

## 1.Introduction

Pour résumer en une phrase ce qu'est **Internet**, on pourrait dire :

«**Ensemble de réseaux interconnectés, utilisant le protocole TCP/IP**»

C'est en effet un **réseau de réseaux**, que l'on trouve dans les universités, les entreprises, les administrations, les centres de recherches, ..., et cela dans le **monde entier**.

Pour pouvoir dialoguer entre eux, ces réseaux doivent tous parler le même «langage»

(on dit «*protocole*» en informatique) : **TCP/IP**

## 2. Historique

## 3.Technique

### 3.1.Les protocoles

**Internet** est un réseau de réseaux. Pour que ces réseaux puissent communiquer entre eux, il faut qu'ils parlent **la même langue**, qu'on appelle «protocole» en informatique. Ce protocole définit des règles précises de structures des données qui vont circuler.

Quand deux machines veulent dialoguer entre elles sur Internet, leurs phrases vont commencer par dire «je parle la langue Internet» (**IP** = *Internet Protocol*), puis annoncer le sujet de conversation

(transfert de fichiers, messagerie, «WEB», ...)

### 3.2.L'adressage

De plus, comme les «dialogues» se font à distance, en utilisant des voies communes, chaque «phrase» doit indiquer en plus à qui elle est destinée, et qui en est l'expéditeur.

Toute machine connectée à Internet est identifiée par une **adresse unique**, dite «**adresse IP**», sous la forme d'un nombre binaire codé sur **32 bits** (à l'heure actuelle). Cela veut dire qu'il ne peut pas y avoir aujourd'hui plus de  $2^{32}$  (2 à la puissance 32) machines, soit 4 294 967 296. (On est en train de prévoir un adressage plus grand, **IPv6**, car cette limite sera vite atteinte, vu le rythme de développement d'Internet)

Comme le maniement d'un tel nombre n'est pas très aisé, on l'épelle communément sous la forme de **4 nombres successifs**, obtenus par «tronçonnage» en 4 morceaux égaux du nombre initial. Et comme le binaire n'est pas pratique non plus, on exprime ces 4 nombres en décimal.

Par exemple, les serveur Web suivants ont pour adresses respectives :

EDF	192.54.193.134	Gaz de France	212.208.192.66
Microsoft	207.46.130.14	l'Elysée	193.252.69.53
La CIA	198.81.129.99	Ministère du Travail	62.161.44.6

Ainsi, quand un internaute se connecte depuis son PC sur le site de Microsoft, tous les «paquets» de données ou de requêtes qu'il émet contiennent l'adresse «207.46.130.14», plus sa propre adresse (pour que Microsoft puisse lui répondre)

### 3.3.Les noms de domaine

Mais il est évident que la manipulation de tels nombres n'est pas pratique non plus. On va donc donner à chaque machine un nom symbolique, appelé «**nom de domaine**», beaucoup plus parlant et facile à retenir par un interlocuteur humain, et c'est un «service d'**annuaire**» qui effectuera la traduction nom symbolique-adresse. Ce service, géré automatiquement par des serveurs dialoguant entre eux est appelé «**DNS**» (*Domain Name Service*)

Pour reprendre les exemples précédentes, les serveur Web suivants ont pour noms symboliques respectifs :

EDF	<a href="http://www.edf.fr">www.edf.fr</a>	Gaz de France	<a href="http://www.gazdefrance.com">www.gazdefrance.com</a>
Microsoft	<a href="http://www.microsoft.com">www.microsoft.com</a>	l'Elysée	<a href="http://www.elysee.fr">www.elysee.fr</a>
La CIA	<a href="http://www.cia.gov">www.cia.gov</a>	Ministère du Travail	<a href="http://www.travail.gouv.fr">www.travail.gouv.fr</a>

Chaque point, dans le nom symbolique, sépare ce qu'on appelle un «domaine» d'un «sous-domaine».

