Introduction au fonctionnement d'Internet : infrastructure, protocoles, applications

Attribution - Partage dans les Mêmes Conditions :
http://creativecommons.org/licenses/by-sa/4.0/fr/

Table des matières

Objectifs 3

Introduction 4

I - Centralisation 5

II - Aux origines du modèle décentralisé d'Internet 9

III - Principe de la communication sur Internet 11

IV - Infrastructure réseau d'Internet 14

V - Ordinateurs clients 15

VI - Ordinateurs serveurs 16

VII - Ordinateurs routeurs 17

VIII - Protocoles réseaux TCP/IP 18

IX - Adresses IP et noms de domaine 21

X - Protocoles applicatifs d'Internet : SSH, mail, Web... 23

XI - Les clients et les serveurs web parlent HTTP 26

XII - Quelques logiciels qui font fonctionner Internet 28

XIII - Synthèse 29

XIV - Cours-émission 30

Glossaire 31

Abréviations 32

Crédits des ressources 33

Contenus annexes 35

Objectifs

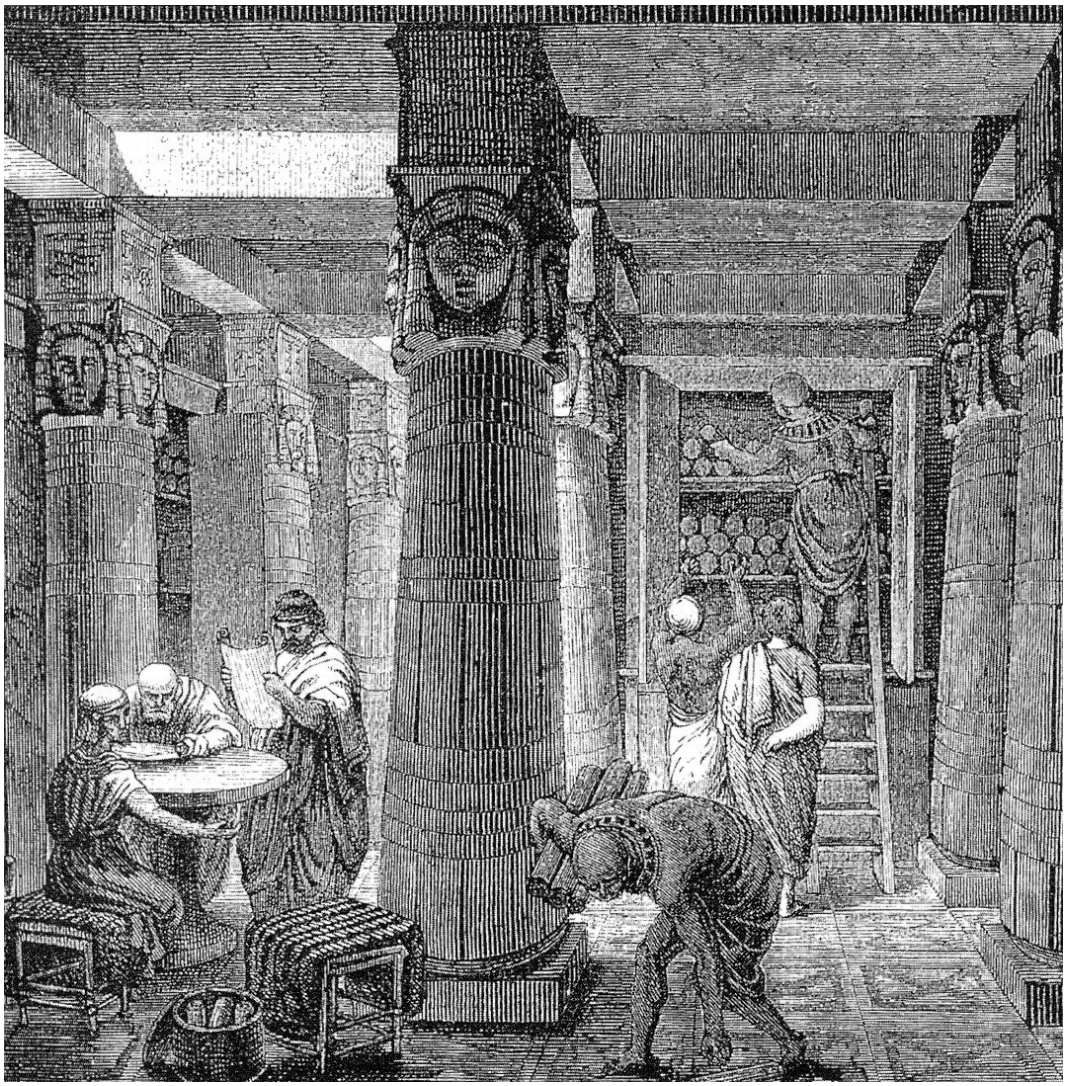
- Savoir définir ce qu'est un client, un serveur, un routeur, un protocole
- Connaître les base du protocole TCP/IP
- Savoir ce qu'est une adresse IP et un nom de domaineq
- Savoir associer les principales applications d'Internet (SSH, mail et Web a minima), leurs principaux protocoles (SSH, IMAP, SMTP et HTTP) et leurs principaux outils (terminal, client mail, navigateur web)

Introduction

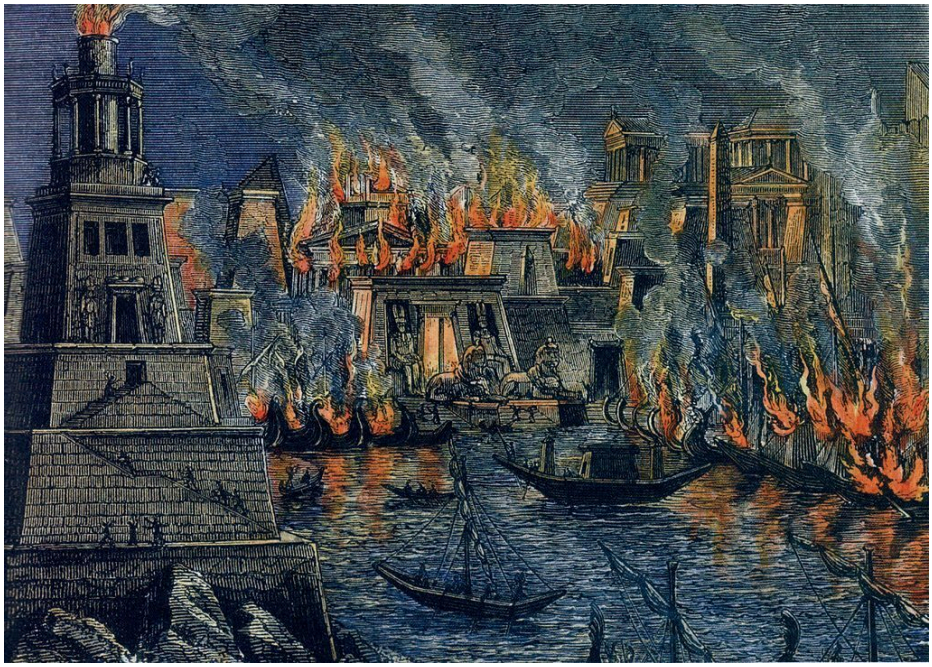
Ce cours est une introduction technique à la notion de communication sur Internet : quels protocoles et outils permettent de communiquer via Internet ?

I Centralisation

Stockage



Nineteenth-century artistic rendering of the Library of Alexandria



The fire of Alexandria

Indexation

The Mundaneum

Drawers in the Mundaneum

Accès

IBM Mainframe computer 7090 (NASA, 1962)

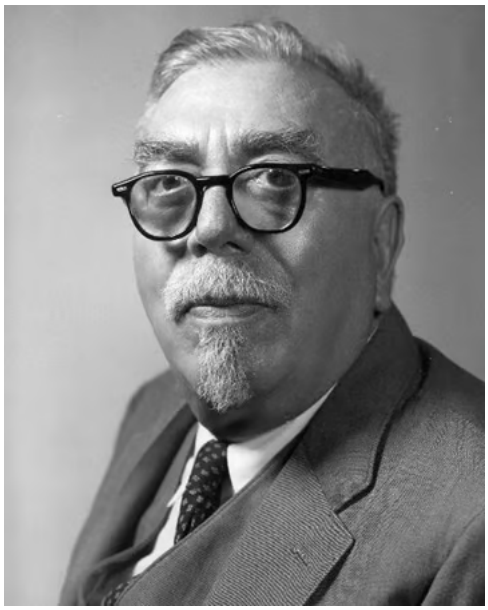
IBM 7090 console

II Aux origines du modèle décentralisé d'Internet

Enjeux militaire

La menace nucléaire plaide pour davantage de résilience.

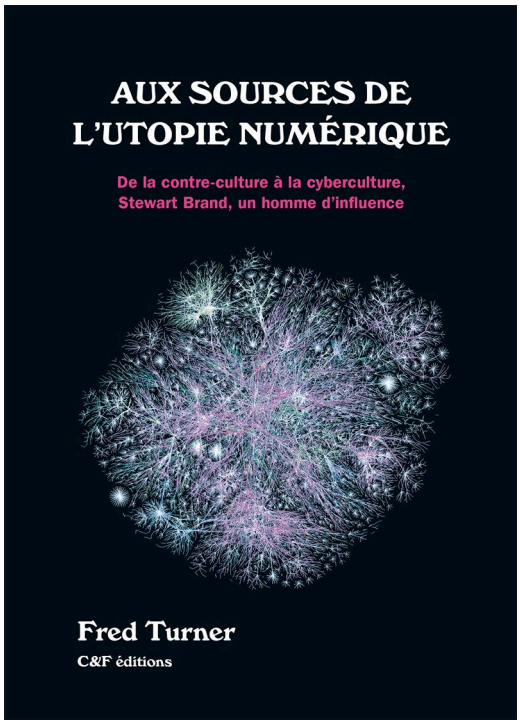
Théorie des systèmes



Norbert Wiener, à travers ses travaux sur la cybernétique, propose une nouvelle vision du monde, dont l'information et la communication sont les éléments fondamentaux.

Contre-culture

La contre-culture hippie rejette la bureaucratie et la hiérarchie. Elle recherche des fonctionnements plus horizontaux.



Les échanges et influences mutuelles entre contre-culture et universitaires d'une part, universitaires et militaires d'autre part, sont nombreux.

III Principe de la communication sur Internet

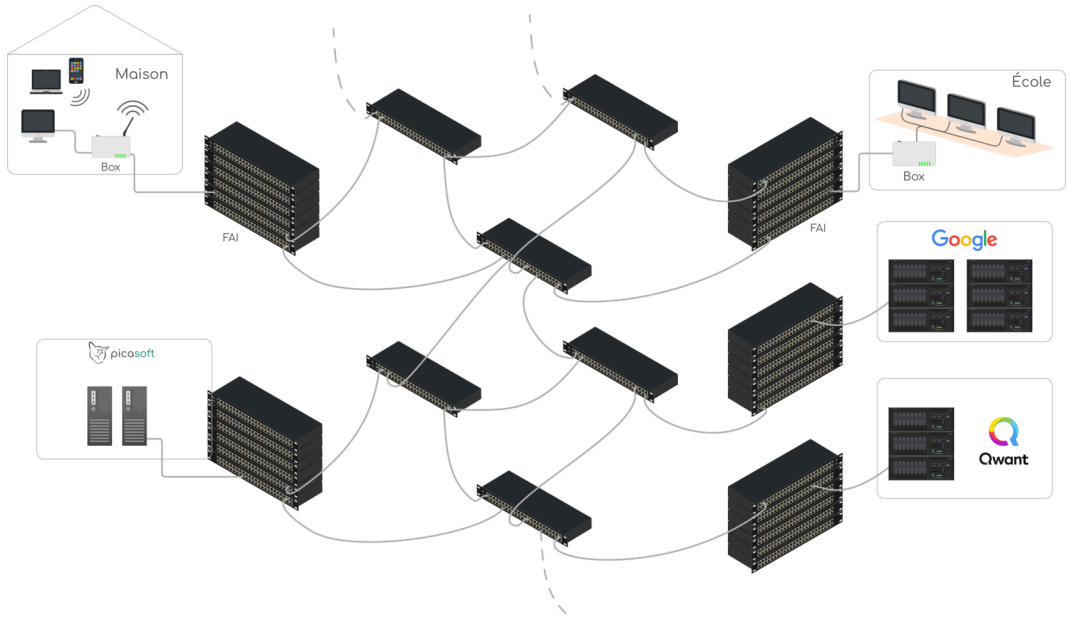
Une communication via Internet implique :

- les ordinateurs des personnes qui communiquent, on les appelle des **clients** ;
- les **routeurs**, ce sont des ordinateurs dont le rôle et de faire passer l'information d'un ordinateur à un autre sur Internet ;
- en général des **serveurs**, ce sont des ordinateurs en charge de gérer la communication pour un ensemble de clients ;
- des moyens physiques pour **relier** ces ordinateurs (câbles et ondes) ;
- des **protocoles** qui permettent à des logiciels installés sur ces machines de communiquer les unes avec les autres.

Internet est un vaste réseau d'ordinateurs qui communiquent entre eux Fondamental

Pour que les **clients**, les **routeurs** et les **serveurs** puissent communiquer entre eux, ils faut qu'ils soient **reliés** physiquement et qu'ils hébergent des logiciels qui respectent des **protocoles** communs.

Exemple



Des machines et des logiciels (confusing !)

Complément

Les termes de serveur et de client sont utilisés pour des machines, mais également pour des logiciels.

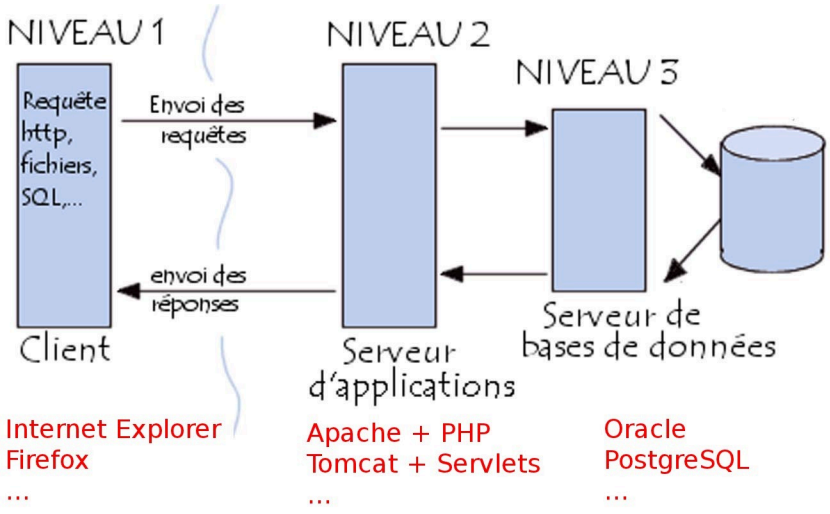
Ainsi Apache est un serveur web, c'est à dire un logiciel destiné à tourner sur une machine serveur et donc à répondre à des requêtes HTTP, tandis que Firefox est un client web, c'est à dire un logiciel destiné à tourner sur une machine cliente et donc à envoyer des requêtes HTTP.

Des clients qui sont aussi des serveurs...

Complément

Il est possible qu'une machine soit à la fois un client et un serveur :

- Si vous avez un ordinateur chez vous, vous pouvez tout à fait héberger un serveur web (qui est accessible quand la machine est allumée et connectée à Internet) et vous pouvez aussi vous en servir comme ordinateur personnel (donc par exemple pour consulter des sites web).
- Dans les architectures à plusieurs couches (architectures *n-tiers*) un ordinateur peut à la fois répondre à des requêtes et en formuler : il est ainsi courant qu'un serveur web ait besoin de demander des informations à un serveur de base de données, dans ce cas la machine est serveur du point de vue du navigateur qui lui demande une page web et client du point de vue de la base de données à laquelle il demande les informations (qui vont lui permettre de créer la page web à renvoyer au client).



