

# **INTRODUCTION**

**à la technologie**

**WWW**



Eric Gressier

Cnam-Cedric

## Bibliographie

"Spinning the Web, How to Provide Information on the Internet"  
A. Ford. International Thomson Publisher.

"Le monde Internet"  
E. Krol. O'Reilly. International Thomson Publisher.

"TCP/IP illustrated, Volume 1 : The protocols"  
W.R. Stevens. Addison Wesley. 1994.

"Initiation à l'Internet - Le réseau au CNAM : Environnement  
planétaire & Environnement local"  
L. Duchien, E. Gressier. Polycopié TP-ACB2. 94-95

"L'Internet : Concepts, Outils et Services"  
L. Duchien. Support de cours TC- TPIP. 94-95.

"Introduction à la communication IPC Internet"  
E. Gressier. Support de cours TC -ProjetCDI. 95-96.

*Eric Gressier*

# Plan

1. Introduction
2. Communication
3. World Wide Web

# Introduction

Copyright

Copyright

Eric Gressier

## Internet - Services Offerts

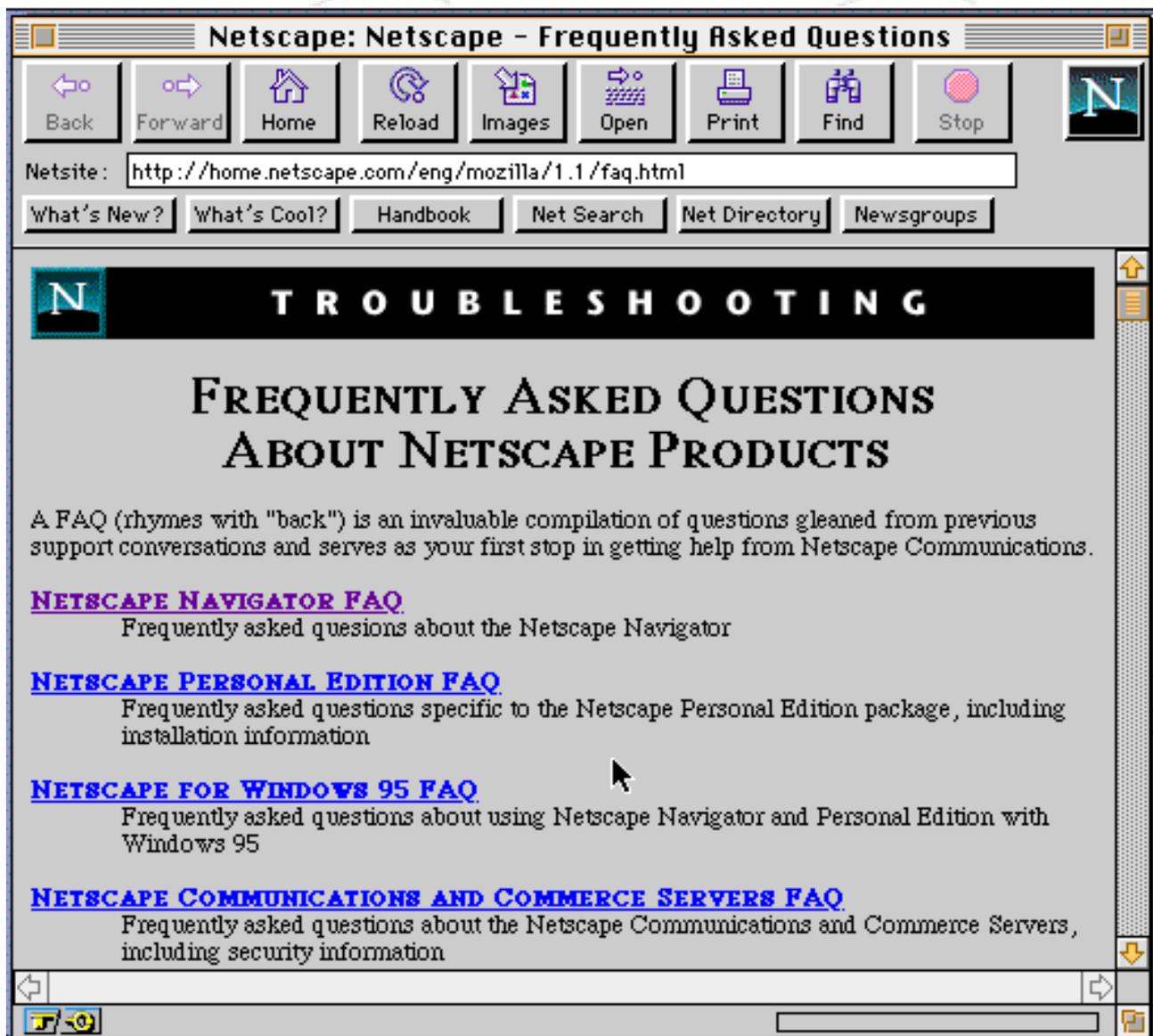
Ensemble de protocoles spécialisés :

- > Transfert de fichiers : **"ftp"**, **"tftp"**
- > Terminal virtuel : **"telnet"**, **"rlogin"**
- > Exécution à distance : **"rsh"**, **"rcp"**
- > Serveur de noms : **Bind**
- > Authentification : **Kerberos**
- > Gestion de Réseaux : **SNMP**
- > Messagerie (e-mail): **SMTP**
- > Partage de fichiers **NFS** : données, sources, binaires, bibliothèques, utilitaires ...

Plus ou moins rustiques :-)

# Systemes d'information : WWW

## WWW : World Wide Web toile d'araignée planétaire



## Aspects du WWW

. impact sociologique des informations :

### **Comité de Rédaction**

. contenu et présentation des informations mise à la disposition du public :

### **Editeur de revue et maquettiste**

. mise en oeuvre :