Les domaines de **1er niveau** sont définis à l'échelle mondiale par l'**ICANN**

(Internet Corporation for Assigned Names and Numbers). Il existe ainsi :

**.com** entreprises commerciales (américaines au départ mais ce n'est pas une obligation)

**.edu** organismes d'enseignement américains

**.gov** organisations gouvernementales des États-Unis

**.mil** armée américaine

**.org** autres organisations

**.net** sites techniques, ressources propres du réseau

**.int** organismes internationaux

De plus, chaque pays possède son propre domaine («*fr*» pour la France, «*de*» pour l'Allemagne,

«*ru*» pour la Russie, ...)

Le domaine de **2ème niveau** définit l'entreprise (ou organisme, université, ...).

Par exemple «.edf», «.microsoft», «.bellamyjc» ,...

Les domaines de **3ème niveau** et au-delà sont **facultatifs**, et sont gérés par l'entreprise pour distinguer plusieurs serveurs éventuels.

Quand une entreprise (organisme, ..) décide de «déposer un nom de domaine», elle s'adresse à un prestataire habilité pour le faire (en France ou aux USA), qui va vérifier si le nom n'existe pas déjà, et ensuite va lui attribuer une **adresse IP** disponible. Le couple *domaine-adresse IP* est ensuite communiqué à des serveurs **DNS**, tous connectés à **Internet**, qui vont dialoguer entre eux pour «répliquer» (si besoin est) cette mise à jour d'annuaire.

#### Exemple :

Un internaute américain (abonné à AOL) veut accéder au serveur d'EDF, dont il connaît uniquement le nom *www.edf.fr*. Son fournisseur d'accès à Internet (dont on parlera plus loin), ici AOL, ignore vraisemblablement l'adresse IP du serveur EDF.

Le serveur DNS d'AOL va interroger le serveur DNS principal du domaine pour savoir s'il connaît l'adresse du site EDF. Si la réponse est négative, le serveur DNS **.com**, par examen du nom, va voir qu'il fait partie du domaine de 1er niveau **.fr**,

Comme tous les serveurs DNS de 1<sup>er</sup> niveau du monde entier se connaissent entre eux, il va interroger le serveur DNS principal du domaine .fr (géré par l'**AFNIC[16]**) lequel a dans ses tables l'adresse du serveur DNS du sous-domaine «*edf*», lequel va lui indiquer l'adresse IP de *www.edf.fr*, et la réponse va remonter jusqu'à l'utilisateur initial.

Si par hasard la chaîne est brisée quelque part (parce qu'un domaine n'existe pas, suite à une faute de frappe, par exemple «*www.def.fr*» au lieu de «*www.edf.fr*»), un message d'erreur va être propagé jusqu'à l'ordinateur de l'utilisateur initial.

Si le nom du serveur de destination a bien été «résolu» (c'est ainsi qu'on désigne l'opération de conversion *Nom-Adresse IP*), la connexion va pouvoir s'établir. Elle n'est jamais effectuée directement (entre le demandeur et le destinataire), mais passe par une série de «serveurs relais». Le chemin (on dit «*route*») est souvent très varié, souvent étonnant, dépendant de l'état du réseau à l'instant considéré.

Par exemple, pour se connecter depuis la région parisienne sud au serveur du Ministère du travail, la connexion est passée successivement par Rennes, Quimper et Brest! De même, pour atteindre le serveur de la CIA (Washington), on est passé par Londres, New-York, Chicago, .... (plus de 20 tronçons)

### **3.4. Les connexions à Internet**

Il existe **plusieurs** façons physiques de se «relier» au «Net» :

#### 3.4.1. La liaison spécialisée

Il s'agit d'un câble à haut débit (cuivre ou fibre optique), connectant une machine ou plus généralement un réseau, à Internet. Ce type de connexion est utilisé par les entreprises ou les fournisseurs d'accès Internet. La connexion est **permanente**. Le débit maximal peut dépasser 30 Mbps[17].

