



Internet

UN RÉSEAU UNIVERSEL

Le développement des moyens de communication tout au long du siècle passé, ainsi que l'avènement des technologies de traitement de l'information ont abouti à un bouleversement technologique que certains ont pu comparer à la révolution industrielle du ^{XX} siècle : les domaines des télécommunications et de l'informatique ont convergé pour permettre le développement d'un réseau mondial d'ordinateurs, non centralisé et non contrôlé : Internet. Le réseau Internet présente une particularité : ses protocoles sont normalisés, mais aucun organisme de certification n'existe. Ils sont en fait produits, de manière indépendante mais coordonnée, par l'ensemble des programmeurs s'intéressant au fonctionnement du réseau. De même, aucune autorité centrale de régulation du trafic ou du contenu n'existe : l'architecture du réseau, conçu au départ pour des applications militaires, ne permettrait pas à une telle autorité de se mettre en place.

LES ORIGINES D'INTERNET

Le développement des ordinateurs date des années 1950 et on a très tôt cherché à les connecter entre eux. Les premières liaisons directes étaient souvent transitoires, servant à une transmission ponctuelle de données. La plupart du temps, cela n'était possible qu'entre ordinateurs de même marque, voire de même modèle, ou demandait une « traduction » (passage d'un système de codage à un autre) partiellement automatisable. Le premier réseau non militaire d'ordinateurs fut mis en place dans les années 1960 par la firme IBM pour le système de réservation de la compagnie d'aviation American Airlines. Tous les ordinateurs des guichets de vente de billets étaient reliés à l'ordinateur central de la compagnie, permettant à celui-ci de traiter « en temps réel » (c'est-à-dire instantanément) les réservations pour indiquer de manière fiable les places encore disponibles.

LE RÉSEAU ARPANET

En 1964, au cœur de la guerre froide, l'ARPA, organisme de recherche de l'armée américaine, réfléchit aux caractéristiques techniques d'un réseau informatique pouvant rester fonctionnel même après une attaque nucléaire. Un tel réseau ne doit pas avoir de « centre », dont la destruction signifierait celle du réseau lui-même : il doit pouvoir fonctionner entre les parties subsistantes même après des

destructions partielles importantes, et ne pas bloquer les informations en cas de saturation des voies de communications. L'Arpanet est créé en 1969 ; il ne rassemble au début que quatre ordinateurs, mais les scientifiques des universités américaines s'y accordent bientôt, et deviennent de plus en plus nombreux. Le réseau est constitué d'un ensemble de « nœuds » (c'est-à-dire de points de jonction des lignes de communication) indépendants, chacun capable de recevoir des données depuis tout le réseau et de les réexpédier vers le nœud suivant le plus proche de sa destination, ou celui dont la voie est la moins encombrée. En cas de panne d'un élément du réseau, les données peuvent ainsi transiter par d'autres chemins pour être délivrées à leur destinataire.

LA NAISSANCE D'INTERNET

Jusqu'en 1977, le réseau Arpanet ne connectait que quelques dizaines de nœuds aux États-Unis. La véritable naissance d'Internet se date de la création du « protocole » (langage commun de transmission de données) TCP/IP (pour *Transfert Control Protocol*, protocole de régulation des transmissions, et *Internet Protocol*, protocole d'Internet) par Vinton Cerf, Robert Kahn, David Clark et Louis Pouzin. En 1983, l'ARPA accorde à l'université de Berkeley des subsides importants pour le développement d'une version d'Unix (un système d'exploitation très utilisé par les informaticiens) intégrant ce protocole. Dès lors, les réseaux locaux existants, principalement ceux des organismes de recherche, vont pouvoir communiquer directement. Au fur et à mesure que des chercheurs s'y connectent, d'énormes banques de données documentaires deviennent accessibles à tous.

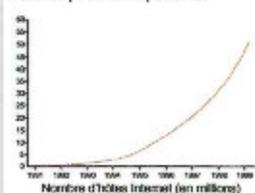
LE DÉVELOPPEMENT D'INTERNET

TCP/IP, plus qu'un protocole, est une suite intégrée de protocoles conçus pour différents usages : transfert de fichiers, courrier électronique,...



En 1990, l'invention par **Tim Berners-Lee**, au Centre européen de recherches nucléaires à Genève, du protocole HTTP (*Hyper Text Transfert Protocol*, protocole de transmission de fichiers hypertextes) permet l'affichage de fichiers hypertextes (c'est-à-dire comportant,

en plus du texte, un certain nombre de liens vers d'autres fichiers textes, images, sons...). C'est la naissance de la Toile mondiale (*World Wide Web* ou *Web*), permettant à tout utilisateur connecté de consulter les pages créées et mises en ligne (c'est-à-dire rendues disponibles sur le réseau) par toutes sortes d'organismes et de particuliers - un peu comme le faisait le Minitel en France, mais avec des structures et surtout un coût beaucoup moins importants.



