

PARTIE PRÉLIMINAIRE – LE CONTEXTE DES TECHNOLOGIES DE L'INFORMATION

PAR Pierre-Paul LEMYRE

Introduction

L'expression « technologies de l'information » est un concept qui évolue en fonction du développement des innovations et des transformations sociales qui en découlent. Elle désigne en principe l'ensemble des matériels, logiciels et services qui permettent la collecte, le traitement et la transmission de l'information. À ce titre, elle englobe sans distinction l'électronique, l'informatique et les communications. L'émergence de la réseautique au milieu des années 1980 a cependant transformé l'expression. Graduellement, les technologies de l'information désignèrent plus particulièrement les nouveaux moyens de communication informatique tel que l'Échange de Documents Informatisés (EDI) et les réseaux privés. Le concept de « nouvelles technologies de l'information » fit alors son apparition mais fut rapidement associé, au cours des cinq dernières années, à un réseau informatique d'envergure mondiale, Internet.

Internet a initialement été développé à la fin des années 1960, dans le cadre du projet militaire américain ARPANet. L'objectif de ce projet était de concevoir un réseau de communication non-centralisé capable de résister à une destruction partielle. Les chercheurs universitaires chargés de la tâche y arrivèrent en développant la technique de l'acheminement de l'information par paquets (*packet switching*). Cette technique a comme principal avantage de permettre aux paquets d'information d'atteindre leurs destinations de façon autonome. Ainsi, les communications peuvent être maintenues même si une partie du réseau est inutilisable en raison d'une attaque militaire ou d'un dysfonctionnement. En effet, les paquets parviennent à leur destination en empruntant les voies de communication demeurées intactes. Grâce à cette réussite, d'autres réseaux fondés sur l'acheminement par paquets furent créés, tel que PRNet et SATNet. Internet prit réellement forme lorsque les militaires décidèrent d'interconnecter ces différents réseaux. La standardisation d'un protocole de communication, TCP/IP, et la mise en place d'une norme d'adressage unique permit d'y arriver. Une fois cette étape franchie, Internet ne cessa de gagner en popularité. Le réseau pris d'abord de l'expansion dans le milieu de l'éducation, puis auprès des entreprises suite à sa privatisation. Finalement, l'incorporation de l'ensemble des réseaux privés et la mise en place du Web confirmèrent la suprématie d'Internet.

Tel qu'il existe aujourd'hui, Internet est un système mondial et décentralisé de réseaux informatiques : un réseau de réseaux. Son architecture appartient à une multitude d'entreprises privées oeuvrant dans les domaines de la téléphonie, de la câblodistribution, des satellites et de l'informatique. Pour cette raison, il n'existe aucun organisme en charge

d'Internet. Bien que certaines organisations se partagent son développement technique, aucune n'est en mesure d'exercer un contrôle central.

Les technologies qui caractérisent Internet peuvent néanmoins être structurées en trois grandes catégories. D'abord, un certain nombre d'entre elles sont indissociables du réseau, il s'agit de ses éléments constitutifs. Ensuite, l'invention du Web a permis le développement de nombreux outils qui sont à l'origine de l'incroyable popularité d'Internet. Finalement, plusieurs autres technologies complémentaires contribuent à accroître ses capacités et son efficacité. Le présent chapitre expose successivement ces trois catégories.

Les éléments constitutifs du réseau Internet

Contrairement aux réseaux téléphoniques et de câblodistribution qui reposent sur une structure matérielle spécifique, Internet n'est pas un réseau physique. Il s'agit plutôt d'un réseau logique fondé sur des normes d'échange et d'adressage. Internet est donc composé de diverses techniques permettant la transmission d'information, et cela indépendamment du support physique utilisé. Ainsi, l'architecture client/serveur permet de répartir efficacement l'information dans le contexte décentralisé d'Internet. De leur côté, les protocoles TCP/IP établissent la forme des communications et les adresses IP permettent l'identification des ordinateurs sur le réseau. D'autres éléments constitutifs, comme les noms de domaine et le système DNS, facilitent l'échange d'information en substituant des adresses mnémoniques aux adresses IP.

► L'architecture client/serveur

L'architecture client/serveur est l'une des idées maîtresses de la réseautique moderne. Elle désigne le type de relation qui existe entre deux ordinateurs distants qui communiquent ensemble. De façon générale, il s'agit d'offrir des services centralisés sur un serveur et de les rendre accessibles aux usagers à partir de leur propre ordinateur grâce à l'entremise de logiciels clients. Selon ce mode de fonctionnement, le logiciel client effectue une requête de service en utilisant un protocole prédéterminé et le serveur la remplit. Une fois la réponse obtenue, le logiciel client a également pour tâche de présenter le résultat.

Cette façon de concevoir les communications informatiques a largement remplacé les unités centrales qui étaient en place jusqu'à la fin des années 1980. Selon ce modèle, toutes les opérations étaient effectuées par un ordinateur central auquel les usagers accédaient par le biais de terminaux. Comme ces terminaux ne faisaient qu'afficher l'information fournie par l'ordinateur central, cela impliquait d'importants transferts de données sur le réseau. Ceci empêchait, entre autres, l'utilisation d'interfaces graphiques. L'architecture client/serveur a résolu ce problème en permettant un partage efficace des ressources entre les deux parties à la communication.

Habituellement, le serveur, aussi appelé démon, est activé en permanence et attend les requêtes des clients. Le serveur étant un logiciel, il est possible qu'un ordinateur opère plusieurs serveurs à la fois, afin d'offrir différents services. Dans le même ordre d'idées, un logiciel client peut utiliser plusieurs protocoles et, par conséquent, interagir avec plusieurs types de serveurs. De plus, rien n'interdit d'installer des logiciels clients et serveurs sur le même ordinateur. Par exemple, il est possible qu'un ordinateur personnel

servant principalement de poste de travail et possédant de multiples logiciels clients héberge un serveur FTP (File Transfer Protocol).

► Les protocoles TCP/IP

Un protocole est un ensemble de règles qui décrivent, en termes techniques, la façon d'exécuter une action. Dans le contexte particulier d'Internet, les protocoles servent à définir les méthodes que doivent utiliser les deux parties à une communication lorsqu'elles s'échangent des données. Il s'agit de standards internationaux qui doivent être respectés par l'ensemble de la collectivité. Leur principal objectif est d'assurer l'uniformité des communications, ce qui permet l'interopérabilité des systèmes. TCP/IP (*Transmission Control Protocol / Internet Protocol*) est le protocole de base d'Internet sur lequel repose tous les autres. Il prend en charge l'organisation et le transport des données sur le réseau.

Tout d'abord, TCP s'occupe de l'organisation en morcelant l'information en paquets. Un paquet comprend une petite partie des données transférées (environ 500 caractères), de l'information quant à la séquence, son adresse d'origine et l'adresse de sa destination. Une fois les paquets construits, ils sont acheminés au module IP qui se charge du transport. TCP a également pour tâche de reconstituer l'information sur l'ordinateur du destinataire en remplaçant les paquets en ordre. Enfin, si certains paquets n'ont pas été reçus, TCP se charge de réclamer un nouvel envoi.

IP, pour sa part, permet le transport des paquets d'un ordinateur à un autre grâce à un système d'adressage. En effet, chaque ordinateur connecté à Internet possède sa propre adresse IP. Lorsqu'un paquet est envoyé sur le réseau, il transite par une passerelle (*gateway*) qui l'achemine un peu plus loin vers sa destination. Ce processus se répète jusqu'à ce qu'une passerelle reconnaisse le paquet comme étant destiné à un ordinateur lui étant directement relié. Toutefois, chaque paquet circule de façon indépendante. Aussi, différentes routes peuvent être empruntées, selon l'achalandage des voies de communication.

Il existe aujourd'hui des centaines de protocoles compatibles avec TCP/IP. Ceux-ci constituent les différents services disponibles sur Internet. Les principaux sont le HTTP (*HyperText Transfer Protocol*) qui constitue le Web, le SMTP (*Simple Mail Transfer Protocol*) qui permet l'échange de courriers électroniques, le FTP (*File Transfer Protocol*) qui est utilisé pour l'échange de fichiers et le NNTP (*Network News Transfer Protocol*) qui permet l'échange d'articles USENET.

La vaste majorité de ces protocoles sont ouverts, ce qui signifie qu'ils sont définis par des documents publics et que toute personne peut les utiliser pour produire des logiciels compatibles. Ceux-ci sont principalement élaborés par des comités de travail, au sein d'organisations chargées du développement technique d'Internet, telle que l'IETF (*Internet Engineering Task Force*) (<http://www.ietf.org>). Ces comités regroupent généralement les principaux intervenants de l'industrie et assurent l'évolution du réseau. Il peut aussi arriver qu'une entreprise divulgue un protocole qu'elle a elle-même développé, estimant être en mesure de bénéficier de sa mise en œuvre. Par exemple, c'est ce qu'a fait la société Sun avec la technologie Java (<http://java.sun.com/docs>). Malgré tout, un certain nombre de protocoles restent fermés, ce qui implique qu'ils ne peuvent être utilisés que par leurs propriétaires. Par exemple, il est impossible de développer des

logiciels compatibles avec le protocole du réseau Napster puisque cette technologie est leur propriété (<http://www.napster.com>).

