

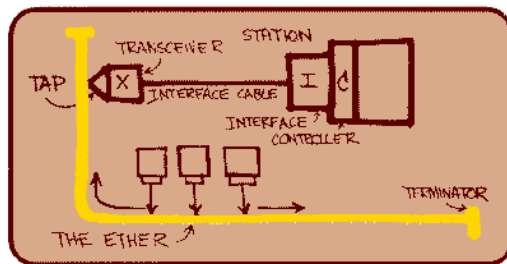


..... La couche Ethernet IEEE 802.3 09.12.2023

l'ancêtre du GPS®

C'est quoi l'Ether ?

Auteur : Pascal Fougeray



1 Préambule

Si vous voulez tout savoir sur Ethernet : <http://grouper.ieee.org/groups/802/3/>

2 Histoire

L'histoire de l'Ethernet remonte presque à celle de l'informatique et le protocole réseau est issu de travaux réalisés par Robert Metcalfe au Xerox Park avec des premiers au niveau des réseaux locaux. En 1976, Robert Metcalfe et David Boggs publient un document intitulé ***Ethernet : Distributed Packet-Switching For Local Computer Networks***.

Au début, Ethernet était en compétition avec 2 systèmes propriétaires, ***Token Ring*** d'IBM, et ***ARCnet*** de TRW-Matra. Mais au fil du temps, Ethernet a fini par s'imposer et a été standardisé sous le nom **IEEE 802.3**.

Le premier LAN Ethernet fut conçu au milieu des **années 1970** par **Robert Metcalfe** et son assistant **David Boggs**.

Le débit original était de 2,94 Mbps. **Robert Metcalfe** était un membre de la direction de recherche pour Xerox. Il travaillait au centre de recherche Palo Alto au USA (PARC : Palo Alto Research Center) où certains des premiers PC ont été construits. Robert Metcalfe quitta Xerox en 1979 pour promouvoir l'utilisation du PC et du LAN (***Local Areas Network***).

Il a réussi à convaincre les entreprises *Digital Equipment*, *Intel* et *Xerox* de travailler ensemble pour promouvoir **l'Ethernet comme un standard**.

Les premières versions d'Ethernet offraient un débit de 10 Mbits/s dans les années 1990, actuellement... 1,6 bits/s ...

<http://grouper.ieee.org/groups/802/3/dj/index.html>



3 Ethernet

La couche Ethernet est constituée de 2 sous couches :

1. La sous-couche Contrôle de la liaison logique (*Logical Link Control - LLC*)

La plus haute, elle multiplexe les protocoles fonctionnant au-dessus de la couche de liaison de données, et optionnellement, fournit le **contrôle de flux**, l'**accusé réception** et la **correction d'erreur**. Le LLC fournit l'adressage et le contrôle de la liaison de données. Il spécifie quels mécanismes doivent être utilisés pour adresser des stations sur le support de transmission et pour le contrôle de l'échange des données entre la machine de l'expéditeur et du destinataire.

2. La sous-couche de Contrôle d'accès au support (*Media Access Control - MAC*)

Les protocoles appartenant à cette sous-couche ont pour fonction de réguler les émissions sur un support donné, en particulier lorsque plusieurs stations indépendantes sont susceptibles d'émettre à tout moment sur le même support (tel qu'un bus Ethernet, habituellement géré par le protocole **CSMA/CD**). Appartiennent également à cette sous-couche la description des formats de trame (cellule élémentaire du transport d'information) et des méthodes de repérage des stations émettrices et réceptrices (**adressage**). Un des formats définis dans ce cadre utilise les **adresses** dites **MAC**, de **6 octets**, attribuées de manière (normalement) univoque à des stations émettrices/réceptrices (cartes), notamment celles utilisées par **ethernet**, **bluetooth** et **Wi-Fi**.

