

Ethernet Commuté

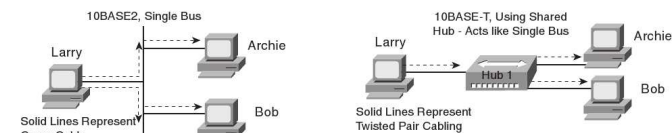
- Objectifs et plan:
 - Comprendre les évolutions d'Ethernet depuis la version historique
 - Abandon du coaxial, répéteurs multiports (*hubs*)
 - Ponts et commutateurs: principes, fonctionnement, algorithmes
 - Augmentation des débits
 - Généralisation du full-duplex, algorithme de contrôle de flot
 - VLAN : réseaux locaux virtuels

Bibliographie et sources

- Les cours de Stéphane Lohier <http://igm.univ-mlv.fr/~lohier>
- **CCNA ICND1 et CCENT**, 2^eed., Wendell Odom Pearson Education 2007.
- **Réseaux**, 4^{ème} éd. Andrew Tanenbaum, Pearson Education 2003.
- **Analyse structurée des réseaux**. Kurose & Ross, Pearson Education 2003.
- **Java et Internet** Roussel, Duris, Bedon et Forax. Vuibert 2002.
- **Les réseaux**, 3^{ème} éd. Guy Pujolle, Eyrolles 2000.
- **Interconnections**, 2nded. Radia Perlman. Addison Wesley 2000.
- **Réseaux haut débits**, T 1, 2^{ème} éd. D Kofman et M. Gagnaire, Dunod 1999.
- **Guide Pratique des Réseaux Ethernet**. Charles Spurgeon, Vuibert 1998.
- **Les réseaux locaux virtuels**. Gilbert Held, InterEditions 1998.
- **Gigabit Networking**. Craig Partridge. Addison Wesley 1994.

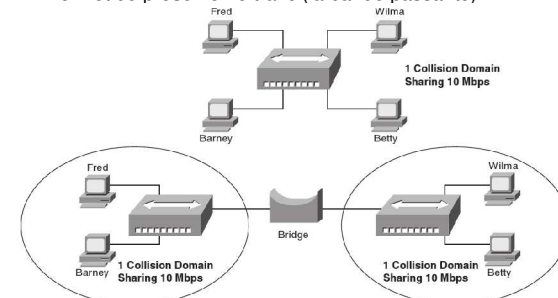
De coaxial vers répéteur multi-ports (hub)

- L'architecture en bus évolue vers une architecture en étoile
 - Les paires torsadées sont plus pratiques à manipuler
 - Le CSMA/CD est toujours en vigueur
 - Le répéteur (*hub*) joue le rôle de « support partagé »
 - À chaque extrémité d'un câble, on « prétend » qu'il y a collision si on émet en même temps qu'on reçoit
 - Même s'il n'y a pas de collision physique
 - La collision a en réalité lieu au niveau du *hub*



Séparation des domaines de collisions

- Un pont (*bridge*)
 - Permet de sortir du domaine de collision (CSMA/CD)
 - Permet de préserver le trafic (la bande passante)



Les ponts (*bridge*) IEEE 802.1 D (couche 2)

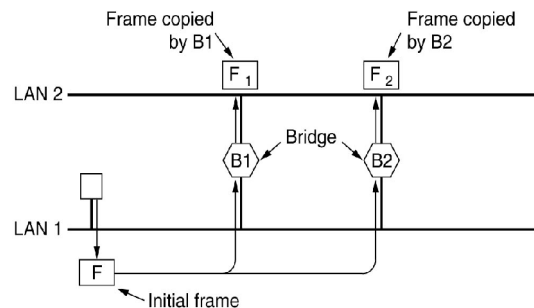
- Permet de déplacer des trames d'un domaine de collision à un autre
- **Sépare** différents **domaines de collisions**
- Ponts transparents – le plus courant
 - Supports différents, mais même protocole
 - Optimise le trafic en apprenant où sont les machines
 - Apprentissage (*learning bridge*)
- Ponts à traduction (*translation*) – plus rare
 - Protocoles différents: modifient l'encapsulation

