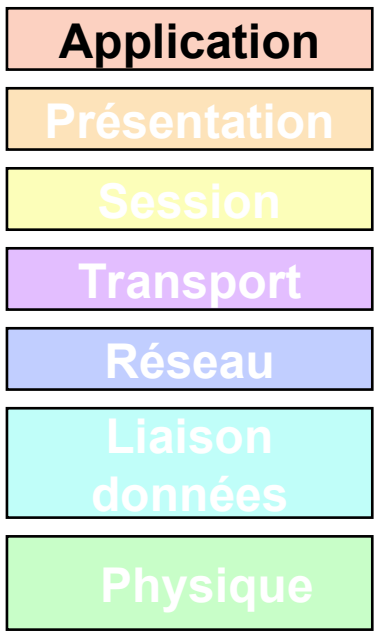
- Services évolués aux applications
- Responsable de la synchronisation
- Fonctions de type
  - Gestion du dialogue (bi- ou unidirectionnel) : qui est en train de parler ?
  - Points de reprise,
  - Retour arrière
- Orchestration
- Gestion des transactions

Couche présentation



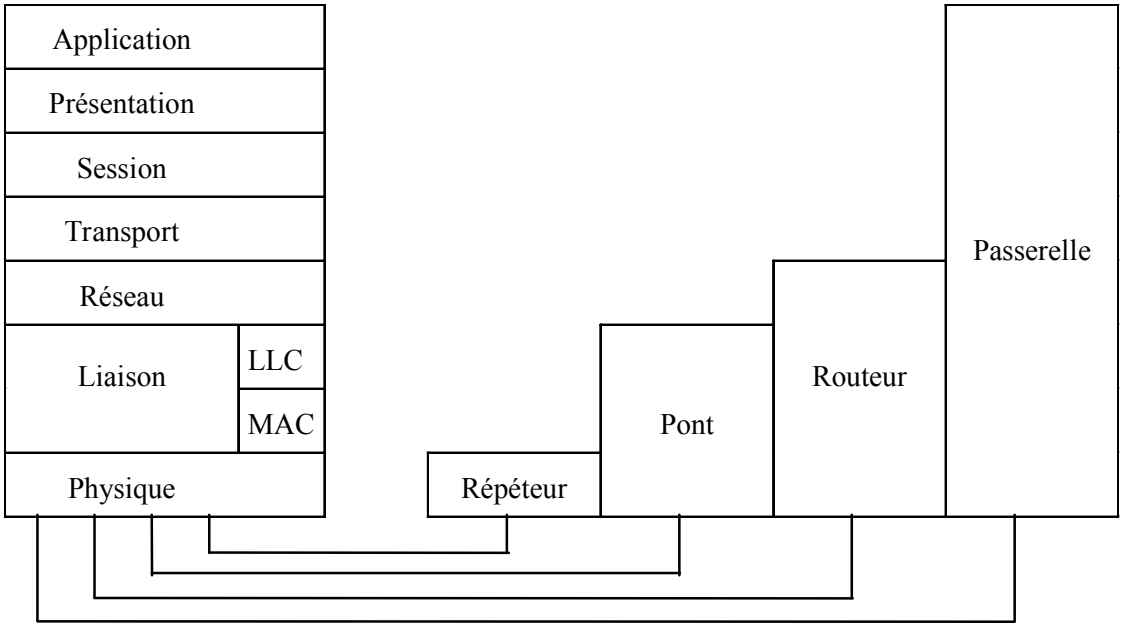
- S'intéresse à la syntaxe et à la sémantique des informations
  - Représentation des données transférées entre entités d'application,
    Représentation de la structure de données et représentation de l'ensemble des actions effectuées sur cette structure de données.
  - encodage des données dans une norme de présentation agrée permettant à des équipements de communiquer, par exemple :
    - ASCII, EBCDIC, Unicode
    - compression des données,
    - chiffrement.
- Exemple:
  - La syntaxe abstraite ASN.1 (ISO 8824, UIT X208) normalisée par l'ISO. utilisée dans la messagerie X400 et les annuaires X500.
  - XML, XSLT, SGML...

Couche application



- Elle offre aux processus d'application le moyen d'accéder à l'environnement OSI.
- Les processus d'application échangent leurs informations par l'intermédiaire des entités d'application
  - exemple : terminal de réseau virtuel, transfert de fichiers, courrier électronique, consultation des annuaires.

Correspondance Couches / éléments actifs



Répéteur

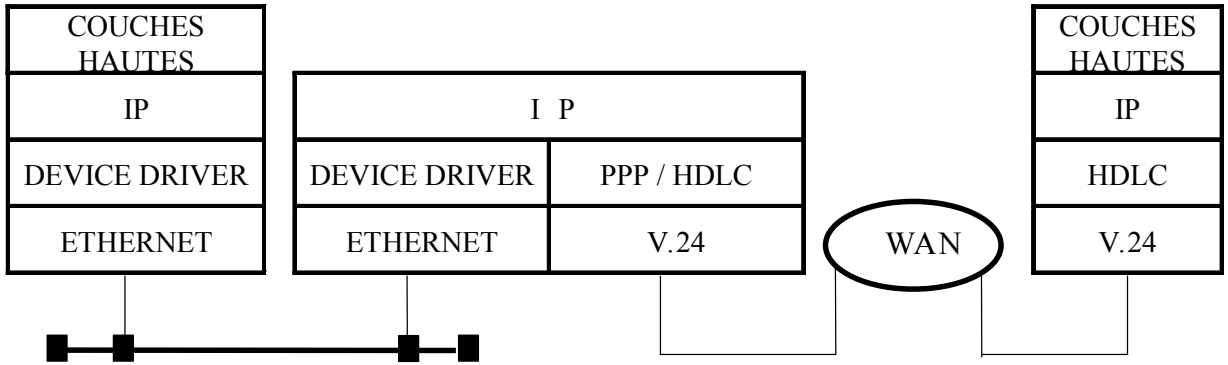
- Media identiques
- Régénération du signal :
  - Simple amplification
  - Régénération et restitution horloge
- Méthode d ’accès identique
- Pas / peu de paramétrages

Pont

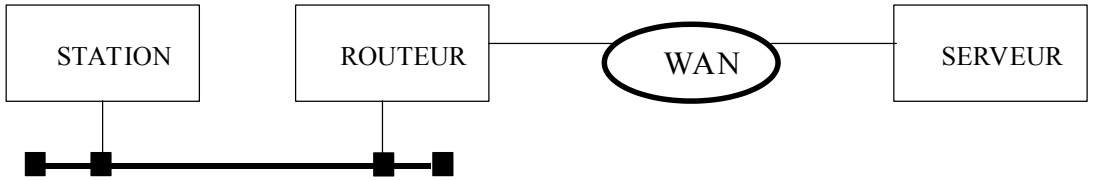
- Élément actif pour interconnexion locale ou à distance des réseaux
- ISO niveau 2
- Transparent aux protocoles des niveaux supérieurs
- Informations vérifiées :
  - Origine
  - Destination
  - Taille
- Possibilité de créer des segments et filtrer le trafic entre segments
- Faculté à apprendre topologie -> configuration automatique (@ MAC, spanning tree…)

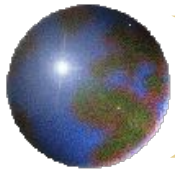
Pont/Routeur

Vue architecturale d'un pont / routeur distant



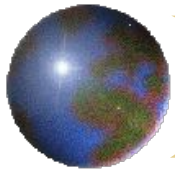
Vue physique





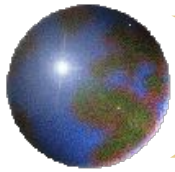
Routeurs

- Agissent au niveau 3 ISO
- Mécanismes évolués pour le transport des paquets
  - Routage
  - Contrôle de flux



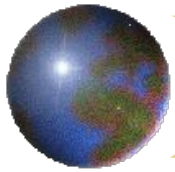
Passerelles

- Pour réseaux de protocoles différents
- Conversion de protocoles
- Agit sur couches supérieures du modèle ISO (couches 3 à 6)
- Services supérieurs à ceux des routeurs
- Utilisés pour les réseaux d ’architectures/protocoles différents
- Environnement dédiés
- Système d ’interconnexion lent et sophistiqué



Ethernet : Sommaire

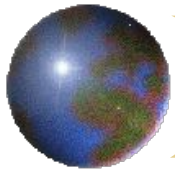
- principes de bases
- Ethernet
- Fast-ethernet
- Giga-ethernet



Ethernet, principes de base

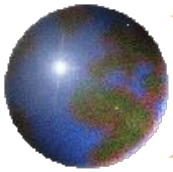
- Principes pour l'émission :
  - écoute préalable du support physique (CSMA : Carrier Sense Multiple Access)
  - détection des collisions (CD : Collision Detection)
  - ajournement persistant si collision : hypothèse de faible charge
  - émission si pas de collision avec écoute pour détecter une collision éventuelle en cours de transmission

- Gestion des collisions (BEB : Binary Exponential Backoff) :
  - utilisation du « slot-time » : temps maximum pour détecter une collision. Notion liée à la « longueur » du réseau et à la longueur minimale d'une trame (64 octets pour 802.3)
  - détection des collisions par comparaison entre le signal émis et le signal effectif sur le câble (mesure de puissance moyenne)
  - si collision, envoi d'un brouillage pour que toutes les stations prennent note de la collision (Jam)



Ethernet

- Trame IEEE 802.3
  - en half duplex : méthode d accès CSMA/CD
  - full duplex : pas de collision
  - taille des trames : 64 octets a 1500 octets
  - gap inter-trame : 96 bits
  - RTD est divise par 100 (par 10 pour Ethernet 100)
- Pour maintenir un domaine de collision de 200 m
  - la taille minimale de la trame doit être de 512 octets (au lieu de 64)
    - « extra carrier extension » si taille inférieure a 512 octets.
    - « packet bursting) : agrégation des petites trames.
- Mêmes adresses que IEEE 802.3
- IEEE 802.3z : standard gigabit en Juin 1998
- IEEE 802.3ab : Gigabit sur paires torsadées (1000 base T) sur une distance de 100m (4 paires à 125Mhz)

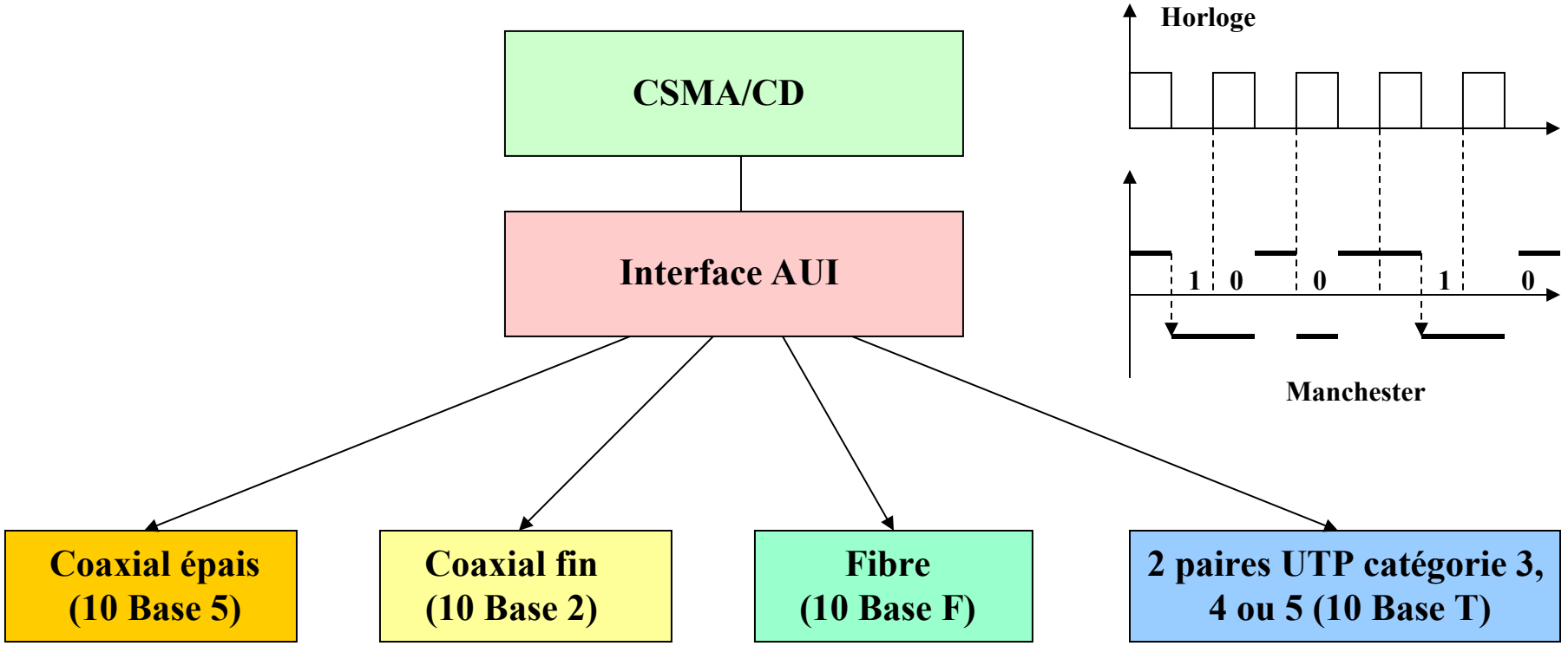


Ethernet : mise en œuvre

- Différents type de supports :
  - câble coaxial
  - paires torsadées : blindées ou non blindées
  - fibres optiques : mono-modes ou multi-modes

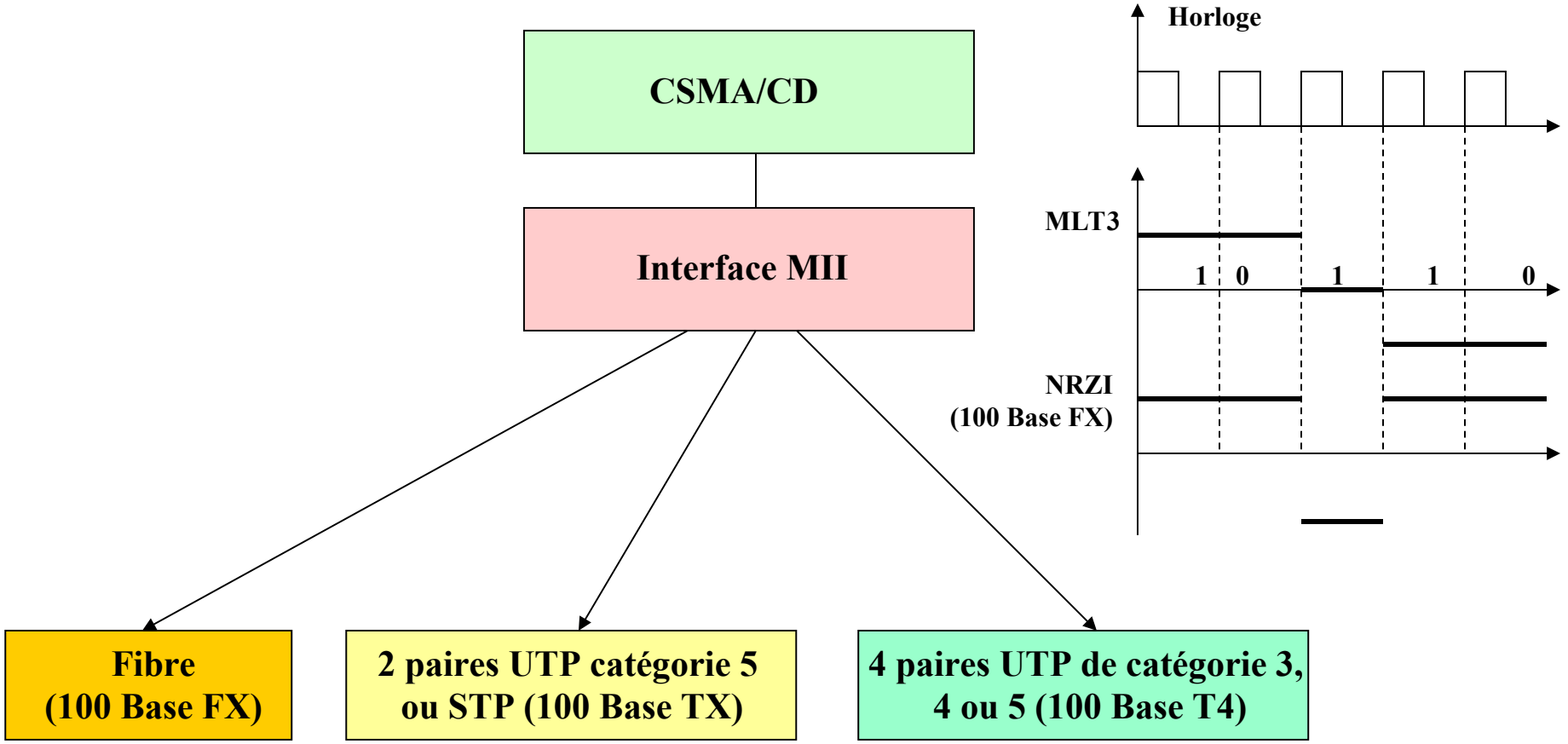
- Différents modes de transmission du signal :
  - Manchester : une transition du signal pour chaque bit transmis (Ethernet)
  - MLT3 : seuls les 1 font changer l'état du signal (100 Base TX et 100 base T4)
  - NRZI (Non Return to Zero Inverted) : 100 Base FX
  - PAM 5 : modulation d'amplitude à 5 niveaux (GigaEthernet)
  - etc.

Ethernet : 802.3



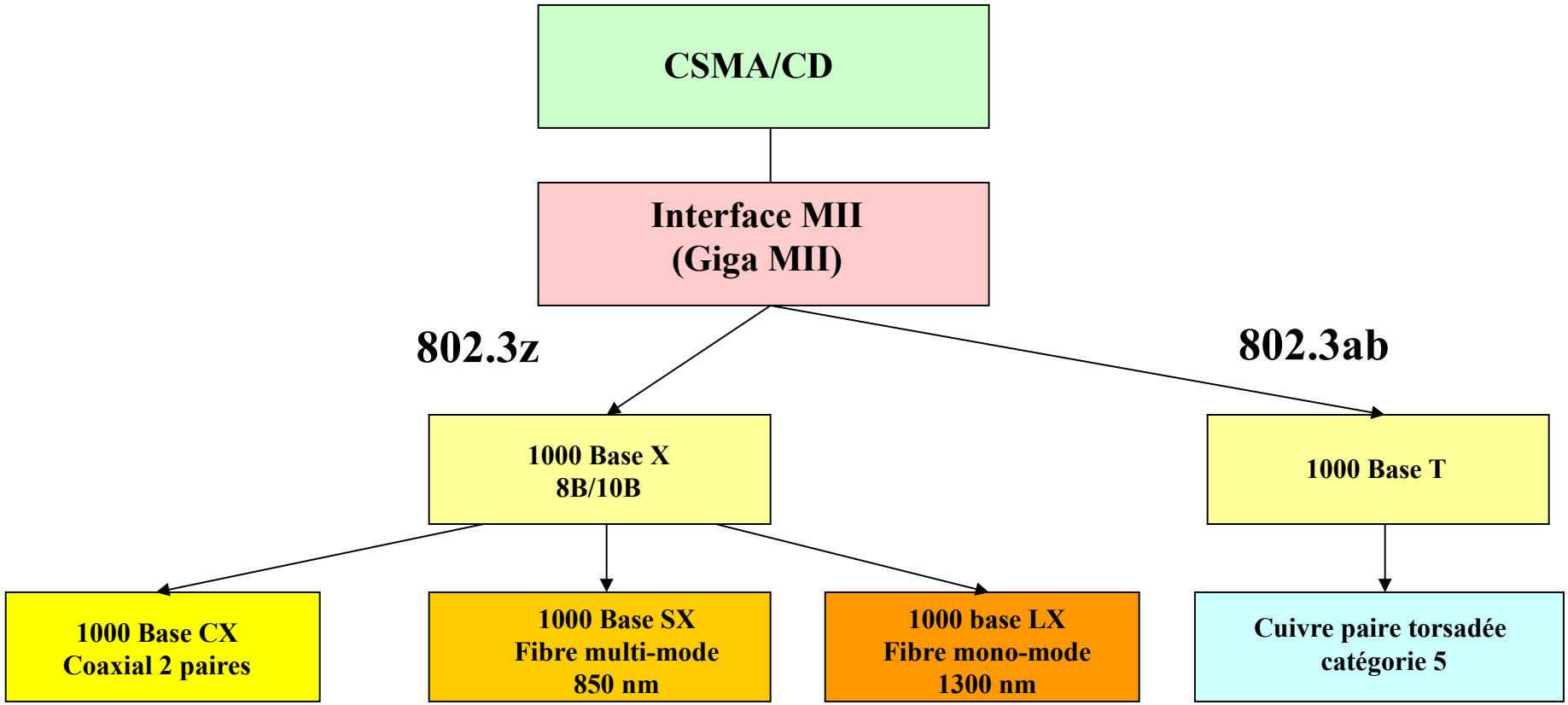
ST : 51,2 µs
L = 500 m (UTP Cat. 5) et 2000 m (fibre multi-mode)
v = 10 Mbit/s
AUI = Attachment Unit Interface

Fast Ethernet : 802.3u

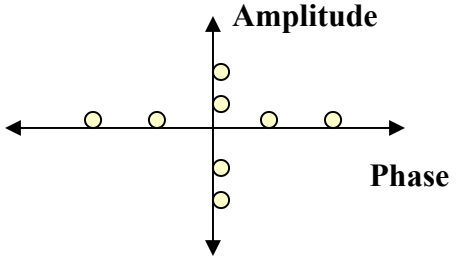


L = 100 m (UTP Cat. 5), 412m (100FX half-duplex) et 2000 m (100FX full-duplex)
v = 100 Mbit/s
MII = Media Independant Interface

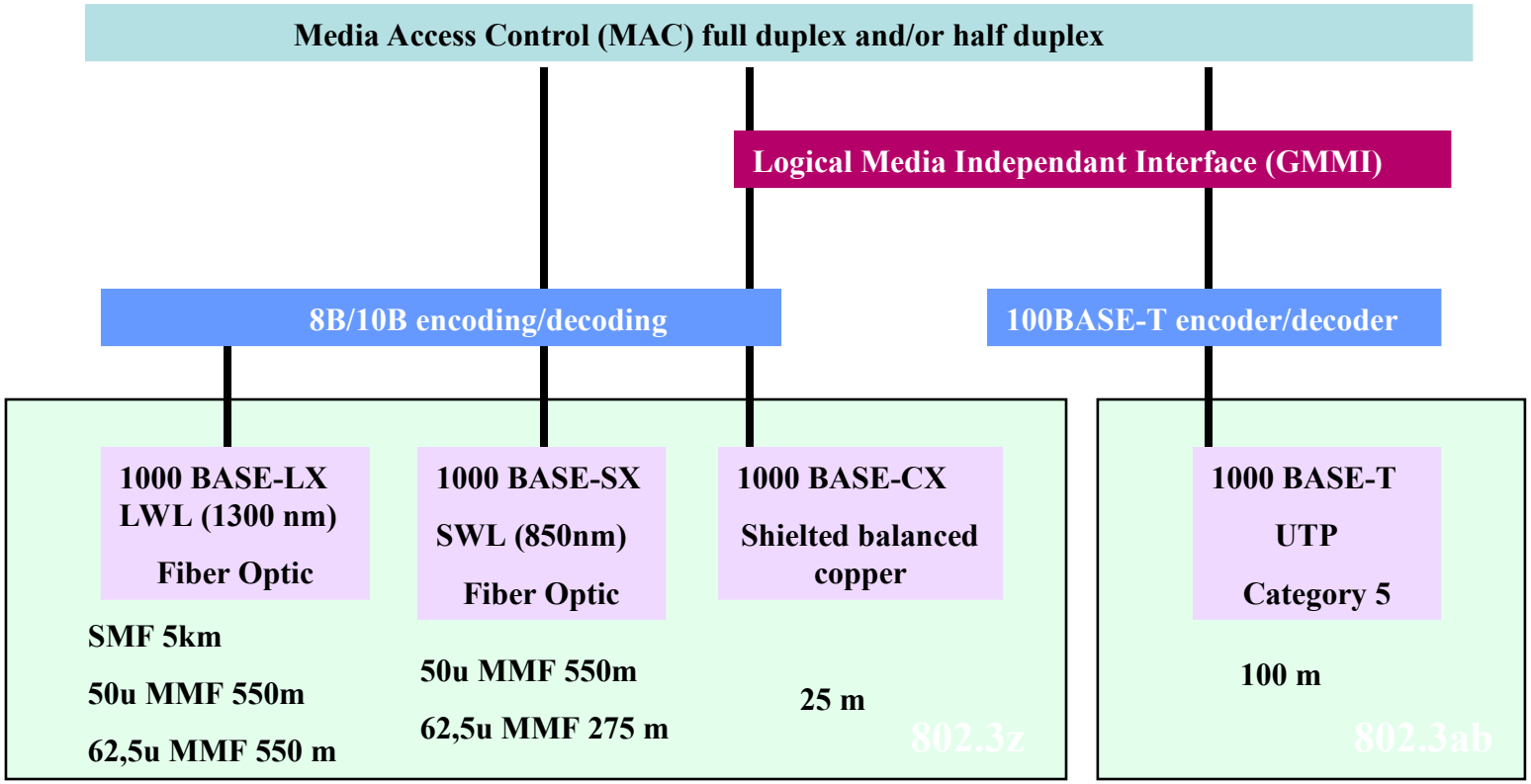
Gigabit Ethernet : 802.3z et 802.3ab



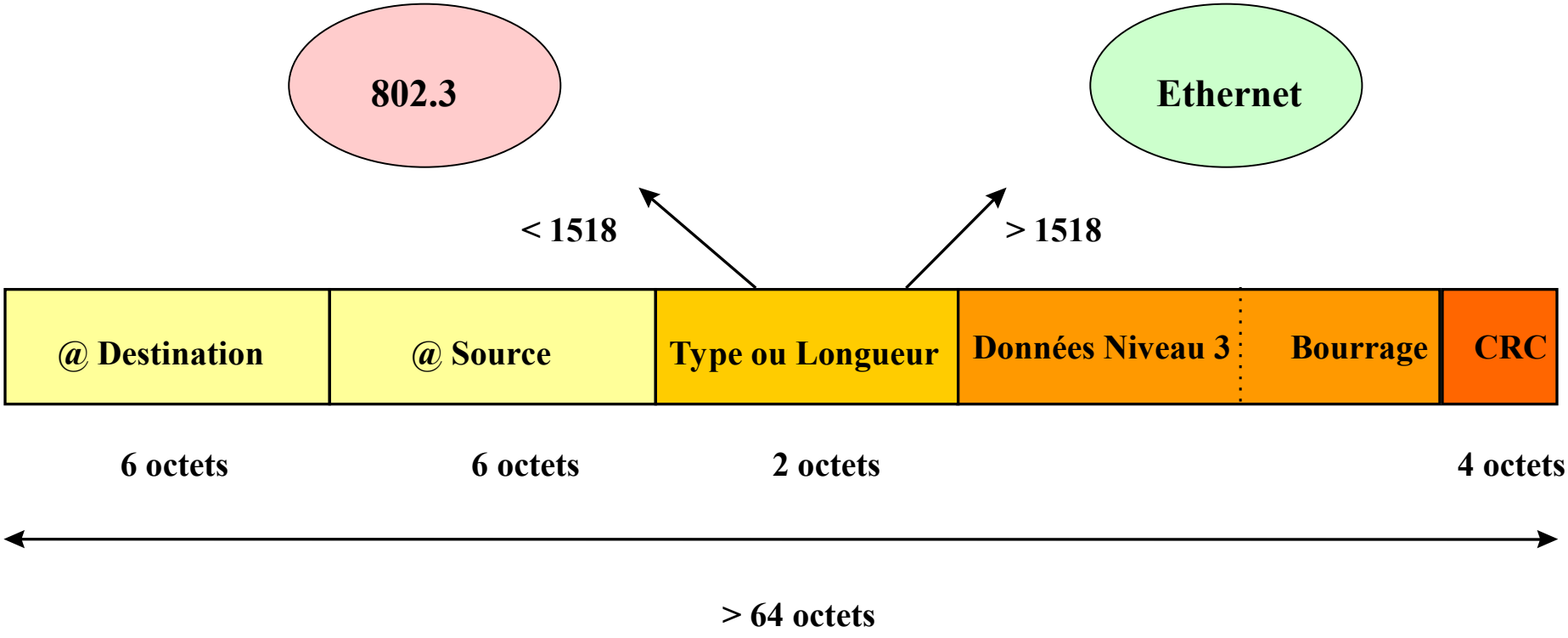
L = 25 m (coaxial) - 25 à 100 m (UTP Cat. 5) - 200 à 550 m (fibre multi-mode)
2000 à 5000 m (fibre mono-mode) voire 70 km avec des équipements spécifiques
v = 1000 Mbit/s
Modulation d'amplitude à quadrature de phase (MAQ): 250 Mbit/paire