### **Administrateur/Ingénieur Système et Réseau**

*Eric Gressier*

# Communication

Copyright

CNAM

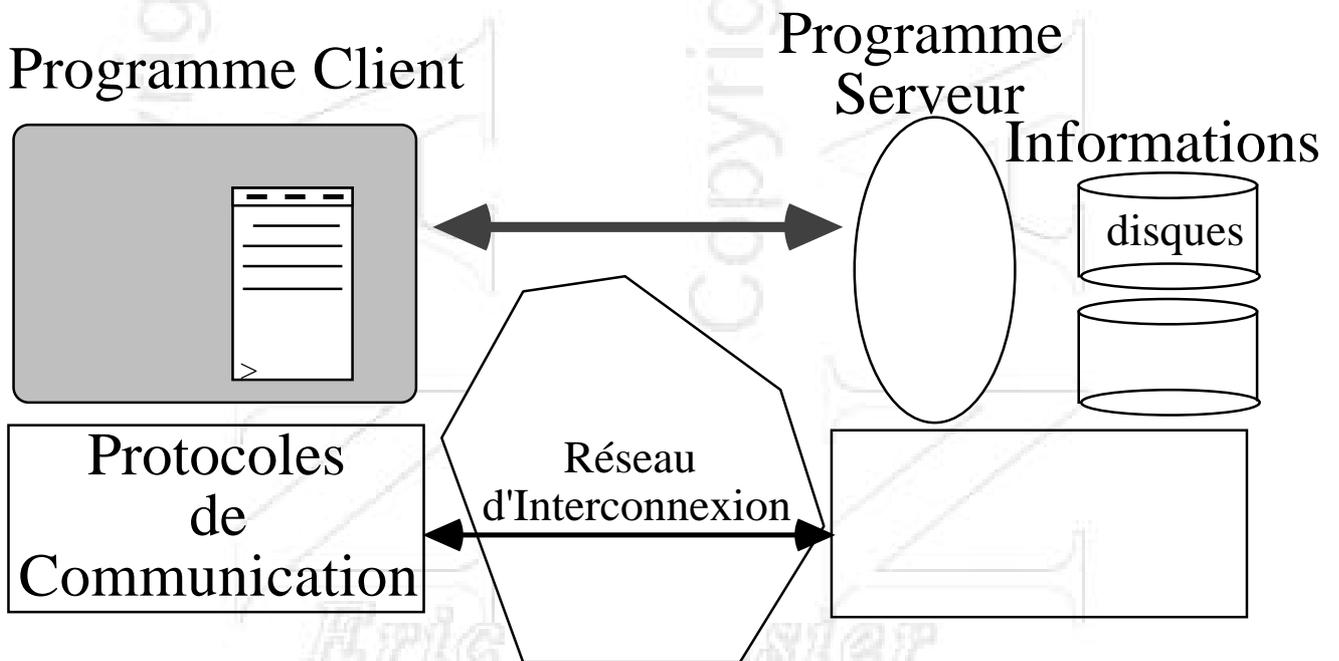
Copyright

CNAM

*Eric Gressier*

# Modèle d'échange des applications sur l'Internet

## Modèle Client/Serveur

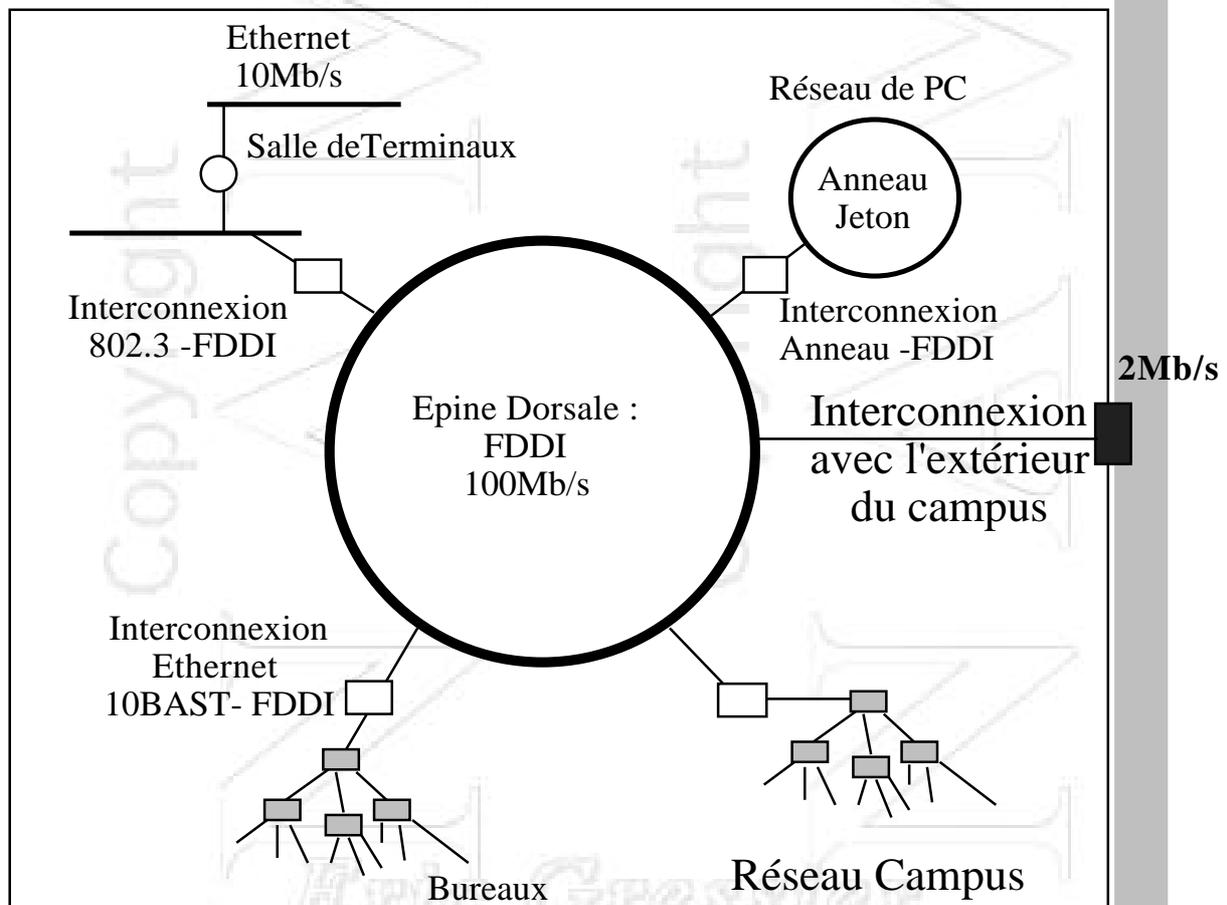


# Protocoles de communication associés à l'Internet

Applications Réseaux

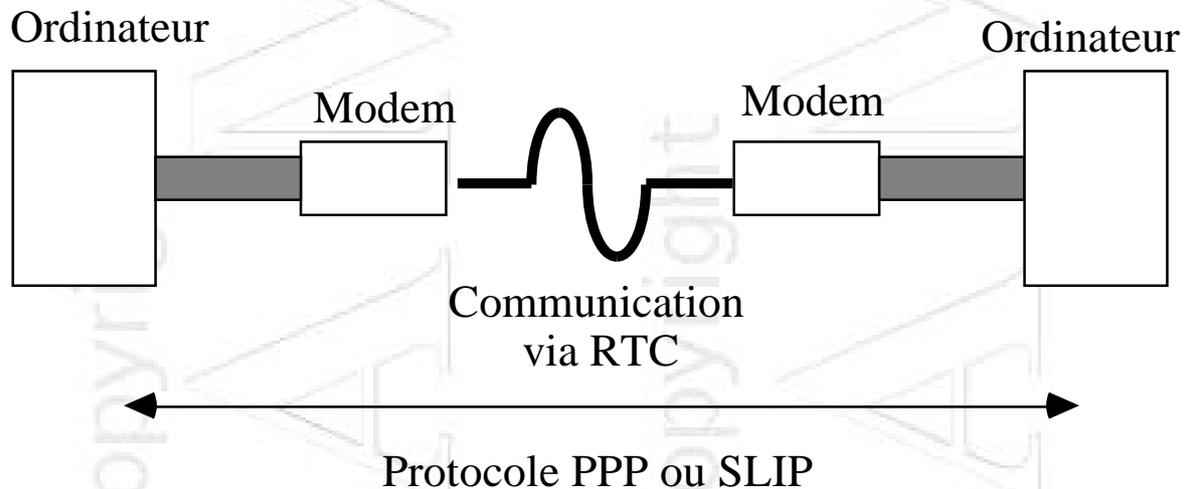
<b>7. Application</b>	ftp, rsh, rlogin rcp, http ...	NFS, NIS	tftp,time, talk ...
<b>6. Présentation</b>		XDR	
<b>5. Session</b>		RPC	
<b>4. Transport</b>	TCP		UDP
<b>3. Réseau</b>	IP		
<b>2. Liaison</b>	Réseaux	Lignes	Réseaux
<b>1. Physique</b>	Locaux	Point à Point	Publiques

## Réseau de liaison : Réseau Local



On peut avoir des réseaux plus complexes avec plusieurs technologies : Ethernet avec Appletalk par exemple.

## Réseau de liaison : RTC + modem



PPP : Point-to-Point Protocol

SLIP : Serial Link Internet Protocol

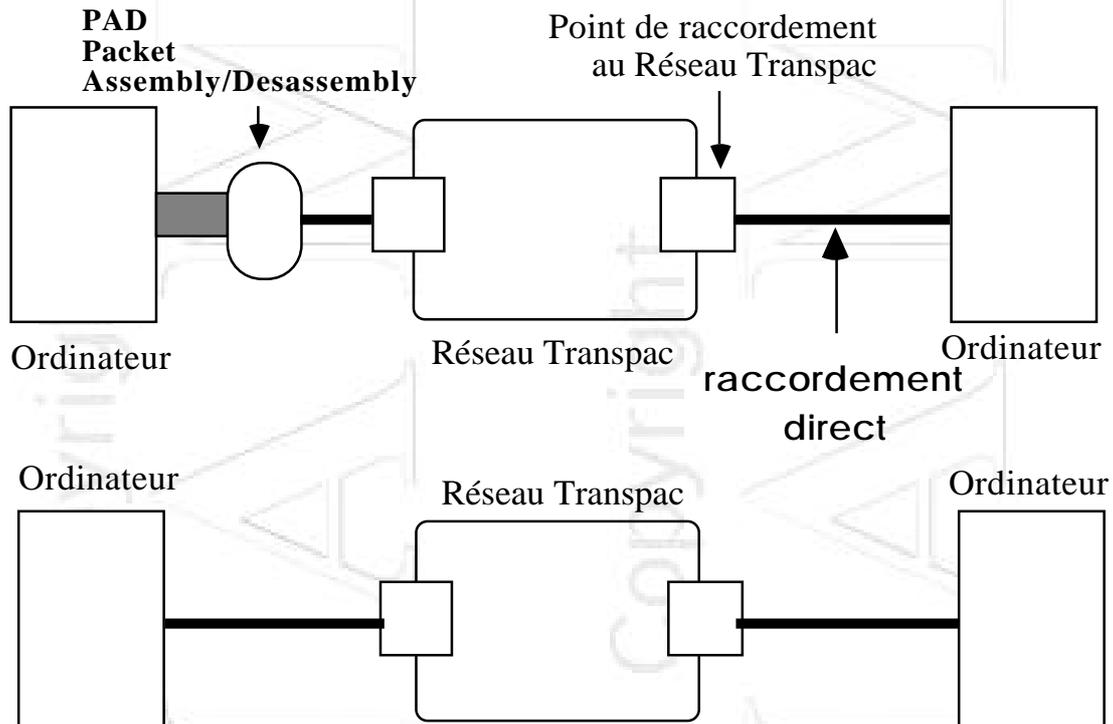
Appel du modem distant par Tel

puis dialogue entre les deux ordinateurs suivant un protocole de liaison

vitesse 14 400b/s ou 28 800b/s sans compression des données par le modem

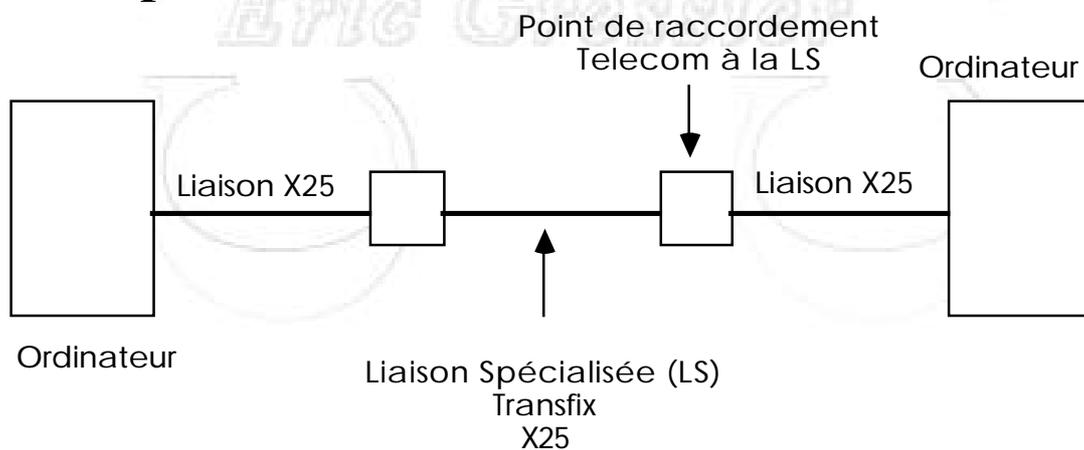
# Réseau de liaison : Réseaux publics et Lignes Spécialisées