La croissance d'Internet est alors exponentielle : des quelques dizaines de nœuds d'Arpanet - limités aux États-Unis - en 1970, le réseau passe à 30 000 nœuds dans une dizaine de pays en 1988 pour atteindre 6 millions en 1995, 30 millions en 1998 et 60 millions de nœuds sur toute la planète à l'aube du ^{XX} siècle.

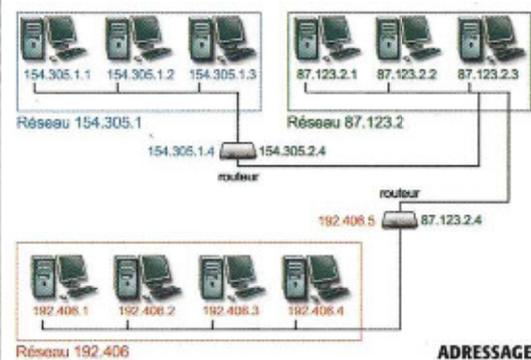
FONCTIONNEMENT ET ADRESSAGE

LES CANAUX DE TRANSMISSION
D'un point de vue physique, Internet utilise tous les modes de transmission disponibles puisqu'il relie des réseaux de natures diverses. Les serveurs de réseaux locaux, les réseaux d'entreprises utilisent des lignes à haut débit comme les systèmes ADSL, DNS, etc., tandis que les internautes individuels se contentent encore souvent d'une liaison par ligne téléphonique normale. Le développement de la transmission de fichiers volumineux « en temps réel »

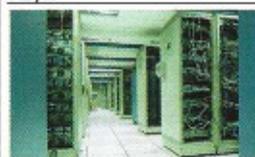


(communications vidéo, par exemple) pousse cependant de plus en plus d'utilisateurs individuels vers les connexions haut débit.

FRACTIONNEMENT



TCP/IP



Les protocoles de la suite TCP/IP permettent aux **serveurs** (les machines où sont stockées les informations disponibles sur le réseau) et aux routeurs (les machines « relais » d'Internet, qui sont les intermédiaires entre émetteurs et destinataires) l'accomplissement de quatre tâches : fractionnement, adressage, acheminement et contrôle des erreurs.

• **Fractionnement** : le fractionnement des messages en « paquets » permet, pour un même message, d'utiliser plusieurs chemins entre l'émetteur et le destinataire. Au départ, les données à transmettre sont découpées en fragments numérotés (appelés paquets), puis la machine réceptrice les reconstitue selon les instructions contenues dans l'en-tête (données ajoutées lors de l'application du protocole, permettant la reconstitution et la vérification).

• **Adressage** : le protocole TCP/IP attribue à chaque machine connectée un identifiant numérique, dit « adresse IP », composé de quatre nombres compris entre 0 et 255. Cette adresse est hiérarchisée : les deux ou trois premiers nombres désignent un « sous-réseau », un réseau indépendant connecté, ou une boucle locale d'Internet, tandis que le dernier indique précisément la machine à l'intérieur du réseau. Les routeurs sont munis de plusieurs

cartes réseau et reçoivent une adresse IP par carte ; ils ont ainsi une adresse IP dans chacun des sous-réseaux auxquels ils sont connectés.

• **Acheminement** : le premier chiffre de l'adresse indique à chaque intermédiaire le réseau sur lequel se trouve le destinataire ; si l'intermédiaire est directement connecté à ce réseau, il lui délivre directement les données. Dans le cas le plus fréquent où le réseau du destinataire est relié par l'intermédiaire d'autres machines, il consulte une « table d'adressage » lui indiquant à quel intermédiaire transmettre les données en fonction du destinataire.

• **Contrôle des erreurs** : Le protocole intègre une procédure de vérification par somme de contrôle (*checksum*). Après l'encodage, l'émetteur additionne toutes les données numériques de l'en-tête et communique le résultat à la suite de celui-ci. Le récepteur vérifie ensuite que la somme des nombres qu'il reçoit est égale à celle calculée par l'émetteur. Dans le cas contraire, une erreur est survenue. L'adresse IP n'est pas nécessairement fixe : si les routeurs et les serveurs connectés en permanence se doivent de conserver leur adresse pour être repérables, les terminaux clients, qui la plupart du temps ne se connectent qu'épisodiquement, reçoivent de leur fournisseur d'accès une adresse IP nouvelle à chaque connexion.

