Le web, ses normes, sa gouvernance

Guillaume Sire

MCF – Université Toulouse 1 Capitole

Objectif de la formation

- Comprendre comment Internet et le web fonctionnent, sur quels protocoles ce fonctionnement repose, quels acteurs instituent ces protocoles et quelles sont les principales controverses et les perspectives concernant la gouvernance

Introduction : la dialectique des réseaux

La dialectique du réseau

- Quoi
  - Le dur : câbles, routeurs, serveurs, modem, satellites, ordinateurs
  - Le mou : logiciels, langages de programmation, langages de description, normes
  - Le contenu : vidéo, texte, son, images, fichiers

- Qui
  - Les producteurs de normes
  - Les producteurs de contenants
  - Les producteurs de logiciels
  - Les producteurs de contenus
  - Les états, les juges, les avocats, la police
  - Les utilisateurs

Utopies/ dystopies

« La table de travail ne serait plus chargée d'aucun livre. »,

*Traité de documentation (*Paul Otlet, 1934)

« Des formes entièrement nouvelles d'encyclopédies vont apparaître, prêtes à l'emploi avec un réseau de pistes fonctionnant par association les traversant, prêtes à être insérées et amplifiées dans le memex. […] »

As We may think (Vannevar Bush, 1945)

« *un réseau de tels [ordinateurs], connectés les uns aux autres par des lignes de télécommunications large bande, fournirait les fonctions de bibliothèques actuelles couplées avec les avancées faites dans le stockage et la récupération d'informations et [d'autres] fonctions symbiotiques.»*

Man-Computer Symbiosis (Licklider, 1960)

« Il n'y aura plus qu'une culture, comme si le monde n'était qu'un seul et même village, une seule et même communauté, où l'on vivra dans un même temps, au même rythme et donc dans un même espace »

The Medium is the Message (McLuhan, 1967)

« Vos concepts légaux de propriété, d'expression, d'identité, de mouvement, de contexte, ne s'appliquent pas à nous. Ils sont basés sur la matière, et il n'y a pas ici de matière. »

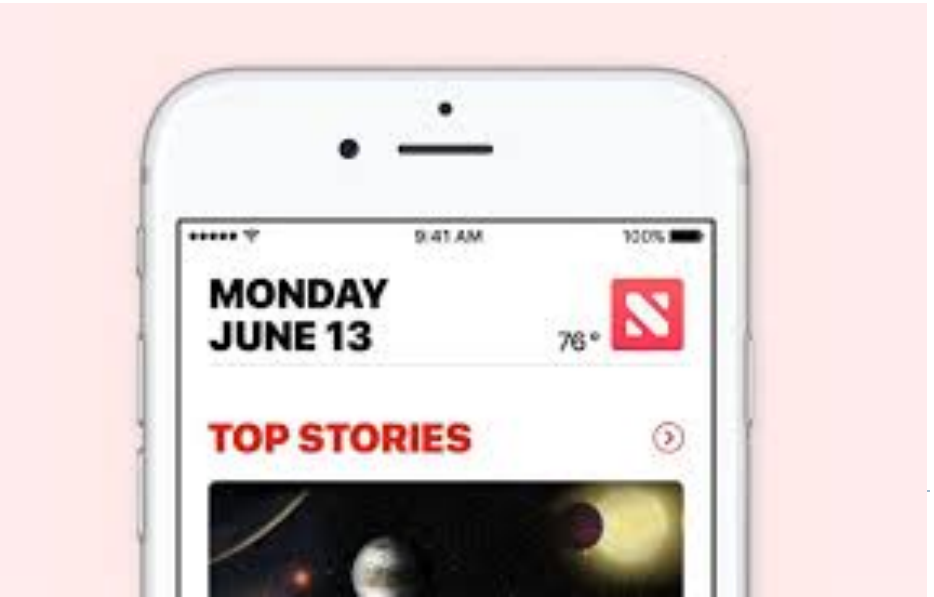
Déclaration d'indépendance du cyberespace (Barlow, 1996)

Matériel / Immatériel

- Désolidarisation du contenu et du contenant

Désintermédiation / Réintermédiation

- Publication (gatekeepers) / Publicisation (gatewatchers)
- Producteurs de contenus / producteurs de logiciels automatisés hiérarchisant le contenu

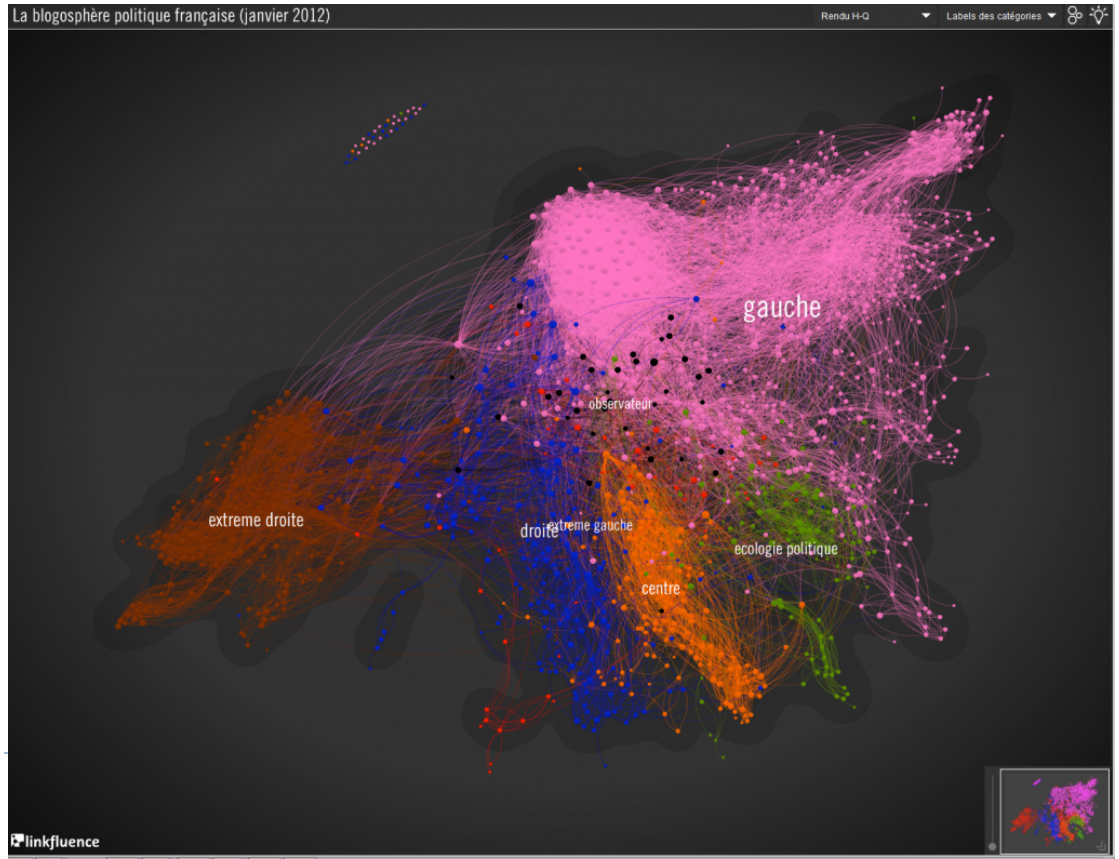


Google news

Accès/ Propriété

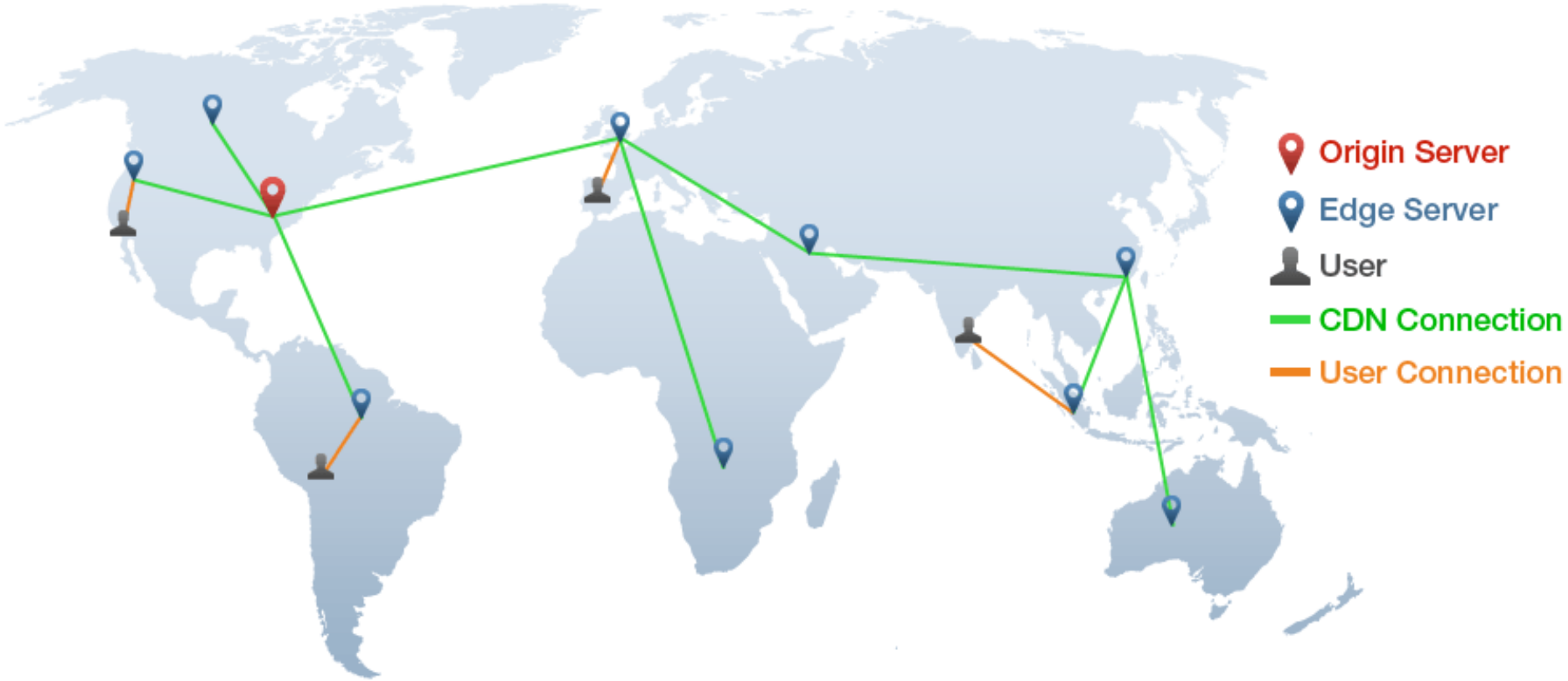
Mondial / Local

- La frontière de la langue
- Des bulles hypertextes
- Les « jardins murés »

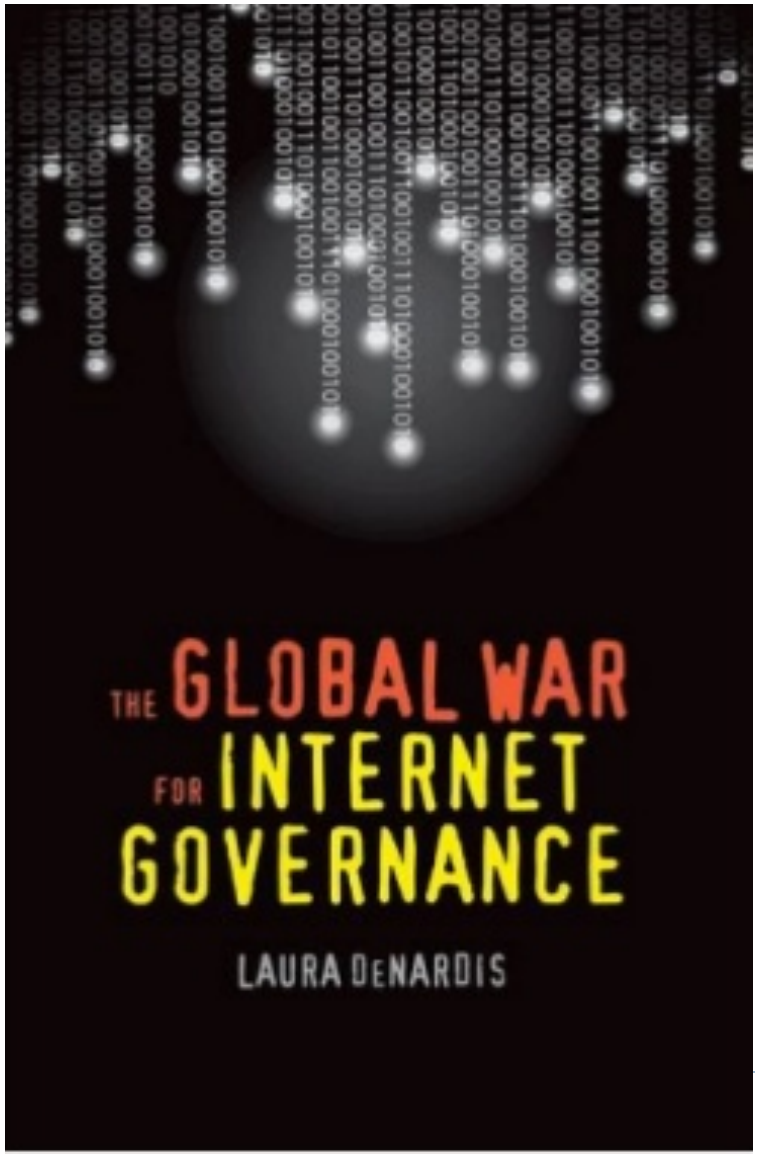


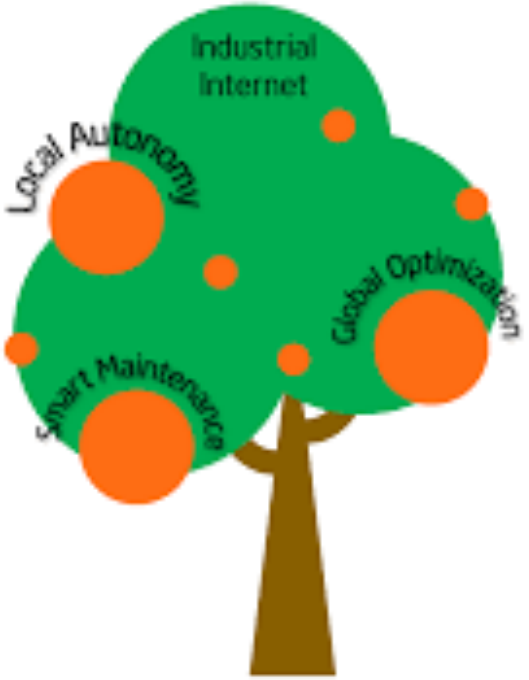
Partout / Nulle part

- Où sont les serveurs ? Quel droit appliquer ?



