

INTERNET

Stage – Semaine 5

Est-ce qu'internet = web ?

- NON
 - Internet = réseau physique mondial d'ordinateurs (ou plus précisément réseau de réseaux)
 - Internet à plus de 50 ans !
 - Le 29 octobre 1969, premier transfert (du texte « lo ») entre machines connectées à distance dans le cadre du projet ARPANET, pré-version d'internet
 - Voir par exemple le [site de l'ICANN](#)
 - Web (World Wide Web) = réseau d'informations constitués des documents mis à disposition sur les serveurs d'internet et reliés par les hyperliens qu'ils contiennent.
 - Créé au début des années 1990

RAPPELS : CLIENT-SERVEUR

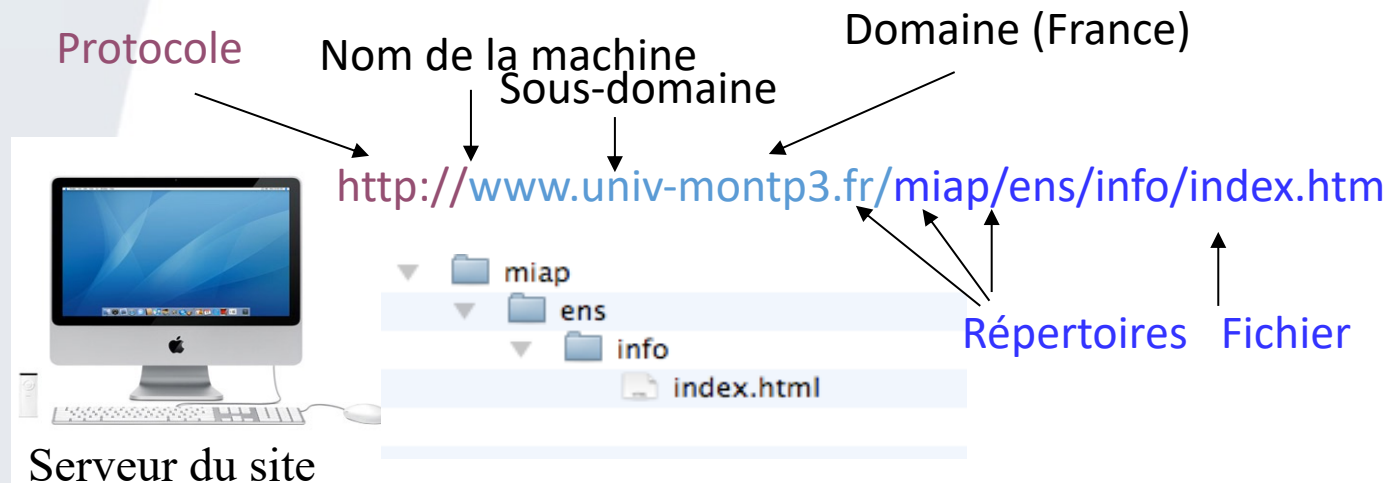
- Principe de base :
 - Un ordinateur (le client) demande un service (exemple : une page web) à un serveur
 - L'échange est effectué en suivant un ou des **protocoles, langages** gérant le dialogue entre machines
- Quelques protocoles :
 - **IP** (Internet Protocol) : gestion d'internet
 - **http** (HyperText Transfert Protocol) : gestion web
 - **https** : idem sécurisé (crypté)
 - **imap** (Interactive Message Access Protocol) : accès à des messageries électroniques à partir de logiciels clients
 - **smtp** (Simple Mail Transfer Protocol) : envoi de mails
 - **ftp** (File Transfert Protocol) : transfert de fichiers (utile quand vous hébergez un site web chez un fournisseur de services)
- Remarque : pare-feu : permet de gérer les protocoles autorisées

ADRESSAGE IP

- Nécessité de localisation : chaque ordinateur est associée à une adresse l'identifiant gérée par le protocole IP
 - Adresse IP = suite de 32 bits (représentés sous forme de 4 blocs de nombres entre 0 et 255 (8 bits)) en IPv4 ou 128bits en IPv6 en cours de déploiement)
 - Exemples : 193.52.137.213, 10.3.7.12 ...

ADRESSAGE HTTP

- Pour le protocole http, localisation d'une page web via URL (Uniform Resource Locator)



- Un site web = ensemble de pages web dont l'URL débute par un même début, appelé racine du site ou URL du site.
 - Exemples :
 - <http://www.univ-montp3.fr/> : site de l'université
 - <http://www.univ-montp3.fr/miap/ens/info/> site des enseignements d'informatique

- Lors d'une navigation sur le web, les données échangées avec les serveurs interrogés passent par des serveurs intermédiaires
 - Les données peuvent être lues par d'autres.
 - Des données peuvent être stockées :
par exemple, la LCEN, Loi pour la confiance dans l'économie numérique, modifiée par décret en juillet 2021, oblige les fournisseurs de services à conserver des informations relatives à l'identité civile des utilisateurs
- Autres traces
 - Cookies (avec RGPD – Règlement général sur la protection des données – vous connaissez !)
 - Historique de navigation
 - Métadonnées dans les fichiers
 - ...

Que se passe-t-il quand on saisit l'adresse IP d'une machine au lieu de l'URL d'un site ?

- Si l'IP correspond à l'adresse d'un site web, l'adresse IP est remplacée par l'URL du site et la page d'accueil du site est affichée (d'autres mécanismes peuvent être mis en place).
- Exemple : 193.52.137.213 ... serveur web de l'université (mais des redirections qui mettent en alerte Firefox)

QUELS OUTILS UTILISEZ-VOUS LORS D'UNE RECHERCHE D'INFORMATION SUR LE WEB ?

1. Un ordinateur (de bureau, portable, smartphone, ...) connecté à internet
2. Un navigateur Web :
 - Chrome, Firefox, Edge, Safari, Opéra, ...



3. Un moteur de recherche :
(site contenant un index
d'une partie du web)



...

Qwant

ou un méta-moteur de recherche
(site construisant sa réponse
en recoupant les résultats
de requêtes à différents
moteurs de recherche)



DuckDuckGo



...

- Remarque : moteur de recherche \neq navigateur web

EST-CE QU'UN MOTEUR DE RECHERCHE INDEXE TOUT LE WEB ?

- Non
 - Le web est dynamique :
 - Des pages se créent (et disparaissent) régulièrement
 - Les crawlers ou spiders, robots d'indexation, ne peuvent pas tout recenser en temps réel
 - Des sites non reliés aux autres (web invisible)
 - Des pages payantes ou privatisées par mot de passe
 - Potentiellement un problème de place dans la base de données d'indexation (et donc des choix sur ce que conserver)
 - ...

DANS LE TD

- Un zoom sur des requêtes dans un moteur de recherche
 - But = réduire le nombre de réponses et obtenir des réponses plus précises
- Réflexion sur la qualité des sites consultés :

La qualité des ressources est très variable.
C'est l'internaute qui doit la juger.