

LT École de Commerce et de Gestion



Nino Silverio

Support de cours provisoire pour l'unité de valeur
"Technologies de l'information et de la
communication" destiné aux classes du BTS de l'ECG.
Il est basé sur le projet Wikipedia [1].

Concepts de base des technologies de l'information et de la communication

Sommaire :

- *Concepts de base sur le fonctionnement d'un ordinateur et son environnement*
- *La sécurité informatique*
- *Les technologies multimédia*
- *Références*



Concepts de base sur le fonctionnement d'un ordinateur et son environnement

Sommaire

- le principe de fonctionnement d'un ordinateur
- les éléments constitutifs usuels d'un PC moderne
- un abécédaire technologique
- l'installation informatique du LTECG

A. Principe de fonctionnement d'un ordinateur [2]

Les technologies utilisées ont énormément changé depuis les années 1940. Toutefois, pour la plupart, elles utilisent les concepts définis par John von Neumann.

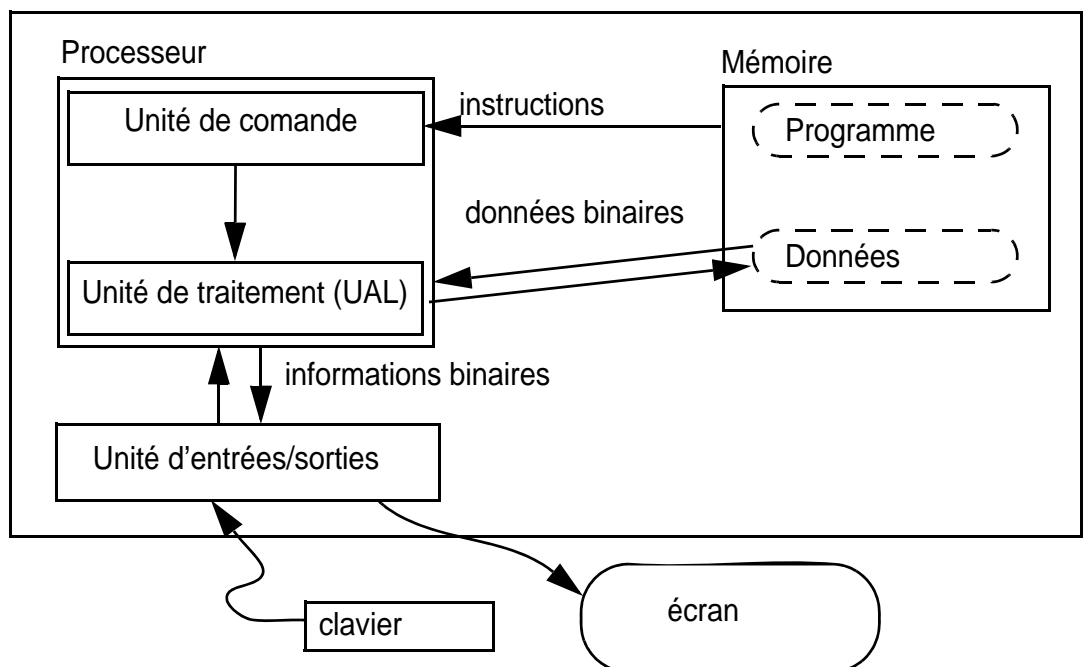
John von Neumann (1903-1957) était un mathématicien américain d'origine hongroise ayant apporté d'importantes contributions autant en physique quantique, qu'en analyse fonctionnelle, en théorie des ensembles, en informatique, en sciences économiques et encore dans beaucoup d'autres domaines.

Il est le père de la théorie des jeux et a publié "Theory of games and economic behavior" avec Oskar Morgenstern en 1944. Von Neumann a conçu l'architecture de von Neumann utilisée dans tous les ordinateurs modernes et a étudié les automates cellulaires afin de construire les premiers exemples d'automates auto-reproductibles (cf. le jeu de la vie).

L'architecture de von Neumann

L'architecture de von Neumann décompose l'ordinateur en 4 parties distinctes

- l'unité arithmétique et logique (UAL) ou unité de traitement : son rôle est d'effectuer les opérations de base ;
- l'unité de contrôle ;
- la mémoire qui se divise entre mémoire volatile et mémoire permanente. Cette dernière est séparée entre mémoire de base au système de contrôle et mémoire demeurant disponible après un arrêt d'alimentation. Elle va contenir à la fois les données et le programme qui dira à l'unité de contrôle quels calculs faire sur ces données ;
- les dispositifs d'entrée-sortie, qui permettront de communiquer avec le monde extérieur.



Mémoire

Au sein du système, la mémoire est une suite de cellules numérotées et contenant chacune une petite quantité d'informations. Cette information peut servir à indiquer à l'ordinateur ce qu'il doit faire (instructions) ou contenir des données à traiter. Dans la plupart des architectures, c'est la même mémoire qui est utilisée pour les deux fonctions.

Cette mémoire peut être réécrite autant de fois que nécessaire. La taille de chacun des blocs de mémoire, ainsi que la technologie utilisée ont varié selon les coûts et les besoins.

Un consensus a fini par se réaliser autour de l'**octet** comme unité adressable, et d'instructions sur format de 4 ou 8 octets.

Les techniques utilisées pour la réalisation des mémoires ont compris des relais électromécaniques, des transistors individuels, des tores de ferrite, et enfin des circuits intégrés incluant des millions de transistors.

Processeur

- L'unité arithmétique et logique ou UAL est l'élément qui réalise les opérations élémentaires (additions, soustractions...), les opérateurs logiques (ET, OU...) et les opérations de comparaison (par exemple la comparaison d'égalité entre deux zones de mémoire). C'est l'UAL qui effectue les calculs de l'ordinateur.
- L'unité de contrôle prend ses instructions dans la mémoire. Celles-ci lui indiquent ce qu'elle doit ordonner à l'UAL, et comment elle devra éventuellement agir selon les résultats que celle-ci lui fournira. Une fois l'opération terminée, l'unité de contrôle passe soit à l'instruction suivante, soit à une autre instruction à laquelle le programme lui ordonne de se brancher.

Entrées-Sorties

Les dispositifs d'entrée/sortie permettent à l'ordinateur de communiquer avec l'extérieur. Le nombre de ces dispositifs est très important, du clavier à l'écran.

Le point commun entre tous les périphériques d'entrée est qu'ils convertissent l'information qu'ils récupèrent de l'extérieur en données compréhensibles par l'ordinateur. À l'inverse, les périphériques de sortie décodent l'information fournie par l'ordinateur afin de la rendre utilisable par l'utilisateur.

Instructions

Les instructions que l'ordinateur peut comprendre ne sont pas celles du langage humain. Le matériel sait juste exécuter un nombre limité d'instructions bien définies. Des instructions typiques comprises par un ordinateur sont « copier le contenu de la cellule 123 et le placer dans la cellule 456 », « ajouter le contenu de la cellule 321 à celui de la cellule 654 et placer le résultat dans la cellule 777 » et « si le contenu de la cellule 999 vaut 0, exécuter l'instruction à la cellule 345 ». Mais la plupart des instructions se composent de deux zones : l'une indiquant quoi faire, qu'on nomme le code opération, et l'autre indiquant où le faire, qu'on nomme opérande.

Au sein de l'ordinateur, les instructions correspondent à des codes - le code pour une copie étant par exemple 001. L'ensemble d'instructions qu'un ordinateur supporte se nomme son langage machine ou langage binaire car les instructions qui sont uniquement comprises par l'ordinateur sont constituées uniquement que de 0 (zéro) et de 1.

En général, les programmeurs n'utilisent plus ce type de langage mais passent par ce que l'on appelle un langage de haut niveau qui est ensuite transformé en langage binaire par un programme dédié (interpréteur ou compilateur selon les besoins). Les programmes ainsi obtenus sont des programmes compilés compréhensibles par l'ordinateur dans son langage natif.

Architecture

La miniaturisation permet d'intégrer l'UAL et l'unité de contrôle au sein d'un même circuit intégré connu sous le nom de **microprocesseur**.

Typiquement, la **mémoire** est située sur des circuits intégrés proches du processeur, une partie de cette mémoire, la **mémoire cache**, pouvant être située sur le même circuit intégré que l'UAL.

L'ensemble doit être complété par une **horloge** qui règle le processeur. Bien sûr, on souhaite que ce soit le plus vite possible, mais on ne peut pas augmenter impunément cette vitesse pour deux raisons :

- plus l'horloge est rapide et plus il chauffe toutes choses égales par ailleurs. Une trop grande température peut le détériorer ;
- il existe une cadence où le processeur devient instable, ce qui signifie que tout va si vite qu'il n'a plus le temps de s'y retrouver.

Un compromis doit donc être trouvé entre :

- vitesse nominale, qui est le choix recommandé par le constructeur ;
- surcadencement, qui augmentera la vitesse de calcul au prix de chauffage plus grand (donc bruits de ventilateurs plus importants à prévoir) et d'une diminution de la durée de vie de la puce; plus un risque de « plantage » dû à l'instabilité ;
- sous-cadencement, où on bride la vitesse, diminue la température et le bruit, et assure une longue durée de vie au processeur.

La tendance est aujourd'hui à regrouper plusieurs UAL dans le même processeur, voire plusieurs processeurs dans la même puce. En effet, la miniaturisation progressive (voir [Loi de Moore](#)) le permet sans grand changement de coût.

B. Les éléments constitutifs usuels d'un PC moderne

Le clavier d'ordinateur [3]

Un clavier d'ordinateur est un périphérique d'entrée constitué de touches et qui permet de saisir du texte.

Claviers nationaux

Plusieurs dispositions des touches existent : clavier AZERTY, clavier QWERTY, clavier QWERTZ et clavier DVORAK. Pour chacune de ces dispositions, des variantes nationales existent. Par exemple, l'AZERTY français n'est pas le même que l'AZERTY belge, et le QWERTZ allemand n'est pas le même que le QWERTZ suisse.

Existent notamment des claviers AZERTY, QWERTY, français, belge, espagnol, états-unien, 102 touches, 105 touches.

Historique

Les claviers informatiques sont similaires en apparence, et parfois dans leur fonctionnement, aux claviers des machines à écrire.

Les claviers ont été créés de manière à être similaires aux claviers des machines à écrire, afin de ne pas dérouter les utilisateurs. Dans les années 1980, chaque ordinateur familial avait le clavier intégré dans l'unité centrale. Ceci signifie que chaque ordinateur avait potentiellement un clavier différent. Cependant, des particularités nationales ont fini par apparaître.

Le clavier PC a été conçu par IBM. Les claviers des machines fonctionnant avec Mac OS et Sun ont été conçus par leurs firmes respectives (Apple Computer et Sun).

Des ajouts successifs ont eu lieu :

- le pavé numérique ;
- les touches de fonctions.

La souris [5]

Une souris est un dispositif de pointage manuel pour ordinateur ; elle est composée d'un petit boîtier fait pour tenir sous la main, sur lequel se trouvent des boutons.

La souris a été inventée en 1963 par Douglas Engelbart du Stanford Research Center après des tests d'utilisation, basés sur le trackball.

Les premières souris étaient en fait de simples trackballs inversées, où l'utilisateur déplaçait l'appareil. La friction de la boule contre la table permettait le mouvement du pointeur sur l'écran. Plus tard, les souris ont utilisé des mécanismes optiques pour détecter les mouvements.

Connecteurs de souris

Les souris les plus récentes pour Mac et PC utilisent le port [USB](#) ; c'est le type de connexion qui deviendra probablement le standard pour toutes les souris à câble.

Le sans-fil

Les technologies actuelles permettent de s'affranchir d'une connexion physique entre la souris et l'ordinateur, en passant par une liaison infra-rouge ou radio. Un boîtier est relié au port classique destinée à la souris et transforme les signaux reçus par le capteur infra-rouge ou radio en signaux compréhensibles par le protocole standard de la souris. La technologie radio est plus sûre (passe par-dessus les obstacles). On utilise un système de canaux radio pour ne pas mélanger les signaux de différents appareils.

L'avenir semble à la technologie *Bluetooth*, standardisée pour tout type de périphérique, qui évite la profusion d'émetteurs/récepteurs.

Notons que cette technique est surtout utile sur les PC. Sur les Macintosh, la souris est connectée au clavier et ne nécessite pas de long fil vers l'unité centrale.

Les boutons (et leur utilisation)

Les souris standard pour PC ont aujourd'hui une molette en plus de leurs deux boutons ; la molette (un bouton spécial) qui peut aussi bien être tournée (molettes mécaniques) que pressée (Trackpoint, donnant un troisième degré de liberté à la souris) s'est répandue. Les souris avec plus de deux boutons (voire deux molettes) remplissent différentes fonctions assignées à chacun par les applications, le pilote ou le système d'exploitation.

Par exemple, un utilisateur du bureau Windows ou KDE utilisera le bouton de gauche dans le navigateur Web pour suivre les liens, alors que celui de droite fera apparaître un menu permettant à l'utilisateur de copier des images ou un lien pour imprimer, etc.

Apple continue de produire des ordinateurs avec des souris ne comptant qu'un seul bouton, car leurs études montreraient que les souris à un bouton sont plus efficaces à l'usage. (Pour simuler le « bouton droit », il faut maintenir la touche spéciale Contrôle - souvent Ctrl-appuyée pendant le clic).

Représentation graphique : le pointeur

Le pointeur de la souris est un graphisme sur l'écran. Il peut prendre nombre de formes, celles-là pouvant dépendre du contexte. Lorsque l'utilisateur déplace la souris, le curseur se déplace. L'utilisateur peut ainsi sélectionner un élément (caractère, mot, bouton, image...)

Le moniteur d'ordinateur [6]

Un moniteur est un périphérique de sortie usuel d'un ordinateur. Il permet de visualiser les informations générées par l'ordinateur, sous forme de texte et d'image.

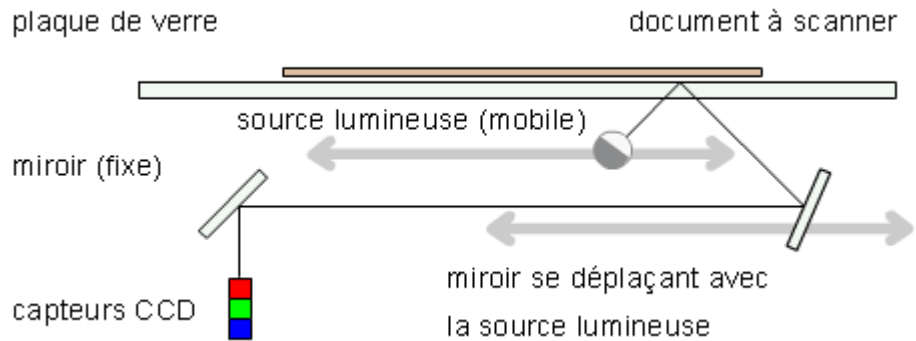
Différentes technologies existent :

- écrans à tube cathodique ; ce sont ceux qui ont un angle de vision le plus large et jusqu'à 2005 le meilleur rendu des couleurs, mais ils sont lourds, encombrants et gourmands en énergie.
- écrans LCD, légers et commodes, mais dotés d'un moins bon rendu des couleurs et, pour certains modèles d'entrée de gamme, d'une rémanence parfois gênante pour les jeux très rapides.
- écrans à plasma, de très bon rendu, mais chers et de durée de vie plus limitée.

Le scanner [4]

Le scanner, en français scanneur, ou encore numériseur est un périphérique d'informatique professionnelle ou domestique qui permet de transformer une image (dessin sur papier, imprimé, photographie) ou un objet réel en une image « électronique », c'est-à-dire stockée dans la mémoire d'un ordinateur ou sur un support informatique (*disque dur*, *CD-ROM*...). L'objet ou l'image est soumis au balayage d'un rayon lumineux, un capteur transforme la lumière réfléchie en un signal électrique qui est transféré à un ordinateur, pour y être ensuite sauvegardée, traitée ou analysée. L'appareil prend souvent la forme d'une tablette sur

laquelle le document doit être posé, mais il existe aussi des numériseurs à main et des stylos numériseurs.



Un capteur CCD convertit la lumière en signaux électriques qui peuvent être numérisés pour obtenir une image numérique.

L'imprimante [7]

Les imprimantes ont été conçues dès l'apparition des premiers ordinateurs, pour permettre la consultation et la conservation sur support papier des résultats produits par les programmes informatiques. En effet, à l'époque des premiers calculateurs, les écrans n'existaient pas encore et les méthodes de stockage de l'information étaient très rudimentaires et très coûteuses.

Avec le temps, les imprimantes ont énormément évolué dans leur méthode d'impression et de traction du papier, mais également dans leur qualité d'impression, leur encombrement et leur coût.

Méthodes d'impression

- Imprimante à aiguilles ou Imprimante matricielle

Sur les imprimantes à aiguilles, la tête d'impression est constituée d'une série d'aiguilles, alignées verticalement de façon à couvrir la hauteur d'une ligne de texte et propulsé par des électroaimants. Le nombre d'aiguilles peut varier d'une imprimante à l'autre (de 9 à 32 en général), la qualité d'impression est proportionnelle au nombre d'aiguilles. Cette tête se déplace le long de la ligne à imprimer.

L'encre est fournie par un ruban encreur, similaire aux rubans de machines à écrire (tissu imprégné d'encre), qui circule en boucle entre la tête d'impression et la feuille de papier. Chaque aiguille permet d'imprimer un minuscule point sur la feuille ; chaque caractère est donc constitué de multiples points.

- Imprimante laser

Sur ce système, l'encre se présente sous la forme d'une poudre extrêmement fine, le toner. L'imprimante laser utilise une technologie proche de celle utilisée dans les photocopieurs. Une imprimante laser est ainsi principalement constituée d'un tambour photosensible (en anglais « drum »)

qui, chargé électrostatiquement, est capable d'attirer l'encre afin de former un motif qui sera déposé sur la feuille de papier.

Le fonctionnement global est le suivant : un ioniseur de papier charge les feuilles positivement. Le laser charge le tambour positivement en certains points grâce à un miroir pivotant. Ainsi, l'encre sous forme de poudre (toner), chargée négativement, se dépose sur les parties du toner ayant été préalablement chargées par le laser.

En tournant, le tambour dépose l'encre sur le papier. Un fil chauffant (appelé coronaire) permet enfin de fixer l'encre sur le papier.

Ainsi, l'imprimante laser, n'ayant pas de tête mécanique, est rapide et peu bruyante.

Cette technique, bien que sophistiquée, permet une impression rapide (non plus ligne par ligne, mais page par page) très fine et très souple (impression de tous types de textes, de graphismes, de photos...) avec une qualité irréprochable pour le noir et blanc. Cependant, elle est peu adaptée aux niveaux de gris, et de ce fait, à l'impression en couleur. Les évolutions technologiques et des techniques du début du XXI^e siècle ont permis d'adapter la couleur à ce système d'impression.

- Imprimante à jet d'encre

Les têtes d'impressions jet d'encre utilisent de l'encre liquide contenue dans un réservoir. La tête proprement dite est percée de fins canaux remplis d'encre, et un système piezo-électrique ou de chauffage électrique produit des variations de pression qui expulsent des gouttelettes sur la feuille, formant des points.

Comme avec les têtes à aiguilles, les caractères sont formés par des concentrations de points, et l'impression se fait donc ligne par ligne. Néanmoins, la finesse de ces gouttelettes est contrôlable, et la technologie permet un mélange des couleurs, si bien que la plupart des imprimantes jet d'encre récentes permettent des impressions « qualité photo ».

Le disque compact [10]

Un disque compact ou CD de l'anglais "compact disc", est un disque optique utilisé pour stocker des données numériques (suite de 0 et de 1).

Principe de fonctionnement

Le disque compact repose sur une méthode de lecture optique : un faisceau de lumière cohérente (laser) vient frapper le disque en rotation. Les irrégularités (cavités) dans la surface réfléchissante de celui-ci produisent des variations binaires (suite de 0 et de 1). Le rayon réfléchi est enregistré par un capteur, et l'information binaire est ensuite transformée en un signal analogique par un convertisseur.

Histoire

Le disque compact fut inventé conjointement par les firmes Philips et Sony, pour l'audio numérique (CD audio) en 1980.

Philips développa le processus de fabrication basé sur leur expérience de la technologie du Laserdisc tandis que Sony contribua à la méthode de correction d'erreurs. Les premiers prototypes produits par Philips mesuraient 115 mm de diamètre, avec un codage sur 14 bits et une capacité de 60 minutes. Sony insista pour qu'on adopte un codage sur 16 bits et une durée de 74 minutes, ce qui a augmenté la taille du disque à 120 mm. Selon les rumeurs, la capacité du CD 12 centimètres a été augmentée à 74 minutes pour que la version la plus lente de la 9e symphonie de Beethoven tienne sur un seul CD.

Types de disques

On distingue plusieurs types de disques compacts :

- CD audio : disque compact audio
- CD-ROM (Compact Disc Read-Only Memory), officiellement cédérom en français : support de stockage informatique
- CD-R : Compact Disc Recordable
- CD-RW : Compact Disc Rewritable

Les appareils de lecture pour CD-audio ne sont pas conçus pour lire les CD-ROM ; a contrario, les lecteurs de CD-ROM (couramment présents sur les ordinateurs personnels) peuvent aussi lire les CD-audio. Il existe aussi des CD « hybrides » contenant de l'information audio (lisible par un lecteur audio) et des informations d'autres types (texte, vidéo, images, etc.), lisibles par un lecteur de CD-ROM.

Capacité de stockage

Les spécifications du disque compact recommandent une vitesse linéaire de 1,22 m/s et un pas entre les pistes de 1,59 μm . Cela conduit à un CD audio de 74 minutes sur un disque de 120 mm ou environ 650 Mo de données sur un CD-ROM. Néanmoins, afin d'autoriser des variations dans la fabrication des supports, il y a une tolérance dans la densité des pistes. En fabriquant délibérément des disques de plus haute densité, on peut augmenter la capacité et rester très proche des spécifications du CD. En utilisant une vitesse linéaire de 1,1975 m/s et un pas entre les pistes de 1,497 μm , on atteint une nouvelle capacité maximale de 79 minutes et 40 secondes ou 702 Mo. Bien que ces disques possèdent une légère variation de fabrication, ils sont très souvent lus par les lecteurs et seul un très faible nombre de lecteurs les rejettent.

Le DVD-Rom [9]

Le DVD-Rom ou DVD-ROM (Digital Versatile Disc Read Only Memory) est un format de DVD successeur du [Compact Disc](#). C'est un support de masse dont les dimensions et l'apparence le font ressembler au CD-Rom comme deux gouttes d'eau, mais il bénéficie d'une densité d'écriture nettement supérieure qui lui permet de disposer d'une capacité de stockage de données au moins sept fois supérieure.

Le support DVD a été mis au point par plusieurs grandes entreprises influentes dans le domaine du multimédia. Actuellement, il permet de stocker tout type d'informations : données, vidéos et musiques.

Tableau de capacité

Face	
Simple	Double

Tableau de capacité

		Face	
Densité	Simple	4,7 Go	9,4 Go
	Double	8,5 Go	17 Go

Formats

Il existe différents formats de DVD (avec, pour chacun, des supports différents) :

- DVD-ROM : ceux sur lesquels on retrouve par exemple les films. Ils sont généralement « pressés » c'est-à-dire qu'il existe une matrice de base qui sert de moule pour les copies... et ne sont donc pas enregistrables.
- DVD-R : aussi noté -R (pour Recordable : enregistrable) cette norme est la première à avoir vu le jour et était principalement destinée à la vidéo. Les informations sauvées sur le support le sont par altération d'une couche inscriptible à l'aide du laser du graveur.
- DVD+R : comme pour le -R mais la norme est plus récente et est plus adaptée que le -R pour le stockage de données. Il possède aussi de meilleures caractéristiques techniques que son cousin. Il n'existe cependant presque aucune différence visible à l'œil nu entre les -R et le +R.
- DVD-RW et DVD+RW : sont les pendants des CD-RW c'est-à-dire les réinscriptibles (ReWritable) avec les mêmes caractéristiques que leurs homologues -R et +R. Les informations sauvées sur le support le sont par réorganisation de la couche enregistrable à l'aide du laser du graveur. C'est pourquoi un formatage est requis avant d'écrire ou pour effacer le disque.
- DVD-Ram : est un format de DVD réinscriptibles, au même titre que les DVD+/-RW, dont le principal atout est qu'il permet d'enchaîner aléatoirement lectures et écritures. Parmi les trois technologies actuelles concurrentes de DVDs réinscriptibles (DVD-RAM, DVD+RW et DVD-RW), le DVD-RAM est considéré comme un format fortement fiable, car les disques ont un contrôle d'erreur intégré et un système de gestion de défaut. Par conséquent, un DVD-RAM est perçu pour être meilleur que les autres technologies de DVD pour une utilisation dans des tâches informatiques traditionnelles comme le stockage de données en général, la copie de secours et l'archivage.

Ces différents formats créaient une certaine confusion. En 2005, de nouveaux types de graveurs permettent d'enregistrer sous plusieurs formats.

Futur

Pour succéder au DVD, deux formats sont en compétition : le Blu-Ray Disc (sa capacité de base est de 25 Go) et le HD-DVD (15 Go en simple couche, 30 Go en double couche). Il est probable que ces deux formats cohabiteront pendant quelques années sur le marché.

Le disque dur [11]

Le disque dur est un périphérique de stockage magnétique. Il a remplacé efficacement les tambours (aujourd'hui obsolètes) et les bandes, seulement utilisées de nos jours pour l'archivage et la sauvegarde. Inventés dans les années 1950 par IBM, leur capacité augmente très rapidement tandis que leur encombrement se réduit.

Historique

Les ingénieurs d'IBM n'étaient pas satisfaits des systèmes de stockage sur tambours magnétiques : l'efficacité volumétrique était très faible, les tambours occupaient beaucoup d'espace pour peu de capacité. En 1953, un ingénieur récemment embauché eut l'idée de superposer des plateaux le long d'un axe et d'y adjoindre une tête de lecture/écriture mobile, située sur un axe parallèle à celui des plateaux. Cette tête venait s'insérer entre les

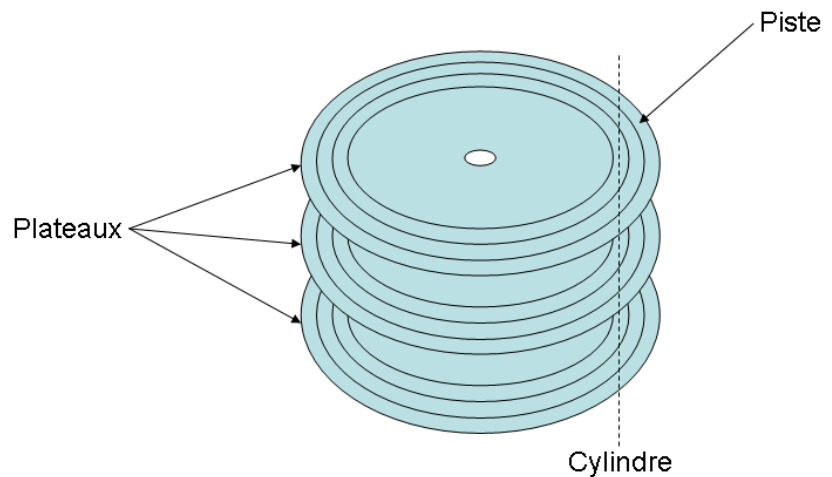
plateaux pour lire les informations, mais devait se retirer complètement pour passer d'un plateau à un autre. Un prototype fut construit avec une vitesse de rotation d'environ 1000 tours/minute. À cette vitesse il était compliqué de maintenir les têtes au-dessus de la surface des plateaux. L'idée fut alors d'injecter de l'air sous-pression au travers de la tête de lecture, ce qui la maintenait au dessus du plateau. La distance tête-plateau était de 20 μ m.

En 1955 le premier système de ce type à été dévoilé au public par IBM, il fut baptisé RAMAC (Random Access Method of Accounting and Control), modèle 305, et la production commerciale commença en juin 1957. Jusqu'à 1961 plus d'un millier d'unités furent vendues. Son prix : 10 000 dollars (de l'époque) par *megaoctet*.

En juin 1954 J. J. Hagopian, ingénieur IBM, a l'idée de faire « voler » les têtes de lecture/écriture au dessus de la surface des plateaux, sur un coussin d'air. Il propose le design de la forme de ces têtes. En septembre 1954 il dessine l'équivalent des disques durs actuels : des plateaux superposés et un axe sur lequel sont fixés les têtes de lecture/écriture. Cela deviendra un produit commercial en 1961 sous la dénomination « IBM 1301 Disk Storage ».

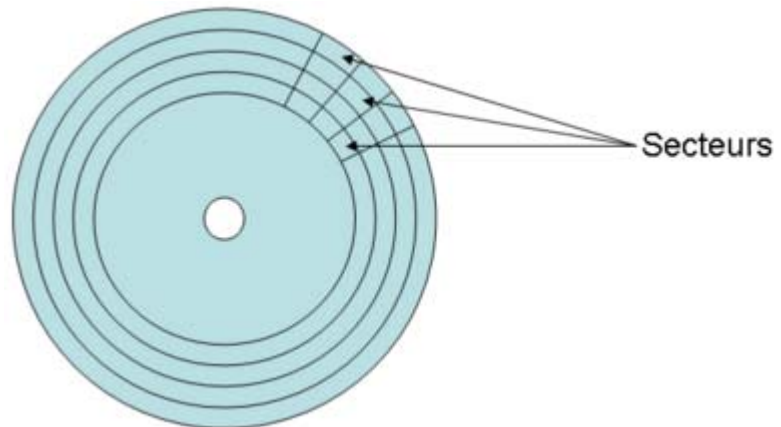
Géométrie

Chaque plateau (2 surfaces) est composé de pistes concentriques. Les pistes situées à un même diamètre forment un cylindre.



Géométrie d'un disque dur

Sur une piste les données sont délimitées en secteurs, aussi appelés blocs.



Géométrie d'un disque dur. Les pistes sont concentriques, les secteurs contigus.

Il faut donc trois coordonnées pour accéder à un bloc :

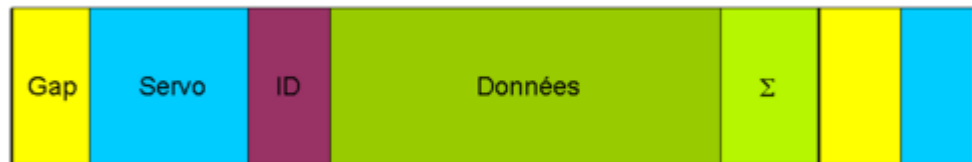
1. le numéro de la tête (choix de la surface)
2. le numéro de la piste (détermine le déplacement de la tête)
3. le numéro du bloc sur cette piste (détermine à partir de quand il faut commencer à lire les données).

Cette conversion est faite par le contrôleur du disque à partir de l'adresse absolue du bloc (un nombre compris entre 0 et le nombre total de blocs (moins 1) contenu sur le disque).

On notera que les secteurs extérieurs et intérieurs n'ont pas la même taille.

Sur les premiers disques, une surface était formatée en usine et contenait les informations permettant au système de se synchroniser (de savoir quel était la position des têtes à tout moment). Cette surface était dénommée « servo ». Par la suite, ces zones de synchronisation ont été mixées entre les blocs de données, mais elles sont toujours formatées en usine. Typiquement donc, on trouvera sur une piste une succession de :

1. un petit « blanc » ou « espace » (« gap » en anglais),
2. une zone servo,
3. un entête contenant le numéro du bloc qui va suivre,
4. les données,
5. une somme de contrôle permettant de corriger des erreurs.



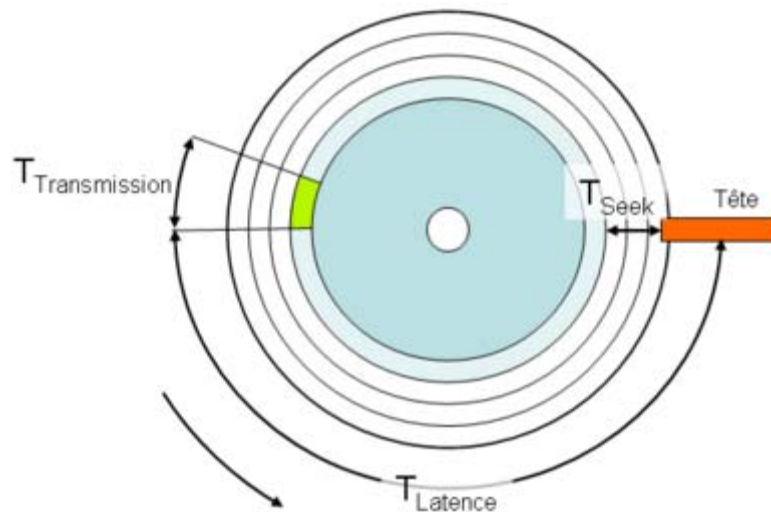
Format d'un secteur. Il ne contient pas seulement les données stockées, mais aussi un préambule permettant de synchroniser le système d'asservissement du disque, une entête avec l'identifiant du bloc et enfin une somme de contrôle Σ permettant de détecter d'éventuelles erreurs.

Performances

Le temps d'accès et le débit d'un disque dur permettent d'en mesurer les performances. Les facteurs principaux à prendre en compte sont :

1. le temps de latence, facteur de la vitesse de rotation des plateaux. Le temps de latence moyen (en seconde) est égal à 60 divisé par la vitesse de rotation en tours par minute. Le temps de latence moyen est égal au temps de latence divisé par 2 (car on estime que statistiquement les données sont à un demi-tour près des têtes).
2. le temps de recherche, ou seek time en anglais, est le temps que met la tête pour se déplacer jusqu'au cylindre choisi. C'est une moyenne entre le temps piste à piste, et le plus long possible.

3. le temps de transfert est le temps que vont mettre les données à être transférées entre le disque dur et l'ordinateur par le biais de son interface.



Pour lire le secteur (en vert) situé sur une piste interne à l'opposée de la tête de lecture (en rouge), il faut déplacer la tête vers l'intérieur, attendre que le bloc arrive sous la tête puis lire la totalité du bloc.

Pour estimer le temps de transfert total, on additionne ces trois temps. On pourra rajouter le temps de réponse du contrôleur, etc. Il faut souvent faire attention aux spécifications des constructeurs, ceux-ci auront tendance à communiquer les valeurs de pointe au lieu des valeur soutenues (par exemple pour les débits).

Voici deux disques comparés. Le premier, le DEC RP07 équipait les ordinateurs DEC des années 70-80, tandis que le Maxtor est un disque de 3,5 pouces récent (2004). Ils peuvent tous les deux être considérés comme des disques haute de gamme au moment de leur mise sur le marché.

	DEC RP07	Maxtor Atlas 15k
Hauteur (cm)	118	2,6
Largeur (cm)	67,3	10,1
Profondeur (cm)	83,8	14,7
Poids (Kg)	181	0,81
Capacité (Mo)	516	74 752 (73 Go)
Vitesse de rotation (t/m)	3 633	15 000
Temps de latence moyen (ms)	8,3	2
Seek time piste à piste (ms)	5	0,3/0,5
Seek time maximum (ms)	-	9
Seek time moyen	23	3,4/3,8
Taux de transfert maximum (Mo/s)	2,1	100
Taux de transfert soutenu (Mo/s)	-	75

	DEC RP07	Maxtor Atlas 15k
Nombre de surfaces	16 + 1 servo	8
Nombre de plateaux	9	4
Secteur/piste	-	50
Octets/secteur	512	512
Interface	MASSBUS	SCSI Ultra 320

Capacité de stockage

Les capacités actuelles (2005) s'échelonnent entre 20 **Go** et 400 **Go**. La capacité des disques durs a augmenté beaucoup plus vite que leur rapidité, limitée par la mécanique.

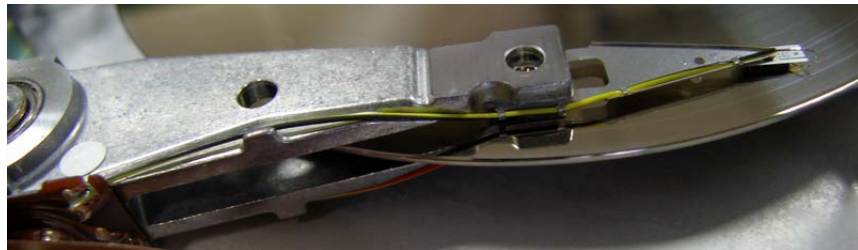
Plateaux

Les plateaux sont solidaires d'un axe sur roulements à billes. Cet axe est maintenu en mouvement par un moteur électrique. La vitesse de rotation est actuellement (2005) comprise entre 3600 et 15000 tours/minute (l'échelle typique des vitesses est 3600, 4200, 5400, 7200, 10000 et 15000 tours/minute). La vitesse de rotation est conservée constante.

Les plateaux sont composés d'un substrat, autrefois en aluminium, de plus en plus souvent en verre, traitées par diverses couches dont une ferromagnétique recouverte d'une couche de protection. L'état de surface doit être le meilleur possible.

Tête de lecture/écriture

Fixées au bout d'un bras, elles sont solidaires d'un second axe qui permet de les faire pivoter en arc de cercle sur la surface des plateaux. Toutes les têtes pivotent donc en même temps. Il y a une tête par surface. Leur géométrie leur permet de voler au dessus de la surface du plateaux sans le toucher : elles reposent sur un coussin d'air créé par la rotation des plateaux. En 1997 les têtes volaient à 25 nanomètres de la surface des plateaux, aujourd'hui (2005) cette valeur est d'environ 10 nanomètres.



Le bras supportant les deux têtes de lecture/écriture. Les rayures visibles sur la surface du plateau indique que le disque dur était en panne, victime d'un «atterrissage».

Le moteur qui les entraîne doit être capable de fournir des accélérations et décélérations très importantes. Un des algorithmes de contrôle des mouvements du bras porte-tête est d'accélérer au maximum puis de freiner au maximum pour que la tête se positionne sur le bon cylindre. Il faudra ensuite attendre un court instant pour que les vibrations engendrées par le freinage s'estompent.

À l'arrêt les têtes doivent être parquées, soit sur une zone spéciale, soit en dehors des plateaux.

Si une ou plusieurs tête rentre en contact avec la surface des plateaux, cela s'appelle un atterrissage et provoque la destruction des informations situées à cet endroit. La mécanique des disques durs est donc assemblée en salle blanche et toutes les précautions (joints

etc.) sont prises pour qu'aucune impureté ne puisse pénétrer à l'intérieur du boîtier (appelé « HDA » pour « Head Disk Assembly » en anglais).

Les dimensions des disques durs sont normalisées :

- 5 pouces 1/4 format apparu dans les années 1980, exista aussi en demie-hauteur.
- 3 pouces 1/2 est la taille standard à ce jour (2005).
- 2 pouces 1/2 pour les ordinateurs portables.
- 1 pouce 8 pour les baladeurs mp3, certains disques durs externes.

