

INTERNET ET RESEAUX

1. Un peu d'histoire (source pixees.fr)

La DARPA (Defense Advanced Research Projects Agency) voit le jour en 1958, cette agence gouvernementale américaine a pour but de veiller à la constante suprématie des États unis en matière technologique et scientifique. En 1962 la DARPA soutient le projet du professeur Licklider qui a pour but de mettre en réseau les ordinateurs des universités américaines afin que ces dernières puissent échanger des informations plus rapidement (même à des milliers de kilomètres de distance). En 1968, ARPAnet, 1er réseau informatique à grande échelle de l'histoire voit le jour. Le 29 octobre 1969, le 1er message (le mot "login") est envoyé depuis l'université de Californie à Los Angeles vers l'université de Stanford via le réseau ARPAnet (les 2 universités sont environ distantes de 500 Km). C'est un demi-succès, puisque seules les lettres "l" et "o" arriveront à bon port. En 1972, 23 ordinateurs sont connectés à ARPAnet (on trouve même des ordinateurs en dehors des États unis). En parallèle au projet ARPAnet, d'autres réseaux voient le jour, problème, ils utilisent des protocoles de communication hétéroclite (UUCP, NCP ou encore X.25) et 2 ordinateurs appartenant à 2 réseaux différents sont incapables de communiquer entre eux puisqu'ils n'utilisent les mêmes protocoles. En 1974 Vint Cerf et Bob Khan vont mettre au point le protocole TCP qui sera très rapidement couplé au protocole IP pour donner TCP/IP. TCP/IP, grâce à sa simplicité, va très rapidement s'imposer comme un standard : les différents réseaux (ARPAnet et les autres) vont adopter TCP/IP. Cette adoption va permettre d'interconnecter tous ces réseaux (2 machines appartenant à 2 réseaux différents vont pouvoir communiquer grâce à cette interconnexion). Internet était né (le terme Internet vient de "internetting" qui signifie "Connexion entre plusieurs réseaux"). TCP/IP est donc au coeur d'Internet, voilà pourquoi aujourd'hui, la plupart des machines utilisent TCP/IP.

2. Modèles théoriques

Les réseaux sont construits sur deux modèles qui se recoupent en grande partie : OSI et TCP/IP.

Dans le modèle OSI (Open System Interconnection), il y a 7 couches indépendantes, chacune ne communiquant qu'avec une couche adjacente.

Dans le modèle TCP/IP, il y a entre 2 et 5 couches...qui correspondent plus ou moins ou couches OSI.

	Nom dans le modèle OSI	Rôles primaire/secondaire	Matériel Protocoles	Structure de données Encapsulées (trame à l'extérieur)
Ordinateur	7 - Application		Port HTTP(S) FTP DNS SMTP POP IMAP...	Données / Requête HTTP
	6 – Présentation OSI	Non utilisé dans TCP/IP		
	5 – Session OSI	Non utilisé dans TCP/IP		
Réseau	4 -Transport	Gérer les connexions applicatives	Port (non matériel, c'est le n° d'application) TCP UDP NAT	Message /segment
	3 - Réseau	Interconnexion des réseaux locaux / fragmentation des paquets	Routeur IP – ARP -ICMP	Paquets /datagrammes
	2 - Liaison de données	Connexion locale des machines / détection des erreurs de transmission	Switch = commutateur MAC	Trame
	1 - Physique	Support de transmission	Hub = concentrateur Ethernet, Wi-Fi, ATM, PPP...	Bits

L'adresse IP est inséparable de son masque de sous-réseau. Dans le masque, les bits à 1 représentent la partie réseau de l'adresse IP, les bits à 0 représentent la partie machine. Les machines du même sont réseau peuvent communiquer directement entre elles sans passer par un routeur (réseau local).

Exemple : IP 192.168.0.1 masque 255.255.0.0 (en notation CIDR 192.168.0.1/16, les 16 derniers bits sont à 0)

La partie réseau est 192.168, la partie machine est 0.1. C'est à dire que les adresses machines iront de (192.168.) 0.0 à (192.168.)255.255, ou presque (en fait 0.1 et 255.254, cf. ci-dessous « adresses réservées »).