Gigabit-Ethernet



Ethernet : Le niveau trame



Ethernet : l'adressage MAC

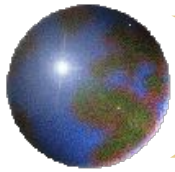
Format des adresses MAC (Medium Access control) :

- 3 octets sont attribués par l'IEEE aux différents vendeurs de matériels réseaux
- 3 octets correspondent au numéro de série dans la production du vendeur

L'adresse Mac est donc constituée de 6 octets soit 48 bits. Une liste d'attribution des 3 premiers octets par l'IEEE se trouve dans le RFC 1340

Exemple : O8:00:5A qui correspond à IBM
08:00:20 pour SUN
00:00:0C pour CISCO

Il existe aussi des adresses « non IEEE »



Ethernet : le niveau trame

Longueur des données limitée à 1518 octets pour Ethernet 10 Mbits/s (limite arbitraire fixée afin qu'une station ne monopolise pas le médium indéfiniment)

- toute valeur supérieure peut servir à référencer le code du protocole de niveau supérieur (niveau 3 du modèle OSI)

Bourrage parfois rendu nécessaire pour obtenir une trame d'une longueur minimale de 64 octets.

- Permet de conserver l'accès au médium pendant au moins un délai de 51,2 μs, délai qui correspond au délai de propagation maximal du signal sur une longueur de 2500 m à une vitesse d'environ 0,7xC selon les supports (en prenant en compte le délai aller-retour dans le cas de stations à chaque extrémité)

Le CRC (Cyclic Redondancy Check) est calculé sur 32 octets et permet de faire de la détection d'erreur (polynôme de degré 32)

Ethernet sans fil : 802.11

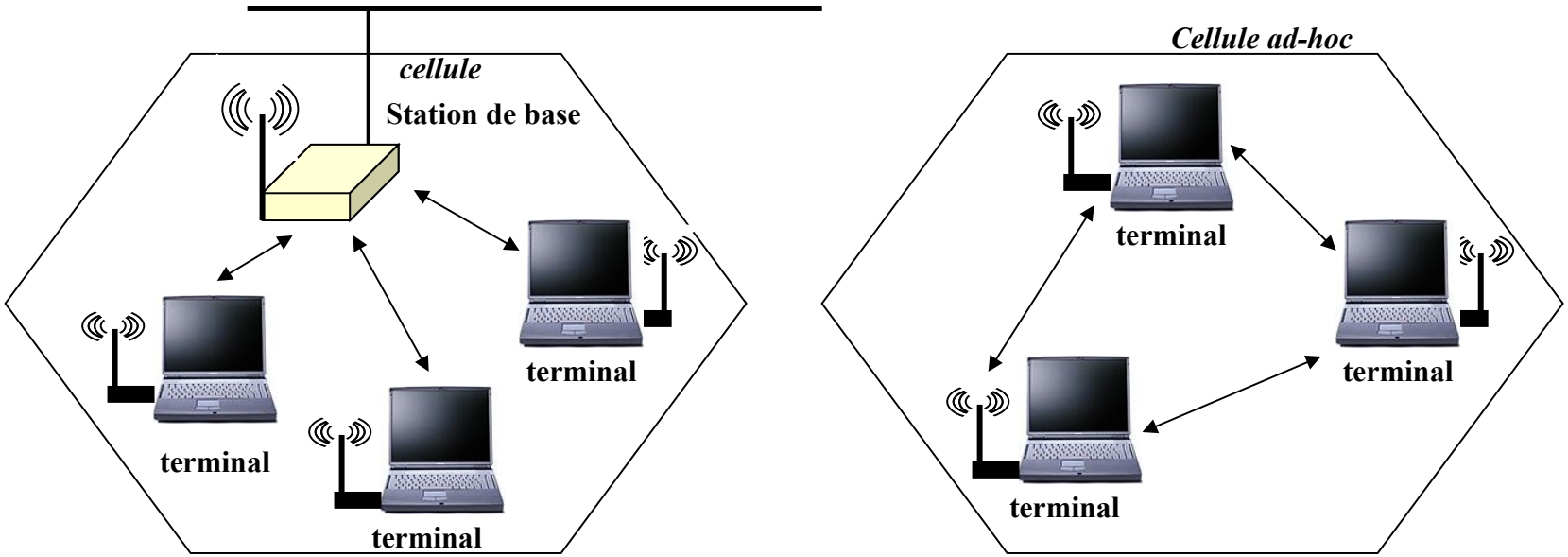
Groupe de travail WLAN (Wireless Local Area Network) créé en 1990

Publication de la norme IEEE 802.11 en 2001

Utilise une bande de fréquence hertzienne dans la gamme de 2,4 GHz

Architecture cellulaire - 1 cellule mesure quelques dizaines de mètres

2 modes de fonctionnement sont possibles : mode infrastructure et mode ad-hoc



Ethernet sans fil : implications

Méthode CSMA/CA - écoute du support partagé - évitement des collisions

Accès au support contrôlé par l'utilisation d'espace inter trame (IFS) et de trames de contrôle (ACK,RTS,CTS)

Un terminal « voit » toutes les trames de la cellule à laquelle il appartient.

Notion d'association entre la station de base et le terminal - la station de base agit comme un pont Ethernet.

La sécurité est « optionnelle » grâce à la possibilité de chiffrer les transmissions à l'aide d'une clé secrète partagée de 40 bits.

L'authentification des terminaux est basée sur la possession de cette clé et « éventuellement » sur une liste d'adresses MAC (ACL)

CONCLUSION :
la sécurité de l'ethernet sans fil varie de nulle à très faible...
pourtant on y échappera pas !!!

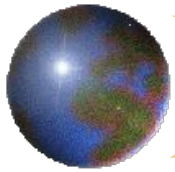
Ethernet sans fil : les technologies

BLUETOOTH : Permet la communication entre PC/portables, téléphones portables, assistants personnels etc… sur une distance de quelques centimètres à quelques mètres pour un débit max de 500 Kbit/s

802.11b : Cette norme permet un débit de 11 Mbits partagé, c'est ce qui est diffusé actuellement dans les matériels Lucent ou Airport

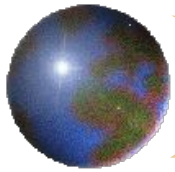
802.11a, HiperLAN : la famille des protocoles HiperLAN visent à accroître le débit et la couverture géographique des cellules.

HiperLAN 1-2	**accès : 200 m**	**débit : 23 Mbit/s**
HiperLAN 3	**accès : 5 km**	**débit : 20 Mbit/s (boucle local radio)**
HiperLAN 4	**accès : 200 m**	**débit : 155 Mbit/s (interconnexion ATM)**



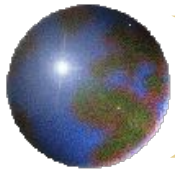
Protocoles de niveau 2 : retransmission trames

- l ’@ destination est confrontée à une table d ’adresses :
  - @ pas dans la table : le pont la rajoute (apprentissage) puis envoie la trame vers tous les autres ports (broadcast)
  - @ dans la table : le pont connaît où est le destinataire, il envoie donc la trame sur le port où est connecté le destinataire



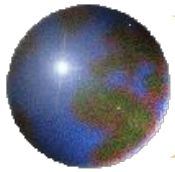
Protocoles de niveau 2 : auto-apprentissage

- L’adresse source de toutes les trames reçues est confrontée à la table :
  - à l ’ initialisation : table de retransmission vide, retransmission des trames jusqu’à ce que les tables soient suffisamment remplies
  - **pas dans la table** : elle est rajoutée dans la liste des destinations possibles pour ce port
  - **déjà dans la table :** mais si elle ne correspond pas à ce port, la table est mise à jour.



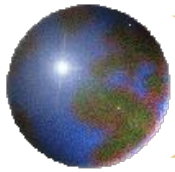
Protocoles de niveau 2 :spanning-tree (802.1)

- chemin entre 2 machines pas unique
- Possibilité bouclage -> diffusion ininterrompue ou une multiplication de chaque trame
- Conséquence : saturation inévitable du réseau
- Rend le chemin unique entre deux dispositifs et met en attente les ponts (ou les ports d'un pont) ayant créé un bouclage.



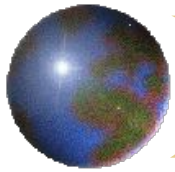
Protocoles de niveau 2 : PPP (1)

- Point to Point protocol (RFC xxxx)
- Utilise une variante de High Speed Data Link Control (HDLC)
- transport de données en full-duplex sur liaison analogique asynchrone (RTC) ou numérique synchrone (RNIS), sur supports cuivre, fibre optique, µ-ondes, satellites...
- Fournit un service de détection d ’erreur



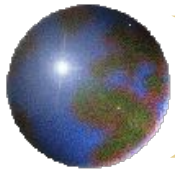
PPP : Encapsulation, LCP et NCP

- Encapsulation : protocoles Ethernet, Token Ring dans ISDN ou connexions modem
- NCP (*Network Control Protocol*) pour établir et configurer les différents protocoles réseaux
- LCP (*Link Control Protocol*) pour établir, configurer et tester la liaison entre les deux routeurs à travers une ligne série



PPP : LCP (Link Control Protocol)

- Établit et gère la connexion point à point
- Établissement connexion en 4 phases :
  - établissement du lien et négociation (obligatoire)
  - authentification (optionnelle)
  - détermination de la qualité du lien (optionnelle)
  - lien prêt (obligatoire)
- Permet aux routeurs de négocier des options pour les couches supérieures (ex : n° nœud, adresse réseau, taille max paquet, mot de passe…)

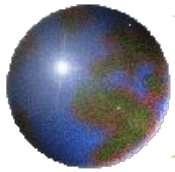


LCP : Caractéristiques

- négociation de la taille des trames (MRU : Maximum Receive Unit)
- vérification de la qualité du lien,
- il permet l'authentification (RFC 1334 : PAP et CHAP),
  - PAP : demande un couple login + mot de passe à la connexion
  - CHAP : login + mot de passe, utilisation d'un challenge et l'un secret partagé entre les 2 équipements pour chiffrer. Vérification périodique du challenge.
- il peut transporter plusieurs protocoles réseaux en même temps,
- il autorise la compression de données.

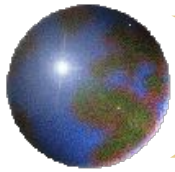
PPP : NCP (Network Control Protocol)

- fournit à chaque protocole réseau un moyen de négocier la prise en charge de chacun d'entre eux par PPP
- si les deux routeurs sont capables de supporter les protocoles IP, OSI, DECnet, IPX, Apple Talk et XNS, alors la phase de négociation de PPP déterminera les points communs entre les deux routeurs.



Caractéristiques logiques des routeurs

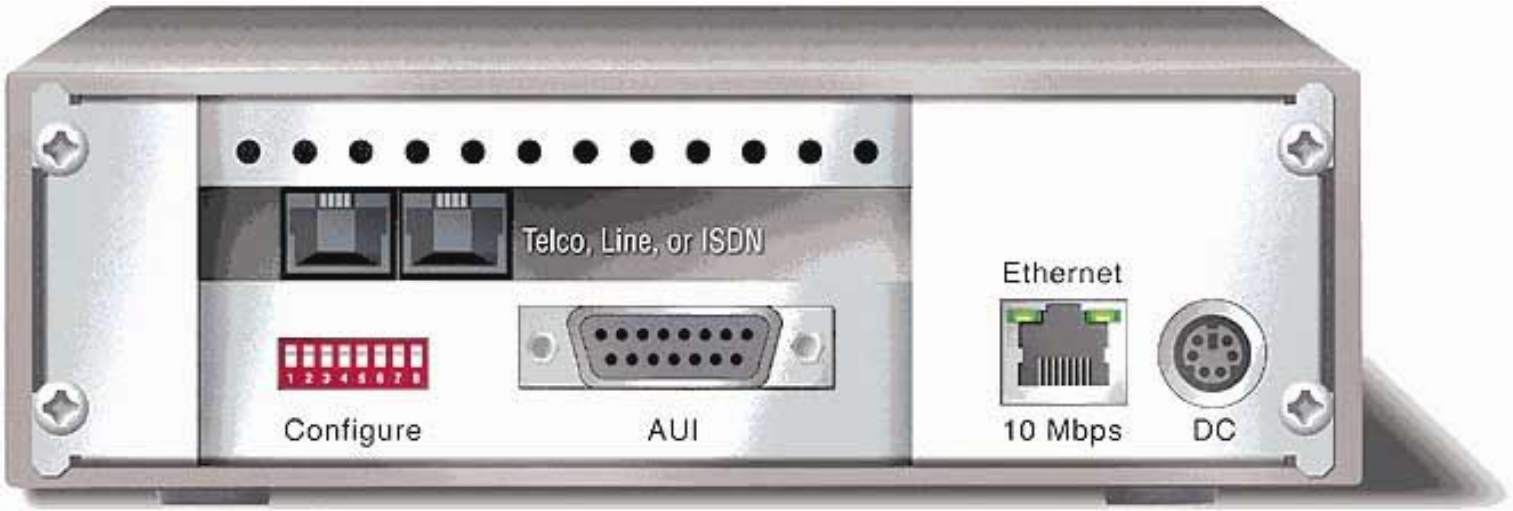
- aptitude à former des réseaux logiques
  - autonomie des sous-réseaux
  - partages des ressources sans restriction géographique liée aux protocoles
- recherche de la meilleure route
  - trouver le meilleur chemin
  - fonction de plusieurs critères (disponibilité, encombrement, priorités…)



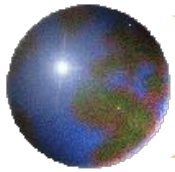
Caractéristiques physiques des routeurs

- Les caractéristiques des interfaces (Ethernet, ATM, RNIS…)
- les débits supportés (2MBits/s, 10MBits/s…)
- le(s) protocole(s) réseaux supportés
- le(s) algorithmes de routage supportés.
- Paramétrage
- Administration et surveillance

Routeurs : caractéristiques des interfaces



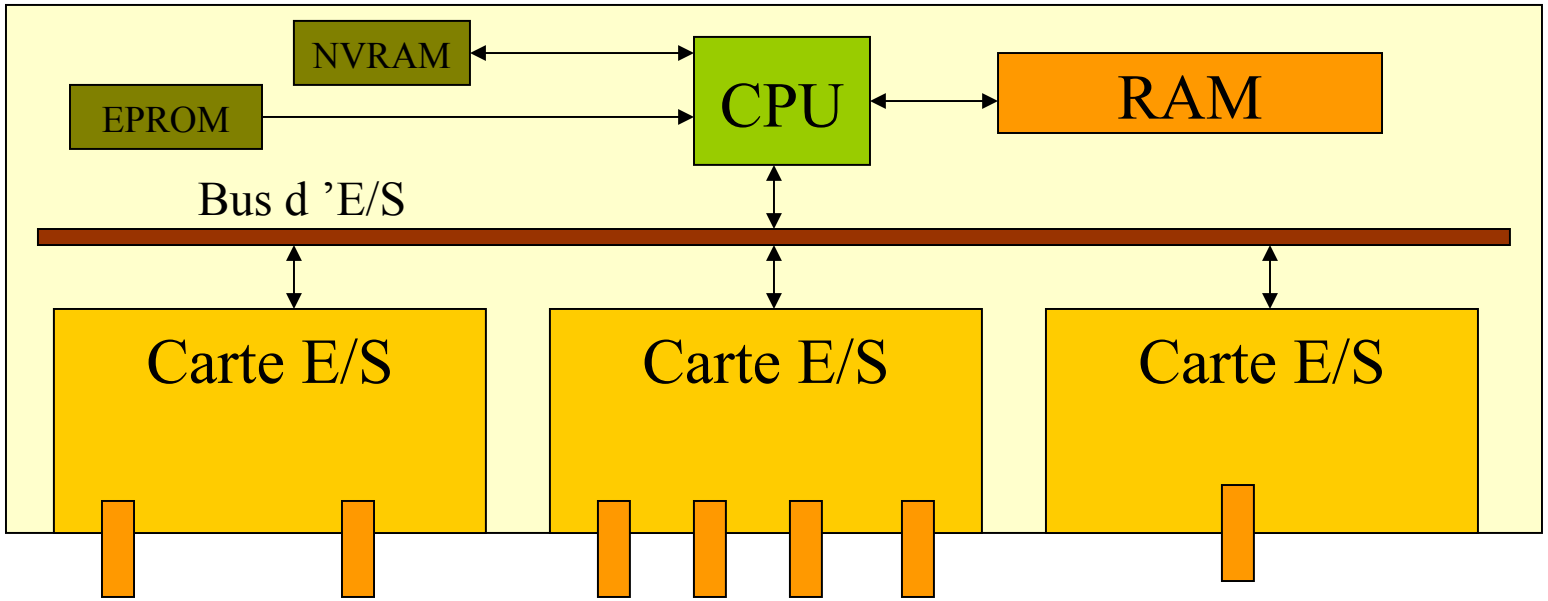
- Toujours au moins 2 interfaces…
- Routeurs orientés « WAN »
  - Interfaces de type V24, RNIS (T0 (2B+D), T2 (30B+D)), X21, xDSL, ATM
  - + Une ou plusieurs interfaces LAN (Ethernet, token ring, ATM)
- Routeurs orientés « LAN »
  - interfaces Ethernet ou ATM cuivre ou optique



Routeurs : débits supportés

- Pas nécessairement le débit théorique des interfaces
- En fonction :
  - De la puissance de routage (« Wire speed »...)
  - La configuration (nombre d ’interfaces...)
  - Des protocoles routés
  - Des éventuels filtrages
  - L ’architecture interne (crossbar, bus fond de panier...)
- Distinction : routeurs != commutateurs routeurs

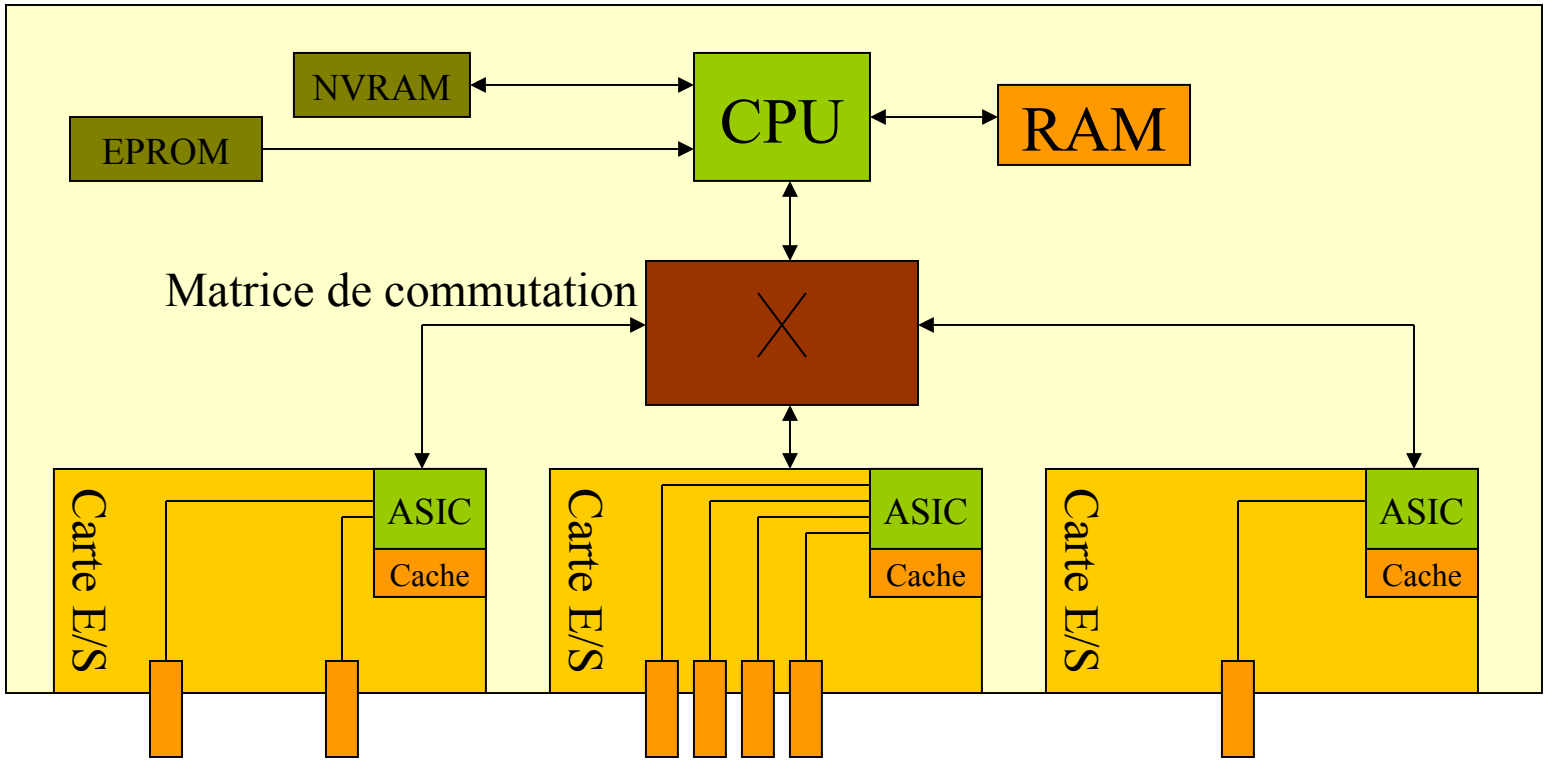
Architecture des routeurs



CPU de type RISC : ex MIPS 4400

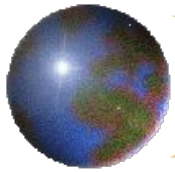
= Interfaces réseau

Architecture des commutateurs / routeurs (commutateurs niveau 3)



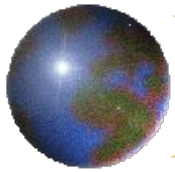
= Interfaces réseau

ASIC : Application Specific Integrated Circuit



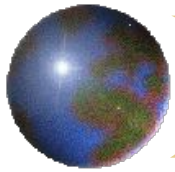
Routeurs : protocoles et algorithmes supportés

- Fonction de la politique commerciale du fabriquant.
- En général : prix licence proportionnelle au nombre de protocoles supportés
  - de base : TCP/IP : routes statiques, RIP
  - en supplément :
    - TCP/IP (OSPF, BGP, Multicast...)
    - IPX/SAP,
    - Appletalk



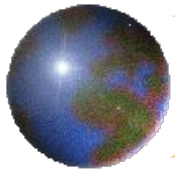
Routeurs : Paramétrage (1)

- Caractéristiques :
  - Paramétrage plus complexe (par rapport aux ponts)
  - Optimisation bande passante
  - réduction coûts (RNIS)
  - protection (sécurité)
- Local
  - interface RS232 : connecteur DB9, DB25 ou RJ45
  - câble croisé (null modem)
  - émulation terminal VT100 ou logiciel spécifique
  - CLI : Command Line Interface



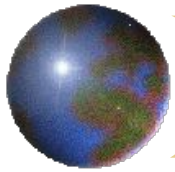
Routeurs : Paramétrage à distance (2)

- telnet : accès en ligne de commande (commandes IOS pour CISCO)
- logiciel spécifique (uniquement pour ce routeur)
- serveur http intégré (sous-ensemble de fonctions)
- TFTP ou FTP (téléchargement configuration)
- logiciel généraliste de gestion/supervision de réseaux
  - HP Open View…



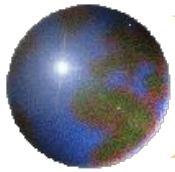
Le routage (sommaire)

- principes du routage
- routage statique
- routage dynamique
- routages interne et externe



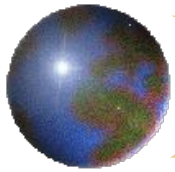
Protocoles et algorithmes de routage (sommaire)

- Introduction
- 2 buts principaux : détermination chemin et transport de paquets
- classes d ’algorithmes de routage



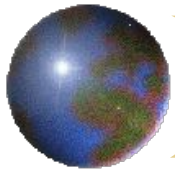
Routage : Introduction

- Routage : déplacer de l'information à travers une interconnexion de réseaux,
  - d 'une source à une destination.
  - au moins un nœud intermédiaire est traversé.
- activités principales :
  - la détermination du chemin optimal
  - transport de paquets à travers le réseau.
  - Maintient de sa table de routage



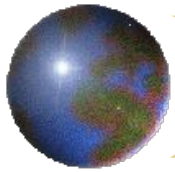
Routage : Détermination du chemin optimal

- unité utilisée (ex : longueur)
- création et maintient d ’une table de routage : informations sur les routes
  - disponibilité du lien
  - priorités
  - débit
  - temps de réponse
  - nombre de sauts



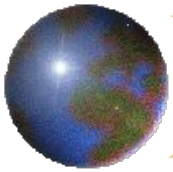
Routage : transport de paquets

- si destinataire pas sur une adresse locale -> envoi au prochain routeur en changeant l ’adresse physique de destination
- si le routeur ne sait pas où envoyer le paquet : destruction
- algorithmes décident de l ’action en fonction de plusieurs paramètres
  - état des liens, disponibilité, encombrement...
  - Priorités



Algorithmes de routage statiques et dynamiques

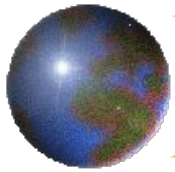
- statiques : pas un algorithme, mais seulement un enchaînement de règles de routage
- dynamiques : chemin est adapté en temps réel d’après les circonstances et les changements qui interviennent sur le réseau



Routage statique (1)

- Commande *route*
  - Permet d'indiquer une route :
    - vers un réseau (net) ou vers un équipement (host)
    - ou une route par défaut (default).
- Syntaxe (dépend des systèmes) :

  route add |delete [net|host] destination |default gateway metric



Problèmes du routage statique

Statique =

- mise à jour manuelle de tous les équipements du réseau
- Difficile de gérer la redondance de routes

(connaissance détaillée de la topologie)

=> vital en cas de rupture de lien

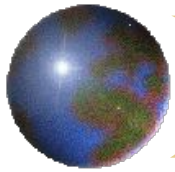
Boucle possible quand un lien est coupé

On recommande en général :

Stations	=>	Routage statique
Routeurs	=>	Routage dynamique

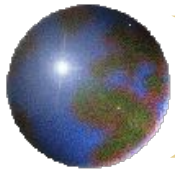
Avec un routage statique

Une station ne peut atteindre que les réseaux qu'on lui indique par les commandes route.