Les plus petits disques rentrent dans la catégorie des microdrives, avec une taille de 1 pouce

La carte mère [14]

La carte mère (en anglais, motherboard) est le composant de base d'un ordinateur, elle sert à interconnecter tous les composants de l'ordinateur et reçoit tous les éléments essentiels à la bonne marche de l'ordinateur. Elle en est le cœur, le cerveau et le système nerveux.

Elle est composée de nombreux circuits permettant une bonne gestion de tous les flux de données qui transitent par elle. C'est elle, en particulier, qui porte le *chipset*, jeu de composants électroniques essentiel, qui fait le lien entre le *processeur*, la *mémoire* et les périphériques.

Les composants supportés

Ce circuit imprimé de haut niveau supporte :

- le *microprocesseur*,
- le *chipset*,
- la *mémoire RAM*,
- les connecteurs d'extension (bus *PCI*, bus *AGP*) pour carte d'extension
- le *BIOS*
- les connecteurs IDE (Le protocole IDE, acronyme de Integrated Drive Electronics, est un des standards les plus répandus en ce qui concerne les disques durs.)
- *l'horloge*
- un connecteur d'alimentation
- les différents *bus*
- les connecteurs standards pour le clavier, le moniteur, la souris
- divers autres composants.

Le microprocesseur [15]

Le microprocesseur est le cœur des micro-ordinateurs. À strictement parler, il s'agit d'un processeur ou "Central processing unit" qui a été réduit en taille suffisamment pour tenir sur un seul circuit intégré (puce).

Comme tous les processeurs, il prend en charge l'exécution des instructions contenues dans les programmes informatiques.

Néanmoins, la distinction entre "Central Processing Unit", CPU, processeur et microprocesseur est souvent abandonnée au profit d'une banalisation de ces termes.

La distinction se fait désormais dans sa fonction entre celle centrale et celle prenant en charge des fonctions comme le graphisme ou la compression/décompression audio-vidéo.

Histoire

Le microprocesseur a été inventé par Marcian Ted Hoff en 1971, alors que celui-ci était ingénieur chez Intel. À l'époque, ni la direction d'Intel ni le client japonais à qui était destiné le microprocesseur, n'ont été intéressés par l'invention.

Le premier microprocesseur commercialisé est le 4004 4-bits le 15 novembre 1971. Il fut suivi par le 8008. Ces processeurs sont les précurseurs des Intel 8080 et Zilog Z80 et de la future famille des Intel x86.

Familles de microprocesseurs

Il existe plusieurs familles de microprocesseurs :

- La plus connue par le grand public est celle de la famille x86, développée principalement par Intel (Pentium), AMD (Athlon), CYRIX, NEC... Les deux premiers constructeurs fournissent la plus grande partie des processeurs actuellement utilisés dans la construction des ordinateurs de type PC (2005).
- Les PowerPC d'IBM et Motorola équipaient jusqu'en 2006 les Macintosh (Apple Computer) et sont utilisés dans divers systèmes embarqués. Une puce différente équipée de 10 processeurs périphérique intégrés a été choisie pour les consoles de jeu : Playstation 3, la Xbox 360 et probablement la future Nintendo Wii...
- La famille 68000 de Motorola animait les anciens Macintosh, les Atari ST et les Commodore Amiga. Leurs dérivés (Dragonball, ColdFire) sont toujours utilisés dans des systèmes embarqués.

Parmi les familles moins connues du grand public :

- La famille Sparc anime la plus grande partie des stations de travail de Sun Microsystems.
- La famille MIPS anime les stations de travail de Silicon Graphics, des consoles de jeux, comme les PSOne et des systèmes embarqués, ou des [routeurs](#) Cisco.
- La famille StrongARM est de nos jours utilisée uniquement dans les systèmes embarqués, elle a précédemment été utilisée par Acorn pour ses Archimedes et RiscPC.

Fonctionnement

Les microprocesseurs sont cadencés par une [horloge](#) qui fonctionne à plusieurs mégahertz (MHz). Au milieu des années 1980, les microprocesseurs fonctionnaient à 4,77 et 8 MHz. Courant 2004, cette vitesse d'horloge atteint 3,4 GHz sur des modèles commerciaux (5 GHz en laboratoire). Plus la vitesse de l'horloge est élevée, plus le microprocesseur sera capable d'exécuter à un rythme élevé les instructions de base des programmes.

Les microprocesseurs actuels sont optimisés pour exécuter plus d'une instruction par cycle d'horloge, ce sont des microprocesseurs avec des unités d'exécution parallélisées. En plus ils sont dotés de procédures qui « anticipent » les instructions suivantes avec l'aide de la statistique.

Dans la course à la puissance des microprocesseurs, il y a deux méthodes d'optimisation. Faire des microprocesseurs au jeu d'instructions simplifié (RISC, Reduced Instruction Set Computer), plus rapide pour les instructions simples, faciles à fabriquer et dont on peut monter la vitesse de l'horloge sans trop de difficultés techniques. L'autre famille de microprocesseurs s'appelle le CISC (Complex Instruction Set Computer) dont le nombre de cycle d'horloge par instruction est plus long, mais qui a en son cœur plus d'instructions pré-câblées que le RISC.

Néanmoins, avec la considérable augmentation de la taille des puces électroniques et la gigantesque accélération des fréquences d'horloge, la distinction entre RISC et CISC a quasi complètement disparu. Là où des familles tranchées existaient, on observe aujourd'hui des microprocesseurs où une structure interne RISC apporte de la puissance tout en restant compatible avec une utilisation de type CISC (la famille Intel x86 a ainsi subi discrètement une transition entre une organisation initialement très typique d'une structure CISC et l'utilisation actuelle d'un cœur RISC très puissant plus un réarrangement du code à la volée).

Les processeurs actuels contiennent des *mémoires cache* qui vont jusqu'à trois niveaux et sont de plus en plus étendues.

Fabrication des microprocesseurs

La fabrication d'un microprocesseur est essentiellement identique à celle de n'importe quel circuit intégré. Elle suit donc un procédé complexe. Mais l'énorme taille de la plupart des microprocesseurs a tendance à augmenter encore le coût de l'opération.

La *loi de Moore*, qui indique que le degré d'intégration des microprocesseurs double tous les 18 mois, indique également que les coûts de production doublent en même temps que le degré d'intégration.

La fabrication des microprocesseurs est aujourd'hui considérée comme l'un des deux facteurs d'augmentation de la capacité des unités de fabrication (avec les contraintes liées à la fabrication des mémoires à grande capacité). La finesse de la gravure industrielle va atteindre bientôt 65 nm (en ce moment 90 nm).

Fonctions à développer

- Organisation parallèle
Dépendant aussi du *système d'exploitation* la tendance actuelle est l'installation de plusieurs processeurs parallèles et de multiples tâches d'où l'importance grandissante des fonctions d'arbitrages entre processus (par exemple l'hyper threading).
- Sécurité et location
Sous l'impulsion des distributeurs de musique et de vidéo qui veulent garder cette formidable source de revenus à forte marge l'industrie met en place toutes les fonctions (par exemple LaGrande et Palladium) qui leur permettra d'obliger les utilisateurs à passer plusieurs fois à la caisse.

Le problème de l'échauffement

L'échauffement des microprocesseurs reste grosso modo et malgré l'usage de techniques de gravures de plus en plus fines, proportionnel au carré de leur fréquence à architecture donnée :

- Un i686 à 500 MHz (1,1V) consommait typiquement 9W.
- Un i686 à 1 GHz (1,7V), deux fois plus rapide, consomme typiquement 34 W, ce qui n'est pas loin du quadruple.
- À 2 GHz un Opteron dissipe 107 W et un G5 55 W.

Ce problème est lié à un autre, celui de la dissipation thermique et donc souvent des ventilateurs, sources de nuisances sonores difficilement compatibles avec un environnement de bureau. Le refroidissement liquide (à eau) est proposé.

La mémoire RAM [16]

La mémoire RAM (Random Access Memory en anglais) est aussi appelée mémoire vive ou mémoire volatile car toutes les données sont perdues à l'extinction de l'alimentation.

Technologie

La mémoire informatique est un composant qui fut d'abord magnétique (tores de ferrite), puis devint électronique dans les années 1970, et qui permet de stocker et relire des informations binaires. Son rôle est notamment de stocker les données qui vont être traitées par l'unité centrale (ou le microprocesseur); elle n'a rien de commun en temps d'accès (quelques dizaines ou centaines de nanosecondes) avec le disque dur (quelques millisecondes, soit dix mille à cent mille fois plus).

La RAM a la particularité de pouvoir être accédée en lecture et en écriture.

Divers types de mémoire vive

Il existe deux familles de mémoires : les RWM et les ROM.

- les RWM, pour Read Write Memory, sont par abus de langage appelées RAM (à tel point que ce nom abusif est passé dans l'usage courant).
- les ROM, pour Read Only Memory, ont gardé leur nom.

La différence entre ces deux types réside dans le rapport entre les temps d'écriture et de lecture sur la mémoire.

- Une RWM a un temps d'écriture proche de son temps de lecture, ce qui simplifie le cadencement.
- D'autres mémoires, utilisées par exemple pour stocker les **BIOS** ont une durée d'écriture bien plus longue que la durée de lecture. C'est abusivement que certaines sources les désignent sous le nom de ROM. Il existe pour elles un sigle approprié, et utilisé par des constructeurs comme IBM ou Intel, qui est NVRAM : non-volatile RAM.
- Les RWM les moins chères sont dynamiques elles ne conservent l'information que lorsque l'information y est en permanence rafraîchie, ce qui implique qu'elles restent alimentées électriquement (sous tension).
- Les mémoires flash (clés USB, et autres) sont statiques.

Une RAM dynamique (DRAM) ne conserve ses informations que si elle est « rafraîchie » régulièrement, c'est-à-dire si un signal lui est transmis de manière régulière (toutes les x millisecondes) afin de remettre au bon niveau les charges électriques représentant l'information, et qui sinon s'affaibliraient progressivement jusqu'à disparaître. Pourquoi compte-tenu de ces contraintes de rafraîchissement et de consommation utiliser quand même de la DRAM ? Parce qu'elle est à la fois bon marché et rapide.

On distingue sur les machines actuelles (2004) les types de mémoire RAM :

- SDRAM (Synchronous Dynamic RAM) pour les machines de la génération Pentium II, Pentium III. On distingue la SDRAM 66, 100 et 133 (fréquence d'accès en MHz).
- RDRAM (Rambus Dynamic RAM). Pour les machines de génération Pentium III et 4. Développées par la société RAMBUS, elles souffrent notamment d'un prix beaucoup plus élevé que les autres types de mémoires et de brevets trop restrictifs de la part de cette société.
- DDR2-SDRAM (Double Data Rate 2-SDRAM). Pour les machines de génération Pentium 4 et plus. On distingue les DDR2 533 et DDR2 667. Le numéro représente la vitesse maximum d'accès. Certains constructeurs privilégient encore la technique d'appellation basée sur la quantité de données théoriquement transportables (PC4300, PC4500, etc), mais la plupart semblent retourner à la vitesse réelle de fonctionnement afin de distinguer plus clairement la DDR2 de la génération précédente.

Il existe aussi des mémoires Flash. Ce sont des mémoires NVRAM effaçables électriquement (EEPROM), qui par conséquent gardent la mémoire sans être alimentée. On les utilise dans les appareils mobiles (appareils photo, téléphones portables etc.). Les utilisateurs de *PDA* auront déjà remarqué que leur temps d'accès, même en lecture seule, est pour le moment bien plus lent que celui de la mémoire dynamique.

La mémoire ROM [44]

La mémoire ROM (Read Only Memory, Mémoire à lecture seule) est une mémoire informatique impossible à modifier.

Elle est aussi appelée mémoire morte car son contenu n'est pas volatile en l'absence de courant électrique. Contrairement aux mémoires vives, les données stockées ne peuvent pas être modifiées en temps réel.

Les mémoires mortes sont utilisées, entre autres, pour stocker les informations vitales d'un ordinateur (ex: *BIOS*, instructions de démarrage, microcode). Puisqu'elles ne peuvent pas être modifiées, il n'y a pas de risque d'effacement accidentel par l'utilisateur.

Les temps d'accès à ce type de mémoire étant relativement lents (pour information les RAM ont un temps d'accès moyens de 45 nanosecondes), les données stockées sont généralement copiées au démarrage, dans une mémoire RAM plus rapide. On appelle cette opération le shadowing.

La carte graphique [13]

Une carte graphique est un composant électronique destiné aux ordinateurs personnels.

Auparavant, ce type de carte s'insérait dans un emplacement *PCI*, mais avec l'évolution de l'électronique, il utilise actuellement un port *AGP* qui permet un accès beaucoup plus rapide, qui sera prochainement remplacé par le port *PCI Express* encore plus rapide.

Sa fonction principale est de traiter les signaux vidéo et de les envoyer à l'écran. Lorsqu'un développeur souhaitera utiliser les fonctionnalités des cartes graphiques, il utilisera des bibliothèques graphiques telles que *DirectX* qui permettent de rendre la programmation d'un logiciel plus facile et permet de rendre le code source plus indépendant de la carte graphique.

Toutes les cartes graphiques ont deux moyens d'accès à leur mémoire (ou buffer), pendant que l'un est utilisé pour recevoir des informations en provenance du reste du système, l'autre est sollicité pour l'affichage à l'écran. Le premier est un accès aléatoire conventionnel (*RAM*) comme pour les mémoires centrales, le deuxième est généralement un accès séquentiel à la zone représentant l'écran (pixel buffer). Les jeux demandant de plus en plus de puissance, il fallait trouver un moyen pour que tous les calculs spécifiques à la 3D ne se fassent plus sur le processeur de la carte mère. Depuis une dizaine d'années, les cartes graphiques prennent en charge ces calculs.

Une des évolutions majeures récentes des cartes graphiques est le fait qu'elles soient devenues programmables. Elles possèdent maintenant leur propre processeur spécialisé dans les calculs d'affichage (multiplication de matrice...). Ces processeurs allègent la charge de calcul du ou des processeurs de la carte mère. Celui-ci peut donc se spécialiser dans d'autres tâches.

La clef USB [8]

On nomme clef USB (ou clé USB) un petit périphérique de stockage de données qui utilise une mémoire flash et un *connecteur USB*.

Les clefs USB sont alimentées en énergie par la *connexion USB* d'un ordinateur, sur lequel la clef est branchée. Une clef USB ne contient donc pas de batterie. Elles sont insensibles à la poussière et aux rayures, contrairement à leurs prédécesseurs : la disquette et le cédérom, ce qui leur donne un indéniable avantage au niveau de la fiabilité.

Leur capacité peut varier de quelques *mégaoctets* à quelques *gigaoctets*. Fin 2004, on trouvait des clés avec un minimum de 64 Mo, toutefois la taille standard était plutôt aux alentours de 256Mo. Ces périphériques sont toujours alimentés par le bus USB. Par rapport aux clefs USB traditionnelles les débits sont généralement meilleurs mais les temps d'accès sont plus importants. Ils sont aussi un peu plus fragiles et peuvent avoir tendance à chauffer en cas d'utilisation intensive, de plus leur taille est légèrement plus importante ; toutefois ils tiennent toujours facilement dans la poche.

La mémoire flash est une mémoire à semiconducteurs, non volatile et réinscriptible, c'est-à-dire une mémoire possédant les caractéristiques d'une *mémoire vive* mais dont les données ne disparaissent pas lors d'une mise hors tension. Ainsi, la mémoire flash stocke les bits de données dans des cellules de mémoire, mais les données sont conservées en mémoire lorsque l'alimentation électrique est coupée.

La carte réseau [22]

Utilité

La carte réseau assure le rattachement d'un équipement informatique à un ensemble d'autres ressources connectées sur le même réseau. Les équipements communiquent sur le réseau au moyen de signaux qui doivent absolument respecter des normes.

Média de transmission des informations

Le média ou support de l'information est généralement un réseau filaire. La carte réseau est, dans ce cas, munie d'un connecteur sur lequel on branche un câble réseau. Ce dernier est relié au réseau par l'intermédiaire d'une prise murale ou directement sur un équipement d'interconnexion de réseau comme un *concentrateur (Hub)* ou un *commutateur (Switch)*.

Récemment est apparue sur le marché domestique la transmission sans fil (*Wi-Fi*). On s'affranchit alors complètement du réseau filaire qui est alors remplacé par un réseau d'ondes électro-magnétiques.

Parmi les autres média de transmission, nous pouvons citer la fibre optique, largement utilisée dans les interconnexions à grand débit.

Les standards

Le standard Ethernet est le standard le plus répandu. Ethernet est un protocole de réseau informatique à commutation de paquets. On le trouve aussi bien en entreprise que chez le particulier. Pour le réseau sans fil, le standard Wi-fi est le plus courant.

Les débits

Les débits s'expriment généralement en Mbps (Mégabits par seconde). C'est la capacité d'un équipement réseau à émettre et recevoir plus ou moins de bits d'informations par unité de temps.

Les débits actuels du standard Ethernet sont :

- 10 Mbps
- 100 Mbps (FAST ETHERNET)
- 1000 Mbps (GIGABIT ETHERNET)

Les cartes réseau peuvent communiquer en 'half duplex'. Dans ce cas, une carte ne peut qu'émettre ou recevoir des informations à un instant donné. Le mode 'full duplex' permet à une carte réseau d'émettre et recevoir simultanément. Deux équipements réseau doivent communiquer dans le même débit. Un paramétrage logiciel de la carte réseau permet soit de forcer le débit ou de le positionner en 'auto-négociation'. Dans ce cas, les cartes réseau négocient un débit commun lors de leur premier échange d'informations.

Types de cartes réseau

On peut relier les ordinateurs de bureau au réseau selon les types de cartes ci-dessous :

- Réseau filaire :
Carte PCI à insérer dans un connecteur PCI libre sur la carte mère

Certains modèles de *cartes mère* disposent d'une interface réseau. Dans ce cas, on branche directement le câble réseau sur le connecteur RJ45 fixé à la carte mère.
- Réseau sans fil :
Carte PCI équipée d'une antenne.

Pour les ordinateurs portables ne disposant pas de connecteurs PCI, d'autres solutions existent.

- Carte réseau au standard *PCMCIA*
- Interface réseau déjà intégrée au portable

routage de router la trame correctement et à la machine destinataire de connaître l'origine des informations qu'elle reçoit, donc d'y répondre si besoin est.

ADSL [72]

ADSL signifie Asymmetric Digital Subscriber Line. La définition française est « Ligne d'abonné numérique à débit asymétrique ». La traduction officielle est : « raccordement numérique asymétrique » (RNA) ou « liaison numérique à débit asymétrique ».

ADSL est une technologie de communication haut débit permettant d'utiliser les lignes téléphoniques déjà existantes afin d'accéder à [Internet](#) et d'autres services.

Utilisant des fréquences supérieures à celles d'un signal voix, la connexion est toujours établie et ne demande pas une phase de synchronisation pour être établie. Afin d'optimiser le débit disponible pour une utilisation courante, le débit est asymétrique : le débit descendant (téléchargement) est plus important que le débit montant (upload). Les données et le signal voix circulent simultanément sur la même ligne sans se gêner (utilisation de fréquences différentes).

API [45]

Une Interface de programmation (en anglais Application Programming Interface ou API) définit la manière dont un composant informatique peut communiquer avec un autre.

BIOS ou Basic Input Output System [29]

Le Basic Input Output System ou BIOS (système de base d'entrée/sortie) est un programme contenu dans la mémoire morte ([ROM](#)) de la carte mère s'exécutant au démarrage de l'ordinateur. Il déclare les disques, configure les composants et recherche un système d'exploitation avant de le lancer. Sa tâche principale est de fournir un support de bas niveau pour communiquer avec les périphériques.

Le BIOS contient également des outils diagnostics pour vérifier sommairement l'intégrité des composants critiques comme la mémoire, le clavier, le disque dur, les ports d'entrée/sortie, etc.

Bluetooth [20]

Bluetooth est une spécification de l'industrie des télécommunications. Elle utilise une technologie radio courte distance destinée à simplifier les connexions entre les appareils électroniques. Elle a été conçue dans le but de remplacer les câbles entre les ordinateurs et les [imprimantes](#), les [scanners](#), les [souris](#), les téléphones portables, les [PDAs](#) et les appareils photos numériques.

Historique

Son nom est directement inspiré du roi danois Harald II surnommé Harald II Blåtand (« à la dent bleue »), connu pour avoir réussi à unifier les états du Danemark, de Norvège et de Suède. Le logo de Bluetooth est d'ailleurs composé des runes nordiques H et B.

- 1994 : Création par le fabricant suédois Ericsson
- 1998 : Plusieurs grandes sociétés (IBM, Intel, Microsoft, Motorola, Nokia et Toshiba) s'associent pour former le Bluetooth Special Interest Group (SIG)
- juillet 1999 : Sortie de la spécification 1.0

- Le 28 mars 2006, le « Bluetooth Special Interest Group » (SIG) annonce la prochaine génération de la technologie sans fil Bluetooth, qui sera capable d'assurer des débits cent fois supérieurs par rapport à la version actuelle passant donc de 1Mb/s à 100Mb/s (soit 12,5 Mo/s). Cette technologie - utilisée dans les téléphones mobiles, périphériques informatiques et autres appareils portables comme les assistants personnels (PDA) - va voir sa vitesse de transmission augmenter dans les années à venir, lui permettant alors d'être utilisée pour les vidéos haute définition et l'échange de fichiers avec son baladeur MP3 par exemple.

Bus informatique [21]

Un bus informatique ou bus local sert à interconnecter différents matériels informatiques. Le terme bus désigne l'ensemble des lignes de communication véhiculant les données numériques entre le processeur, la mémoire vive et les divers périphériques.

Description

Les bus sont situés à l'intérieur d'un même ordinateur et permettent de connecter les différentes parties fonctionnelles de cet ordinateur entre elles. Un bus peut interconnecter plus de deux dispositifs.

Les informations transmises peuvent être les informations utiles à échanger entre les dispositifs ou des informations de contrôle permettant de gérer l'état du bus lui-même.

Un bus est souvent caractérisé par une fréquence et le nombre de bits d'informations qu'il peut transmettre simultanément. Lorsqu'un bus peut transmettre plus d'un bit d'information simultanément on parlera d'une bus parallèle, sinon d'un bus série. La fréquence donnée est tantôt la fréquence du signal électrique sur le bus, tantôt la cadence de transmission des informations, qui peut être un multiple de la fréquence du signal.

Ainsi un bus de 32 bits dont le signal a une fréquence de 33 mégahertz peut transmettre $32 \times 33 \times 10^6$ bits par seconde soit 1056×10^9 bits par seconde, soit 132 méga-octets par seconde.

Les technologies utilisées pour fabriquer les bus sont variées, conducteurs électriques gravés sur un circuit imprimé, câble, fibre optique, etc.

Les bus parallèles

D'un point de vue physique, ce type de bus est un ensemble de conducteurs électriques parallèles. À chaque cycle de temps, chaque conducteur transmet un *bit*.

Ces bus ont donc une taille en nombre de conducteurs, et une taille en bits. Les tailles les plus courantes (en bits) sont : 8, 16, 32, 64 ou plus. Lorsque l'on parle de la taille d'un bus, cela signifie qu'il s'agit du nombre d'informations (ou bits) que le bus peut transmettre en un cycle, non compté les informations de contrôle.

Certains conducteurs supplémentaires sont affectés à la transmission des signaux de contrôles de l'état du bus.

Le bus sert à transmettre un entier informatique de la taille du bus. Les différents bits du bus ont chacun un poids différents numéroté de zéro à N-1 ou N est la taille du bus. Par exemple pour un bus 4 bits on peut transmettre 16 valeurs différents ($2^4 = 16$).

Deux bus distincts sont utilisés, un bus de données et un bus d'adresse. Le bus d'adresse est utilisé pour sélectionner la ou les cellules mémoires qui doivent être lues ou écrites, le bus de données servant à transmettre le contenu de la mémoire lui-même.

Bus série

Un bus série permet de transmettre les informations bit par bit. Toutefois il comporte plus d'une ligne permettant de transmettre des informations par l'addition d'éventuel signaux de contrôle et généralement par l'utilisation de deux lignes distinctes permettant ainsi à ces bus d'être bidirectionnels afin de permettre la transmission d'information dans les deux directions simultanément.

Le bus série transmettant les données bit par bit, il est nécessaire lorsque l'on veut par exemple transmettre un mot de 32 bits de sérialiser l'information pour sa transmission. Le lecteur devra effectuer l'opération inverse pour reconstruire le mot de 32 bits à partir des bits entrant.

Certains périphériques informatiques tels que les *souris* utilisent un bus série. Les *disques durs* de générations récentes aussi.

Bus série versus bus parallèles

A priori l'idée d'utiliser un bus parallèle semble plus efficace : un bus série, transmettant les bits d'informations un par un, peut-il être plus rapide qu'un bus parallèle les transmettant 32 par 32 ?

Les bus parallèles sont limités en cadence par des difficultés techniques et physiques. À des fréquences de fonctionnement élevées les bus parallèles produisent plus d'interférences électromagnétiques qu'un bus série ce qui perturbe la qualité des signaux électriques transmis jusqu'à les rendre inutilisables.

Chipset [17]

Chipset est un terme générique (signifiant "jeu de circuits") qui désigne un groupe de microprocesseurs et de périphériques digitaux intégrés dans un circuit (« chips » en anglais) utilisé dans un ordinateur, une console de jeux, un téléphone portable...

Son rôle est de contrôler les échanges de données entre les divers composants d'une *carte mère* tels le *processeur*, la *RAM*, la *mémoire cache* ou le *disque dur* ; il s'occupe aussi de la gestion d'énergie. Plus le chipset sera performant et plus les performances globales de l'ordinateur seront élevées.

Le terme de chipset est aussi communément utilisé pour désigner les puces de la *carte mère* dans un PC : Northbridge et Southbridge. Le Northbridge manipule typiquement des communications entre *l'unité centrale*, la *RAM*, les ports *AGP* ou *PCI Express* et le Southbridge. La puce Southbridge définit et commande le fonctionnement de tous les bus et dispositifs qui ne sont pas pris en charge par le Northbridge. Ceci inclut presque toujours le bus *PCI*, l'interface pour le *clavier* et la *souris*. Sur des machines récentes, il inclura également généralement le support de l'interface parallèle ATA et/ou du Serial ATA (pour connecter les *disques durs*, le lecteur de *CD-ROMs*, etc.), d'une interface *Ethernet*, *USB*, et IEEE 1394 (*Firewire*).

Le fabricant du chipset peut-être différent du fabricant de la carte mère. (Exemples : des PC ont des cartes mères incluant des chipset NVIDIA nForce, Intel, VIA...).

Commerce électronique [79]

Le commerce électronique ou e-commerce désigne l'échange de biens et de services entre deux entités sur les réseaux, notamment *Internet*.

On peut distinguer :

- L'échange électronique entre professionnels, souvent appelé B2B (se prononce bitoubi), acronyme anglais de Business-to-business
- Le commerce électronique à destination des particuliers, ou B2C (se prononce bitouci), acronyme anglais de Business-to-consumer. Il s'agit de sites web marchands, type télé-achat.
- Le commerce électronique entre particuliers, ou C2C (se prononce sitousi), acronyme anglais de Consumer-to-consumer. Il s'agit de sites web permettant la vente entre particuliers.

Vente à distance

Lorsque un bien est vendu dans le cadre du commerce électronique, il s'agit aussi de vente à distance, et les lois qui y prévalent s'appliquent.

En France, les lois diffèrent selon que ce sont des biens ou des services qui sont achetés en ligne. - Lors de l'achat d'un bien manufacturé vous avez un délai de réflexion de 7 jours (ouvrés) pour renvoyer votre commande. - Lors de l'achat d'un service (un voyage ou un billet d'avion) vous ne disposez pas de ce délai de rétractation.

Commerce électronique trans-frontalier

À l'heure de la mondialisation, Internet est devenu un vecteur formidable du commerce électronique. Pourtant, les questions relatives à l'achat de produits à l'étranger révèle des difficultés, notamment en cas de litige.

Les pays européens doivent pour leur part transcrire dans leurs codes nationaux les directives touchant ce domaine, ce qui rendra homogènes les règles applicables entre chaque pays de l'UE.

Dans le cadre du B2C et lorsqu'un achat a lieu en dehors de l'Union Européenne, il convient d'être prudent, de savoir à qui l'on a affaire, et de bien connaître les conditions de la vente. En cas de litige grave, le seul recours pourrait être le dépôt d'une plainte et dans le pays de l'acheteur, et dans le pays du vendeur. Le droit français protège les consommateurs en indiquant qu'un acheteur ne saurait être privé de son droit à déposer plainte dans son pays de résidence.

Il semble qu'il vaille mieux aussi avoir des notions du droit du pays dans lequel se situe le vendeur.

Lorsqu'il s'agit de B2B, le droit de la consommation laisse plutôt la place au droit du commerce international.

Quand un produit est acheté à l'étranger, les droits de douane et la TVA (ou son équivalent) sont à acquitter, comme si le produit était acheté sur le sol national. En pratique :

- pour tous les achats effectués à l'intérieur de l'Union européenne, il n'y a pas de droits de douane et la TVA qui s'applique est celle du pays d'achat du produit. Il peut donc être intéressant d'acheter dans les pays européens dont la TVA est plus faible (par exemple, celle de l'Allemagne est de 15%).
- pour tous les achats effectués en dehors de l'Union européenne, les droits de douane et la TVA sont à acquitter à l'entrée sur le territoire. Comme l'acheteur n'est généralement pas présent au moment où la commande passe la frontière (le plus souvent il s'agit d'un aéroport), les services postaux sont assermentés pour encaisser ces taxes. En général ces taxes sont appliquées sous la forme de forfait ou de manière globale (coût du produit + port par exemple) ce qui peut renchérir de beaucoup le coût final de l'achat. Les sociétés privées sont mieux organisées pour ce travail que les services postaux traditionnels.

Les produits électroniques sont souvent stoppés et taxés aux frontières. Seuls les livres, qui bénéficient d'une TVA et de droits de douane très faibles, ne sont jamais bloqués par les services postaux car le coût du recouvrement serait plus élevé que les taxes elles-mêmes.

Commutateur réseau [42]

Un commutateur réseau ou switch est un équipement qui connecte plusieurs segments dans un réseau informatique. Il utilise la logique d'un pont mais permet une topologie physique et logique en étoile. Les commutateurs sont souvent utilisés pour remplacer des [concentrateurs](#).

Chaque nœud connecté à un concentrateur reçoit les trames des autres par diffusion (broadcast), même celles qui ne lui sont pas adressées. Un commutateur, quant à lui, connecte des segments et maintient les connexions aussi longtemps que des données sont envoyées.

Concentrateur [41]

En général, un concentrateur (en anglais, hub - cette traduction est souvent utilisée en français, mais c'est un anglicisme) est le nœud central d'un réseau informatique. Il s'agit d'un dispositif électronique servant de commutateur réseau, et permettant de créer un réseau informatique local de type [Ethernet](#). Ce dispositif connecte entre eux plusieurs ordinateurs au moyen de câbles Ethernet, en répercutant les données de l'un sur l'autre, les faisant fonctionner comme s'ils ne formaient qu'un seul raccordement et qu'ils étaient directement connectés ensemble.

En utilisant un concentrateur, chaque équipement attaché à celui-ci partage le même domaine de diffusion ainsi que le même domaine de collision. Comme dans tout segment de réseau Ethernet, une seule des machines connectées peut y transmettre à la fois. Dans le cas contraire, une collision se produit et les machines doivent retransmettre leurs trames après avoir attendu un temps aléatoire.

Ce dispositif est un simple appareil de connexion ne permettant pas de protection particulière des données, par opposition au [switch](#), qui permet de diriger les données uniquement vers la machine destinataire. Il possède deux types de « ports », ou prises physiques :

- Les ports pour la connexion des ordinateurs du (sous)-réseau ;
- Le port pour extension du réseau auquel se connectent d'autres concentrateurs (il n'y en a en général qu'un seul par concentrateur). Ce type de port est en fait identique au précédent, à l'exception du fait que le câblage y est inversé (on pourrait tout aussi bien utiliser un câble Ethernet croisé sur les ports destinés aux ordinateurs).

DirectX [35]

Microsoft DirectX est une suite d'[APIs](#) multimédia intégrée au [système d'exploitation](#) Windows permettant d'exploiter les capacités matérielles d'un ordinateur. Elle est apparue pour la première fois en 1995.

DirectX fournit un ensemble de bibliothèques de fonctions essentiellement dédiées aux traitements audio/vidéo (carte vidéo, carte son, carte 3D, etc.) et aux périphériques d'entrée/sortie (joystick, [carte réseau](#), [souris](#), etc.).

L'avantage des fonctions de DirectX pour les programmeurs est que celles-ci utilisent (si possible) un algorithme alternatif (confié au processeur) quand le matériel installé ne gère pas ce type de traitement. Il fonctionne comme une surcouche de Windows, évitant théoriquement aux programmeurs de devoir s'occuper des différences matérielles qui existent entre les différents PCs. Par exemple, si une carte vidéo n'a pas de fonctions dédiées à la 3D, DirectX demandera au processeur de s'occuper du rendu d'une image de synthèse ou le rendu 3D en temps réel.

DirectX est la propriété de la société Microsoft. Ce produit n'étant pas libre, les sources ne sont pas rendues publiques, contrairement à la bibliothèque OpenGL, concurrente de Direct3D. Malgré cela, il devient de plus en plus incontournable notamment dans le domaine de la programmation des jeux vidéo 3D, Microsoft passant des accords technologiques avec les constructeurs de cartes 3D grand public.

Domain Name System [49]

Domain Name System (DNS) est un système permettant d'établir une correspondance entre une *adresse IP* et un *nom de domaine*.

Associer une adresse IP et un nom de domaine

Les ordinateurs connectés à un réseau, par exemple Internet, possèdent tous une *adresse IP*. Ces adresses sont numériques afin d'être plus facilement traitées par une machine. Selon IPv4, elles prennent la forme xxx.yyy.zzz.aaa, où xxx, yyy, zzz et aaa sont quatre nombres variant entre 0 et 255 (en système décimal).

Il n'est évidemment pas simple pour un humain de retenir ce numéro lorsque l'on désire accéder à un ordinateur d'*Internet*. C'est pourquoi le Domain Name System (ou DNS, système de noms de domaine) fut inventé en 1983 par Paul Mockapetris. Il permet d'associer à une adresse IP, un nom intelligible, humainement plus simple à retenir, appelé *nom de domaine*. fr.wikipedia.org, par exemple, est composé du domaine générique org, du domaine déposé wikipedia et du nom d'hôte fr.

Quand un utilisateur souhaite accéder à un site, comme par exemple www.free.fr, son ordinateur émet une requête spéciale à un serveur DNS, demandant 'Quelle est l'adresse de www.free.fr ?'. Le serveur répond en retournant l'adresse IP du serveur, qui est dans ce cas-ci, 213.228.0.42.

Il est également possible de poser la question inverse, à savoir 'Quel est le nom de domaine de telle adresse IP ?'. On parle alors de résolution inverse.

Plusieurs noms de domaine peuvent pointer vers une même adresse IP. Mais une adresse IP ne peut pointer que sur un unique nom de domaine.

Lorsqu'un service génère un trafic important, celui-ci peut faire appel à la technique du DNS Round-Robin, qui consiste à associer plusieurs adresses IP à un nom de domaine. Les différentes versions de Wikipedia, comme fr.wikipedia.org par exemple, sont associées à plusieurs adresses IP : 207.142.131.247, 207.142.131.248, 207.142.131.235, 207.142.131.236, 207.142.131.245 et 207.142.131.246. Une rotation circulaire entre ces différentes adresses permet ainsi de répartir la charge générée par ce trafic important, entre les différentes machines, ayant ces adresses IP.

Un système réparti

Il existe en fait des centaines de serveurs DNS dans le monde entier. Chaque serveur DNS n'a en réalité à sa disposition qu'un ensemble d'informations restreint.

Quand notre serveur DNS (par exemple, celui de notre fournisseur d'accès à Internet) doit trouver l'adresse IP de fr.wikipedia.org, une certaine communication s'instaure alors avec

d'autres serveurs DNS. Tout d'abord, notre serveur demande à des serveurs DNS peu nombreux appelés root-servers quels serveurs peuvent lui répondre pour la zone org. Parmi ceux-ci, notre serveur va en choisir un pour savoir quel serveur est capable de lui répondre pour la zone wikipedia.org. C'est ce dernier qui pourra lui donner l'adresse IP de fr.wikipedia.org.

Cependant, ce processus de résolution de nom est long. C'est pourquoi, la plupart des serveurs DNS (et notamment ceux des Fournisseurs d'accès à Internet) font aussi office de DNS cache : ils gardent en *mémoire* la réponse d'une résolution de nom afin de ne pas effectuer ce long processus à nouveau ultérieurement.

Un nom de domaine peut utiliser plusieurs serveurs DNS. Généralement, les noms de domaines en utilisent deux : un primaire et un secondaire. Seuls ces derniers peuvent fournir une réponse valable à tout instant. On parle de réponse faisant autorité (authoritative answer en anglais). Les serveurs des Fournisseurs d'accès à Internet quant à eux fournissent des réponses qui ne sont pas nécessairement à jour, à cause du cache mis en place. On parle alors de réponse ne faisant pas autorité (non authoritative answer en anglais).

Pour trouver le nom de domaine d'une IP, on utilise le même principe. Dans un nom de domaine, la partie la plus générale est à droite: org dans fr.wikipedia.org. Dans une adresse ip, c'est le contraire: 213 est la partie la plus générale de 213.228.0.42. Pour conserver une logique cohérente, on inverse l'ordre des quatre termes de l'adresse et on la concatène au pseudo domaine in-addr.arpa.. Ainsi, par exemple, pour trouver le nom de domaine de l'adresse IP 213.228.0.42, on résout 42.0.228.213.in-addr.arpa, qui est un pointeur vers www1.free.fr.

Cette architecture garantit au réseau Internet une certaine sécurité. Quand un serveur DNS tombe, le bon fonctionnement de la résolution de nom n'est pas remise en cause. De plus, elle permet de mettre à jour l'adresse IP associée à un nom de domaine dans le monde entier facilement et assez rapidement (un délai de 48 heures est généralement suffisant).

FireWire [31]

FireWire est le nom d'une norme d'interface série multiplexée, aussi connue sous le nom IEEE 1394, également aussi appelée interface iLink. Il s'agit d'un bus rapide véhiculant à la fois des données et des signaux de commandes des différents appareils qu'il relie.

