

# Réseaux Locaux Ethernet

Université de Marne-la-Vallée

duris@univ-mlv.fr

## Motivation

- Les réseaux locaux **LAN (Local Area Network)** sont au coeur même du réseau
  - la plupart des machines sont directement sur des LAN
  - les LAN supportent les protocoles des couches supérieures
- Ethernet est représentatif des évolutions des technologies de réseaux LAN (débits, supports)
- C'est aujourd'hui la technologie la plus largement utilisée
  - jusqu'en 1990/1995, d'autres concurrents coexistaient
  - pratiquement tous relégués au rang de techno. secondaire

## Plan

- Origines, besoins, historique des LAN
  - 30 ans d'histoire
- Principes basiques de fonctionnement et normalisation
  - Trame, protocole, supports, signalisation, composants
  - Les standards de l'IEEE et les règles de configuration
- Évolutions en terme de débit et de support
  - De 3 Mégabit/s sur coaxial en 1973 à plusieurs Gigabit/s sur fibre optique aujourd'hui
- Élargissement des compétences d'Ethernet
  - LAN virtuels (VLAN), Ethernet sans fil (WiFi, IEEE802.11)

## Bibliographie

- **Réseaux**. 4<sup>ème</sup> éd. Andrew Tanenbaum, Pearson Education 2003.
- **Analyse structurée des réseaux**. Kurose & Ross, Pearson Education 2003.
- **Java et Internet** Concepts et programmation. Roussel et Duris, Vuibert 2002.
- **High Speed LAN Technology Handbook**. Chowdhury. Springer 2000.
- **Les réseaux**. 3<sup>ème</sup> éd. Guy Pujolle, Eyrolles 2000.
- **Interconnections**. 2<sup>nd</sup>ed. Radia Perlman. Addison Wesley 2000.
- **Réseaux haut débits**. Tome 1. 2<sup>ème</sup> éd. D Kofman et M. Gagnaire, Dunod 1999.
- **Guide Pratique des Réseaux Ethernet**. Charles Spurgeon, Vuibert 1998.
- **Gigabit Ethernet**. Kadambi, Krayford & Kalkunte. Prentice Hall 1998.
- **Les réseaux locaux virtuels**. Gilbert Held, InterEditions 1998.
- **Les réseaux locaux commutés et ATM**. Alexis Ferréro, InterEditions 1998.
- **Gigabit Networking**. Craig Partridge. Addison Wesley 1994.

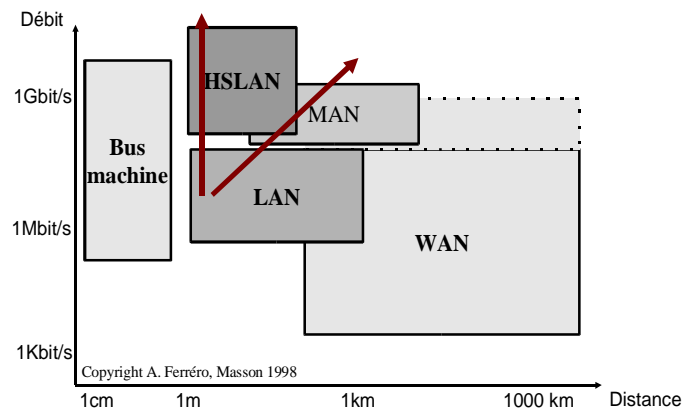
## Besoins des LAN

- Relier différentes machines sur le même médium, à l'échelle d'un bâtiment, voire d'un site.
  - De l'ordre de la centaine de connexions
  - Débits de l'ordre de la dizaine de mégabit/seconde
  - Couverture géographique de l'ordre du kilomètre
- Le terme LAN (*Local Area Network*) regroupe les termes francophones suivants:
  - RLE: Réseaux Locaux d'Entreprise
  - RLI: Réseaux Locaux Industriels

## Les autres dénominations de réseaux

- MAN *Metropolitan Area Network*
  - Augmentation de couverture ( $\approx 100$  km) et du débit maximal ( $\approx 100$  Mbit/s). Ex. FDDI, DQDB, ATM.
- HSLAN *High Speed LAN*
  - Quelques dizaines d'accès, sur quelques dizaines de mètres, de l'ordre de 100 Mbit/s à plusieurs Gigabits/s. (Fibre Channel, HiPPI, Fast Ethernet/Gigabit, ATM)
- WAN *Wide Area Network*
  - Longues distances (mondial). Lignes louées, X.25 Frame Relay, RNIS, ATM.

## Les positions relatives des technologies



## Protocoles d'accès multiples

- Principe: relier différentes machines sur un même canal
  - Soit on ne relie que 2 machines entre elles: liaison **point à point** (PPP *Point to Point Protocol*, HDLC *High-level Data Link Control*)
  - Soit on doit relier plus de 2 machines: liaison à **diffusion**
    - Protocoles à partage de canal
      - Multiplexage par répartition dans le temps (TDM), en fréquence (FDM) ou par répartition de code (CDMA)
    - Protocoles à accès aléatoire
      - ALOHA à allocation temporelle (discrétisé) ou ALOHA pur, mais aussi avec détection de porteuse [et de collision]
    - Protocoles à partage de ressource
      - Sondage ou invitation à transmettre, passage de jeton

## Méthode d'accès aléatoire: Aloha

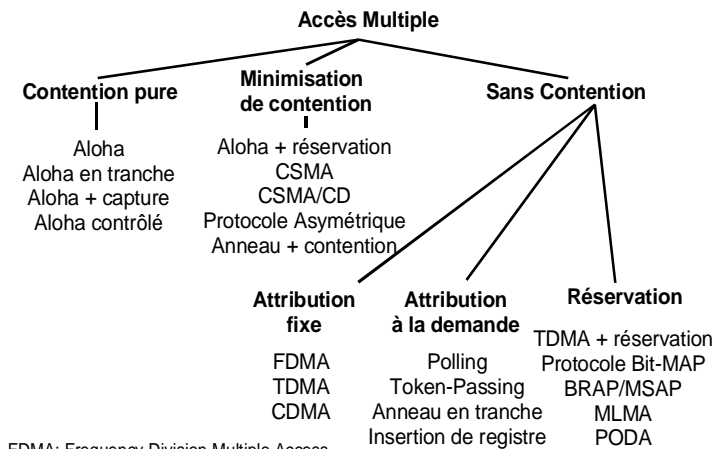
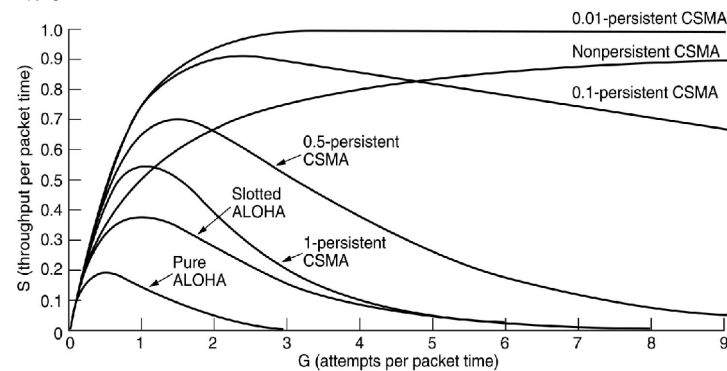
- 1970. Norman Abramson (Université de Hawaï)
  - ALOHA discrétisé (*slotted*) ou à allocation temporelle (72):
    - Toutes les trames ont une taille fixe
    - Le temps est divisé en intervalles permettant la transmission d'une trame
    - Les expéditeurs débutent les émissions au début d'un intervalle de temps (ils sont synchrones)
    - Les expéditeurs sont informés de toute collision avant la fin de l'intervalle de temps auquel elle se produit
    - Si il y a une collision, elle est détectée avant la fin de l'intervalle de temps en cours: l'expéditeur la retransmet au cours des intervalles de temps suivant avec une probabilité  $p$ , jusqu'à un succès.

## ALOHA et compagnie

- ALOHA discrétisé (*slotted*)
  - Possibilité d'utiliser la totalité du débit si on est seul
  - Requier la synchronisation entre machines
- ALOHA pur (version initiale)
  - Pas de synchronisation ni de découpage en intervalle
- CSMA (*Carrier Sense Multiple Access*)
  - Détection de porteuse: ne pas émettre si une autre station est déjà en train d'émettre
    - Persistant (attente après collision et émission dès que libre)
    - Non-persistant (écoute, attente, écoute... jusqu'à libre)
- CSMA/CD (Collision Detection)
  - Interrompt l'émission dès qu'une collision est détectée

## Comparaison CSMA et Aloha

Copyright A. Tanenbaum, Pearson Education, 2003



FDMA: Frequency Division Multiple Access  
 TDMA: Time Division Multiple Access  
 CDMA: Code Division Multiple Access

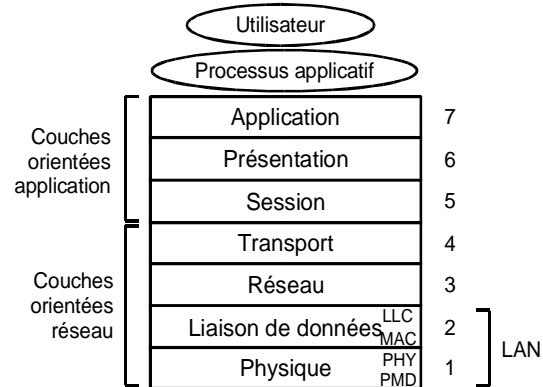
## Caractéristiques principales des LAN

- Débit nominal: quantité d'information/seconde
- Topologie logique de connexion: bus, étoile, anneau...
- Méthode de partage des accès: droit de parole
- Standardisation (état des normes)
- Format des trames (unité de donnée)
- Distance de couverture (longueur max. des liens)
- Nombre d'accès (limite sup. des connexions)
- Médias et connectique (câbles & topologie)

## Les différentes notions de débit

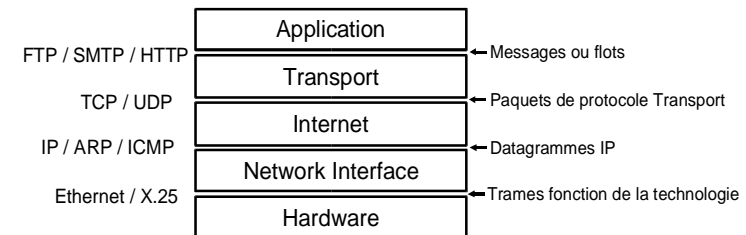
- La notions de « débit » est contextuelle:
  - Utilisateur: débit de données informatiques (octets):  
-> octets par seconde **Ko/s** (*bytes per second Kb/s*).
  - Codage de l'information: débit d'information unitaire transmis sur un médium (bits):  
-> bits transmis par seconde **Kbit/s**
  - Signalisation physique pour représenter l'information  
-> signaux distincts par secondes = **Bauds**
    - Rapidité de modulation du signal  
i.e. nombre de changements d'état par seconde

## Open System Interconnection (OSI)



## Modèle TCP/IP

- En parallèle du modèle OSI (DoD/ DARPA)



- Atouts: Unix, IETF et IRTF (RFC)

- *Internet Engineering Task Force et Internet Research Task Force*

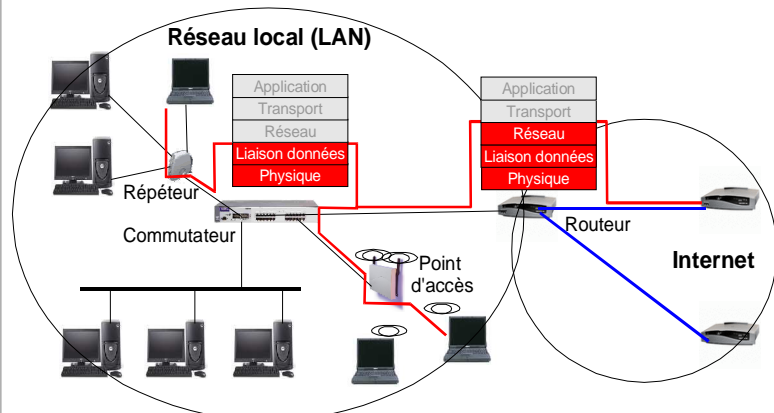
## Couche physique

- Transmission du signal
  - Codage, détection de présence de signal, gestion des horloges, synchronisation
  - Valeurs relatives au médium, paramètres électriques ou optiques et mécaniques des liaisons / connectique.
- PMD (partie basse) *Physical Medium Dependent*
  - Dédicée à un médium ou à un débit spécifique
- PHY (partie haute)
  - Gère les fonctionnalités communes