3.1 Les @ MAC

Aussi appelée @physique, est une séquence composée de chiffres et de lettres codée (Base 16!) sur 48 bits, soit 6 octets. Elle est présentée au format hexadécimal, en séparant les octets par un double point ou un tiret, exemple : **16 :64 :33 :51 :AB :CD**.

IL existe virtuellement 2^{48} adresses MAC disponibles ce qui fait plus de 281 000 milliards ...

Cette adresse MAC est découpée en 4 parties!

- 1 bit **I/G** : indique si l'adresse est
 - **Individuelle**, auquel cas le bit sera à **0** pour une machine unique on parle d'**unicast**
 - **Groupe** on parle de **multicast** ou **broadcast**, en passant le bit à **1**
- 1 bit **U/L** : indique
 - 0 si l'adresse est **Universelle** (conforme au format de l'IEEE)
 - **Locale**, 1 pour une adresse administrée localement
- 22 bits **réservés** : tous les bits sont à zéro pour une adresse locale, sinon ils contiennent l'adresse du **constructeur** ;
- 24 bits : adresse unique. Ils différencient les différentes cartes réseaux d'un même constructeur. L'IEEE donne des préfixes de 24 bits (appelés *Organizationally Unique Identifier - OUI*) aux fabricants, ce qui offre 2^{24} environ 16 millions d'adresses MAC disponibles par préfixe. **Un constructeur peut posséder plusieurs préfixes**

Remarque : Tous les périphériques connectés à un réseau ont une adresse MAC. En raison d'un conflit avec un autre matériel ou par souci de confidentialité, un utilisateur **peut souhaiter la changer**. Une commande réalisera une **correction temporaire** tandis qu'un fichier de configuration ou un logiciel dédié autorisera un changement durable. **Dans tous les cas, la modification ne sera pas physique mais logicielle.**

On ne peut pas changer l'adresse MAC d'une interface physique, on peut changer l'adresse MAC d'une interface virtuelle!

Quelques adresses particulières! Juste savoir pas à connaître ou juste la première!

FF :FF :FF :FF :FF :FF	Adresse broadcast
01 :00 :0C :CC :CC :CC	<i>Cisco Discovery Protocol</i>
01 :80 :C2 :00 :00 :00	<i>Spanning Tree Protocol</i>
33 :33 :xx :xx :xx :xx	Adresses <i>multicast</i> IPv6
01 :00 :5E :xx :xx :xx	Adresses <i>multicast</i> IPv4
00 :00 :0c :07 :ac :xx	Adresses HSRP
00 :00 :5E :00 :01 :XX	Adresses VRRP



Remarque : HSRP et VRRP sont 2 protocoles permettant à 2 routeurs d'être jumeaux ☺
 Nous verrons (peut-être...) cela dans la partie **roulage internet**

3.2 Le MTU

Lors d'une transmission de données, le MTU (*Maximum Transmission Unit*) est la taille maximale d'un paquet pouvant être transmis en une seule fois donc sans fragmentation sur une interface.

Sa taille varie selon le type d'interface et va de 1500 octets à 9216 octets pour les *jumbo frames*. On parle de norme 802.3. Si vous voulez tout savoir et bien tout est là !