Exemples d'architecture Web

Complément

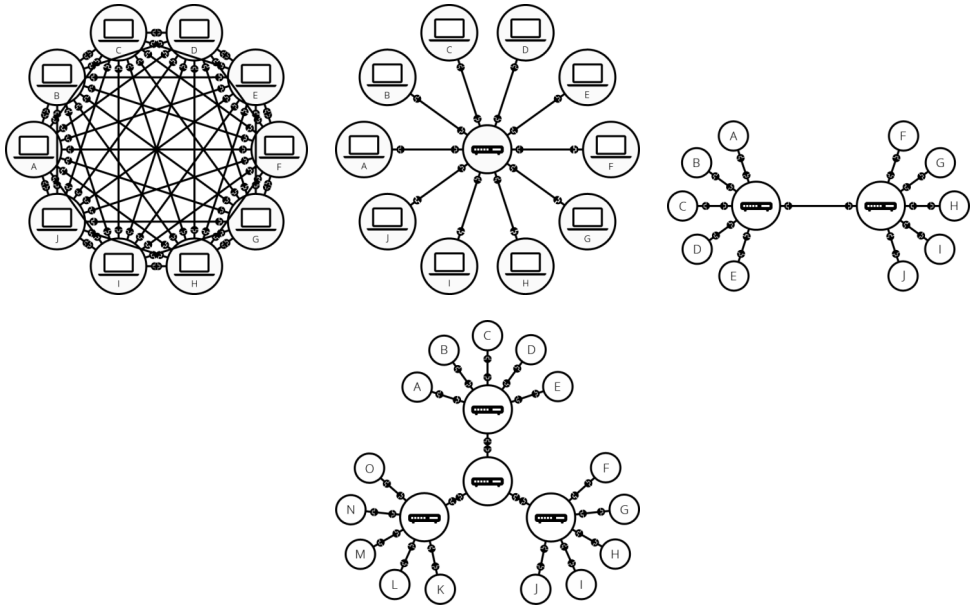
Comment fonctionne Internet ? (cf. p.35)

Les machines qui font fonctionner Internet (cf. p.37)

Le fonctionnement de l'Internet Complément

https://developer.mozilla.org/fr/docs/Learn/Common_questions/How_does_the_Internet_work

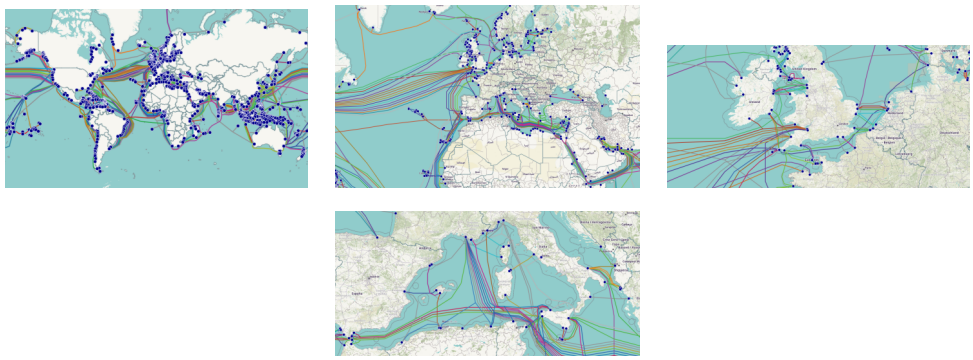
IV Infrastructure réseau d'Internet



Internet est un réseau de réseaux

Fondamental

Les composants fondamentaux du réseau Internet sont les câbles très haut débit par lesquels transitent les données.



Carte des câbles sous-marins du réseau Internet

Complément

Les acteurs du réseau : FAI, FCA, hébergeur, transitaires (cf. p.41)

V Ordinateurs clients

Client Az Définition

Un ordinateur client est un ordinateur qui demande un service à un autre ordinateur, par exemple : envoie ce mail à telle adresse de ma part.

Tout ordinateur qui utilise des services sur Internet est un client.

Envoyer un mail Exemple

Si j'envoie un mail à *contact@picasoft.net* depuis mon mail *stephane.crozat@utc.fr* :

- Mon ordinateur est le client, via le logiciel de mail Thunderbird il demande au serveur mail de l'UTC d'envoyer mon mail au serveur mail de Picasoft (il utilise pour cela le protocole SMTPS).
- Ensuite quand quelqu'un chez Picasoft consultera les mails du compte *contact*, il demandera à son serveur quels nouveaux mails il a reçu et il y aura le mien parmi ceux-là (il utilise pour cela le protocole IMAPS).

Demander une page web Exemple

Si je consulte la page page *https://fr.wikipedia.org/wiki/Internet* :

- Mon ordinateur est le client, via le navigateur web Firefox il demande au serveur web de Wikipédia de m'envoyer une page HTML (il utilise pour cela la requête GET du protocole HTTPS).
- Le serveur web de Wikipédia renvoie la page HTML demandée (il utilise pour cela une réponse HTTPS).

VI Ordinateurs serveurs

Serveur Az Définition

Un serveur est un ordinateur dont le rôle est de proposer des services, par exemple : stocker des pages web ou acheminer des mails.

Exemple

- Les serveurs sont souvent des ordinateurs sans écran et sans clavier situés dans des datacenters, par exemple Picacoft dispose de deux serveurs dans un datacenter de Toulouse géré par l'association Tetaneutral. Aucun membre de Picasoft n'a jamais rencontré physiquement ces machines.
- Chacun peut installer un serveur chez lui ou transformer son ordinateur personnel en serveur, par exemple *https://pi.crzt.fr* est un serveur web que j'héberge sur un Raspberry Pi chez moi.

https://tecnologiahabitual.blogspot.com/2014/07/mantenimiento-en-datacenters.html

Exemple de serveur Exemple

L'UTC propose un serveur pour envoyer des mails en utilisant le protocole SMTPS, il s'appelle *smtps.utc.fr*.

```
stph@mypc:~$ ping smtps.utc.fr
PING smtps.utc.fr (195.83.155.8) 56(84) bytes of data.
64 bytes from smtps.utc.fr (195.83.155.8): icmp_seq=1 ttl=62 time=0.454 ms
64 bytes from smtps.utc.fr (195.83.155.8): icmp_seq=2 ttl=62 time=0.752 ms
64 bytes from smtps.utc.fr (195.83.155.8): icmp_seq=3 ttl=62 time=0.739 ms
64 bytes from smtps.utc.fr (195.83.155.8): icmp_seq=4 ttl=62 time=0.726 ms
```

La commande *ping* permet de de tester l'accessibilité d'une machine sur Internet.

VII Ordinateurs routeurs

Routeur Az Définition

Un routeur est un ordinateur dont le rôle est d'acheminer des paquets d'information sur Internet, on peut voir cela comme un coursier.

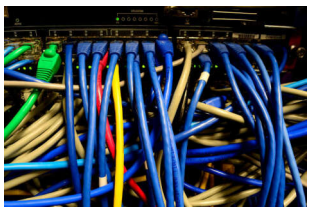
Remarque

Pour qu'un message se déplace d'un ordinateur à un autre sur Internet, il devra passer par plusieurs routeurs, de plus en plus proches de sa destination.

traceroute from pic.crzt.fr to www.utc.fr Exemple

```
1  h7.tetaneutral.net (91.224.148.1)
2  renater.par.franceix.net (37.49.236.19)
3  te0-0-0-1-ren-nr-compiegne-rtr-091.noc.renater.fr (193.51.177.53)
4  rrtp-vl221-te0-0-0-0-compiegne-rtr-011.noc.renater.fr (193.51.181.189)
5  * * *
6  * * *
7  gw-utc.utc.fr (195.83.154.242)
```

Exemple



Routeurs

VIII Protocoles réseaux TCP/IP

Afin de pouvoir communiquer entre elles toutes les machines d'Internet respectent des règles communes. Ces règles sont standardisées au sein de *protocoles* [p.31].

Le fonctionnement d'Internet repose fondamentalement sur les protocoles réseaux IP et TCP.

Protocole IP Az Définition

Le protocole (ou couche) IP gère les communications entre plusieurs machines d'un réseau.

Elle est basée sur l'**adressage**, le **routage** et la **fragmentation**.

- Pour s'adresser à une machine à laquelle on n'est pas directement connecté, il faut un moyen de l'identifier : c'est son **adresse**.
- Pour arriver jusqu'à cette machine, il faut trouver le chemin qui nous sépare d'elle : c'est le **routage**.
- Si le message à transmettre est trop gros pour les couches inférieures, il faut le découper : c'est la **fragmentation**.

Paquet IP Az Définition

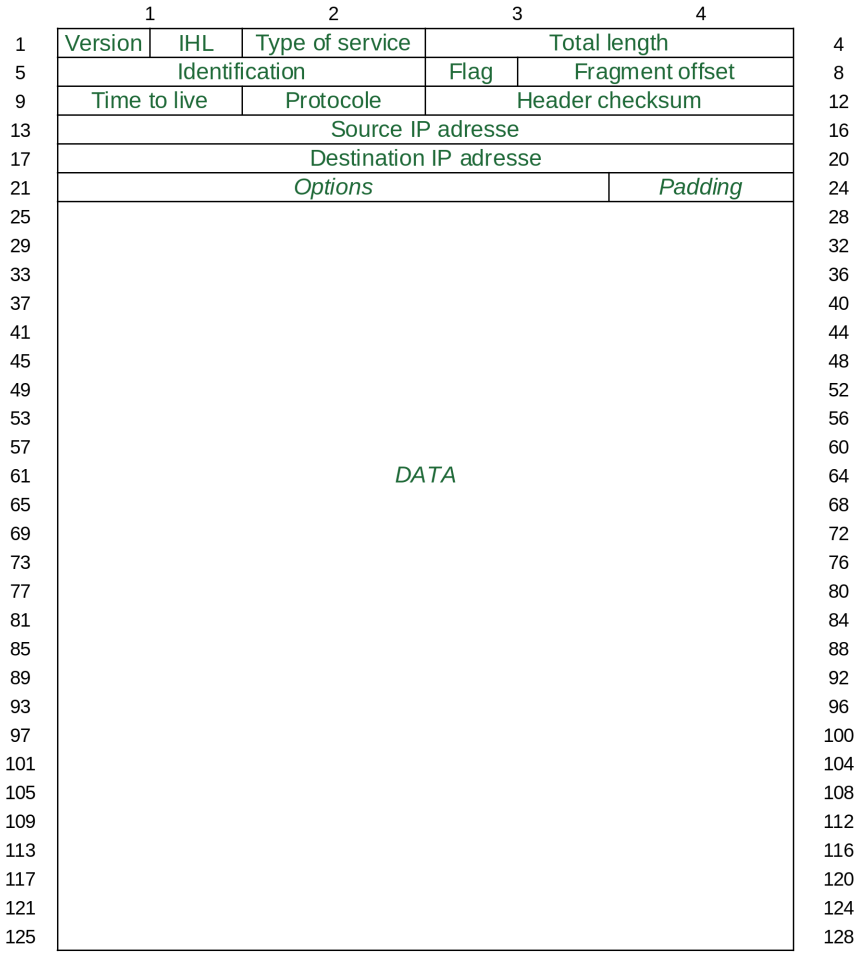
Les entités transmises par le protocole IP sont des **paquets** (ou datagramme).

Un message sur Internet est généralement découpé en plusieurs paquets qui sont transmis indépendants par le protocole IP.

Le protocole TCP est en charge de s'assurer que tous les paquets sont transmis et remis dans le bon ordre.

Exemple

Exemple



Codage d'un paquet IP

Protocole TCP

Définition

Le protocole (ou couche) TCP gère les communications **de bout en bout**.

Elle est basée sur la **correction des erreurs**, le **ré-ordonnancement** et le **contrôle de flux**.

- Plutôt que de raisonner en terme de paquets, elle raisonne en terme de messages complets.
- Elle est capable de **corriger des erreurs** dans la transmission des messages (congestion du réseau...) et de **ré-ordonnancer** des paquets qui auraient pris une route différente et seraient arrivés dans le mauvais ordre.
- Ses mécanismes de **contrôle de flux** permettent d'adapter la vitesse d'envoi des données à la capacité de réception du destinataire.

Message

Définition

Les entités transmises par la couche TCP sont des **messages**.

Remarque

Sur la base du protocole de communication TCP/IP on peut bâtir des applications, tels que HTTP pour consulter des fichiers sur le Web ou FTP pour transférer des fichiers.

Complément

Les couches réseaux du modèle OSI *(cf. p.44)*

IX Adresses IP et noms de domaine

Adresses IP

Az Définition

Une adresse IP est un **numéro d'identification d'une machine**.

Il est attribué de façon à chaque appareil relié à Internet.

Les adresses ont un format différent selon la version d'IP utilisée :

- pour IPv4, les adresses sont sur 4 octets (exemple : `192.168.3.200`),
- pour IPv6, les adresses sont sur 16 octets (exemple : `2001:0db8:0000:85a3:0000:0000:ac1f:8001`).

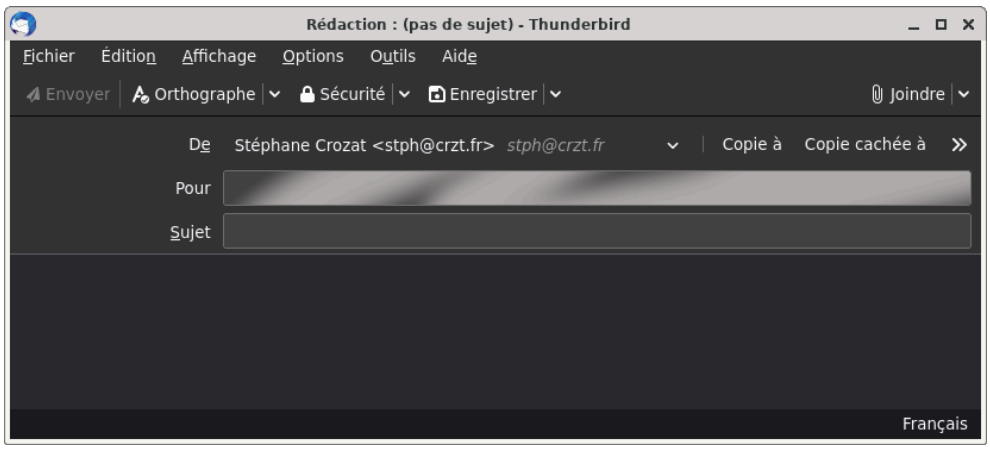
Nom de domaine

Az Définition

Un nom de domaine est une **adresse symbolique** permettant d'éviter de devoir retenir des adresses IP.

Ainsi, à un nom de domaine tel que `wikipedia.org` on peut associer une adresse telle que `91.198.174.192`.

Exemple



DNS

Az Définition

Un **serveur DNS** (pour *Domain Name System*) est un serveur qui a connaissance de la **correspondance entre noms de domaines et adresses IP**. Il est capable de répondre à des requêtes qui lui demandent de **résoudre** un nom de domaine en une adresse IP.

Complément

Adresses IP *(cf. p.50)*

Noms de domaine *(cf. p.52)*

X Protocoles applicatifs d'Internet : SSH, mail, Web...

Protocole Az Définition

Un protocole est un langage standard que tous les ordinateurs qui veulent communiquer entre eux respectent.

Les protocoles sont la condition de fonctionnement d'Internet, ils permettent à des ordinateurs différents qui utilisent des logiciels différents de communiquer entre eux.

Qui établit les protocoles ? Fondamental

Internet est géré de façon décentralisée, il n'y a pas une structure qui gère l'ensemble, les protocoles sont souvent le résultat de discussions ouvertes à toutes et tous.

En pratique il y a quelques organismes de standardisation, comme l'*IETF* p.32 qui publie les RFC ou le W3C qui standardise les langages du Web, comme HTML ou CSS.

Internet ≠ Web (métonymie) Attention

Le protocole principal du Web est HTTP.

Il existe d'autres protocoles sur Internet : FTP pour le transfert de fichier, SSH pour l'accès à des ordinateurs distants, SMTP et IMAP pour le mail, etc.

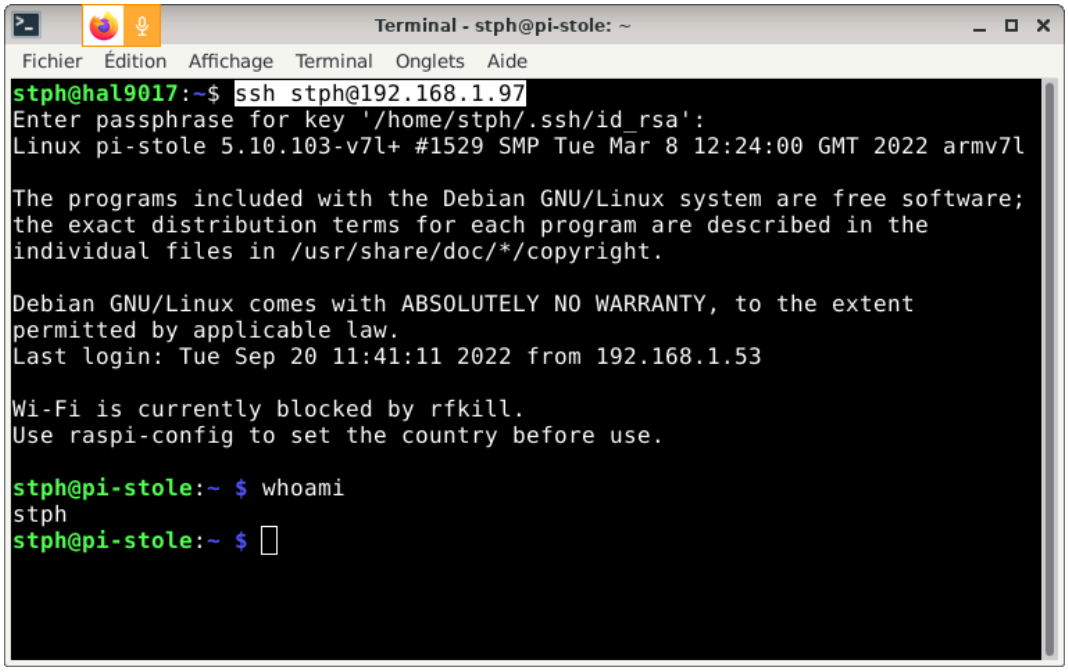
Il y a donc plusieurs applications sur Internet qui ne sont pas du Web.