Plug and Play, on peut l'utiliser pour brancher toutes sortes de périphériques gourmands en bande passante, notamment disques durs et caméscopes numériques. Elle permet l'alimentation du périphérique, ainsi que le raccordement de 63 périphériques par bus et leur branchement/débranchement à chaud.

FireWire a été inventé par Apple au début des années 1990 et peut atteindre des débits de plusieurs dizaines de Mo/s. Son objectif clairement affiché était de remplacer à terme le bus *USB*, en tout cas pour les périphériques par lesquels circulent des flux importants de données.

Le FireWire permet de disposer de débits théoriques atteignant :

- 400 Mb/s en version 1 (s400 ou IEEE 1394a)
- 800 Mb/s en version 2 (s800 ou IEEE 1394b)

Firmware [28]

En informatique, le firmware est un type de *logiciel* qui est intégré dans un composant matériel (hardware).

Le firmware réside dans une mémoire *ROM*. Le rôle du firmware est typiquement de contrôler le fonctionnement de certains composants matériels qui forment un système électronique.

D'une manière générale, le firmware est une couche intermédiaire entre le logiciel (flexibilité maximale) et le matériel (flexibilité minimale). Cette organisation apparaît clairement dans les noms en anglais: soft > firm > hard (-ware). Dans ce contexte, quand on oppose logiciel et firmware (qui est un type de logiciel) on considère que logiciel signifie logiciel de haut niveau complètement indépendant des composants matériels avec lequel il travaille. De son côté, le firmware interagit avec des composants matériels qui ne peuvent plus être modifiés une fois fabriqués, et donc le firmware ne peut subir que des évolutions limitées.

Bien que l'utilisateur final n'ait pas accès directement au firmware, il est cependant souvent possible de le modifier par l'installation de mises-à-jour pour des améliorations mineures ou des corrections de bugs. Pour cela il faut que le firmware réside dans certains types de mémoires ROM reprogrammables, le plus souvent il s'agit de mémoire flash.

En français, on utilise parfois l'expression logiciel interne pour parler de firmware. Exemples de firmware :

- les BIOS présents dans les ordinateurs de type PC ;
- les programmes qui contrôlent les différents composants d'un téléphone mobile GSM, d'un lecteur de DVD ou d'un baladeur MP3 ;
- la gestion de l'injection électronique des moteurs de voitures ;

Freeware [86]

Un graticiel ou gratuiciel (anglais freeware) est un *logiciel* qui est mis gratuitement à disposition par son créateur, mais qui est soumis à certaines contraintes quant à sa diffusion. Les graticiels sont soit des logiciels complets, soit des logiciels commerciaux qui sont diffusés de manière bridée en termes de fonctionnalités (version réduite). Ils sont parfois financés par la publicité qu'ils contiennent (Adware).

Il ne faut pas confondre graticiel et partagiciel (anglais shareware), où on peut utiliser le logiciel complet ou bridé gratuitement mais pendant une durée déterminée : un traitement de texte pourrait par exemple interdire la sauvegarde des fichiers créés, ou fonctionner uniquement pendant les 2 mois qui suivent son installation.

Il faut aussi distinguer le graticiel du logiciel libre. Le logiciel libre, même s'il est souvent gratuit, offre des libertés que la gratuité ne prend pas en compte. Notamment, un graticiel sera la plupart du temps diffusé sans les sources des programmes. La licence de distribution peut être restrictive (pas de diffusion sur cédérom, ou uniquement sur certains sites internet).

Fréquence d'horloge [36]

La fréquence d'horloge est le nombre de cycles effectués par le *processeur* en une seconde. Il est indiqué en Hertz (unité) (Hz).

Les microprocesseurs sont cadencés par une horloge qui fonctionne à plusieurs mégahertz (MHz). Le hertz (symbole: Hz) est l'unité dérivée de fréquence du système international (SI). Elle est équivalente à une oscillation par seconde. Un mégahertz est égal à un million de hertz. Au milieu des années 1980, les microprocesseurs fonctionnaient à 4,77 et 8 MHz. Courant 2004, cette vitesse d'horloge atteint 3,4 GHz sur des modèles commerciaux (5 GHz en laboratoire). Un gigahertz est égal à 1000 mégahertz. Plus la vitesse de l'horloge est élevée, plus le microprocesseur sera capable d'exécuter à un rythme élevé les instructions de base des programmes.

Les microprocesseurs actuels sont optimisés pour exécuter plus d'une instruction par cycle d'horloge, ce sont des microprocesseurs avec des unités d'exécution parallélisées. En plus ils sont dotés de procédures qui « anticipent » les instructions suivantes avec l'aide de la statistique.

FTP [119]

Le File Transfer Protocol (protocole de transfert de fichiers), ou FTP, est un protocole de communication dédié à l'échange informatique de fichiers sur un réseau TCP/IP. Il permet, depuis un ordinateur, de copier des fichiers depuis ou vers un autre ordinateur du réseau, d'administrer un site web, ou encore de supprimer ou modifier des fichiers sur cet ordinateur.

La variante sécurisée de FTP avec les protocoles SSL ou TLS s'appelle FTPS.

FTP obéit à un modèle client-serveur, c'est-à-dire qu'une des deux parties, le client, envoie des requêtes auxquelles réagit l'autre, appelé serveur. En pratique, le serveur est un ordinateur sur lequel fonctionne un logiciel lui-même appelé serveur FTP, qui rend publique une arborescence de fichiers similaire à un système de fichiers Unix. Pour accéder à un serveur FTP, on utilise un logiciel client FTP (possédant une interface graphique ou en ligne de commande).

Deux ports sont standardisés (well known ports) pour les connexions FTP : le port 21 pour les commandes et le port 20 pour les données.

HTML [76]

Le HTML, abréviation de l'anglais Hypertext Markup Language (Hypertext est parfois écrit HyperText pour marquer le T de l'abréviation), aussi appelé langage HTML, rarement traduit littéralement en langage de balisage hypertexte, est le langage informatique créé et utilisé pour écrire les [pages Web](#). HTML permet en particulier d'insérer des hyperliens dans du texte, donc de créer de l'hypertexte, d'où le nom du langage.

Langage de balisage hypertexte

HTML est un langage de description de documents à balises.

Pour expliquer les balises HTML, voici un exemple :

```
<p>Lorem ipsum dolor sit amet, consectetur adipiscing elit.</p>
```

La phrase Lorem ipsum dolor (...) est délimitée par une balise d'ouverture <p> et une autre de fermeture </p>. La seule différence entre ces deux balises est le slash. Les balises servent de délimitation, <p></p> délimitant un paragraphe.

Certaines balises ne sont pas doubles, et ne contiennent pas de balise de fermeture. Sont concernées la balise d'image img, la balise de ligne de séparation hr, et la balise de retour à la ligne br.

On assiste quelquefois à des oublis de balises de fermeture en HTML, alors qu'elles sont nécessaires. Cette erreur s'est répandue sur le web, et les navigateurs sont devenus très permissifs à une certaine époque, ce qui a tendu à les complexifier encore davantage.

Les navigateurs récents ont aujourd'hui deux niveaux de compréhension d'un fichier HTML : le niveau strict, qui affiche exactement ce que le fichier HTML demande, et le niveau brouillon, qui affiche une page en supportant les erreurs les plus fréquentes. C'est au vu de l'en-tête du document (le doctype plus précisément) que le navigateur choisit le mode à adopter.

Accéder à une page HTML

C'est le protocole HTTP qui, sur le web, permet de transférer à partir d'un [Serveur Web](#), un fichier HTML. Lorsque le serveur web reçoit une demande concernant un fichier, il est possible qu'il ait à générer une partie du fichier suivant les indications qu'aura eu soin de lui laisser l'auteur de la page. (cf langages spécialisés web.)

Les documents HTML sont identifiés par une adresse URL, et sont interprétés par un logiciel appelé [Navigateur Web](#). Grâce à ce dernier, le fichier HTML apparaît à l'écran comme l'auteur l'a voulu. Sont ainsi représentés texte, typographie, couleurs, tableaux, images, parfois du son, etc.

HTTP [77]

Le Hypertext Transfer Protocol, abrégé HTTP, littéralement « protocole de transfert hypertexte », est un protocole de communication informatique client-serveur développé pour le World Wide Web. Il est utilisé pour transférer les documents (document [HTML](#), image, feuille de style, etc.) entre le [serveur HTTP](#) et le [navigateur Web](#) lorsqu'un visiteur consulte un site Web.

HTTPS (Secured) est la variante du HTTP sécurisé avec les protocoles [SSL](#) ou [TLS](#). Il permet au visiteur de vérifier l'identité du site auquel il accède grâce à un certificat d'authentification. Il permet également de chiffrer la communication. Il est généralement utilisé pour les transactions financières en ligne : commerce électronique, banque en ligne, courtage en ligne, etc.

Hub cf. [Concentrateur](#)

Internet [26]

Internet est le nom donné au réseau informatique mondial, reposant sur le système d'adresses global des protocoles de communication [TCP/IP](#) (Transmission Control Protocol/Internet Protocol) et qui rend accessible au public des services comme le courrier électronique et le World Wide Web.

Comme Internet a été popularisé par l'apparition du World Wide Web (le Web), un système hypertexte fonctionnant sur Internet, les deux sont parfois confondus par le public non averti.

Un système hypertexte est un système contenant des documents liés entre eux par des hyperliens permettant de passer automatiquement (en pratique grâce à l'informatique) du document consulté à un autre document lié. Un document hypertexte est donc un document qui contient des hyperliens. Un hyperlien ou lien hypertexte ou simplement lien, est une référence dans un système hypertexte permettant de passer automatiquement d'un document consulté à un document lié.

Le Web est une application informatique fonctionnant sur Internet parmi d'autres, comme le courrier électronique, la messagerie instantanée ou les systèmes de partage de fichiers *peer to peer*.

Terminologie

Le terme d'origine anglaise Internet est dérivé du concept d'internetting (interconnecter des réseaux) dont la première utilisation documentée remonte à octobre 1972 par Robert Kahn (<http://www.cnri.reston.va.us/bios/kahn.html>) au cours de la première ICCC (International Conference on Computer Communications) à Washington.

Au cours de l'histoire de la création d'Internet, on trouve différents noms qui sont parfois considérés comme ancêtres du terme Internet : internetting, interconnected networks, internetworking, internetwork, international inter-connected networks, Inter Net, inter-net, International Network. Toutefois les origines exactes du terme Internet restent à déterminer. Ce flou a favorisé l'apparition de multiples explications faisant office d'origine. Aujourd'hui ceux qui prétendent détenir la véritable origine du terme sont légion (un exemple courant est de dire qu'Internet est l'acronyme d'interconnected networks). Toutefois on sait que c'est le 1er janvier 1983 que le nom Internet est devenu officiel.

La définition de ce qu'est Internet n'est pas évidente à expliciter de manière précise sans entrer dans les détails techniques, ce qui tend à une vulgarisation de la définition et facilite les confusions et imprécisions en français. Une des confusions les plus courantes porte sur le Net (en français « réseau ») et le Web (en français « toile » dans le sens « toile d'araignée »). En réaction à l'importance croissante du « phénomène Internet » et la prolifération de termes relatifs à ce phénomène dans le langage, il y a eu diverses publications au Journal Officiel de la République française (<http://www.journal-officiel.gouv.fr/>). L'une d'elle indique qu'il faut utiliser le mot Internet comme un nom commun, c'est-à-dire sans majuscule. L'Académie française recommande de dire « l'internet », comme on dit souvent « le web ».

En anglais, on utilise un article défini et une majuscule pour parler d'Internet. Cet usage vient du fait qu'Internet est de loin le plus étendu (mondial) et le plus grand internet du monde. Un internet (avec un i minuscule) est un terme anglais utilisé pour désigner une interconnexion de réseaux informatiques par internetworking (voir l'article anglais internetworking).

Dans l'usage courant, on fait référence à Internet de différentes manières. Outre les recommandations officielles, il n'est pas rare de rencontrer les termes suivants : « le Net » ou « le net », « Internet », « l'Internet », « le réseau des réseaux » ou plus simplement « le réseau » ou « le Réseau » décliné parfois en « Le réseau ». Certains termes sont utilisés à tort pour faire référence à Internet, comme par exemple : « la Toile », « le web » ou « le Web » (the Web en anglais), mais cela désigne le Web et non pas Internet. Cette confusion entre web et net existe aussi en anglais.

Internet a été conçu pour relier des réseaux informatiques hétéroclites sur des distances intercontinentales : universitaires, d'entreprises, gouvernementaux, domestiques, etc., qui peuvent eux-mêmes relier des sous-réseaux et finalement des ordinateurs.

Gouvernance/ Gestion

Un certain nombre d'organismes est en charge de la gestion d'Internet, avec des attributions spécifiques. Ils participent à l'élaboration des standards techniques, l'attribution des noms de domaines, des adresses IP, etc. :

- [ICANN](#) (L'ICANN est une organisation internationale sans but lucratif dont le rôle premier est d'allouer l'espace des adresses de protocole Internet (IP), d'attribuer les identificateurs de protocole, de gérer le système de nom de domaine de premier niveau

pour les codes génériques (*gTLD*) et les codes nationaux (*ccTLD*), et d'assurer les fonctions de gestion du système de serveurs racines. Par le contrôle qu'elle exerce sur l'affectation des noms de domaines de premier niveau, l'ICANN dérive en pratique un droit de délégation sur la vente des noms de domaines à différentes organisations, comme VeriSign pour les domaines .com et .net ou de RESTENA pour le domaine .lu.;

- [IETF](#) (L'Internet Engineering Task Force, abrégée IETF, littéralement traduit de l'anglais en « Détachement d'ingénierie d'Internet » est un groupe informel, international, ouvert à tout individu, qui participent à l'élaboration de standards pour Internet. L'IETF produit la plupart des nouveaux standards d'Internet.) ;
- [ISOC](#) (L'Internet Society est une association de droit américain à vocation internationale créée en janvier 1992 par les pionniers de l'Internet pour promouvoir et coordonner le développement des réseaux informatiques dans le monde. Elle est en 2005 l'autorité morale et technique la plus influente dans l'univers du réseau Internet.)

Technique

Internet est composé d'une multitude de réseaux répartis dans le monde entier. Chaque réseau est rattaché à une entité propre (université, fournisseur d'accès à Internet, armée) et se voit attribuer un identifiant unique appelé Autonomous System (AS). Afin de pouvoir communiquer entre eux, les réseaux s'échangent des données, soit en établissant une liaison directe, soit en se rattachant à un nœud d'échange (point de peering).

Chaque réseau est donc connecté à plusieurs autres réseaux. Lorsqu'une communication doit s'établir entre deux ordinateurs appartenant à des AS différents, il faut alors déterminer le chemin à effectuer parmi les réseaux. Aucun élément d'Internet ne connaît le réseau dans son ensemble, les données sont simplement redirigées vers un autre nœud selon des règles de routage. Environ 50 % du trafic mondial d'internet passe par l'État de Virginie.

Requis

Faire partie d'Internet, en tant que réseau de réseaux, nécessite d'être connecté à un réseau IP. Pour le grand public, du matériel et des logiciels sont nécessaires :

- Canal de communication :
 - lignes téléphoniques analogiques ou numériques
 - fibre optique
 - câble
 - satellite
- Fournisseur d'accès à Internet (FAI) (en anglais ISP pour Internet Service Provider)
- Client pour le protocole réseau utilisé (PPP, PPPoX, Ethernet, etc.)

D'autres logiciels sont eux nécessaires pour exploiter Internet suivant les usages.

- World Wide Web : un navigateur Web
- Messagerie électronique : un client SMTP et POP(POP3) ou IMAP / IMAP4

D'autres encore assurent la sécurité, par exemple : [Pare-feu](#)

Protocoles

Internet fonctionne suivant un modèle en couches. Les éléments appartenant aux mêmes couches utilisent un protocole de communication pour s'échanger des informations.

Un protocole est un ensemble de règles qui définissent un langage afin de faire communiquer plusieurs ordinateurs. Ils sont définis par des normes ouvertes, les RFC.

Chaque protocole a des indications particulières et, ensemble, ils fournissent un éventail de moyens permettant de répondre à la multiplicité et à la diversité des besoins sur Internet.

Les principaux sont les suivants :

- IP (Internet Protocol) : protocole réseau qui définit le mode d'échange élémentaire entre les ordinateurs participant au réseau en leur donnant une adresse unique sur le réseau.
- TCP : responsable de l'établissement de la connexion et du contrôle de la transmission. C'est un protocole de remise fiable. Il s'assure que le destinataire a bien reçu les données, au contraire d'UDP.
- [HTTP](#) (HyperText Transfer Protocol) : protocole mis en œuvre pour le chargement des pages Web.
- [HTTPS](#) : pendant du HTTP pour la navigation en mode sécurisé.
- FTP (File Transfer Protocol) : protocole utilisé pour le transfert de fichiers sur Internet.
- SMTP (Simple Mail Transfer Protocol) : mode d'échange du courrier électronique en envoi.
- POP3 (Post Office Protocol version 3) : mode d'échange du courrier électronique en réception.
- IMAP (Internet Message Access Protocol) : un autre mode d'échange de courrier électronique.
- IRC (Internet Relay Chat) : protocole de discussion instantanée.
- SSL ou SET : protocoles de transaction sécurisée, utilisés notamment pour le paiement sécurisé.
- UDP : permet de communiquer, de façon non fiable mais légère, par petits datagrammes (paquets).
- [DNS](#) (Domain Name System) : système de résolution de noms Internet.

Intranet [43]

Un intranet est un réseau informatique utilisé à l'intérieur d'une entreprise ou de toute autre entité organisationnelle utilisant les techniques de communication d'Internet ([IP](#)).

Logiciel [27]

Un logiciel est l'ensemble des éléments informatiques qui permettent d'assurer une tâche ou une fonction (exemple : logiciel de comptabilité, logiciel de gestion des prêts).

Le terme logiciel est souvent employé pour programme informatique et inversement. Un logiciel peut être composé d'un seul, ou d'une suite de programmes. Ce dernier cas est d'autant plus fréquent que la capacité réduite de calcul de l'ordinateur oblige à une segmentation des tâches en plusieurs modules séparés ; cependant, les énormes capacités des micro-ordinateurs actuels en regard des applications typiques de bureautique ont permis la réalisation d'applications monolithiques. Généralement, les programmes sont accompagnés d'un ensemble de données permettant de les faire fonctionner (par exemple, un jeu viendra avec de nombreuses images, animations, sons...).

Les programmes peuvent être de différentes formes :

- exécutables : ils peuvent être exécutés directement par l'ordinateur ;
généralement, ils ne peuvent être exécutés que sur un type de machine et de système d'exploitation particulier (exemple : Microsoft Windows sur un compatible PC) ;
cependant, il existe des exécutables (en bytecode) exécutables sur une variété de plate-formes (comme ceux du langage Java) ; ils visent en fait l'exécution pour une machine

Diverses présentations des logiciels

virtuelle, qui est elle-même un logiciel disponible sur les diverses plates-formes.

- fichiers sources : il s'agit généralement d'un texte respectant les règles d'écriture d'un langage de programmation particulier ; à titre indicatif, l'ordre de grandeur de la taille d'un logiciel comme Microsoft Word est d'un million de lignes de code ;
pour un langage compilé : ils doivent être traduits en un exécutable par un compilateur ;
pour un interpréteur : ils sont exécutés directement à la lecture (par exemple des scripts Perl ou PHP).
- bibliothèques : il s'agit de programmes exécutables ou source qui, en eux-mêmes, ne sont pas exécutables directement et n'offrent pas de fonctionnalité à l'utilisateur, mais fournissent des services à d'autres programmes (par exemple, on trouvera des bibliothèques permettant à un programme de charger des animations ou de jouer des sons) ; on trouve en particulier des bibliothèques dynamiques (dll Windows ou so GNU/Linux).

Les données associées au logiciel peuvent également être de différents formats : fichiers classiques, bases de données (relationnelles, hiérarchiques, etc.). Les données du logiciel peuvent être éclatées en un grand nombre de fichiers, ou tout le logiciel peut être rassemblé en un seul fichier ; par exemple, sous Windows, la définition de l'interface utilisateur, le dessin des icônes etc., sont souvent intégrés dans le même fichier que l'application principale.

Développement de logiciels

Les logiciels, suivant leur taille, peuvent être développés par une personne seule, une petite équipe, ou un ensemble d'équipes coordonnées. Le développement de grands logiciels par de grandes équipes pose de grands problèmes de coordination, en raison de la quantité importante d'informations à communiquer entre les intervenants : documentation, réunions. Pour ces raisons, le développement de logiciels dans un contexte professionnel suit souvent des règles strictes permettant le travail en groupe et la maintenance du code ; en effet, souvent, les personnes qui doivent opérer des modifications ultérieures dans le code ne sont plus les personnes qui l'ont développé.

Bugs

Des erreurs de conception dans les logiciels peuvent causer des comportements incorrects, souvent appelés bugs (ou bogues). La gravité de ceux-ci peut aller de très mineure (p.ex., apparence légèrement incorrecte d'un élément d'interface graphique), à des événements catastrophiques (explosion de la fusée Ariane vol 501, irradiation incorrecte de patients par une machine de traitement...) en passant par des pertes plus ou moins grandes de données, et, rarement, par une détérioration du matériel.

Il est difficile, pour des raisons fondamentales, de produire des logiciels sans bugs ; cependant, il existe des mécanismes par lesquels on peut limiter la quantité de bugs, voire les supprimer. Citons d'une part des préceptes d'organisation des équipes de programmation et leur méthodologie, d'autre part les technologies de recherche de bugs dans les logiciels. La recherche en informatique a développé un domaine d'étude, la vérification formelle, dont l'objectif est de certifier la qualité des logiciels et de garantir leur fiabilité. Dans l'ensemble, l'obtention de logiciels complexes peu bogués est coûteuse en hommes et en temps.

Ouverture du code source

On classe les logiciels d'après la disponibilité du code source et de la licence qui régit la distribution du programme :

- code ouvert : tout le monde peut lire le code source. Ce terme n'est pas synonyme de logiciel libre ;
- code fermé : le code source n'est disponible que pour une minorité de personnes ;
- logiciel libre : tout le monde peut étudier, copier, distribuer, modifier et distribuer des versions modifiées du logiciel (définition de la free software foundation). Les logiciels libres sont protégés pour la plupart par une licence d'utilisation ;
- logiciel propriétaire : au moins un de ces droits n'est pas rempli pour les utilisateurs. La plupart du temps, acquérir une licence d'utilisation nécessite le paiement d'une certaine somme aux créateurs du logiciel ;
- logiciel commercial : logiciel destiné à la vente, il peut être libre ou propriétaire, même s'il est souvent propriétaire.

Loi de Moore [12]

Il existe en fait deux lois de Moore.

1. La Loi de Moore a été exprimée en 1965 par Gordon Moore, ingénieur de Fairchild Semiconductor, un des deux fondateurs d'Intel. Elle indiquait que la complexité des semiconducteurs proposés en entrée de gamme doublait tous les ans depuis 1959, date de leur invention. Cette augmentation exponentielle fut rapidement nommée Loi de Moore ou, compte-tenu de l'ajustement ultérieur, Première loi de Moore.

2. En 1980, Moore énonça une seconde loi selon laquelle le nombre de transistors des microprocesseurs (et non plus de simples circuits intégrés moins complexes car formés de composants indépendants) sur une puce de silicium double tous les dix-huit mois. Même s'il ne s'agit pas d'une vraie loi physique, cette prédiction s'est révélée incroyablement exacte. Entre 1971 et 2001, la densité des transistors a doublé chaque 1,96 année. En conséquence, les machines électroniques sont devenues de moins en moins coûteuses et de plus en plus puissantes.

Cette deuxième loi est à peu près vérifiée depuis 1973, et pourrait en principe continuer encore jusqu'en 2015 avant qu'on ne bute sur des effets de bruit parasites (effets quantiques, désintégrations alpha). En 2004, ponctuellement, elle souffre un petit ralentissement dû à des difficultés de dissipation thermique, qui empêche une montée en fréquence en dépit de la taille plus faible des composants. On contourne pour le moment cette difficulté de deux façons :

- doublement sur une puce du nombre de processeurs, la fréquence restant pour sa part inchangée
- expérimentation de puces fonctionnant en mode totalement asynchrone. On s'aperçoit en effet que la simple transmission du signal d'horloge à tous les composants consomme la moitié de l'espace et de la puissance des microprocesseurs actuels.

Un autre facteur peut venir la freiner, et qui n'a cette fois-ci plus rien de physique, mais qui est d'ordre financier : il n'y a pas que la puissance des microprocesseurs qui augmente exponentiellement. Il y a aussi le coût de leurs chaînes de production, à un point tel que même des géants concurrents comme IBM et Siemens ont dû grouper leurs investissements pour arriver à suivre le mouvement. La rentabilité des générations futures de machines dépend d'un futur pour le moins incertain (beaucoup d'utilisateurs de PC, par exemple, commencent à prendre comme critère de choix prioritaire non plus la vitesse d'un PC, mais son niveau de silence) et il se pourrait que dans ces conditions ce soit une décision de financiers, et non un palier physique, qui mette fin à cette Loi de Moore.

Mémoire cache [37]

La mémoire cache (ou tout type de cache) est une mémoire intermédiaire dans laquelle se trouvent stockées toutes les informations que le *processeur central* est le plus susceptible de demander.

Elle sert donc à accélérer la communication entre un élément fournisseur (*disque dur* par exemple) plus lent que l'élément demandeur (processeur par exemple). Comme ces informations sont immédiatement disponibles, le temps de traitement se trouve diminué d'autant, ce qui mécaniquement accroît notablement les performances de l'ordinateur.

Il existe souvent plusieurs niveaux de mémoire cache : une interne au processeur, une autre intégrée sur la *carte mère*, mais on peut en avoir aussi sur le disque dur, etc.

Moteur de recherche

Un moteur de recherche est un *logiciel* permettant de retrouver des ressources (pages Web, forums Usenet, images, vidéo, etc.) associées à des mots quelconques. Certains sites Web offrent un moteur de recherche comme principale fonctionnalité ; on appelle alors moteur de recherche le site lui-même.

Fonctionnement	<p>Le fonctionnement d'un moteur de recherche se décompose en trois étapes principales.</p> <ol style="list-style-type: none">1. Le Web est systématiquement exploré par un robot d'indexation suivant récursivement tous les hyperliens qu'il trouve et récupérant les ressources jugées intéressantes. L'exploration est lancée depuis une ressource pivot, comme une page d'annuaire Web.2. L'indexation des ressources récupérées consiste à extraire les mots considérés comme significatifs (presque tous) correspondant à chaque ressource. Les mots extraits sont enregistrés dans une base de données organisée comme un gigantesque dictionnaire inverse. La partie requêtes du moteur de recherche peut ainsi rapidement retrouver les correspondances. Un algorithme gardé secret est généralement appliqué pour donner un poids variable aux correspondances, afin de pouvoir présenter les résultats des recherches par ordre de pertinence supposé. L'algorithme tient généralement compte du contexte du mot clé (titre, paragraphe, hyperlien...) et de la ressource (ressources liées, popularité du site...)3. La ressource indexée peut ensuite être retournée dans les résultats d'une recherche d'un visiteur contenant un mot clé correspondant.
Financement	<p>Les sites dont la recherche constitue le principal service peuvent se financer avec deux sources : la publicité et la vente de technologie.</p>
Publicité	<p>Les moteurs de recherche se financent principalement avec la publicité ciblée. Il s'agit de présenter des publicités correspondant aux mots recherchés par le visiteur. Pour l'annonceur, cela revient à acheter des mots clés : par exemple une agence de voyage peut acheter des mots clés comme « vacances », « hôtel » et « plage » ou « Cannes », « Antibes » et « Nice » si elle est spécialisée dans cette région.</p> <p>Le moteur de recherche peut afficher la publicité de deux manières : en encart séparé ou en l'intégrant aux résultats de la recherche. Pour le visiteur, l'encart séparé se présente comme une publicité classique. L'intégration aux résultats se fait en revanche au détriment de la pertinence des résultats et peut avoir des retombées négatives sur la qualité perçue du moteur. De ce fait, tous les moteurs ne vendent pas de placement dans les résultats.</p>

Vente de technologie

Les grandes organisations (entreprises, administrations) disposent généralement de très nombreuses ressources informatiques dans un vaste intranet. Leurs ressources n'étant pas accessibles depuis Internet, elles ne sont pas couvertes par les moteurs de recherche du Web. Elles doivent donc installer leur propre moteur si elles veulent mener des recherches dans leurs ressources. Elles constituent donc un marché pour les développeurs de moteurs de recherche.

Il arrive également que des sites Web publics utilisent les services d'un moteur de recherche pour étoffer leur offre. Ainsi Yahoo!, spécialiste de l'annuaire Web, a utilisé pendant quelques années la technologie de Google pour la recherche.

Exemples

- le moteur de recherche à mot dans le contexte [Google](#)
- l'annuaire et moteur de recherche [Yahoo!](#)

Navigateur Web [74]

Un navigateur Web est un logiciel conçu pour consulter le World Wide Web.

Le terme navigateur Web (ou navigateur Internet) est inspiré de Netscape Navigator. D'autres termes sont ou ont été utilisés. Le premier terme utilisé était browser, comme en anglais. Par la suite on a vu fureteur (surtout utilisé au Québec), butineur, brouteur, arpenteur, fouineur ou explorateur (inspiré de Internet Explorer).

Fonctionnalités

La fonction principale d'un navigateur Web est la récupération de ressources, identifiées par une [URL](#), sur un serveur Web, en utilisant le protocole HTTP. Les ressources sont ensuite affichées par le navigateur à l'utilisateur final.

Rendu des ressources

Les principales ressources qu'un navigateur peut afficher sont les pages Web au format HTML. Un navigateur graphique (au contraire d'un navigateur en mode texte) peut également afficher des images dans différents formats (PNG, JPEG, etc.)

Nom de domaine [48]

Un nom de domaine est un identifiant unique lié à une entité dont les ordinateurs sont reliés au réseau Internet. Le système est hiérarchique, permettant la définition de sous-domaine(s) d'un domaine existant.

Syntaxe

Ce nom est composé d'au minimum deux mots séparés par un point. Le suffixe sera choisi dans la liste des entités du plus haut niveau, ou Top Level Domain (abrégé TLD). Le mot précédent le suffixe peut être n'importe lequel (à quelques conditions syntaxiques près).

Il existe deux types de TLD :

1. les ccTLD (pour Country Code, code pays, c'est-à-dire les domaines nationaux), composés de deux lettres identifiants un pays (lu (Luxembourg), fr (France), aq (Antartique), pl (Pologne), gr (Grece), de (Allemagne), etc.).
2. les TLD génériques, composés de trois lettres ou plus.

Les règles de dépôt d'un nom de domaine dans ces TLD ne sont pas les mêmes pour tous.

Voici quelques noms de domaines valides : wikipedia.com, paris.fr, paris-france.fr, lions.com, louvre.museum, yahourt.biz, platon.name, gazonk.aq, frobbos.gr, etc.

On dira que wikipedia est un sous-domaine du domaine fr, ou bien, si le contexte le permet, platon est sous-domaine de .name.

Il est possible de définir des domaines dans un sous-domaine, par exemple tour-eiffel.paris.fr, signifiant que tour-eiffel est un sous-domaine du domaine paris.fr.

Nom d'hôte

Chaque domaine est peuplé de sous-domaine ou d'hôtes, c'est-à-dire d'ordinateurs. Le nom de la machine apparaîtra donc avant le premier '.' en partant de la gauche. Par exemple : www.paris.fr désignera la machine www dans le domaine paris.fr.

Système de nom de domaine (DNS)

Le **DNS** (Domain Name System) permet d'associer des adresses Internet à des noms d'hôtes, et inversement. Chaque domaine doit être défini dans deux serveurs DNS au minimum. Ces serveurs peuvent être interrogés pour associer un nom d'hôte à une adresse **IP** ou bien pour récupérer les adresses IP des serveurs de noms associés à un nom de domaine (entre autres requêtes possibles).

Dépot d'un nom de domaine

Si le nom est libre, il peut être acheté auprès d'un registrar : une société dont l'activité est la vente de noms de domaine (exemple Restena). Ces sociétés doivent être accréditées auprès de l'**ICANN**.

Les noms de domaines ne doivent pas dépasser 63 caractères.

Les noms de domaine peuvent donner lieu à des conflits juridiques, chaque partie revendiquant la propriété intellectuelle du nom litigieux.

Octet [18]

Un octet est une unité de mesure en informatique mesurant la quantité de données. Un octet est lui-même composé de 8 **bits**, soit 8 chiffres binaires.

Symboles

En français, l'octet est souvent noté « o » ou « B » de l'anglais byte. L'octet est aussi parfois noté « b », mais en général le « b » minuscule est utilisé pour noter le bit, soit 8 fois moins de données.

Le « o » n'est pas acceptable dans le Système international d'unités (SI) à cause du risque de confusion avec 0 (zéro). Le « B », quant à lui, est le symbole du bel. Cette question n'est toujours pas résolue, les unités d'information ne faisant pas partie du SI.

Usages

L'octet et ses multiples sont généralement utilisés comme mesure de la capacité de mémorisation de la **mémoire informatique**, comme la mémoire vive, les disquettes, les **disques durs** ou les **CD-ROM**. Le débit de données en octets par seconde est souvent utilisés pour indiquer les taux de transfert des bus informatiques entre les périphériques informatiques. En revanche les taux de transfert des réseaux informatiques sont plus souvent donnés en bits par seconde.

Un octet peut prendre $2^8=256$ valeurs différentes, entre 00000000 et 11111111. Par exemple : 11000100 ou 00000001 sont des **représentations binaires** d'octets, C4 et 01 étant des représentations hexadécimales. Le système hexadécimal est un système de numération utilisant la base 16. Il utilise les 10 premiers chiffres arabes puis les 6 premières lettres de l'alphabet latin : 0 1 2 3 4 5 6 7 8 9 A B C D E F.

Ces 256 valeurs permettent notamment de représenter les nombres naturels entre 0 et 255 compris ; on parle alors d'octet non signé. Si on utilise un octet pour représenter un nombre entier entre -128 et 127 compris, on parle d'octet signé.

Multiples conventionnels

Par convention, et de manière erronée selon le SI (Le Système international d'unités, ou SI, est le système d'unités le plus largement employé), les unités dérivées que sont le kilo-octet (ou kilooctet), le mégaoctet, le gigaoctet sont souvent utilisées pour représenter les valeurs suivantes en puissance de 2 :

- 1 kilo-octet (ko ou Ko) = 2^{10} octets = 1 024 octets (et pas 1000 octets comme on pourrait le supposer), soit 2 à la puissance 10..
- 1 méga-octet (Mo) = 2^{20} octets = 1 024 ko = 1 048 576 octets.
- 1 giga-octet (Go) = 2^{30} octets = 1 024 Mo = 1 073 741 824 octets.
- 1 téra-octet (To) = 2^{40} octets = 1 024 Go = 1 099 511 627 776 octets.
- 1 péta-octet (Po) = 2^{50} octets = 1 024 To = 1 125 899 906 842 624 octets.
- 1 exa-octet (Eo) = 2^{60} octets = 1 024 Po = 1 152 921 504 606 846 976 octets.
- 1 zetta-octet (Zo) = 2^{70} octets = 1 024 Eo = 1 180 591 620 717 411 303 424 octets.
- 1 yotta-octet (Yo) = 2^{80} octets = 1 024 Zo = 1 208 925 819 614 629 174 706 176 octets.

Un problème particulier au français est la formation des multiples, à cause de la voyelle initiale. Aussi voit-on aussi bien « kilo-octet », « kilooctet » que « kiloctet ».

Multiples normalisés

Depuis la normalisation de 1998 par la Commission électrotechnique internationale, les préfixes kilo, méga, giga, téra, etc, correspondent aux mêmes multiplicateurs que dans tous les autres domaines, soit des puissances de 10 :

- un kilo-quelque chose = quelque chose $\times 10^3$.
- un méga-quelque chose = quelque chose $\times 10^6$.
- un giga-quelque chose = quelque chose $\times 10^9$.
- un téra-quelque chose = quelque chose $\times 10^{12}$.

Il est à noter que l'impact de cette normalisation reste très faible, l'usage traditionnel restant largement en vigueur chez les informaticiens.

Donc, appliqué à l'informatique, cela donne :

- 1 kilooctet (ko) = 10^3 = 1 000 octets
- 1 mégaoctet (Mo) = 10^6 octets = 1 000 ko = 1 000 000 octets
- 1 gigaoctet (Go) = 10^9 octets = 1 000 Mo = 1 000 000 000 octets
- 1 téraoctet (To) = 10^{12} octets = 1 000 Go = 1 000 000 000 000 octets

Les puissances de 2 sont maintenant représentées par les symboles :

- kibi pour « kilo binaire ».
- mébi pour « méga binaire ».
- gibi pour « giga binaire ».
- tébi pour « téra binaire ».

et ainsi de suite...

L'usage de ces préfixes est très restreint et se répand très lentement.

- 1 kibioctet (Kio) = 2^{10} octets = 1024 octets
- 1 mébioctet (Mio) = 2^{20} octets = 1024 Kio

- 1 gibioctet (Gio) = 2^{30} octets = 1024 Mio
- 1 tébioctet (Tio) = 2^{40} octets = 1024 Gio
- 1 pébioctet (Pio) = 2^{50} octets = 1024 Tio
- 1 exbioctet (Eio) = 2^{60} octets = 1024 Pio
- 1 zébioctet (Zio) = 2^{70} octets = 1024 Eio
- 1 yobioctet (Yio) = 2^{80} octets = 1024 Zio

Cette distinction est d'ailleurs utilisée depuis longtemps par les fabricants de disques durs, qui ne peuvent pas se permettre de confusion dans la signification des unités. Le fait que l'usage de préfixes en puissances de 10 permette d'afficher commercialement des capacités supérieures à celles données par les puissances de 2 n'est certes pas nuisible du point de vue de la mise en marché. Ainsi, un disque dur de 100 gigaoctets (100×10^9 octets) contient le même nombre (arrondi) d'octets qu'un disque de 93,13 gibioctets ($93,13 \times 2^{30}$ octets). Les disques durs courants étant divisés en secteurs de 512 octets, un comptage en unités de 1024 octets serait plus naturel — du moment que les préfixes binaires sont utilisés.

Open source [88]

L'expression Open Source caractérise les *logiciels* dont le code source est visible, modifiable et librement redistribuable sous certaines conditions, ces conditions peuvent être plus ou moins strictes. La formulation de ces conditions constitue d'ailleurs le critère principal qui différencie le logiciel open source du Logiciel libre.