► Les adresses IP

Tout ordinateur connecté à Internet possède une adresse IP qui lui est unique. Il s'agit de son identifiant sur le réseau. Celle-ci consiste en un numéro de 32 bits, soit quatre octets séparés par des points. Bien que l'adresse soit présentée sous la forme décimale, la valeur de chaque segment ne dépasse jamais celle d'un octet, soit 256. Par exemple, 132.204.136.21 (<http://132.204.136.21>) est l'adresse IP du serveur Web principal du LexUM (<http://www.lexum.umontreal.ca>). Ce sont ces adresses que les ordinateurs utilisent lorsqu'ils communiquent entre eux.

La structure de ces adresses est hiérarchique. Ainsi, le premier octet désigne une plage d'adresse plus importante que ceux qui suivent. Dans cette optique, chaque segment désigne une classe d'adresse. Le premier (132..xxx.xxx.xxx) désigne une classe A, réservé aux réseaux importants. Le second (132.204.xxx.xxx) désigne une classe B, attribué aux réseaux de taille moyenne. Le troisième (132.204.136.xxx) désigne une classe C, qui convient aux petits réseaux nécessitant moins de 256 adresses. Finalement, le quatrième (132.204.136.21) désigne une classe D, c'est à dire l'adresse d'un ordinateur particulier. Une fois que l'une de ces plages est obtenue, chaque organisation gère elle-même l'attribution des adresses qu'elle contrôle.

Toutefois, tous les ordinateurs ne possèdent pas une adresse permanente, attribuée de façon statique. Plusieurs fournisseurs d'accès Internet (FAI), pour économiser sur le nombre d'adresses nécessaires, partagent un nombre limité d'adresses entre de nombreux usagers. Ceci se fait en attribuant dynamiquement une adresse différente à chaque connexion. Puisque les usagers ne se connectent pas tous au même moment, un petit nombre d'adresses suffit. Par conséquent, il devient difficile d'identifier les usagers car leurs adresses changent constamment. Seul le fournisseur d'accès peut y arriver en consultant son fichier journal (*log file*). Ce fichier lui permet de faire la correspondance entre l'adresse IP et l'identité véritable de l'utilisateur.

Malgré l'économie engendrée par l'attribution dynamique, le nombre d'adresses IP disponibles est insuffisant pour suffire à la demande. Aussi, sans la mise en place d'une nouvelle architecture, la croissance d'Internet pourrait être freinée par le manque d'adresses IP. Pour cette raison, l'IETF (<http://www.ietf.org>) a développé une nouvelle version du protocole IP, nommé IPv6. Celui-ci allonge l'adresse IP de 32 à 128 bits et multiplie le nombre d'adresses disponibles de façon exponentielle. Ce nouveau protocole est actuellement intégré à la plupart des produits vendus et devrait être mis en application au cours des prochaines années.

► Les noms de domaine

En plus de leur adresse numérique, de nombreux ordinateurs possèdent un nom de domaine. Il s'agit d'une série de mots séparés par des points. De façon générale, chaque nom de domaine correspond à une adresse IP. Ils servent principalement à substituer une adresse mnémorique à l'adresse numérique qui peut être difficile à mémoriser. Ils offrent également un avantage pratique car ils permettent de modifier l'adresse IP d'un service sans modifier l'adresse utilisée par les usagers.

Les noms de domaine sont, eux aussi, hiérarchiques. Par contre, contrairement aux adresses IP, les éléments plus généraux se trouvent à la droite de l'adresse. Les noms de domaine se décortiquent donc à l'envers. Par exemple, www.jurisint.org est le nom de domaine de Juris International. La partie « org » reflète le statut de l'entité en indiquant qu'il s'agit d'une organisation. Cette partie du nom est appelée domaine de tête. La partie « jurisint » désigne spécifiquement l'organisation. Il s'agit du nom de domaine de second niveau. Un troisième niveau peut aussi être défini afin de désigner un serveur en particulier. Ici, « www » réfère au serveur qui gère les requêtes Web. Ce troisième niveau est optionnel. Enfin, il est possible d'ajouter des sous-domaines à cette hiérarchie. Par exemple, www.lexum.umontreal.ca est un sous-domaine de second niveau de www.umontreal.ca.

Les noms de domaine de tête constituent la racine de l'organisation des noms de domaine sur Internet. Ils se regroupent dans deux structures principales. La première est géographique et s'organise par pays d'origine. Il existe, par exemple, une racine « fr » pour la France, « ca » pour le Canada, « ma » pour le Maroc, etc. La seconde est organisationnelle et regroupe les noms de domaine de tête génériques suivants :

- « com » (entreprises),
- « edu » (universités et collèges américains),
- « gov » (gouvernement fédéral américain),
- « mil » (forces armées américaines),
- « net » (ressources globales de nature réseau),
- « org » (organismes à but scientifique, de recherche ou non lucratif)

et prochainement les :

- « .aero » (industrie du transport aérien),
- « .bizz » (affaires),
- « .coop » (coopératives),
- « .info » (information (utilisation illimitée)),
- « .museum » (musées),
- « .name » (noms personnels) et
- « .pro » (compatibles, avocats et médecins).

Pour posséder un nom de domaine de second niveau sous l'une des hiérarchies, un enregistrement est nécessaire auprès du registraire compétent. Généralement, chaque racine nationale est dotée de son propre registraire. Pour ce qui est de la hiérarchie organisationnelle, elle est gérée par l'ICANN (*Internet Corporation for Assigned Names and Numbers*) (<http://www.icann.org>) depuis 1998. Le principal rôle des registraires est de maintenir une base de donnée associant les noms de domaine enregistrés aux adresses IP correspondantes. Ils s'assurent également de l'unicité des noms de domaine de second niveau dont ils ont la responsabilité.

► Le système DNS

Puisque les ordinateurs communiquent entre eux à l'aide d'adresses IP et que les usagers les identifient avec des noms de domaine, un mécanisme de conversion est nécessaire. Ce rôle est tenu par le système de noms de domaine ou DNS (*Domain Naming System*). Ce système est composé de milliers de serveurs DNS. Pour être en mesure de résoudre les noms de domaine en adresse IP, chaque ordinateur doit connaître l'adresse de l'un d'entre

eux. Ainsi, à chaque fois que l'utilisateur utilise un nom de domaine, son ordinateur envoie une requête au serveur DNS qui lui fournit l'adresse IP correspondante.

La caractéristique fondamentale du système de noms de domaine est qu'il repose sur le principe de la délégation. En effet, chaque domaine doit avoir son propre serveur DNS contenant de l'information sur ses ordinateurs et ses sous-domaines. Il n'a pas besoin de contenir d'information sur les autres domaines, puisqu'il peut toujours la réclamer des serveurs DNS responsables de ces domaines. En ce sens, le système de nom de domaine est réellement distribué sur l'ensemble d'Internet. Une autre propriété des serveurs DNS est leur capacité de mémoriser les noms de domaine réclamés pendant un certain temps. En effet, ceux-ci possèdent une cache, c'est-à-dire une mémoire de masse temporaire. De cette façon, si une requête est envoyée pour un nom de domaine ayant été résolu dernièrement, le serveur peut retransmettre l'information immédiatement.

Le système de nom de domaine opère dans l'ombre et est pratiquement invisible pour l'utilisateur. Ainsi, lorsque l'adresse www.juristint.org est entrée dans un navigateur Web, celui-ci commence par contacter son propre serveur DNS. Le serveur fouille d'abord sa cache à la recherche de ce nom de domaine. Si l'information ne s'y trouve pas, il interroge ensuite le serveur DNS du domaine « org ». Celui-ci délègue alors la requête au serveur DNS du domaine « juristint » dont l'adresse est présente dans sa base de données. Enfin, ce dernier est en mesure de répondre à la requête initiale et retourne l'adresse IP du serveur « www » de son domaine. Le navigateur peut alors contacter le serveur du site Web afin d'afficher le document réclamé.

Certains éléments sont susceptibles de modifier quelque peu l'application de ce système. Ainsi, les protocoles TCP/IP peuvent transporter de l'information supplémentaire, y compris le nom de domaine. Par exemple, le protocole HTTP fournit le nom de domaine du serveur Web, ce qui permet d'installer plusieurs sites Web sur une même adresse IP. Inversement, il est possible d'enregistrer plusieurs noms de domaine et de les faire pointer vers une seule adresse IP.

► L'accès au réseau

L'accès à Internet est certainement la première chose à prendre en compte lors de l'élaboration d'un plan de commerce électronique. Les choix effectués à ce niveau sont très importants car ils influencent, entre autres, les services offerts à l'entreprise, sa vitesse de réaction, l'image projetée, etc. Le marché de l'accès à Internet pouvant être très compétitif, les entrepreneurs font face à un vaste éventail de possibilités.

Il existe trois types de fournisseurs de services Internet (FAI). Premièrement, les fournisseurs d'accès individuels offrent une gamme variée de services allant du simple courrier électronique à l'accès total et illimité à Internet. La création d'un compte PPP ou SLIP auprès d'un fournisseur d'accès individuel est généralement la méthode la plus avantageuse pour obtenir un accès complet à Internet pour un faible coût. Le plus souvent, ces services accordent une banque d'heures ou un accès illimité pour un prix forfaitaire. Cette solution est, à l'heure actuelle, la plus utilisée. Deuxièmement, certaines entreprises offrent en contrepartie d'un prix légèrement supérieur un accès non seulement au réseau Internet mais aussi à leur propre réseau, lequel contient en principe plusieurs services à valeur ajoutée. Il peut s'agir, par exemple, de magazines en ligne, de services de cotes boursières ou de groupes de bavardage. America Online (<http://www.aol.com>) est certainement le plus connu de ces services. Troisièmement, il est parfois possible

d'accéder à Internet par le biais de réseau d'accès public. Il s'agit de réseaux locaux ou étatiques offrant des services spéciaux à leur communauté, parfois gratuitement.