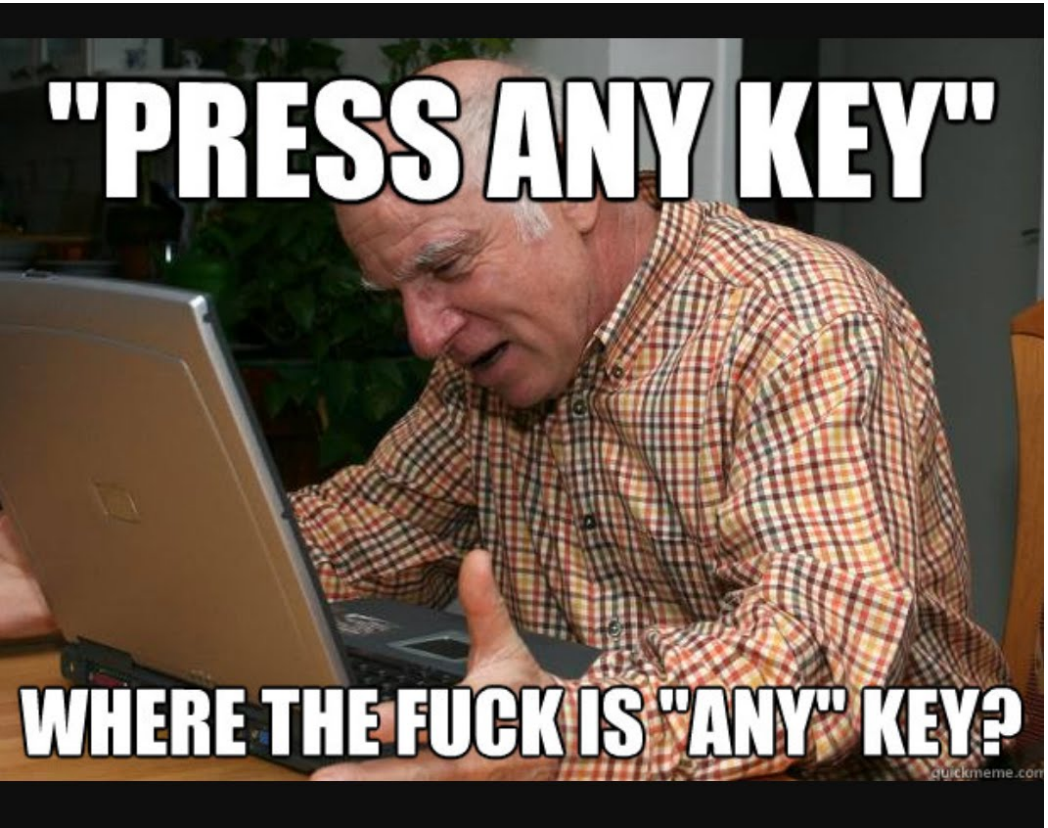
Gouvernance globale/ gouvernance locale





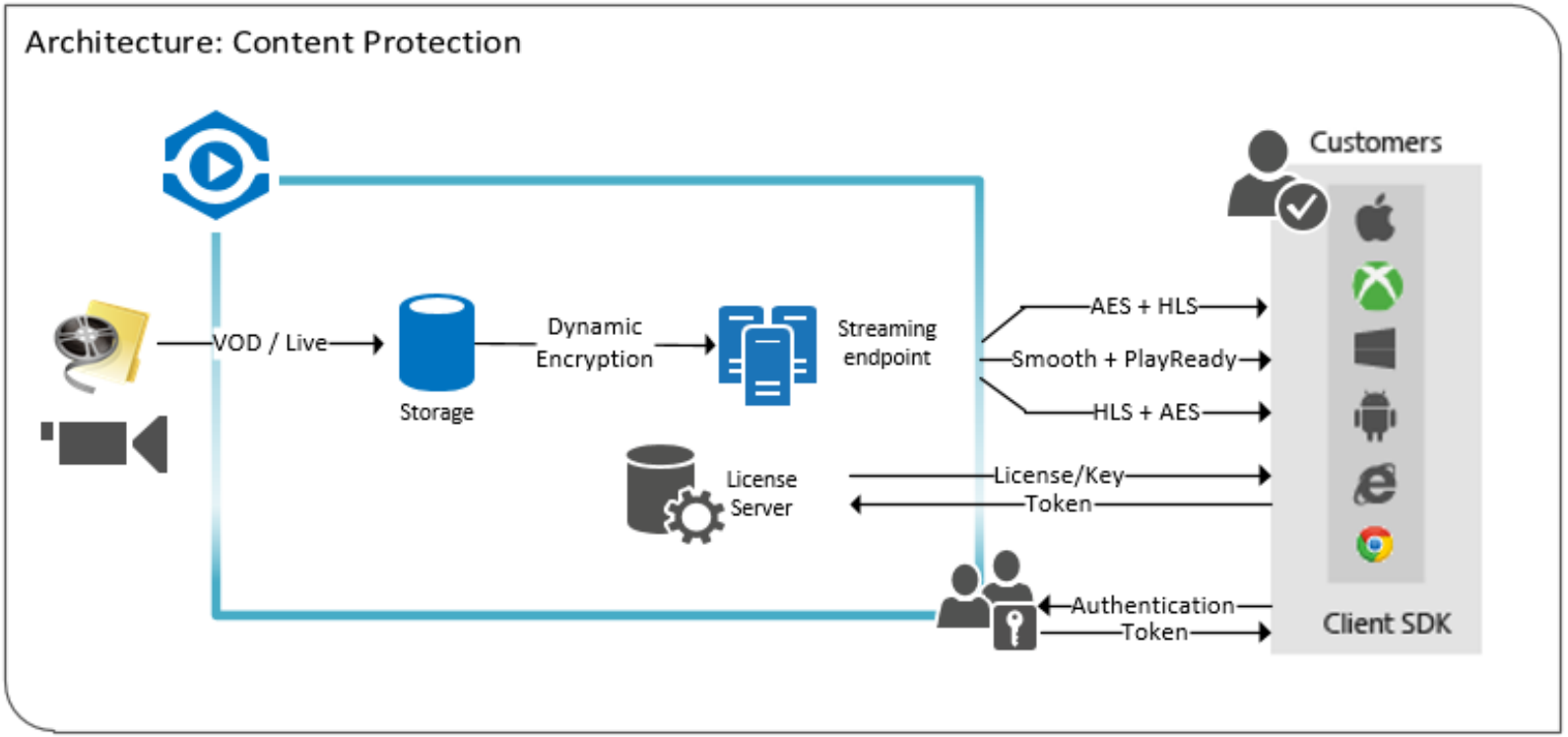
Tous / certains

- La fracture numérique
- La littératie



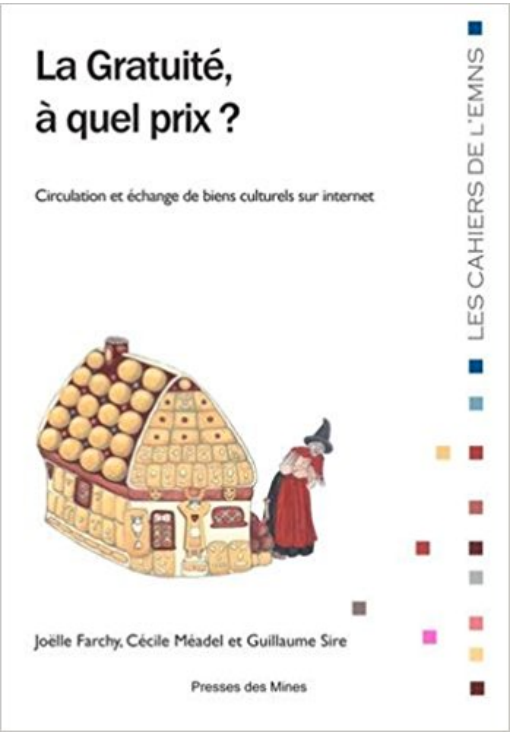
Libre circulation / Sécurisation des données

- Formats ouverts et fermés
- Observabilité / Opacité
- Interopérabilité / Contrôle

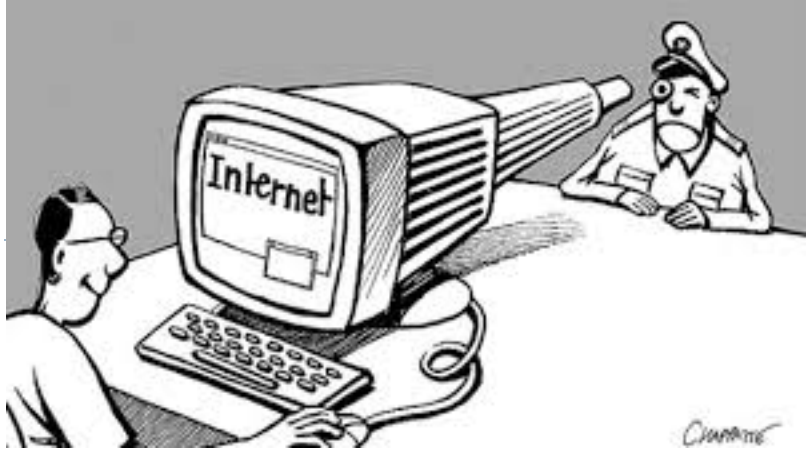


Gratuit / Payant

- Marchés mutliversants
- Répartition de la valeur entre producteurs de contenus, de contenants et OTT



Vie privée / vie public



- Les frontières renégociées
- Que devient la vie privée ?

Normes stables/ normes évolutives

- W3C / Whatwg

Plan de cours

- Introduction : la dialectique des réseaux

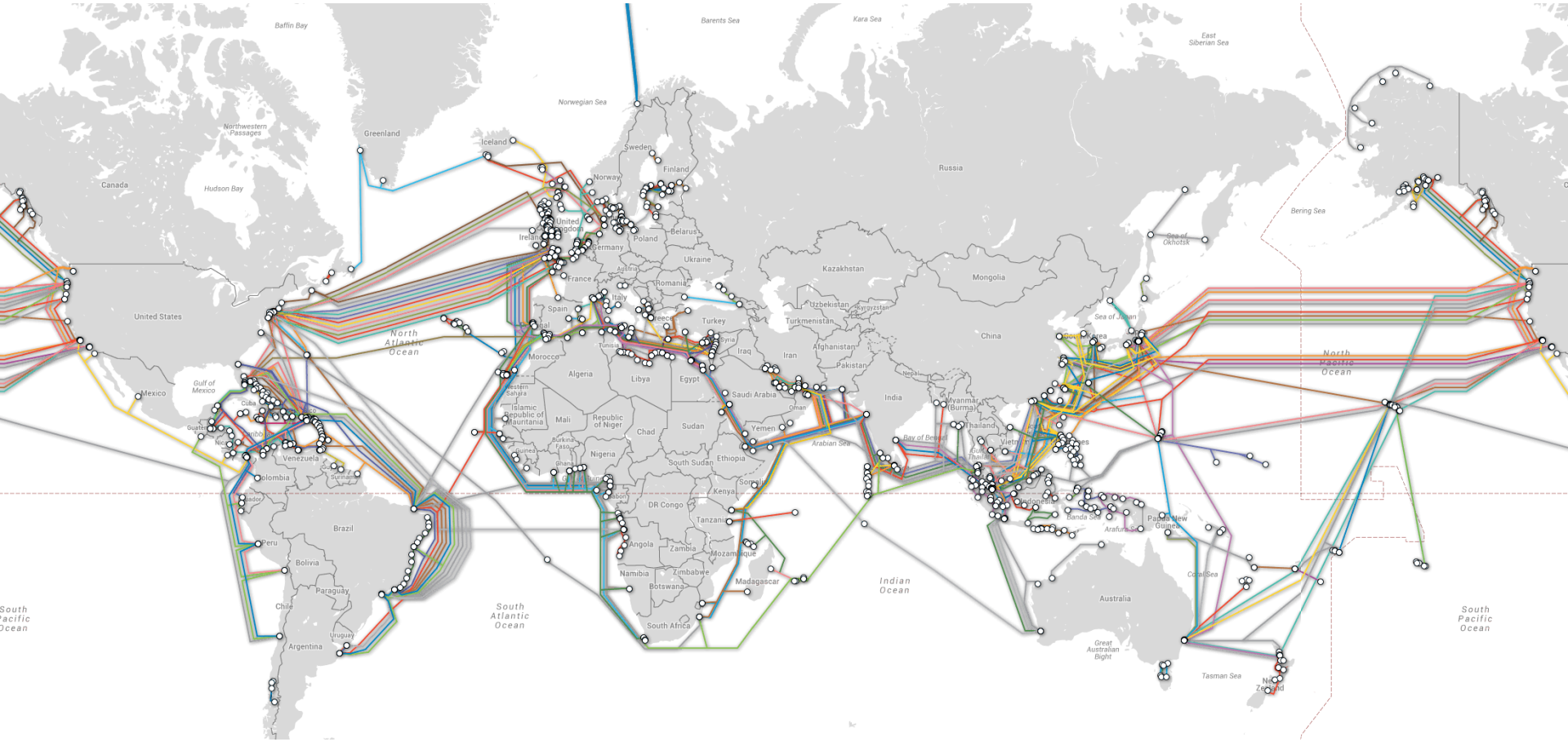
- 1. La configuration Internet
  - 1.1 Entre les réseaux
  - 1.2 Internet Protocol (IP)
  - 1.3 User Datagram Protocol (UDP)
  - 1.4 Transmission Control Protocol (TCP)
  - 1.5 Le modèle Open Systems Interconnection (OSI)
  - 1.6 L'infrastructure physique

- 2. Le dispositif web
  - 2.1 Quelques dates clés
  - 2.2 L'HyperText Transfer Protocol (HTTP)
  - 2.3 L'Uniform Resource Locator (URL)
  - 2.4 L'HyperText Markup Language (HTML)
  - 2.5 Les autres formats
  - 2.6 Les principaux acteurs
  - 2.7 Les profondeurs du web
  - 2.8 Le web sémantique

Plan de cours

- 3. Gouvernance et gouvernementalité
  - 3.1 Les principaux acteurs de la régulation technique
  - 3.2 La fabrique des normes au W3C
  - 3.3 Les stratégies de contournement
  - 3.4 La controverse encrypted media extensions (EME)
  - 3.5 Les « faire-faire » du web
- 4. Controverses et perspectives
  - 4.1 Le pouvoir des USA
  - 4.2 Le pouvoir des navigateurs et des logiciels
  - 4.3 La neutralité du réseau
  - 4.4 Extension du domaine d'Internet
  - 4.5 L'avenir des protocoles du web par rapport aux autres
  - 4.6 L'apprentissage du code

