

## 22 - Exemples d'activités autour de l'internet : structure, indexation et partage des données, sécurité

### I- Internet : Notion de réseaux

Un réseau est un ensemble interconnecté. Dans le cadre d'un réseau informatique, c'est un ensemble de machine électronique (ordinateur, serveur, téléphone, montre connecté,...) qui sont connectées entre elles.

*LAN, MAN, WAN : Kesako ?*

- **Local Area Network** : c'est un réseau de terminaux très proche géographiquement (même pièce, maison, bâtiment)
- **Metropolitan Area Network** : c'est un réseau de terminaux plus vaste pouvant englober une ville
- **Wide Area Network** : Réseau couvrant un pays, un continent ou même une planète, il relie entre eux les MAN et les LAN.

***Le WAN le plus vaste est Internet***

Les ordinateurs communiquant au travers d'Internet sont reliés les uns aux autres par des voies de télécommunications (structures physiques du réseau : câbles, fibre optique, ondes, ADLS) et utilisent un même protocole de communication.

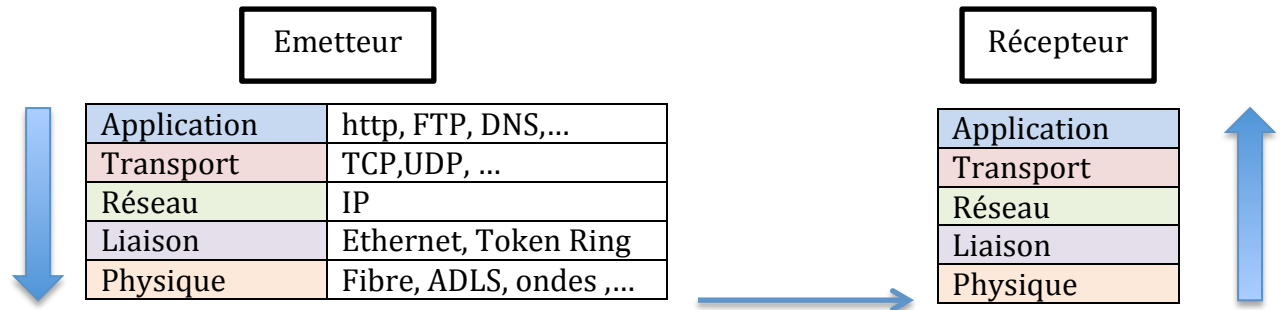
La **notion de protocole** est importante car c'est le protocole qui va dicter les règles de communication entre les terminaux, qu'ils soient basés sur la même technologie ou non. La suite de protocoles utilisée par Internet est **TCP/IP**.

- **Transmission Control Protocol (TCP)**. Ce protocole définit les règles et assure la bonne communication entre les ordinateurs, qu'ils soient basés sur la même technologie ou non. Il décompose en «paquets» les informations envoyées et dirige ces paquets vers leur destination selon leur type (message électronique, transfert de fichiers, etc.)
- Le protocole **IP (Internet Protocol)** est utilisé pour transmettre les données d'une machine à une autre jusqu'à leur destination finale.

Il existe d'autre protocole utilisé par les services offerts par internet :

Services	Protocoles
Visualisation de page web	<b>HTTP (HyperText Transfert Protocol)</b>
Transferts de fichier	<b>FTP (File Transfer Protocol)</b>
Transferts de courriers électronique	<b>SMTP (Simple Mail Transfert Protocol)</b>
Récupération des e-mails sur le serveur	<b>POP (Post Office Protocol)</b>
Forums de discussion en temps différés	<b>NNTP (Network News Transfert Protocol)</b>
Dialogue en temps réel	<b>IRC (Internet Relay Chat)</b>

A chaque utilisation d'internet, l'utilisateur va utiliser ces différents protocoles, la plus part de temps les uns à la suite des autres, on appel ça la **pile de protocole**. Ces enchainements de protocoles sont totalement transparents aux yeux de l'utilisateur et s'effectuent aux différentes couches du modèle TCP/IP d'internet :



Internet relie plusieurs milliards de machines entre elles, il peut donc être représenté comme un graphe où les nœuds représentent les terminaux et où les arrêtes représentent les connexions entre les terminaux.

Pour transmettre une information d'un terminal A à un terminal B, il faut donc trouver le plus court chemin à travers ce graphe, on appelle ça « le routage ». (Cf. Dijkstra programme ES, même si les algorithmes de routage utilisés sont bien plus efficaces et complexes). Le routage est un enjeux primaire sur un réseau comme internet, car les communications y sont extrêmement nombre à un même instant t. Il y a donc des problèmes de blocage.