<http://grouper.ieee.org/groups/802/3/>

Exemple de valeur de MTU selon le type de réseau

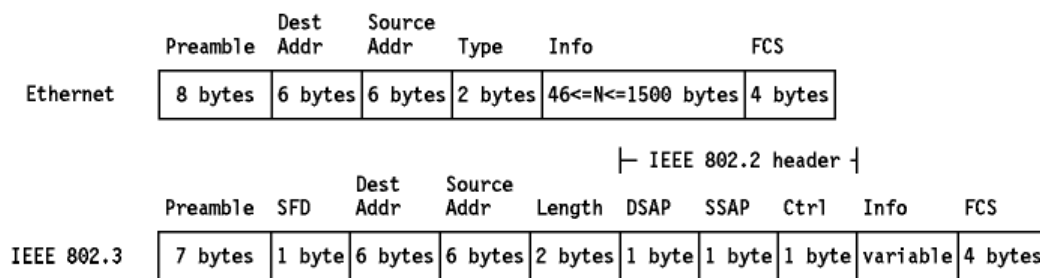
- IPv4 : le MTU minimal est de 68 octets
- IPv6 : le MTU minimal par défaut vaut 1 280 octets
- sur **Ethernet** :
 - MTU = 1 500 octets par défaut.
 - MTU = 1500 à 9000 octets avec Jumbo frame.
 - Loopback MTU 65536
- sur PPPoE : MTU = 1 492 octets. 1500 octets moins les 8 octets du PPPoE
- sur ADSL ou ATM25 : MTU = 1 468 octets.
- sur ADSL en Wi-Fi MTU = 1 468 octets.
- sur Token Ring, le MTU est de 4 500 octets.
- sur Frame Relay le MTU est de 1 564 octets.

3.3 En-tête ETHERNET

L'Ethernet est basé sur un principe de dialogue **sans connexion et donc sans fiabilité**. Les trames sont envoyées par l'adaptateur sans aucune procédure de type "*handshake*" avec le destinataire. Le service sans connexion d'Ethernet est également non-fiable, ce qui signifie qu'aucun acquittement, positif ou négatif, n'est émis lorsqu'une trame passe le contrôle CRC avec succès ou lorsque celle-ci échoue. Cette absence de fiabilité constitue sans doute la clé de la simplicité et des coûts modérés des systèmes Ethernet.

Ce service de couche 2 du modèle OSI est semblable au service en mode datagramme de couche 3 assuré par IP et au service sans connexion de couche 4 d'UDP.

Frame Formats for Ethernet and IEEE 802.3



<http://www.danzig.jct.ac.il/tcp-ip-lab/ibm-tutorial/3376c28.html>

Définition des différents champs

- **Préambule** : codé sur 7 octets, il permet de synchroniser l'envoi.
 Chacun des octets vaut 10101010 et cette série permet à la **carte réceptrice de synchroniser son horloge**.
- **SFD** : codé sur 1 octet, indique à la carte réceptrice que le début de la trame va commencer.
 La valeur de SFD (*Starting Frame Delimiter*) est 10101011.
- **Adresse source** : codé sur 6 octets c'est l'@ MAC de l'émetteur
- **Adresse destination** : codé sur 6 octets c'est l'@ MAC du receptrice
 Dans le cadre d'un *broadcast*, l'adresse utilisée est FF-FF-FF-FF-FF-FF.
- **EtherType** : codé sur 2 octets, indique le type de protocole inséré dans le champ donnée.

Quelques valeurs

0x0800	0x0806	0x86DD	0x8100	0x8847	0x8870			0xCAFE
IPv4	ARP	IPv6	VLAN	MPLS Unicast	Jumbo Frames			oui oui ^^

Jumbo Frames (Frames supérieures à 1500 octets pouvant aller jusqu'à 9216, utilisées par les ISP!!!)



- **DATA** : codé entre 46 et 1500 octets et **contient les données de la couche 3**.

Dans le cas de TCP/IP, c'est ici que vient se loger le datagramme IP. L'unité de transfert maximale est le MTU (*Maximale Transfer Unit*) et sa valeur est classiquement de 1500 octets.

Si la taille des données est **inférieure à 46 octets**, alors elle devra être complétée avec des octets de bourrage (**padding**) et c'est la couche réseau qui sera chargée de les éliminer.

Nous aurons l'occasion de le voir un peu plus loin dans ce cours lors de la capture avec Wireshark d'une trame ARP !

- **FCS** : codé sur 4 octets, il 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.

Les erreurs binaires sont principalement créées par les variations d'**affaiblissement du signal** et l'**induction électromagnétique parasite** dans les câbles Ethernet ou les cartes d'interface.

