

Technologies de l'Internet

Partie 1 : Introduction

Iulian Ober

iulian.ober@irit.fr

Définition

Internet est un *réseau de réseaux* informatiques communiquant entre eux grâce à un ensemble de règles appelées *protocoles*, définies indépendamment des constructeurs d'ordinateurs et de réseaux.

Définition Wikipedia:

« ...le réseau public mondial utilisant le protocole de communication IP (*Internet Protocol*). »

Participants

Internet interconnecte des organisations très diverses :

- universités
- organismes de recherche publique
- services gouvernementaux
- entreprises publiques ou privées
- grand public

...

Historique

- 1961-1965 (US) : premiers travaux sur les réseaux informatiques, les télécommunications par paquets. Premiers essais d'interconnexion à longue distance.
- 1967 : débuts du programme ARPANET du département américain de la Défense (DOD). Difficile de discerner entre les vrais objectifs du programme et les mythes:
 - faciliter les télécommunications entre les chercheurs (militaires) Une des premières applications est le courrier électronique.
 - normaliser les techniques de connexion entre les terminaux et les ordinateurs du DOD
 - <mythe>construire un réseau de communication capable de résister à une guerre nucléaire, à des sabotages, etc., en évitant la centralisation excessive des infrastructures. *Lorsqu'un des centres (nœuds) est virtuellement détruit, les données doivent emprunter d'autres chemins et d'autres nœuds pour atteindre les destinataires désignés.*</mythe>

Historique

- **1969 : 2 nœuds (UCLA, Stanford)**
- **1971 : 21 nœuds**
- 1974 : adoption du protocole IP permettant une architecture ouverte et extensible.
Ce protocole est toujours utilisé en 2007.
- 1979 : création du réseau civil USENET, utilisé pour des discussions thématiques sur des forums (newsgroups).
- 1982 : L'accès au réseau est accordé gratuitement.
- 1983 : La National Science Foundation (NSF) finance la mise en réseau de 60 universités américaines et 3 européennes. Les stations de travail apparaissent.
- 1984 : Scission d'ARPANET entre MILNET (réseau militaire) et INTERNET. La NSF prend le relai d'ARPA et assure le développement d'Internet qui intègre le NSFnet.
- **1984 : ~1000 nœuds**
- 1991 : Apparition des fournisseurs d'accès commerciaux
- 1992 : Apparition du World Wide Web (WWW), développé par le CERN à Genève. Il constitue un moyen simple d'organiser l'information et la navigation sur Internet par des hyperliens, en utilisant le format de fichier HTML.
- **2003 : ~180 mil. nœuds, ~600 mil. utilisateurs**

Qu'est ce que l'Internet

- une interconnexion de réseaux
- une pile de protocoles (IP, TCP, UDP, FTP, HTTP, ...)
- un espace de nommage (.org, .com, .fr → domaines et sous-domaines)
- un ensemble d'organisations de normalisation et les normes associées (IETF, IANA, W3C, ...)

Les protocoles de l'Internet (vue simplifiée)

couche application
(hôtes)

FTP, HTTP, Telnet, SSH,...
DNS

couche transport
(hôtes)

TCP
(Transmission
Control Protocol)

UDP
(User Datagram
Protocol)

couche réseau
(hôtes, routeurs)

IP (Internet Protocol)

IP

- couche « réseau »
- commutation de paquets (datagrammes)
 - 2 hôtes peuvent communiquer des paquets sans avoir connaissance de l'architecture du réseau (routes, sous-réseaux, sur Ethernet/ATM/FDDI/...)
 - chaque hôte est identifiée par son **adresse IP** AA.BB.CC.DD (où XX est une paire de chiffres hexa – IPv4)
- type « meilleur effort »
 - perte de paquets, duplication, arrivée en désordre, modification des paquets
- IPv4 → 2^{32} adresses ≠.
Nouvelle version IPv6 → adresse sur 128 bits.
Normalisé, en cours d'adoption en pratique.

