Protocole et norme Ethernet *de base*Eléments d'interconnexion d'Ethernet
Evolutions d'Ethernet vers très hauts débits
Commutation Ethernet
Spanning Tree

- Norme de réseau local
- Développé par Digital et Xerox
 - à partir de 1972
 - normalisé IEEE802.3 en 1980
- Basé sur le protocole CSMA/CD
- Standard du marché du réseau local
- Technologie simple et modulaire
 - base de son succès historique

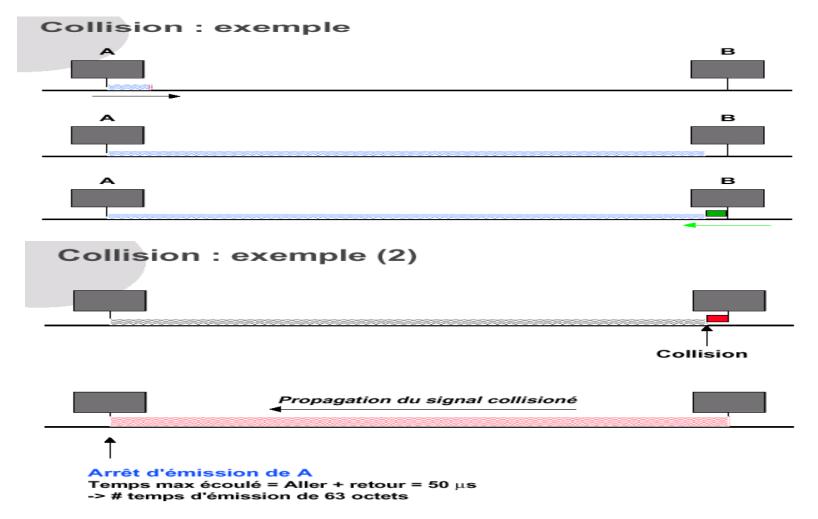
Evolution des débits Ethernet

Année	Technologie	Débit	Données/heure
1970	Ethernet	10 Mb/s	4,5 Go/h
1993	Fast Ethernet	100 Mb/s	45 Go/h
1998	Gigabit Ethernet	1 000 Mb/s (1 Gb/s)	450 Go/h
2002	10 Gigabit Ethernet (10 GbE)	10 000 Mb/s (10 Gb/s)	4.5 To/h
2010	100 Gigabit Ethernet (100 GbE)	100 000 Mb/s (100 Gb/s)	45 To/h
2017	400 Gigabit Ethernet (400 GbE)	400 000 Mb/s (400 Gb/s)	180 To/h

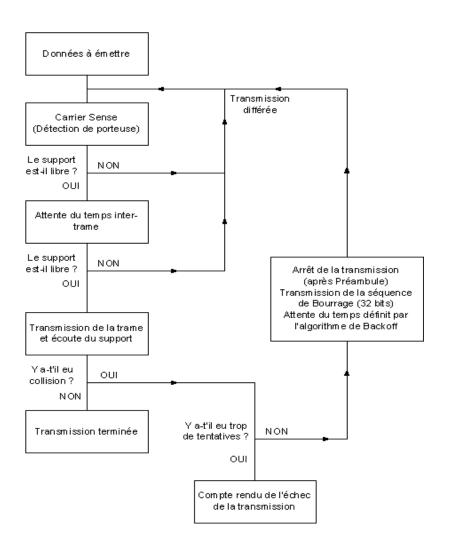
CSMA/CD

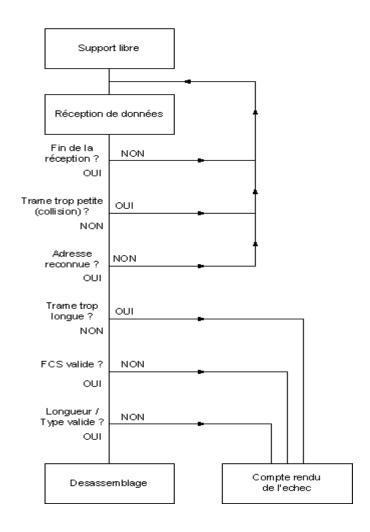
- Carrier Sense Multiple Access with Collision Detection
- Adapté au mode fonctionnement Half Duplex
- Accès multiples et écoute de la porteuse, avec détection de collision
- Principe:
 - une station regarde si le câble est libre avant d'émettre
 - mais le délai de propagation d'une trame sur le support n'est pas nul
 - une station peut émettre alors qu'une autre a déjà commencé à émettre
 - quand ces 2 trames émises presque simultanément se rencontrent
 - il y a **collision**
 - la conséquence d'une collision est une trame trop courte (< 72 octets)
- Sur un réseau d'une certaine taille ...
 - le délai de propagation peut être important
- ... comprenant de nombreuses stations
 - beaucoup d'émissions
- Après une collision il y a arrêt d'émission puis réémission
 - temps d'attente basé sur un nombre aléatoire x slot time (51.2 us)
- Le taux de collisions peut devenir bloquant :
 - 15 réémissions maximum pour une trame, après arrêt de réémission et erreur réseau

