

SUJET:
COMPRENDRE
INTERNET

REFERENCE: ATLANTIC/DOCTECH/
Initiation à l'informatique/Comprendre Internet
Mise à jour: 11/01/11



DOMAINE:
Initiation à l'informatique

COMPRENDRE
INTERNET

VERSIONS & DATES	OBJET	AUTEUR(S)
<i>Version 1.0 – juin 2010</i>	<i>Création</i>	<i>Bernard GIACOMONI - Association A.T.L.A.N.T.I.C</i>
<i>Version 1.1 – Janvier 2011</i>	<i>Amélioration de la présentation</i>	<i>Bernard GIACOMONI - Association A.T.L.A.N.T.I.C</i>

SOMMAIRE

Table des matières

I.INTRODUCTION:	5
I.1.AVANT PROPOS:	5
I.2.BREF RAPPEL HISTORIQUE:	5
I.3.PROGRESSION ADOPTEE POUR CET OUVRAGE:	6
I.4.REMARQUE:	6
II.POINT DE VUE DE L'USAGER:	7
II.1.GENERALITES:	7
II.2.STRUCTURE GLOBALE DU WEB:	7
II.2.1.SHEMA DE PRINCIPE:.....	7
II.2.2.LE FOURNISSEUR D'ACCES:.....	8
II.2.3.LA NAVIGATION SUR INTERNET:.....	8
II.2.4.LE COURRIER ELECTRONIQUE:.....	8
II.2.5.LE TELECHARGEMENT DE FICHIERS:.....	8
II.3.LA COMMUNICATION SUR INTERNET:	9
II.3.1.LES ACTEURS DE LA COMMUNICATION:.....	9
II.3.2.LE MODELE CLIENT-SERVEUR:.....	9
II.3.3.CONCLUSION:.....	10
III.MECANISMES DE COMMUNICATION SUR LE RESEAU:	11
III.1.INTRODUCTION:	11
III.2.LES DIFFERENTS NIVEAUX DE LA COMMUNICATION:	11
III.2.1.TRANSPARENCE DES MECANISMES DE COMMUNICATION POUR L'USAGER:.....	11
III.2.2.LA RONDE DES "DEMONS":.....	11
III.2.3.LES TROIS NIVEAUX DE MECANISMES:.....	12
III.3.LE MODELE O.S.I:	14
III.4.MODELE OSI ET TRAMES D'ENVOI DE DONNEES:	15
III.5.UTILITE DE L'APPROCHE PAR NIVEAUX:	16
IV.MECANISMES DE GESTION DU LIEN PHYSIQUE:	17
IV.1.NOTION PHYSIQUE DE RESEAU INFORMATIQUE:	17
IV.1.1.LES SYSTEMES INFORMATIQUES REPARTIS:.....	17
IV.1.2.NOTION DE TOPOLOGIE:.....	18
IV.1.3.LES TOPOLOGIES DE BASE:.....	19
IV.1.4.RESEAUX LOCAUX:.....	19
IV.1.5.RESEAUX ETENDUS:.....	19
IV.1.6.LE WEB:.....	21
IV.2.ECHANGE DE DONNEES ENTRE DEUX HOTES " ADJACENTS ":	22
IV.2.1.AVANT-PROPOS:.....	22
IV.2.2.MEDIA DE TRANSMISSION:.....	22
IV.2.2.1.GENERALITES:.....	22
IV.2.2.2.INJECTION D'UNE INFORMATION DANS UN MEDIA:.....	22

SUJET:
COMPRENDRE
INTERNET

REFERENCE: ATLANTIC/DOCTECH/
Initiation à l'informatique/Comprendre Internet
Mise à jour: 11/01/11

IV.2.3. TRAME D'ENVOI DE DONNEES:.....	23
IV.2.4. CODAGE ELECTRIQUE D'UNE INFORMATION DIGITALE:.....	23
IV.2.4.1. CODAGE EN BANDE DE BASE:.....	23
IV.2.4.2. TRANSMISSION PAR MODULATION:.....	24
IV.2.4.3. NOTION DE MODEM:.....	24
IV.2.4.4. EXEMPLE DE MODULATION: VOTRE LIAISON A.D.S.L.:.....	25
IV.2.5. ACHEMINEMENT D'UNE TRAME VERS UN DESTINATAIRE SUR UN RESEAU:.....	25
IV.2.5.1. POSITION DU PROBLEME:.....	25
IV.2.5.2. MECANISME D'ADRESSAGE PHYSIQUE:.....	26
IV.2.5.3. LIMITES DU MECANISME D'ADRESSAGE PHYSIQUE:.....	27
V.MECANISMES DES SERVICES RESEAU:.....	28
V.1. NOTION DE PROCESSUS LOGICIEL:.....	28
V.2. NOTION DE PROTOCOLE DE COMMUNICATION:.....	28
V.3. ADRESSAGE DES MACHINES HOTES D'UNE INTERCONNEXION DE RESEAUX:.....	28
V.4. ADRESSAGE DES PROCESSUS SUR UNE INTERCONNEXION DE RESEAUX:.....	28
V.5. LA SUITE DE PROTOCOLES TCP-IP:.....	29
V.5.1. GENERALITES:.....	29
V.5.2. ADRESSAGE D'UN PROCESSUS SOUS TCP-IP:.....	29
V.5.3. PROTOCOLE TCP ET MECANIQUE DE TRANSMISSION " PAR PAQUETS " :.....	31
V.5.3.1. JUSTIFICATION:.....	31
V.5.3.2. PRINCIPES DE FONCTIONNEMENT:.....	31
V.5.3.3. MECANISMES DE SECURISATION:.....	32
V.5.4. PROTOCOLE IP ET ROUTAGE INTER-RESEAUX:.....	32
V.5.5. TCP-IP ET LE MODELE CLIENT-SERVEUR:.....	33
V.5.5.1. RAPPELS SUR LE MODELE CLIENT-SERVEUR ET CONSEQUENCES:.....	33
V.5.5.2. CAS DE TCP-IP ET DU WEB:.....	34
V.5.6. ATTRIBUTION DES ADRESSES I.P. AUX HOTES D'INTERNET:.....	34
V.5.6.1. ATTRIBUTION DES ADRESSES IP AUX POSTES CLIENTS:.....	34
V.5.6.2. ATTRIBUTION DES ADRESSES IP AUX POSTES SERVEURS:.....	34
VI.MECANISMES DES SERVICES APPLICATIFS:.....	35
VI.1. INTRODUCTION:.....	35
VI.2. MECANISMES DE CONTRÔLE DE SESSION:.....	35
VI.3. MECANISMES DE PRESENTATION DES DONNEES:.....	35
VI.4. SYSTEME DES NOMS DE DOMAINES:.....	35
VI.4.1. INTRODUCTION:.....	35
VI.4.2. PRINCIPES GENERAUX DU SYSTEME DES NOMS DE DOMAINES:.....	36
VI.4.3. GESTION DES NOMS DE DOMAINES:.....	37
VI.4.4. RESOLUTION DES NOMS DE DOMAINES:.....	37
VI.5. LES U.R.L.:.....	38
VI.5.1. DEFINITION:.....	38
VI.5.2. STRUCTURE D'UNE U.R.L.:.....	38
VI.5.3. UTILISATION DES U.R.L. CONCEPT D'HYPERLIEN:.....	39
VII.LES SERVICES DU NIVEAU APPLICATIF:.....	41
VII.1. NAVIGATION SUR LE WEB.....	41
VII.1.1. DEFINITIONS:.....	41
VII.1.2. FONCTIONNEMENT DE BASE.....	41
VII.1.3. STRUCTURE ET CONTENU D'UNE PAGE WEB:.....	42
VII.1.3.1. LANGAGE HTML:.....	42

SUJET:
COMPRENDRE
INTERNET

REFERENCE: ATLANTIC/DOCTECH/
Initiation à l'informatique/Comprendre Internet
Mise à jour: 11/01/11

VII.1.3.2.STRUCTURE GENERALE D'UNE PAGE WEB RECUE PAR UN NAVIGATEUR:.....	44
VII.1.4.LE PROTOCOLE HTTP:.....	45
VII.1.5.NOTION DE SITE WEB:.....	45
VII.1.6.NOTION DE MOTEUR DE RECHERCHE:.....	46
VII.1.7.NOTION DE FAVORIS:.....	46
VII.1.8.NOTION DE COOKIE:.....	46
VII.2.MESSAGERIE ELECTRONIQUE:.....	48
VII.2.1.DEFINITION:.....	48
VII.2.2.MECANISME DE BASE:.....	48
VII.2.2.1.MAIL TRANSFER AGENT ET BOITE AUX LETTRES ELECTRONIQUES:.....	48
VII.2.2.2.RECEPTION D'UN MAIL PAR UN M.T.A.:.....	48
VII.2.2.3.EMISSION D'UN MAIL VERS SON DESTINATAIRE PAR UN M.T.A.:.....	48
VII.2.3.MODALITES D'ACCES D'UN UTILISATEUR AU COURRIER ELECTRONIQUE:.....	48
VII.2.3.1.INTRODUCTION:.....	48
VII.2.3.2.ACCES PAR UN CLIENT DE MESSAGERIE LOCAL:.....	49
VII.2.3.3.ACCES PAR UN WEBMAIL:.....	50
VIII.CONCLUSION:.....	52

SUJET: COMPRENDRE INTERNET	REFERENCE: ATLANTIC/DOCTECH/ Initiation à l'informatique/Comprendre Internet Mise à jour: 11/01/11
---	---

I.INTRODUCTION:

I.1.AVANT PROPOS:

L'accès à internet vient en tête des motivations des particuliers lorsqu'ils s'équipent en matériel informatique. Or, toute personne travaillant dans le domaine de l'assistance informatique peut aisément constater que la plupart des problèmes rencontrés par les utilisateurs non spécialistes ont trait à l'accès à internet, qu'il s'agisse de l'installation des matériels et logiciels, de la maintenance des liaisons physiques et logiques ou même de l'utilisation rationnelle des outils et fonctions associés.

L'utilisation d'internet devient de plus en plus nécessaire, non seulement dans la vie professionnelle, mais également dans beaucoup d'activités courantes: qu'il s'agisse de consulter un compte en banque, d'acquérir des titres de transport, de rechercher du travail, de s'inscrire à une formation, et pour bien d'autres activités incontournables, les utilisateurs sont fortement incités à recourir aux accès en ligne. Déjà, beaucoup d'entreprises ont totalement dématérialisé un grand nombre d'activités et il n'est pas déraisonnable de penser qu'à moyen terme ce sera le cas pour de nombreuses démarches de la vie courantes, même si l'on peut déplorer la diminution du lien social qui en résulte parfois.

Une maîtrise minimale de la structure et du fonctionnement d'internet permet à un utilisateur d'éviter la grande majorité des dysfonctionnements, ou d'y remédier par lui-même, économisant ainsi des interventions coûteuses ou des interruptions de services souvent très pénalisantes. Cet ouvrage vise donc principalement à former des **utilisateurs avertis et autonomes**. Son contenu peut également constituer une bonne base de départ pour des personnes débutant dans un cycle de formation d'administrateur ou de développeur WEB.

REMARQUE:

Le lecteur désireux d'acquérir des connaissances plus approfondies pourra avec profit consulter les divers ouvrages mis en ligne par le site ATLANTIC.

I.2.BREF RAPPEL HISTORIQUE:

INTERNET est un concept apparu aux Etats Unis vers le milieu des années 1960. A l'origine, il s'agissait de connecter entre eux des **réseaux d'universités** ou de **centres de recherche** (surtout militaires) par l'intermédiaire d'une infrastructure de communication à haut débit. La destination première du WEB était donc l'interconnexion de **RÉSEAUX LOCAUX** (local networks) et non de particuliers, à des fins non commerciales (recherche, défense, etc). Cette origine étatique, scientifique et non commerciale d'internet est encore sensible dans certaines de ses caractéristiques, comme la gratuité de nombreux services ou un certain parti-pris "libertaire".

La première infrastructure matérialisant ce projet, désignée sous l'acronyme **ARPANet** (Advanced **R**esearch **P**rojects **A**gency **N**etwork) a été développée par le D.A.R.P.A, organisme du Département de la Défense des Etats Unis. Lorsqu'elle est entrée en service en 1969, elle interconnectait une dizaine d'universités et de centres de recherche des Etats Unis.

A partir de 1974, **ARPANet** adopte le protocole de transport de données **TCP-IP**. Ce choix, qui permet une standardisation des mécanismes de communication et des matériels, favorisera beaucoup l'extension d'**INTERNET** au niveau mondial.

A la fin des années 1970, les militaires américains abandonnent aux civils une partie des infrastructures supportant **INTERNET**. Cette décision va permettre l'essor de nouvelles activités sur internet: a partir du début des années 1980 des infrastructures d'interconnexion publiques et privées à destination commerciale, se développent dans la plupart des pays développés. Comme l'ARPANet d'origine, ces infrastructures (réseaux d'interconnexion) sont constitués de lignes à très haut débit, appelées **BACKBONES** en anglais ou **DORSALES** en français.

En France, la dorsale **RENATER** (**RE**seau **N**ational de Télécommunication pour la **T**echnologie, l'**E**nseignement et la **R**echerche), mise en service en 1993, a constitué l'équivalent de l'ARPANet d'origine. D'autres dorsales, à destination

SUJET: COMPRENDRE INTERNET	REFERENCE: ATLANTIC/DOCTECH/ Initiation à l'informatique/Comprendre Internet Mise à jour: 11/01/11
---	---

commerciale, sont également mise en service dans les années 1980-1990 (COLT, LDCOM, etc.). Ces infrastructures, reliées aux dorsales des autres pays, vont progressivement constituer ce que l'on appelle (avec un brin de grandiloquence) le WORLD WIDE WEB, le réseau mondial.

Ces nouvelles infrastructures permettent la mise en réseau de nombreuses entreprises publiques et privées, en particulier dans le domaine des services (banques, assurances, administrations). Cependant, les DORSALES étant constituées d'équipements extrêmement coûteux à l'acquisition, à l'installation et à l'entretien (fibres optiques à haut débit, faisceaux hertziens, liaisons satellitaires...), le prix d'une connexion directe par une "ligne dédiée" (installation et location) est très élevé. De ce fait, jusqu'au milieu des années 1990, l'accès à INTERNET reste réservé aux grandes ou moyennes structures.

Cependant, une solution de connexion bien moins onéreuse était disponible depuis de nombreuses années. En effet, le réseau téléphonique était utilisé depuis le premier tiers du vingtième siècle pour transmettre des données, au moyen d'équipements appelés MODEM. (MODulateurs-DEModulateur). Cette technologie était employée notamment pour alimenter les téléscripteurs (à partir du milieu des années 1930), ou pour réaliser des liaisons distantes entre systèmes informatiques ou entre ordinateurs et consoles de télé-traitement. Le réseau MINITEL Français, basé sur cette infrastructure, était disponible pour les particuliers depuis le début des années 1980.

A partir du milieu des années 1990, de nombreuses sociétés (en particulier, les opérateurs téléphoniques gestionnaires des lignes, comme FRANCE TELECOM), ont donc commencé à proposer des accès pour le public aux dorsales INTERNET par l'intermédiaire du réseau téléphonique. Dès cet instant, le coût d'un équipement d'accès à internet a cessé d'être prohibitif pour des particuliers. Cependant, le débit proposé était très faible (de 9,6 à 56 kbits/s), la tarification à la durée ne permettait pas la connexion illimitée et, de plus, la ligne téléphonique vocale était indisponible pendant la durée de la connexion à internet.

A la fin des années 1990, la technologie **ADSL** est diffusée au grand public. Utilisant toujours la connexion au réseau téléphonique local (boucle locale), mais opérant une séparation des flux vocaux et numériques, elle permet d'atteindre des débits beaucoup plus importants (plusieurs méga-bits/s) et de bénéficier d'une connexion illimitée pour une tarification forfaitaire assez faible tout en laissant disponible la ligne téléphonique. D'autres solutions (câble, réseau EDF, etc.) apparaissent également. Cependant, ADSL est de loin la plus répandue, à cause de sa simplicité, de son faible coût et de ses performances.

Actuellement, la plupart des Fournisseurs d'Accès Internet (F.A.I ou PROVIDERS en anglais) ont développé leur propre dorsale. Les différents BACKBONES sont reliés entre eux. En France, ces points d'interconnexion sont souvent situés dans la région Lyonnaise ou la région Parisienne.

I.3.PROGRESSION ADOPTEE POUR CET OUVRAGE:

Cet ouvrage vise à former ce que l'on pourrait appeler des **utilisateurs éclairés** d'internet, capables de bien **cerner leurs besoins** en relation avec cet outil (notamment en termes de définition de leur poste informatique), de bien **tirer partie des fonctionnalités** offertes par les services disponibles sur la "toile" et de **se comporter d'une manière rationnelle et efficace** face aux menaces et aux dysfonctionnements les plus courants.

Pour acquérir ce type de comportement, il n'est pas forcément nécessaire de connaître d'une manière exhaustive et détaillée les mécanismes mis en oeuvre par internet. Il est souvent possible de se contenter d'une approche plus globale et synthétique, sans tomber dans la simplification à outrance.

C'est pour cette raison que la démarche adoptée dans cet ouvrage n'est pas une étude systématique, mais plutôt une succession d'approches "concentriques" à partir du point de vue global de l'utilisateur.

I.4.REMARQUE:

Les lecteurs qui désireraient acquérir une connaissance plus approfondie et plus rigoureuse du domaine peuvent consulter l'ouvrage "Réseaux Informatiques", rubrique DOCUMENTATION EN LIGNE, sous-rubrique RESEAUX du site ATLANTIC.

II.POINT DE VUE DE L'USAGER:

II.1.GENERALITES:

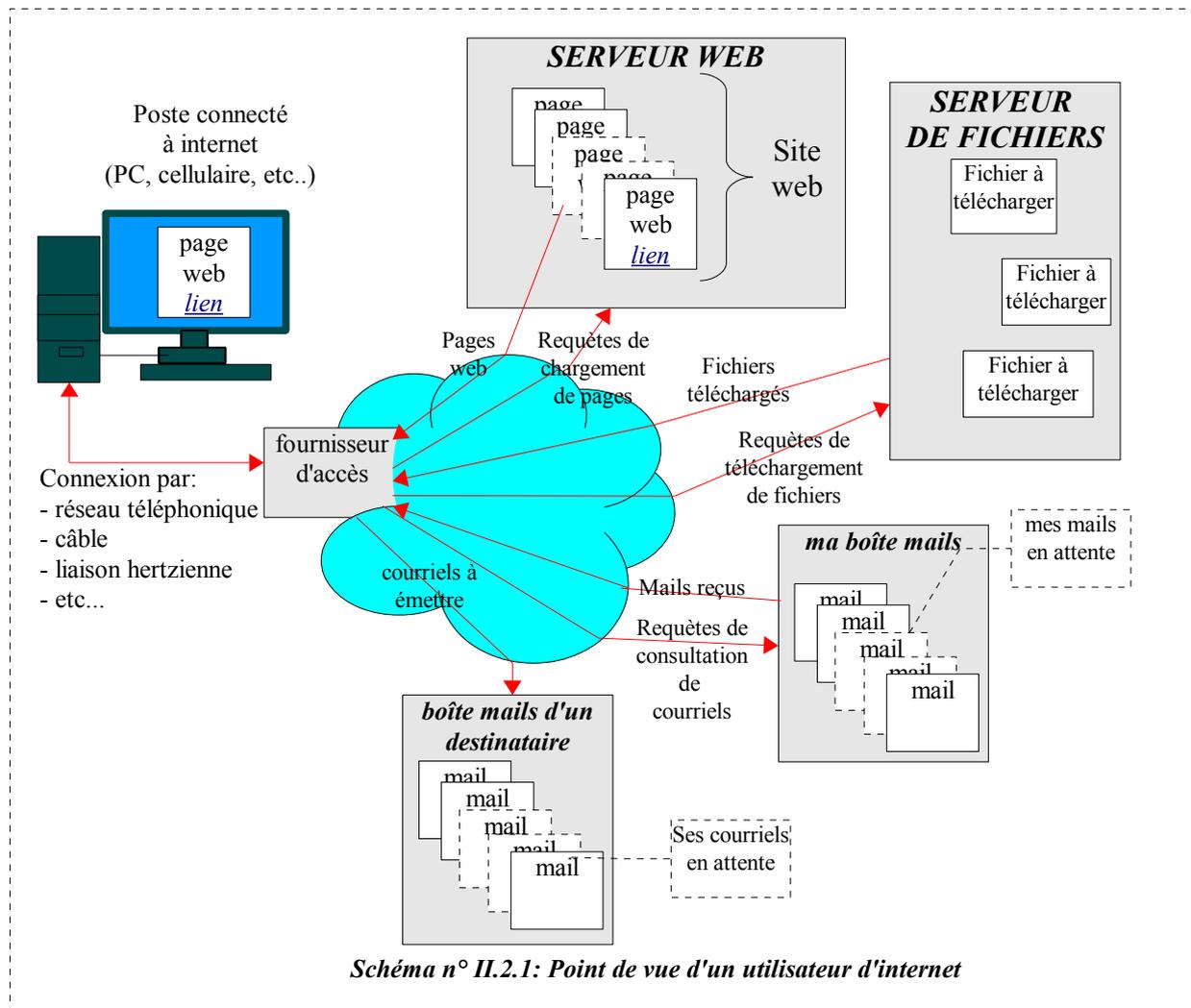
L'utilisateur de base d'internet est celui qui se contente d'utiliser les deux services principaux que le réseau met à sa disposition:

- La " navigation sur le net ", c'est-à-dire la consultation de " sites web ".
- Le courrier électronique (c'est-à-dire l'échange de " mails " ou " courriels ").

Ces usagers, qui sont évidemment les plus nombreux, se soucient peu des détails de l'infrastructure internet et des mécanismes qu'elle met en jeu. Il leur suffit de connaître un certain nombre de concepts et de mécanismes globaux que l'on peut modéliser de la manière suivante:

II.2.STRUCTURE GLOBALE DU WEB:

II.2.1.SHEMA DE PRINCIPE:



SUJET: COMPRENDRE INTERNET	REFERENCE: ATLANTIC/DOCTECH/ Initiation à l'informatique/Comprendre Internet Mise à jour: 11/01/11
---	---

II.2.2.LE FOURNISSEUR D'ACCES:

L'utilisateur a conscience que son accès à l'infrastructure internet s'effectue par l'intermédiaire d'un **FOURNISSEUR D'ACCES**, auquel il doit se **CONNECTER** ou se **DÉCONNECTER** suivant ses besoins. C'est ce fournisseur d'accès qui lui permet d'accéder aux infrastructures de communication d'internet (les DORSALES ou BACKBONES évoquées dans le chapitre précédent), par l'intermédiaire du réseau téléphonique, du câble ou d'autres moyens de raccordements disponibles.

II.2.3.LA NAVIGATION SUR INTERNET:

Pour l'utilisateur de base, le fait de " naviguer sur le web " peut se résumer à afficher successivement des " pages web " sur son navigateur.

La navigation entre les pages web s'effectue en activant dans la fenêtre d'affichage de son navigateur (principalement avec la souris) des objets graphiques: les " liens " ou les " champs de formulaires ". Ces actions déclenchent l'envoi vers des entités appelées **SERVEURS WEB** des requêtes d'envoi de pages web appartenant à des sites administrés par le serveur. Le serveur répond à ces requêtes par l'envoi de la page web demandée.

Un **SITE WEB** est donc vu par l'utilisateur comme une collection de pages web organisées et présentées suivant la logique interne et la nature de ce site.