UDP

- couche « transport » rudimentaire
- commutation de paquets
- ajoute à IP:
 - multiplexage/démultiplexage (ports)
 - somme de contrôle
- utilisé par des applications « légères » qui n'ont pas besoin d'un mode « connexion » (sessions courtes, besoin de rapidité).
E.g., DNS

TCP

- couche « transport » plus solide
- mode « connexion ». Une session est composée de:
 - établissement de la connexion
 - transferts de données (bidirectionnel)
 - fin de la connexion
- services
 - multiplexage/démultiplexage (ports)
 - transfert fiable – ordonné et sans erreurs
(n° de séquence, contrôle d'erreurs, retransmission)
 - contrôle du flux et de la congestion
(adaptation de la vitesse de transmission à la liaison ou à la vitesse du récepteur)

Qu'est ce que l'Internet

- une interconnexion de réseaux
- une pile de protocoles (IP, TCP, UDP, FTP, HTTP, ...)
- un espace de nommage (.org, .com, .fr → domaines et sous-domaines)
- un ensemble d'organisations de normalisation et les normes associées (IETF, IANA, W3C, ...)

DNS

« Domain Name System »

- système / protocole permettant d'établir une **correspondance** entre une adresse IP et un nom de hôte

Exemple : `www.iut-blagnac.fr` ↔ `193.54.227.200`

Domaines – espaces de noms

- Le monde (Internet) est découpé en domaines hiérarchisés
- Le « top level domain », le dernier suffixe dans le nom, désigne
 - le pays (fr = France, de = Allemagne, us = US)
 - ou la « catégorie » du domaine (org = non-profit; com, biz = commercial; edu = education/recherche (US);...)
- Les sous-domaines sont souvent la désignation des entreprises ou des institutions
 - iut-blagnac.fr, sncf.fr
- Il peut y avoir des sous-sous-domaines à plusieurs niveaux
 - comlab.ox.ac.uk
- Enfin, le premier composant d'un nom désigne le hôte
 - www.iut-blagnac.fr → machine « www » du domaine « iut-blagnac.fr »

DNS - suite

■ système réparti – hiérarchique:

- **un ensemble de 13 root-servers**

NASA, DoD, ICANN, VeriSign, ...

- **un/plusieurs serveurs par domaine, faisant autorité**

e.g. ns1.cict.fr pour iut-blagnac.fr

- **serveurs intermédiaires (des FAI) faisant office de « cache »**

e.g., irit-psa.irit.fr résout aussi le nom « www.iut-blagnac.fr », mais sa réponse ne fait pas autorité

Applications (et protocoles de niveau « application »)

Dans la suite, on se limite aux plus courantes / traditionnelles:

- e-mail
- news
- telnet
- ftp
- web

D'autres applications incluent :

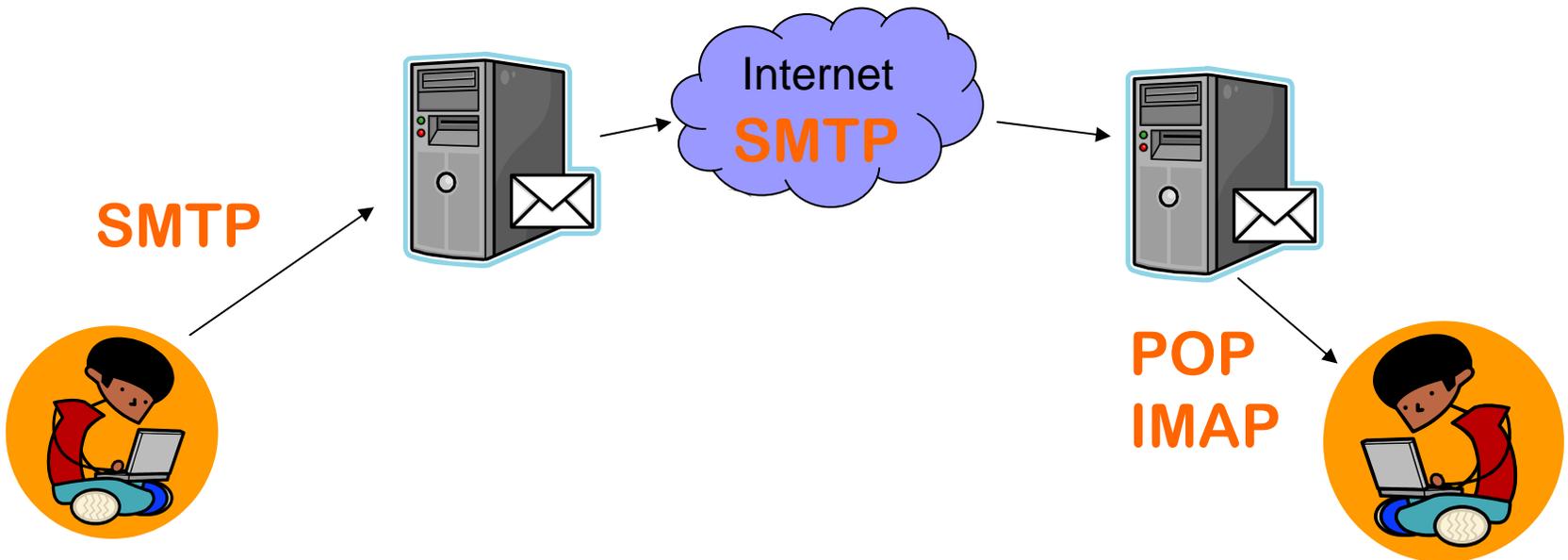
- échange de flots multimédias
- calcul réparti
- instant messaging