Ethernet et CSMA/CD

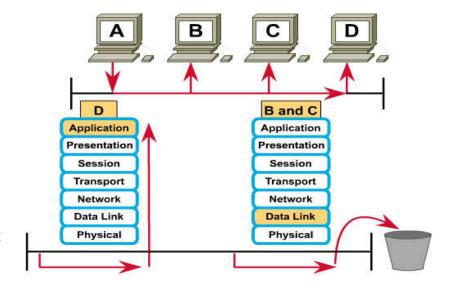


Ethernet et CSMA/CD émission réception

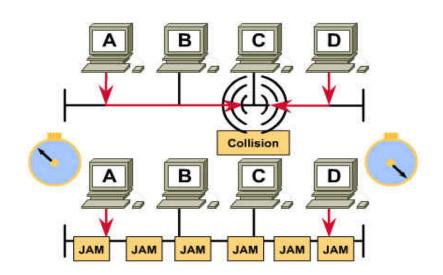




- Une station qui veut émettre
 - regarde si le câble est libre
 - si oui, elle envoie sa trame
 - tout le monde en a connaissance
 - seul le destinataire exploite la trame
 - si non, elle attend que le câble soit libre



- Si deux stations émettent en même temps
- => collision
 - les deux trames sont inexploitables
 - les deux stations détectent la collision
 - réémission ultérieure



Ethernet et CSMA/CD

- Inter Frame Gap (ou Space)
 - Ethernet précise un temps de silence entre 2 transmissions de trames
 - Permet aux équipements électroniques et aux médias de retrouver l'état de repos

- IFS minimum de 96 bits (12 octets)
 - 9,6 μs en Ethernet 10 Mb/s
 - 960 ns en Fast Ethernet
 - 96 ns en Giga Ethernet

- Buts initiaux (cahier des charges)
 - simple, faible coût
 - peu de fonctions optionnelles
 - pas de notion de priorité (pas adapté au MM)
 - réseau égalitaire (pas de maître)
 - débit de 10 Mb/s
- Non Buts initiaux (cahier des charges)
 - full duplex (lié au point à point)
 - contrôle d'erreur
 - sécurité et confidentialité
 - vitesse variable
 - priorité
 - protection contre un utilisateur malveillant

• Support de transmission historique

- segment = bus = câble coaxial (dans la 1ère version)
- pas de boucle, pas de sens de circulation
- diffusion
- passif

• Transceiver = transmitter + receiver

- équipement raccordé au câble par un transceiver
- Adresse unique au monde pour un équipement
 - racine imposée pour chaque constructeur

Trames

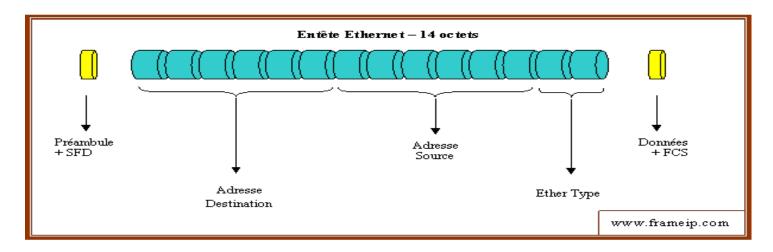
- sur le câble circulent des trames (suite d'éléments binaires = trains de bits)
- à un instant T, une seule trame circule sur le câble
- toute trame émise est reçue par tous les transceivers du segment Ethernet
- une trame contient l'adresse de l'émetteur et l'adresse du destinataire

• Ethernet est un réseau

- probabiliste
- sans chef d 'orchestre
- égalitaire
- Comparaison avec une réunion sans animateur entre gens polis

• Format d'une trame Ethernet II utilisée dans les réseaux IP :





Longueur des trames

- 14 octets utiles réservés au protocole + 8 octets de préambule + 4 octets de FCS
- champ de données variant de 46 (minimum) à 1500 (maximum) octets : un bourrage peut être nécessaire si les données sont inférieures à 46 octets afin de garantir la validité de la trame
- longueur minimale (sans préambule) : 64 octets (46 octets de données)
- longueur maximale (sans préambule) : 1518 octets (1500 octets de données)

Trame Ethernet

Préambule

- Ce champ est codé sur 7 octets et permet de synchroniser l'envoi.
- Chacun des octets vaut 10101010 et cette série permet à la carte réceptrice de synchroniser son horloge.

• SFD

- Ce champ est codé sur 1 octets et indique à la carte réceptrice que le début de la trame va commencer.
- La valeur de SFD (Starting Frame Delimiter) est 10101011.

FCS

- Ce champ est codé sur 4 octets et représente la séquence de contrôle de trame.
- Il permet à l'adaptateur qui réceptionnera cette trame de détecter toute erreur pouvant s'être glissée au sein de la trame.