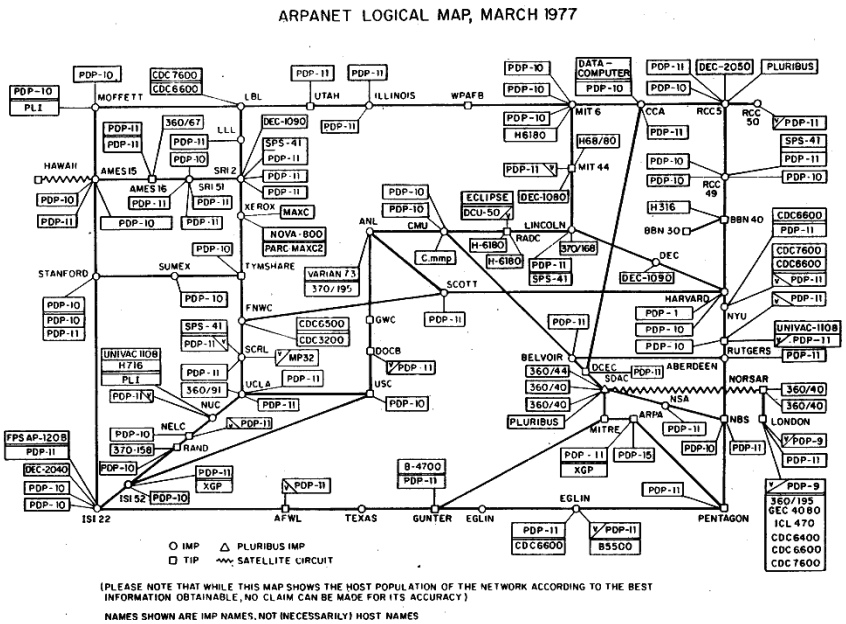
1. La configuration Internet



1.1 Entre les réseaux

- Arpanet (développé par la Defense Advanced Research Projects Agency (DARPA,)), opérationnel à partir de 1969
- Objectif initial : lier les universités, et rendre interopérable les connexion

« Pour chacun de ces trois terminaux, j'avais trois jeux différents de commandes. Si bien que si j'étais en train de parler en direct avec quelqu'un chez SDC et que je voulais discuter de ça avec quelqu'un que je connaissais à Berkeley ou au MIT, il fallait que je me lève de devant le terminal S.D.C., que j'aille m'enregistrer sur l'autre terminal afin d'entrer en contact avec eux.
Je me suis dit, hé, mec, ce qu'il me reste à faire est évident : au lieu d'avoir ces trois terminaux, il nous faut un terminal qui va partout où tu veux et où il existe un ordinateur interactif. Cette idée était l'ARPAnet. » — Robert Taylor,



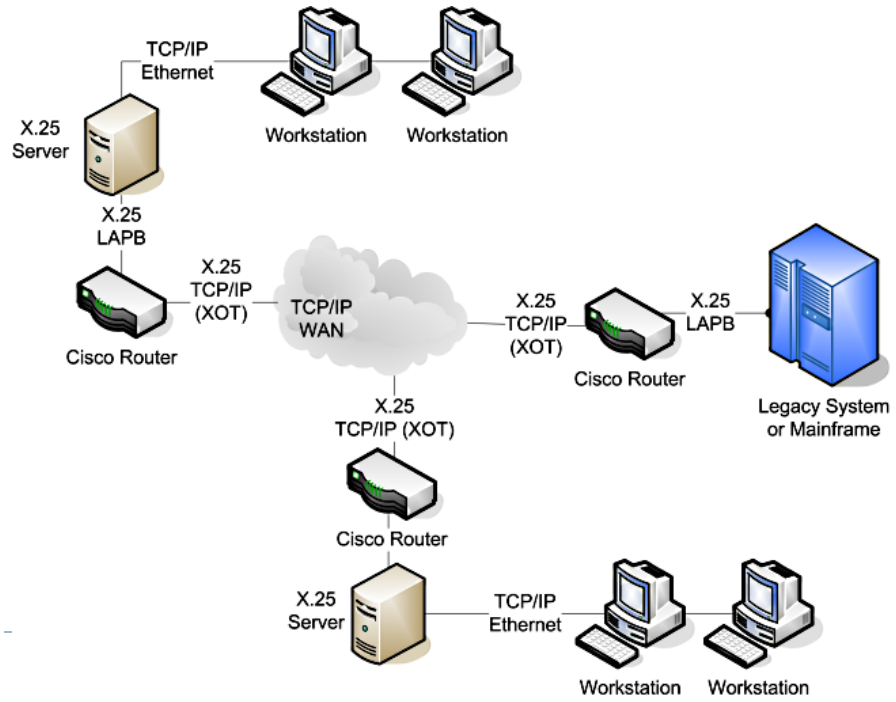
1.1 Entre les réseaux

- En 1980, Arpanet se divise en deux réseaux distincts, l'un militaire (MILNET, de *Military Network*, qui deviendra le DDN — Defense Data Network)
- et l'autre, universitaire (NSFnet)

Le 1er janvier 1983, ARPANET adopte le TCP/IP
En 1984, DDN et NSFnet comptent déjà près de 4 millions de nœuds interconnectés, et plus de 1 000 ordinateurs à travers le monde y sont relié

1.1 Entre les réseaux

- X.25, Vidéotex, etc. : des protocoles utilisés à des fins commerciales
- L'unification implique un choc des cultures

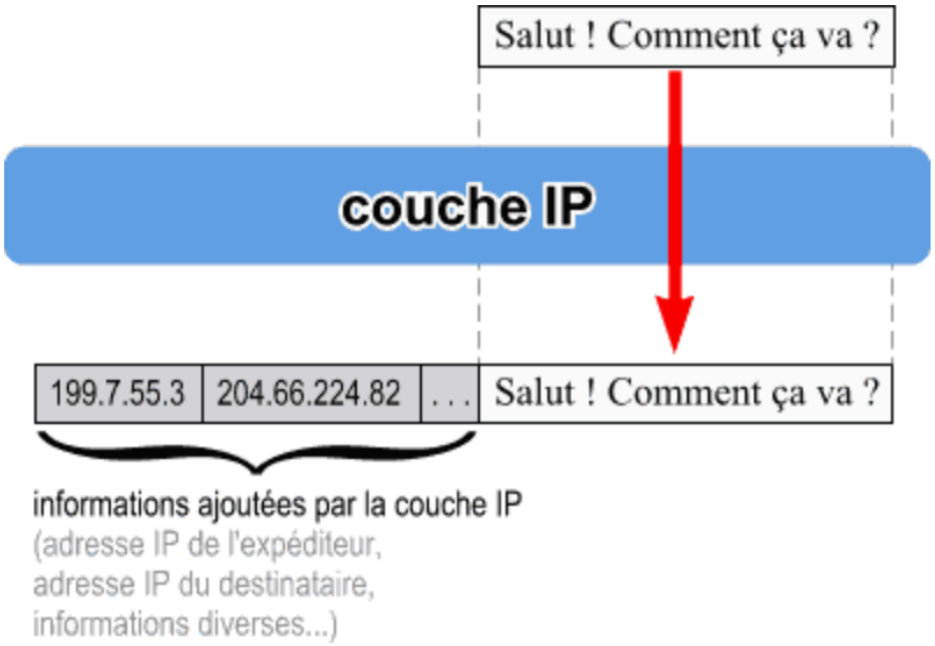


1.1 Entre les réseaux

- La suite de protocole TCP/IP va permettre de rendre interopérables les infrastructures et les protocoles
  - Inventé par Vinton G. Cerf et Bob Kahn
  - Géré par Internet Engineering Task Force (IETF)
  - « Requirements for Internet Hosts -- Communication Layers [archive] », Request for comments no 1122, octobre 1989.

1.2 Internet Protocol (IP)

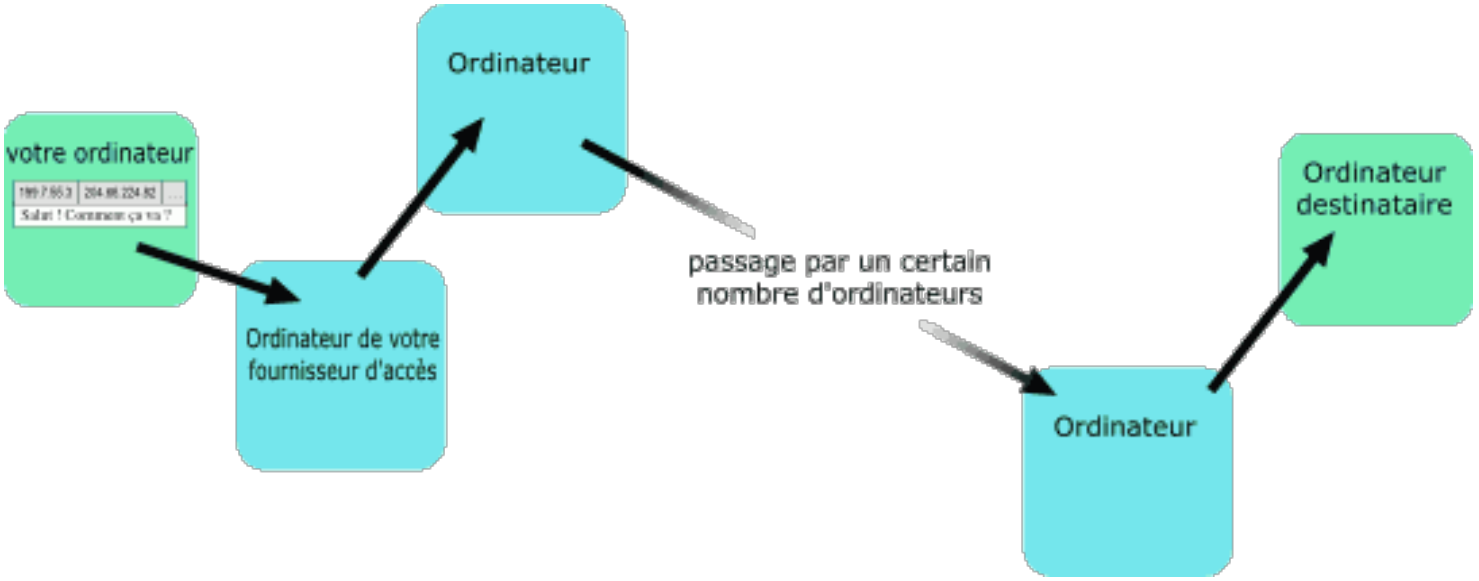
- L'adresse IP est une adresse unique attribuée à chaque ordinateur
- Elle se présente sous la forme de 4 nombres (entre 0 et 255) séparés par des points. (Par exemple: 204.35.129.3)
- Le protocole va permettre que soient ajoutées et comprises les informations relatives à l'IP à un message



Source :
http://sebsauvage.net/comprendre/tcpip/

1.2 Internet Protocol (IP)

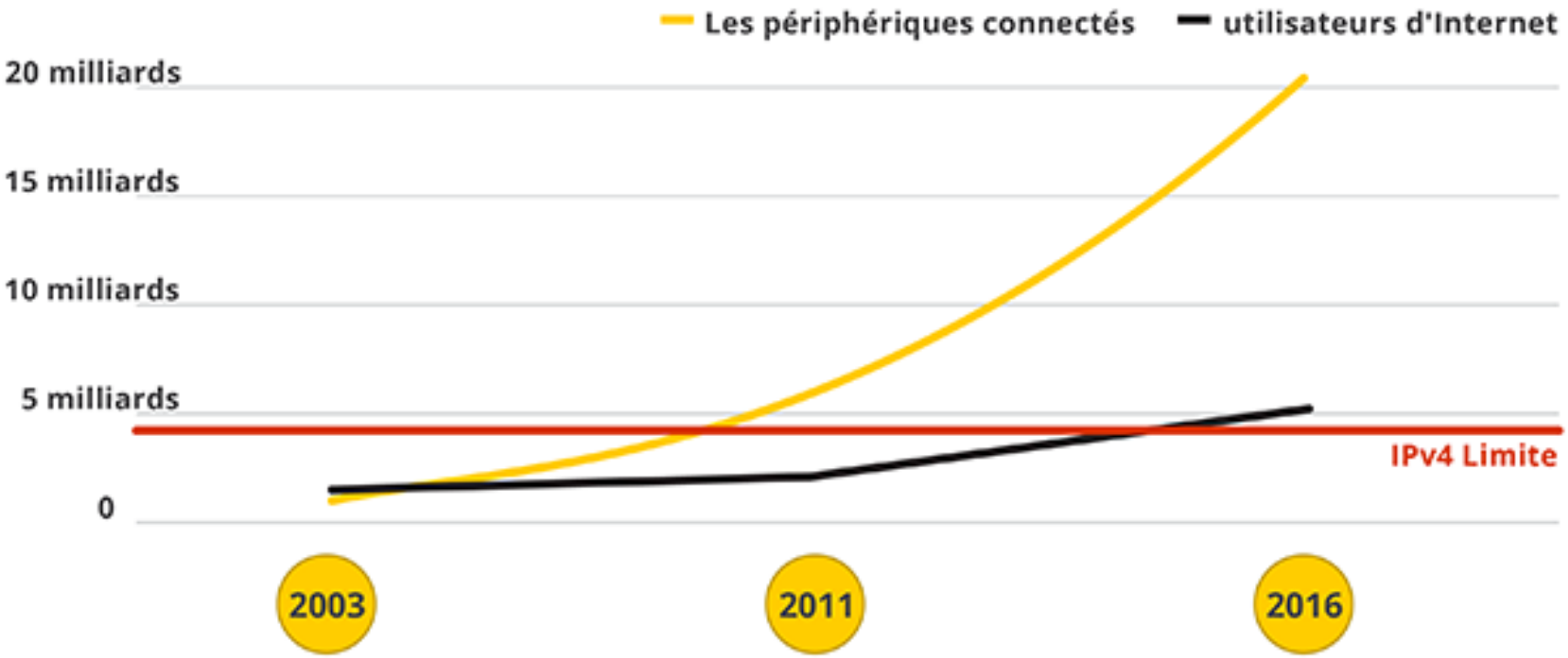
- Le routage



Source : http://sebsauvage.net/comprendre/tcpip/

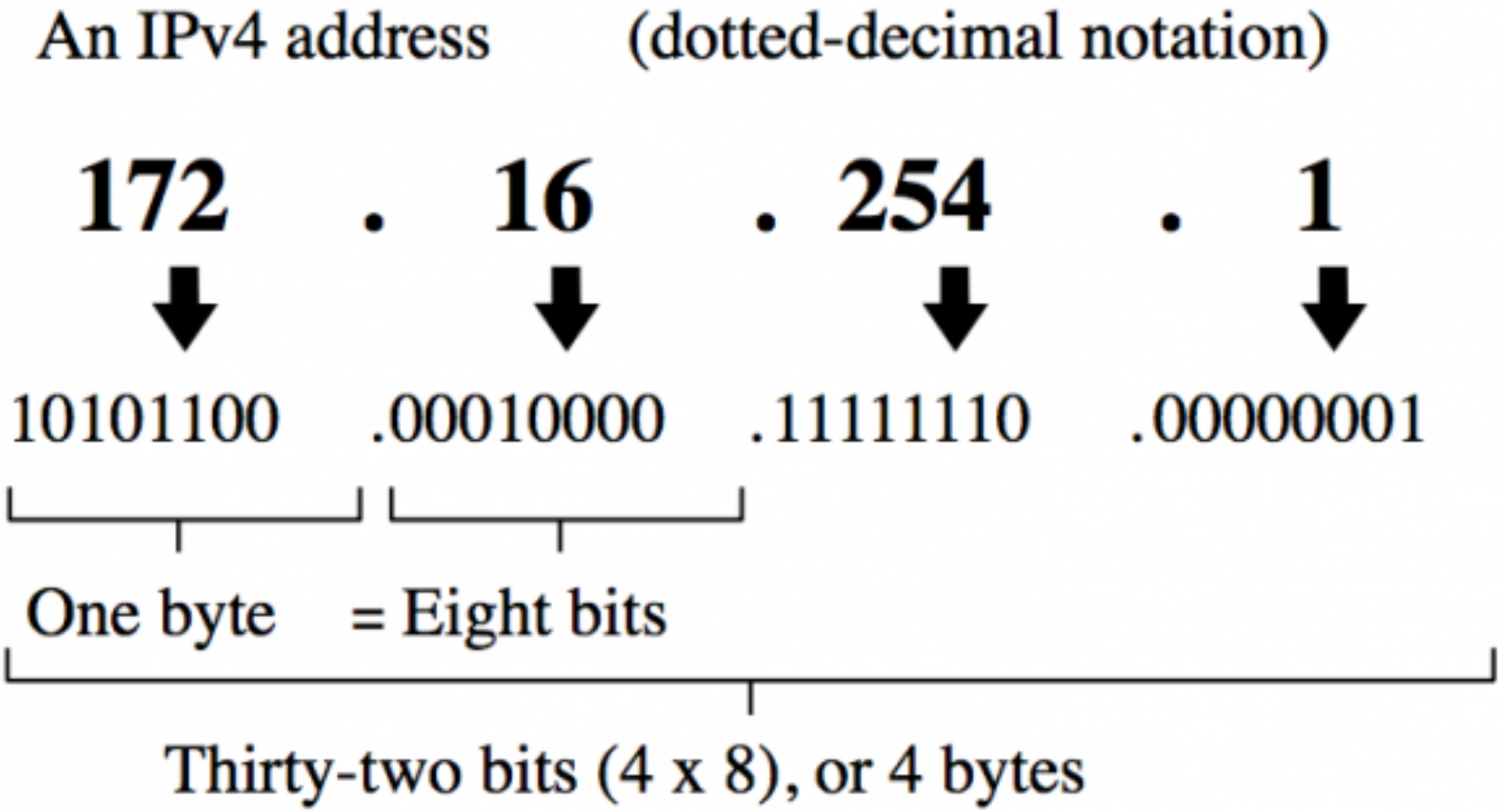
1.2 Internet Protocol (IP)