Principe général d'un pont transparent

- **Éviter de transmettre les trames** là où elles sont **inutiles**
 - Utiliser pour cela les adresses MAC
 - Apprendre les **adresses MAC des émetteurs** reliés à chacun des port du pont (table d'apprentissage)
 - **Aiguiller** (*switch*) les trames vers le « bon » réseau sur la base de cette table d'apprentissage
 - Lorsqu'une trame est reçue sur un port P
 - Apprentissage de l'adresse MAC source SA: associée à P
 - Étude de l'adresse MAC destination DA
 - Si DA associée au port P détruire la trame: c'est le filtrage (*discard*)
 - Si DA associée à un port D, transmission vers D seulement (aiguillage)
 - Si DA n'est pas dans la table, transmission vers tous sauf P (*flooding*)
 - Toute trame incomplète, incorrecte ou ayant subi une collision est **détruite**

Ponts transparents en parallèle

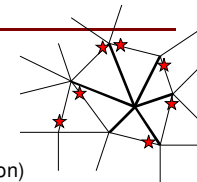
- Il est interdit d'avoir des boucles



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 la diffusion)
 - Superposer au graphe des connexions réseau un graphe sans circuit, dit « arbre recouvrant ».
 - Trouver une racine (selon ≠ critères, @MAC minimum, etc.)
 - Calculer la distance du plus court chemin à la racine
 - Pour chaque sous-réseau, élire un pont « désigné » (le + près)
 - Pour chaque pont, choisir un port « racine » qui offre le meilleur chemin
 - Placer dans le *Spanning Tree* le port racine et quelques autres...



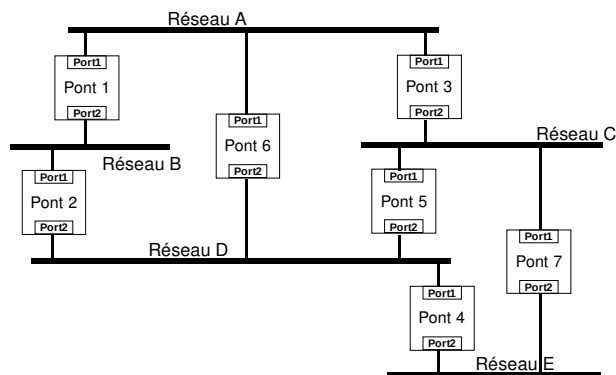
Spanning Tree Algorithm (principe)

- Communication entre ponts par transmission de BPDU (*Bridge Protocol Data Unit*)
 - Par multicast (01-80-C2-00-00-00) de trame LLC (SAP 0x42)
 - Les ponts sont identifiés par une valeur unique (ID) composée de 2 octets de priorité + 6 octets d'adresse MAC
- Au début, tout port se considère racine et émet dans ses BPDU
 - 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)
 - Cela 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 doit transmettre, il ne transmet plus sur ce port
- Stabilité: seul le pont "désigné" émet sur le LAN

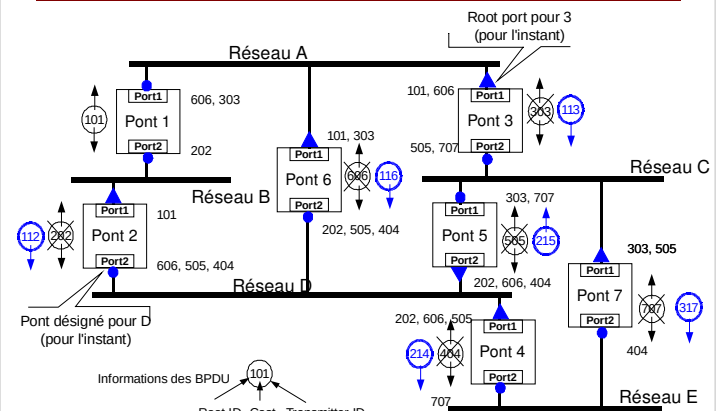
Choix des ponts et des ports

- 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
- Pendant le calcul: aucune trame de données n'est relayée
 - Un port bloqué ne relaye plus les trames de données