Pour accéder à un site web particulier, l'utilisateur doit pouvoir attendre ce site à partir de son **NAVIGATEUR**, qui n'est autre qu'un programme informatique supporté par son terminal (PC fixe ou mobile, téléphone cellulaire, etc.). Il effectue cette opération en utilisant un identifiant complexe appelé **UNIFORM RESOURCE LOCATOR (U.R.L)** en anglais, ce qui peut se traduire en français par " localisateur de ressource uniformisé " (ex: www.atlantic-83.fr).

L'U.R.L est en fait un moyen de localiser n'importe quelle **RESSOURCE** sur le net, et en particulier des fichiers: (l'URL d'un site " pointe " en fait sur le fichier contenant la première page web du site).

II.2.4.LE COURRIER ELECTRONIQUE:

Pour recevoir un mail, l'utilisateur doit accéder au contenu de sa **BOÎTE DE MESSAGERIE**. Celle-ci est donc vue comme une entité permettant de stocker les mails adressés à l'utilisateur. Celui-ci émet vers sa boîte aux lettres une requête de téléchargement des mails en attente. En réponse, le gestionnaire de la boîte aux lettres émet vers l'utilisateur les mails en attente.

Lorsque l'utilisateur envoie un mail, il a conscience qu'il l'adresse à la **BOÎTE DE MESSAGERIE DU DESTINATAIRE** et non au poste informatique de ce dernier. Son mail sera stocké en attente dans cette boîte jusqu'à ce que le destinataire y accède.

L'utilisateur localise la boîte mail d'un destinataire sur le réseau par son **ADRESSE MAIL** (qui est de la forme <id.destinataire>@<id. serveur messagerie destinataire>, par exemple: webmestre@atlantic-83.fr).

NOTA:

Pour émettre ou recevoir des mails, l'utilisateur dispose de deux solutions:

- Soit il utilise un logiciel **CLIENT DE MESSAGERIE** (ex: outlook), hébergé par son poste de travail.
- Soit il se connecte à des sites particuliers, appelés **WEBMAIL** qui se chargent de lui présenter les fichiers en attente dans sa boîte aux lettres.
- Les **WEBMAIL** permettent aussi d'envoyer des mails vers la boîte aux lettres des destinataires.

II.2.5.LE TELECHARGEMENT DE FICHIERS:

La plupart du temps, le téléchargement de fichiers est effectué par l'utilisateur en activant un lien ou un bouton de formulaire sur une page web. Cette fonction ne lui paraît donc pas réellement distincte de la " navigation ". Il s'agit pourtant d'un troisième service offert par le réseau, supporté par des entités appelées **SERVEURS DE FICHIERS**.

II.3.LA COMMUNICATION SUR INTERNET:

II.3.1.LES ACTEURS DE LA COMMUNICATION:

Pour communiquer sur internet, vous êtes toujours obligé d'utiliser un programme s'exécutant sur votre terminal informatique (PC, téléphone cellulaire, etc.). Ce logiciel est:

- Un **NAVIGATEUR**, si vous naviguez sur le web (y compris lorsque vous prenez connaissance de votre messagerie par l'intermédiaire d'un site WEBMAIL comme GMAIL).
- Un **CLIENT DE MESSAGERIE** (OutLook, Lotus, etc.) si vous voulez accéder directement à votre serveur de messagerie sans passer par le web.
- Un **CLIENT DE TELECHARGEMENT DE FICHERS** (FileZilla, etc.) si vous voulez télécharger des fichiers sans passer par les fonctionnalités d'un site web.

D'autre part, c'est une **machine** qui vous répond, ou plus précisément un **logiciel** hébergé dans un type de machine particulier appelé **SERVEUR**:

- Lorsque vous naviguez dans un **SITE WEB**, qui n'est autre qu'un programme informatique hébergé par un serveur particulier appelé **SERVEUR WEB**, c'est le logiciel du serveur qui répond à vos requêtes par l'envoi des pages web que vous avez sélectionnées en cliquant sur des liens ou sur des formulaires..
- Lorsque vous accédez à votre boîte de messagerie, c'est un logiciel particulier, appelé **MAIN TRANSFER AGENT (M.T.A)**, situé sur un **SERVEUR DE MESSAGERIE** hébergeant votre boîte aux lettres qui satisfait votre requête en vous communiquant les messages en attente. De même, lorsque vous envoyez un message à un correspondant, c'est en fait au **M.T.A** de votre correspondant que vous l'adressez: ce dernier n'accédera au message que lorsqu'il se connectera à celui-ci.

Dans nombre de domaines (transactions boursières, surveillance du trafic à buts policiers ou antiterroristes, contrôle automatique à distance, etc), des logiciels informatiques communiquent sur internet sans intervention humaine en temps réel.

De ce fait, bien que certaines applications comme les messageries instantanées (chats, forums, téléphonie sur internet etc..), donnent l'impression d'une communication écrite, orale ou visuelle **en temps réel** entre correspondants humains, la communication sur internet se produit toujours **entre** ou **par l'intermédiaire** de deux **programmes** s'exécutant sur des machines distantes.

II.3.2.LE MODELE CLIENT-SERVEUR:

Une autre caractéristique de la communication sur internet est que les rôles des deux programmes communicants sont asymétriques:

Par exemple, dans le cas de la communication d'un utilisateur (ou plutôt, du navigateur de l'utilisateur) avec un site web:

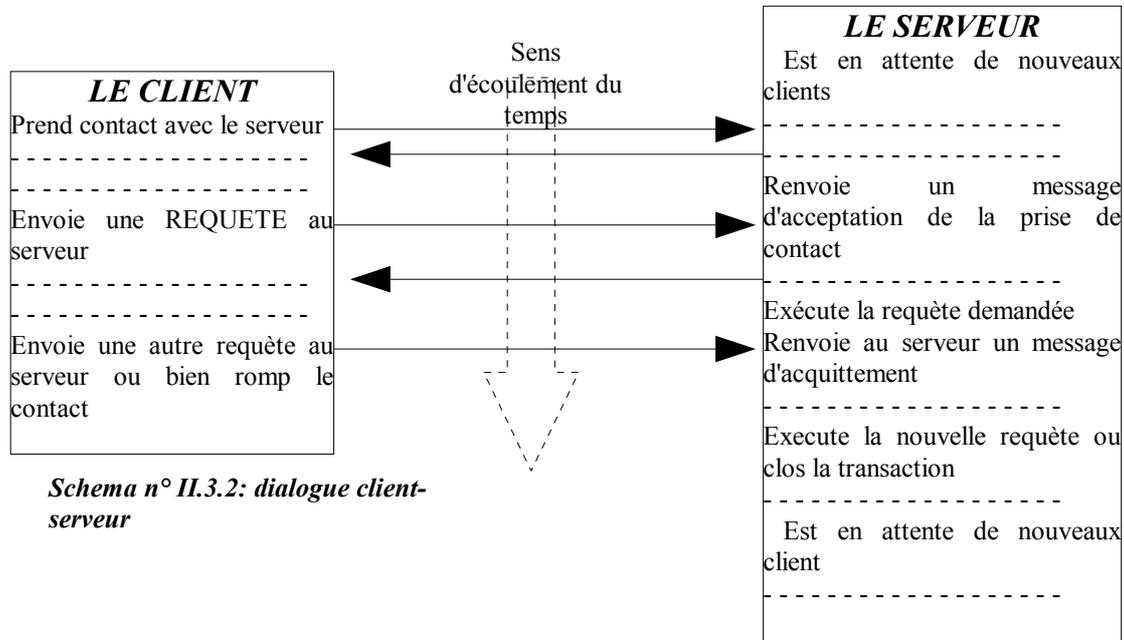
- Le programme de l'utilisateur (navigateur) n'est connecté qu'à un site à la fois. A l'inverse, le site web peut être connecté à un grand nombre d'utilisateurs à la fois.
- Le programme gérant le site doit rester en permanence à l'écoute des demandes éventuelles des utilisateurs, alors que ceux-ci peuvent apparaître et disparaître à tout moment.
- Le navigateur de l'usager adresse au site des demandes d'envoi de pages web (en cliquant sur des liens ou des formulaires). Le site web répond en adressant à l'utilisateur le contenu des pages web correspondant aux demandes.

Nous pouvons voir que cette forme de transaction asymétrique présente beaucoup d'analogies avec celles qui se déroulent entre des **CLIENTS** s'adressant à un **SERVEUR** (dans un bar, un restaurant, un magasin...): chaque client attend du serveur la fourniture d'un produit ou d'un service bien défini. Il n'intervient en aucune manière dans sa réalisation. C'est le serveur qui assure cette réalisation en employant les méthodes et procédés qui lui sont propres, et que le client ne désire pas forcément connaître.

De ce fait, les machines hébergeant les sites web sont appelées **SERVEURS WEB**. Ces machines sont équipées de logiciels **SERVEURS** (logiciel serveur **APACHE**, par exemple), gérant les différents sites hébergés. Les navigateurs sont

appelés **CLIENTS WEB**. Nous verrons plus tard que le même modèle est utilisé pour la messagerie électronique (mail).

Le schéma suivant décrit le fonctionnement d'une transaction Client-Serveur:



Schema n° II.3.2: dialogue client-serveur

REMARQUE:

Cette orientation vers un **dialogue de type CLIENTS-SERVEUR** a des conséquences techniques importantes: en effet, le plus souvent, la requête adressée par le client est beaucoup plus courte, en terme de quantité d'informations que la réponse du serveur (Par analogie: il est plus court de **commander une pizza**, c'est-à-dire transmettre une requête au pizaiolo, que de la **confectionner**, ce qui constitue une partie de la réponse à cette requête). De ce fait, **la transmission du client vers le serveur exige le plus souvent un débit beaucoup plus faible que la transmission du serveur au client.**

II.3.3.CONCLUSION:

D'un point de vue technique, on peut considérer qu'une communication sur INTERNET se fait toujours **entre deux PROGRAMMES INFORMATIQUES** s'exécutant dans des **MACHINES DISTANTES**.

Cette communication prend la plupart du temps la forme d'un dialogue entre un programme **SERVEUR** et un programme **CLIENT**: le serveur est capable de répondre aux **REQUETES** que le client lui adresse en lui fournissant un **SERVICE** (qui peut consister en l'envoi de données), puis renvoie au client un **MESSAGE D'ACQUITTEMENT** (compte-rendu de l'échange).

SUJET: COMPRENDRE INTERNET	REFERENCE: ATLANTIC/DOCTECH/ Initiation à l'informatique/Comprendre Internet Mise à jour: 11/01/11
---	---

III.MECANISMES DE COMMUNICATION SUR LE RESEAU:

III.1.INTRODUCTION:

Nous avons vu précédemment que dans sa pratique courante, un utilisateur d'internet peut se contenter de connaître un petit nombre de concepts d'une manière très globale et synthétique (sites web, boîtes de messagerie, fournisseur d'accès, etc.), sans entrer dans la complexité de ceux-ci. Il possède, en quelque sorte, une connaissance " en surface " de l' " océan " de composants et de mécanismes qui supportent le web.

Avant de pousser plus avant dans ces profondeurs, et pour éviter de nous y perdre, nous allons hiérarchiser les nombreux concepts qui les occupent plusieurs " niveaux ", dont voici la description:

III.2.LES DIFFERENTS NIVEAUX DE LA COMMUNICATION:

III.2.1.TRANSARENCE DES MECANISMES DE COMMUNICATION POUR L'USAGER:

Nous avons vu précédemment que lorsque nous utilisons internet, nous le faisons toujours par l'intermédiaire d'une application logicielle (c'est-à-dire d'un programme: un navigateur, un logiciel de messagerie, etc.). Ces **APPLICATIONS RESEAUX** nous permettent d'accéder à un certain nombre de services disponibles en ignorant totalement les mécanismes logiciels et les infrastructures matérielles que ces services utilisent: pour naviguer sur internet, il nous suffit la plupart du temps de savoir utiliser l'interface du navigateur que nous utilisons.

EXEMPLE:

Certains logiciels navigateurs internet (Mozilla firefox, Safari, Opera, etc.) possèdent des versions adaptées à plusieurs systèmes d'exploitation (Window, Linux, systèmes Unix divers, etc.). Un utilisateur pourra les utiliser de la même manière quel que soit le système d'exploitation et quelle que soit l'infrastructure matérielle (ordinateur, connexion réseau) dont il dispose.

En fait, les spécificités du poste de travail sont rendues **TRANSPARENTES** pour l'utilisateur par l'action de divers mécanismes inclus dans le " software " et le " hardware " de la machine. Cette transparence est obtenue par l'organisation de ces mécanismes en plusieurs niveaux dont nous allons décrire le principe.

III.2.2.LA RONDE DES "DEMONS":

Dans le domaine des systèmes d'exploitation du type UNIX/LINUX, le terme DAEMON (démon en français) est employé pour désigner des processus logiciels chargés de gérer les tâches dévolues au fonctionnement de base des systèmes informatique, et en particulier les tâches d'entrée-sortie. Ces tâches s'effectuant **en arrière plan** de l'activité visible de l'ordinateur, les daemons agissent un peu comme une armée de serviteurs invisibles.

Tous les systèmes informatiques utilisent des équivalents des daemons : sous Windows, ils sont appelés "services", cependant, l'appellation daemon me semble beaucoup plus imagée et évocatrice : qui d'entre nous n'a pas été un jour intimement persuadé (surtout dans des moments de grand désarroi technologique), que son ordinateur est animé par une armée de petits démons plus ou moins malfaisants ? (ne niez pas: j'ai même vu certains d'entre vous leur parler...). Ce point de vue n'est d'ailleurs pas sans avantage psychologique, car il constitue un moyen pratique et peu onéreux de détourner nos colères et frustrations " informatiques " vers des boucs émissaires aussi disponibles qu'endurants.

Pour humaniser et imagier un peu l'univers très abstrait des mécanismes de communication, j'ai pris le risque d'évoquer ces petits démons aussi discrets qu'industriels. Je vous demande de supposer qu'ils se mettent en action chaque fois que nous activons l'interface graphique de notre application réseau (en cliquant sur un lien, en validant un formulaire, en utilisant un menu...).

Ces petits démons (qui, je le rappelle avant d'être convaincu de sorcellerie, ne sont autres que les programmes informatiques ou les mécanismes de logique électronique installés sur le poste de travail ou les organes de liaison), se chargent de traduire les **commandes** saisie sur votre interface en **actions** correspondant à la fonctionnalité activée. La plupart de ces commandes aboutissent à l'envoi de messages à destination d'un hôte éloigné du réseau (serveur web,

SUJET: COMPRENDRE INTERNET	REFERENCE: ATLANTIC/DOCTECH/ Initiation à l'informatique/Comprendre Internet Mise à jour: 11/01/11
---	---

serveur de messagerie, etc.), qu'il s'agisse de l'envoi de **REQUETES** destinées à obtenir un service quelconque de la part du destinataire ou de la transmission de **DONNEES** nécessaires au fonctionnement de celui-ci.

Nos démons uniront donc leurs efforts pour réaliser les tâches suivantes:

- **E1:** Interpréter la commande passée par l'utilisateur, en déduire les actions à effectuer, et en particulier les messages à transmettre, puis élaborer ces messages.
- **E2:** Conditionner ces messages pour la transmission: les morceler en segments de longueur acceptable, y inclure l'adresse de la **machine destinataire** et du **programme destinataire** dans cette machine (le message est-il destiné au navigateur? au programme de messagerie électronique?), etc.
- **E3:** Transformer ces segments de données en signaux physiques capables de se propager sur le média du réseau, puis injecter ces signaux sur le média.

Remarquons que les tâches E1, E2 et E3 doivent se succéder pour aboutir à l'émission d'un message vers un destinataire, qui est un programme hébergé par une machine distante connectée au réseau.

De même, lorsqu'un message destiné à votre application arrive sur la connexion physique de votre poste de travail, ces mêmes démons infatigables tenteront l'impossible pour mettre à votre disposition les informations qu'il contient. Ils devront, pour cela:

- **R3:** Examiner les signaux circulants sur le média, en extraire ceux qui sont destinées à la machine locale et en vérifier le contenu.
- **R2:** Transformer ces signaux en segments de messages, détecter le programme réseau de la machine qui en est destinataire (est-ce le navigateur ou le client de messagerie? les deux peuvent fonctionner en même temps), ré-assembler ces segments pour reconstituer le message complet.
- **R1:** Enfin, interpréter le message pour accomplir l'action à laquelle il est destiné (affichage d'une page web sur le navigateur ou d'une liste de courriels sur la fenêtre du client de messagerie, enregistrement des données reçues dans un fichier local, etc.).

Remarquons que les tâches R3, R2 et R1 doivent se succéder pour aboutir à la réception et au traitement d'un message par une des applications réseaux lancées sur la machine.

Pour poursuivre notre analogie, nous dirons que l'activité d'une telle " armée " nécessiterait obligatoirement une très forte organisation. En particulier, il pourrait sembler judicieux de spécialiser chaque lutin dans certaines des activités que nous avons énumérées ci-dessus. Par exemple:

- Certains d'entre eux seraient spécialisés dans les activités E1 et R1.
- D'autres seraient spécialisés dans les activités E2 et R2.
- D'autres enfin seraient spécialisés dans les activités E3 et R3.

III.2.3.LES TROIS NIVEAUX DE MECANISMES:

Si nous avons gardé en tête que nos démons symbolisent en fait des mécanismes matériels et logiciels intervenants dans la communication réseau, les trois spécialisations ainsi définies correspondent à trois niveaux de mécanismes que l'on peut représenter par le schéma suivant:



COMMENTAIRES SUR LE SCHEMA:

Les **UTILISATEURS** ont accès aux fonctionnalités offertes par le réseau en activant les **SERVICES APPLICATIFS**. Ils le font par l'intermédiaire d'applications informatiques. Pour les utilisateurs basiques d'internet, ces applications seront essentiellement les Navigateurs (internet explorer, Mozilla, Opera, etc..), les Clients de Messagerie (outlook, lotus, etc..) ou les programmes de transfert de fichiers (Filezilla...).

Les mécanismes du niveau des **SERVICES APPLICATIF** ont pour rôle:

- Soit d'interpréter les commandes passées par l'utilisateur, d'en déduire les actions à effectuer, et en particulier les messages à transmettre, puis d'élaborer ces messages.
- Soit d'interpréter un message reçu du réseau afin accomplir l'action à laquelle il est destiné (affichage d'une page web sur le navigateur ou d'une liste de courriels sur le client de messagerie, enregistrement des données dans un fichier local, etc.).

Pour réaliser les fonctions activées par les utilisateurs, les applications réseaux ont besoin d'échanger des messages avec d'autres utilisateurs du réseau (sites web, serveurs de messagerie, etc). Pour ce faire, ils utilisent les mécanismes du niveau des **SERVICES RESEAU**, qui ont donc pour rôle:

- Soit de conditionner les messages élaborés par le niveau des services applicatifs pour la transmission (par exemple, les morceler en segments de longueur acceptable, y inclure l'adresse de la machine destinataire et du programme destinataire dans cette machine, etc.).
- Soit de transformer les données acquises par le niveau gestion du lien physique en segments de messages, détecter le programme réseau qui en est le destinataire, puis ré-assembler ces segments pour reconstituer le message initial.

Cependant, les mécanismes du niveau services réseaux n'ont pas connaissance de la structure matérielle de la machine, ni des mécanismes qui permettent de transformer les messages en **signaux physiques** injectables sur le média du réseau. Ce niveau fera donc appel aux mécanismes du niveau **GESTION DU LIEN PHYSIQUE** pour réaliser ces tâches. Ce niveau aura donc pour rôle:

- Soit de transformer les segments fournis par le niveau des services réseaux en signaux physiques capables de se propager sur le média, puis d'envoyer ces signaux vers le nœud adjacent du réseau qui constitue la première étape de la route à suivre pour atteindre le destinataire.
- Soit de récupérer les signaux circulants sur le média et destinées à l'utilisateur de la machine, d'en vérifier le contenu, puis de les transmettre au niveau des services réseau.

SUJET:
COMPRENDRE
INTERNET

REFERENCE: ATLANTIC/DOCTECH/
Initiation à l'informatique/Comprendre Internet
Mise à jour: 11/01/11

III.3.LE MODELE O.S.I:

L'I.S.O (**International Standard Organisation**), a diffusé dès les années 1980 un modèle d'organisation pour les " systèmes d'interconnexion ouverts ". Celui-ci, appelé **Modèle O.S.I** (pour Open System Interconnexion) hiérarchise les mécanismes de communication en sept **COUCHES** (dites " fonctionnelles ", car elles décrivent les mécanismes indépendamment de la manière dont ils sont réalisés matériellement). Ces couches peuvent être succinctement définies comme suit:

<i>NUMERO DE COUCHE</i>	<i>NOM</i>	<i>FONCTION</i>
7	APPLICATION	Mécanismes supportant les fonctions externes des applications.
6	PRESENTATION	Résolution des problèmes liés aux différences de codage des données entre systèmes connectés.
5	SESSION	Synchronisation des flux de données échangés, points de reprise en cas d'erreurs, etc.
4	TRANSPORT	Conditionnement des messages pour le transport: découpage des messages en " paquets de données " de longueur compatible avec une bonne transmission et ré-assemblage à la réception.

<i>NUMERO DE COUCHE</i>	<i>NOM</i>	<i>FONCTION</i>
3	RESEAU	Routage de " paquets de données " vers leur destinataire à travers des réseaux interconnectés (pouvant être de technologie hétérogène).
2	LIAISON	<ul style="list-style-type: none"> • Sous-couche LLC: établissement, entretien et contrôle des liaisons de données entre deux nœuds " adjacents " du réseau, • Sous couche MAC: contrôle de l'accès au média de transmission..
1	PHYSIQUE	Transformation des données en signaux physiques inversement.

REMARQUES:

- L.L.C. = Logical Link Control.
- M.A.C. = Media Acces Control.

Nous pouvons voir que la décomposition effectuée par le modèle O.S.I. est plus détaillé que la décomposition en trois niveaux présentée plus haut. Nous pouvons établir la correspondance suivante:

<i>NIVEAU</i>	<i>CORRESPONDANCE O.S.I</i>	<i>PRODUITS R ESEAUX CORRESPONDANTS</i>
GESTION DU LIEN PHYSIQUE	Couches OSI 1 et 2 (physique, liaison)	Cartes réseaux, connectique, médias, pilotes de périphériques
SERVICES RESEAU	Couches OSI 3 et 4 (réseau, routage)	Logiciels contrôleurs de réseaux (TCP-IP pour internet)
SERVICES APPLICATIFS	Couches OSI 5, 6 et 7 (session, présentation, application)	Applications réseaux (navigateurs, clients de messagerie, etc.)

SUJET:
COMPRENDRE
INTERNET

REFERENCE: ATLANTIC/DOCTECH/
Initiation à l'informatique/Comprendre Internet
Mise à jour: 11/01/11

En fait, très peu de solutions techniques respectent entièrement le modèle O.S.I.:

- Les matériels de connexion et les pilotes de périphériques disponibles sur le marché traitent en général la totalité des mécanismes des couches 1 et 2.
- Les mécanismes correspondant aux couches 3 et 4 sont traités par des produits logiciels qui englobent ces deux couches (c'est le cas, en particulier de TCP-IP, utilisé sur internet).
- Enfin, la plupart des applications réseaux disponibles assurent elles-même la gestion des sessions et la présentation des données.

Le modèle O.S.I. a donc surtout une valeur théorique. En fait, la plupart des architectures de réseaux existantes s'appuient sur une décomposition (implicite) correspondant aux trois niveaux définis plus haut.