Notons néanmoins que certaines applications non web à l'origine ont aussi une couche web aujourd'hui :

- si vous utilisez Thunderbird, vous faites du mail sur Internet (vous utilisez les protocoles STMP et IMAP), mais pas du Web ;
- si vous faites du webmail, vous faites du Web et du mail (HTTP, IMAP, SMTP) ;
- si vous communiquez via la messagerie de Twitter, vous faites du Web.

Connexion distante (1969) Exemple

- Naissance : 1969
- Protocole : Telnet (1969, RFC 15), SSH (1995, RFC 4251)



Mail (1971)

Exemple

- Naissance : 1971
- Protocoles : SMTP (1982, RFC 8211[1]) — POP (1985, RFC 937[2]) — IMAP (2003, RFC 3561[3])

Forum (1979)

Exemple

- Naissance : *usenet*, 1979
- Application web : Discourse (décentralisé, libre)
- → Médias sociaux

Messagerie instantanée (1988)

Exemple

- Naissance : commande Unix `talk` (années 1970)
- Protocole : IRC, 1988
- Applications web de messagerie instantanée :
  - Slack (centralisé, propriétaire)
  - Mattermost (décentralisé, libre)
  - Element/Matrix (décentralisé, libre, fédéré, chiffré E2E)

1. https://tools.ietf.org/html/rfc821
2. https://tools.ietf.org/html/rfc937
3. https://tools.ietf.org/html/rfc3501

Web (1990) Exemple

- Naissance : 1990
- Protocole : HTTP (1990, W3C, RFC 1945[4]) (et le langage de formatage de documents HTML)

RFC Complément

https://www.ietf.org/standards/rfcs[5]

https://www.bortzmeyer.org[6]

Complément

Internet et ses protocoles (cf. p.54)

FTP : File Transfer Protocol (cf. p.57)

HTTP : HyperText Transfert Protocol (cf. p.59)

SSH - Secure Shell (introduction) (cf. p.67)

4. https://www.w3.org/Protocols/rfc1945/rfc1945.txt
5. https://www.ietf.org/standards/rfcs/
6. https://www.bortzmeyer.org/

XI Les clients et les serveurs web parlent HTTP

Serveur web — Az Définition

Un serveur web est un logiciel qui **traite les requêtes HTTP** de clients.

Il est installé sur une machine hébergeant des documents (HTML, CSS, JavaScript, etc.) et est accessible depuis Internet.

Navigateur web — Az Définition

Un navigateur web est client HTTP qui a trois rôles.

1. Effectuer des requêtes HTTP (GET, POST, etc.).
2. Comprendre les réponses à ces requêtes.
3. Interpréter les fichiers web (HTML, CSS, JavaScript) retournés.

Méthodes HTTP — Az Définition

Les méthodes HTTP sont les commande que le client (navigateur) envoie au serveur.

Exemple

Les plus courantes sont :

- GET (pour demander une ressource, une page HTML par exemple)
- POST (pour transmettre des données au serveur, par exemple quand on remplit un formulaire en ligne)

Code de retour HTTP — Az Définition

Un code HTTP est un code à 3 chiffres qui est présent dans chaque réponse HTTP du serveur.

Exemple

Les plus courants sont :

- `200 OK` : « Tout est bon »,
- `401 Unauthorized`: « Vous n'avez pas le droit d'accéder à cette ressource, il faut une authentification »,
- `404 Not Found`: « Ressource non trouvée »,
- `500 Internal Server Error`: « Erreur du serveur ».

Complément

Serveurs web (cf. p.62)

Navigateurs web (cf. p.64)

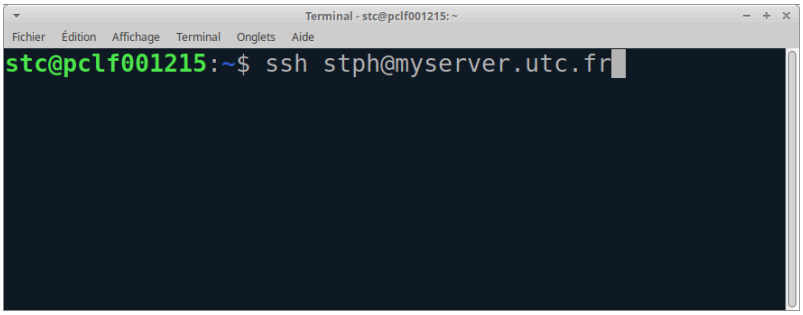
URL : Uniform Resource Locator (cf. p.72)

Codes de retour HTTP (cf. p.73)

XII Quelques logiciels qui font fonctionner Internet

Logiciels côté clients

- Navigateur web : Mozilla Firefox ou Google Chrome (clients HTTP)
- Client Mail : Mozilla Thunderbird (client IMAP ou POP)



Client SSH sous Linux

Logiciels côté serveur

- Serveur web : Apache, Nginx
- Serveur mail : Postfix (SMTP), Dovecot (IMAP)

Typologie de logiciels ⊕ Complément

- propriétaire + centralisé (Zoom, Discord, Youtube)
- libre + centralisé (Signal)
- libre + décentralisé (Mattermost, BigBlueButton, Etherpad)
- libre + décentralisé + fédéré (Mastodon, Element, Mobilizon, Peertube)

XIII Synthèse

Questions

- Quelles sont les différences entre Internet et le web ?
- Retracer le chemin d'une requête entre un navigateur web et un site web en nommant les principaux protocoles utilisés.
- Est-ce possible pour un dirigeant de couper l'accès à un site web ? Citer des stratégies et leurs limites.

Concepts

- Contre-culture et Internet
- Décentralisation d'Internet
- Standardisation des protocoles

XIV Cours-émission

Première partie (jusque 35:40)

Glossaire

Protocole

Un **protocole informatique** est un ensemble de règles techniques permettant à des ordinateurs d'échanger des informations via un langage commun.

Abréviations

IETF : Internet Engineering Task Force

Crédits des ressources

Nineteenth-century artistic rendering of the Library of Alexandria p. 5
Licence : Domaine Public - O. Von Corven, fr.wikipedia.org/wiki/Bibliothèque d'Alexandrie[7]

The fire of Alexandria p. 6
Licence : Domaine Public - Hermann Göll, 1876 (gravure sur bois).

The Mundaneum p. 6
Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions - Stefaan Van der Biest, fr.wikipedia.org[8]

Drawers in the Mundaneum p. 7
Par fdecomite — Flickr: Drawers, CC BY 2.0, commons.wikimedia.org/w/index.php?curid=15888697

IBM Mainframe computer 7090 (NASA, 1962) p. 7
Universel - Transfert dans le Domaine Public - en.wikipedia.org/wiki/IBM_7090

IBM 7090 console p. 8
Universel - Transfert dans le Domaine Public - en.wikipedia.org/wiki/IBM_7090

La menace nucléaire plaide pour davantage de résilience. p. 9
Universel - Transfert dans le Domaine Public - Burnt Pineapple Productions — flickr.com[9]

Norbert Wiener, à travers ses travaux sur la cybernétique, propose une nouvelle vision du monde, dont l'information et la communication sont les éléments fondamentaux. p. 9
Universel - Transfert dans le Domaine Public - Garry Olsh[10]

La contre-culture hippie rejette la bureaucratie et la hiérarchie. Elle recherche des fonctionnements plus horizontaux. p. 10
Attribution - Partage dans les Mêmes Conditions - Derek Redmond and Paul Campbell

p. 11, 37
Universel - Transfert dans le Domaine Public - Picasoft

Exemples d'architecture Web p. 12
Licence de documentation libre GNU - Pillou, commentcamarche.net

p. 36
Universel - Transfert dans le Domaine Public - Picasoft

p. 36
Universel - Transfert dans le Domaine Public - Picasoft

7. https://fr.wikipedia.org/wiki/Biblioth%C3%A8que_d%27Alexandrie
8. https://fr.wikipedia.org/wiki/Mundaneum#/media/Fichier:Mundaneum_Bergen_14.jpg
9. https://www.flickr.com
10. https://commons.wikimedia.org/w/index.php?title=User:Garry_Olsh&action=edit&redlink=1

p. 37
Universel - Transfert dans le Domaine Public - Picasoft

Clients, box et serveurs p. 37
Attribution - Picasoft

Réseau avec deux clients et wikipedia.org comme extrémités p. 38
Attribution - Partage dans les Mêmes Conditions - Ferran91 — Article Wikipédia en anglais (Internet access diagram)

traceroute www.icann.org. depuis la France p. 41
https://stefansundin.github.io/traceroute-mapper

Rhizome, un FAI de la FFDN — rhizome-fai.net p. 43
db.ffdn.org

https://tecnologiahabitual.blogspot.com/2014/07/mantenimiento-en-datacenters.html p. 16
Attribution - Partage dans les Mêmes Conditions - Mariano Simón Lázaro

Codage d'un paquet IP p. 19
Attribution - Partage dans les Mêmes Conditions - Stéphane Crozat

Modèle OSI p. 45
Licence : Domaine Public

Répartition géographique des différents RIR p. 52
Attribution - Partage dans les Mêmes Conditions - Dork, Canuckguy et al., Sémhur

Contenu d'une requête HTTP et de sa réponse p. 62
Attribution - Partage dans les Mêmes Conditions - framasoft.org

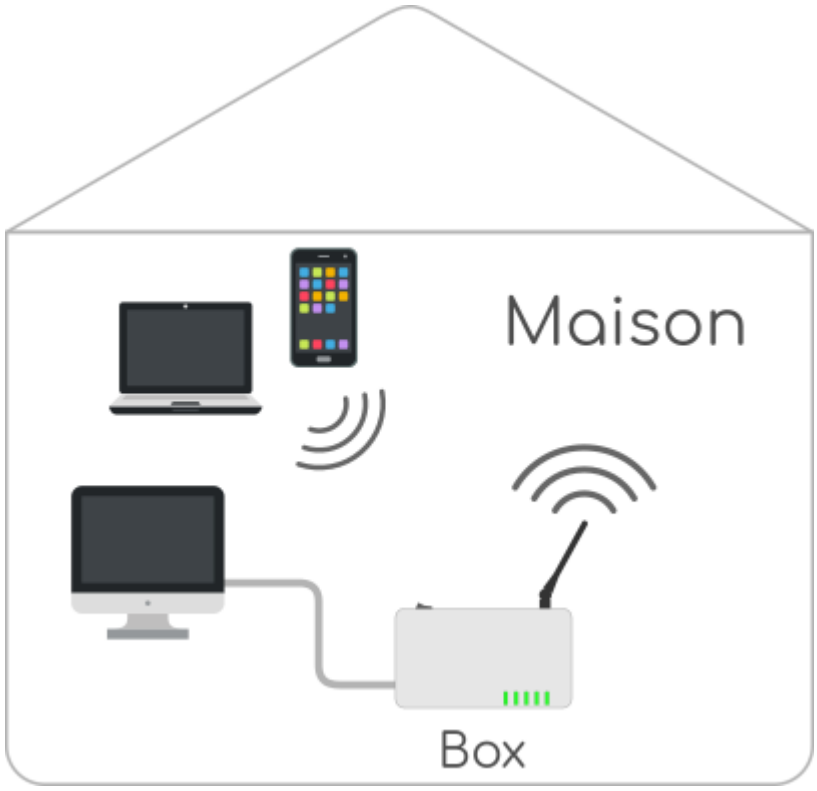
Outils de développement Firefox (onglet requête) p. 67
Licence : Domaine Public

Décomposition d'une URL p. 72
Licence : Domaine Public

Contenus annexes

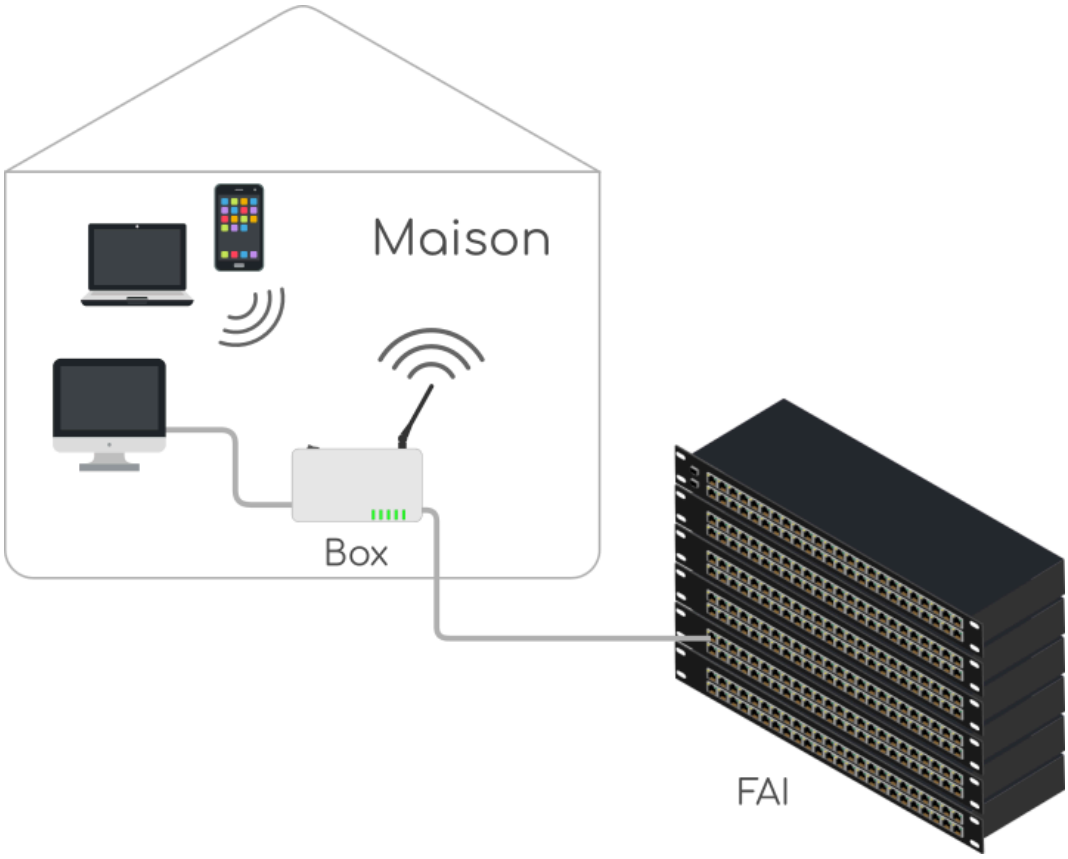
1. Comment fonctionne Internet ?

La box, le point de sortie de la maison



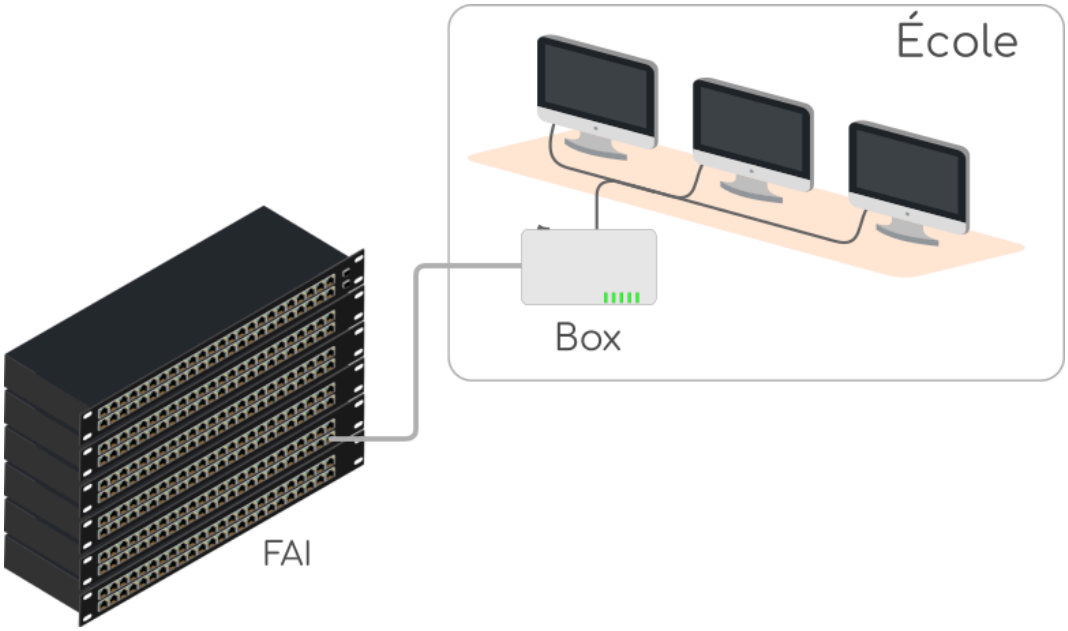
Pour se connecter depuis son domicile on relie les ordinateurs de la maison à sa box (on peut les relier par un fil ou bien en wifi).