#### 3.4.2. La liaison spécialisée par satellite

Elle est utilisée lorsque d'importants débits sont nécessaires mais elle peu répandue, en raison de son coût. De plus, ce type de liaison ne fonctionne que dans le sens de la réception (Internet vers Satellite vers machine), l'émission (Machine vers Internet) étant réalisée via une deuxième connexion (par modem ou liaison spécialisée). Cette méthode est effectivement **très coûteuse**, car elle nécessite d'utiliser des satellites en orbite basse (quelques centaines de km d'altitude), donc non géostationnaires, donc en nombre important. Le débit est de l'ordre de 1 à 10 Mbps

#### 3.4.3. Le satellite de télévision

Les satellites de télévision (Astra, HotBird,...) sont géostationnaires et donc sont situés à une altitude très élevée (36000 km). Cela a pour conséquence d'introduire un délai non négligeable de transmission du signal (même à 300000 km/s, il faut presque ¼ s pour parcourir l'aller et retour entre la Terre et le satellite. Or ¼ de seconde est une «éternité» à l'échelle informatique, où l'on raisonne souvent en milliardièmes de seconde)

Des expériences on déjà eu lieu avec les bouquets numériques français (*TPS* et *Canal Satellite*). Les fonctionnalités sont assez limitées (téléchargements de fichiers chez *Canal Satellite*, E-mails chez *TPS*), et ce système ne peut pas être utilisé dans les domaines professionnels. Et là aussi, l'utilisation d'un modem (intégré au démodulateur numérique) est nécessaire pour l'émission. Le débit est de l'ordre de 1 Mbps

#### 3.4.4. Le câble

L'utilisation du réseau de télévision par câble pour se connecter à Internet est désormais possible dans certaines régions (urbaines). La connexion est **permanente**. Le débit est de l'ordre de 1 à 10 Mbps. Ce genre de connexion est géré par des compagnies privées.

### 3.4.5. La connexion dite par accès distant (modem et RNIS)

C'est le **mode le plus courant** de connexion chez les particuliers . Pour cela on utilise une ligne téléphonique classique (**RTC**), dédiée en temps normal à acheminer des signaux électriques analogiques (voix). Cela nécessite l'utilisation «modem». La connexion téléphonique est établie avec un **FAI** .Le FAI est relié à Internet via une (ou plusieurs) liaisons spécialisées.

La vitesse maximale ne dépasse pas actuellement 56 kbps].

Il existe un autre type de lignes téléphoniques : Les liaisons **RNIS** , lesquelles véhiculent des informations purement numériques. Il n'y a plus besoin de modem (seulement d'une carte adaptatrice), et le débit est nettement supérieur à celui des lignes RTC (au minimum 64 kbps, garantis, voire 128, 256 ou davantage). La connexion est **quasi immédiate** (2 secondes). Sinon le principe de connexion à Internet est le même qu'avec les liaisons RTC, en faisant appel à un FAI.

NB: **Numéris** est le nom **commercial** donné par France Télécom à son service de connexion RNIS.

### 3.4.6. ADSL(Asymetric digital subscriber line)

C'est un nouveau moyen de connexion, surtout destiné aux particuliers (et petites entreprises), qui commence seulement maintenant à se déployer en France. Comme son nom le suggère, il est asymétrique, c'est à dire que les débits sont différents en réception et en émission. Il est basé sur l'utilisation de lignes téléphoniques spéciales (haut débit) entre centraux téléphoniques, et de lignes traditionnelles chez l'abonné. Le débit est nettement plus important que dans le cas d'une liaison traditionnelle (entre 500 kbps et 1Mbps) . En plus de la vitesse, l'avantage le plus important d'ADSL est la **permanence** de la connexion Internet (sans taxation d'unités téléphoniques) ! Mais l'abonnement mensuel est relativement onéreux (environ 40 euros par mois)

## 4. LES UTILISATIONS

Il y a autant d'utilisations possibles **d'Internet** qu'il existe de sous-protocoles (soit, en théorie, **65536**)

Mais les plus courantes sont en nombre beaucoup plus réduit (moins d'une dizaine).

