

Introduction au Web et à HTML



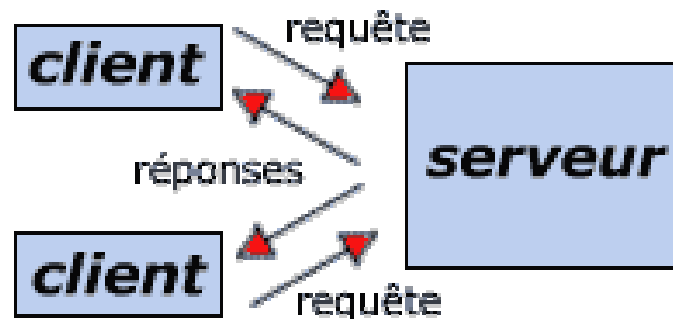
Daniel Gautheret
ESIL, Université de la Méditerranée

Concept client-serveur

Une machine client contacte une machine serveur, généralement plus importante

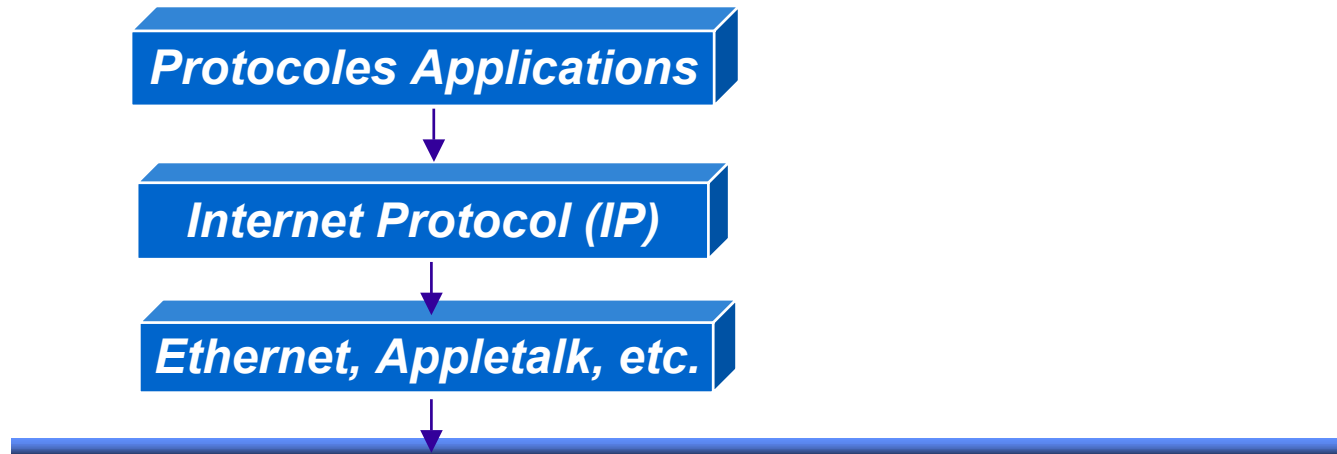
Le serveur fournit des services: données, programmes

Programmes client/programmes serveur



Protocoles de Communication

- ★ Protocoles Niveau 1: Ethernet, Appletalk
- ★ Protocoles Niveau 2: Internet Protocol (IP)
- ★ Protocoles Niveau 3: Application



Internet Protocol

Historique

- ★ Arpanet (69). Armée US
- ★ TCP/IP (72)
- ★ Inclus dans Unix (80)
- ★ Backbone NSFnet (88)
- ★ Ebone, Renater (92)

Aujourd 'hui plusieurs centaines de millions de machines

Universalité du protocole

- ★ Non lié à une machine
- ★ Transporte n'importe quoi

Gratuité, égalité

- ★ A partir de votre point d'entrée, vous pouvez lire et diffuser dans le monde entier

Quelques Propriétés d'IP

Adresses IP: uniques au monde

- ★ 192.134.235.250
- ★ classes A,B,C. 4 milliards d'adresses possibles

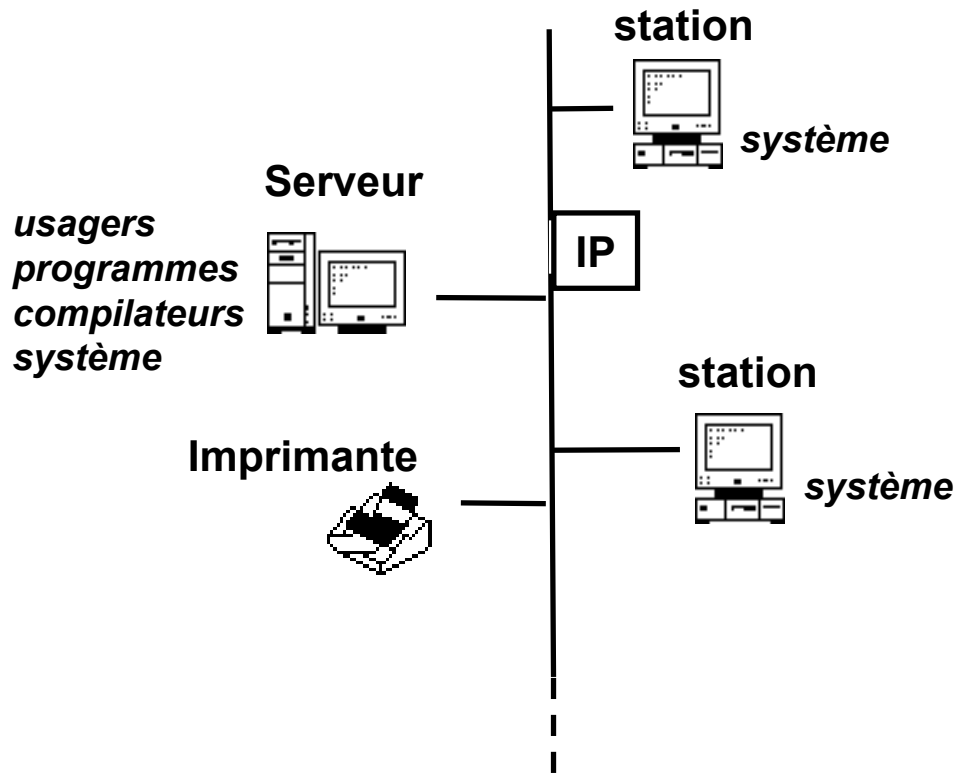
1 à 126	0 à 255	0 à 255	0 à 255
---------	---------	---------	---------