La valeur de **FCS** (*Frame Check Sequence*) est le résultat d'un calcul polynomial appelé **CRC** (*Cyclic Redundancy Code*).

À la réception de la trame, la couche liaison effectue le même calcul et compare les 2 résultats qui doivent être égaux afin de valider la conformité de la trame reçue.

Pour les réseaux locaux tels ceux que vous utilisez le polynôme utilisé est :

$$X^{32} + X^{26} + X^{23} + X^{22} + X^{16} + X^{12} + X^{10} + X^8 + X^7 + X^5 + X^4 + X^2 + 1$$

Remarque : Les 8 octets du préambule et du SFD ainsi que le FCS ne sont pas visibles avec Wireshark, parce qu'ils sont enlevés par la carte réseau !

C'est la carte réseau qui fait le calcul du CRC!!!

Voici sur l'image suivante 4 de trames Ethernet avec 4 ethertypes différents !

<i>Préambule</i>	<i>SPD</i>	<i>@ MAC Dest</i>	<i>@ MAC Src</i>	<i>EtherType</i>	<i>DATA : donc paquet IP</i>		<i>FCS</i>
------------------	------------	-------------------	------------------	------------------	------------------------------	--	------------

<i>Préambule</i>	<i>SPD</i>	<i>@ MAC Dest</i>	<i>@ MAC Src</i>	<i>EtherType</i>	<i>TAG</i>	<i>DATA : donc paquet IP</i>	<i>FCS</i>
------------------	------------	-------------------	------------------	------------------	------------	------------------------------	------------

<i>Préambule</i>	<i>SPD</i>	<i>@ MAC Dest</i>	<i>@ MAC Src</i>	<i>EtherType</i>	<i>PPPoE</i>	<i>DATA : donc paquet IP</i>	<i>FCS</i>
------------------	------------	-------------------	------------------	------------------	--------------	------------------------------	------------

<i>Préambule</i>	<i>SPD</i>	<i>@ MAC Dest</i>	<i>@ MAC Src</i>	<i>EtherType</i>	<i>SHIM</i>	<i>DATA : donc paquet IP</i>	<i>FCS</i>
------------------	------------	-------------------	------------------	------------------	-------------	------------------------------	------------

3.4 Le CSMA/CD

Je vais vous en parler rapidement,

CSMA/CD (*Carrier Sense Multiple Access / Collision Detection*) pour

- **CS** : **Écouter** le Canal
- **MA** : **Accès Multiple**
- **CD** : **Détecter** les collisions

Principe : Plusieurs stations peuvent tenter d'accéder simultanément au support, ce qui impose pour chaque station :

- L'écoute et la détection du signal sur le réseau avant d'émettre ses trames (Si silence, attente de 96 temps bits (IFG) avant émission → Délai de silence obligatoire entre 2 trames)
- Une fois sa trame émise, la station écoute le support pendant un temps : $t \geq 2 \cdot tp$ avec tp : temps de propagation jusqu'au point le plus éloigné du bus.

Au bout de ce temps, 2 cas peuvent se présenter :

1. la trame émise n'est pas altérée (pas de collision) => la station peut poursuivre sa transmission.
2. la station détecte une trame altérée (collision = tension moyenne double de la normale) => la station reprend la transmission de la trame suivant une procédure définie à l'avance (émission d'une séquence de brouillage + Attente pseudo aléatoire).