On se limitera ici à en décrire quatre.

### 4.1. La messagerie (smtp et pop)

C'est l'application **la plus utilisée** et une des plus anciennes. Elle sert à envoyer des **messages** (protocole **SMTP**) et à en recevoir (protocole **POP**). . Par exemple, si l'utilisateur A veut envoyer un message à l'utilisateur B, il va rédiger ce message . Puis il indique l'adresse de courrier électronique de B (appelée communément *adresse E-mail*).

Cette adresse est en **2 parties**, séparée par le caractère **@** (arobas), devenu **symbole de l'Internet** :

la première (à gauche) identifie le «**compte**» du destinataire dans son domaine,

la deuxième (à droite) désigne le **domaine**.

Le domaine est celui de l'entreprise, université,.. dans le cas de liaisons permanentes (p.ex. *edf.fr*), ou celui d'un FAI, dans le cas d'un particulier ou association (p.ex. *wanadoo.fr*)

Ce message est ensuite envoyé vers le serveur **SMTP** dont dépend l'expéditeur (p.ex., le serveur de messagerie sortante du FAI Wanadoo est *smtp.wanadoo.fr*, ...).

Ce serveur joue le rôle d'un bureau de tri postal. Il va identifier le domaine du destinataire, puis transmettre le message au serveur de messagerie de ce domaine.

Le destinataire B va utiliser un logiciel de messagerie analogue à celui de l'expéditeur A (ce n'est pas forcément le même, mais tous les deux utilisent les mêmes règles). Il va interroger son serveur de boîte postale (p.ex. *pop.free.fr*), en s'identifiant par un couple nom/mot de passe (de la même façon que pour récupérer du courrier dans une boîte postale, il faut obligatoirement posséder une clef, fournie par le bureau de poste au départ, ou dans le cas de poste restante présenter une pièce d'identité).

Si des messages existent, ils vont être téléchargés depuis le serveur POP de B vers sa machine, et ils seront affichés dans le logiciel de messagerie de B.

**Au départ**, les messages ne contenaient que du **texte pur**.

Suite à l'amélioration du protocole de messagerie, il est possible désormais d'inclure des **pièces-jointes** de n'importe quel type (documents bureautiques, images, sons, vidéos, ..).

Pour transférer des données volumineuses, il existe un autre protocole, prévu pour cela (**FTP**).

## **4.2. Les forums de discussion (les «News») (nntp)**

La messagerie permet de communiquer individuellement avec d'autres personnes connectées, par contre elle n'est pas adaptée à des forums de discussions. Il existe un protocole pour cela (**NNTP** = *Network News Transfer Protocol*), permettant à un nombre **illimité** de personnes, situées géographiquement dans le **monde entier**, de discuter publiquement entre elles. Les forums sont classés par thème et il en existe plusieurs dizaines de milliers!

Les messages sont stockés/dupliqués sur des serveurs (dit serveurs de News), en principe il en existe un chez tout FAI (celui de Wanadoo s'appelle *news.wanadoo.fr*, celui de FREE est *news.free.fr*,...). Ces serveurs répliquent entre eux en permanence tous les messages qu'ils possèdent.

Les forums ont une **structure hiérarchisée, arborescente**.

La première est la **langue** (au début des News, seul l'anglais était pratiqué, mais avec le développement d'Internet sont apparus des groupes de discussion en autres langues)

Ainsi il existe une hiérarchie «**fr.**»

«**fr.**» signifie ici «**francophone**», et non pas «français». Les groupes «fr» concernent aussi bien les Québécois, Sénégalais, Suisses, Belges, ..., que les Français! Il n'y a pas de place dans ces forums pour les chauvins, nationalistes ou racistes!

Ensuite, on trouve une série de grands **domaines**.

Par exemple «*fr.comp*» est l'ensemble des groupes concernant les «computers» (on a gardé l'appellation américaine afin de reproduire une arborescence analogue).

