

IFT 6800

Atelier en Technologies d'information

Chapitre 4

Technologies Internet



Plan du cours

1. Introduction
2. Composantes matérielles d'un ordinateur
3. Systèmes d'exploitation (exemple Unix) et Base de données
4. Technologies Internet
5. Introduction à la programmation Internet
6. Introduction à Java script
7. Java



Internet (2)

- Permet de faire "parler" deux ordinateurs différents.
- Communication passe par divers moyens.
- Gérer le réseau entre les deux ordinateurs
 - Choix du chemin (routage: liste des nœuds à emprunter)
 - Gestion des pertes de message
 - Reconfiguration du réseau



3

Internet : Infrastructure de Communication

- Protocoles de communication
 - ✱ TCP/IP, UDP/IP
- Accès aux protocoles
 - ✱ BSD sockets, TLI, Sockets Streams
- Pour désigner un partenaire :
 - ✱ numéro IP, numéro de port (socket)
- Notions de services / protocoles



4

Internet / Services

- Besoin de services standards
 - Service de nom (DNS: Domain Name System)
 - * www.machin.com 129.175.45.33
 - Service d'heure (horloge parlante)
 - Service de transfert de fichier (FTP)
 - Service de transfert de message: e-mail (SMTP)

→ De nombreux autres services



5

Pages Web (1)

- Inventé au CERN (Genève)
- Service Internet
 - Permet le transfert de *pages* d'hyper-texte
 - Informations hypermédia:
 - * texte, des images, des sons, des animations
 - * des liens vers d'autres pages.
- Chaque page est définie par une adresse (URL)
- Protocole de Transfert de pages (HTTP)
 - HyperText Transport Protocol
- Langage de Description de page
 - HyperText Markup Language

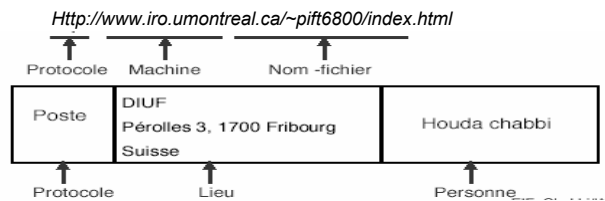


6

Uniform Resource Locator

Permet d'identifier la localisation d'une ressource Internet

- Forme URL même principe qu'une adresse postale
- < Protocole > : // < machine > [: port] < nom-fichier >



7

HyperText Transfer Protocol

Un des protocoles de communication sur le Web

- Basé sur le paradigme de requête/réponse.
- Utilise des connexions TCP/IP (Transmission Control Protocol : service fiable de livraison des flux d'octets) pour transférer les données.
- Principe:
 - Connexion
 - Requête du client
 - Réponse du serveur
 - Déconnexion



8

HyperText Markup Language (1)

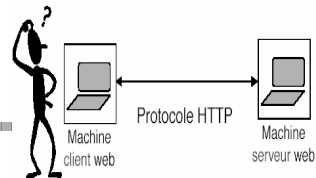
Décrit le format d'un documents Web

- Hypertext
 - Les documents ne sont plus séparées mais reliés les uns aux autres: on peut se déplacer d'un document vers un autre en suivant des liens.
- Markup
 - C'est un texte contenant des balises qui conditionnent l'aspect physiques aux textes.
Ex: `Gras` → **Gras**
- Plusieurs versions
 - dont la plus récente HTML-4.0.1 intégration des scripts, applets, frame, etc.



9

WEB, Comment ça marche?



- Client
 - Une plate forme
 - Du logiciel: *Navigateur/browser*
 - Une connexion au réseau.
- Serveur
 - Une plate-forme: ordinateur
 - Du logiciel: *daemon http Apache par exemple*
 - De l'information

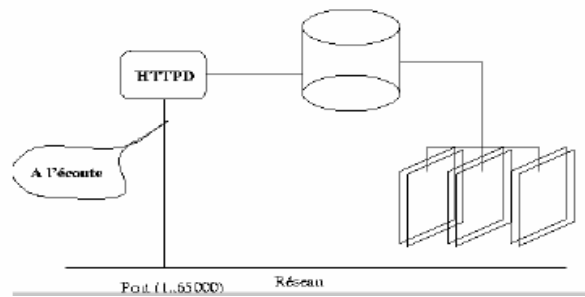


10

WEB, Comment ça marche ?