Trame Ethernet

- Adresse destination et adresse source :
 - ce champ est codé sur 6 octets et représente l'adresse MAC (Medium Access Control) de l'adaptateur destinataire ou émetteur.
 - cette adresse est ce que l'on appelle l'adresse physique d'une carte Ethernet (Hardware address) ou adresse MAC (medium Access Control).
 - cette adresse est divisée en deux parties égales :
 - Les trois premiers octets désignent le constructeur. C'est l'IEEE, qui référence ces préfixes.
 - Les trois derniers octets désignent le numéro d'identifiant de la carte, dont la valeur est laissée à l'initiative du constructeur qui possède le préfixe.
 - ce principe d'adressage assure ainsi l'unicité de l'attribution des numéros d'adresse MAC.

Trame Ethernet

- Champ Code Protocole ou Ether Type
 - Ce champ est codé sur 2 octets et indique le type de protocole inséré dans le champ donnée.
 - Voici quelques exemples de codes protocoles :
 - 0x6000 DEC
 - 0x0600 XNS
 - 0x0800 IPv4
 - 0x0806 ARP
 - 0x8035 RARP
 - 0x809B AppleTalk
 - 0x8100 802.1Q
 - 0x86DD IPv6

Adresses Ethernet: particularités

- Adresse source = adresse physique de la station émettrice
 - 1er bit transmis = 0 => adresse d'une station (obligatoire)
- Adresse unique et non programmable = 6 octets (48 bits)
 - notation hexadécimale suivant équipementiers :
 - 08:00:20:06:d4:e8
 - 0800.2066.d4e8
 - 8:0:20:66:d4:e8
- Certaines cartes permettent de changer l'adresse MAC
 - peut poser des problèmes
 - duplication intempestive => problème de fonctionnement
 - usurpation => problème de sécurité

Adresses Ethernet: particularités

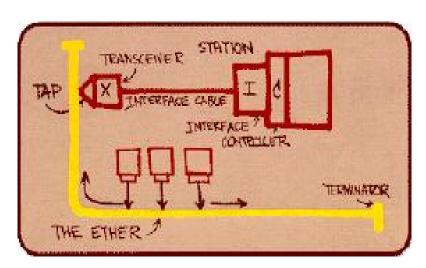
- Adresse destinataire
 - 1er bit transmis
 - = 0 adresse d'une station unique
 - = 1 adresse d'un groupe de stations (multicast)
 - tous les bits à 1 :
 - adresse de broadcast => toutes les stations du réseau sont concernées
- - toutes les stations de tous les segments d'un réseau
- Multicast = désigne un groupe de station
 - applications spécifiques, exemple :
 - management d'équipements
 - téléchargement de gros volumes de données

Ethernet (types de base : pour l'histoire)

- Dans l'ordre d'apparition :
 - 10 Base 5
 - Câble coaxial gros
 - 10 Base 2
 - Câble coaxial fin
 - − 10 Base T



- − 10 Base F
 - Fibre optique



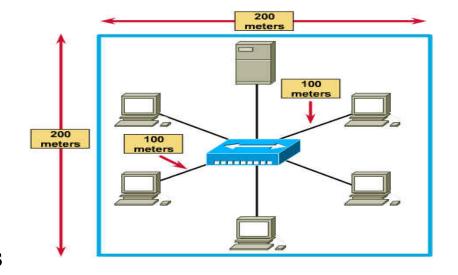
Ethernet: interconnexion

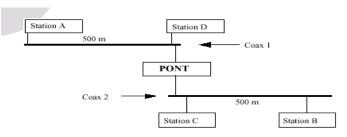
Hub = répéteur / concentrateur et Pont

- Répéteur et Hub (répéteur / concentrateur) :
 - remise en forme du signal bit par bit
 - pas d'action sur le contenu des trames
 - remplacés par des commutateurs dans les réseaux modernes

• Pont :

 Principe de fonctionnement des ponts repris par les commutateurs des réseaux modernes





- Les trames A <---> D ne vont pas sur Coax 2
- Les trames C <---> B ne vont pas sur Coax 1
- · ---> il faut que le Pont sache où sont A, B, C, D

Fast Ethernet

- Topologie:
 - étoile,
 - bus logique à diffusion recréé par les concentrateurs, comme Ethernet 10 Base T
 - longueur branches (câble cuivre) : 100 m
- Supports:
 - 100 Base TX : 2 paires torsadées UTP/FTP cat. 5
 - 100 Base T4: 4 paires torsadées UTP cat. 3, 4
 - 100 Base FX : 1 paire fibre optique
- Débit :
 - 100 Mb/s théorique
- Norme: 802.3u
- Méthode d'accès : CSMA/CD
- Trame standard Ethernet conservée : 64 à 1518 octets
- Full-Duplex : norme 802.3x
 - apparaît avec Fast Ethernet
 - généralisé depuis avec la commutation
- Autonégociation valable pour les supports cuivres uniquement
- Interconnexion
 - pas plus de 2 répéteurs par domaines de collisions => longueur maximum = 200 m
 - fibre optique allonge ces distances (> km voire +> kms)