## Réseau Transpac :

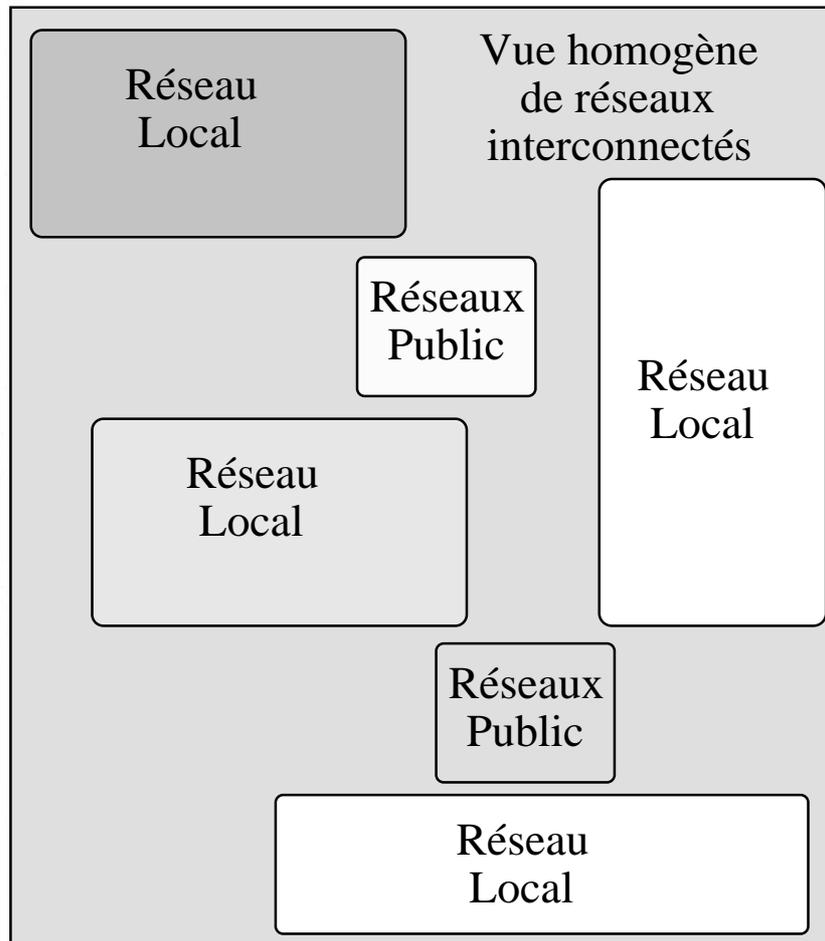


Numéris offre un schéma assez proche

## Liaison spécialisée :



## Internet: Fédération de Réseaux de liaison



en plus du transfert d'information :

- **Homogénéisation** -> Adressage et Adaptation de la transmission à la liaison traversée
- **Routage** -> intra-domaine et inter-domaine
- **Contrôle de congestion** -> Gestion des Ressources de l'ensemble du réseau

## Réseau Internet : IP (Internet Protocol)

- \* Adressage Internet
- \* Conversions d'Adresses (IP:192.200.3.45<->LAN: 08:00:20:06:4b:8e)
- \* Routage entre Réseaux
- \* Fragmentation/Réassemblage, Adaptation de la taille des messages soumis par la couche Transport suivant les possibilités offertes par la couche Liaison.
- \* Communications dans le mode minimal : **DATAGRAM** (mode non connecté) -> Envois de paquets **sans Acquittements** => la détection des messages erronés ou perdus et leurs réémissions sont à la charge de l'émetteur des messages (couche Transport).

## Adressage Internet (1)

### Adressage uniforme des réseaux de liaison

qui masque ->Adresses LAN  
->Adresses WAN

l'Internet définit des **Adresses Uniques**  
**Universelles**

T . U . V . W  
192.200.25.1

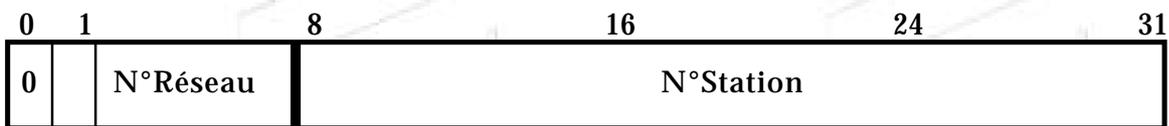
(N°Réseau, N°station)

l'Adresse IP est sur **32 bits- 4 octets**

## Adressage Internet (2)

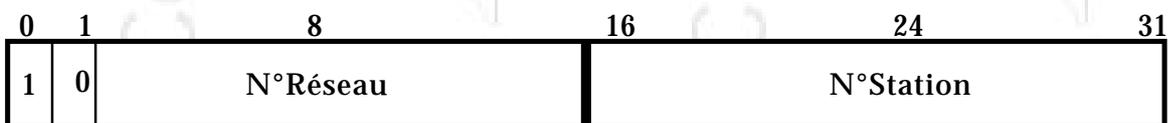
=> 3 Classes d'Adresses :

**Classe A** : Peu de Réseaux, de nombreuses Stations par Réseau



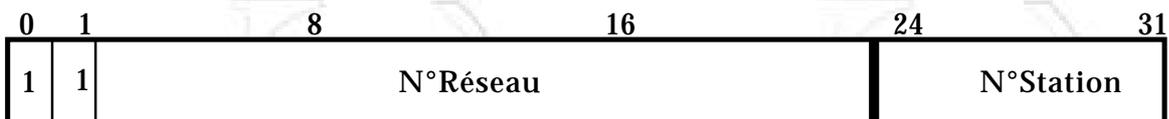
N°de Réseau : 1 à 126

**Classe B** :



N°de Réseau : 128.1 à 191.254

**Classe C** : Beaucoup de Réseaux, Peu de Stations par Réseau

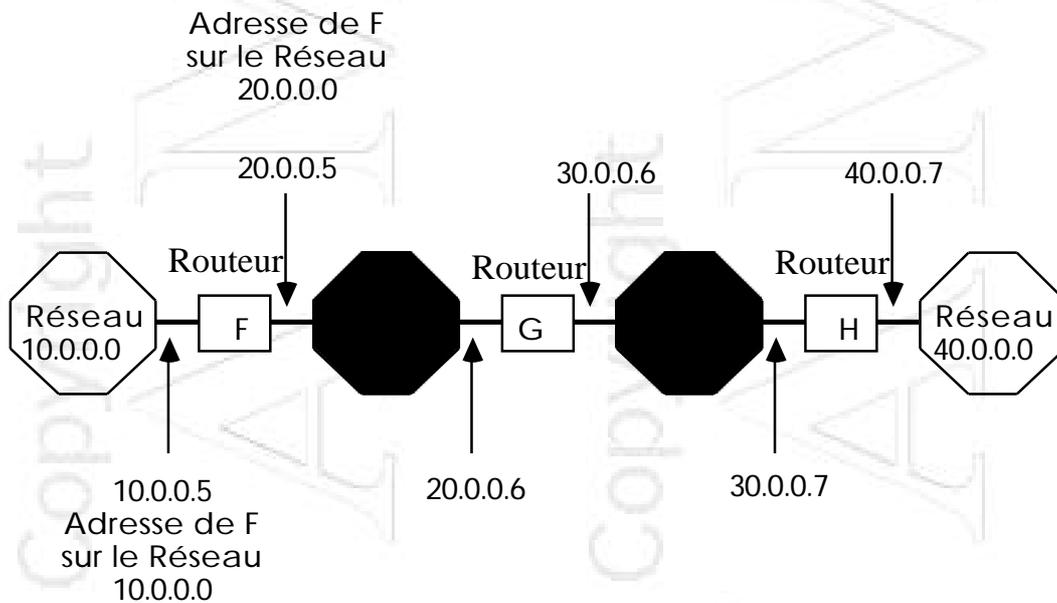


N°de Réseau : 192.0.1 à 223.255.254

N°de Station : 1 à 254    **Broadcast : 255** dans le N° de Station

# Routage IP

-> atteindre un destinataire en masquant la traversée d'une série de réseaux hétérogènes



Pour envoyer à un site sur le Réseau suivant	Envoyer à cette Adresse
20.0.0.0	Délivrer Direct
30.0.0.0	Délivrer Direct
10.0.0.0	20.0.0.5
40.0.0.0	30.0.0.7

Tables de Routage de G

Indiquer le routeur par défaut qui saura diriger les paquets vers le bon réseau

## Exemple de traversée de réseaux

/usr/etc/traceroute cyr.culture.fr (départ du cnam)

traceroute to cyr.culture.fr (143.126.201.251), 30 hops max, 40 byte packets

1	internet-gw	(163.173.128.2)	0 ms	0 ms	0 ms
2	renater-gw	(192.33.159.1)	0 ms	10 ms	0 ms
3	danton1.rerif.ft.net	(193.48.58.113)	110 ms	80 ms	90 ms
4	stlamb3.rerif.ft.net	(193.48.53.49)	100 ms	130 ms	100 ms
5	stamand1.renater.ft.net	(192.93.43.115)	90 ms	60 ms	50 ms
6	stamand3.renater.ft.net	(192.93.43.17)	70 ms	100 ms	90 ms
7	rbs1.renater.ft.net	(192.93.43.170)	130 ms	120 ms	*
8	Paris-EBS2.Ebone.NET	(192.121.156.226)	110 ms	90 ms	100 ms
9	icm-dc-1.icp.net	(192.121.156.202)	220 ms	110 ms	220 ms
10	icm-dc-1-F0/0.icp.net	(144.228.20.101)	200 ms	230 ms	290 ms
11	Vienna1.VA.Alter.Net	(192.41.177.249)	250 ms	210 ms	240 ms
12	Falls-Church4.VA.ALTER.NET	(137.39.100.33)	330 ms	220 ms	180 ms
13	Falls-Church1.VA.ALTER.NET	(137.39.8.2)	270 ms	290 ms	230 ms
14	Amsterdam2.NL.EU.net	(134.222.35.1)	380 ms	410 ms	460 ms
15	Amsterdam1.NL.EU.net	(193.242.84.1)	350 ms	380 ms	310 ms
16	134.222.30.2	(134.222.30.2)	150 ms	490 ms	530 ms
17	Rocquencourt.FR.EU.net	(193.107.192.18)	340 ms	340 ms	330 ms
18	143.126.200.203	(143.126.200.203)	300 ms	410 ms	*
19	cyr.culture.fr	(143.126.201.251)	460 ms	220 ms	290 ms

Situation corrigée depuis ... normalement !

## Couche Transport

Acheminer des données d'un bout à l'autre d'un réseau composé lui-même d'un ensemble de sous-réseaux.