Le FAI, le point d'entrée sur Internet



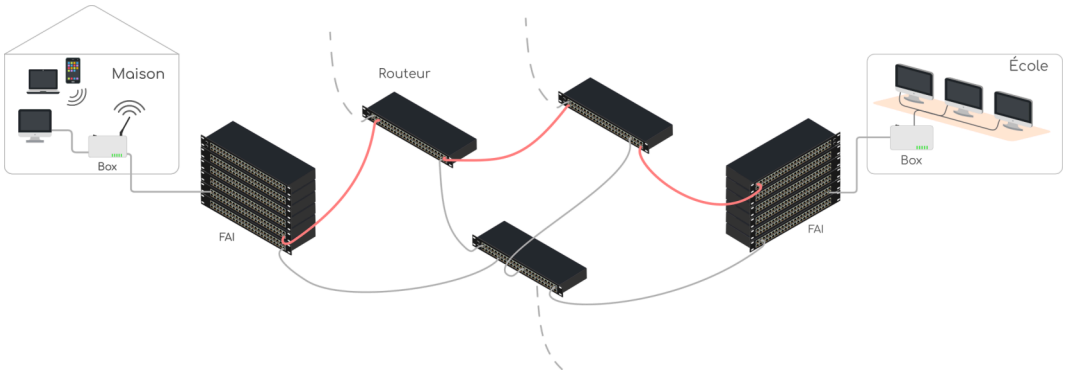
La box est reliée au FAI qui est lui même connecté à Internet. Toutes nos communications sur Internet passent donc par la box et le FAI.

L'école, comme à la maison (ou presque)



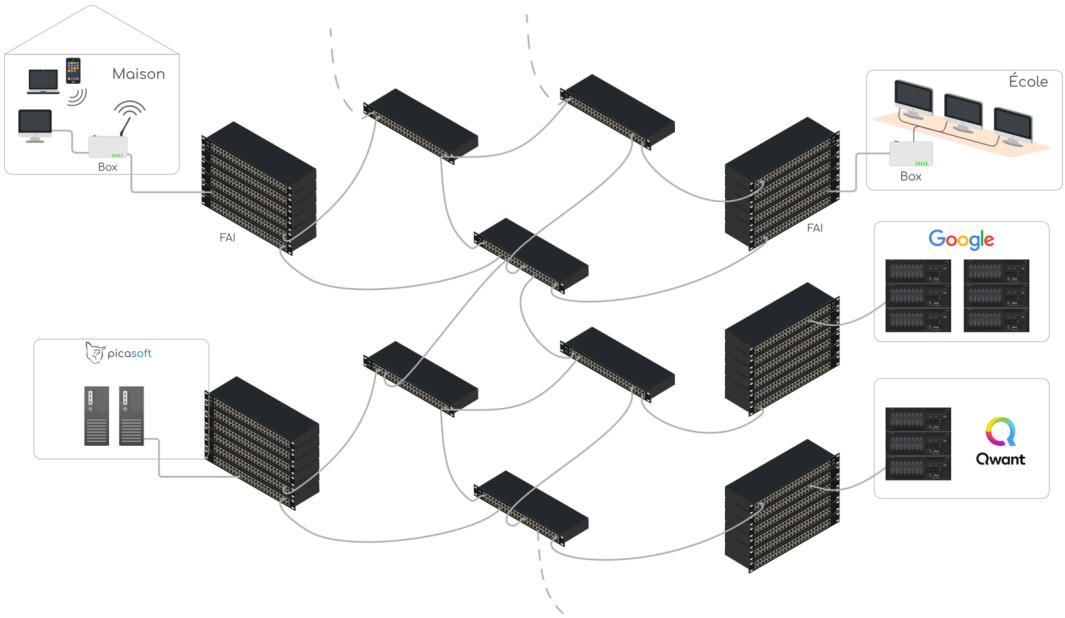
Le réseau de l'école ressemble à celui de la maison, simplement la box n'est pas toujours tout à fait la même.

Les routeurs



Il n'existe pas de fil qui relie directement le FAI de la maison et celui de l'école, sur Internet les communications sont acheminées via des ordinateurs dont c'est le rôle, les routeurs.

Internet

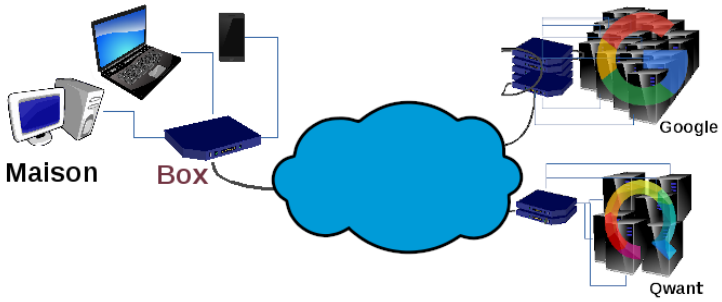


Et tout Internet fonctionne de la même façon, différents FAI se communiquent les informations que nous échangeons via des routeurs.

2. Les machines qui font fonctionner Internet

Aux extrémités du réseau : serveurs et clients

Az Définition



Clients, box et serveurs

Aux extrémités d'un chemin sur Internet il y a deux acteurs : l'utilisateur et le FCA (Fournisseur de Contenu et d'Applications).

- L'utilisateur a besoin d'un ordinateur et d'une box Internet. L'ordinateur doit disposer d'une carte réseau qui s'interface avec un câble Ethernet ou émet des ondes électromagnétiques (Wi-Fi, réseau cellulaire, etc.). Un tel ordinateur est appelé un **client**.
- Le FCA va quant à lui disposer d'une machine qui héberge son contenu capable de traiter les requêtes des clients. Pour recevoir ces requêtes, cette machine doit être reliée à Internet et son adresse doit être connue des clients. Une telle machine est appelée **serveur**.

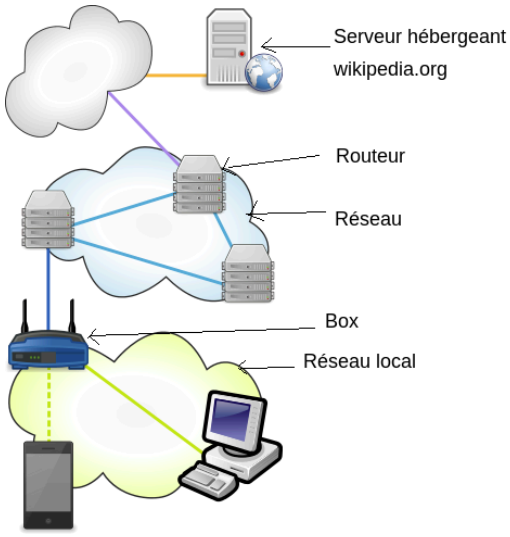
Au milieu du réseau : routeurs

Az Définition

Pour relier ces extrémités, Internet a besoin d'un autre type d'ordinateur : les **routeurs**.

- Un routeur est un ordinateur qui transmet les données qu'il reçoit (venant d'un réseau quelconque) à d'autres routeurs ou aux destinataires quand cela est possible.
- Pour savoir vers où transférer les données, le routeur maintient une **table de routage** qui contient les destinations possibles, appelées **routes**.

Exemple



Réseau avec deux clients et wikipedia.org comme extrémités

Comme une lettre à la poste

Remarque

Une analogie avec le courrier postal permet de mieux comprendre le routage :

- Les clients sont les personnes voulant envoyer une lettre,
- Les serveurs sont les destinataires de la lettre,
- Les facteurs sont les câbles par lesquels transitent les données,
- Les centres de tri sont les routeurs qui orientent les lettres vers d'autres centres de tri ou vers le bureau de poste final.

Comme une lettre à la poste

Exemple

Soit un envoi de courrier depuis Alice habitant à Compiègne vers Bob habitant à Soissons.

1. Le courrier part de chez Alice et est acheminé jusqu'au centre de tri de Compiègne.
2. Le courrier est acheminé depuis le centre de tri de Compiègne vers le centre de tri Aisne2.
3. Le courrier est acheminé depuis le centre de tri Aisne2 vers le centre de tri de Soissons.
4. Le courrier est acheminé depuis le centre de tri de Soissons jusque chez Bob.

Destination	Envoyer vers
Senlis	Centre de tri Senlis
Soissons	Centre de tri Aisne2
Paris	Centre de tri IDF1
Nîmes	Centre de tri Sud2
...	

Extrait de la table de routage du centre de tri de Compiègne

Destination	Envoyer vers
Senlis	Centre de tri Oise1
Compiègne	Centre de tri Oise2
Laon	Centre de tri Laon
Soissons	Centre de tri Soissons
Paris	Centre de tri IDF1
...	

Extrait de la table de routage du centre de tri Aisne2

Fondamental

L'interconnexion de tous les réseaux d'Internet est réalisée par des routeurs qui ne connaissent que leur environnement proche.

Les box Internet sont des routeurs

Remarque

Elles gèrent un **réseau local** composé des appareils de l'utilisateur (ordinateur, smartphone, etc.) et transmet les données en provenance de ces appareils à d'autres routeurs.

Depuis le réseau Internet, seules les box Internet sont visibles : les appareils du réseau local ne peuvent pas être contactés directement.

Réseau de Diffusion de Contenu (CDN)

Remarque

Afin d'améliorer les chemins empruntés pour accéder à du contenu sur Internet, un FCA peut conserver des copies de ses contenus dans un **Réseau de Diffusion de Contenu**, plus communément appelé CDN (pour *Content Delivery Network*).

Un CDN est constitué de serveurs situés à des points stratégiques qui **répliquent** le contenu du serveur d'origine. Par exemple, un FCA mexicain pourrait utiliser un CDN présent en France pour accélérer l'accès aux données pour les utilisateurs français et libérer de l'espace sur les fibres optiques entre la France et les États-Unis.

Identifier les routes utilisées

Méthode

Il est possible d'identifier le parcours exact d'une requête avec les commandes **traceroute** (disponible sous GNU/Linux et MacOS) et **tracert** (disponible sous Windows).

Identifier les route utilisées

Exemple

Pour connaître la route empruntée par une requête en direction du site de l'ICANN (https://www.icann.org), on peut exécuter la commande suivante :

```
traceroute www.icann.org.
```

```
traceroute to www.icann.org. (192.0.32.7), 30 hops max, 60 byte packets
 1  * * *
 2  10.24.0.1 (10.24.0.1)  72.344 ms  72.283 ms  72.305 ms
 3  192.168.253.30 (192.168.253.30)  72.306 ms  72.302 ms  77.540 ms
 4  192.168.255.6 (192.168.255.6)  77.569 ms  77.566 ms  78.655 ms
 5  ppp18-metz.isdnet.net (194.149.166.145)  78.681 ms  80.408 ms  80.434 ms
 6  ppp6-metz.isdnet.net (194.149.166.133)  78.654 ms  38.479 ms  38.291 ms
 7  * * *
 8  be4204.ccr32.par04.atlas.cogentco.com (149.11.115.13)  42.493 ms  42.424 ms
42.369 ms
 9  be2067.ccr31.par04.atlas.cogentco.com (130.117.3.33)  41.388 ms  41.284 ms
41.180 ms
10  ae-29.r04.parsfr01.fr.bb.gin.ntt.net (129.250.8.125)  39.545 ms  39.478 ms
39.708 ms
11  ae-23.r24.amstnl02.nl.bb.gin.ntt.net (129.250.4.137)  49.371 ms  40.029 ms
42.071 ms
12  ae-3.r25.amstnl02.nl.bb.gin.ntt.net (129.250.4.69)  39.909 ms  46.701 ms
49.333 ms
13  ae-5.r23.asbnva02.us.bb.gin.ntt.net (129.250.6.162)  139.280 ms  145.858 ms
140.580 ms
14  ae-0.r22.asbnva02.us.bb.gin.ntt.net (129.250.3.84)  126.531 ms  123.791 ms
115.589 ms
15  ae-5.r23.lsanca07.us.bb.gin.ntt.net (129.250.3.189)  185.790 ms  177.418 ms
183.428 ms
```

```
16  ae-5.r01.lsanca20.us.bb.gin.ntt.net (129.250.6.49)  180.727 ms  204.185 ms
    201.918 ms
17  ge-100-0-0-16.r01.lsanca20.us.ce.gin.ntt.net (204.1.254.150)  199.037 ms
    186.855 ms  185.762 ms
18  www.icann.org (192.0.32.7)  187.313 ms  239.217 ms  232.851 ms
```

La dernière ligne correspond au serveur de l'ICANN ; les autres lignes correspondent aux routeurs intermédiaires (adresse IP et nom de domaine si disponible).

Pour fonctionner, le programme **traceroute** suppose que tous les routeurs intermédiaires enverront un message au client pour leur indiquer que leur requête a été traitée par eux. Les symboles * * * correspondent aux routeurs ne souhaitant pas décliner leur identité pour des raisons de sécurité. En effet, la sortie de traceroute permet de deviner la table de routage des routeurs, ce que certains préfèrent éviter.

Visualisez les routes sur le globe !

Complément

Vous pouvez visualiser les routes prises par vos requêtes à l'aide du site *Traceroute mapper* : stefansundin.github.io/traceroute-mapper/

traceroute www.icann.org. depuis la France

À retenir

- Un client est un ordinateur voulant accéder à des contenus et applications via Internet.
- Un serveur est un ordinateur qui héberge le contenu et les applications disponibles via Internet.
- Un routeur s'occupe de transférer la requête à un autre routeur voisin.
- L'ensemble des réseaux de routeurs forme le réseau Internet.

3. Les acteurs du réseau : FAI, FCA, hébergeur, transitaires

Carte des câbles sous-marins du réseau Internet

Les composants fondamentaux du réseau Internet sont les câbles très haut débit par lesquels transitent les données.

Cette carte montre les câbles sous-marins qui relient les différents continents. Si leur nombre est assez faible, la quantité de données échangées est très élevée.

FAI Az Définition

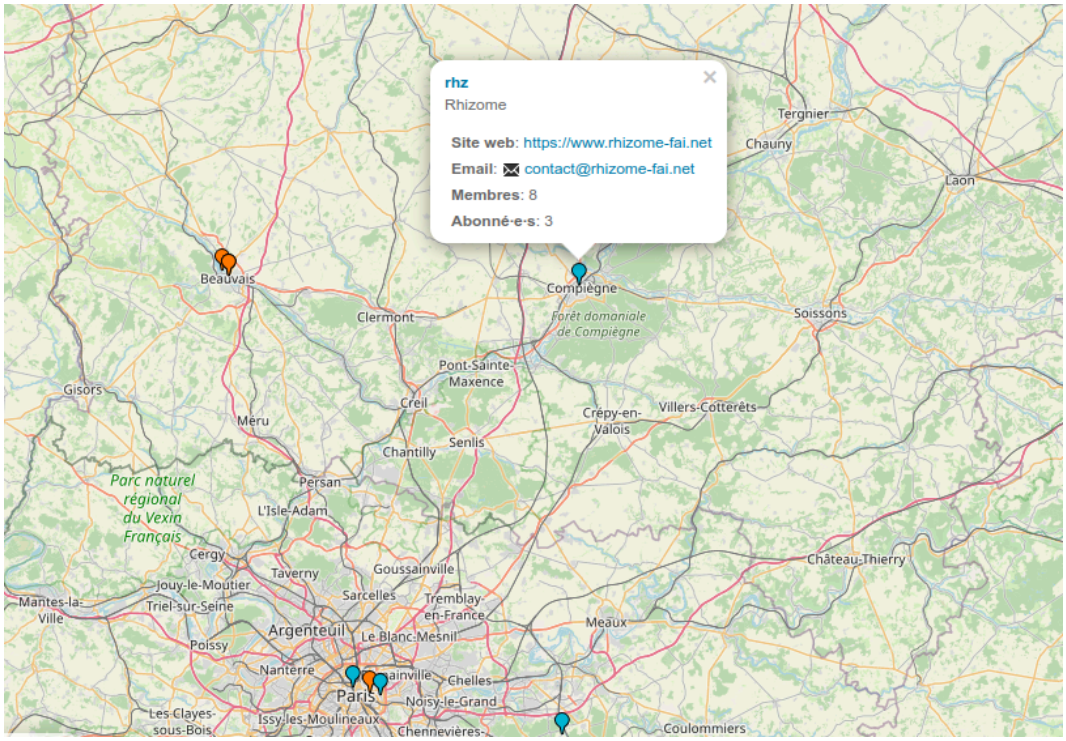
Un FAI (pour Fournisseur d'Accès à Internet) est un acteur qui propose une connexion Internet à ses utilisateurs. En anglais, on parle d'*Internet Service Provider* (abrégé ISP). Les FAI sont le dernier intermédiaire entre les câbles composant le réseau et les utilisateurs finaux.

Plusieurs types de FAI 💬 Remarque

- Les FAI existent sous forme d'entreprises ou d'association. Les entreprises en France représentent plus de 95% du marché. On y retrouve Orange, Bouygues, SFR, Free, etc.
- Moins connus mais plus nombreux, les FAI associatifs proposent une connexion Internet à leurs adhérents. En France, la FFDN[11] est une fédération qui regroupe plus de 30 FAI associatifs partageant des valeurs de « bénévolat, solidarité, fonctionnement démocratique et à but non lucratif ».

11. https://www.ffdn.org/

Exemple



Rhizome, un FAI de la FFDN — rhizome-fai.net

FCA

Az Définition

Un FCA (pour Fournisseur de Contenu et d'Applications) est un acteur qui propose du contenu accessible via Internet et font appels aux autres acteurs pour acheminer ce contenu vers les utilisateurs. Tous les acteurs proposant un site web sont des FCA.

Synonyme : fournisseur de service

FCA

Exemple

- Wikipédia, qui propose le site éponyme,
- Framasoft, qui propose un service de visioconférence appelé Framatalk,
- ProtonMail, qui propose un service de mails,
- Les chaînes de télévision par Internet.