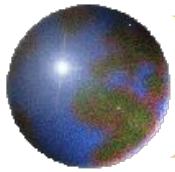
Caractéristiques des algorithmes de routage

- Résultat du processus pas immédiat : mécanisme d ’apprentissage des routes
- temps de convergence de l’algorithme : temps nécessaire à un algorithme pour aboutir à une description complète du réseau
  - avant convergence : risque de **bouclage** ou **perte** de paquets
- génèrent un trafic sur le réseau lié à l ’échanges des informations de routage (tables de routage)
- Chaque routeur procède de la même façon



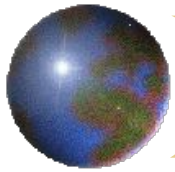
Algorithmes de routage : Distribués et centralisés

- Distribués : chemin est élaboré par chaque routeur (plus courant)
- Centralisés : chemin est calculé par un routeur central
  - simple,
  - énorme inconvénient : peu fiable et vulnérable



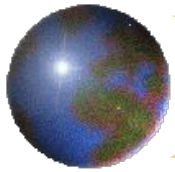
Algorithmes de routage : multi-chemins

- *multi-path* : opposition aux algorithmes *single-path*,
- supportent plusieurs chemins dans la même direction
  - multiplexage du trafic sur plusieurs liaisons
  - sécurité accrue dans le cas d’une liaison défaillante



Algorithmes de routage : source routing

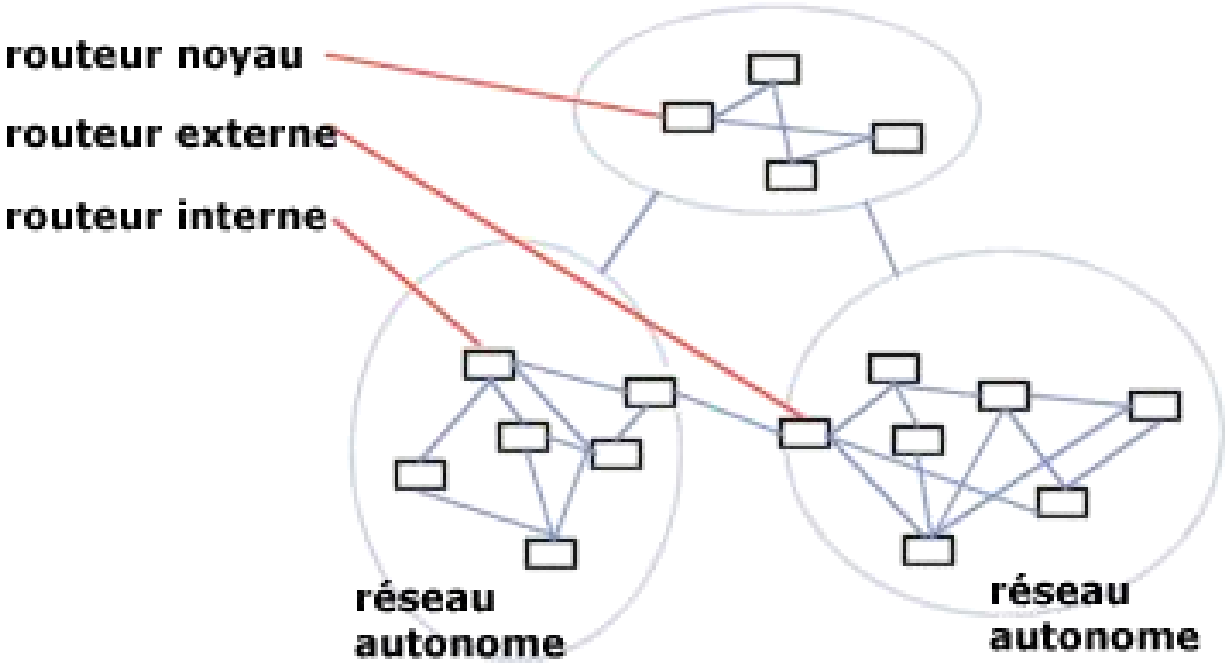
- nœud terminal: détient l'intelligence et détermine le chemin
  - meilleurs résultats, mais temps de convergence important
- reverse : l'intelligence est confiée aux routeurs intermédiaires

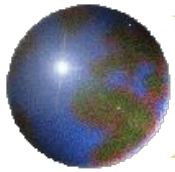


Algorithmes de routage : intra et inter domaines

- domaine : zone sous l ’autorité d ’une même administration (aussi appelé « Autonomous system »)
- algorithmes intra-domaine : contrôlent des routeurs situés dans le même domaine : IGP (*Interior Gateway Protocol)*
- algorithmes inter-domaines : permettant la communication entre routeurs n’ayant aucune appartenance à un même domaine : EGP (*Exterior Gateway Protocol)*

Domaines et routeurs

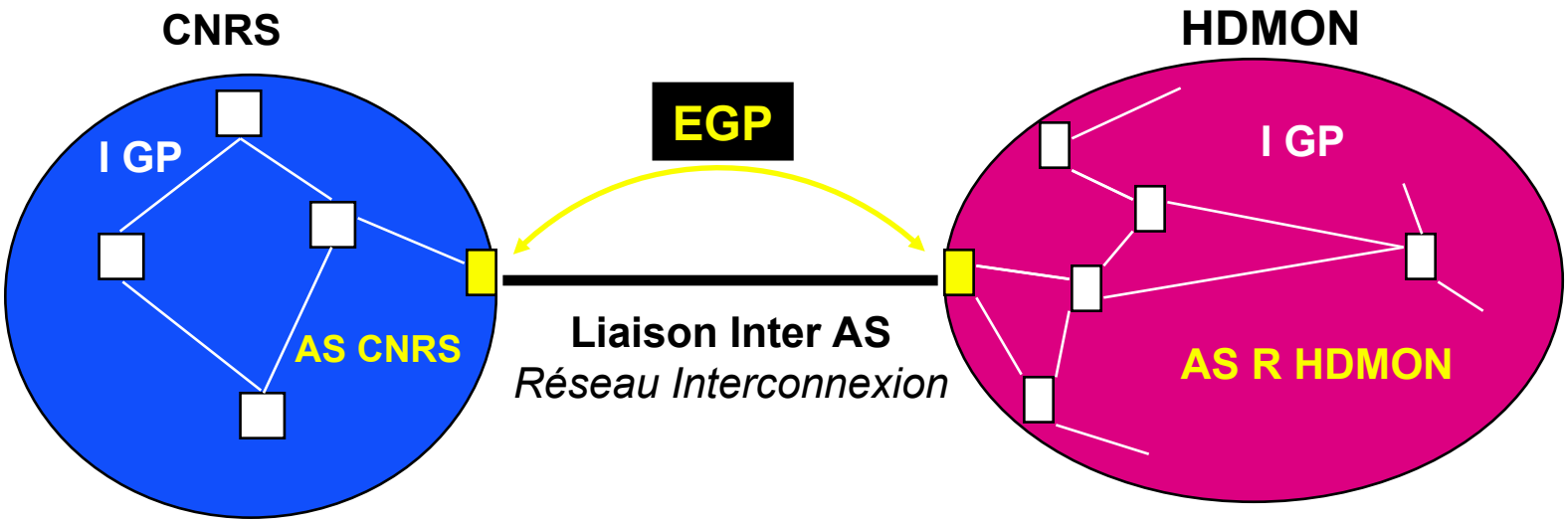




Algorithme de routage : inter domaines

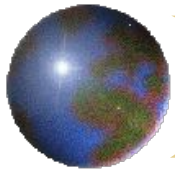
- Autonomous System (AS)
  - "Ensemble de réseaux et de routeurs sous une administration unique"
    - ex.: entreprise, campus, réseau régional, cœur d'un réseau national
    - toutes les parties d'un AS doivent être connexes
  - Les numéros d'AS sont délivrés par le NIC-France (hostmaster@nic.fr)
    - Un numéro = 16 bits (ex:Renater a pour numéro d'AS 1717)
  - Utilisé par différents protocoles de routage pour l'échange d'informations : EGP, BGP...

Autonomous system : exemple



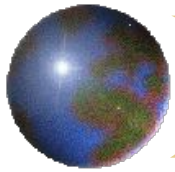
Routeur Frontière (gateway)

Routeur interne



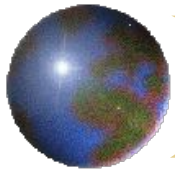
Algorithmes de routage : Distance-Vector

- Algorithme de Bellman-Ford
- vecteur de distance : exprime la distance en nombre de sauts (1 saut=1 routeur) :
  - construit un embryon de la table de routage
  - diffuse intégralement sa table à travers le réseau.
  - plusieurs routes : choix de la route qui a la meilleure métrique
  - convergence = plusieurs minutes
  - charge réseau : assez importante



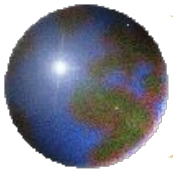
Algorithmes de routage : Link-State

- Algorithme de Dijkstra
- État de liaison : le routeur envoie au(x) voisin(s) que informations sur les liens actifs qu ’il a avec eux
- Propagation d’ informations complémentaires :
  - débit
  - coût télécom associé
  - temps de transit (ex : liaison satellite ou cryptée)
  - pondération (poids associé à cette route) : informations prioritaires ou sensibles
- temps de convergence court, charge réseau faible (seules les modifications sont transmises)



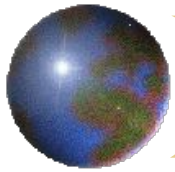
Protocoles de routage

- = implémentation des algorithmes de routage, typiquement réalisés par des logiciels calculant le(s) chemin(s) de routage à travers des réseaux interconnectés
- protocoles de routage != protocoles supportés par les réseaux locaux avec lesquels les routeurs opèrent



Protocole de routage : RIPv1

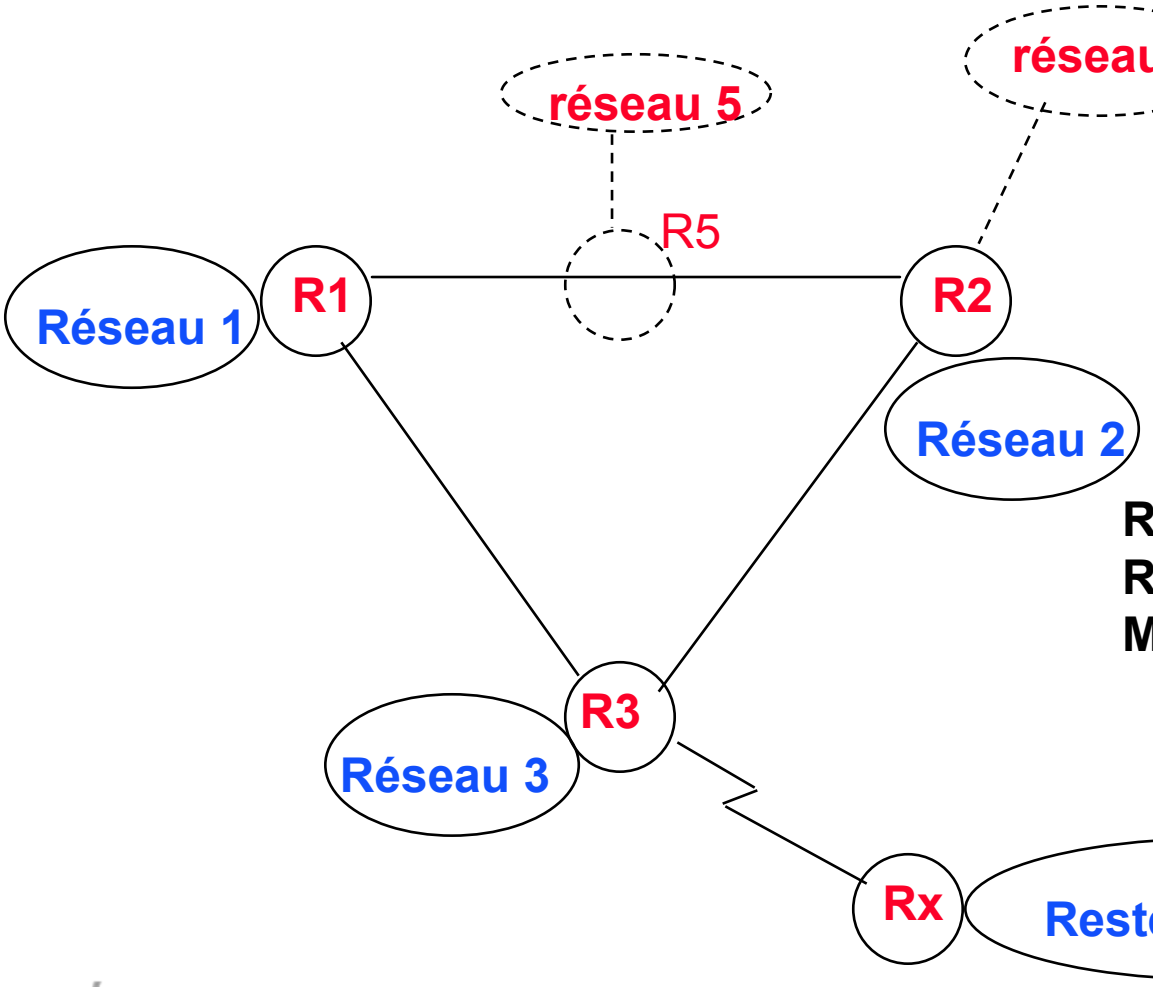
- RFC 1058) RIP (*Routing Information Protocol*) pour IP, RIP-IPX, RIP-XNS
- **intra-domaine** de type **vecteur de distance**,
- comptabilise nombre de sauts, en déduire le chemin le plus court.
- premier protocole à avoir été implémenté sur TCP/IP.
- Plus simple que OSPF,
- 15 sauts maximum entre les réseaux source et destination
- temps convergence = quelques minutes
- génère un trafic important sur le réseau
- diffuse l'intégralité de sa table de routage
- ne supporte pas les sous-classes IP



Protocole de routage : RIPv1 (2)

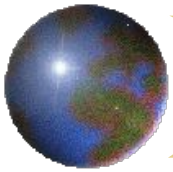
- L' équipement de routage diffuse (broadcast)
  - toutes les 30 secondes
  - la liste des réseaux qu'il peut atteindre
  - avec leur distance (nombre de sauts)
- Un message RIP est contenu dans un datagramme UDP
  - N° de port = 520
- Daemon routed ou gated sous Unix
- router rip (Cisco)
- RIPv2 (RFC 1387, RFC 1388) supprime certains de ces inconvénients
  - gestion sous-réseaux
  - diffusion multicast
  - authentification

RIP : Exemple



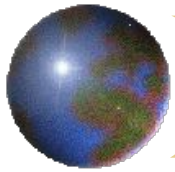
Tables de routage

R1	R2	R3
Res2, R2	Res1, R1	Res1, R1
Res3, R3	Res3, R3	Res2, R2
Monde,R3	Monde R3	Monde Rx



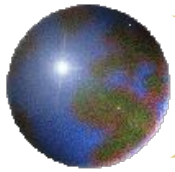
RIP : Exemple (1)

- Après le boot :
  - la table de routage de R1 :
    - 130.190.4.0 d = 0
    - 130.190.5.0 d = 0
  - la table de routage de R2 :
    - 130.190.4.0 d = 0
    - 130.190.6.0 d = 0
  - la table de routage de R3 :
    - 194.57.137.0 d = 0
    - 193.64.203.0 d = 0



RIP : Exemple (2)

- R2 envoie un message broadcast sur les réseaux
  193.64.203.0 et 204.27.1.0 contenant la table de routage :
  - 193.64.203.0 d = 0
  - 204.27.1.0 d = 0
- R1 envoie un message de broadacst sur les réseaux
  195.132.92.0 et 204.27.1.0
- R3 envoie un message de broadcast sur les réseaux
  194.57.137.0 et 193.64.203.0



RIP : Exemple (3)

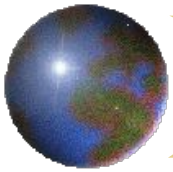
- Après 30 secondes (ou par mise à jour déclenchée) :
  - R1 envoie un message broadcast
  - R2 met sa table de routage à jour :
    - 130.190.4.0 d = 0 (R1, d = 0 + 1) local
    - 130.190.6.0 d = 0 local
    - 130.190.5.0 d = 1 (R1, d = 0 + 1) R1
  - R2 envoie un message broadcast
  - R1 met sa table de routage à jour :
    - 130.190.4.0 d = 0 (R2, d = 0 + 1) local
    - 130.190.5.0 d = 0 (R2, d = 1 + 1) local
    - 130.190.6.0 d = 1 (R2, d = 0 + 1) R2

RIP : Exemple (4)

- A la réception des messages de broadcast chaque routeur met sa table de routage à jour :
- R2:

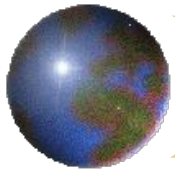
193.64.203.0	d=0	
204.27.1.0	d=0	
195.132.92.0	(d=0 +1)	d=1
204.27.1.0	(d=0 +1)	d=1
194.57.137.0	(d=0 +1)	d=1
193.64.203.0	(d=0 +1)	d=1

- de même pour les routeurs R1 et R3



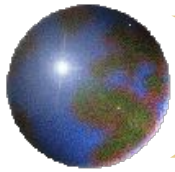
Protocole de routage : OSPF (1)

- RFC 1247 : OSPF (*Open Shortest Path First*) pour IP
- protocole intra-domaine du type Link-State
- technique : trouver le chemin optimal
  - bande passante (la plus élevée possible),
  - débit potentiel disponible,
  - de coût (le moins élevé possible)
  - Règle : débit élevé = coût faible
- temps de convergence plus long que RIP, trafic généré moins important
- définit des zones contiguës (areas) : communication d ’informations de routage inter-zone par un EGP



Protocole de routage : OSPF (2)

- chaque AREA se comporte comme un réseau indépendant
- elle ne connaît que l'état des liaisons internes à l'AREA
- Deux niveaux de routage :
  - intra Area
  - inter Area
- On distingue 3 classes d'Aires :
  - l'Aire backbone (Area 0)
    - chemin obligatoire pour passer d'une aire à l'autre
  - les Aires secondaires
    - Tous les nœuds de routage ont une vue complète de la carte du réseau
    - ils calculent localement la meilleure route entre une source et une destination.
  - les Aires terminales (stub area)
    - même comportement que les aires secondaires
    - sauf : ne mémorisent pas les informations sur les routes externes
    - Toutes les routes externes sont récapitulées dans une route par défaut



Protocole de routage : OSPF (4)

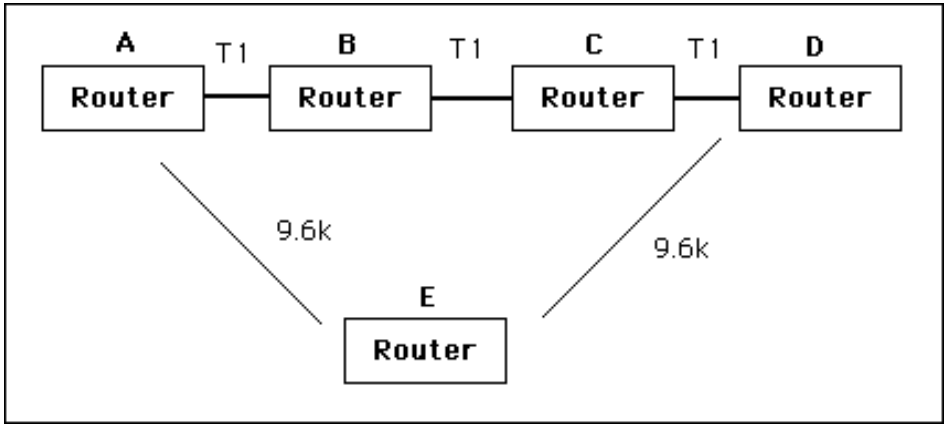
- OSPF
  - Calcule des coûts en guise de métriques
  - Sait router les sous-réseaux,
  - par types de service,
  - permet le load balancing

  - Inclut un système d'authentification des messages échangés

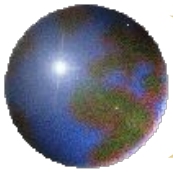
  - Envoie un LSA (Link state Announcement)
    - quand l'état d'une ligne change
    - ou toutes les 30 minutes.

Comparaison RIP / OSPF

Quel chemin pour aller de A à D ?

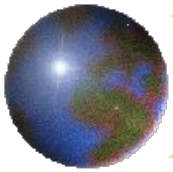


- RIP : A->E->D (à 9,6Kb/s max)
  - car nb sauts (A->E->D) < à (A->B->C->D)
- OSPF : A->B->C->D
  - car débit liaison (A-B) > débit liaison (A-E)



Protocole de routage : EGP

- EGP (*Exterior Gateway Protocol*) pour IP
- dynamique inter-domaines implémenté dans les backbones d'interconnexion des réseaux TCP/IP universitaires
- peu à peu remplacé par BGP (plus performant)



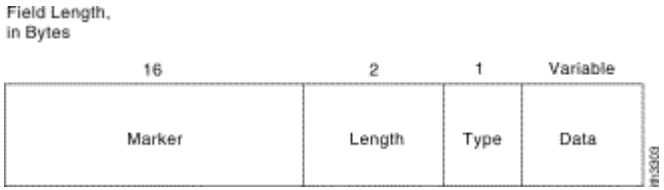
Protocole de routage : BGP (1)

- RFC 1105 : BGP-1,RFC 1163 : BGP-2, RFC 1267 : BGP-3
- RFC 1771 (A *Border Gateway Protocol 4)*, successeur d ’EGP

  Ni vraiment distance vector ni vraiment link state
  - transmet le "chemin d'AS" entre la source et la destination
    - détection simple et efficace des boucles
- Adapté à des topologie complexes (maillées)
  - Échange des informations de routage par une connexion TCP (port 179)
  - inclut un système d'authentification des messages échangés
  - Peut aussi être utilisé comme protocole de routage interne (I-BGP)

Protocole de routage : BGP (2)

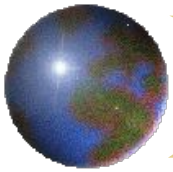
- algorithme *policy-based* (prise en compte dans le processus de routage, de l'importance du trafic sur le réseau)
- 1ere update : totalité de la table de routage, puis mises à jour incrémentale
- toutes 30 secondes : messages de test de liaison « keep alive »
- Support des CIDR
- paquet BGP :



- *marker* : identifiant d 'authentification
- *type* : 4 types de message : open, update, notification, keep-alive
- *data* : données pour les couches supérieures

Principaux protocoles de niveau 3

	MODÈLE OSI	TCP / IP	IPX (NOVELL)	APPLE TALK (APPLE)		
LOGICIEL	APPLICATION	FTP, XWindow,NFS, Telnet	NETWARE FILE SHARING PROTOCOL (NFSP)	APPLE TALK FILING PROTOCOL (AFP)		
	PRÉSENTATION					
	SESSION	REMOTE PROCEDURE CALL (RPC)		APPLE TALK SESSION PROTOCOL (ASP)		
	TRANSPORT	UDP / (TCP) TRANSMISSION CONTROL PROTOCOL	SEQUENCED PACKET EXCHANGE (SPX)	APPLE TALK TRANSACTION PROTOCOL (ATP)		
MATERIEL	RÉSEAU	INTERNET PROTOCOL (IP)	INTERNET PACKET (IPX)	DATAGRAM DELIVERY PROTOCOL (DDP)		
	LIAISON	DEVICE DRIVER	DEVICE DRIVER	DEVICE DRIVER		
	PHYSIQUE	TYPES DE MATÉRIELS VARIES	TYPES DE MATÉRIELS VARIES	LOCAL TALK	ETHER TALK	TOKEN TALK

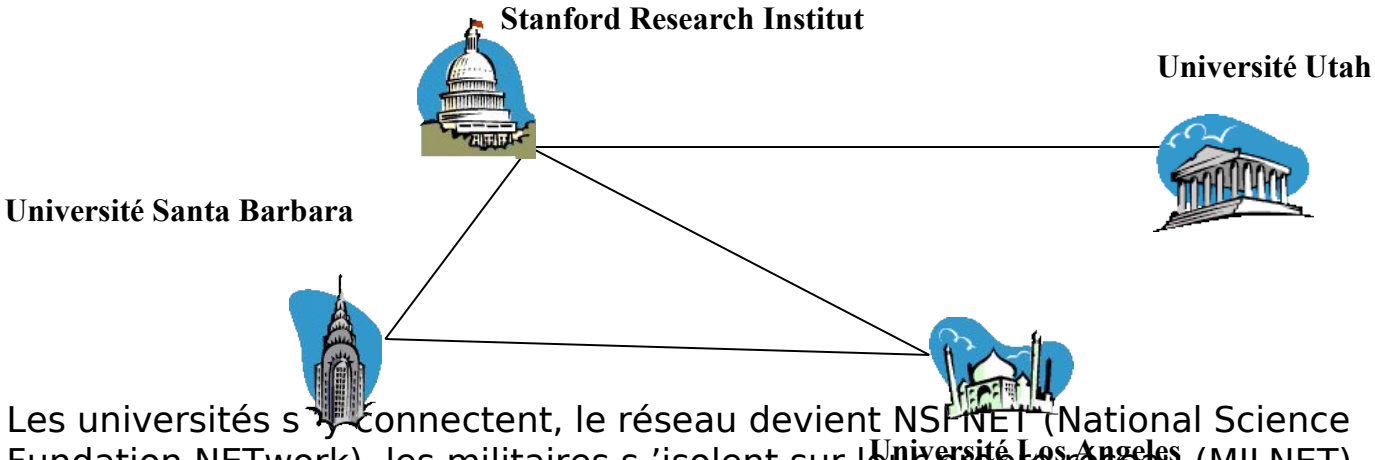


Internet Protocol (sommaire)

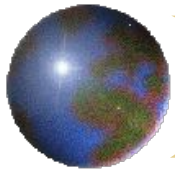
- Historique
- adressage
- fonctions et contenu des paquets
- couche transport (TCP et UDP)

IP : Historique

- Issu de travaux de recherche par (et pour) les militaires
- but : garder le réseau opérationnel même en cas de destruction d ’une partie.
- 1969 : l ’agence américaine D.A.R.P.A (*Défense Advanced Research Projects Ajency*) lance un projet de réseau à commutation de paquets : ARPANET.