L'ADRESSE IP ET LE PORT

De nombreux programmes TCP/IP peuvent être exécutés simultanément sur Internet (il est possible par exemple d'ouvrir plusieurs navigateurs simultanément, de naviguer sur des pages HTML tout en téléchargeant un fichier). L'ordinateur doit alors pouvoir distinguer les différentes sources de données. Pour cela, chacune de ces applications se voit attribuer une adresse unique sur la machine, codée sur 16 bits : un port (la

Internet en Europe

Part de la population connectée au réseau mondial

Belgique



44 %

Royaume-Uni



43 %

Italie



38 %

Allemagne



30 %

France



29 %

Espagne



29 %

Évolution du Nasdaq, indice boursier des sociétés de haute technologie (valeur 100 en 1971)
2 200 points en mars 1999
5 000 points en mars 2000



Éclatement de la bulle : le 4 avril 2000, le Nasdaq perd 13,5 % de sa valeur en quelques minutes.

combinaison adresse IP + port est alors une adresse unique au monde, elle est appelée *socket*.

De cette manière, lorsque l'ordinateur reçoit des informations destinées à un port, les données sont envoyées vers l'application correspondante. S'il s'agit d'une requête à destination de l'application, l'application est appelée « application serveur ». S'il s'agit d'une réponse, on parle alors d'« application cliente ».

Le processus qui consiste à pouvoir faire transiter sur une connexion des informations provenant de diverses applications s'appelle le « multiplexage ». De la même façon, le fait d'arriver à mettre en parallèle (donc répartir sur les diverses applications) le flux de données s'appelle le « démultiplexage ».

LE SYSTÈME DNS ET LES URL

Le système d'adressage du protocole TCP/IP n'est guère convivial, même s'il convient parfaitement au repérage par les programmes. Pour rendre les adresses plus faciles à mémoriser et à utiliser, notamment pour les applications grand public, une correspondance est établie entre les adresses IP et des adresses alphabétiques nommées URL (*Universal Resource Locator*, localisateur universel des données). Ce système est appelé DNS pour *Domain Name System* (système de noms de domaines). Chaque adresse URL est composée, dans l'ordre de lecture :

- d'un terme indiquant le protocole utilisé (*http* : pour le Web, *mailto* : pour le courrier électronique, *ftp* : pour la transmission de fichiers...);
- éventuellement, de l'identifiant et du mot de passe de l'utilisateur (option déconseillée car le mot de passe est alors lisible à l'écran);
- du nom du serveur et de son nom de domaine proprement dit (le fameux *.com* ou *.fr*);
- enfin, le cas échéant, du chemin d'accès au fichier désiré (les sous-dossiers ou sous-catalogues sont séparés par des barres obliques). Les adresses de sites Web et de messagerie électronique obéissent à une syntaxe légèrement différente, car l'adresse électronique spécifie un destinataire particulier parmi les machines connectées au serveur.

LES PROTOCOLES TCP/IP

La suite de protocoles TCP/IP comprend plusieurs types de protocoles : d'une part les protocoles de base (PPP, Telnet...) et d'autre part les protocoles spécifiques (FTP, HTTP, ...).

• Au-dessus des protocoles de base se trouvent les protocoles correspondant à des utilisations spécifiques : transfert de fichier (protocole FTP), courrier électronique (protocoles SMTP et POP), Toile mondiale (protocole HTTP) sont les plus utilisés.

L'ACCÈS AU RÉSEAU : PROTOCOLE PPP

Beaucoup d'ordinateurs individuels sont reliés à leur fournisseur d'accès par une ligne téléphonique et un *modem* ; on parle de liaison « point

à point », c'est-à-dire de liaison simple entre deux machines, sans partage de la bande passante. Cette liaison s'effectue selon un protocole particulier, nommé PPP (pour *Point to Point Protocol*).

TELNET

Telnet est le protocole de base d'Internet, qui permet l'« interfaçage » de terminaux (un terminal est un ordinateur qui se contente d'afficher les informations qu'il reçoit ; c'est la plus simple expression du « client ») et de serveurs possédant des configurations différentes.

La transmission de données par Telnet est très fruste : les données, non codées, sont transmises dans l'ordre où elles se présentent, sans réorganisation d'aucune sorte. Ce protocole sert de base à des protocoles de niveau supérieur, dédiés à des utilisations particulières d'Internet.