- Pénurie d'adresses IPv4

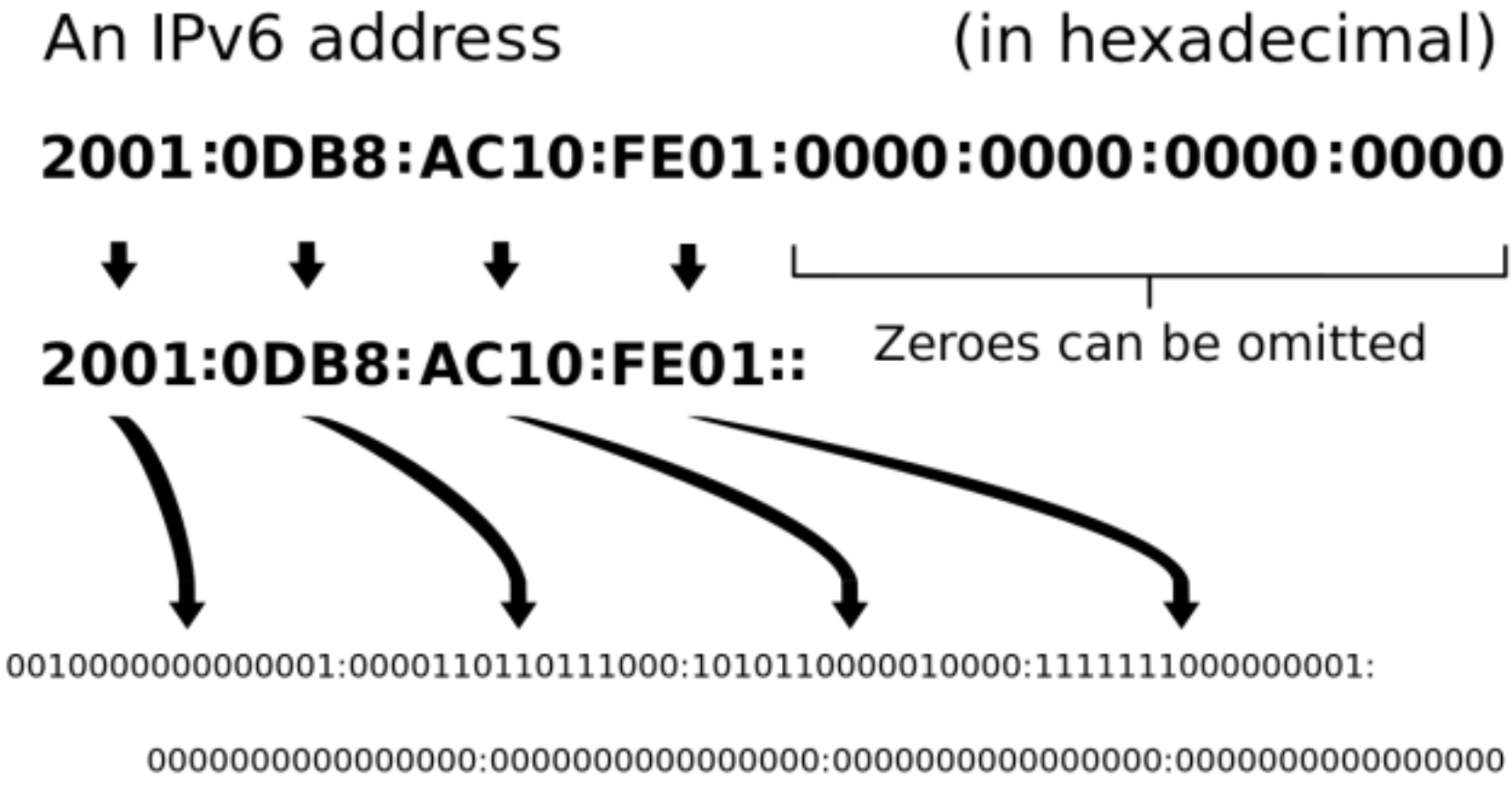


Source : ZDnet

1.2 Internet Protocol (IP)



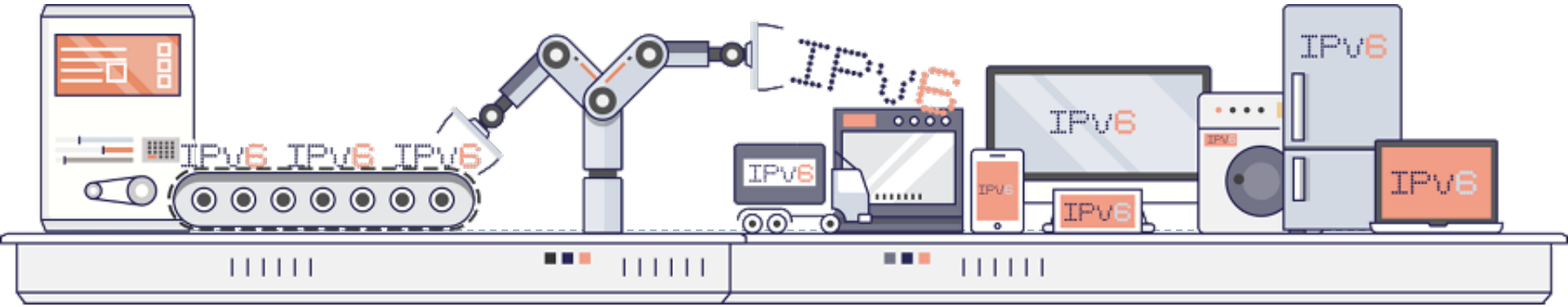
1.2 Internet Protocol (IP)



1.2 Internet Protocol (IP)

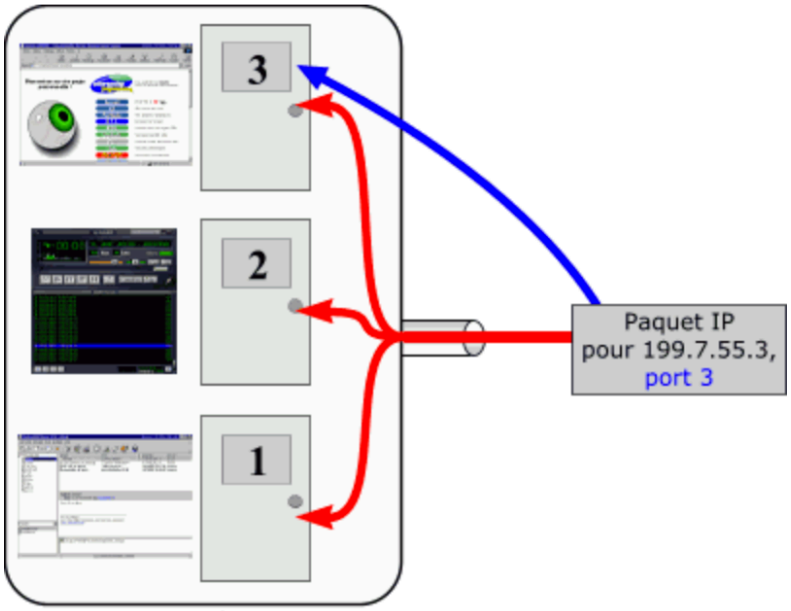
- IPv4 vs IPv6

	Internet Protocol version 4 (IPv4)	Internet Protocol version 6 (IPv6)
Deployed	1981	1999
Address Size	32-bit number	128-bit number
Address Format	Dotted Decimal Notation: 192.149.252.76	Hexadecimal Notation: 3FFE:F200:0234:AB00: 0123:4567:8901:ABCD
Prefix Notation	192.149.0.0/24	3FFE:F200:0234::/48
Number of Addresses	2^{32} = ~4,294,967,296	2^{128} = ~340,282,366, 920,938,463,463,374, 607,431,768,211,456



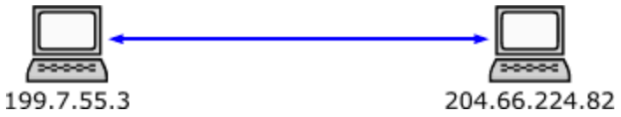
1.3 User Datagram Protocol (UDP)

- Comment savoir à quel logiciel, sur la machine destinataire, adresser le parquet ?
  - On ajoute un « port » à l'IP destinataire, l'équivalent d'un numéro d'appartement à une adresse



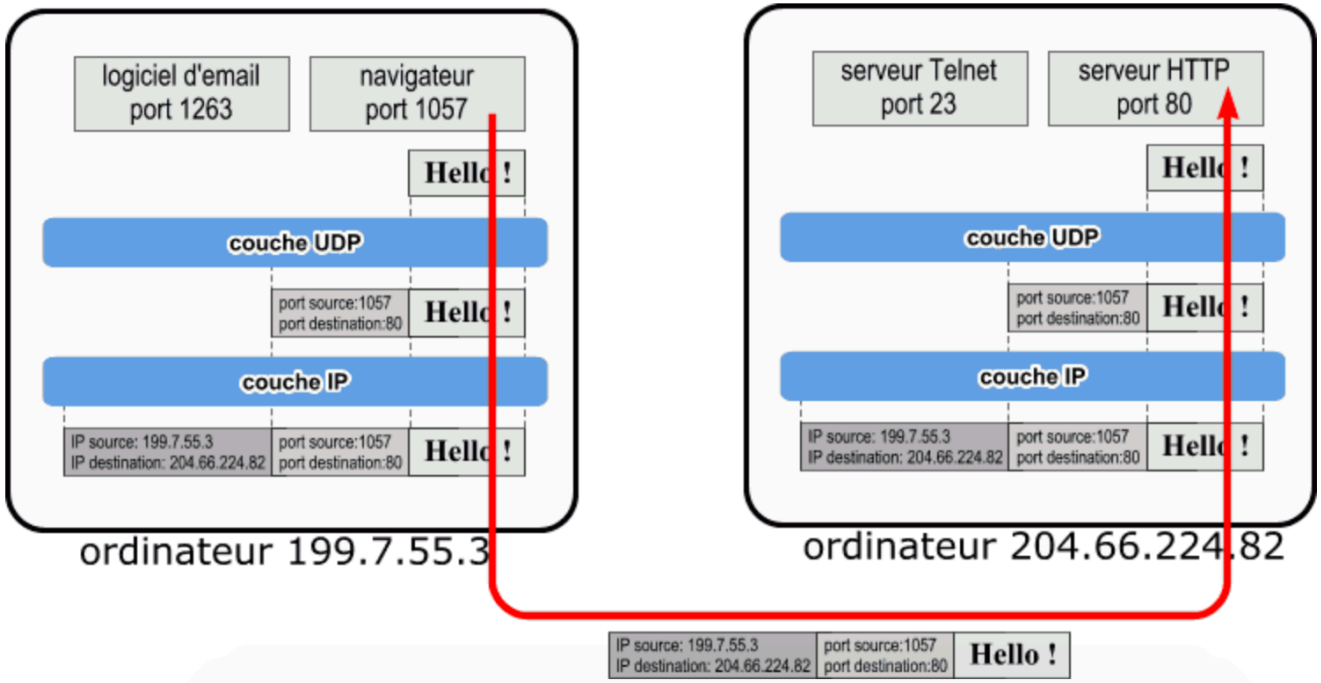
1.3 User Datagram Protocol (UDP)

Avec le protocole **IP** on pouvait envoyer des données d'un ordinateur A à un ordinateur B.



Avec **UDP/IP**, on peut être plus précis: on envoie des données d'une **application x** sur l'**ordinateur A** vers une **application y** sur l'**ordinateur B**.

