

Chapitre 2

Interaction humain machine sur le Web

2.1 Internet et son fonctionnement

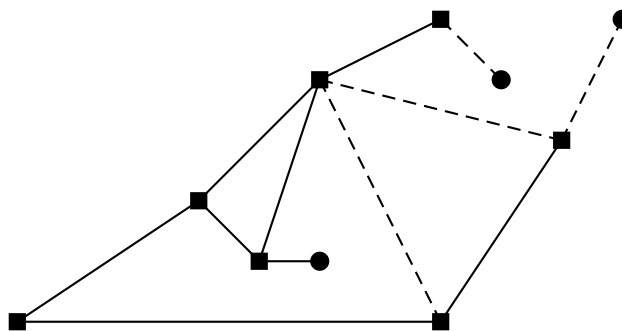
★ Vidéo de Monsieur Bidouille partie 1 et partie 2.

2.1.1 Les réseaux informatiques

Un réseau informatique est un ensemble de machines reliées entre elles et s'échangeant des informations. Par exemple le réseau au sein d'un domicile, d'une entreprise, d'un lycée, etc.

Internet désigne le réseau reliant l'ensemble de ces réseaux informatiques. Les ordinateurs reliés entre eux au sein de ces réseaux et donc d'Internet le sont par des moyens filaires (fibre optique, ADSL, etc.) ou non filaires (Wifi, Bluetooth, etc.). Internet est donc indépendant du moyen physique de connexion, il est défini par ses protocoles.

On pourrait dessiner Internet de cette façon :



où ■ désigne un réseau informatique, ● une machine seule, un trait pointillé une connexion non filaire et un trait plein une connexion filaire.

2.1.2 Circulation des données sur internet

Les données transférées sur internet sont découpées en paquets. Les paquets sont expédiés de l'expéditeur vers le destinataire en passant par des routeurs aux nœuds du réseaux qui décident vers quel autre nœud il doit les envoyer ; un routeur choisit parmi les différents chemins possibles en fonction de l'encombrement du réseau, des pannes, etc.

L'adressage des machines sur le réseau est assuré par le protocole IP (Internet Protocol). Une adresse IP est constitué de 4 nombres compris entre 0 et 255, par exemple 66.178.234.0. Les paquets circulant sur le réseau possèdent dans leurs en-têtes les adresses IP de leurs expéditeur et destinataire. Les routeurs possèdent des tables d'adresses IP qui permettent de guider les paquets vers le routeur suivant. L'adresse IP est fournie à la machine par le fournisseur d'accès au réseau, elle n'est donc pas propre à la machine même si elle permet de l'identifier.

L'intégrité des données et leur assemblage est assurée par le protocole TCP (Transmission Control Protocol). Le paquet contient un numéro qui permettra de savoir comment l'assembler avec les autres et on vérifie que les données qu'il contient n'ont pas été altérées lors des transferts.

La durée de vie d'un paquet est limitée afin que celui-ci ne circule pas éternellement sur le réseau.

Le système IP est actuellement en cours de transition de l'IPv4 (sur 4 octets) vers l'IPv6 (sur 16 octets) en raison de l'épuisement des adresses IPv4 à cause de l'augmentation constante de machines connectées à internet.

2.1.3 L'annuaire d'internet

Une adresse IP est un outil pour les machines mais pas pour les humains ; en effet, il est difficile de retenir que son site préféré a pour adresse 113.45.666.1 ! Les humains utilisent des adresses symboliques tels que mouette.org qui sont plus faciles à retenir.

La correspondance entre les adresses symboliques et IP est effectuée par les serveurs DNS (Domain Name System). L'annuaire DNS est organisé en domaines et sous-domaines, par exemple .org est un domaine et wikipedia.org un sous domaine de celui-ci. Lorsque l'on saisit une adresse symbolique, une requête est envoyée à un serveur DNS qui renvoie l'IP correspondante et notre ordinateur se connecte alors à l'aide de l'IP.

Exemple de lecture d'une adresse symbolique

https :// $\underbrace{\text{fr}}_3$. $\underbrace{\text{wikipedia}}_2$. $\underbrace{\text{org}}_1$ / $\underbrace{\text{wiki}}_4$ / $\underbrace{\text{Loutre}}_5$

Le https : ne fait pas partie de l'adresse symbolique, il désigne en fait le protocole de connexion https (hypertext transfer protocol secure). Une adresse symbolique ne se lit pas tout à fait de gauche à droite mais comme suit.