## Couche liaison de données

- Transfert des données sur le support physique
  - Formatage, contrôle d'erreurs et gestion de flux
- MAC (partie basse) *Medium Access Control*
  - Méthode d'accès et format des trames, adresses
  - Tous les LAN sont orientés et sans connexion
- LLC (partie haute) *Logical Link Control*
  - Interface d'abstraction de MAC sous-jacente
    - Type 1. Sans acquittement ni connexion (IP, IPX)
    - Type 2. Orienté connexion (NetBEUI, MS-LAN Manager)
    - Type 3. Acquittement sans connexion

## Couches OSI et LAN



## Historique d'Ethernet

- 1973 : naissance d'Ethernet
  - Bob Metcalfe et David Boggs (Xerox PARC)
  - Initialement 2,94 Mbit/s, câble coaxial, ≤256 ordinateurs
- 1979 : création du Consortium DIX (DEC, Intel, Xerox)
- Puis création de l'IEEE Projet 802 (*Institute of Electrical and Electronic Engineers*)
  - 802.1 : High Level Internetwork Interface (HLI)
  - 802.2 : Logical Link Control (LLC)
  - Et un groupe DLMAC...

## Data Link & Medium Access Control

- Différents modes d'accès au support physique :
  - **802.3** : Carrier Sense Multiple Access with Collision Detection (**CSMA/CD**) : Ethernet
    - > Initialement dirigé par DEC, Intel et Xerox
    - Rudimentaire, mais simple à administrer
  - **802.4** : **Token Bus**
    - > Sponsorisé par Burroughs, Concord Data Systems, Honeywell et Western Digital
  - **802.5** : **Token Ring**
    - > Exclusivité d'IBM

## Standardisations

- 1982. IEEE et DIX convergent
  - Standards IEEE 802.3 et Ethernet V.2 (ou Ethernet II)
  - Des différences mineures subsistent
  - Autres organismes de normalisation...
    - NIST (National Institute of Standards and Technology)
    - ECMA (European Computer Manufacturers Association)
    - ANSI (American National Standards Institute)
    - ISO (International Standardization Organization)
- 1990. Standardisation ISO/IEC 8802-3

## Normes principales IEEE 802

- 802.1 High Level Interface, Network Management, Bridging, Glossary
- 802.2 Logical Link Control
- **802.3 CSMA/CD Ethernet**
- 802.4 Token Bus
- 802.5 Token Ring (LAN IBM)
- 802.6 Metropolitan Area Network (DQDB : Double Queue Dual Bus)
- 802.7 Broadband LAN Technical Advisory Group
- 802.8 Fiber Optic Technical Advisory Group
- 802.9 Integrated Service LAN (IsoEthernet), pour isochrone (temps réel)
- 802.10 LAN Security (SILS : Standard for Interoperable LAN Security)
- **802.11 Wireless LAN**
- 802.12 Demand Priority LAN (100VG - AnyLAN)
- 802.14 Cable TV MAN
- **802.15 Wireless Personal Area Network (WPAN)**, bluetooth
- **802.16 Fixed Broadband Wireless Access** (sans fil large bande)

## L'accès au réseau

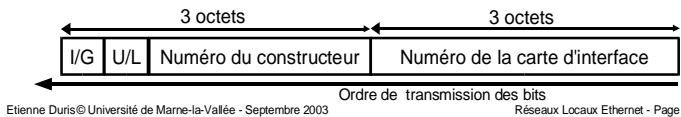
- Carte d'Interface Réseau (NIC)
  - Identification d'une machine sur le support
  - Adresse **physique** (MAC)
  - Permet d'émettre et de recevoir des trames
    - format dépendant du protocole
    - nous nous intéresserons à Ethernet et/ou IEEE 802.3
  - Intermédiaire pour l'adressage **logique**
    - Par exemple de type IP par les mécanismes ARP ou DHCP

## Adresses MAC

- Identifie chaque carte d'interface sur le réseau local
- Nécessairement unique (pour un réseau donné)
  - Partie dépendant du constructeur et numéro de série
- Adressage standardisé IEEE 802
  - CSMA/CD, Token Bus, Token Ring, DQDB
  - Également pour FDDI et ATM
- Longueur 6 octets, représentés en hexadécimal
  - Classiquement sous l'une des formes
    - **00:0B:DB:16:E7:8A** ou **00:0B-DB-16-E7-8A**

## Adresses MAC (Ethernet)

- Les octets sont transmis le bit de poids faible en premier (*LSB, Least Significant Bit first*)
  - Premier bit transmis: I/G : Individuel (0) / Groupe (1)
  - Second bit transmis: U/L : adresse administrée de manière Universelle (0) / Locale (1)
  - Les 24 premiers bits constituent l'OUI (*Organizationally Unique Identifier*). Ex: **00:00:0C** (Cisco) **00:C0:4F** (DELL)
  - Les 24 bits restant sont des numéros de série



## Exemples d'adresse MAC

- Si on dispose de l'OUI **AC-DE-48**, il est possible de construire une carte ayant l'adresse suivante:
  - **AC-DE-48-00-00-80**, soit **1010 1100-1101 1110-...**
  - Ce qui correspond, lors d'une transmission LSB, à la suite:
- premier octet | second octet | ...
 

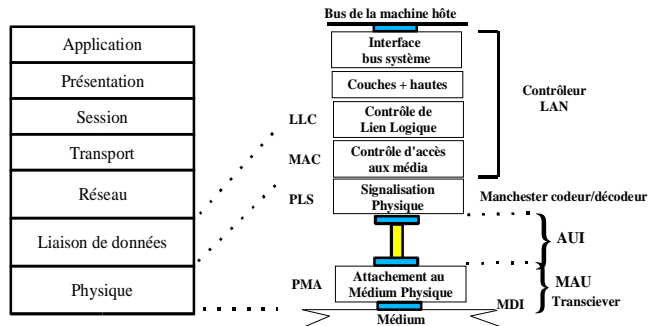
lsb	msb																		
0011	0101	0111	1011	0001	0010	0000	0000	0000	0000	0000	0000	0000	0001						
C	A	E	D	8	4	0	0	0	0	0	0	0	8						
- Ou l'adresse de groupe **AD-DE-48-00-00-80**

## Le détail d'Ethernet à 10 Mégabit/s

- Éléments de base d'Ethernet
  - La trame
    - Formats standards Ethernet II ou IEEE 802.3
  - Le protocole d'accès au média
    - CSMA/CD : détection de porteuse et des collisions
  - Les composants de signalisation
    - Interface Ethernet, transceiver (+ câble), répéteur
  - Le médium physique
    - Coaxial, paire torsadée, fibre optique...

## OSI versus IEEE

- Correspondance entre les 7 couches OSI et 802.3



## La trame Ethernet

- Trame Ethernet II (selon le consortium DIX)

8 octets	6 octets	6 octets	2 octets	46 à 1500 octets	4 octets
Préambule	Adresse MAC destination	Adresse MAC source	Type	Données / Bourrage	Séquence Contrôle de Trame

- Trame standardisée IEEE 802.3

7 octets	1 octet	6 octets	6 octets	2 octets	46 à 1500 octets	4 octets
Préambule	Délim. Début Trame	Adresse MAC destination	Adresse MAC source	Long./Type	LLC / Données / Bourrage	Séquence Contrôle de Trame

## En-tête de trame

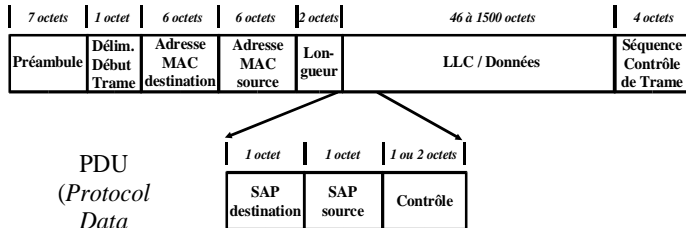
- Préambule: 8 octets ou 64 bits pour synchroniser
  - 1010101010 ... .. 10101011
  - Pratiquement, DIX et 802.3 ne font pas de différence
  - Crée une fréquence de 5 MHz reconnue dans le PLS
- Adresses destination et source (6 octets chacune)
  - Adresses physiques ou adresses MAC
    - Attribuées par constructeurs de cartes d'interface Ethernet
    - 24 bits propres au constructeur (OUI, assigné par IEEE)
    - Adresses multicast, dont broadcast (**FF-FF-FF-FF-FF-FF**)
    - 2 premiers bits spéciaux: @ globale/locale et indiv/groupe

## Type ou Longueur

- Champ Type d'Ethernet II (16 bits = 2 octets)
  - Indique le protocole de haut niveau transporté
    - 0x0800 trame IP, 0x0805 X.25, 0x0806 ARP...
- Champ Longueur/Type de IEEE 802.3 (16 bits)
  - Indique le nombre d'octets dans le champ de données
    - Éventuellement moins de 46 octets (caractères de bourrage)
    - Si  $\leq 1518$ , c'est la taille des données en octets
    - Si  $\geq 1536$  (0x600), c'est utilisé comme le type des données
- Les trames Ethernet et 802.3 sont « compatibles »

## Données LLC et SAP

- SAP (*Service Access Point*)
  - Utilisés en cas de besoin de service de fiabilité, par ex
    - Remplace alors le champ type



## PDU de LLC 802.2

- Un PDU (*Protocol Data Unit*) de LLC contient:
  - Un point d'accès (SAP) au service destination;
  - Un point d'accès au service source
    - Identifiants du protocole de + haut niveau auquel les données sont destinées (idem au champ Type d'Ethernet II)
    - Ce sont des sortes de « ports » ou boîtes aux lettres
    - Leur valeur est choisie par l'IEEE
  - Un champ de contrôle: « information » et « supervision » sur 2 octets ou « non numéroté » sur 1 octet
  - Variation: Ethernet SNAP (multi.prot), NetWare Ethernet (IPX/SPX)

## Données

- Ethernet (DIX)
  - 46 octets ≤ données dans une trame ≤ 1500 octets
  - Cela incombe au logiciel de réseau
- IEEE 802.3
  - 0 octets ≤ données dans une trame ≤ 1500 octets
    - Nombre spécifié dans le champ longueur + pad éventuel
  - Si longueur utilisée, on peut décrire un type LLC
    - Il est fourni par le protocole de la couche LLC (802.2)
    - Permet de faire du démultiplexage (DSAP/SSAP)

## Contrôle de trame

- FCS (Frame Contrôle Sequence)
  - 4 octets (32 bits) pour le Cyclic Redundancy Code
    - CRC calculé sur les champs dest, src, type/lgr et données
    - Calculé/inséré à l'émission et calculé/vérifié à la réception
- Détection de fin de trame
  - Après toutes ces valeurs, silence sur la trame
    - Possibilité de bits de « bavure » (bits additionnels post FCS)
    - Le récepteur tronque à l'octet complet le plus proche

## Codes détecteurs / correcteurs

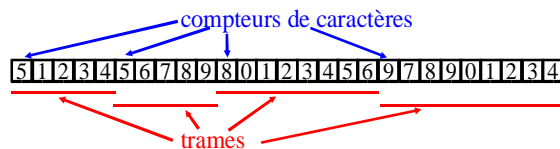
- Un code *détecteur* d'erreur, tel que le CRC utilisé pour Ethernet, permet d'assurer qu'une trame est reçue sans erreur de transmission
  - La probabilité d'une erreur bit non détectée dans le cas d'Ethernet est de 1 pour 4,3 milliards
- Un code *correcteur* d'erreur, tel que celui de Hamming, permet non seulement de détecter une erreur, mais de plus de savoir où elle s'est produite et de donner les moyens de la corriger.
- Faisons un aparté sur les codes...