Par exemple, votre navigateur peut envoyer un message à un serveur HTTP (un serveur Web):



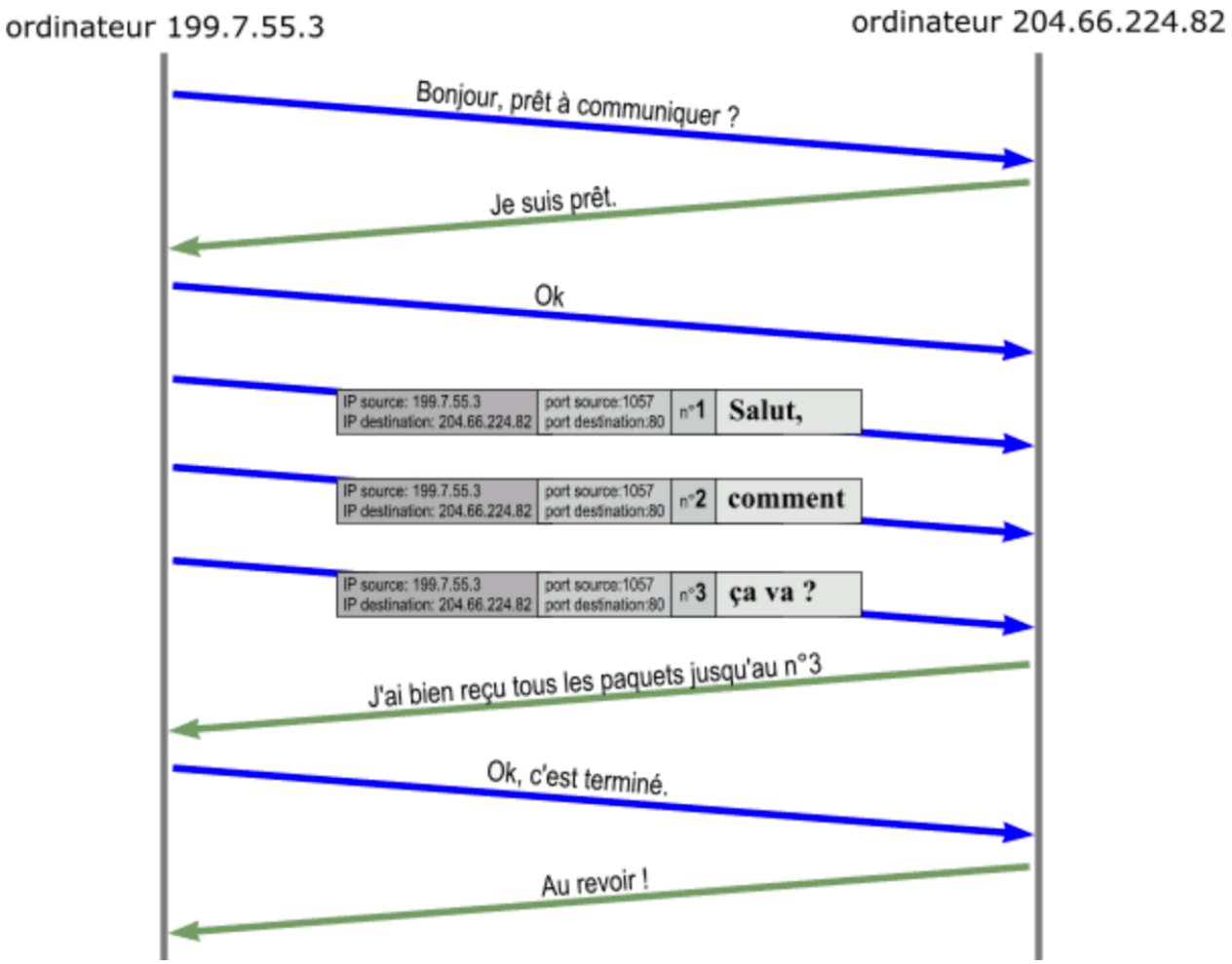
1.3 User Datagram Protocol (UDP)

- Problèmes du UDP/IP :
  - Quand on envoie un paquet IP sur Internet, il passe par des dizaines d'ordinateurs. Et il arrive que des paquets se **perdent** ou arrivent en **double exemplaire** (imaginez votre ordre de virement arrivant deux fois…)
  - Même si le paquet arrive à destination, rien ne vous permet de savoir si le paquet est bien arrivé (aucun accusé de réception)
  - **La taille des paquets IP est limitée** (environ 1500 octets). Comment faire pour envoyer une photo JPEG qui fait 62000 octets ?

1.4 Transmission Control Protocol (TCP)

- Le TCP est capable :
  - de faire tout ce que UDP sait faire (ports).
  - de vérifier que le destinataire est prêt à recevoir les données.
  - de **découper** les gros paquets de données en paquets plus petits pour que IP les accepte
  - de **numéroter** les paquets, et à la réception de **vérifier** qu'ils sont tous bien arrivés, de **redemander** les paquets manquants et de les **réassembler** avant de les donner aux logiciels. Des accusés de réception sont envoyés pour prévenir l'expéditeur que les données sont bien arrivées.

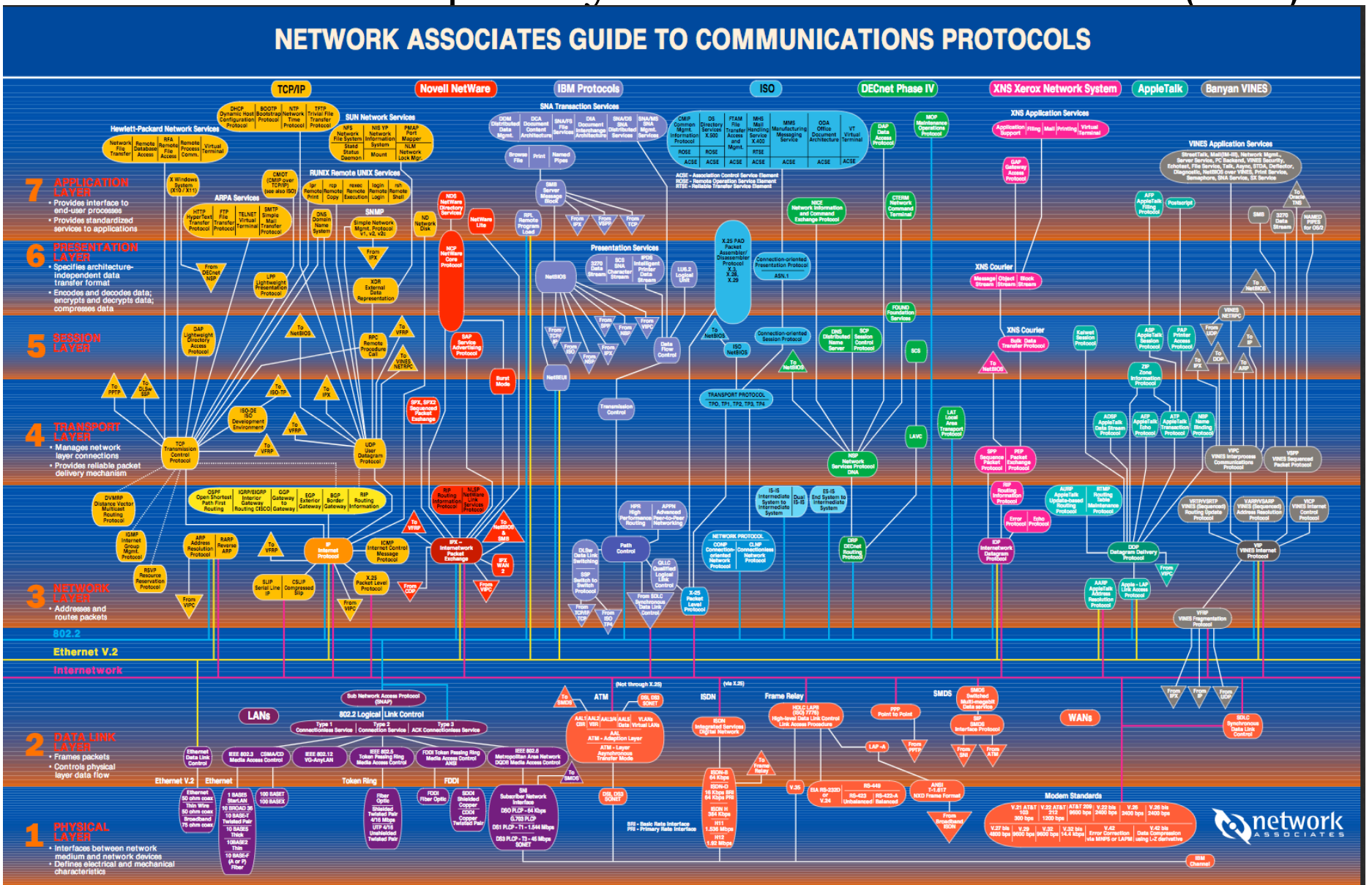
1.4 Transmission Control Protocol (TCP)



1.5 Le modèle Open Systems Interconnection (OSI)

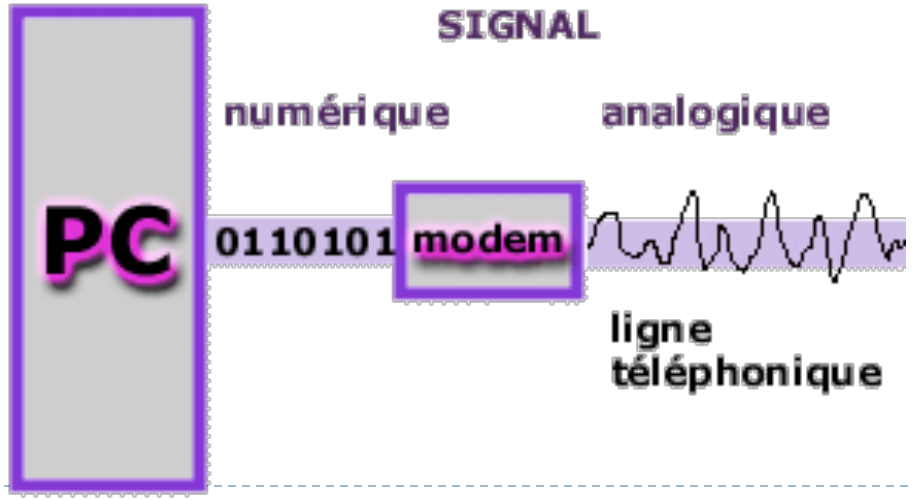
- Le modèle ***O**pen **S**ystems **I**nterconnection* (OSI) a vocation à décrire la totalité des protocoles existants et leurs rapports

1.5 Le modèle Open Systems Interconnection (OSI)



1.6 L'infrastructure physique

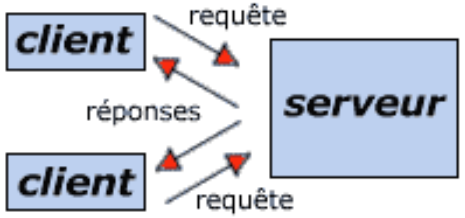
Modem (Modulateur/DEModulateur) : module les informations numériques en ondes analogiques; en sens inverse il retranscrit les données sous forme analogique en données numérique.



1.6 L'infrastructure physique

Serveur : fournit des données à celui qui lui en demande. Ce demandeur peut être un autre serveur ou l'ordinateur d'un utilisateur final. Les données servies peuvent être de toute nature : sons, images, texte

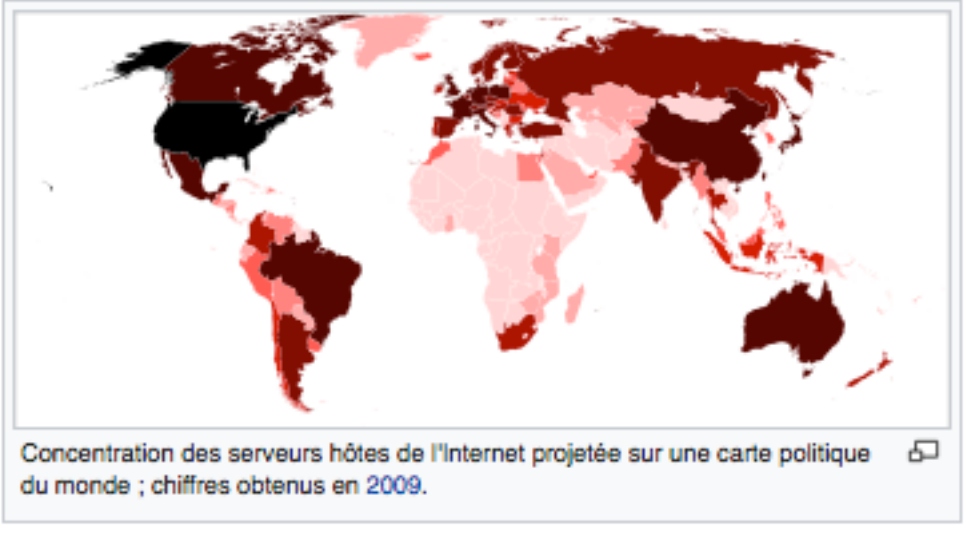
(Le logiciel qui fait fonctionner le serveur s'appelle « serveur HTTP »)



- Le client émet une requête vers le serveur grâce à son adresse IP et le port, qui désigne un service particulier du serveur
- Le serveur reçoit la demande et répond à l'aide de l'adresse de la machine cliente et son port

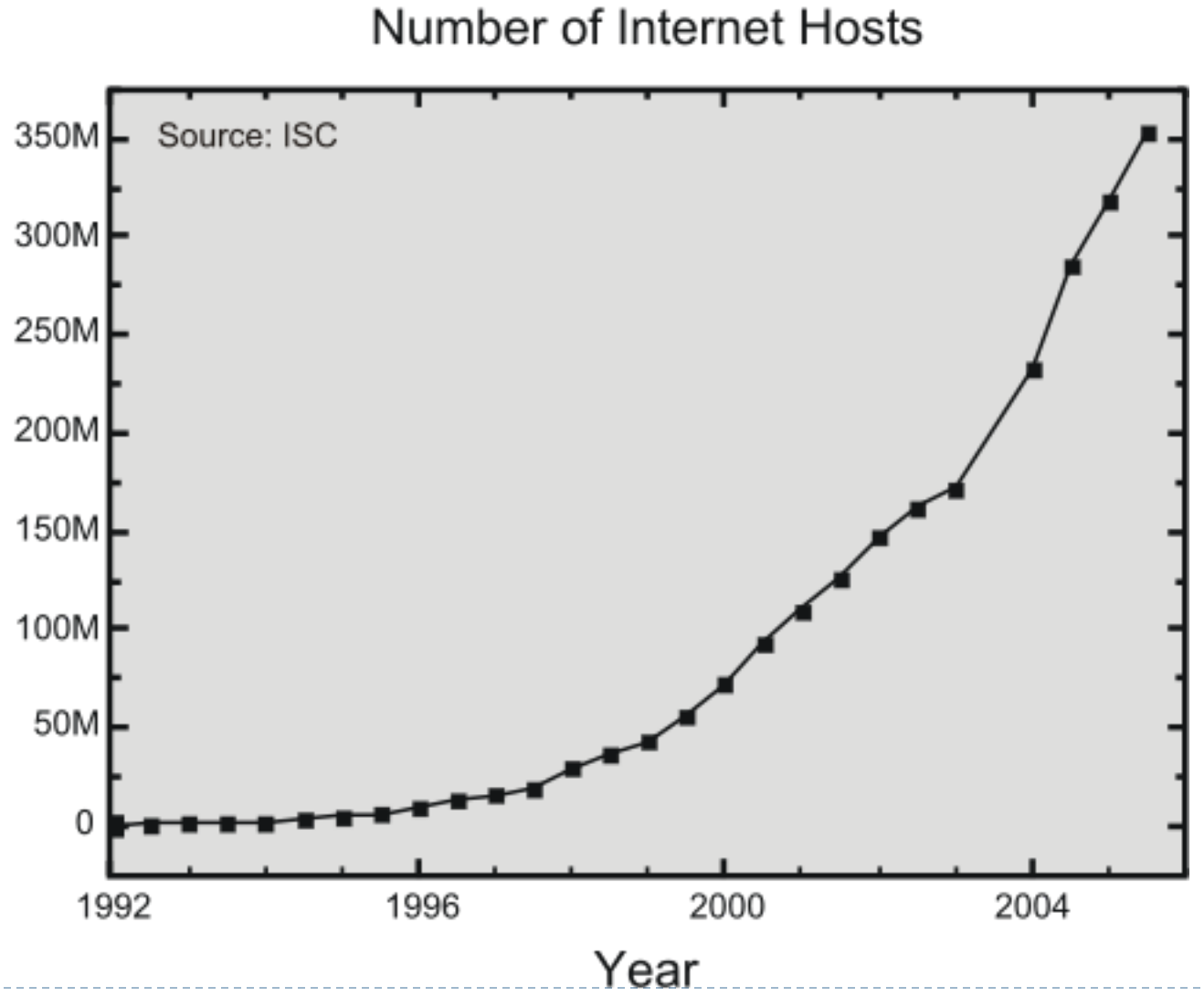
1.6 L'infrastructure physique

Hébergeur : désigne l'entreprise qui détient les fermes de serveurs sur lesquels sont stockées les informations relatives aux sites web