Jusqu'à maintenant, ces divers services exploitent principalement le marché des liaisons téléphoniques. Pour se connecter à Internet de cette façon, il est donc nécessaire de posséder un modem. Le modem est un MODulateur DÉModulateur qui transforme le signal numérique en signal analogique afin de le transmettre par les lignes téléphoniques. Les modems atteignent une vitesse maximale de 56000 bits par seconde, ce qui permet de transférer 1 mégabit de données en 8 à 10 minutes.

Il existe plusieurs autres technologies permettant de se connecter à Internet tout en bénéficiant d'une plus grande vitesse de transmission. Tout d'abord, il existe des liaisons numériques qui utilisent deux canaux à 64 kilo-octets par seconde chacun. Puisqu'elles laissent passer le signal numérique, ces connexions ISDN (USA) ou RNIS (France) nécessitent un adaptateur plutôt qu'un modem. Son utilisation est toutefois restreinte compte tenu de l'important coût d'installation associé à cette technologie. D'ailleurs, les connexions par câble et par ADSL (*Asynchronous Digital Subscriber Lines*) sont beaucoup plus populaires. Les connexions par câble sont généralement disponibles là où existe déjà un réseau de câblodistribution. Cette technologie permet d'obtenir des vitesses variant entre 30 et 40 méga-octets par seconde. L'ADSL est, quant à elle, moins rapide que le câble puisque sa vitesse maximum ne dépasse pas 7 mégabits par seconde. Son implantation est limitée aux régions métropolitaines car elle implique d'énormes investissements. Néanmoins, elle repose sur l'utilisation de lignes téléphoniques traditionnelles et permet ainsi d'offrir un accès Internet rapide là où aucun réseau de câblodistribution n'existe. Enfin, quelques multinationales offrent maintenant des connexions Internet par satellite. Les DSS (*Digital Satellite Service*) ne procurent cependant qu'une vitesse maximum de 400 kilo-octets par seconde. De plus, la transmission étant unidirectionnelle, il n'est pas possible d'envoyer de données par le biais de ces services. L'utilisateur doit donc recourir à une technologie complémentaire, telle que le modem.

Dans tous les cas, le choix d'un fournisseur de service Internet implique de prendre certains critères en considération. La qualité du service à la clientèle est un élément déterminant. Celui-ci devrait être disponible 7 jours sur 7 et préférablement 24h sur 24. La qualité de la connexion est également primordiale. Un fournisseur d'accès doit permettre à ses usagers d'utiliser leur matériel à leur plein potentiel. Un modem 56 kilo-octets est inutile si la connexion s'effectue toujours à une vitesse de 28 kilo-octets. Dans le même ordre d'idées, l'accès devrait être disponible en tout temps. Certains fournisseurs d'accès ne possèdent pas assez de lignes téléphoniques et il devient impossible de les rejoindre lors des heures les plus achalandées. À ce sujet, un ratio de 15 usagers par modem est acceptable. La vitesse du lien entre le fournisseur et Internet est un autre élément à prendre en compte, particulièrement pour les liaisons téléphoniques. À cet égard, la norme est aujourd'hui de deux connexions T1 (1,5 méga-octets par seconde chacune) ou plus.

Le recours à un fournisseur de service Internet n'est pas l'unique façon d'accéder à Internet. Il est également possible de devenir son propre fournisseur d'accès en se branchant directement au réseau. Cette possibilité est généralement réservée aux entreprises disposant d'importants moyens financiers car elle nécessite des investissements matériels conséquents. Ce type de connexion à Internet s'effectue généralement par le biais de lignes téléphoniques dédiées T1 ou T3 (45 méga-octets par

seconde). Ces dernières servent souvent à la mise en place d'intranet ou d'extranet. Un intranet est un réseau local compris à l'intérieur d'une entreprise. Ces réseaux utilisent les mêmes technologies qu'Internet mais seuls les employés des entreprises concernés y ont accès. Il s'agit, en quelque sorte, d'un Internet privé. L'extranet, quant à lui, est une partie de l'intranet qui est rendu accessible à certains usagers se trouvant à l'extérieur de l'entreprise. Ainsi, il peut servir à partager, de façon sécuritaire, de l'information d'affaire ou à effectuer des opérations avec des fournisseurs et partenaires.

L'établissement de tels réseaux n'est toutefois pas l'exclusivité des grandes entreprises. En effet, il est également possible de connecter plusieurs ordinateurs à Internet par le biais d'une seule connexion à un fournisseur de service Internet. Ceci se réalise par la création d'un réseau local où l'ordinateur connecté est utilisé en tant que routeur IP. Selon ce modèle, seul le routeur est doté d'une adresse IP sur Internet. Pourtant, dans les faits, le routeur partage les paquets IP qui lui parviennent entre les différents postes de travail, en utilisant l'adresse du réseau local qui figurent sur ceux-ci. Cette façon de procéder permet donc de partager une adresse IP entre plusieurs ordinateurs.

Le Web

Si Internet est aussi populaire aujourd'hui, c'est en grande partie grâce au Web. La « toile » est la portion la plus facilement navigable d'Internet. Développé au CERN (*European Particle Physics Lab*) en Suisse par Tim Berners-Lee, le Web est basé sur le principe de l'hypertextualité. Cette technologie permet de lier des documents entre eux, indépendamment de leur location. Ainsi, en créant des liens hypertextes, on permet à l'utilisateur de voyager d'un serveur à l'autre par de simples clics. Le Web possède aussi l'avantage d'unifier l'ensemble des services qui étaient antérieurement offerts sur Internet (FTP, Gopher, Usenet, courrier électronique) dans un espace global d'information et de communication. Trois technologies de bases sont utilisées pour obtenir ces résultats, soit la norme d'adressage des URL, le protocole HTTP et le langage HTML. Cependant, l'hébergement d'un site Web et sa navigabilité impliquent la connaissance de notions particulières. Plusieurs technologies secondaires se sont aussi développées autour du Web, telles que les plugiciels, le langage Java, les fichiers témoins et les fichiers journaux.

► Les URL

L'URL (*Uniform Ressource Locator*) est une norme d'adressage qui permet de faire référence à toutes les ressources présentes sur Internet. Elle sert à spécifier l'adresse Internet des services offerts par des ordinateurs connectés au réseau. Par exemple, il peut s'agir de l'accès à un document spécifique, d'une connexion en mode terminal sur un ordinateur distant ou de communiquer avec un internaute. La forme de l'URL est toujours la même :

Service://<user>:<password>@<host>:<port>/<path>

Le « service », aussi appelé « préfixe », correspond au protocole utilisé pour établir la connexion. Le protocole le plus utilisé est le HTTP, qui permet d'accéder aux documents hypertextes. Évidemment, l'ordinateur contacté doit posséder un serveur de ce type pour que la connexion puisse être établie. Les deux arguments <user> et <password> sont optionnels. Ils sont utilisés uniquement lorsque le serveur requiert une identification. Ceci

se présente, notamment, lorsqu'un usager désire accéder à son courrier électronique. Généralement, le préfixe sera directement suivi de la section <host> qui désigne l'adresse IP de l'ordinateur contacté. Celle-ci peut être remplacée par un nom de domaine. L'argument <port> sert, pour sa part, à spécifier le lieu logique de connexion qu'utilise le client pour communiquer avec le serveur. Il est représenté par un nombre variant entre 0 et 65536. Toutefois, il n'est habituellement pas nécessaire de le mentionner car un port par défaut est associé à chaque protocole. Ainsi, le port 80 est assigné au protocole HTTP. Finalement, le <path>, ou suffixe, désigne l'emplacement d'un fichier sur le disque dur de l'ordinateur contacté. Il est composé d'une série de répertoires et du nom du fichier en question. En somme, un URL complet ressemble à ceci :

`http://lemyrep:xxx@www.jurisint.org:80/pub/05/fr/index.htm`

Cependant, puisque les logiciels clients prennent souvent l'adressage en charge, peu d'utilisateurs utilisent directement les URL. En général, une interface intuitive se substitue à ceux-ci. Par exemple, les signets permettent d'accéder aux sites Web gardés en mémoire par le biais de boutons. Également, les navigateurs Web modernes sont conçus pour corriger certaines erreurs présentes dans les URL. Ainsi, si le protocole n'est pas mentionné, le navigateur présume qu'il s'agit d'une requête HTTP. De la même façon, si le domaine de tête manque, le domaine « .com » sera consulté. Le résultat est qu'en entrant un URL incomplet tel que « ibm », on aboutit tout de même au site Web de l'entreprise du même nom, soit à l'adresse `http://www.ibm.com`.

► Le protocole HTTP

Le protocole HTTP (*HyperText Transfer Protocol*) constitue le mode de communication utilisé pour la diffusion de contenu sur le Web. Il définit le format des messages qui sont échangés entre les logiciels clients (navigateur Web) et les serveurs Web. Il entre en jeu lorsqu'un URL commençant par `http://` est sollicité. À ce moment, une requête est envoyée au serveur qui retourne une réponse comprenant l'objet réclamé.