## Notion de Trame

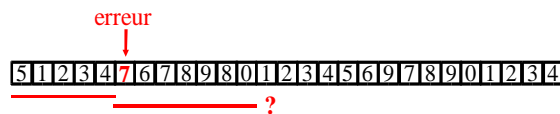
- Objectif: découper le « train » de bits (flot de données) à émettre en morceaux délimités, i.e. en **trames**.
- Différentes manières d'effectuer ce découpage:
  - Compter les caractères;
  - Utiliser des caractères de début et de fin (caractères de transparence);
  - Utiliser des fanions de début et de fin (bits de transparence);
  - Violier le codage utilisé dans la couche physique.

## Comptage de caractères

- Chaîne à transmettre



- En cas d'erreur: perte de synchronisation



## Caractères de transparence

- Données envoyées par la Couche Réseau (émetteur)

A B C D E F G  
DLE STX C DLE E DLE ETX

- Ajout par la Couche Liaison de Données

A B C D E F G  
DLE STX DLE DLE STX C DLE DLE E DLE DLE ETX DLE ETX  
début DLE de transparence fin

- Données remises à la Couche Réseau (destinataire)

A B C D E F G  
DLE STX C DLE E DLE ETX

## Bits de transparence

- Train de bits à envoyer

011011111111111111110010



- Ajout des bits de transparence et des fanions

01111110 0110111111011111011111010010 01111110

début bits de transparence fin

- Retrait des bits de transparence et des fanions

011011111111111111110010

## Violation du codage

- Codage Manchester (simple)

- Inclus le signal d'horloge

- $\frac{1}{2}$  temps bit à l'inverse de la valeur  
+  $\frac{1}{2}$  temps bit à la valeur.



- Violation du codage:  
peut être utilisé pour délimiter les trames



## Erreurs de transmission

- Elles existent et existeront encore:

- Bruit **thermique**: agitation des électrons (cuivre)  
=> bruit large spectre. Cf. rapport Signal/Bruit dans le Théorème de Shannon.
- Bruit **impulsif**: étincelles rupture (relais), surtensions  
=> impulsions avec des périodes de 10ms.
- **Amplitude, vitesse, propagation** et **phase** des signaux dépendent de leur fréquence (téléphone).
- **Diaphonie**: proximité physique de deux lignes.
- Etc.

## Rapport Signal/Bruit

- La part de **bruit** dans un signal reçu (rapport S/B) est exprimée en **Décibels (db)**.

- Unité logarithmique:

quand S/B croit exponentiellement, le débit en décibels croit linéairement.

$$\bullet (S/B)_{db} = 10 \log_{10} (S/B)_{bits/s}$$

- **Nyquist et Shannon**:

$$D_{bits/s} = F_{Hz} \log_2 (1 + S / B)$$

où D = débit binaire maximal d'un canal bruité,  
F = bande passante utilisable dans le canal.

## Erreurs

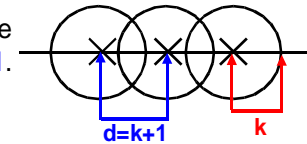
- Taux d'erreurs moyen: simples ou par paquets?
- Codes correcteurs (beaucoup de redondance): être capable de reconstituer les données originales
- Codes détecteurs (juste un peu de redondance): être seulement capable de détecter, pas corriger
- **Distance de Hamming:** (XOR)  
nombre de bits différents entre 2 mots du code  
Ex: 10001001  $\wedge$  10110001 = 3

```

10001001
10110001
-----
00111000
    
```

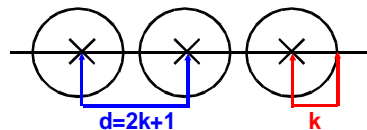
## Détecter

- Trame: longueur  $n$  bits =  $m$  données +  $r$  contrôle
  - l'ensemble des  $n$  bits est un **mot du code**
- Distance de Hamming d'un code est  $d$   
=>  $d$  erreurs simples pour passer d'un mot à un autre
- $2^m$  combinaisons de bits de données mais seule une partie des  $2^n$  mots du code est accepté
- Détecter  $k$  erreurs nécessite une distance de Hamming de  $d=k+1$ .



## Corriger

- Pour corriger  $k$  erreurs, il faut que la distance de Hamming soit d'au moins  $d=2k+1$



- Le nombre minimal de bits de contrôle  $r$  permettant de corriger toutes les erreurs *simples* doit vérifier:

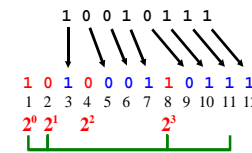
$$(m+r+1) \leq 2^r$$

- Comme on connaît  $m$ , on peut trouver  $r$ .

## Code de Hamming

- Permet d'atteindre la limite théorique. (1950)

- Bits de données
- Bits aux **puissance de 2** = bits de **contrôle**.  
Les **autres** = bits de **données**.



- Chaque bit de donnée est contrôlé par les bits de contrôle qui entrent dans sa décomposition en somme de puissances de 2.
- Ex: le bit 11=8+2+1 est vérifié par les bits **8, 2 et 1**.

## Code de Hamming (suite)

- Vérification de la parité des bits de contrôle:
  - Les positions de ceux qui sont faux sont cumulées et donnent la **position du bit erroné**.

- Correction de paquets d'erreurs de longueur k

H:	1001000	0	0	1	1	0	0	1	0	0	0	0	0	
a:	1100001	1	0	1	1	1	0	0	1	0	0	1		
m:	1101101	1	1	1	0	1	0	1	1	0	1	0	1	
m:	1101101	1	1	1	0	1	0	1	1	0	1	0	1	
i:	1101001	0	1	1	0	1	0	1	1	1	0	0	1	
n:	1101110	0	1	1	0	1	0	1	0	1	1	1	0	
g:	1100111	1	1	1	1	1	0	0	1	1	1	1		

K=7

- **kr** bits de contrôle pour corriger un seul paquet d'erreurs d'au plus **k** bits sur un bloc de **km** bits de données.

## Codes polynomiaux

- Code de redondance cyclique (CRC)
  - Arithmétique polynomiale modulo 2 + th. algébrique
    - pas de retenue; addition = soustraction = XOR
  - Générateur G(x) de degré r (suite de bits = coef. polynôme)
    - On divise  $x^i.M(x)$  par G(x) modulo 2 (reste R(x))
    - On transmet le résultat  $T(x) = x^i.M(x) - R(x)$
    - A réception, on divise la trame reçue par G(X)
    - Reste nul si OK, sinon  $(T(x)+E(x)) / G(x) = E(x)/G(x)$
  - Si erreur, le récepteur reçoit  $T(x)+E(x)$ .
    - nombre de bits inversés = nombre de coef. à 1 dans E(x)
    - Si  $G(x)=(x+1)$ , G'(x) alors les trames ayant un nombre impair d'erreurs seront détectées.

## Un exemple de calcul de CRC

- À émettre: 1101011011; générateur G(x): 10011
  - Ou encore  $x^9+x^8+x^6+x^4+x^3+x^1+x^0$  divisé par  $x^4+x^1+x^0$

1101011011	0000		10011
10011			1100001010
10011			
10011			
00001			
00000			
00010			
...			
	1110		

- Trame émise: 1101011011**1110**

## CRC Ethernet

- Il existe différents polynômes générateurs (G) normalisés:
  - CRC-12  $G(x) = x^{12}+x^{11}+x^3+x^2+x+1$ 
    - Pour les caractères codés sur 6 bits
  - CRC-16  $G(x) = x^{16}+x^{15}+x^2+1$
  - CRC-CCITTG(x) =  $x^{16}+x^{12}+x^5+1$ 
    - Pour les caractères codés sur 8 bits; produisent 16 bits
- Dans le cas Ethernet, il génère un code sur 32 bits
 
$$G(x) = x^{32}+x^{26}+x^{23}+x^{22}+x^{16}+x^{12}+x^{11}+x^{10}+x^8+x^7+x^5+x^4+x^2+x+1$$

## Contrôle d'accès au support physique

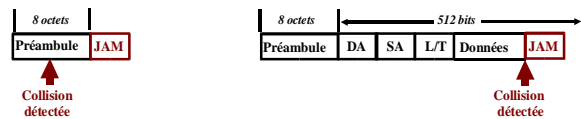
- Porteuse = un signal est transmis sur le canal
  - Sinon, le canal est silencieux
- Pour émettre = attendre l'absence de porteuse
  - Le canal doit être au repos
- Canal au repos + espace inter-trame -> émission
  - Court délai après le passage au repos du canal
- Si émission simultanée => détection de collision
  - Les deux émetteurs réorganisent leur transmission

## CSMA/CD

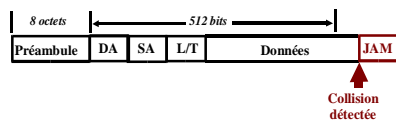
- Écouter avant d'agir
  - Support écouté en permanence (signal = émission)
    - Si silence, attente de 96 temps bits (IFG) avant émission
    - Délai de silence obligatoire entre 2 trames
  - Si collision (tension moyenne double de la normale)
    - => émission d'une séquence de brouillage, puis repli
    - 32 bits (différents du CRC) sont émis (96 bits min au total)
    - Attente pseudo aléatoire (multiple de 512 temps bits)
  - Délai d'insertion, *slot time* ou temps d'acquisition du canal
    - 512 temps bits (51,2µs) pour Ethernet/803.2 (10 Mbit/s)

## Collisions et brouillage

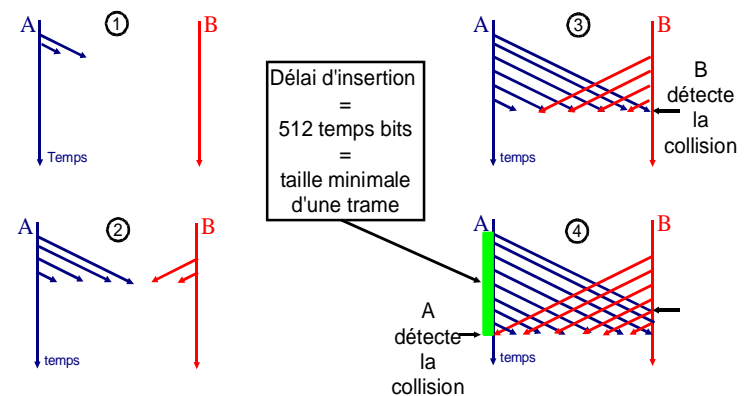
- Collisions « in-window »



- Collision « out-window »



## Délai d'insertion ou slot time

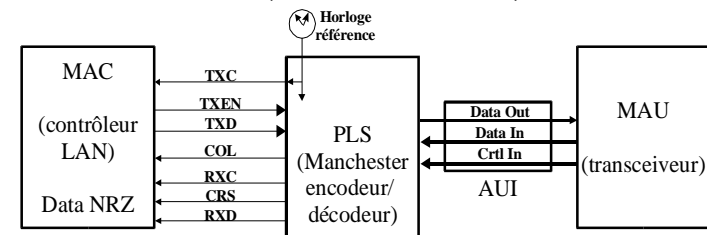


## Algorithme de repli (T.B.E.Backoff)

- Multiple du délai d'insertion ( $r * 512$  temps bits)
  - $r$  est un nombre aléatoire vérifiant  $0 \leq r < 2^k$  où  $k = \min(n, 10)$  et  $n$  est le nombre de tentatives de retransmissions
- Calcul individuel pour chaque émetteur
  - Minimise les probabilités de collisions successives
- 16 tentatives de retransmission maximum

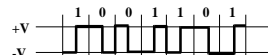
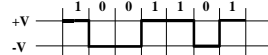
## Signalisation physique (PLS)

- Couche PHY, mais réside dans l'équipement
  - Fonctions: émission/réception, encodage/décodage, de écoute du support et détection des erreurs
- Interface MAC/AUI (i.e. NRZ/Manchester)



## Signal et codage

- Codage unipolaire sans retour à zéro (NRZ)
  - Machine (horloge)
- Codage Manchester (simple)
  - Inclus le signal d'horloge
    - $\frac{1}{2}$  temps bit à l'inverse de la valeur
    - $+\frac{1}{2}$  temps bit à la valeur.
- Codage Manchester différentiel
  - Bit 0 = Changement de polarité
  - Bit 1 = Polarité du début temps bit identique à précédente
    - Le sens des fils n'a plus d'importance.