Exemple pas à pas :

étape I: En attente d'une requête

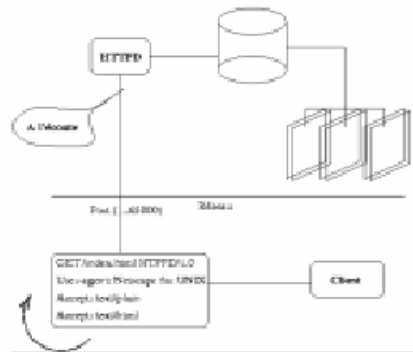


11

WEB, Comment ça marche ?

Exemple pas à pas :

étape II: Émission d'une requête par un client



URL: <http://medit.epfl.ch/index.html>
(Localisation du serveur,
initialisation d'une connexion
réseau bi-directionnelle)

Émission de la requête en ASCII :
GET/index.html HTTP/1.0
(cette chaîne est sauvegardée en
mémoire)



12

WEB, Comment ça marche ?

- **étape III: Analyse de la requête par le serveur**

Le serveur déchiffre la requête et détermine les actions à réaliser (GET, recherche index.html,...)

- **étape IV: Lecture des autres informations sur le réseau**

(Si besoin) : par exemple le type du navigateur, ...

- **étape V: Exécution de la méthode demandée**

Pour exécuter GET, httpd recherche dans ces documents (système de gestion des fichiers)

- **étape VI: Acte final: fermeture du fichier, fermeture de la connexion réseau.**

- **étape VII: Retour à l'étape I.**



13

Le Mail

- **Service de Messagerie**

- Permet d'envoyer des messages (texte).
- Permet de joindre toute sorte de fichiers (attachement).

- **Adresse e-mail**

Serhanim @ iro.umontreal.ca

Nom d'utilisateur

Nom du Serveur



14

Newsgroup

- Groupe de discussion informatique.
 - Groupes locaux / mondiaux.
 - On peut poster des questions / réponses.
 - Très utile pour trouver des informations
 - Chaque groupe est dédié a un sujet précis.
 - news://comp.lang.postscript
 - * Ordinateurs (comp)
 - * Langages (lang)
 - * Postscript
- Type du groupe



15

Intégration Web

- E-mail, newsgroup et d'autres protocoles sont aujourd'hui intégrés dans le web
 - Messages des Newsgroup accessibles par des pages web.
 - Possible de lire son mail par des pages web.
- Aujourd'hui on désigne sous le nom de Web l'ensemble:
 - Pages web
 - Newsgroup
 - Services de mail



16

Philosophie WEB



17

Le Web : une idée de départ

- Réseau de ressources d'information
- Clients de consultation
 - Navigateur, browser...
 - Netscape Communicator, Microsoft Internet Explorer
- Serveurs de documents
 - HTTP, démon HTTP, Serveur Web
 - Apache (www.apache.org), Microsoft IIS, Netscape, Cern
- => Consultation de documents

18

Les acteurs du web

- Cern : ?
- Ncsa : ?
- W3c : MIT + INRIA + Université de Keio
 - Unifie les développements,
 - Propose les standards
- IETF : crée des standards
- Netscape, Microsoft, IBM, Sun...
 - Standard de faits (frames, cookies, plug-ins 's...)

19

Exécution sur les clients

- Les langages Web
 - JavaScript, Jscript
- Les langages d'animation
 - Flash, VRML...
- Les langages de programmation
 - Java
 - tcl/tk (safe_tcl), Caml, Python, Perl,