Deux propriétés particulières caractérisent le protocole HTTP. Premièrement, ce protocole n'établit pas de connexion entre le client et le serveur. Le serveur ne fait que recevoir une requête à laquelle il répond immédiatement. Dès qu'une requête est satisfaite le serveur ignore complètement le client qui en est l'auteur pour se consacrer à d'autres requêtes. Cette façon de faire simplifie grandement la structure des serveurs et augmente leurs performances. Par contre, elle rend le suivi des usagers impossible. Deuxièmement, le protocole HTTP est sans état, ce qui signifie qu'il n'existe aucun lien entre les requêtes. Chaque requête est traitée comme si elle était la toute première provenant de ce client, c'est-à-dire sans tenir compte du contexte. En somme, lorsqu'un usager désire consulter un document contenant des images, son navigateur envoie plusieurs requêtes, c'est-à-dire pour le document-même et pour chaque image. En effet, le serveur ne réalise pas qu'existe un lien entre le document et ses images ou entre les différentes requêtes. Toutefois, cette affirmation doit être tempérée puisque certaines applications, employées de concert avec le protocole HTTP, permettent de tenir compte des requêtes antérieures.

La requête HTTP permet de passer plusieurs types de commandes aux serveurs Web. Les principales commandes sont « GET », « HEAD » et « POST ». La commande la plus utilisée est « GET » et sert à demander au serveur de renvoyer le contenu existant à l'URL spécifié. En général, il s'agit d'un simple fichier mais l'URL peut également pointer vers un programme à exécuter. La commande « HEAD » est semblable mais seule l'entête

associée à la ressource est retournée. Ceci permet, entre autres, de vérifier la date de la dernière modification effectuée sur un fichier. La commande « POST » sert plutôt à envoyer des données au serveur dans le but, par exemple, de retourner un formulaire rempli par un usager.

La requête, en plus de contenir une commande, est constituée de différentes informations. D'abord, elle comprend l'URL où la commande doit être exécutée et la version du protocole utilisée par le client. Un certain nombre de champs optionnels peuvent également être ajoutés. Il s'agit, par exemple, de l'adresse de courrier électronique définie dans les options du navigateur (« From: »), d'une date qui spécifie au serveur que les documents qui ont été modifiés depuis celle-ci ne doivent pas être transmis (« If-Modified-Since: ») ou du type de navigateur utilisé (« User-Agent: »). Un corps de message peut ensuite suivre s'il s'agit d'une commande « POST ».

La réponse du serveur débute toujours par une ligne de statut indiquant la version du protocole utilisée ainsi qu'un code de réponse. Ce code indique si la requête a pu être traitée correctement et, si la réponse est négative, pourquoi. Le code d'erreur le plus commun est le 404, qui signifie que la ressource demandée n'existe pas. Ce code survient, entre autres, lorsqu'un lien hypertexte qui n'a pas été vérifié depuis un certain temps pointe vers une ressource qui a été déplacée ou supprimée. Tout comme pour la requête, des champs optionnels peuvent suivre. Par exemple, « Date: » indique la date de la réponse, « Content-Type: » précise le type du document envoyé et « Server: » permet de préciser le nom et la version du serveur. Vient ensuite le corps du message qui contient le document demandé.

Encadré 01 : Exemple de conversation HTTP

Requête :

```
GET /INDEX.HTML HTTP/1.0  
From : lemyrep@lexum.umontreal.ca  
If-Modified-Since : Sunday, 11-May-1997 9:33:11 GMT  
User-Agent : Mozilla/3.0 (Win95; I)
```

Réponse :

```
HTTP/1.0 200 OK  
Date : Sunday, 11-May-1997 19:33:14 GMT  
Server : Apache/1.1  
Content-Type : text/html  
Content-Length : 465  
Last-Modified : Sunday, 11-May-1997 10:54:42 GMT
```

<html> ...

► Le langage HTML

Les normes de communication établies par le protocole HTTP ne sont pas suffisantes pour créer un environnement universel fondé sur l'hypertextualité tel que le Web. Il est également essentiel qu'utiliser un langage commun. Sur le Web, ce langage est le HTML (*HyperText Markup Language*). On peut définir le HTML comme étant un langage de balisage de document qui permet d'obtenir l'affichage d'une structure et d'une présentation adéquate sur n'importe quel type d'ordinateur. En fait, les documents HTML sont de simples fichiers textes. Ce sont des balises ajoutées au texte et permettant aux navigateurs d'afficher le résultat escompté par l'auteur.

Ainsi, pour créer un document HTML, un simple éditeur de texte est suffisant. Une fois le balisage effectué, il suffit de changer le suffixe « .txt » par « .html ». Cependant, très peu de Webmestres créent leurs documents de cette façon en raison du temps et de la complexité que cela peut impliquer. Pour y arriver, il faut en effet connaître l'ensemble des balises à utiliser et les écrire au long. La vaste majorité des créateurs de site Web utilisent donc des logiciels spécialisés. Certains d'entre eux permettent de générer automatiquement du code HTML à partir d'autres documents ou à partir d'une interface de mise en forme. Ces logiciels devraient cependant être mis de côté puisque le code HTML qui en résulte est de piètre qualité. En effet, celui-ci est en général trop compliqué, limite les possibilités de modification ultérieure et comprend de nombreuses balises inutiles qui alourdissent considérablement la taille du fichier et entraînent souvent des erreurs de lecture. Les logiciels spécialisés permettant de travailler directement sur le code source devraient leur être préférés.

À l'intérieur d'un document HTML, tous les éléments structurels ou de présentation doivent être placés dans des balises. Celles-ci s'ouvrent toujours par le caractère « < » et se ferment toujours par le caractère « > ». Par exemple, la balise « » permet de placer du texte en caractères gras. De cette façon, tout texte situé à l'extérieur des crochets est considéré comme un élément de contenu et sera affiché sur la page Web alors que tout ce qui se situe à l'intérieur des crochets reste invisible pour l'utilisateur. Lorsqu'un élément structurel ou de présentation doit être appliqué à une portion du contenu, les balises viennent généralement par paires : les balises d'ouverture (« <> ») indiquent le début de l'application et les balises de fermeture (« </> ») en indiquent la fin. Ainsi, la balise « » se ferme par une balise « ». Lorsque plusieurs balises sont ouvertes les unes après les autres, elles s'additionnent. Également, la façon dont le contenu est présenté à l'intérieur d'un document HTML n'a aucune importance puisque la présentation est prise en charge par les balises. Par exemple, même si des sauts de ligne étaient insérés régulièrement, ceux-ci n'apparaîtraient pas sur la page Web à moins d'ajouter la balise «
 » (qui implique un saut de ligne) à la fin de chaque ligne.

En plus des balises, tous les documents HTML ont la même structure en commun. D'abord, pour confirmer la nature du document au navigateur, ceux-ci doivent toujours commencer par une balise « <html> » et se terminer par sa correspondante « </html> ». À l'intérieur de cette balise se trouvent deux sections principales, soit l'entête et le corps du document. L'entête débute par « <head> » et se termine par « </head> ». Elle contient de l'information relative au document tel que son titre, son auteur, une description, des mots clefs, etc. Le corps du document suit immédiatement l'entête et débute par « <body> » pour se terminer par « </body> ». On y retrouve l'ensemble du contenu.

Pour assurer l'uniformité stylistique d'un groupe de pages Web, il est également possible de contenir les éléments de présentation à l'intérieur d'un fichier autonome, appelé feuille de style ou CCS (Cascading Style Sheet). Cette façon de faire est d'ailleurs préférable car elle rend les modifications ultérieures de la présentation d'un site plus facile. En effet, il est beaucoup plus aisé de modifier une seule feuille de style plutôt que des dizaines, voir des centaines de fichiers HTML. Sur le plan technique, la feuille de style associe des éléments de présentation aux éléments structurels qui sont susceptibles de se retrouver dans les fichiers HTML visés. Par exemple, chaque titre de section peut être associé à une police, une taille de caractère, une couleur, etc. Pour que les navigateurs en tiennent compte, il suffit que l'URL de la feuille de style soit mentionnée dans l'entête du fichier HTML.

Le langage HTML peut être insuffisant pour la mise sur pied de sites complexes. Par exemple, les Webmestres qui ont à gérer de grande quantité d'informations ont besoin de bases de données, ce que le langage HTML est incapable de prendre en charge. Pour cette raison, de nombreux serveurs d'applications se sont développés. Ces derniers se superposent aux serveurs Web afin de traduire les requêtes HTTP dans un format compréhensible par les bases de données ou d'autres types de logiciels. Ils effectuent ensuite l'opération inverse en créant du code HTML à partir de la réponse reçue. Ces outils sont aujourd'hui très répandus parmi les sites Web commerciaux.