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

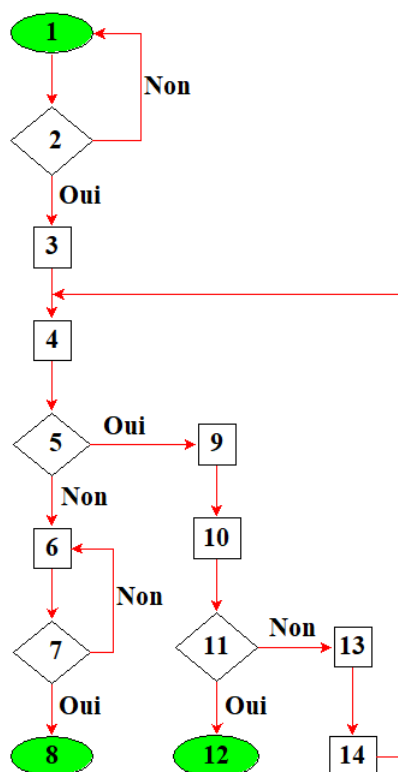
- Multiple du délai d'insertion ($r \dots 512$ temps bits)



- r est un nombre aléatoire vérifiant $0 \leq r < 2P_k P$ 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

Remarques :

- Le temps d'émission doit être supérieur au double du temps de propagation entre les 2 points les plus éloignés du réseau.
- Le nombre de collisions augmente avec le nombre de stations voulant émettre et le débit réel diminue.
- Cette méthode réduit les temps d'attente, lorsque le nombre de stations voulant émettre est faible.



1. L'hôte veut transmettre
2. Porteuse détectée?
3. Mise en trame
4. Début d'émission
5. Collision détectée?
6. Poursuivre l'émission
7. Transmission terminée?
8. Transmission terminée
9. Diffusion générale d'un signal de bourrage
10. Essais = Essais + 1
11. Essais > trop nombreux?
12. Trop de collisions; annuler transmission
13. Algorithme calcule réémission temporisée
14. Attendre t secondes.

Aucun intérêt d'en connaître davantage dans votre cursus. Si vous retenir que CSMA/CD est :
Elle détermine quand et comment un paquet de données est situé dans le câble.

Avant de pouvoir transmettre des données, un matériel Ethernet doit écouter le support pour s'assurer qu'il est « libre », c'est-à-dire qu'il n'y a pas d'autre matériel effectuant une transmission. Ainsi, quand le réseau est **libre**, les matériels sont **autorisés** à commencer la transmission.

Pendant le processus de transmission, le matériel doit continuer d'écouter le réseau pour savoir si un autre matériel est également en train de transmettre. S'il n'y en a pas d'autre, les données sont considérées comme envoyées au destinataire sans interruptions.

En revanche, s'il détecte pendant la transmission qu'un autre matériel est également en train de transmettre (on parle de collision), les deux matériels arrêtent leur transmission et effectuent un processus connu comme back-off, consistant à attendre une durée aléatoire avant d'essayer de reprendre la transmission.

Cela était problématique, ce n'est plus le cas actuellement, car avec les **Switchs**, on peut considérer que les machines sont reliées 2 à 2...!!!