- Les universités s ’y connectent, le réseau devient NSFNET (National Science Fundation NETwork), les militaires s ’isolent sur leur propre réseau (MILNET)
- en 1990, NSFNET devient Internet



Pourquoi IP a-t-il supplanté la « concurrence » ?

Le modèle OSI avait de nombreuses faiblesses :

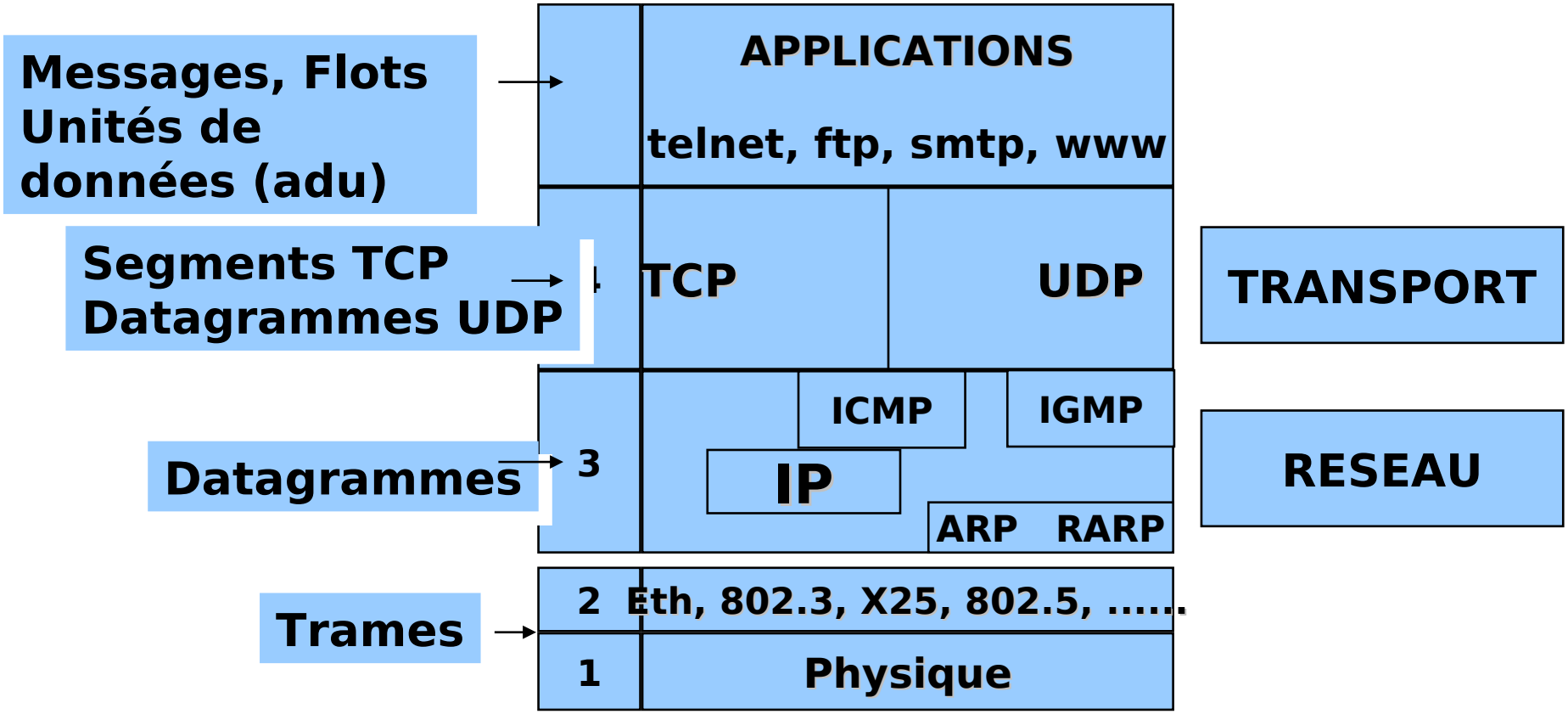
- Trop de couches dans le modèle … avec une forte influence de SNA d'IBM
- Complexité dans la description
- Difficile à implémenter et peu efficace
- Redondance des mécanismes : contrôle d'erreur, contrôle de flux, etc. à chaque couche ou presque
- Mauvaise prise en compte des modèles sans connexion et des réseaux locaux (apparition d'Ethernet)

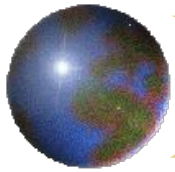
Le modèle TCP/IP s'est imposé au détriment du modèle OSI

Le modèle orienté « sans connexion» (datagramme) s'est imposé face au modèle orienté connexion

La notion de couches ou niveaux est restée comme vision structurante des fonctions réseaux

IP : modèle en couches



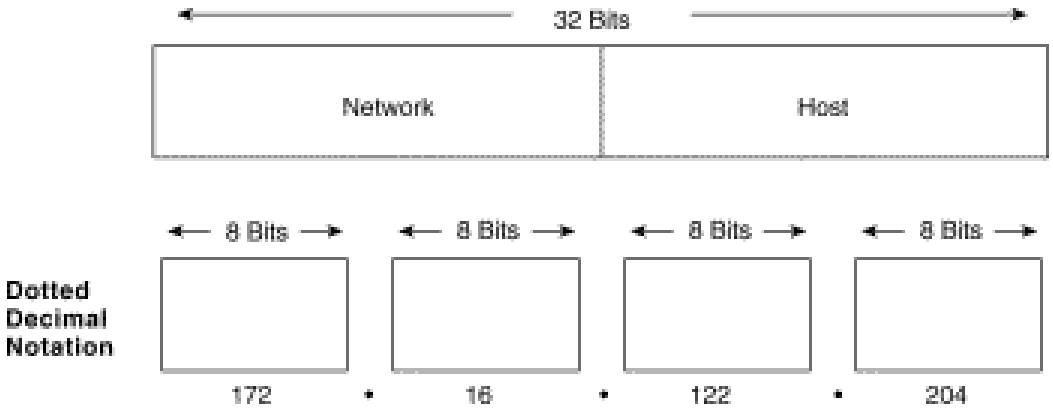


IP

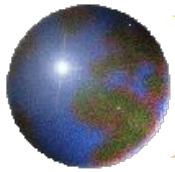
- 3 couches principales :
  - Réseau : IP :
    - ICMP : Internet Control Message Protocol
    - Mapping adresses IP-adresses physiques : ARP, RARP (Address Resolution Protocol, Reverse Address Resolution Protocol)
  - Transport : TCP et UDP :
    - TCP : *Transmission Control Protocol* : Service de transport fiable
      - reprise automatique en cas d ’erreur de transmission, de perte de paquets ou de panne de lisison entre emmeteur et recepteur
      - Mode connecté
    - UDP : transport non fiable
  - Application : ...

IP : adressage (1)

- IPv4 : 32 bits (4 octets), 4 chiffres de 0 à 255 séparés par un point
- Unique au monde

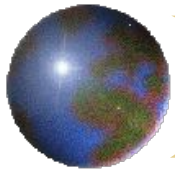


http://www.iana.org/assignments/ipv4-address-space (allocation @ip)



IP : adressage (2)

- Découpée en deux :
  - adresse de réseau, ou *network id*
    assigné par une autorité, identifie le réseau
  - identificateur local de machine, ou *host id*
    assigné par l ’administrateur du réseau, identifie la machine sur le réseau
  - le découpage précis dépend de la classe d ’adresses...

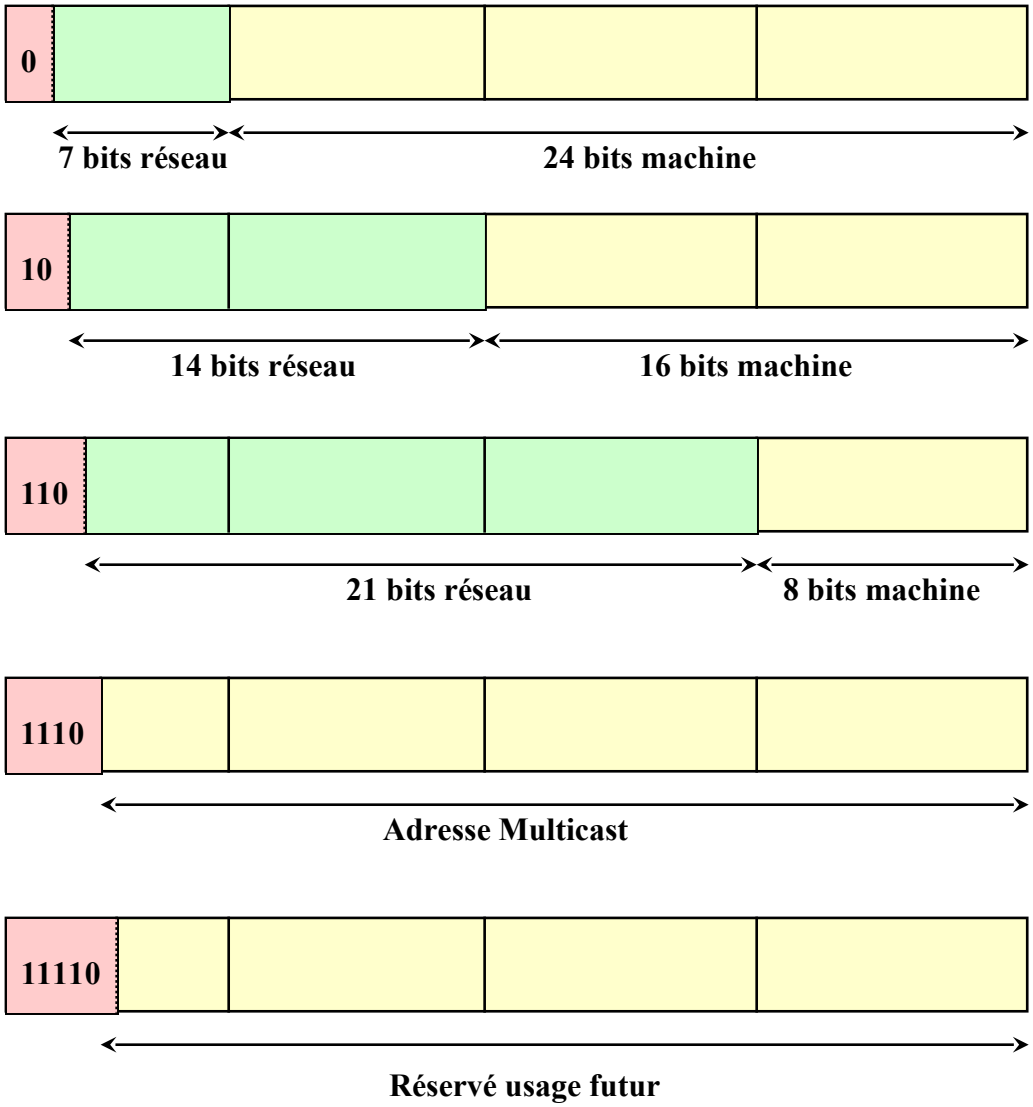


IP : adressage (3)

- Classification :
  - classe A : N.H.H.H
  - classe B : N.N.H.H
  - classe C : N.N.N.H
  - N = adresse réseau
    H = adresse locale
  - classe D : cas particulier, pas de distinction network/host
- L'espace d'adressage n'est pas hiérarchisé ou arborescent
  - à la différence du téléphone, de Transpac, d'ATM, d'IPv6

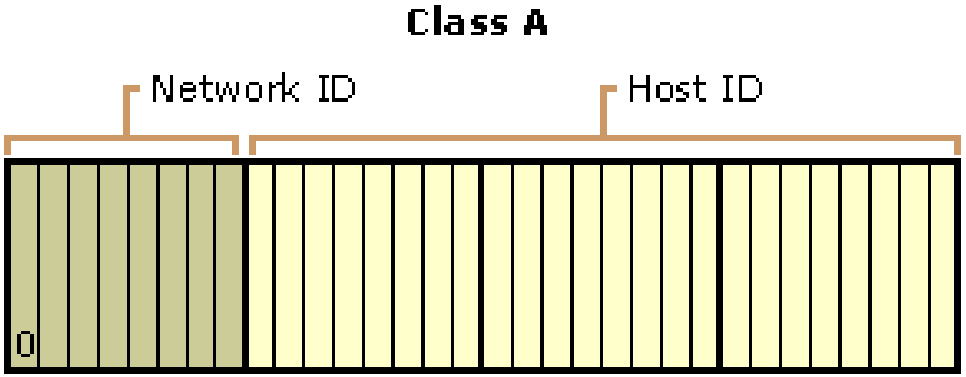
Schéma...

IP : les classes



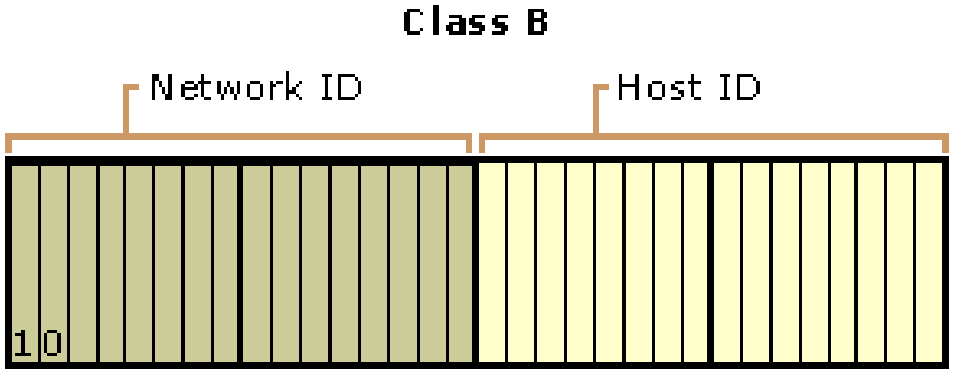
IP : adressage : classe A

- 7 bits pour le numéro de réseau
  - 1.0.0.0 à 126.0.0.0
- 24 bits pour l'adressage local
  - 2^{24}-2 @ locales possibles (16,277,214)
- En France, pas de réseau de classe A
  - Ex : 16.0.0.0 (DEC) 18.0.0.0 (MIT)



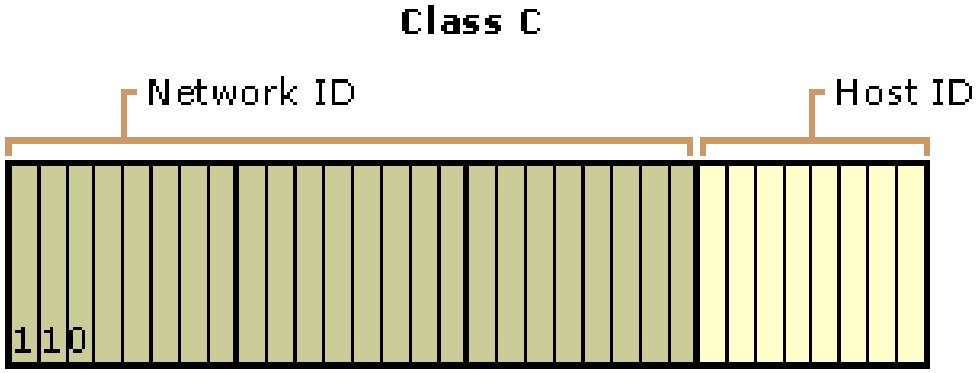
IP : adressage : classe B

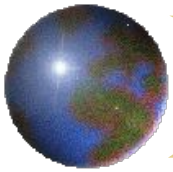
- 16 bits pour le numéro de réseau
  - 128.1.0.0 à 191.254.0.0
- 16 bits pour l'adressage local
  - 2^{16}-2 @ locales possibles (65,534)



IP : adressage : classe C

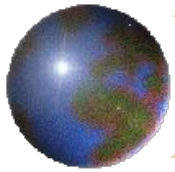
- 24 bits pour le numéro de réseau
  - 192.0.1.0 à 223.255.255.0
- 8 bits pour l'adressage local
  - 2^8-2 @ locales possibles





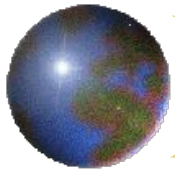
IP : adressage : classe D

- Adresses multicast (RFC 1700)
  - Transmissions point à multipoint; Exemple vidéo-conférence
- Réseaux 224 à 231
- Pas de structuration : utilisée de façon très spéciale, ponctuelle, sans contrainte d ’unicité, sans organisation gérant leur attribution
- Adresse : Pré allocation - utilisation - libération



IP : adresses particulières (1)

- Classe E :239 à 254 Réservés pour une utilisation future
- Adresses particulières
  - soi-même : 127.0.0.1 (loopback ou localhost)
    - test logiciels, communication inter-processus sur la station
  - tous les bits de la partie machine à 0 => le réseau
    - 130.190.0.0 désigne le réseau de classe B : 130.190
  - tous les bits de la partie machine à 1 => tous les hosts du réseau
    - diffusion, broadcast IP
    - 130.190.255.255 désigne toutes les machines du réseau 130.190
  - 0.0.0.0 : une machine ne connaît pas son adresse
    - (station sans disque qui utilise RARP)
    - client DHCP



IP : adresses particulières (2)

- RFC 1918 : Address allocation for private networks
  - 10.0.0.0-10.255.255.255 (10/8 préfixe)
  - 172.16.0.0-172.31.255.255 (172.16/12 préfixe)
  - 192.168.0.0-192.168.255.255 (192.168/16 préfixe)

  Utilisées pour :
  - un adressage interne à une organisation
    - non connectée à Internet
    - Connectée à Internet avec translation d ’adresse (NAT ou PAT)
  - L ’interconnexion de réseaux : interface WAN de 2 routeurs
- réseau test : 192.0.2.0-192.0.2.255 (192.0.2/24 préfixe)
- Subnet réseau et sunbet broadcast pas utilisé (RFC 950) : bizarre...

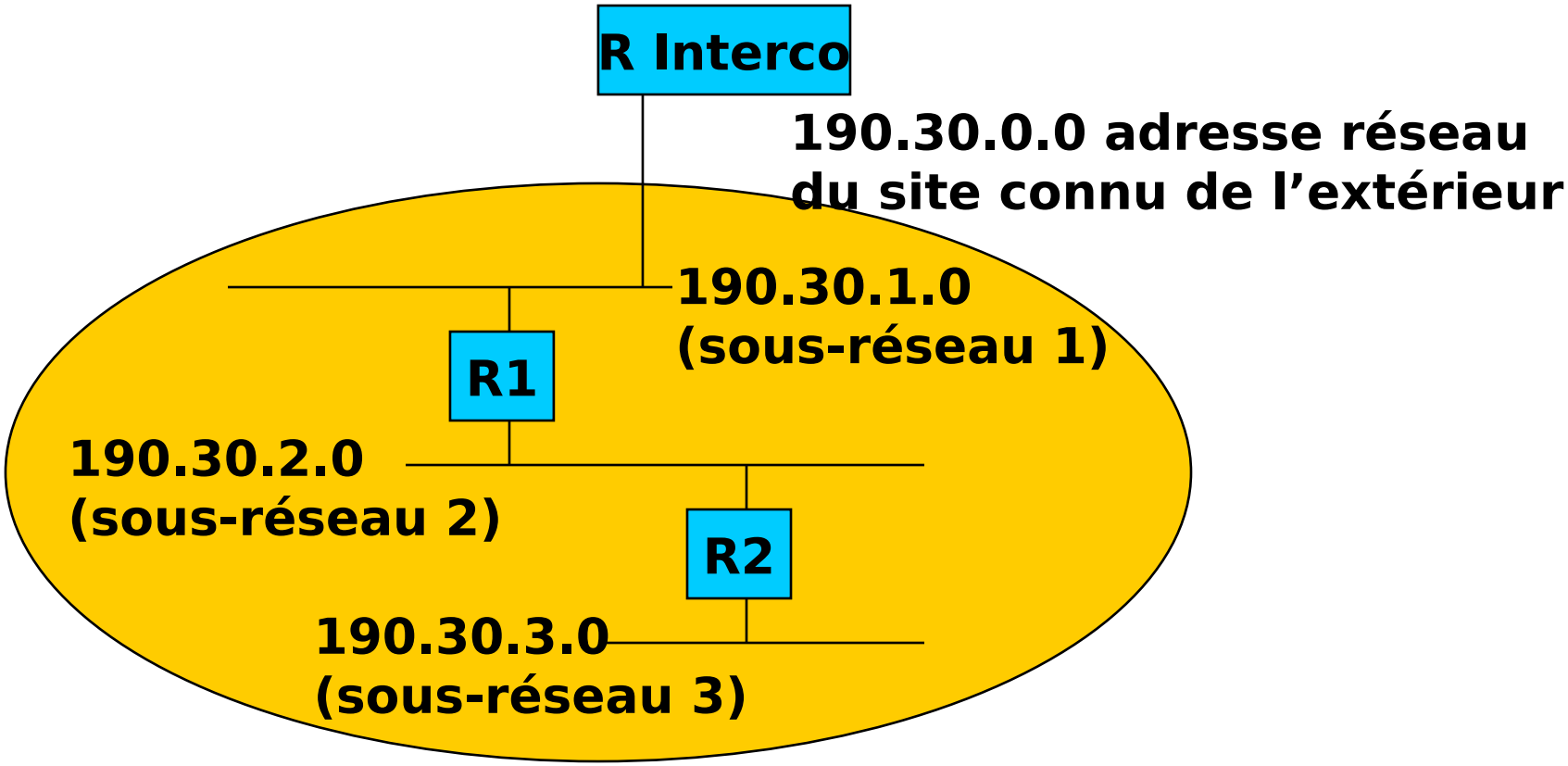
Sous réseaux IP : subnets

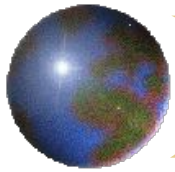
- Découpage d'un réseau en entités plus petites
  - sous-réseau ou "subnet"
  - permet meilleure structuration du réseau du site
  - décidé par l'administrateur du site
  - adresse de sous-réseaux prélevé sur la partie « host-id »
  - longueur comptée en bits décidée par l ’administrateur

  - interconnexion des sous-réseaux impérativement par des routeurs

Sous réseaux IP (2)

- Exemple : découpage en 3 sous-réseaux, numérotation par le 3ème octet





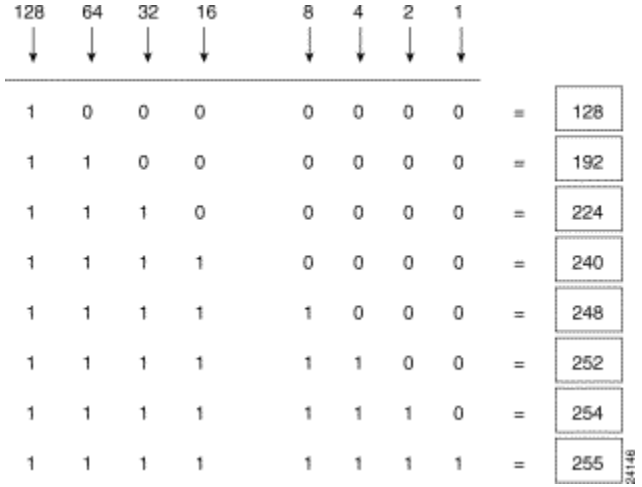
Sous réseaux IP (3)

- le découpage est inconnu de l'extérieur !
- passe par l ’utilisation d ’un subnet-mask
  - même notation que l'adresse IP:
    - bits réseau à 1
    - bits de la partie sous-réseau à 1
    - bits de la partie "host" à 0

Classe	Subnet Mask	Subnet Mask	Préfixe réseau
Class A	11111111 00000000 00000000 00000000	255.0.0.0	/8
Class B	11111111 11111111 00000000 00000000	255.255.0.0	/16
Class C	11111111 11111111 11111111 00000000	255.255.255.0	/24

- utilisation :
  - @IP & subnet_mask = adresse network +subnet
    - utilisée pour le routage local au site
  - @IP & !(subnet_mask) = host id effectif

Sous réseaux IP (4)

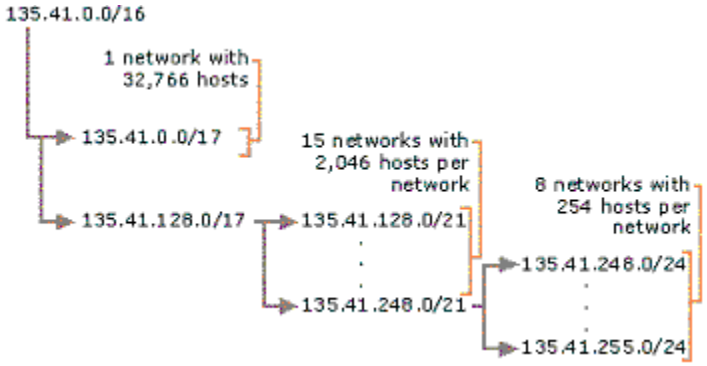


Nombre de subnets nécessaires	Subnet Mask	Numbre de hosts par subnet
1-2	255.255.255.128 ou /25	126
3-4	255.255.255.192 ou /26	62
5-8	255.255.255.224 ou /27	30
9-16	255.255.255.240 ou /28	14
17-32	255.255.255.248 ou /29	6
33-64	255.255.255.252 ou /30	2

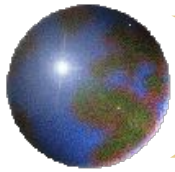
- Je veux router 1 réseau avec 32 machines, solution :
  - diviser une classe C en 8 sous-classes de 32 adresses
  - il me reste 7 sous réseaux de 32 adresses...
- Efficacité ?
- Variable Lenght Subnet Mask, CIDR

Sous réseaux IP : VLSM

- Variable Lenght Subnet Mask
- Possibilité de differents masks dans une même classe
- Exemple / classe B :



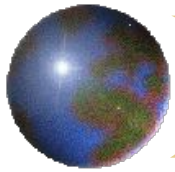
- Autre exemple classe C : 192.168.10/24



IP : classes et CIDR

- CIDR (classless inter domain routing) : plusieurs réseaux peuvent être décrits par une seule route
- RFC 1466, 1518,1519
- réduire taille des tables de routage
- Hiérarchisation (géographique) des réseaux
  - regrouper les réseaux affectés à une même zone géographique dans une même classe
  - les décrire par une seule route
- Ex : RIPE a affecté les classes C : 193/8, 194/8 et 195/8 (193.0.0.0 à 195.255.255.255) à Renater

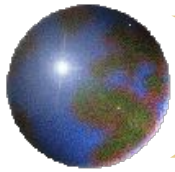
  ces classes peuvent être décrites par 2 entrées dans les tables de routage au lieu de 3 :
  - 193/8 : 193.0.0.0 netmask : 255.0.0.0
  - 194/7 : 194.0.0.0 netmask : 254.0.0.0



IP : classes et CIDR (2)

- Correspondance Netmask / nombre de réseaux de l'agrégat :
  - 255 => / 24 R = 1
  - 254 => / 23 R = 2
  - 252 => / 22 R = 4
  - 248 => / 21 R = 8
  - 240 => / 20 R = 16
  - ...

