

Le Web

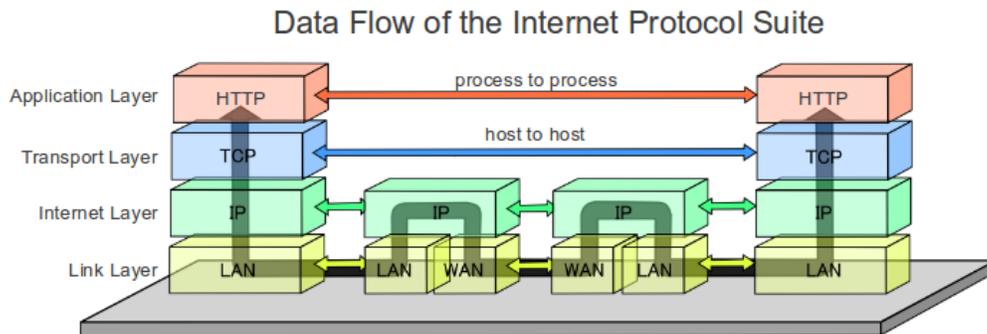


FIGURE 1 – Source :Wikipédia

Le Web désigne l'ensemble des informations disponible sur Internet par l'application **HTTP** pour Hyper text Transfert Protocol. Dans le modèle en couches, l'application HTTP est au-dessus de la couche Transport TCP

Qu'est ce qu'un Hypertexte ?

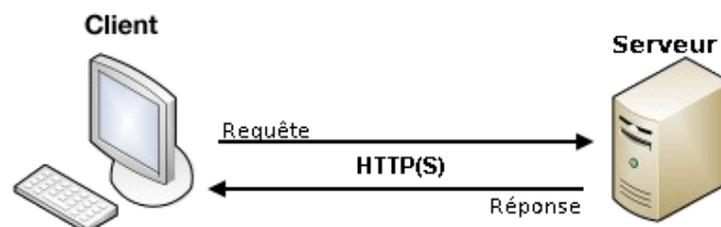
Un texte se lit de manière séquentielle, un hypertexte est un texte "enrichi" avec des **liens** qui permettent d'avoir accès à d'autres textes ou des sons ou des images et ce parcours dépend du lecteur suivant sa curiosité

On peut donc regarder le Web comme un ensemble d'hypertextes ou pages connectés entre eux.

Le Web est apparu au début des années 1990.

1 Modèle client-serveur, navigateurs Web et moteurs de recherche

La relation entre deux ordinateurs connectés sur Internet n'est pas symétrique. L'un a des informations (**le serveur**) que l'autre veut consulter (**le client**) La plupart



du temps quand on surfe sur le Web de lien en lien **en tant que client** on sollicite une information (texte, image, vidéo, son) qui se trouve sur un **serveur** distant suivant un **protocole** de communication nommé **HTTP** pour (Hyper Text Transfert Protocol) ou encore HTTPS (S pour Secure) lorsque les données sont chiffrées si on fait des achats en ligne par exemple

Sur l'ordinateur du client le **navigateur Web** est un logiciel permettant au client de mettre en oeuvre le protocole HTTP de manière graphique

A l'heure actuelle les navigateurs Web les plus utilisés (La Chine et la Russie mises à part) sont Chrome, Safari, Mozilla Firefox et Internet Explorer

Si j'utilise un navigateur (Firefox par exemple) je peux entrer l'**URL** (pour Uniform Resource Locator) de la page d'un site que je connais sans forcément passer par un **moteur de recherche**

Un moteur de recherche est un logiciel me permettant d'avoir un choix d'URL en fonction de ma recherche.