Pour des raisons de calcul binaire (les « 1 » du masque sont contigus), les masques de sous réseau ne peuvent prendre que les valeurs suivantes : 0 ; 128 ; 192 ; 224 ; 240 ; 248 ; 252 ; 254 ; 255

Adresses réservées :

- la première adresse d'une plage est l'adresse du réseau proprement dit, elle ne peut pas être utilisée par une machine. La dernière adresse est l'adresse de broadcast, elle permet d'envoyer un message à toutes les machines du réseau.
- IP réservées pour des usages spéciaux : https://en.wikipedia.org/wiki/Reserved_IP_addresses

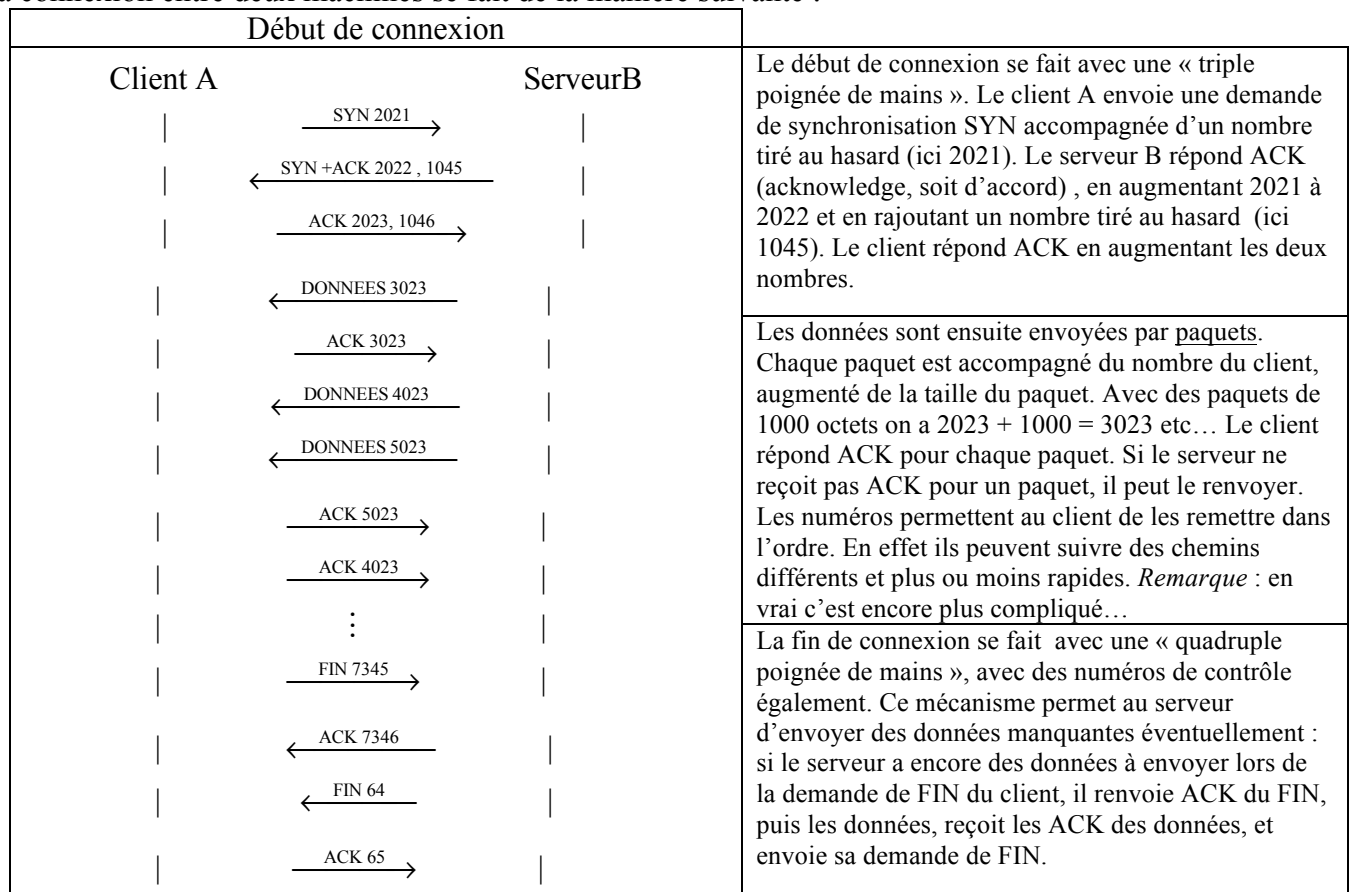
Remarque : la notion de masque de sous-réseau remplace la notion de classe d'adresse IP, que vous pouvez rencontrer dans vos recherches, mais qui n'est plus utilisée (notion instructive ceci dit).

5. Protocole IP et protocole TCP : TCP/IP (couches 2 et 3 du modèle TCP/IP)

Le protocole TCP (transmission control protocol) est un protocole de communication fiable entre applications (comme un navigateur internet et le site que l'on contacte). IP est un protocole de communication entre machines (l'ordinateur sur lequel tourne le navigateur, et le serveur web sur lequel est le site). Les deux sont étroitement liés, c'est pourquoi on parle souvent du protocole TCP/IP.