R = (255 - netmask) + 1



CIDR

- Solutions à court terme :
  - "découper" les réseaux de classe A
    - en utilisant la technique des masques (*subnet mask*) .
    - 1 réseau de classe A permet d'adresser > 16 millions de *hosts*

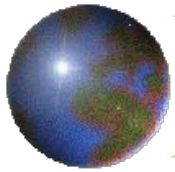
  => allocation de réseaux sans classe (classless)

  - agréger les tables de routage :

  => allouer aux "utilisateurs" des réseaux de classe C contigus
  des réseaux contigus ont les mêmes bits de poids fort :
  ils ont même préfixe

  => grouper les préfixes par région, prestataires ...

  => router les préfixes des supernets (ou agrégats)
  une seule entrée par agrégat dans la table de routage suffit

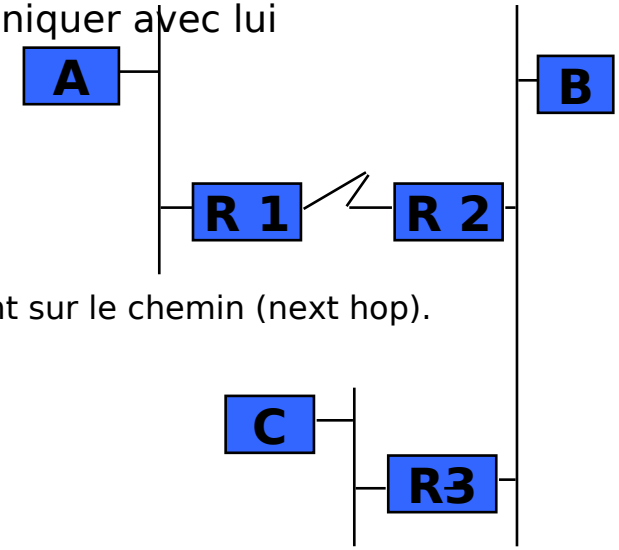


Adressage IP : efficacité ?

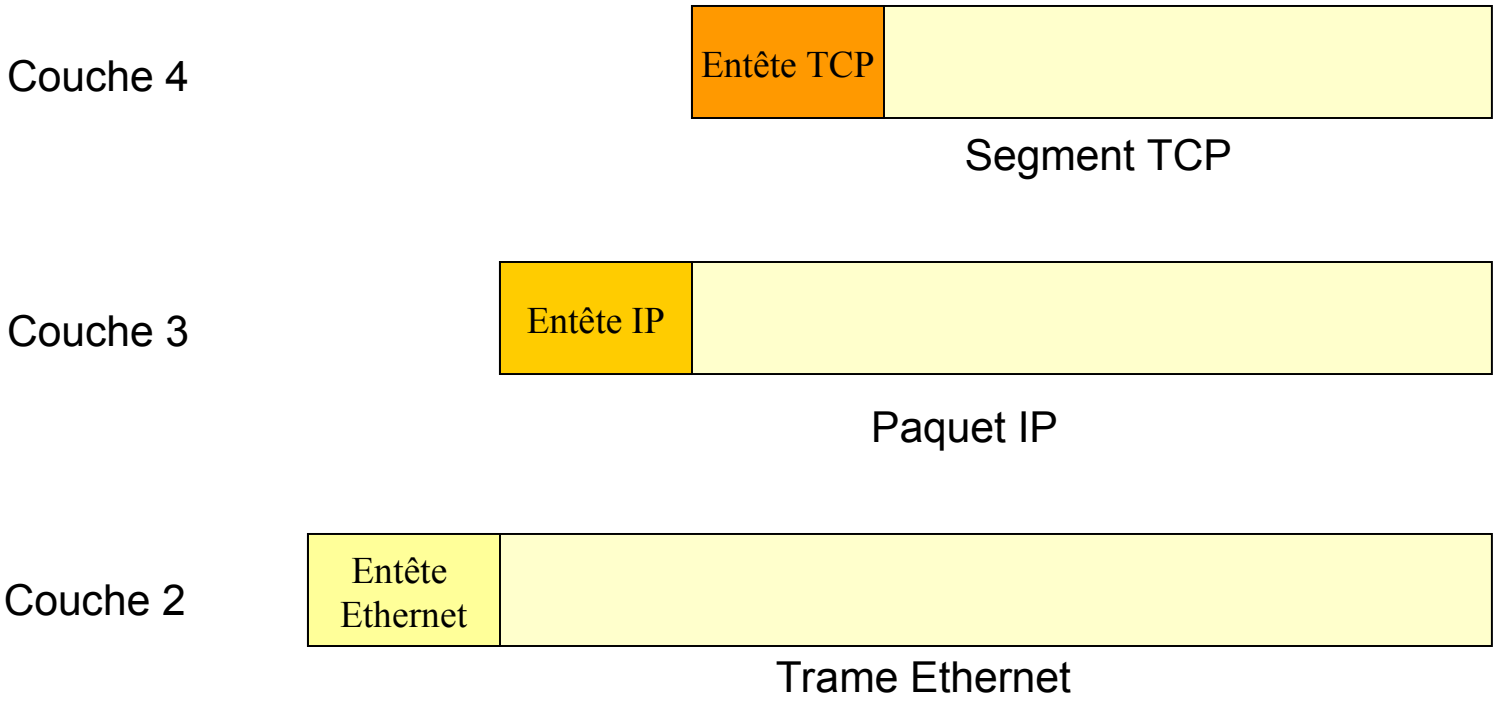
- Pour chaque réseau perte de 2 adresses
  - réseau x.y.z.0
  - broadcast : x.y.z.255
- de nombreuses adresses sont spéciales
- manque d ’adresses :
  - classes A : épuisées : en cours de ré affectation
  - Classes B : quasi-épuisées
  - Classes C : 220-223 restent pour l ’instant disponibles
- techniques d ’économie d ’adresses
  - adressage dynamique : DHCP
  - translation d ’adresse (NAT, PAT)

IP : Fonctions

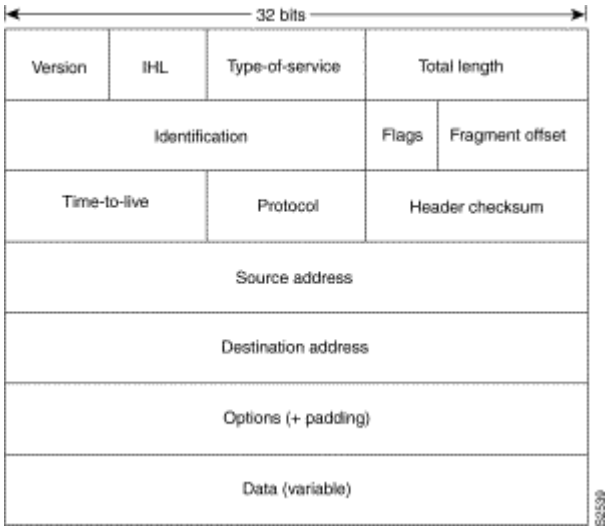
- Transporte des datagrammes de bout en bout
- Pour aller de l' équipement A à l'équipement C, le datagramme passe par R1, R2, R3
- Chaque datagramme contient :
  - l'adresse IP (Internet) de l'émetteur
  - l'adresse IP (Internet) du destinataire
  - chaque interface d'un équipement a une adresse IP
- Il faut connaître l'adresse IP d'un équipement pour communiquer avec lui
- C'est un mode sans connexion
  - chaque datagramme est traité indépendamment des autres
- Sans garantie de remise des datagrammes (unreliable)
  - IP fait au mieux (Best Effort)
- Assure le routage : savoir où envoyer le datagramme.
  - Les équipements IP ne connaissent que le prochain équipement sur le chemin (next hop).



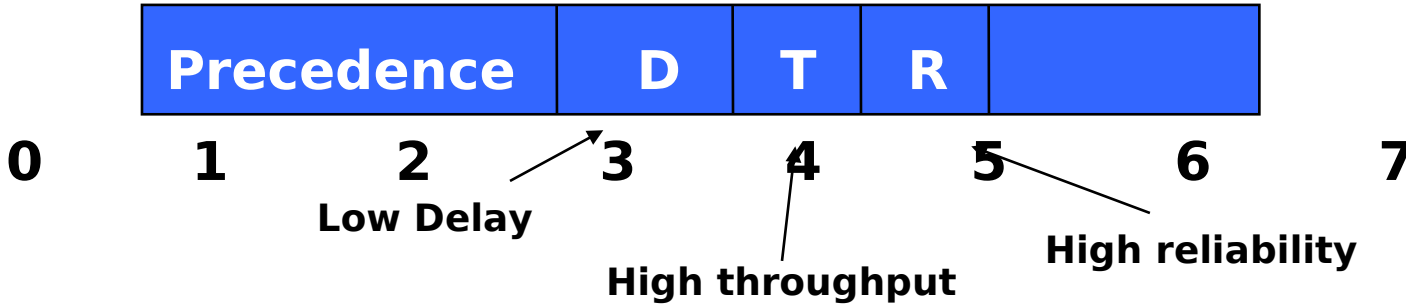
Paquet IP

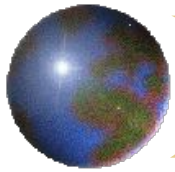

Couche 4
Entête TCP
Segment TCP
Couche 3
Entête IP
Paquet IP
Couche 2
Entête Ethernet
Trame Ethernet

IP : contenu du paquet (1)



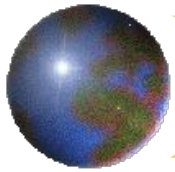
- Version : 4 (IPv4)
- IHL : Internet Header Length
  - En nombre de mots de 32 bits
  - Généralement 5, sans options l'en-tête est de 20 octets
- TOS : Type Of Service
  - Etait prévu pour routage avec contrainte de qualité de service... mais n 'a pas été utilisé !
  - Precedence (0-7 indiquant l'importance du datagramme)





IP : contenu du paquet (2)

- TL : Total Length
  - Longueur du datagramme, incluant l'entête
  - Unité = octet
  - Maximum : 64 Koctets
  - Recommandé : moins de 576 octets
- TTL : Time To Live
  - Théoriquement exprimé en incréments de 1 seconde
  - L'expéditeur le met à une certaine valeur
  - Décrémenté de 1 (ou plus) à chaque traversée de routeur
  - Le datagramme est détruit quand TTL=0
  - Évite au datagramme de circuler éternellement en cas de boucle



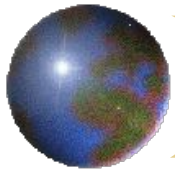
IP : contenu du paquet (3)

- Protocol : Identifie le protocole de la couche supérieure

Numéro	Nom court	libellé
1	ICMP	Internet Control Message [RFC792]
2	IGMP	Internet Group Management [RFC1112]
4	IP	IP in IP (encapsulation) [RFC2003]
6	TCP	Transmission Control [RFC793]
8	EGP	Exterior Gateway Protocol [RFC888,DLM1]
17	UDP	User Datagram [RFC768,JBP]
41	IPv6	Ipv6 [Deering]
46	RSVP	Reservation Protocol [Bob Braden]
47	GRE	General Routing Encapsulation [Tony Li]
103	PIM	Protocol Independent Multicast [Farinacci]

source : ftp://ftp.isi.edu/in-notes/iana/assignments/protocol-numbers

- Header Checksum
  - couvre l ’entête IP uniquement
  - but: vérifier son intégrité
  - recalculé par chaque routeur (puisque le champ TTL est modifié)
  - mais ne couvre pas les données (de la responsabilité du transport)

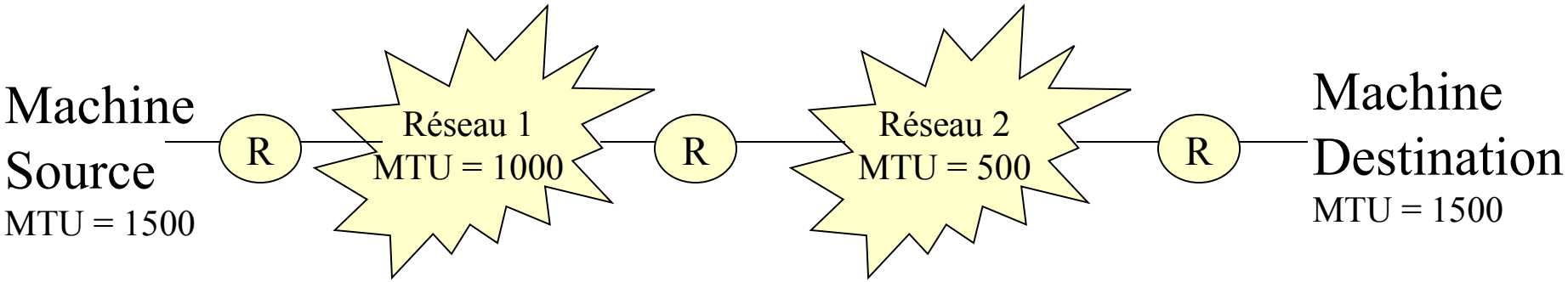


IP : contenu du paquet (4)

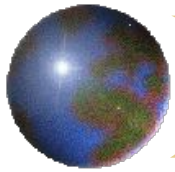
- Champs liés à la fragmentation IP :
  - ID : Identification du datagramme
    - Utilisé par l'émétteur et le destinataire pour identifier le datagramme
    - Numérotation faite par l'émetteur
    - Uniquement utilisé pour la fragmentation
  - Flags : Flags pour la fragmentation
    - 001 : il y a encore des fragments
    - 000 : dernier fragment (ou pas encore fragmenté)
    - 01X : ne pas fragmenter
  - FO : Fragment offset
    - position du fragment dans le datagramme d ’origine,
    - calculé en unité de 8 octets,
    - premier fragment = 0
    - Le destinataire doit récupérer tout les fragments, si un fragment est perdu tout le datagramme est jeté.

IP : MTU

- Maximum transmission unit
- Ethernet : 1518 octets, FDDI : 4500 octets, Token Ring : 2 à 4 Ko
- TU = datagramme IP (en-tête + données)
- Chaque sous réseau peut faire transiter des datagrammes IP d ’une longueur donnée : MTU
- C'est la machine destinataire qui ré assemble non le routeur à la frontière d'un type de réseau



Dessin paquets fragmentés



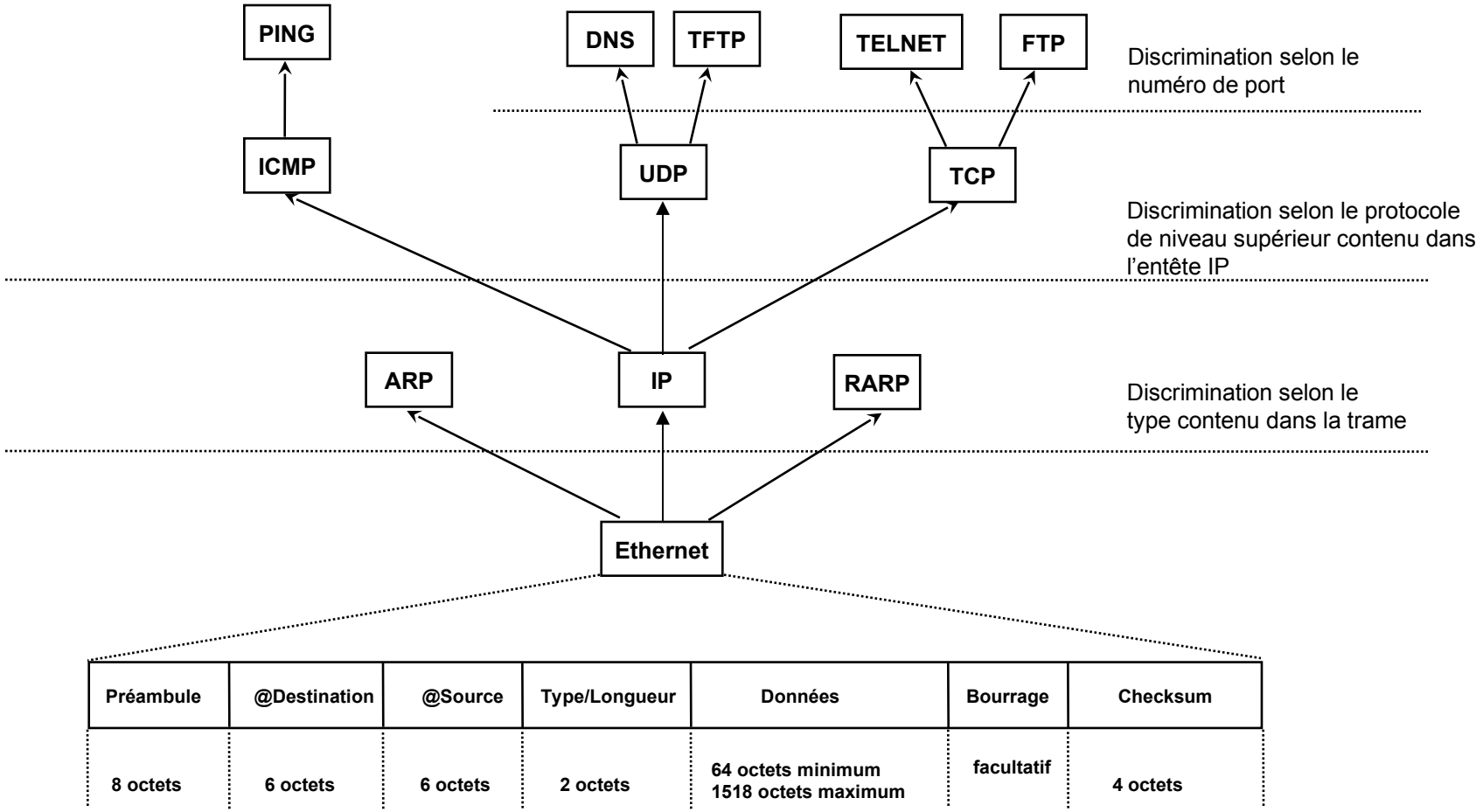
IP : contenu du paquet (5)

- Options : variables en taille, permettent des extensions
  - Certaines options sont standards (décrites dans des RFCs).
    - Exemples : niveau de sécurité, time stamp (chaque routeur ajoute l'heure de passage)
  - Elles se composent
    - du code de l'option 1 octet
    - de la longueur de l'option 1 octet
    - des données associées
- Padding : Compléte le champs options
  - Pour que la longueur de l'en-tête soit un multiple de 32.

  Remarque :

  Taille de l'entête est importante (> 20 octets)
  - IP ne peut pas utiliser un lien < 4.8 Kb/s
  - Certains protocoles permettent la compression de l'entête IP

IP et les protocoles

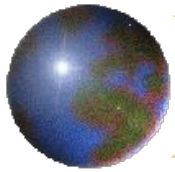


Nom	Port/protocole	Description
Echo	7/tcp et 7/udp	Echo
Discard	9/tcp et 9/udp	Discard
ftp-data	20/tcp	File Transfer protocol (data)
ftp	21/tcp	File Transfer protocol (control)
Ssh	22/tcp	Secure Shell (SSH) remote login protocol
telnet	23/tcp	Telnet
Smtp	25/tcp	Simple Mail Transfert Protocol
Time	37/tcp et 37/udp	Time
Domain	53/tcp et 53/udp	Domain Name System (DNS)
Whois	63/tcp et 63/udp	Whois++ protocol
Bootps	67/tcp et 67/udp	Bootstrap protocol server
Bootpc	68/tcp et 68/udp	Bootstrap protocol client
Tftp	69/tcp et 69/udp	Trivial File Transfer
Gopher	70/tcp et 70/udp	Gopher (ancètre du web)
Finger	79/tcp et 79/udp	Finger
http	80/tcp	World Wide Web (http)
Pop2	109/tcp	Post Office Protocol –v2
Pop3	110/tcp	Post Office Protocol –v3
Sunrpc	111/tcp et 111/udp	Sun Remote Procedure call
Ident auth	113/tcp	Authentification service
nntp	119/tcp	Network News Transfer Protocol
Netbios-ns	137/tcp et 137/udp	NETBIOS Name Service
netbios-dgm	138/tcp et 138/udp	NETBIOS Datagram Service
netbios-ssn	139/tcp et 139/udp	NETBIOS Session Service
Imap	143/tcp	Internet Message Access Protocol
xdmcp	177/tcp et 177/udp	X Display Manager Control Protocol
Bgp	179/tcp et 170/udp	Border Gateway Protocol
z39.50	210/tcp	ANSI Z39.50
Exec	512/tcp	remote process execution
Login	513/tcp	remote login a la telnet
Shell	514/tcp	Remote command
Syslog	514/udp	Log (évennements) à distance
Printer	515/udp	Spooler
Talk	517/tcp et 517/udp	Dialogue en direct

- Liste non exaustive des ports TCP ou UDP les plus utilisés

Référence :

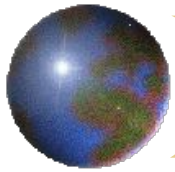
http://www.iana.org/assignment/port-numbers



ARP (RFC826)

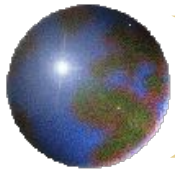
- L'adresse IP est totalement indépendante de l'adresse physique
  - exemple Ethernet : l'adresse MAC Ethernet unique est fournie par le constructeur (l ’IEEE fournit au constructeur un bloc d ’ adresses).
  - Une machine peut avoir plusieurs adresses IP indépendamment du nombre d ’interfaces
- Problème : Trouver une adresse MAC à partir de l'adresse IP
- Address Resolution Protocol
  - Permet de trouver l'adresse physique d'une machine sur le même réseau en donnant uniquement son adresse IP
- Stockage des adresses physiques dans une table ARP (cache).
  - Le cache est remis à jour périodiquement
  - Sous Unix, pour Ethernet, visualisation de la table par la commande :

    arp -a



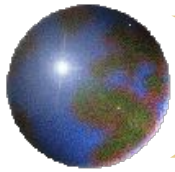
ARP

- Soit deux équipements sur le même segment Ethernet.
- La machine A veut envoyer un datagramme à la machine B.
- Elle connait son adresse IP, mais pas son adresse Ethernet :
  - A envoie une trame de broadcast Ethernet qui demande l'adresse Ethernet de B :
    - adresse destinataire FF.FF.FF.FF.FF.FF avec Type = 0806
    - en indiquant l'adresse IP de B.
  - Toutes les machines reçoivent la requête.
  - Seul B répond à A en lui donnant son adresse Ethernet.
  - Si c'est une autre machine qui répond à la place de A on parle alors de "Proxy ARP".
    - utilisation par les routeurs quand le destinataire est dans un autre réseau IP



RARP (RFC903)

- Problème :
  - Trouver une adresse IP à partir de l'adresse Ethernet
- Reverse Address Resolution Protocol
  - Permet de demander une adresse IP en indiquant l'adresse Ethernet.
  - Utilisé au moment du "boot" par certains équipements, envoie son adresse MAC dans le champ "Target HA"
- Type = 8035 dans la trame Ethernet
- Ethernet type = 32821 dans la trame IEEE802.3
- Utilisé par
  - les Macintosh avec une boite Kinetics
  - les stations sans disque
  - les terminaux X
- Même format de message que ARP
- Problème : diffusion broadcast :
  - nécessité d ’un Proxy ARP pour traverser les routeurs
  - nécessité d ’un serveur RARP sur chaque réseau
  - Utilisation d ’un serveur BOOTP (protocole UDP qui est routé)



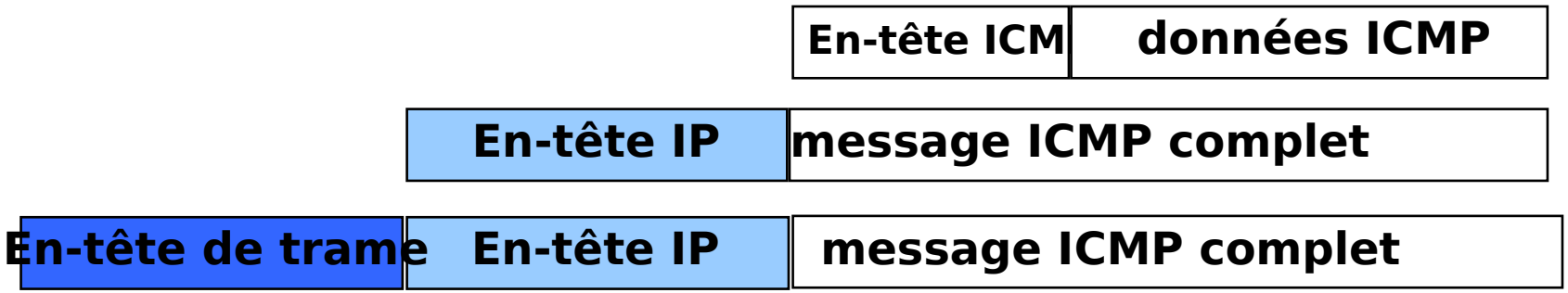
ICMP (RFC792)

- Internet Control Message Protocol
  - Protocole de "gestion" de réseau = *mécanisme de rapport d'erreur.*
- Implémenté sur tous les équipements IP : stations, routeurs.
- Message envoyé par l'équipement destinataire ou un routeur intermédiaire
  - Quand il s'aperçoit d'un problème dans un datagramme
  - Pour avertir l'émetteur afin qu'il modifie son comportement.
  - ex : routeur qui a une mauvaise information de routage.