Le plus utilisé et le plus célèbre est Google mais il en existe d'autres comme Qwant (Européen) Par exemple imaginons que **je recherche un musée d'art moderne à Rome**, dans le moteur de recherche je peux entrer "*musée art moderne Rome*" et **celui ci va me proposer plusieurs réponses possibles** dans lesquelles je dois **faire un tri** et je finirai par avoir le lien sur lequel je vais cliquer et/ou noter l'URL du site du musée <http://lagallerianazionale.com> et utiliser plus tard un navigateur Web dans lequel je vais entrer l'URL ci-dessus

Dans la réalité de nos jours le navigateur Web est couplé avec un moteur de recherche à tel point que la différence entre les deux devient pour beaucoup imperceptible

L'association la plus courante est Chrome+Google mais d'autres personnes vont choisir par exemple Firefox+Qwant pour éviter le traçage publicitaire intempestif

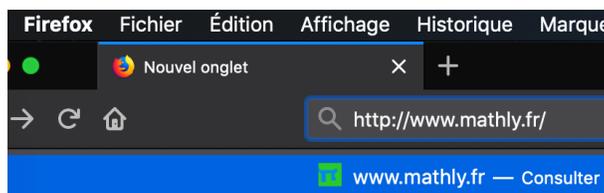
Ex 1

Paramétrer Firefox pour que la page d'accueil soit le moteur de recherche Qwant

2 Protocole HTTP

Regardons de plus près le protocole HTTP

Par l'intermédiaire du navigateur Firefox on demande à consulter la page d'accueil du site [mathly.fr](http://www.mathly.fr)



Une fois la page chargée, je peux avec l'aide de la console Web avoir des informations réseau sur le chargement de la page On voit que :

État	Mé...	Domaine	Fichier	Source	Type	Transfert	Taille	0 ms	160 ms	320 r	En-têtes	Cookies	P
200	GET	www.ma...	/	document	html	1,24 Ko	2,07...	14 ms			URL de la requête : http://www...		
												Méthode de la requête : GET	
												Adresse distante : 85.236.158.8	
												Code d'état : 200 OK	
												Version : HTTP/1.1	
												Modifier et renvoyer	
												Filter les en-têtes	
												En-têtes de la réponse (214 o)	
🕒	1 requête	2,07 Ko / 1,24 Ko transférés	Terminé en : 511 ms	DOMContentLoaded: 49 ms	load: 463 ms								

1. La méthode de la requête est GET
2. Le code est 200 (la page a bien été chargée) (un code d'erreur connu est le fameux 404)
3. On a l'adresse IP du serveur distant

Maintenant pour être plus précis nous utilisons le logiciel Wireshark qui permet d'analyser les paquets d'information

Voici la demande (frame 284) par le client 192.168.1.21 au serveur 85.236.158.88 de la page d'accueil du site mathly.fr

```

▶ 284 37.315619 192.168.1.21 85.236.158.88 HTTP 414 GET / HTTP/1.1
▶ Frame 284: 414 bytes on wire (3312 bits), 414 bytes captured (3312 bits) on interface 0
▶ Ethernet II, Src: Apple_18:68:79 (a0:99:9b:18:68:79), Dst: Sagemcom_3e:ba:9e (40:c7:29:3e:ba:9e)
▶ Internet Protocol Version 4, Src: 192.168.1.21, Dst: 85.236.158.88
▶ Transmission Control Protocol, Src Port: 62750, Dst Port: 80, Seq: 1, Ack: 1, Len: 348
▼ Hypertext Transfer Protocol
  ▶ GET / HTTP/1.1\r\n
    Host: www.mathly.fr\r\n
    User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10.14; rv:68.0) Gecko/20100101 Firefox/68.0\r\n
    Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8\r\n
    Accept-Language: fr,fr-FR;q=0.8,en-US;q=0.5,en;q=0.3\r\n
    Accept-Encoding: gzip, deflate\r\n
    Connection: keep-alive\r\n

```

Ex 2

1. Que signifie "bytes" ?
2. Quelle est la taille en octet du paquet 246 (Frame 246) ?
3. Ce paquet contient une en-tête, dans l'ordre d'apparition sur le canal de transmission, Liaison (Wi-Fi), Réseau (IP) et Transport (TCP).
Rechercher sur le Web la taille en octets de ces en-têtes et en déduire la taille restante du message HTTP

Quelques millièmes de seconde après (frame 286) arrive la réponse

284	37.315619	192.168.1.21	85.236.158.88	HTTP
285	37.321841	85.236.158.88	192.168.1.21	TCP
286	37.416540	85.236.158.88	192.168.1.21	HTTP
287	37.416614	192.168.1.21	85.236.158.88	TCP