REMARQUE:

La connaissance détaillée des couches de l'O.S.I. n'est pas nécessaire dans le cadre du présent ouvrage. Pour approfondir cette notion, vous pouvez consulter l'ouvrage RESEAUX INFORMATIQUES, disponible dans la rubrique Réseaux de la documentation en ligne du site ATLANTIC. Sinon, je vous conseille de vous contenter de retenir le tableau ci-dessus.

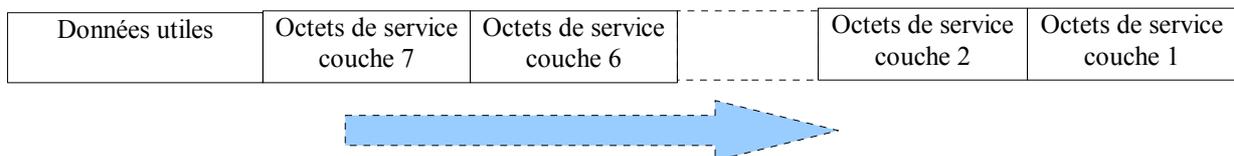
III.4.MODELE OSI ET TRAMES D'ENVOI DE DONNEES:

A l'intérieur d'un ordinateur, les informations se présentent sous la forme d'une suite d'octets (c'est-à-dire de groupes de 8 " bits ") contenant le codage de l'information en binaire. Ces octets vont constituer la " partie utile " d'un message à transmettre.

Cependant, pour qu'une transmission puisse s'effectuer en toute sécurité et atteindre le bon destinataire, il est nécessaire de rajouter un certain nombre d'octets "de service" aux octets utiles. Ces octets, ajoutés par les mécanismes de communication de l'émetteur sont destinés à être exploités par les mécanismes de communication du récepteur. Ils permettront à ceux-ci:

- De s'assurer que le message leur est bien adressé.
- De vérifier qu'il a été transmis intégralement et sans erreur.
- De connaître le format des données.
- D'identifier les différents segments d'un message qui a subi une segmentation afin de les remettre dans l'ordre à la réception.
- Etc.

Lors de l'émission d'un message, chacune des couches (au sens de l'OSI) de mécanismes de communication traversée ajoute ses propres octets de service. Lorsque toutes les couches d'émission auront été traversées, les segments de messages à injecter sur le média correspondront chacun à une structure de ce type:



Le contenu, le format et la position des octets de service de chaque couche dépend de la technologie de transmission utilisée par cette couche. Par exemple, sur internet, le contenu des couches 3 et 4 (niveau des services réseaux) correspond aux spécifications du protocole TCP-IP.

C'est cet ensemble d'octets (octets utiles+octets de service) qui est injecté sur le média de liaison, sous forme d'impulsions physiques dont la nature dépend du média (impulsions électriques, optiques, électromagnétiques, infrarouges...). Le train d'impulsion engendré, circulant sur le média, constitue une **TRAME DE TRANSMISSION DE DONNEES**.

Lors de la réception d'une trame par un hôte du réseau, les mécanismes de communication de cet hôte exploiteront ces différentes enveloppes d'octets de service dans l'ordre inverse de l'émetteur:

- En premier lieu, les octets de service des couches 1 et 2 seront utilisés par les mécanismes de gestion du lien physique pour déterminer si la trame leur est bien adressée et si elle a été transmise sans erreur.

<p>SUJET: COMPRENDRE INTERNET</p>	<p>REFERENCE: ATLANTIC/DOCTECH/ Initiation à l'informatique/Comprendre Internet Mise à jour: 11/01/11</p>
--	--

- Puis, la trame, dépouillée de ses deux premières couches d'octets de service sera transmise aux mécanismes des services réseaux pour y être traitée par eux.
- Et ainsi de suite....

III.5.UTILITE DE L'APPROCHE PAR NIVEAUX:

Chaque niveau cache au niveau qui lui est supérieur la complexité des traitements qu'il accomplit pour le compte de ce niveau supérieur. Il permet donc à celui-ci de faire abstraction de ces traitements. C'est grâce à cette architecture en plusieurs niveaux d'abstraction qu'un utilisateur d'internet n'a, à priori, besoin de connaître que le mode d'emploi du navigateur qu'il utilise.

D'autre part, la décomposition en niveaux permet de créer des architectures composites: chaque niveau ou couche peut être supporté par des composants d'origines différentes, pourvu que ces composants respectent les interfaces entre les niveaux. Cette possibilité est à la base de la création de " systèmes ouverts ", c'est-à-dire acceptant des produits de n'importe quel fournisseur, pourvu qu'il respecte les normes et standards d'interconnexion.

Les chapitres suivant exposent plus en détail les différents concepts attachés à ces trois niveaux.

IV.MECANISMES DE GESTION DU LIEN PHYSIQUE:

IV.1.NOTION PHYSIQUE DE RESEAU INFORMATIQUE:

IV.1.1.LES SYSTEMES INFORMATIQUES REPARTIS:

Actuellement, le système informatique d'une entreprise est le plus souvent constitué de plusieurs postes de travail équipant les différents services (comptabilité, ressources humaines, accueil, secrétariat, etc.). Un tel système est dit REPARTI. Les postes sont le plus souvent des systèmes de type "PC" ou "stations de travail".

Ces postes sont, en général, connectés entre eux en fonction des besoins d'accès de leurs utilisateurs aux informations. Ils peuvent également partager des connexions vers divers périphériques (disques partagés, imprimantes, scanners, etc.). Le schéma suivant donne un exemple de ce type d'organisation:

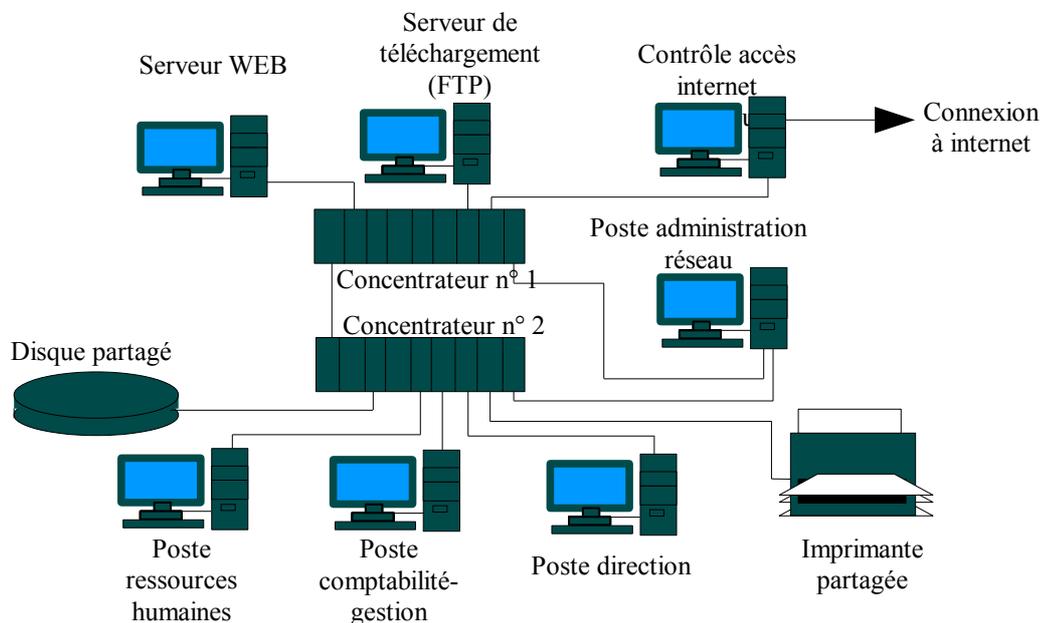


Schéma n° IV.1.1: Exemple de système informatique d'entreprise.

COMMENTAIRE:

Ce schéma représente UN EXEMPLE d'interconnexion de systèmes informatiques pour une entreprise. Nous pouvons distinguer ainsi:

- Des postes de travail correspondant à différentes fonctions: Direction, Administration réseau, Ressources humaines, Comptabilité-gestion.
- Des systèmes périphériques offrant des services ou des ressources partagés par ces postes de travail (disque partagé, Imprimante partagée).
- La machine supportant le serveur WEB de l'entreprise.
- Celle qui supporte le serveur de fichiers offerts au téléchargement (par exemple: documentation sur l'entreprise, notices, catalogues, etc.)
- La machine contrôlant la connexion à internet, munie des logiciels de sécurité nécessaires (pare-feux, translation d'adresse, etc.).

Les composants énumérés ci-dessus sont ceux qui supportent les différents services offerts aux utilisateurs. Nous les appellerons les HÔTES du système réparti.

D'autres composants ont pour unique fonction de permettre l'interconnexion des hôtes entre eux: il s'agit des médias d'interconnexion (conducteurs électriques, fibres optiques, liaisons sans fil, etc.) et des concentrateurs 1 et 2. Les

concentrateurs sont des systèmes matériels qui offrent aux utilisateurs un certain nombre de points de connexion (prises RJ45, prises coaxiales, etc), et permettent ainsi d'interconnecter différents médias (de même type ou de type différent, suivant la classe du concentrateur).

IV.1.2. NOTION DE TOPOLOGIE:

Si l'on s'intéresse uniquement aux interconnexions entre éléments, on obtient le schéma suivant:

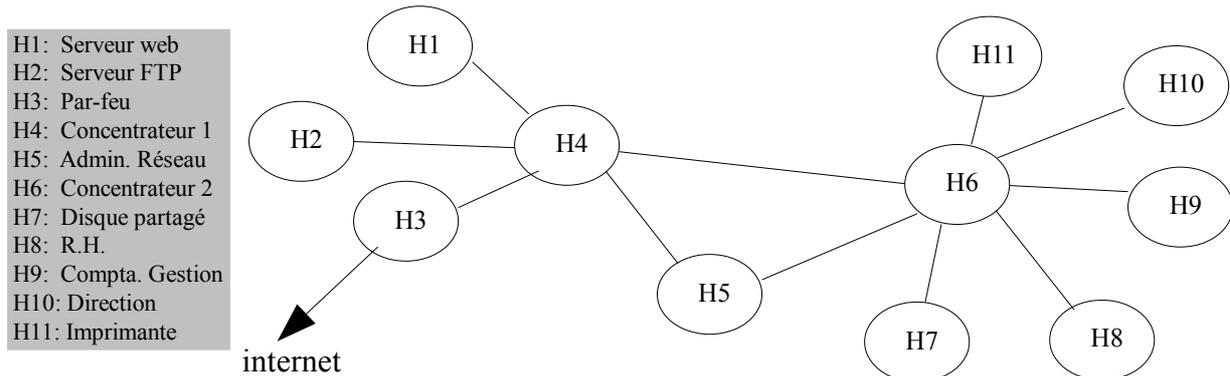


Schéma n° IV.1.2: Topologie

Ce deuxième schéma ne s'intéresse plus à la fonction des différents éléments. Il fait apparaître uniquement leurs interconnexions physiques. Nous avons affaire ici à un **RESEAU**, composé de NOEUDS (les différents éléments interconnectés), reliés par des **LIENS** (les interconnexions physiques).

REMARQUES:

- Le mot français "réseau" vient du latin "rete" qui signifiait "filet". Il n'est donc pas étonnant que les anglophones désignent un réseau par les termes "NET" (filet) ou "NETWORK" (maillage), qui évoquent le même concept.
- Les liens sont souvent appelés **LIAISONS**, car ils permettent de réaliser la liaison **POINT À POINT** de deux éléments (qui sont dits alors adjacents).
- Les nœuds sont appelés **HOTES** du réseau. Les nœuds non terminaux sont appelés aussi **COMMUTATEURS**, car ils permettent de créer des chemins entre des nœuds non adjacents.

Ce schéma permet de faire ressortir la **TOPOLOGIE** du réseau., c'est-à-dire la manière dont ses interconnexions sont organisées.

La topologie d'un réseau est évidemment une donnée très importante pour l'étude de son organisation, de son administration et de son fonctionnement (et en particulier, la sécurité de ce fonctionnement face aux dysfonctionnements et aux agressions possibles).

Par exemple:

- l'analyse de l'exemple ci-dessus permet de mettre en évidence deux structures "en étoile", organisées autour des nœuds H4 et H6.
- L'étoile formée autour de H4 forme un "sous-réseau" regroupant les hôtes dont la fonction exige de communiquer avec l'extérieur, alors que l'étoile formée autour de H6 forme un autre sous-réseau regroupant les postes dédiés aux fonctions internes de l'entreprise, mais ne pouvant communiquer directement avec l'extérieur: ce choix peut donc faire partie d'une stratégie de protection contre les intrusions.
- Le poste d'administration du réseau, qui doit pouvoir accéder à tous les hôtes, est relié directement aux deux domaines, par l'intermédiaire des concentrateurs. Il lui est ainsi possible de gérer les droits d'accès des différents postes, et en particulier la possibilité d'accès d'un sous-domaine à l'autre.

ATTENTION:

Le schéma topologique ne reflète pas forcément l'organisation **spatiale** des composants du réseau. Ainsi, dans le schéma ci-dessus, les nœuds H5, H7, H8, H9, H10 et H11 peuvent être physiquement situés dans n'importe quelle position les uns par rapport aux autres et par rapport à H6. Seule compte la manière dont leurs connexions au réseau sont réalisées, c'est-à-dire le fait que chacun d'eux est relié uniquement à H6 (structure en étoile autour de H6).

IV.1.3.LES TOPOLOGIES DE BASE:

Outre la topologie en étoiles, que l'on voit apparaître ci-dessus, les autres structures de base sont la topologie en anneau (l'interconnexion des hôtes forme une boucle fermée) et la structure en "BUS" (les hôtes sont tous connectés directement au même média):

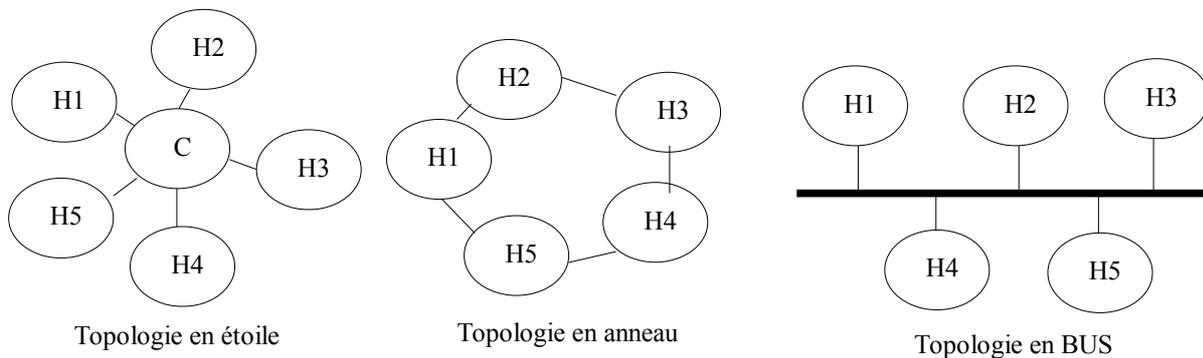


Schéma n° IV.1.3: Topologies de base

En pratique, pour des raisons tenant à la disponibilité des matériels de connexions, les topologies de réseau existantes sont pratiquement toujours des combinaisons de ces trois topologies de base. La topologie en étoile est de loin la plus utilisée pour les réseaux locaux. En effet, bien qu'assez onéreuse, car nécessitant, en supplément de l'acquisition du média et des matériels de connexion à ce média, l'acquisition de concentrateurs qui n'existent pas dans d'autres topologies, elle présente beaucoup d'avantages du point de vue de la sûreté, de la maintenabilité et de l'évolutivité.

IV.1.4.RESEAUX LOCAUX:

Un réseau est dit **LOCAL** lorsque tous ses composants sont internes à une organisation ou une entreprise: ces composants sont des **propriétés** de l'entreprise et sont entièrement situés dans une emprise géographique appartenant à celle-ci, quelle que soit la grandeur de cette emprise (c'est le cas du schéma n° IV.1.1). L'entreprise a donc la possibilité de **maîtriser entièrement** son architecture, son fonctionnement, sa maintenance, son évolution et son accès. Le terme anglais pour un réseau local est **LOCAL AREA NETWORK (L.A.N)**.

IV.1.5.RESEAUX ETENDUS:

Lorsqu'une organisation locale a besoin de se connecter à des systèmes informatiques distants ou bien ne lui appartenant pas, il ne lui est en général pas possible d'installer des équipements d'interconnexion privés pour assurer la liaison.

EXEMPLE:

à moins d'être EDF ou la SNCF, une entreprise parisienne ne peut pas envisager d'installer elle-même une connexion filaire privée avec une agence marseillaise: d'une part, l'investissement serait énorme et d'autre part, il faudrait traverser le domaine public ou des domaines privés qui ne lui appartiennent pas.

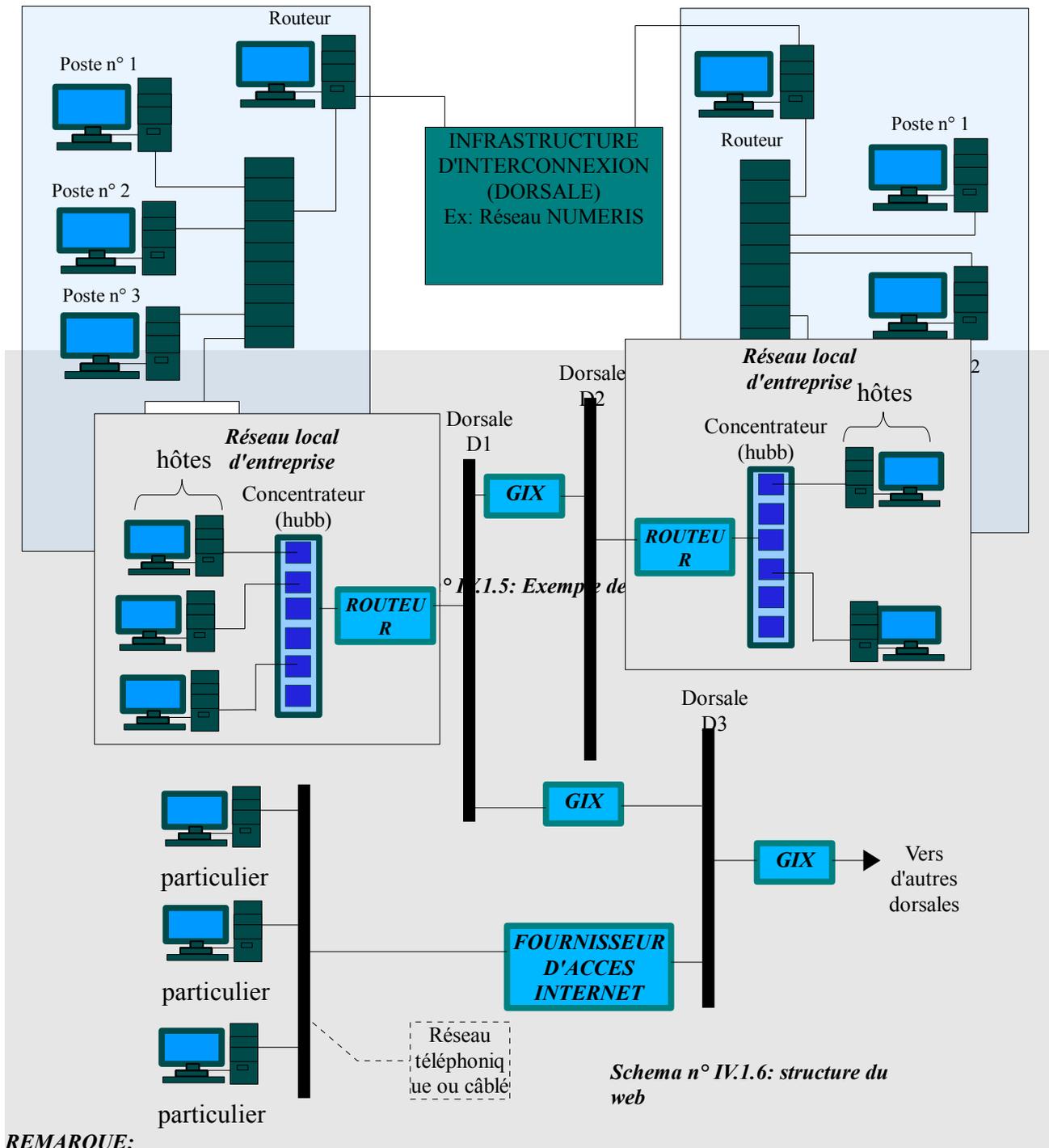
De ce fait, ce type d'interconnexion ne peut être réalisé qu'en utilisant des équipements mis à disposition par des organismes gestionnaires de réseaux (France Telecoms, les opérateurs internet, etc.). Les réseaux obtenus par ces interconnexions de réseaux locaux sont appelés **RESEAUX ETENDUS** en Français, ou **WIDE AREA NETWORKS (W.A.N)** en anglais. Suivant leurs dimensions, on parlera aussi de **RESEAU METROPOLITAIN** (à l'échelle d'une ville), de **RESEAU NATIONAL** (à l'échelle d'un pays), voire de **RESEAU MONDIAL**.

SUJET:
COMPRENDRE
INTERNET

REFERENCE: ATLANTIC/DOCTECH/
Initiation à l'informatique/Comprendre Internet
Mise à jour: 11/01/11

Le schéma suivant représente un exemple de réseau étendu:

- Nous pouvons distinguer à gauche un SITE PRINCIPAL, que nous supposons localisé en région Marseillaise. Ce site est équipé d'un réseau local (L.A.N), hébergeant 3 postes de travail.
- A droite du schéma est représenté le système informatique d'une agence éloignée (à Paris, par exemple). Celui-ci comprend deux postes de travail. Il s'agit également d'un réseau local.
- Les deux L.A.N. distants sont interconnectés grâce à une DORSALE à haut débit (par exemple, le réseau NUMERIS de France Telecom). L'accès à ce média est assuré sur les deux sites par des équipements appelés ROUTEURS qui assurent la transformation des messages circulant sur les deux LAN en messages compatibles avec la technologie de transmission utilisée sur la dorsale, et inversement.



REMARQUE:

SUJET:
COMPRENDRE
INTERNET

REFERENCE: ATLANTIC/DOCTECH/
Initiation à l'informatique/Comprendre Internet
Mise à jour: 11/01/11

Il ne faut pas croire qu'un tel schéma ne s'applique qu'à des organisations importantes. En effet, le schéma de l'installation d'un particulier équipé d'une "BOX" ADSL (livebox, freebox, etc.), est assez proche, en terme de complexité, de ce qui est représenté ci-dessus:

EXEMPLE:

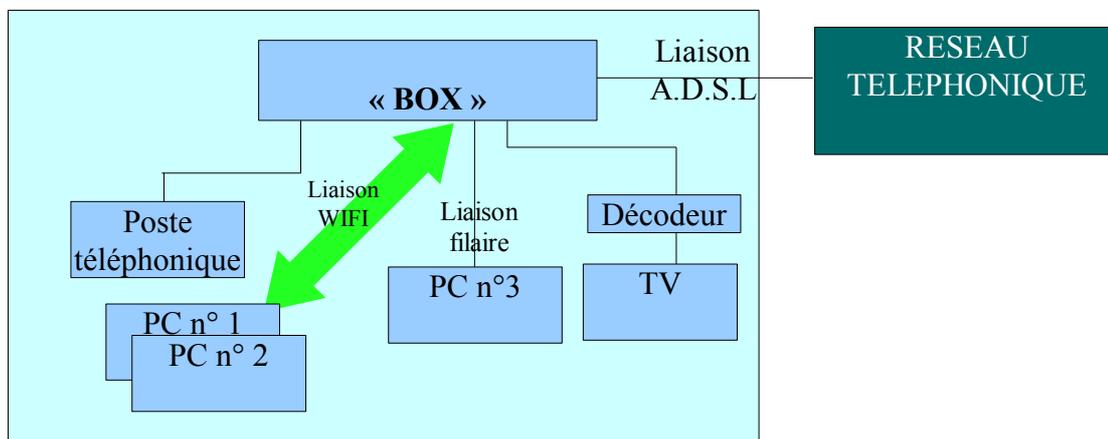


Schéma n° IV.1.5.bis: Connexion d'un particulier avec une BOX ADSL

COMMENTAIRE:

Ce schéma représente l'installation d'un particulier équipé d'une " BOX " ADSL. Le foyer dispose d'un poste PC fixe (PC n°3), connecté à la BOX par une liaison filaire (câble RJ45). Deux membres de la famille sont également équipés de PC portables qui se connectent grâce à la liaison WIFI. Nous voyons également apparaître les liaisons utilisées pour la TV numérique et la téléphonie numérique par ADSL. Nous pouvons voir qu'une BOX est l'équivalent d'un routeur d'agence associé à un concentrateur assez complexe puisqu'il est équipé de ports permettant la connexion à des technologies différentes (ports ethernet, USB, WiFi).