- ★ Numéros attribués par l'IANA (internet assigned numbers agency)

Adresses symboliques

- ★ organisation.domaine
- ★ noms de domaines
 - edu, gov, com, mil, net, fr, uk, it, de, es
- ★ igs.cnrs-mrs.fr
- ★ yahoo.com

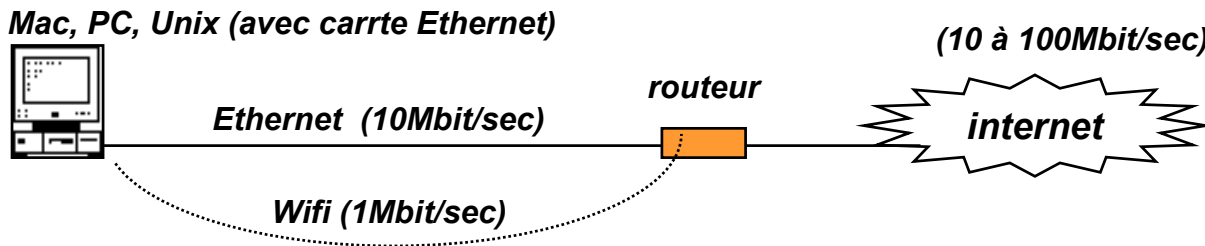
Réseau Internet: ensemble des clients/serveurs IP