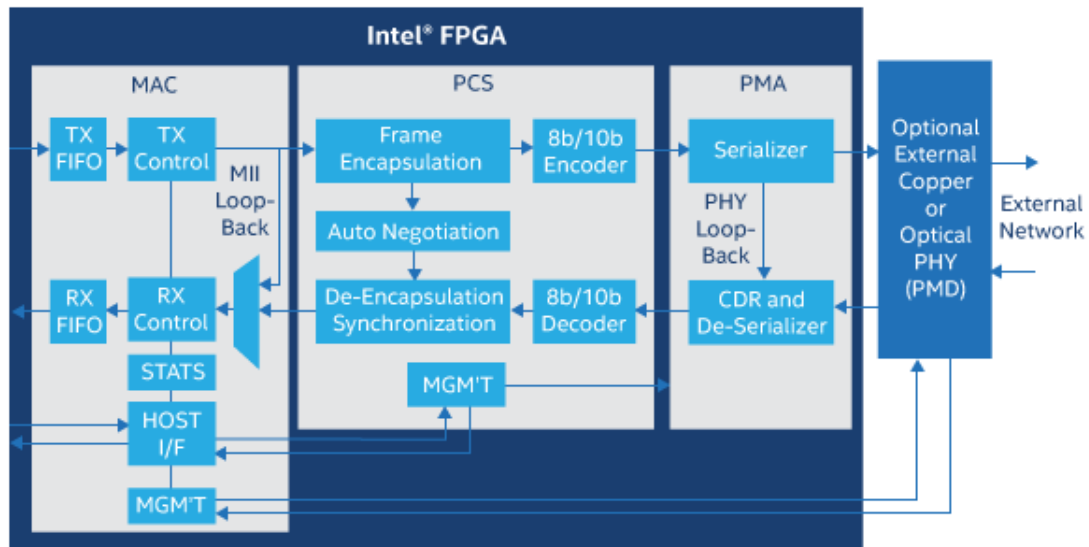
4 Technologies Ethernet moderne

Ethernet a pratiquement tout remplacé dans le monde moderne. Il reste des bribes d'ADSL qui utilisent ATM mais à part cela plus rien.

Attention ne pas confondre Fibre optique qui est la couche 1 et ADSL qui n'est pas que la couche 1... l'ADSL s'appuie sur du cuivre est une belle « connerie » due à la puissance de France Télécom au début des années 2000...

4.1 La carte réseau

D'un ordinateur, d'un Switch, d'un routeur etc... gère la couche 2 Liaisons de données et est constituée d'une puce (FPGA : sorte de μP dédié) qui fait tout ou presque. Du moins tout ce qui est nécessaire à la couche 2 et qui doit être très très très rapide!!!



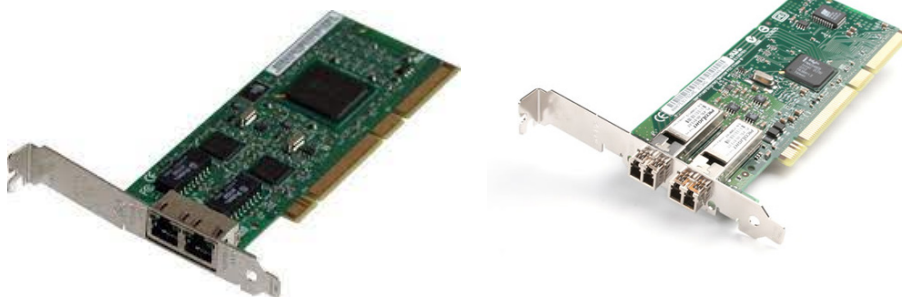
Source :

<https://www.intel.com/content/www/us/en/programmable/products/intellectual-property/ip/interface-protocols/m-alt-ethernet-mac.html>

Attention ne pas confondre Ethernet et RJ45!!!

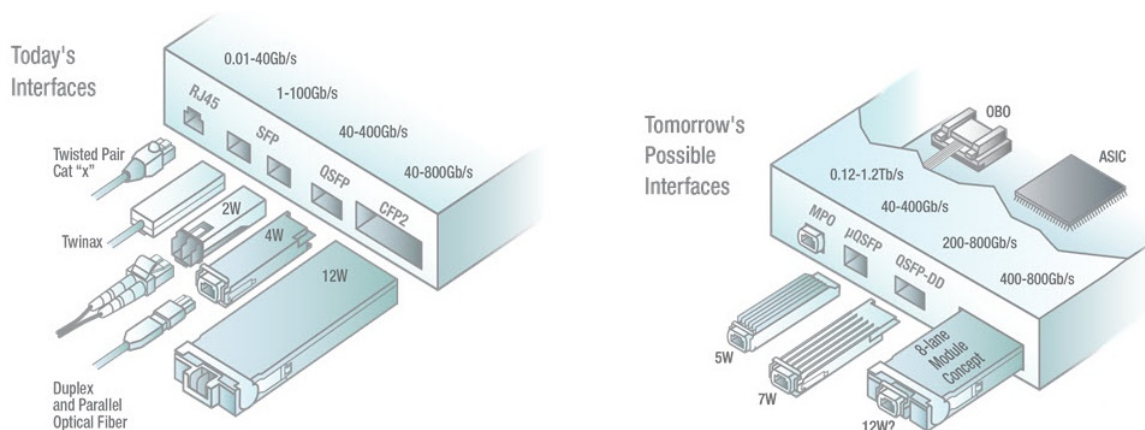
RJ45 c'est le connecteur!!!