Deux modes de communication :

. **type lettre** : mode non connecté/datagramme (celui de la couche IP)

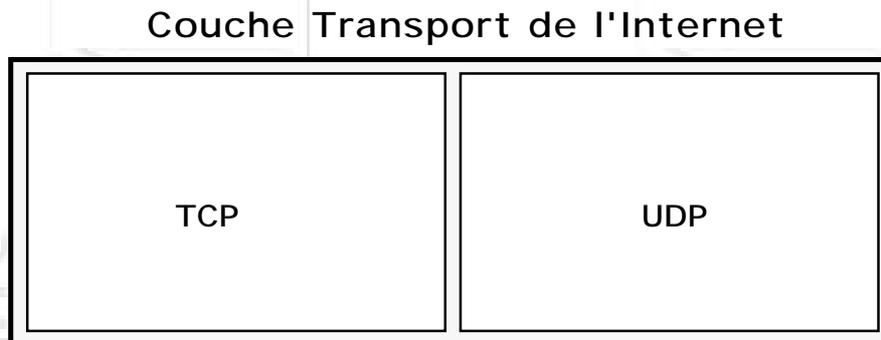
- messages non ordonnés,
- pertes et duplications de messages non gérées
- saturation du récepteur non gérée

. **type téléphone** : mode connecté -> assurer un service fiable

- messages ordonnés
- pas de pertes de messages
- élimination des doublons
- contrôle de flux sur l'échange
- support des données urgentes

Le protocole http utilise le mode téléphone.

## Procoles TCP & UDP



mode connecté : TCP - Transmission Control Protocol

mode non-connecté : UDP - User Datagram Protocol

Cette couche ne manipule pas des Adresses de sites IP mais des "**N° de Port**" pour atteindre un service distant<sub>1</sub>.

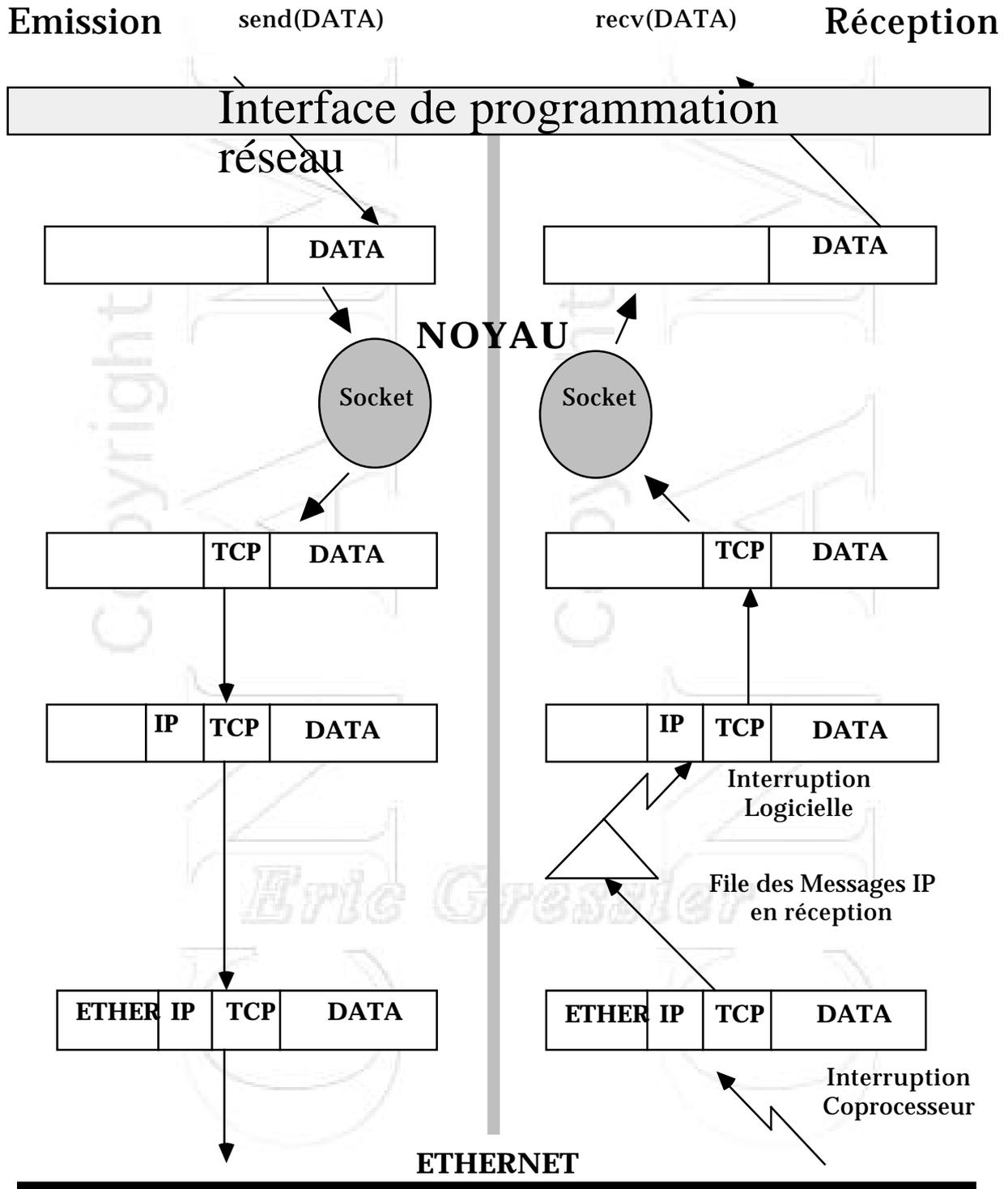
Le N° de port distingue un service parmi l'ensemble des services accessibles à travers la couche Transport sur une machine distante.

### HTTP utilise TCP

---

<sub>1</sub> En fait, la désignation d'un processus unix qui réalise le service distant se fait à l'aide de la notion d'extrémité. Hors, une extrémité est un N° de port + une adresse IP.

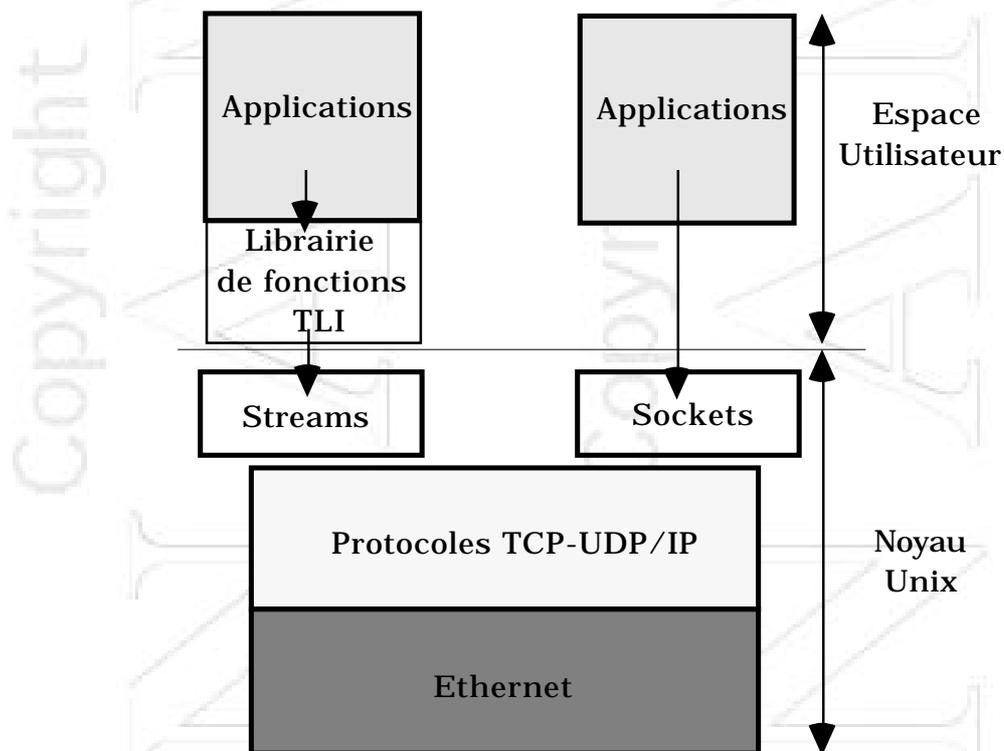
# Interface avec les protocoles (1)



## Interfaces avec les protocoles (2)

Spécifique à chaque système d'exploitation.

Sur Unix :



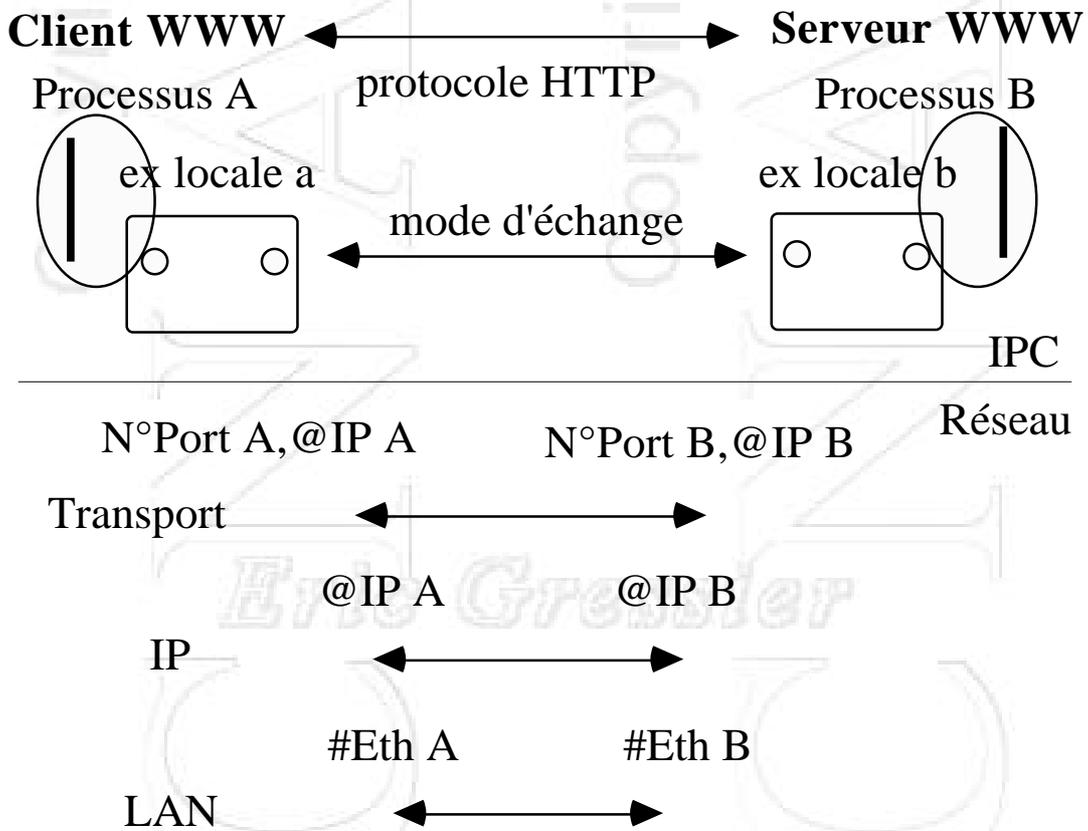
# Communication entre Client/Serveur

Association (Dialogue) IPC:

Local (un Client)

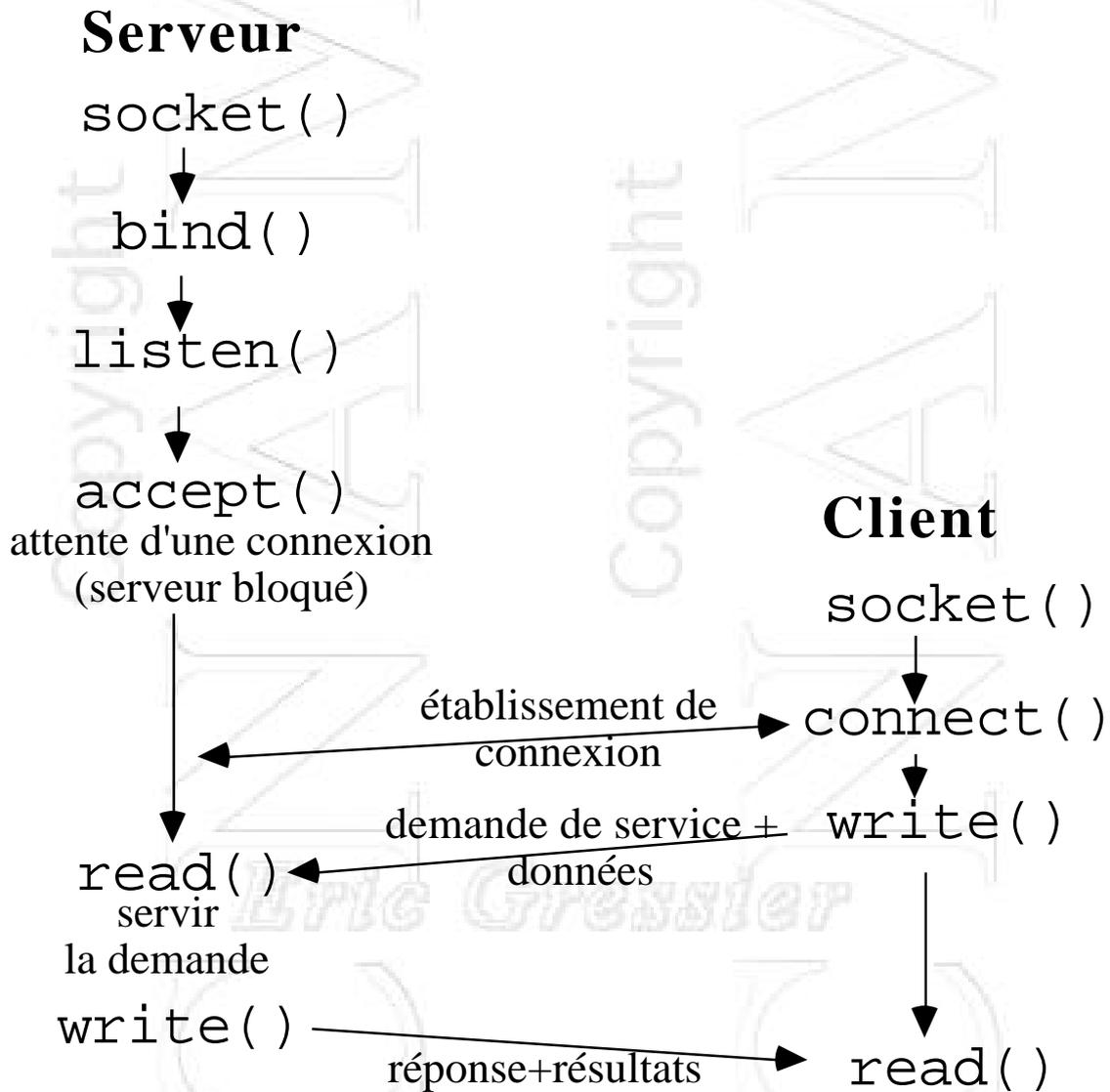
Distant (un Serveur)

{processus,extrêmité,mode d'échange,extrêmité,processus}



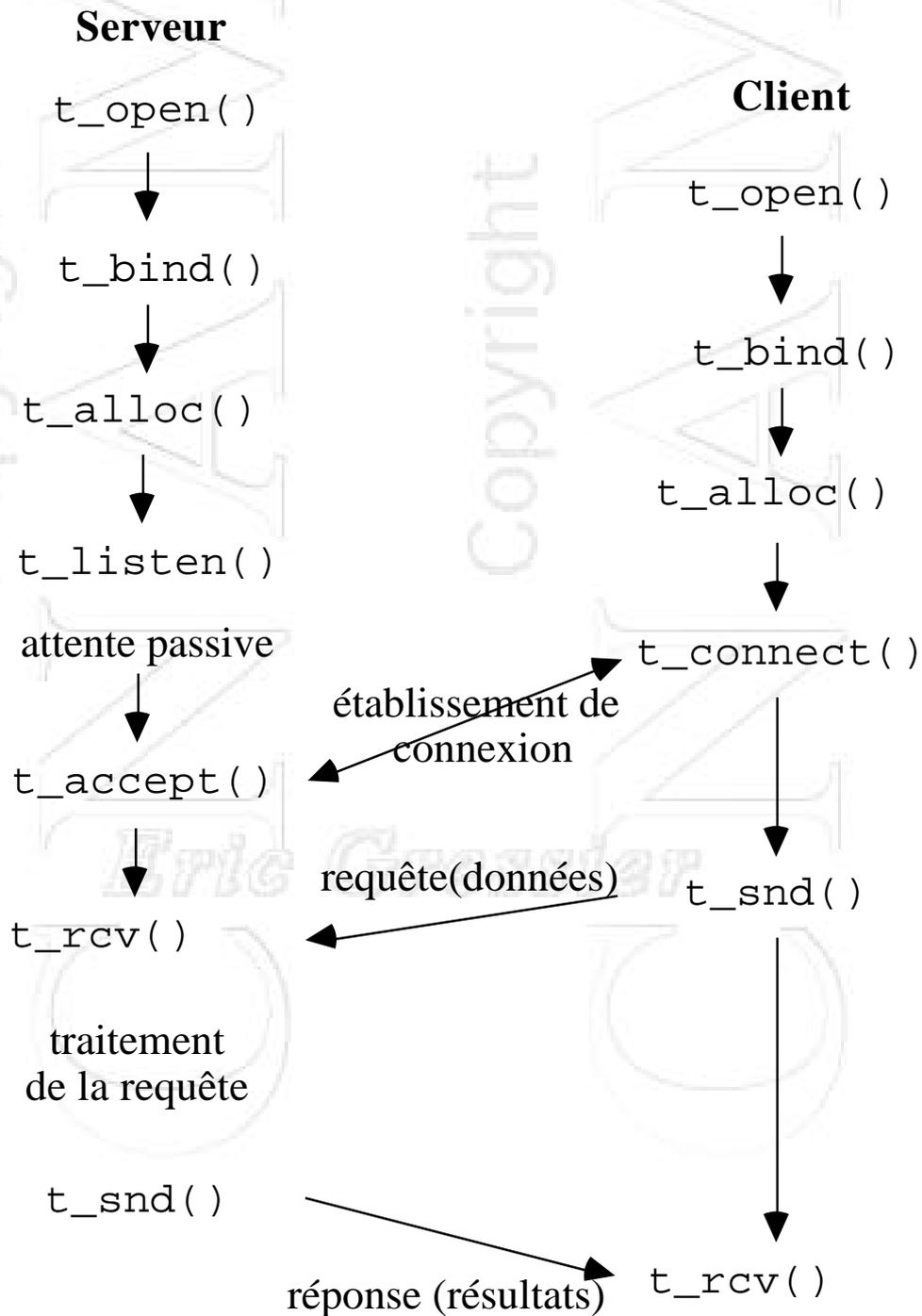
# Programmation réseau - sockets

Client/Serveur mode connecté :



# Programmation réseau - TLI

Client-Serveur en mode connecté :



## Désignation des objets

Les machines sont désignées de manière unique au niveau mondial.

Le mécanisme de nommage est réparti. Le plan de nommage est hiérarchique. On partitionne l'espace des noms, et chaque sous-espace peut être lui-même subdivisé à son tour.

### Exemple : bertrand@asterix.iat.cnam.fr

objet@[nomdomaineobjet].[nomsous-domaine].nomsous-domaine.nomdomaine

"bertrand" : objet (utilisateur)

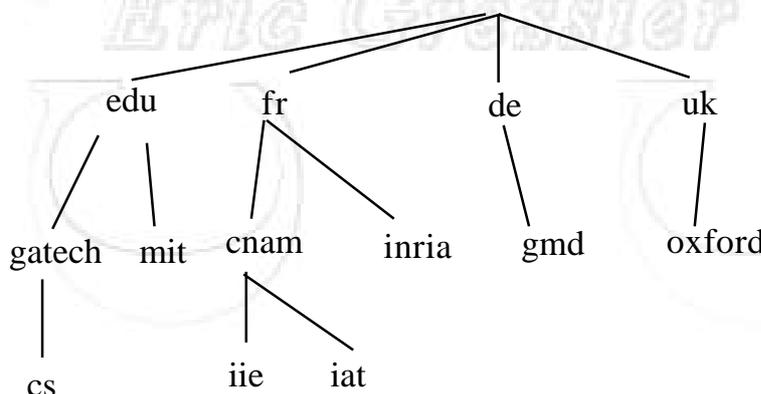
"asterix" : nomdomaineobjet (ressource qui contient l'objet, nom d'une machine souvent)

"iat" : nomsous-domaine (correspond à un organisme, par exemple

"cnam" : nomsous-domaine suit iat (reflète l'organisation de l'institution IAT institut du Cnam)

"fr" : nomdomaine (France)

Le nom de domaine peut être politique (fr, de, jp, nl) ou institutionnel (com, mil, gov, edu) :



## Serveurs de noms

L'utilisateur manipule un nom textuel, la machine utilise une adresse IP.