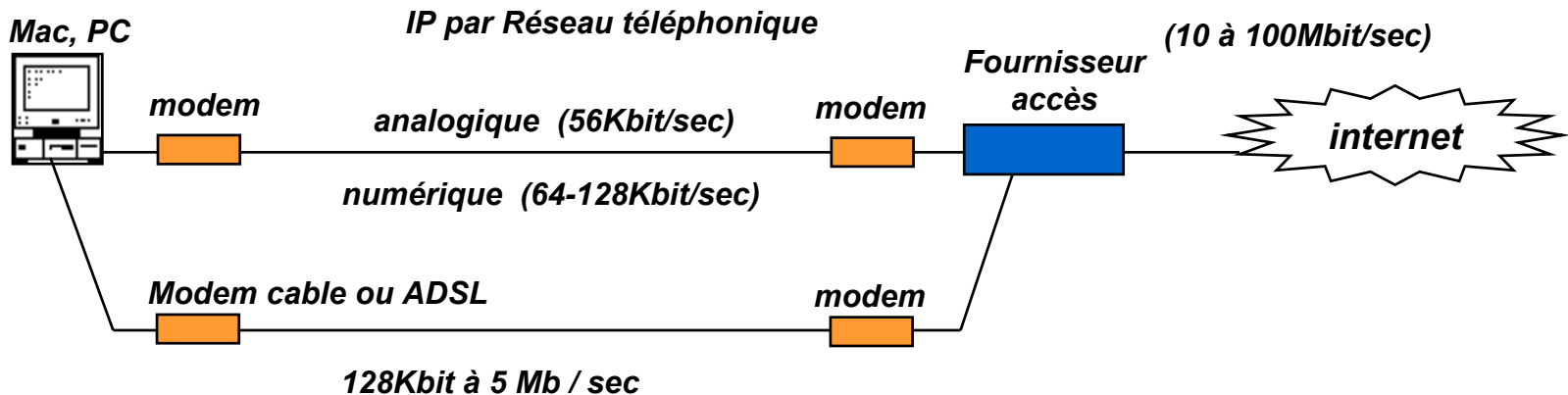
Installation d'une Connexion Internet

Ex. d'installations avec accès Internet

- **Entreprise/Université: toutes les machines sur Ethernet ou Wifi**



- **Maison via Modem**

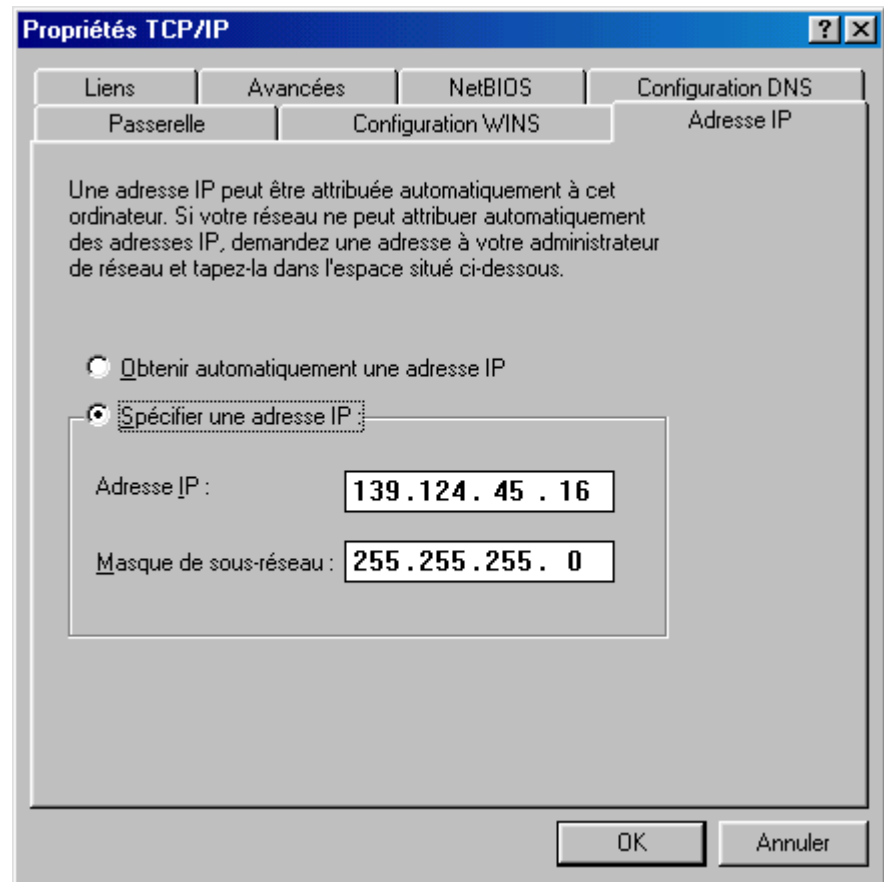


Paramètres d'une connexion...

Spécifier l'adresse IP (« IP fixe »)

ou:

Adresse obtenue automatiquement (cas le + fréquent)



Protocoles applicatifs (3ème niveau)

smtp et pop (mail)

ftp (transfert de fichier)

telnet (connexion à distance)

nntp (forums-news)

http (world wide web)

Exemples

mail (protocole *smtp* et *pop*): envoi et lecture de courrier électronique

★ Adresse d'une personne

- nom@adresse-symbolique
- lbc2@ylumin.univ-mrs.fr

★ Exemple d'utilisation sous Unix

```
mail lbc2@ylumin.univ-mrs.fr  
salut!
```

.

Les 2 formes de courrier

Deux façons de gérer le courrier électronique via le Web.

Serveur classique (pop3)

- ★ 1: Votre fournisseur d'accès garde votre courrier pour vous. Il vous l'envoie lorsque vous vous connectez avec votre navigateur.
- ★ Avantage: vous pouvez-vous déconnecter et lire le courrier localement
- ★ Inconvénient: Il faut configurer votre client mail pour que le fournisseur d'accès le reconnaisse
- ★ Clients: Outlook, Thunderbird, Eudora

Webmail

- ★ 2: Vous utilisez un site de courrier sur le Web (hotmail, gmail)
- ★ Avantage: vous n'avez rien à configurer. Vous pouvez immédiatement lire votre courrier de n'importe quel coin du globe.
- ★ Inconvénient: Il faut rester connecté durant toute la lecture. Réaction plus lente.

ftp (protocole ftp)

ftp: transfert de fichier

```
ftp <adresse-  
symbolique>  
get <fichier>  
put <fichier>  
ls  
cd <répertoire>  
quit
```

ftp anonyme

nom d'utilisateur:

anonymous

mot de passe:

ce que vous voulez

```
ftp ftp.loria.fr
```

```
Connected to mirabelle.loria.fr.
```

```
220 mirabelle.loria.fr FTP server
```

```
Name (ftp.loria.fr:gauthere): anonymous
```

```
331 Guest login ok, send your complete e-mail address as password.
```

```
Password:
```

```
230-*      Bienvenue au CRIN-Cnrs / INRIA Lorraine      *
```

```
ftp> cd pub/mac/umich/game
```

```
ftp> ls
```

```
drwxr-xr-x 2 garti ftp      1536 Jul 20 04:16 adventure
```

```
drwxr-xr-x 2 garti ftp      3584 Jul 20 04:16 arcade
```

```
drwxr-xr-x 2 garti ftp      1536 Jul 20 04:16 space
```

```
ftp> cd space
```

```
ftp> ls
```

```
-r--r--r-- 1 garti ftp      482509 Nov 12 1994 asterax1.2.cpt.hqx.gz
```

```
-rw-rw-r-- 1 garti ftp       58414 Feb 18 1994 bazfaz1.1.sit.hqx.gz
```

```
-r--r--r-- 1 garti ftp     1288215 Mar 31 1996 cyclonetwo1.02.sit.hqx.gz
```

```
ftp> get bazfaz1.1.sit.hqx.gz
```

```
local: bazfaz1.1.sit.hqx.gz remote: bazfaz1.1.sit.hqx.gz
```

```
200 PORT command successful.
```

```
150 Opening BINARY mode data connection for .....
```

```
226 Transfer complete.
```

```
58414 bytes received in 1.44 seconds (39.54 Kbytes/s)
```

```
ftp> quit
```

```
221 Goodbye.
```

session ftp

Telnet (protocole telnet)