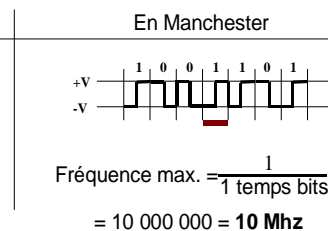
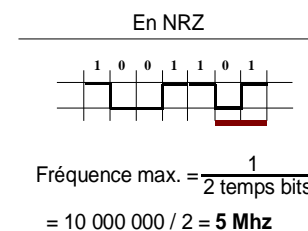
## Débits, codages et fréquences

- Le codage Manchester requiert une fréquence double du débit binaire.

Débit = 10.000.000 bits / seconde

=> 1 temps bit = 1/10.000.000 seconde

$$\text{Fréquence} = \frac{1}{\text{Période}}$$



## « Signal Quality Error » (SQE Test)

- Salve de 10 bits après chaque trame émise
  - Du transceiver vers l'interface Ethernet
    - Très courte salve de bits sur CI à une fréquence de 10Mhz
  - Aucune incidence sur le réseau
    - Seulement du transceiver vers la NIC, rien vers le réseau
    - Émis pendant l'IFG, donc pas d'influence sur le débit
  - Permet de tester l'électronique et le fil de signalisation de collision
    - Interprété par l'interface comme un test réussi et non comme une collision

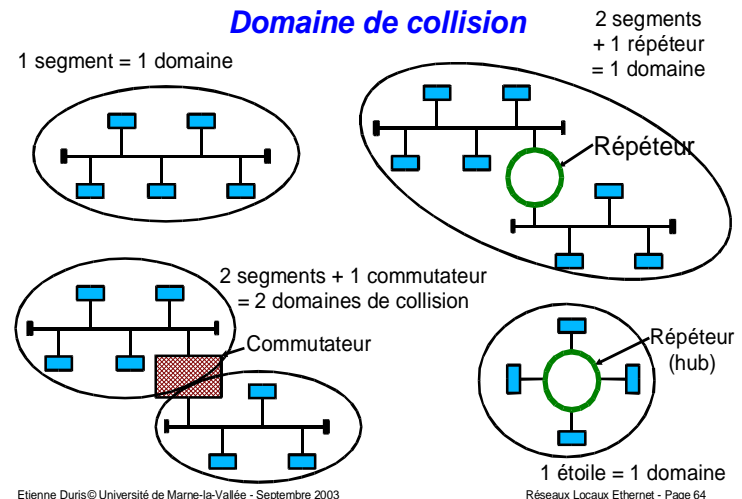
## SQE Test (ou « heart beat »)

- Les transceivers disposent d'un interrupteur
  - Ne pas confondre SQE Test, normal s'il fonctionne, avec SQE qui est le signal d'une collision
- **Activé** sur le transceiver d'un **ordinateur**
  - L'absence de battement peut être interprété comme un dysfonctionnement du transceiver ou du réseau
- **Désactivé** sur un transceiver relié à un **répéteur**
  - Les répéteurs ne disposent pas de l'IFG. Ils doivent réagir immédiatement aux signaux sur les segments.

## Topologie CSMA/CD

- Le partage du support impose des normes
  - Slot time de 512 temps bits
  - Temps d'inter-trame de 96 temps bits
- Étroitement liées aux caractéristiques du réseau
  - Longueur maximale (distance max entre 2 machines)
  - Qualité du support physique (vitesse, propagation...)
- Différents « segments » peuvent être reliés
  - À condition de respecter les normes

## Domaine de collision



## Répéteur

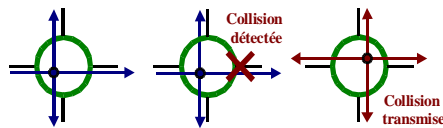
- Rôle principal: régénérer un signal reçu
- Équipement de réseau de couche physique (PLS)
  - Dispose le plus souvent d'une carte Ethernet
  - Considéré comme périphérique de la couche PHY
  - Néanmoins, comportement indépendant du support
- Même protocole, même vitesse, différents média
  - La liaison de deux segments de vitesses ou de protocoles distincts nécessite un pont, un commutateur ou un routeur

## Fonctions d'un répéteur

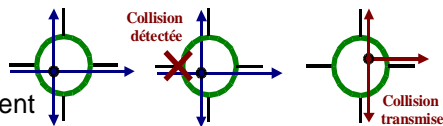
- Restauration du signal
  - Amplitude
  - Synchronisation
  - Symétrie
- Restauration du préambule
- Détection des collisions
  - *Transmit/receive collisions* et extension de fragment
- Partitionnement

## Détection des collisions

- Transmit collision



- Receive collision



- Extension de fragment

- Trames inférieures à 96 bits étendues à 96 bits

## One Port Left

- Garanti qu'un lien inter-répéteur ne peut être source de brouillage indéfini



- Après les 96 bits de jam, un répéteur qui n'a plus qu'un **seul port** en collision émet ce jam sur tous les ports **sauf** celui qui est en collision.

## Limitation des répéteurs

- Pas plus d'un lien actif entre 2 répéteurs
  - Plus généralement, pas de boucle en CSMA/CD
- Nombre de répéteurs limité à 4 pour ≠ raisons
  - Temps maximal de parcours du signal
    - round-trip delay = 512 temps bit
  - Rétrécissement de l'espace inter-trame (IFG)
    - Consécutif à la restauration du préambule
    - Variation possible d'une trame à l'autre
    - IFG peut devenir nettement inférieur à 96 temps bit

## Partitionnement

- Capacité à isoler un segment défectueux
  - Automatique et temporaire (réversible)
  - Optionnel sur 10Base5 et 10Base2
  - Obligatoire sur 10Base-T
- Soit si un signal de collision dure trop longtemps
- Soit si 30 échecs consécutifs de transmission de trame

## Supports physiques à 10 Mbit/s

- **10Base5** ou thick Ethernet (épais) câble coaxial
  - 10 pour le débit (Mbit/s), Base pour signal en bande de base et 5 pour la taille max. du segment (x100m)
- **10Base2** ou thin Ethernet (fin) câble coaxial
- **10Base-T** : paire torsadée non blindée (UTP)
  - 100m maximum, catégories : 3, 5 les plus fréquentes
- **10Base-F** ou fibre optique
  - 2 km entre répéteurs, immunité électro-magnétique
  - Supporte facilement l'augmentation de débit

## 10Base5

- « Thick Ethernet » ( Ethernet épais)
  - Coaxial (1cm de Ø, jaune, peu flexible, 50 ohms)
    - Jusqu'à 500 mètres de câble pour 1 segment
    - Max 100 machines, tous les 2,5 mètres (bandes noires)
    - Connectique N-type entre câbles
  - MAU (transceiver) = prise vampire
  - AUI = câble entre transceiver et machine
    - Prise 15 broches (DB15)
    - Longueur maximale du câble = 50 mètres
  - Carte Ethernet avec connecteur 15 broches



## 10Base2

- « Thin Ethernet » (Ethernet fin) ou « Cheapernet »

- Coaxial (0,5cm de Ø, noir/gris, + flexible, 50 ohms)

- Jusqu'à 185 mètres de câble pour un segment
- Maximum de 30 machines
- Espacées de 50 cm

- Transceiver peut être intégré à la carte Ethernet

- Dans ce cas, pas besoin de câble de transceiver (AUI)

- Connectique T-BNC et barillet



- Pour raccorder une machine directement au câble coaxial

## 10Base-T

- Ethernet « à paire torsadée »

- Paire torsadée non blindée (Unshield Twisted Pair)

- 4 paires polarisées (1 fil Ø, 1 fil ⊖);
- 100 mètres par segment;
- Peut supporter le mode optionnel de fonctionnement full duplex (802.3x);
- Impédance 100 ohms;

- Topologie en étoile (répéteur)

- Transceiver le + souvent intégré à la carte Ethernet

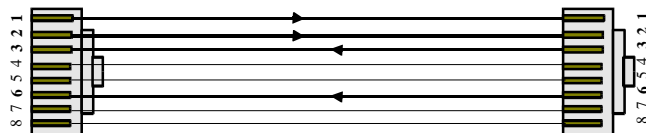
- Connectique type RJ45 (8 broches)

## Câble de liaison

- La liaison d'une station à une autre par un lien équipé de la même prise aux deux extrémités pose un **problème** (symétrie).

- 2 fils sont dédiés à l'émission

- 2 fils sont dédiés à la réception

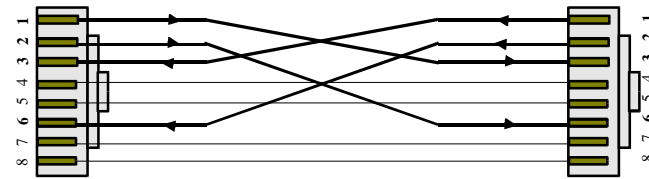


MDI (Media Dependant Interface)

## Câblage croisé

- Pour relier deux machines entre elles, il faut donc inverser les connexions des fils aux broches de la prise entre les deux extrémités (MDI-X)

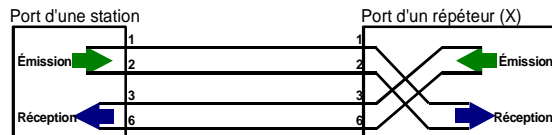
- La broche d'émission doit être reliée à une broche de réception (et vice versa)



MDI-X

## Câblage des ports d'équipements

- Pour éviter d'avoir à sertir des câbles croisés, les répéteurs, hubs, etc. fournissent des ports dont les fils sont croisés en interne (marqués d'un « X »)
  - Câble croisé entre deux machines ou deux hubs (MDI-X)
  - Câblage « droit » entre une machine et un hub (MDI), car les fils sont croisés à l'intérieur des ports des répéteurs



## 10Base-F (Fibre optique)

- FOIRL Fiber Optic InterRepeater Link (original)
  - Seulement entre 2 répéteurs (ancien), 1000m max.
- 10Base-F (nouveau) 3 types de segments:
  - 10Base-FL (Fiber Link) liaison optique (2000m)
  - 10Base-FB (Fiber Backbone) inter répéteurs
  - 10Base-FP (Fiber Passive) segment optique passif
    - (Passive star Fiber Ethernet) Concentrateur qui peut connecter jusqu'à 33 machines en étoile
- Codage Manchester avec lumière (LED ou laser)
  - allumée (1) ou éteinte (0)

## Câble à fibre optique (10 Mbit/s)

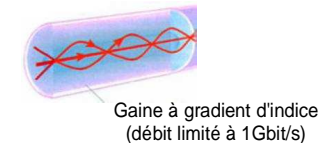
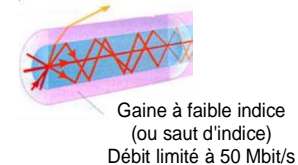
- Multimode (MMF) : 62,5/125
  - Noyau de 62,5µm, revêtement externe de 125µm
  - 2 brins ou un groupement
  - Longueur d'onde 850 nm, perte max. 12,5 dB
  - Connecteur ST



- Monomode (SMF)
  - Noyau compris entre 2 et 10 µm
  - Distance maximale bien supérieure (beaucoup + cher)

## Fibres optiques

Fibres multimode: diamètre de coeur de 50 à 80 microns



Fibres monomode: diamètre de coeur de 8 à 10 microns

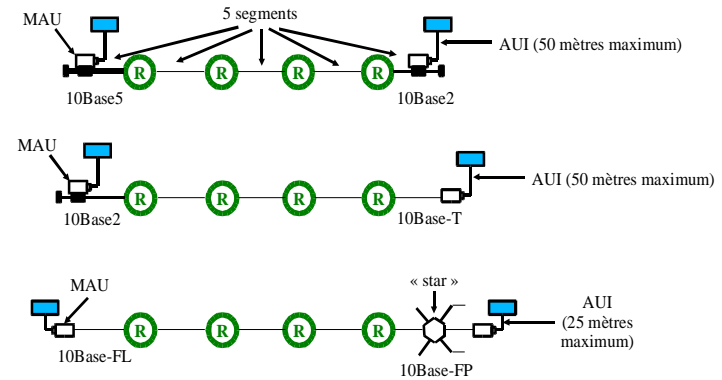


## Règles de configuration (10 Mbit/s)

### • Modèle 1 (configuration « sur étagère »)

- Entre 2 équipements d'1 même domaine de collision
  - un maximum de 5 segments,
  - de 4 répéteurs,
  - de 2 MAU et de 2 AUI
- La longueur d'un câble AUI pour 10Base-FL et 10Base-FP doit être inférieure à 25 m.  
Pour tout autre câble AUI, cette longueur doit être inférieure à 50 m.