Puis dans *fr.comp* on trouve «*fr.comp.os*» (concernant les «Operating Systems», = Systèmes d'exploitation), dans lequel, à son tour, contient *fr.comp.os.ms-windows*, *fr.comp.os.msdos*, *fr.comp.os.linux*, ..; et ainsi de suite.

«*fr.soc*» contient les groupes discutant de problèmes de société

«*fr.rec*» concernent les domaines «récréatifs» (cuisine, photo, plongée,..)

Il existe aussi (éventuellement) des groupes «locaux», propres à un serveur de News, et qui ne seront pas répliqués. Ils ont alors un caractère **privé**. Par exemple Wanadoo a ses propres forums (réservés à ses membres, ...)

Les forums sont **très utiles** pour résoudre, en particulier, des problèmes techniques .

## **4.3. Le transfert de fichiers (ftp)**

C'est une des plus anciennes applications **d'Internet**, car utilisée au départ sur des machines **UNIX** pour transférer des fichiers entre elles (dans un réseau local utilisant le protocole **TCP/IP**). (Personnellement, ce fut mon premier contact avec **TCP/IP**, et donc **Internet**, il y a bientôt 10 ans)

Elle est basée sur le protocole **FTP**[26], et consiste à établir une liaison permanente (durant toute la connexion) entre un serveur et un poste client. Le client dispose d'un certain nombre de commandes (normalisées), qui vont lui permettre de recevoir ou d'envoyer un ou plusieurs fichiers, de n'importe quelle nature (texte, logiciels, images, sons, ..) et de n'importe quelle taille.

Il existe un grand nombre de logiciels basés sur ce protocole, certains très rustiques (dans aucune interface, tout s'effectuant par des commandes tapées au clavier), d'autres dotés d'une interface graphique, visuelle, beaucoup plus intuitive.

*Exemple de connexion au serveur FTP de Microsoft à l'aide du logiciel WS\_FTP*

La connexion à ces serveurs nécessite dans tous les cas une **authentification**, par un couple *nom d'utilisateur/mot de passe*. Mais certains serveurs FTP ont un caractère **public** (p.ex. serveur de téléchargement mettant à la disposition du public toute sorte de documents).

Dans ce cas, une convention **universelle** a été établie, dans laquelle le nom d'utilisateur est «**anonymous**», et le mot de passe *l'adresse électronique de l'utilisateur*. On parle alors de serveurs FTP *anonymes*.

Pour un serveur donné, la connexion peut être anonyme ou non. Suivant le cas, l'accès aux fichiers est plus ou moins réduit (généralement les connexions anonymes ne donnent accès qu'aux documents publics).

#### 4.4. Le WEB, ou «World Wide Web» (http)

**Web** (en anglais) = **Toile** (d'araignée)

C'est l'application qui a **popularisé** Internet par excellence, à un point tel que souvent (dans les médias!) ce terme est **confondu** avec **Internet**, mais ce n'est qu'une partie.

Elle date de **1989**, développée à l'époque pour les besoins internes du **CERN** (Centre Européen de Recherche Nucléaire). Son succès est dû à sa simplicité d'utilisation (apparente parfois!), son ergonomie intuitive (navigation par clic de souris), l'utilisation abondante du multimédia (animation graphique, sons, vidéos, ...)

##### *4.4.1. Principe*

Par **WEB** on désigne un ensemble de **serveurs** mettant à disposition différents documents (textes, images, sons,...) via le protocole **HTTP**.

Côté **client** (=utilisateur), les logiciels qui permettent de se connecter à ce type de serveurs, avec le même protocole HTTP, s'appellent des **navigateurs web** (nos cousins Québécois disent joliment «*butineurs*»). Les navigateurs les plus connus et utilisés sont **Communicator** (ou *Navigator*), de **Nestcape** (racheté par **AOL**), et **Internet Explorer**, de **Microsoft**. Ces deux outils sont **gratuits**.