FILE TRANSFER PROTOCOL

La plus ancienne utilisation du réseau Internet est la transmission de fichiers d'un ordinateur à un autre. Le MIT (*Massachusetts Institute of Technology*, Institut de technologie du Massachusetts) mit au point le protocole FTP (*File Transfer Protocol*, protocole de transfert de fichier) en 1971. Depuis 1973, aucune modification importante n'a été apportée à ce protocole.

Aujourd'hui, malgré l'explosion du courrier électronique et de la Toile mondiale, le protocole FTP reste un des plus utilisés aussi bien par les usagers professionnels (transferts de fichiers d'une entreprise à une autre, par exemple) que par le grand public (transmission de musique, de vidéos...). La transmission de données par FTP nécessite que les deux machines demeurent connectées durant tout le temps de la transmission. Des programmes permettent cependant, en cas d'interruption de la communication, de reprendre le chargement à l'endroit où il avait été interrompu. Le système de découpage par paquets caractéristique de l'Internet permet même de charger un même fichier depuis des serveurs différents, autorisant ainsi l'utilisation préférentielle des canaux les moins encombrés.

COURRIER ÉLECTRONIQUE

C'est l'application d'Internet la plus utilisée, tant pour des usages professionnels que personnels et ludiques. Elle permet d'envoyer des messages écrits (codés en « texte pur », mais des programmes de messagerie de plus en plus nombreux supportent du texte enrichi, voire de l'hypertexte) et d'y joindre des fichiers de tout type. Chaque serveur Internet, identifié par son URL (*masociete.com*), dirige les messages vers le destinataire à l'intérieur de la boucle locale, identifié par son nom d'utilisateur, lequel est séparé de l'adresse du serveur de messagerie par l'arobe, ou arrobe, @ (nommé aussi « arabesque », à cause d'une confusion avec une ancienne unité espagnole de capacité). L'adresse électronique se présente donc ainsi : *jean.dupont@masociete.com* (les adresses électroniques n'acceptent pas les accents sur les voyelles).

Le caractère @, énoncé *at*, c'est-à-dire

« à », chez les Anglo-Saxons, est souvent énoncé « chez » en Français, car un simple « à » risquerait d'être parfois mal interprété, par exemple lorsqu'on dicte une adresse au téléphone.

La transmission du courrier électronique utilise plusieurs protocoles selon les cas :

- Le protocole SMTP (pour *Simple Mail Transfer Protocol*, protocole simple de transmission de courrier) est un protocole de transmission de courrier point à point, c'est-à-dire entre deux machines connectées au même moment. Pour l'utilisateur individuel, qui souvent n'est qu'épisodiquement connecté, c'est le protocole d'envoi du courrier.
- Le protocole POP (pour *Post Office Protocol*, protocole de bureau de poste) permet de recevoir du courrier qui a été envoyé alors que le destinataire n'était pas connecté, et qui est stocké en attendant sur le serveur du fournisseur d'accès. Le protocole POP gère l'identification du destinataire et bloque la boîte aux lettres pendant la consultation, empêchant ainsi une deuxième machine de relever parallèlement le courrier.
- Un troisième protocole, IMAP (pour *Internet Mail Access Protocol*, protocole d'accès au courrier Internet), est une alternative à POP qui permet la consultation simultanée de la même boîte par plusieurs machines, la gestion de plusieurs boîtes aux lettres, etc.

LA TOILE MONDIALE (WORLD WIDE WEB)

Si le courrier électronique a fait le succès d'Internet, c'est assurément la Toile qui lui a donné l'essentiel de son potentiel commercial. La Toile est constituée de documents HTML (c'est-à-dire en hypertexte), soit de pages simples, soit de « sites », c'est-à-dire d'ensembles de pages interconnectées, plus ou moins fréquemment remises à jour.



Les premiers sites Internet n'étaient que de simples vitrines, présentant les réalisations de l'auteur. Or, l'augmentation des débits, qui permet d'enrichir le contenu du site en animations, et l'amélioration de la sécurité des transmissions, autorisant des transferts de liquidités, ont transformé la conception des sites. Les concepteurs disposent ainsi de véritables galeries virtuelles, et les sociétés commerciales de véritables boutiques. La SNCF, par exemple, qui offrait de réserver son billet en ligne, propose à présent de l'imprimer chez soi : le site Internet est réellement devenu l'équivalent d'un guichet... Le fonctionnement de la Toile repose sur le protocole HTTP. Ce protocole fonctionne en deux temps :

- envoi par le client d'une « requête HTTP » ;
- envoi en retour par le serveur d'une « réponse HTTP ».

La requête est la demande effectuée par le client, le plus souvent d'afficher une page précise ; elle peut être accompagnée de données (« l'en-tête »

de la requête), par exemple lors du remplissage de formulaires, de bons de commandes, etc.