## Exemples de configurations

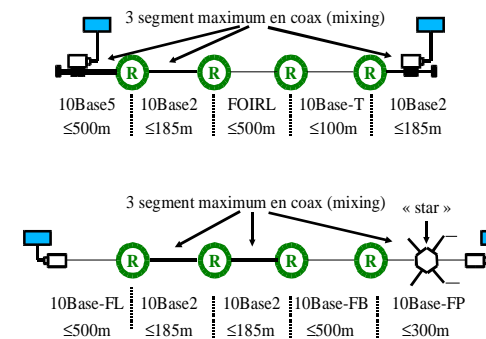


## Règles 10 Mbit/s Modèle 1 (suite)

### • Si chemin = 4 répéteurs et 5 segments

- Au maximum 3 segments peuvent être mixés (segments hybrides ou « mixing segments » ou coax)  
les 2 autres doivent alors être des segments de lien (« link segment »), i.e., point à point.
- Chaque segment de lien fibre optique (FOIRL, 10Base-FL ou 10Base-FB) doit être  $\leq 500\text{m}$
- Chaque segment 10Base-FP doit être  $\leq 300\text{m}$

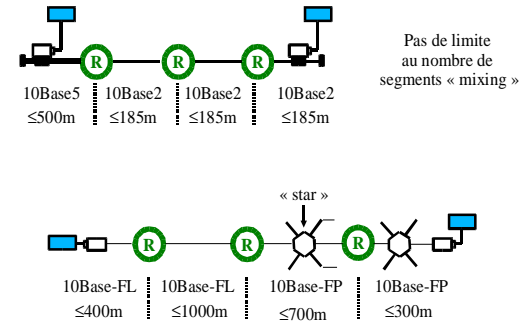
## Exemple 4 répéteurs / 5 segments



## Règles 10 Mbit/s Modèle 1 (fin)

- Si chemin = 3 répéteurs et 4 segments
  - Le nombre de segments mixés n'est pas limité
  - La longueur de chaque segment inter-répéteur ne peut excéder :
    - 1000 mètres pour FOIRL, 10Base-FL et 10Base-FB
    - 700 mètres pour 10Base-FP
  - La longueur de chaque segment répéteur-DTE ne peut excéder :
    - 400 mètres pour 10Base-FL (ou FOIRL)
    - 300 mètres pour 10Base-FP

## Exemple 3 répéteurs / 4 segments

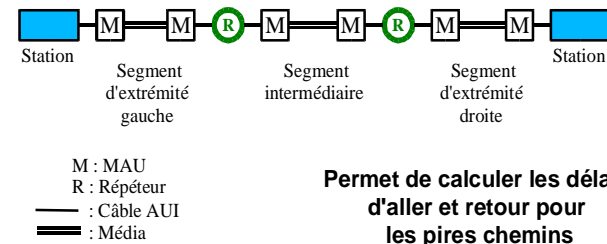


## Règles de configuration 10 Mbit/s

- **Modèle 2** (cas différents de « sur étagère »)
  - Modèles de réseaux
    - Segments d'extrémité gauche, intermédiaires et d'extrémité droite
  - Trouver les « pires chemins »
    - Pour lesquels les segments sont les plus longs et avec le plus grand nombre de répéteurs.
  - Calcul des délais d'aller et retour (+ 5 bits)  $\leq 575$
  - Calcul du rétrécissement de l'espace inter-trame  $\leq 49$

## Modèle de réseau

- Modélisation du réseau sous cette forme:



## Table des délais d'aller et retour

Type de segment	Long Max	Extr. gauche Base   Max	Seg. Interm. Base   Max	Extr. Droite Base   Max	Délai RT par mètre
10Base5	500	11,75   55,05	46,5   89,8	169,5   212,8	0,0866
10Base2	185	11,75   30,731	46,5   65,48	169,5   188,48	0,1026
FOIRL	1000	7,75   107,75	29   129	152   252	0,1
10Base-T	100	15,25   26,55	42   53,3	165   176,3	0,113
10Base-FP	1000	11,25   111,25	61   161	183,5   284	0,1
10Base-FB	2000	N/A   N/A	24   224	N/A   N/A	0,1
10Base-FL	2000	12,25   212,25	33,5   233,5	156,5   356,5	0,1
Excess AUI	48	0   4,88	0   4,88	0   4,88	0,1026

## Rétrécissement inter-trame

- On doit distinguer l'extrémité émettrice de l'extrémité réceptrice, puis échanger
  - Rétrécissement total doit être inférieur ou égal à 49 bits

Type de segment	Extrémité émettrice	Segment intermédiaire
Coaxial	16	11
Lien sauf 10Base-FB	10,5	8
10Base-FB	N/A	2
10Base-FP	11	8

## Un exemple simple

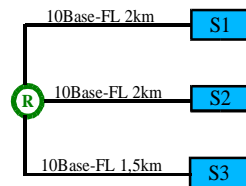
- Exemple :  
(pire chemin S1<->S2)

- Délai A/R:

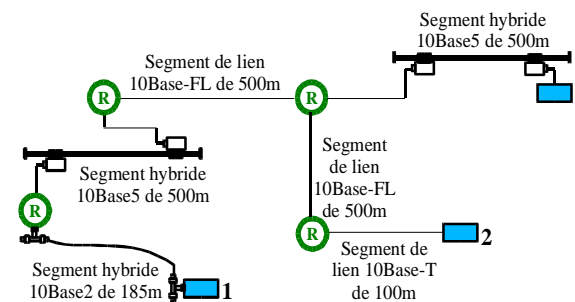
- Extr. G : 212,25 +  
Seg. Inter (absent)  
Extr. D : 356,5 +  
Marge : 5 = **573,75** ≤ 575

- Rétrécissement de l'intervalle inter trame :

- Extr. Émettrice : 10,5 ≤ **49**



## Exemple de configuration (conforme aux modèles 1 et 2)



## Ponts et commutateurs

- Unités de couche 2 (MAC) 802.1D :
  - « déplacer des trames d'un réseau local à un autre »
  - Autrement dit « interconnecter » des LAN
  - En Ethernet, crée plusieurs domaines de collision distincts
  - Permet de relier des réseaux Ethernet à  $\neq$  vitesses
- Apprentissage d'adresse et filtrage (promiscuité)
  - Associe l'@MAC d'1 station à son port d'attachement
  - 1 buffer / port: permet certaines vérifications des trames
  - Filtrage = élimination des trames déjà sur le bon port

## Principe général d'un pont transparent

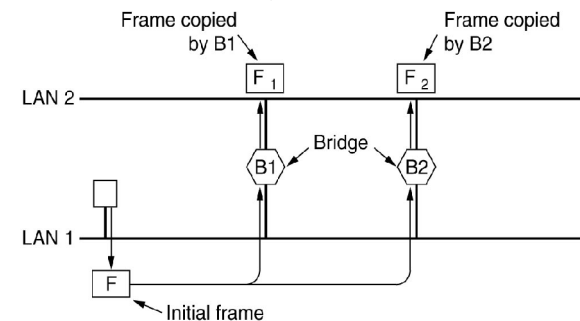
- Initialement: limiter la transmission des trames
  - Utiliser pour cela les adresses MAC
  - Apprendre les adresses MAC des émetteurs reliés à chacun des port du pont (table d'apprentissage)
  - Une trame reçue sur un port P: récupération de @Dest
    - Si elle est destinée à une machine accessible via P, détruire la trame: c'est le filtrage (*discard*)
    - Si DA correspond à un port D (table), on transmet vers D
    - Si on ne trouve pas l'adresse, transmission vers tous sauf P
  - Toute trame incomplète, incorrecte ou ayant subi une collision est détruite

## Ponts à translation

- Par opposition aux ponts transparents, qui relient des supports différents mais utilisent le même format de trame sur la couche liaison de données
- Ponts à translation : permettent de relier différents protocoles de communication de la couche 2 (Ethernet avec Token Ring, par exemple, mais aussi Ethernet II avec Ethernet SNAP)
  - Modification de l'encapsulation = « translation » (i.e., traduction)

## Ponts transparents en parallèle

- Le principe d'apprentissage exclut les boucles



Copyright A. Tanenbaum, Pearson Education, 2003

## Algorithme « Spanning Tree »

- Défini par IEEE 802.1D
  - Correspond aux ponts transparents
  - En cas de multiples ponts entre deux réseaux
  - Risque de boucle (en particulier pour @ multicast)
- Superposer au graphe des connexions réseau un « arbre recouvrant », sans cycle.
  - 1. Trouver une racine (w.r.t. ≠ critères, @MAC min)
  - 2. Calculer la distance du plus court chemin à la racine
  - 3. Pour chaque LAN, élire un pont « désigné » (le + près)
  - 4. Choisir un port « racine » qui offre le meilleur chemin
  - 3. Placer dans le ST le port racine et quelques autres...