Gigabit Ethernet

- Topologie:
 - étoile (point à point), uniquement disponible en commutation
- Supports :
 - 1000 Base T : 4 paires torsadées cat. 5e, 6, 7 (100 m)
 - 1000 Base SX et ZX : 1 paire fibre optique multimode (courte et longue distance)
 - 1000 Base LX et LH: 1 paire fibre optique monomode (courte et longue distance)
- Mécanisme de full-duplex généralisé
- Débit :
 - 1000 Mb/s théorique
 - débit pratique proche du théorique car utilisé uniquement en commutation
- Norme: 802.3z et 802.3ab
- Méthode d'accès : CSMA/CD
- Trame standard Ethernet conservée : 64 à 1518 octets
- Autonégociation valable pour les supports cuivres uniquement
- Migration et interconnexion :
 - cartes et commutateurs 10/100/1000 avec ou sans autonégociation
- Utilisation
 - au départ construction d'artères centrales (backbones) et raccordement de serveurs
 - Aujourd'hui raccordement généralisé des stations (coûts en baisse constante)
- Avantages
 - conservation de la technologie et de l'expérience Ethernet
- Inconvénients
 - pas de qualité ni de priorité de service native

10 Gigabit Ethernet

- Topologie:
 - étoile (point à point) uniquement disponible en commutation
- Supports:
 - 10 Gbase-SR: 300 m fibre multimode
 - 10 Gbase-LR et ER: 40 kms fibre monomode
 - 10 Gbase-T : 100 m câble cat 6 ou 7 (4 paires)
- Débit :
 - 10 000 Mb/s théorique
 - débit pratique proche du théorique car utilisé uniquement en commutation et full duplex
- Norme principale: 802.3ae
- Méthode d'accès : CSMA/CD maintenu pour compatibilité
- Trame standard Ethernet conservée : 64 à 1518 octets
- Migration et interconnexion
 - avec 10 /100 /1000 via des commutateurs possédant tous types d'interfaces
 - cartes coupleurs spécifiques sans auto-négociation
- Utilisation
 - construction d'artères centrales de réseaux locaux ou de campus (backbones)
 - raccordement de serveurs de calcul ou de stockage
- Avantages
 - conservation de la technologie et de l'expérience Ethernet
- Inconvénients
 - pas de qualité ni de priorité de service native

100 Gigabit Ethernet

- Topologie :
 - étoile (point à point)
 - uniquement disponible en commutation
- Supports et distances :
 - 100 m fibre multimode
 - 40 kms fibre monomode
 - 10 m câble cat 7 (4 paires)
- Débit :
 - 100 000 Mb/s théorique
 - uniquement en full duplex
- Débit intermédiaire de 40 Gb/s également normalisée
- Norme principale : 802.3ba normalisée fin 2010 (évolutions depuis)
- Méthode d'accès : CSMA/CD à priori pas utile mais conservé
- Trame: 64 à 1518 octets
 - la trame standard Ethernet est conservée

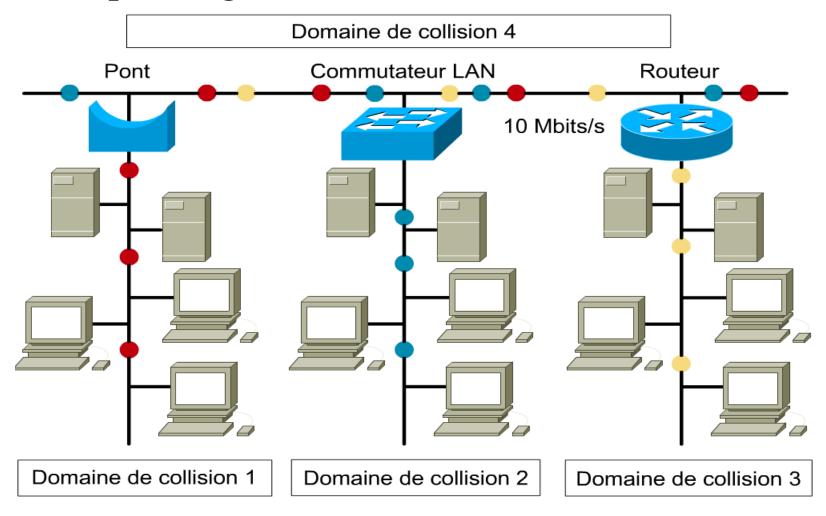
200 et 400 Gigabit Ethernet

- Topologie :
 - étoile (point à point)
 - uniquement disponible en commutation
- Supports et distances :
 - câble de cuivre twinax uniquement pour 200 Gbs (3 m)
 - fibre multimode (100 m) et monomode (500 m à 2000 m voire 10000 m)
- Débit :
 - 200 000 Mb/s ou 400 000 Mb/s théorique
 - uniquement en full duplex
- Débits intermédiaires de 25 Gb/s et 200 Gb/s également proposés
- Norme principale : 802.3bs validée en 2017
- Méthode d'accès : CSMA/CD à priori pas utile mais conservé
- Trame: 64 à 1518 octets
 - la trame standard Ethernet est conservée