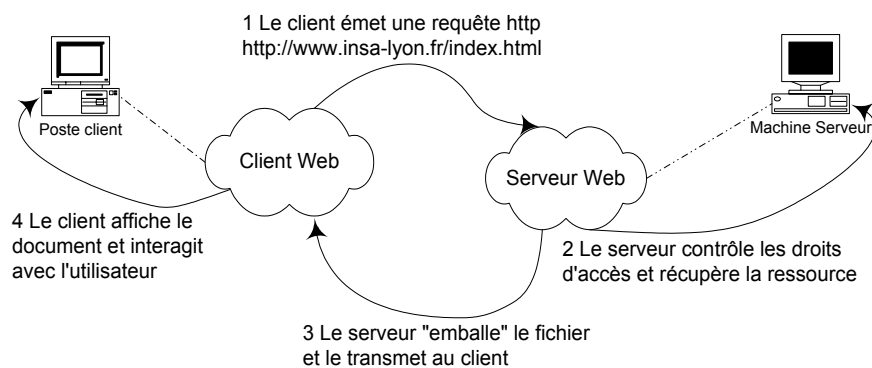
20

Exécutions sur le serveur

- Mécanisme : CGI (Common Gateway Interface)
- Des ressources particulières déclenchent l'exécution d'un programme sur le serveur
- Le document qui correspond à la requête est le flux standard de sortie de ce programme
- Possibilité de passer des paramètres
- Exemple : le Web applicatif !

21

WEB : Architecture documentaire



22

Architecture WEB

Les standards

23

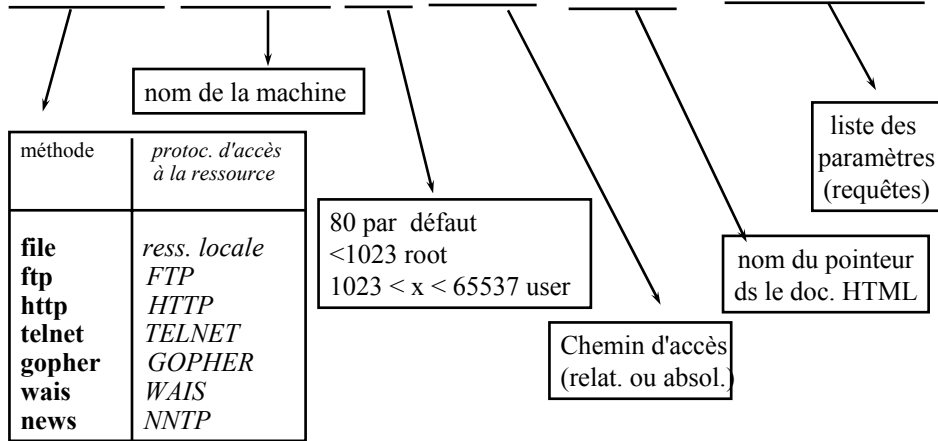
Pourquoi faire des standards ?

- Standardiser les formats des documents :
 - Comment présenter une ressource : MIME
- Identifier de manière unique les documents
 - Comment identifier une ressource : URL
- Standardiser les échanges
 - Comment récupérer une ressource : HTTP
- Comment décrire une ressource
 - Description interne : ML
 - Description des relations entre ressources HT

24

Identification des ressources URI/URL (rfc 1738 : 1994)

méthode://machine:port/fichier[#ancre | ?paramètres]



25

HTTP

- HyperText Transfer Protocol
 - 1/0 rfc 1945 (05 1996) : Internet Informal
 - Tim Berners-Lee, Roy T. Fielding, Henry Frystyk
 - 1/1 rfc 2068 (01 1997) : Internet Proposed Standard
 - Fielding, Getty, Mogul, Frystyk, Berners-Lee
 - Connexions persistantes
- Protocole de type Remote Procedure Call sur TCP
 - Connexion
 - Requête du client
 - Réponse du serveur
 - Déconnexion

26

HTTP : les acteurs

- **Clients** : émet la requête
 - browser, aspirateurs, indexeur, proxy...
- **Serveurs** : renvoie la réponse (interprète la req.)
- **Proxys** : Joue le rôle de serveur pour un client et de client pour un serveur. *Cache* les ressources
- **Gateway** : Idem que proxy mais invisible pour le client
- **Tunnel** : intermédiaire entre un client et un serveur

27

HTTP : Les requêtes

- Format d'une requête
 <Méthode><Chemin><Version_http>
 [<Champ_optionnel>: <Valeur>]
- Méthodes
 - GET, HEAD, POST, PUT
- Champs optionnels
 - User-Agent, If-Modified-Since, Authorization=XXX

28

HTTP Les réponses

- Format d'une réponse
 - <Version_HTTP><Code_Réponse><Texte>
 - Content-Type: <Type_MIME>
 - [<Champ_optionnel>: <Valeur>]