Encadré 02 : Exemple de fichier HTML



<html>

<head>

<title>Cour suprême du Canada - Menu principal</title> (1)

<meta NAME="Keywords" CONTENT="Cour suprême du Canada, histoire,

compétence, juridiction, juges, procédure de nominations des juges, procédure d'appel, administration, édifice de la Cour suprême du Canada, visites guidées">
<meta NAME="description" CONTENT="Table des matières pour l'information sur l'histoire, l'administration, les juges, l'édifice et les visites guidées de la Cour suprême du Canada">
</head>

```
<body text="#000000" background="scmark.jpg">
<img SRC="line.gif" ALT="Decorative line graphic" height="10" width="100%"
align="CENTER"> (2)
<table align="center" WIDTH="701">
<tr><td width="565" colspan="3">
<p align="center"><b><font size="+2" color="#800000">Cour suprême du
Canada.<br></font></b> (4)
</td></tr>
<tr><td width="159" valign="top" align="left">
<img SRC="scc-csc.jpg" alt="Sceau de la Cour suprême du Canada" WIDTH="159"
HEIGHT="159"> (2)
</td>
<td VALIGN="middle" NOWRAP width="185" align="left">
<ul>
<li><a href="What-new-Quoi-neuf/menu_f.htm">Quoi de neuf</a></li> (3)
<li><a href="Role/menu_f.htm">Rôle</a></li>
<li><a href="History_Histoire/menu_f.htm">Histoire</a></li>
<li><a href="Judges_Juges/menu_f.htm">Juges</a></li>
<li><a href="Administration/menu_f.htm">Administration</a></li>
<li><a href="Visits_Visites/visite_f.htm">Visites de la Cour</a></li>
</ul>
</td>
<td VALIGN="middle" NOWRAP width="337" align="left">
<ul>
<li><a href="Judgments_Jugements/menu_f.htm">Jugements</a></li> (3)
<li><a href="News_Release/presse_f.htm">Communiqués de presse</a></li>
<li><a href="Case_Information_Dossier/menu_f.htm">Renseignements sur les
dossiers de la Cour</a></li>
<li><a href="http://204.19.231.171/accueil.htm">Bibliothèque et recherche</a></li>
<li><a href="Act_Rules_Loi_Règles/menu_f.htm">Loi et Règles</a></li>
<li><a href="FAQ/menu_f.htm">FAQ</a></li>
</ul>
</td></tr></table>
<p align="center">[ <a href="index_e.htm">English</a> | <a
href="Contacts/contacts_f.htm">Contacts</a> | <a href="photos_f.htm">Sources</a>
| <a href="sitemap/plan-site_f.html">Plan du site</a> ] (3)
<p align="center"><img SRC="line.gif" ALT="Decorative line graphic" height="10"
width="100%"> (2)
</body>

</html>
```

(1) Titre du document

(2) Images

(3) Liens hypertextes

(4) Titre de section

► L'hébergement

Une fois qu'un site Web a été mis sur pieds, encore faut-il le rendre accessible au public. Pour y arriver, celui-ci doit être placé sur un ordinateur disposant d'un serveur HTTP et d'une connexion permanente au réseau. Le type d'hébergement choisi est très important car il influence directement l'image projetée par l'entreprise sur Internet. Plusieurs éléments peuvent être considérés, dont principalement :

- ❑ La taille de la bande passante. Généralement, une ligne T1 ou T3 est suffisante. La bande passante doit être évaluée en tenant compte du nombre et du type de sites hébergés en commun. Ainsi, une ligne T1 peut être insuffisante pour 5 sites fortement achalandés et suffisante pour une cinquantaine de sites peu fréquentés ;
- ❑ L'espace disque. Celui-ci doit pouvoir contenir l'ensemble du site. Le nombre et la taille des fichiers utilisés ainsi que l'évolution probable du site doivent alors être pris en considération ;
- ❑ Le support technique. En cas de panne, le site doit pouvoir être remis sur pieds rapidement. Également, un technicien devrait être disponible pour résoudre les difficultés techniques qui sont susceptibles de se présenter ;
- ❑ La possibilité d'utiliser un nom de domaine et des alias de courrier électronique propres ;
- ❑ Les services complémentaires, tel que des outils d'administration du site (statistiques).

La nature des activités de l'hébergeur joue également son rôle. Par exemple, plusieurs fournisseurs d'accès Internet proposent des services d'hébergement à leurs clients. Bien que la vaste majorité d'entre eux soient compétents, leurs efforts et investissements risquent d'être accordés en priorité aux services d'accès à Internet. De la même façon, la plupart des développeurs de site Web proposent des services d'hébergement satisfaisant. Toutefois, le coût de ces services est élevé et la bande passante limitée. Pour ces raisons, il est souvent préférable d'avoir recours à des hébergeurs spécialisés.

Parmi les hébergeurs spécialisés, de nombreuses possibilités sont envisageables. Pour la vaste majorité des entreprises, la location d'une quantité d'espace disque sur un serveur partagé est suffisante. Cette solution possède l'avantage d'être facile à mettre en place en raison de la prise en charge du support technique par le fournisseur de services. Les entreprises qui souhaitent développer des sites Web plus complexes peuvent aussi envisager la co-location. Cette technique consiste à fournir le matériel et les logiciels nécessaires au fournisseur de service, lequel s'engage à l'entreposer et à fournir la connexion à Internet. Cette solution accorde beaucoup de flexibilité à l'entreprise mais implique un coût élevé et requiert une expertise technique. À ces deux façons de faire s'ajoute la possibilité de louer un serveur dédié, c'est-à-dire entièrement consacré au site Web de l'entreprise.

Finalement, les entreprises disposant des connaissances techniques nécessaires peuvent placer leur site Web en ligne elles-mêmes. Il faut toutefois disposer d'un accès rapide et continu à Internet, d'une adresse IP permanente, d'un ordinateur assez performant pour répondre au besoin du site et d'un logiciel serveur Web.

► La navigation

Le succès d'un site Web marchand ne dépend pas uniquement des biens et des services qui y sont offerts. Le site doit être bien conçu afin de répondre adéquatement aux besoins de rapidité et de facilité des internautes. À ce titre, toutes les pages HTML d'un site devraient être accessibles en moins de 2 ou 3 clics de souris. Pour atteindre ce but, il faut tout d'abord s'assurer que le site peut être lu par les principaux navigateurs Web graphiques, soit Netscape (<http://www.netscape.com>) et Internet Explorer (<http://www.microsoft.com/windows/IE/>). Cette vérification implique que l'utilisation de balises HTML inconnues de l'un des deux navigateurs doit être évitée. Le site devrait également être conçu en fonction des utilisateurs de navigateurs textuels, tel que Lynx (<http://lynx.browser.org/>), en ajoutant une légende à chaque image.

D'autres critères de conception doivent être respectés afin d'assurer une lecture agréable du site. Le principal critère concerne sans aucun doute la taille des documents HTML. Une page devrait en principe se charger en moins d'une vingtaine de secondes. Il est donc important de ne pas alourdir les pages avec des images et des objets de tailles importantes. Habituellement, il est recommandé de séparer le contenu d'une page lorsqu'elle fait plus de 60 kilo-octets. Parallèlement, il est possible de réduire le nombre de documents transférés en tenant compte des caches utilisées par les navigateurs Web. En effet, les navigateurs conservent une copie des documents récemment téléchargés afin d'éviter la répétition du transfert lorsqu'un usager réclame le même document. Il est donc possible d'accélérer la navigation des usagers en faisant, par exemple, systématiquement référence aux documents de la même façon ou en créant un répertoire unique réservé aux images.

De nombreux autres critères peuvent être considérés. Par exemple, les liens hypertextes et les scripts qui ne fonctionnent pas sont des éléments qui nuisent à la navigation. Également, en matière d'URL, le nom des fichiers principaux de chaque niveau hiérarchique devrait être « index.html » afin de permettre aux usagers de passer simplement d'une section à l'autre.

Mêmes en tenant compte de tous ces critères, l'information demeure difficile d'accès si le site n'est pas doté d'une structure adéquate. Aujourd'hui, la majorité des sites Web sont conçus de façon similaire. La page préliminaire qui consiste à offrir aux usagers le choix parmi plusieurs versions du site (légère ou à large bande passante) est optionnelle. Il peut s'agir aussi d'une introduction Flash (<http://www.macromedia.com/software/flash/>). Dans tous les cas, la page préliminaire mène directement à la page d'accueil qui se situe au sommet de l'architecture pyramidale du site. Celle-ci constitue le pivot central autour duquel tous les autres documents sont rattachés. Elle contient des liens vers les principales sections du site, elles-mêmes subdivisées en sous-sections. Pour se déplacer aisément à l'intérieur de cette hiérarchie, chaque page devrait contenir une barre de navigation contenant des liens vers les principales sections. Dans la même optique, les sites complexes devraient présenter leur structure hiérarchique. Enfin, les coordonnées de l'entreprise devraient être placées bien en vue.

► Les plugiciels

À l'origine, les navigateurs Web ont été conçus pour afficher seulement du texte et des images. Les plugiciels (*plug-ins*) ont été développés pour contrer cette limite et permettre d'accéder à toutes sortes de contenus à partir d'un simple navigateur. Il s'agit de petits

logiciels pouvant être greffés au navigateur afin d'étendre ses capacités. Une fois cette opération effectuée, le plugiciel prend en charge la lecture de certains types de fichiers. Ainsi, lorsqu'un usager emprunte un lien vers l'un de ces fichiers, son navigateur est en mesure d'en afficher le contenu.

Les navigateurs comprennent toujours une certaine quantité de plugiciels installés par défaut afin d'être en mesure de lire un minimum de fichiers multimédias. L'utilisateur peut également télécharger d'autres plugiciels, lesquels sont en principe gratuits. Par exemple, une personne intéressée par les environnements en trois dimensions devra probablement se procurer un plugiciel VRML (*Virtual Reality Modeling Language*).

La prudence est cependant de mise lors du développement d'un site Web nécessitant l'utilisation de plugiciels. Puisque les usagers du site ne possèdent pas toujours les plugiciels en question, il est possible que certains d'entre eux n'aient pas accès au contenu diffusé. Dans ces circonstances, il est essentiel d'offrir une solution de rechange pour aux usagers ou de les diriger vers un site de téléchargement des plugiciels manquants.

► Le langage Java

Java est un langage public développé par Sun Microsystems (<http://www.sun.com>). L'objectif principal du langage Java est d'être indépendant des plates-formes matérielles. Ceci signifie qu'un logiciel conçu avec Java devrait fonctionner sur n'importe quel ordinateur, peu importe son fabricant ou son système d'exploitation. Même si cet objectif n'est toujours pas atteint, le langage Java permet aujourd'hui aux serveurs HTTP d'envoyer sur le Web des programmes pouvant être exécutés par la plupart des logiciels clients.