800 Gigabit Ethernet et Terabit Ethernet en projet à horizon 2025

Problèmes d'Ethernet partagé

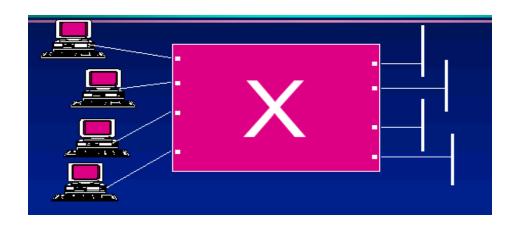
Pourquoi segmenter un réseau local:



Commutation Ethernet

- Accroissement important des besoins en débit
 - évolution des applications
 - nouveaux services
 - habitudes des utilisateurs
 - augmentation du parc informatique
- Problèmes à résoudre sur le réseau local
 - charge
 - collisions
 - broadcast
- Réponse par les technologies classiques et la segmentation
 - ponts : technologie dépassée
 - routeurs : traitement ralenti le transit des trames
- Réponse par les nouvelles technologies
 - augmentation de débit (FastEthernet, GigaEthernet)
 - si partagé pas encore satisfaisant
- Objectifs
 - rapprocher le débit pratique du débit théorique
 - diminuer voire éliminer les collisions
- C'est la **commutation**

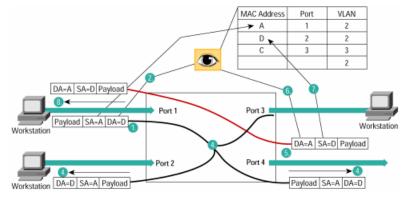
Commutation: principes



- Fonctionnement type « ponts »
- Ports avec bande passante dédiée et non partagée
- Communication port à port, parallèles entre elles
- Liens point à point créés à la demande
- Conservation de l'infrastructure Ethernet de base

Commutation: chemins

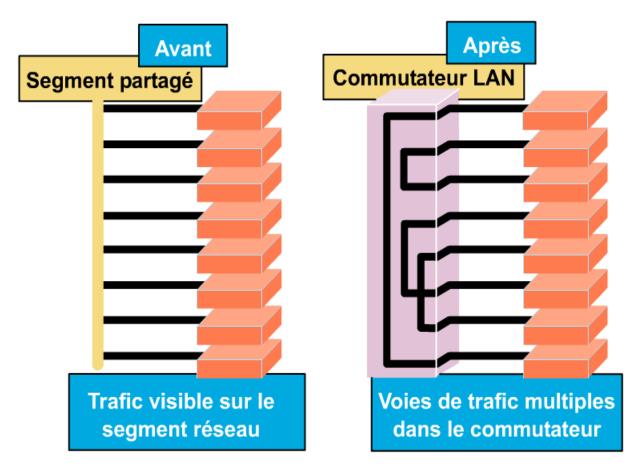
Création de tables de commutation



- Mise à jour des tables de commutation
 - Signalisation implicite ou commutation dynamique
 - le plus couramment utilisé
 - vérification et ajout d'une ligne dans la table de commutation à chaque passage d'une nouvelle trame Ethernet
 - recherche par diffusion de la destination
 - commutateur ouvre un *chemin virtuel* en fonction de l'adresse Ethernet de destination
 - Signalisation explicite ou commutation statique
 - définition d'une matrice figée permettant un chemin figé
 - très peu employé, pas de souplesse
 - ports définis en relation permanente

Micro Segmentation et Commutation

Routes dédiées entre les hôtes émetteurs et récepteurs



Apprentissage des adresses

- Pour fonctionner un commutateur doit connaître les adresses MAC (Ethernet) de tous les ordinateurs connectés au réseau.
- A partir de cette connaissance le commutateur peut construire la topologie du réseau.
- Il existe deux façons de connaître les adresses MAC :
 - configuration statique
 - auto apprentissage (learning mode)
- Apprentissage automatique
 - Un commutateur Ethernet peut apprendre les adresses de chaque dispositif du réseau en lisant les entêtes de chaque trame Ethernet transmise, en relevant les adresses source et en les associant au port du commutateur sur lequel la trame est entrée.
- L'adresse apprise est stockée dans une mémoire appelée CAM:
 - Content Addressable Memory

Apprentissage des adresses

- Les adresses sont donc apprises de manière dynamique.
- A mesure que le commutateur lit de nouvelles adresses, il les apprend et les stocke dans la CAM.
- Même si une trame est détruite, l'adresse source est stockée pour usage futur éventuellement.
- A chaque nouvelle entrée dans la CAM un marqueur de temps est associé à l'adresse nouvellement apprise.
- Cela permet de garder les adresses durant un temps limité (en général configurable).
- Chaque fois que le commutateur relit une adresse source déjà stockée il met à jour le marqueur de temps associé à cette adresse.
 - Exemple :





Marqueur de temps décrémenté

Trame relue => marqueur de temps mis à jour

Apprentissage des adresses

 Les adresses dont le marqueur de temps arrive à expiration sont éliminées de la mémoire.