Serveur Internet

Az Définition

Un **serveur** Internet est un ordinateur accessible via Internet qui rend des **services** aux utilisateurs. Il **héberge** du contenu (pages web, mails, vidéo, etc.). On dit qu'il **sert** ces contenus aux utilisateurs en réponse à des demande d'accès, ou **requêtes**.

Hébergeur Az Définition

Un hébergeur est le propriétaire de serveurs utilisés par un ou plusieurs FCA pour distribuer son contenu. Si le FCA est aussi le propriétaire de ses serveurs, il est à la fois FCA et hébergeur.

En France Exemple

Les entreprises françaises Gandi, OVH ou Scaleway sont des hébergeurs : elles permettent aux particuliers et entreprises de louer des serveurs et d'héberger du contenu.

Transitaire Az Définition

Un transitaire est un acteur chargé d'acheminer les données entre différentes parties du globe via le réseau physique. Ils sont en charge du cœur de l'infrastructure nécessaire à la circulation des données. Les transitaires sont des intermédiaires entre les FCA et les FAI connus du grand public.

France Exemple

En France, on peut citer France-IX, Zayo France, Nerim, Metro Optic, Open Transit, etc. Ces noms sont inconnus du grand public, car les clients des transitaires sont les FAI ou d'autres acteurs, mais pas les citoyens. Certaines entreprises sont à la fois des transitaires et des FAI, comme Orange avec Open Transit.

À retenir

Le réseau physique d'Internet est composé d'une multitude d'acteurs qui travaillent en commun pour interconnecter leurs réseaux respectifs.

Les principaux acteurs sont :

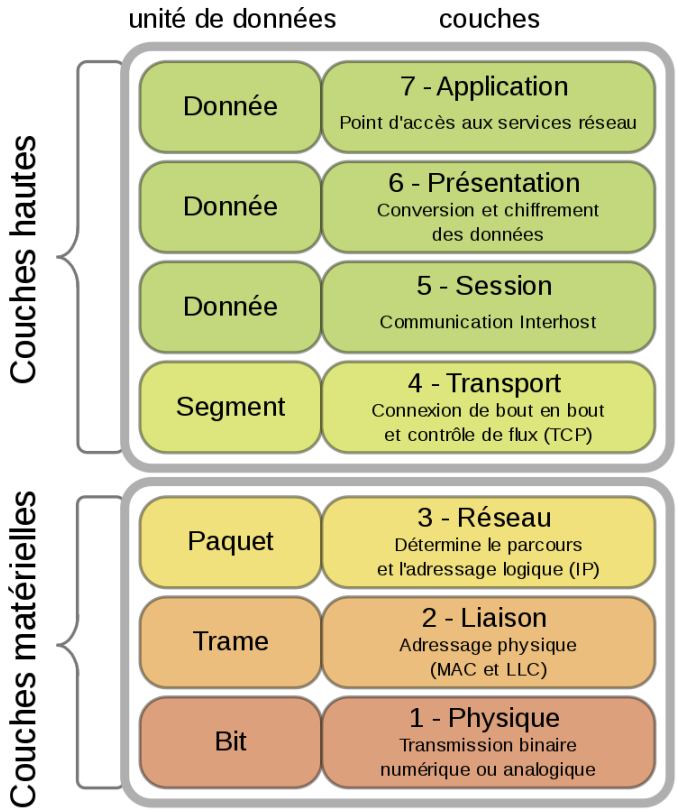
- les Fournisseur d'Accès à Internet (FAI)
- les pour Fournisseur de Contenu et d'Applications (FCA)
- les hébergeurs (qui gère les serveurs de FCA)
- les transitaires (intermédiaires entre les FAI et le les FCA)

4. Les couches réseaux du modèle OSI

Modèle OSI Az Définition

Le **modèle OSI** (pour *Open Systems Interconnection*) est une norme de communication en réseau de tous les systèmes informatiques.

C'est un modèle conceptuel qui décrit le fonctionnement théorique de n'importe quel réseau en 7 couches.



Modèle OSI

Une norme... pour créer des normes

Remarque

Le modèle complet est décrit par la norme ISO7498 et porte le nom de "Modèle basique de référence pour l'interconnexion des systèmes ouverts". Cette norme spécifie un cadre général propice à la création de **normes futures**.

En d'autres termes, le modèle OSI n'impose aucune technologie physique ou logicielle, mais définit une architecture que les technologies doivent respecter afin qu'un réseau global puisse émerger.

Le découpage en couche permet de "séparer les préoccupations", c'est-à-dire de traiter un aspect particulier de la problématique générale (identifier une machine, envoyer des données, formater les données, etc.)

Service, protocole et interface

Az Définition

Dans la terminologie OSI, chaque couche est délimitée par un **service**, un **protocole** et une **interface**.

- Un service correspond aux fonctions réalisables par une couche : c'est un ensemble de commandes, ou **primitives**.
- Un protocole correspond aux règles d'échange entre les couches, dans le but de réaliser un service.
- L'interface est le point d'accès concret au service. Elle spécifie comment utiliser le protocole permettant d'accéder au service.

Les couches basses **rendent des services** aux couches plus hautes.

Une cuisine et des étagères

Exemple

Le modèle OSI peut être difficile à conceptualiser. On peut se le représenter grossièrement comme le placard d'une cuisine :

- Chaque étagère (couche) contient des objets ayant des buts communs (protocoles) : de haut en bas, l'étagère des aliments, l'étagère des ustensiles, l'étagère des plats, l'étagère des condiments, etc.
- Il y a un ordre : on commence par sélectionner les aliments, les découper, puis les faire cuire et enfin assaisonner le plat.
- On peut dire que chaque étagère plus basse **rend service** aux étagères les plus hautes : la casserole va faire cuire les aliments, les condiments vont assaisonner le plat, etc.
- L'interface dans ce cas est l'être humain qui cuisine : il interface les objets des différentes étagères.

Niveau 1 : couche physique

Fondamental

La couche physique est chargée d'émettre et de recevoir des données brutes entre deux machines.

Plus compliqué qu'il n'y paraît...

Exemple

Si envoyer un 0 ou un 1 à une machine peut paraître trivial, il existe une grande diversité de techniques pour ce faire : fibre optique, câble Ethernet, ondes électromagnétiques, ondes radio, etc., et pourquoi pas des écriteaux tenus par un humain ?

Niveau 2 : couche liaison de données

Fondamental

Cette couche gère la communication entre deux machines directement connectées entre elles. Elle répond à plusieurs questions problématiques de la couche physique : où commence un message donné et où s'arrête-t-il ? Un support physique n'est jamais parfait, comment détecter les erreurs ? Peut-on réparer les erreurs ?

Pour ce faire, elle structure les données transmises à l'aide de marqueurs (début, fin, etc.), de champs (ici se trouvent les données, ici se trouve l'identifiant du destinataire, etc.) et elle met en place des mécanismes de contrôle.

Les entités transmises par la couche 2 sont des **trames**.

Détection d'erreur

Exemple

En supposant que la couche 2 transmette une suite de nombres, elle pourra rajouter la somme de ces nombres à la fin du message. Le destinataire pourra refaire cette somme et vérifier qu'il n'y a pas eu d'erreur de transmission si le résultat est exact.

Niveau 3 : couche réseau

Fondamental

La couche réseau gère les communications entre plusieurs machines d'un même réseau.

Elle exploite les notions d'**adressage**, de **routage**, de **fragmentation**.

- Pour s'adresser à une machine à laquelle on n'est pas directement connecté, il faut un moyen de l'identifier : c'est son adresse.
- Pour arriver jusqu'à cette machine, il faut trouver le chemin qui nous sépare d'elle : c'est le routage.
- Si le message à transmettre est trop gros pour les couches inférieures, il faut le découper : c'est la fragmentation.

Les entités transmises par la couche 3 sont des **paquets**, ou **datagrammes**.

Les adresses postales d'Internet

Exemple

L'équivalent des adresses postales sur Internet sont les **adresses IP** (pour *Internet Protocol*). Le protocole IP permet l'adressage unique de toutes les machines du réseau et prend en charge la fragmentation.

Niveau 4 : couche transport

Fondamental

La couche transport gère les communications **de bout en bout**, sans se soucier du mode de transport.

Plutôt que de raisonner en terme de paquets, elle raisonne en terme de messages complets.

Elle est capable de **corriger des erreurs** dans la transmission des messages (congestion du réseau...) et de **ré-ordonnancer** des paquets qui auraient pris une route différente et seraient arrivés dans le mauvais ordre.

Ses mécanismes de **contrôle de flux** permettent d'adapter la vitesse d'envoi des données à la capacité de réception du destinataire.

Les entités transmises par la couche 4 sont des **messages**.

Niveau 5 : couche session

Fondamental

La couche session a un rôle de chef d'orchestre des discussions entre les différents acteurs. Elle gère par exemple les tours de parole.

Elle synchronise les communications émanant de différentes sources en organisant les connexions de niveau 4.

Aussi, si une communication est interrompue et que la couche 4 ne peut pas corriger les erreurs (exemple : coupure longue d'électricité), la couche 5 permet de reprendre la communication à l'endroit où elle s'est arrêtée.

Diffusion vidéo

Exemple

Dans une application de diffusion vidéo, la voix et l'image proviennent de deux sources différentes. Il faut organiser leur transmission de telle sorte que les données de la voix soient synchronisés avec les données de l'image. En respectant le modèle OSI, la couche session combinerait ces deux flux en veillant à leur synchronisation.

Niveau 6 : couche présentation

Fondamental

La couche présentation chargée du **codage** des données à communiquer. Les couches 1 à 5 transportent une donnée brute sans se préoccuper de leur signification. La couche présentation a pour rôle de transformer les données (image, texte, vidéo, etc.) en format transportable par les couches inférieures.

Elle négocie également avec le destinataire pour élire un format commun et assurer que les deux systèmes se comprennent.

Optionnellement, cette couche assure chiffrement des données, les rendant inintelligibles pour tous sauf le destinataire.

Une analogie avec les langages

Exemple

Si l'ordinateur A parle français et anglais et que l'ordinateur B parle anglais et espagnol, alors la couche présentation négociera une communication en anglais.

Niveau 7 : couche application

Fondamental

La couche application ne contient pas les véritables applications mais leur rend des services. Elle offre un ensemble de fonctionnalités standard, comme du transfert de fichier.

Des protocoles bien connus

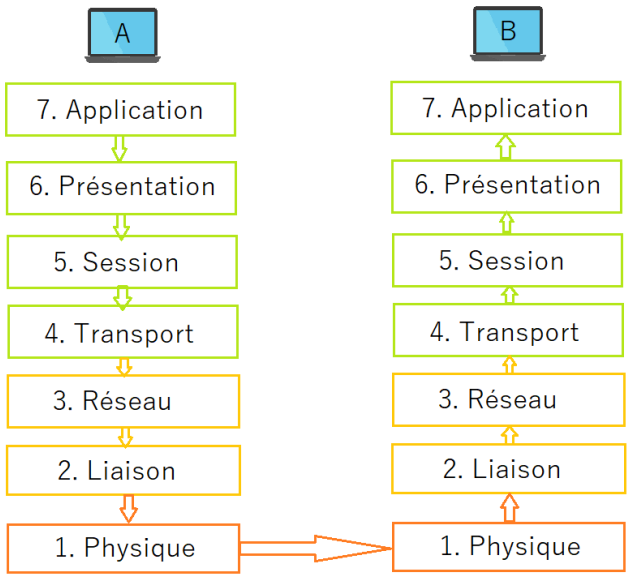
Exemple

On peut rattacher à la couche 7 les protocoles de communication tels que HTTP pour consulter des fichiers sur le Web ou FTP pour transférer des fichiers.

Chemin d'une donnée dans le modèle OSI

Fondamental

Une application sur un ordinateur A voulant communiquer avec une application sur un ordinateur B fera passer les données par toutes les couches de haut en bas. La série de bits sortant de la couche 1 va circuler sur Internet via les routeurs pour arriver à destination, où le chemin inverse sera fait.



Communication selon le modèle OSI

Le parcours d'une requête web

Exemple

Suivons le parcours d'une requête web d'un client à un client. On descend d'abord en couche :

- la requête est émise par l'utilisateur via son client, via un navigateur Web par exemple (application),
- la requête est chiffrée par HTTPS (présentation),
- s'en suit la mise en place de la connexion (session) et son transport de données (transport),
- ces paquets sont passés de proche en proche sur le réseau via IP (couche réseau),
- au niveau du réseau local ces données sont acheminées vers à la machine en question grâce à son adresse MAC (couche liaison),
- ces données sont transférées via un médium, ondes ou signaux électriques par exemple (couche physique).

Puis on remonte dans les couches :

- ce médium est écouté par l'interface réseau de la machine du serveur (couche physique),
- les données sont reçues par la machine du serveur (couche liaison),
- la machine récupère les paquets (couche réseau) et les ordonne selon leur ordre de départ (couche transport),
- la connexion avec le client est fermée ou maintenue (couche session),
- les données sont déchiffrées pour obtenir la requête (couche présentation),
- la requête est lue puis traitée par le serveur,

Modèle OSI et TCP/IP Complément

L'architecture du modèle OSI est très stricte et s'efforce de séparer les fonctionnalités qui sont séparables.

Si ce modèle forme une excellente base théorique pour créer de nouveaux protocoles et de nouveaux réseaux, l'architecture n'est pas intégralement utilisée en pratique.

Internet repose essentiellement sur les protocoles IP et TCP, que l'on peut associer aux couches 3 et 4. Par contre, les couches 5, 6 et 7 n'ont pas vraiment d'équivalent dans le monde Internet : leurs services sont intégrés soit dans les couches plus basses, soit dans la couche 8.

À retenir

- Le modèle OSI est une base théorique pour former des réseaux dont les machines sont capables de communiquer entre elles.
- Il définit 7 couches indépendantes qui remplissent plusieurs services.
- Le modèle OSI décrit tous les aspects des communications mais ne spécifie pas de protocole concret.

5. Adresses IP

IPv4 et IPv6 Rappel

Il existe deux versions du protocole IP : IPv4 et IPv6.

Adresses IP Définition

Une adresse IP est un **numéro d'identification d'une machine**.

Il est attribué de façon à chaque appareil relié à Internet.

Les adresses ont un format différent selon la version d'IP utilisée :

- pour IPv4, les adresses sont sur 4 octets (exemple : `192.168.3.200`),
- pour IPv6, les adresses sont sur 16 octets (exemple : `2001:0db8:0000:85a3:0000:0000:ac1f:8001`).

Remarque

On dit souvent une IP pour parler d'une adresse IP

IP statique ou dynamique

Attention

L'adresse IP attribuée à un appareil peut être statique (fixe) ou dynamique (temporaire).

- Seuls les appareils avec des IP statiques peuvent correctement fonctionner en temps que serveur (il est difficile de savoir comment les contacter sinon).
- Les appareils avec IP dynamiques sont donc essentiellement des consommateurs de contenus (c'est typiquement le cas des smartphones).

Pour un Internet plus équilibré il serait intéressant que la majorité des utilisateurs disposent d'IP fixes pour pouvoir proposer du contenu, mais la pénurie d'IPv4 rend cet objectif difficile à atteindre.

IP locales

Définition

Les adresses IP sont des ressources (devenue rares) qui ont une valeur. Pour limiter les coûts, à l'intérieur d'un réseau local, les ordinateurs ont des adresses IP spéciales réservées aux réseaux locaux appelées adresses **IP locales**.

Ces adresses sont libres d'utilisation et ne nécessitent pas de demander d'autorisation à qui que ce soit.

Synonyme : IP privée

Translation d'adresse

Méthode

Comme les adresses privées ne sont pas utilisées de manière unique, une technique appelée **translation d'adresse** associe une adresse IP publique unique à un ensemble d'adresse IP privées d'un réseau local donné.

Ainsi, les appareils à l'intérieur du réseau local peuvent tout de même être contactés depuis Internet à travers l'IP publique.

IPv4 et pénurie

Complément

Aujourd'hui, la quasi-totalité des 4 milliards d'adresses IPv4 sont attribuées. Cet état de fait pousse le monde à se tourner vers IPv6 qui dispose d'un espace d'adressage bien plus vaste.

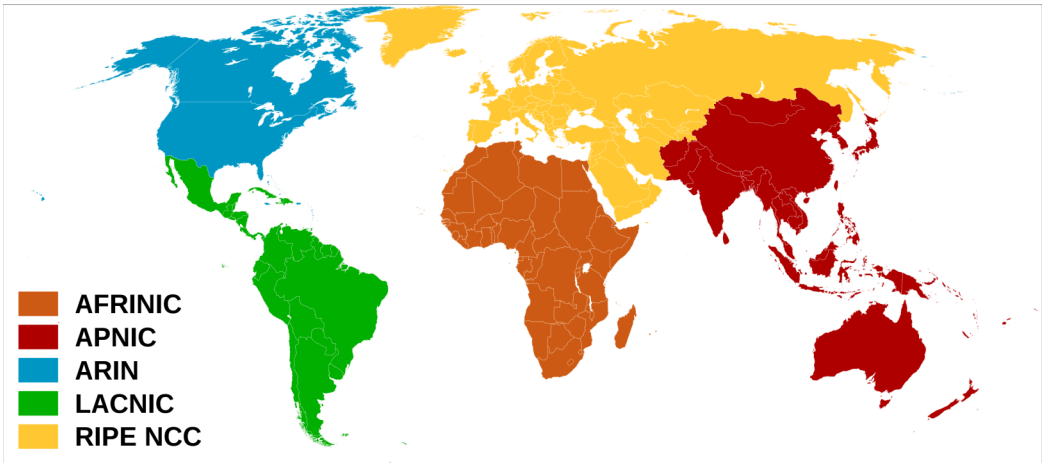
La translation d'adresses est un palliatif temporaire mais très utilisé pour le manque d'adresses IPv4.