TCP/IP découpe les données en paquets (appelés aussi segments, datagrammes, ou trames suivant le niveau auquel on se place). Ainsi une image lourde sera découpée en milliers de paquets. L'avantage étant que si un des paquets se perd, alors le serveur n'a pas besoin de renvoyer toute l'image mais juste le segment manquant.

La connexion entre deux machines se fait de la manière suivante :



Remarques :

- TCP identifie les applications qui communiquent entre elles via des ports. Par exemple, le port 80 est utilisé pour le Web en http, le port 443 pour le Web sécurisé (https), 143 pour le mail, les jeux utilisent 6112 (sauf Counter Strike qui a son port dédié 27015), etc.
- Le protocole UDP est un protocole de communication plus rapide et moins fiable. Il n'y a pas de vérification de bonne réception (phase acknowledge ACK). Il est utilisé pour le streaming de voix ou d'images. En effet, même avec une perte éventuelle de segments de données, le message restera compréhensible : l'image sera pixélisée, ou bien il y aura de la friture sur la ligne, mais cela restera compréhensible.
- Les deux protocoles TCP et UDP peuvent être utilisés simultanément. C'est ce qui se passe lorsque l'on joue en réseau (en protocole TCP) tout en étant connecté à un serveur de discussion, un tchat (en protocole UDP). Ainsi le jeu peut laguer, alors que le tchat ne le fera pas : TCP est un protocole plus lent que UDP. Certains jeux de tirs à la première personne utilisent UDP pour palier à ce problème (ce qui entraîne d'autres, comme des délocalisations intempestives).

Le protocole IP envoie des paquets de données, sous le format simplifié suivant:

En-tête	Adresse IP SRC (source)	Adresse IP DST (destinataire)	Données à envoyer
---------	-------------------------	-------------------------------	-------------------

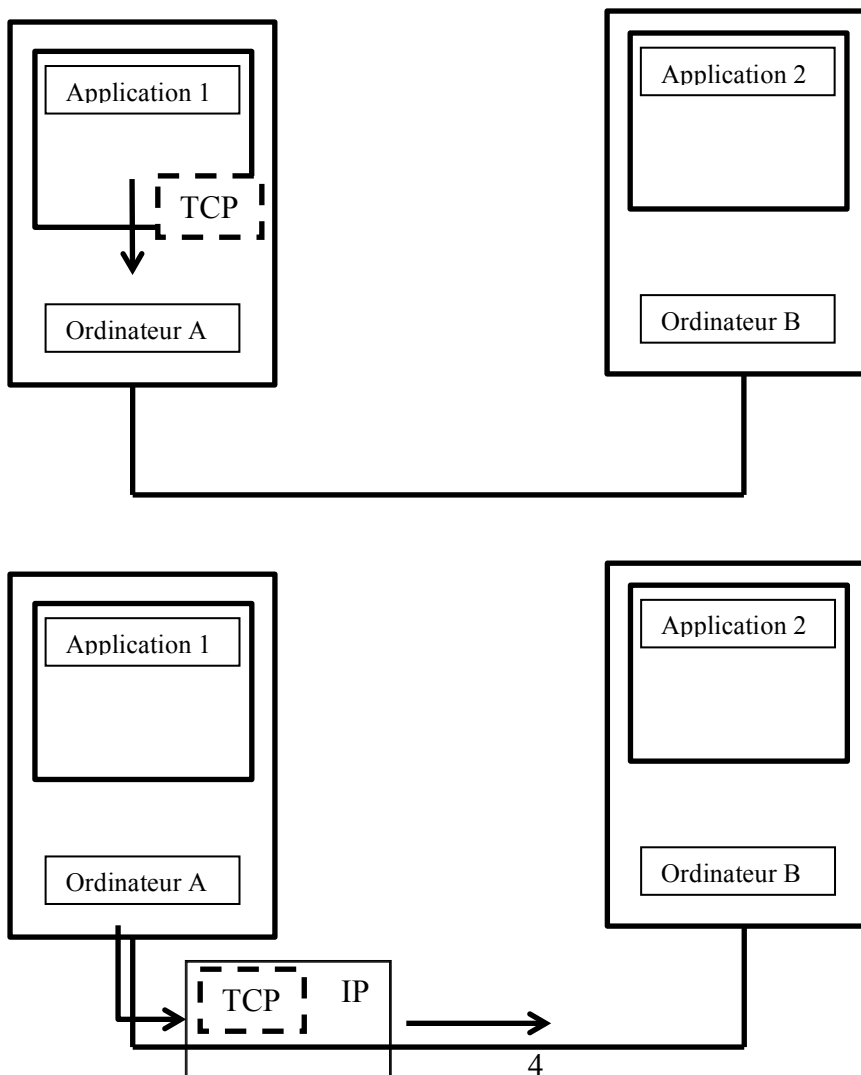
Dans les données, il y a le segment TCP (ou UDP), de format également simplifié :