Telnet: connection à une machine éloignée

```
telnet ens12.univ-mrs.fr
```

nom, mot de passe

...session normale...

logout

Les news (forums): nntp

The screenshot shows a Netscape Newsreader window titled "Re: je cherche un logiciel d'analyses statistiques - fr.bio.logiciel - Netscape Newsgroup". The interface includes a menu bar (File, Edit, View, Go, Message, Communicator, Help), a toolbar with icons for Get Msg, New Msg, Reply, Reply All, Forward, File, Next, Print, Mark, and Stop, and a sidebar for navigation.

The sidebar shows a tree view of folders: Local Mail (Inbox, Unsent Messages, Drafts, Templates, Sent, Trash) and news (news.univ-aix.fr, bionet.general, bionet....wanted, bionet.software, bionet....offered, fr.bio.general, fr.bio.genome, **fr.bio.logiciel**, fr.bio.medecine, fr.bio.pharmacie).

The main pane displays a list of messages with columns for Subject, Sender, and Date:

Subject	Sender	Date
Re: je cherche un logiciel d'analyses statistiq...	Charles Plessy	04/02/00 09:30
mazurka13	philippe.lillo	dim. 02:18
Re: traitement de photo	Betram Sultan	dim. 19:16
Re: OMNIPAGE	Betram Sultan	dim. 19:21
Re: recherche des cours de java et html	chehem	dim. 21:13
Re: winamp	Morgan YVARS	lun. 20:24
Enquête sur le BioNetbook	Catherine Letondal	mar. 17:35
statistiques avec excel 97	Eric C	mar. 01:48

The selected message content is as follows:

Subject: Re: je cherche un logiciel d'analyses statistiques
Date: Fri, 04 Feb 2000 09:30:28 +0100
From: [Charles Plessy <charles.plessy@evc.net>](mailto:charles.plessy@evc.net)
Organization: UUNet
Newsgroups: [fr.bio.logiciel](#)
References: [1](#), [2](#)

http://www.quelest.co.jp/Download/KyPlot/kyplot_e.htm

est gratuit mais je ne sais pas s'il sait faire tout ce que tu veux.

Charles

The status bar at the bottom indicates "Total messages: 17 Unread messages: 15".

Le protocole http

Historique

- 1965: Hypertext (1987: Hypercard)
- 1991: Web (toile) du CERN (Tim Berners-Lee)
- 1993: 50 serveurs
- 1999: 3 000 000 serveurs (800 millions de pages)

Information transmise sous forme de *pages hypertexte*

- Liens

Multimédia

- texte, image, son, vidéo

Les clients http acceptent aussi ftp, smtp, telnet

Le *World Wide Web*

WWW = "le Web" = "la Toile"

Ensemble des clients/serveurs http

Le Web n'est pas Internet

Adresses URL (Uniform Resource Locator)

★ protocole://adresse-symbolique/chemin-du-fichier

★ Exemples

- <http://www-biol.univ-mrs.fr/index.html>
- gopher://ndb.rutgers.edu:17105/11/etc/ndb_link_files
- <ftp://ftp.pdb.gov/pub>
- <file://mon-fichier>

★ Ce sont les adresses que l'on utilise avec les navigateurs et en écrivant des pages HTML

Les clients WWW

- ★ Netscape, Mozilla, Firefox
- ★ Microsoft Explorer
- ★ Opera

**adresse URL
courante** ←



Navigateurs: Aspects pratiques

★ Home page, Page d'accueil

- La page chargée par défaut lors du lancement de Netscape
- Initialement: page de Msn ou Netscape (USA) => lent!
- changement de la page d'accueil:
Preferences/options

★ Bookmarks, signets, favoris

- Les "marque-page"
- Servent à mémoriser tout URL intéressant
- Menu « signets » ou « bookmarks » ou « favoris »

Chargement des pages, cache

Les pages sont *téléchargées* après le premier appel d'un lien

- Affichage à l'écran
- Sauvegarde sur disque (cache)

Lors d'un deuxième appel, le navigateur vérifie si la page a été modifiée sur le serveur

- ★ Si modifiée: nouveau téléchargement
- ★ Sinon: lecture locale dans cache
 - Plus rapide
 - Libère le réseau

Dans Mozilla, la taille maxi du cache est réglable dans le menu "Edition/préférences/avancées".