L'utilisation des termes Open Source a été suggérée par Christine Peterson du Foresight Institute, pour lever l'ambiguïté sémantique du mot anglais *free* qui signifie libre au sens de « liberté », mais également libre au sens « libre accès, gratuité », et signifier aux utilisateurs qu'un logiciel a un coût ; le but était aussi d'être plus Business friendly, le terme Logiciel libre ayant une connotation libertaire qui risquait de rebuter quelques entreprises. D'un point de vue économique, la marque Open Source contribuait à la création d'une nouvelle forme de marché et d'économie, bien plus sexy que les termes « Logiciel Libre », dans le contexte de la nouvelle économie, où l'on parlait de Business model, angel, bourse, etc.

Afin d'éviter que ce terme ne soit galvaudé, utilisé à mauvais escient et dilué, Eric Steven Raymond a essayé initialement de le déposer, sa tentative ayant échoué, l'Open Source Initiative créée par Bruce Perens et Eric Steven Raymond délivre désormais le label OSI approved pour les licences qui satisfont aux critères définis dans L'Open Source Définition. L'Open Source Définition est une adaptation des Free Software Guidelines du projet Debian écrite par Bruce Perens, les conditions qui définissent comment les modifications faites au logiciel doivent être reversées dans le pot commun sont moins strictes que celles définies par la FSF pour le Logiciel libre.

L'Open Source Initiative a défini les critères nécessaires afin de pouvoir utiliser l'appellation Open Source. Il y a donc une définition de l'Open Source dont voici les critères essentiels :

- Libre redistribution
- Code source disponible
- Travaux dérivés possibles

Le fait de disposer des sources d'un logiciel ne suffit pas à dire qu'il est Open Source™. Dans tous les cas, on se référera à la licence d'utilisation du logiciel.

Par ailleurs, une confusion peut exister entre les termes logiciel libre (free software) et open source. Bien que tous les logiciels libres rentrent dans les critères de l'Open Source™, l'esprit qui anime les logiciels libres (copyleft) et les règles plus restrictives qui les composent incitent à clairement les différencier.

Page Web [75]

La page Web est l'unité de consultation du World Wide Web. Ce terme a une signification pratique ; il n'a pas de définition technique formelle. C'est un document informatique qui peut contenir du texte, des images, des formulaires à remplir et divers autres éléments multimédia et interactifs.

Une page Web peut être téléchargée et consultée à l'aide d'un logiciel appelé *navigateur Web*. Les navigateurs les plus utilisés affichent la page sur écran dans une fenêtre et permettent de l'imprimer. Cependant, les technologies de base utilisées par les page Web ont été conçues pour permettre la consultation sur du matériel varié : écrans de toute taille, terminal en mode texte, imprimante, dispositif braille, synthétiseur vocal. Le terme page est donc réducteur.

Techniquement, une page Web est constituée d'une ou plusieurs ressources distinctes. La principale ressource est généralement un document écrit en langage *HTML* qui contient le texte et définit la disposition des autres ressources : images, animations, sons, programmes ou autres documents. Chaque ressource est identifiée par une *URL*. Ceci permet d'intégrer dans une page des ressources provenant de n'importe quel serveur Web.

PCI [32]

Le Peripheral Component Interconnect (PCI) est un standard de *bus local* (interne) d'un ordinateur, situé sur la *carte mère*.

Un des intérêts du bus PCI est que deux cartes PCI peuvent dialoguer entre elles sans passer par le processeur et que la norme assure une parfaite compatibilité entre cartes et unité centrale.

Historique

La spécification de ce bus est due à Intel. Elle a été publiée en 1992 et implantée pour la première fois en 1994 sur des cartes mères pour processeur Intel 80486.

Il est généralement utilisé dans les ordinateurs personnels (PC ou Mac notamment).

Le bus PCI (ainsi que l'*AGP*) va être remplacé pour les cartes graphiques par une version plus rapide, le *PCI Express*.

Spécification

La spécification du bus PCI décrit la taille du bus (dont l'espacement des conducteurs), les caractéristiques électriques, les chronogrammes du bus et les protocoles.

Il existe plusieurs variantes de ce bus :

- bus 32 bits à 33 MHz (soit une bande passante maxi de 133 Mo/s) (la plus répandue) ;
- bus 32 bits à 66 MHz (soit une bande passante maxi de 266 Mo/s) ;
- PCI 2.2 : bus 64 bits à 66 MHz (soit une bande passante maxi de 533 Mo/s) ;
- PCI-X : 133 MHz (soit une bande passante maxi de 1066 Mo/s), utilisé principalement dans les machines professionnelles ;
- PCI-X 2.0 : 266 MHz (soit une bande passante maxi de 2133 Mo/s) ;

- PCI Express : norme dérivée du PCI, destinée à le remplacer dans les ordinateurs personnels ;
- Mini PCI : dérivé destiné à être intégré dans les ordinateurs portables.

Utilisations

Cartes généralement connectées à ce port :

- carte son
- carte graphique
- carte réseau

PCI Express [33]

Le PCI Express (anciennement 3GIO, 3rd Generation Input/Output) est un bus local série (ne pas confondre avec le PCI-X) destiné à remplacer à partir de 2004 tous les bus internes d'un PC, dont le PCI et l'AGP. Il a été développé à l'origine par Intel et est devenu depuis une norme officielle. Il est présent sur la carte mère et sert à connecter des cartes filles.

Son avantage est d'être non seulement plus rapide que les bus existant, tout en étant dérivé de la norme PCI, ce qui permet aux différents constructeurs d'adapter très simplement leur cartes d'extension existantes sans modifications importantes.

PCMCIA [51]

PCMCIA (pour Personal Computer Memory Card International Association), ou PC Card, est un format de carte d'extension ultra-plat, dédié aux ordinateurs portables et à d'autres périphériques. Le standard a été développé par une association de constructeurs.

Présentation

Le format PCMCIA est *plug-and-play*, c'est-à-dire qu'on peut brancher et débrancher les cartes sans couper l'ordinateur ou le périphérique. Ce bus informatique a une longueur de 32 bits (au format CardBus, et 16 bits au format PC Card) et est cadencé à 33 MHz. Il peut transmettre $32 \times 33 \times 10^6$ bits par seconde soit 132 méga-octets par seconde de débit théorique.

Il en existe deux sortes, numérotées 1 et 2. La deuxième comporte trois sous-types différenciés par leur épaisseur :

- Type I, épaisseur 3,3 mm. C'est la carte la plus fine, elle est utilisée pour la mémoire flash.
- Type II, épaisseur 5 mm. Elle est utilisée pour le modem et la carte réseau.
- Type III, épaisseur 10,5 mm. C'est la carte la plus volumineuse, elle est utilisée pour le disque dur et les cartes combinées modem/réseau.

PDA ou Assistant personnel [39]

Un assistant personnel est un appareil numérique portable, souvent appelé par son acronyme anglais PDA pour Personal Digital Assistant. Le concept est inventé par la société Apple avec son Newton (1993-1998), mais arrivé trop tôt, celui-ci ne connaîtra pas le succès escompté.

Il s'agit d'un petit boîtier de la taille d'une calculatrice, qui tient dans la main, abritant une architecture informatique et doté d'un écran tactile et parfois d'un clavier incorporé avec des petites touches. Un stylet est souvent associé à l'écran tactile.

Le PDA est utilisé principalement pour ses fonctions d'agenda, de répertoire téléphonique et de bloc-notes, mais les avancées technologiques ont permis de lui adjoindre des fonctionnalités multimédia, telles que le dictaphone, le lecteur de mp3, d'images, de vidéo, et parfois le téléphone (avec une puce GSM intégrée) et/ou le GPS.

Un stylet permet d'écrire directement sur l'écran du PDA pour y enregistrer ou extraire des informations, au moyen soit d'un langage écrit simplifié (à chaque caractère correspond un mouvement particulier du stylet), soit d'un clavier émulé.

La mémoire interne (en général plusieurs méga-octets) de certains PDA peut être considérablement augmentée en lui adjoignant une mémoire externe sous la forme d'une carte-mémoire que l'on enfiche dans le PDA (selon les formats, les cartes mémoires peuvent stocker entre 16 Mo et 4 Go).

Quelle que soit la marque du PDA, des milliers de logiciels gratuits ou payants sont disponibles sur l'Internet et peuvent y être installés dans la limite de la mémoire du PDA (calculatrices scientifiques, bases de données, jeux, gestionnaires de comptes bancaires, etc.).

Système d'exploitation les plus répandus :

- Palm OS édité par la société PalmSource;
- Pocket PC aussi appelé Windows Mobile, édité par Microsoft.

L'extension des capacités techniques a permis aux PDA de devenir communicants : au-delà des liaisons infrarouge d'origine, les PDA embarquent des liaisons [WiFi](#) ou [Bluetooth](#), et peuvent faire fonction de téléphone mobile. Ils se trouvent alors en concurrence frontale avec les appareils téléphoniques mobiles qui dans le même temps ont également accru leurs capacités techniques et fonctionnelles jusqu'à être couramment appelés Smartphones (téléphones intelligents).

Les PDA disposent d'un écran plus grand, mais sont de ce fait légèrement moins maniables que les Smartphone. Ceux-ci sont en revanche handicapés dans leur faible capacité d'écriture. Pour le reste, les différents systèmes d'exploitation proposent des fonctionnalités, des performances et des compatibilités comparables.

Plug and Play [46]

Le PNP ou Plug and Play, littéralement branche et joue, est une procédure permettant aux périphériques récents d'être reconnus rapidement et automatiquement par le système d'exploitation dès le redémarrage après l'installation matérielle. Cette procédure permet l'installation en requérant un minimum d'intervention de la part de l'utilisateur et donc en minimisant les erreurs de manipulation et de paramétrage.

Poste-à-poste ou P2P ou Peer to peer [47]

Le terme poste-à-poste est la traduction (initialement adoptée au Canada) de l'anglais peer-to-peer, laquelle est souvent abrégée P2P. On peut aussi traduire par « pair à pair » ou « égal à égal ».

P2P désigne un type de protocole de communication sur réseau informatique dont les éléments (les nœuds) ne jouent pas exclusivement les rôles de client ou de serveur mais fonctionnent des deux façons, en étant à la fois clients et serveurs des autres nœuds de ces réseaux, contrairement aux systèmes de type client-serveur, au sens habituel du terme.

Les logiciels de téléchargement P2P suscitent une vive polémique à l'heure actuelle. En effet, selon les sociétés de droits d'auteur, les réseaux P2P servent presque exclusivement au transfert illégal de fichiers protégés par le copyright. Selon les pays, ceci peut ou non entraîner l'interdiction de logiciels de P2P (l'interdiction se basant sur le fait que l'utilisation principale qui est faite du logiciel est contraire à la loi). Cependant, les divers jugements rendus varient énormément d'une affaire à l'autre et d'un pays à l'autre.

En revanche, dans la plupart des pays occidentaux, les utilisateurs de logiciels P2P sont de plus en plus souvent la cible de procès de la part des majors du disque et du cinéma. En France, la loi sur le droit d'auteur interdit strictement toute exploitation d'une œuvre sans l'autorisation de ses ayants-droit, à quelques exceptions près, parmi lesquelles figure le droit à la copie privée. En pratique, cela se traduit globalement par la condamnation des personnes poursuivies pour avoir partagé des fichiers sous copyright, tandis que, généralement, le téléchargement de fichiers audios ou vidéos pour une utilisation personnelle n'est pas condamnée. Ici encore, les jugements ne sont pas parfaitement homogènes et varient d'un tribunal à l'autre. Par ailleurs, il est à noter que la notion de copie privée n'existe pas pour les logiciels.

Proxy

Un serveur mandataire est un serveur qui a pour fonction de relayer différentes requêtes et d'entretenir un cache des réponses. Connu en anglais sous le terme de « Proxy server », il a été inventé par le Centre européen de recherche nucléaire en 1994. Il a été prévu à l'origine pour relier à Internet des réseaux locaux n'utilisant pas le protocole TCP/IP; il a été depuis doté de nouvelles fonctions concernant :

- le cache ;
- la journalisation des requêtes (« logging ») ;
- la sécurité du réseau local ;
- le filtrage et l'anonymat.

Aujourd'hui, les réseaux locaux utilisent le protocole TCP/IP et peuvent être reliés à Internet via une simple passerelle ou un routeur, mais l'utilité des serveurs mandataires est toujours aussi importante, notamment dans le cadre de la sécurisation des systèmes d'information.

Réseau local [23]

Un réseau local (en abrégé RLE pour réseau local d'entreprise, en anglais LAN pour Local Area Network) est un réseau informatique à une échelle géographique relativement restreinte, par exemple une salle informatique, une habitation particulière, un bâtiment ou un site d'entreprise.

Le terme de LAN pouvant couvrir des réalités fort différentes, géographiques, techniques, etc., on emploiera le terme de réseau intra-site pour la partie du réseau d'une entreprise qui dessert un site, par opposition au réseau inter-site qui relie les différents sites de l'entreprise.

Pour le particulier, le grand public, les LAN peuvent être vus comme un moyen de partager une connexion Internet, pour travailler sur un même serveur ou pour jouer à des jeux en réseau.

Routeur [89]

Un routeur est un matériel de communication de réseau informatique. Son travail est de déterminer le prochain nœud du réseau auquel un paquet de données doit être envoyé, afin que ce dernier atteigne sa destination finale le plus rapidement possible.

Le routage est souvent associé au protocole de communication *IP*, même si d'autres protocoles routables moins populaires existent.

Bien que les ordinateurs ordinaires peuvent être utilisés pour faire le routage, les routeurs modernes sont plus souvent des ordinateurs très spécialisés, habituellement avec du matériel supplémentaire pour accélérer des fonctions, comme le transfert (acheminement) de paquets.

Les routeurs actuels jouent donc pour les données un rôle proche de celui des commutateurs téléphoniques pour la voix. Certaines fonctions de ces derniers sont d'ailleurs de plus en plus reprises par les routeurs dans la convergence appelée voix ou téléphonie sur IP (VoIP).

Un routeur doit être connecté à au moins deux réseaux informatiques pour être utile, sinon il n'aura rien à router. L'appareil crée ou maintient une table, appelée table de routage, qui contient les meilleures routes vers d'autres réseaux via les métriques associées à ces routes.

Serveur [80]

Un serveur informatique, appelé serveur lorsque le contexte s'y prête, est un ordinateur ou un programme informatique qui partage des ressources -- comme ses périphériques et ses disques durs -- avec d'autres ordinateurs clients sur un réseau informatique. Il est possible pour un ordinateur d'être client et serveur en même temps.

Serveur FTP [25]

Un serveur *FTP* (pour File Transfer Protocol) permet, comme son nom l'indique de transférer des fichiers par Internet. Si vous en avez l'autorisation, vous pouvez télécharger et envoyer des fichiers sur un ordinateur distant faisant fonctionner un tel serveur.

Serveur HTTP [24]

Un serveur *HTTP* ou démon HTTP ou HTTPd (HTTP daemon) ou (moins précisément) serveur Web, est un logiciel servant des requêtes respectant le protocole de communication HyperText Transfer Protocol (HTTP), qui a été développé pour le World Wide Web.

Un ordinateur sur lequel fonctionne un serveur HTTP est appelé serveur Web. Le terme « serveur Web » peut aussi désigner le serveur HTTP (le logiciel) lui-même. Les deux termes sont utilisés pour le logiciel car le protocole HTTP a été développé pour le Web et les pages Web sont en pratique toujours servies avec ce protocole. D'autres ressources du Web comme les fichiers à télécharger ou les flux audio ou vidéo sont en revanche fréquemment servies avec d'autres protocoles.

Les serveurs HTTP les plus utilisés sont :

- Apache HTTP Server de la Apache Software Foundation, successeur du NCSA httpd ;
- Internet Information Services de Microsoft ;

Le plus populaire est Apache qui sert environ 67% des sites Web en 2004 selon Netcraft (voir http://news.netcraft.com/archives/web_server_survey.html en anglais).

Historiquement, d'autres serveurs HTTP importants furent CERN httpd, développé par les inventeurs du Web, abandonné le 15 juillet 1996 et NCSA httpd, développé au NCSA en même temps que Mosaic, abandonné mi-1994.

Serveur Web

Le terme serveur Web désigne :

- un ordinateur tenant le rôle de serveur informatique sur lequel fonctionne un logiciel *serveur HTTP* ;
- le logiciel *serveur HTTP* lui-même.

La plupart des ordinateurs utilisés comme serveur Web sont reliés à *Internet* et hébergent des sites Web du World Wide Web. Les autres serveurs se trouvent sur des *intranets* et hébergent des documents internes d'une entreprise, d'une administration, etc.

Shareware [87]

Un partagiciel ou shareware est un *logiciel*, protégé par le droit d'auteur, dont l'usage peut être limité dans le temps, à moins d'en retribuer l'auteur.

Étymologie

Le mot partagiciel est un calque du mot anglais shareware également très souvent utilisé en français. Shareware est lui-même une contraction de share et software. Share se traduit ici par contribution. On peut aussi trouver le terme contribuciel, mais plus rarement.

Erreurs fréquemment commises

Un partagiciel peut facilement être confondu avec un abandonware ou avec un *logiciel libre*.

Un partagiciel n'est pas un logiciel libre car:

- il est souvent livré sans son code source,
- il n'est pas possible de le distribuer sans que celui qui en fait l'acquisition n'ait à payer une licence.

Généralités

Un partagiciel peut être utilisé gratuitement et librement pendant une durée ou un nombre d'utilisations qui sont indiqués par l'auteur. Cela permet de tester les fonctionnalités et voir si elles correspondent à ses besoins.

Au bout de cette période d'essai, il est possible soit de payer une contribution (souvent modique) et continuer à utiliser le logiciel, soit de le désinstaller. Il est également permis de distribuer le logiciel à une autre personne, toujours pour essai.

Hormis l'utilisation légale du produit, le paiement de la licence peut aussi débloquent un certain nombre de fonctionnalités jusqu'alors inaccessibles comme la sauvegarde, la réception régulière de mises à jour et, parfois, la possibilité de prendre contact avec l'auteur.

Certains auteurs ne demandent que l'envoi d'une carte postale comme paiement de la licence, dans ce cas on parle de Carticiel (ou Cardware en anglais).

Généralement conçus par des passionnés, les programmes diffusés en partagiciel sont souvent de bon niveau.

Évolution du concept

Les premiers partagiciels étaient disponibles en version complète et non limitée dans le temps. Ce mode de distribution n'a pas vraiment fonctionné : les clients continuaient d'utiliser

le logiciel sans le payer. Depuis lors, les logiciels ont évolué en trois branches selon la conception qu'en a l'auteur :

- L'auteur peut choisir de maintenir la disponibilité de son logiciel en version complète et non limitée, mais ajoute un message récurrent pour rappeler à l'utilisateur qu'il doit payer ce logiciel s'il l'utilise régulièrement.
- Il peut préférer limiter l'usage de son logiciel, en le distribuant comme une version de démonstration : l'usage du logiciel ou de certaines fonctionnalités sont bloqués après une période d'essai, afin de forcer l'utilisateur à payer le logiciel. Certains de ces logiciels (en réalité, ce sont des versions de démonstration) ont des fonctionnalités bloquées même pendant la période d'essai ; dans ce cas ce sont souvent des fonctions de confort. Elles sont débloquées lors du paiement de la contribution.
- Enfin, il peut au contraire décider que les utilisateurs sont libres de choisir de payer ou non le logiciel, selon l'utilisation qu'ils en font. Ce mode de distribution, de plus en plus répandu, se nomme donationware en anglais.

Switch cf. [Commutateur Ethernet](#)

Système binaire [38]

Le système binaire est un système de numération utilisant la base 2. On nomme couramment bit (de l'anglais binary digit, soit « chiffre binaire ») les chiffres de la numération binaire. Ceux-ci ne peuvent prendre que deux valeurs, notées par convention 0 et 1.

Usage

Le système binaire est souvent utilisé pour représenter des valeurs telles que « vrai » et « faux », « tout » et « rien », « marche » et « arrêt » (on et off en anglais). Il convient notamment pour représenter le fonctionnement de l'électronique numérique utilisée dans les ordinateurs, d'où son usage en informatique. S'il se montre peu efficace pour l'usage humain (il faut 16 chiffres binaires pour représenter un nombre décimal de 5 chiffres !), il permet d'utiliser en électronique des circuits de commutation, dont le coût unitaire est si faible (quelques picoeuros) que la charge des traductions depuis et vers le décimal ne constitue plus un problème.

Codage binaire

Le codage le plus courant est l'équivalent en base deux de la numération de position que nous utilisons quotidiennement en base 10.

Pour trouver la représentation binaire d'un nombre, on le décompose en somme de puissances de 2. Par exemple avec le nombre dont la représentation décimale est 59.

$$59 = 1 \times 32 + 1 \times 16 + 1 \times 8 + 0 \times 4 + 1 \times 2 + 1 \times 1$$

$$59 = 1 \times 2^5 + 1 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0$$

$$59 = 111011 \text{ en binaire}$$

Système d'exploitation [19]

Un système d'exploitation (SE ou OS en anglais pour Operating System) est un ensemble cohérent de [logiciels](#) permettant d'utiliser un ordinateur et tous ses éléments (ou périphériques). Il assure le démarrage de celui-ci et fournit aux programmes applicatifs les interfaces pour contrôler les éléments de l'ordinateur. Les programmes applicatifs n'ont

traditionnellement pas vocation à être considérés comme partie intégrante du système, mais ce point de vue est en train d'évoluer.

Composition

Typiquement, un SE est composé de :

- un noyau ;
- un ensemble d'outils système.

Le noyau assure les fonctionnalités suivantes :

- pilotes de périphériques ;
- gestion des processus :
 - gestion de la mémoire (il la distribue aux processus) ;
 - ordonnancement des processus (répartition du temps processeur).
- systèmes de fichiers ;
- protocoles réseau (TCP/IP, etc.).

Historique

Les systèmes d'exploitation existent depuis le milieu des années 1950 ; ils ont considérablement évolué depuis.

- Beaucoup ont disparu, d'autres ont été ré-écrits suite à l'évolution du matériel.
- La plupart des systèmes d'exploitation actuels proposent un environnement graphique pour interagir simplement avec l'utilisateur.

Liste

Les plus connus sont :

- systèmes d'exploitation édités par Microsoft :
 - Windows 1 à 3.11, 95, 98, 98 SE, NT, Me, 2000, XP ;
- Mac OS : le premier système d'exploitation des ordinateurs Apple Macintosh, qui succéda aux systèmes Lisa et Apple II, et fut suivi de Mac OS X ;
- Dérivés d'UNIX (sous différentes déclinaisons : BSD, System V, etc.) dont :
 - GNU/Linux : un système d'exploitation libre s'appuyant sur le noyau Linux et les outils GNU (distributions: Debian, Mandriva (MandrakeLinux), Red Hat, Fedora, SuSE, Slackware, EduLinux...)
 - la famille BSD : un effort réussi pour rendre sa liberté au système de Berkeley comprenant : NetBSD, OpenBSD, FreeBSD et ses dérivés, PicoBSD et DragonFly BSD, Darwin (sur lequel est construit Mac OS X, semi-propriétaire).
 - les UNIX propriétaires : AIX (IBM, SystemV), A/UX (Apple, SystemV), BOS (Bull Operating System), Irix (Silicon Graphics, SystemV), HP-UX (Hewlett Packard, SystemV), NeXTSTEP (NeXT, BSD), Sinix (Siemens), Solaris (Sun, SystemV), SunOS (Sun, BSD), Tru64 (Compaq).
- les systèmes d'exploitation grands systèmes (mainframes) : IBM: MVS, VM, DOS/VSE, TPF, Bull: GCOS ;

URL ou Universal Resource Locator [73]

Un URL fournit une méthode standardisée pour localiser un document et n'importe quel type de ressources sur Internet. Un URL est toujours représenté sous le format `protocole://machine.domaine[:port]/[chemin_d'accès]`

Voici deux des protocoles acceptés les plus importants :

- `http` pour Hyper Text Transport Protocol
- `ftp` pour un transfert de fichier via FTP

Elle est informellement appelée une adresse Web ou URL, sigle signifiant Uniform Resource Locator en anglais, littéralement « repère uniforme de ressource ». L'usage du féminin ou du masculin pour l'abréviation « URL » (un ou une URL ?) semble assez flottant.

Quelques exemples pratiques :

- URL de Wikipédia :
`http://fr.wikipedia.org/`
- URL de la page "URL" sur Wikipédia :
`http://fr.wikipedia.org/wiki/URL`
- URL d'un fichier sur un site FTP :
`ftp://ftp.rfc-editor.org/in-notes/rfc2396.txt`

USB [30]

L'Universal Serial Bus (USB) est un bus qui permet de connecter des périphériques externes à un ordinateur (hôte dans la littérature USB). Il supporte 127 périphériques simultanés. Le bus supporte les branchements et débranchements à chaud, fournit l'alimentation électrique des périphériques.

La version 1.1 du bus peut communiquer dans deux modes : lent (1,5 Mbit/s) ou rapide (12 Mbit/s).

- Le mode lent permet de connecter des périphériques qui ont besoin de transférer peu de données, comme les *claviers* et *souris*.
- Le mode rapide est utilisé pour connecter des *imprimantes*, *scanners*, *disques durs*, graveurs de *CD*, et autres périphériques ayant besoin de plus de rapidité. Néanmoins il est insuffisant pour beaucoup de périphériques de stockage de masse (par exemple, il ne permet que la vitesse 4x sur les lecteurs/graveurs de CD).
- La nouvelle version de ce bus, USB 2.0, comporte un troisième mode permettant de communiquer à 480 Mbit/s. Il est utilisé par les périphériques rapides : disques durs, graveurs...

Attention ! On appelle maintenant (depuis fin 2002) « USB 2.0 » tout dispositif USB, même à la norme 1.1 :

- USB 2.0 Full Speed pour un dispositif transmettant au maximum à 12 Mbit/s (ex USB 1.1).
- USB 2.0 High Speed pour un dispositif transmettant jusqu'à 480 Mbit/s (ex USB 2.0).

Les logos apposés sur le dispositif diffèrent. Voir les détails sur le site <http://www.usb.org>



Connecteur USB de type A



Connecteur USB de type B

VoIP [120]

La voix sur réseau IP, parfois appelée téléphonie IP ou téléphonie sur Internet et souvent abrégée en « VoIP » (Voice over IP), est une technique qui permet de communiquer par la voix via Internet ou tout autre réseau acceptant le protocole TCP/IP.

Wi-Fi [40]

Le Wi-Fi (également orthographié Wi-fi, WiFi, Wifi ou encore wifi) ou l'ASFI (pour Accès Sans Fil à Internet) est une technologie de réseau informatique mise en place pour fonctionner en réseau interne et depuis devenue un moyen d'accès à haut débit à Internet.

Présentation

La norme IEEE 802.11 (ISO/CEI 8802-11) est un standard international décrivant les caractéristiques d'un réseau local sans fil (WLAN, Wireless LAN). Le nom Wi-Fi (contraction de Wireless Fidelity, Fidélité sans fil) correspond initialement au nom donné à la certification délivrée par la WECA (Wireless Ethernet Compatibility Alliance), l'organisme chargé de maintenir l'interopérabilité entre les matériels répondant à la norme 802.11. Par abus de langage (et pour des raisons de marketing) le nom de la norme se confond aujourd'hui avec le nom de la certification. Ainsi un réseau Wi-Fi est en réalité un réseau répondant à la norme 802.11.

Grâce au Wi-Fi il est possible de créer des réseaux locaux sans fil à haut débit. Dans la pratique le Wi-Fi permet de relier des ordinateurs portables, des machines de bureau, des assistants personnels (*PDA*) ou même des périphériques à une liaison haut débit (11 Mbit/s) sur un rayon de plusieurs dizaines de mètres en intérieur (généralement entre une vingtaine et une cinquantaine de mètres). Dans un environnement ouvert la portée peut atteindre plusieurs centaines de mètres voire dans des conditions optimales plusieurs dizaines de kilomètres.

Ainsi des fournisseurs d'accès Internet commencent à irriguer des zones à fortes concentration d'utilisateurs (gares, aéroports, hôtels, trains, etc.) avec des réseaux sans fil connectés à Internet. Ces zones d'accès sont appelées « hot spots ».

Les iBooks d'Apple furent en 1999 parmi les premiers ordinateurs grand public à proposer un équipement Wi-Fi intégré (sous le nom d'Airport), bientôt suivis par le reste de la gamme. À partir de 2003, on voit aussi apparaître des modèles de PC portables bâtis autour de la technologie Intel Centrino, qui leur permettent une intégration similaire. Les autres modèles de PC doivent encore s'équiper d'une carte d'extension adaptée (*PCMCIA*, *USB*, Compact Flash, *PCI*, MiniPCI, etc.).

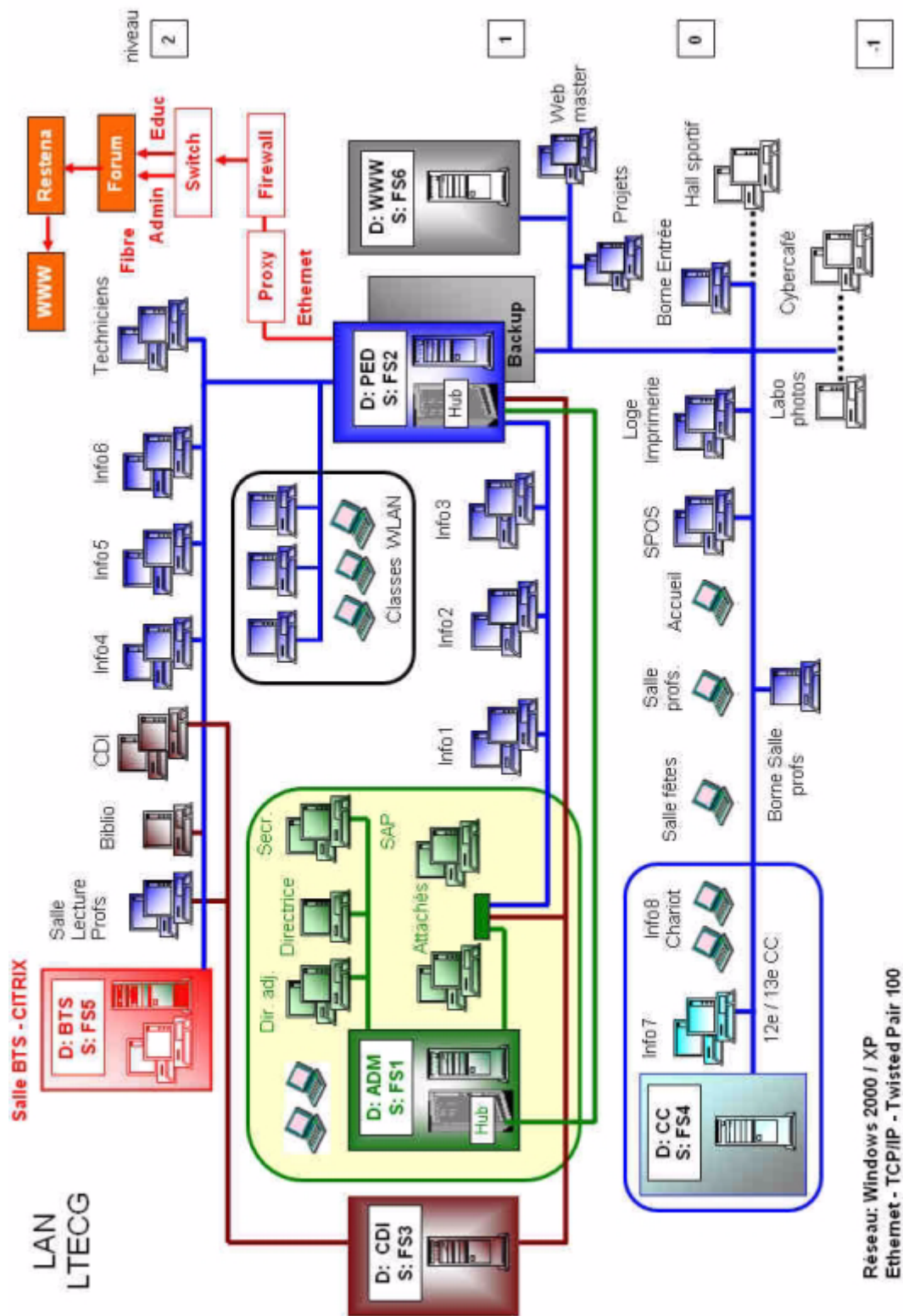
Les différentes normes Wi-Fi

La norme IEEE 802.11 est en réalité la norme initiale offrant des débits de 1 ou 2 Mbit/s. Des révisions ont été apportées à la norme originale afin d'optimiser le débit (c'est le cas des normes 802.11a, 802.11b et 802.11g, appelées normes 802.11 physiques) ou bien préciser des éléments afin d'assurer une meilleure sécurité ou une meilleure interopérabilité. Voici un tableau présentant quelques différentes révisions de la norme 802.11 et leurs significations :

Norme	Nom	Description
802.11a	Wi-Fi 5	La norme 802.11a (baptisé Wi-Fi 5) permet d'obtenir un haut débit (54 Mbit/s théoriques, 30 Mbit/s réels). La norme 802.11a spécifie 8 canaux radio dans la bande de fréquence des 5 GHz.
802.11b	Wi-Fi	La norme 802.11b est la norme la plus répandue en base installée actuellement. Elle propose un débit théorique de 11 Mbit/s (6 Mbit/s réels) avec une portée pouvant aller jusqu'à 300 mètres dans un environnement dégagé. La plage de fréquence utilisée est la bande des 2,4 GHz, avec en France 13 canaux radio disponibles.
802.11d	Internationalisation	La norme 802.11d est un supplément à la norme 802.11 dont le but est de permettre une utilisation internationale des réseaux locaux 802.11. Elle consiste à permettre aux différents équipements d'échanger des informations sur les plages de fréquence et les puissances autorisées dans le pays d'origine du matériel.
802.11e	Amélioration de la qualité de service	La norme 802.11e vise à donner des possibilités en matière de qualité de service au niveau de la couche liaison de données. Ainsi cette norme a pour but de définir les besoins des différents paquets en terme de bande passante et de délai de transmission de telle manière à permettre notamment une meilleure transmission de la voix et de la vidéo.
802.11f	Itinérance (roaming)	La norme 802.11f est une recommandation à l'intention des vendeurs de points d'accès pour une meilleure interopérabilité des produits.
802.11g		La norme 802.11g est la plus répandue dans le commerce actuellement. Elle offre un haut débit (54 Mbit/s théoriques, 30 Mbit/s réels) sur la bande de fréquence des 2,4 GHz. La norme 802.11g a une compatibilité descendante avec la norme 802.11b, ce qui signifie que des matériels conformes à la norme 802.11g peuvent fonctionner en 802.11b. Cette aptitude permet aux nouveaux équipements de proposer le 802.11g tout en restant compatible avec les réseaux existants qui sont souvent encore en 802.11b.

Norme	Nom	Description
802.11h		La norme 802.11h vise à rapprocher la norme 802.11 du standard Européen (HiperLAN 2, d'où le h de 802.11h) et être en conformité avec la réglementation européenne en matière de fréquence et d'économie d'énergie.
802.11i		La norme 802.11i a pour but d'améliorer la sécurité des transmissions (gestion et distribution des clés, chiffrement et authentification). Cette norme s'appuie sur l'AES (Advanced Encryption Standard) et propose un chiffrement des communications pour les transmissions utilisant les technologies 802.11a, 802.11b et 802.11g.

D. L'installation informatique du LTECG



La sécurité informatique

Sommaire :

- la sécurité informatique
- un abécédaire de la sécurité informatique



A. La sécurité informatique [52]

Il existe de nombreux risques en sécurité informatique. De plus, ceux-ci évoluent d'année en année.

Il importe de mesurer ces risques, non seulement en fonction de la probabilité ou de la fréquence de leurs survenance, mais aussi en mesurant leurs effets possibles. Ces effets, selon les circonstances et le moment où ils se manifestent, peuvent avoir des conséquences négligeables ou catastrophiques. Parfois, le traitement informatique en cours échoue, il suffit de la relancer, éventuellement par une autre méthode si on craint que la cause en réapparaisse ; parfois l'incident est bloquant et on doit procéder à une réparation ou une correction avant de poursuivre le travail entrepris. Mais ces mêmes incidents peuvent avoir des conséquences beaucoup plus fâcheuses :

- données irrémédiablement perdues ou altérées ce qui les rend inexploitable,
- données ou traitements durablement indisponibles, pouvant entraîner l'arrêt d'une production ou d'un service,
- divulgation accidentelle d'informations confidentielles ou erronées pouvant profiter à des sociétés concurrentes ou nuire à l'image de l'entreprise,
- déclenchement d'actions pouvant provoquer des accidents physiques ou induire des drames humains,
- etc.

À l'ère de la généralisation des traitements et des échanges en masse, on imagine par exemple assez bien l'impact que pourraient avoir des événements majeurs comme, par exemple, une panne électrique de grande ampleur ou la saturation du réseau [Internet](#) pendant plusieurs heures.

Hormis ces cas exceptionnels, beaucoup de risques peuvent être anticipés et il existe des parades pour la plupart d'entre eux. On peut citer en exemple les précautions prises peu avant l'an 2000 qui, même si la réalité du risque a parfois été et reste aujourd'hui controversée, ont peut-être évité de graves désagréments.

Chaque organisation, mais aussi chaque utilisateur particulier, a tout intérêt à évaluer, même grossièrement, les risques qu'il encourt et les protections raisonnables qu'elle ou il peut mettre en œuvre. Dans le monde professionnel, les risques et les moyens de prévention sont essentiellement évalués en raison de leurs coûts. Il est par exemple évident qu'une panne qui aurait pour conséquence l'arrêt de la production d'une usine pendant une journée mérite qu'on consacre pour la prévenir une somme égale, justement, à une fraction de la valeur de sa production quotidienne ; cette fraction sera d'autant plus importante que la probabilité et la fréquence d'une telle panne sont élevées.

Risques humains

Les risques humains sont les plus importants, même s'ils sont le plus souvent ignorés ou minimisés. Ils concernent les utilisateurs mais également les informaticiens eux-mêmes.

- La maladresse : comme en toute activité, les humains commettent des erreurs ; il leur arrive donc assez fréquemment d'exécuter un traitement non souhaité, d'effacer involontairement des données ou des programmes ...
- L'inconscience et l'ignorance : de nombreux utilisateurs d'outils informatiques sont encore inconscients ou ignorants des risques qu'ils encourrent aux systèmes qu'ils utilisent, et introduisent souvent des programmes malveillants sans le savoir. Des manipulations inconsidérées (autant avec des logiciels que physiques) sont aussi courantes.