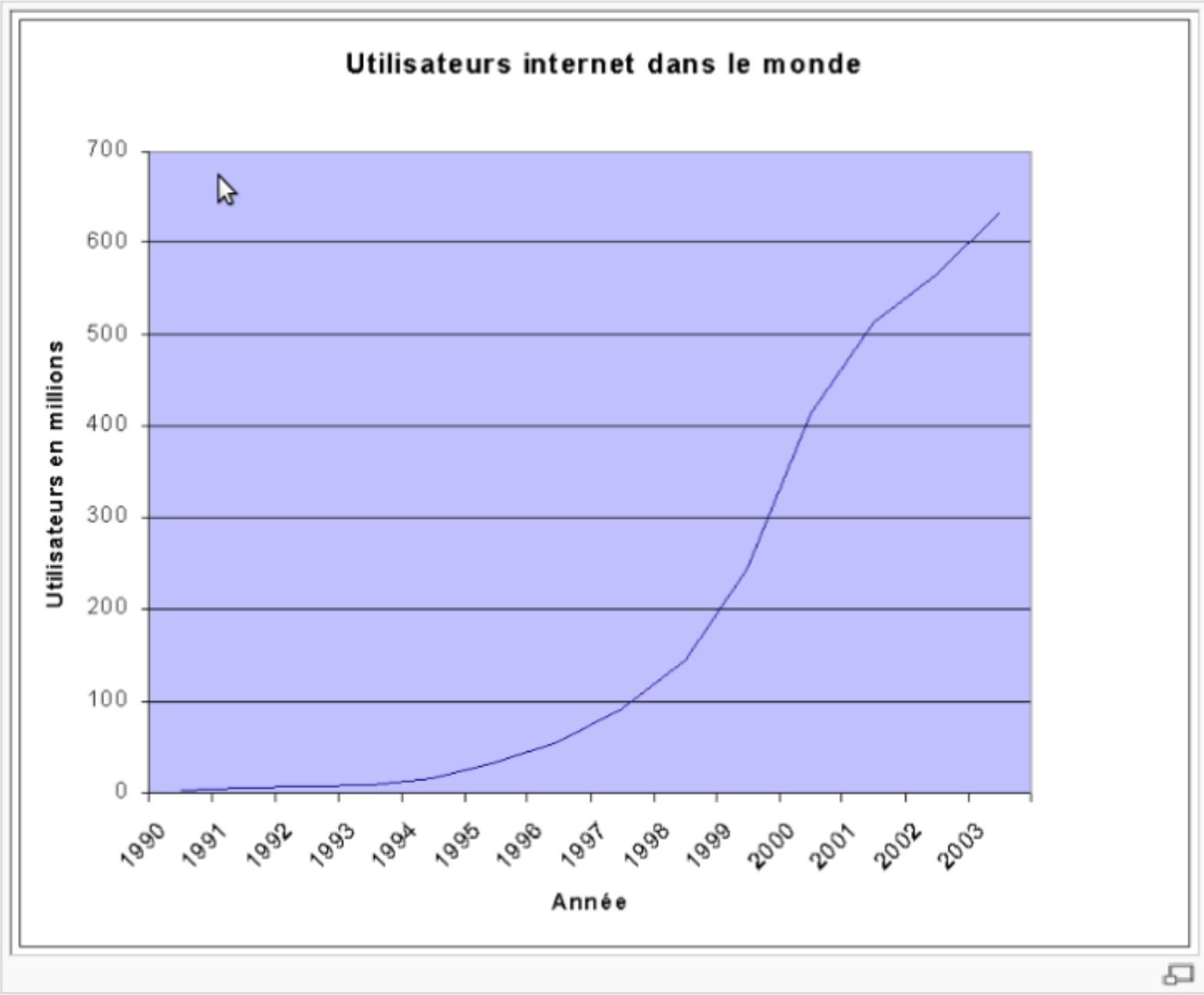


Concentration des serveurs hôtes de l'Internet projetée sur une carte politique du monde ; chiffres obtenus en 2009.

1.6 L'infrastructure physique



Selon Internet World Stats , il a été recensé 1,59 milliard d'internautes en mars 2009.



1.6 L'infrastructure physique

4 familles d'acteurs composent l'infrastructure (tous ont des serveurs en réseau, et ils sont interdépendants) :

- FAI
- Hébergeurs de services et de contenus
- CDN : Content Delivery Network, assurent la copie et la mise à disposition de contenu en réseau.
- Opérateurs de transit IP : permettent de faire le lien entre les différents opérateurs. Ce sont eux qui fournissent le service de transport. Pour cela, ils passent des accords de connexion avec les différents fournisseurs. Les grandes entreprises peuvent éventuellement s'adresser à des opérateurs de transit IP pour se connecter à Internet sans passer par un fournisseur d'accès à Internet.

1.6 L'infrastructure physique

Modifier

Peering et transit

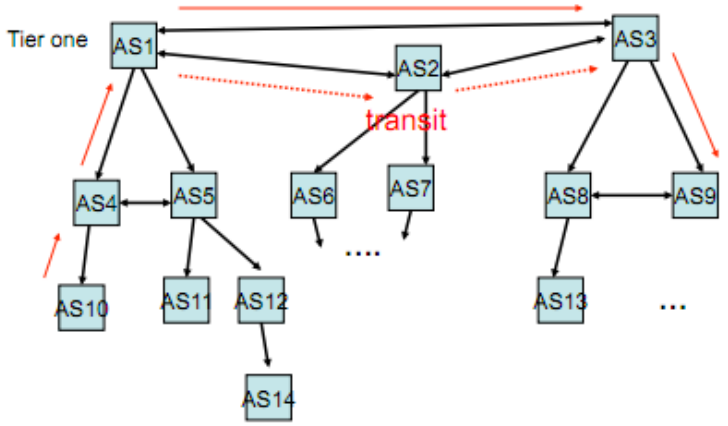
Afin qu'une machine de l'*opérateur A* puisse communiquer avec une machine de l'*opérateur B*, il faut donc qu'il existe une route de *A* vers *B*. Deux cas de figure se présentent :

Peering

Considérons (voir schéma) qu'*AS1* et *AS4* sont présents dans une même salle machine (cette salle est nommée **Point de Présence** ou **POP**) ; ils se connectent alors directement en reliant leurs routeurs de bordure par un câble. Ils passent un *contrat de peering* qui stipule généralement qu'ils s'échangent leur trafic gratuitement.

Transit

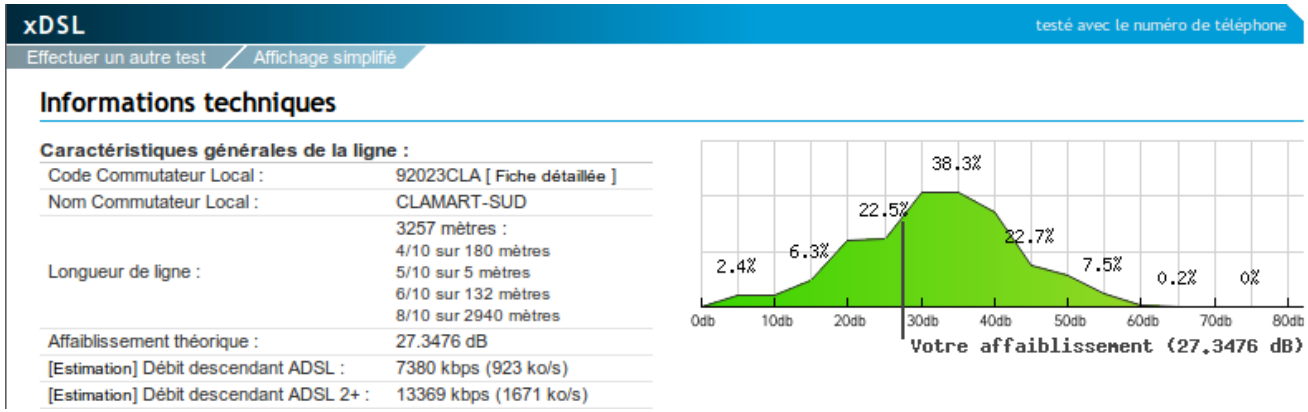
Prenons maintenant sur le schéma *AS1* qui est en Chine et *AS3* au Brésil, il est peu probable qu'ils aient des bordures localisées dans un même *POP*. Chacun doit alors payer les services d'un **opérateur tiers** (ici *AS2*) qui se chargera de transporter leur trafic d'une bordure de *AS1* à une bordure de *AS3*. *AS3* facture alors au Mbits transporté le trafic écoulé à travers son réseau.

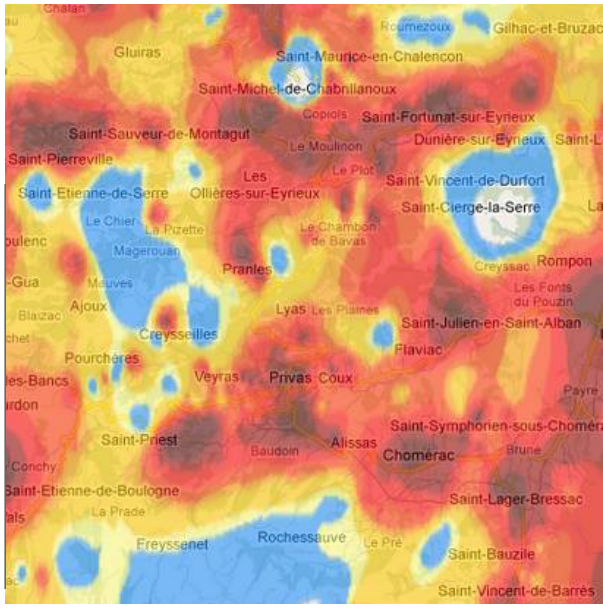


Source : https://doc.rhizome-fai.net/doku.php?id=theorie_jargon:internet:fonctionnement_d_internet

1.6 L'infrastructure physique

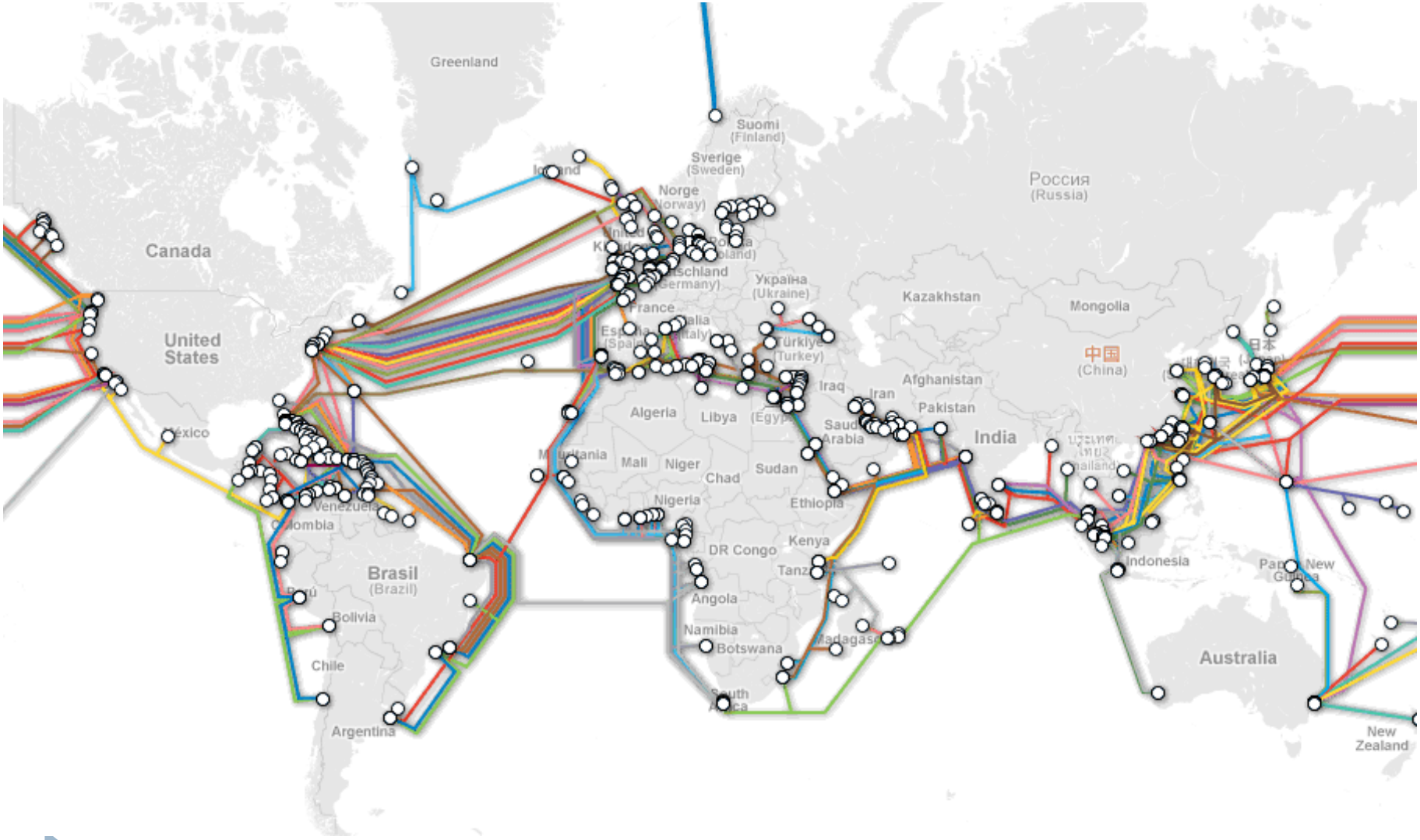
(L'affaiblissement est fonction de la distance entre chez vous et le répartiteur téléphonique)