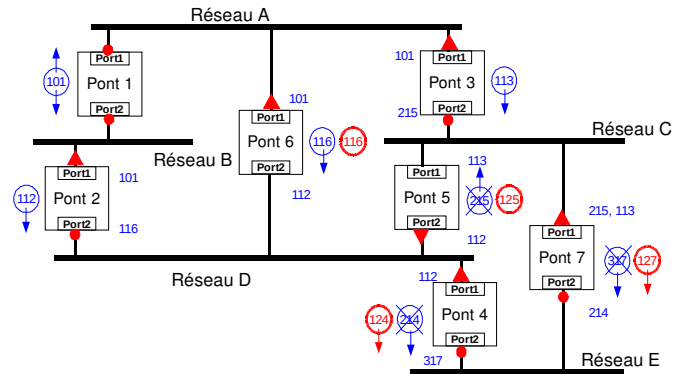
Exemple de réseau ponté (avant STA)



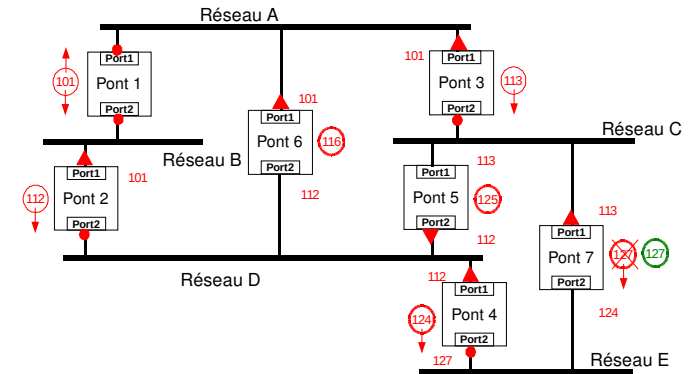
Exemple de réseau ponté (pendant STA) - (1)



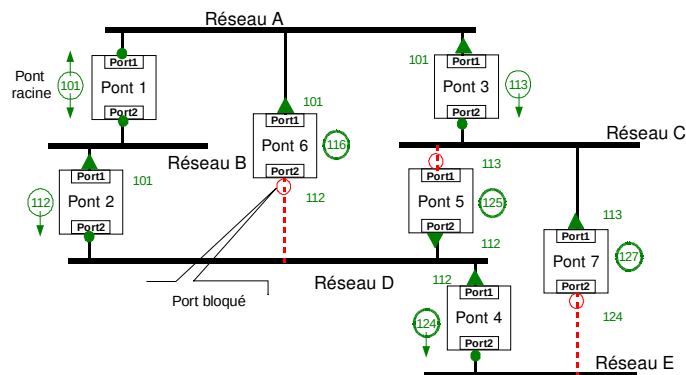
Exemple de réseau ponté (pendant STA) - (2)



Exemple de réseau ponté (pendant STA) - (3)

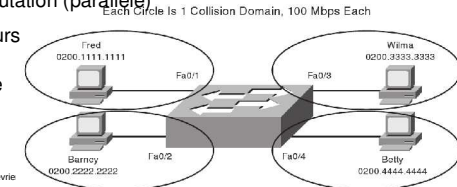


Exemple de réseau ponté (après STA)



Commutateur (Switch) (couche 2)

- 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
 - Basés sur les PABX de la téléphonie
- Hub intelligent (*smart*) *Intelligent Switching Hub*
 - Matrice de commutation (parallèle)
 - Autorisent plusieurs communications simultanées entre différents ports