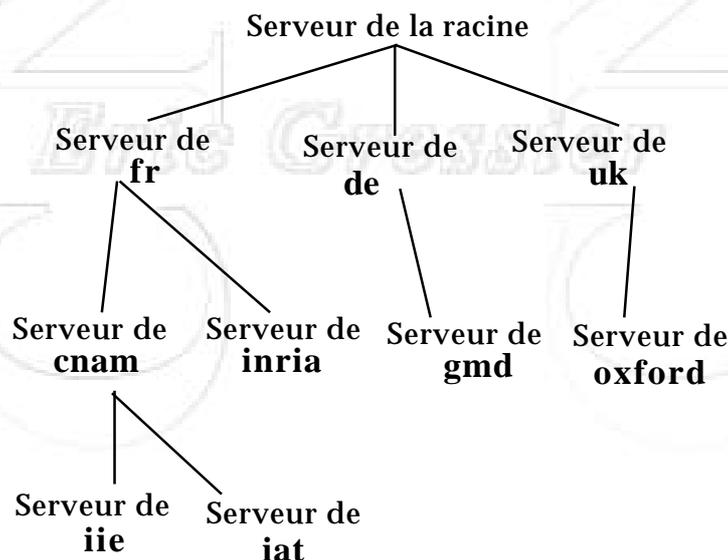
On utilise un annuaire réparti : le “système de noms de domaine” (DNS) qui met en correspondance les noms et les adresses.

Le service de noms réparti repose sur une organisation hiérarchique de serveurs de résolution de noms.

décomposition hiérarchique des noms

<->

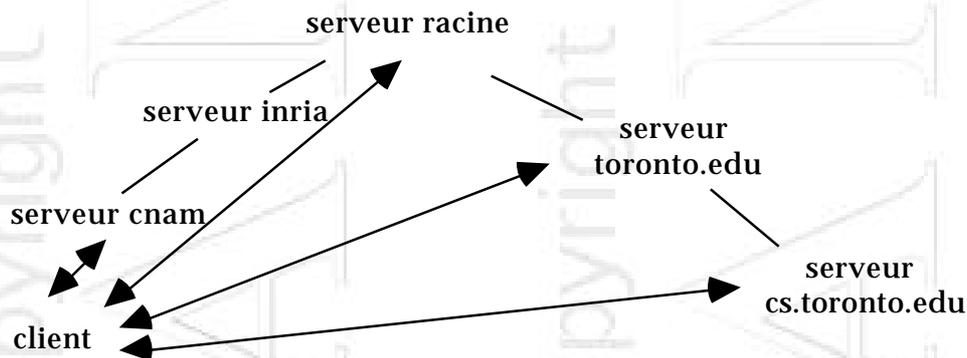
décomposition en serveurs capables de résoudre les noms en adresses.



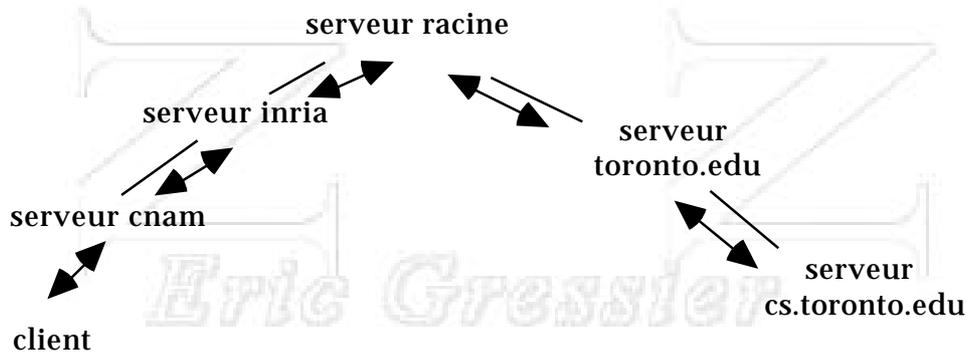
## Résolution de noms

A partir du nom d'une machine récupérer son adresse IP, par exemple : recherche de **ftp.cs.toronto.edu** depuis le **cnam**.

La résolution peut être "itérative"