Les applications Java que l'on retrouve sur le Web empruntent habituellement la forme d'applets. Il s'agit de petits programmes introduits à l'intérieur d'une page Web par la balise HTML « `<applet></applet>` ». Les navigateurs Web modernes sont en mesure de comprendre ces programmes parce qu'une machine virtuelle Java (*Java Virtual Machine*) spécifique à chaque système d'exploitation est intégrée à leur code. Cette machine permet d'interpréter les applets Java en fonction de la plate-forme matérielle utilisée par l'utilisateur.

Grâce au langage Java, il est possible d'ajouter de l'interactivité aux pages HTML statiques traditionnelles. Par exemple, Java peut servir à créer des objets mobiles à l'intérieur d'une page, à mettre en place des calculateurs en ligne ou à réaliser des formulaires dynamiques. Aujourd'hui, de nombreux sites offrent des bibliothèques publiques d'applets Java. Cependant, ces applets ne devraient pas être utilisés inutilement puisqu'ils alourdisent considérablement la taille des documents HTML.

► Les fichiers témoins

Les fichiers témoins (*cookies*) constituent un autre exemple de technologie complètement intégrée au Web. Bien qu'ils passent souvent inaperçus aux yeux des usagers, ils sont utilisés à grande échelle sur le Web. Leur principale fonction est d'attribuer aux usagers un identificateur unique. Concrètement, un fichier témoin est une simple ligne d'information que le serveur Web inscrit sur le disque dur de l'ordinateur d'un usager par le biais de sa réponse HTTP. Cette ligne de texte est conservée dans un endroit prévu à cet effet par le navigateur Web et est retournée au serveur lorsque l'utilisateur fait contact de nouveau. Grâce à l'utilisation de cette technique, le serveur Web est en mesure de tenir

compte du contexte de la requête et de fournir ainsi un contenu personnalisé. Par exemple, les fichiers témoins permettent au serveur de mémoriser les articles sélectionnés par un consommateur même si celui-ci navigue entre les différentes pages d'un site.

L'utilisation de fichiers témoins est encadrée par des normes techniques strictes en raison des risques qu'ils présentent en matière de vie privée. Leur contenu est restreint à :

- ❑ un numéro d'identification ;
- ❑ une date d'expiration ;
- ❑ l'URL du serveur expéditeur, et à
- ❑ une mention indiquant si une communication sécurisée est nécessaire pour la transmission de l'information.

La ligne d'information ne doit pas dépasser la taille de 4 kilo-octets. Un client Web ne peut emmagasiner plus de 300 fichiers témoins au total, la limite par domaine ou serveur étant fixée à 20. Toutefois, il est toujours possible pour un usager de bloquer les fichiers témoins par la configuration des préférences de son navigateur Web. Cette possibilité doit être considérée lors de la conception d'un site Web dont le fonctionnement repose substantiellement sur l'utilisation de fichiers témoins.

► Les fichiers journaux

Lorsque des documents sont réclamés d'un serveur Web, celui-ci enregistre habituellement une trace des communications dans un fichier journal (*log file*). Ce fichier peut être extrêmement utile pour l'administrateur d'un site Web car il contient des informations sur les usagers. Les principaux éléments pouvant être retrouvés à l'intérieur d'un fichier journal sont :

- ❑ l'adresse IP d'où provient la requête ;
- ❑ l'heure de la requête ;
- ❑ le nom du document réclamé ;
- ❑ le nom d'utilisateur utilisé (si un enregistrement était nécessaire) ;
- ❑ la taille de la réponse en bits ;
- ❑ le logiciel et la plate-forme utilisés pour effectuer la requête ;
- ❑ l'URL de la page de référence si l'utilisateur a cliqué sur un lien hypertexte ;
- ❑ le numéro du fichier témoin envoyé s'il y a lieu, et
- ❑ le code de réponse HTTP.

Encadré 03 : Exemple d'entrée dans un fichier journal

```
132.204.136.36 - - [19/Oct/2000:00:00:04 -0600] "GET
/articles/archives.html HTTP/1.1" 200 20607 "http://www.lex-
electronica.org" "Mozilla/4.0 (compatible; MSIE 5.0; Windows 98;
DigExt)"
```

L'accumulation de cette information requiert beaucoup d'espace disque, particulièrement pour les serveurs achalandés. Pour cette raison, le fichier journal ne comporte pas toujours l'ensemble de ces éléments.

En général, les fichiers journaux servent à créer des statistiques d'utilisation pour les sites Web. Cette tâche est effectuée à l'aide de logiciels spécialisés qui analysent les nombreuses entrées du fichier journal et génèrent des statistiques, lesquelles sont essentielles à l'évaluation de la performance du site. Elles indiquent, entre autres, le nombre de requêtes pour chaque page, une vue d'ensemble de la situation géographique des usagers (déterminée à partir des adresses IP) ainsi que les URL de leur provenance. Certaines entreprises en ligne offrent des services similaires grâce à l'emploi de compteurs plus ou moins complexes.

Les technologies complémentaires

Plusieurs autres technologies doivent être considérées dans le cadre du commerce électronique. La présente section expose leur fonctionnement et traite des différents types de licences dont elles peuvent faire l'objet. Il est notamment question de la diffusion en continu, des certificats électroniques, de la cryptographie, des coupes-feux, des serveurs mandataires et de technologies moins récentes. En effet, certaines d'entre elles ont été créées avant même la création du Web. Par exemple, le courrier électronique remonte aux origines du réseau et les groupes de nouvelles USENET sont utilisés depuis les années 1980.

► Le courrier électronique

Le courrier électronique est la technologie qui représente le plus important échange de données sur Internet. Il s'agit, en quelque sorte, d'une adaptation de la poste puisque le courrier électronique permet une communication en différé. Par contre, cette technologie est à certains égards beaucoup plus avantageuse que la poste traditionnelle : elle est instantanée, gratuite et permet d'envoyer des documents volumineux sans problème.

La messagerie électronique est basée sur l'utilisation de deux protocoles distincts. Le premier, SMTP (*Simple Mail Transfer Protocol*), est un protocole point à point qui sert à établir une communication entre deux serveurs. Ce protocole spécifie l'ensemble des éléments d'un message, soit le format des adresses, les champs du message, la gestion des heures, etc. Ainsi, pour expédier un courrier électronique, il est essentiel de posséder l'adresse du serveur SMTP de son fournisseur d'accès. Lors de l'envoi, le logiciel client fait simplement passer le message à ce serveur qui s'assure ensuite son acheminement vers le serveur du destinataire, où il sera entreposé dans un répertoire réservé à chaque usager. Toutefois, la vaste majorité des individus n'ont pas accès à ce répertoire car il est situé sur un ordinateur distant. Un second protocole, POP3, est donc nécessaire afin de récupérer le message. Ainsi, il devient essentiel de posséder l'adresse du serveur POP3 de son fournisseur d'accès Internet afin d'être en mesure de recevoir un courrier électronique. Le protocole POP3 garantit également la confidentialité des messages présents sur le serveur en les protégeant à l'aide d'un mot de passe. Toutefois, la confidentialité n'est aucunement assurée lors du transfert sur le réseau puisque SMTP et POP3 ne chiffrent pas les messages.

Pour ce qui est de la forme, le message est toujours conçu de la même façon. Il est composé d'une entête comprenant les informations nécessaires au transport du message et d'un corps formé par le texte lui-même. Les principaux champs de l'en-tête susceptibles d'être présents sont :

- ❑ « X-sender » : indique l'adresse de provenance, qui est généralement l'adresse de l'émetteur ;
- ❑ « Date » : il s'agit de la date de début de composition du message ;
- ❑ « To » : indique l'adresse du destinataire ;
- ❑ « From » : indique l'adresse de l'émetteur ;
- ❑ « Subject » : il s'agit du sujet du message ;
- ❑ « Cc » : indique les adresses des autres usagers à qui le message a été envoyé (carbon copy) ;
- ❑ « Bcc » : semblable à « cc », sauf que les destinataires n'en auront pas connaissance les uns des autres ;
- ❑ « Attachment » : permet de joindre un fichier au message, et
- ❑ « Reply-to » : indique l'adresse à laquelle les réponses devront être retournées.

Encadré 04 : Exemple de courrier électronique

*X-Sender: lemyrep@urbino.lexum.umontreal.ca
 X-Mailer: QUALCOMM Windows Eudora Version 5.0
 Date: Tue, 19 Sep 2000 10:39:32 -0400
 To: labbee@lexum.umontreal.ca
 From: Pierre-Paul Lemyre <lemyrep@lexum.umontreal.ca>
 Subject: Juris International
 Cc: daniel.poulin@umontreal.ca
 Bcc: lefebvre@lexum.umontreal.ca
 Attached: techno.doc;*

Bonjour,

Voici la dernière version de mon texte « Le contexte des technologies de l'information »

*Pierre-Paul Lemyre,
 LexUM, Centre de recherche en droit public
 Faculté de droit, Université de Montréal
 lemyrep@lexum.umontreal.ca*

Les adresses de courrier électronique ont la forme suivante :

nom@organisation.domaine

En fonction de cette structure, le nom d'utilisateur doit être unique pour chaque organisation. Toutefois, rien n'empêche qu'une adresse de courrier corresponde à un groupe d'individus. En effet, les listes de discussion et de diffusion permettent d'envoyer un message à une seule adresse afin que celui-ci soit redirigé vers tous les abonnés de la liste. Il devient alors beaucoup plus facile de rejoindre une grande quantité de destinataires.