Techniques de commutation

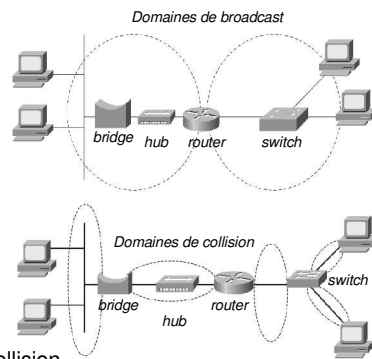
- À la volée (*cut-through*): correspond à la plupart des commutateurs / switchs actuels
 - On récupère l'adresse MAC destination
 - Une fois qu'on a identifié le port vers lequel on doit transmettre, on « commute » le flot vers ce port
 - Pas de correction ni d'assurance de qualité des trames
- Enregistrement & acheminement (*store & forward*): 802.1D
 - Ressemble plus au fonctionnement des ponts
 - La trame est reçue en entier avant d'être transmise
 - Si erreur, elle est jetée
- Adaptatif (mixage des 2 méthodes)

Utilisation et limitation des switchs

- Utilisations (exemples)
 - Isoler le trafic, allonger les distances, inter-connecter entre eux des protocoles différents
 - Créer des réseaux locaux virtuels (VLAN), ce qui permet d'améliorer la sécurité et la souplesse sans perdre de vitesse
- Limitations (« officielles »)
 - Pas plus de 7 « sauts » entre deux machines (délais, en 802.1D)
 - Dans la pratique, aucun problème si *cut-through*
 - Problèmes spécifiques en tant qu'équipements inter-protocoles
 - par exemple, tailles de trames maximales différentes

Domaines de broadcast versus collision

- Répéteurs (*hub*) étendent les domaines de collision
- Ponts (*bridge*) et commutateurs (*switch*) étendent les domaines de broadcast
- Routeurs délimitent les domaines de broadcast
- Chaque port d'un commutateur/pont délimite un domaine de collision



Evolutions d'Ethernet

- Augmentation du débit
 - 100Mbit/s **Fast Ethernet** 802.3u (1995)
 - 1000Mbit/s **Gigabit Ethernet** 802.3z (1998)
 - 100000Mbit/s **10Gbit/s** 802.3ae (2002)
- **Full duplex** et **contrôle de flot** (IEEE 802.3x, 1997)
 - Amélioration de la bande passante ($100 + 100 = 200\text{Mbit/s}$)
 - Suppression du principe CSMA/CD (plus de collision)
 - Gestion du droit de parole (goulets d'étranglement)
- Réseaux Locaux Virtuels (**VLAN**)
 - Souplesse d'utilisation, gain de bande passante et sécurité

Fast Ethernet - 100 Mbit/s (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)
- Fréquence élevée : Manchester remplacé par 4B/5B
- **Auto-négociation** 10/100 Mbit/s et full/half duplex
- Diamètre max. d'un domaine de collision divisé par 10
 - Débit multiplié par 10 et taille minimale de trame inchangée

Full duplex (802.3x)

- Communication simultanée possible dans les 2 sens
 - 10Base-T Full duplex = 20Mbit/s de bande passante
 - 100Base-T full duplex = 200Mbit/s bande passante
 - Propriété du support et "compétence" des cartes Ethernet
 - Ex: 10BaseT, 10BaseF, 100BaseX (fibre, paire torsadée)
 - Contre-exemple: 10Base5, 10Base2 (coaxial) mais aussi 100BaseT4
- Plus de collision: CSMA/CD ne s'applique plus
 - Plus de limite liée au délai d'aller et retour (*slot time*)
 - Exemple: un segment 100Base-FX en full duplex peut atteindre 2km, contre 412m en half duplex (MMF)
 - Problèmes pour gérer le droit de parole...

Contrôle de flot (802.3x)

- Problème: gérer les flots sur les liens full-duplex
 - MAC Control (trames spécifiques)
 - Transparent aux fonctions normales de contrôle
 - Remplace la régulation fournie par CSMA/CD
 - Permet aux stations d'interagir en temps réel
 - Valeur de type 0x8808 + « opcodes »
 - Adresse destination : 01-80-C2-00-00-01 (multicast)
 - Interprétées par la couche MAC
 - Support du système PAUSE (opcode 00-01)

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

Trame PAUSE

- Spécifique (et optionnel) pour full duplex
 - Permet de suspendre ponctuellement l'émission.
 - Les données qui accompagnent l'op-code 00-01 représentent un 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.