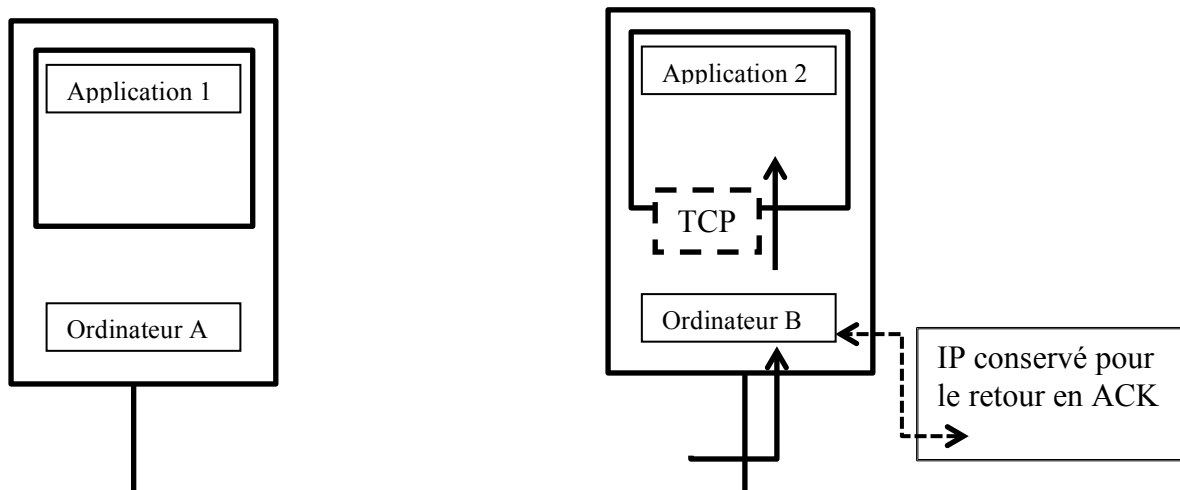
Port source	Port destination	N° SYN/FIN/DATA	N° ACK	Longueur totale	Données
-------------	------------------	-----------------	--------	-----------------	---------

Il y a donc *encapsulation* des segments TCP à l'intérieur des paquets IP.

Dans l'en-tête du paquet IP figure un élément important, qui est la durée de vie TTL (time to live). Si le paquet n'est pas arrivé à destination au bout du TTL, alors il est détruit. Sans le TTL, Internet serait submergé de paquets fantômes !

Ci-dessous le schéma pour l'envoi d'un segment de données





Le retour pour confirmation (acknowledgment) se fait suivant le même principe.

Au niveau matériel, les routeurs gèrent la communication entre réseaux locaux distincts, suivant le protocole IP.

6. Survol de la couche 1 : adresses MAC et protocole Ethernet

L'adresse MAC est unique à chaque objet connecté. Elle est codée hexadécimal sur 6 octets (soit 280 mille milliards d'adresses). Un ordinateur peut avoir plusieurs adresses MAC (une pour la carte réseau, une pour la carte Wi-Fi, et Windows en rajoute quelques-unes pour le plaisir).

Les 3 premiers octets sont caractéristiques du constructeur. Par exemple, 60:8B:0E est le « OUI » (organization unique identifier) d'Apple, 38:94:ED celui de Netgear, etc...

L'adresse ff:ff:ff:ff:ff:ff est particulière, c'est l'adresse de broadcast, qui permet d'envoyer un message à toutes les machines du réseau local.

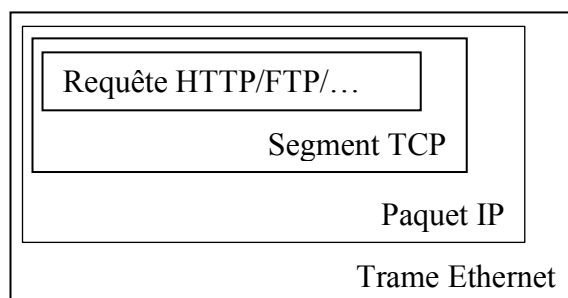
Le protocole Ethernet envoie des *frames* sur le réseau, de ce format simplifié :

Adresse MAC DST (destinataire)	Adresse MAC SRC (source)	Paquet IP	CRC (code de correction des erreurs)
-----------------------------------	-----------------------------	-----------	--------------------------------------

On constate que le paquet IP est enchâssé dans la trame Ethernet.

Remarque : le protocole Wi-Fi est très semblable au protocole Ethernet.