Il en existe au moins une dizaine d'autres, dont **Lynx**, (texte uniquement), **Opera** (très léger, gratuit maintenant, mais avec bandeau publicitaire), ...

Les documents transmis par les serveurs Web et affichés par les navigateurs affichés sont composés de **texte brut**, auquel on a ajouté différentes **balises**, ou marqueurs (qui ne seront pas affichés à l'écran), dont le but est de **mettre en forme** le texte.

Par exemple, un mot ou une phrase encadrée respectivement par les balises **<B>** et **</B>** apparaîtra en gras (*B*= initiale de «**bold**»).

Le couple de balises **<CENTER>** et **</CENTER>** sert à centrer du texte à l'écran. On peut aussi définir la taille, couleur, police des caractères, introduire des sauts de paragraphes, retraits, listes numérotées, ...

Toutes ces règles sont parfaitement codifiées au niveau **mondial**, définies par le consortium **W3C** et constituent un langage commun, appelé **HTML**.

Mais **le plus important**, en plus de la mise en forme du texte, est la possibilité d'insérer des **liens** vers d'autres documents (situés sur le même serveur ou sur des serveurs différents), l'opération de branchement vers ces liens étant réalisée par un simple clic de souris. Ces liens sont appelés «**liens hypertexte**».

Vu que les documents ainsi adressés peuvent à leur tour comporter d'autres liens, l'ensemble de tous les documents constitue une véritable **toile**, d'où le nom donné à cette application.

les adresses **DNS** des serveurs web commencent souvent par le préfixe «**www**» (World Wide Web), mais ce n'est absolument pas une obligation. Il permet seulement, de façon simple et mnémotechnique, de distinguer plusieurs serveurs dans un même domaine. (par exemple, chez Microsoft, le site principal s'appelle [www.microsoft.com](http://www.microsoft.com), et le serveur dédié aux développeurs s'appelle [msdn.microsoft.com](http://msdn.microsoft.com))

##### *4.4.2. L'adressage des documents*

L'adressage d'un document défini soit dans un navigateur, soit sous forme de lien dans un autre document HTML, est réalisé en respectant une certaine syntaxe, elle aussi universelle : l'**URL**

Un URL comporte (au maximum) 5 parties successives :

le **protocole** utilisé (généralement **http**, mais ce n'est pas obligatoire, les navigateurs http sachant généralement traiter d'autres protocoles Internet, tels que **FTP**, pour les transferts de fichiers - cf. plus loin), suivi des caractères **«//:»**. La plupart des navigateurs tolèrent l'omission de cette chaîne, supposant que le protocole par défaut est **http**.

le **nom** (DNS) du serveur (p.ex. **www.bellamyjc.net**), suivi éventuellement du caractère **«/»**

le **n° de port** du serveur, précédé du caractère **«:»**. C'est un nombre qui indique en quelque sorte le «n° de canal» sur lequel le serveur «émet». Par défaut, ce port a le n° **80**, et dans ce cas il est **facultatif** de le préciser. Mais il peut parfois y avoir plusieurs serveurs Web sur la même machine. (p.ex. le serveur Web normal, et un serveur d'administration, d'usage réservé). Dans ce cas, on distinguera chaque serveur en indiquant le n° de port (80, 81, 7922, 8080, ..)

le **répertoire** (éventuel) du document sur le serveur, suivi du caractère **«/»** (généralement, vu le grand nombre de documents que peut héberger un serveur Web, ces derniers ne sont pas mis en vrac à un seul endroit, mais classés suivant une structure arborescente) (p.ex. **/fr/**)

le **nom du fichier** contenant le document lui-même (p.ex. **internetstory.html**). Ce nom peut éventuellement être omis, dans le cas où il s'agit du document par défaut du serveur (généralement, il s'appelle **index.html** ou **default.html**).

En reprenant l'exemple précédent, l'URL complet sera donc : <http://www.bellamyjc.net:80/fr/internetstory.html> ou encore : [bellamyjc.net/fr/internetstory.html](http://bellamyjc.net/fr/internetstory.html)