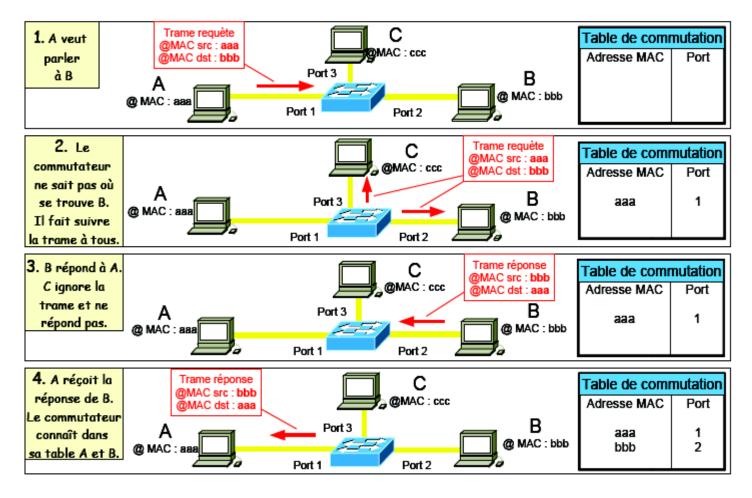




Le port 3 a une information plus récente que le port 1 car son TTL est plus élevé 91s plus tard la machine 23 n'a pas parlé (ni la machine 25) la table CAM évolue

- Ce système permet de maintenir une base de données des adresses et une topologie du réseau à jour et fonctionnelle.
- Les commutateurs évolués permettent de modifier tous ces paramètres gérant leur fonctionnement pour l'adapter à tout type de situation et de réseau.

Commutation : exemple de fonctionnement



• Les tables de commutation sont volatiles pour permettre les changements dans un réseau (topologie, machines, etc)

Commutation: méthodes

- A la volée ou on the fly ou cut through
 - arrivée de la trame
 - lecture des 1 ers octets de la trame (entête)
 - commute la trame vers le destinataire en fonction de l'adresse de destination

Avantages

- temps de latence très faible
 - $< 1 \mu s$
 - indépendant de la longueur de la trame

Inconvénients

- retransmission des erreurs (CRC, fragments de collision)
- pas adaptée aux variations de débits (10/100 => bufferisation)

Commutation: méthodes

Store & forward

- arrivée de la trame
- stockage puis analyse de la trame
- commutation vers le destinataire en fonction de l'adresse de destination

Avantages

- traitement des erreurs
- possibilité de traitement plus fin
- adaptée aux variations de débits (10/100 => bufferisation)

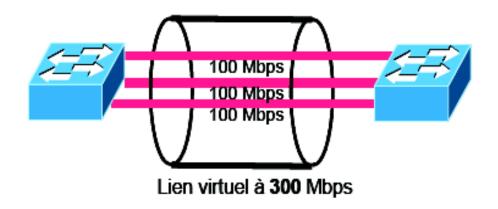
• Inconvénients

- plus lent que la commutation à la volée
- temps de latence fonction de la longueur de la trame

Méthodes de commutation

- Avantages et désavantages des différentes méthodes :
 - La latence induite par chaque mode de commutation dépend de la manière selon laquelle le commutateur traite les trames.
 - Plus le mode de commutation est rapide, moins la latence sera importante.
 - Pour gagner en rapidité d'envoi des trames la seule solution est de réduire le temps de vérification des erreurs.
 - A contrario le manque de vérification des erreurs peut générer une quantité importante de retransmissions inutiles.

Commutation: apport qualitatif

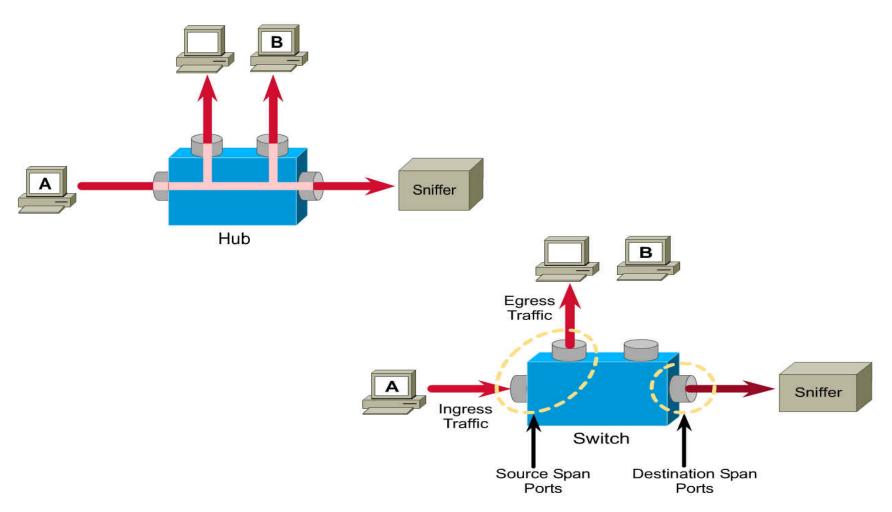


• Agrégation de liens :