Attribution des adresses IP

Complément

L'attribution des adresses IP est gérée par l'IANA (pour *Internet Assigned Numbers Authority*), un département de l'ICANN, qui administre les ressources numériques d'Internet.

L'IANA a segmenté les adresses IP disponibles en 256 blocs distribués aux RIR (Registres Internet Régionaux). Les RIR gèrent les blocs d'adresses au niveau local en définissant des politiques d'attributions d'adresses IPv4 à ceux qui le demandent.



Répartition géographique des différents RIR

6. Noms de domaine

Nom de domaine

Az Définition

Un nom de domaine est une **adresse symbolique** permettant d'éviter de devoir retenir des adresses IP.

Ainsi, à un nom de domaine tel que `wikipedia.org` on peut associer une adresse telle que `91.198.174.192`.

Les noms de domaines ont une structure bien particulière : ils forment une arborescence dont la racine est un point, et dont les niveaux sont séparés par des points. Cette arborescence permet de connaître l'endroit où chercher pour trouver l'adresse IP correspondante au nom de domaine concerné.

Utilité d'un nom de domaine

Remarque

- Un nom de domaine est destiné à être utilisé par des utilisateurs humains. Il n'est fondamentalement pas nécessaire mais permet de donner du sens et est aussi plus facile à retenir.
- De plus un nom de domaine peut être changé là où une adresse IP ne peut généralement pas l'être.

Pour autant, cela ajoute un niveau de complexité : l'appareil qui veut communiquer avec le serveur de Wikipédia a besoin de transformer `www.wikipedia.org.` en une adresse IP.

Pour cela, il fait appel à un **serveur de noms de domaine**.

DNS

Az Définition

Un **serveur DNS** (pour *Domain Name System*) est un serveur qui a connaissance de la **correspondance entre noms de domaines et adresses IP**. Il est capable de répondre à des requêtes qui lui demandent de **résoudre** un nom de domaine en une adresse IP.

DNS, un système hiérarchique

Les serveurs DNS sont organisés hiérarchiquement et sont habilités à établir la correspondance entre IP et nom de domaines pour seulement certains noms de domaines et un certain niveau de l'arborescence. Quand un DNS est autorisé à établir cette correspondance, il est **fait autorité** pour un ensemble de noms de domaines donnés.

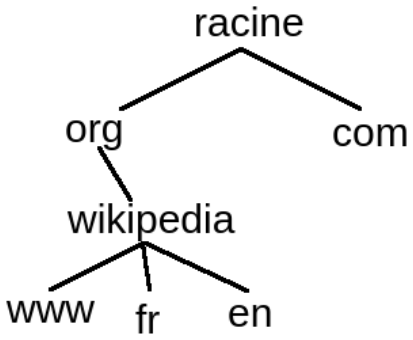
Si un serveur DNS n'a pas l'adresse IP correspondante pour un nom de domaine, il va faire appel à un autre serveur DNS.

Les serveurs DNS responsables du premier niveau d'arborescence (.fr, .com, .org, etc.) sont appelés TLD, pour *Top-Level Domain*.

Résolution d'une requête DNS

Exemple

Les navigateurs web ont besoin de résoudre des noms de domaine. Lorsque `www.wikipedia.org` est consulté, un **résolveur** DNS entre en jeu.



Chaque nœud de cet arbre correspond à un serveur DNS, responsable d'une partie de l'arborescence.

Pour résoudre `www.wikipedia.org`, le résolveur commence donc le haut de l'arborescence : le **serveur racine**. Ce dernier n'est pas responsable de l'adresse IP de Wikipédia mais connaît les serveurs DNS TLD pour `.org`.

Les serveurs DNS autoritaires pour `.org` ne sont pas non plus responsables de l'adresse IP de Wikipédia. À leur tour, ils connaissent les adresses IP des serveurs DNS faisant autorité pour Wikipédia.

Ces serveurs fourniront finalement l'adresse du sous-domaine `www` de `wikipedia.org`.

Résoudre une IP

Méthode

Il existe des résolveurs DNS en ligne de commande, comme `host`.

```
$ host www.wikipedia.org
www.wikipedia.org is an alias for dyna.wikimedia.org.
dyna.wikimedia.org has address 91.198.174.192
dyna.wikimedia.org has IPv6 address 2620:0:862:ed1a::1
```

`host` retourne les enregistrements DNS, ici on voit l'adresse IPv4 actuelle de Wikipédia, qui est `91.198.174.192`

Réponse ne faisant pas autorité

Complément

L'exemple ci-dessus mentionne que la réponse ne fait pas autorité, ce qui signifie qu'elle émane d'un serveur n'ayant pas autorité sur le nom de domaine.

Ceci s'explique par un système de **cache** : pour éviter que les serveurs DNS de Wikipédia soient surchargés, d'autres serveurs hébergent une copie des correspondances entre IP et nom de domaine. Ces serveurs ne sont en revanche pas autorisés à modifier ces correspondances.

Comment est-ce que les noms de domaines sont enregistrés ?

Fondamental

Un nom de domaine peut être acheté auprès d'organismes en charge de l'administration des nom de domaines, appelés **registrar**. Le registrar va s'occuper d'ajouter le nom de domaine à un de ses DNS faisant autorité puis va aller indiquer aux DNS TLD correspondants qu'un nouvel enregistrant a eu lieu.

Résolution locale de nom de domaine via le fichier hosts

Complément

Les systèmes d'exploitation maintiennent un fichier `hosts` contenant les informations de résolution de certains nom de domaines. Ce fichier a généralement cette forme :

```
127.0.0.1       localhost
127.0.0.1       monPC
208.80.154.224  www.wikipedia.org.
```

Ce fichier est interrogé en premier : il est ainsi possible de forcer la résolution d'un nom de domaine en n'importe quel adresse, au niveau de l'ordinateur.

À retenir

- Un nom de domaine est un alias pour une adresse IP.
- Les noms de domaines décrivent une arborescence.
- Le système permettant d'associer adresse IP et nom de domaine s'appelle DNS.
- Les serveurs DNS sont organisés de manière hiérarchique et s'occupent d'un ensemble donné de noms de domaines.

7. Internet et ses protocoles

Objectifs

- Rappeler le fonctionnement d'Internet ;
- Découvrir la notion de protocole ;
- Découvrir les différents protocoles d'Internet.

Mise en situation

Internet est un réseau reliant des ordinateurs du monde entier. Il rend possible des applications, comme : le Web, le transfert de fichiers ou les mails. Comment ces différents ordinateurs, de marques et de types différents, peuvent-ils arriver à communiquer entre eux ?

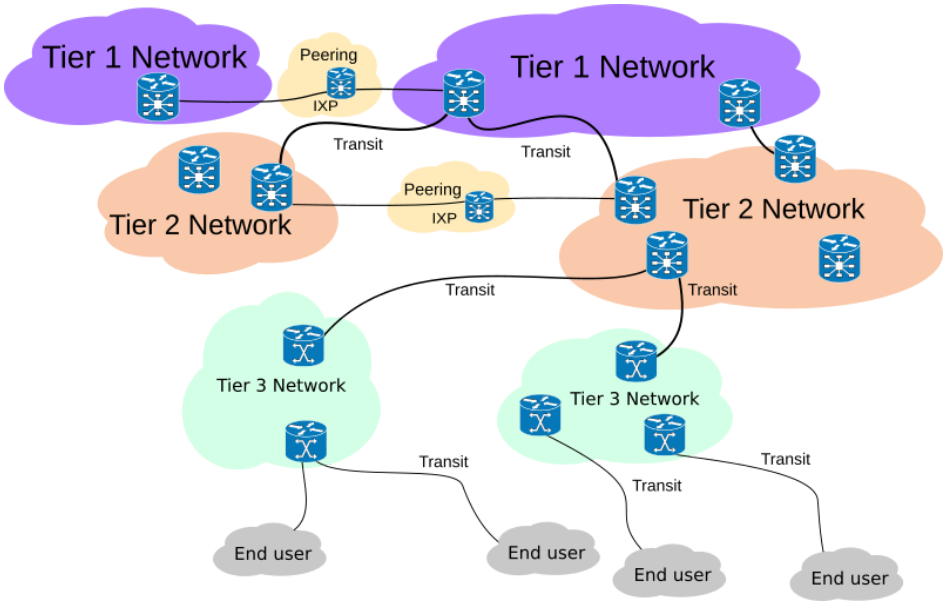
Si des PC Windows, des téléphones Android, des tablettes Apple ou même des voitures Tesla, parviennent à communiquer avec des serveurs Linux, c'est parce que, *in fine*, tous ces ordinateurs utilisent des protocoles communs.

Structuration d'Internet

Rappel

Internet fait intervenir plusieurs acteurs :

- les **utilisateurs** qui utilisent le réseau,
- les **fournisseurs d'accès à Internet** (FAI) qui offrent aux utilisateurs finaux un accès aux réseau (catégorisés en *Tier*),
- les **fournisseurs de contenus d'application** (FCA) qui souhaitent proposer du contenu sur Internet,
- les **hébergeurs** qui possèdent les machines louées par les FCA,
- les **transitaires** qui permettent aux hébergeurs ou aux FAI de se relier au reste d'Internet.



Transit et peering entre opérateurs Internet de niveau 1 à 3

Pour assurer la communication entre des paires d'ordinateurs appelés **extrémités du réseau**, des messages sont envoyés à travers le réseau et passent par les différents acteurs.

Modèle OSI

Rappel

Le **modèle OSI** (pour *Open Systems Interconnection*) est une norme de communications, en réseau, de tous les systèmes informatiques. C'est un modèle conceptuel qui définit le fonctionnement d'un réseau global en **7 couches**. Chaque couche réalise un **service** accessible via l'**interface** de la couche. Ce service est réalisé grâce à un **protocole**.

1. La couche « **physique** » est chargée de la transmission effective des signaux entre les interlocuteurs : elle est limité à l'émission et la réception de bits.
2. La couche « **liaison de données** » gère les communications entre deux machines connectées entre elles.
3. La couche « **réseau** » gère les communications de proche en proche (routage et adressage des paquets).
4. La couche « **transport** » gère les communications de bout en bout entre programmes en cours d'exécution.
5. La couche « **session** » gère la synchronisation des échanges.
6. La couche « **présentation** » est chargée du codage des données applicatives (conversion entre données manipulées au niveau applicatif et chaînes d'octets transmises).
7. La couche « **application** » est le point d'accès aux services réseaux.

https://fr.wikipedia.org/wiki/Mod%C3%A8le_OSI

Protocole

Az Définition

Un protocole informatique est un **ensemble de règles** qui régissent les **échanges de données** entre plusieurs processus ou entre des ordinateurs en réseaux.

Protocole dans un bureau de vote

Exemple

Les règles dans un bureau de vote peuvent être vues sous l'angle d'un protocole : le votant se présente, récupère les différents bulletins, réalise un choix tenu secret, vote en soumettant son bulletin, et signe pour attester du vote.

Applications et protocoles d'Internet

Fondamental

Parmi les protocoles et applications les plus importants de la couche applicative d"Internet on retrouve :

- les **connexions à des machines distantes** avec **SSH** (*Secure Shell*),
- la consultation de documents sur le **Web** avec **HTTP** (*HyperText Transfert Protocol*),
- le **transfert de fichiers** assuré par le protocole **FTP** (*File Transfert Protocol*),
- l'**envoi et la réception d'e-mails** avec **SMTP** (*Simple Mail Transfert Protocol*), **IMAP** (*Internet Message Access Protocol*) et **POP** (*Post Office Protocol*).

Remarque

Chaque protocole dispose d'un **port** par défaut qui lui est réservé.

À retenir

- Le modèle OSI est un modèle conceptuel d'Internet composé de sept couches.
- Les protocoles des applications d'Internet se trouvent sur la septième couche et disposent d'un port réservé.
- Les protocoles principaux des applications d'Internet sont SSH, HTTP, FTP et SMTP/POP/IMAP.

8. FTP : File Transfer Protocol

Objectif

- Se familiariser avec FTP.

Mise en situation

Si le mail permet l'échange de messages, il est mal adapté à l'échange de fichiers, surtout s'ils sont nombreux ou volumineux. Le protocole FTP est historiquement le premier à avoir rempli ce rôle. Un client FTP peut déposer et récupérer des fichiers sur un serveur FTP.

Le principal défaut de FTP est qu'il n'est pas chiffré. Ainsi il est remplacé aujourd'hui par ses variantes sécurisées SFTP et FTP/S. On utilise aussi parfois HTTP pour échanger des fichiers, via des sites comme *Firefox Send*.

FTP — File Transfer Protocol

Az Définition

FTP est le protocole de référence pour le **transfert de fichiers**. Il respecte une architecture client-serveur entre un **client FTP** et un **serveur FTP**.

Voici quelques fonctions permises par le protocole :

- établir une connexion
- envoyer son identifiant/nom d'utilisateur
- envoyer son mot de passe
- afficher le dossier courant
- changer de dossier courant
- récupérer une copie d'un fichier (téléchargement)
- déposer une copie d'un fichier (téléversement)
- supprimer un fichier
- ...

FTP a besoin d'un protocole de transport fiable tel que TCP (on ne peut pas se permettre de perdre des paquet quand on transfère des fichiers en général).

Il fonctionne par défaut sur le port 21 pour recevoir les commandes FTP et le port 20 pour recevoir les données.

Gérer les fichiers d'un site web

Exemple

Un développeur a besoin de transférer sur son serveur les fichiers qui constituent son site web. Il va pouvoir utiliser le protocole FTP pour transférer ses fichiers depuis son poste client jusqu'au serveur fourni par l'hébergeur.

Échanges avec un serveur FTP : informations non chiffrées

Attention

- La connexion à un serveur FTP se fait via un couple nom d'utilisateur/mot de passe. Le client envoie une commande USER pour le nom d'utilisateur et une commande PASS pour le mot de passe. Cependant, aucun protection n'est offerte par le protocole et donc, par défaut, le mot de passe circule en clair sur Internet. Ce protocole permet aussi l'authentification anonyme sur un serveur.
- On utilise SFTP ou FTP/S aujourd'hui plutôt que FTP (ces protocoles fonctionnement de la même façon en utilisant le chiffrement).

Conseil

- Ne jamais utiliser FTP avec un mot de passe qui est également utilisé par ailleurs. : si le mot de passe était écouté, seul le serveur FTP serait compromis.
- N'utiliser que des mots de passe générés aléatoirement (il n'apprendront rien sur "vous" à celui qui écoute).

Les méthodes sécurisées

Fondamental

FTP pose un important problème de sécurité par l'absence de chiffrement des communications.

C'est pourquoi il existe aujourd'hui des protocoles alternatifs qui chiffrent les communications par défaut.

Ce sont ces protocoles qui sont recommandés.

- **SFTP** (SSH File Transfer Protocol) utilise une session SSH pour transférer les fichiers et implémentent les commandes FTP. C'est le protocole à privilégier lorsqu'on dispose d'un accès SSH à un serveur.
- **FTP/S** est une extension de FTP lui permettant de supporter le protocole TLS. TLS fournit un chiffrement des communications sur un réseau.

SFTP en ligne de commande

Méthode

```
1 sftp répertoire-d-arrivée utilisateur@adresse-ip
```

Une fois connecté, plusieurs commandes analogues aux commandes Linux sont disponibles pour interagir avec le système de fichiers.

Par exemple :

- `ls` pour lister les fichiers sur le serveur
- `lls` (local ls) pour liste les fichiers sur le cient
- pwd pour voir le répertoire courant
- `cd` pour changer de répertoire
- `get` pour télécharger un fichier depuis le serveur
- `put` pour téléverser un fichier sur le serveur

SFTP avec FileZilla

Méthode

FileZilla est un client graphique qui permet de communiquer avec un serveur en FTP, SFTP et FTP/S.

Sur Linux (Debian et Ubuntu), on pourra installer FileZilla avec la commande suivante :

```
1 sudo apt install filezilla
```

Sinon, on pourra télécharger FileZilla depuis cette page : https://filezilla-project.org/download.php?show_all=1

Alternatives à FTP

Complément

Il existe d'autres protocoles et outils pour échanger des fichiers, comme les programmes en ligne de commande `scp` et `rsync` (qui s'appuient sur le protocole SSH).

```
1 scp chemin/local/fichier user@<adresse-IP>:chemin/de/réception/
```


```
1 scp chemin/local/fichier user@<adresse-IP>:chemin/de/réception/
```


Serveur de téléchargement

Complément

Un serveur FTP peut aussi choisir de bloquer certaines commandes.

Une entreprise ou association peut décider de mettre en téléchargement un logiciel en proposant un accès anonyme à un serveur FTP. Dans ce cas, elle ne donnera droit qu'à la commande permettant récupérer une copie du logiciel (RETR).