Ci-dessus 2 cartes réseau de serveurs, à gauche en RJ45 donc cuivre et à droite en fibre optique.









4.2 Les connecteurs

Un petit dessin est plus rapide à comprendre qu'un long texte...



Source : <https://www.informatiquenews.fr/wp-content/uploads/2016/03/30-Ethernet-3.jpg>

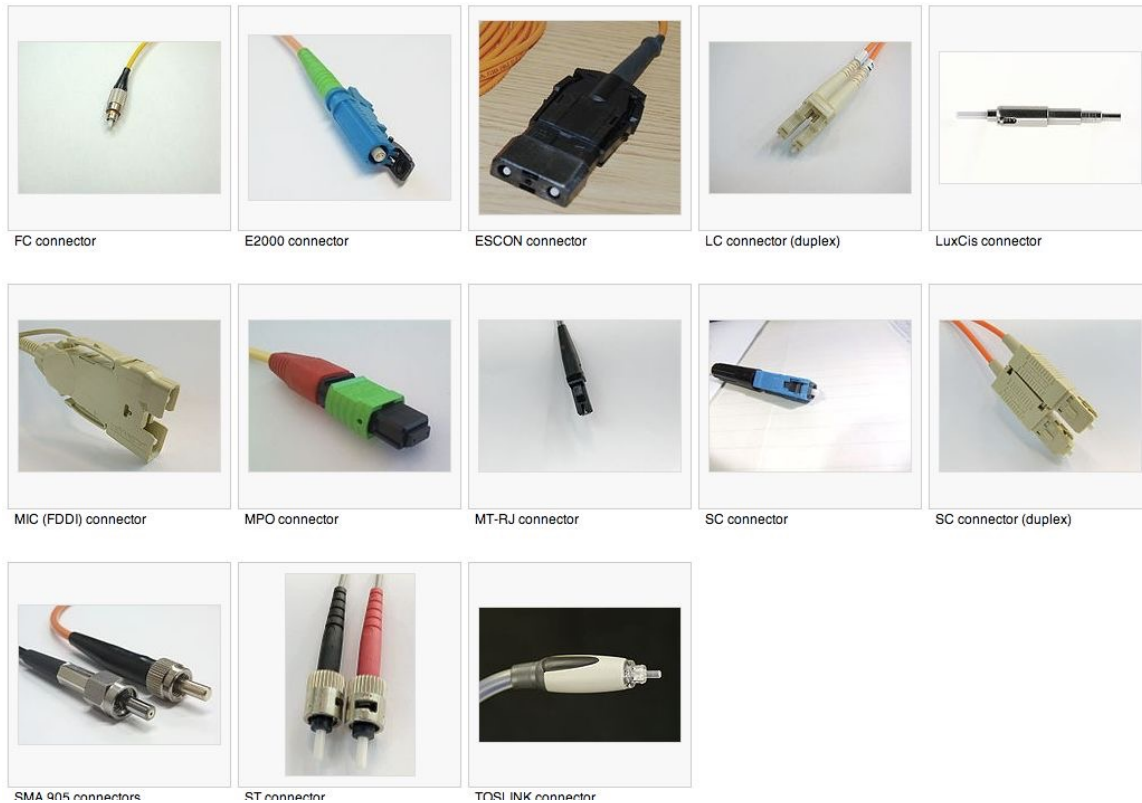
Selon les lieux (Entreprise, **ISP** ou **DC**) d'utilisation et la vitesse, la connectique n'est pas la même comme le montre la figure suivante.