La résolution peut être "récursive",



**Pour chaque campus indiquer la machine dédiée qui sert à résoudre les noms. Vital!**

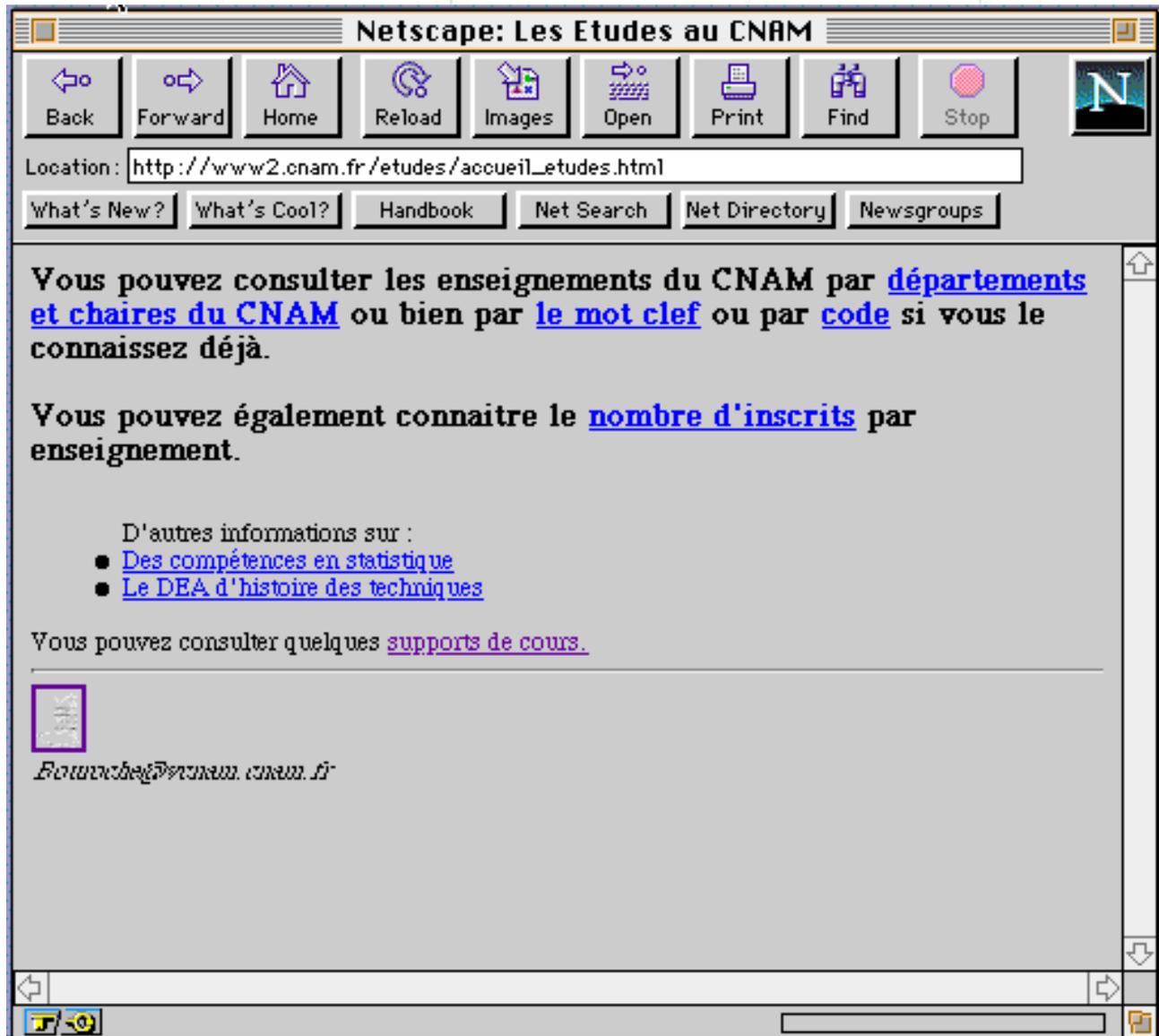
# World Wide Web

Copyright

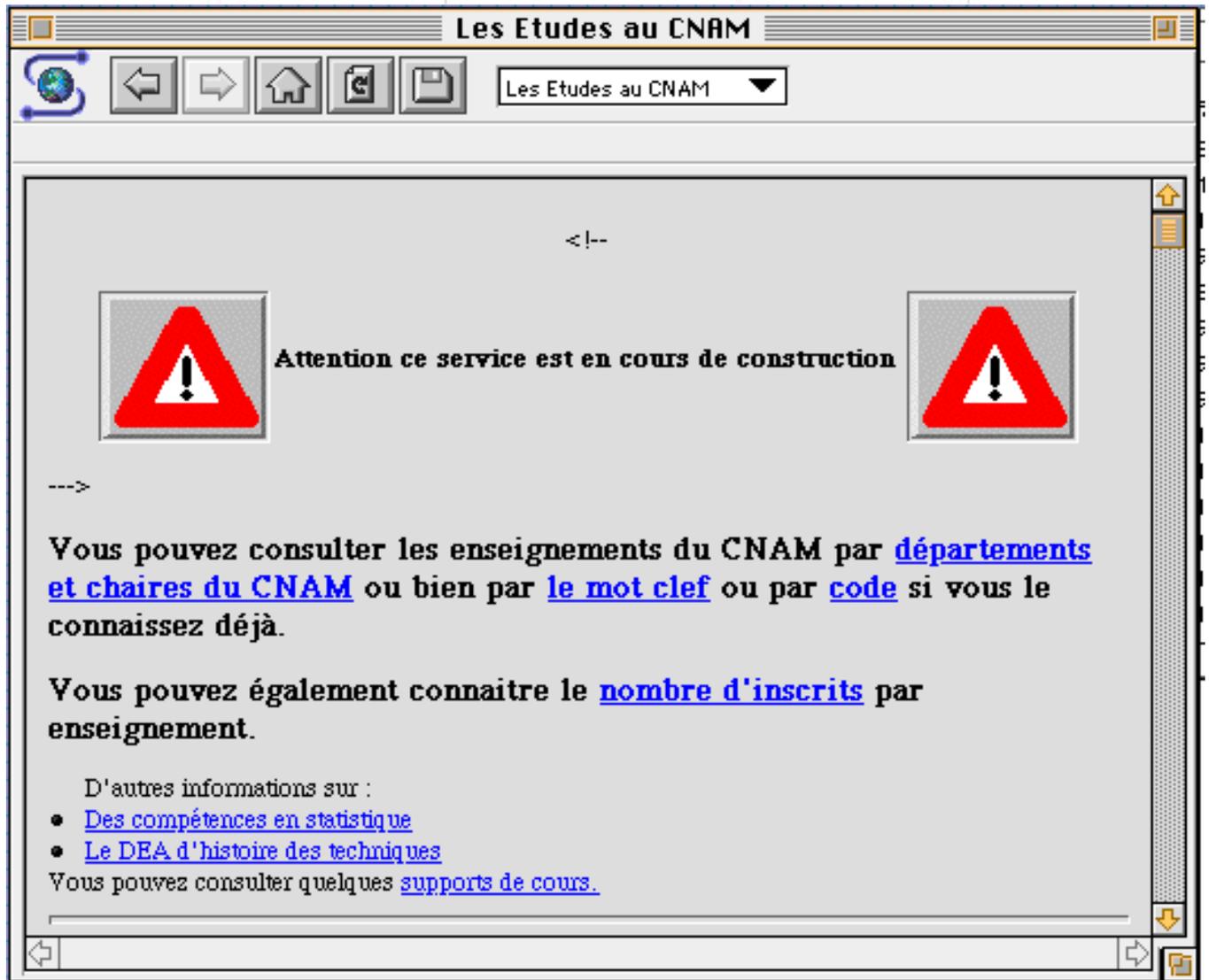
Copyright

Eric Gressier

## Vue : Consultation à partir d'un Client WEB - Netscape



## Vue : Consultation à partir d'un Client WEB - NCSA Mosaic



Notez <!-- --> qui indique un commentaire, et qui n'a pas été compris par le client Mosaic.

## Vue : Description en langage source des Données à afficher

```

<Html>
<Head>
<Title>Les Etudes au CNAM</Title>
</Head>
<Body>
<!--
<p align=center>

<blink>
<b>Attention ce service est en cours de construction</b>
</blink>

<p>
--->
<H3>
Vous pouvez consulter les enseignements du CNAM par
<a href="http://web2.cnam.fr/htbin/sig/?R=">départements et chaires
  du CNAM</a> ou bien par
<a href="http://web2.cnam.fr/~sig/motif.html">le mot clef</a> ou par
<a href="http://web2.cnam.fr/~sig/code.html">code</a> si vous le
connaissez d'ailleurs;.
<p>
Vous pouvez également connaître le
<a href="http://www2.cnam.fr/~madec/cptdep.html">
nombre d'inscrits</a> par enseignement.
</H3>
<br>
<ul>
D'autres informations sur :
<li><a href="/stat/home.html">Des compétences en statistique</a>
<li><a href="/instituts/cdht/dea.html">Le DEA d'histoire des techniques</a>
</ul>
Vous pouvez consulter quelques <a href="/cours/accueil_cours.html">
supports de cours.</a>
<hr>
<a href=""></a>
<address>Bouroche@vcnam.cnam.fr</address>
</Body>
</Html>

```

# WEB : Système d'Information Multimédia d'accès

## Définition officielle :

"wide-area hypermedia information retrieval initiative aiming to give universal access to a large universe of documents".

- Concept d'**hypermédia** qui est l'extension de la notion d'**hypertexte**. Une base de données universelle où les documents de tous types sont identifiés de manière unique, et pointent les uns vers les autres par des liens.
- Interrogation en mode Client/Serveur : 1er client largement répandu, Mosaic a été développé par NCSA (National Center for Supercomputing Applications)

WWW Outils le plus flexible pour explorer l'Internet.

WWW a été développé par le CERN (Laboratoire européen de recherche en physique des particules) à Genève. Le projet a démarré en mars 1989.

## avantages du WWW pour l'utilisateur

A travers le Client :

- Interface graphique avec la souris.
- Possibilités de sauvegarder des documents de différents formats, d'annoter des documents, de laisser des marques (pour un accès ultérieur), d'avoir un historique de la navigation ...
- Accès DIRECT aux serveurs www, ftp anonyme, gopher, telnet, nntp et wais. Il inclut donc 6 clients.
- Visualisation de documents HTML.
- Restitution de différents formats de fichiers sons, video, graphiques ...

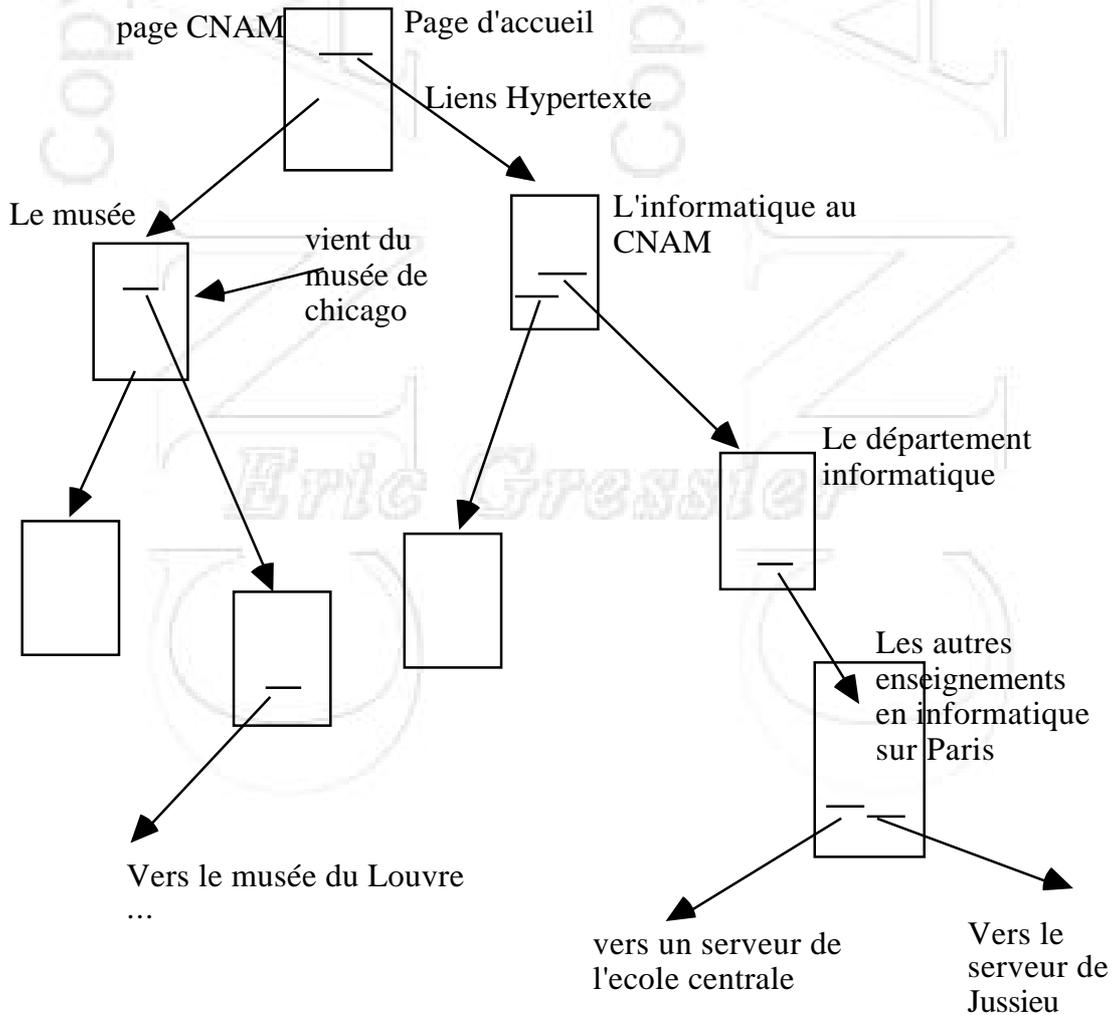
C'est concrètement un outil multimédia disponible.

La mise à disposition de logiciels puissants et gratuits a largement contribué au développement de WWW.