- Un message ICMP ne doit pas engendrer un autre message ICMP
  - Il ne demande pas de réponse

ICMP

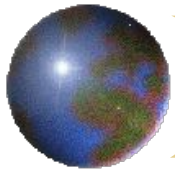
- Un message ICMP est contenu dans un datagramme IP
  - Utilise IP comme un protocole de couche supérieure.
  - Champ protocole du datagramme IP = 1
- Chaque message ICMP a son format (3 champs communs)
  - TYPE (1 octet), 22 types définis
  - CODE (1 octet), plus d'information sur le champ type.
  - CHECKSUM (2 octets), sur le message ICMP.



options

Type	Nom	Code
0	Echo Reply	
3	Destination Unreachable	0 Net Unreachable 1 Host Unreachable 2 Protocol Unreachable 3 Port Unreachable 4 Fragmentation Needed and Don't Fragment was Set 5 Source Route Failed 6 Destination Network Unknown 7 Destination Host Unknown 8 Source Host Isolated 9 Communication with Destination Network is Administratively Prohibited 10 Communication with Destination Host is Administratively Prohibited 11 Destination Network Unreachable for Type of Service 12 Destination Host Unreachable for Type of Service 13 Communication Administratively Prohibited [RFC1812] 14 Host Precedence Violation [RFC1812] 15 Precedence cutoff in effect
4	Source Quench (demande ralentissement)	
5	Redirect	0 Redirect Datagram for the Network (or subnet) 1 Redirect Datagram for the Host 2 Redirect Datagram for the Type of Service and Network 3 Redirect Datagram for the Type of Service and Host
8	Echo	
11	Time Exceeded (ttl expiré)	0 Time to Live exceeded in Transit 1 Fragment Reassembly Time Exceeded
12	Parameter Problem (problème de parametrage)	0 Pointer indicates the error 1 Missing a Required Option [RFC1108] 2 Bad Length
30	Traceroute	

- ftp://ftp.isi.edu/in-notes/iana/assignments/icmp-parameters

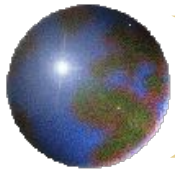


ICMP : options les plus courantes

- Type d'indication d'un message ICMP

	(type message, code)
- Impossible d'atteindre le réseau (Network unreachable)	(3,0)
- Host non atteignable (Host unreachable)	(3,1)
- Port TCP or UDP (service) indisponible (Port unreachable)	(3,3)
- Demande de ralentir l'émission (Source quench)	(4,0)
- Durée de vie dépassée (Time to Live exceeded)	(11,0)
- Redirection (Redirect, change a route)	(5,0-3)
- Echo et réponse à echo (Echo Echo reply) , commande ping	(8)
- Demande de "subnet mask" (Address Mask request)	(17)
- Réponse de " subnet mask" (address mask Reply)	(18)

- Permet de palier aux manques de services de IP



ICMP Redirect

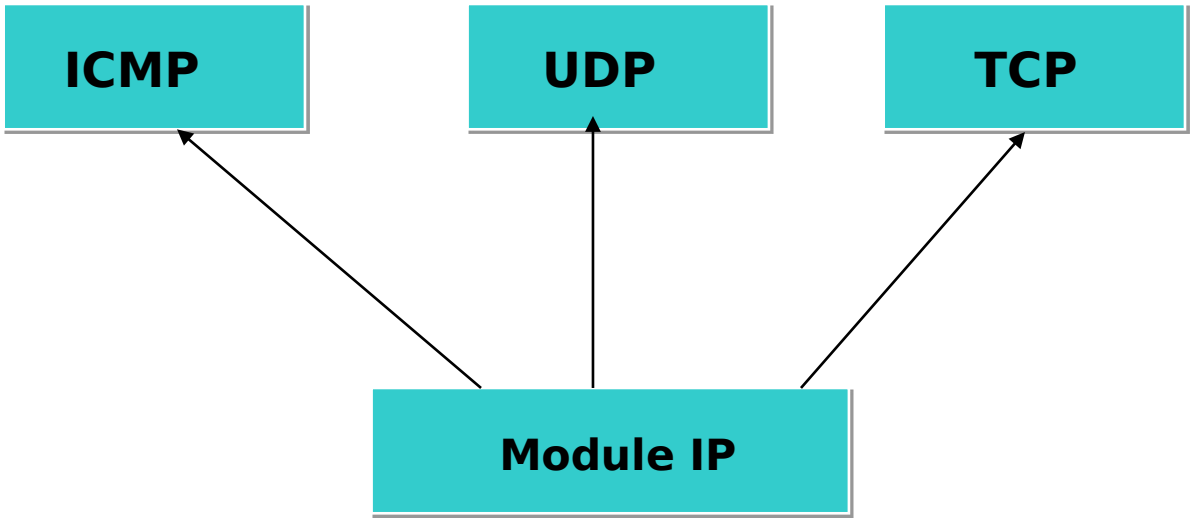
L e routage statique n'est pas complètement statique :

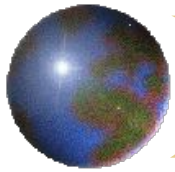
- si la machine B (mal configurée) envoie un datagramme IP pour la machine A au routeur R3
- R3 envoie ce datagramme à R2 qui le transmet à R1 qui le délivre à A
- Puis R3 envoie un message *ICMP redirect* à B disant que pour atteindre A il faut passer par R2.
- B ajoutera cette information dans sa table de routage
  - s'il supporte ICMP redirect
- Ce mécanisme évite la mise à jour manuelle de toutes les machines quand on ajoute un routeur

  => par contre, il faudra mettre à jour les routeurs

IP : couche transport (1)

- Deux protocoles pour la communication entre applications
  - TCP : Transmission Control Protocol
    - protocole en mode orienté connexion
  - UDP: User Datagram protocol
    - protocole en mode sans connexion



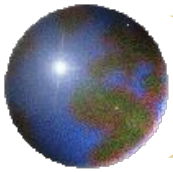


IP : couche transport (2)

- Identification d'une application : numéro de port
  - le port est une destination abstraite utilisé par le protocole

- socket = Combinaison @ IP - Numéro de port
  - 130.190.5.1 - 23 est le démon telnetd sur la station 130.190.5.1

- La combinaison de 2 sockets définit complètement une connexion TCP ou un échange UDP
  - Exemple 130.190.5.1 - 23 et 147.171.150.2 - 1094
    - Connexion entre un processus client qui a pris le numéro 1094 sur la machine 147.171.150.2 et le démon telnetd sur la machine 130.190.5.1
    - Un utilisateur sur 147.171.150.2 a fait un telnet 130.190.5.1
    - C'est ce que l'on peut voir avec la commande Unix *netstat -a*

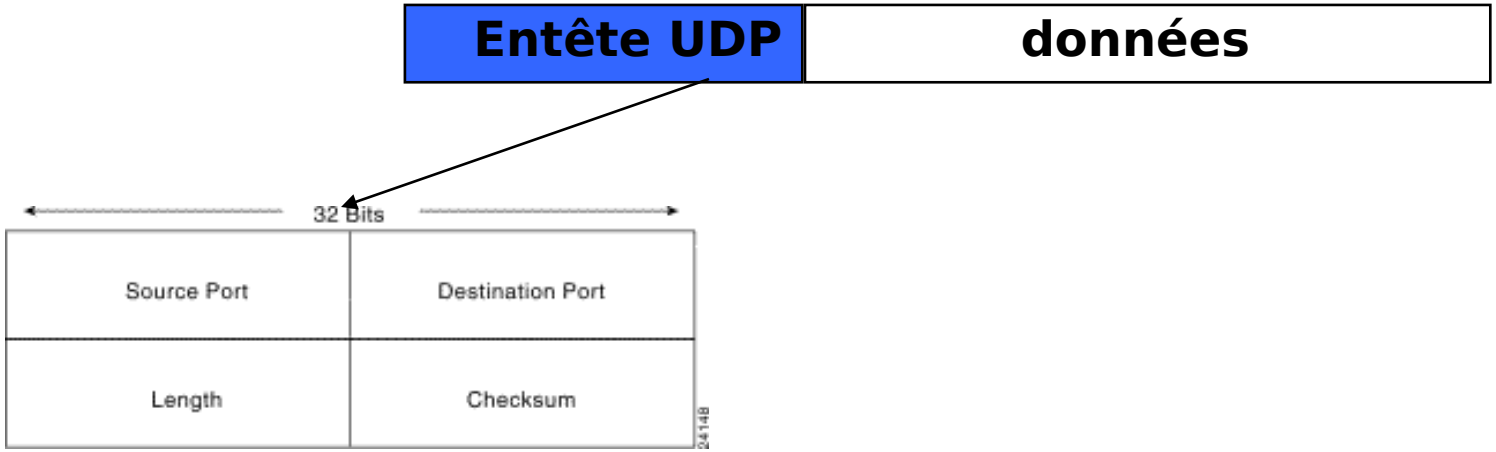


IP : Couche transport (3)

- Mode client serveur
  - serveur , on parle de démons (daemon) dans Unix : démon = processus offrant un service réseau
  - le client se voit attribué un numéro de port non affecté (>1024) pour éviter toute confusion avec les ports "officiels"
- Tous les équipements TCP/IP respectent cette attribution de ports pré-définis
  - fichier /etc/services d'Unix

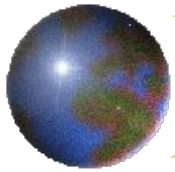
UDP (RFC768)

- User Datagram Protocol
  - service sans connexion, sans garantie, utilisant IP pour le transport de messages entre machines



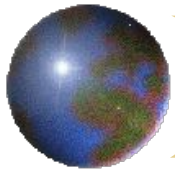
UDP (1)

- Entête de 8 octets
- Source Port : numéro de port
  - optionnel, identifie un port pour la réponse.
- Destination Port : numéro de port
- Length, taille de l'entête et des données
  - Unité = octet
  - Taille maximale = 64 Koctets (- entête IP)
- Checksum : fonction de l'entête et des données
  - optionnel
  - c'est la seule garantie sur la validité des données qui arrivent à destination
- Un datagramme UDP est contenu dans un datagramme IP



UDP (2)

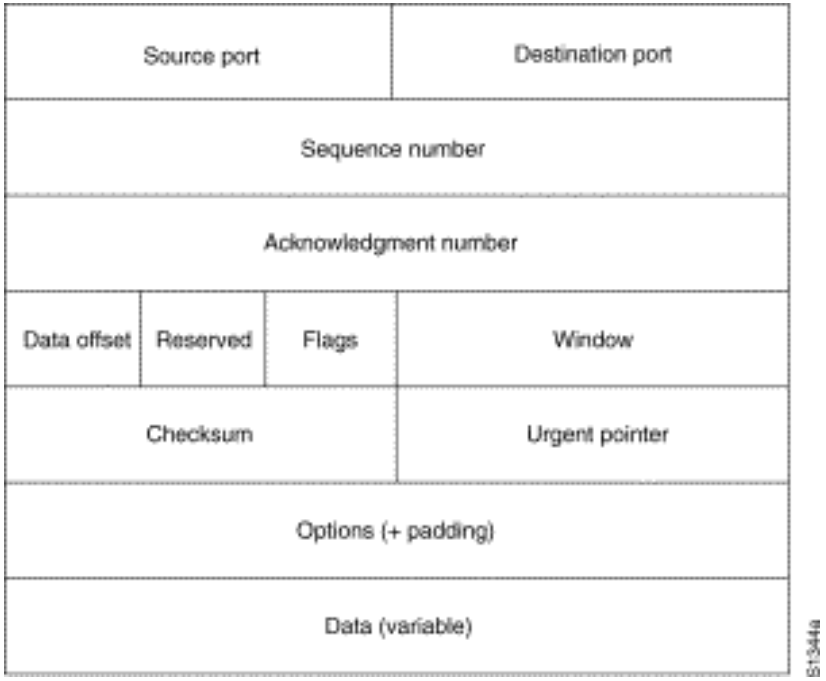
- Ne fait pas
  - mode connecté,
  - retransmission si erreur ou perte,
  - séquencement,
  - contrôle de flux
- C'est un protocole de transport non fiable
- Utilisé pour la diffusion, exemple
  - rwho
  - tftp (Trivial FileTransfert Protocol) , port 69
  - ntp (Network Time Protocol), port 123
  - dès que le multicast est nécessaire
  - multimédia



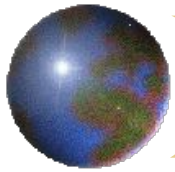
TCP (RFC793)

- Transmission Control Protocol
- Transport
  - De bout en bout entre applications
  - En mode connecté : ouverture ... fermeture (circuit virtuel)
  - Sans erreur : contrôle et retransmission si nécessaire
  - Sans perte : "numérotation" et retransmission
  - Ordonné : bon séquencement
  - Système d'acquittement
  - Avec contrôle de flux (fenêtre d’émission)
  - Full duplex
  - Indication du service par numéro de port

TCP : paquet (1)

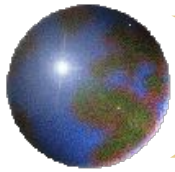


- Taille minimale de 20 octets
  - plus 20 octets en-tête IP => en-tête TCP/IP : 40 octets
- Source Port (16 bits)
  - Numéro du port TCP qui identifie l'application du coté émetteur du segment TCP
- Destination Port (16 bits)
  - Numéro du port TCP qui identifie l'application du coté destinataire du segment TCP



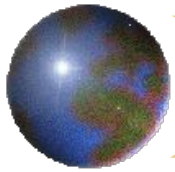
TCP : paquet (2)

- Sequence Number (32bits)
  - Nombre donné en octets.
  - Par rapport au début de la connexion; référence au flux dans la même direction que le segment
  - Assure le bon séquencement.
- Acknowledgment Number (32 bits)
  - Nombre donné en octets, c'est une information en retour.
  - Identifie la position du dernier octet reçu en donnant le numéro du prochain octet que lui,destinataire, espère recevoir.
  - Les octets précédents ont été reçus sans erreur ni perte



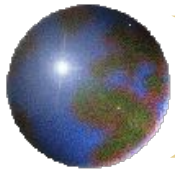
TCP : paquet (3)

- Data Offset ou HLEN, longueur de l'en-tête (4bits)
  - Donné en multiple de 32 bits , souvent = 5
- URG (1bit) : Urgent : Indique que l'application destinataire doit arrêter tout traitement pour interpréter les données urgentes
  - exemple interruption, Ctrl C dans telnet
- ACK (1bit) : acquittement : tenir compte de l'Acknowledgment Number
- PSH (1bit) : délivrer immédiatement les données à la couche supérieure
  - l'émetteur ne prévoit pas d'envoyer d'autres données dans l'immédiat
  - exemple : après une fin de ligne sous telnet en mode ligne)
- RST (1bit) : reset : reprise d' une connexion au départ
  - après plusieurs SYN incompréhensifs ou un crash, ...
- SYN (1bit) synchronisation des séquences : Désire établir une connexion
- FIN (1bit) : Termine la connexion, plus de données à transmettre



TCP : paquet (4)

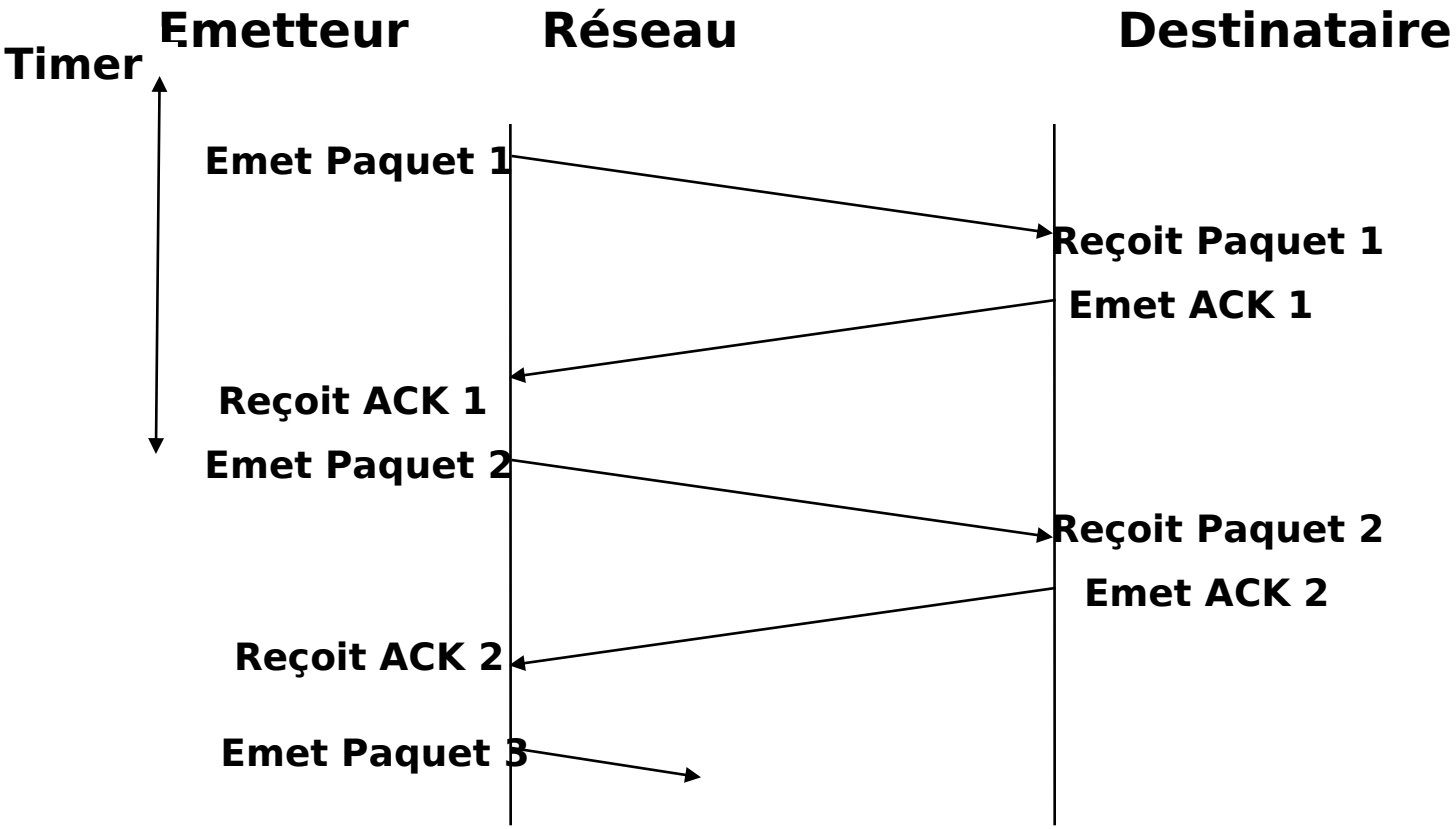
- Window :
  - Nombre d'octets que l'émetteur peut envoyer par rapport à l'Acknowlegment Number sans recevoir d'acquittement (espace libre dans le buffer de réception).
    C'est une information de retour.
  - Permet le contrôle de flux
- Checksum :
  - Fonction sur l'entête et les données
  - Permet de vérifier que le transport s'est effectué sans erreur
  - Si le récepteur s'aperçoit d'une erreur, il fait comme si le segment avait été perdu, il ne l'acquitte pas



TCP : paquet (5)

- ACK et retransmission
  - le schéma d'ACK de TCP est cumulatif
    - référence au nombre d'octets déjà reçu
  - avantages
    - facile à générer et sans ambiguïté.
- Les délais de retransmission
  - pour chaque segment envoyé il y a un timer de déclenché, mais la structure de l'Internet impose des timers variable.
  - algorithme adaptatif
    - ajustement automatique et dynamique, tout au long de la connexion,
    - en fonction des délais d'acquittement des segments précédents ("segment round-trip time") .
      Ceci permet à TCP de s'adapter sans paramétrage, à tous les débits et à tous les temps de réponse, donc à tous les réseaux

TCP : contrôle d ’erreurs (1)



TCP : contrôle d ’erreurs (2)

- Mécanisme « Send and Wait » : La technique la plus simple : on transmet un segment puis on attend l ’acquittement avant de transmettre le suivant
  - si l’on n’a pas d’acquittement à l’arrivée à échéance du timer, on retransmet puis on attend de nouveau...
  - exploite très mal le réseau :
    - il n ’est utilisé que lors de la transmission !

      tt: temps de transmission de la trame d ’information

      tp: temps de propagation

      efficacité = temps utile / temps total = tt / (tt + 2*tp) = 1 / (1 + 2*tp/tt)
    - si la taille du réseau augmente, tp augmente, l ’efficacité diminue !

- On introduit la notion de fenêtre d ’anticipation...