En complétant avec le cours sur le protocole http, on donc cette structure de données :



Au niveau matériel, les switch gèrent la distribution des données aux adresses MAC sur le réseau local, en utilisant une table de routage. Si des données doivent être envoyées à une machine qui n'est pas sur le réseau local, alors elles sont envoyées au routeur.

7. Un exemple de protocole de contrôle

Le protocole du bit alterné est un protocole simple de la couche 1b (liaison) de TCP/IP, qui permet la retransmission de trames perdues ou corrompues.

Une fois la connexion établie, lors de l'envoi de données d'un serveur B à un client A, B rajoute un bit de séquence SEQ aux données, soit SEQ0 soit SEQ1. A dispose de deux ACK : ACK0 et ACK1. Ces bits uniques sont appelés drapeaux (flag en anglais).

- B envoie une trame de données continuellement avec par exemple SEQ0, jusqu'à ce qu'il reçoive un ACK0. Une horloge interne permet de déterminer la fréquence d'envoi de la trame.
- Quand A reçoit les données avec SEQ0, il envoie continuellement ACK0, jusqu'à ce qu'il reçoive des données avec SEQ1. Si les données sont corrompues, il ne renvoie pas ACK0. Là aussi une horloge donne la fréquence d'envoi du ACK.
- Quand B reçoit un ACK0, il change son numéro de séquence en SEQ1, et envoie continuellement la trame suivante de données jusqu'à recevoir un ACK1.
- Tout message reçu avec le mauvais SEQ ou le mauvais ACK sera considéré comme un « négative acknowledge », et ignoré.

Le protocole peut être initialisé en envoyant des messages (données ou ACK) sans réel contenu, avec SEQ1 et ACK1. Le premier message SEQ0 sera considéré comme un vrai message.

Ce protocole n'est plus utilisé actuellement, il est remplacé par des protocoles plus complexes, et plus efficaces surtout.

8. En pratique : invite de commande et logiciels

a. *Sous Windows :*

En ligne de commande :

- `ipconfig` pour connaître sa configuration/carte réseau
- `ping` pour joindre une machine du réseau ou une adresse IP externe (qui peut être donnée sous sa forme DNS `www....`)
- `tracert` pour voir les routeurs par lesquels on passe pour joindre une adresse
- Table de routage :
 - `arp -a` pour afficher la table.
 - `arp -a @ip` pour afficher uniquement les entrées associées à @ip.
 - `arp -s @ip @MAC` pour ajouter manuellement une entrée.
- `route print` pour connaître sa table de routage. Les tables de windows contiennent des routes propres à l'implémentation que fait Windows du routage, en plus de 0.0.0.0 et de son propre réseau local.

Interface graphique : panneau de configuration > connexions réseau > protocole internet (TCP/IP), où l'on peut modifier son IP (soyez prudents...)

b. *Wireshark*

Ce logiciel permet d'observer les paquets (sniffer, renifleur en français). Faire une capture (start, puis stop, puis filter avec l'IP à observer). On a ensuite le détail couche par couche. On peut remarquer que sur les premiers paquets il n'y pas de couche applicative, puisque la connexion n'est pas encore établie entre les applications.

c. *Plus joli que traceroute avec en prime un « whois » et un sniffer.*

Une application fort sympathique permet d'observer visuellement le trajet emprunté par un paquet données pour rejoindre une URL lointaine. Il s'agit de Open Visual Traceroute, téléchargeable ici : <https://sourceforge.net/projects/openvisualtrace/>. Et en plus c'est libre.

9. En pratique : avec Python

La bibliothèque `socket` permet de programmer des applications client-serveur. Une socket (prise en français, inutilisé) est une interface logicielle, présente entre http (sur le navigateur) et TCP. Par défaut, la socket de Python utilise IPv4 et TCP. Voir TP pour une approche basique de `socket`.

10. Protocole DNS (rappels de 2nde)

Le protocole DNS (domain name system) permet de relier les adresses IP aux adresses URL (universal resource locator). En français, URL se traduit par adresse symbolique ou adresse universelle. Les URL sont les adresses du type www.sitemachin.soussitebidule.org.

Ce protocole fait appel à des ordinateurs spécialisés, les serveurs DNS. Un serveur DNS ne connaît qu'une petite partie des adresses. Lorsque l'on tape une URL dans la barre du navigateur, la première étape réalisée est cette conversion, qui se fait en plusieurs étapes.

Par exemple, lorsque l'on tape www.wikipedia.fr :

- Le serveur DNS (situé dans la box chez vous) ne reconnaît pas cette adresse. Il va donc transmettre sa demande au serveur DNS spécialisé dans le Top Level Domain (TLD) « .fr ». Votre box connaît les IP de tous les serveurs de TLD (.com, .eu, .fr, ...).
- Le serveur DNS de « .fr » ne connaît pas www.wikipedia.fr, mais il connaît « wikipedia.fr ». Il va donc demander au serveur DNS « wikipedia.fr » l'adresse.
- Le serveur DNS « wikipedia.fr » connaît l'IP de www.wikipedia.fr, il renvoie cette adresse au serveur « .fr », qui la renvoie à la box, qui la renvoie à l'ordinateur.