► Les groupes de nouvelles

Bien que le courrier électronique soit le moyen de communication en différé le plus populaire sur Internet, il n'en demeure pas moins que les groupes de nouvelles (*newsgroups*) sont également très utilisés. Ces babillards électroniques sont supportés par un réseau de serveurs appelé USENET. Généralement, chaque fournisseur d'accès possède son propre serveur USENET, hébergeant ainsi une multitude de groupes de discussion spécialisés. Comme il existe plusieurs dizaines de milliers de groupes et que ce nombre croît sans cesse, les thèmes abordés sont innombrables.

Contrairement au courrier électronique, les articles des groupes de nouvelles ne sont pas envoyés aux destinataires. Ils sont plutôt entreposés sur un serveur afin d'être téléchargés par les usagers. Toutefois, pour diminuer l'utilisation d'espace disque, cet entreposage est limité dans le temps et en nombre. Par exemple, un serveur USENET peut conserver les articles d'un groupe pour une période d'un mois jusqu'à un maximum de 3000 articles. En ce qui concerne le transport de ces articles sur le réseau, il est pris en charge par le protocole NNTP (*Network News Transfer Protocol*). En plus de spécifier la structure des articles, ce protocole permet aux différents serveurs USENET de rafraîchir régulièrement le contenu de leurs groupes de discussion en le comparant entre eux. Ainsi, lorsqu'un usager envoie un article à son propre serveur, celui-ci est relayé graduellement d'un serveur à l'autre, jusqu'à ce que l'ensemble du réseau en ait pris connaissance. Un article peut aussi être supprimé de la même façon.

Enfin, les adresses USENET sont uniquement constituées de mots clefs séparés par des points. Par exemple, « fr.misc.droit.internet » est l'adresse d'un groupe de discussion francophone sur le droit et Internet. Les adresses sont cependant comprises dans l'une ou l'autre des hiérarchies existantes. Les « comp » (sujets reliés aux ordinateurs), « news » (diffusion de nouvelles) et « rec » (sujets récréatifs) s'affichent parmi les principales. Il existe également des hiérarchies nationales, régionales et organisationnelles.

► La diffusion en continu

Lorsqu'un serveur Web reçoit une requête, il transmet l'information vers le client le plus vite possible afin de compléter la communication et de passer à la requête suivante. Le document doit donc être téléchargé en entier avant de pouvoir être visionné. Cette approche est idéale pour les documents classiques du Web qui sont principalement composés de texte et d'images. Cependant, les extraits sonores et vidéos posent le problème de la taille des fichiers. En effet, ces extraits sont nécessairement plus volumineux et leur temps de téléchargement est beaucoup plus important. La diffusion en continu a été développée afin d'éliminer cette attente. Selon ce procédé, le document est affiché à l'écran dès son arrivée dans l'ordinateur de l'utilisateur. En fait, seule une petite portion du document doit être téléchargée lors de l'établissement de la communication afin d'assurer une certaine marge de manœuvre entre l'entrée et la lecture des données. La connexion avec le serveur est donc continue, contrairement à ce qui se produit avec un serveur Web. Le système développé par l'entreprise RealNetworks (<http://www.real.com>) en est un excellent exemple.

La taille des fichiers audio et vidéo demeure néanmoins problématique. En effet, leur visualisation requiert que les données arrivent au moins à une vitesse équivalente à celle de la lecture. La solution consiste à compresser les documents au maximum. La qualité

du son et de l'image est alors réduite. En conséquence, la diffusion en continu constitue un compromis entre la qualité et l'accès au contenu.

► La cryptographie

La cryptographie est une technologie qui ne concerne pas particulièrement Internet. En effet, le but initial de son développement est la sécurisation des communications militaires. Toutefois, en l'absence de mécanisme de protection des communications sur Internet, il est devenu nécessaire d'adapter la cryptographie au réseau Internet. Cette technologie consiste à transformer un message lisible en message chiffré à l'aide d'opérations mathématiques afin que seules les personnes autorisées puissent avoir accès à son contenu. Ainsi, la confidentialité des communications devient possible. La cryptographie assure également l'intégrité et l'authenticité des messages car elle empêche leur altération et permet au destinataire de vérifier l'identité de l'expéditeur.

Les systèmes cryptographiques symétriques, ou à clef secrète, sont les plus anciens. Le plus connu est le système DES (*Data Encryption Standard*). Selon ce système, la même clef sert au chiffage et au déchiffage du message. Le principal problème relié à cette technique est que les deux parties à la communication doivent connaître la clef. Dans le contexte des rapports dématérialisés et momentanés d'Internet, il est difficile de convenir d'une telle clef sans la révéler.

Pour cette raison, les systèmes de cryptographie asymétrique, ou à clef publique, ont été développés. Cette technologie fonctionne par l'attribution d'une paire de clefs propres à chaque partie. Cette paire de clef est créée automatiquement à l'aide de logiciels spécialisés. L'une d'elles, la clé publique, sert à chiffrer les messages. Cette clef doit être accessible à tous et peut être publiée et distribuée. L'autre clef, la clef secrète, sert à déchiffrer les messages. Elle doit donc être conservée précieusement et ne jamais circuler sur le réseau. Ces clefs sont complémentaires, ce qui signifie que tout ce qui est chiffré avec une clef publique ne peut être déchiffré que par la clef secrète correspondante. Cette complémentarité est rendue possible grâce à l'utilisation de fonctions mathématiques à sens unique.

La cryptographie asymétrique est utilisée entre autre par le protocole SSL (*Secure Socket Layer*) mis de l'avant par Netscape (<http://www.netscape.com>). Ce protocole est aujourd'hui intégré à tous les navigateurs Web afin de permettre des communications HTTP sécurisées lorsque la situation le requiert. L'opération, complètement transparente pour l'utilisateur, se déroule de la façon suivante:

1. Un usager entre en communication avec un serveur Web, lequel possède déjà sa paire de clefs publique/privée.
2. Le logiciel client génère une paire de clefs publique/privée.
3. Le logiciel client réclame la clef publique du serveur.
4. La clef publique du client est chiffrée avec la clef publique du serveur et envoyée au serveur.
5. Le serveur déchiffre le message avec sa clef privée.
6. Le serveur envoie une confirmation du bon déroulement de l'opération au client en la chiffrant avec la clef publique du client.
7. Par la suite, toutes les informations circulant entre le client et le serveur sont chiffrées.

Deux autres fonctions peuvent être combinées au chiffrement afin d'augmenter le niveau de sécurité de la communication. Tout d'abord, une signature peut être ajoutée au message afin d'assurer son authenticité. Le mécanisme de la signature fonctionne à l'inverse de celui utilisé pour le chiffrement : la signature est effectuée en utilisant la clef privée et la lecture de la signature requiert la clef publique correspondante. De cette façon, seul l'émetteur du message est en mesure de le signer, alors que toute personne peut l'authentifier. Il est également possible d'assurer l'intégrité du message en ayant recours à une fonction de hachage. Il s'agit d'une opération mathématique qui permet de réduire un message en une série de caractères d'une longueur fixe. L'empreinte résultant de ce hachage doit être jointe au message et signée par l'expéditeur. Lors de la réception, le destinataire peut ainsi effectuer le même calcul et procéder à une comparaison. Si les deux empreintes correspondent, le message n'a pas été falsifié.

► Les certificats électroniques

La cryptographie asymétrique est toutefois imparfaite. En effet, il demeure possible de générer un paire de clefs au nom d'une autre personne. En conséquence, l'identité des parties à une communication, même chiffrée, ne peut être garantie. Dans cette situation, un mécanisme permettant la vérification du lien entre une clef publique et une personne est nécessaire.

Cette problématique a contribué à l'apparition des certificats électroniques. Le certificat est un document qui établit les relations existantes entre une clef publique, son propriétaire et l'application pour laquelle il est émis. Dans le cas d'une personne, il sert à prouver son identité. Il peut également servir à prouver qu'une application n'a pas été détournée de ses fonctions ou qu'un site est bel et bien celui auquel on désire accéder. La valeur des certificats varie en fonction des démarches effectuées afin d'établir l'identité. Ainsi, un certificat nécessitant une constatation physique est nécessairement plus fiable qu'un certificat accordé en ligne. Leur forme, quant à elle, est précisée par la norme X.509. Ils doivent être infalsifiables et comprennent généralement les éléments suivants :

- ❑ le nom, prénom et adresse de courrier électronique du propriétaire ;
- ❑ la clef publique du propriétaire ;
- ❑ la date d'expiration du certificat ;
- ❑ le nom de l'autorité de certification ;
- ❑ un numéro de série ;
- ❑ la signature numérique de l'autorité de certification, et
- ❑ des informations spécifiques supplémentaires.