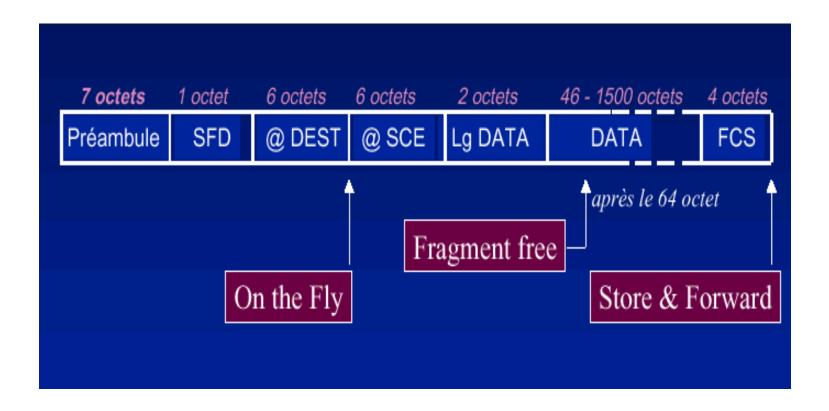
- Permet à plusieurs liens physiques d'être vus comme un seul
- Permet de palier à la saturation (augmentation du débit)
- Assure la redondance en cas de coupure d'un des liens
 - valable autant pour les équipements de réseau ...
 - ... que pour les serveurs ...
 - ... si les systèmes implémentent ce protocole!

Commutation: introduction

• Commutation Ethernet : meilleure sécurité ...



Commutation: récapitulatif



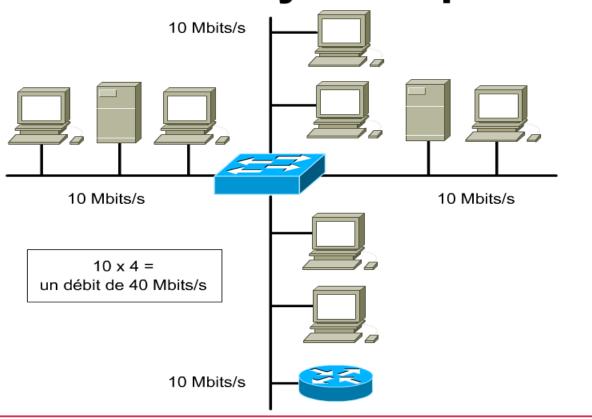
• Commutation = niveau 2 du modèle ISO

Gestion des buffers

- Un commutateur Ethernet peut utiliser une technique de buffering pour stocker les paquets avant de les envoyer via le bon port.
- Le buffering peut aussi être utilisé quand le port de destination est occupé afin de ne pas générer de collision inutile.
- La zone tampon dans laquelle le commutateur stocke les données est appelé un « buffer de mémoire »
- Le commutateur conserve une carte des ports de destination auxquels sont envoyés les trames en cas de réémission immédiate.
- Le commutateur efface cette carte uniquement après une transmission réussie.
- Cette gestion des buffers est importante pour la commutation 10/100, où un port à 100 Mbs peut envoyer des trames à un port à 10Mbs.
- Ce fonctionnement, appelé asymétrique, est important dans un réseau.

Gestion des buffers

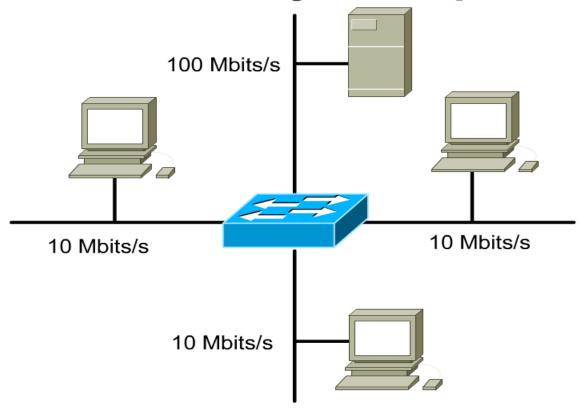
Commutation symétrique



- Assure la commutation entre des bandes passantes similaires (10/10 ou 100/100 Mbits/s)
- Plusieurs communications simultanées augmentent le débit.

Gestion des buffers

Commutation asymétrique



- Assure la commutation entre des largeurs de bande différentes (10/100 Mbits/s)
- ◆ Le commutateur doit utiliser la mise en mémoire tampon.