Rentrer une IP ou rentrer une adresse symbolique dans le champ de saisie d'un navigateur revient au même. Par contre, pour créer un réseau local, il est nécessaire de connaître les IP de chaque machine, qui n'ont pas d'adresse symbolique.

11. Modèle client-serveur (rappels de 2nde)

Les communications Internet sont souvent basées sur le modèle client-serveur. Le serveur, qui contient par exemple un site web, un service de messagerie électronique, etc. attend une demande des clients en écoutant le réseau. Dès qu'une demande d'un client lui parvient, comme une demande de consultation du site, ou de réception du courrier, une connexion est établie entre le client et le serveur via les protocoles TCP et IP. Ces protocoles sont universels, et permettent ainsi à des machines très différentes de communiquer (exemple : une caméra de vidéo-surveillance et un téléphone).

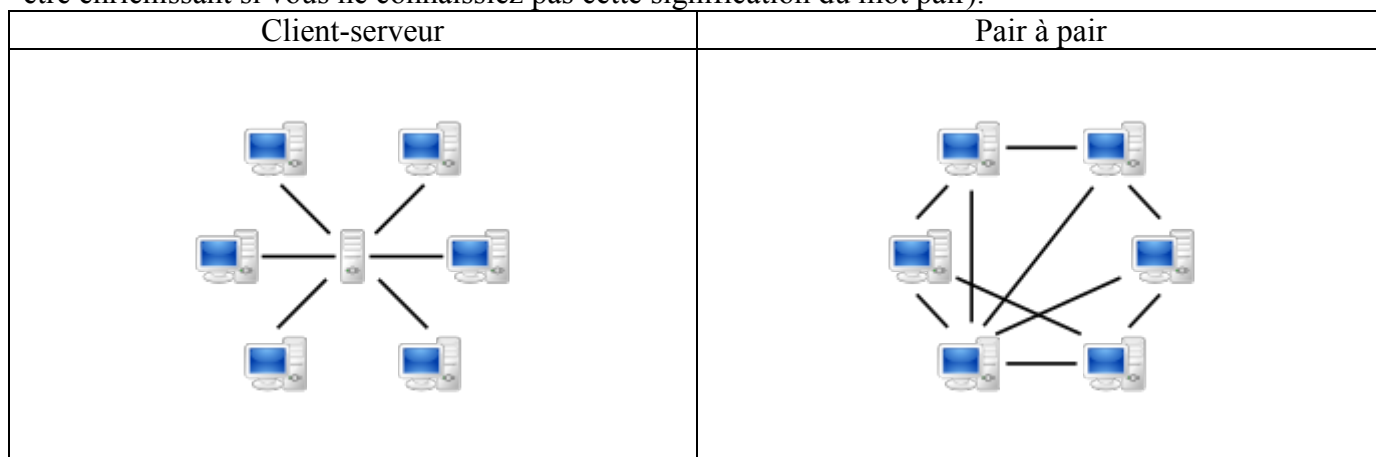
Les serveurs sont en général des ordinateurs spécialisés, même si n'importe quel ordinateur peut faire office. Les grosses entreprises, notamment les GAFAM, qui reçoivent des dizaines de milliers de requêtes par seconde, disposent de centaines de serveurs dans des data centers (fermes ou centres de données). Par exemple, IBM a en France 10 data centers, dont un à Montpellier, pour une surface de 20000 m². Mais aussi de 6 fermes géantes dans le monde, chacune grande comme 10 terrains de football. Google aurait eu en 2016 à peu près 2,5 millions de serveurs.

Les centres de données sont très énergivores, ils consomment actuellement 6 % de l'énergie mondiale et ce chiffre augmente constamment. Non seulement il faut un approvisionnement constant en électricité, sans aucune micro-coupure, mais il faut également refroidir ces fermes.

Remarque : en France, les box consomment représentent à peu près 1% de la consommation électrique totale.

12. Réseaux pair-à-pair (rappels de 2nde)

Il existe d'autres architectures de communication que le système client-serveur. Le réseau pair à pair (peer to peer en anglais, abrégé p2p) est une architecture où chaque machine est à la fois client et serveur. C'est de là que vient le terme « pair », dans le sens « tous égaux » (ouvrir le dictionnaire peut être enrichissant si vous ne connaissiez pas cette signification du mot pair).