IV.1.6.LE WEB:

Le réseau internet, que l'on nomme communément WEB (pour World Exchange Bus) ou encore WWW (pour World Wide Web), n'est autre qu'un réseau mondial à accès public, constitué par l'interconnexion de nombreux autres réseaux publics ou privés.

L'infrastructure d'interconnexion d'internet est constituée par les DORSALES évoquées au premier chapitre. Ces dorsales sont bâties autour de médias à très haut débits (fibres de verre, faisceaux hertziens, etc). Certaines dorsales sont gérées par des administrations, d'autres le sont par des organisations privées, et en particulier par les opérateurs internet (les fournisseurs d'accès).

Les dorsales sont reliées entre elles par des points d'interconnexion appelés G.I.X. pour Global Internet eXchange, qui sont en fait des ROUTEURS à haute performance.

Les différents systèmes interconnectés sont reliés aux dorsales, soit directement, par des " lignes dédiées ", soit par l'intermédiaire de " Fournisseurs d'Accès ", avec lesquels ils communiquent par une autre infrastructure publique (réseau téléphonique, câble, etc.).

Le schéma ci-dessous donne une idée de cette structuration:

<p>SUJET: COMPRENDRE INTERNET</p>	<p>REFERENCE: ATLANTIC/DOCTECH/ Initiation à l'informatique/Comprendre Internet Mise à jour: 11/01/11</p>
--	--

IV.2.ECHANGE DE DONNEES ENTRE DEUX HOTES " ADJACENTS ":

IV.2.1.AVANT-PROPOS:

Nous allons d'abord nous intéresser à l'échange de données entre deux hôtes " ADJACENTS " d'un réseau, c'est-à-dire reliés directement par un média homogène. Cet exposé existe sous une forme beaucoup plus détaillée dans l'ouvrage " COMPRENDRE SON ORDINATEUR ", figurant dans la documentation en ligne offerte par le site ATLANTIC (www.atlantic-83.fr), dans la rubrique " initiation informatique ".

IV.2.2.MEDIA DE TRANSMISSION:

IV.2.2.1.GENERALITES:

L'échange de données entre deux systèmes informatiques, s'effectue par l'intermédiaire d'un **MEDIA DE TRANSMISSION**, c'est-à-dire un milieu conducteur dans lequel des signaux peuvent se propager. Les catégories de médias les plus utilisés sont:

- Les conducteurs électriques qui permettent la propagation de signaux électriques.
- Les fibres optiques, qui permettent la propagation de signaux lumineux
- Les liaisons hertziennes (appelées encore transmissions sans fil ou WIFI), qui utilisent la propagation d'ondes électromagnétiques modulées.

Dans un média de transmission, les informations sont véhiculées sous la forme de trains d'**IMPULSIONS**, c'est-à-dire de variations du signal propagé par le média. Ces impulsions codent les informations binaires à transmettre. Évidemment, chacune des catégories ci-dessus n'est capable de propager qu'un certain type de signal:

- Pour un conducteur électrique, ces signaux sont des **variations de tension**.
- Pour une fibre optique, les signaux seront des **variations d'intensité lumineuse**.
- Pour une transmission "sans fil" (WIFI), ce sont les **variation simultanées d'un champ électrique et d'un champ magnétique dans l'espace** qui codent et transportent le signal.

En général, les lignes exigeant de hauts débits, comme les DORSALES INTERNET sont constituées par des fibres optiques ou des liaisons hertziennes à grande puissance (faisceaux hertziens). Les conducteurs électriques (paires torsadées, câbles coaxiaux) ou les liaisons WIFI sont plutôt réservés à des usages locaux sur de courtes distances.

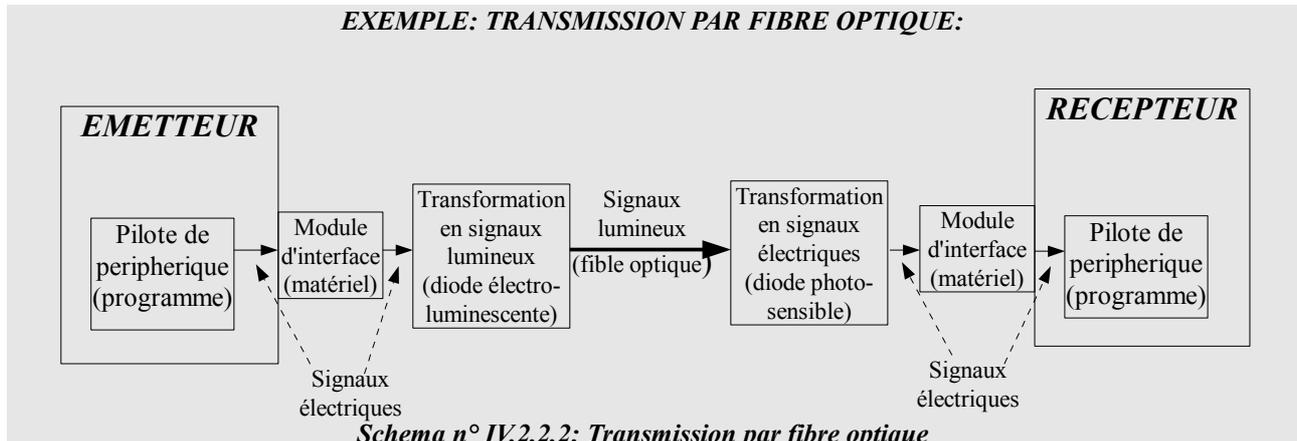
IV.2.2.2.INJECTION D'UNE INFORMATION DANS UN MEDIA:

SUJET:
COMPRENDRE
INTERNET

REFERENCE: ATLANTIC/DOCTECH/
Initiation à l'informatique/Comprendre Internet
Mise à jour: 11/01/11

Un système informatique ne traitant que des signaux électriques, c'est uniquement sous cette forme qu'il pourra délivrer les informations à émettre sur le média. Si celui-ci n'accepte pas ce type de signal, ceux-ci devront, avant leur injection, être transformés en signaux compatibles avec sa nature. Le récepteur devra transformer de nouveau les informations reçues en signaux électriques.

EXEMPLE: TRANSMISSION PAR FIBRE OPTIQUE:



COMMENTAIRE:

Pour transmettre une information issue d'un ordinateur sur une fibre optique, il faut la transformer en impulsions lumineuses, c'est-à-dire faire varier l'intensité du faisceau lumineux d'une façon analogue aux variations du signal électrique qui la code.

Ce passage du signal électrique au signal lumineux est obtenu grâce à une diode électroluminescente (diode LED). En effet, l'intensité de l'émission de lumière d'une telle diode croît et décroît dans le même sens que l'intensité du courant électrique qui la traverse. Un dispositif optique permet d'injecter cette lumière sous la forme d'un " rayon " fin dans l'âme de la fibre.

A l'autre extrémité de la fibre, les impulsions lumineuses doivent être retransformées en impulsions électriques. On utilise pour cela une diode photosensible qui produit un courant électrique lorsqu'elle est frappée par un rayon lumineux. L'intensité du courant électrique ainsi produit croît et décroît dans le même sens que l'intensité du rayon lumineux qui la frappe.

IV.2.3. TRAME D'ENVOI DE DONNEES:

Nous avons abordé, au chapitre précédent, la notion de TRAME de transmission de données. Nous avons vu en particulier:

- Qu'une trame était constituée par un train d'impulsions injecté sur le média de transmission.
- Que son contenu correspondait au codage en impulsions physiques compatibles avec la nature des médiums d'un segment de message à émettre.
- Que chacun de ces segments était constitué de " données utiles " accompagnées d'une succession d'enveloppes d'octets de service correspondant aux différentes couches de mécanismes de transmission traversés.

Le contenu, le format et la position des octets de service des couches 1 et 2 dans une trame dépendent de la technologie de liaison utilisée. On parlera ainsi d'une TRAME ETHERNET, d'une TRAME BSC, d'une TRAME HDLC, etc.

IV.2.4. CODAGE ELECTRIQUE D'UNE INFORMATION DIGITALE:

IV.2.4.1. CODAGE EN BANDE DE BASE:

La figure suivante représente le codage électrique en **BANDE DE BASE** de l'octet 10110110 par la méthode dite

SUJET:
COMPRENDRE
INTERNET

REFERENCE: ATLANTIC/DOCTECH/
Initiation à l'informatique/Comprendre Internet
Mise à jour: 11/01/11

"NRZ". Nous voyons que la forme du signal **traduit directement la configuration binaire** de la donnée transmise, sous forme de variations de tension appelées "signaux carrés":

- Les bits dont la valeur est 1 sont transmis sous la forme d'un palier de tension à +5 volts, maintenu pendant une durée ΔT .
- Les bits de valeur 0 sont traduits par des paliers de même durée, mais de tension -5 volts.
- Une absence de transmission est traduite par un palier 0 volts.

La variation du signal électrique en fonction du temps (courbe rouge) correspond à la transmission de la valeur binaire 10110110 (soit 182 en décimal, s'il s'agit d'un entier naturel):

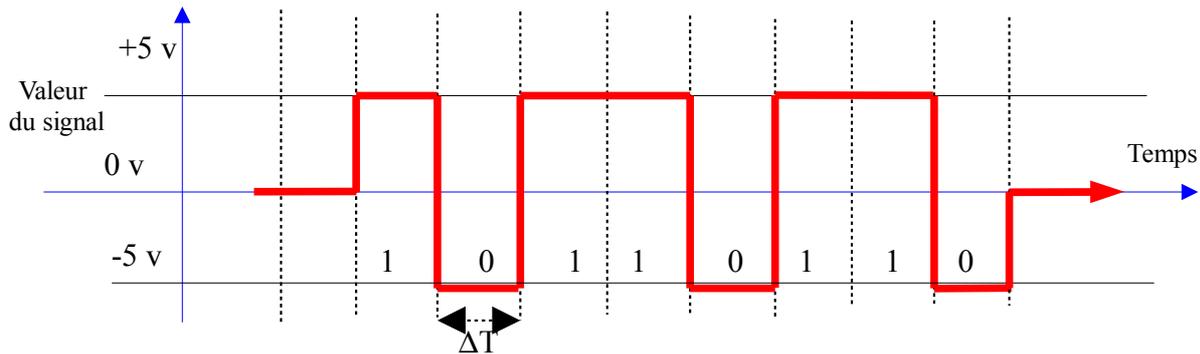


Schéma n° IV.2.4.1: codage électrique en bande de base

Le codage NRZ n'est qu'une des nombreux codages électriques existants. Il n'est pas question d'en faire une étude approfondie dans le cadre de ce cours. Nous avons choisi un exemple en NRZ parce que c'est la méthode la plus intuitive.

Remarquons que la durée ΔT détermine la **vitesse de transmission** du message, donc le **débit** de la ligne. Dans le cas présent, si ΔT vaut 0,1 milliseconde, le débit de la ligne pourrait atteindre:

$$1 / 0,0001 = 10000 \text{ bits par seconde, soit } 10 \text{ kbits/s}$$

IV.2.4.2. TRANSMISSION PAR MODULATION:

Dans ce cas, la trame est injectée sur le média sous la forme de variations d'un signal périodique à haute fréquence, appelé "**PORTEUSE**" (la dénomination exacte est "onde porteuse"), en fonction du codage électrique obtenu.

Le schéma suivant représente un codage de l'octet 10110010 par une **modulation de l'amplitude** de la porteuse: Un bit à 0 est représenté par une amplitude moindre qu'un bit à 1, maintenue pendant une durée Δt :

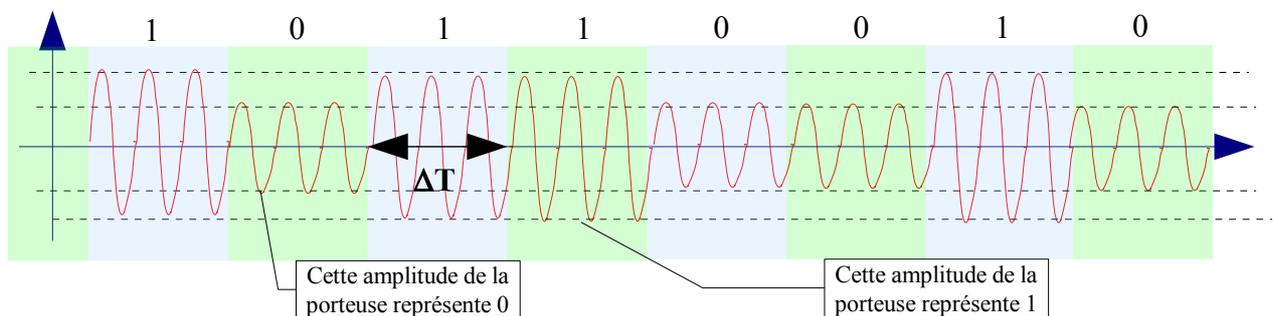


Schéma n° IV.2.4.2: transmission en modulation d'amplitude

REMARQUE:

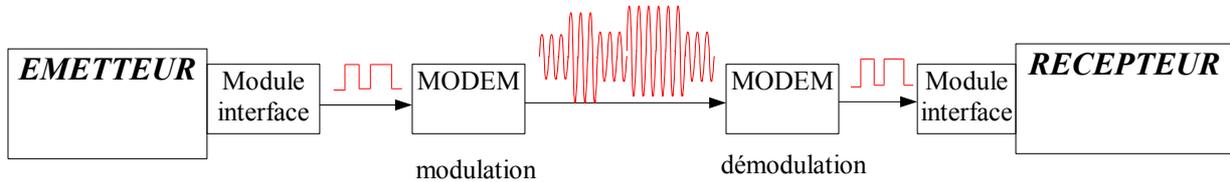
Il est également possible de moduler la **FRÉQUENCE** de la porteuse (**modulation de fréquence**) ou encore sa **PHASE** (**modulation de phase**).

IV.2.4.3. NOTION DE MODEM:

SUJET:
COMPRENDRE
INTERNET

REFERENCE: ATLANTIC/DOCTECH/
Initiation à l'informatique/Comprendre Internet
Mise à jour: 11/01/11

La transmission par modulation est réalisée en intercalant dans le circuit de transmissions deux composants appelés **MODEM (Modulateur-DEModulateur)**. Ce sont eux qui, à l'émission, modulent les porteuses avec les signaux électriques en bande de base et qui, à la réception, filtrent les porteuses, récupèrent les informations et les codent de nouveau en bande de base:



Schema n° IV.2.4.3: Transmission par modems

IV.2.4.4.EXEMPLE DE MODULATION: VOTRE LIAISON A.D.S.L:

Les liaisons ADSL sont de bons exemples d'utilisation de la modulation. Une liaison A.D.S.L. (Asymétric Digital Suscribed Line) désigne un procédé de transmission de signaux numériques entre deux postes à partir des liaisons téléphoniques locales de ces deux postes ("boucle de cuivre" locale, reliant l'abonné au central téléphonique).

Les informations ADSL utilisent donc **le même média que les signaux téléphoniques**. Pour séparer ces deux flux d'information, on utilise des composants appelés **FILTRES EN FREQUENCE** qui permettent de séparer les signaux modulés sur deux porteuses de fréquences différentes (c'est ce que fait le séparateur ADSL que vous avez placé sur vos prises téléphoniques). Cette technique, qui permet de faire transiter simultanément sur le même média plusieurs flux de données indépendants, s'appelle un **MULTIPLEXAGE** (dans ce cas, il s'agit d'un multiplexage en fréquence).

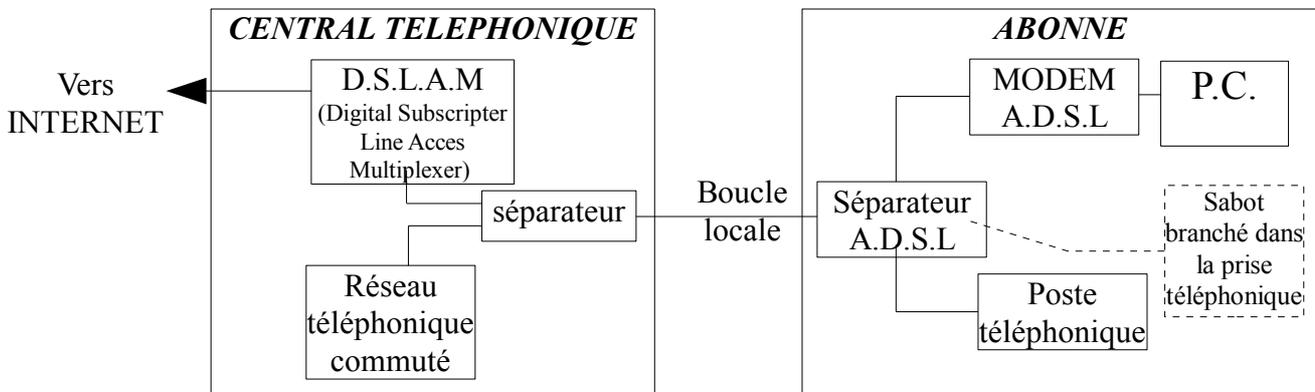


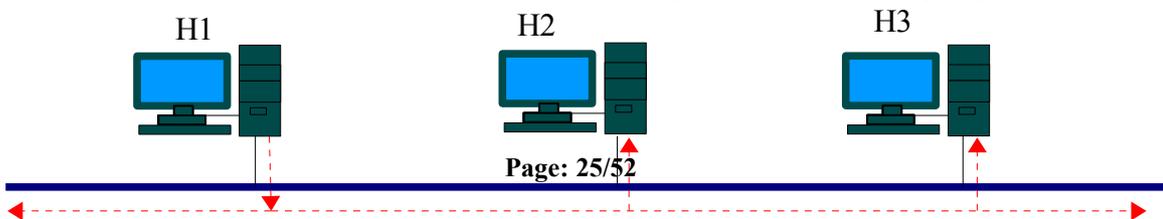
Schéma n° IV.2.4.4: liaison A.D.S.L

C'est donc le **multiplexage en fréquences** qui permet de séparer flux ADSL et téléphonie classique (en utilisant un filtre ADSL). Les différents canaux attribués au flux ADSL peuvent transmettre simultanément les informations du web, la télévision numérique et la téléphonie numérique, en utilisant des porteuses différentes. Les modems utilisés sont, dans ce cas des **MODEMS ADSL**, en général intégrés aux diverses BOX (LifeBox, NeufBox...) fournies par les fournisseurs d'accès à internet.

IV.2.5.ACHEMINEMENT D'UNE TRAME VERS UN DESTINATAIRE SUR UN RESEAU:

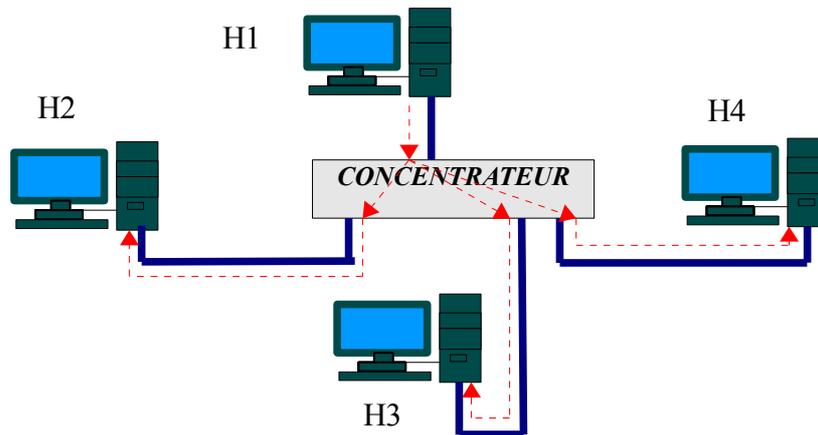
IV.2.5.1.POSITION DU PROBLEME:

Une particularité d'un réseau est que lorsqu'un hôte émet une trame sur le média de ce réseau, cette trame peut se trouver distribuée à l'ensemble des hôtes connectés au média. C'est toujours le cas dans une topologie en " bus ":



La trame émise par H1 est propagée vers tous les hôtes du réseau (H2, H3, etc...)

Cela peut être également le cas dans les topologies en étoile (si le concentrateur ne dispose pas de fonctions de filtrage):

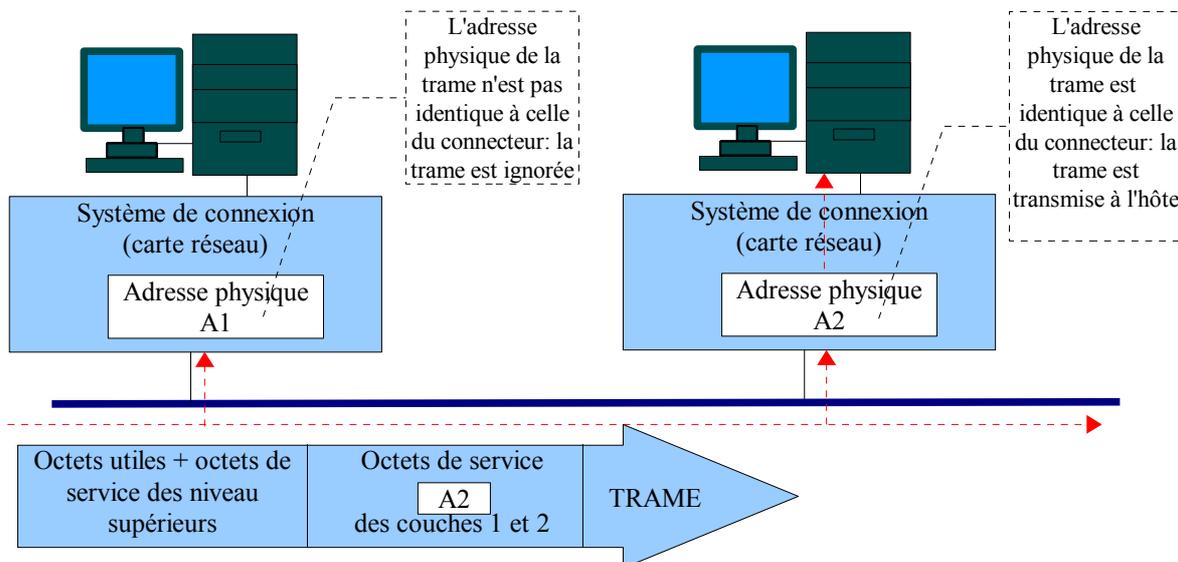


Il est donc nécessaire de mettre en place un mécanisme d'ADRESSAGE qui permette aux hôtes du réseau de reconnaître les TRAMES qui leur sont destinées.