- Cache disque: sur le disque dur
- Cache en mémoire: mémoire vive (RAM)

Logiciels associés aux Navigateurs

Les logiciels clients ne sont pas prévus pour traiter *tous* les types d'information pouvant être transmis sur le Web

- ★ lecteurs de sons (Real Audio, mp3)
- ★ lecteurs de films (QuickTime player, real audio)
- ★ lecteurs de fichiers image (PostScript, pdf)

Les navigateurs peuvent toutefois lancer automatiquement certaines applications lorsque certains fichiers sont téléchargés

Naviguer sur le Web

Par liens (hypertexte)

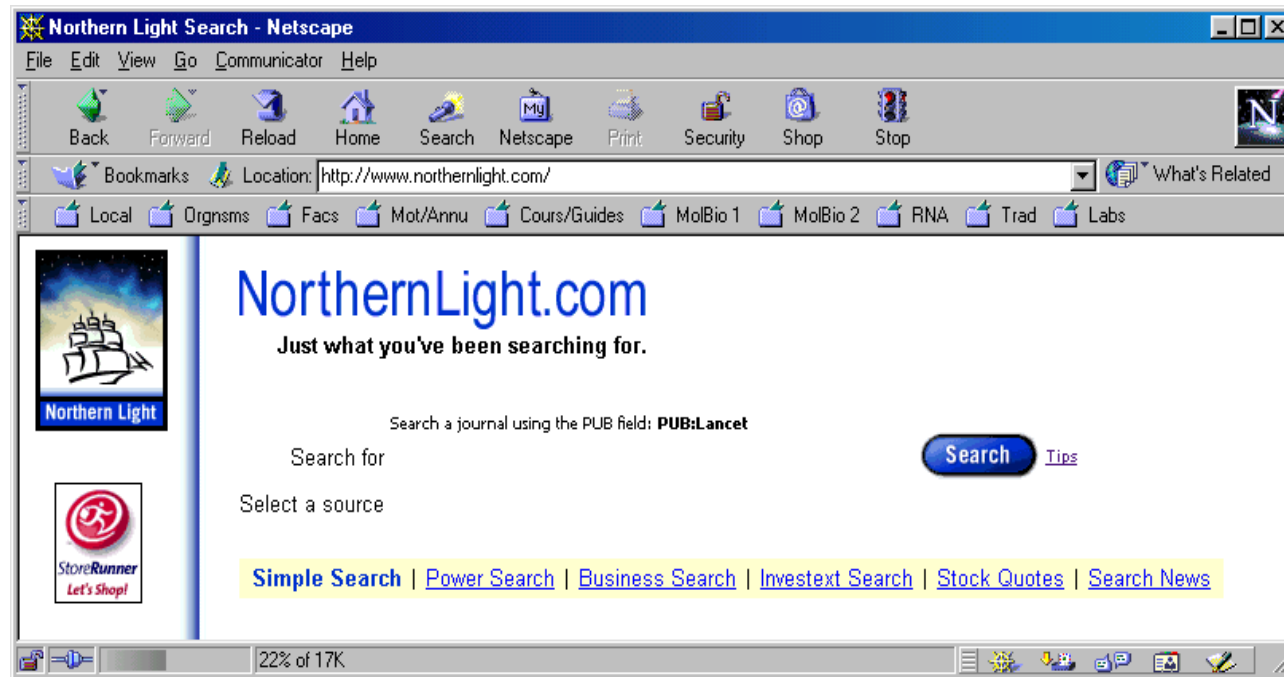
En connaissant l'URL

Index de serveurs
(INFOBIOGEN en
biologie, etc.)

Moteurs de recherche



Les moteurs de recherche



★ Internationaux

- Google
- Altavista
- Hotbot
- Fast/alltheweb

★ Français

- Voilà
- Nomade

★ Catalogues

- Yahoo

★ Meta-moteurs

- Profusion
- Savvy-search
- surfwax

Expressions booléennes

Indispensables pour une recherche efficace...

Exemple:

★ ELISA
(le test HIV)



The screenshot shows a Netscape browser window titled "Recherche Google: elisa - Netscape". The address bar contains the URL: <http://www.google.fr/search?hl=fr&ie=UTF-8&oe=UTF-8&q=elisa&btnG=Recherche+Google&meta=>. The search bar contains the word "elisa". The search results show 1 - 10 results out of approximately 2,240,000. The first result is for "Elisa.fi" with a description in Finnish. The second result is for "Elisa" with a description in French. The third result is for "ELISA vêtements grandes tailles, mode femme, prêt à porter" with a description in French. The fourth result is for "ELISA - [Traduire cette page]" with a description in English.

Expressions booléennes

AND (ET)

- ★ ELISA **AND** HIV
- ★ ELISA **+**HIV
- ★ ELISA HIV

La plupart des moteurs interprètent l'absence d'expression comme « AND »

[Evaluation of a direct ELISA HIV-1/HIV-2 EIA using synthetic ...](#) - [Traduire cette page]

... Evaluation of a direct **ELISA HIV-1/HIV-2** EIA using synthetic peptides: detect-**HIV** (version 2). Abstr Gen Meet Am Soc Microbiol. ...

www.aegis.com/pubs/aidsline/1994/oct/M94A0851.html - 9k - [En cache](#) - [Pages similaires](#)

[Elisa hiv test website,Elisa hiv test search,Elisa hiv test ...](#) - [Traduire cette page]

Elisa hiv test. Related Searches: Email. Powered by gnumatic.com

www.gnumatic.com/Elisa_hiv_test%09.html - 3k - [En cache](#) - [Pages similaires](#)

[AIDSmeds Community Forums Forums - PCR / Elisa HIV Test Combo](#) - [Traduire cette page]

... Topic Title: PCR / **Elisa HIV** Test Combo Created On Monday September 08, 2003 10:42 AM, ... PCR / **Elisa HIV** Test Combo - driwauz - Monday September 08, 2003 10:42 AM. ...