Les réseaux p2p servent au partage de fichiers et au calcul partagé principalement. Leur réalisation nécessite un logiciel particulier, appelé de manière générique « serveur », en tant qu'abréviation de *serveur-client*.

Les applications les plus connues pour le partage de fichier sont eMule et BitTorrent. Elles ne servent pas qu'au piratage de musique, films et séries...en effet de nombreux fichiers de travail sont trop lourds pour être envoyé par email, une des possibilités pour les distribuer est de passer par un réseau p2p.

Remarque : l'utilisation de BitTorrent n'a rien d'illégal en soi. Télécharger des ressources qui ne sont pas sous licence libre l'est par contre. Dans ce cas, les risques de piratage informatique (à distinguer du piratage de contenu) de votre ordinateur ne sont pas négligeable : virus dans les fichiers, hameçonnage (phishing), etc...

Le calcul partagé est l'utilisation d'une partie du temps de calcul des ordinateurs personnels des internautes pour collaborer à des projets scientifiques. Les ordinateurs d'aujourd'hui sont tellement puissants que la majeure partie du temps, une grande partie de leur processeur est disponible pour effectuer des calculs. Le projet BOINC a saisi cette opportunité pour créer un parc informatique réparti dans le monde afin d'utiliser cette puissance de calcul totale pour effectuer des calculs trop complexes pour être réalisés dans un laboratoire.

Le projet BOINC demande donc au particulier de permettre l'usage de la puissance de calcul dont il n'a pas immédiatement besoin pour contribuer à la recherche sur le repliement des protéines (Folding@Home) et même la recherche d'intelligence extra-terrestre par analyse de spectre électromagnétique (SETI@home).

EXERCICES INTERNET ET RESEAUX

1. Quelles sont les valeurs possibles pour les masques de sous-réseau, sachant que les 1 doivent être contigus ?
2. Soit l'adresse 192.16.5.133/29. Combien de bits sont utilisés pour identifier la partie réseau ? Combien de bits sont utilisés pour identifier la partie hôte ?
3. Soit l'adresse 201.16.5.10/28. Quel est le masque réseau correspondant ?
4. Quel est le plus petit masque de sous-réseau permettant de connecter 1000 machines ? 1100 machines ?
5. Quelle est la plage d'adresses du réseau pour les IP suivantes, combien de machines peut-on connecter, quelle est l'adresse du réseau et quelle est l'adresse de broadcast ?
 - a. 192.168.0.42/24
 - b. 192.168.0.23/27
 - c. 192.168.0.168/27
 - d. 172.200.202.16/22
 - e. 196.115.110.0/255.255.255.240
 - f. 134.168.215.18/255.255.192.0
6. Quelle adresse doit-on utiliser pour envoyer un message à toutes les machines du réseau 145.1.0.0/16 ? Du réseau 32.1.0.0/12 ?
7. Les couples suivants indiquent-ils des adresses de réseau, de machine ou de broadcast ?

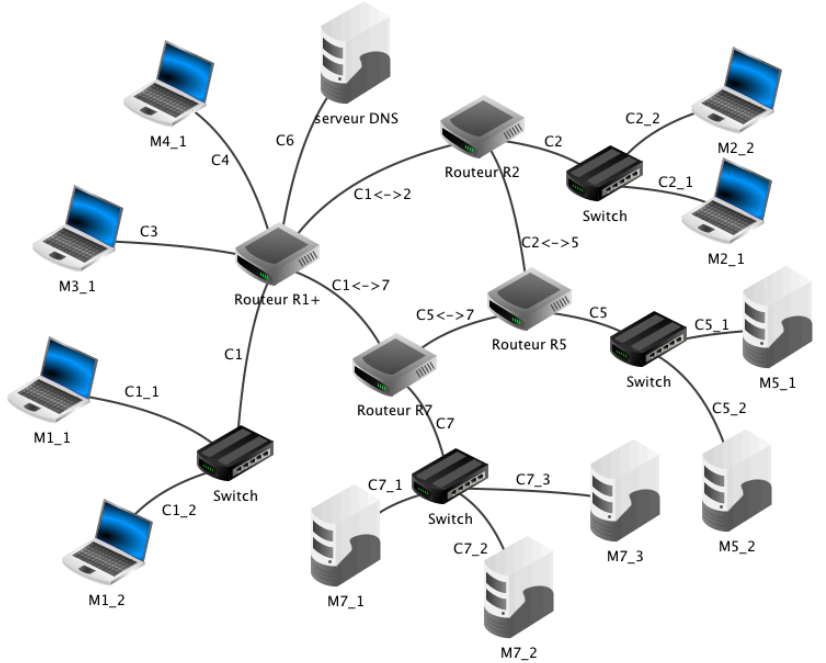
192.192.15.15/255.255.255.0	192.192.65.0/255.255.255.0
201.221.21.0/255.255.254.0	201.221.1.0/255.255.254.0
199.81.82.29/255.255.255.224	199.81.82.31/255.255.255.224
10.0.0.255/255.255.254.0	
8. Les trois IP : 192.168.135.200 ; 192.168.147.203 ; 192.168.192.168 font-elles partie du même sous-réseau de masque 255.255.248.0 ?
2^{ème} version : Les trois IP : 201.202.128.5 ; 201.202.135.200 ; 201.202.145.1 font-elles partie du même sous-réseau de masque 255.255.248.0 ?
9. On donne le masque de sous-réseau 255.255.252.0
Parmi les machine suivantes, lesquelles sont des machines du même sous-réseau ?