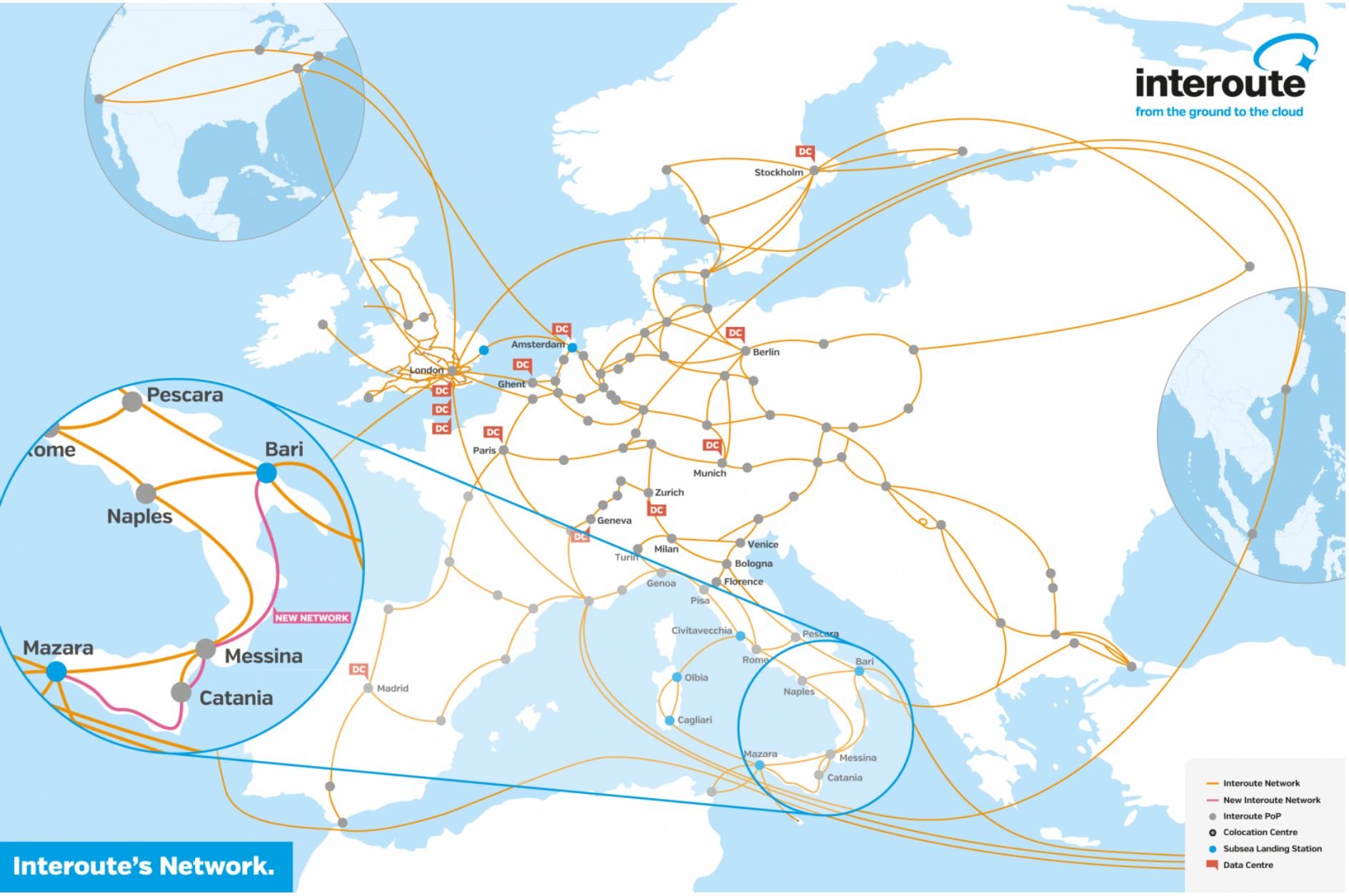




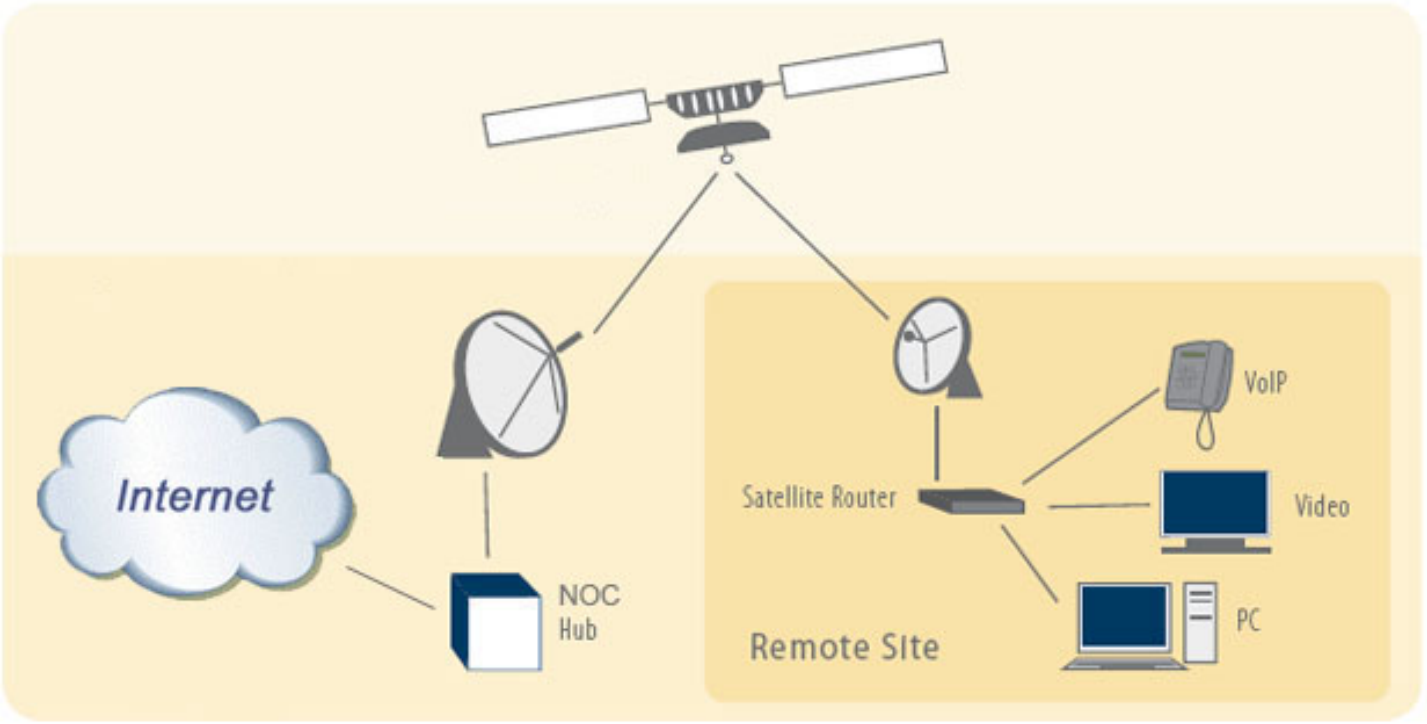
Source : http://www.ariase.com/fr/news/lignes-telephoniques-adsl-affaiblissement-record-article-2470.html

Les câbles


Greenland
Canada
United States
México
Suomi (Finland)
Sverige (Sweden)
Norge (Norway)
Россия (Russia)
Україна (Ukraine)
France
Türkiye (Turkey)
Kazakhstan
Mongolia
中国 (China)
Iraq
Iran
Afghanistan
Pakistan
Algeria
Libya
Saudi Arabia
India
Mali
Niger
Chad
Sudan
Nigeria
Ethiopia
Kenya
DR Congo
Tanzania
Angola
Namibia
Botswana
Indonesia
Papua New Guinea
Australia
New Zealand
Venezuela
Colombia
Brasil (Brazil)
Bolivia
Chile
Argentina


interoute
from the ground to the cloud
Stockholm
Amsterdam
London
Ghent
Berlin
Paris
Munich
Zurich
Geneva
Milan
Turin
Venice
Bologna
Florence
Genoa
Pisa
Civitavecchia
Pescara
Rome
Bari
Naples
Olbia
Cagliari
Mazara
Messina
Catania
Madrid
DC
Pescara
Bari
Naples
NEW NETWORK
Mazara
Messina
Catania
Interoute Network
New Interoute Network
Interoute PoP
Colocation Centre
Subsea Landing Station
Data Centre
Interoute's Network.

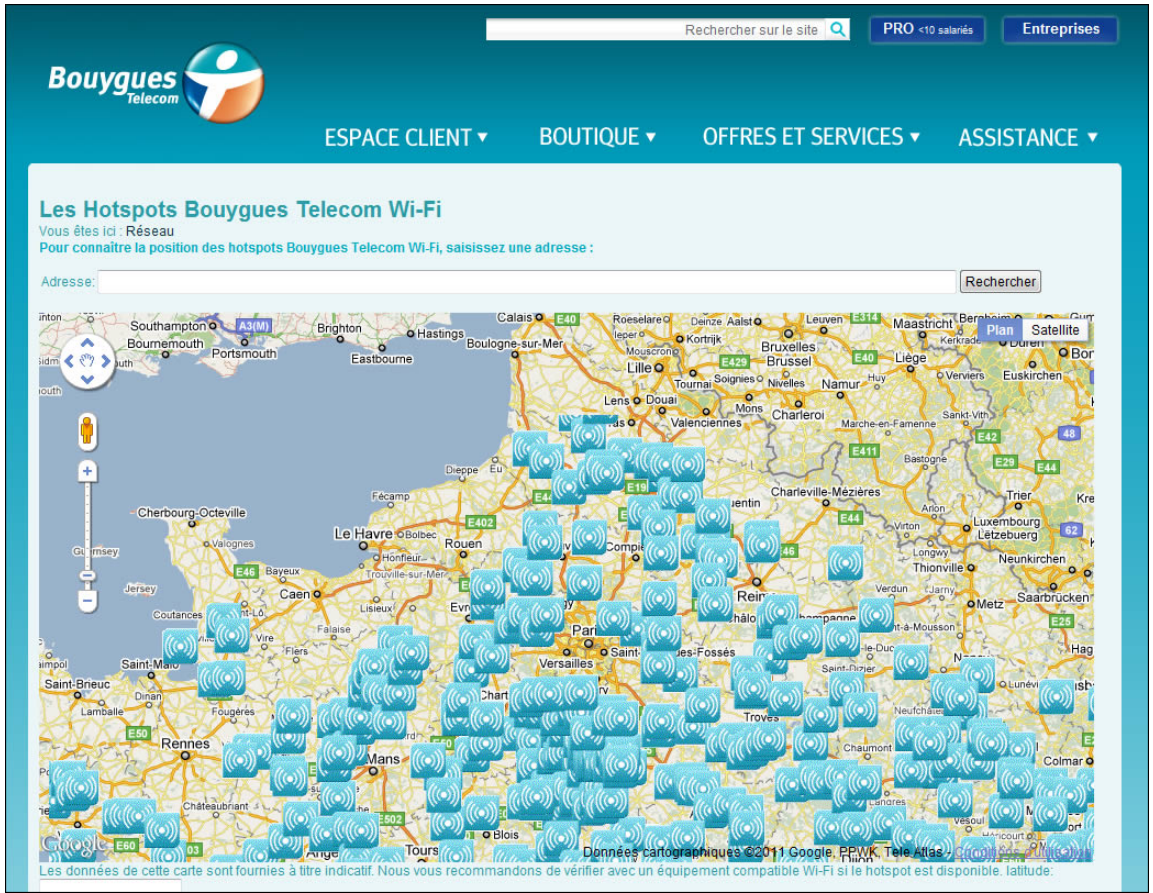
Les satellites



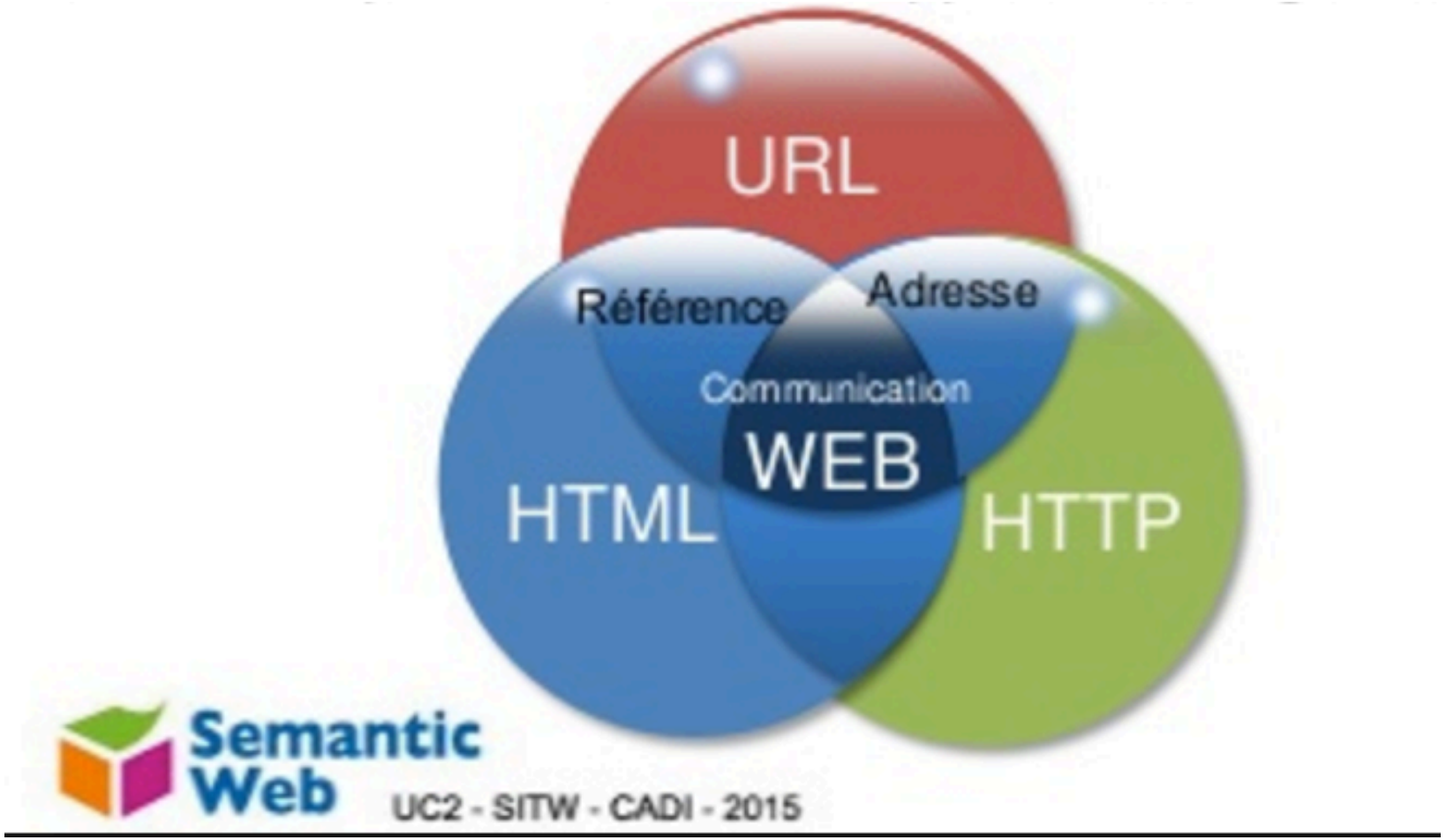
On se connecte par satellite au serveur d'un opérateur, lui-même connecté par câble
Les connexions internet par satellite utilisent la norme DVB (Digital Video Broadcasting)
Cela peut être unidirectionnel (la requête est envoyée par téléphone, et l'information reçue par satellite) ou bidirectionnel (la requête est envoyée par satellite)

Les hotspots / wifi partagé

l'abonné partage sa connexion en transformant sa box en *hotspot* public, ce qui lui permet en retour d'accéder à Internet au travers des boxs d'autres abonnés, mais aussi aux hotspots installés par son opérateur dans les lieux publics