Enterprise	Service Provider		Data Center		
Server Line Card 	CFP 	CFP2/8 	QSFP 	SFP 	OSFP and QSFP-DD 
nx10G PHY	100G Gearbox	nx25G CDR	nx25G CDR	nx25G CDR	nx50G PAM
		nx50G PAM	nx50G PAM	nx50G PAM	nx100G PAM
			nx100G PAM		

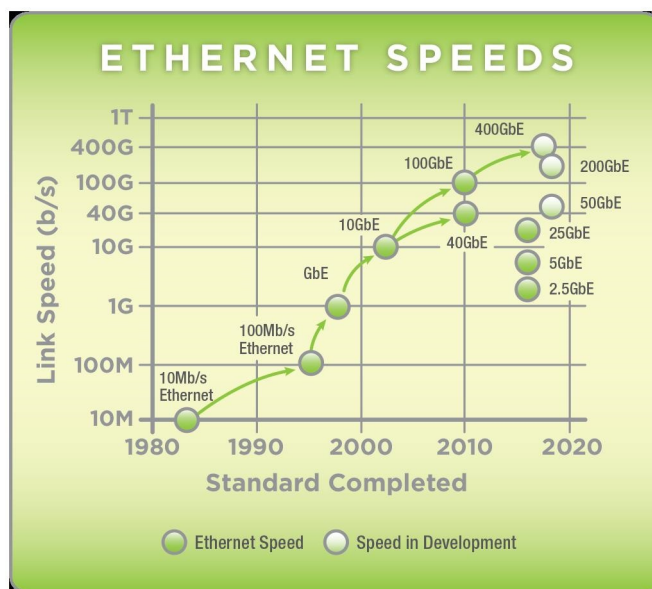
Source :

<https://www.inphi.com/products/optical-phy/>

Plus d'informations ici : https://fr.wikipedia.org/wiki/Connecteur_fibre_optique



4.3 La vitesse et les débits !!!



4.4 Les principaux standards de la norme IEEE 802.3x

Classe	Codage / Support / connectique	L.M. / D
Fast Ethernet (802.3u)		
100BaseT4	Codage Manchester / 4 paires Torsadées UTP/STP cat 3,4,5	100m
100BaseTX	2 paires Torsadées UTP cat 5 (supporte 200Mbit/s en full duplex)	100m
100BaseFX	Fibre Multimodes (62.5/125µm)	400m
Gigabit Ethernet (802.3z 802.3ab)		
1000BaseLX	Laser grandes ondes Fibre monomodes et multimodes	3km
1000BaseSX	Laser ondes courtes Fibre multimodes	500m
1000BaseCX	Codage Manchester / Paires Torsadées blindés 150W	25m
1000BaseT (802.3ab)	Paires Torsadées FTP cat 5,5PeP(100Mhz), 6(200Mhz) et 7(600Mhz)	100m
10 Gigabit Ethernet (802.3ae/ak)		
10GBaseX (LX4-CX4)	Codage 8B/10B - Fibre	40km
10GBaseR (SR-ER)	Codage 64B/66B Fibre (LAN)	40km
10GBaseW (SW-LW-EW)	Codage 64B/66B Fibre (WAN)	40km

Remarque : Je n'ai pas mis ceux à 100G et plus

5 Conclusion

Ce cours est vaste et impossible de tout connaître!!!

À retenir :

- **@MAC qu'il est impossible de changer et qui est unique par carte réelle !**
- Ne confondez pas Ethernet et Internet!!!
- On utilisera VRRP dans le TP : VRRP *load Balancing* dans quelques semaines.
- Ethernet c'est la couche 2 du modèle OSI, la couche liaison!
- Ethernet c'est les Switchs!
- Internet c'est... un cours prochain, patience et peut-être une question du CT ☺