134.65.153.10	134.65.151.10	134.65.155.111
---------------	---------------	----------------

134.65.160.1 134.65.147.0 134.65.143.55
 134.65.148.210 134.65.159.254

10. On considère deux sous-réseaux A et B reliés par un routeur. Pourquoi une machine du sous-réseau A d'adresse 125.48.0.20/255.255.0.0 ne peut pas communiquer directement à une machine du sous-réseau B d'adresse 125.48.45.21/255.255.0.0 ? *remarque* : il y a un bout de phrase très important dans l'énoncé.
11. Décrire sous forme de schéma la session TCP entre un client A et un serveur B, sachant que B envoie 5 paquets de données de taille 100, 100, 100, 100 et 80, et que A demande la fin de session alors que B a encore deux paquets de données à envoyer.
12. Pourquoi le protocole de bit alterné peut-il être mis en défaut ? Donner plusieurs exemples.
13. On donne le réseau suivant. Les câbles et les machines sont tous numérotés de la manière suivante :

- Ordinateur $M_{x,y}$ d'adresse 192.168.x.y, appartenant au réseau d'adresse 192.168.x.0/255.255.255.0
- Câbles C1 à C7 vers les réseaux 1 à 7 (réseaux locaux parfois sans commutateur pour simplifier)
- Câbles $C_{a \leftrightarrow b}$ reliant deux routeurs de numéros a et b
- Routeur :
 - interface vers les réseaux : IP 192.168.x.1
 - interface entre deux routeurs IP 192.168.z.1 ou 2, z n'étant pas attribué à un réseau.

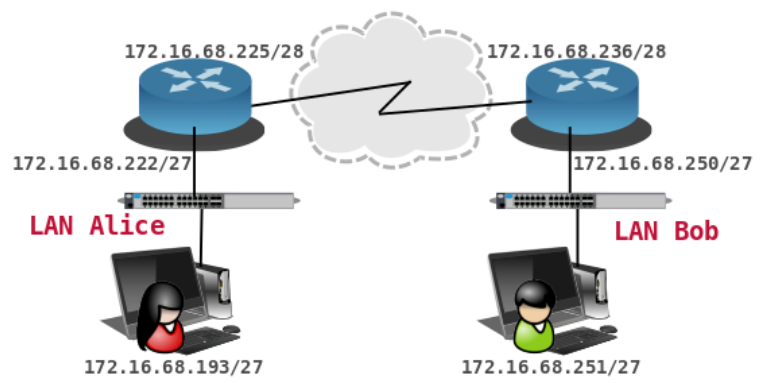


Le serveur DNS est, comme son nom l'indique, un serveur DNS pour l'ensemble du réseau (les ordinateurs étant tous identifiés par leur nom $M_{x,y}$).

On effectue une commande sur la machine M1_1. Que se passe-t-il dans les cas suivants :

- a. ping M5_1 avec le câble C6 coupé ?
- b. ping 192.168.5.1 avec le câble C6 coupé ?
- c. traceroute M7_1 avec le câble C1<->7 coupé ?
- d. ping 192.168.1.2 avec le câble C1 coupé ?
- e. ping M5_1 avec un TTL de 3 ? de 7 ?
- f. ping M5_2 avec le câble C2_1 coupé ?

14. Un exercice de inetdoc.net/
 Alice est au bord de la crise de nerfs !
 Aucun des messages envoyés à Bob n'est arrivé à destination. Bob est lui aussi sur le point de craquer ! Il essaie désespérément d'envoyer des messages à Alice sans succès. Il faut absolument faire quelque chose pour les aider.



Quelle erreur a été commise dans l'affectation des adresses (et/ou) des masques réseau ?
 Proposer une solution pour rendre les communications possibles.

15. Reprendre le paragraphe 5 et faire un échange avec 4 paquets de données de 1000, 800, 1200 et 500 octets, où l'échange se fait dans l'ordre arrivée du paquet 800, puis 1000, demande de fin de connexion, arrivées des paquets 1200 puis 500.