IV.2.5.2.MECANISME D'ADRESSAGE PHYSIQUE:

Les systèmes de connexion des hôtes aux réseaux (les " cartes réseaux ") renferment dans leur logique interne un identifiant sous la forme d'une valeur numérique binaire (en général codée sur 6 octets, soit 48 bits). Cet identifiant est appelé **ADRESSE PHYSIQUE**, car elle identifie un **POINT DE CONNEXION PHYSIQUE** d'un hôte à un réseau (dans la terminaison O.S.I., on l'appelle **ADRESSE M.A.C.** pour Media Access Control).

Pour adresser un hôte donné, il suffit d'inclure dans les octets de service de la trame cette adresse physique: lorsque le système de connexion d'un hôte reçoit un message, sa logique interne extrait cette donnée de la trame reçue et la compare à sa propre adresse physique. Si les deux données sont identiques, il transmet la trame à son hôte, sinon, il n'en tient pas compte:



Schema n°VI.2.5.2: mécanisme d'adressage physique

IV.2.5.3.LIMITES DU MECANISME D'ADRESSAGE PHYSIQUE:

Les mécanismes d'adressage physique sont, en général, liés à une technologie de liaison particulière, qui détermine l'emplacement et le format des octets de services dans les trames: en particulier, chaque technologie peut utiliser un format d'adresse qui lui est propre. Certaines adresses physiques sont libellées sur 16 bits au lieu de 48. Il n'est donc pas possible avec une adresse physique d'adresser des hôtes reliés par un chemin composé de plusieurs médias successifs utilisant des technologie de liaison différentes.

D'autre part, l'adresse physique est indissolublement liée au **composant** de liaison: par exemple, chaque cartes réseau de technologie ETHERNET possède une adresse physique qui lui est propre et qu'il est (pratiquement) impossible de changer. De ce fait, lorsqu'on change une carte réseau, on change également l'adresse physique de l'hôte qu'elle connecte.

L'adressage physique seule ne peut donc suffire que dans un réseau local et de technologie homogène (Il serait tout de même dommage que lorsque Google procède au remplacement d'une carte réseau défectueuse sur un de ses serveurs, il soit obligé d'en avertir tous ses utilisateurs !!!).

Du fait de ces inconvénients, des mécanismes d'adressage moins liés au matériel se superposent à l'adressage physique dans les autres niveaux de mécanismes d'internet: nous verrons qu'il s'agit de l'adressage IP et des URL.

SUJET: COMPRENDRE INTERNET	REFERENCE: ATLANTIC/DOCTECH/ Initiation à l'informatique/Comprendre Internet Mise à jour: 11/01/11
---	---

V.MECANISMES DES SERVICES RESEAU:

V.1.NOTION DE PROCESSUS LOGICIEL:

Pour être utilisable, votre ordinateur doit être équipé d'un certain nombre de PROGRAMMES INFORMATIQUES (navigateurs, traitements de textes, etc.). Ces programmes sont enregistrés dans des fichiers sur des supports de masse (disque dur, clef USB, etc.). Ces fichiers contiennent donc les instructions et les données qui constituent le CODE EXECUTABLE du programme.

Pour utiliser ce programme, il est nécessaire, avant d'en lancer l'exécution, de copier son code exécutable dans la mémoire vive de l'ordinateur, puis de lui attribuer un PROCESSEUR capable de l'exécuter. C'est le système d'exploitation qui réalise ces opérations lorsqu'il reçoit une requête de lancement d'un programme provenant soit d'un utilisateur (double clic sur le lien d'un fichier exécutable), soit d'un autre programme en cours d'utilisation.

Rien n'empêche de faire plusieurs copies du code exécutable d'un programme en mémoire vive et de lancer simultanément l'exécution de chacune de ces copies (il suffit de disposer de suffisamment de processeurs ou de disposer d'un système d'exploitation capable de partager le processeur entre les différentes copies). A un instant donné, plusieurs copies (ou **instances**) d'un programme donné peuvent donc être en cours d'exécution dans la même machine: vous pouvez vérifier vous-même qu'il est possible de lancer simultanément plusieurs instances de votre navigateur ou de la calculatrice, (si vous êtes sous windows), et que chaque instance est utilisable.

Chaque instance en cours d'exécution d'un programme donné constitue un **PROCESSUS LOGICIEL** (notion à ne pas confondre avec la notion de PROCESSEUR, qui désigne le circuit intégré qui exécute les programmes). Par la suite, nous utiliserons le terme **processus** pour désigner un programme en cours d'utilisation, car cela allège beaucoup la rédaction.

V.2.NOTION DE PROTOCOLE DE COMMUNICATION:

Dans le domaine technologique, un **PROTOCOLE DE COMMUNICATION** est un ensemble de règles destinées à régir la communication entre deux entités (par exemple, deux hôtes d'un réseau). Ces règles peuvent concerner:

- La représentation des informations échangées (codage, format).
- Les mécanismes de transmission (adressage, séquençement et synchronisation des échanges, segmentation et ré-assemblage, etc.).
- Les mécanismes de sécurisation de la transmission (contrôle de l'intégrité des données transmises, mécanismes d'acquiescement, de signalisation ou de correction d'erreurs, etc.).

V.3.ADRESSAGE DES MACHINES HOTES D'UNE INTERCONNEXION DE RESEAUX:

Nous avons vu au chapitre précédent que les mécanismes d'ADRESSAGE PHYSIQUE présentaient les inconvénients d'être liés à la fois à la technologie de liaison utilisée et au composant matériel assurant la liaison. Ces deux caractéristiques les rendent inappropriés pour adresser des machines situées sur des réseaux interconnectés pouvant être de technologie différente et très éloignés dans l'espace.

De ce fait, il est nécessaire d'utiliser au niveau des services réseaux (et en particulier dans la couche OSI RESEAU) un autre système d'adressage non lié à la technologie de liaison, l'ADRESSAGE LOGIQUE.

V.4.ADRESSAGE DES PROCESSUS SUR UNE INTERCONNEXION DE RESEAUX:

Nous avons établi précédemment que la communication sur internet se déroulait toujours entre deux programmes informatiques en cours d'exécution, c'est-à-dire entre deux **processus logiciels**. Pour que ces processus puissent échanger des messages, il est nécessaire qu'ils disposent d'un moyen de s'identifier mutuellement sans ambiguïté sur le réseau. Il faut donc disposer d'un **mécanisme d'identification des processus** valide pour l'ensemble du réseau.

L'adresse d'un processus sera donc constituée:

- De l'ADRESSE LOGIQUE de la machine sur laquelle il s'exécute.

SUJET: COMPRENDRE INTERNET	REFERENCE: ATLANTIC/DOCTECH/ Initiation à l'informatique/Comprendre Internet Mise à jour: 11/01/11
---	---

- D'un identifiant interne à la machine appelé PORT RESEAU. Un port réseau est la plupart du temps un simple nombre entier naturel (entier positif).

REMARQUES:

- Les PORTS RESEAU sont attribués aux processus lorsque ceux-ci en ont besoin pour ouvrir une communication réseau. Lorsque la communication est terminée, ils peuvent être réattribués à un autre processus en fonction des besoins. Un port réseau est donc une entité éphémère (de même, d'ailleurs, que les processus).
- Si un processus communique simultanément avec plusieurs autres processus, il peut se voir attribuer plusieurs ports de communication. De ce fait, si un numéro de port permet d'identifier un processus, l'inverse n'est pas vrai.

V.5.LA SUITE DE PROTOCOLES TCP-IP:

V.5.1.GENERALITES:

TCP-IP est en fait une **suite** de deux protocoles:

- Le protocole I.P. (Internet Protocol), qui correspond à la couche 3 du modèle OSI (routage de paquets sur une interconnexion de réseaux).
- Le protocole T.C.P. (Transmission Control Protocol), qui correspond à la couche 4 du modèle OSI (transport de messages constitués de plusieurs segments sur un réseau).

L'ensemble correspond aux mécanismes du NIVEAU DES SERVICES RESEAUX défini au chapitre précédent. TCP-IP, qui était à l'origine un protocole " propriétaire ", c'est-à-dire lié à un constructeur particulier (XEROX), est petit à petit devenu un standard, en particulier sur internet, en particulier grâce au bon niveau de sécurité qu'il permet d'obtenir.

V.5.2.ADRESSAGE D'UN PROCESSUS SOUS TCP-IP:

TCP-IP inclue un mécanisme d'identification des processus basé sur un couple de données:

- La première de ces données permet d'identifier la **machine** dans laquelle le processus s'exécute: c'est l'ADRESSE LOGIQUE, communément appelée **ADRESSE IP**. Vous avez certainement rencontré ce type d'adresse, qui se note le plus souvent sous la forme de quatre nombres séparés par des points (Exemple: 94.23.63.201 est l'adresse IP du serveur hébergeant le site de l'association ATLANTIC). Il est important de se souvenir que la donnée d'une adresse IP permet de **repérer sans ambiguïté une machine connectée sur internet**.
- La deuxième donnée est un simple nombre entier naturel, que l'on nomme **NUMERO DE PORT**. Un PORT du protocole TCP définit un " point d'accès " à un processus logiciel **à l'intérieur d'une machine donnée**. Contrairement à l'adresse IP, il s'agit donc d'une donnée valable localement. A l'intérieur d'une même machine, l'attribution des numéros de ports est contrôlée par le système d'exploitation. Deux ports ne peuvent donc avoir le même numéro. De ce fait, à un instant donné, un numéro de port permet d'identifier sans ambiguïté un des processus actifs d'une machine.

Le couple {ADRESSE-IP + Numéro de PORT TCP} définit donc sans ambiguïté un processus donnée tournant sur une machine connectée à internet. Il constitue un système d'adressage des processus valable sur tout l'espace du web et depuis n'importe quel point de cet espace.

REMARQUES:

- En fait, une adresse IP repère non pas une **machine**, mais une des **connexions physiques** de cette machine sur le réseau. Une machine peut donc avoir plusieurs adresses IP, si elle est connectée à plusieurs réseaux. En revanche, une adresse IP correspond à une connexion et une seule.
- De même, un processus donné peut déclarer plusieurs ports (correspondant à plusieurs services que ce processus peut offrir). Ces ports auront des numéros différents. En revanche, un numéro de port correspond à un processus local et un seul.
- Le couple [ADRESSE LOGIQUE+NUMÉRO DE PORT] est appelé SOCKET par les développeurs (mot

SUJET:
COMPRENDRE
INTERNET

REFERENCE: ATLANTIC/DOCTECH/
Initiation à l'informatique/Comprendre Internet
Mise à jour: 11/01/11

anglais que l'on peut traduire par "support de connexion").

- Un PORT n'existe que pendant le temps où le processus qui l'a déclaré existe. Dès qu'un processus est terminé, le système d'exploitation élimine tous les ports que ce processus a déclaré et peut réallouer leurs numéros à d'autres ports dès l'instant que d'autres processus en font la demande.
- Il existe actuellement deux versions d'adressage IP, appelées respectivement IP-V4 et l'IP-V6. Une adresse IP-V4 est représentée par un groupe de 4 octets (c'est-à-dire 32 bits) alors qu'une adresse IP-V6 en comprend 16 (128 bits). L'IP-V6 permet donc de disposer de beaucoup plus de capacité d'adressage, alors que la capacité de l'IP-V4 commence à être insuffisante pour les besoins du web. Cependant, l'IP-V4 constitue encore l'adressage de base sur internet et le constituera sûrement encore pendant quelques années.

Le schéma suivant représente le mécanisme TCP-IP d'adressage des processus sur internet:

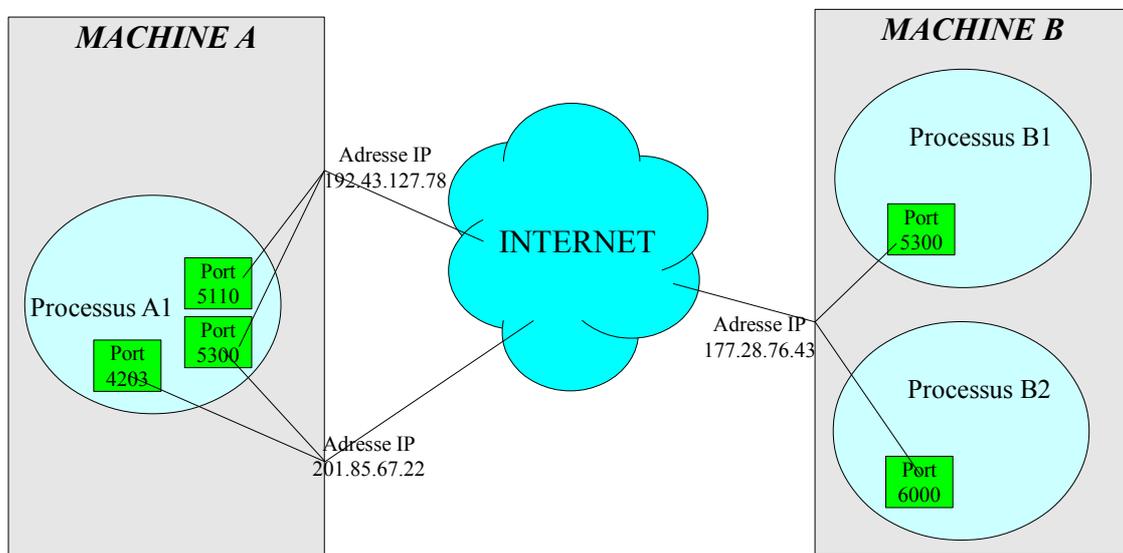


Schéma n° V.5.2: Adressage des processus

COMMENTAIRES:

- Le schéma ci-dessus représente deux machines connectées à internet (machines A et B). Le reste du réseau internet est volontairement représenté par un "nuage", car on n'a pas besoin de connaître sa structure interne.
- La machine A est connectée à internet par deux points de connexion physique différents. Ces connexions sont repérées par les adresses IP 192.43.127.78 et 201.85.67.22. La machine B est connectée à internet par une seule connexion physique, dont l'adresse IP est 177.28.76.43.
- Dans la machine A s'exécute le processus logiciel A1. Celui-ci a déclaré 3 ports, de numéros 5110, 5300 et 4263. Le processus A1 est en cours de communication sur les sockets [192.43.127.78 + 5110], [192.43.127.78 + 5300], [201.85.67.22 + 5300], et [201.85.67.22 + 4263]. On peut supposer que chacun des trois ports ouverts par le processus A1 correspond à un service particulier que ce processus peut offrir.
- Dans la machine B s'exécutent les processus logiciel B1 et B2. Le processus B1 est en cours de communication sur son port 5300, via l'adresse IP 177.28.76.43 (donc, sur le socket réseau [177.28.76.43+5300]). Il communique également sur le socket [177.28.76.43+6000].
- Pour envoyer un message vers le port 5110 du processus A1, un processus s'exécutant sur la machine B devra donc utiliser soit le socket [192.43.127.78, 5110], soit le socket [201.85.67.22, 5110].

V.5.3. PROTOCOLE TCP ET MECANISME DE TRANSMISSION " PAR PAQUETS " :

V.5.3.1. JUSTIFICATION:

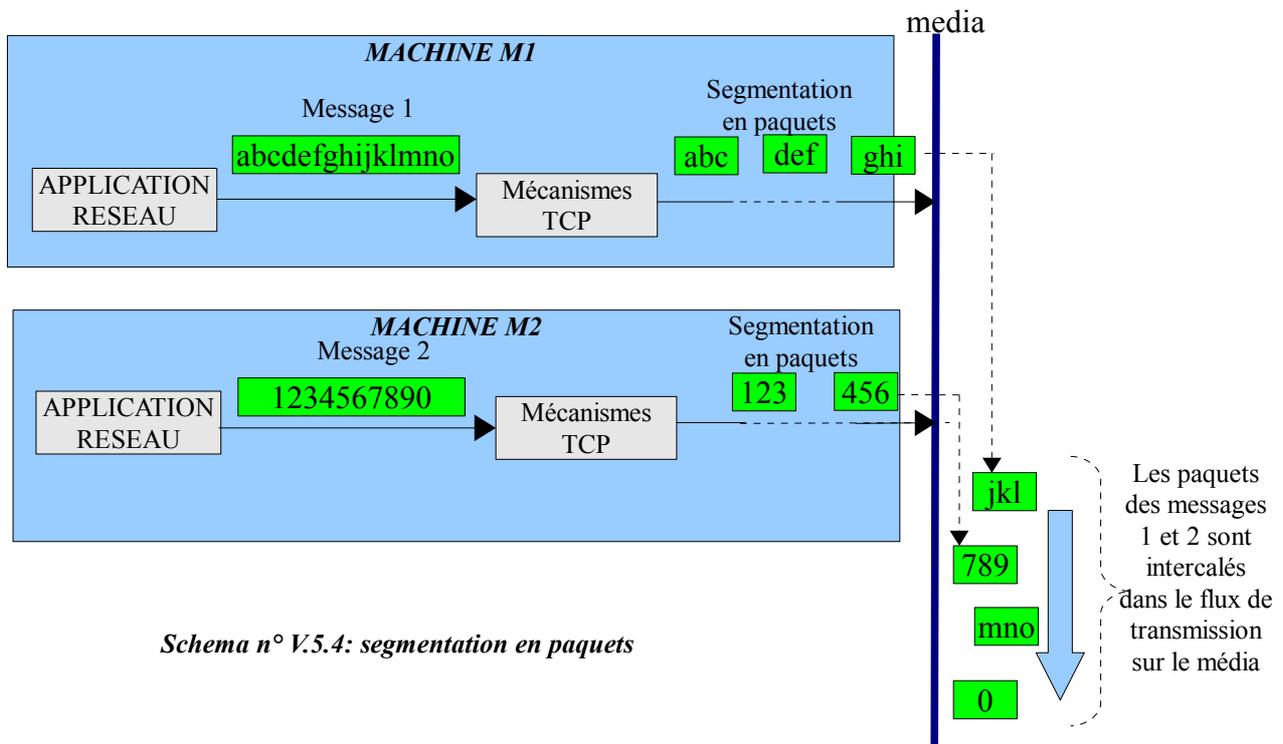
Lorsqu'un processus du niveau des services applicatifs doit transmettre un message à un processus distant, il utilise des mécanismes du niveau des services réseau: sur internet, ces mécanismes sont ceux qui supportent la suite de protocoles TCP-IP.

Cette activation se traduit par le lancement de logiciels accompagné de la fourniture des données à transmettre. Celles-ci sont parfois très volumineuses: elles peuvent, par exemple, correspondre au contenu d'un fichier photographique de plusieurs méga-octets. Si de tels volumes de données pouvaient être transmis en un seul bloc sur internet, occupant ainsi certains segments de média pendant plusieurs secondes, ils bloqueraient pendant cette durée les transmissions en provenance d'autres machines, y compris les messages courts ou urgents. Le résultat en termes de rendement serait catastrophique.

De ce fait, le protocole TCP inclut un mécanisme qui segmente les messages volumineux en **PAQUETS** de dimensions plus modestes (moins de 65536 octets).

V.5.3.2. PRINCIPES DE FONCTIONNEMENT:

Ce procédé permet d'intercaler sur le média des paquets appartenant à différents messages. Ainsi, un message court M2 peut être transmis sans attendre la fin d'un autre message M1, même si celui-ci est en cours de transmission, en intercalant ses paquets dans les paquets de M1.



Schema n° V.5.4: segmentation en paquets

Ce mécanisme de segmentation implique la mise en œuvre d'un mécanisme de ré-assemblage des paquets chez le récepteur. Pour contrôler ce procédé, éviter les erreurs (perte de paquets, mauvais ré-assemblage, etc.) ou recommencer la transmission si nécessaire, un certain nombre d'octets de service sont ajoutés aux données de chaque paquet:

- Un " numéro de séquence " qui permet de repérer l'ordre de chaque paquet dans le message).
- Le numéro de port du destinataire
- Et d'autres paramètres qu'il est inutile de détailler ici...

SUJET:
COMPRENDRE
INTERNET

REFERENCE: ATLANTIC/DOCTECH/
Initiation à l'informatique/Comprendre Internet
Mise à jour: 11/01/11

Ces octets de service sont nommés " enveloppe TCP ". En résumé, à la " sortie " des mécanismes TCP:
un MESSAGE application => Une suite de PAQUETS
Un PAQUET = [Un SEGMENT DE MESSAGE] + [une ENVELOPPE TCP (octets de service)]

V.5.3.3. MECANISMES DE SECURISATION:

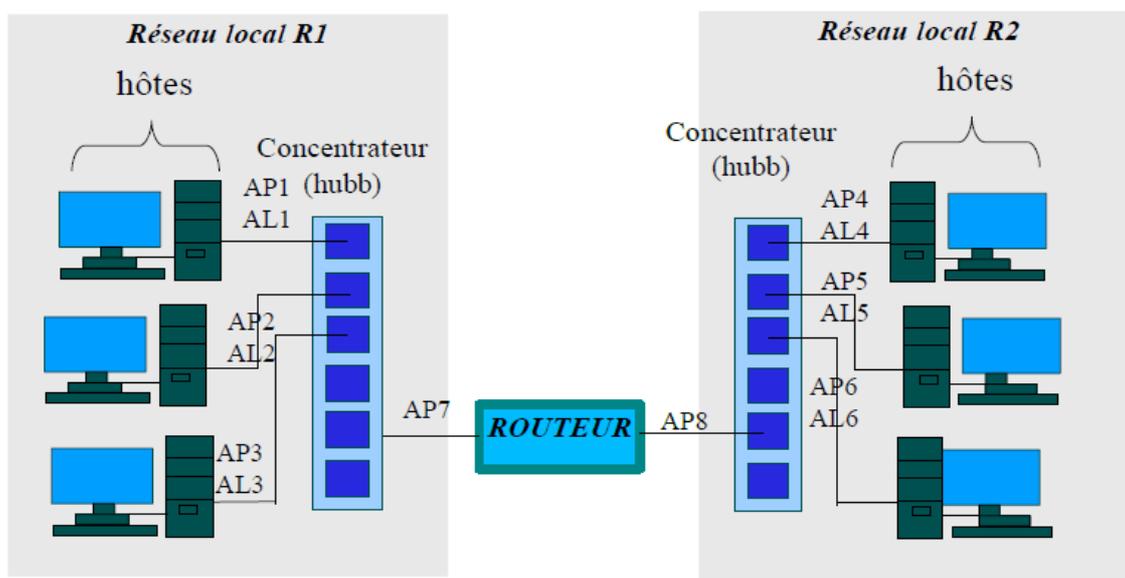
Le protocole TCP de transmission par paquets est très sécurisé. En effet, la réception de chaque paquet par le destinataire donne lieu à l'envoi d'un message d'acquittement vers l'émetteur. Ce mécanisme, associé à la gestion des numéros de séquences, permet à l'émetteur de s'assurer que ces paquets ont bien été transmis, ou bien, en cas de dysfonctionnement, de répéter les envois défectueux. Le récepteur d'un paquet doit donc pouvoir identifier son émetteur durant toute la durée de la transmission. De ce fait, on dit que TCP est un protocole de transport en MODE CONNECTE.