À retenir

- FTP est un protocole utilisé pour transférer des fichiers sur Internet.
- Il peut être utilisé pour envoyer des pages web sur le serveur d'un hébergeur.
- SFTP et FTP/S sont des protocoles similaires qui fonctionne comme FTP mais avec des communications chiffrées.

9. HTTP : HyperText Transfert Protocol

Objectif

- Se familiariser avec HTTP.

Mise en situation

HTTP est le protocole du Web : il permet à un client (le navigateur) d'obtenir des pages et des ressources depuis un serveur web (le site).

On connaît ce protocole car les adresses web commencent par `http://`.

Par exemple, `http://en.wikipedia.org/wiki/Tim_Berners-Lee` signifie : je souhaite communiquer avec le serveur `en.wikipedia.org` en utilisant le protocole HTTP, je lui demande de me donner la page qui s'appelle `Tim_Berners-Lee`. On appelle cela une requête HTTP.

Une courte histoire du Web

Rappel

HTTP est conçu au début des années 90 par Tim Berners-Lee[12] dans le cadre de ses travaux au CERN. L'objectif est de mettre en place un système qui permette l'échange d'informations et de ressources, entre les chercheurs du CERN, en s'appuyant sur le réseau interne. Le projet prend rapidement le nom de *World Wide Web* et abouti à la création du protocole HTTP et du premier navigateur.

HTTP

Az Définition

HTTP (pour *HyperText Transfert Protocol*) est le protocole utilisé par le Web pour le transfert de ressources.

Ce protocole fonctionne sous la forme de **requête-réponse** dans une architecture **client-serveur** : le client adresse une requête HTTP à un serveur qui lui enverra une réponse appropriée.

Ce protocole repose généralement sur un protocole de transport fiable, tel que TCP.

Port 80

Fondamental

Le port **80** est le port réservé par les serveurs Web pour HTTP.

Les requêtes HTTP

Il existe différent types de requêtes HTTP, voici les plus importantes :

- GET : c'est la plus commune et elle demande à recevoir une certaine ressource disponible sur un serveur web. Quand un client fait une requête GET, c'est qu'il veut uniquement recevoir une ressource et qu'il ne cherche pas à la modifier.
- POST: une requête POST est utilisée dans l'intention de modifier ou ajouter une ressource sur le serveur web.
- DELETE : elle est utilisée par le client lorsqu'il veut supprimer une ressource d'un serveur web.

12. Tim Berners-Lee - https://fr.wikipedia.org/wiki/Tim_Berners-Lee

Format des requêtes

Les requêtes HTTP suivent le format suivant :

- le type de la requête (GET, POST, etc.) avec la référence de la ressource en question sur le serveur,
- des en-têtes donnant diverses informations au serveur,
- un corps optionnel contenant des données utiles à la requête ; dans le cas de l'ajout d'une ressource, le corps contiendra la ressource en elle-même.

Un message HTTP

Exemple

```
1 GET /wiki/Hypertext_Transfer_Protocol HTTP/1.1
2 Host: www.wikipedia.org
```

On demande à recevoir (GET) la ressource /wiki/Hypertext_Tansfert_Protocol se trouvant sur le serveur www.wikipedia.org.

La directive `Host:` est une en-tête.

Réaliser une requête GET

Méthode

- Les navigateurs web exécute une requête GET à chaque fois qu'un utilisateur demande à consulter une page web.
- Il existe des programmes tels que `wget` ou `curl` pour effectuer des requêtes HTTP en ligne de commande.

```
1 curl https://fr.wikipedia.org/wiki/Hypertext_Transfer_Protocol
```

```
1 wget https://fr.wikipedia.org/wiki/Hypertext_Transfer_Protocol
```

Cette commande enregistre le contenu de la page `Hypertext_Transfer_Protocol` de Wikipédia. Contrairement à un navigateur web, la page n'est pas affichée ni mise en page.

Les en-têtes HTTP

Les en-têtes (ou *headers*) HTTP permettent d'ajouter de nombreuses informations aux requêtes, ce qui facile l'ajout de fonctionnalités au protocole : gestion du cache, de l'authentification, des formats de données à échanger, etc.

Exemple

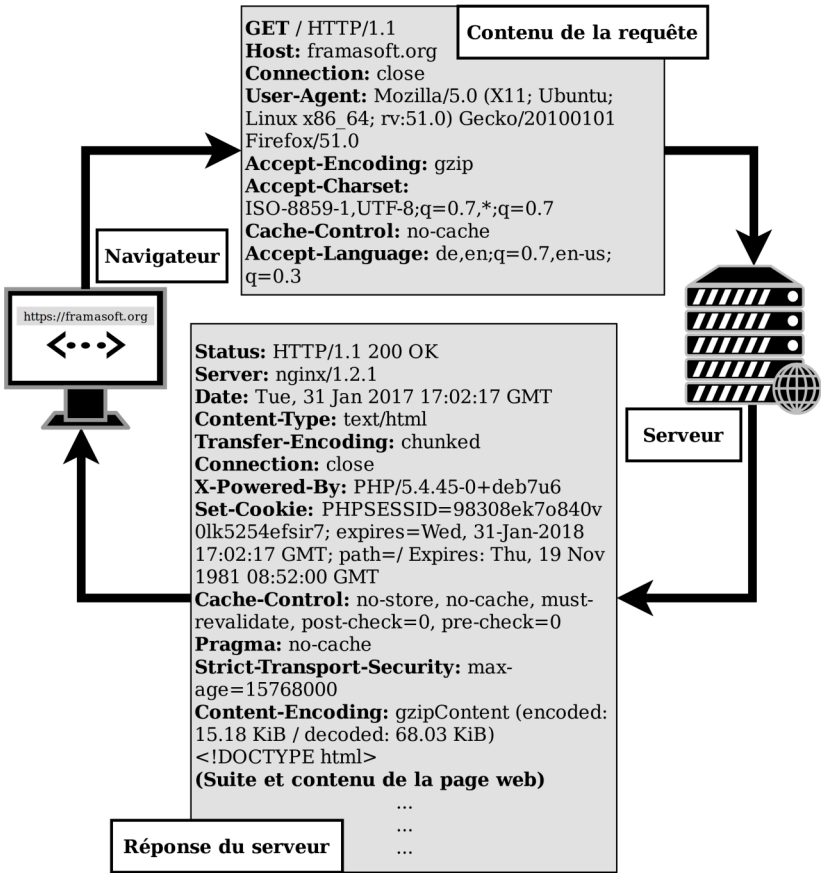
Nous avons vu plus haut une requête utilisant le *header* `Host`. Un autre très utilisé est le *header* `Authorization` qui est utilisé pour gérer l'accès avec des identifiants à une ressource. Sans rentrer dans le détail, voici un exemple :

```
1 GET /internal/secret_file.txt HTTP/1.1
2 Authorization: Basic a3lhbmU6
```

Ici on demande l'accès à la ressource /internal/secret_file.txt en utilisant de l'authentification. L'en-tête `Authorization` permet de spécifier que l'on utilisera la méthode d'authentification "Basic" (il en existe plusieurs) et on passe comme valeur d'identifiant "a3lhbmU6".

Récapitulatif

Fondamental



Contenu d'une requête HTTP et de sa réponse

À retenir

- HTTP est le protocole du Web qui permet d'échanger des ressources, le plus souvent des fichiers HTML.
- Il permet de faire des requêtes pour demander à consulter, ajouter, modifier ou supprimer une ressource sur un serveur web.

Complément

- *Serveurs web* (cf. p.62)
- *Navigateurs web* (cf. p.64)

10. Serveurs web

Rappel

HTTP : HyperText Transfert Protocol (cf. p.59)

Serveur web

Az Définition

Un serveur web est un logiciel qui **traite les requêtes HTTP** de clients.

Il est installé sur une machine hébergeant des documents (HTML, CSS, JavaScript, etc.) et est accessible depuis Internet ou un réseau local (un intranet).

Les ressources servies par le serveur peuvent être **statiques**, c'est-à-dire sans qu'elles existent préalablement à la requête (un fichier HTML, une image PNG, etc.) ou **dynamiques**, c'est-à-dire construites à chaque requête faite au serveur.

Serveur Web et serveur HTTP

Remarque

Les serveurs web utilisent presque exclusivement le protocole HTTP, créé spécifiquement pour le Web. On parle par abus de langage de serveur web, mais le nom **serveur HTTP** est aussi possible.

Serveurs HTTP

Exemple

- **Apache** : le serveur web le plus utilisé avec 44,3% des parts de marché en février 2019.
- **Nginx** : le deuxième serveur web le plus utilisé mais est le premier parmi les 1000 sites les plus actifs.
- **Node.js** : un environnement d'exécution JavaScript en dehors du navigateur qui fait aussi office de serveur HTTP.

Installation d'un serveur Apache sous GNU/Linux

Méthode

Sur les distributions de famille Debian, `apt` est utilisé par défaut pour la gestion des logiciels. Il faut commencer par mettre à jour les dépôts de logiciels sur son serveur.

```
1 apt update
```

Puis installer la dernière version d'Apache ainsi :

```
1 apt install apache2
```

Après l'installation, le serveur est prêt à recevoir des requêtes HTTP sur son port 80.

Fondamental

Par défaut, Apache va placer ses fichiers de configuration dans le dossier `/etc/apache2` et sert les fichiers du répertoire `/var/www/html`.

Installation d'un serveur Nginx sous GNU/Linux Complément

On pourra se référer aux documentations suivantes pour installer le serveur web Nginx :

- https://doc.ubuntu-fr.org/nginx
- Installer Nginx (nginx.com)[13]

À retenir

- Un serveur hébergeant des fichiers web a besoin d'un serveur web pour les rendre disponibles à travers le protocole HTTP.
- Apache et Nginx sont les serveurs web les plus utilisés.

11. Navigateurs web

Navigateurs web Définition

Un navigateur web est client HTTP qui a trois rôles.

1. Effectuer des requêtes HTTP (GET, POST, etc.).
2. Comprendre les réponses à ces requêtes.
3. Interpréter les fichiers web (HTML, CSS, JavaScript) retournés.

Remarque

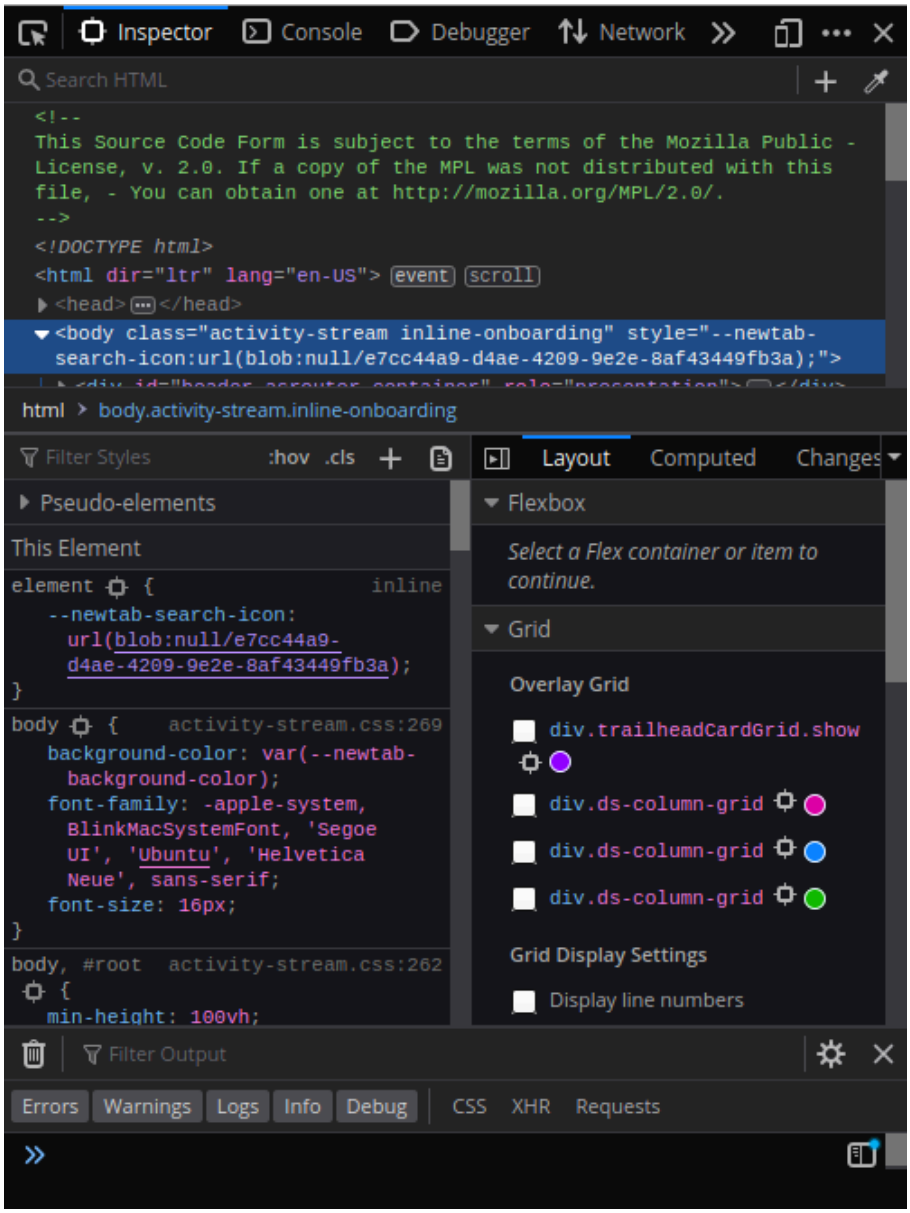
Tous les navigateurs web possèdent un interpréteur JavaScript permettant d'exécuter ce langage sur l'ordinateur client.

Les outils de développement sous Firefox Fondamental

Les navigateurs possèdent aussi une partie « Outils de développement » destinée à aider les développeurs web dans leur activité.

Sous Firefox, ces outils sont disponibles en appuyant sur `F12` ou `clic droit > Inspecter l'élément`.

13. https://docs.nginx.com/nginx/admin-guide/installing-nginx/installing-nginx-open-source/



Outils de développement Firefox (F12)

Voilà à quoi ressemble l'inspecteur depuis le site `mozilla.org`.

On retrouve plusieurs onglets en haut de la fenêtre. En voici les plus importants :

- L'inspecteur :

  C'est celui que l'on voit sur l'image ci-dessus, on y retrouve l'HTML de la page et son CSS. Ils sont modifiables depuis cet onglet.

- La console :

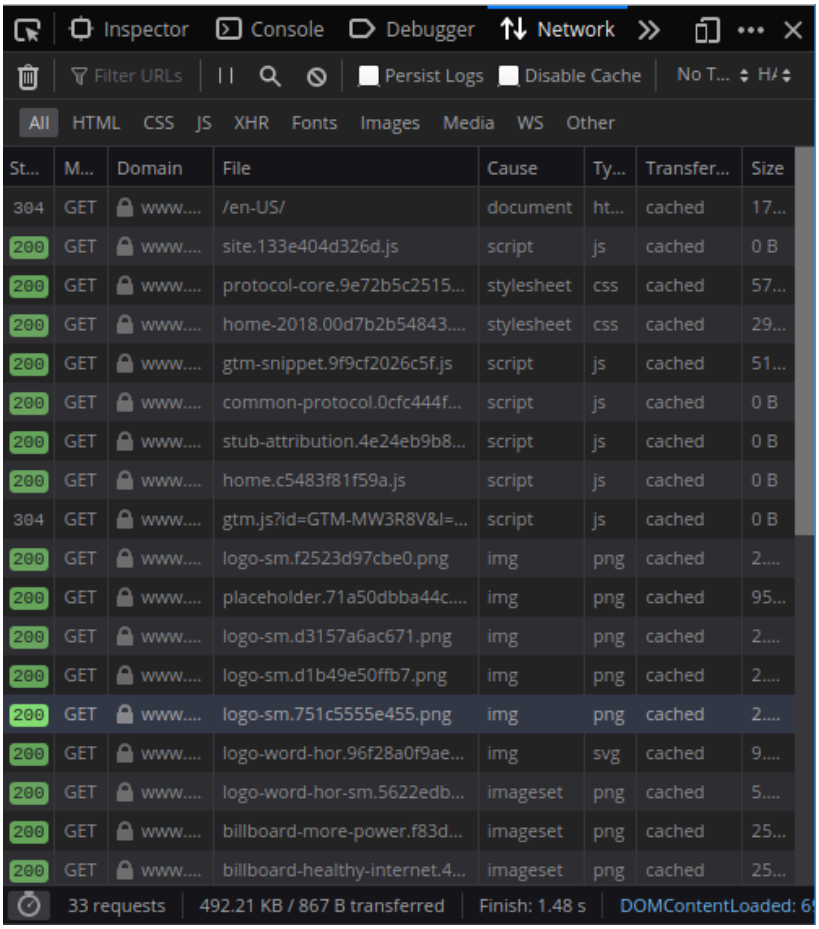
  C'est une console JavaScript. Un utilisateur peut donc y écrire des lignes de JavaScript qui seront interprétées à la volée.

- Le déboguer :

  Cet outil permet un suivi très fin de ce qu'il se passe pendant l'interprétation de la page et l'exécution de JavaScript.