 - <Document>
- Code réponse
 - 100 - 199 : Informatif : 100 Continue
 - 200 - 299 : Requête client réussie : 200 OK, 201 Created
 - 300 - 399 : Requête client redirigée : 301 Moved Permanently, 302 Temporary
 - 400 - 499 : Erreur client : 400 Bad, 401 Unauthorized, 403 forbidden, 404 not found
 - 500 - 599 : Erreur du serveur : 500 Internal Server Error, 501 not implemented

29

HTTP

- Texte libre : description en anglais du code de réponse
- Champs optionnels du serveur
 - Date de la requête : date
 - Date de modification : Last-Modified
 - Identification du serveur : Server
 - Taille du document : Content-Length
- HTTP/1.1
 - Connexion « Keep-Alive »
 - Gestion fine des caches Web
 - Négociation du contenu
 - Virtual Hosts

30

Exemple HTTP

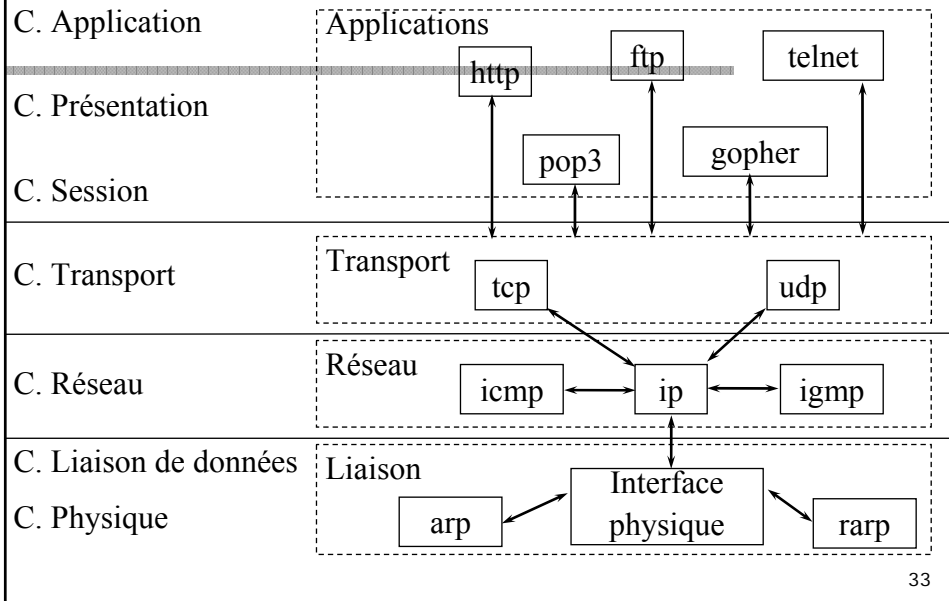
```
telnet lisaix0 80
>GET /index.html HTTP/1.0
>User-Agent: libwww/2.12
>
=> HTTP/1.1 200 OK
=> Date: Fri, 11 Feb 1998 18:30:27 GMT
=> Server: Apache/1.2.4
=> Content-Length: 381
=> Content-Type: text/html
=>
=> <HTML>
=> <BODY>
=>   Bonjour tout le monde ....
=> </BODY>
=> </HTML>
```