L'avantage de TCP est la grande sécurisation des échanges de données. L'inconvénient est que l'on ne peut garantir la durée d'une transmission. De ce fait, TCP possède une variante non sécurisée, sans mécanisme de segmentation ni d'acquittement, le sous-protocole Unconnected Datagram Protocol (U.D.P). Certaines applications réseau n'exigeant pas une grande sécurité de transmission ou bien assurant elles-même cette sécurité, utilisent U.D.P. plutôt que TCP.

V.5.4.PROTOCOLE IP ET ROUTAGE INTER-RESEAUX:

Nous avons vu que, d'un point de vue physique, INTERNET était constitué par l'interconnexion de différentes DORSALES à haut débit. Chacune de ces dorsales, qui constitue un sous-réseau, peut utiliser des solutions technologiques qui lui sont propres. De ce fait, les modalités de transmission des informations (codage, protocole de liaison entre les hôtes, etc.) peuvent être totalement différentes d'une dorsale à l'autre. Il en est de même en ce qui concerne les réseaux locaux raccordés à internet. De ce fait, le passage d'un sous-réseau à un autre par un point d'interconnexion nécessite un travail de transcodage des informations.

D'autre part, les systèmes d'adressage physique entre les hôtes d'un réseau dépendent aussi de la technologie utilisée. Ils peuvent donc être différents et incompatibles d'un réseau à l'autre. Un hôte donné ne peut donc pas, en utilisant le mécanisme d'adressage propre à la technologie de son réseau, adresser des informations à un hôte d'un réseau différent. TCP-IP résout le problème en utilisant un mécanisme de **ROUTAGE INTER-RESEAUX**, supporté par des équipements d'interconnexion appelés **ROUTEURS**. Les routeurs sont équipés de logiciels mettant en œuvre le mécanisme de routage du protocole **IP** (Internet Protocol), dont voici la description sommaire:



Schema n° V.5.3.1: Routage inter-reseaux

SUJET: COMPRENDRE INTERNET	REFERENCE: ATLANTIC/DOCTECH/ Initiation à l'informatique/Comprendre Internet Mise à jour: 11/01/11
---	---

COMMENTAIRES:

Le schéma ci-dessus représente l'interconnexion de deux réseaux locaux par l'intermédiaire d'un routeur.

A chaque point de connexion d'un hôte, l'adresse physique et l'adresse logique de la connexion sont représentées par les symboles AP1 (adresse physique 1), AL1 (adresse logique 1), AP2, AL2, etc.

A l'intérieur de chaque réseau local, les hôtes peuvent communiquer entre eux en utilisant les adresses physiques (ex: à l'intérieur du réseau R1, tous les hôtes sont repérés par leurs adresses physiques AP1, AP2, AP3. Ils peuvent donc communiquer en utilisant ces adresses, par les mécanismes du niveau gestion du lien physique). En revanche, les hôtes du réseau R1 ne peuvent pas forcément connaître les adresses physiques du réseau R2. En effet, ces adresses dépendent des composants de liaison dont la gestion ne dépend pas forcément de l'administrateur de R1. De plus, le format de ces adresses, dépendant de la technologie de liaison employée pour R2, n'est pas forcément compatibles avec les mécanismes de liaison disponibles sur R1.

Un hôte de R1 ne peut donc pas utiliser l'adresse physique d'un hôte de R2 pour lui adresser des données. En revanche, il peut utiliser son adresse logique, qui ne dépend pas de la technologie de liaison et qui est valable sur tout l'ensemble des réseaux interconnectés. Le mécanisme est le suivant:

- L'hôte de R1 expédie sur le média de R1 un message contenant l'adresse logique du destinataire (ex: AL4) et l'adresse physique du routeur sur le réseau R1 (AP7).
- Le routeur reconnaît son adresse physique dans l'enveloppe de la couche liaison des trames. Il récupère alors (dans l'enveloppe de la couche réseau) l'adresse logique du destinataire (AL4).
- Le routeur maintient en permanence une table de correspondance (**TABLE DE ROUTAGE**) entre les adresses physiques et les adresses logiques de chaque réseau, grâce à un protocole particulier appelé **A.R.P.(Address Resolution Protocol)**. Il est donc capable de déterminer quelle adresse physique correspond à l'adresse logique AL4 (c'est-à-dire AP4). Il remplace donc dans les trames l'adresse physique AP7 par AP4, puis il émet le message sur le réseau R2.
- L'hôte d'adresse physique AP4 est donc en mesure de recevoir ce message, puisqu'il est bien adressé à son adresse physique.
- S'il existait plusieurs routeurs entre l'émetteur et le destinataire, chaque routeur intermédiaire, constatant que le destinataire n'est pas dans sa table de routage, adresserait le message au routeur suivant, jusqu'à ce que le réseau du destinataire soit atteint.

V.5.5.TCP-IP ET LE MODELE CLIENT-SERVEUR:

V.5.5.1.RAPPELS SUR LE MODELE CLIENT-SERVEUR ET CONSEQUENCES:

Nous avons vu que sur internet, la communication fonctionnait suivant le modèle **CLIENT-SERVEUR**, dont nous rappellerons les particularités suivantes:

- Lors d'une transaction entre client et serveur, c'est toujours le client qui prend contact avec le serveur et non l'inverse.
- Chaque client, en fonction de ses besoins, établit un contact avec le serveur afin de lui adresser une demande de service appelée **REQUETE**. Lorsque le serveur a satisfait cette requête, le client peut rompre le contact et il n'est pas nécessaire que le serveur conserve une trace de lui.

De ce qui précède, on peut déduire qu'un client peut être une **entité éphémère**: un processus est lancé dans une machine, ce processus établit le contact avec le processus serveur, demande un service, puis rompt le contact et disparaît. En revanche, le processus serveur doit être une **entité durable**, car il faut qu'il reste à l'écoute des demandes des clients sur ses différents ports.

De ce fait, la machine supportant le serveur doit impérativement posséder à tout moment une **adresse logique unique et fixe** sur le réseau: dans le cas contraire, aucun client ne pourrait la contacter. En revanche, un client n'a besoin de posséder une adresse logique que **durant la transaction avec le serveur**. Celui-ci aura alors besoin de cette adresse pour lui renvoyer les messages nécessaires à la transaction). L'adresse du client peut donc être envoyée au serveur uniquement au moment de la prise de contact. Elle peut varier à chaque transaction.

SUJET: COMPRENDRE INTERNET	REFERENCE: ATLANTIC/DOCTECH/ Initiation à l'informatique/Comprendre Internet Mise à jour: 11/01/11
---	---

Au moment de l'établissement du contact du client avec le serveur, celui-ci a besoin de connaître le numéro de port sur lequel on peut établir la communication avec ce serveur. Ce numéro de port doit donc rester disponible en permanence. En revanche, le numéro de port du processus client peut être créé uniquement au début de la transaction et transmis au serveur lors de l'établissement du contact avec celui-ci.

V.5.5.2.CAS DE TCP-IP ET DU WEB:

TCP-IP permet la communication suivant un modèle CLIENT-SERVEUR. En effet:

- Lors de l'ouverture d'une communication, un processus donné peut être déclaré CLIENT ou SERVEUR.
- Un processus déclaré CLIENT peut communiquer uniquement avec un SERVEUR.
- Un SERVEUR peut traiter simultanément plusieurs CLIENTS.

Lorsqu'un processus se déclare en tant que SERVEUR TCP, il se met immédiatement en attente de DEMANDES DE CONNEXION d'éventuels CLIENTS TCP. Lorsqu'il détecte une telle demande, il renvoie à ce client un message d'acceptation, en utilisant l'adresse IP et le numéro de port trouvé dans la demande de connexion. La communication entre le client et son serveur peut alors se dérouler, jusqu'à la survenue d'un message de déconnexion de la part du client. Après déconnexion, le serveur perd toute trace du client, et celui-ci devra se reconnecter pour pouvoir de nouveau communiquer avec le serveur. Plusieurs clients peuvent être connectés simultanément au même serveur.

Les hôtes internet qui offrent des services permanents, tels que l'accès à des sites web, le téléchargement de fichier ou la messagerie électronique sont déclarés en tant que SERVEURS (serveur web, serveurs de messagerie, etc.), alors que les logiciels des usagers fonctionnent en tant que CLIENTS (votre navigateur est un CLIENT WEB, outlook est un CLIENT DE MESSAGERIE, etc.).

V.5.6.ATTRIBUTION DES ADRESSES I.P. AUX HOTES D'INTERNET:

V.5.6.1.ATTRIBUTION DES ADRESSES IP AUX POSTES CLIENTS:

En vertu des remarques exprimées plus haut, les hôtes internet abritant uniquement des logiciels CLIENTS ne sont obligés de disposer d'une adresse IP **valable sur le net** que durant leur connexion à un SERVEUR. En dehors de ces périodes, elle leur est inutile. Une adresse IP leur est fournie par leur **Fournisseur d'Accès Internet (F.A.I.)** lors de leur connexion à celui-ci.

Dès que le poste client s'est vu attribuer cette adresse IP par son F.A.I, il devient visible sur le net comme n'importe quel hôte, et ceci, jusqu'à ce qu'il se déconnecte de son F.A.I. Il peut donc dialoguer sur le net en utilisant TCP-IP, et même, à la rigueur, lancer un serveur chez lui.

Une fois le client déconnecté, il perd son adresse IP et le F.A.I. peut réattribuer celle-ci à un autre hôte. Un client change donc d'adresse IP chaque fois qu'il se connecte (ou reconnecte) à son F.A.I.

REMARQUE:

On estime qu'un poste informatique quelconque créant une connexion TCP-IP sur internet dispose d'une dizaine de minutes en moyenne avant de subir une attaque quelconque. Or, mener une telle attaque exige d'utiliser l'adresse IP du poste. Il n'est donc pas très raisonnable, même si l'on dispose d'un pare-feu et d'un anti-virus performant et à jour, de laisser son ordinateur connecté en permanence à internet: quelques déconnexions-reconnexions de temps à autres permettent de changer l'adresse IP du poste, ce qui complique considérablement le travail d'éventuels " hackers ".

V.5.6.2.ATTRIBUTION DES ADRESSES IP AUX POSTES SERVEURS:

En revanche, les postes abritant des logiciels serveurs doivent obligatoirement disposer d'adresses IP fixes. Il est donc difficile de connecter ces postes par l'intermédiaire de fournisseurs d'accès. De ce fait, les postes serveurs sont la plupart du temps reliés directement aux dorsales internet par des lignes dédiées.

L'attribution d'adresses IP permanentes sur internet est contrôlée par des organismes nationaux et internationaux de régulation. Chaque fournisseur d'accès dispose d'un certain nombre d'adresses disponibles pour les clients qui accèdent à internet par son intermédiaire. Ce nombre étant forcément limité, il est possible qu'à certaines heures de forte utilisation du réseau, votre F.A.I. n'ait plus d'adresse disponible: ceci peut expliquer certains refus de connexion ou même certaines déconnexions intempestives.

VI.MECANISMES DES SERVICES APPLICATIFS:

VI.1.INTRODUCTION:

Une partie des mécanismes de ce niveau (qui correspond aux couches 5, 6 et 7 du modèle O.S.I) dépend de la nature des applications utilisées: les navigateurs englobent des mécanismes évidemment différents de ceux d'un client de messagerie ou d'un logiciel de téléchargement de fichiers. Cependant il existe des mécanismes communs à ces différentes applications. Leur étude fait l'objet du chapitre en cours.

VI.2.MECANISMES DE CONTRÔLE DE SESSION:

Le modèle O.S.I. assigne ces mécanismes à la couche n° 5 (Couche SESSION). Ils concernent notamment:

- L'organisation et la synchronisation des échanges entre les processus en cours de communication (qui doit parler, qui parle, etc).
- L'insertion dans le flot des données échangées de POINTS DE REPRISE permettant de reprendre un échange interrompu par un dysfonctionnement au point où il en était avant l'interruption (ou tout au moins, pas très loin de ce point).

Ces mécanismes sont le plus souvent inclus dans les applications réseau, qui peuvent cependant faire appel à des produits existants (protocoles de gestion de session).

VI.3.MECANISMES DE PRESENTATION DES DONNEES:

La manière dont une donnée est représentée sous forme binaire peut varier d'un type de machine à un autre, mais aussi, d'un système d'exploitation un autre. Par exemples:

- Un nombre entier n'est pas représenté de la même manière dans un processeur INTEL et dans un processeur RISC.
- Il existe différentes façons de représenter un nombre fractionnaire.
- Les fins de lignes d'un fichier de texte ne sont pas codées de la même manière sous windows et sous linux.
- Les données transmises peuvent être cryptées ou compressées.

Les mécanismes de présentation (couche n° 6 de l'OSI) permettent de rendre les informations échangées compréhensibles par chacun des processus communicants, grâce à des opérations de conversion de format, cryptage-décryptage, compression-décompression, etc.

Comme pour la couche session, ces mécanismes sont le plus souvent inclus dans les applications réseau, qui peuvent cependant faire appel à des produits existants (logiciels de cryptage, par exemple).

VI.4.SYSTEME DES NOMS DE DOMAINES:

VI.4.1.INTRODUCTION:

Nous avons vu au chapitre précédent que le protocole TCP-IP utilisait un système d'adressage logique des HOTES, indépendant de la technologie des réseaux: l'ADRESSAGE IP. Une adresse IP peut être utilisée telle quelle pour adresser un HOTE internet à partir d'une application réseau (par exemple, taper l'adresse IP 209.85.229.147 dans la barre d'adresse de votre navigateur vous connecte au serveur de www.google.fr).

Or, lorsqu'il est utilisé au niveau applicatif, ce système présente au moins deux inconvénients:

- La capacité d'adressage des adresses IP V4 ne permet pas d'attribuer une adresse spécifique et fixe à tous les hôtes du réseau mondial. Le système IP V6 en serait capable, mais il ne sera pas entièrement déployé avant quelques années.
- Le format des adresses I.P. est entièrement numérique et basé sur des impératifs techniques. De ce fait, leur utilisation est peu pratique pour les utilisateurs des services du web: elles ne se mémorisent pas facilement et

SUJET:
COMPRENDRE
INTERNET

REFERENCE: ATLANTIC/DOCTECH/
Initiation à l'informatique/Comprendre Internet
Mise à jour: 11/01/11

n'obéissent pas à une logique en rapport avec les besoins des utilisateurs.

De ce fait, le niveau des services applicatifs superpose au système IP un système d'adressage au format beaucoup plus compréhensible et facile à mémoriser, qui, de plus, n'est pas limité en capacité: le **SYSTEME DES NOMS DE DOMAINES** (ou, en anglais, **DOMAIN NAMES SYSTEM: D.N.S.**).

VI.4.2.PRINCIPES GENERAUX DU SYSTEME DES NOMS DE DOMAINES:

REGLE N° 1:

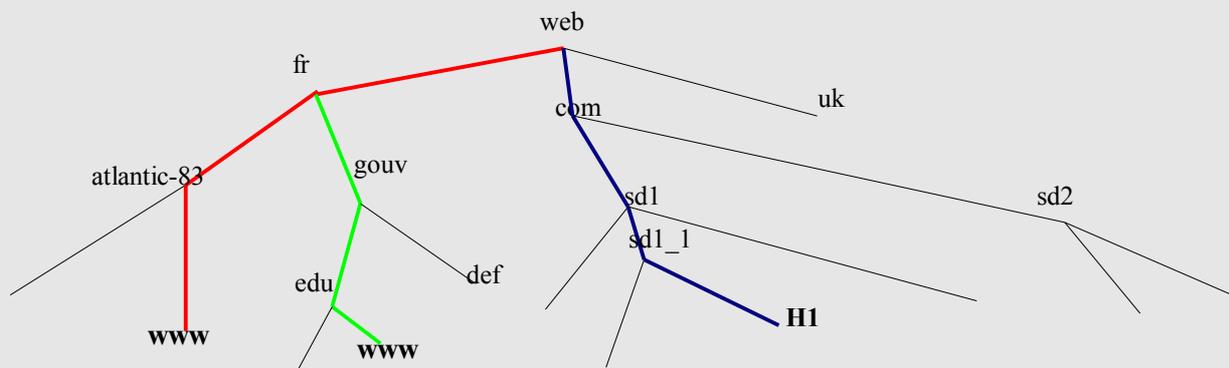
Un **NOM DE DOMAINE** définit toujours un **HÔTE** (ordinateur connecté) ou un **GROUPE D'HÔTES** d'internet.

EXEMPLE: www.atlantic-83.fr est un nom de domaine qui désigne l'hôte du serveur hébergeant le site internet ATLANTIC.

REGLE N° 2:

La notion de domaine est hiérarchique: elle permet de structurer l'ensemble des hôtes du web suivant une arborescence de domaines et sous-domaines inclus les uns dans les autres, jusqu'à ce qu'on arrive au niveau de chaque hôte. Le domaine racine de cette arborescence est l'ensemble des hôtes du web. Les nœuds non terminaux sont les noms des sous-domaines successifs englobant l'hôte ou le groupe d'hôtes ciblé. Le nœud terminal correspond à un hôte ou à un groupe d'hôtes.

EXEMPLE:



Schema n° VI.4.2: le système des noms de domaines.

REGLE N° 3

Un **NOM DE DOMAINE** se présente sous la forme d'une chaîne de caractères constituée de labels littéraux séparés par des points.

EXEMPLE: www.atlantic-83.fr est composé de 3 labels: fr, atlantic-83 et www.

La suite des labels décrit, **de droite à gauche**, le chemin d'accès à l'hôte ou au groupe d'hôtes désigné par le nom de domaine dans l'arborescence du web.

EXEMPLES: Dans l'arborescence du schéma n° VI.4.2:

- sd1_1.sd1.com désigne le sous-domaine sd1_1, inclus dans le sous-domaine sd1, lui-même inclus dans le sous-domaine com du web.
- H1.sd1.sd1_1.com désigne l'hôte H1 du sous-domaine sd1_1.
- www.atlantic-83.fr désigne l'hôte www du sous-domaine atlantic-83.
- www.edu.gouv.fr désigne l'hôte www appartenant au sous-domaine edu (éducation nationale) du domaine gouv, dédié aux administrations françaises.

SUJET: COMPRENDRE INTERNET	REFERENCE: ATLANTIC/DOCTECH/ Initiation à l'informatique/Comprendre Internet Mise à jour: 11/01/11
---	---

VI.4.3.GESTION DES NOMS DE DOMAINES:

Chaque domaine est géré par une organisation administrative. Cette organisation peut créer des sous-domaines de son domaine, qu'elle peut gérer elle-même ou dont elle peut déléguer la gestion à d'autres autorités.

Les domaines de premier niveau (ceux dont le nom s'écrit tout à droite de la chaîne de caractères) sont appelés **Top Level Domains (TLD)**. La gestion des TLD est supervisée par l'**ICANN** (Internet Corporation for Assigned Names and Numbers), qui délègue la gestion des domaines associés à chaque TLD à des bureaux d'enregistrement (REGISTRY).

REMARQUE:

Certains TLD sont appelés **CODES COUNTRY**, car ils font, en principes, référence à des hôtes rattachés à un pays déterminé (fr pour France, uk pour United Kingdom, etc.). D'autres sont dits **GENERIQUES** (comme .com ou .org).

EXEMPLE:

L'organisation en charge du domaine .fr (AFNIC-Association Française pour le Nommage Internet en Coopération) tient cette délégation de l'ICANN. L'AFNIC délègue la gestion du sous-domaine gouv du domaine fr à une administration gouvernementale.

VI.4.4.RESOLUTION DES NOMS DE DOMAINES:

Lorsqu'un nom de domaine correspond à un hôte connecté à internet, on peut lui associer l'ADRESSE IP de cet hôte.

REMARQUE: HOTES VIRTUELS

Il n'y a cependant pas identité entre nom de domaine et adresse IP. En effet, les logiciels serveurs peuvent associer plusieurs noms de domaines à la même adresse IP. Chacun de ces domaines constitue alors un **HOTE VIRTUEL**. C'est le cas, en particulier, lorsqu'un même serveur web héberge plusieurs sites (cas assez fréquent pour des hébergements peu onéreux): chacun de ces sites pourra avoir un nom de domaine indépendant des autres, mais qui pointerait sur la même adresse IP. Chacun des domaines correspondra à un **HOTE VIRTUEL**.

Pour les raisons exposées plus haut, le système des noms de domaines est utilisé, de préférence à l'adressage IP, par les applications réseaux " grand public " (lorsque vous vous connectez à un site web, c'est bien le nom de domaine du serveur qui le supporte que, la plupart du temps, vous entrez dans la barre d'adresse de votre navigateur). Cependant, lorsqu'une de ces application veut échanger des informations avec un hôte distant, elle est obligée d'utiliser les mécanismes de la couche 3 de l'O.S.I (couche réseaux), qui, eux, n'exploitent que les adresses IP. Il faut donc disposer d'un moyen pour déterminer, à partir d'un nom de domaine, l'adresse IP qui lui est éventuellement associée. Cette opération est appelée **RESOLUTION** du nom de domaine.

La résolution d'un nom de domaine est effectuée grâce à l'utilisation de serveurs particuliers, les **SERVEURS DE NOMS DE DOMAINES** (en anglais: **DOMAIN NAMES SERVER, D.N.S**) et d'un protocole particulier appelé **DOMAIN NAME SERVICE**.

Les serveurs de noms de domaines sont des logiciels qui maintiennent à jour des tables de correspondance entre les noms de domaines attribués et les adresses IP correspondantes, pour une zone du web donnée. Les serveurs des différentes zones peuvent communiquer entre eux. Lorsqu'une application réseau a besoin de résoudre un nom de domaine, elle interroge un des serveurs de noms de domaines de sa zone. Si celui-ci est capable de résoudre le nom du domaine en question, il renvoie l'adresse IP vers le demandeur. Sinon, il met en jeu une procédure d'interrogation en cascade des autres serveurs disponibles (à l'échelle de tout le web !) jusqu'à ce que le nom de domaine soit résolu et l'adresse IP correspondante retransmise au demandeur, ou que la recherche soit déclarée infructueuse.

D'autre part, chaque hôte entretenant sa propre table de correspondance entre les noms de domaines qu'il a résolu et les adresses IP correspondantes, l'appel aux serveurs DNS n'est en général nécessaire que lors de la première utilisation du nom de domaine dans la session d'exploitation en cours.

SUJET: COMPRENDRE INTERNET	REFERENCE: ATLANTIC/DOCTECH/ Initiation à l'informatique/Comprendre Internet Mise à jour: 11/01/11
---	---

VI.5.LES U.R.L.:

VI.5.1.DEFINITION:

Les **U.R.L.** (de l'anglais Uniform Resource Locator), sont des identifiants textuels complexes qui permettent de localiser une **RESSOURCE** sur le web.

*Dans cette définition, la notion de **RESSOURCE** concerne toute entité qu'une application web peut être amenée à utiliser dans le cadre de son exécution: documents (textes, images, vidéo, etc), sites web, serveurs ftp, boîtes aux lettres électroniques, etc.*

Une U.R.L. peut donc être définie comme l'adresse web d'une ressource. Les appellations officiellement admises par l'administration française sont **ADRESSE UNIVERSELLE** ou **ADRESSE RETICULAIRE**, ce qui explique qu'en français, URL soit du genre féminin.

REMARQUE:

Nous verrons par la suite qu'une U.R.L.est beaucoup plus qu'une simple adresse, car elle définit également la manière d'utiliser la ressource et peut même transporter des paramètres. Il serait donc plus juste de définir une U.R.L. Comme un " activateur de ressource ".