Encadré 05 : Exemple de certificat électronique

Certificate:

Data:

Version: 0 (0x0)

Serial Number: 0 (0x0)

Signature Algorithm: md5withRSAEncryption

Issuer: C=ZA, SP=Western Cape, L=Cape Town, O=Thawte Consulting cc,

OU=Certification Services, CN=www.thawte.com,

Email=webmaster@thawte.com

```
Validity
Not Before: Nov 14 17:15:25 1996 GMT
Not After : Dec 14 17:15:25 1996 GMT
Subject: C=ZA, SP=Western Cape, L=Cape Town, O=Thawte Consulting cc,
OU=Certification Services, CN=www.thawte.com,
Email=webmaster@thawte.com
Subject Public Key Info:
Public Key Algorithm: rsaEncryption
Modulus:
00:9a:92:25:ed:a4:77:69:23:d4:53:05:2b:1f:3a:
55:32:bb:26:de:0a:48:d8:fc:c8:c0:c8:77:f6:5d:
61:fd:1b:33:23:4f:f4:a8:2d:96:44:c9:5f:c2:6e:
45:6a:9a:21:a3:28:d3:27:a6:72:19:45:1e:9c:80:
a5:94:ac:8a:67
Exponent: 65537 (0x10001)
Signature Algorithm: md5withRSAEncryption
7c:8e:7b:58:b9:0e:28:4c:90:ab:20:83:61:9e:ab:78:2b:a4:
54:39:80:7b:b9:d9:49:b3:b2:2a:fe:8a:52:f4:c2:89:0e:5c:
7b:92:f8:cb:77:3f:56:22:9d:96:8b:b9:05:c4:18:01:bc:40:
ee:bc:0e:fe:fc:f8:9b:9d:70:e3
```

Le mécanisme requiert inévitablement un certain niveau de confiance, puisque l'identité de l'autorité de certification doit elle-même être certifiée par une autorité supérieure. En pratique, les principales autorités de certification se certifient réciproquement, ce qui permet d'obtenir un degré de confiance relativement élevé. De plus, les logiciels clients sont souvent munis d'une liste d'autorités de certification approuvées, ce qui implique qu'ils sont en mesure de vérifier eux-même la fiabilité d'un bon nombre de certificats.

► Les coupe-feux

Bien que les coupe-feux (*firewalls*) ne soient pas nécessaires au fonctionnement d'Internet, ils sont aujourd'hui indispensables aux entreprises qui désirent commercer en ligne tout en protégeant leurs données et services. Un coupe-feu est un dispositif informatique, matériel ou logiciel, qui filtre l'information circulant entre deux réseaux, permettant ainsi la mise en place d'une politique d'accès. Ce dispositif est toujours installé au point de jonction entre deux réseaux, appelé passerelle. Généralement, il sert à protéger un réseau privé des agressions provenant d'un réseau public, tel Internet. Dans ce cadre, les étrangers se voient refuser l'accès aux données privées et il devient possible de limiter les ressources extérieures accessibles aux usagers du réseau privé.

De façon générale, un coupe-feu intercepte tous les paquets IP qui franchissent la passerelle, peu importe leur direction, et détermine s'il doit les bloquer ou les acheminer à leur destination. Pour y arriver, il consulte une liste de règles établies par l'administrateur du réseau. Les critères sur lesquels reposent ces règles varient d'un coupe-feu à l'autre, dont principalement : le type de communication, l'adresse de la source ou de la destination et le port utilisé. Certains coupes-feu sophistiqués sont également en mesure d'analyser les données transférées.

Le niveau de sécurité fournit par un coupe-feu dépend donc de l'utilisation qui en est faite. Ainsi, un coupe-feu qui laisse uniquement passer les courriers électroniques est très sécuritaire car il limite les intrusions possibles à celles causées par une faille du système

de transmission du courrier électronique. Cependant, une telle limitation restreint considérablement l'autonomie des usagers, car ceux-ci n'ont pas accès aux autres ressources d'Internet, soit le Web, les groupes de nouvelles, etc. De plus, un coupe-feu ne protège jamais un réseau contre l'ensemble des intrusions. Par exemple, un transfert de données peut toujours avoir lieu par le biais d'un support physique comme un disque compact.

► Les serveurs mandataires

Les serveurs mandataires (*proxy*) sont généralement utilisés de concert avec les coupe-feu. D'ailleurs, puisque ces deux outils sont installés au niveau de la passerelle d'un réseau et que leur fonctionnement est similaire, ils sont souvent distribués sous forme d'ensemble logiciel. Le serveur mandataire, tout comme le coupe-feu, filtre les paquets qui circulent entre les réseaux. Toutefois, les fonctions du serveur mandataire sont totalement différentes. Il sert principalement à améliorer les performances d'un réseau, à partager une connexion à Internet ou à assurer l'anonymat des usagers.

Un serveur mandataire améliore les performances d'un réseau en réduisant l'utilisation de bande passante à l'aide d'une cache. Ainsi, les documents réclamés par un usager sont gardés en mémoire pendant une certaine période. En conséquence, le serveur est en mesure de les retransmettre immédiatement lorsqu'ils sont réclamés par un autre usager.

Un serveur mandataire est utilisé comme routeur IP afin de partager une connexion à Internet entre plusieurs ordinateurs d'un réseau local. Selon cette stratégie, l'ordinateur qui héberge le serveur mandataire est le seul à posséder une adresse IP sur le réseau Internet. En conséquence, les paquets d'information sont redirigés vers les usagers en fonction de leurs propres adresses sur le réseau privé.

Un serveur mandataire assure l'anonymat des usagers en cachant aux tiers leur adresse IP. En effet, certains serveurs mandataires sont en mesure d'éliminer des paquets toute information à propos de l'ordinateur duquel ils proviennent. Ainsi, seul le serveur mandataire est visible aux yeux des tiers.

► Les logiciels et leurs licences

L'exploitation des technologies précédentes est rendue possible grâce à la diffusion de logiciels spécialisés. Ces logiciels étant sous forme numérique, il est par conséquent aisé de les distribuer en ligne. Pour cette raison, Internet a constitué un élément déterminant dans l'évolution du cadre juridique des logiciels. Depuis de nombreuses années, ce cadre est assuré par des contrats de licence qui accordent aux licenciés certains droits. Ces licences comprennent également des clauses concernant la propriété du logiciel, l'exclusivité de la licence, les restrictions quant au nombre d'usagers, aux droits de copie ou de désassemblage. Avec l'avènement d'Internet, de nouvelles licences sont apparues. Les logiciels concernés se regroupent aujourd'hui en trois grandes catégories : les logiciels propriétaires, les logiciels libres et les logiciels du domaine public.

La vaste majorité des logiciels sont propriétaires, c'est-à-dire que leur licence restreint considérablement leur utilisation, leur redistribution ou la possibilité de les modifier. Pour assurer l'efficacité de ces restrictions, ceux-ci sont distribués sous forme d'exécutables binaires. Le code source nécessaire à la compréhension de leur fonctionnement est donc gardé secret. Par ailleurs, les logiciels propriétaires ont recours à différentes méthodes de

distribution sur le réseau. Ainsi, la redistribution des gratuiciels (*freewares*) est presque toujours permise. Toutefois, leur utilisation et leur modification restent limitées. De la même façon, les partagiiciels (*sharewares*) peuvent être redistribués et utilisés gratuitement pendant une certaine période de temps. Cependant, une fois ce délai atteint, l'utilisateur doit payer le développeur s'il désire continuer à utiliser le logiciel.

Les logiciels libres (*free software*) accordent beaucoup plus de libertés aux licenciés. Leurs licences permettent généralement à toute personne d'utiliser, de copier et de redistribuer le logiciel avec ou sans modification. Ces droits impliquent nécessairement la disponibilité du code source, lequel peut être joint à la version binaire ou rendu disponible sur Internet. Malgré le fait que la désignation anglaise « *free software* » porte à confusion, les logiciels libres ne sont pas nécessairement gratuits. Pour cette raison, certains préfèrent utiliser les termes « *open source* ».

Finalement, certains logiciels appartiennent au domaine public, ce qui signifie que les auteurs abandonnent leurs droits au public. Dans ces conditions, les logiciels ne sont soumis à aucun droit de propriété intellectuelle. Ils accordent donc aux usagers un niveau de liberté semblable aux logiciels libres. Toutefois, rien n'empêche un individu de construire un logiciel totalement propriétaire à partir d'un code du domaine public.

Conclusion

Internet est le résultat d'une multitude de technologies différentes. Celles-ci ont été développées tout au long de l'évolution du réseau par des milliers de chercheurs. Certains des outils ainsi créés sont publics, alors que d'autres demeurent propriétaires. Dans bien des cas, ils ont simplement été empruntés à d'autres champs de connaissance. Aujourd'hui, toutes ces technologies se superposent les unes aux autres pour former le vaste réseau que nous connaissons.

Le futur nous réserve certainement plusieurs autres bouleversements majeurs, puisque Internet est encore très jeune. En effet, les réseaux informatiques ont à peine plus de trente ans et le Web n'a pas encore dix ans. D'ailleurs, une nouvelle norme d'adressage et un nouveau protocole HTTP devraient multiplier les capacités d'Internet au cours des années à venir. Les progrès en matière de communications sans fil sont également prometteurs. Dans tous les cas, il est probable que les prochaines grandes innovations concerneront l'augmentation de la bande passante, laquelle constitue encore aujourd'hui un obstacle à l'évolution d'Internet.

Bibliographie sélective

- ❑ COHEN, Laura, Internet Tutorials, Albany University, 2000, <http://www.albany.edu/library/internet/>.
- ❑ CONNER-SAX, Kiersten, KROL, Ed, The Whole Internet: The Next Generation, O'reilly, Sebastopol, 1999.
- ❑ INTERNET.COM, Web Developer's Virtual Library, Internet.com, 2000, <http://wdvl.com/WDVL/>.
- ❑ MAIRE, Gilles, « Un nouveau guide Internet », (1999) UNGI, <http://www.imagnet.fr/ime/toc.htm>.
- ❑ SOHIER, Danny J., Le guide de l'internaute 2000, Montréal, Éditions Logiques, 1999.