# Les concepts de l'hypertexte

Un document WWW est un document hypertexte:

- Il s'agit d'une méthode de présentation de l'information où certains groupes de mots sont des **liens** vers d'autres documents.
- Les liens vers d'autres documents sont mis en évidence (soulignés, inverse vidéo, couleur).
- Les liens entraînent vers d'autres documents hypertexte qui sont soit stockés localement, soit sur un autre ordinateur de l'Internet.
- La vision du monde de l'information est un graphe:



# Le langage des pages Hypertexte : HTML

**HTML** :Hyper Text Markup Langage :

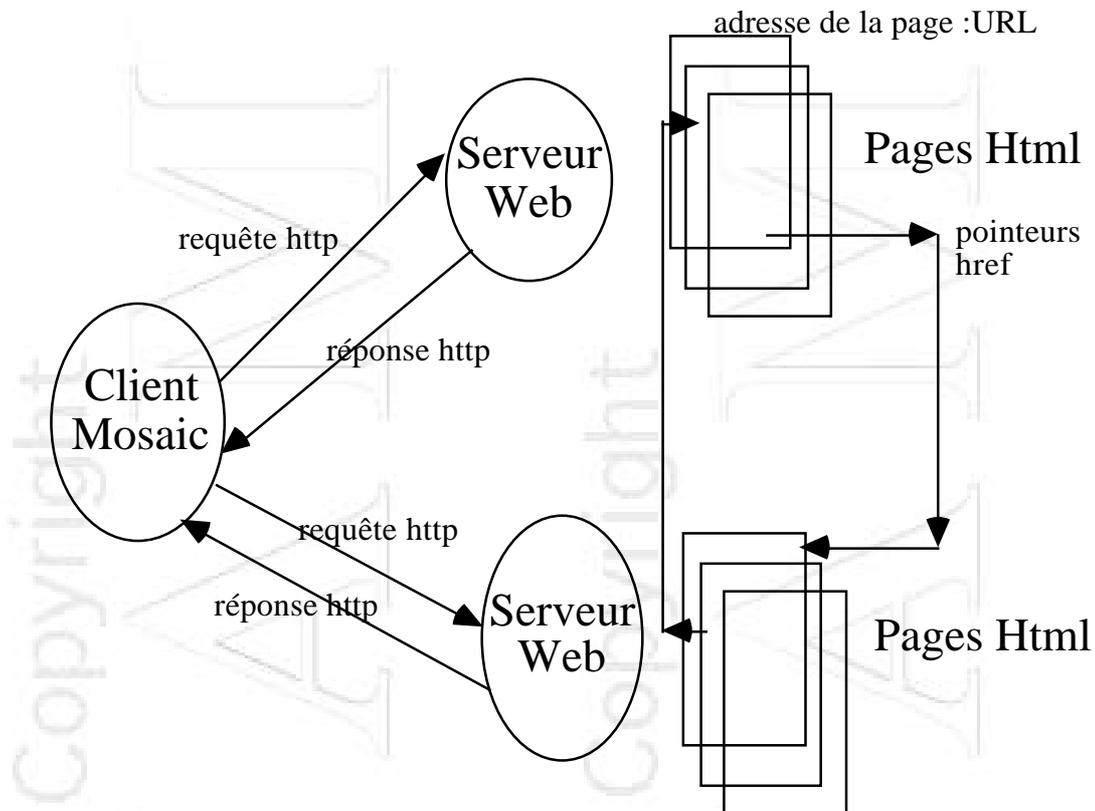
- Généré par les serveurs et interprété par les clients, ce langage est similaire à un sous ensemble de SGML (Standard Generalized Markup Language).
- Il offre les fonctions de base pour permettre un affichage agréable (paragraphe, listes, indentations, titres de paragraphe, styles et polices de caractères, insertion d'images, ...) en restant très simple.
- HTML inclut évidemment les liens hypermedia.

HTML est un sous-ensemble de SGML (Standard Generalized Markup Language)

Exercices sur <http://tulipe.cnam.fr/cours/cours1.html> par Pierre Cubaud et Laurence Duchien.

*Eric Gressier*

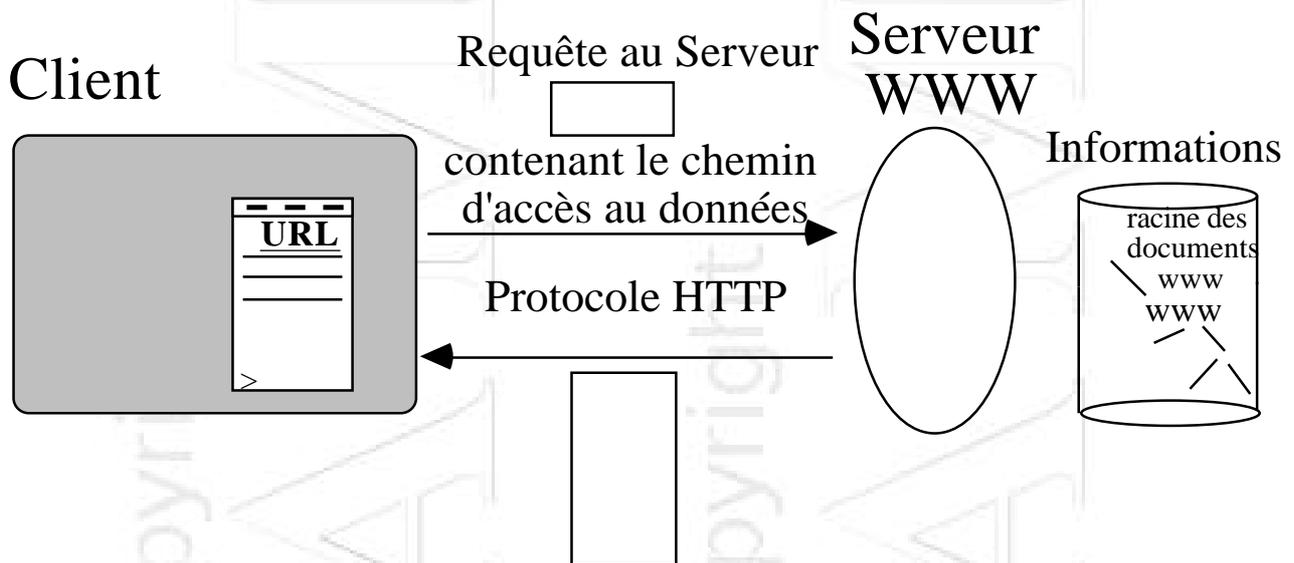
## Intéraction client/serveur WWW



- Un client WWW interroge un serveur WWW quelque part sur la planète.
- Le client propose une interface multimédia.
- Il présente plus que des pages de texte formatées avec des liens vers d'autres documents, c'est un outil multimédia qui gère des sources audio, des images et de la vidéo.
- De plus, il propose une interface simplifiée vers d'autres services (ftp, gopher, news, wais,...).

Lors d'une interrogation, le serveur retourne une page formatée (en html) suivant un protocole de transport (http ou autre). Le client local interprète cette page formatée et l'affiche.

## Transaction WWW Client-Serveur



Message contenant une spécification MIME des données, en particulier le type des données

pour un document décrit en HTML :

Content-type: text/html

MIME (Multipurpose Internet Mail Extensions) spécification qui permet d'échanger d'autres types de document que du texte ascii, par exemple : image, video, audio, message, binaire (application).

## Un peu de vocabulaire

**HTTP** : HyperText Transfert Protocol : protocole de transport de ces pages hypertexte

HTTP est le protocole de communication entre les clients et les serveurs www. Il fonctionne de manière très simple : deux possibilités : requête contenant un pointeur hypertexte (URL) et réponse sous forme de fichiers de données.

**URL** : Uniform Ressource Locator : adresse planétaire d'une ressource.

Un URL, Uniform Resource Locator, est une manière unique et universelle de désigner un document. Il contient la méthode d'accès au document, le nom du serveur et le chemin d'accès au document. Un lien hypermédia est simplement un URL.

*Eric Gressier*

## format d'un URL

protocole : chemin-d'accès

forme BNF :

nomprotocole:nomserveur[:port]{/répertoire}\*/ressource

nomprotocole :

http	HyperText Transfer Protocol
ftp	File Transfer Protocol
gopher	le protocole Gopher
mailto	adresse électronique
news	News du réseau Usenet
nntp	News de Usenet en accès local NNTP
prospero	le protocole Prospero
telnet, rlogin et tn3270	références à des sessions interactives
wais	Wide Area Informations Server

Exemples d'URL :

`http://www.cnam.fr/Marillion/Marillion.html`

`news:fr.rec.cuisine`

`ftp://ftp.cnam.fr/pub/Atari/0new0`

`gopher://gopher.inria.fr`

URL du cnam si nécessaire : `http://www.cnam.fr/`

## Logiciels Clients et Serveurs

### Clients Netscape ou Mosaic :

- . PC (Windows 3.1, OS-2, Windows95, Windows NT) ,
- . Macintosh,
- . Postes avec système X-Window (Unix, VMS)

**Mosaic 2.0.1** retrait sur l'archive ftp.ncsa.uiuc.edu

**Netscape 1.12** infos sur les distributions par l'URL <http://home.netscape.com/comprod/mirror/index.html>, Netscape version 2.0 est annoncé.

Il y a d'autres Clients.

Pour l'enseignement et la recherche, les licences sont souvent gratuites.

### Serveurs :

**HTTPD de W3C** pour les plateformes suivantes: un bon nombre d'Unix et VMS.

#### autres versions

- . Windows/NT : **HTTPS** archive emwac.ed.ac.uk dans pub/https
- . VMS : **HTTP server** accessible par l'URL [http://www.eng.ohio-state.edu/soft/http\\_server\\_1-9a.tar](http://www.eng.ohio-state.edu/soft/http_server_1-9a.tar)
- . Macintosh : **MacHTTP** archive oac.hsc.uth.tmc.edu dans public/mac/MacHTTP

Pour plus d'informations sur les serveurs, consulter :

<http://www.w3.org/pub/WWW/Servers.html>

Pour les outils de conversion ou d'édition HTML, consulter :

<http://www.w3.org/pub/WWW/Tools/Overview.html>

## Projet W3

Le consortium W3C est composé d'industriels et ses produits sont disponibles gratuitement.

Le projet W3 se poursuit, des travaux sont effectués dans un grand nombre de sites. Il est géré par l'INRIA et par le MIT en collaboration avec le CERN.

Pour toute information sur W3C, consulter :

**<http://www.w3.org/pub/WWW/>**

Pour les logiciels associés à WWW, consulter :

**<http://www.w3.org/pub/WWW/Status.html>**

Le projet W3C centralise un grand nombre d'informations sur tout ce qui se passe sur WWW. Quand on cherche qqch, c'est le premier site à consulter. En second, regarder les serveurs NCSA et Netscape.

## Aspects non traités

- . Sécurité :
  - Sécurité du serveur contre les intrusions réseau
  - Confidentialité des données échangées
  - Authentification des clients WWW
  
- . Administration du serveur et de son contenu
  - Droits d'accès
  - Configuration du serveur au lancement
  
- . Le lancement de requête par un serveur, ou les scripts CGI (Common Gateway Interface)
  
- . Methodologie de mise en place d'un service WWW
  - choix du Webmaster, son rôle
  - ouverture des répertoires aux auteurs de pages
  - compte rendu d'expérience des usagers
  - coordination des acteurs
  - dynamique du contenu du serveur