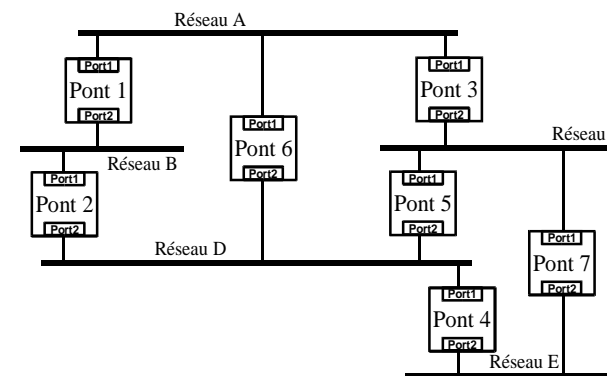
## Spanning Tree Algorithm (principe)

- Au début, tout port se considère comme racine
  - Id = 2 octets de priorité + 6 octets d'adresse MAC
- Communication entre ponts par transmission de BPDU par multicast (01-80-c2-00-00-00) de trame LLC (SAP 0x42)
  - Root ID: ID du pont présumé racine
  - Coût du « meilleur » chemin du pont émetteur vers la racine
  - Transmitter ID: ID du pont émetteur du BPDU (+ port)
- Chaque pont garde le meilleur BPDU reçu (y compris lui-même)
  - Identifie le port racine (allant au pont racine au meilleur coût)
- Lorsqu'un pont reçoit sur un port un meilleur BPDU que celui qu'il devait transmettre, il ne transmet plus sur ce port
  - Stabilité: seul le pont « désigné » émet sur le LAN

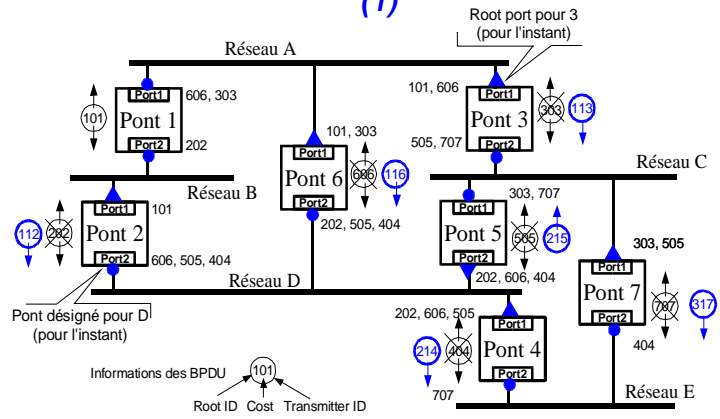
## Choix des ponts et des ports

- La notion de coût intègre différents critères:
  - Somme de l'inverse des débits des réseaux traversés
  - Somme des délais d'émission
  - Si litige, niveau de priorité (ID), port le + faible, etc.
- Si un pont P continue à émettre des BPDU vers un ou plusieurs LAN, P est dit « pont désigné » pour ces LAN
  - Au final, les ports d'un pont P qui sont placés dans le ST:
    - Le port racine « root » du pont P
    - Tous les ports des LAN pour lesquels P est le pont désigné.
  - Les autres ports sont bloqués

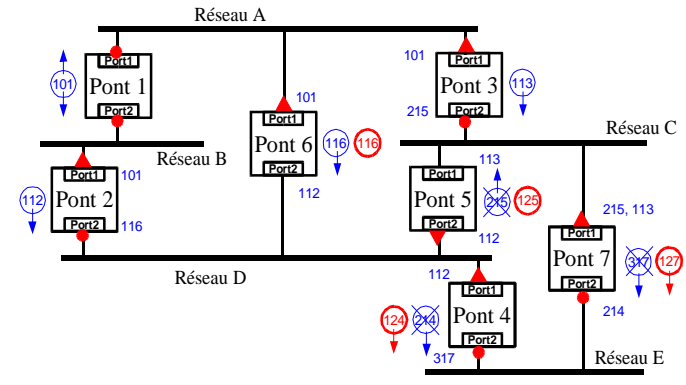
## Exemple de réseau ponté (avant STA)



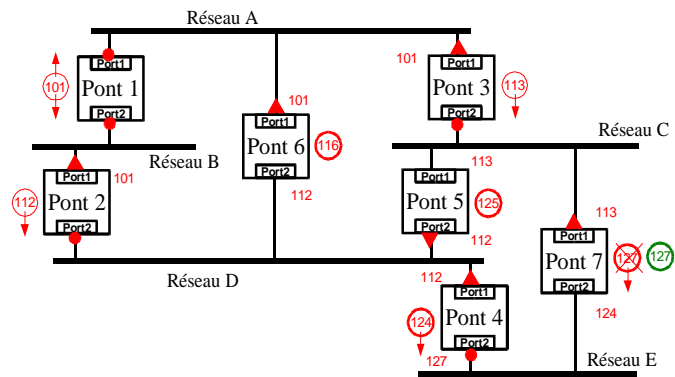
### Exemple de réseau ponté (pendant STA) – (1)



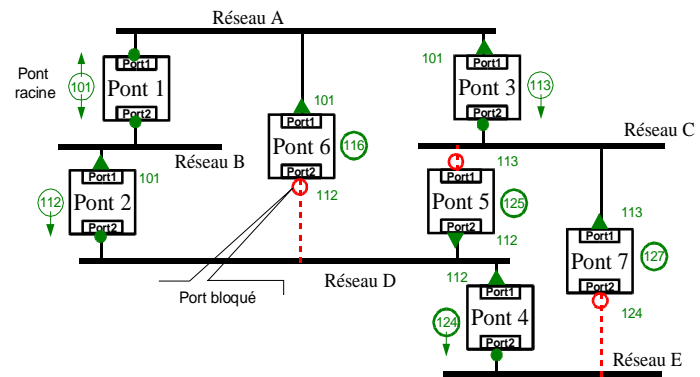
### Exemple de réseau ponté (pendant STA) – (2)



### Exemple de réseau ponté (pendant STA) – (3)



### Exemple de réseau ponté (après STA)



## Autres techniques de pontage

- Routage par la source (*Source Routing - SR*)
  - Proposé par IBM pour Token Ring, puis FDDI
  - Idée: la source donne le chemin, le pont ne fait qu'aiguiller en fonction des indications dans la trame MAC.
  - Ce sont les stations qui savent comment atteindre leurs interlocuteurs
    - Utilisation de trames de recherche (exploration)
- *Source Route Transparent - SRT*
  - Rendre le pontage à translation plus aisé

## Commutateurs (switch) – Couche 2 (MAC)

- Essentiellement, c'est la même chose qu'un pont.
- Historiquement, les commutateurs ont fait dans l'électronique ce que les ponts faisaient en soft.
- Hub de base = répéteur
- Hub intelligent (*smart*) *Intelligent Switching Hub*
  - Basés sur les PABX de la téléphonie
  - Matrice de commutation (parallèle)
  - Autorisent plusieurs communications simultanées entre différents ports

## Techniques de commutation

- À la volée (*cut-through*)
  - On récupère l'adresse MAC destination
  - Une fois qu'on a identifié le port vers lequel on doit transmettre, on « commute » le flot vers ce port
  - Pas de correction ni d'assurance de qualité des trames
- Enregistrement & acheminement (*store & forward*)
  - Ressemble plus au fonctionnement des ponts
  - La trame est reçue en entier avant d'être transmise
  - Si erreur, elle est jetée
- Adaptif (mixage des 2 méthodes)

## Utilisations des ponts/commutateurs

- Isolement de trafic
- Extension de distances
- Vitesses différentes
- Commutation hybride et répéteur
- Création de LAN Virtuels (VLAN)
- Interface de connexion à un backbone ATM (Émulation de LAN ou LANE)

## **Limitations des ponts/commutateurs**

- Pas plus de 7 « sauts » entre 2 stations (diamètre)
  - Nécessité de limiter le temps de réponse
  - Cumul des différents délai de transmission
- Problèmes spécifiques pour utilisation en tant qu'équipement inter-protocoles
  - Entre Ethernet et FDDI ou Token Ring, par exemple
  - Tailles max de trames ( $\approx 1500$  vs  $\approx 4500$ )

## **Routeurs**

- Couche réseau (niveau 3)
- Relient différentes technologies réseau (AppleTalk, TokenRing, ...) et différents protocoles (IP, IPX, DECnet...)
- Décompacte les trames et les recompacte
- Ne transmet pas automatiquement les broadcasts
- Incorporent des algorithmes de routage, fragmentation, i.e., cherchent le meilleur chemin

## **Ethernet et ses évolutions**

- **Ethernet et IEEE 802.3** (10 Mégabits par seconde)
  - Très largement répandu (installations, équipements)
  - Garder la compatibilité et la simplicité d'administration
- **Fast Ethernet** (100 Mbit/s) IEEE 802.3u
  - 100Base-T, MII, négociation 10/100 Mbit/s
- **Ethernet Gigabit** (1000 Mbit/s) IEEE 802.3z
  - 1000Base-T, GMII, négociation half/full-duplex
- **...Ethernet 10 Gigabit/s** (10.000Mbit/s) IEEE 802.3ae

## **Fast Ethernet (100 Mbit/s)**

- 1995. Ethernet 100 Mbit/s normalisé IEEE 802.3u
- Supports normalisés IEEE 100Base-T
  - 100Base-X, normes issues de FDDI
    - 100Base-TX (2 paires torsadées Cat. 5)
    - 100Base-FX (fibre optique)
  - Autres normes dédiées à la « vieille » paire cuivrée
    - 100Base-T4 (4 paires torsadées Cat. 3)
    - 100Base-T2 (2 paires torsadées Cat. 3)
- MII, auto-négociation 10/100 Mbit/s, 4B/5B

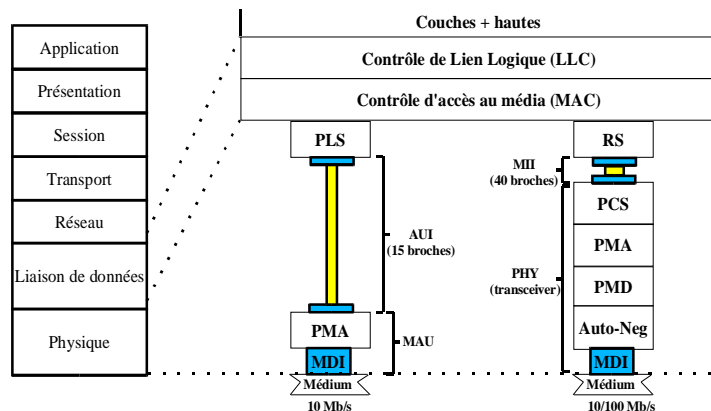
## Ethernet 100VG-AnyLAN

- IEEE 802.12 (Voice Grade)
  - Utiliser la paire torsadée (blindée ou non) de qualité vocale (Cat. 3) à des débits de 100 Mbit/s.
- Méthode d'accès: Demand Priority (DPAM)
  - Niveau haut: forte contrainte temps réel (voix, vidéo)
  - Niveau bas: trafic de données asynchrone
- Demande de transmission -> planifiées par le concentrateur 100VG-AnyLAN
  - Topologie en arbre (concentrateur ≈ pont transparent)

## Principales différences 10/100 Mbit/s (Fast Ethernet)

- Media Independant Interface (MII)
  - Remplace la notion d'AUI
- Sous couche de réconciliation (RS)
  - Transmission par série de bits / par quartets de bits
- Suppression du codage Manchester
  - Trop hautes fréquences => NRZ + clock dans le MII
- Vitesse duale 10/100 Mbit/s et half/full duplex
  - Protocole d'« auto-négociation » pour paire torsadée

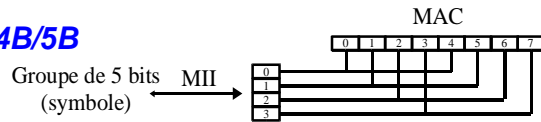
## 802.3 / 802.3u et 10 / 100Mbit/s



## Supports et codages

- Codage 4B/5B (4 bits -> symbole sur 5 bits)
  - 100Base-TX. ≤ 100 mètres. Full duplex OK.
  - 100Base-FX. De 100 à 2000 m. Full duplex OK.
    - Longueur maximale en half entre 2 machines de 412 m.
- Codage 8B/6T (1 octet -> 6 signaux ternaires)
  - 100Base-T4. ≤ 100 mètres. Full duplex KO.
- Codage PAM5x5 (Modulation d'amplitude)
  - 100Base-T2. ≤ 100 mètres. Full duplex OK

## Codage 4B/5B



- Chaque ½ octet transformé en groupes de 5 bits
  - 16 symboles représentent des données (0 à F)
  - 1 symbole représente le support libre (IDLE)
  - 15 symboles pour le contrôle (« J », « K », etc.)
- Fréquence élevée :
  - 100Mbit/s + 4 bits en // = 25Mhz (cadence MII)
  - 25Mhz + 4B/5B => 125 Mbauds sur le support
- En 100Base-TX, signalisation ternaire MLT-3. En 100Base-FX, c'est inutile.