31

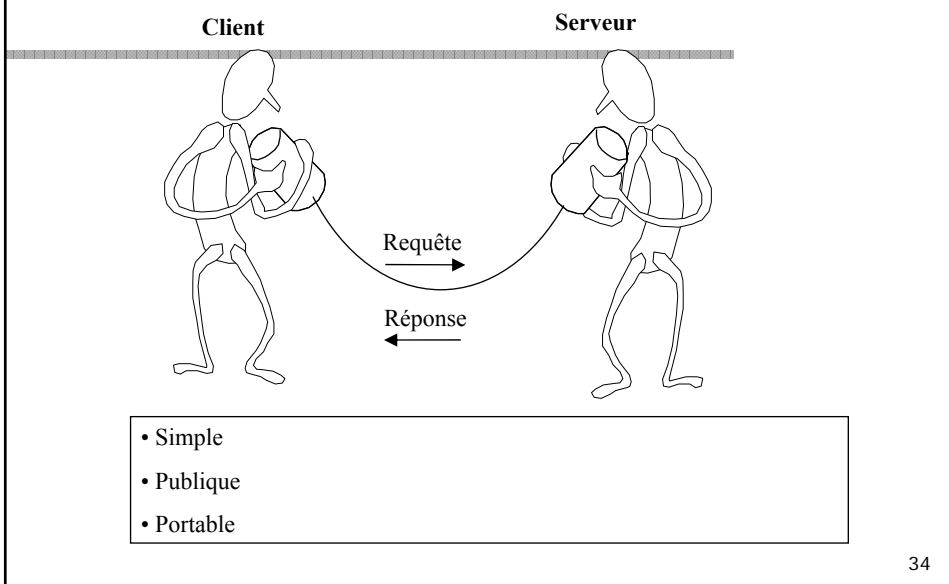
Les Protocoles

32

Protocoles : une approche technique



Protocoles



Caractéristiques d'un protocole

- **Nom** : Simple Mail Transfert Protocol
- **Rfc** : 1830, 1845, 1846
- **Mode de fonctionnement** : Connecté
- **Port de connexion** : 25
- **Commande / Requêtes** : EXPN, QUIT, HELO...
- **Client** : Eudora
- **Serveur** : Sendmail

35

Protocoles : Exemples

Famille	Nom	Client	Serveur	Port
Courrier	SMTP	Sendmail	Sendmail	25
	POP3	Eudora	Popper	110
	IMAP	Eudora	Imapd	143
Transfert de fichiers	FTP	ftp	Ftpd	20/21
Forums	NNTP	Tin	Nntpd	119
Web	HTTP	Netscape	Httpd	80
Conversion IP/Nom	DNS	Resolver	BIND in.named	42/udp

36

Fonctionnement d'un serveur

- Démon
 - C1 Attendre l'arrivée d'une requête émise par un client
 - C2 Démarrer un nouveau serveur traitant la requête client
 - C3 Retour à C1
- Principe de lancement
 - Manuel
 - Inetd
 - StandAlone

37

Exemple de connexion

```
telnet ifhpserv 25
=>220 ifhpserv.insa-lyon.fr ESMTP Sendmail 8.7.1/8.6.9 ready
at Sun, 11 Oct 1998 13:52:07 +0200 (METDST)

EXPN stephane.frenot
=>250 Stephane FRENOT stephf@lisiflory.insa-lyon.fr
```

38

Les intermédiaires

39

Pourquoi des intermédiaires

- Approche par les flux.
- Comment améliorer le flux ?
 - Sécurité
 - Cryptage C/S
 - Vitesse
 - Caches
 - Vitesse + Sécurité ?

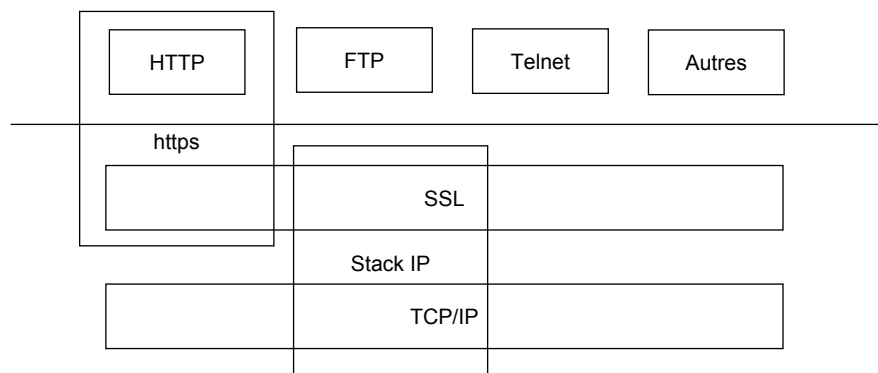
40

Sécuriser le flux

- Man in the Middle :
 - Intermédiaire qui simule le serveur Web (ie Proxy)
- Eavesdropping :
 - Sniffer (protège l'accès aux ressources)
- ==> Sécuriser la connexion
 - 1) Soit sécuriser le transport (HTTPS)
 - 2) Soit sécuriser les applications (S-HTTP)