- Le moniteur réseau :

  Cet outil permet d'analyser avec une précision fine toutes les requêtes envoyées et leur réponse associée.

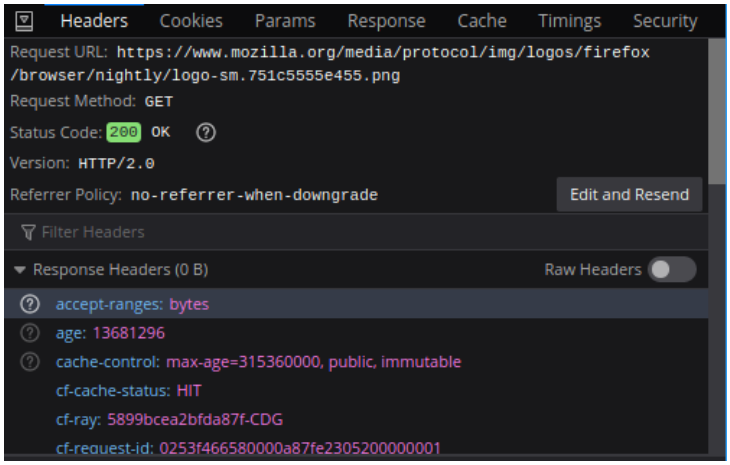


Extrait du moniteur réseau quand on accède à www.mozilla.org

Il y pour chaque requête :

- le code HTTP de la réponse (304 pour signifier que le document demandé n'a pas été modifié depuis la dernière demande et 200 pour dire que tout va bien),
- la requête complète, ici il s'agit d'accéder à la page d'accueil, il n'y a donc que des GET pour récupérer le HTML, le CSS, les script JavaScript et les images,
- l'adresse du serveur web (www.mozilla.org),
- la ressource demandée,
- le type du fichier (HTML, CSS, PNG, etc),
- la quantité transférée, cette valeur est à `cached` lorsque que le navigateur l'avait déjà en mémoire,
- la taille de la ressource.

En cliquant sur une requête, toutes ses informations apparaissent.



Outils de développement Firefox (onglet requête)

À retenir

- Les navigateurs web sont des clients HTTP.
- Ils interprètent les langages du Web (HTML, CSS, JS).
- Ils disposent d'outils de développement permettant d'inspecter les fichiers des pages web et les requêtes et réponses HTTP.

12. SSH : Secure SHell

Objectif

- Se familiariser avec le protocole SSH (Secure SHell).

Mise en situation

SSH est une application d'Internet qui permet de se connecter à un serveur à distance, puis d'exécuter des commandes sur ce serveur.

C'est aujourd'hui le mode de contrôle privilégié des informaticien·nes sur les serveurs qui font fonctionner Internet.

Terminal et shell

Rappel

Le **terminal** (aussi appelé invité de commandes ou console) est le moyen historique pour interagir avec un ordinateur. Ce programme offre un **shell** (une interface système) donnant accès aux programmes de l'ordinateur. Il suffit d'un écran et d'un clavier pour utiliser ce shell.

- *Terminal et shell* (cf. p.68)

Le programme SSH (Secure SHell)

Az Définition

Le programme SSH (*Secure Shell*) permet d'interagir de manière sécurisée avec un ordinateur distant via un shell. C'est le programme de référence pour effectuer des opérations à distance.

Toutes les commandes tapées depuis un clavier d'ordinateur à un emplacement A sont exécutées dans le shell d'un ordinateur à un emplacement B. Les informations qui transitent sur Internet via SSH sont les chaînes de caractères représentant les commandes à exécuter et les chaînes de caractères représentant les résultats de ces exécutions.

Le protocole SSH

Az Définition

SSH est aussi le nom du protocole de communication utilisé par le programme SSH. Ce protocole fonctionne sur une **architecture client-serveur** ; l'ordinateur qui fournit les commandes est le **client** et l'ordinateur qui exécute les commandes sur son système est le **serveur**.

SSH repose sur TCP pour le transport et est par défaut associé au port **22**.

L'établissement d'une connexion SSH se fait en deux étapes :

- Établissement d'une communication sécurisée ;
- Authentification du client.

Établir une connexion SSH

Méthode

On utilise la commande `ssh` depuis un terminal :

```
1 ssh alice@80.67.182.69
```

La machine se trouvant à l'adresse `80.67.182.69` doit avoir un utilisateur `alice`.

Attention

Si `ssh` se connecte pour la première fois au serveur, l'utilisateur devra examiner l'« empreinte » du serveur pour des raisons de sécurité. En pratique, sauf besoin très spécifique, on répond `yes` systématiquement.

Remarque

Si un serveur est associé à un nom de domaine, on peut utiliser ce domaine à la place de l'IP.

```
1 ssh alice@picasoft.net
```

À retenir

- Le protocole SSH permet à un client d'ouvrir un shell sur un serveur distant.
- Ce protocole sécurise les communications en employant plusieurs méthodes de chiffrement, de l'authentification au transfert de données.

13. À la découverte du terminal

Objectif

- Savoir ouvrir un terminal et exécuter une commande.

Mise en situation

On a l'habitude d'interagir avec les ordinateurs en mode graphique, c'est-à-dire en utilisant essentiellement la souris pour cliquer sur des éléments d'interface pour effectuer des actions. Mais il existe une autre manière d'interagir avec un ordinateur, très utilisée en informatique : le mode texte ou **interface en ligne de commande** (CLI en anglais).

Ce mode d'interaction est très utile pour pouvoir **utiliser un ordinateur à distance**, ce qui est généralement le cas lorsque l'on souhaite administrer un serveur web.

À travers ce module vous allez découvrir le **shell** qui permet de dialoguer avec le système d'exploitation d'une machine. Il existe beaucoup de shells, le plus connu vient du monde Linux et se nomme bash. Vous allez découvrir les commandes de base pour parcourir des répertoires ou éditer des fichiers.

Terminal ou CLI

Az Définition

Un **terminal**, ou **interface en ligne de commande** (*CLI* en anglais), est une interface homme-machine dans laquelle l'utilisateur interagit avec la machine en mode **texte**. L'utilisateur écrit des **lignes de commande**, la machine les exécute et affiche le résultat des commandes.

Le terme terminal est très général : un terminal peut servir à dialoguer avec un programme informatique, à donner des ordres à un ordinateur, ou à exécuter d'autres programmes.

Shell

Az Définition

Un **shell** est une interface en ligne de commande permettant de dialoguer avec le système d'exploitation de la machine. On dit qu'il **interprète** les commandes.

Langages CLI

Exemple

Il existe au moins autant de shells que de systèmes d'exploitation. Chaque shell propose des commandes spécifiques. Le choix d'un shell se fait surtout par rapport à des critères pratiques (de quoi ai-je besoin ?) et des critères subjectifs (quel shell me semble le plus ergonomique ?).

- Sur Linux, le shell le plus connu et installé par défaut sur la plupart des systèmes s'appelle `bash` (pour "Bourne Again Shell").
- Sur Windows, il existe trois shells :
  - `cmd`, le shell historique,
  - `powershell`, une version plus moderne,
  - et il est possible d'utiliser `bash` avec le sous-système Linux[14].
- Sur macOS, le shell installé par défaut s'appelle `zsh`. Il partage de grandes similarités avec `bash`.

14. https://learn.microsoft.com/fr-fr/windows/wsl/install

Shells et Windows

Remarque

Le shell historique `cmd` de Windows n'est presque plus utilisé. `powershell` utilise une syntaxe assez différente de la plupart des autres shells. Il est possible d'installer `bash` sous Windows depuis peu, mais il n'est pas inclus par défaut dans les installations.

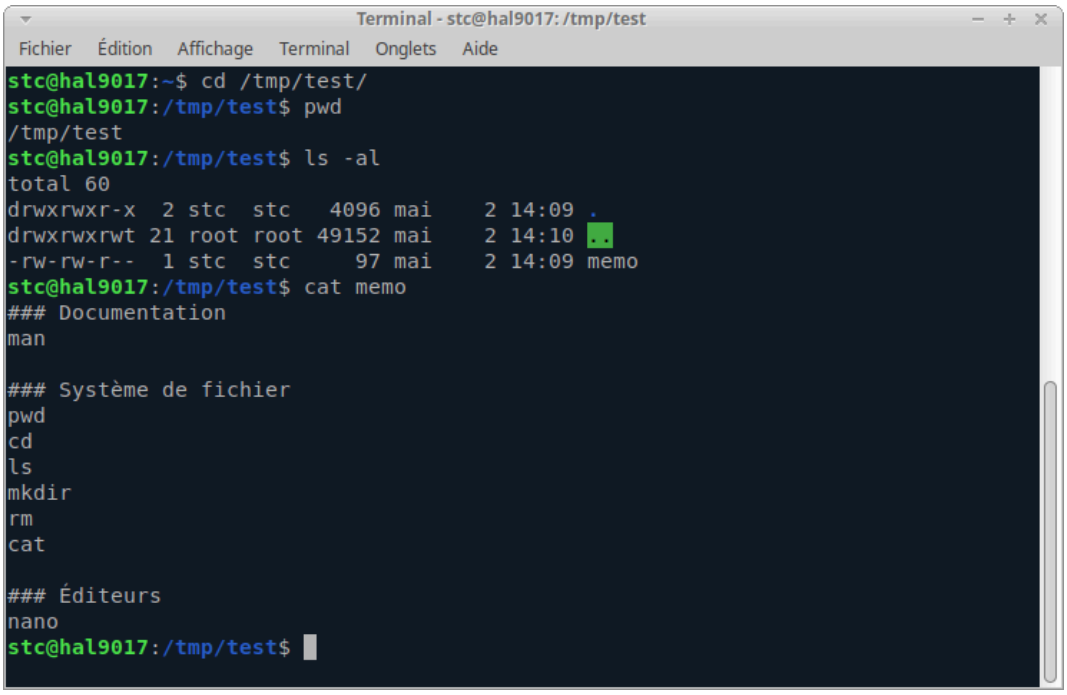
Bash et Windows

Conseil

La plupart des shells adhèrent à des standards communs, mais ce n'est pas le cas des shells disponibles sur Windows. Pour suivre les exemples et exercices qui suivent, il est recommandé **d'activer le sous-système Linux pour Windows** et d'utiliser `bash`, en suivant la partie "Installer le sous-système Windows pour Linux" de cette documentation : https://learn.microsoft.com/fr-fr/windows/wsl/install

Commandes et Bash

Exemple



Cette image montre quelques commandes de base, exécutées par un shell `bash` sur un système Linux.

- `cd /tmp/test` se rend dans le dossier `test` qui se trouve dans le dossier racine `/tmp`.
- pwd affiche le dossier où on est situé dans le système de fichier.
- `ls -al` affiche la liste des fichiers dans le dossier courant.
- `cat memo` affiche le contenu du fichier `memo`.

Terminal et environnement graphique

Remarque

- Un ordinateur personnel moderne PC dispose d'une interface graphique et d'une interface terminal : les deux permettent d'effectuer à peu près les mêmes opérations : visualiser des fichiers, les supprimer, ouvrir des applications, etc.
- Un serveur n'offre en général qu'une possibilité d'accès à distance via un terminal. C'est une des raisons pour lesquelles savoir utiliser le terminal est utile.

Ouvrir un terminal sous Linux

Méthode

En général (et en particulier sur les systèmes Ubuntu), le raccourci `Ctrl+Alt+T` ouvre un terminal. Une alternative consiste à chercher `Terminal` dans la liste des applications.

Ouvrir un terminal sous Windows

Méthode

Ouvrir la fenêtre `Exécuter` à l'aide du raccourci `Super`+R (la touche `Super` est en général représentée par un logo Windows sur le clavier). Entrer :

- `powershell` dans la fenêtre qui s'est ouverte pour démarrer un terminal Windows
- `bash` pour ouvrir un shell Bash s'il a été installé

Ouvrir un terminal sous macOS

Méthode

Depuis le Launchpad, chercher `Terminal` et cliquer sur l'icône qui s'affiche.

Quelques commandes de base

Syntaxe

Ces commandes sont des commandes Bash qui fonctionnent également sous macOS et avec la plupart des autres shells Linux.

- `ls -al` : lister les fichiers dans le répertoire courant
- pwd : afficher le répertoire courant
- `cat` *`fichier`* : afficher le contenu d'un fichier
- `cd` *`dossier`* : se rendre dans un dossier fils du dossier courant
- `cd ..` : se rendre dans le dossier parent
- `echo` *`message`* : afficher un message
- `man` *`commande`* : afficher la documentation détaillée d'une commande

Copier/coller dans un terminal

Méthode

Les actions « copier » et « coller » sont accessibles en effectuant un clic droit sur la fenêtre d'un terminal. Il est plus rapide d'utiliser des raccourcis clavier pour copier et coller, respectivement :

- Sur Bash et la plupart des shells Linux : `Ctrl+Shift+C` et `Ctrl+Shift+V`
- Sur macOS : `Command+C` et `Command+V`
- Sur Windows (`powershell`) : `Ctrl+C` et `Ctrl+V`

À retenir

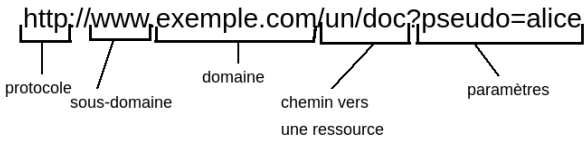
- Un shell, ou par abus de langage un terminal, permet de dialoguer avec le système d'exploitation d'une machine.
- Il existe beaucoup de shells. Le plus connu vient du monde Linux et se nomme `bash`.
- Un shell permet de se passer du mode graphique, ce qui est souvent indispensable pour travailler sur une machine à distance.

14. URL : Uniform Resource Locator

Les URL

Définition

Les URL sont nées en même temps que le World Wide Web et permettent d'identifier une ressource sur un serveur web.



Décomposition d'une URL

Elles contiennent :

- Le **protocole** à utiliser pour accéder à la ressource, ici HTTP,
- L'**adresse du serveur** à rejoindre, ici le nom de domaine `www.exemple.com`,
- Le **chemin de la ressource** dans le serveur,
- Éventuellement des paramètres que l'on veut envoyer au serveur (utilisés par les requêtes HTTP GET).

Exemple

Lorsque `https://www.wikipedia.org/wiki/URL` est rentrée dans un navigateur web, ce dernier fait une requête GET sur la ressource « `/wiki/URL` » du serveur pointé par `www.wikipedia.org`, en sécurisant la connexion grâce à HTTPS.

Éléments de syntaxe d'une URL Syntaxe

Certains caractères sont réservés pour un usage spécifique. Par exemple :

- le caractère `/` est réservé pour indiquer le chemin des fichiers à consulter,
- le caractère ? est réservé pour annoncer les paramètres à transmettre au serveur,
- le caractère = est réservé pour indiquer la valeur des paramètres,
- le caractère # est réservé pour indiquer un **fragment**, souvent une partie spécifique de la page web (par exemple, un paragraphe).

Lorsqu'un caractère réservé doit être utilisé pour autre chose dans l'URL, il est remplacé par un symbole pourcent suivi de son code ASCII au format hexadécimal ('/' devient %2F et '?' devient %3F).

À retenir

- Une URL permet d'identifier une ressource précise sur un serveur web donné.
- Les URL permettent de passer des paramètres aux serveurs web.

15. Codes de retour HTTP

Mise en situation

Les réponses d'un serveur web à une requêtes HTTP peuvent être variées : renvoyer la ressource demandée, indiquer que la ressource envoyée par la requête a bien été reçue, refuser l'accès à une ressource, etc.

Pour permettre de déterminer le type d'une réponse, HTTP se base sur ce que l'on appelle les codes de retour (ou code d'état).

Code HTTP Az Définition

Un code HTTP est un code à 3 chiffres qui est présent dans chaque réponse HTTP du serveur. Ce code est destiné au client HTTP pour lui indiquer le statut de sa requête.

Les codes HTTP sont séparés en 5 familles, qui se distinguent par le premier digit du code retour :

- `1xx` : Informations
- `2xx` : Succès
- `3xx` : Redirections
- `4xx` : Erreur du client
- `5xx` : Erreur du serveur

Les codes HTTP courants Exemple

Il existe plusieurs dizaines de code HTTP, mais certains sont plus utilisés que d'autres, en voici quelques-uns :

- `200 OK` : « Tout est bon »,
- `201 Created`: « Ressource ajoutée avec succès »,

- `401 Unauthorized`: « Vous n'avez pas le droit d'accéder à cette ressource, il faut une authentification »,
- `404 Not Found`: « Ressource non trouvée »,
- `500 Internal Server Error`: « Erreur du serveur ».

Observer les codes HTTP

Exemple

La commande `curl` permet de montrer le code de retour HTTP reçu suite à une requête. Pour cela il suffit de rajouter l'option `-I`.

La première ligne affichée contient le code de retour HTTP.

```
$ curl -I https://fr.wikipedia.org/wiki/Hypertext_Transfer_Protocol
HTTP/2 200
```

Si on demande une page qui n'existe pas on obtient le code 404.

```
$ curl -I https://fr.wikipedia.org/fausse_page
HTTP/2 404
```

À retenir

- Chaque réponse HTTP comporte un code qui permet d'indiquer au client le statut de la requête.
- Il existe 5 familles de code différents (informations, succès, redirections, erreurs client, erreurs serveur)