## Deux types de répéteurs 100 Mbit/s

- **Classe I** (1 par domaine de coll. si câbles max)
  - Relier des types de supports différents
  - Décodage du signal sur port entrant, puis codage pour envoi sur port sortant
  - Temps de propagation assez important
- **Classe II** (2 par domaine de coll. si câbles max)
  - Relier des segments de même type uniquement
  - Pas de décodage / codage du signal
  - Temps de propagation plus court

## Règles de configuration (100 Mbit/s)

- **Modèle 1** (configuration « sur étagère »)
  - Longueur des segments UTP: 100 mètres
  - Longueur des segments de fibre: 412 mètres
  - Longueur des câbles MII: 0,5 mètre
  - Diamètre max. d'un domaine de collision:

Type de répéteur	Tout cuivre	Tout fibre	Cuivre (T2 T4 et fibre (FX)	Cuivre (TX) et fibre (FX)
1 segment DTE-DTE	100	412	N/A	N/A
1 répéteur de classe I	200	272	231 (a)	260,8 (a)
1 répéteur de classe II	200	320	N/A	308,8 (a)
2 répéteur de classe II	205	228	N/A	216,2 (b)

(a) 100m de câble paires torsadées et 1 lien fibre

(b) 105m de câble paires torsadées et 1 lien fibre

## Règles de configuration (100 Mbit/s)

- **Modèle 2** (autre que « sur étagère »)
  - Diamètre maximum
    - Aller retour ≤ 512 bits (temps de propagation par défaut)
  - Délai de segment
    - Dépendant du câble. Ex: Câble AT&T 1061 (cat. 5) (NPV 70%) de 100m:  $0,477 * 2(a/r) * 100 = 95,4 \text{ bits}$  + marge de 4 bits

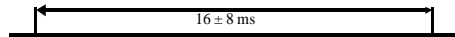
Composant	Délai aller/retour	Délai max.
2 DTE FX/TX		100
Câble cat. 5	1,112 bit/mètre	111,2
Fibre optique	1bit/mètre	412
R. classe II TX/FX		92

Vitesse relative à c	ns/m	Délai aller/retour
0,4 (40% NVP)	8,34	0,834 bit/mètre
0,6 (60% NVP)	5,56	0,556 bit/mètre
0,7 (70% NVP)	4,77	0,477 bit/mètre

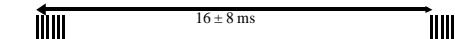
NVP: Nominal Velocity Propagation (relatif à c)

## Auto négociation en 100 Mbit/s

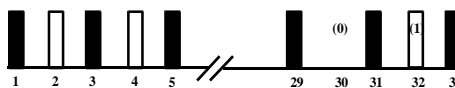
- Le 10 Mbit/s émet des « Normal Link Pulse »



- Le 100 Mbit/s émet des « Fast Link Pulse »



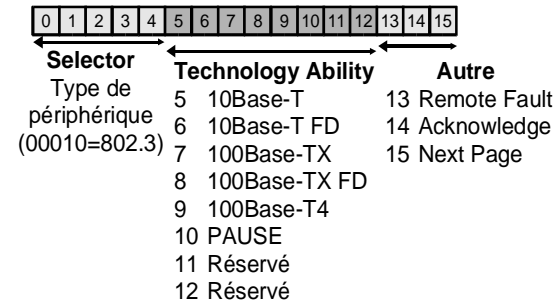
- Positions d'horloge (■) / positions de données (□) 2ms



- Pulsation de donnée: présent=1, absent=0
- Mot de 16 bit: (Base) Link Code Word (LCW)

## Base Link Code Word

- Transmis dans un Fast Link Pulse :

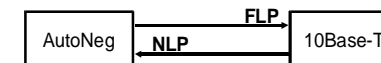


## Auto négociation en 100 Mbit/s

- Émission de FLP
  - Lors de la mise sous tension, de « reset » ou requête explicite de renégociation
- Détection NLP/FLP
  - Si NLP reçu: l'autre est 10Mbit/s seulement
  - Si FLP reçu: le LCW contient les informations pour configuration optimale
- Détection parallèle si pas d'auto négociation
  - Seulement les vieux 100Base-T4 et 100Base-TX

## Quelques exemples

- Interopérabilité avec 10Base-T

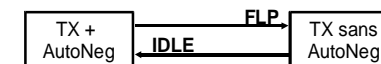


- Protection contre supports incompatibles



- Interopérabilité en 100Base-T sans auto négociation

- Détection //



## Full duplex et contrôle de flot

- IEEE 802.3x (1997)
  - Full duplex
    - 10Base-T full duplex = 20Mbit/s de bande passante
    - 100Base-T full duplex = 200Mbit/s bande passante
    - Un segment 100Base-FX en full duplex peut atteindre 2km, contre 412 m en half duplex (MMF)
  - Contrôle de flot sur liens full duplex:
    - MAC Control
    - PAUSE

## Full Duplex

- Émission et réception simultanée sur le support
  - Ex: 10Base-[T,FL], 100Base-[TX,FX,T2], 1000Base-X
  - Contre-ex: 10Base[5,2],10Base-[FP], 100Base-T4
- Exactement 2 extrémités au lien full duplex
- Pas d'utilisation de CSMA/CD
- Les deux cartes Ethernet pour un lien sont configurées en full duplex
- Respect du format des trames et de l'IPG

## Contrôle de flot

- MAC Control (trames spécifiques)
  - Transparent aux fonctions normales de contrôle
  - Remplace la régulation fournie par CSMA/CD
  - Permettre aux stations d'interagir en temps réel
- Valeur de type 0x8808 + « opcodes »
- Interprétées par la couche MAC
- Support du système PAUSE

## MAC Control

- Format de la trame MAC Control

6 octets	6 octets	2 octets	2 octets	44 octets	4 octets
Dest. Adr	Src. Adr	Type	op-code	Données Op-code ou padding	FCS

- Dest Adr : 01-80-C2-00-00-01 (multicast) pour PAUSE
- Src Adr : adresse unicast assignée au port émetteur
- Type : 8808
- MAC Control op-code : 00-01 pour une trame PAUSE
- Données : spécifiques en fonction de op-code
- Taille de la trame: taille minimale d'une trame

## Trame PAUSE

- Spécifique (et optionnel) pour full duplex
  - Permet de suspendre ponctuellement l'émission.
  - Les données qui accompagnent l'op-code 00-01 représentent une valeur « PAUSE timer » (16 bits).
  - Si PAUSE Timer non nul, l'émission est « inhibée »
  - L'émetteur (MAC) décrémente le PAUSE timer tous les 512 temps bits; à zéro, il recommence à émettre.
  - Émission de trames PAUSE pas inhibée.
  - Réception d'une nouvelle trame PAUSE possible.

## Switched Ethernet (commuté)

- Commence à devenir particulièrement intéressant avec le débit de Fast Ethernet
  - Fast Ethernet, Full Duplex et VLAN Tagging
  - Passage de « médium partagé » vers « commutation »
- Commutateur de niveau 2 (Switch 2) = pont
  - Problèmes de noms liés au marketing
  - Switch (hard, ASIC) + efficace que les vieux ponts
  - Multiports actifs simultanément, divers autoconfigurations et de nombreuses options.

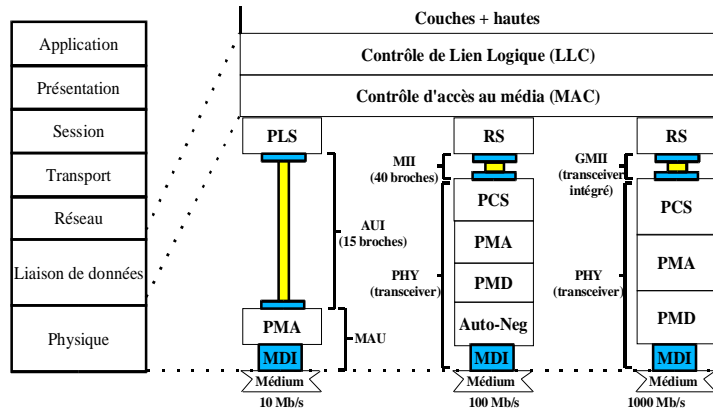
## Gigabit Ethernet

- 1998 : Standardisation IEEE 802.3z
  - 1000 Mbit/s = 1 Gbit/s: naissance d'Ethernet Gigabit
- Supports et codages inspirés de Fibre Channel
  - 1000Base-X
    - 1000Base-LX (paire de fibres grande longueur d'onde)
    - 1000Base-SX (paire de fibres longueur d'onde courte)
    - 1000Base-CX (2 paires de câble cuivré blindés 150 Ohm)
  - 1000Base-T (IEEE 803.2ab 1999) 4 paires UTP Cat 5
- GMII, autonegociation full/half-duplex, 8B/10B

## Principales différences entre 1 Gbit/s et 100 Mbit/s

- 8 bits (« octets wide ») / 4 bits (nibble-wide)
- Gigabit MII (GMII) / MII
- Adoption de « Fibre Channel » pour la signalisation (8B/10B)
- Une seule classe de répéteur
- Modification et généralisation de l'auto-négociation pour la fibre
- Extension de portuse en CSMA/CD, mais nette préférence pour le Full Duplex

## 802.3 / 802.3u (Fast) / 802.3z (Giga)

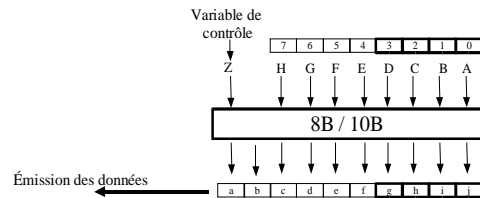


## Supports physiques

- 1000Base-CX: STP, 150 Ω. Taille max. 25 mètres.
- 1000Base-LX: (*long*) longueur d'onde: 1300 nm.
  - Soit mono-mode (SMF): 3 km max.
  - Soit multi-mode (MMF): 550m (50μ) et 440m (62,5μ).
- 1000Base-SX: (*short*) longueur d'onde: 850 nm.
  - Multi-mode (MMF): 500m (50μ) et 220m (62,5μ).
- 1000Base-T: 4 paires UTP, Cat. 5. 100m max.

## Signalisation: 8B/10B

- 1000Base-X fibre: inspiré de Fibre Chanel
  - Au lieu de 4 bits simultanés comme pour 100Mbit/s, il s'agit là de transmettre 8 bits simultanément.



- 1000Base-T: *pulse amplitude modulation (PAM)*

## Répéteurs Gigabit Ethernet

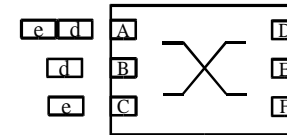
- Plus qu'une seule « classe » de répéteur
  - Tous les segments sont point à point (liens)
- Deux « types » de répéteurs pour Gigabit:
  - Unidirectionnels (principe half-duplex)
    - Leur fonctionnement est classique CSMA/CD
    - Limité à un seul par domaine de collision
  - Bidirectionnels (principe full-duplex)
    - « Diffuseur à mémoire répartie » (1 mémoire tampon / port)
    - Domaine de collision = intérieur du répéteur (bus Ethernet)

## Couche MAC Gigabit

- Principe fondamental d'Ethernet: compatibilité
  - À 10, 100 et 1000 Mbit/s conserver le format de trame
- En 10 et 100: taille minimale = 64 octets
- En Gigabit, possibilité d'extension de cette taille minimale, en fonction du type de fonctionnement:
  - Accès full-duplex
  - Accès half-duplex

## Gigabit Full-Duplex

- Pas de risque de contention
  - CSMA/CD n'est pas appliqué
- Format des trames conformes aux normes
- Contrôle de flux
  - Éviter la congestion dans les switches
  - Phénomène de « HOL » (Head of Line)
  - Notion de « backpressure »



## Gigabit Half-Duplex

- C'est le CSMA/CD qui gère les collisions/accès
- La vitesse, les fréquences imposent des restrictions.
- Pour autoriser des diamètres de 200 mètres, il faudrait augmenter la taille minimale des trames.
- Pour compatibilité, 2 techniques sont utilisées:
  - **Carrier Extension** (extension de porteuse)
  - **Frame Bursting** (pipelining de transmission)

## Extension de porteuse

- Afin d'autoriser de plus grands segments en half duplex: augmenter la taille minimale de trame, au sens présence d'un signal sur le support.
- *Slot time* ou délai d'insertion classique Ethernet de 512 temps bits devient de 512 temps octets, soit 4096 temps bits.
  - C'est le temps d'un aller et retour sur le domaine de collision.
  - Implique que cette valeur soit également utilisée par l'algorithme de repli (TBEB).