...

E-mail

Utiliser l'Internet comme la poste:

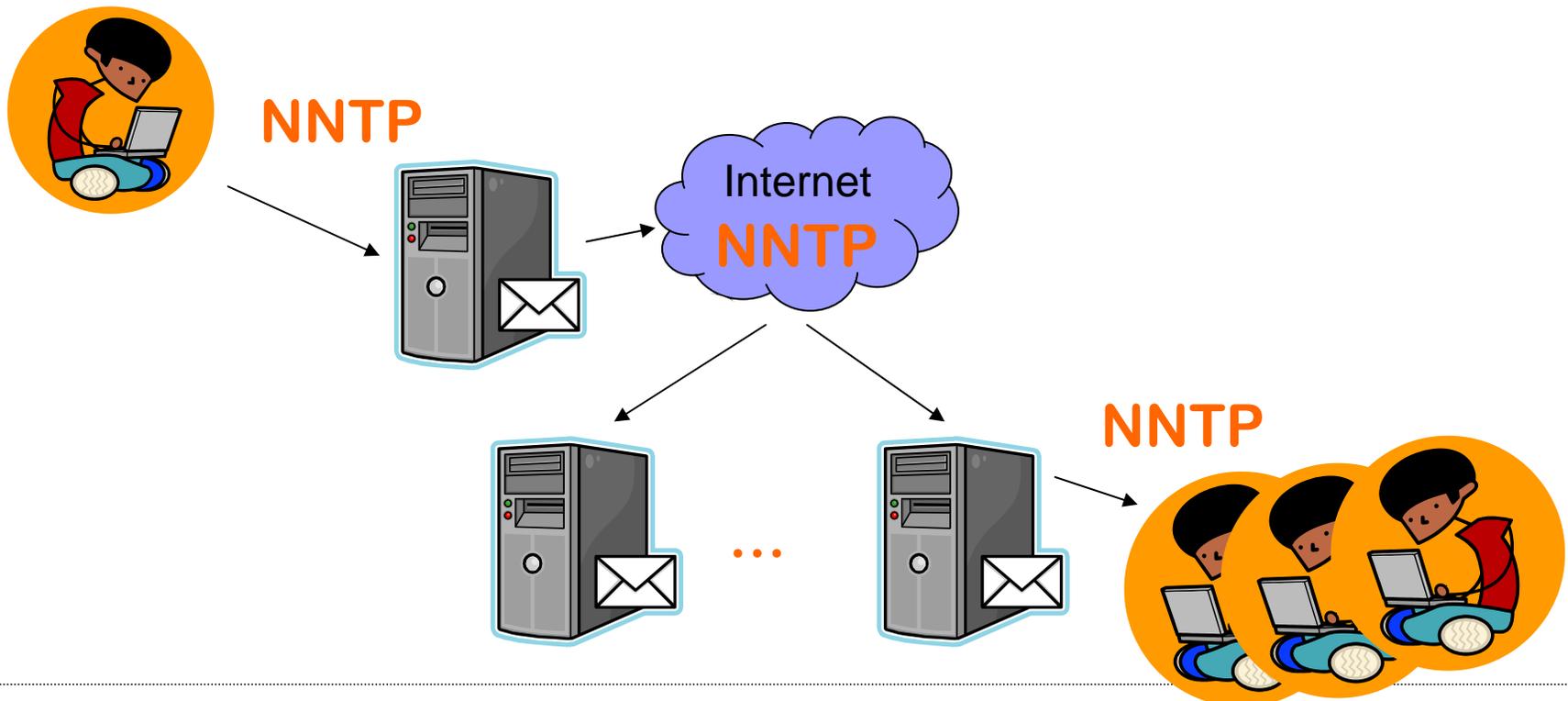
- déposer un message dans la boîte aux lettres de son correspondant, *qu'il soit ou non devant une machine.*
- le correspondant sera capable de le lire à sa prochaine connexion.