- La malveillance : Aujourd'hui, il serait quasiment inconcevable de prétexter l'ignorance des problèmes sus-cités, tant les médias ont pu parler des différents problèmes de [virus](#) et de [vers](#) ces dernières années (même s'ils ont tendance, en vulgarisant, à se tromper sur les causes et les problèmes). Ainsi, certains utilisateurs, pour des raisons très diverses, peuvent volontairement mettre en péril le système d'information, en y introduisant en connaissance de cause des virus (en connectant par exemple un ordinateur portable sur un réseau d'entreprise), ou en faisant exprès d'introduire de mauvaises informations dans une base de données. De même il est relativement aisé pour un informaticien d'ajouter délibérément des fonctions cachées leur permettant, directement ou avec l'aide de complices, de détourner à leur profit de l'information et particulièrement de l'argent.
- L'ingénierie sociale : l'ingénierie sociale (social engineering en anglais) est une méthode pour obtenir d'un système informatique des informations confidentielles, que l'on n'est pas normalement autorisé à obtenir, en vue de les exploiter à d'autres fins (publicitaires par exemple). Elle consiste à se faire passer pour quelqu'un que l'on est pas (en général un administrateur) et de demander des informations personnelles (nom de connexion, mot de passe, données confidentielles ...) en inventant un quelconque motif (problème dans le réseau, modification de celui-ci, heure tardive ...). Elle peut se faire soit au moyen d'une simple communication téléphonique, soit par mail, soit en se déplaçant directement sur place.
- L'espionnage : l'espionnage, notamment industriel, emploie les mêmes moyens, ainsi que bien d'autres, pour obtenir des informations sur des activités concurrentes, procédés de fabrication, projets en cours, futurs produits, politique de prix, clients et prospects, etc.

Risques techniques

Les risques techniques sont tout simplement ceux liés aux défauts et pannes inévitables que connaissent tous les systèmes matériels et logiciels. Ces incidents sont évidemment plus ou moins fréquents selon le soin apporté lors de la fabrication et des tests effectués avant que les ordinateurs et les programmes ne soient mis en service. Cependant les pannes ont parfois des causes indirectes, voire très indirectes, donc difficiles à prévoir.

- Incidents liés au matériel : si on peut le plus souvent négliger la probabilité d'une erreur d'exécution par un processeur (il y eut néanmoins une exception célèbre avec l'une des toutes premières générations du [processeur Pentium d'Intel](#) qui pouvait produire, dans certaines circonstances, des erreurs de calcul), la plupart des composants électroniques, produits en grandes séries, peuvent comporter des défauts et bien entendu finissent un jour ou l'autre par tomber en panne. Certaines de ces pannes sont assez difficiles à déceler car intermittentes ou rares.
- Incidents liés au logiciel : ils sont de très loin les plus fréquents ; la complexité croissante des systèmes d'exploitation et des programmes nécessite l'effort conjoint de dizaines, de centaines, voire de milliers de programmeurs. Individuellement ou collectivement, ils font inévitablement des erreurs que les meilleures méthodes de travail et les meilleurs outils de contrôle ou de test ne peuvent pas éliminer en totalité.
- Incidents liés à l'environnement : les machines électroniques et les réseaux de communication sont sensibles aux variations de température ou d'humidité (tout particulièrement en cas d'incendie ou d'inondation) ainsi qu'aux champs électriques et magnétiques. Il n'est pas rare que des ordinateurs connaissent des pannes définitives ou intermittentes à cause de conditions climatiques inhabituelles ou par l'influence d'installations électriques notamment industrielles (et parfois celle des ordinateurs eux-mêmes !).

Pour s'en prémunir, on recourt généralement à des moyens simples bien que parfois onéreux :

- Redondance des matériels : la probabilité ou la fréquence de pannes d'un équipement est représentée par un nombre très faible (compris entre 0 et 1, exprimé sous la forme 10^{-n}) ; en doublant ou en triplant cet équipement, on multiplie leurs probabilités ou fréquences de panne respectives, le résultat est donc un nombre beaucoup plus faible ; autrement dit l'ensemble est beaucoup plus fiable (ce qui le plus souvent reporte le risque principal ailleurs).
- Dispersion des sites : pour les mêmes raisons un accident environnemental a très peu de chance de se produire simultanément en deux endroits distants.
- Programmes ou procédures de contrôle indépendants : ils permettent bien souvent de déceler les anomalies avant qu'elles ne produisent des effets dévastateurs.

Programmes malveillants

Un *logiciel malveillant* (malware en anglais) est un logiciel développé dans le but de nuire à un système informatique. Voici les principaux types de programmes malveillants :

- Le *virus* : programme se dupliquant sur d'autres ordinateurs
- Le *ver* (worm en anglais) : exploite les ressources d'un ordinateur afin d'assurer sa reproduction
- Le *cheval de Troie* (trojan en anglais) : programme à apparence légitime (voulue) qui exécute des routines nuisibles sans l'autorisation de l'utilisateur
- La *porte dérobée* (backdoor en anglais) : ouvre d'un accès frauduleux sur un système informatique, à distance
- Le *logiciel espion* (spyware en anglais) : collecteur d'informations personnelles sur l'ordinateur d'un utilisateur sans son autorisation, et en envoyant celles-ci à un organisme tiers
- L'*exploit* : programme permettant d'exploiter une faille de sécurité d'un logiciel
- Le *rootkit* : logiciel invisible permettant soit d'obtenir les droits d'administrateur sur la machine infectée, soit truquer les informations système aux yeux d'un utilisateur

Techniques d'attaque par messagerie

En dehors des nombreux programmes malveillants qui se propagent par la messagerie électronique, il existe des attaques spécifiques à celle-ci :

- Le *pourriel* (spam en anglais) : un courrier électronique non sollicité, la plupart du temps de la publicité. Ils encombrant le réseau, et font perdre du temps à leurs destinataires
- L'*hameçonnage* (phishing en anglais) : un courrier électronique dont l'expéditeur se fait généralement passer pour un organisme financier et demandant au destinataire de fournir des informations confidentielles

Attaques sur le réseau

Voici les principales techniques d'attaques sur le réseau :

- Le *sniffing* : technique permettant de récupérer toutes les informations transitant sur un réseau (on utilise pour cela un logiciel sniffer). Elle est généralement utilisée pour récupérer les mots de passe des applications qui ne chiffrent pas leurs communications, et pour identifier les machines qui communiquent sur le réseau.
- La *mystification* (en Anglais spoofing) : technique consistant à prendre l'identité d'une autre personne ou d'une autre machine. Elle est généralement utilisée pour récupérer des informations sensibles, que l'on ne pourrait pas avoir autrement.
- Le *déni de service* : technique visant à générer des arrêts de service, et ainsi d'empêcher le bon fonctionnement d'un système.

Attaques sur les mots de passe

Les attaques sur les mots de passe peuvent consister à faire de nombreux essais jusqu'à trouver le bon mot de passe.

Dans ce cadre, notons les deux méthodes suivantes:

- L'attaque par dictionnaire : le mot testé est pris dans une liste prédéfinie contenant les mots de passe les plus courants et aussi des variantes de ceux-ci (à l'envers, avec un chiffre à la fin, etc.). Ces listes sont généralement dans toutes les langues les plus utilisées, contiennent des mots existants, ou des diminutifs (comme par exemple "powa" pour "power", ou "G0d" pour "god").
- Attaque par force brute : toutes les possibilités sont faites dans l'ordre jusqu'à trouver la bonne solution (par exemple de "aaaaaa" jusqu'à "ZZZZZZ" pour un mot de passe composé strictement de six caractères alphabétiques).

Techniques d'intrusion système

Voici quelques techniques d'intrusion utilisées fréquemment :

- Connexion à une ressource partagée
- Attaque par force brute
- Attaque par débordement de tampon
- Accès à un interpréteur de commande interactif
- Élévation des privilèges
- Installation d'un rootkit
- Effacement des traces

Autres types d'attaques

On peut noter qu'il existe d'autres types d'attaques, souvent moins connues :

- Le cassage de logiciel (Cracking en anglais) : cette technique a pour but la modification d'un programme pour déjouer sa protection (en général pour permettre une utilisation complète, ou à durée illimitée).

B. Un abécédaire sur la sécurité informatique

Antivirus [63]

Un antivirus est un logiciel censé protéger un micro-ordinateur contre les programmes néfastes appelés *virus*, *vers*, *macrovirus*, etc.

Les principaux antivirus du marché se fondent sur des fichiers de signatures et comparent alors la signature du virus aux codes à vérifier. Certains programmes appliquent également la méthode dite heuristique tendant à découvrir un *code malveillant* par son comportement. L'analyse de forme repose sur du filtrage basé sur des règles. Cette dernière méthode peut être très efficace pour les serveurs de courriels supportant les règles puisqu'elle ne repose pas sur un fichier de signatures.

Les antivirus peuvent examiner (scanner) le contenu d'un disque dur, mais également la mémoire de l'ordinateur. Pour les plus modernes, ils agissent en amont de la machine en scrutant les échanges de fichiers avec l'extérieur, aussi bien en flux montant que descendant. Ainsi, les courriels sont examinés, mais aussi les fichiers copiés sur ou à partir de supports amovibles tels que *cédéroms*, disquettes, connexions réseau...

Botnet [84]

Les botnets ou ordinateurs « zombies » forment des réseaux de PC infectés par des *virus informatiques* ou par des *chevaux de Troie*, contrôlés via *Internet* le plus souvent à des fins malveillantes.

En plus de servir à paralyser le trafic (attaque par *déni de service*), de moteur à la diffusion de *spam*, les botnets peuvent également être utilisés pour commettre des délits comme le vol de données bancaires et identitaires à grande échelle. Les botnets sont parfois loués à des tiers peu scrupuleux.

Selon *Symantec* fin 2004, le parc le plus important de PC contaminés se trouve au Royaume-Uni (avec 25,2 % des machines). Les États-Unis décroche la deuxième place avec 24,6 %. Ils sont suivis respectivement par la Chine (7,8 %), le Canada (4,9 %) et l'Espagne (3,8 %). La France est au sixième rang avec 3,6 % d'ordinateurs victimes d'une prise de contrôle sauvage à distance.

Canular informatique [55]

Les canulars sont appelés hoax en anglais et sont souvent utilisés pour le *pourriel* ou de simples lettre-chaînes. Dans ce dernier cas, internet ne fait qu'amplifier un phénomène qui existait déjà à travers le courrier

Les canulars sont souvent bâtis sur les mêmes modèles que les légendes urbaines. Dans ce cas ils en exploitent les caractéristiques de diffusion par colportage, ce qui renforce à la fois leur impact et leur audience.

Très souvent on peut s'apercevoir que :

- les fausses alertes au virus qui circulent par courriel sont destinées à faire paniquer les utilisateurs novices, parfois à leur faire commettre des manipulations dangereuses de leur système informatique et souvent à congestionner le réseau par leur diffusion hors de tout contrôle ;

- on essaie de vous prendre par les sentiments de manière assez grossière (sauvez Brian !)
- les faits relatés sont généralement très flous (au Brésil, par exemple, sans plus de détail),
- les références sont généralement inexistantes ou au contraire trop énormes (le Pentagone, Microsoft...).
- parfois, on vous fait des promesses disproportionnées (devenir milliardaire vite et aisément, gagner un bateau...)
- parfois on vous assure à maintes reprises que ce n'est pas un canular.
- on vous demande de renvoyer le message à toutes vos connaissances, ou à une adresse de courrier électronique bien précise

Enfin une variante appelée le virhoax associe le *virus* et le hoax. Elle profite de la crédulité du destinataire, le pousse à effacer un fichier de son ordinateur, en lui faisant penser que c'est un virus.

Le site www.hoaxbuster.com contient une base de données à jour de tous les canulars en circulation !

Certificat électronique [82]

Tel qu'on l'utilise en cryptographie et en sécurité informatique, un certificat électronique ou certificat numérique est un bloc de données contenant, dans un format spécifié, les parties suivantes :

- la partie publique d'une paire de *clés asymétriques*,
- des informations sur le porteur de cette paire de clés, telles que son nom, son adresse électronique, son titre, son numéro de téléphone, le nom de l'entité qui a délivré ce certificat, etc.
- et enfin la signature numérique des données ci-dessus par une personne ou entité prenant en charge la création de ce certificat et ayant autorité de certification.

Applications

Une des applications majeures des certificats est la signature numérique et le chiffrement des données, au travers des Infrastructures à Clés Publiques.

Les certificats respectent des standards spécifiant leur contenu de façon rigoureuse. On trouve parmi les plus connus et les plus utilisés : le standard X.509 en version 1, 2, et 3, sur lequel se fondent les infrastructures à clés publiques.

Cheval de Troie [58]

En informatique, un cheval de Troie est un logiciel d'apparence légitime, mais conçu pour subrepticement exécuter des actions nuisibles à l'utilisateur.

Un cheval de Troie n'est pas un *virus informatique* dans le sens où il ne se duplique pas par lui-même, fonction essentielle pour qu'un logiciel puisse être considéré comme un virus. Un cheval de Troie est conçu pour être dupliqué par des utilisateurs naïfs, attirés par les fonctionnalités vantées.

Les chevaux de Troie servent très fréquemment à introduire une porte dérobée sur un ordinateur. L'action nuisible à l'utilisateur est alors le fait qu'un pirate informatique peut à tout moment prendre à distance (par *Internet*) le contrôle de l'ordinateur.

Pour éviter les infections de chevaux de Troie, la règle la plus simple est d'installer un minimum de logiciels, de provenance sûre. Après infection, on peut détecter un cheval de Troie avec un [antivirus](#) à jour. Enfin, on peut utiliser un [pare-feu](#) pour limiter et surveiller les connexions au réseau que pourrait utiliser le [pirate](#). Toutefois, une fois installé, le cheval de Troie peut désactiver les antivirus et pare-feu existants.

Il est difficile voire impossible de définir exactement ce qu'est un cheval de Troie, car la légitimité d'un logiciel dépend aussi du contexte dans lequel il est employé. Les [portes dérobées](#) par exemple peuvent s'avérer utiles pour un administrateur réseau ; en revanche dans les mains d'un pirate elles sont clairement illégitimes.

Cookie [81]

En informatique, un cookie est défini par le protocole de communication [HTTP](#) comme étant une suite d'informations envoyée par un [serveur HTTP](#) à un client HTTP, que ce dernier retourne lors de chaque interrogation du même serveur HTTP.

Étymologie

Cookie provient d'un mot anglais signifiant « biscuit ». Les francisations témoin (OLF, 1996) et témoin de connexion (avis officiel français, 1999) ont été proposés mais ne sont pas aussi connus que cookie en 2005. Le terme mouchard est parfois utilisé avec une connotation négative, à cause des controverses sur l'usage des cookies.

Usage

Les cookies permettent de donner un état au protocole de communication HTTP, donc de définir des transactions. Ils ont été inventés par Lou Montulli alors qu'il travaillait chez Netscape Communications, dans le but d'adapter dynamiquement le contenu des sites Web aux habitudes de navigation de l'internaute.

Cracker [69]

Un cracker est une sorte de [pirate informatique](#) spécialisé dans le cassage des sécurités des logiciels, notamment les partagiciels (shareware en anglais), qui nécessitent généralement des clés d'enregistrement. On trouve parfois le terme de craqueur, casseur et déplombeur.

Controverse sur l'usage du terme

Pour certaines personnes, cracker et pirate informatique sont synonymes.

Le cracker est généralement – à tort – considéré comme un hacker nuisible. Cette assertion est fausse car :

- Le cracker n'est pas un [hacker](#).
- Le cracker n'agit pas sur [Internet](#).

Toutefois il peut arriver qu'un ancien cracker devienne hacker. En tant que cracker il acquiert une partie des connaissances techniques nécessaires à un hacker. De plus l'utilisation d'Internet peut-être un outil puissant pour un cracker. L'amalgame fait par certain médias entre crackers et hackers est vigoureusement combattu par les hackers.

Les moyens techniques utilisés par le cracker

Le cracker a de bonnes, voire de grandes, connaissances en assembleur (langage machine) qui lui permettent de comprendre le code envoyé au processeur et par conséquent le fonctionnement de la partie du programme observé. Un cracker avec un bon niveau écrira aussi ses propres programmes pour s'en servir comme outils dans son activité. Ces outils peuvent être génériques ou spécifiques au programme à cracker, généralement, ce sont les patches, des programmes qui « crackent » automatiquement un programme.

Les créateurs de logiciel utilisent diverses techniques pour empêcher les cracks. Dans la pratique il n'y a pas de techniques miracles, l'éditeur de logiciel peut seulement rendre plus difficile le travail du cracker, ce qui est presque impossible, car il faut bien que le processeur interprète le code, et si le processeur le peut, le crackeur également. Des logiciels permettant de compresser (UPX par exemple) ou de chiffrer les programmes peuvent décourager les amateurs mais ne feront que retarder les spécialistes.

Cependant le cracking ne se limite pas à l'attaque ni à l'étude de logiciels, mais plus souvent aux codes secrets, c'est-à-dire à la cryptographie, domaine dans lequel leurs connaissances mathématiques et informatiques font d'eux d'excellents casseurs de code.

À quoi aboutit-il ?

À un crack, par exemple une clef d'enregistrement, un générateur de clef ou une version modifiée du programme cracké dont certaines restrictions d'utilisation sont supprimées, on parle alors de patch.

Les modifications de comportement d'un programme les plus prisées sont :

- la suppression d'une vérification de licence, dans ce cas la motivation du cracker est la même que lorsque il crée une clef d'enregistrement ;
- la tricherie dans les jeux vidéo :
 - permettre dans un jeu de passer une étape difficile par divers moyens. Actuellement la plupart des jeux vidéo intègrent cette possibilité, les éditeurs considèrent que la prolifération de telles techniques exprime un besoin réel des utilisateurs ;
 - donner à un joueur un avantage dans les jeux multi-utilisateurs ; les créateurs de logiciels et beaucoup d'utilisateurs combattent vigoureusement cette utilisation des cracks. Certains des utilisateurs d'un jeu peuvent parfois, bien que rarement, utiliser de tels patchs, afin de rendre leur partie plus intéressante. Le jeu est alors détourné : l'objectif passe de savoir qui est le meilleur à ce jeu à qui est capable de trouver le meilleur patch pour tricher. Tant qu'une partie classique ne devient pas le terrain de jeux de crackers, on ne peut pas considérer cela comme de la tricherie car l'objectif est clairement annoncé.

Les motivations du cracker

Les motivations des crackers sont assez difficile à cerner, deux discours sont entendus suivant de quel côté de la barrière on se place :

- les crackers ont tendance à mettre l'emphasis sur la difficulté technique et la compétition qui s'instaure entre crackers ou entre crackers et créateurs de logiciel. Ils considèrent que quelqu'un qui utilise un de ces moyens pour utiliser un logiciel ne l'aurait de toute façon pas acheté donc qu'il n'y a pas de perte financière pour les éditeurs de logiciels, ceci est aussi le point de vue des utilisateurs de crack. Dans cette optique le crack n'est vu que comme un défi technique n'ayant pas d'impact économique (la « perte » se calculant sur les ventes théoriques du logiciel), et cette « course à l'armement » technique renforce la sécurité des logiciels du marché, qui deviennent d'autant plus difficiles à casser.
- les créateurs et éditeurs de logiciel eux considèrent le point de vue purement économique, un crack utilisé par 50 000 personnes sur un logiciel valant 100 € leur coûte 5 000 000 €

On ne peut pas nier non plus l'importance de la tricherie dans les jeux vidéo. Son utilisation a un impact économique si on considère que l'utilisation massive de tricherie diminue l'intérêt d'un jeu.

Il est difficile de considérer la motivation première des crackers comme étant d'ordre économique bien qu'une fois devenus techniquement compétents, l'aspect économique devient prépondérant.

Cryptographie asymétrique ou cryptographie à clé publique [83]

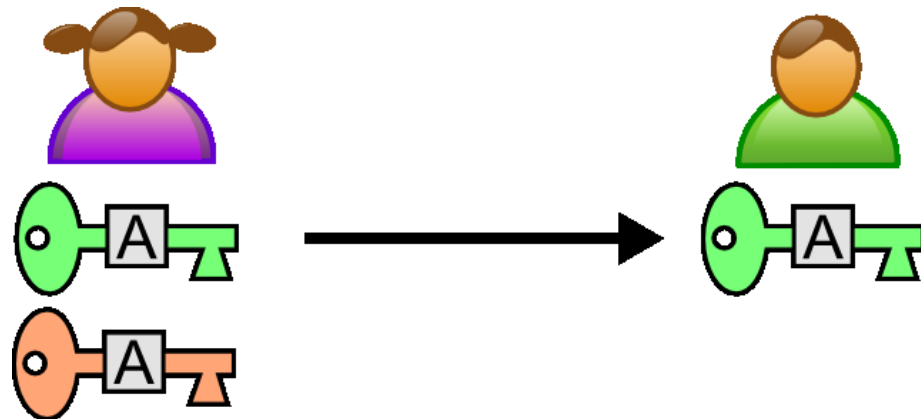
Principe

La cryptographie asymétrique, ou cryptographie à clé publique est fondée sur l'existence de fonctions à sens unique — c'est-à-dire qu'il est simple d'appliquer cette fonction à un message, mais extrêmement difficile de retrouver ce message à partir du moment où on l'a transformé.

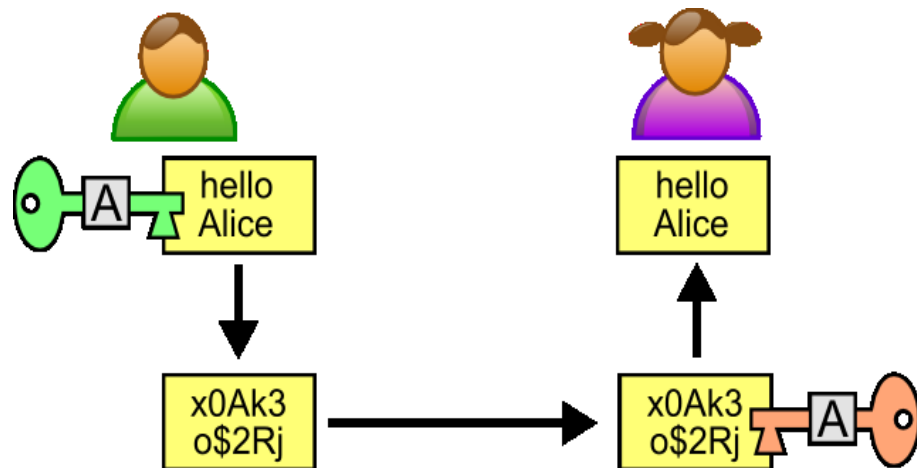
En réalité, on utilise en cryptographie asymétrique des fonctions à sens unique et à brèche secrète. Une telle fonction est difficile à inverser, à moins de posséder une information particulière, tenue secrète, nommée clé privée.

À partir d'une telle fonction, voici comment se déroulent les choses : Alice souhaite pouvoir recevoir des messages chiffrés de n'importe qui. Elle génère alors une valeur à partir d'une fonction à sens unique et à brèche secrète à l'aide d'un algorithme de chiffrement asymétrique, par exemple RSA.

Elle la diffuse, mais garde secrète l'information permettant d'inverser cette fonction. On parle de clé publique pour celle qu'on diffuse (sans avoir à se préoccuper de sa sécurité) et de clé privée pour l'information secrète (qui doit rester la propriété exclusive d'Alice).



1^{re} étape : Alice génère deux clés. La clé publique (verte) qu'elle envoie à Bob et la clé privée (rouge) qu'elle conserve précieusement sans la divulguer à quiconque.



2e et 3e étapes : Bob chiffre le message avec la clé publique d'Alice et envoie le texte chiffré.
Alice déchiffre le message grâce à sa clé privée..

Chiffrement

Un des rôles de la clé publique est de permettre le chiffrement ; c'est donc cette clé qu'utilisera Bob pour envoyer des messages chiffrés à Alice. L'autre clé – l'information secrète – sert à déchiffrer. Ainsi, Alice, et elle seule, peut prendre connaissance des messages de Bob, à condition que la brèche ne soit pas trouvée.

Analogie

Alice a choisi un coffre-fort. Elle l'envoie ouvert à Bernard, et en garde la clé. Lorsque Bernard veut écrire à Alice, il y dépose son message, ferme le coffre, et le renvoie à Alice. À sa réception, seule Alice peut ouvrir le coffre, puisqu'elle seule en possède la clé, à supposer le coffre inviolable, et que personne ne puisse retrouver la clé.

Authentification

D'autre part, l'utilisation par Alice de sa clé privée sur le condensat d'un message, permettra à Bernard de vérifier, en appliquant au résultat la clé publique d'Alice (et en retrouvant donc ce condensat) que nulle autre qu'Alice a pu signer ce message. C'est ainsi que Bernard sera rassuré sur l'origine du message reçu : il appartient bien à Alice.

Transmission sécurisée de la clé symétrique

La cryptographie asymétrique répond à un besoin majeur de la cryptographie symétrique : le partage sécurisé d'une clé entre deux correspondants, afin de prévenir l'interception de cette clé par une personne tierce non autorisée, et donc la lecture des données chiffrées sans autorisation.

Les mécanismes de chiffrement symétrique étant moins coûteux en temps de calcul, ceux-ci sont privilégiés aux mécanismes de chiffrement asymétrique. Cependant toute utilisation de clé de chiffrement symétrique nécessite que les deux correspondants se partagent cette clé, c'est-à-dire la connaissent avant l'échange. Ceci peut être un problème si la communication de cette clé s'effectue par l'intermédiaire d'un medium non sécurisé, « en clair ». Afin de pallier cet inconvénient, on utilise un mécanisme de chiffrement asymétrique pour la seule phase d'échange de la clé symétrique, et l'on utilise cette dernière pour tout le reste de l'échange.

Mécanismes d'authentification

Un inconvénient majeur de l'utilisation des mécanismes de chiffrement asymétriques est le fait que la clé publique est distribuée à toutes les personnes, Bob, Carole... souhaitant échanger des données de façon confidentielle, et donc lorsque la personne possédant la clé privée, Alice, déchiffre les données chiffrées, celle-ci n'a aucun moyen de vérifier avec certitude la provenance de ces données (Bob, ou Carole...). On parle ici de problèmes d'authentification. Afin de résoudre ce problème, on utilise des mécanismes

d'authentification permettant de garantir la provenance des informations chiffrées. Ces mécanismes sont fondés sur le chiffrement asymétrique.

**Principe
d'authentification par chiffrement
asymétrique :**

Objectif : Bob souhaite envoyer des données chiffrées à Alice en garantissant la provenance de celles-ci (authentification des messages).

1. Alice crée une paire de clés asymétriques : clé privée KprA, clé publique KpuA
2. Alice envoie sa clé publique à Bob pour que celui-ci puisse lui envoyer des données chiffrées par la suite
3. Lorsque Bob va envoyer des informations chiffrées à Alice, celui-ci désire signer numériquement ces informations afin de garantir à Alice que celles-ci proviennent effectivement de lui :
 1. Bob doit donc, avant d'envoyer ces données chiffrées, créer une paire de clés asymétriques : clé publique KpuB, clé privée KprB
 2. Bob envoie sa clé publique KpuB à Alice
 3. Bob chiffre son message avec sa clé privée (signature numérique du message par chiffrement) : KprB, puis chiffre une seconde fois le message précédent avec la clé publique d'Alice (chiffrement réel du message) : KpuA.
 4. Alice reçoit le message chiffré de Bob, le dé-chiffre avec sa clé privée : KprA. À ce stade le message n'est pas encore « lisible » car celui-ci a été chiffré deux fois de suite.
 5. Alice déchiffre une seconde fois le message avec la clé publique de Bob : KpuB

Si le message reçu et déchiffré par Alice provient effectivement de Bob, le message déchiffré deux fois sera un message « lisible », signe que l'expéditeur du message est bien Bob. En effet, seule la clé privée de Bob qui est unique : KprB, chiffre le message de telle façon qu'une fois dé-chiffré par Alice à l'aide de la clé publique de Bob : KpuB, celui devient lisible. Dans le cas contraire le message ne sera pas « lisible » car il n'aura pas été chiffré avec la clé privée de Bob et cependant dé-chiffré avec la clé publique de Bob (possédée par Alice). On en conclut que Bob n'est pas l'expéditeur du message.

On remarque ici, que la clé publique a un rôle double puisqu'elle permet de chiffrer des données (lors du chiffrement conventionnel) mais aussi de déchiffrer des données (lors d'une authentification).

Certificats

La cryptographie asymétrique est également utilisée avec les [certificats numériques](#), celui-ci contenant la clé publique de l'entité associée au certificat. La clé privée est quant à elle stockée au niveau de cette dernière entité. Une application des certificats est par exemple la mise en œuvre d'une infrastructure à clés publiques PKI pour gérer l'authentification et la signature numérique d'une entité, par exemple un [serveur web](#) (Apache avec le module [SSL](#) par exemple), ou simplement un client souhaitant signer et chiffrer des informations à l'aide de son certificat de la façon décrite dans les sections précédentes.

**Une clé privée
inviolable ?**

Un chiffrement symétrique au moyen d'une clé de 128 bits propose 2^{128} (un nombre à trente-huit chiffres) façons de chiffrer un message. Un pirate qui essaierait de décrypter le message par la force brute devrait donc les essayer une par une.

Pour les systèmes à clé publique, il en va autrement. Tout d'abord les clés sont plus longues (par exemple 1024 bits minimum pour RSA) ; ceci est dû au fait qu'elles possèdent une structure mathématique très particulière (on ne peut choisir une suite de bits aléatoire comme en clé secrète.) Ensuite, il y a clairement mieux à faire qu'une recherche exhaustive

sur, par exemple, 1024 bits, à savoir exploiter la structure mathématique de la clé (pour RSA, cela mène à la factorisation.)

Il faut noter le développement actuel de la cryptographie utilisant les courbes elliptiques, qui permettent (au prix d'une théorie et d'implémentations plus complexes) l'utilisation de clés nettement plus petites que celles des algorithmes classiques (une taille de 160 bits étant considérée comme très sûre actuellement), pour un niveau de sécurité équivalent.

Déni de service [68]

Fonctionnement général

D'une manière générale, l'attaque par déni de service ou Denial of Service (DoS) vise à rendre une application informatique incapable de répondre aux requêtes de ses utilisateurs.

Une machine serveur offrant des services à ses clients (par exemple un serveur web) doit traiter des requêtes provenant de plusieurs clients. Lorsque ces derniers ne peuvent en bénéficier pour des raisons délibérément provoquées par un tiers il y a déni de service.

Fonctionnement détaillé

Une *machine serveur* offrant des services à ses clients (un *serveur web* par exemple) doit traiter des requêtes provenant de plusieurs clients, et ce suivant un débit de traitement dépendant de la configuration matérielle et logicielle de celle-ci. La durée de traitement des requêtes clientes est supérieure à la durée de libération des ressources créées pour le traitement de celles-ci. Ainsi, si le débit des requêtes que doit traiter cette machine est supérieur à la vitesse de libération des ressources de celle-ci pendant une période trop prolongée, la machine sature et est rendue inactive puisqu'elle ne peut plus répondre aux demandes de services de ses clients : on parle de déni de services.

Les différents types d'attaques

Il existe un grand nombre de type d'attaques par déni de service (le simple fait de débrancher la prise d'un serveur peut être qualifiée d'attaque par déni de service).

DDoS

En particulier, il existe le Distributed Denial of Service:

Une attaque déni distribué de service ou Distributed Denial of Service (DDoS), est une attaque contre un ordinateur, situé en réseau par un ensemble de machine de façon simultanée. À la différence des attaques de type DoS où l'attaquant est constitué d'une unique machine, plusieurs machines sont utilisées pour l'attaque et donc celle-ci est plus dévastatrice et la saturation de la machine serveur est plus rapide. Dans ce type d'attaque des machines rebonds sont souvent utilisées à l'insu des propriétaires de celles-ci, permettant d'augmenter le nombre de machines hostiles participant à l'attaque et de se dissimuler derrière ces machines pour le pirate. Un ensemble de machine-rebonds, également appelé *botnet*, est contrôlable par un *pirate* après infection de chacune d'elles par un programme de type *porte dérobée*.

Responsables de ces attaques

- Ces attaques sont souvent utilisées par les *lamers*, les script kiddies (En informatique, script kiddie est un terme péjoratif pour désigner un pirate qui utilise des programmes qu'il n'a pas créés lui-même. Les script kiddies, bien que ne possédant pas de compétences informatiques particulières, peuvent causer des dégâts assez importants) ou les concepteurs de *virus*. Ils peuvent transiter par des *botnets*.
- Ce genre d'attaque est également très utile à un pirate qui cherche à prendre le contrôle d'une machine en utilisant l'*IP spoofing*. En effet en cas de demande de connexion avec une adresse *IP* « spoofée », le véritable propriétaire de l'adresse IP (auquel le *serveur* fait confiance) recevrait le paquet émis par le serveur, il réinitialiserait alors automatiquement la connexion et le *pirate* perdrait systématiquement la possibilité de créer une connexion unidirectionnelle avec le serveur.

Exemples d'attaques DoS et DDoS

Attaque sur le serveur de mise à jour de Microsoft

Attaque contre des sites internet connus tel que Google, Microsoft, Apple.

Attaques de type d'octobre 2002 sur les serveurs racines [DNS](#)

...

Comment éviter ces attaques

- Les attaques de type DoS peuvent être évitées en repérant l'adresse de la machine hostile, dans le cas d'une attaque à distance, et de bannir celle-ci. Les prochaines demandes de services provenant de cette machine seront donc rejetées directement sans être traitées.
- Les attaques de type DDoS sont beaucoup plus difficiles à éviter, on peut seulement limiter leur effet dévastateur. Ceci peut se faire en repérant les machines hostiles effectuant trop de requêtes en un temps limité (comportement client anormal), et en bannissant ces machines au fur et à mesure de leur détection. Cependant une attaque massive et rapide ne sera sans doute pas arrêtée avec cette méthode. Une architecture répartie, composée de plusieurs machines serveurs offrant le même type de services permet de répartir les points d'accès pour accéder aux services voulus et offre des solutions alternatives si une de ces machines devait être rendue inactive suite à une attaque.

Exploit [61]

Dans le domaine de la sécurité informatique, un exploit est un programme permettant à un individu d'exploiter une faille de sécurité dans un [système d'exploitation](#), un programme, et ceci à distance (remote exploit) ou sur la machine sur laquelle cet exploit est exécuté (local exploit). Une Proof of concept (preuve d'un concept ou POC) est la publication d'une faille logiciel afin de démontrer son existence et de favoriser le processus de correction de celle-ci.

Les failles logicielles sont présentes dans la plupart des [logiciels](#) que nous utilisons. Souvent découvertes par des laboratoires ou chercheurs indépendants elles sont, suivant les intentions de la personne, publiées sous forme de proof of concept afin de démontrer son existence et favoriser le Mais les POC sont bien souvent déclinés en exploits permettant ainsi l'exécution de code arbitraire.

Firewall Cf. [Pare-feu](#)

Hacker [70]

Le hacker (bidouilleur en anglais) est un expert dans son domaine. Ce terme est souvent employé en informatique, et la plupart du temps à mauvais escient : un hacker n'est pas un pirate informatique.

Domaine informatique

Dans le jargon des programmeurs, le mot hacker est utilisé pour désigner une personne maîtrisant totalement l'art de la programmation et la connaissance détaillée des systèmes, sans être forcément orientée vers la sécurité. On utilise d'ailleurs le terme de kernel hackers pour ceux qui sont passés maîtres dans l'art de coder ou recoder le noyau de Linux. Ainsi sont considérés comme de véritables hackers des gens comme Linus Torvalds (le père de Linux), Larry Wall (le père de Perl). [Eric Steven Raymond](#) définit parfaitement dans ses ouvrages et écrits le terme de hacker. Il se présente comme hacker et avocat de la communauté des hackers.

Dérive linguistique

Certains hackers spécialistes dans le domaine de la sécurité informatique, comme [Kevin Mitnick](#), sont des pirates informatiques.

Il faut se garder de perpétuer l'erreur souvent répétée par les médias (écrits, télévisés, radiophoniques...) d'amalgamer les termes hacker avec *cracker* ou *pirate informatique*.

IP spoofing [67]

L'IP spoofing est une technique de hacking consistant à utiliser l'*adresse IP* d'une machine, ou d'un équipement, afin d'en usurper l'identité. Elle permet de récupérer l'accès à des informations en se faisant passer pour la machine dont on spoofe l'adresse IP. De manière plus précise, cette technique permet la création de paquets IP avec une adresse IP source appartenant à quelqu'un d'autre.

L'en-tête de chaque paquet IP contient son adresse source ; cette adresse est normalement celle d'où le paquet provient. Toutefois, cet en-tête peut être modifié pour contenir une adresse IP différente. Cette technique peut ainsi servir à attaquer des réseaux en se faisant passer pour quelqu'un d'autre.

En effet certains services peu sécurisés se basent sur l'adresse IP pour identifier l'émetteur. L'exemple typique est d'utiliser une relation de confiance. Un *pirate* utilisera donc l'adresse IP d'une machine de confiance (autorisée) pour obtenir une connexion à un *serveur*.

Pour éviter ce genre d'attaques, il ne faut pas utiliser de service se basant sur l'adresse IP pour identifier les clients. L'IP spoofing peut également être utilisé pour causer des *dénis de services* et lors d'attaque par rebond.

Lamer [85]

L'expression Lamer est dérivé de l'anglais lame, qui est traduit en français par boiteux, faible ou piètre. L'utilisation du terme de Lamer se retrouve dans presque tous les domaines du Net qu'il s'agisse de jeux vidéo, de hacking ou de sites web...

Mais le terme Lamer est plus courant dans le monde du hacking. Il s'agit alors de pirates qui n'ont généralement presque aucun savoir dans le domaine du hacking, mais se pavanent en réalisant des *exploits* très faciles à reproduire. Aux yeux des *hackers* véritables, ils sont des « rigolos » ou des amateurs un peu vantards.

Dans le domaine des jeux vidéo, le Lamer est un tricheur (« cheater » en anglais) qui ne sait même pas tricher discrètement.

Le sous-entendu général, quand on utilise le terme, est que la personne ainsi désignée ne connaît rien au domaine où elle affirme avoir une expertise, mais qu'au contraire elle est désespérément à la recherche d'une reconnaissance sociale qui lui manque.

On pourra remarquer que l'attitude de ceux qui utilisent le terme relève souvent du simple mécanisme social d'autoprotection d'une élite ou d'un groupe fermé, en assurant une catégorisation de ses marges pour renforcer la séparation entre l'intérieur du groupe (par exemple, les hackers) et son extérieur.

Logiciel espion (spyware) [60]

Un logiciel espion est un *logiciel malveillant* qui infecte un ordinateur dans le but de collecter et de transmettre à des tiers, des informations de l'environnement sur lequel il est installé sans que l'utilisateur n'en ai conscience.