Le contenu de la réponse est affiché sur le terminal client par un programme adéquat, nommé navigateur, selon les spécifications de mise en page contenues dans le codage de l'hypertexte (langage de description de page HTML ou XML). D'un simple clic, l'utilisateur peut passer d'une page à l'autre, sur un même site ou sur des sites différents. Cette facilité d'utilisation, le développement de moteurs de recherche rapides et performants (généralistes ou spécialisés), l'absence totale de compétences informatiques nécessaires à la navigation, en ont fait le vecteur principal de l'expansion d'Internet.

LA DISCUSSION EN LIGNE (CHAT)

Le protocole IRC (*Internet Relay Chat*) permet aux internautes de discuter entre eux « en temps réel ». Il existe plusieurs réseaux de discussion et, pour communiquer, deux internautes doivent impérativement être connectés sur le même réseau. Chacun d'entre eux propose plusieurs « canaux » ou « salles » de discussion (*Chat room*), consacrés à des sujets précis, ou à des regroupements de toute nature.

UTILISATIONS

Le développement de la Toile mondiale et l'accroissement des possibilités de téléchargement ont multiplié les utilisations possibles d'Internet. Les activités ludiques ou culturelles, le travail intellectuel, le commerce de masse disposent à présent d'un outil universel et interactif.

LES UTILISATIONS LUDIQUES ET CULTURELLES



L'interactivité et la rapidité d'Internet permettent à des *joueurs* distants de jouer plusieurs milliers de kilomètres de jouer ensemble et d'organiser des tournois. Outre les jeux de société ou de plateau classiques (échecs, « wargames »...) et les jeux électroniques conçus au départ pour une utilisation solitaire, sont apparus sur le modèle des jeux de rôles des jeux interactifs, où tous les joueurs (qui peuvent être des milliers en réseau) participent à l'environnement et à l'action. Les qualités graphiques et sonores et les nouveaux algorithmes de compression (divX pour la vidéo, MP3 pour la musique...) permettent aujourd'hui des téléchargements massifs d'œuvres musicales ou audiovisuelles, à partir d'un centre serveur (comme pour le système Napster) ou en « peer to peer » (c'est-à-dire que les utilisateurs mettent leurs ressources en commun, chacun téléchargeant ce qu'il désire à partir des autres machines connectées et laissant les autres télécharger ce qu'ils désirent depuis son propre ordinateur). Ces téléchargements s'effectuent en général de manière « sauvage », c'est-à-

dire en dehors des règles du droit d'auteur et du copyright. Ce piratage à grande échelle est combattu par les industries du disque et de l'audiovisuel, mais son incidence sur leur chiffre d'affaires est sans doute à relativiser (les internautes piratant beaucoup d'œuvres qu'ils n'achèteraient pas s'ils devaient les payer).

LE TÉLÉTRAVAIL

Dans les professions intellectuelles et administratives, le travail s'effectue de nos jours presque entièrement sur ordinateur. L'utilisation d'Internet permet à une entreprise d'envoyer à ses partenaires les fichiers en un simple clic, réduisant les délais et les frais de poste ou de courrier.



Cela permet également à des *travailleurs indépendants* et même à des salariés de travailler chez eux, dans les conditions qui leur conviennent, et d'envoyer à leur client ou employeur le résultat de leur travail.

Le commerce électronique



Les systèmes de codage d'envoi des données permettent des transmissions confidentielles avec des risques minimes d'interception par des tiers ; cette possibilité est mise à profit par les entreprises pour le télépaiement (par carte bancaire) qui permet aux internautes de commander biens et services en ligne. Internet est en outre un irremplaçable outil de publicité et de marketing. Offre de services gratuits, échange de liens entre deux sites accroissant la fréquentation de chacun d'eux, courrier électronique publicitaire... permettent à une entreprise de toucher à moindre coût un très large public. À la fin des années 1990, la mise en place de ces outils s'est faite rapidement et l'effet de mode a poussé certaines entreprises à beaucoup investir dans leur présence sur Internet. Le nombre de sociétés spécialisées



(phénomène des « *start-up* ») et l'importance des capitaux investis a explosé. Toutefois, si les possibilités d'Internet en matière de marketing sont patentes, il s'en faut de beaucoup que le commerce électronique rapporte autant que l'on avait cru (les internautes étant très attachés à la gratuité). L'adaptation de l'investissement à cette réalité a fait dégonfler la « bulle Internet », et de nombreuses sociétés de ce secteur ont dû fermer leurs portes.