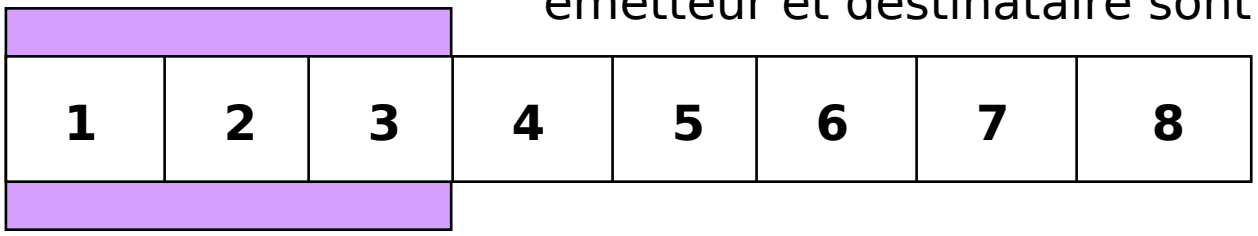
  (ou fenêtre de transmission)

TCP : contrôle d 'erreurs (3)

- Mécanisme du glissement de fenêtre (sliding window)

fenêtre initiale

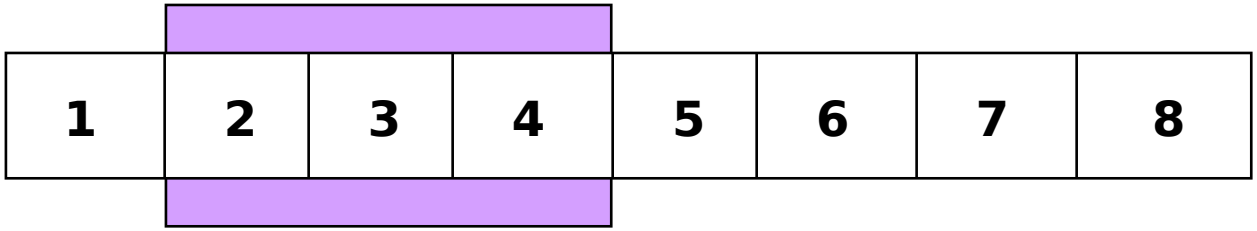
émetteur et destinataire sont concernés



L'émetteur peut envoyer 3 paquets avant de recevoir un acquittement

glissement →

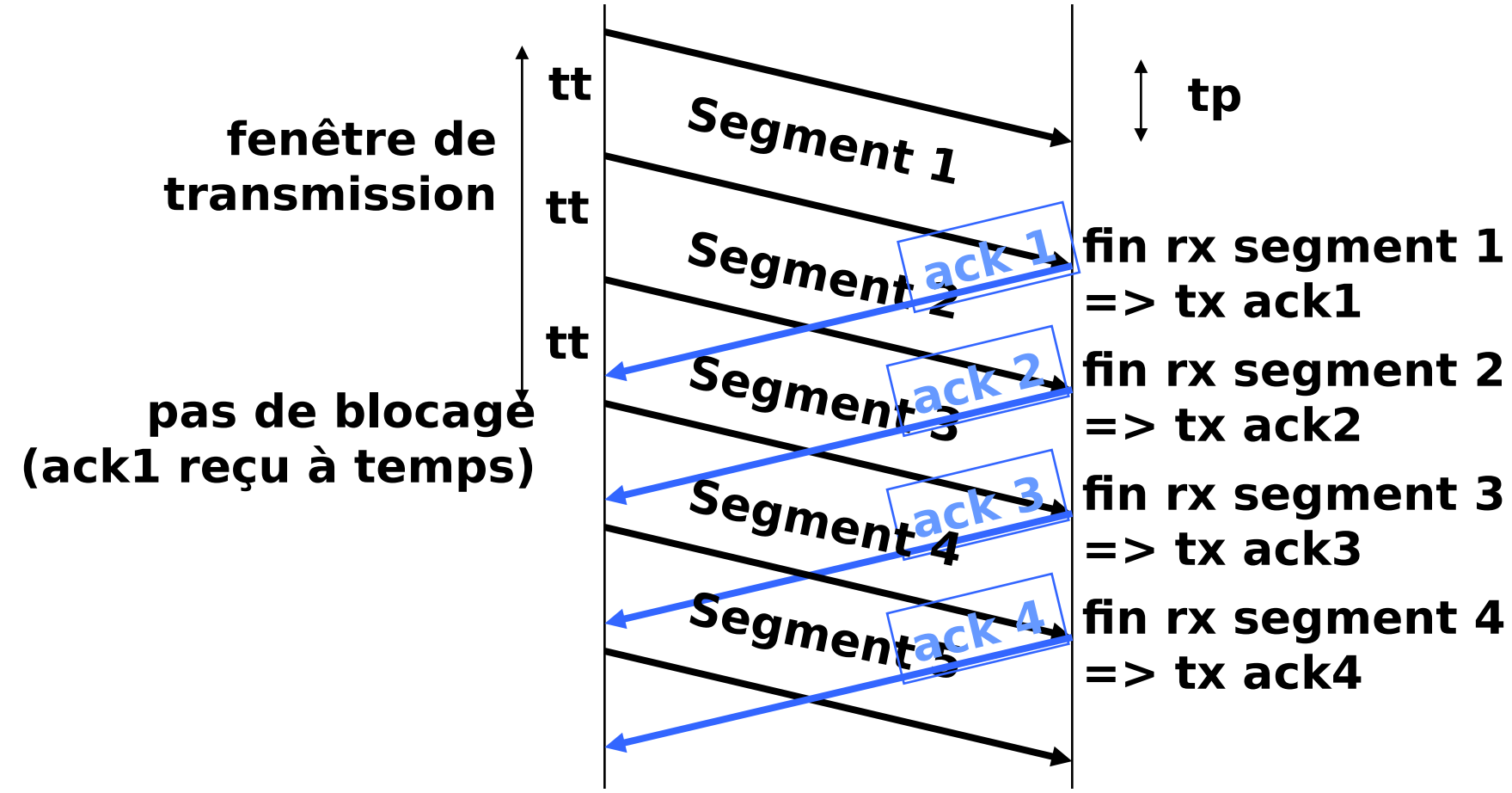
L'acquittement du paquet 1 arrive, la fenêtre glisse

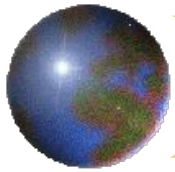


Les performances sont fonctions de la taille de la fenêtre et de la vitesse à laquelle le réseau accepte les paquets.

TCP : contrôle d ’erreurs (4)

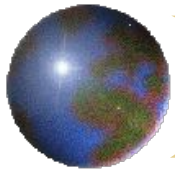
Si la fenêtre a une taille suffisante, il n’y a pas de blocage !





TCP : Contrôle de flux / congestion

- Le destinataire joue avec la place disponible dans les buffers pour réduire la transmission (où l'augmenter)
  - champ « window advertisement » dans les ACK
  - donne le nombre d'octets que le receveur est prêt à accepter à ce moment
  - on modifie la fenêtre d’émission en conséquence
- Contrôle de flux/congestion sont indispensables à l'Internet
  - hétérogénéité des machines pour les communications de bout en bout...
    - TCP résout ce problème avec l ’algorithme de contrôle de congestion du « slow start »
      - perte de segment interprétée comme signe de congestion
      - => on réduit drastiquement la fenêtre de congestion



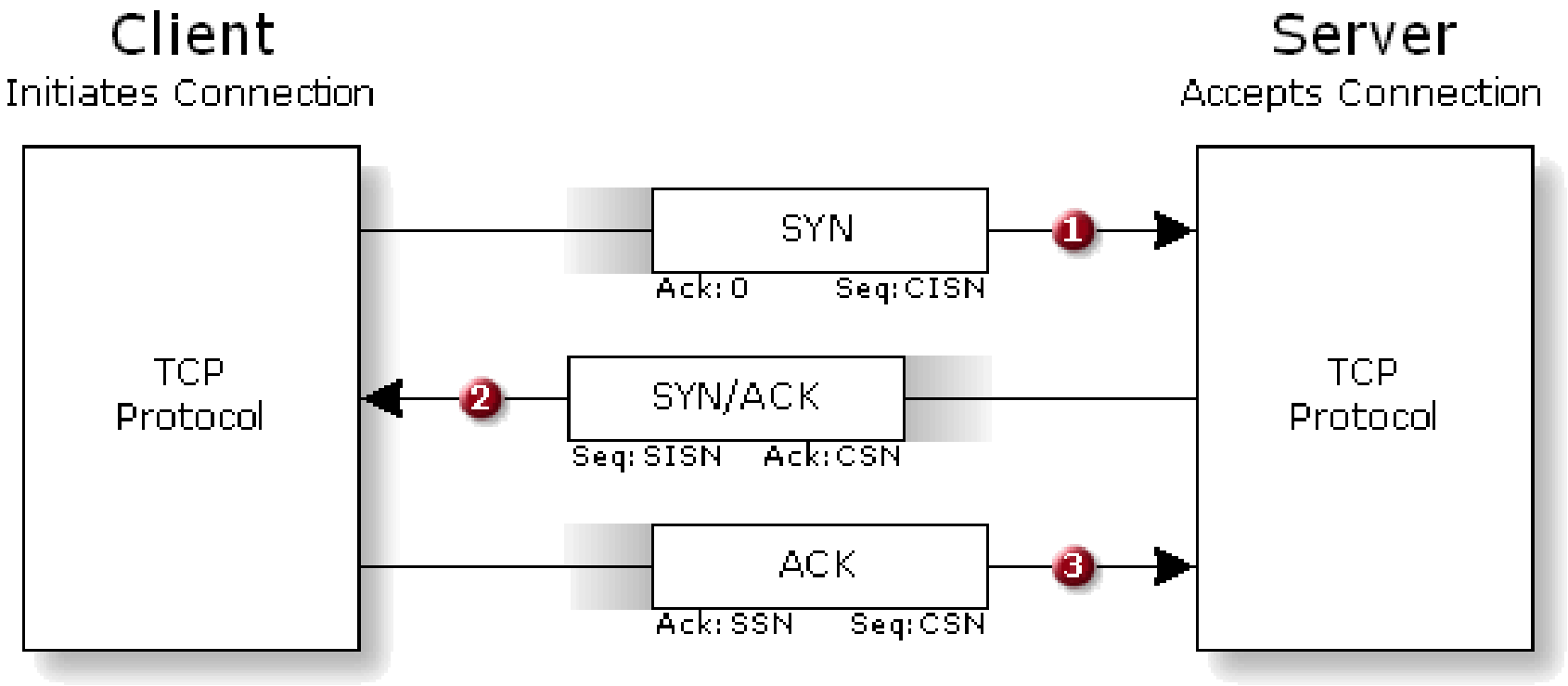
Primitives de haut niveau de gestion de connexion

Primitive	Signification
SOCKET	Création d'un nouveau point terminal
BIND	Attache une adresse locole au socket
LISTEN	Annonce la volonté d'accepter des connexions (non bloquant)
ACCEPT	Bloque l'appelant jusqu'à l'arrivée d'une tentative de connexion
CONNECT	Tentative d'établissement d'une connexion
SEND	Envoi de données sur la connexion
RECEIVE	Réception de données sur la connexion
CLOSE	Libération de la connexion.

- Exemple, des sockets unix
  - serveur : Socket + Bind + Listen + Accept (loop avec Receive + Send) + Close
  - client : Socket + Connect + Loop (Send + Receive) + Close

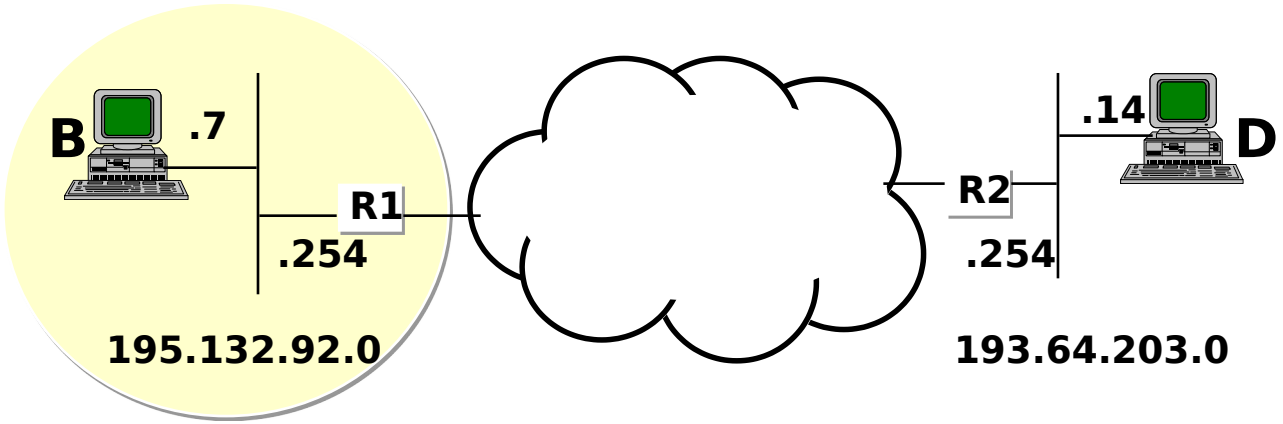
  - socket unix : listen non bloquant
  - socket ISO TPDU : bloquant

TCP : ouverture d 'une connexion (3 way handshake)



Spoofing...

TCP/IP : Entre 2 stations (1)



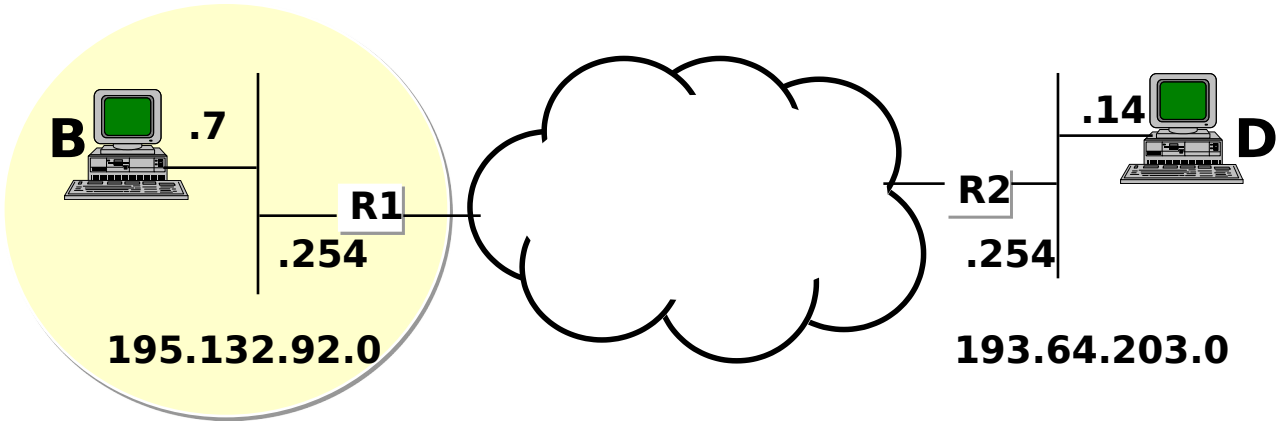
Sur la machine B, "telnet Machine D" Que se passe-t-il ?

- Nom --> @ IP
  - La machine B traduit le nom "machine D" en 193.64.203.14
    - table hosts ou cache ou interrogation d'un DNS.

  Si elle ne trouve pas, message "Host unknown".
- Comment atteindre 193.64.203.14 (1er saut) ?
  - Ce n'est pas un numéro 195.132.92.X
  - Il faut donc passer par un routeur.
  - B consulte sa table de routage : il faut passer par 195.132.92.254

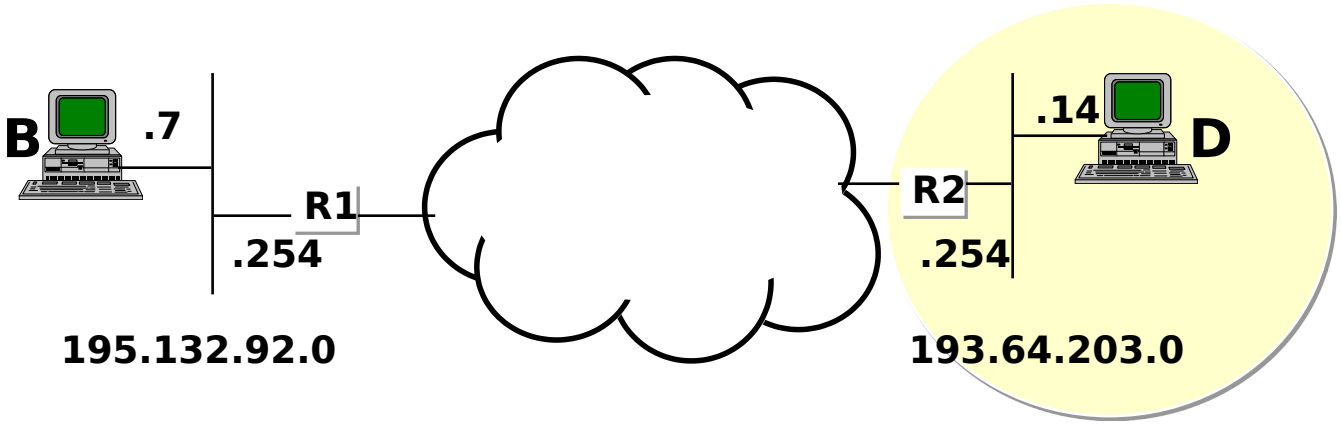
Si elle n'a pas ce renseignement, message "Network unreachable"

TCP/IP : Entre 2 stations (2)



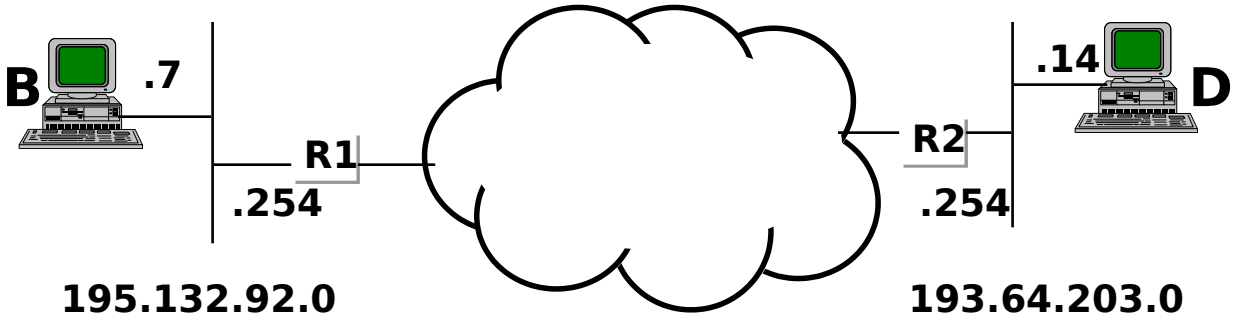
- @IP de R1 --> @ Ethernet de R1
- B émet une trame de "broadcast" Ethernet après avoir vérifié qu'elle n'a pas déjà l'information dans sa table ARP
- Contenant une trame ARP avec la question :
- quelle est l'adresse Ethernet de 195.132.92.254 ?
- R1 répond à B : l'@ Ethernet de 195.132.92.254 est 0:0:c:0:5b:37
  - B envoie une trame Ethernet avec l'@ destination 0:0:c:0:5b:37
    - Incluant un datagramme IP (@ orig 195.132.92.7 et @ dest 193.64.203.14
      - Contenant un segment TCP
        - Avec un numéro de port destinataire 23 (telnetd)
        - SYN=1 (ouverture d'une connexion TCP)

TCP/IP : Entre 2 stations (3)



- R1 reçoit la trame Ethernet
  - Extrait le datagramme IP, l'@ IP du destinataire (193.64.203.14) et cherche où l'envoyer.
  - Il a une interface sur l'Internet (par Renater ou un autre opérateur)
  - Les protocoles de routage font leur travail et le datagramme IP arrive sur R2
- R2 recherche alors l'@ Ethernet de 193.64.203.14
  - S'il ne la trouve pas dans sa table ARP, il envoie un broadcast ARP
  - Il peut ensuite envoyer le datagramme IP à la machine D
- D reçoit le datagramme IP
  - Extrait le segment TCP.
  - Ouvre une session TCP.
  - Avec l'indication de port numéro 23, il appelle la partie telnetd du démon inetd.

TCP/IP : Entre 2 stations (4)

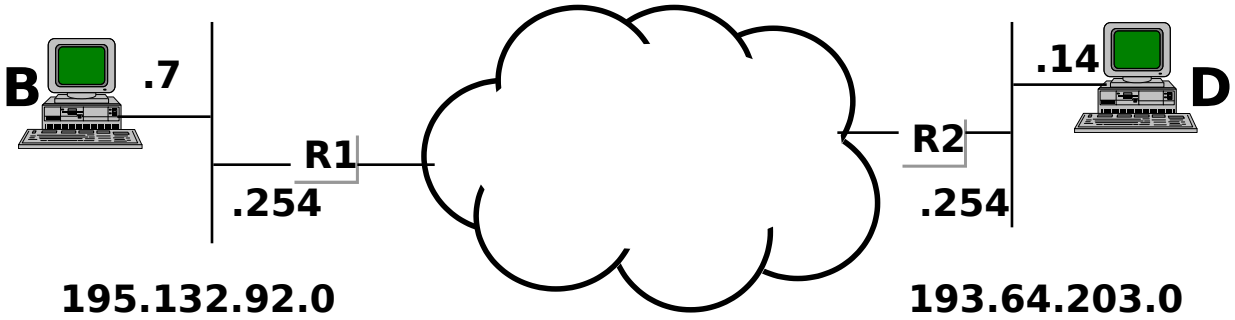


- telnetd demande le nom de l'utilisateur.
  - La question est transportée par un segment TCP, dans un datagramme IP
    - @origine 193.64.203.14, @ destination 195.132.92.7.

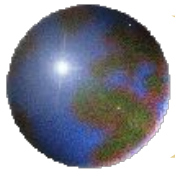
Pour envoyer ce datagramme

- *la machine D cherche l'itinéraire avec la même méthode que la machine B au départ*
  - table de routage, ARP, ...

TCP/IP : Entre 2 stations (5)



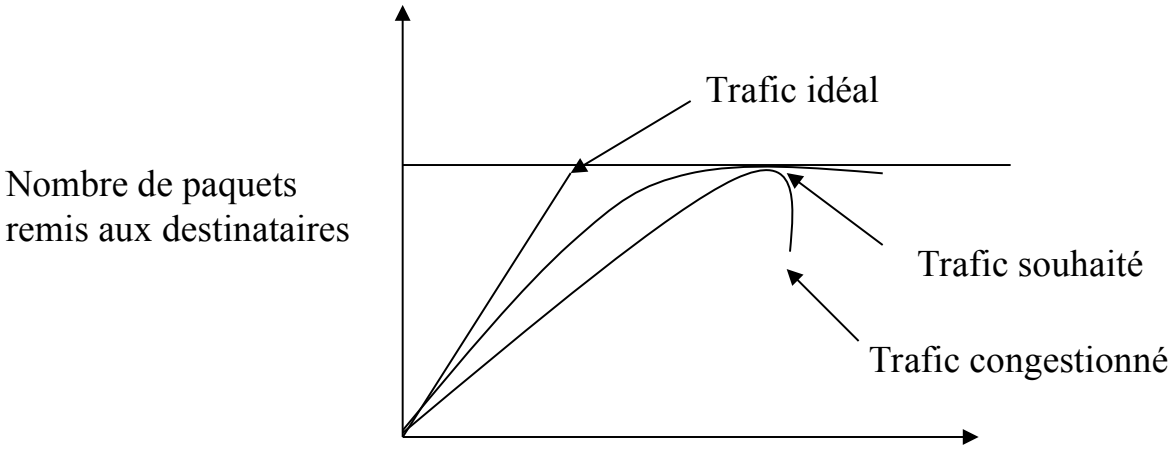
- Remarques
  - D ne tient pas compte de la précédente arrivée d'un datagramme IP pour trouver l'itinéraire de la réponse. Il refait le raisonnement, comme s'il n'avait rien reçu
  - Pour tester la connectivité IP, il n'est pas utile de tester un appel de B vers A, si on a déjà testé un appel de A vers B
  - Le broadcast ARP n'est utilisé que lors de la première recherche d'adresse Ethernet
  - Quand il y a un problème, la recherche d'erreur est difficile quand on n'a pas ce mécanisme en tête



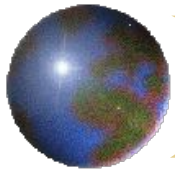
Qualité de service

- Congestion
- Gestion de congestion, de flux
- Qualité de service
- Réservation de bande passante

Congestion

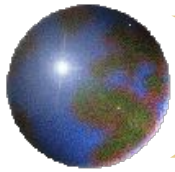


- Causes :
  - flot important arrivant à toutes les interfaces d ’entrée d ’un routeur, flot devant être transmis sur une seule interface de sortie
  - Place insuffisante dans la file d ’attente
  - Augmenter la taille des files d ’attente augmente le coût de traitement d ’un paquet



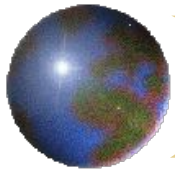
Comparaison contrôle de flux / contrôle de congestion

- Contrôle de congestion
  - doit s ’assurer que le sous-réseau est capable d ’acheminer le trafic
  - problème global qui tient compte du comportement des ordinateurs, routeurs, mécanismes de transmission des paquets qui ont pour effet d ’amoindrir la capacité de transport du sous-réseau…
- Contrôle de flux :
  - prise en compte du trafic point à point entre un émetteur et un récepteur. S ’assure qu ’un émetteur ne transmette pas les paquets de façon continue plus rapidement que le récepteur ne peut les absorber
  - retour d ’informations du récepteur vers l ’émetteur



Contrôle de congestion (1)

- Boucle ouverte
  - résolution du problème à la conception du protocole en s ’assurant que le problème n ’arrivera pas
  - augmentation du trafic, détruire les paquets (lesquels ?), ordonnancer à différents endroits.
- Boucle fermée : basé sur la rétro-action
  - surveiller le système pour détecter une congestion (quand et où)
  - envoyer des informations aux endroits où des actions doivent être prises
  - ajuster le comportement du système pour corriger le problème
- Informations : % de paquets détruits par manque de place dans les buffers, longueur des files d ’attente, nombre de paquets hors délai à retransmettre, temps moyen d ’acheminement des paquets...

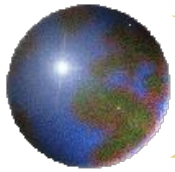


Contrôle de congestion (2)

- Problème : l ’envoi des informations d ’indication de congestion augmentent le trafic -> congestion
- Solutions :
  - utilisation d ’un champs particulier dans les paquets pour indiquer qu ’un routeur est en situation de congestion
  - transmission périodique de paquets de sondage
- Boucle fermée
  - boucle explicite : informations de congestion retransmises en retour du lieu de congestion vers la source
  - boucle implicite : la source détruit l ’existence du phénomène en effectuant des observations locales (exemple : temps A/R d ’un ACK)

Contrôle de congestion boucle ouverte dans les différentes couches

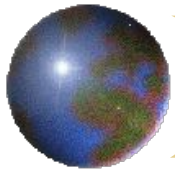
Couche	Voies d'action
Transport	Politique de retransmission Politique de masquage sélectif des anomalies Politique de l'accusé de réception Politique du contrôle de flux Définition des intervalles de temps hors délais
Réseau	Circuits virtuels ou data grammes Politique de mise en file d'attente et de distribution des paquets Politique de destruction des paquets Politique de routage Gestion de la durée de vie des paquets
Liaison de données	Politique de retransmission Politique de masquage sélectif des dérangements Politique de l'accusé de réception Politique du contrôle de flux



Contrôle de congestion boucle ouverte

- Traffic shaping (canalisation de trafic)
  - Algorithme du sceau percé (Turner 1986)
    - débit non régulé à l ’entrée : le sceau se remplit en fonction du débit d ’entrée
    - le débit de sortie est constant et est régulé par le sceau
    - variantes :
      - à compteur d ’octets
      - à jeton (compteur de jetons incrémenté par top d ’horloge)

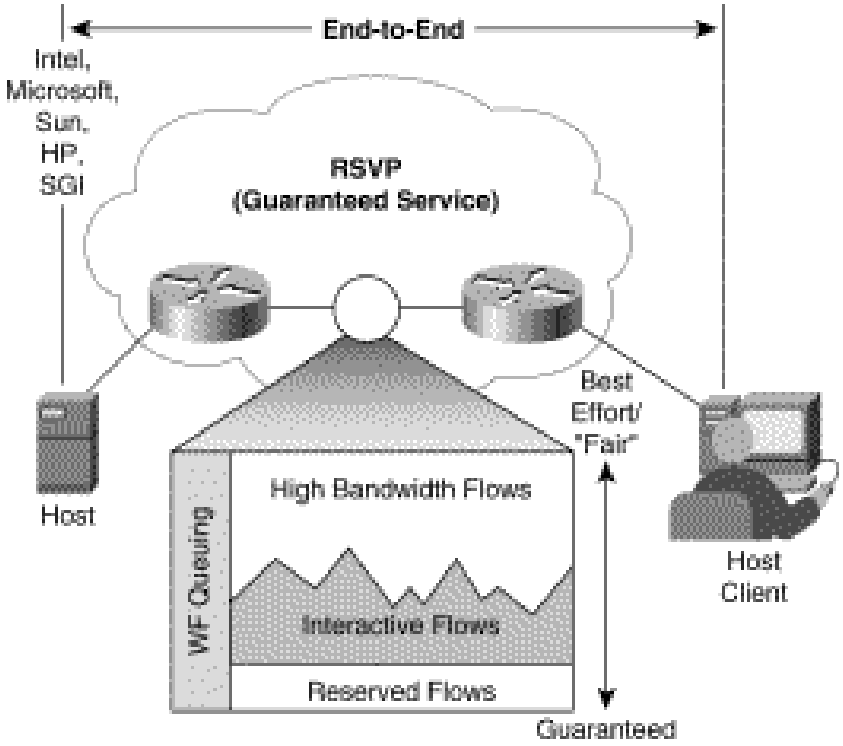
Dessins...