News

Forums de discussions basées sur le protocole NNTP

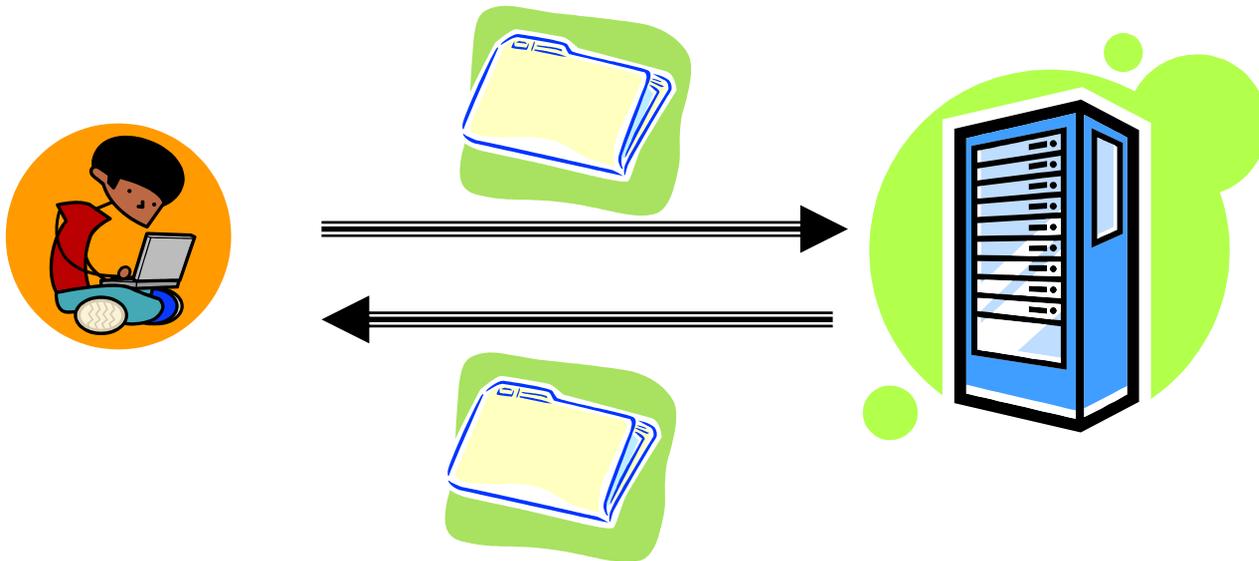
- forums actuels souvent basés sur le web (HTTP)
- ...mais des puristes survivent, surtout dans les académiques, professionnelles ou techno-geek
- exemples : comp.lang.c++,