www.aidsmeds.com/Fusetalk/messageview.cfm?catid=4&threadid=6326 - 75k - [En cache](#) - [Pages similaires](#)

[elisa hiv test reliable at 28 days - PROSPECTOR](#) - [Traduire cette page]

elisa hiv test reliable at 28 days. Mrs. Test - the best home test kits. Description: MrsTest offers a variety of self-tests for ...

www.dr-test.com/catalog/elisa-hiv-test-reliable-at-28-days.htm - 25k - [En cache](#) - [Pages similaires](#)

Expressions booléennes

OR (OU)

- ★ encryption **OR** cryptographie

NOT (NON)

- ★ elisa **NOT** HIV
- ★ elisa **-**HIV



Google

Expression exacte

- ★ « to be or not to be »

Le langage HTML

HyperText Markup Language

- ★ Langage pour définir les différents composants d'un document WWW
- ★ indépendant du logiciel client (Netscape, Explorer)
- ★ indépendant de la machine

Cours HTML disponibles sur le Web

- ★ COMMENTCAMARCHE :
 - <http://www.commentcamarche.net>
- ★ MANAS (Nice)
 - <http://nephi.unice.fr/CoursHTML/>

Source

La "source" d'un document est le "programme" html pour ce document

Le nom du fichier source se termine par .html

Pour l'afficher à partir de Netscape: Menu "View", option "Document Source"

Source d'un document HTML:

```
<TITLE> ABIM Infobio home</TITLE>
<BODY BGCOLOR=#FFFFFF>
<PRE><H1><IMG ALIGN=MIDDLE SRC=/icons/soirlogo.gif> Info - BioInfo - BioPhys</H1></PRE>
<P>
<IMG SRC="/icons/redball.gif"> <B>Vous avez demande la page "Info - BioInfo - BioPhysique"</B>
<P>
<img Src=/icons/barre.njaune.gif><P>
<BR>
Page bioinformatique orient&eacute;e vers :
<P>
<H2>Informatique</H2>
<UL>
<LI>Les compagnies ayant un serveur <A HREF="/bioinfo/compagnie.html"><B>www</B></A>
<P>
<LI><A HREF="/bioinfo/logiciel.html"><B>Logiciels </B></A>
<P>
<LI><A HREF="/bioinfo/infodoc.html"><B>Documentation</B></A> <P>
<LI>Ma <A HREF="/bioinfo/make.html"><B>make</B></A> &agrave;
</UL>
```

Commandes HTML

- ★ Le langage HTML est un langage à *balises*:
 - ★ Pour faire une action sur un groupe de mots, Il faut *baliser* ce groupe de mots.
 - ★ Par exemple: mettre en gras le mot "bonjour »
 - ★ `Bonjour`
 - ★ En général, les commandes sont de la forme :
 - ★ `<marqueur> texte </marqueur>`
 - ★ ou
 - ★ `<marqueur attribut=argument> texte </marqueur>`
 - ★ ou
 - ★ `<marqueur>`
- Majuscules et minuscules sont sans importance

Commandes HTML

- | | | |
|---|---|--|
| <code><HTML> bla bla </HTML></code> | | - Commandes de début et de fin du document HTML (facultatif) |
| <code><HEAD> bla bla </HEAD></code> | | - Informations non affichées concernant le document Date et auteurs par exemple (facultatif) |
| <code><TITLE> bla bla </TITLE></code> | | - Titre du document Cette information apparait dans la barre de titre du client WWW |
| <code><BODY></code>
bla
bla
<code></BODY></code> | } | - Corps du document Toutes les informations affichées par le client WWW y sont contenues. |
| <code><H1>bla bla</H1></code>
<code><H2>bla bla</H2></code>
..
<code><H6>bla bla</H6></code> | } | - H1 à H6: caractères de moins en moins gros |
| <code>bla bla</code> | | - Gras (bold) |
| <code><BLINK>bla bla</BLINK></code> | | - Clignotant |
| <code><CENTER>bla bla</CENTER></code> | | - Centré |
| <code>bla</code> | | - Texte rouge |

Commandes HTML

`<P>`

– Saut de paragraphe

`
`

– Saut à la ligne (sans écartement des lignes)

`<PRE>`

 bla bla

 bla bla

– Garde les sauts de ligne tels-quels

`</PRE>`

`<UL TYPE=SQUARE>`

– Liste. Résultat:

`` premier

□ premier

`` second

□ second

``

`<HR>`

– Insertion d'une ligne horizontale de séparation

`<TABLE>`

– Table. Résultat:

`<TR>`

`<TD>`Article`</TD>`

`<TD>`Prix`</TD>`

`</TR>`

`<TR>`

`<TD>`ordinateur`</TD>`

`<TD>`10,00F`</TD>`

`</TR>`

`</TABLE>`

Article	Prix
Ordinateur	10,00F

Commandes HTML

Accents

- ★ **utiliser le caractère &**, la lettre, le nom de l'accent **et** un point virgule
- ★ **par ex: é se code** é

Les liens

- ★ `texte-à-afficher`
- ★ `texte-à-afficher`
- ★ `texte-à-afficher`
- ★ nécessite un signet: ``

Les images

★ Images des pages html: format jpg, gif ou png.