Étymologie	Logiciel espion est tiré de l'Anglais spyware et se dit également espiogiciel ou plus rarement espioniciel.
Utilisation	Il est permis de penser que les logiciels espions sont développés principalement par des sociétés proposant de la publicité sur Internet . En effet, pour permettre l'envoi de publicité ciblé, il est nécessaire de bien connaître sa cible. Cette connaissance peut être facilement obtenue par des techniques de profilage dont le logiciel espion fait partie.
Fonctionnement	<p>Un logiciel espion est composé de trois mécanismes distincts:</p> <ul style="list-style-type: none">• Le mécanisme d'infection. Par exemple, le spyware Cydoor utilise le logiciel grand public Kazaa.• Le mécanisme assurant la collecte d'information. Pour le même exemple, la collecte consiste à enregistrer tous ce que l'utilisateur recherche et télécharge via le logiciel Kazaa.• Le mécanisme assurant la transmission à un tiers. Ce mécanisme est généralement assuré via le réseau Internet. Le tiers peut être le concepteur du programme ou une entreprise. La transmission de données à un état est très spéculatifs et relève peut-être plus de la théorie du complot.
Où trouve-t-on des logiciels espions ?	<p>Le logiciel espion est très répandu sur les systèmes Microsoft, sur lesquels il n'est généralement pas détectable par les logiciels anti-virus ni les anti-spywares actuels, car fonctionnant avec le même principe qu'un rootkit.</p> <p>Ils sont souvent présents dans des graticiels (freeware, différents des logiciels libres), ou des partagiciels (shareware), afin de rentabiliser leur développement. En général les logiciels Open Source comme Mozilla FireFox n'en contiennent aucun.</p> <p>On en trouve également dans les logiciels d'installation de pilotes fournis avec certains périphériques et, plus généralement, tous les logiciels propriétaires peuvent en cacher puisque leurs sources ne sont pas accessibles.</p> <p>Enfin, certains administrateurs systèmes ou réseaux installent eux-mêmes ce type de logiciel pour surveiller à distance l'activité de leurs ordinateurs, sans avoir à se connecter dessus.</p>
Logiciels espions connus	Une très complète se trouve ici : www.spywarewarrior.com/rogue_anti-spyware.htm
Comment lutter contre les logiciels espions ?	<p>Il faut utiliser un logiciel anti-spyware. Il existe plusieurs logiciels spécifiques pour détecter et supprimer ces logiciels avec Microsoft Windows. Ces logiciels, comme les antivirus, utilisent des bases de données fréquemment mises à jour.</p> <p>On peut installer plusieurs de ces logiciels, car souvent ils ne détectent pas les mêmes espions. Même tous installés, ils ne détecteraient qu'une partie réduite des logiciels espions existants.</p>
Faire attention aux programmes utilisés	<p>Puisque certains types de programmes ont une fâcheuse tendance à inclure des logiciels espions, un moyen de lutter le plus efficacement est de réduire au maximum l'installation de logiciels dont les sources ne sont pas disponibles (à la différence d'un logiciel libre).</p> <p>Ainsi, même si vous n'êtes pas vous-même capable de comprendre les sources du programme, d'autres personnes compétentes et bien intentionnées le feront et lanceront l'alerte s'il y a un logiciel espion inclus dedans. Si les sources ne sont pas disponibles, alors uniquement l'auteur ou la société éditrice pourra vous le garantir, avec ses enjeux</p>

économiques actuels, et les lois en vigueur dans son pays ; ce qui n'est évidemment pas obligatoirement une garantie suffisante.

De plus, à l'instar de Gator, ou d'autres sociétés comme des constructeurs d'imprimantes, ont pignon sur rue, certaines se vantant même d'exercer ses activités douteuses. Elles jouent de surcroît sur un vide ou un flou juridique dans la loi de leur pays.

Contrôler les flux sortants

Le contrôle des flux sortants est la plupart du temps réalisé par l'administrateur réseau. Par l'intermédiaire d'un pare-feu, ce contrôle des flux sortants repose sur le principe de bloquer toute connexion qui tente de s'effectuer à partir de l'ordinateur (ou du réseau interne) vers l'extérieur (généralement *Internet*), sauf les connexions autorisées préalablement (on autorise généralement les connexions vers des sites internet, mais on n'autorise moins souvent le *Peer-to-peer*).

Même si le contrôle des flux sortants est encore peu mis en place à l'heure actuelle, il est primordial dans la compréhension et le blocage de certains problèmes, comme la présence de logiciels espions, car ils vont être amenés à se connecter à l'extérieur pour envoyer les informations qu'ils auront recueillies.

Logiciel malveillant [54]

Un logiciel malveillant (de l'anglais malware) est un logiciel développé dans le but de nuire à un système informatique. Les virus et les vers sont les deux exemples de logiciels malveillants les plus connus.

Ethymologie

Logiciel malveillant est une traduction de l'anglais malware qui est lui même un mot-valise, contraction de malicious (malicieux) et software (logiciel). Il s'agit donc littéralement d'un logiciel malicieux.

En France, l'usage du terme logiciel malveillant est préconisé par la commission générale de terminologie et de néologie.

Différentes classes de logiciels malveillants

Les *virus* et les *vers* sont deux grandes classes de logiciels malveillants. Leur caractéristique commune est qu'ils sont tous les deux capables de se répliquer eux-mêmes. Ils peuvent générer des copies d'eux-mêmes, parfois modifiées. Toutefois, tous les programmes qui se répliquent ne sont pas forcément des virus ou des vers. Par exemple, un logiciel peut se copier à d'autres médias en tant qu'élément de protection du système. Pour être classifié comme virus ou ver, il faut qu'au moins certaines de ces copies puissent se répliquer elles-mêmes aussi. La différence entre un virus et un ver est qu'un ver fonctionne plus ou moins indépendamment, tandis qu'un virus dépend des autres hôtes du réseau pour se propager.

- Les *virus*
- Les *vers* (worm). Ils se répandent dans le courrier électronique en profitant des failles des différents logiciels de messagerie (notamment Microsoft Outlook). Dès qu'ils ont infecté un ordinateur, ils s'envoient eux-mêmes dans tout le carnet d'adresses, ce qui fait que l'on reçoit ce virus de personnes connues. Certains d'entre eux ont connu une expansion fulgurante (I Love You ou Welcome to the matrix). Les experts n'arrivent pas à se mettre d'accord sur l'appartenance ou non des vers à la classe des virus informatiques.
- Les *chevaux de Troie* (trojan horses). Ce nom vient de la célèbre ruse imaginée par Ulysse. Ces programmes prétendent être légitimes (souvent de petits jeux ou utilitaires), mais comportent des routines nuisibles exécutées sans l'autorisation de l'utilisateur. On

confond souvent les chevaux de Troie avec les *backdoors*. Ces derniers sont en effet une catégorie de chevaux de Troie, mais pas la seule. Les backdoors prennent le contrôle de l'ordinateur et permettent à quelqu'un de l'extérieur de le contrôler par le biais d'Internet. Les chevaux de Troie ne sont pas des virus car il leur manque la fonction de reproduction, essentielle pour qu'un programme puisse être considéré comme un virus.

- Les *portes dérobées* (backdoor)
- Les *logiciels espion* (spyware). Ils peuvent accompagner certains gratuits (mais pas les *logiciels libres*), *partagiciels* et pilotes de périphériques, s'installant discrètement sur l'ordinateur, sans prévenir l'utilisateur, et collectant et envoyant des informations personnelles à des organismes tierces.
- Les *exploits*
- Les *rootkits*

Techniques de lutte contre les logiciels malveillants

Voici différentes techniques de lutte contre les logiciels malveillants, celles-ci pouvant être cumulées car n'agissant pas sur les mêmes risques :

- Utilisation d'un *antivirus* mis à jour quotidienne ou à chaque connexion sur Internet
- Utilisation d'un *pare-feu* (verrouillage des ports et des protocoles dont l'utilisation n'est pas requise, de l'ordinateur vers l'extérieur et de l'extérieur vers l'ordinateur).
- Utilisation d'un anti-spyware
- Utilisation d'un anti-rootkit
- Utilisation d'un HIDS (Host-based Intrusion Detection System), un système de détection d'intrusion sur l'hôte
- L'utilisation des outils plus sécurisés mais surtout moins monopolistes de navigation, de mails etc. tels que Mozilla Firefox, Opera, Thunderbird ou Foxmail augmentent l'efficacité des moyens antivirus.

Dans les entreprises, la diversification des systèmes d'exploitations est une solution complémentaire et raisonnable à la lutte antivirus comme l'introduction de systèmes logiciels ou matériels Unix, Linux ou propriétaires non standards dans la gestion des réseaux.

Pare-feu (firewall) [64]

En informatique, un pare-feu est un dispositif logiciel ou matériel qui filtre le flux de données sur un réseau informatique. Il est parfois appelé coupe-feu ou encore firewall.

Fonctionnement général

Le pare-feu est aujourd'hui considéré la pierre angulaire de la sécurité d'un réseau informatique. Il permet d'appliquer une politique d'accès aux ressources réseau (serveurs).

Le filtrage se fait selon divers critères. Les plus courants sont :

- l'origine ou la destination des paquets (*adresse IP*, ports TCP ou UDP, interface réseau, etc.)
- les options contenues dans les données (fragmentation, validité, etc.)
- les données elles-mêmes (taille, correspondance à un motif, etc.)
- les utilisateurs pour les plus récents

Un pare-feu fait souvent office de *routeur* et permet ainsi d'isoler le réseau en plusieurs zones de sécurité appelées *zones démilitarisées* ou *DMZ*. Ces zones sont séparées suivant le niveau de confiance qu'on leur porte.

Catégories de pare-feu	Les pare-feu sont le plus vieil équipement de sécurité et comme tel, ils ont été soumis à des nombreuses évolutions. Suivant la génération du pare-feu ou son rôle précis, on peut les classer en différentes catégories.
Pare-feu sans états (stateless firewall)	C'est le plus vieux dispositif de filtrage réseau, introduit sur les routeurs. Il regarde chaque paquet indépendamment des autres et le compare à une liste de règles. La configuration de ces dispositifs est souvent complexe et l'absence de prise en compte des machines à états des protocoles réseaux ne permet pas d'obtenir une finesse du filtrage très évoluée. Ces pare-feu ont donc tendance à tomber en désuétude mais restent présents sur certains routeurs ou systèmes d'exploitation.
Pare-feu à états (statefull firewall)	Certains protocoles dits « à états » comme TCP ou FTP introduisent une notion de connexion. Les pare-feu à états vérifient la conformité des paquets à une connexion en cours.
Pare-feu applicatif	Dernière génération de pare-feu, ils vérifient la complète conformité du paquet à un protocole attendu. Par exemple, ce type de pare-feu permet de vérifier que seul du HTTP passe par le port TCP 80.
Pare-feu authentifiant	Un pare-feu authentifiant réalise l'authentification des connexions passant à travers le filtre IP. L'administrateur peut ainsi définir les règles de filtrage par utilisateur et non plus par IP, et suivre l'activité réseau par utilisateur.
Pare-feu personnel	Les pare-feu personnels, généralement installés sur une machine de travail, agissent comme un pare-feu à états. Bien souvent, ils vérifient aussi quel programme est à l'origine des données. Le but est de lutter contre les virus informatiques et les logiciels espions .

Phishing [66]

En informatique, l'hameçonnage, ou phishing en anglais, est un terme désignant l'obtention d'informations confidentielles (comme les mots de passe ou d'autres informations privées), en se faisant passer auprès des victimes pour quelqu'un digne de confiance ayant un besoin légitime de l'information demandée. C'est une forme d'attaque informatique de type ingénierie sociale.

Deux origines possibles du terme phishing :

- Le terme phishing viendrait de la contraction de phone et fishing. Originellement le phishing c'est l'arnaque téléphonique qui consiste à se faire passer pour quelqu'un d'autre comme un policier ou un banquier pour tenter d'extirper des informations ou plus. La pratique a ensuite été adaptée au web.
- Le terme phishing aurait été inventé par les pirates qui essayaient de voler des comptes AOL. C'est un terme anglais qui serait construit sur l'expression password harvesting fishing, soit « pêche aux mots de passe » (?). Un attaquant se faisait passer pour un membre de l'équipe AOL et envoyait un message instantané à une victime potentielle. Ce message demandait à la victime d'indiquer son mot de passe, afin de, par exemple, « vérifier son compte AOL » ou pour « confirmer ses informations bancaires ». Une fois que la victime avait révélé son mot de passe, l'attaquant pouvait accéder au compte et l'utiliser à des fins malveillantes, comme l'envoi de pourriel.

Hameçonnage contemporain	De nos jours, les criminels informatiques utilisent l'hameçonnage à des fins plus axées sur le vol d'argent. Les cibles les plus populaires sont les services bancaires en ligne, et les sites de ventes aux enchères tels que eBay. Les adeptes de l'hameçonnage envoient habituellement des courriels à un grand nombre de victimes potentielles. Ces courriels
---------------------------------	---

redirigent les personnes qui les reçoivent vers une page Web qui semble faire partie de leur banque en ligne (par exemple), mais qui en réalité intercepte les informations confidentielles et les transmet au fraudeur.

Typiquement, les messages ainsi envoyés semblent émaner d'une société digne de confiance et sont formulés de manière à ne pas alarmer le destinataire, afin qu'il effectue une action en conséquence. Une approche souvent utilisée est d'indiquer à la victime que son compte a été désactivé à cause d'un problème quelconque, et que la réactivation ne sera possible qu'en cas d'action de sa part. Le message fournit alors un hyperlien qui dirige l'utilisateur vers une page Web qui ressemble à s'y méprendre au vrai site de la société digne de confiance. Arrivé sur cette page trompeuse, l'utilisateur est invité à saisir des informations confidentielles qui sont alors capturées par le criminel.

Parades

La vérification de l'URL dans la barre d'adresse du navigateur Web peut ne pas être suffisante pour détecter la supercherie, car certains navigateurs n'empêchent pas l'adresse affichée à cet endroit d'être contrefaite. Il est toutefois possible d'utiliser la boîte de dialogue « propriétés de la page » fournie par le navigateur pour découvrir la véritable adresse de la fausse page.

Une personne contactée au sujet d'un compte devant être « vérifié », doit chercher à régler le problème directement avec la société concernée, ou se rendre sur le site Web en tapant manuellement l'adresse dans son navigateur.

En règle générale, il est recommandé de faire suivre le message suspect à l'adresse spoof (par exemple, si le phishing concerne societe.com, ce sera spoof@societe.com), ce qui permettra à la société de faire une enquête.

Il faut être particulièrement vigilant lorsque l'on rencontre une adresse contenant le symbole « @ », par exemple <http://www.google.com@members.tripod.com/>. Ce genre d'adresse va essayer de connecter l'internaute en tant qu'utilisateur « www.google.com » sur le serveur « members.tripod.com ». Il y a de fortes chances que cela se réalise même si l'utilisateur indiqué n'existe pas réellement sur le serveur, mais par cette méthode la première partie de l'adresse semble être tout à fait innocente (www.google.com). De même, certains attaquants utilisent des adresses de sites contenant une faute de frappe, ou bien des sous-domaines, par exemple <http://www.mabanquefavorite.com.spamdomain.net/>. Des navigateurs récents, tels que Firefox, possèdent un système permettant d'avertir l'utilisateur du danger et de lui demander si il veut vraiment utiliser de telles adresses douteuses. Netscape 8 intègre également des technologies permettant de tenir à jour une liste noire de sites dangereux de ce type.

Les fraudes concernant les banques en ligne visent à obtenir l'identifiant et le mot de passe du titulaire d'un compte. Il est alors possible au fraudeur de se connecter sur le site web de la banque et d'effectuer des virements de fonds vers son propre compte. Pour parer à ce type de fraude, la plupart des sites bancaires en ligne n'autorisent plus l'internaute à saisir lui-même le compte destinataire du virement : il faut, en règle générale, téléphoner à un service de la banque qui reste seul habilité à saisir le compte destinataire dans une liste de comptes. La conversation téléphonique est souvent enregistrée et peut alors servir de preuve.

Pirate informatique [71]

Un pirate informatique est une personne commettant des actes considérés comme des délits ou des crimes dont l'objet ou l'arme est lié à l'informatique. Ce terme fait référence aux pirates du milieu maritime.

Le terme pirate a été choisi par la commission générale de terminologie et de néologie et définit une personne qui contourne ou détruit les protections d'un *logiciel*, d'un ordinateur ou d'un réseau informatique. Il a été choisi pour remplacer aussi le terme *cracker*.

Porte dérobée (Backdoor) [59]

Dans un *logiciel*, une porte dérobée (de l'anglais backdoor, littéralement porte arrière) est une fonctionnalité inconnue de l'utilisateur légitime, qui donne un accès secret au logiciel.

Une porte dérobée peut être introduite soit par le développeur du logiciel, soit par un tiers, typiquement un *pirate informatique*. La personne connaissant la porte dérobée peut l'utiliser pour surveiller les activités du logiciel, voire en prendre le contrôle. Enfin, selon l'étendue des droits que le système d'exploitation donne au logiciel contenant la porte dérobée, le contrôle peut s'étendre à l'ensemble des opérations de l'ordinateur.

La généralisation de la mise en réseau des ordinateurs rend les portes dérobées nettement plus utiles que du temps où un accès physique à l'ordinateur était la règle.

Parmi les motivations amenant les développeurs de logiciel à créer des portes dérobées, il y a :

- L'intérêt pratique d'un accès facile et toujours ouvert au logiciel pour pouvoir mener efficacement les actions de maintenance.
- La possibilité de désactiver subrepticement le logiciel en cas de désaccord avec son client (non paiement de licence).

Parmi les motivations amenant les pirates informatiques à installer une porte dérobée :

- 1. La possibilité de surveiller ce que fait l'utilisateur légitime et de copier ou détruire des données ayant une valeur (mots de passe, coordonnées bancaires, secrets commerciaux).
- 2. La possibilité de prendre le contrôle d'un ordinateur et de pouvoir l'utiliser pour mener des actions malveillantes (envoi de *spam* notamment pour le *phishing*, de *virus informatiques*, *déni de service*).
- 3. Le contrôle d'un vaste réseau d'ordinateurs (voir *botnet*), qui peut être utilisé pour du chantage au *déni de service* distribué (DDoS), ou revendu à des criminels.

Pour installer des portes dérobées en masse, les pirates informatiques utilisent des *virus*. Ceux-ci se répandent automatiquement et installent un serveur informatique sur chaque ordinateur infecté. Ensuite le pirate peut se connecter à travers *Internet* au *serveur*.

Rootkit [62]

Un « rootkit » est un programme ou un ensemble de programmes permettant à un pirate de maintenir -dans le temps- un accès frauduleux à un système informatique. Le prérequis du rootkit est une machine « déjà » piratée.

La fonction principale du « rootkit » est de simplifier, voire automatiser, la mise en place d'une ou plusieurs « *backdoors* ». Ces « portes dérobées » (utilisables en local ou à distance) permettent au pirate de s'introduire à nouveau au cœur de la machine sans pour

autant exploiter une nouvelle fois la faille avec laquelle il a pu obtenir l'accès frauduleux initial.

De plus, certains « rootkit » opèrent une suite de modifications, notamment au niveau du noyau (kernel) permettant de cacher des fichiers, des processus... Rien à voir donc avec un virus ou ver de nouvelle génération. Un « rootkit » ne se réplique pas.

L'installation d'un « rootkit » nécessite des droits administrateurs sur la machine, notamment à cause des modifications profondes du système qu'il engendre. Cela signifie que le pirate doit initialement disposer d'un accès frauduleux, avec les droits du « root » sous Linux par exemple, afin de mettre en place son « rootkit ».

À aucun moment un « rootkit » ne permet de s'introduire de manière frauduleuse sur une machine saine. En revanche, certains « rootkit » permettent la collecte des mots de passes qui transitent par la machine « corrompue ». Ainsi, un « rootkit » peut indirectement donner l'accès à d'autres machines.

Certains « rootkit » sont également livrés avec des collections d'« *exploits* », ces petits bouts de code dédiés à l'exploitation d'une faille bien déterminée. Le but est d'aider les pirates dans leur conquête de machines encore vierges.

Le rootkit automatise l'installation d'une *porte dérobée* ou d'un *cheval de Troie*. Le ver automatise l'exploitation d'une vulnérabilité à travers le réseau et peut accessoirement installer une backdoor une fois au cœur d'une machine.

Le « rootkit » n'a de raison d'être que si une faille est présente, si les conditions sont réunies pour que son exploitation soit réussie et si elle permet un accès avec les droits administrateur. Par transitivité, pas de faille, pas de rootkit.

La discrétion est l'essence même du « rootkit ». Il permet à un pirate de cacher son intrusion et sa présence sur une machine. Le meilleur moyen de se protéger des rootkit est donc de se prémunir des failles.

Spam [56]

Le spam, mot anglais du jargon informatique, désigne les communications électroniques massives, notamment de courrier électronique, sans sollicitation des destinataires, à des fins publicitaires ou malhonnêtes.

Les mots pourriel (de « poubelle » et « courriel »), pollurriel (de « pollution » et « courriel ») et merdiel (à l'étymologie transparente), ainsi que d'autres variantes formées sur des racines exprimant l'exaspération des utilisateurs, ont également été formés pour désigner le spam. Le mot pourriel est d'usage assez courant (probablement par euphonie avec pourri), pollurriel est plus rarement utilisé, alors que les autres termes apparaissent ou disparaissent selon la fantaisie et l'irritation des victimes.

Le verbe spammer est souvent utilisé pour qualifier l'action d'envoyer du spam, le spamming. Le mot spammeur désigne celui qui envoie du spam. Les mots polluposter, pollupostage et polluposteur sont également utilisés.

Origine du mot

Le mot spam provient d'un sketch des Monty Python dans lequel le même mot, désignant un jambon en boîte de basse qualité, envahit la conversation et le menu d'un petit restaurant. Spam est l'acronyme de Shoulder of Pork and hAM (épaule de porc et jambon), ou selon d'autres sources Spiced Pork and hAM (porc épicé et jambon), Spiced Pork And Meat ou simplement SPiced hAM. Ce sketch parodiait d'ailleurs une des premières formes de spam.

En effet c'est une publicité radiophonique pour le Spam, pendant laquelle le terme était répété de nombreuses fois, qui est à l'origine du sketch des Monty Python.

Contenu du spam

Le spam contient généralement de la publicité. Des escrocs envoient également des propositions malhonnêtes de cette façon. Les lettres en chaînes peuvent aussi être qualifiées de spam.

Les produits les plus vantés sont les services pornographiques, les médicaments, le crédit financier ou des escroqueries prétendant enrichir rapidement.

Parfois aussi il s'agit de messages d'entreprises ignorantes de la [Netiquette](#) qui y voient un moyen peu coûteux d'assurer leur promotion.

Enfin la dernière forme de spam, le [phishing](#), consiste à tromper le destinataire en faisant passer le message pour un message de sa banque ou d'un quelconque service protégé par mot de passe. Le but est de récupérer les données personnelles des destinataires (notamment des mots de passe) en les attirant sur un site factice enregistrant toutes leurs actions.

Cibles du spam

Le spam peut s'attaquer à divers médias électroniques : les courriels, les forums de discussion de Usenet, les moteurs de recherche, les wikis, les messageries instantanées.

Par courrier électronique

Le spam par courrier électronique (pourriel) est le type de spam le plus répandu. Le coût d'envoi d'un courrier électronique étant négligeable, il est facile d'envoyer un message à des millions de destinataires. Les destinataires assument le coût de réception et de stockage en boîte aux lettres, ce qui peut causer des coûts non négligeables aux prestataires de services, à cause du volume pris par le spam.

Contrairement aux promotions commerciales pour lesquelles les utilisateurs peuvent avoir donné leur accord, le spam n'est pas sollicité. Il est souvent rédigé spécialement pour contourner les filtres anti-spam. Un mot clé tel que Viagra (souvent vanté par le spam) peut être ainsi écrit « v1@gr@ » ou « v|agra » de manière à tromper une règle de filtrage basée sur ce mot.

Les spammeurs redoublent d'imagination pour masquer leurs activités et ne pas être démasqués, que ce soit en falsifiant les adresses d'expéditeur ou en utilisant des serveurs SMTP (serveur de courrier électronique) non sécurisés qui permettent des envois anonymes.

Les adresses à spammer sont généralement collectées par robot d'indexation. Il existe un marché pour les listes d'adresses.

Pour éviter d'être spammé, les internautes font souvent figurer leurs adresses d'une manière masquée lorsqu'elle doit apparaître dans un site Web ou dans Usenet. Par exemple:

- Jean@NOSPAM.exemple.fr pour Jean@exemple.fr.
- Jean à exemple point fr pour Jean@exemple.fr

Par message de forum de discussion

Ce type de spam est apparu sur Usenet avant le spam par courrier électronique. Les forums de discussion de Usenet sont une cible facile du spam. En effet, un message envoyé à un forum touche tous les lecteurs du forum. Certains groupes de discussion ne reçoivent pratiquement plus que du spam (c'est l'une des raisons pour lesquelles de nombreux forums sont modérés, c'est-à-dire surveillés par un humain ou un robot qui effectue un tri parmi les articles proposés). D'autre part, les usenaute faisant généralement figurer leur adresse électronique dans leurs articles, les spammeurs peuvent facilement récolter des milliers

d'adresses au moyen d'un robot, puis spammer les auteurs de ces articles par courrier électronique.

Le phénomène est rendu encore plus pénible et par la publication croisée ou la publication multiple, qui consistent respectivement à destiner un message à plusieurs groupes simultanément ou à envoyer le message dans plusieurs de groupes de suite.

Pour cette raison, tout message de promotion, même d'un livre, est prohibé dans les forums Usenet, à l'exception de news:alt.business.

Dans Usenet, on parle également de spam lorsqu'un article, quel que soit son contenu, et même s'il n'appartient pas aux catégories usuelles de messages abusifs (publicités commerciales, escroqueries, insultes...) est publié en un nombre d'exemplaires excessif: tous les exemplaires d'un tel article peuvent être annulés par les utilisateurs (les critères numériques exacts permettant d'identifier de tels spams dans la principale hiérarchie francophone sont donnés dans la documentation du forum news:fr.usenet.abus.d). Les diverses hiérarchies possédant des critères différents pour identifier et annuler les articles de spam, il existe une certaine incertitude quant aux traitements qui peuvent être appliqués aux spams diffusés simultanément dans plusieurs de ces hiérarchies. Une règle particulièrement simple décide du sort des escroqueries manifestes, c'est-à-dire des articles proposant de « gagner de l'argent rapidement et sans rien faire » (habituellement appelés « MMF », de l'anglais Make money fast: « gagnez de l'argent rapidement »): ces articles peuvent être annulés immédiatement par n'importe quel utilisateur.

Les actions des usenautes spécialisés dans la lutte contre le spam donnent souvent lieu à des accusations de censure et de cabale.

Par systèmes vocaux

Le développement de la voix sur IP ([VoIP](#), téléphonie par Internet) va avoir pour conséquence l'arrivée en 2005 sur nos combinés d'un nouveau type de spam, le spam vocal, baptisé SpIT (Spam over Internet Telephony).

Spamdexing

Le spam destiné aux robots d'indexation de moteur de recherche consiste à modifier des pages Web pour augmenter les chances d'avoir un bon classement dans le moteur de recherche.

Parmi les techniques utilisées :

- La manipulation des mots clés consistant à ajouter une longue liste de mots souvent recherchés (comme « sexe » ou « piratage ») répétitivement dans une page (« sexe à UneVille », « sexe à UneAutreVille », « sexe à EncoreUneAutreVille », « sexe à UnVillageTropPetitPourÊtreUneVille », et toute autre variation possible) pour apparaître immédiatement si on fait une recherche avec ces mots. Parfois une page ne contient que les résultats d'une recherche, mis sur le Web pour être trouvé et classé par les moteurs de recherche et affiché aux usagers cherchant avec les mêmes mots.
- Le bourrage de mots clés populaires, dissimulés au visiteur mais pas au robot, soit en les imprimant blanc sur fond blanc, ou en utilisant la police de caractères la plus petite, ou encore en les utilisant en lignes « commentaire » et « méta » qui ne sont pas affichées à l'utilisateur, ou encore en changeant le contenu de la page après que le logiciel « robot » ou « araignée » l'avait lu ou en modifiant le serveur pour envoyer une page au moteur de recherche et une autre aux usagers ordinaires.
- Le spam de liens consiste à mettre les liens vers un site qu'on veut promouvoir dans autant d'autres sites externes que possible, incluant les forums publics et les pages de commentaires d'autres sites.

- Une ferme de liens (link farm) est un site hébergeant des listes de liens vers tous les autres sites qu'on contrôle pour améliorer le classement de ces derniers en les faisant apparaître populaires. Google compte notamment la quantité et l'importance des liens vers un site pour déterminer l'importance du site. Parfois on construit aussi les sites multiples (simulant des sites indépendants et pas simplement des sous domaines du même site) avec presque le même contenu; chacun contient un tas de liens vers tous les autres pour améliorer leur classement.
- La technique du Googlebombing en sa forme originale consiste à placer des hyperliens vers George W. Bush avec des phrases comme « l'idiot du monde » dans autant de sites Web que possible. La destination de cet hyperlien est normalement un site externe (dans cet exemple, la page de Bush). Si ce genre de lien figure dans un assez grand nombre de pages Web, une recherche pour « l'idiot du monde » va diriger l'utilisateur immédiatement sur Bush, peu importe s'il désire ça et peu importe si ces mots figurent sur son site ou pas.
- Une autre variation est le Spam des affiliés (affiliate spam) ou une compagnie paye pour chaque visiteur ou chaque client envoyé par des liens affichés par des autres, du genre «affiliez-vous et devenez riche, mettez un lien vers www.arnaqueur.porno.example.com et pour chaque victime qui nous donne tous ses numéros de carte de crédit nous vous donnons un sou ». Les liens venant de ces programmes d'affiliation contiennent le code d'identification d'un affilié de façon www.arnaqueur.porno.example.com/donnemoiargent?MonsieurLeSpammeur pour laisser savoir qui doit être payé pour avoir posté tous ces liens partout.

Les opérateurs de sites de recherche comme Google cherchent toujours des moyens de détecter ce genre de choses et les rendre plus difficiles à utiliser effectivement; si on a un tas de liens venant des sites du genre «www.ferme-aux-liens-inutiles.spam.example.com» on peut se faire enlever de l'importance dans les recherches au lieu d'en gagner.

Lutte contre le spam

- Lutte technique

Les techniques pour lutter contre le spam mettent en œuvre diverses techniques de classification automatique pour trier entre le spam et le courrier légitime. Ces techniques peuvent être mises en œuvre soit au niveau des fournisseurs de service Internet qui protègent leur messagerie, soit au niveau des utilisateurs par des outils appropriés.

Ces techniques peuvent être soit préventives (marquage du courrier pour indiquer qu'il s'agit de spam) soit curatives (blocage, voire renvoi des messages incriminés vers l'expéditeur). À noter que cette dernière comporte des inconvénients puisque le destinataire doit pouvoir être maître des courriers qu'il souhaite recevoir. De plus renvoyer un message ne peut que faire empirer la situation en occupant un peu plus le réseau, avec de fortes probabilités que l'auteur du spam ait maquillé sa véritable adresse ou utilisé l'adresse d'un tiers (tout à fait innocent) comme adresse de retour.

Plusieurs techniques de lutte contre le spam sont possibles et peuvent être cumulées : analyse statistique (méthode bayésienne), filtrage par mots clés ou par auteur, listes blanches (désignation de personnes ou de machines autorisées à publier dans certains lieux), listes noires

(désignation de personnes ou de machines auxquelles il est interdit de publier dans certains lieux), interrogation en temps réel de serveurs spécialisés dans la lutte anti-spam.

Ces techniques de lutte, tout comme les *antivirus*, doivent s'adapter en permanence car de nouveaux spams réussissent à contourner ces défenses.

- Méthode bayésienne

Cette méthode d'analyse statistique utilise l'inférence bayésienne formulée par le mathématicien Thomas Bayes. Celle-ci permet d'associer des probabilités aux mots contenus dans les courriers. En fonction du pointage obtenu, la probabilité qu'il s'agisse vraiment de spam augmente ou diminue. Cette méthode requiert une phase d'apprentissage de mots autorisés et interdits pour être réellement efficace.

La méthode bayésienne sert également à d'autres classifications automatiques du courrier, en particulier dans Lotus Notes.

- Filtrage par mots clés ou adresses

Cette méthode est très limitée car elle se base sur le rejet ou le tri du courrier en fonction de règles de vocabulaire préalablement établies, définissant des mots comme interdits. Certains mots clés revenant souvent dans le spam, tels que « sexe », « viagra » ou « money » pourront servir de base pour la constitution de ces règles. De même on pourra décider de bloquer tous les messages en provenance d'un expéditeur précis, d'un domaine spécifique, voire d'un pays entier.

Cette méthode engendre de fortes probabilités d'erreur et s'avère également peu efficace lorsque les spammeurs maquillent les mots utilisés (« vi@gr@ », « s3x », etc.). Il convient alors d'utiliser les expressions régulières.

- Rendre les courriels payants

Mettre un prix sur l'envoi de courriel, symbolique pour les envois légitimes mais dissuasif pour les envois massifs (à 2 centimes d'euros par courrier, celui-ci reste toutefois du même ordre de coût pour l'expéditeur qu'une publicité radio ; or elle peut être bien mieux ciblée selon l'endroit où a été récoltée l'adresse). Et à 20 centimes d'euros il sera nécessaire de mettre une franchise sinon c'est l'accès à l'envoi de courrier pour le particulier au budget le plus serré qui commence à s'estomper.

- Modération

Dans les forums Internet et Usenet, ainsi que sur les listes de diffusion, on a souvent recours à la modération: une personne de confiance (« modérateur ») lit les messages dont la publication est proposée, et refuse éventuellement de

les diffuser (modération a priori); ou bien cette personne lit les messages qui ont déjà été diffusés, et efface ceux qui lui semblent hors de propos. Comme cette méthode nécessite des moyens humains importants, et que de plus les modérateurs sont souvent accusés de censurer à outrance, il existe aussi une modération par robot (généralement appelée « robot-modération »): n'importe qui peut publier un message par l'intermédiaire du robot, même si cet article est dépourvu d'intérêt (et même s'il constitue effectivement un spam), mais le robot ne laisse passer le message que s'il répond à un critère simple et connu de tous, comme la présence d'un certain mot dans son titre. Cette protection est surtout efficace contre les robots qui émettent automatiquement des messages identiques dans des dizaines de forums, et qui n'ont pas été programmés pour produire des messages conformes aux exigences spécifiques de tel ou tel forum.

- Lutte judiciaire et législative

En France, le pourriel est réglementé, d'autant plus qu'il implique la possession, la conservation (et souvent le commerce) de listes d'adresses électroniques récupérées automatiquement (dans des forums de discussion, des sites Web), ce en contradiction avec la loi Informatique et libertés. Une loi impose l'accord des destinataires pour tout type d'adresse comportant le nom d'une personne.

Aux États-Unis, le spam est réglementé depuis 2003 par une loi appelée CAN-SPAM Act. Elle autorise le spam, à condition que le sujet du courrier soit descriptif, que l'adresse d'expédition soit valide et qu'une méthode de désinscription (hyperlien) soit fourni.

Dans de nombreux pays, aucune réglementation spécifique au spam n'existe.

Quelques poursuites judiciaires ont été amorcées en utilisant des lois existantes :

- Si on utilise une fausse adresse de retour et cette adresse appartient à quelqu'un d'autre, cela peut être considéré comme une usurpation d'identité.
- Si on promeut une action de compagnie en bourse, on peut être accusé de pratiquer le courtage boursier sans licence.
- Si on continue à utiliser un serveur après que son propriétaire ait demandé à la cour une injonction de désistement, c'est du vol de temps de processeur d'ordinateur qui peut être interdit par des lois conçues pour arrêter d'autres attaques contre les systèmes informatiques.
- Si on fait la promotion de produits médicaux, on risque d'être trouvé coupable de pratique de pharmacien, médecin ou infirmière sans licence.
- L'envoi de publicités pornographiques vers des boîtes à lettres d'enfants ne sera pas une bonne idée.
- Si on commet d'autres crimes, comme la fraude ou le sabotage des pages Web ou sites informatisés, on peut se retrouver en prison.

L'emprisonnement est rare mais cela arrive : Dave Rhodes, qui envoyait des arnaques du genre Ponzi ou « pyramide » intitulées « Make Money Fast » au début des années 1990, se retrouva en prison pour quelques années, trouvé coupable de fraude.

La République Populaire de Chine a déjà condamné à mort et exécuté des personnes dont le seul crime était l'envoi de spams.

Le plus souvent, les poursuites judiciaires qui se sont déjà déroulées étaient des procès civils coûteux amorcés par les grands fournisseurs comme AOL ou Yahoo! contre les spammeurs les plus abusifs du réseau, ceux qui envoient des millions de courriers. Le site cyberpromo.com a dû fermer ses portes à cause de poursuites judiciaires de ce genre et à cause de difficulté à trouver un fournisseur d'accès à Internet prêt à donner l'accès au réseau à une telle compagnie.

Comme le problème est international, les lois nationales ont assez peu d'effet sur le volume du spam.

SSL [78]

Secure Socket Layer (SSL) est un protocole de sécurisation des échanges sur Internet, développé à l'origine par Netscape (SSL version 2 et SSL version 3). Il a été renommé en Transport Layer Security (TLS) par l'[IETF](#) (TLS version 1). Il y a très peu de différence entre SSL version 3 et TLS version 1. Par abus de langage, on parle de SSL pour désigner indifféremment SSL ou TLS.

SSL fonctionne suivant un mode client-serveur. Il fournit quatre objectifs de sécurité :

- l'authentification du serveur;
- la confidentialité des données échangées (ou session chiffrée);
- l'intégrité des données échangées
- de manière optionnelle, l'authentification du client.

Avec le développement d'[Internet](#), de nombreuses sociétés commerciales proposent des achats en ligne pour les particuliers. L'offre croît tous les jours, mais le chiffre d'affaires dégagé par le [commerce électronique](#) (e-commerce) reste encore modeste car le client n'a pas encore une confiance totale dans le paiement par carte bancaire. Une des façons de sécuriser ce paiement est d'utiliser des protocoles d'authentification et de chiffrement tels que SSL, mais cela ne dispense pas le commerçant de veiller soigneusement aux numéros de carte bancaire parfois stockés sur ses serveurs.

La session chiffrée est utilisée généralement lors de l'envoi du numéro de carte bancaire, mais elle peut l'être dans d'autres cas. Le chiffrement est réalisé par à la fois un chiffrement asymétrique (qui va permettre une authentification) comme par exemple l'algorithme RSA et à la fois par un chiffrement symétrique (qui est plus léger qu'un chiffrement asymétrique) et qui va assurer la transmission des informations (comme par exemple le DES). RSA et DES sont des méthodes d'encryptage de données.