▶ Frame 286: 1334 bytes on wire (10672 bits), 1334 bytes captured (10672 bits) on interface 0
▶ Ethernet II, Src: Sagemcom_3e:ba:9f (40:c7:29:3e:ba:9f), Dst: Apple (08:00:27:00:00:00)
▶ Internet Protocol Version 4, Src: 85.236.158.88, Dst: 192.168.1.21
▶ Transmission Control Protocol, Src Port: 80, Dst Port: 62750, Seq: 312112000, Win: 65535, Len: 0
▼ Hypertext Transfer Protocol
▶ HTTP/1.1 200 OK\r\n
Content-Type: text/html; charset=UTF-8\r\n
Content-Length: 1054\r\n
Content-Encoding: gzip\r\n
Vary: Accept-Encoding,User-Agent\r\n
Date: Sun, 18 Aug 2019 20:05:49 GMT\r\n
Server: LiteSpeed\r\n
Connection: close\r\n
\r\n
[HTTP response 1/1]
[Time since request: 0.100921000 seconds]

Ces éléments d'information sont ensuite "lus" par le navigateur du client et l'utilisateur peut voir la page Web demandée sur son écran

Ex 3

La taille maximale d'un paquet est 1500 octets .

Supposons que l'on veuille envoyer sur le réseau un fichier de 3 Ko.

En combien de paquets sera fragmenté le fichier par le protocole TCP (tenir compte de la taille des en-têtes) ?

On peut affiner la requête en ajoutant des paramètres dans l'URL.

Regardons sur un exemple.

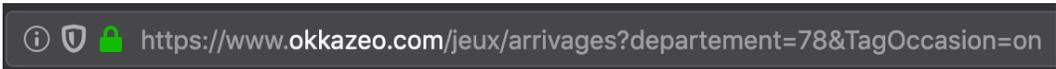
Si vous allez sur le site www.okkazeo.com , vous pouvez voir en fonction du département des jeux de société d'occasion

Ainsi par exemple on peut cliquer sur la carte de France et choisir le département 91 on obtient alors dans la barre d'URL Après le point d'interrogation la valeur du paramètre



`departement` est 91, ainsi si on veut maintenant choisir le département 78, on peut soit revenir à la carte soit directement modifier dans la barre URL le département en remplaçant 91 par 78

Une fois que l'on a choisi le département 78 on peut faire un second choix et ne vouloir que les jeux d'occasion en cochant une case pour cela et ensuite en cliquant sur Filtrer on observe alors dans la barre d'URL Dans la barre URL on observe que la requête est

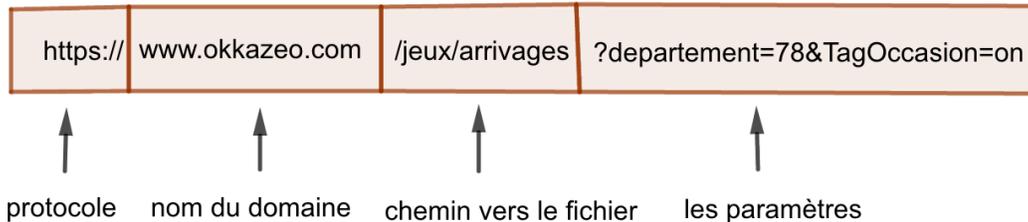


devenu `departement=78&TagOccasion=on`

On peut affiner la requête en ajoutant un paramètre avec le symbole `&` pour et

Ainsi si on voulait uniquement les jeux neufs on peut modifier l'URL ainsi `departement=78&TagNeuf=on`

En conclusion on retiendra la structure de l'URL d'une page



Ex 3

Décomposer les URL suivantes :

1. `https://www.cloudflare.com/learning/dns/dns-server-types/`
2. `https://www.cloudflare.com/learning/dns/dns-server-types/#tld-nameserver`

3 Langages HTML et CSS

Tout navigateur Web permet de "regarder à l'intérieur" d'une page Web.