Afficher une image:

★ ``

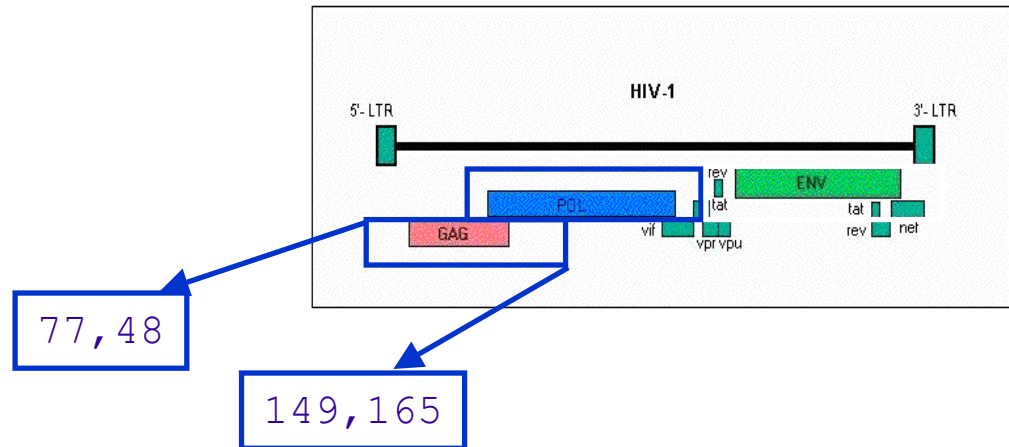
options: align=center, align=bottom, border=0 , border=1

Image faisant lien:

★ ``

Image "active" (contenant plusieurs liens)

- ★ But: l'utilisateur doit pouvoir se rendre à des endroits différents en cliquant sur différentes parties de l'image.



Code HTML:

```
<MAP NAME="VIRUS">  
  <AREA SHAPE="RECT" HREF= "gag.html" COORDS="77,48,149,165">  
  <AREA SHAPE="RECT" HREF= "pol.html" COORDS="144,30,275,144">  
  <AREA SHAPE="RECT" HREF= "env.html" COORDS="325,113,452,130">  
</MAP>  
<IMG SRC="hiv.gif" USEMAP="#VIRUS" >
```

Frames (cadres)

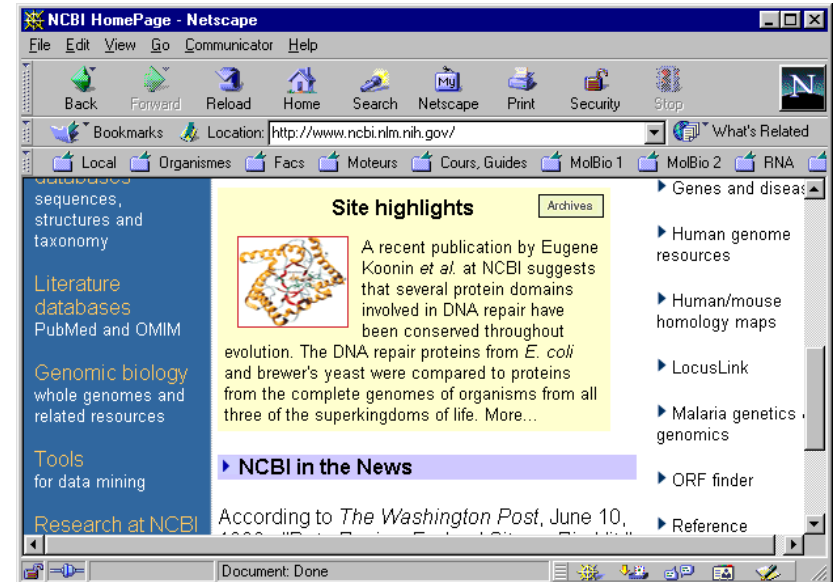
- une page html pour le cadre:

```
<FRAMESET COLS="25%,*" FRAMEBUFFER=0 BORDER=0 FRAMESPACING=0>  
  <FRAME NAME="gauche" SRC="sommaire.html" SCROLLING="VERTICAL">  
  <FRAME NAME="droite" SRC="welcome.html" SCROLLING="auto">  
</FRAMESET>
```