Commutation: récapitulatif

- Commutation permet de diminuer voire d'éviter les collisions ...
- ... Ethernet implémente le full-duplex
- Fonctionnement global du réseau grandement amélioré
- Trafic émis et reçu n'est pas propagé sur tous les ports
 - aspect très important pour la sécurité
 - il devient très difficile d'espionner tout le réseau (sniffing)
 - visibilité limitée sauf à recopier le trafic (mirroring)
- L'idéal:
 - commutation systématique jusqu'au poste de travail
- Généralisation de la commutation
- Au cœur des réseaux modernes (locaux ou étendus) à hauts et très hauts débits

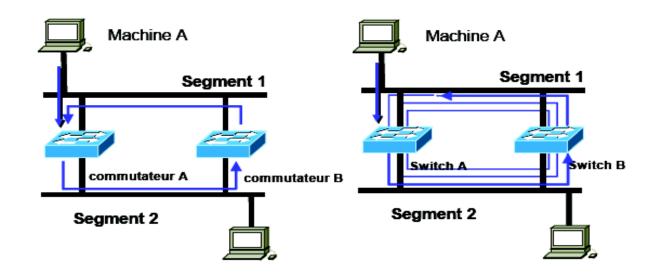
Commutateurs : critères de choix

- Bus fond de panier à très haut débit (> n Gb/s)
- Commutation matérielle :
 - processeurs spécialisés (ASIC) => gain en performance
- Débit de traitement : plusieurs millions trames/s
- Veiller au temps de latence très faible ($< 1 \mu s$)
- Niveau 3 possible => routage IP
- Respect du format des trames Ethernet_II et 802.3
- Attention à la taille des buffers : entrée, sortie, partagés
- Nombre d'adresses Ethernet gérées (par port et au total)
- Mécanisme de contrôle de congestion
- Bande passante globale :
 - N/2 x d Mb/s pour N ports (d=10, 100, 1000)
- Tolérance aux pannes, redondance, réparation à chaud (hotplug)
- Conformité à SNMP : logiciels d'administration dédiés
- Richesse du système d'exploitation : matériel hautement administrable
- Gestion des Réseaux Virtuels (Vlans)
- Gestion de la Qualité de Service (QoS)
- Présence de ports ATM, FDDI, Gigabit, 10 Gb/s, 100 Gb/s, ...

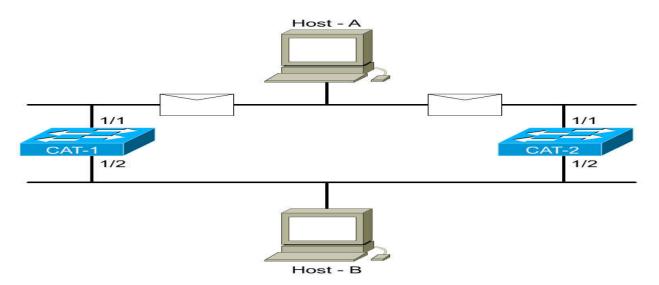
Commutation: conclusion

- Très bien adapté au trafic unicast
- 1er niveau de sécurisation d'un réseau
- Problèmes possibles avec le broadcast
 - tempêtes de broadcast
 - paralysie du réseau par saturation
 - solutions :
 - confiner le broadcast en domaines de collision
 - création de réseaux logiques
- Réseaux logiques sur une même infrastructure physique : VLAN
 - concept au cœur des réseaux modernes

- Description générale du protocole :
 - Toujours en tenant compte qu'Ethernet est un bus logique et non un anneau.
 - La fonction principale du Spanning Tree est de permettre la présence de routes (commutées ou pontées) dupliquées sans engendrer les effets de latence induits par les boucles sur un réseau.



- Les normes du Spanning Tree :
 - IEEE 802.1d, 802.1s et 802.1w
- Le Spanning Tree est transparent pour les machines du réseau.
- Le protocole utilisé est le BPDU :
 - Bridge Protocol Data Unit



• Rappel:

- Les ponts et les commutateurs prennent leurs décisions d'acheminement des trames de type unicast uniquement sur la base de l'adresse MAC de destination de la trame.
- Si l'adresse MAC de destination est inconnue l'équipement achemine la trame à travers tous les ports (inondation) et espère qu'elle arrivera à destination.
- Ce même principe est appliqué pour toutes les trames de broadcast.

• L'algorithme de Spanning Tree :

- implémenté par le protocole Spanning Tree original, IEEE 802.1d, évite les boucles en calculant une topologie de réseau de Spanning Tree stable.
- En créant des réseaux tolérant aux fautes, une route sans boucle doit exister entre tous les nœuds d'un réseau Ethernet.

- L'algoritme du Spanning Tree est utilisé pour calculer cette route sans boucle.
- Les trames dédiées au Spanning Tree, sont envoyées et reçues par tous les commutateurs du réseau à intervalle régulier et servent à déterminer la topologie du Spanning Tree.
- Le protocole du Spanning Tree détecte et élimine les boucles en positionnant certaines connexions en mode *attente*, qui s'activeront en cas de défaillance d'une connexion active.