1. Le .org est le domaine de premier niveau, il indique ici que nous sommes une adresse d'une organisation non gouvernementale.
2. Le sous-domaine wikipedia.
3. Une fois dans le domaine wikipedia.org, le fr indique que nous sommes sur la partie française du site.
4. Dans la partie française de wikipedia.org, on va dans le dossier wiki.
5. Dans ce dossier, on va sur la page loutre.

2.2 Le Web et son fonctionnement

Le concept du Web est né en 1989 au CERN (Conseil Européen de Recherche Nucléaire) lorsque Tim Berners-Lee proposa à ses collègues chercheurs de mettre en réseau leurs documents de recherche afin de les partager plus facilement et massivement. Le langage HTML permettant la mise en réseau de ces informations apparut en 1991 lors du lancement du Web.

Le Web est donc un réseau d'informations (textes, images, sons...) supportées par Internet ; il est donc à ne pas confondre avec ce dernier : les deux sont aussi différents qu'un dessin et la feuille sur laquelle il est fait. Ces informations ou documents qui les contiennent sont reliées entre elles par le principe de l'hypertexte : texte contenant des liens vers d'autres documents. Ces documents sont hébergés sur des serveurs informatiques.

Un site Web est l'un des formats possibles de ces informations mais on peut directement accéder à un document (par exemple un pdf) sur le Web si on en connaît l'adresse URL (Uniform Resource Locator).

Pour consulter une page Web, un ordinateur client envoie une requête à un serveur qui lui renvoie les informations demandées, celles-ci sont alors mise en forme par le navigateur. Les échanges sont effectués selon le protocole HTTP(S) (Hypertext Transfer Protocol (Secure)) qui permet la communication entre client et serveur.

2.2.1 Les langages HTML, CSS, JavaScript, PHP

Les pages Web sont programmées en langages HTML (Hypertext Markup Language) et CSS (Cascading Style Sheets).

Le HTML sert à programmer le contenu de la page (le texte, les menus, les adresses des fichiers qu'il doit appeler : CSS, JavaScript, images, sons...).

Le CSS sert à programmer l'apparence de la page (couleurs, polices, styles...), le fond est donc séparé de la forme !

Le JavaScript est un langage permettant des interactions de l'utilisateur avec la page Web, notamment par la modification de contenus ou balises HTML, ou par l'obtention d'animations.

Le PHP (PHP : Hypertext Preprocessor) est un langage libre développé par Rasmus Lerdorf en 1994 basé sur le langage C. Le plus souvent sur des serveurs Web, il permet de générer du code (HTML et CSS entre autres) et donc de personnaliser les contenus proposés à l'utilisateur.

HTML et CSS sont des langages de description / balisage. Ils servent à décrire les propriétés d'un contenu placé dans un environnement particulier. JavaScript et PHP sont eux, à l'instar de Python, des langages de programmation. Ils permettent contrairement aux deux précédents de déclarer des variables, d'effectuer des opérations, des tests, d'utiliser des boucles, des fonctions, etc. Ils peuvent être directement intégrés aux fichiers HTML à l'aide de balises dédiées mais ils sont généralement dans des fichiers séparés qui sont appelés lors de la lecture du fichier HTML.

2.2.2 Web statique vs Web dynamique

Bien que chacun des programmes associés aux langages ci-dessus puissent être stocké dans un fichier dédié (.html, .css, .js, .php), les échanges clients-serveurs sur le Web se font essentiellement sous forme de flux. Autrement dit, le navigateur client ne télécharge pas de fichiers depuis le serveur mais reçoit un flux de données de divers types.

Les sites dits statiques – qui furent les premiers du Web – affichent un contenu et une apparence identiques pour tous les utilisateurs. Par ailleurs, ils ne permettent pas réellement d'interaction avec l'utilisateur.

Les sites dits dynamiques personnalisent le contenu et l'apparence en fonction de l'utilisateur, notamment grâce à l'accès à des serveurs de base de données en plus des serveurs HTTP. Ils permettent une réelle interaction entre l'utilisateur et le site. Les langages JavaScript et PHP, permettant la modification et la personnalisation des sites Web, sont donc à la base du Web dynamique.

Il est important de noter toutefois que chacun des langages cités ci-dessus peut se passer de l'architecture client-serveur. Ils peuvent directement être interprétés et exécutés localement par le navigateur.