Nous allons utiliser l'outil "Inspecteur" du navigateur Mozilla sur la page de TP `mathly.fr/\tp_2nd.php` et nous verrons "en direct" comment la page est structurée par les principales balises du langage **HTML** (pour HyperText Markup Language)

Nous verrons aussi à ce moment quelques éléments du langage **CSS** (pour Cascading Style Sheet) pour la mise en forme de la page

Pour le reste voir TP

4 Exemples d'autres protocoles que HTTP "au-dessus de TCP/IP"

"Au-dessus de TCP/IP" il existe d'autres protocoles par exemple qui se servent de TCP/IP

1. Le protocole **FTP** (pour File Transfert Protocol) permet d'envoyer un fichier, des images etc... d'un client vers un serveur distant et vice-versa
2. Le protocole **SMB** qui permet de partager des fichiers dans un réseau local

5 Histoire

Voir ici `https://www.ina.fr/video/CAB95050558/internet-video.html` un reportage de télévision (France2, septembre 1995) au sujet de l'arrivée du Web en France. Vous constaterez la confusion entre Internet et le Web

1. **1990** : Création du World Wide Web par Tim Berners Lee (Prix Turing 2016) au CERN de Genève, en mettant au point le protocole HTTP, le langage HTML et un premier navigateur
2. **1993** : Apparition du navigateur Web Mosaïc qui rend populaire le Web
3. **1995** : Apparition des langages Javascript (client) et PHP (serveur) qui rendent les pages Web dynamiques (le client peut interagir avec la page Web)
4. **1998** : Apparition du moteur de recherche Google

6 Enjeux sociétaux

"Nos maisons sont dans la ligne de mire du capitalisme de surveillance. Des entreprises spécialisées se disputaient en 2017 un marché de 14,7 milliards de dollars pour des appareils ménagers connectés, contre 6,8 milliards l'année précédente. À ce rythme-là, le montant atteindra 101 milliards de dollars en 2021. Commercialisés depuis quelques années, des objets absurdes se tiennent à l'affût dans nos intérieurs : brosse à dents intelligente, ampoule intelligente, tasse à café intelligente, four intelligent, extracteur de jus intelligent, sans oublier les couverts intelligents censés améliorer notre digestion. D'autres semblent plus inquiétants : une caméra de surveillance à domicile avec reconnaissance faciale, un système d'alarme qui repère les vibrations inhabituelles précédant un cambriolage, des GPS d'intérieur, des capteurs qui s'adaptent à tous les objets pour analyser le mouvement et la température, sans oublier des cafards cyborgs qui détectent les sons. Même la chambre du nourrisson est repensée pour devenir une source de surplus comportemental.

Tandis que la course aux profits générés par la surveillance s'exacerbe, les capitalistes s'aperçoivent que les économies de gamme ne suffisent pas. Certes, l'excédent de données doit être abondant et varié ; mais le moyen le plus sûr de prédire le comportement reste d'intervenir à la source : en le façonnant. J'appelle « économies de l'action » ces processus inventés pour y parvenir : des logiciels configurés pour intervenir dans des situations réelles sur des personnes et des choses réelles. **Toute l'architecture numérique de connexion et de communication est désormais mobilisée au service de ce nouvel objectif.** Ces interventions visent à augmenter la certitude en influençant certaines attitudes : elles ajustent, adaptent, manipulent, enrôlent par effet de groupe, donnent un coup de pouce. Elles infléchissent nos conduites dans des directions particulières, par exemple en insérant une phrase précise dans notre fil d'actualités, en programmant l'apparition au moment opportun d'un bouton « achat » sur notre téléphone, en coupant le moteur de notre voiture si le paiement de l'assurance tarde trop, ou encore en nous orientant par GPS dans notre quête de Pokémon. "

Shoshana Zuboff Professeure émérite à la Harvard Business School. Auteure de The Age of Surveillance Capitalism : The Fight for a Human Future at the New Frontier of Power, Public Affairs, New York, 2019. (article dans le monde diplomatique de Janvier 2019)

Questions

1. Que signifie pour vous "capitalisme de surveillance" ?
2. Des sociétés vous proposent des colliers avec balise GPS intégré pour ainsi suivre votre animal de compagnie. Seriez vous prêts à acheter un tel collier ? Pourquoi ?

3. Donner d'autres exemples "d'économies de l'action"
4. Peut-on considérer la publicité faisant partie des économies de l'action ? Qu'est ce qu'un "influenceur" ?