#### **Activité Routage :** [link](#)

Dans le cas d'internet, l'adresse du nœud (le terminal), n'est pas une lettre mais une adresse IP. Chaque machine connectée à internet possède une adresse IP unique qui lui permet d'être authentifiée sur le réseau.

Les adresses IP sont sous la forme xxx.xxx.xxx.xxx où chaque « xxx » est un nombre compris entre 0 et 255 (IPv4).

#### **Activité adresse IP :**

- Sur combien d'octet est stocké une adresse IP ?
- Sur combien de bit est stocké une adresse IP ?
- Combien y'a-t-il d'adresse IP possible ?
- En imaginant qu'aucune adresse IP ne soit encore allouée et que dans le monde, il se vend 45 millions d'objets connectés nécessitant une adresse IP disponible chaque jour. Sachant que cette vente augmente de 2,5% chaque jours, au bout de combien de jours le stock d'adresse IP disponible sera-t-il épuisé ?
- Quelle solution peut-on trouver subvenir au besoin mondial d'adresse IP ?

## II- Le World Wide Web

### Un peu d'histoire :

Il ne faut pas confondre internet et le web. Le web est un service proposé par internet, au même titre que la messagerie. Le web est l'ensemble des pages web et des liens qui relient ces pages.

Le web est né en 1989 à la frontière franco-Suisse, au CERN, lorsque l'informaticien Tim Berners-Lee décide de créer un système hypertexte pour que les collaborateurs s'échangent des informations.

Les technologies du web sont versées dans le domaine public en 1991.

En 1994 on recense 130 sites internet

En 2014, le milliard de site est dépassé.

### L'architecture client-serveur :

Toutes les pages web sont stockées sur des serveurs un peu partout dans le monde, pour y accéder, le client (navigateur internet) fait une requête pour que le serveur lui transmette la page. Cette requête utilise le protocole **HTTP** et le principe **d'URL (Uniform Ressource Locator)**. Cette ressource permet d'identifier la localisation d'une page web afin d'interroger le bon serveur.

Ex :

[https://www4.ac-nancy-metz.fr/capesmath/data/uploads/info\\_2018.pdf](https://www4.ac-nancy-metz.fr/capesmath/data/uploads/info_2018.pdf)

Le protocole, ici **HTTPS** signifie qu'il est plus sécurisé.

Le nom de domaine / serveur quelque part dans le monde

Arborescence disque

Nom du fichier

### Composition du page web

Une page web est écrite avec un langage spécifique au web, le **HTML (HyperText Markup Language)**. C'est un langage dit balisé, c'est à dire qu'on va écrire : « <h1> mon texte </h1> », et au yeux du navigateur « mon texte » sera considéré comme un titre de niveau 1 (<h1>). Ce langage balisé permet de hiérarchiser le texte et ainsi de modifier l'apparence du texte en fonction de sa balise dans un fichier de style **CSS (Cascading Style Sheet)**.

```
1 <!doctype html>
2 <html lang="fr">
3 <head>
4 <title>tp-CSS</title>
5 <link rel="stylesheet" type="text/css" href="base.css" />
6 <meta charset="utf-8">
7 </head>
8 <body>
9 <div class="wrapper">
10 <header>
11 <h1>Mobile Phone Company</h1>
12 <ul id="gsmenu">
13 <li><a href="#"> id="p1" accesskey="1" title="Homepage (Alt-1)">Home</a></li>
14 <li><a href="#startcontent" id="p2" accesskey="2" title="Skip over navigation (Alt-2)">Skip nav</a></li>
15 <li><a href="#"> id="p3" title="News about your site">News</a></li>
16 <li><a href="#"> id="p4" accesskey="9" title="Contact us by email or via an electronic form (Alt-9)">Contact</a></li>
17 <li><a href="#"> id="sitemap" accesskey="s" title="Sitemap (Alt-s)">Sitemap</a></li>
18 <li><a href="#"> id="p5" accesskey="h" title="Help and Accessibility Charter (Alt-h)">Help</a></li>
19 </ul> <!-- This ends the system menu options -->
20 </header> <!-- End of divHeader -->
21 <div id="Contenu_principal">
22 <div id="changer" src="Habilage/fleche-bleue.gif" alt="Changer de modèle">
23 
24 <div id="article">
25 <p>Donec ultricies varius quam. Cum sociis natoque penatibus et magnis dis parturient montes,
26 nascetur ridiculus mus. Aliquam neque magna, varius id, tempor non, molestie ac, magna.
27 Aliquam non mi. Mauris quam purus, imperdiet at, ullamcorper id, tincidunt sed, ligula.
28 Ut nulla nibh, consectetur ac, faucibus quis, adipiscing vitae, nunc. Fusce quis mauris
29 nec erat mattis faucibus.
30 </p>
31 <h2>Title - h2</h2>
32 <p>
33 Donec ultricies varius quam. Cum sociis natoque penatibus et magnis dis parturient montes,
34 nascetur ridiculus mus. Aliquam neque magna, varius id, tempor non, molestie ac, magna.
35 Aliquam non mi. Mauris quam purus, imperdiet at, ullamcorper id, tincidunt sed, ligula.
36 </p>
37 </div>
38 </div>
39 </body>
40 </html>
```

```
body{
  background: #000;
  color:#8a8888;
}

a{
  color:#8a8888;
  text-decoration: none;
}

.wrap{
  width:960px;
  margin: 0 auto;
}

.description{
  background-color: #000;
  color:#8a8888;
  line-height: 50px;
  margin:20px;
}

.rouge{
  color:#d91b3c;
}

.fondRouge{
  background-color:#d91b3c;
  color:#fff;
}

header{
  overflow: hidden;
}

header nav{
  float:left;
  width:180px;
}
```