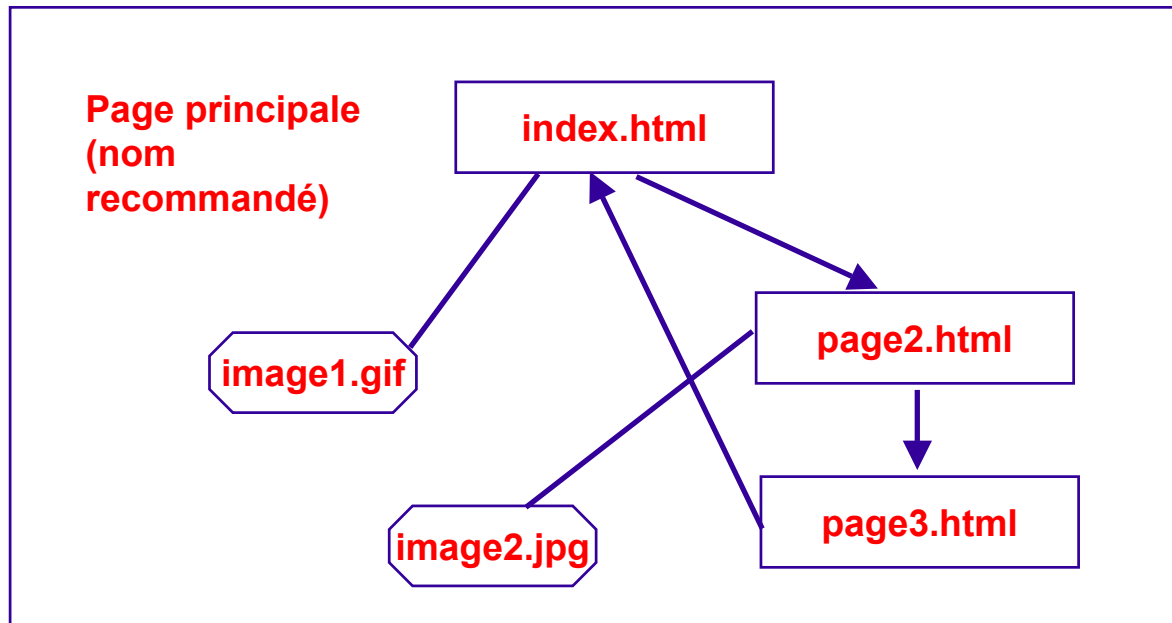
- une page html pour chaque partie du cadre:

```
<A HREF= "page1.html" TARGET= "droite" >  
  cliquer ici pour voir la page 1 </A>  
<A HREF= "page2.html" TARGET= "droite" >  
  cliquer ici pour la page 2 </A>
```

sommaire.html



Organisation de vos pages



Tout dans le même répertoire

→ liens

— utilisation

Les noms des fichiers

En théorie, tout est permis, mais il est très fortement recommandé de suivre les règles suivantes, afin de faciliter le transport de vos pages entre machines et entre systèmes (Windows/Mac/Unix).

- ★ N'utiliser que des minuscules
- ★ Pour les fichiers html, utiliser le suffixe « .html » ou « .htm »
- ★ Pour les graphiques, utiliser le suffixe « .gif » ou « .jpg »
- ★ **Aucun caractère spécial (:/%\$), aucun accent, aucun espace**
- ★ Placer tous les fichiers dans le même répertoire
- ★ Editer les pages html avec Wordpad, pas avec Word. Word convertit automatiquement le type du fichier et vous place en mode « éditeur ». Vous perdez le contrôle sur le code html.

Publication des pages HTML sur un serveur http

A l'ESIL:

- ★ Copier les fichiers dans votre répertoire « public_html »
- ★ Vérifiez le bon fonctionnement avec Netscape / Explorer / différents systèmes (Mac, PC, Linux)
- ★ L'URL à l'esil est: <http://www.esil.univ-mrs.fr/~votrelogin>

Chez un hébergeur

- ★ La plupart des fournisseurs d'accès vous permettent de publier vos pages
- ★ Transférer les pages par ftp sur leur serveur

Déclarer ses pages aux moteurs de recherche

- ★ Les moteurs de recherche explorent le web et mémorisent les mots contenus dans les pages
- ★ Mais ils n'indexent qu'une partie du Web (15-16% pour les plus complets: **Northern Light, Alta Vista, Hotbot, Google**)
- ★ L'indexation d'une page dépend entre autres du **nombre de liens** sur cette page.
- ★ Il est donc plus efficace de déclarer vous même votre site (voir liens du type « register URL » sur les pages d'accueil des principaux moteurs)

Les commandes META

- ★ Les commandes META servent à insérer dans votre page des informations lisibles par les navigateurs mais invisibles par les utilisateurs.
- ★ Se placent dans l'en-tête

Exemple:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2//EN">
<HTML>
<HEAD>
  <TITLE>UTRs</TITLE>
  <META NAME="Author" CONTENT="Daniel Gautheret">
  <META NAME="Description" CONTENT="Information on messenger RNA untranslated regions">
  <META NAME="Keywords" CONTENT="mRNA, UTRs, UTR, messenger, RNA, messenger, ARN">
  <META NAME="Classification" CONTENT="Biology">
</head>
```

Ce que vous ne verrez pas dans ce cours

Les Formulaires (“Forms”)

- ★ Permettent à l'utilisateur d'entrer des informations à traiter, d'envoyer ces informations, puis de recevoir le résultat du traitement.
- ★ La page vide est facile à réaliser en HTML, mais pour traiter l'information et renvoyer le résultat, il faut savoir programmer.

Advanced BLAST

Message of the day ...

Choose program to use and database to search:

Program: blastn Database: nr

The query sequence is filtered for low complexity regions by default.

Enter here your input data as: Sequence in FASTA format

Submit Query

Champ Menu déroulant Bouton

Les Feuilles de style

- ★ Fichier .css définissant les styles par défaut de `<body>`, `<p>`, `<h1>`, etc.
- ★ Une ligne de commande en début de page html indique que la feuille de style doit être utilisée

Vous ne verrez pas... (suite)

JAVA

- ★ Site Web avec beaucoup d'échange d'information entre serveur et client => Réseau trop lent.
- ★ Envoyer au client le programme qui génère l'information. Le programme sera exécuté par le client.
- ★ Tous les clients doivent être capables de recevoir et exécuter ce programme => Définition d'un nouveau standard de programmation: JAVA
- ★ L'interpréteur de JAVA peut être installé sur n'importe quelle machine, ce qui en fait un langage "universel".