Ver informatique [57]

Un ver informatique est un [logiciel malveillant](#) qui se reproduit sur des ordinateurs à l'aide d'un réseau informatique comme l'Internet.

Un ver contrairement à un [virus informatique](#) n'a pas besoin d'un programme hôte pour se reproduire. Il exploite les différentes ressources afin d'assurer sa reproduction et éventuellement avoir d'autres effets directs ou indirects (espionner, offrir un point d'accès caché, faire des dégâts, etc.).

Le plus souvent écrit sous forme de script intégré dans un courriel ou sur une page [HTML](#) sur internet, il est exécuté quand le système y accède.

Virus informatique [53]

Un virus informatique est un logiciel malveillant écrit dans le but de se dupliquer sur d'autres ordinateurs. Il peut aussi avoir comme effet, recherché ou non, de nuire en perturbant plus ou moins gravement le fonctionnement de l'ordinateur infecté. Il peut se répandre à travers tout moyen d'échange de données numériques comme l'[Internet](#), mais aussi les disquettes, les [cédéroms](#), les [clefs USB](#) etc.

Son appellation provient d'une analogie avec le virus biologique puisqu'il présente des similitudes dans sa manière de se propager et de se reproduire.

Le nombre total de virus couverts par Sophos s'élève à 93 875 (tous types confondus, en août 2004 (http://www.mag-secur.com/article.php3?id_article=1208)). Ce chiffre n'est qu'une approximation grossière du nombre réel de virus en circulation, chaque éditeur d'[antivirus](#) ayant intérêt à « gonfler » la réalité, d'autant plus que sur tous les virus identifiés, très peu atteignent le stade de la diffusion massive sur les réseaux. La très grande majorité concerne la plate-forme Windows. Le reste est essentiellement destiné à des systèmes d'exploitation qui ne sont plus distribués depuis quelques années (comme Mac OS 9 et ses prédécesseurs).

Les virus font souvent l'objet de fausses alertes que la rumeur propage, encombrant les messageries. Certaines d'entre elles, jouant sur l'ignorance en informatique des utilisateurs, leur font parfois détruire des éléments de système d'exploitation totalement sains.

Les différents types de virus

- Le virus classique est un morceau de programme, souvent écrit en assembleur, qui s'intègre dans un programme normal (ou dans le Master Boot Record dans le cas d'un virus de boot), le plus souvent à la fin, mais aussi au début ou même au milieu. À chaque fois que l'utilisateur exécute ce programme « infecté » il active le virus qui en profite pour aller s'intégrer dans d'autres programmes exécutables. De plus, lorsqu'il contient une charge virale, il peut après un certain laps de temps (qui peut être très long) ou un événement particulier, corrompre des fonctions du système de l'ordinateur ou des fichiers de l'utilisateur. Cela peut aller d'un simple message anodin à la destruction complète de toutes les données de l'ordinateur. On parle dans ce cas de bombe logique ou de charge utile.
- Les macro-virus qui s'attaquent aux macros de logiciels de la suite Microsoft Office (Word, Excel, etc.) grâce au VBA de Microsoft. Par exemple, en s'intégrant dans le modèle normal.dot de Word, un virus peut être activé à chaque fois que l'utilisateur lance ce programme.
- Les virus de boot

D'autres menaces existent en informatique, s'en distinguant souvent par l'absence de système de reproduction caractéristique des virus : le terme logiciel malveillant est dans ce cas plus approprié.

Caractéristiques

- la résidence : dès son exécution, le virus s'extrait de son hôte et va se loger dans la mémoire vive où il prend le contrôle de la machine ;
- la cryptographie : à chaque répllication, le virus est chiffré (afin de dissimuler les instructions qui, si elles s'y trouvaient en clair, révéleraient la présence de ce virus ou pourraient indiquer la présence de code suspect);

- la furtivité : le virus « trompe » le système d'exploitation (et par conséquent les logiciels antivirus) sur l'état des fichiers infectés ;
- le polymorphisme : le virus est chiffré et la routine de déchiffrement est capable de changer certaines de ses instructions au fil des répliques afin de rendre plus difficile la détection par les antivirus.
- le métamorphisme : contrairement au chiffrement simple et au polymorphisme, où le corps du virus ne change pas et est simplement chiffré, le métamorphisme permet au virus de modifier sa structure même et les instructions qui le composent

Les logiciels antivirus

Les antivirus sont des logiciels capables de détecter des virus, détruire, mettre en quarantaine et parfois de réparer les fichiers infectés sans les endommager. Ils utilisent pour cela de nombreuses techniques, parmi lesquelles :

- la reconnaissance de séquences d'octets caractéristiques (signatures) d'un virus particulier ;
- la détection d'instructions suspectes dans le code d'un programme (analyse heuristique);
- la création de listes de renseignements sur tous les fichiers du système, en vue de détecter d'éventuelles modifications ultérieures de ces fichiers par un virus ;
- la détection d'ordres suspects ;
- la surveillance des lecteurs de support amovible : disquettes, *clé USB*, *CD-ROM*...

Virologie

Le terme virus informatique a été créé par analogie avec le virus en biologie : un virus informatique utilise son hôte (l'ordinateur qu'il infecte) pour se reproduire et se transmettre à d'autres ordinateurs.

Comme pour les virus biologiques, où la diversité génétique ralentit les chances de croissance d'un virus, en informatique ce sont les systèmes les plus répandus qui sont le plus atteints par les virus : (Microsoft Windows, Microsoft Office, Microsoft Outlook, Microsoft Internet Explorer et Microsoft Internet Information Server).

Cependant, des systèmes moins répandus ne sont pas touchés proportionnellement. La majorité des autres systèmes, en tant que variantes de l'architecture UNIX (BSD, Mac OS X ou Linux), utilisent en standard une gestion des droits de chaque utilisateur, qui leur permet d'éviter les attaques les plus simples et les dégâts sont normalement circonscrits à l'utilisateur, laissant la base du système d'exploitation intacte. Les versions professionnelles de Windows (NT/2000/XP pro) permettent cependant de gérer les droits de la même manière.

Le facteur le plus important de la multiplication des virus sous Microsoft Windows est sa grande popularité, qui fait de lui une cible de choix pour les créateurs de virus. De plus, l'ouverture par défaut de ports réseau, non indispensables au fonctionnement standard, mais réclamés par le système de mise à jour automatique et d'autres fonctionnalités très peu documentées. La possibilité d'exécuter automatiquement des scripts dans les courriels est une autre source d'infection.

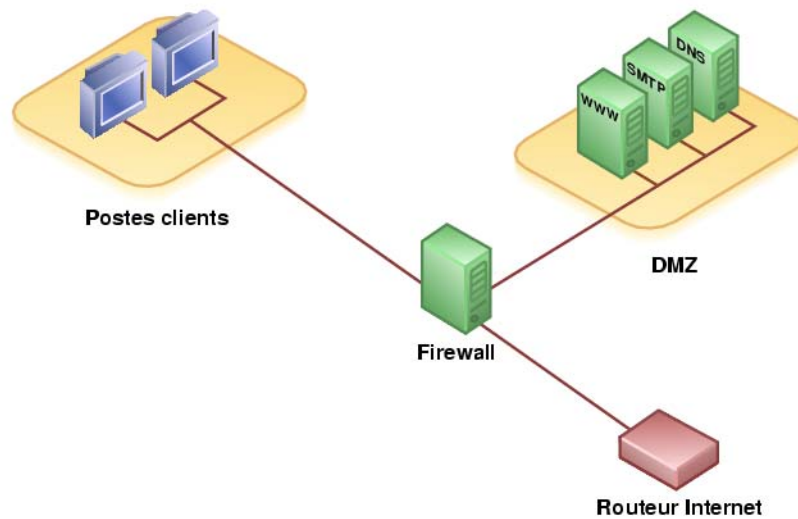
La démocratisation de l'accès à Internet a été un facteur majeur dans la rapidité de propagation à grande échelle des virus les plus récents. Ceci est notamment dû à la faculté des virus de s'approprier des adresses de courriel présentes sur la machine infectée (dans le carnet d'adresses mais aussi dans les messages reçus ou dans les archives de pages web visitées ou de messages de groupes de discussions).

De même, l'interconnexion des ordinateurs en réseaux locaux a amplifié la faculté de propagation des virus qui trouvent de cette manière plus de cibles potentielles.

Zone démilitarisée [65]

En sécurité informatique, une zone démilitarisée (DMZ) est un sous-réseau isolé par un pare-feu. Ce sous-réseau contient des machines se situant entre un réseau interne (LAN) et un réseau externe (typiquement, Internet).

La DMZ permet à ses machines d'accéder à Internet et/ou de publier des services sur Internet. En cas de compromission d'une machine, l'accès vers le réseau interne sera interdit.



Les technologies multimédia

Sommaire :

- un abécédaire multimédia



A. Un abécédaire multimédia [90]

Le mot multimédia est apparu vers la fin des années 1980, lorsque les [CD-ROM](#) se sont développés. Il désignait alors les applications qui, grâce à la mémoire du CD et aux capacités de l'ordinateur, pouvaient générer, utiliser ou piloter différents médias simultanément :

- musique ;
- son ;
- image ;
- vidéo ;
- interface homme-machine interactive.

Ce mot a été créé plus pour des besoins mercatiques que pour refléter de réels progrès techniques. Différents logiciels présentaient toutes les caractéristiques de ces nouveaux logiciels sur différents micro-ordinateurs avant l'arrivée du CD-ROM, même si la qualité en était parfois inférieure. Auparavant, on parlait plutôt de « vidéo interactive » qui consistait au pilotage de magnétoscopes par des ordinateurs.

Il faudra attendre la fin des années 1990 avec l'arrivée de méthodes de compression de son et vidéo, ainsi qu'une certaine montée en puissance des ordinateurs personnels, pour atteindre simultanément des qualités semblables aux différents autres médias réunis.

Aujourd'hui on utilise le mot multimédia pour désigner toute application utilisant ou servant à travailler sur au moins un média spécifique.

Les manettes de jeux, gants tactiles, lunettes 3D avec gyroscope, écrans tactiles et autres interfaces homme-machine sont souvent considérées comme faisant partie du multimédia ; alors qu'elles ne sont en elles-mêmes pas des « média » mais plutôt des extensions de média (une lunette 3D sans image à afficher par exemple).

Par ailleurs, en recherche en informatique, on nomme multimédia l'étude des médias non textuels, principalement les images, les vidéo et les sons.

Dans ce support, nous nous intéressons uniquement aux supports : image et vidéo.

AVI [104]

L'Audio Video Interleave (audio vidéo imbriqués) (dont l'acronyme est AVI), est un format de fichier conçu pour stocker des données audio et vidéo.

AVI utilise un même paquet standard « [fichier conteneur](#) » afin d'être lu simultanément.

Il a été présenté par Microsoft en novembre 1992, en tant qu'élément de la vidéo pour la technologie de Windows.

Dans un fichier AVI, chaque composante audio ou vidéo peut être compressée par n'importe quel codec. Le format [DivX](#) est souvent utilisé comme [codec](#) vidéo, et le format [mp3](#) comme codec audio, mais d'autres codecs peuvent également être utilisés, par exemple [XviD](#) ou [MPEG](#) pour la vidéo, et mp2, [WAV](#) etc. pour l'audio.

Le format AVI permet de réunir en un seul fichier une piste vidéo et jusqu'à 99 pistes audios, ce qui permet de bénéficier, par exemple, de plusieurs langues pour un même film.

Codec [93]

Le terme Codec est construit d'après les mots COMpression et DÉCompression.

Il s'agit d'un procédé permettant de compresser et de décompresser un signal, de l'audio ou de la vidéo, le plus souvent en temps réel. Le procédé peut être sous forme de logiciel ou encore de matériel (hardware). Par extension, c'est le logiciel ou le circuit qui contient cet algorithme.

Ces algorithmes de compression de données permettent, en général, de réduire la taille du fichier original par des facteurs allant de 2 à 100 (voire plus pour certaines applications).

La compression se fait soit avec des algorithmes purement mathématiques de compression de données sans perte d'information (comme un fichier gzip) ou par des algorithmes prenant en compte les caractéristiques des données à compresser et qui peuvent perdre des informations dites « non pertinentes ».

La compression [Ogg Vorbis](#) par exemple, compresse le son suivant des critères « psycho-acoustiques » prenant en compte les fréquences non ou peu audibles du spectre sonore telles que les harmoniques et les fréquences très aiguës. C'est une compression « destructive », car elle perd des informations sonores.

Les « codecs » vidéo [MPEG](#), [DivX](#) ou [XviD](#) par exemple, utilisent des algorithmes prenant en compte la persistance rétinienne, la différenciation des couleurs par l'œil et d'autres « imperfections » de la vue pour enlever de la compression à effectuer des détails que l'être humain ne perçoit presque pas en temps normal. Il s'agit aussi de compressions dites destructives.

Il ne faut pas confondre les « codecs » avec les « [conteneurs](#) », les flux audio et/ou vidéo étant stockés dans ces derniers. Chaque conteneur accepte tel ou tel codecs audios et vidéos, et permet la présence d'une ou plusieurs pistes audios, de sous-titres intégrés ou non, de chapitres, et éventuellement d'autres informations. Certains formats de fichiers désignent à la fois le conteneur et le codec, d'autres uniquement le conteneur, et d'autres uniquement le codec. À cela s'ajoute la notion de support : disque dur, disque optique CD ou DVD, mémoire flash...

Pour rappel, les principaux « conteneurs » sont :

- AVI qui peut contenir une piste vidéo au codec [MPEG-1](#), [MPEG-2](#), [DivX](#), [XviD](#) et une piste audio au codec [MP3](#) ou [WMA](#)
- BivX qui est une variante de l'AVI qui permet d'intégrer une deuxième bande-son.
- [Ogg](#) qui contient du son ogg vorbis et/ou de la vidéo ogg theora
- Ogm qui est un hack de l'AVI qui permet d'intégrer des pistes sons [Ogg Vorbis](#) et des sous-titres.
- QuickTime (c'est le nom de la technologie multimédia d'Apple)
- RealMediaVideo avec l'extension RV9 et RV10, il peut désigner selon le contexte, ou le conteneur, ou le codec, ou les deux.
- Divx6 qui autorise l'incorporation de menu du même type que ceux que l'on trouve sur les DVD vidéo, mais qui n'est pas libre.

Les principaux codecs audio sont :

- [MP3](#)
- [Ogg Vorbis](#)
- [WMA](#)

Les principaux codecs vidéo sont :

- [MPEG-1](#), qui n'est plus utilisé en 2005, à moins de posséder quelques [VCD](#).
- [MPEG-2](#), qui est le codec utilisé en 2005 notamment par les DVD vidéo standard.
- [MPEG-4](#), et ses implémentations [DivX](#), [XviD](#), [WMV](#)...
- [Ogg Theora](#)

Les principaux codecs de compression d'images sont :

- [PNG](#)
- [JPG](#)
- [JPG 2000](#)
- [GIF](#)
- [TIFF](#)

Compression de données [91]

La compression de données traite de la manière dont on peut réduire la quantité d'information, souvent mesurée en bits, utilisée pour représenter une séquence d'information. Elle est une branche de la théorie de l'information.

La compression peut concerner un flux d'information transmis ou un ensemble statique de données (par exemple un fichier). Dans le premier cas les données sont compressées lors de la transmission, dans le second les données sont compressées lors du stockage de l'information.

On peut classer les méthodes de compression en deux types, compression avec perte et compression sans perte.

- Compression sans perte

La compression est dite sans perte lorsqu'il n'y a aucune perte de données sur l'information d'origine. Il y a autant d'information après la compression qu'avant, elle est seulement réécrite d'une manière plus concise (c'est par exemple le cas de la compression gzip). La compression sans perte est dite aussi compactage.

L'information à compresser est vue comme la sortie d'une source de symboles qui produit des textes finis selon certaines règles. Le but est de réduire la taille moyenne des textes obtenus après la compression tout en ayant la possibilité de retrouver exactement le message d'origine (on trouve aussi la dénomination codage de source en opposition au codage de canal qui désigne le codage correcteurs d'erreurs).

Les formats de fichier de compression sans perte les plus courants sont : rar, zip, ace, etc.

- Compression avec pertes

Utilisée pour compresser des photos, des bandes musicales, des films...

Cette technique est fondée sur une idée simple : seul un sous-ensemble très faible de toutes les images possibles (à savoir celles que l'on obtiendrait par exemple en tirant les valeurs de chaque *pixel* par un générateur aléatoire) possède un caractère exploitable et informatif pour l'œil. Ce sont donc ces images-là qu'on va s'attacher à coder de façon courte. Dans la pratique, l'œil a besoin pour identifier des zones qu'il existe des corrélations entre pixels voisins, c'est-à-dire qu'il existe de zones contiguës de couleurs voisines. Les programmes de compression s'attachent à découvrir ces zones et à les coder de la façon aussi compacte que possible. Le *JPEG 2000*, par exemple, arrive typiquement à coder des images photographiques sur 1 bit par pixel sans perte visible de qualité sur un écran, soit une compression d'un facteur 24 à 1.

De même, seul un sous-ensemble très faible de sons possibles est exploitable par l'oreille, qui a besoin de régularités engendrant elles-mêmes une redondance (coder avec fidélité un bruit de souffle n'aurait pas grand intérêt). Un codage éliminant cette redondance et la restituant à l'arrivée reste donc acceptable, même si le son restitué n'est pas en tout point identique au son d'origine.

Il y a moins d'information après la compression qu'avant, l'information retranchée étant sélectionnée d'après des critères fixés selon le type de données traitées. La compression d'une image en format *jpeg* est un exemple de compression avec perte. Puisque l'œil ne perçoit pas nécessairement tous les détails d'une image, il est possible de retrancher des données, dans l'espace des fréquences, de telle sorte que le résultat soit très ressemblant à l'original, voire pareil, pour l'œil. Le tout est de savoir quelles données retrancher. L'image finale n'étant pas, numériquement parlant, identique à l'image initiale, il s'agit d'une compression avec perte.

Récapitulatif

Domaines	sans pertes	avec pertes
Audio	<i>FLAC</i>	<i>MP3, Ogg Vorbis</i>
Image	<i>GIF, PNG, PCX</i>	<i>JPEG, JPEG2000</i>
Vidéo		<i>MPEG, MPEG-2, MPEG-4</i>

Compression MPEG [101]

MPEG en général

MPEG (Moving Picture Experts Groups) est un groupe de travail sous ISO/IEC, fondé en 1988, responsable du développement des normes internationales pour la compression, la

décompression, le traitement et la représentation codée d'images mobiles, de l'audio et de leur combinaison.

Jusqu'ici le MPEG a produit :

MPEG-1 : Une norme pour la mémorisation et la récupération d'images mobiles et de l'audio associé sur des supports de stockage.

MPEG-2 : Une norme pour la télévision digitale.

Deux normes supplémentaires sont actuellement en cours de développement:

MPEG-4 : Une norme pour les applications multimédia.

MPEG-7 : Une norme de représentation du contenu pour la recherche de l'information.

Principes fondamentales des algorithmes MPEG

Les séquences vidéo contiennent une très grande redondance statistique et subjective entre des images successives ; dans le domaine temporel comme dans le domaine spatial.

La propriété statistique de base sur laquelle les techniques de compactage de MPEG comptent est la corrélation entre *pixels*, y compris l'acceptation du mouvement de translation corrélé simple entre les trames consécutives.

Ainsi, on suppose que l'importance d'un pixel particulier de l'image peut être prévue des pixels voisins de la même trame (utilisant des techniques de codage Intra-frame) ou des pixels d'une trame voisine (utilisant des techniques inter-frame). Intuitivement il est clair que dans quelque circonstance, c.-à-d. pendant un changement de scène d'une séquence vidéo, la corrélation temporelle entre pixels dans des trames voisines est petite et peut même disparaître – les scènes vidéo alors semblent une collection non corrélée d'images fixes. Dans ce cas les techniques de codage d'Intra-frame sont appropriées pour explorer la corrélation spatiale pour réaliser une compression de données efficace.

L'algorithme MPEG utilise une technique de codage sur des blocs de 8x8 pixel, pour analyser efficacement les corrélations spatiales entre pixels voisins de la même image.

Cependant, si la corrélation entre pixels dans des trames voisines est grande, c.-à-d. dans les cas où deux trames consécutives ont un contenu semblable ou identique, il est souhaitable d'utiliser la technique de codage d'inter-frame.

Dans le schéma de codage vidéo MPEG une combinaison adaptative entre les deux mouvements (temporel et spatial) de l'information est utilisée pour réaliser une grande compression de données.

Conteneur vidéo [100]

Un conteneur vidéo est un format de fichier permettant de rassembler en un seul fichier :

- un fichier vidéo ;
- un fichier audio ;
- d'autres données, par exemple : des méta-données (auteur, date, etc.), des sous-titres, etc.

Les pistes audio et vidéo peuvent être codées avec un grand nombre de *codecs*.

Les principaux conteneurs vidéo sont : *AVI*, *OGM Matroska*, Quicktime...

Digital rights management [107]

La gestion numérique des droits ou GDN (en anglais Digital Rights Management ou DRM, traduite parfois par « gestion des droits numériques » - et il s'agit là d'un contresens, car ce sont les contenus et leur gestion qui sont numériques, et non les droits eux-mêmes) a pour objectif de contrôler par des mesures techniques de protection l'utilisation qui est faite des œuvres numériques. La technique se voulant suffisante et nécessaire au contrôle, elle prévoit par exemple de :

- rendre impossible la consultation d'une œuvre hors de la zone géographique prévue (les zones des DVD) ;
- rendre impossible l'utilisation de matériel concurrent pour consulter une œuvre (incompatibilité des formats musicaux, dont iTunes) ;
- rendre impossible la consultation d'une œuvre selon ses préférences (désactivation de l'avance rapide sur certains passages publicitaires de DVD) ;
- limiter ou rendre impossible le transfert des œuvres d'un appareil à l'autre (limitation de la copie) ;
- rendre impossible l'extraction numérique de passage de l'œuvre.

Ces protections introduisent un problème majeur : elles créent une incompatibilité potentielle des fichiers protégés avec certains systèmes, certains logiciels ou certains matériels. Ainsi, les mesures de protection sur les fichiers musicaux mises en place par Microsoft interdisent la lecture de ces fichiers protégés sur iPod et vice-versa. Les autres systèmes d'exploitation que Windows risquent également de ne pas pouvoir accéder à ces fichiers protégés.

Ces protections introduisent également un second problème majeur : elles sont illimitées dans le temps. Alors qu'une édition originale de Voltaire est encore lisible aujourd'hui, il n'est pas du tout certain que les DRM seront lisibles plus de quelques années. Si Apple venait à disparaître, les morceaux protégés par DRM Apple pourraient devenir illisibles, ce qui placerait les consommateurs en position de créancier sans espoir de recours. De même, la possibilité de revendre les droits acquis n'existe en général pas, ce qui est une exception au principe de droit commun (on peut revendre ses disques et ses livres), sans parler enfin de ce qui arrivera le jour où l'œuvre tombera dans le domaine public. En pratique, le DRM correspond davantage à une location ou à un droit d'usage provisoire techniquement surveillé, qu'à une vente.

Le logiciel libre est également fortement menacé par ces mesures, puisque seul les lecteurs Microsoft (ou autres) seront capables de lire de tels fichiers médias.

Cela crée également un problème éthique :

- la copie privée, qui est un droit pour lequel le consommateur paye à chaque fois qu'il achète un support vierge (CD, DVD, cassette...), devient beaucoup plus difficile voire impossible (voir DADVSI).
- les DRM peuvent constituer une atteinte à la vie privée, ou du moins sont un pas de plus vers une informatique encore plus restrictive et surveillée.
- les fichiers téléchargés sont dépourvus de cette protection et donc plus sûrs, plus compatibles et d'usage plus étendu que les fichiers légaux mais sous DRM.

DivX [102]

DivX est un codec vidéo créé par DivXNetworks, Inc., connu pour sa capacité à compresser de longs et gros fichiers vidéo en fichiers bien plus légers et a été au centre de controverses à cause de son utilisation pour la copie et la distribution de films extraits de DVD aux droits réservés.

Un film sur support DVD occupe généralement 6 gigaoctets ; avec DivX, ce même film peut être compressé sur environ 600 à 1 400 mégoctets ce qui permet de le stocker sur 1 ou 2 cédéroms à la place. Pour une telle compression, la perte de qualité est pourtant minime, sauf éventuellement pour les scènes d'action rapide. Divers programmes sont disponibles pour créer un fichier DivX à partir d'un DVD. Le fichier peut ensuite être stocké sur un disque dur, gravé sur un CD-R ou un DVD-R, ou bien (en général, c'est illégal) partagé sur un réseau poste-à-poste.

DivX 3.11 et les versions précédentes réfèrent généralement à une version « hackée » du codec MPEG-4 de Microsoft. DivX a été créé aux alentours de 1999 par le Français Jérôme Rota (connu sous le pseudonyme de Gej). Le [codec](#) de Microsoft a été transformé pour permettre la compression de fichiers AVI - c'est le codec DivX 3.11. L'entreprise créée par Jérôme Rota, DivXNetworks, Inc., a par la suite produit une version du codec totalement indépendante du codec de Microsoft afin d'éviter des problèmes avec la firme de Redmond. DivXNetworks a demandé la dépose d'un brevet pour son nouveau codec, qui suit la certification de la norme MPEG-4.

Le codec DivX est disponible en téléchargement depuis le site de DivXNetworks, Inc. pour les systèmes d'exploitations Windows (XP, 2000, ME, 98), GNU/Linux et Mac OS X.

Ce n'est pas un logiciel libre et son code source n'est pas disponible, mais une version publique — appelée OpenDivX — a été mise à disposition par DivXNetworks au début de l'année 2001. Cette version a servi de base pour le codec (logiciel libre cette fois) [XviD](#), qui est à présent maintenu par un groupe indépendant.

Encodage numérique [92]

L'encodage (on dit parfois aussi codage) est le procédé qui consiste à transformer une source vidéo ou audio en un format informatique déterminé, à l'aide de [codecs](#).

Exemples de codecs audio : [MP3](#), MP3Pro, [OGG](#), [WMA](#)...

Exemples de codecs vidéo : [DivX](#), [XviD](#), RealVideo...

Ces formats permettent généralement une forte compression de données à caractère destructif et irréversible, dont on essaie de limiter au maximum la perception par l'être humain.

Entrelacement (en anglais : Interlace) [118]

Cette technique est issue de la diffusion des signaux vidéo par voie hertzienne dont la bande passante est plus limitée. Elle a aussi pour effet d'éliminer le scintillement.

Elle consiste à diviser l'image en 2 trames, 1 trame pour les lignes impaires (ou trame supérieure) et 1 trame pour les lignes paires (ou trame inférieure), puis de les afficher en les entrelaçant. La seconde trame étant affichée avec un décalage d'un 50e de seconde en PAL ou d'un 60e de seconde en NTSC. C'est la persistance de la trame précédente dans l'œil

qui donne l'illusion de voir des images complètes. Un léger lignage horizontal reste perceptible, car l'œil et le cerveau ne se laissent pas bernier si facilement.

Lors de l'enregistrement, une caméra PAL va enregistrer 50 trames pour 1 seconde de film et chaque trame va représenter 1/50e de l'action. Dans le cas de scènes rapides, la trame inférieure sera très légèrement différente de la trame supérieure ce qui va occasionner cet effet de peigne (fines bandes noires ou blanches qui apparaissent pendant les mouvements) lors de la conversion vidéo.

FLAC [116]

FLAC, acronyme de Free Lossless Audio Codec, est un codec de compression audio sans perte. À l'inverse de codecs tels que *MP3* ou *Vorbis*, il n'enlève aucune information du flux audio.

Comparaisons

FLAC se distingue d'algorithmes sans perte (tels que ZIP et gzip) en ce qu'il a été créé spécifiquement pour compresser des données audio. La méthode ZIP réduit la taille d'un fichier audio de qualité CD de 20 à 40%, alors que FLAC obtient des taux de 30 à 70%.

Bien que des codecs à perte puissent atteindre des ratios de 80-90%, voire plus, ils le font en éliminant des données du flux original. FLAC utilise une technique similaire, mais il ajoute également des données « résiduelles » permettant de restaurer l'original sans déformation.

FLAC est disponible pour pratiquement tous les *OS* existants.

GIF [114]

Le Graphics Interchange Format (littéralement « format d'échange de graphiques »), plus connu sous l'acronyme GIF, est un format d'image numérique couramment utilisé sur le World Wide Web.

Origine

GIF a été mis au point par CompuServe en 1987 pour permettre le téléchargement d'images en couleur. Ce format utilise l'algorithme de compression LZW, nettement plus efficace que l'algorithme RLE utilisé par la plupart des formats alors disponibles.

Caractéristiques

GIF supporte 16 777 216 nuances de couleur : 8 bits par composante *RVB*, soit 2^{24} nuances. GIF n'enregistre pas directement la couleur de chaque *pixel*. Pour chaque image, une palette de 2 à 256 couleurs est construite. Ensuite chaque pixel de l'image référence une entrée de la palette. Cette méthode limite donc à 256 le nombre maximal de couleurs différentes présentes dans une image. On parle de format 8 bits car à chaque pixel correspond un nombre inférieur à 256, donc représentable avec 8 bits. L'usage d'une palette permettait un affichage beaucoup plus rapide sur les ordinateurs de l'époque dont les cartes graphiques contenaient elle-même une palette d'au plus 256 couleurs.

La limitation à 256 couleurs n'est pas gênante pour les logos, les graphiques et la plupart des images synthétiques, ainsi que les photographies noir et blanc. En revanche une photographie couleur de qualité nécessite plus de nuances.

GIF permet de spécifier qu'une entrée de la palette est transparente. C'est notamment utile lorsqu'une image non rectangulaire est intégrée à un document comme une page Web : on voit le document à travers les pixels transparents. GIF propose un mode *entrelacé* permettant de commencer par transmettre quelques lignes d'une image, puis les lignes placées entre elles. Ce mode permet de donner plus rapidement un aperçu de l'image lorsque la transmission est lente.

Usage sur le Web

En 1989, le format GIF a été étendu (format GIF89a au lieu de GIF87a) pour permettre le stockage de plusieurs images dans un fichier. Ceci permet de créer des diaporamas, voire des animations si les images sont affichées à un rythme suffisamment soutenu. Chaque image d'une animation peut avoir sa propre palette.

En 1993, le [navigateur Web](#) NCSA Mosaic a été le premier à permettre l'intégration d'images aux pages Web : les formats GIF et XBM étaient supportés. Le support du format [JPEG](#), utile aux photographies, a été introduit en 1994 par Netscape Navigator.

En décembre 1994, Unisys, détenteur de brevets sur la compression LZW, a soudainement annoncé que les auteurs de logiciel produisant des images GIF devaient payer des royalties. Ceci a motivé le développement du format [PNG](#), basé sur la compression gzip libre et qui améliore toutes les fonctionnalités de GIF, sauf les animations pour lesquels le format MNG a été prévu.

10 ans plus tard, le format GIF est majoritairement utilisé sur le Web pour les images synthétiques, tandis que JPEG est utilisé pour les photographies et que Macromedia Flash tend à s'imposer pour les animations. Le format XBM est tombé dans l'oubli, bien que supporté par les navigateurs. Le support du format PNG par les navigateurs a été lent et émaillé de problèmes techniques, ce format est nettement moins utilisé que GIF.

Il est à noter que les brevets d'Unisys sont arrivés à expiration le 20 juin 2003 aux États-Unis, le 18 juin 2004 dans la plupart des pays d'Europe, le 20 juin 2004 au Japon et le 7 juillet 2004 au Canada.

Image numérique [108]

On désigne sous le terme d'image numérique toute image (dessin, icône, photographie...) acquise, créée, traitée, stockée sous forme binaire (suite de 0 et de 1) :

- Acquis par des dispositifs comme les [scanners](#), les appareils photo ou caméscopes numériques, les cartes d'acquisition vidéo (qui numérisent directement une source comme la télévision).
- Créée directement par des programmes informatiques, via la souris, les tablettes graphiques ou par la modélisation 3D (ce que l'on appelle par abus de langage les « images de synthèse »).
- Traitée grâce à des outils informatiques. Il est facile de la modifier en taille, en couleur, d'ajouter ou supprimer des éléments, d'appliquer des filtres variés, etc.
- Stockée sur un support informatique (disquette, [disque dur](#), [CD-ROM](#)...)

Types d'images

On distingue deux types d'images à la composition et au comportement différent :

- Images matricielles (ou images bitmap)

Elle est composée comme son nom l'indique d'une matrice (tableau) de points à plusieurs dimensions, chaque dimension représentant une dimension spatiale (hauteur, largeur, profondeur), temporelle (durée) ou autre (par exemple, un niveau de résolution).

Images 2D

Dans le cas des images à deux dimensions (le plus courant), les points sont appelés [pixels](#).

Ce type d'image s'adapte bien à l'affichage sur écran informatique (lui aussi orienté pixel) ; il est en revanche peu adapté pour l'impression, car la résolution des écrans informatiques, généralement de 72 à 96 ppp (« points par pouce », en anglais dots per inch ou dpi) est bien inférieure à celle atteinte par les imprimantes, au moins 300 ppp aujourd'hui. L'image imprimée, si elle n'a pas une haute résolution, sera donc plus ou moins floue ou laissera apparaître des pixels carrés visibles.

- Images vectorielles

Le principe est de représenter les données de l'image par des formules géométriques qui vont pouvoir être décrites d'un point de vue mathématique. Cela signifie qu'au lieu de mémoriser une mosaïque de points élémentaires, on stocke la succession d'opérations conduisant au tracé. Par exemple, un dessin peut être mémorisé par l'ordinateur comme « une droite tracée entre les points (x1,y1) et (x2,y2) », puis « un cercle tracé de centre (x3,y3) et de rayon 30 de couleur rouge ».

L'avantage de ce type d'image est la possibilité de l'agrandir indéfiniment sans perdre la qualité initiale, ainsi qu'un faible encombrement. L'usage de prédilection de ce type d'images concerne les schémas qu'il est possible de générer avec certains logiciels de CAO (Conception Assistée par Ordinateur) comme AutoCAD. Ce type d'images est aussi utilisé pour les animations Flash, utilisées sur Internet pour la création de bannières publicitaires, l'introduction de sites web, voire des sites web complets.

Résolution

La résolution d'une image est définie par un nombre de pixels par unité de longueur de la structure à numériser (classiquement en ppp). Ce paramètre est défini lors de la numérisation (passage de l'image sous forme binaire), et dépend principalement des caractéristiques du matériel utilisé lors de la numérisation. Plus le nombre de pixels par unité de longueur de la structure à numériser est élevé, plus la quantité d'information qui décrit cette structure est importante et plus la résolution est élevée. La résolution d'une image numérique définit le degré de détail de l'image. Ainsi, plus la résolution est élevée, meilleure est la restitution.

Cependant, pour une même dimension d'image, plus la résolution est élevée, plus le nombre de pixels composant l'image est grand. Le nombre de pixels est proportionnel au carré de la résolution, étant donné le caractère bidimensionnel de l'image : si la résolution est multipliée par deux, le nombre de pixels est multiplié par quatre. Augmenter la résolution peut entraîner des temps de visualisation et d'impression plus longs, et conduire à une taille trop importante du fichier contenant l'image et à de la place excessive occupée en mémoire.

La résolution de qualité d'image se distingue de la résolution du format de l'image, correspondant au nombre de pixels qui compose l'image en hauteur (axe vertical) et en largeur (axe horizontal) : 200 pixels par 450 pixels par exemple, abrégé en « 200x450 ».

Représentation des couleurs

Il existe plusieurs modes de représentation numérique de la couleur, le plus utilisé pour le maniement des images est l'espace colorimétrique Rouge, Vert, Bleu (*RVB* ou *RGB*). Cet

espace est basé sur une synthèse additive des couleurs, c'est-à-dire que le mélange des trois composantes R, V, et B à leur valeur maximum donne du blanc, à l'instar de la lumière. Le mélange de ces trois couleurs à des proportions diverses permet quasiment de reproduire à l'écran toutes les couleurs du spectre visible, sans avoir à spécifier une multitude de fréquences lumineuses.

Il existe d'autres modes de représentation des couleurs : Cyan, Magenta, Jaune, Noir (CMJN ou CMYK) utilisé principalement pour l'impression, et basé sur une synthèse soustractive des couleurs.

Les images bitmap en couleurs peuvent être représentées soit par une image dans laquelle la valeur du pixel est une combinaison linéaire des valeurs des trois composantes couleurs, soit par trois images représentant chacune une composante couleur. Dans le premier cas, selon le nombre de bits (unité d'information élémentaire qui peut prendre deux valeurs distinctes) alloués pour le stockage d'une couleur de pixel, on distingue généralement les différents types d'images suivants :

Images 24 bits (ou « couleurs vraies »)

Il s'agit d'une appellation trompeuse car le monde numérique (fini, limité) ne peut pas rendre compte intégralement de la réalité (infinie). Le codage de la couleur est réalisé sur trois octets, chaque octet représentant la valeur d'une composante couleur par un entier de 0 à 255. Ces trois valeurs codent généralement la couleur dans l'espace RVB. Le nombre de couleurs différentes pouvant être ainsi représenté est de $256 \times 256 \times 256$ possibilités, soit près de 16 millions de couleurs. Comme la différence de nuance entre deux couleurs très proches mais différentes dans ce mode de représentation est quasiment imperceptible pour l'œil humain, on considère commodément que ce système permet une restitution exacte des couleurs, c'est pourquoi on parle de « couleurs vraies ».

R	V	B	Couleur	
0	0	0	noir	
11	11	11	nuance de noir	
255	0	0	rouge	
0	255	0	vert	
0	0	255	bleu	
128	128	128	gris	
255	255	255	blanc	

Les images bitmap basées sur cette représentation peuvent rapidement occuper un espace de stockage considérable, chaque pixel nécessitant trois octets pour coder sa couleur.

Images à palettes, images en 256 couleurs (8 bits)

Pour réduire la place occupée par l'information de couleur, on utilise une palette de couleurs « attachée » à l'image. On parle alors de couleurs indexées : la valeur associée à un pixel ne véhicule plus la couleur effective du pixel, mais renvoie à l'entrée correspondant à cette valeur dans une table (ou palette) de couleurs appelée look-up table ou LUT en anglais, dans laquelle on dispose de la représentation complète de la couleur considérée.