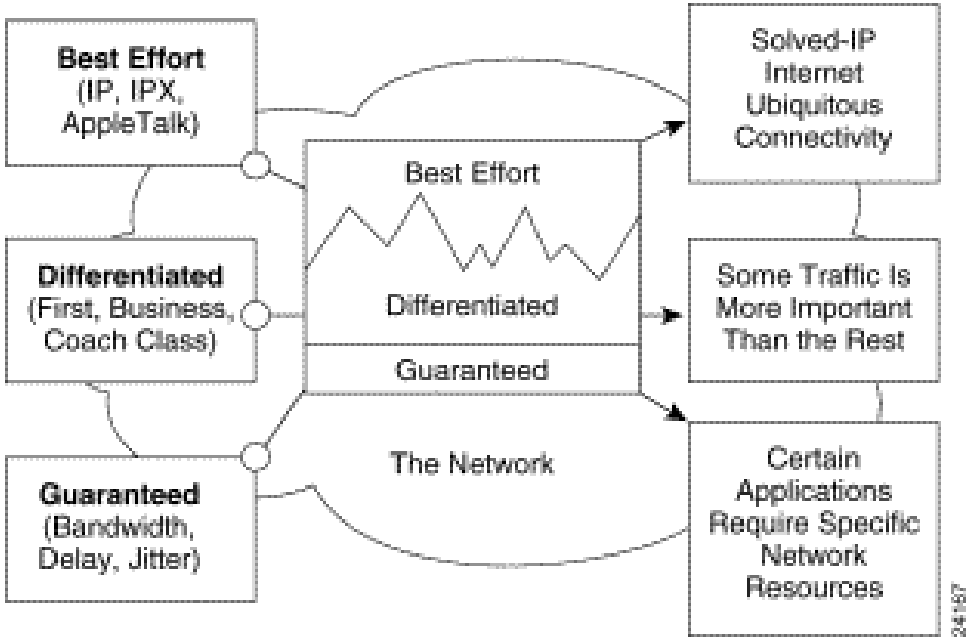


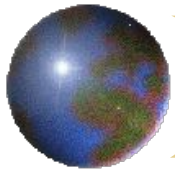
Qualité de service de bout en bout

- Classes de service (type ATM)
  - CBR (Constant bit rate) : débit constant garanti, téléphone
  - VBR (Variable bit rate) : realtime, non realtime, flux multimédia, gigue faible
  - ABR (Available bit rate) : débit disponible, web, sporadique
  - UBR (Unsiecified bit rate) : aucun débit spécifié, pas d ’infos de congestion, plus basse priorité.
- 3 niveaux de qualité de service
  - « best effort » : sans qualité de service, service basique, sans garantie
  - « differencied service » : qualité de service logicielle, une partie du trafic est mieux traitée que le reste (meilleure prise en charge, plus de bande passante en moyenne, moins de perte en moyenne)
  - « guaranteed service » : service garanti par une réservation de ressources réseau pour un trafic donné.

Qualité de service : comparaison



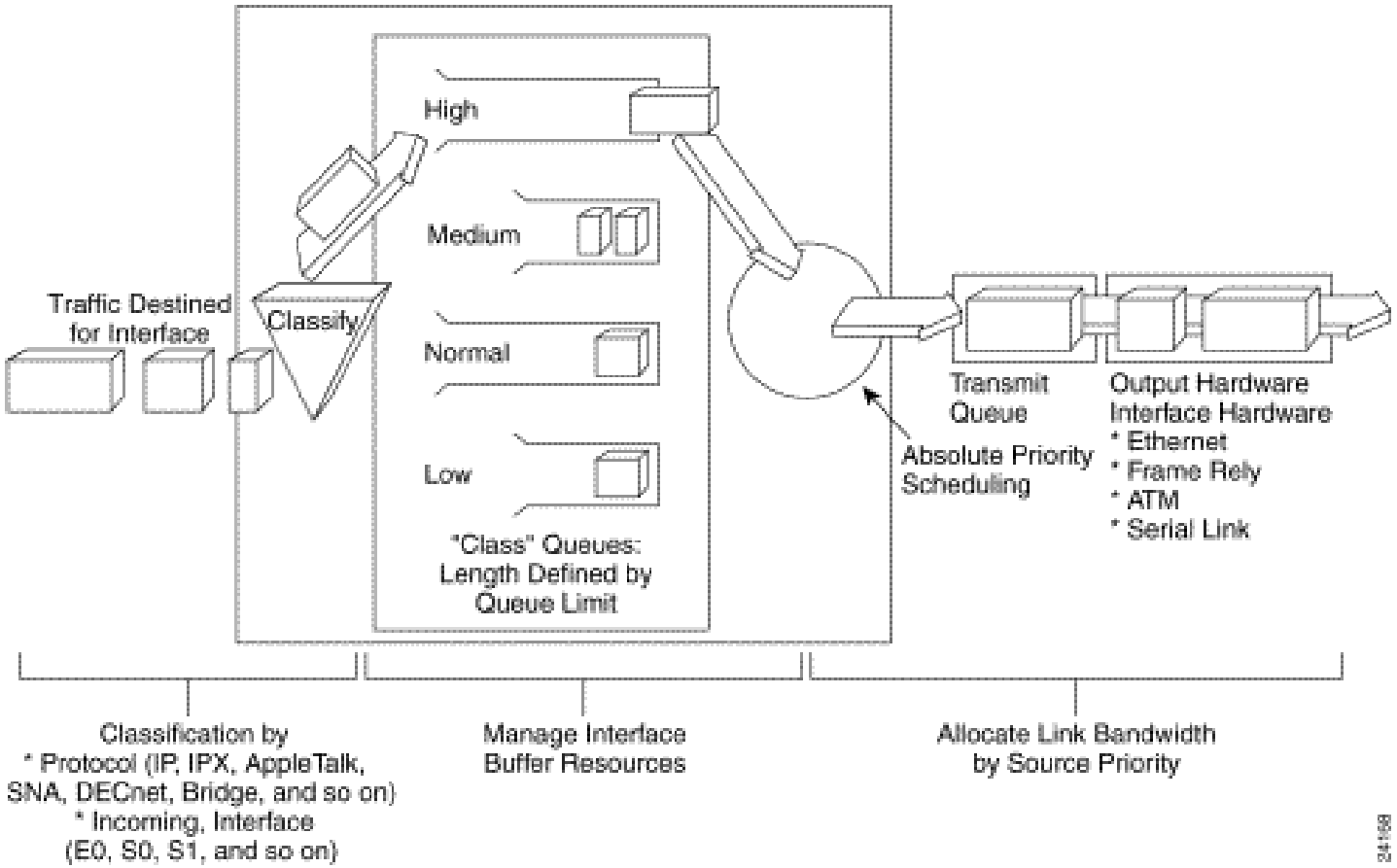


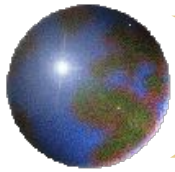


Qualité de services: algorithmes

- Chaque constructeur implémente ses algorithmes
- Exemple Cisco :
  - First-in, first-out (FIFO) queuing
  - Priority queuing (PQ)
  - Custom queuing (CQ)
  - Weighted fair queuing (WFQ)
- Identification du flux selon :
  - type de flux, niveau 4 ISO, exemple service TCP ou UDP
  - protocole (niveau 3) : adresse IP, subnet
  - niveau 2 : VLAN, adresse MAC (ethernet)
  - niveau 1 : port physique
- Mécanisme de priorité : mise en file d ’attente
  - nombre de files d ’attente
  - capacité des files
  - efficacité de l ’algorithme

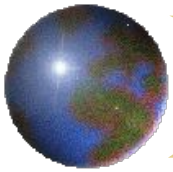
QoS : exemple d 'algorithme





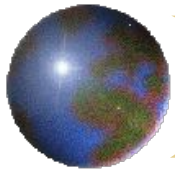
Réservation de bande passante

- Protocole RSVP (Resource reSerVation Protocol)
  - Autorise plusieurs émetteurs à transmettre vers plusieurs groupes de destinataires
    - chaque groupe de destinataires à une adresse de groupe
    - chaque destinataire peut appartenir à un ou plusieurs groupes
    - émission d ’un paquet : l ’émetteur place l ’adresse de groupe dans le paquet
    - l ’algorithme de routage multi-destinataire (ne fait par partie du protocole RSVP) construit un arbre recouvrant sur tous les membres du groupe



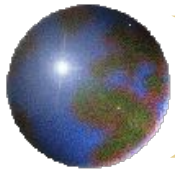
IP : résumé fonctionnalités

- Service sans connexion
- adressage
- fragmentation / ré assemblage
- support de datagrammes de taille variable
- service « best effort » :
  - délai de transmission, perte de paquets, corruption : les protocoles de plus haut niveau doivent gérer ces problèmes
- Fournit seulement des services de contrôle d'envoi et de distribution de messages par le protocole ICMP (
  Internet Control Message Protocol)



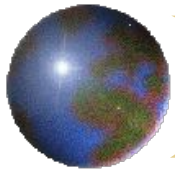
IP ne fournit pas

- Un service de bout en bout et de contrôle de flux (service fournit par TCP et les couches supérieures)
- le séquencement (TCP)
- Le rapport d ’erreur (fait par ICMP)
- la gestion des tables de routage (fait par RIP, OSPF, BGP…)
- la gestion des connexions (c ’est un protocole sans connexion, fait par TCP)
- La résolution d ’adresse ou de nom (fait par ARP, RARP, DNS)
- la configuration (pour IPv4, fait par DHCP ou BOOTP)
- le multicast (fait par IGMP, DVMRP, PIM, MBGP)



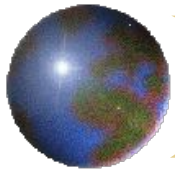
IP : avantages

- Indépendance vis à vis des technologies réseau et des constructeurs
  - testé et fonctionnel dans réseaux LAN, WAN, RTC, haut débit...
  - micro-ordinateurs, serveurs, équipements réseau
- Connectivité universelle : chaque nœud a une adresse unique
- bien documenté, documentation accessible par tous (RFC)
- protocoles simples mais efficace



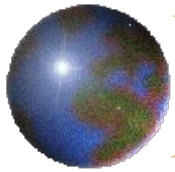
IP : inconvénients

- majorité des standards édites aux USA
- lobying des grosses sociétés US sur les protocoles
- Sécurité n’a pas été prise en compte à la conception du protocole
  - Il est possible de générer des paquets invalides, fausses adresses IP source
  - filtrage du protocole FTP est un cauchemar (mode actif)



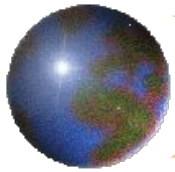
Économie d ’adresses IP

- Plusieurs mécanismes permettent d ’économiser des adresses
  - Traduction d ’adresse privées (nombre illimité) en adresses publiques (routées, et limitées) : NAT, PAT
  - d ’affectation d ’adresse IP dynamique : DHCP : Utiliser une adresse IP uniquement en cas de besoin
  - PPP : mécanisme d ’affectation d ’IP dans la définition du protocole
    - à la connexion d ’un modem sur un serveur de modem, après authentification, affectation d ’une adresse IP en fonction de plusieurs critères :
      - login/mot de passe
      - modem appelé...



Translation d ’adresse

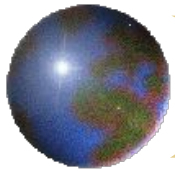
- Translation des adresses IP de l’en-tête
- Re-calcul et vérification du checksum
- Re-calcul et modification du checksum TCP
- Transparence au niveau application
- Conversion de certains paquets
  - FTP :
    - commandes PORT et PASV : ASCII
    - checksum et numérotation des séquences
    - table spécifique : ACK, SEQ#
  - ICMP, SNMP, SMTP ...



Translation d ’adresse : conséquences

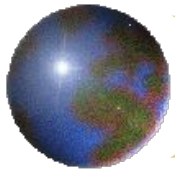
- La translation introduira plus ou moins de délais pendant sa commutation .
- NAT fait que certaines applications spécifiques qui utilisent les adresses IP fonctionneront difficilement ou seront impossibles à utiliser.
- NAT cache l’identité “réelle” des hosts.
- Tout paquet qui doit être translaté doit passer par le routeur NAT.
- L’adressage interne ne doit pas être diffusé vers l’extérieur
- L’adressage externe peut être distribué vers l’intérieur

- Problèmes :
  - Multicast
  - Routage
  - protocoles véhiculant les adresses IP dans leurs champs de control, exemple : FTP.



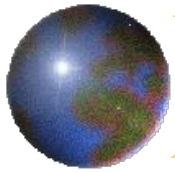
NAT, outil de sécurité

- Limitation des SYN Flooding, spoofing ...
- inconvénient mode statique
- pas (peu) de log, nécessité d ’outils supplémentaires
- perte de la traçabilité de bout-en-bout
- problématique de l’encryption
- paquets applicatifs
- en-tête TCP/IP



Évolution IPV6 (sommaire)

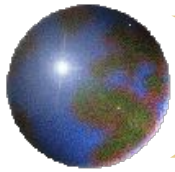
- étend la fonction d'adressage et de routage
- tend à résoudre les problèmes qui vont devenir critiques (applications temps réel, multipoint, sécurité...)
  - permet le contrôle de flux
- cherche à faciliter la migration de IPX et OSI vers IP (intégration des adresses des autres protocoles)
- Optimisation du protocole pour haut débits
- Ajoute des nouvelles fonctionnalités
  - mobilité
  - autoconfiguration
  - sécurité



IPv6 : Caractéristiques (1)

- Adresse plus longue : 128 bits (16 octets)
  - adressage de 340 x 10 e36 équipements
  - adressage hiérarchique
  - une partie peut-être l'adresse MAC (IEEE802)
    => autoconfiguration
- 3 types d'adresses :
  - Unicast
  - Multicast
  - Anycast

  plus d'adresse de broadcast



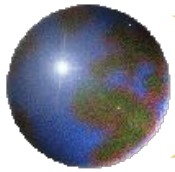
IPv6 : Caractéristiques (2)

- En-tête simplifié
  - nombre de champs réduit de moitié

  => augmente l'efficacité de commutation des équipements de routage

- Extension de l'en-tête pour les options
  - Les options IPv6 sont placées dans des en-têtes séparés, intercalés entre l'en-tête IPv6 et l'en-tête de la couche transport

    => introduction aisée de nouvelles fonctionnalités
  - la longueur des options n'est plus limitée à 40 octets

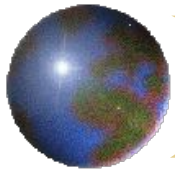


IPv6 : nouvelles fonctionnalités (1)

- Autoconfiguration : "*plug and play* » (vu plus loin)
  - Gestion de la mobilité
  - Renumérotation facile si changement de prestataire
  - Serveurs d'adresses (DHCP : *Dynamic Host Configuration Protocol)*
  - et SAA : *Stateless Address Autoconfiguration (RFC 1971)*

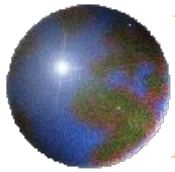
- Multipoint (*Multicast*) inclus de base
  - pour les routeurs et les clients
  - "scope" = meilleur routage des paquets multicast

  => plus besoin de Mbone ni de mrouted



IPv6 : nouvelles fonctionnalités (2)

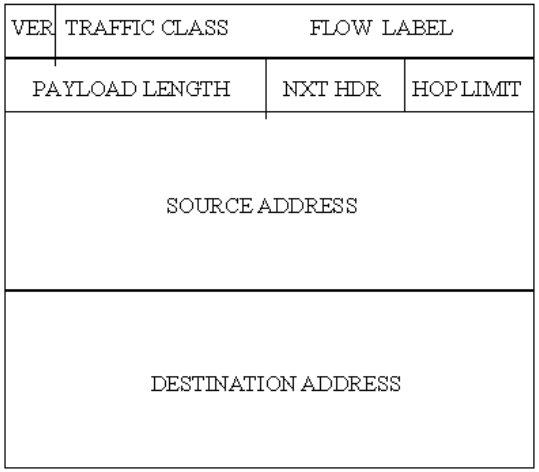
- "Marquage" des flux particuliers : (*Flow Label*)
  - applications temps réel, Qualité de Service (QoS)
  - Priorité du trafic de contrôle
- Sécurité :
  - authentification et intégrité des données
  - *en option* : confidentialité
- Routage à partir de la source
  - Source Demand Routing Protocol



paquet IPv4 / IPv6

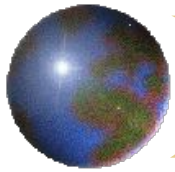
- Champs supprimés
  - Header Length (IHL)
  - Header CS
  - ID, Flags et Fragment Offset (FO) : pas de fragmentation dans IPv6
  - Total control : les couches supérieures contrôlent et les réseaux actuels sont fiables
- Champs renommés (avec des différences)
  - ToS --> Flow Label
  - Total Length (TL) --> Payload Length
  - TTL --> Hop Limit
  - Protocol --> Next header (mêmes valeurs que dans IPv4)
- Adresses : 32 --> 128 bits (4 --> 16 octets)
- Alignement 32 --> 64 bits

Paquet IPv6



- Vers : Version Number (= 6)
- Flow label : marquage des paquets "spéciaux"
- Payload length : longueur du paquet après en-tête (en octets).
  - autorise des paquets > 64 Koctets => Payload length = 0

  + longueur du paquet dans l'en-tête de l'option "Hop-by-hop"
- Next header : indique le type d'entête suivant immédiatement l'entête IPv6
  - On utilise les mêmes valeurs que dans le champ "Protocol" de IPv4 pour référencer les protocoles de niveau 4 **(TCP=6 , UDP=17, ICMP=1)**
- Hop limit :
  - -1 chaque fois que le paquet est commuté par un équipement
  - si hop_limit = 0 => le paquet est détruit.
  - permet de réduire l'effet des boucles de routage.
- Source address : @ de l'émetteur du paquet
- Destination address : Une adresse de destination

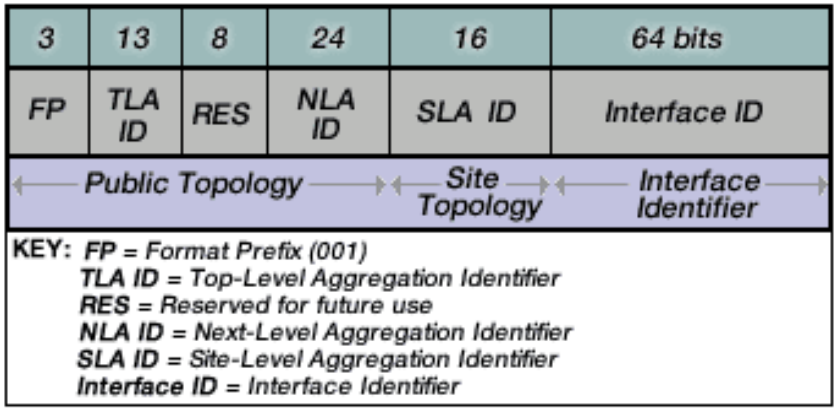


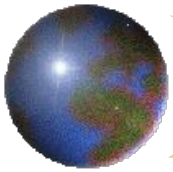
Paquet IPv6 : en têtes optionnelles

- Hop-by-Hop Header :
  - transport d'information qui doit être examinée sur chaque noeud du chemin suivi par le datagramme IP.
- End-to-end Header :
  - transport d'information qui n'est à examiner que par le destinataire du datagramme.
- Routing Header :
  - routage à partir de la source
  - liste un ou plusieurs noeuds intermédiaires "à visiter" au cours de l'acheminement du datagramme
  - "Reverse bit" :
    - si = 1 => utiliser l'information de routage pour le retour
    - sinon => résoudre le routage à partir de l'extrémité destinataire
- Fragment header :
  - envoi de paquets plus long que le MTU
    - Maximum Transmission Unit : 512 -> 1500 octets
    - minimum : MTU = 576 octets.
  - *nota* : dans IPv6, la fragmentation n'est réalisée que par la source.
- Authentication Header :
  - authentification et intégrité des données
- Privacy Header :
  - encryptage des données à protéger
    - datagramme TCP/UDP ou datagramme IPv6 entier, à la demande

IPv6 : allocation des adresses

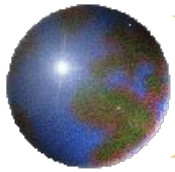
- Allocation initiale des adresses : # 15%
  - @ Unicast des prestataires de connectivité (*ISP*)
  - ISO NSAP
  - IPX
  - @ Multicast
  - @ réservées pour allocation géographique
  - @ à usage local
  - @ compatibles IPv4
- Pas de réallocation pour les adresses IPv4 actuelles
- Réserve pour la croissance : # 85%





IPv6 : espace d ’adressage

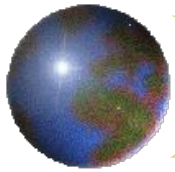
- 2^{138} adresses ($3x10^{38}$) : Avogadro !
- $7x10^{23}$ adresses par m^2 !
- Calcul de C. Huitema :
  - Pire des scénarios : >1000 @ / m^2
  - scénario prévisible : > 10^{15} @ / m^2



IPv6 : représentation des adresses

- Format de Base (16 octets):
  - Adresse IPv6 Globale :
    - FEDC:BA98:7654:3210:FEDC:BA98:7654:3210
- Format compressé :
    - FF01:0:0:0:0:0:0:43 => FF01::43
  - Adresse Link Local :
    - FE80::*IID* *IID=@ IEEE-802*
  - Adresse compatible IPv4 :
    - 0:0:0:0:0:0:134.157.4.16 => ::134.157.4.16
  - Adresse de Loopback :
    - 0:0:0:0:0:0:0:1 => ::1
  - Adresse non specifiée :
    - 0:0:0:0:0:0:0:0 => ::

      Ne peut jamais être adresse destination



IPv6 : préfixes

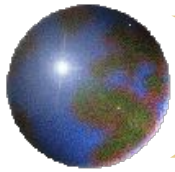
- Notion issue de CIDR
- On les note sous la forme :

  Adresse IPv6 / longueur du préfixe

  - Exemples :

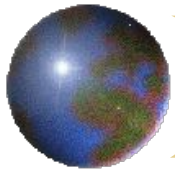
    5F00::/8

    5F06:B500::/32



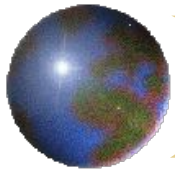
IPv6 : nouveaux protocoles

- Neighbor Discovery (ND)
  - *RFC 1970*
- Autoconfiguration
  - Stateless Address Autoconfiguration *(RFC 1971)*
  - DHCPv6 :Dynamic Host Configuration Protocol Path MTU discovery (pMTU)
  - *RFC 1981*
- Resource ReSerVation Protocol (RSVP)



IPv6 : Autoconfiguration à mémoire d'état (stateful)

- DHCPv6 : *Dynamic Host Configuration Protocol*
  - s'appuie sur le protocole BOOTP (RFC 951)
  - et la version IPv4 de DHCP (RFC 1541)
- le serveur
  - mémorise l'état du client
  - fournit les @IPv6 et des paramètres de configuration du client
- le client
  - émet des requêtes et des “acquittements” (selon le protocole DHCP)

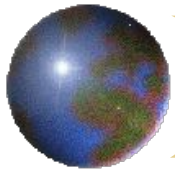


IPv6 : Autoconfiguration sans état (stateless)

Construire une adresse globale à partir de :

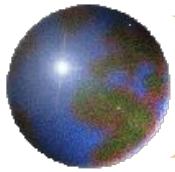
- l ’adresse MAC de la machine
- des annonces de préfixes faites sur le même câble *(par les routeurs)*

=> Ce mécanisme ne s'applique pas aux routeurs



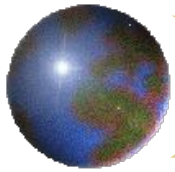
IPv6 : principes de routage

- Topologie :
  - Régions IPv4 "pures"
  - Régions IPv6 "pures"
  - Régions IPv4 et IPv6
- Méthodes :
  - double pile logicielle => routage des paquets IPv4 et IPv6
  - tunnels manuels
  - tunnels automatiques
- Mécanismes :
  - encapsulation des paquets IPv6 dans des paquets IPv4 (IPv4 -> IPv6 ?)
  - traduction des en-têtes IPv6 <--> IPv4 à la frontière des régions



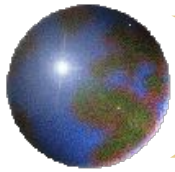
IPv6 : Sécurité

- RFC 1825 à 1829 et RFC 1851, 1852
  - IPv6 et IPv4
  - Authentification
  - Intégrité
  - Confidentialité
- Indépendant des algorithmes de chiffrement
  - champ SAID : Security Association Identifier :
    Type de clé, durée de vie, algorithme, ...
- Administration des clés : séparée
- Les fonctions de sécurité sont optionnelles
- elles n’affectent pas les autres utilisateurs



IPv6 : Plan de transition

- Philosophie générale :
  - Compatibilité de IPv6 avec IPv4 (postes de travail et routeurs)
    - conserver les @IPv4 déjà allouées
  - Facile à installer et faible coût initial
  - Acquérir de l'expérience le plus tôt possible
- Objectifs :
  - Évolution progressive des machines et des routeurs : PAS de JOUR J !!!
  - Peu ou pas de dépendances (phases multiples, état des autres équipements...)
  - Terminer la transition avant l'épuisement des @ IPv4



Bibliographie

- http://www.urec.cnrs.fr/cours