Il existe deux types de pages web :

- Les **pages statiques** : Une page web statique est une page web dont le contenu ne varie pas en fonction des caractéristiques de la demande, c'est-à-dire qu'à un moment donné tous les internautes qui demandent la page reçoivent le même contenu ( site du jury )
- Les **pages dynamiques** : À l'inverse, une page web dynamique est générée à la demande et son contenu varie en fonction des caractéristiques de la demande (heure, adresse IP de l'ordinateur du demandeur, formulaire rempli par le demandeur, etc.) qui ne sont connues qu'au moment de sa consultation (ex : réseau sociaux)

Ces pages dynamiques utilisent un autre langage de programmation, **le PHP**, plus orienté algorithme, au contraire du HTML et du CSS qui sont orientés description.

Ce langage permet de communiquer avec une base de données et ainsi récupérer des informations dans cette base pour les afficher sur la page. Par exemple sur *Youtube*, en fonction des abonnements, la page web envoie une requête à la base pour récupérer les dernières vidéos des abonnements et ainsi les afficher sur la page.

Activité :

Création d'une page web statique avec au début, juste le langage HTML, et ensuite mettre en forme avec du CSS.

Activité :

Réalisation d'un page web dynamique en python avec le framework CherryPy. Cela permet de communiquer avec une base de donnée locale sans apprendre le langage PHP. Néanmoins, il est nécessaire de connaître les instructions de base du langage **MySQL**.

Sur le web, pour aller sur une page il n'est pas utile de connaître tout les URL par cœur, il existe un outil très pratique appeler **moteur de recherche**. Il en existe des dizaines comme *Google, Bing, Yahoo, ...* Pour trouver une page web, il suffit de taper un ou plusieurs mots clés et le moteur va afficher une liste de liens vers des sites contenant le mot clé. La question est : Comment classer les pages entre elles ? Laquelle mérite d'être première ? Pourquoi ?

Depuis la première version du page rank, les algorithmes se sont améliorés mais sont souvent propriétaires. Le premier algorithme rendu public par les créateurs de Google est celui qu'on peut décrire en développement.

Développement : [LE PAGE RANK](#)

### III- Sécurité, bonnes pratiques et enjeux

#### Chiffrement et authentification :

Pour des question de respect de la vie privé, de secret professionnels, de données sensibles ou classées secret défense, il est nécessaire de crypter les données et les communications.

Activité et développement de chiffrement simples : César, ROT13, affine, Vigenère, stéganographie

Description brève de systèmes de cryptographie symétrique/asymétrique/hybride

*Voir leçon 29. Exemples d'algorithmes de chiffrement et de déchiffrement.*

Garantir l'intégrité des données : bit de contrôle, exemple de l'ISBN.

Activité : [bit de parité](#)

Authentification et chiffrement sur le WWW : cas de [SSL](#)

Le Big data et les questions de sécurité que cela engendre.

Activité utilisant Wireshark : regarder le contenu de paquets de données chiffrées et non- chiffrées

Activité [Man-in-the-Middle](#)

#### Vie Privée et identité numérique

Informations personnelles sur les réseaux sociaux. Ce que dit la loi.

Identité numérique : ensemble des traces laisser sur internet permettant de mieux cerner un individu : Ce que l'ont dit ou publie, ce que l'on aime, ce qu'on partage, ...  
L'anonymat ne nous protège pas aux yeux de la loi, d'autant plus que tous ce qui est dit sur internet est stocké.

#### Ouverture

[Net neutrality](#)