VI.5.2.STRUCTURE D'UNE U.R.L.:

Globalement, une URL se présente comme une chaîne de caractères séparée en différents champs littéraux par des caractères de séparation. Le format général d'une U.R.L. peut être décrit comme suit (les différents champs sont représentés entre les signes < et >). Les caractères séparateurs sont figurés en rouge:

```
<protocole>://<id. utilisateur>:<mot de passe>@<nom de domaine ou adresse IP>:  
<n° de port>/<chemin d'accès >?<données optionnelles>#<signet ou ancre>
```

- **Champ <protocole>:** il définit le protocole de communication à utiliser pour accéder à la ressource (exemple: pour accéder à une page web, le protocole est http). S'il est omis, le séparateur " : " peut être également omis. Certaines applications réseau peuvent déterminer le protocole à utiliser en fonction de la forme de l'URL.
- **Champs <id.utilisateur et mot de passe>:** ils précisent le login et le mot de passe de l'utilisateur, quand l'accès à la ressource le nécessite. Si seul le login est transmis, on omettra le séparateur " : ".
- **Champ <nom de domaine ou adresse IP>:** ce champ permet d'identifier l'hôte hébergeant la ressource.
- **Champ <numero de port>:** permet de préciser le numéro de port TCP-IP correspondant au serveur auquel le protocole de communication s'adresse. S'il est omis, un port standard caractéristique du type de serveur sera utilisé (port 80 pour un site web, port 21 pour un serveur FTP) et on omettra également le séparateur " : ".
- **Champ <chemin d'accès>:** l'accès à la ressource se traduisant la plupart du temps par l'accès à un fichier, ce paramètre décrira le chemin d'accès à ce fichier depuis la racine du serveur.
- **Champ <données optionnelles>:** en fonction de la nature de la ressource, ce champ va permettre de transmettre des données optionnelles (par exemple, des paramètres destinés à être appliqués à cette ressource).
- **Champ <signet ou ancre>:** permet de spécifier un emplacement particulier à l'intérieur de la ressource (exemple: un endroit particulier d'une page web).

En fait, la plupart des champs ne sont pas utilisés par les utilisateurs courants du web (ils concernent surtout les développeurs d'applications web ou les administrateurs réseau). Pour la plupart des utilisateurs, une U.R.L. se résume aux champs suivants:

```
<protocole>://<nom de domaine ou adresse IP>/<chemin d'accès à la ressource>
```

Une U.R.L. peut même se résumer à un simple nom de domaine (exemple: www.atlantic-83.fr) ou au chemin d'accès à un fichier de la machine locale (exemple sous windows: c:\donnees\infos.doc).

VI.5.3.UTILISATION DES U.R.L. CONCEPT D'HYPERLIEN:

Les U.R.L. sont très utilisées dans les applications réseau sur internet, chaque fois qu'il s'agit de localiser une ressource. Elles sont à la base du mécanisme des **HYPERLIENS**.

DEFINITION:

Un **HYPERLIEN** est constitué par:

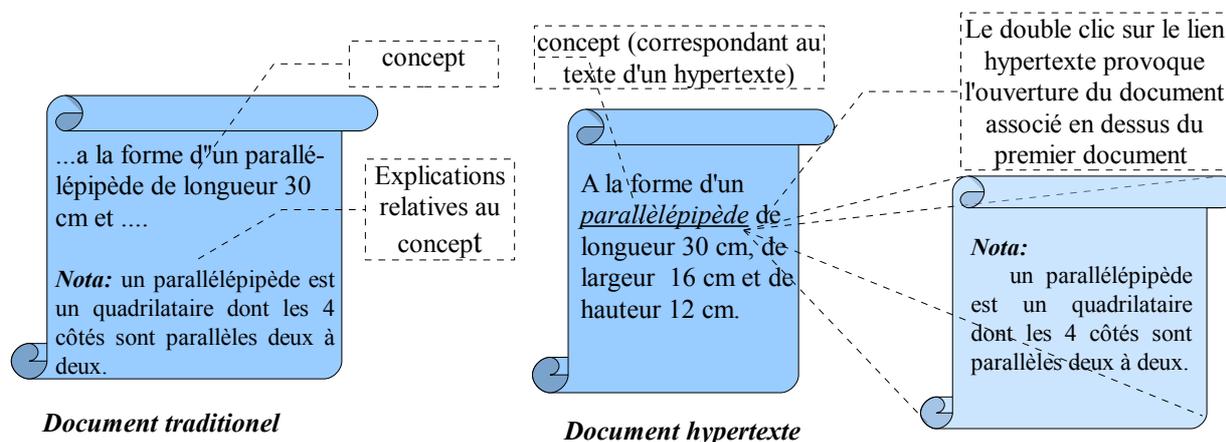
- Un **objet graphique** inséré dans une fenêtre d'interface graphique affichée par un logiciel interactif (navigateur web, traitement de texte, etc.).
- Une **U.R.L.** qui lui est associée.

Un hyperlien peut être une chaîne de caractère, une zone d'une photographie, un dessin, ou tout autre élément graphique affichable sur l'écran. L'U.R.L. associée peut pointer vers n'importe quelle ressource et pas uniquement vers un document texte.

COMPORTEMENT:

Un double clic de la souris sur la zone de l'écran qu'occupe l'hyperlien provoque l'ouverture de la ressource pointée par l'U.R.L. et l'affichage du contenu de cette ressource, en fonction de son type MIME (type d'extension du fichier: .doc, .odt, .pdf, etc.). Il faut, évidemment que le poste local soit équipé des logiciels nécessaires au traitement de ce type de ressource).

En général, le contenu de cette ressource est choisi parce qu'il apporte des renseignements supplémentaires sur l'entité ou le concept représenté par l'hyperlien. De ce fait, par rapport à un document traditionnel que l'on ne peut parcourir que dans deux dimensions (la largeur et la hauteur de chaque page, y compris dans la séquence des pages), les liens hypertextes permettent, en quelque sorte, d'accéder à une dimension supplémentaire, la " profondeur" représenté par le contenu de la ressource associée. Ceci explique leur nom.



Schema n° VI.5.3: lien hypertexte.

Si l'hyperlien pointe sur une ressource externe à la machine locale, le navigateur internet par défaut sera lancé par le système d'exploitation, puis l'URL associée à l'hyperlien sera chargée dans la barre d'adresse.

EXEMPLES:

- Pour accéder à l'image de fond du portail du site atlantic, qui est le fichier photographique Cote.jpg, il suffit de saisir l'URL <http://www.atlantic-83.fr/SiteAtlantic/images/Cote.jpg> dans la barre d'adresses de votre navigateur, ou bien de cliquer sur le lien: <http://www.atlantic-83.fr/SiteAtlantic/images/Cote.jpg>. Le fichier Cote.jpg est une ressource web hébergée par l'hôte dont le nom de domaine est www.atlantic-83.fr. Le fichier se trouve dans le

SUJET:
COMPRENDRE
INTERNET

REFERENCE: ATLANTIC/DOCTECH/
Initiation à l'informatique/Comprendre Internet
Mise à jour: 11/01/11

sous-répertoire images, lui-même inclus dans le sous-répertoire SiteAtlantic du répertoire racine de l'hôte.

- Pour accéder au site atlantic lui-même, il suffira d'activer l'URL suivante: <http://www.atlantic-83.fr>. En effet, en l'absence du chemin d'accès de la ressource à l'intérieur de l'hôte, cette URL sera interprétée par le serveur comme s'adressant à un fichier de nom " index ", se trouvant dans le répertoire racine de l'hôte. L'URL sera donc, dans ce cas, équivalente à <http://www.atlantic-83.fr/index.php>.
- Pour accéder au fichier Fic.txt situé sur la racine du disque dur local d'un système windows, il suffira de créer un lien dont l'URL sera tout simplement: <c:\Fic.txt>.

VII.LES SERVICES DU NIVEAU APPLICATIF:

VII.1.NAVIGATION SUR LE WEB

VII.1.1.DEFINITIONS:

Le terme familier " Naviguer sur le web " correspond à l'action de consulter le contenu de certaines des **PAGES WEB** accessibles sur internet en les affichant sur l'écran du terminal que l'on utilise, dans le but d'y rechercher des services bien précis, mais aussi sans objectif défini, avec pour simple motivation la curiosité.

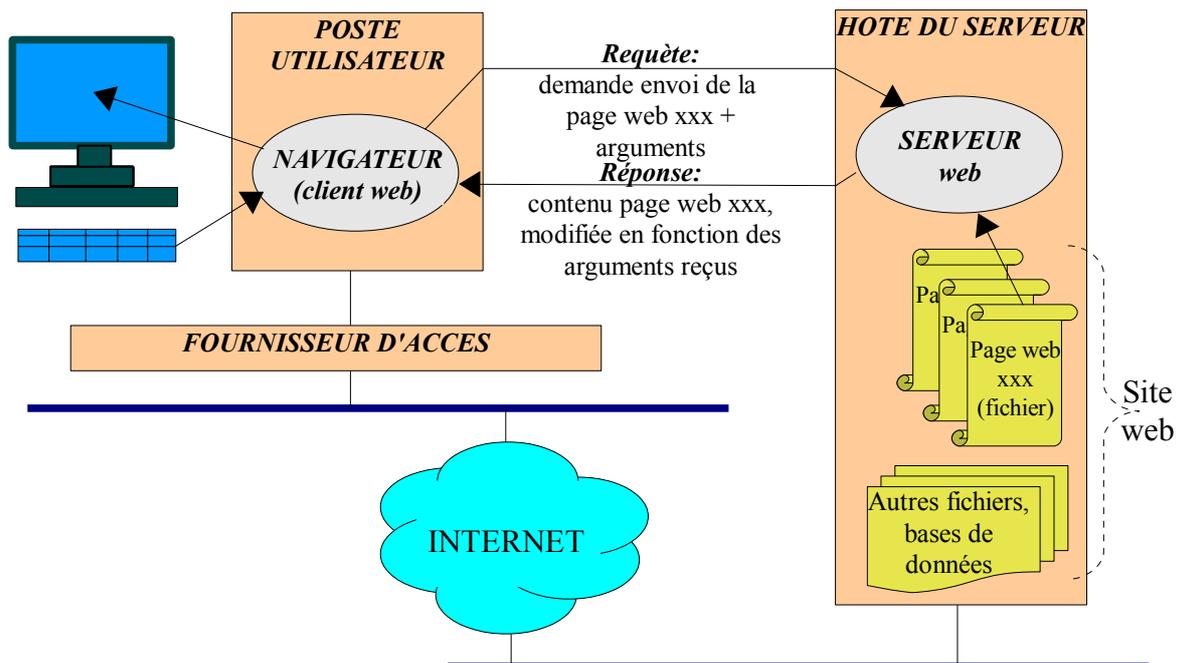
La notion de **NAVIGATION** recouvre ainsi un certain nombre de mécanismes qui permettent d'instaurer trois modes d'accès aux pages web:

1. Un mode que l'on peut qualifier de **DETERMINISTE**, dans lequel l'utilisateur fournit directement l'U.R.L. de la page qu'il veut consulter à son navigateur (soit en saisissant directement cette U.R.L., soit en utilisant le mécanisme des FAVORIS).
2. Un autre mode que l'on peut qualifier d'**ALEATOIRE** dans lequel l'utilisateur accède à différentes pages web en activant des hyperliens liés aux objets graphiques qu'il rencontre dans ces pages, au fil de sa recherche.
3. Enfin, un troisième mode, que nous pourrions qualifier de **THEMATIQUE**, dans lequel l'utilisateur fait appel à des mécanismes particuliers appelés **MOTEURS DE RECHERCHE** pour sélectionner les pages web correspondant à ses **CRITÈRES DE RECHERCHE**.

VII.1.2.FONCTIONNEMENT DE BASE

Naviguer sur le web nécessite l'utilisation sur le terminal d'accès d'un type particulier de logiciel appelé **NAVIGATEUR** en français. Le terme **BROWSER** (to browse = feuilleter, regarder, brouter...) est utilisé en anglais).

Un **NAVIGATEUR** permet de communiquer avec un autre type de logiciel appelé **SERVEUR WEB** suivant un schéma de communication du type CLIENT-SERVEUR illustré par le schéma ci-dessous. Un **NAVIGATEUR** est donc un **CLIENT WEB**.



Schema n° VII.1.2 : Mécanisme de consultation des pages web

SUJET:
COMPRENDRE
INTERNET

REFERENCE: ATLANTIC/DOCTECH/
Initiation à l'informatique/Comprendre Internet
Mise à jour: 11/01/11

COMMENTAIRES:

- Pour consulter une page web particulière, l'utilisateur doit rapatrier dans son terminal le contenu de cette page (qui n'est autre qu'un fichier hébergé par un hôte du web), afin de l'afficher sur son écran. Pour ce faire, le navigateur envoie vers l'hôte qui héberge cette page une requête particulière. Pour que ce traitement puisse aboutir, il faut que l'hôte hébergeur soit muni d'un logiciel **SERVEUR WEB**, capable d'interpréter et de satisfaire ce type de requête.
- En réponse à la requête du client web (navigateur), le serveur web renvoie vers celui-ci le contenu de la page web demandée, éventuellement modifié en fonction de la valeur des arguments contenus dans la requête.
- Le navigateur peut alors afficher le contenu de la page web reçue dans sa fenêtre d'affichage.

Les informations contenues dans les requêtes expédiées comprennent essentiellement:

- L'URL de la page web demandée
- Le type de la requête (il en existe plusieurs types).
- Des données liées au type de la requête.

L'URL de la page web demandée comprend:

- Le nom de domaine du site web auquel la page appartient
- Le chemin d'accès à la page web à partir de la racine du site
- Dans la partie " données optionnelles ", une liste d'arguments d'appel: ce sont eux qui vont permettre de " customiser " le contenu de la page web avant d'envoyer celui-ci au client, permettant ainsi de créer des pages " interactives ".

L'envoi des requêtes est déclenché par les actions de l'utilisateur sur l'interface du navigateur. Une requête d'appel de page web peut ainsi résulter des actions suivantes:

1. Saisie de l'URL de la page dans la **BARRE D'ADRESSES** du navigateur (suivie d'un appui sur ENTREE).
2. Activation d'un **HYPERLIEN** inclus dans la pages web affichée dans la fenêtre du navigateur (au moyen du classique " double-clic ")
3. Activation d'autres objets graphiques que sont les boutons de type **SUBMIT** des **FORMULAIRES** inclus dans la page affichée. Comme les hyperliens , les boutons SUBMIT (validation et envoi des valeurs saisies dans le formulaire) sont associés à l'U.R.L. d'une page web.
4. Double-clic sur un **FAVORIS** (un favoris n'est autre qu'un hyperlien).
5. etc.

NOTA: D'autres actions peuvent également déclencher indirectement l'envoi de requêtes d'appel de pages.

VII.1.3.STRUCTURE ET CONTENU D'UNE PAGE WEB:

Une page web reçue par un navigateur est un fichier de texte. Ceci signifie qu'il ne contient que des CARACTERES, affichables ou non sur l'écran. Ce texte décrit dans un langage particulier le **CONTENU TEXTUEL OU GRAPHIQUE** et l'**ASPECT** (disposition, taille et police de caractères, couleurs, etc.) de la page web à afficher sur l'écran.

VII.1.3.1.LANGAGE HTML:

Le langage utilisé pour décrire la structure et le contenu d'une page web est appelé **Hyper Text Markup Language (H.T.M.L)**. Les principales caractéristiques de ce langage sont qu'il utilise des entités appelées **BALISES (MARKUP** en anglais) et qu'il permet de définir des **HYPERLIENS** renvoyant à d'autres documents situés sur le net. Le parcours du lecteur s'effectue donc dans trois dimensions : les deux dimensions de la page en cours et la troisième " dimension " matérialisée par les hyperliens (C'est là probablement que réside la justification de l'appellation " hypertexte ").

Notion de balise :

Une balise se présente sous la forme d'une chaînes de caractères délimitée par des " chevrons " ouvrants et fermants. La chaîne comprise entre les " chevrons " < et > comprend obligatoirement le NOM correspondant au TYPE de la balise, mais peut aussi inclure divers paramètres permettant de préciser le rôle de la balise, suivant le schéma général suivant :

SUJET: COMPRENDRE INTERNET	REFERENCE: ATLANTIC/DOCTECH/ Initiation à l'informatique/Comprendre Internet Mise à jour: 11/01/11
---	---

`<[type de balise] [nom paramètre 1]=[valeur paramètre 1]
[nom paramètre 2]=[valeur paramètre 2]...nom paramètre n]=[valeur paramètre n]>`

EXEMPLES:

A-La balise: ``

est une balise de type "IMG" (image). Elle permet d'inclure dans le document l'image dont l'U.R.L. est définie par l'initialisation du paramètre "SRC" avec la chaîne de caractères:

`"http://www.atlantic-83.fr/SiteAtlantic/images/Cote.jpg".`

B-La balise : `<p class="TypeParagraphe1">`

délimite le début d'un paragraphe de texte dont les attributs d'affichage (police de caractère, taille, couleur de ces caractères, etc.) sont définis par le "style" appelé "TypeParagraphe1" (défini lui-même dans une "feuille de style", mais celles-ci ne sont pas abortées dans cet ouvrage).

Balises ouvrantes et fermantes :

La plupart des balises vont par paires, car elles sont destinées à agir sur un segment du contenu affichable (texte, image, vidéo) qui compose le document. Une balise dite "ouvrante" délimite le début du segment, alors que l'autre balise (dite "fermante"), délimite la fin de ce segment.

EXEMPLE:

Les balises `<p class="TypeParagraphe1">` et `</p>` délimitent le début et la fin d'un paragraphe de texte de style "TypeParagraphe1".

Un couple de balises (ouvrante + fermante) peut définir un "bloc", c'est-à-dire une région de l'écran (en fait, un rectangle), ou bien simplement s'appliquer à une chaîne de caractères. Lorsqu'il définit un bloc, les paramètres de la balise ouvrante (paramètres "class", "id", "style" ou autres) peuvent permettre de spécifier ses dimensions et sa position sur l'écran.

EXEMPLE:

Les balises `<p style="position: absolute; top: 100px ; left: 200px ;">Paragraphe</p>` permettent de positionner un bloc contenant le mot "Paragraphe" à 100 pixels par rapport au haut de l'écran et à 200 pixels par rapport à la gauche de celui-ci.

Un couple de balises ouvrantes et fermantes peut lui-même contenir d'autre balises. Le système des balises permet donc de décrire les contenus sous une forme arborescente.

Balises singulières :

Certaines balises ne vont pas par couple. C'est par exemple le cas de la balise `` (inclusion d'images) que nous avons vue plus haut. C'est aussi le cas de la balise `
` qui permet d'inclure un retour à la ligne dans le texte affiché.

Quelques balises remarquables :

L'utilisation du langage HTML n'entre pas dans le cadre du présent document. La description des balises suivantes n'est donnée que pour matérialiser un peu dans l'esprit du lecteur la notion de page web et de document HTML. Dans le paragraphe suivant, un exemple simple de page web utilisant ces balises sera donné.

SUJET:
COMPRENDRE
INTERNET

REFERENCE: ATLANTIC/DOCTECH/
Initiation à l'informatique/Comprendre Internet
Mise à jour: 11/01/11

FORME	FONCTION
<!doctype>	Définit la version de langage utilisée.
<html> ... </html>	Délimite le début et la fin d'un document HTML
<head> ... </head>	Délimite le début et la fin de l'en-tête d'une page HTML. Le contenu de l'en-tête n'est normalement pas affichable : son rôle est de définir divers paramètres destinés à être exploités par les navigateurs , les moteurs de recherche, etc. (balises <meta>). On peut y trouver également le couple <title></title> qui permet de préciser le nom de la page web .
<body> ... </body>	Délimite le début et la fin de la partie affichable du document.
<p> ... </p>	Délimite le début et la fin d'un paragraphe de texte.
	Permet d'inclure une image dans le document.
<div> ... </div>	Délimite un bloc d'affichage (rectangle) sur le contenu duquel lequel on pourra appliquer des attributs de style, de positionnement, etc.
 	Permet de passer à la ligne dans un texte.
 ... 	Délimite un segment de texte sur le contenu duquel lequel on pourra appliquer des attributs de style.

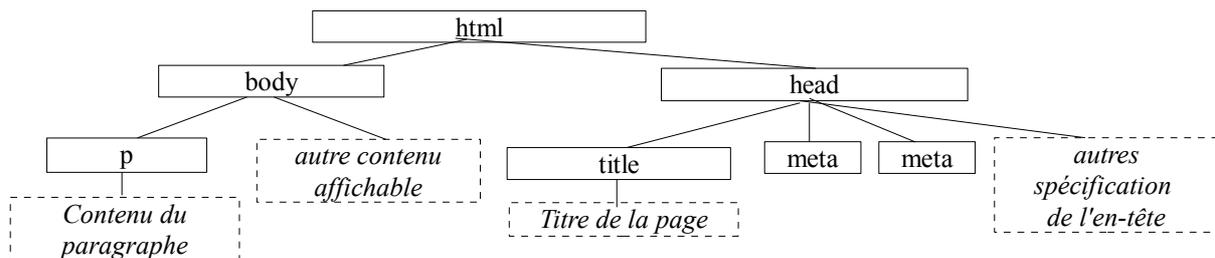
VII.1.3.2.STRUCTURE GENERALE D'UNE PAGE WEB RECUE PAR UN NAVIGATEUR:

Lorsqu'elle est reçue par un navigateur, une page HTML présente en général la structure suivante:

```
<!DOCTYPE .....>
<html>
  <head>
    <title>.....</title>
    <meta.....>
    <meta.....>
    .....
    .....
  </head>
  <body>
    .....
    <p style=.....>
    .....
    </p>
    .....
  </body>
</html>
```

- La balise DOCTYPE permet de définir la version HTML utilisée.
- Les balises <html> et </html> délimitent le début et la fin du contenu.
- Les balise <head> et </head> délimitent les données qui ne sont pas destinées à l'affichage. On y trouve notamment le titre de la page (entre les balises<title> et</title>) ou les balises <meta>, destinées à renseigner les navigateurs des visiteurs ou les moteurs de recherche (référencement, accessibilité, etc.).
- Les balises <body> et </body> permettent de délimiter la partie de la page qui décrit les données à afficher.

Cette description correspond à la structure arborescente suivante:



Schema n° VII.1.3.2 : Arborescence d'une page web

VII.1.4.LE PROTOCOLE HTTP:

Comme son nom l'indique, le protocole **H.T.T.P (Hyper Text Transport Protocol)** a pour utilité principale de permettre l'échange de **DOCUMENTS HTML** entre un **SERVEUR HTTP** et des **CLIENTS HTTP** (Navigateurs). Il définit les requêtes d'appel de pages web effectuées par les navigateurs et celui des réponses correspondantes adressées par les serveurs web aux clients. Le protocole HTTP définit également d'autres requêtes permettant de contrôler le dialogue entre clients et serveur.

Avant la transmission, les requêtes et réponses HTTP sont encapsulées par le niveau applicatif dans des messages au format TCT/IP, qui sont eux-mêmes transmis suivant les modalités de ce protocole. Pour créer ces messages TCP/IP, il est indispensable de connaître l'adresse IP du destinataire des messages. De ce fait, les navigateurs doivent faire appel au mécanisme du Domain Name Server, afin de convertir le nom de domaine contenu dans l'URL en adresse IP. A l'arrivée, les requêtes et réponses HTTP sont extraites des messages TCP-IP pour être interprétées par le niveau applicatif.

Un **SERVEUR WEB** est donc (au minimum) un **SERVEUR HTTP**, tandis qu'un **NAVIGATEUR** est (au minimum) un **CLIENT HTTP**.