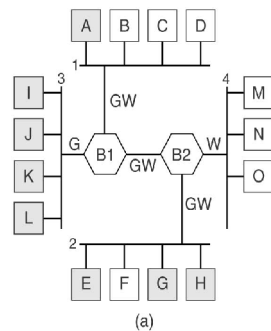
Gigabit Ethernet (802.3z)

- Supports et codages inspirés de Fibre Channel
 - 1000Base-X
 - 1000Base-LX (fibre, grande long. d'onde): 3km SMF, 400/500m MMF
 - 1000Base-SX (fibre, long. d'onde courte): 200/500m en MMF
 - 1000Base-CX (2 paires de câble cuivré blindé 150 Ohm): 25m
 - 1000Base-T (IEEE 803.2ab 1999) 4 paires UTP Cat 5: 100m
 - 4B/5B remplacé par 8B/10B, auto-négociation étendue
 - Diamètre max. du dom. collision encore div/10 si CSMA/CD?
 - Non: mécanismes ad hoc (*carrier extension/frame bursting*)
 - Nette préférence pour Full Duplex
 - En 10Gb, full duplex obligatoire

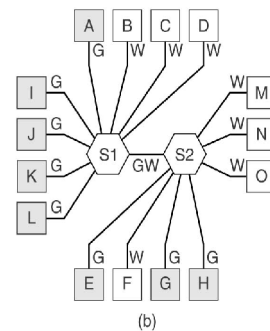
Réseaux Locaux Virtuels (VLAN)

- 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



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

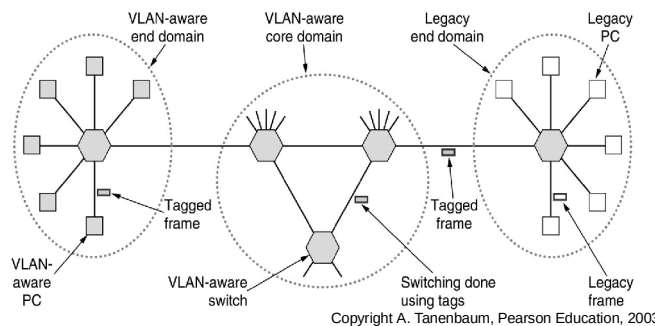
- Niveau 1: *VLAN par port*
 - L'administrateur associe un VLAN à chaque port
 - La plupart des configurations de commutateurs sont réalisées sur la base de VLAN par ports
- Niveau 2: *VLAN par adresse MAC*
 - L'appartenance d'une trame à un VLAN est déterminée par l'adresse MAC de l'émetteur et / ou 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

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
- A l'intérieur d'un switch, n'importe quelle solution peut faire l'affaire pour identifier les VLANs et les affecter aux trames
- Entre différents switches, il faut assurer une **cohérence** entre les configurations de VLANs
 - Les switches ajoutent des « étiquettes » (*tags*) aux trames
 - Ces tags permettent d'aiguiller les trames entre switches
 - Ils peuvent être retirés avant de délivrer les trames 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

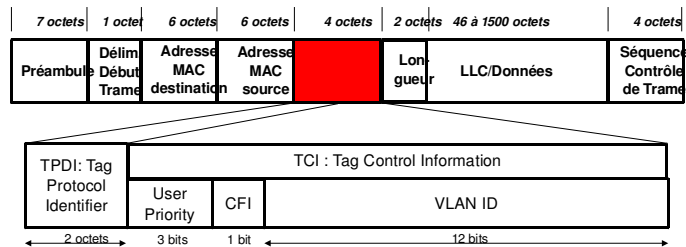


Extension des trames

- 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)