41

SSL : Secure Socket Layer



42

SSL : moyens

- Communication sécurisé sur média non sécurisé
 - Protection de la connexion
 - Handshake initial pour définir le codage
 - Cryptage symétrique
 - Authentification optionnelle
 - L'entité pair peut s'authentifier en utilisant un cryptage asymétrique (Clé public/privée)
 - Sécurisation des échanges
 - Authentification du message échangé (MAC)
Message Authentication Code

43

Cryptage SSL (HTTPS)

- Encodage sur une clé unique connue du client et du serveur
- Le client génère aléatoirement un nombre
- Puis le transfère de manière cryptée au serveur
- Celui-ci encode les messages à l'aide de ce nombre

44

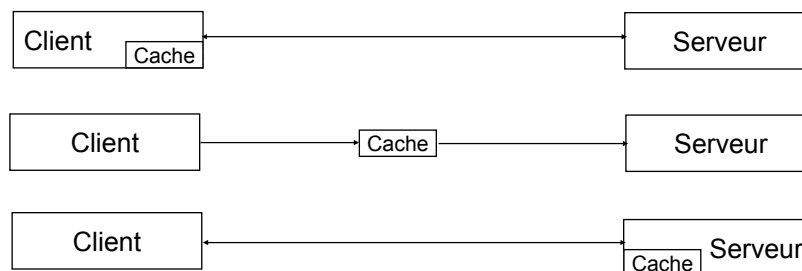
FireWall

- FireWall : Pare-Feu
- Filtrage des paquets
 - Table de filtrage des ports de connexion
 - Inbound HTTP www.interne.com
 - Outbound HTTP www.externe.com
 - Inbound telnet telnet.interne.com
 - Sinon interdit
- Tunneling Protocol
 - Encodage d'une session
 - Réseaux privés virtuels

45

Caches

■ Optimisation des flux



46

Proxy

- Cache des documents transférés
 - Cache mémoire sur le client
 - Cache disque sur le client
 - Serveur Cache local
 - Serveurs Cache nationaux
- Fonctions
 - Disponibilité,
 - Maintenabilité
 - Pre-caching
 - Baisse de la charge
- Difficulté :?

47

Serveur Proxy

```
proxy.univ-lyon1.fr 3128
function FindProxyForURL(url, host) {
  if (isPlainHostName(host))      return "DIRECT";
  if ( dnsDomainIs( host,"univ-lyon1.fr")||
      dnsDomainIs(host,"cpe.fr") ||
      dnsDomainIs(host,"enssib.fr") ||
      dnsDomainIs(host,"cermep.fr") ||
      dnsDomainIs(host,"dr7.cnrs.fr") ||
      dnsDomainIs(host,"www.dsi.cnrs.fr") ||
      dnsDomainIs(host,"insa-lyon.fr"))
      return "DIRECT";
  if (url.substring(0, 5) == "http:" ||url.substring(0, 7) == "gopher:")
      return "PROXY proxy.univ-lyon1.fr:3128; DIRECT";
  if (url.substring(0, 5) == "wais:") return "PROXY web.univ-
lyon2.fr:8001";
  else return "DIRECT";
}
```

48

Caches

- Coopérations de caches
 - Hiérarchies (parent)
 - Transmission au parent si on n'a pas le document
 - Récupération dans le cache fils
<http://www.serveurs-nationaux.jussieu.fr/cache>
 - Coopération d'égal à égal
 - Requête au voisin si on a pas le document
- Caches Web
 - Cache commerciaux ...
 - Cache gratuits : Squid : <http://squid.nlanr.net/Squid>

49

Miroirs (Aspirateurs)

- But :
 - Accélérer les accès aux sites les plus demandés
 - Diminuer le trafic global de l'Internet
- Moyen : copie des sites stratégiques sur différents endroits
 - Logiciel miroir = robot qui browse le Web automatiquement
 - Racine + Règles d'arrêt
 - Extraction des <A...>....
- webcopy, wget, w3mir

50

Anneaux de diffusion

- Interconnexion de sites Web
- Mettre en commun des utilisateurs sur un sujet
- Définir des sous-réseau logiques d'information
- Site central qui diffuse les entêtes des documents
- Diffusion aux abonnés d'un cgi à insérer dans les documents