FTP

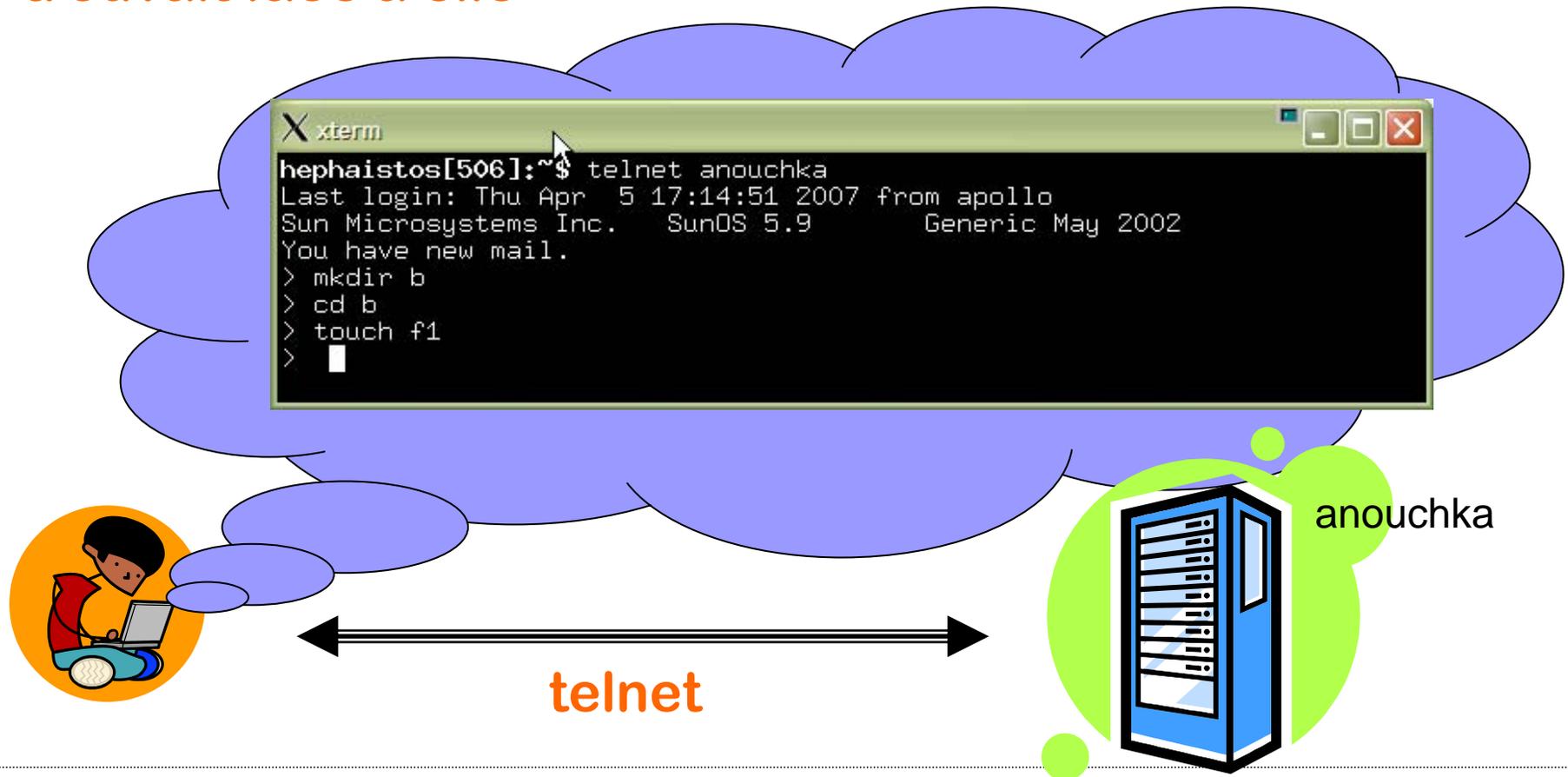
« File Transfer Protocol »

- service d'échange de fichiers
- permet de **déposer** des fichiers sur une machine distante, ou de **télécharger** des fichiers sur sa machine



Telnet

Permet à un utilisateur d'Internet de se connecter et donc d'utiliser à distance une machine **comme si il se trouvait face à elle**



Le World-wide Web (www)

- un système **hypertexte** public fonctionnant sur Internet et qui permet de consulter, avec un **navigateur**, des **pages** mises en ligne dans des **sites**
- service le plus connu de l'Internet, maintenant
- relativement récent
 - 1989-1992, projet **hypertexte** du CERN, Genève à l'origine du **langage HTML** et du **protocole HTTP**
 - 1992 : le premier navigateur graphique – Mozaic (NCSA, US)

Hypertexte

Référence à la capacité d'inclure dans un document

- du texte, des images, des sons, des vidéos, etc..., bref du multi-média
- des **liens non-ambigus** vers d'autres documents, qui permettent de les **trouver automatiquement** et les **ouvrir**

Le format hypertexte « standard » du Web est le HTML (*Hypertext Markup Language*)

Le **Navigateur** est l'outil qui permet de visualiser les document et ouvrir les liens.

URL (Uniform Resource Locator)

Chaîne de caractères construite selon un « standard » pour désigner une **ressource**, notamment dans un lien hypertexte

- une autre page hypertexte (via http ou ftp)
- un flux vidéo ou audio (via mms, rtsp)
- la boîte mail d'une personne (via smtp)
- le téléphone classique (RTC) de quelqu'un
- ...

Structure: « **protocole:chemin** ».