NOTA:

Les personnes désireuses d'approfondir ces mécanismes peuvent consulter l'ouvrage: " Réseaux Informatiques ", dans la rubrique " réseaux " de la documentation en ligne du site ATLANTIC.

VII.1.5.NOTION DE SITE WEB:

DEFINITION :

Un **SITE WEB** est constitué par un ensemble **COHERENT** et **ORGANISE** de **PAGES WEB**, hébergées et contrôlées par un serveur web, et **SE PARTAGEANT** un certain nombre de **RESSOURCES** (fichiers de données, bases de données, etc.).

En effet, pour qu'un ensemble de **PAGES WEB** constituent un **SITE WEB**, il est indispensable que la conception globale de cet ensemble prenne en compte certaines exigences spécifiques. Celles-ci dépendent en partie d'impératifs techniques et ergonomiques généraux, mais aussi de la destination du site. Elles concernent (au minimum) les domaines suivants :

- L'aspect général et l'accessibilité (et en particulier, la "charte graphique" du site).
- Le contrôle de l'accès au site (gestion et contrôle des droits d'accès des visiteurs).
- La navigation à l'intérieur de ses pages web (ergonomie d'utilisation).
- Le contrôle des accès aux ressources partagées.

Remarquons qu'une **PAGE WEB** hébergée par un serveur n'est pas forcément un **DOCUMENT HTML** (fichier texte HTML). En effet, nous avons vu plus haut que lors de la transmission d'une page web à un client, le serveur pouvait transformer le contenu de cette page web en fonction des paramètres d'appel contenus dans la requête du client : en fait, dans ce cas, la page web contient, en plus du code HTML, un autre type de code qui, interprété par des logiciels appelés **PREPROCESSEURS**, permet d'adapter le code HTML de la page (supprimer ou modifier ce code ou en créer du nouveau) en fonction de la valeur de ces paramètres, **avant d'expédier cette page vers le client** : c'est ce mécanisme qui permet de rendre un site **INTERACTIF**. Un des codes de préprocesseurs les plus utilisés est le **P.H.P.** (acronyme approximatif de **Hypertext Pre Processor**).

Remarquons également qu'une fois la page web traitée par le préprocesseur, celle-ci se trouve débarrassée du code de préprocesseur : c'est donc bien un document HTML qui est expédié vers le client.

HEBERGEMENT:

Pour pouvoir être consulté, un site web doit être hébergé par un hôte internet (un ordinateur) muni d'un logiciel serveur (serveur HTTP). Le plus répandu de ces logiciels est le produit libre **APACHE**. Le serveur d'un hôte donné peut gérer l'accès à un ou plusieurs sites web hébergés par son hôte.

SUJET: COMPRENDRE INTERNET	REFERENCE: ATLANTIC/DOCTECH/ Initiation à l'informatique/Comprendre Internet Mise à jour: 11/01/11
---	---

ACCES A UN SITE WEB:

A un site web doit être associé un nom de domaine (ou de sous-domaine). En effet, la requête d'appel à la page portail d'un site se fait en spécifiant uniquement son nom de domaine, sans ajouter de chemin d'accès à une page. Par exemple, pour ouvrir la page portail du site ATLANTIC, il suffit de spécifier le nom de domaine du site: www.atlantic-83.fr sans indication sur la page à ouvrir. C'est le serveur qui aiguillera la demande vers la page portail, dont l'URL est: www.atlantic-83.fr/index.php.

HEBERGEMENT COLLECTIF-NOTION D'HOTE VIRTUEL:

Nous avons vu dans les paragraphes précédents qu'un nom de domaine doit correspondre à un ordinateur ou à un groupe d'ordinateurs. Or, nous avons vu également ci-dessus qu'un hôte physique pouvait héberger plusieurs sites, ce qui semble contradictoire. Cette difficulté est résolue de la manière suivante:

Lorsque plusieurs sites sont hébergés par un même ordinateur, chacun de ces sites se voit tout de même attribuer un nom de domaine, correspondant à un **HÔTE VIRTUEL (VIRTUAL HOST** en anglais). Les requêtes destinées aux différents sites hébergés par l'hôte physique sont adressées à la même adresse IP. Le logiciel serveur se charge d'aiguiller ces requêtes vers les hôtes virtuels correspondant à chaque site.

La technique des hôtes virtuels permet donc d'héberger plusieurs sites sur une même machine serveuse, offrant ainsi des hébergements à coût réduit.

VII.1.6. NOTION DE MOTEUR DE RECHERCHE:

Un **MOTEUR DE RECHERCHE** est un logiciel permettant de rechercher sur le net des pages web dont le contenu est susceptible d'aborder un sujet donné.

Concrètement, un moteur de recherche analyse le contenu des pages web afin d'y détecter des **MOTS CLEF**. En fonction de l'ordre dans lequel ces mots clef lui ont été communiqués et du nombre d'occurrences trouvées dans les pages web, celles-ci sont classées, puis présentées sous la forme d'une liste.

Les moteurs de recherche sont en général mis en oeuvre par des sites particuliers, tels que google, yahoo, etc. dont ils ne constituent par l'unique raison d'être. Les mots-clefs sont saisis par l'utilisateur, puis le moteur de recherche est activé. Les résultats sont présentés sous la forme d'une ou plusieurs pages web contenant une liste de liens (chaque lien correspondant à un site trouvé), accompagnés d'informations trouvées dans l'en-tête de ces pages.

VII.1.7. NOTION DE FAVORIS:

Un moyen d'accéder facilement à une page web donnée est d'activer sur son écran un hyperlien vers cette page. Le mécanisme des favoris permet de créer à la demande de tels hyperliens. A cet effet, les navigateurs sont munis de menus spéciaux appelés " favoris " (Internet Explorer), " marque page " (Firefox), etc. qui permettent de créer de tels hyperliens ou de visualiser les hyperliens créés afin de les utiliser.

En général, ces menus permettent de créer un hyperlien vers la **page en cours d'affichage** dans la fenêtre du navigateur. D'autres fonctions permettent de gérer ces favoris (les organiser en répertoires et sous répertoires, les supprimer, les importer d'un autre navigateur, etc.).

VII.1.8. NOTION DE COOKIE:

Un **COOKIE** est un fichier qui peut être créé sur votre disque dur par votre navigateur, lors de la manipulation d'une page web. Il est destiné à conserver un certain nombre de données (sur les sites que vous visitez, sur les rubriques que vous consultez dans ces sites, sur les informations que vous entrez dans les formulaires, etc.).

Lorsque vous vous connectez à une page web, les informations contenue dans vos cookies peuvent être lues par le serveur, qui peut les utiliser pour personnaliser les pages web qu'il vous renvoie. Il peut ainsi orienter votre navigation vers certaines rubriques ou certains sites. L'avantage est que votre navigation peut être accélérée, puisqu'on vous guide

<i>SUJET:</i> <i>COMPRENDRE</i> <i>INTERNET</i>	<i>REFERENCE: ATLANTIC/DOCTECH/</i> <i>Initiation à l'informatique/Comprendre Internet</i> <i>Mise à jour: 11/01/11</i>
--	--

vers vos centres d'intérêt. Les inconvénients sont la possibilité de " manipulation " de vos choix ou de collecte illégale d'informations, mais aussi la saturation de l'espace disque local et le ralentissement de la vitesse d'exécution qui peuvent résulter d'une trop grande accumulation de cookies.

Les navigateurs possèdent tous un menu qui permet à l'utilisateur d'effacer ces cookies ou même d'interdire leur création. Pour ne pas trop dégrader les performances de votre machine, il est recommandé d'effacer les cookies à intervalles réguliers.

VII.2.MESSAGERIE ELECTRONIQUE:

VII.2.1.DEFINITION:

La messagerie électronique est le plus ancien service de niveau applicatif offert par internet. A l'origine, il s'agissait uniquement d'échanger des **MESSAGES DE TEXTES** (appelés **MAIL** en anglais, **COURRIEL** en français) entre des hôtes d'internet. Le service s'est par la suite enrichi de la possibilité d'**ATTACHER** à ces messages des **PIECES JOINTES**, qui peuvent être des fichiers de tous types.

VII.2.2.MECANISME DE BASE:

VII.2.2.1.MAIL TRANSFER AGENT ET BOITE AUX LETTRES ELECTRONIQUES:

Le mécanisme de la messagerie électronique est fondé sur l'existence de serveurs particuliers appelés **Mail Transfer Agents (M.T.A.)** en anglais (que l'on peut traduire en français par agents de transfert de courriel).

Les M.T.A hébergent des entités informatiques appelées **BOITES AUX LETTRES ELECTRONIQUES** ou, plus simplement, **BOÎTES MAIL**. Une BOITE MAIL permet de stocker les courriels destinés à un internaute donné, repéré par un nom d'utilisateur assorti d'un mot de passe.

VII.2.2.2.RECEPTION D'UN MAIL PAR UN M.T.A.:

La transmission d'un mail à un M.T.A. s'effectue en utilisant le protocole **S.M.T.P (Simple Mail Transfert Protocol)**. Ce protocole s'appuie lui-même sur le protocole TCP/IP pour le transport des messages sur le web. Comme l'émetteur d'un mail cible en fait une **BOITE MAIL particulière** hébergée par un M.T.A., l'U.R.L. contenue dans le message comprendra non seulement le **NOM DE DOMAINE** du M.T.A., mais aussi le **NOM DE L'UTILISATEUR** de la boîte mail ciblée. De ce fait, une **ADRESSE MAIL** aura le format suivant:

<nom d'utilisateur>@<nom de domaine>
(Exemple: napoleon.bonaparte@gmail.com)

RAPPEL: dans une U.R.L., le caractère @ sépare les identifiants de connexion de l'utilisateur associé à l'U.R.L du nom de domaine. Une adresse mail n'est donc qu'une U.R.L. partielle.

Lorsqu'un M.T.A. reçoit un courriel destiné à une boîte mail qu'il héberge, il stocke ce mail dans celle-ci. Sinon, il réémet ce courriel vers un autre M.T.A. en utilisant le protocole S.M.T.P.

VII.2.2.3.EMISSION D'UN MAIL VERS SON DESTINATAIRE PAR UN M.T.A.:

Pour retirer un courriel d'une boîte mail hébergée par un M.T.A., il faut utiliser le protocole **P.O.P. (Post Office Protocol)**. Un M.T.A. est un **SERVEUR P.O.P.** Lorsqu'un **CLIENT P.O.P.** envoie au M.T.A. une requête P.O.P. d'interrogation de sa boîte mail, celui-ci lui répond en lui renvoyant les courriels stockés dans cette boîte.

Le protocole P.O.P. comprend également des requêtes permettant la gestion par le client de sa boîte mail (par exemple, suppression d'un mail en attente).

VII.2.3.MODALITES D'ACCES D'UN UTILISATEUR AU COURRIER ELECTRONIQUE:

VII.2.3.1.INTRODUCTION:

Pour un utilisateur d'internet, il existe deux manières d'accéder au courrier électronique:

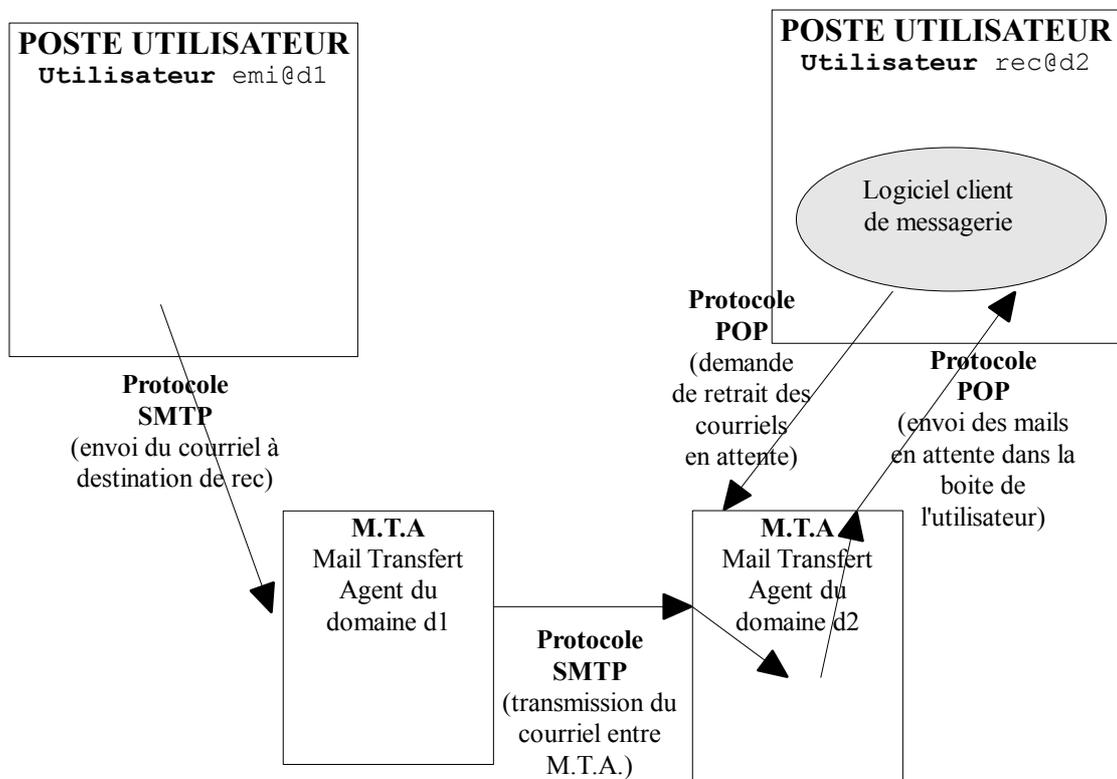
1. Par l'intermédiaire d'un **CLIENT DE MESSAGERIE** local.
2. Par l'intermédiaire d'un **WEBMAIL**.

VII.2.3.2. ACCES PAR UN CLIENT DE MESSAGERIE LOCAL:

Dans ce cas, l'utilisateur accède à la messagerie par l'intermédiaire d'un programme chargé dans son poste internet. Il s'agit d'un **CLIENT DE MESSAGERIE**. Le client de messagerie le plus utilisé est sans doute **OUTLOOK**, sous WINDOWS. Un CLIENT DE MESSAGERIE est à la fois un CLIENT S.M.T.P et un CLIENT P.O.P. Il est donc capable:

- D'une part de transmettre un courriel au M.T.A. de son domaine.
- D'autre part d'envoyer au M.T.A. de son domaine une requête d'interrogation de sa boîte mail, puis de récupérer les courriels en attente dans sa boîte.
- Enfin, de gérer le contenu de sa boîte mail par l'intermédiaire des requêtes P.O.P. adéquates.

Le mécanisme d'utilisation du courrier électronique par l'intermédiaire d'un client de messagerie correspond au schéma suivant:



Schema n° VII.2.3.2 : Accès au mail par l'intermédiaire d'un client de messagerie

COMMENTAIRES:

- Dans ce schéma, le poste utilisateur de gauche (utilisateur emi) veut émettre un courriel vers la boîte de messagerie de l'utilisateur rec (poste utilisateur de droite). Pour cela, il adresse une requête SMTP contenant ce courriel et l'adresse mail du destinataire (rec@d2) au M.T.A. de son domaine (d1).
- A la réception de ce message, le M.T.A. du domaine d1 analyse l'adresse mail du destinataire. Il en déduit que la boîte mail de rec n'est pas hébergée chez lui. Il route donc la requête S.M.T.P. vers le M.T.A. du domaine d2.
- A la réception de la requête, le M.T.A. du domaine D2 détecte qu'elle est adressée à une boîte mail qu'il héberge. Il va donc placer le courriel dans cette boîte.
- Lorsque l'utilisateur rec désire consulter ses courriels, il envoie au M.T.A. qui héberge sa boîte mail (M.T.A. du domaine D2), une requête P.O.P. de demande d'envoi des courriels en attente.
- En réponse, le M.T.A. du domaine D2 lui transmet les courriels en attente dans sa boîte mail.

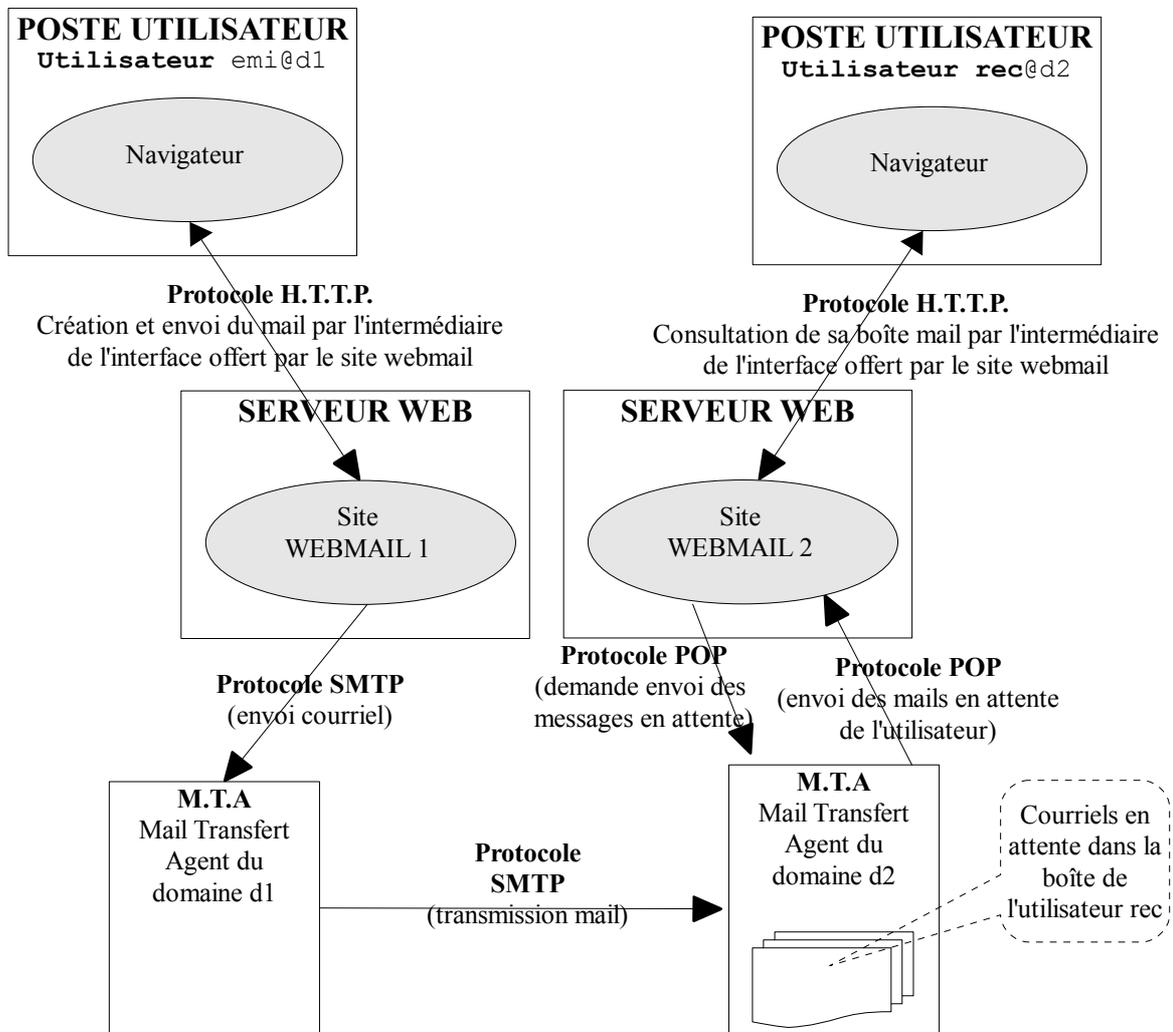
VII.2.3.3. ACCES PAR UN WEBMAIL:

Dans ce cas, l'utilisateur accède à la messagerie par son NAVIGATEUR, en se connectant à un type de site web particulier appelé **WEBMAIL**. GMAIL est un exemple de webmail.

Comme un CLIENT DE MESSAGERIE, un WEBMAIL est à la fois un CLIENT S.M.T.P et un CLIENT P.O.P. Il est donc capable:

- D'une part de transmettre au M.T.A. de son domaine les courriels saisis par ses visiteurs sur leur navigateur.
- D'autre part d'envoyer au M.T.A. de son domaine une requête d'interrogation de la boîte mail d'un visiteur, lorsque celui-ci active cette fonction sur son navigateur, puis de récupérer les courriels en attente et de les afficher sur le navigateur du visiteur.
- Enfin, de permettre à un visiteur de gérer le contenu de sa boîte mail à partir de son navigateur. Les commandes de gestion activées par le navigateur du visiteur sont transformées en requêtes P.O.P.

Le mécanisme d'utilisation du courrier électronique par l'intermédiaire d'un client de messagerie correspond au schéma suivant:



Schema n° VII.2.3.3 : Accès au mail par l'intermédiaire d'un webmail

SUJET: COMPRENDRE INTERNET	REFERENCE: ATLANTIC/DOCTECH/ Initiation à l'informatique/Comprendre Internet Mise à jour: 11/01/11
---	---

COMMENTAIRES:

- Dans ce schéma, comme dans le précédent, le poste utilisateur de gauche (utilisateur emi) veut émettre un courriel vers la boîte de messagerie de l'utilisateur rec (poste utilisateur de droite). Pour cela, il se connecte à un site webmail sur lequel il s'est créé une boîte mail (webmail n° 1), puis il crée le courriel et commande son envoi vers le destinataire par l'intermédiaire d'un formulaire offert par le site.
- Le site webmail n° 1 adresse alors au M.T.A. de son domaine (d1) une requête SMTP contenant ce courriel et l'adresse mail du destinataire (rec@d2) .
- A la réception de ce message, le M.T.A. du domaine d1 analyse l'adresse mail du destinataire. Il en déduit que la boîte mail de rec n'est pas hébergée chez lui. Il route donc la requête S.M.T.P. vers le M.T.A. du domaine d2.
- A la réception de la requête, le M.T.A. du domaine D2 détecte qu'elle est adressée à une boîte mail qu'il héberge. Il va donc placer le courriel dans cette boîte.
- Lorsque l'utilisateur rec désire consulter ses courriels, il se connecte à un site webmail qui gère sa boîte mail (il peut être différent du premier webmail) et demande à consulter ses courriels en attente par l'intermédiaire d'un formulaire offert par le site.
- Le webmail n° 2 envoie alors au M.T.A. qui héberge la boîte mail de rec (M.T.A. du domaine D2), une requête P.O.P. de demande d'envoi des courriels en attente.
- En réponse, le M.T.A. du domaine D2 transmet au site les courriels en attente dans sa boîte mail. Le site peut alors les afficher dans une page web à destination de rec.

SUJET:
COMPRENDRE
INTERNET

REFERENCE: ATLANTIC/DOCTECH/
Initiation à l'informatique/Comprendre Internet
Mise à jour: 11/01/11

VIII.CONCLUSION:

Nous espérons que la lecture de cet ouvrage pourra contribuer à ordonner et à préciser les notions souvent confuses et contradictoires qu'un utilisateur d'internet possède en général sur cet outil.

Tout en adoptant volontairement une démarche qui s'appuie plus sur l'intuition que sur une étude rigoureuse, nous avons tout de même abordé la plus grande partie des problématiques liées aux réseaux et à internet. Les lecteurs qui désiraient approfondir leurs connaissances dans ces domaines pourront consulter l'ouvrage " RESEAUX INFORMATIQUES ", disponible dans la rubrique "réseaux"documentation en ligne offerte par le site ATLANTIC (www.atlantic-83.fr).

Un cours interactif basé sur l'ouvrage présent sera prochainement mis en ligne sur ce même site.