Selon le nombre de couleurs présentes dans l'image, on peut ainsi gagner une place non négligeable : on considère en pratique que 256 couleurs parmi les 16 millions de couleurs 24 bits sont suffisantes. Pour les coder, on aura donc une palette occupant 24 bits x 256

entrées, soit 3 x 256 octets, et les pixels de l'image seront associés à des index codés sur un octet. L'occupation d'une telle image est donc de 1 octet par pixel plus la LUT, ce qui représente un peu plus du tiers de la place occupée par une image en couleurs 24 bits (plus l'image contient de pixels, plus le gain de place est important, la limite étant le tiers de la place occupée par l'image en couleurs vraies).

Une autre méthode existante consiste à se passer de palette, et de coder directement les trois couleurs en utilisant un octet : chaque composante couleur est codée sur deux bits, le bit restant peut servir soit à gérer plus de couleurs sur une des composantes, soit à gérer la transparence du pixel. Avec cette méthode, on obtient des images bitmap avec un codage couleur effectivement limité à 8 bits, bien que la plage des couleurs possibles soit très réduite par rapport à celle qu'offre la méthode utilisant une palette.

Dans le cas des images en couleurs indexées, il est possible de spécifier que les pixels utilisant une des couleurs de la palette ne soient pas affichés lors de la lecture des données de l'image. Cette propriété de transparence est très utilisée (et utile) pour les images des pages web, afin que la couleur de fond de l'image n'empêche pas la visualisation de l'arrière-plan de la page.

Images en teintes (ou niveaux) de gris

On ne code ici plus que le niveau de l'intensité lumineuse, généralement sur un octet (256 valeurs). Par convention, la valeur zéro représente le noir (intensité lumineuse nulle) et la valeur 255 le blanc (intensité lumineuse maximale).

Ce procédé est fréquemment utilisé pour reproduire des photos en noir et blanc ou du texte dans certaines conditions (avec utilisation d'un filtre pour adoucir les contours afin d'obtenir des caractères plus lisses).

Ce codage de la simple intensité lumineuse est également utilisé pour le codage d'images couleurs : l'image est représentée par trois images d'intensité lumineuses, chacune se situant dans une composante distincte de l'espace colorimétrique (par exemple, intensité de rouge, de vert et de bleu).

Images avec gestion de la translucidité

On peut attribuer à une image un canal supplémentaire, appelé canal alpha, qui définit le degré de transparence de l'image. Il s'agit d'un canal similaire aux canaux traditionnels définissant les composantes de couleur, codé sur un nombre fixe de bits par pixel (en général 8 ou 16). On échelonne ainsi linéairement la translucidité d'un pixel, de l'opacité complète à la transparence.

Formats d'images

Un format d'image est une représentation informatique de l'image, associée à des informations sur la façon dont l'image est codée et fournissant éventuellement des indications sur la manière de la décoder et de la manipuler.

La plupart des formats sont composés d'un en-tête contenant des attributs (dimensions de l'image, type de codage, LUT, etc.), suivi des données (l'image proprement dite). La structuration des attributs et des données diffère pour chaque format d'image. De plus, les formats actuels intègrent souvent une zone de métadonnées (metadata en anglais) servant à préciser les informations concernant l'image comme :

- la date, l'heure et le lieu de la prise de vue,
- les caractéristiques physiques de la photographie (sensibilité ISO, vitesse d'obturation, usage du flash...)

Ces métadonnées sont par exemple largement utilisées dans le format EXIF (extension du format JPEG), qui est le format le plus utilisé dans les appareils photo numériques.

Quelques précautions à prendre concernant les formats d'images :

- les formats dits « propriétaires », peuvent différer selon le logiciel qui les manipule. De plus, leur pérennité n'est pas garantie : réaliser de nouveaux programmes pour les lire peut s'avérer difficile (surtout si leurs spécifications n'ont pas été rendues publiques), cela peut même s'avérer illégal si les algorithmes utilisés sont protégés par des brevets.
- Il faut prêter attention aux différentes versions que peut recouvrir un format particulier. Notamment pour le format TIFF qui varie selon les versions ; certaines d'entre elles ne sont pas reconnues par certains logiciels.

Formats propriétaires

Le format **TIFF** est considéré comme un format propriétaire, le brevet étant contrôlé par la firme Aldus.

Dans le passé, le format **GIF** était soumis au brevet Unisys contrôlé par la société CompuServe, c'était donc un format propriétaire. Mais il est à noter que les brevets d'Unisys sont arrivés à expiration le 20 juin 2003 aux États-Unis, le 18 juin 2004 dans la plupart des pays d'Europe, le 20 juin 2004 au Japon et le 7 juillet 2004 au Canada. Il est donc devenu depuis un format libre de droits.

Image numérique et droits d'auteur

Pour tenter de faire respecter le droit d'auteur (en France) et le copyright (dans presque tous les autres pays), il existe des techniques de marquage numérique d'une image. Ces techniques, que l'on nomme empreinte, sont de plus en plus utilisées. L'empreinte est supposée conserver une preuve de l'origine de l'image, sous la forme d'une signature visible ou invisible, qui doit résister aux traitements susceptibles d'être appliqués à l'image. Ce « tatouage » peut se faire selon deux méthodes, généralement désignées par le même terme de filigrane :

Protection par signature visible

Cette technique consiste à intégrer une indication sur l'image, par exemple l'organisme ou l'auteur à qui appartient l'image, afin de dissuader les pirates de s'en servir. L'inconvénient de cette méthode est qu'il est très facile d'éliminer ce type de tatouage avec un outil de traitement d'images, puisque le tatouage est visible.

Protection par signature cryptée

Cette technique consiste à cacher le tatouage dans les données de l'image. Cette approche a l'avantage de ne pas gêner la lecture de l'image par le simple spectateur tout en permettant une facile identification. L'auteur en tire un avantage complémentaire : l'éventuel pirate inattentif ne sera pas tenté de retirer ou modifier la signature ; le pirate plus volontaire verra son activité illégale rendue un peu plus difficile ou facilement prouvable (par la seule présence du tatouage).

JPEG ou JPG [109]

Le JPEG est un format de compression de données numériques utilisés pour les images. C'est un acronyme de Joint Photographic Expert Group qui lui-même vient du nom d'un comité créé en 1986, résultant de la fusion de plusieurs groupes (d'où le joint) de professionnels de l'industrie de l'image. Ce comité a donné son nom à la norme ouverte de compression d'images numériques JPEG. Ensuite cette norme a donné son nom au format de données définis et au format de fichier le plus utilisé pour contenir ces données.

Généralement le terme JPEG fait référence à la norme JPEG, formellement ISO/CEI 10918-1 ou UIT-T Recommandation T.81. Cette norme décrit une **méthode de compression** basée sur la transformée en cosinus discrète. Elle spécifie uniquement la transformation d'une image brute en une suite de bits. Elle ne spécifie donc pas directement comment stocker les informations sur ses dimensions, son auteur, etc. Ceci est le rôle d'un format de fichier.

Le JPEG est un format à perte, qui élimine donc des informations, mais un des points forts de JPEG est que son taux de compression est réglable. Un compromis doit cependant être

fait entre le taux de compression et la qualité de l'image compressée. En d'autres termes, le taux de compression ne doit pas être trop élevé, ni l'opération de compression être trop souvent répétée, sous peine de nuire gravement à la qualité générale de l'image. Certains logiciels offrent plusieurs choix pré-programmés de compression et d'autres permettent de l'affiner très précisément.

MP3 [94]

Présentation

Le MP3 est l'abréviation de MPEG-1/2 Audio Layer 3, la spécification sonore du standard *MPEG1* (Motion Picture Expert Group).

L'extension d'un fichier audio compressé au format mp3 est .mp3

Utilité

MP3 (ou, plus précisément, MPEG-1/2 Audio Layer 3) est un algorithme de compression capable de réduire drastiquement la quantité de données nécessaire pour restituer de l'audio, mais qui, pour l'auditeur, ressemble à une reproduction du son original non compressé, c'est-à-dire avec perte significative mais acceptable de qualité sonore pour l'oreille humaine.

Ce format populaire de compression audio permet une compression approximative de 1:4 à 1:12. Un fichier audio occupe ainsi 4 à 12 fois moins d'espace une fois transcodé en format MP3. Une spécificité intéressante qui facilite le téléchargement et permet d'engranger énormément de données musicales sur un disque dur ou une mémoire flash.

Technique de l'encodage

La compression peut être plus grande que ce ratio en choisissant un débit binaire (en anglais bitrate) plus faible. On considère en général qu'il faut au moins 128 kilobits par seconde (kbit/s) pour bénéficier d'une qualité audio acceptable pour un morceau de musique, sachant que 8 kbit/s est la qualité audio d'un téléphone.

- Ce format de données utilise un modèle psycho-acoustique, c'est-à-dire qu'il s'agit d'un système de compression avec perte. Il ne retransmet pas intégralement le spectre des fréquences audio, l'oreille n'étant de toute façon pas capable de distinguer de trop petites différences.
- Le MP3 supprime notamment les ultrasons et autres fréquences que notre oreille ne peut entendre (mais comme ils ne sont pas présents dans un CD audio il ne supprime que très peu de choses à cet endroit). C'est en général une légende urbaine que la suppression des fréquences inaudibles pour une source venant d'un CD audio.
- De plus, il exploite le mécanisme psycho-acoustique de « masque » : si l'on écoute des gazouillis d'oiseaux, et que soudainement, un coup de klaxon retentit, les gazouillis des oiseaux nous deviennent imperceptibles. Cette information peut donc être supprimée.
- En théorie, on ne devrait pas pouvoir remarquer si un fichier a été compressé en MP3 ou non du moins le but étant de s'en rapprocher. Mais en pratique, diverses formes de compression existent, avec des qualités variables. Si l'on compresse bien un fichier à plus de 128 kbit/s à l'aide d'un bon programme, la différence sera presque inaudible. A contrario, si on le fait à moins de 128 kbit/s, ou avec un mauvais programme, il y aura des défauts perceptibles.

Mais d'une manière générale on admet que la compression avec perte entraîne une différence de qualité audible vis à vis des compressions sans perte.

- On peut améliorer la qualité à débit moyen égal en utilisant un débit binaire variable (VBR ou Variable Bit Rate par opposition à un débit constant CBR, Constant Bit Rate). Dans ce cas, les instants peu complexes comme les silences seront codés à un taux plus faible, par exemple 64 kbit/s, laissant ainsi une réserve de bits supplémentaires

pour coder les parties plus problématiques comportant des sauts de fréquences rapides jusqu'à 320 kbit/s.

Étiquettes

Outre le fait de stocker la musique de façon très compacte tout en conservant une qualité acceptable, le MP3 apporte une fonctionnalité rarement présente sur les formats audio qui l'ont précédé : les métadonnées (données sur les données). En clair, le fichier mp3 ne contient pas seulement la musique mais peut également apporter des informations sur celles-ci (telles que l'interprète, le titre, le nom de l'album, voire la pochette, les paroles ou du karaoké). Ces informations sont stockées sous forme d'étiquettes (tag en anglais) dont il existe plusieurs versions.

Le format MP3 initial ne permettait pas de stocker des étiquettes, tout au plus, il permettait de préciser certains paramètres binaires comme le fait que le morceau soit protégé ou non par copyright ou le fait qu'il s'agisse d'un original ou d'une copie.

Les étiquettes MP3 sont enregistrées au format ID3 (version 1 ou 2).

Licence

Bien que le MP3 soit souvent perçu par l'utilisateur final comme une technologie gratuite (il peut en effet encoder ou décoder gratuitement sa musique de manière tout à fait légale pour peu que l'enregistrement original lui appartienne), cette technologie fait l'objet d'une licence. L'algorithme « MPEG-1 Layer 3 » décrit dans les standards ISO/IEC IS 11172-3 et ISO/IEC IS 13818-3 est soumis à des royalties à Fraunhofer IIS et Thomson (les détenteurs du brevet) pour toute utilisation commerciale ou implémentation physique (notamment sur les baladeurs mp3).

Alternatives

La popularité du format MP3 a rapidement conquis de très nombreux utilisateurs tant par sa facilité d'utilisation que par le fait que pour la première fois, elle permettait de transmettre de l'information multimédia par internet. Néanmoins, les limites de cette technologie aussi bien quantitativement (taux de compression donc taille des fichiers et temps de téléchargement) que qualitativement (perte de qualité par rapport à l'enregistrement non compressé, gestion numérique des droits) ont motivé plusieurs initiatives proposant des alternatives :

- MP3-pro développé par Fraunhofer IIS et Thomson
- Advanced Audio Coding développé par les laboratoires Dolby et exploité par Apple pour l'iPod et le Musicstore
- [Ogg Vorbis](#), une solution libre développée par Xiph.org
- [Windows Media Audio](#) développé par Microsoft.

MPEG-1 [95]

Le MPEG-1 est une norme de compression pour la vidéo numérique, élaborée par le groupe MPEG en 1988. Ce groupe a pour but de développer des standards internationaux de compression, décompression, traitement et codage d'image animées et de données audio.

La norme MPEG-1 représente chaque image comme un ensemble de blocs 16 x 16. Elle permet d'obtenir une résolution de:

- 352x240 à 30 images par seconde en NTSC
- 352x288 à 25 images par seconde en PAL/SECAM

Le MPEG-1 permet d'obtenir des débits de l'ordre de 1,2 Mbit/s (exploitable sur un lecteur de CD-ROM).

Le MPEG-1 permet d'encoder une vidéo grâce à plusieurs techniques :

- Intra coded frames (Frames I, correspondant à un codage interne) : les images sont codées séparément sans faire référence aux images précédentes
- Predictive coded frames (frames P ou codage prédictif) : les images sont décrites par différence avec les images précédentes
- Bidirectionally predictive coded frames (Frames B) : les images sont décrites par différence avec l'image précédente et l'image suivante
- DC Coded frames : les images sont décodées en faisant des moyennes par bloc

Le MPEG-1, comme ses descendants [MPEG-2](#) ou [MPEG-4](#) comporte plusieurs parties, dont la partie vidéo (Part.2), et la partie audio (Part.3). Cette partie audio se décompose en 3 couches (layers) de complexité et d'efficacité de compression croissantes. La couche MPEG-1 Audio Layer 3, la plus efficace donc, a donné naissance au format de compression audio [MP3](#) (à ne pas confondre avec MPEG-3).

MPEG-2 [97]

MPEG-2 est la norme de seconde génération (1994) du Motion Picture Experts Group qui fait suite à [MPEG-1](#). MPEG-2 définit les aspects compression de l'image et transport à travers des réseaux pour la télévision numérique.

Les aspects transport sont définis dans la norme ISO/CEI 13818-1 (Codage générique des images animées et du son associé - Partie Systèmes). Les aspects compression, quant à eux, sont définis dans les normes ISO/CEI 13818-2, 3 et 4 (Codage générique des images animées et du son associé - Parties vidéo, audio).

Ce format vidéo est utilisé pour les DVD, CVD et SVCD avec différentes résolutions d'image. Ce format est également utilisé dans la diffusion de télévision numérique par satellite, câble, réseau de télécommunication ou hertzien (TNT) et avec des petites modifications sur les DVD commerciaux.

MPEG-3 [98]

MPEG-3 désigne un ensemble de normes audio et vidéo introduites par MPEG (Motion Picture Expert Group). MPEG-3 a été conçu pour la prise en charge des signaux HDTV à des débits de 20 à 40Mbits/s.

Il a vite été observé que des résultats semblables pouvaient être obtenus par de légères modifications de [MPEG-2](#). Perdant de son intérêt, MPEG-3 a été abandonné.

MPEG-3 ne doit pas être confondu avec MPEG-1 Part 3 Layer 3 (ou MPEG-1 Audio Layer 3), plus connu sous le nom de [MP3](#).

MPEG-4 [99]

MPEG-4, introduit en 1998, est la désignation pour un groupe de normes de codage audio et vidéo acceptées par MPEG (Motion Picture Experts Group). MPEG-4 est d'abord conçu pour gérer le contenu en bas-débit, depuis 4800 bit/s jusqu'à approximativement 4 Mbit/s. L'usage principal est le web (internet) en streaming, ou flux vidéo, le CD, le vidéophone et même la télévision.

MPEG-4 reprend plusieurs des fonctions de [MPEG-1](#) et [MPEG-2](#), en ajoutant de nouvelles comme les graphismes en 2 dimensions, et les mondes virtuels en 3 dimensions, le support pour la gestion des droits numériques et plusieurs types d'interactivités.

La plupart des fonctions proposées par MPEG4 sont facultatives. Ceci signifie qu'il n'y a sans doute pas beaucoup d'implémentations exhaustives de la norme MPEG-4. Le standard offre donc la possibilité de regrouper une partie seulement de ces fonctions (profiles, level), selon le domaine d'application.

Ogg [112]

Ogg est le nom du principal projet de la fondation Xiph.Org dont le but est de proposer à la communauté des formats et codecs multimédias ouverts, libres et dégagés de tout brevet.

C'est aussi le nom du format de fichier conteneur proposé par ce même projet. Il existe d'autres conteneurs libres développés par d'autres projets, tels MCF ou son rejeton Matroska.

.ogg (prononcer « augue ») est une des extensions possibles pour les fichiers au format Ogg. Par abus de langage, on appelle couramment « fichier Ogg » un fichier audio au format Ogg contenant des données audio compressées en Ogg Vorbis, l'un des codecs du projet Ogg.

Les principaux travaux du projet Ogg sont les suivants :

- le format de fichier conteneur Ogg, qui peut contenir des pistes audio (en général Ogg Vorbis), vidéo (en général Ogg Theora) et texte (sous-titres). Il peut y avoir plusieurs pistes de chaque type pour, par exemple, proposer des médias multilingues. L'extension habituelle du format Ogg est .ogg1, que le fichier contienne uniquement de l'audio, ou de l'audio et de la vidéo ;
- le format de compression audio déjà populaire Ogg Vorbis et les codecs associés ;
- le format de compression vidéo Ogg Theora dont le codec est basé sur celui de VP3, libéré par la société On2 Technologies ;
- le format de compression audio sans pertes [FLAC](#).

Pixel [111]

Le pixel ou point est l'unité de base d'une image numérique. Son nom provient de l'expression anglaise picture element, c'est-à-dire, « élément d'image » ou « point élémentaire ».

C'est le point minimal adressable par le contrôleur vidéo. Par exemple, pour les résolution d'affichage :

- la résolution du VGA est de 640 x 480, soit 307 200 points ;
- la résolution du Super-VGA est de 800 x 600, soit 480 000 pixels ;
- la résolution du XGA est de 1 024 x 768, soit 786 432 pixels.

Chaque Pixel est en fait composé d'une triade de composants électroluminescents, rendants des tons rouge, vert et bleu ([RVB](#)) une fois bombardés par le canon à électron du tube cathodique.

PNG [115]

Le PNG (Portable Network Graphics) est un format d'images numériques libre de droit, qui a été créé pour remplacer le format propriétaire [GIF](#), dont la compression était soumise à un brevet. Le PNG est un format non destructeur spécialement adapté pour publier des images simples comprenant des aplats de couleurs.

Utilisation

- Pour les images synthétiques
PNG est particulièrement approprié lorsqu'il s'agit d'enregistrer des images synthétiques destinées au Web comme des graphiques, des icônes, des images représentant du texte (bonne conservation de la lisibilité), ou des images avec peu de dégradés. Le PNG surpasse régulièrement le format GIF tant en ce qui concerne la taille (avec une palette de couleurs bien choisie) que la qualité puisqu'il n'est pas limité à 256 couleurs.
- Pas pour les photos
Les caractéristiques de PNG lui permettent certes d'enregistrer des photographies sans perte de données, mais la taille du fichier résultant reste très supérieure à celle de formats spécifiques aux photographies comme [JPEG](#) ou JPEG2000. Il n'est donc pas destiné à cet usage.

Détails sur le format

PNG permet principalement d'enregistrer les images bitmap sous trois formes différentes :

- 8 bits en niveaux de gris (256 niveaux)
- 8 bits permettant de choisir parmi une palette de maximum 256 couleurs contenues dans le fichier (équivalent au format GIF)
- 24 bits en 16 millions de couleurs (couleurs vraies)

Après l'application d'un filtre prédictif qui permet généralement d'obtenir de plus hauts niveaux de compression, le tout est compressé sans pertes suivant l'algorithme deflate (RFC 1951 [1]), généralement avec gzip.

Les composantes des pixels ou les entrées de palette sont données soit au format [RVB](#) (rouge, vert, bleu), soit au format RVBA (avec un canal alpha supplémentaire pour la transparence). Dans ce cas, 8 bits supplémentaires sont utilisés par pixel ou par entrée de palette, ce qui fait 16 bits pour une image en niveaux de gris et 32 bits pour une image en couleurs.

La transparence

La présence d'un canal alpha définissant différents niveaux de transparence le rend idéal pour la composition sur les pages Web. Cette caractéristique est bien implémentée par la majorité des navigateurs Web actuels (2004) à l'exception d'Internet Explorer 6.

La transparence

Lorsque l'image PNG utilise une palette de 256 couleurs maximum, il n'est alors possible d'utiliser qu'un seul niveau de transparence (totalement transparent ou totalement opaque).

C'est le même comportement qu'avec le format GIF et cela fonctionne même avec Internet Explorer 6. Par conséquent, les images Web au format GIF peuvent être converties en cette version de PNG sans crainte d'incompatibilité avec la majorité des navigateurs Web actuels (2004), et avec l'avantage d'une taille de fichier souvent réduite et sans souci de brevet (le brevet GIF est néanmoins tombé en 2004 dans le domaine public).

Autres comparaisons avec GIF

Le PNG, d'ailleurs parfois appelé par boutade PNG's Not GIF (PNG n'est pas GIF), peut faire tout ce que le format GIF peut faire et même plus, comme la transparence. Il ne permet cependant pas de faire des images animées, mais le format dérivé MNG a été créé par ses auteurs pour pallier ce manque.

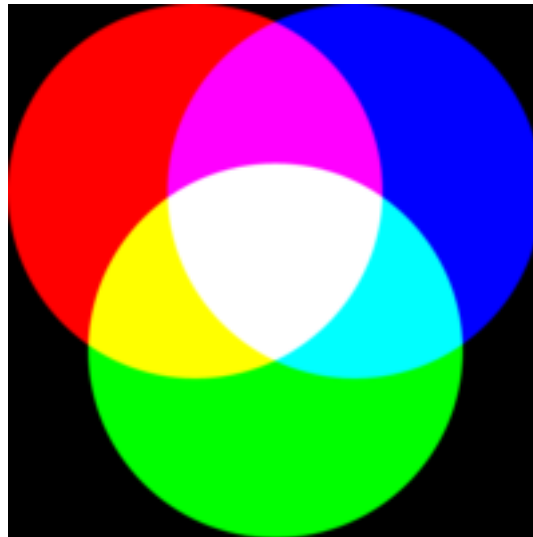
RVB ou RGB [117]

RVB (rouge vert bleu) ou RGB en anglais (red green blue), est un format de codage des couleurs.

Ces trois couleurs sont les couleurs primaires en synthèse additive. Elles correspondent en fait aux trois longueurs d'ondes auxquelles répondent les trois types de cônes de l'œil humain. Additionnées, elles permettent d'obtenir du blanc, lumière parfaite pour l'œil humain.

Elles sont utilisées en vidéo, pour l'affichage sur les écrans, et dans les logiciels d'imagerie.

Couplées deux à deux ces couleurs donnent les couleurs cyan, magenta et jaune secondaires en synthèse additive et primaires en synthèse soustractive.



On parle de synthèse soustractive quand — pour construire une couleur — on utilise plusieurs couleurs de base (le plus souvent le cyan, le magenta et le jaune) en retirant directement leurs couleurs de la lumière (par exemple avec des filtres colorés). En imprimerie, on ajoute le noir aux trois couleurs de base pour des raisons d'économie (une seule couleur au lieu de trois, et de surcroît l'une des moins chères à fabriquer), ainsi que pour améliorer le contraste des images et celui du texte simple. Le noir du texte doit être pur, c'est-à-dire, que les composantes CMJ doivent être à 0. Si ce n'est pas le cas un léger décalage lors des aplats de couleur chez l'imprimeur produira un effet de flou très désagréable autour du texte. Ce système de couleur est nommé CMJN (Cyan-Magenta-Jaune-Noir) ou CMYK (Cyan-Magenta-Yellow-black).

Les systèmes de photo sur papier chimiques utilisent une synthèse soustractive.

Streaming [121]

Le streaming est un principe utilisé principalement pour l'envoi de contenu en « direct » (ou en léger différé). Très utilisé sur Internet, il permet la lecture d'un flux audio ou vidéo à mesure qu'il est diffusé. Il s'oppose ainsi à la diffusion par téléchargement qui nécessite par exemple de récupérer l'ensemble des données d'un morceau ou d'un extrait vidéo avant de pouvoir l'écouter ou le regarder.

Principe

Le programme lecteur de contenu, ou client, streaming va récupérer une partie du contenu qu'il met dans une mémoire tampon (dite buffer). Lorsque le programme estime qu'il a suffisamment de données dans sa mémoire tampon pour lui permettre de lire le contenu

audio ou vidéo sans accroche, même en cas de petit ralentissement réseau, la lecture démarre, et le téléchargement du flux continue.

En raison des latences créées par le réseau (Internet ou LAN) et des opérations de codage/décodage effectuées, un délai de 5 à 40 secondes (voire plus) peut intervenir entre le signal émis par la source et le signal reçu sur le lecteur.

Applications

Les webradios sont un exemple pratique d'utilisation du streaming.

Les webTVs utilisent aussi la technologie du streaming que ce soit en direct ou pour des fichiers à la demande.

Le streaming video était à l'origine un format non enregistrable sur le PC qui y accède mais des logiciels comme Net Transport permettent maintenant d'enregistrer la video sur son disque dur.

TIFF [110]

Le Tag(ged) Image File Format généralement abrégé TIFF est un format de fichier pour *image numérique*.

Usages

TIFF est un format extrêmement flexible qui fait que TIFF est utilisé dans des applications très diverses, des *scanners* industriels aux appareils photo numériques en passant par les imprimantes. En revanche cela fait également qu'il n'existe pas de logiciel capable d'afficher n'importe quelle image TIFF. En outre il est possible de créer des extensions propriétaires et secrètes à TIFF.

Histoire

TIFF a été développé par Microsoft et Aldus. Aldus a été racheté par Adobe Systems. Depuis Adobe possède les droits sur le texte de la spécification TIFF et la marque TIFF. La révision 6.0 de TIFF date du 3 juin 1992.

Vidéo CD [96]

Le Vidéo CD ou VCD est un format standard de stockage vidéo sur disque compact nommé habituellement CD. Un vidéo CD est jouable sur un lecteur dédié, sur un ordinateur personnel (PC), mais surtout sur la plupart des lecteurs DVD de salon.

Spécifications techniques

La résolution du VCD est de 352x240 pixels pour la norme NTSC et de 352x288 pixels pour la norme PAL, soit approximativement un quart seulement de la résolution maximale d'une TV. La partie vidéo du VCD est encodée au format *MPEG-1*; la partie audio au format MPEG Layer 2 (MP2). Les débits maximum sont de l'ordre de 1150 kbit/s pour la vidéo, et 224 kbit/s pour l'audio. La qualité générale de l'image est souvent comparée à celle de la vidéo VHS.

La durée maximale d'enregistrement sur un VCD est d'environ 74 minutes (la même que celle d'un disque compact audio classique).

Le SVCD (Super-VCD) est un standard amélioré du VCD, qui utilise le format de compression *MPEG-2* à débit variable et qui permet d'obtenir une meilleure qualité vidéo.

Adoption

Les VCDs commerciaux n'ont jamais été diffusés aux États-Unis ou en Europe, mais ont en revanche été très populaires en Asie à cause essentiellement de leur prix modique. Ils sont toutefois utilisés par les vidéaste-amateurs comme moyen de stockage puisqu'un simple ordinateur muni d'un graveur de CD permet d'en fabriquer.

WAV [105]

WAV (ou WAVE), une contraction de WAVEform audio format, est un standard pour stocker l'audio digitale de Microsoft et IBM. C'est le format le plus courant pour l'audio non-comprimé sur les plateformes de Microsoft, mais il est bien courant sur les systèmes GNU/Linux aussi.

WMA [106]

Windows Media Audio aussi appelé WMA est un format de compression audio de type « lossy » (avec perte). Le format WMA offre pour spécificité la possibilité de protéger dès l'encodage les fichiers de sortie par une technique nommée *Digital Rights Management* (ou DRM).

Pour réduire de plus en plus la taille d'un fichier, on n'utilise plus le Constant BitRate (ou CBR), mais le Variable BitRate (ou VBR). Le principe est simple, lorsque l'encodeur juge qu'on peut utiliser un bitrate plus faible à certains moments dans la piste, il diminue le bitrate.

Avec le WMA, on trouve les VBR quality 98, 90, 75, 50, 25, 10. Les chiffres n'ont aucun rapport avec le bitrate. C'est en fait le pourcentage de qualité « théorique » par rapport au fichier original.

- 98 correspond environ à un CBR de 320 kbits/s
- 90 correspond environ à un CBR de 192 kbits/s
- 75 correspond environ à un CBR de 112-128 kbits/s
- 50 correspond environ à un CBR de 64-80 kbits/s
- En dessous, pour de la musique, la qualité est très moyenne.

WMV [113]

Codec vidéo développé par Microsoft. très utilisé en *streaming*. Sa version Haute Définition a été désignée par le DVD-Forum pour être implantée sur les futur HD-DVD de Nec et Blu-Ray Disc de Sony, qui sont les probables remplaçants du DVD vidéo. Principaux défauts: son coût, l'incompatibilité avec des systèmes d'exploitation autres que Windows.

XviD [103]

XviD est un codec vidéo compatible *MPEG-4* et distribué sous licence publique générale GNU (GPL). À l'origine basé sur OpenDivX, XviD fut développé par un groupe de volontaires après que les sources de OpenDivX ne soient plus disponibles.

Certaines de ces propriétés sont protégées par des brevets logiciels (notamment aux États-Unis d'Amérique et au Japon). À cause de cela, les version 0.9.x de XviD ont été déclarées illégales dans les pays où les brevets de ce type ont cours (ce n'est pas le cas en Europe). Dans les versions 1.0.x, XviD est distribué sous GNU/GPL sans restriction géographique.

Le principal concurrent de XviD est *DivX*. Alors que XviD est OpenSource (entièrement gratuit et modifiable), DivX est distribué comme gratuitiel (juste les fichiers binaires et il est illégal de les modifier) ou alors en version pro, payante.



Références

Ce support est publié sous licence FDL GNU. Voir licence de documentation libre GNU d'après l'article de Wikipédia.

1. <http://fr.wikipedia.org/wiki/Accueil>
2. <http://fr.wikipedia.org/wiki/Ordinateur>
3. http://fr.wikipedia.org/wiki/Clavier_informatique
4. http://fr.wikipedia.org/wiki/Scanner_de_document
5. http://fr.wikipedia.org/wiki/Souris_%28informatique%29
6. http://fr.wikipedia.org/wiki/Moniteur_d%27ordinateur
7. <http://fr.wikipedia.org/wiki/Imprimante>
8. http://fr.wikipedia.org/wiki/Clef_USB
9. <http://fr.wikipedia.org/wiki/DVD-Rom>
10. http://fr.wikipedia.org/wiki/Disque_compact
11. http://fr.wikipedia.org/wiki/Disque_dur
12. http://fr.wikipedia.org/wiki/Loi_de_Moore
13. http://fr.wikipedia.org/wiki/Carte_graphique
14. http://fr.wikipedia.org/wiki/Carte_m%C3%A8re
15. <http://fr.wikipedia.org/wiki/Microprocesseur>
16. http://fr.wikipedia.org/wiki/M%C3%A9moire_RAM
17. <http://fr.wikipedia.org/wiki/Chipset>
18. http://fr.wikipedia.org/wiki/M%C3%A9moire_cache
19. http://fr.wikipedia.org/wiki/Syst%C3%A8me_d%27exploitation
20. <http://fr.wikipedia.org/wiki/Bluetooth>
21. http://fr.wikipedia.org/wiki/Bus_informatique
22. http://fr.wikipedia.org/wiki/Carte_r%C3%A9seau
23. http://fr.wikipedia.org/wiki/R%C3%A9seau_local
24. http://fr.wikipedia.org/wiki/Serveur_HTTP
25. http://fr.wikipedia.org/wiki/Serveur_FTP
26. <http://fr.wikipedia.org/wiki/Internet>
27. <http://fr.wikipedia.org/wiki/Logiciel>
28. <http://fr.wikipedia.org/wiki/Firmware>
29. <http://fr.wikipedia.org/wiki/BIOS>

30. <http://fr.wikipedia.org/wiki/USB>
31. <http://fr.wikipedia.org/wiki/Firewire>
32. http://fr.wikipedia.org/wiki/Peripheral_Component_Interconnect
33. http://fr.wikipedia.org/wiki/PCI_Express
34. http://fr.wikipedia.org/wiki/Accelerated_Graphics_Port
35. <http://fr.wikipedia.org/wiki/DirectX>
36. http://fr.wikipedia.org/wiki/Fr%C3%A9quence_d%27horloge
37. http://fr.wikipedia.org/wiki/M%C3%A9moire_cache
38. http://fr.wikipedia.org/wiki/Syst%C3%A8me_binaire
39. <http://fr.wikipedia.org/wiki/PDA>
40. <http://fr.wikipedia.org/wiki/Wifi>
41. <http://fr.wikipedia.org/wiki/Concentrateur>
42. http://fr.wikipedia.org/wiki/Commutateur_r%C3%A9seau
43. <http://fr.wikipedia.org/wiki/Intranet>
44. http://fr.wikipedia.org/wiki/M%C3%A9moire_ROM
45. http://fr.wikipedia.org/wiki/Application_programming_interface
46. http://fr.wikipedia.org/wiki/Plug_and_Play
47. <http://fr.wikipedia.org/wiki/P2p>
48. http://fr.wikipedia.org/wiki/Nom_de_domaine
49. <http://fr.wikipedia.org/wiki/DNS>
50. http://fr.wikipedia.org/wiki/Adresse_IP
51. <http://fr.wikipedia.org/wiki/PCMCIA>
52. http://fr.wikipedia.org/wiki/Risques_en_s%C3%A9curit%C3%A9_informatique
53. http://fr.wikipedia.org/wiki/Virus_informatique
54. http://fr.wikipedia.org/wiki/Logiciel_malveillant
55. http://fr.wikipedia.org/wiki/Canular_informatique
56. <http://fr.wikipedia.org/wiki/Pourriel>
57. http://fr.wikipedia.org/wiki/Ver_informatique
58. http://fr.wikipedia.org/wiki/Cheval_de_Troie_%28informatique%29
59. <http://fr.wikipedia.org/wiki/Backdoor>
60. http://fr.wikipedia.org/wiki/Logiciel_espion
61. http://fr.wikipedia.org/wiki/Exploit_%28informatique%29
62. <http://fr.wikipedia.org/wiki/Rootkit>
63. <http://fr.wikipedia.org/wiki/Antivirus>
64. <http://fr.wikipedia.org/wiki/Pare-feu>
65. http://fr.wikipedia.org/wiki/Zone_d%C3%A9militaris%C3%A9e
66. <http://fr.wikipedia.org/wiki/Hame%C3%A7onnage>
67. <http://fr.wikipedia.org/wiki/Spoofing>
68. http://fr.wikipedia.org/wiki/D%C3%A9ni_de_service
69. <http://fr.wikipedia.org/wiki/Cracker>
70. <http://fr.wikipedia.org/wiki/Hacker>
71. http://fr.wikipedia.org/wiki/Pirate_informatique
72. <http://fr.wikipedia.org/wiki/ADSL>
73. <http://fr.wikipedia.org/wiki/URL>
74. http://fr.wikipedia.org/wiki/Navigateur_Web
75. http://fr.wikipedia.org/wiki/Page_Web
76. <http://fr.wikipedia.org/wiki/HTML>
77. <http://fr.wikipedia.org/wiki/HTTP>
78. <http://fr.wikipedia.org/wiki/SSL>
79. <http://fr.wikipedia.org/wiki/E-commerce>
80. <http://fr.wikipedia.org/wiki/Serveur>
81. <http://fr.wikipedia.org/wiki/Cookie>
82. http://fr.wikipedia.org/wiki/Certificat_%C3%A9lectronique
83. http://fr.wikipedia.org/wiki/Cryptographie_asym%C3%A9trique

84. <http://fr.wikipedia.org/wiki/Botnet>
85. <http://fr.wikipedia.org/wiki/Lamer>
86. <http://fr.wikipedia.org/wiki/Freeware>
87. <http://fr.wikipedia.org/wiki/Partagiciel>
88. http://fr.wikipedia.org/wiki/Open_source
89. <http://fr.wikipedia.org/wiki/Routeur>
90. <http://fr.wikipedia.org/wiki/Multim%C3%A9dia>
91. http://fr.wikipedia.org/wiki/Compression_de_donn%C3%A9es
92. http://fr.wikipedia.org/wiki/Encodage_num%C3%A9rique
93. <http://fr.wikipedia.org/wiki/Codec>
94. <http://fr.wikipedia.org/wiki/MP3>
95. <http://fr.wikipedia.org/wiki/MPEG-1>
96. http://fr.wikipedia.org/wiki/Vid%C3%A9o_CD
97. <http://fr.wikipedia.org/wiki/MPEG-2>
98. <http://fr.wikipedia.org/wiki/MPEG-3>
99. <http://fr.wikipedia.org/wiki/MPEG-4>
100. http://fr.wikipedia.org/wiki/Conteneur_vid%C3%A9o
101. http://fr.wikipedia.org/wiki/Compression_MPEG
102. <http://fr.wikipedia.org/wiki/DivX>
103. <http://fr.wikipedia.org/wiki/XviD>
104. http://fr.wikipedia.org/wiki/Audio_Video_Interleave
105. <http://fr.wikipedia.org/wiki/WAV>
106. <http://fr.wikipedia.org/wiki/WMA>
107. http://fr.wikipedia.org/wiki/Digital_Rights_Management
108. http://fr.wikipedia.org/wiki/Image_num%C3%A9rique
109. <http://fr.wikipedia.org/wiki/JPEG>
110. <http://fr.wikipedia.org/wiki/TIFF>
111. <http://fr.wikipedia.org/wiki/Pixel>
112. <http://fr.wikipedia.org/wiki/Ogg>
113. <http://fr.wikipedia.org/wiki/WMV>
114. <http://fr.wikipedia.org/wiki/Gif>
115. http://fr.wikipedia.org/wiki/Portable_Network_Graphics
116. <http://fr.wikipedia.org/wiki/FLAC>
117. <http://fr.wikipedia.org/wiki/RGB>
118. <http://www.ripp-it.com/glossaire/mot-Entrelacement-44-lettre-e-Categorie-toutes.html>
119. http://fr.wikipedia.org/wiki/File_Transfer_Protocol
120. <http://fr.wikipedia.org/wiki/Voip>
121. <http://fr.wikipedia.org/wiki/Streaming>