La forme du chemin dépend du protocole:

- **ftp://ftp.netscape.com/pub/communicator**
chemin: // nom ou IP de l'hôte / chemin dans le système de fichiers
- **http://fr.wikipedia.org/wiki/IP**
chemin: idem
- **mailto:iober@irit.fr**
chemin: adresse mail
- **callto:+33562747540**
chemin: no. téléphone
- **news:comp.theory**
chemin: nom du groupe de news

Internet et sécurité

Hypothèse réaliste : tout paquet IP peut être **corrompu par une personne malveillante.**

Conséquences:

- HTTP: la page web que vous visualisez peut ne pas être celle envoyée par le serveur
- HTTP, FTP: le fichier téléchargé peut être corrompu (e.g., virus)
- DNS : l'adresse IP avec laquelle vous communiquez ne correspond pas vraiment au nom de hôte que vous avez spécifié

...

Sécurité : introduction aux solutions

Il appartient aux participants (couches supérieures) d'un protocole basé sur IP d'assurer la sécurité de leur conversation.

⇒ la majorité des protocoles classiques vus présentent maintenant une **version sécurisée**

- **IPsec** : sécurité au niveau réseau
- **SSL, TLS** : sécurité au niveau transport, remplace TCP, basé sur IP (pas sur IPsec !)
- **HTTPS** : HTTP au dessus d'une liaison TLS (au lieu de TCP).
Le même principe peut s'appliquer à POP, IMAP, SMTP, ...
- **SSH, SFTP** : versions sécurisées de Telnet, FTP (mais pas basées sur TLS !)

Qu'est ce que l'Internet

- une interconnexion de réseaux
- une pile de protocoles
(IP, TCP, UDP, FTP, HTTP, ...)
- un espace de nommage
(.org, .com, .fr → domaines et sous-domaines)
- un ensemble d'organisations de normalisation et les normes associées
(IETF, IANA, W3C, ...)

Acteurs de la gestion et de la normalisation

- **Internet Society (ISOC)**

- **Internet Architecture Board (IAB)**

- Mission: orientations architecturales à long terme

- **Internet Engineering Task Force (IETF)**

- Mission: standards de l'internet (e.g., protocoles)

- **Internet Research Task Force (IRTF)**

- Mission: recherche à + long terme sur les protocoles

- **Internet Assigned Numbers Authority (IANA)**

- Mission: gestion des noms de domaine top-level, de l'espace d'adressage IP,...

- **World Wide Web Consortium (W3C)**

- Mission: normalisation des technologies du web (HTML, XML, CSS, SOAP,...)

Les documents de l'IETF

- **Internet Drafts** – documents de travail
- **RFCs (Request For Comments)**
 - ⇒ processus de maturation des standards
 - experimental
 - informational
 - best current practice
 - standards track
 - proposed standard
 - draft standard
 - internet standard
- **Internet Standards**
 - version finale d'un RFC
 - exemple : RFC 0791 définissant le protocole IP
<http://www.ietf.org/rfc/rfc0791.txt>

Les documents du W3C

Processus de maturation de standards:

- Working Draft (WD)
- Last Call Working Draft
- Candidate Recommendation (CR)
- Proposed Recommendation (PR)
- W3C Recommendation (REC)

Exemple : la recommandation définissant le langage XML:

<http://www.w3.org/TR/2006/REC-xml-20060816/>

Structure du cours

- fichiers structurés : notions du langage **XML**
- document hypertexte : langage **HTML**
- feuilles de style: langage **CSS**
- le protocole **HTTP**
- sécurité : la **PKI** (public key infrastructure) et ses applications

XML

- langage standard (recommandation W3C) permettant de **définir des données structurées**
- un fichier XML est un fichier texte respectant un format particulier
- structuré en **éléments** à l'aide de **balises** :

```
<unelement>  
contenu  
</unelement>
```
- La **balise ouvrante** peut contenir des **attributs** :

```
<unelement attr="valeur attribut">
```

exemple XML

```
<?xml version="1.0" encoding="iso-8859-1" ?>
<!-- commentaire : voici mon curriculum vitae -->
<cv>
  <!-- commentaire : état civil -->
  <identité naissance="1980">
<nom>Moustique</nom>
<prénom>Jules</prénom>
<prénom>Édouard</prénom>
<nationalité>grolandaise</nationalité>
</identité>
  <!-- commentaire : mes diplômes maintenant -->
<diplome année="2005" intitulé="Master ID" mention="TB" />
<diplome année="2003" intitulé="Licence " mention="AB" />
</cv>
```