## Extension de porteuse

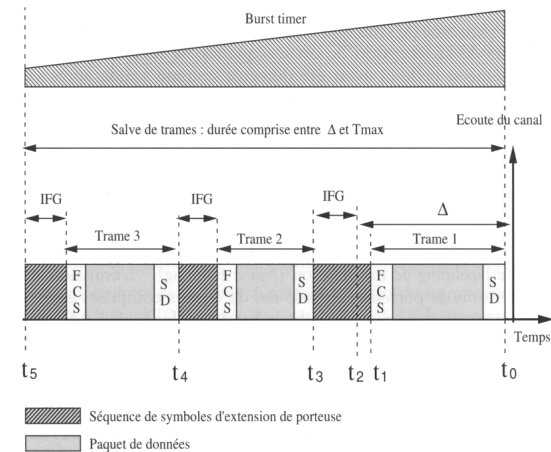
- Pour assurer la compatibilité avec 10 et 100 Mbit/s
  - Garder la taille min. de trame de 64 octets (512 bits)
  - Ajout d'un mécanisme de bourrage systématique, mais spécifique:
    - Lors d'émission d'1 trame entre 64 et 512 octets, cette trame est inchangée mais on ajoute (pour atteindre 512) un *symbole d'extension de porteuse* artificiel.
    - Ces symboles ne sont pas émis comme du bourrage, mais après le champ FCS de la trame.
    - Ainsi, le récepteur peut les supprimer aisément pour retransmettre la trame « courte » sur les brins 10, 100 Mbit/s

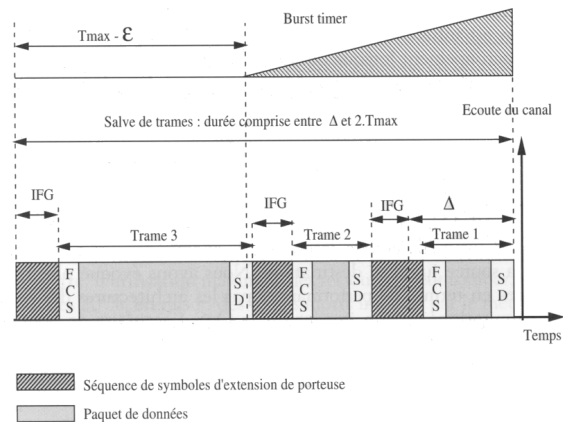
## Extension de porteuse

- En émission:
  - Si détection de collision après l'émission d'une trame courte (< 512 octets) et pendant l'extension => **repli**.
- En réception:
  - Si une trame courte est reçue correcte, il faut attendre la fin de la réception des symboles d'extension de porteuse avant de transmettre la trame à la couche supérieure
    - Si une collision se produisait alors, la trame risquerait d'être émise et reçue deux fois.

## Frame Bursting

- But: limiter la perte de bande passante liée aux émissions des symboles d'extension de porteuse
- Principe: envoyer plusieurs trames « courtes » (une salve ou *burst* de trames) dans une seule et même « trame » au sens CSMA/CD
  - C'est à dire lors d'une même acquisition de canal
- *Burst timer* : il s'agit d'une limitation du temps ou l'émetteur garde le droit d'émettre
  - Études actuelles sur la durée optimale du burst timer



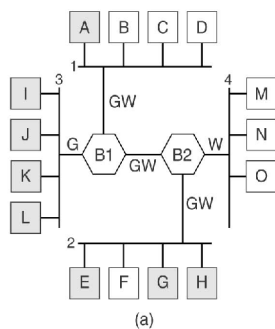


Copyright Dunod - Réseaux haut débit - Kofman&Gagnaire

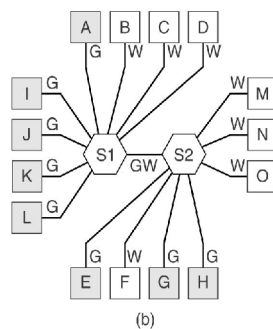
## Réseaux Locaux Virtuels (VLAN)

- Différents avantages des LAN
  - Communication rapide, broadcasts, plug & play
- Différents problèmes des LAN
  - Sécurité, *broadcast storms*, connectique
- Volonté de garder le meilleur en se passant du pire
  - Fonctions de la couche 3 avec la vitesse de la couche 2
  - Faciliter la gestion de la mobilité des postes (logicielle)
  - Différencier clairement les domaines d'activités (sécurité)
  - Conserver la compatibilité ascendante

## Exemples de topologies



(a)



(b)

Copyright A. Tanenbaum, Pearson Education, 2003

## VLAN: les avantages

- **Segmentation** du réseau local flexible
  - regrouper les utilisateurs / ressources qui communiquent le plus fréquemment indépendamment de leur emplacement
- **Organisation** virtuelle, **gestion** simple des ressources
  - modifications logique ou géographiques facilitées, et gérées via la console (plutôt que dans l'armoire de brassage)
- **Efficacité** de bande passante / utilisation des serveurs
  - limitation de l'effet des inondations de broadcasts. Partage possible d'une même ressource par plusieurs VLAN
- **Sécurité** réseau améliorée
  - VLAN = frontière virtuelle, franchissable avec un routeur

## Mise en oeuvre des VLAN

- Niveau 1: *VLAN par port*
  - l'administrateur associe un VLAN à chaque port
- Niveau 2: *VLAN par adresse MAC*
  - L'appartenance d'une trame à un VLAN est déterminée par l'adresse MAC de l'émetteur et / ou du destinataire
- Niveau 3: *VLAN par protocole*
  - L'appartenance d'une trame à un VLAN est déterminée par le protocole, le sous-réseau ou l'adresse de niveau 3

## Port-based VLAN

- L'administrateur configure le switch
  - il affecte chaque port à un VLAN (voir à plusieurs).
- Le switch détermine le VLAN des trames
  - à partir des ports sur lesquels elles arrivent
- Quand un utilisateur se déplace vers un autre port
  - il suffit d'affecter son VLAN au nouveau port
  - le changement du réseau est transparent pour l'utilisateur
  - **mais** il faut gérer le changement manuellement (admin.)
- **Mais** si un port est relié à un segment (ou répéteur)
  - tout ce segment est associé à ce même VLAN

## MAC address-based VLAN

- Chaque carte MAC est gérée individuellement
  - chaque switch maintient une table @ MAC <-> VLAN
  - **pbme**: il faut les initialiser (solution: VLAN par défaut)
- Le switch détermine le VLAN de chaque trame
  - à partir de l'adresse MAC source ou destination
- Quand un utilisateur se déplace vers un autre port
  - son VLAN reste le même, sans aucune modification
  - modification physique => aucune modification logique
- Différents VLANs possibles sur un même segment
  - mais les machines peuvent communiquer entre-elles

## Protocol-based VLAN

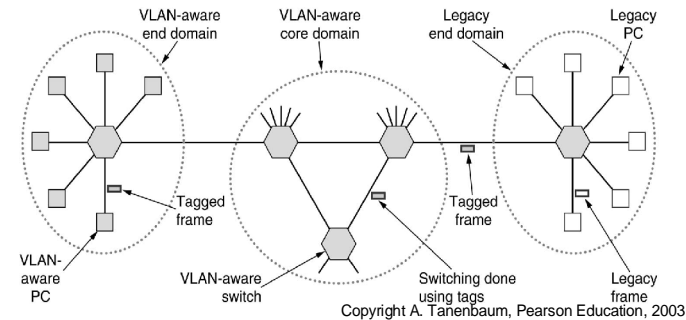
- Regroupement par protocole
  - un VLAN pour IP, un autre pour IPX, etc.
  - le switch associe le VLAN en fonction du protocole transporté
- Par adresse réseau (niveau 3)
  - un VLAN pour un ensemble d'adresses (IP, par exemple)
  - ou un VLAN pour une adresse de sous-réseau donnée
- Très flexible, et peut éliminer le frame tagging
  - organisation logique sans besoin de ré-affecter d'@ IP
  - néanmoins, pas de fonction de routage (≠ d'un routeur)
- Assez coûteux et viole les principes OSI
  - nécessite d'aller inspecter à l'intérieur de la trame

## Compatibilité (1)

- Comment une machine fait elle pour savoir à quel VLAN elle appartient?
  - La machine peut ne pas le savoir, ou ne pas pouvoir l'exprimer
  - L'administrateur le spécifie au niveau des switches
- Les switches ajoutent des « étiquettes » (*tags*) aux trames, et masquent leur présence aux machines
  - Ces tags permettent d'aiguiller les trames entre switches
  - Ils sont retirés de la trame avant qu'elle soit délivrée aux machines qui ne savent pas les interpréter.

## Compatibilité (2)

- *A priori*, seul le commutateur ou le pont a besoin de connaître l'appartenance d'une trame à un VLAN

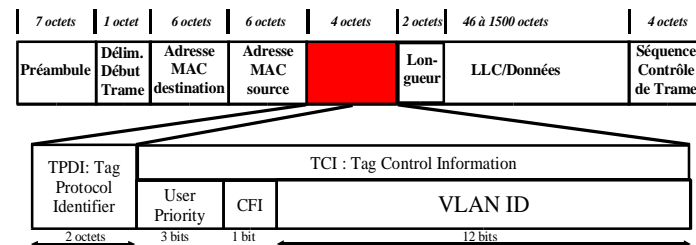


## Extension des trames

- Norme IEEE **802.1Q** (802.1p et 802.1q)
  - Définissent respectivement des extensions pour :
    - La priorité, ou Qualité de Service au niveau MAC, qu'on appelle alors Classe de Service ou (**CoS**)
    - L'interopérabilité des réseaux locaux commutés (*Virtual Bridged Local Area Networks*): les **VLAN**
  - Étendent le format d'Ethernet/IEEE 803.2
    - Ajoutent 4 octets pour spécifier ces informations (**tag**)
    - La longueur maximale d'une trame était de **1518** octets
    - Elle passe à **1522** octets (attention à l'inter-opérabilité)

## Format de trame (802.1p & 802.q)

- Ajout d'un champ de 4 octets (**tag**) après l'adresse source



## Champs IEEE 802.1p et 802.1q

- TPID (*Tag Protocol Identifier*). 2 octets
  - 0x8100 pour les trames « taggées »
- TCI (*Tag Control Information*). 2 octets
  - 3 premiers bits: *user priority* (802.1p) de 0 à 7
  - 1 bit CFI (*Canonical Format Indicator*). Ethernet:0
  - 12 bits: VID (*VLAN Identifier*)
    - 0 => l'en-tête ne contient que des infos de priorité
    - 1 => valeur par défaut d'identificateur de VLAN
    - FFF => réservé (implémentation)

## Trois types de trames VLAN

- pouvant traverser un switch « *VLAN-aware* »
  - Trame non taggée
    - absence d'en-tête (TPID+TCI) après adresse MAC source
  - Trame taggée de priorité
    - VID=0 => Pas signifiant. Ne transporte que des infos de priorité. Du point de vue du processus de transmission, elle est considérée comme une trame non taggée
  - Trame taggée de VLAN
    - TIPD=0x8100, CFI=0 et VID entre 1 et 4094.

## Politique de transmission

- 802.1Q (1998) a choisi de considérer:
  - Switchs *VLAN-aware* utilisent par défaut le tag
    - Si une trame est taggée de VLAN, on utilise son VID
    - Sinon (trame taggée de priorité ou non taggée), on peut utiliser d'autres moyens pour déterminer le VLAN d'appartenance, en inspectant les données comme l'adresse MAC ou le champ type (solutions propriétaires).
  - En l'absence d'un autre moyen, c'est le port de réception qui donnera l'appartenance de la trame à un VLAN: son PVID (*Port VLAN Identifier*)