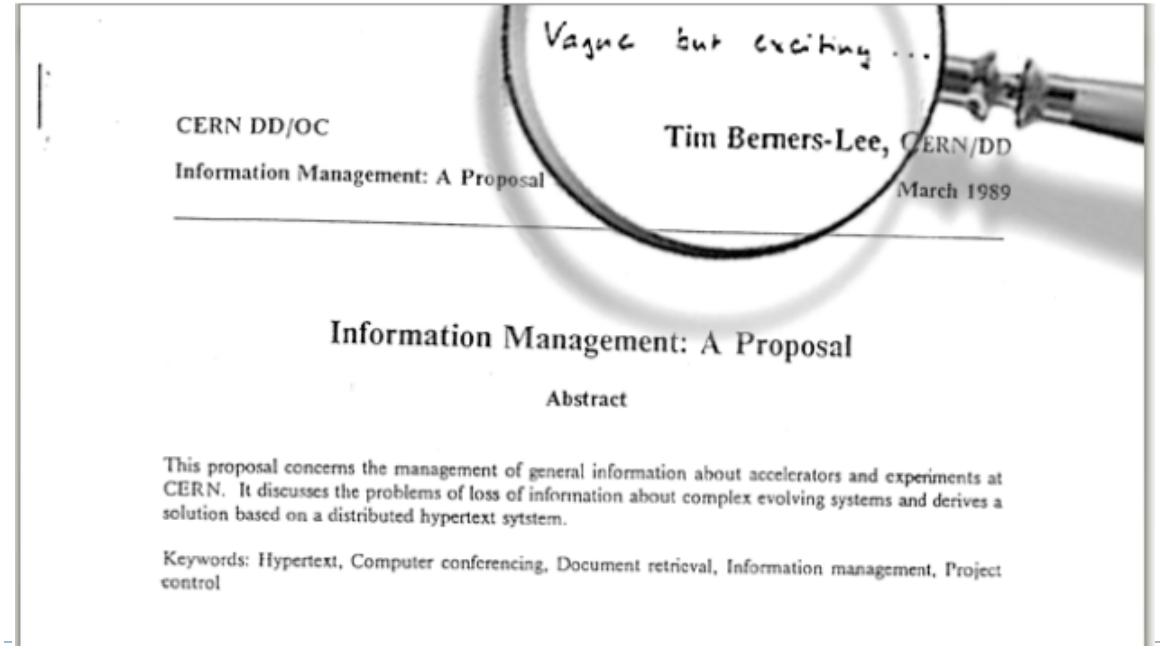


2. Le dispositif web



2.1 Quelques dates clés

- 1989 : Tim Berners Lee propose au CERN un *draft* de quelques pages, appelé « *gestion de l'information : une proposition* » dans lequel ilimagine une réorganisation des bases de données du centre de recherche nucléaire.



Vague but exciting ...

CERN DD/OC

Information Management: A Proposal

Tim Berners-Lee, CERN/DD

March 1989

Information Management: A Proposal

Abstract

This proposal concerns the management of general information about accelerators and experiments at CERN. It discusses the problems of loss of information about complex evolving systems and derives a solution based on a distributed hypertext sytstem.

Keywords: Hypertext, Computer conferencing, Document retrieval, Information management, Project control

2.1 Quelques dates clés

- 1990 : Le premier site et tout premier serveur Web voient le jour sur l'ordinateur NeXT de Tim Berners-Lee. Le site, qui n'est consultable que sur le réseau interne du CERN, est affiché sur un outil baptisé « navigateur ». Son adresse : info.cern.ch. Tim Berners-Lee travaille désormais avec le belge Robert Caillau et met en place les deux protocoles de base :
  - HyperText Transfer Protocol (HTTP, pour lier les documents)
  - le HTML (pour décrire les documents)
    - [Auquel s'ajoutera l'Uniform Resource Locator en 1994 (URL, pour localiser les documents)]

- 6 août 1991 : ouverture à tout le monde

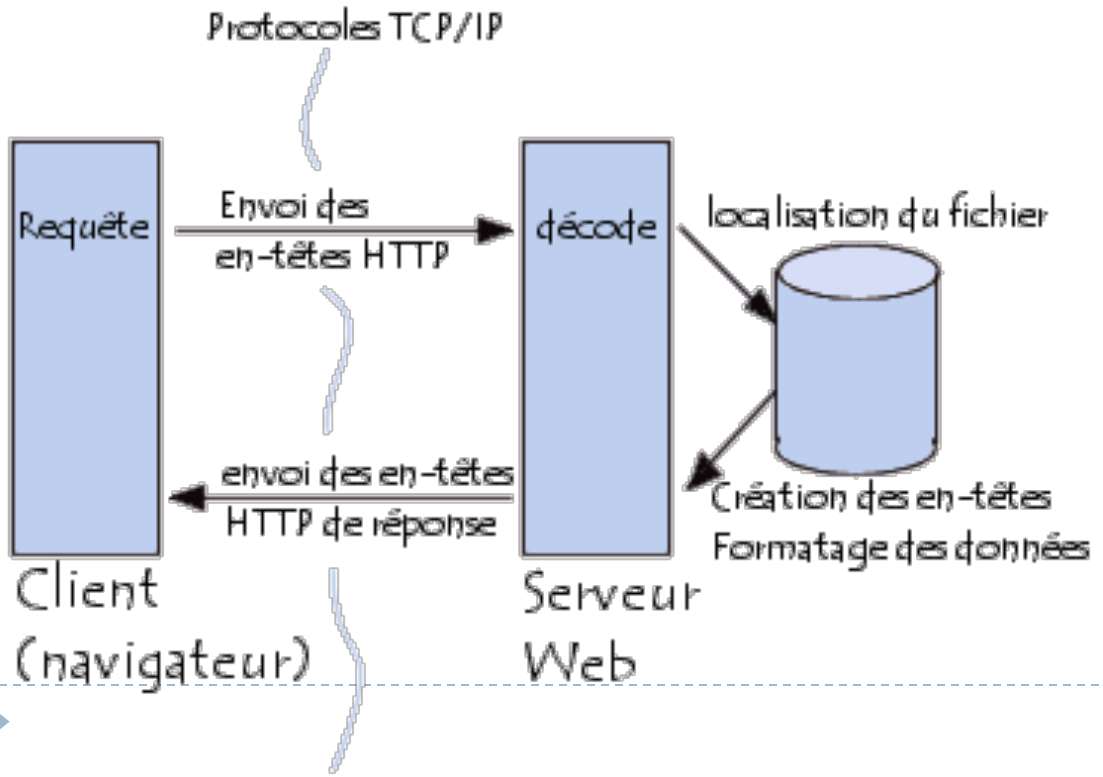
2.1 Quelques dates clés

- 1994 : Tim Berners-Lee quitte l'IETF pour créer le W3C
- 1994 : création de Yahoo

- 1995 : création d'Amazon
- 1998 : création de Google
- 2005 : les propriétaires de navigateurs crée le Whawg

2.2 L'HyperText Transfer Protocol (HTTP)

- Utilise le TCP / IP (le port sera celui du navigateur)
- Une requête est envoyée au serveur par le navigateur
- Une réponse est envoyée au navigateur par le serveur



2.2 L'HyperText Transfer Protocol (HTTP)

- L'HyperText Transfer Protocol Secure, plus connu sous l'abréviation HTTPS — littéralement « protocole de transfert hypertexte sécurisé » — est la combinaison du HTTP avec une couche de chiffrement comme SSL ou TLS.

- En février 2017, le protocole de sécurité HTTPS était utilisé par environ 16,28 % de l'Internet français (source : Statistiques sur l'internet français. udomo.fr »)

Le serveur HTTP

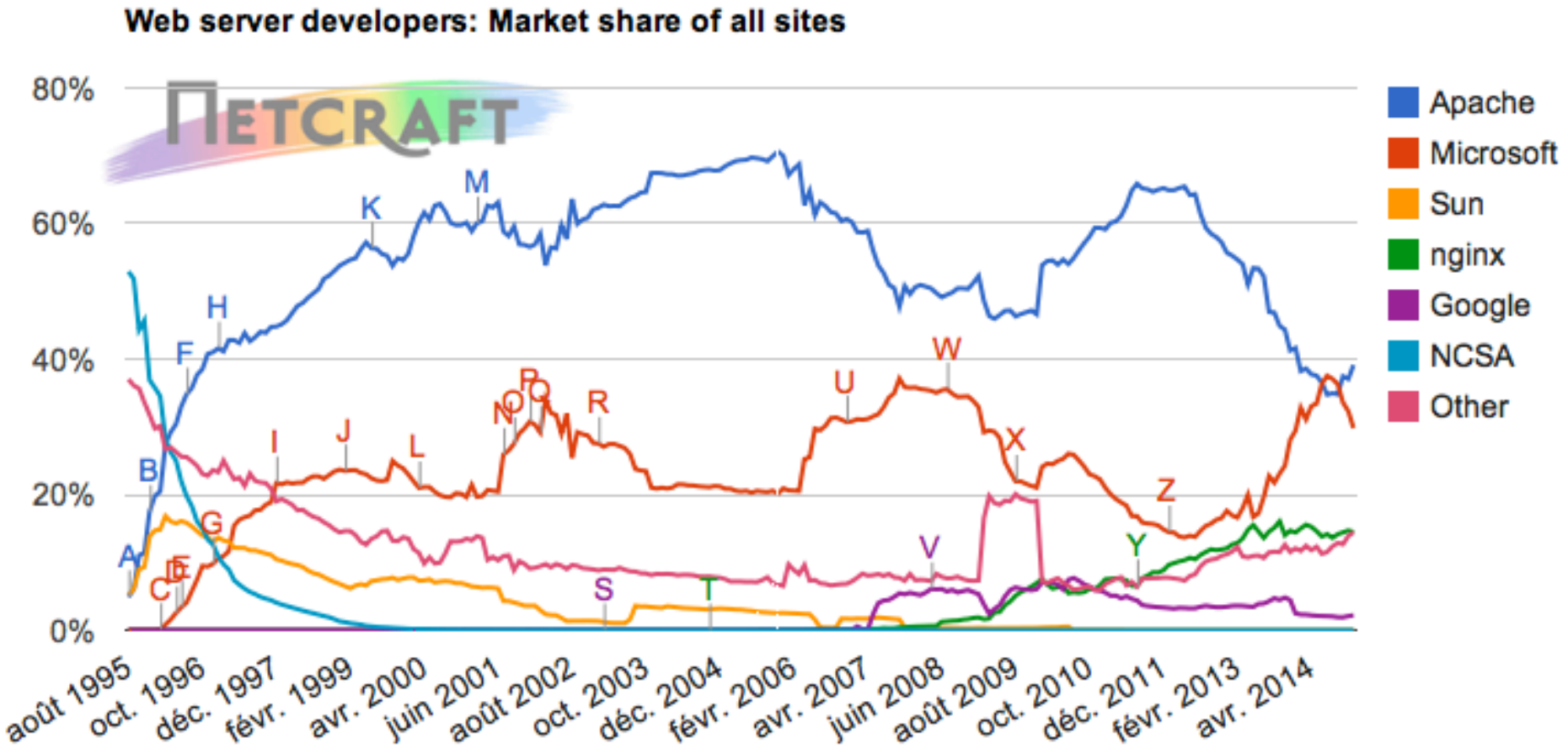
- Le serveur HTTP est le logiciel installé sur le serveur et permettant de gérer les requêtes/réponses

Quelques serveurs HTTP :

- *Apache HTTP Server* de la *Apache Software Foundation*, successeur du NCSA HTTPd ;
- *Apache Tomcat* de la *Apache Software Foundation*, évolution de Apache pour J2EE ;
- *BusyBox* httpd, utilisé dans le domaine de l'embarqué, et notamment avec *OpenWRT*[2] ;
- *Google Web Server* de *Google* ;
- *Internet Information Services* (IIS) de Microsoft ;
- *lighttpd* de Jan Kneschke ;
- *Monkey web server* de Eduardo Silva Pereira, dédié au noyau Linux, permettant d'utiliser pleinement ses fonctionnalités ;
- nginx d'Igor Sysoev ;
- NodeJS sous MIT Licence conçu par Ryan Lienhart Dahl en lignes de programmation en JavaScript ;
- *Sun Java System Web Server* de Sun Microsystems (anciennement iPlanet de Netscape, puis Sun ONE de Sun Microsystems) ;
- Tengine, fork de nginx, de Taobao (9^{e} rang mondial Alexa en juillet 2014) ;
- *Zeus Web Server* de Zeus Technology ;
- Gunicorn est un serveur web HTTP WSGI écrit en Python pour Unix.

Le serveur HTTP le plus utilisé est *Apache HTTP Server* qui sert environ 55 % des sites web en janvier 2013 selon Netcraft[3].

Le serveur HTTP

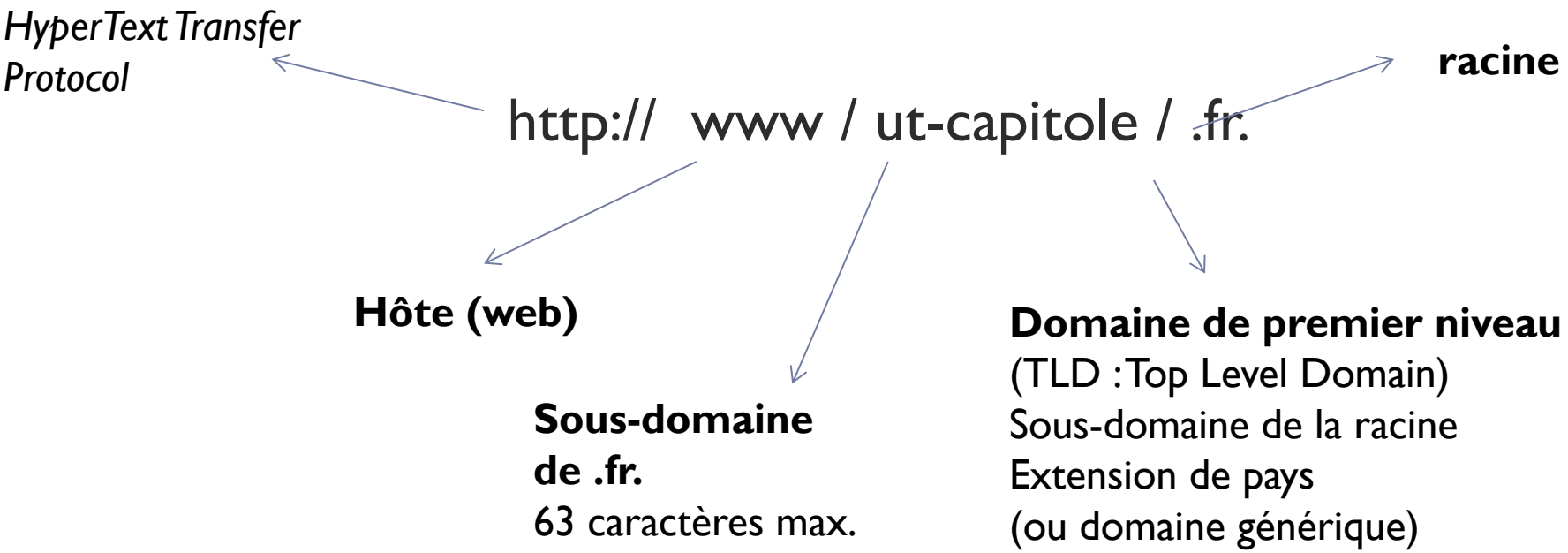


2.3 L'Uniform Resource Locator (URL)

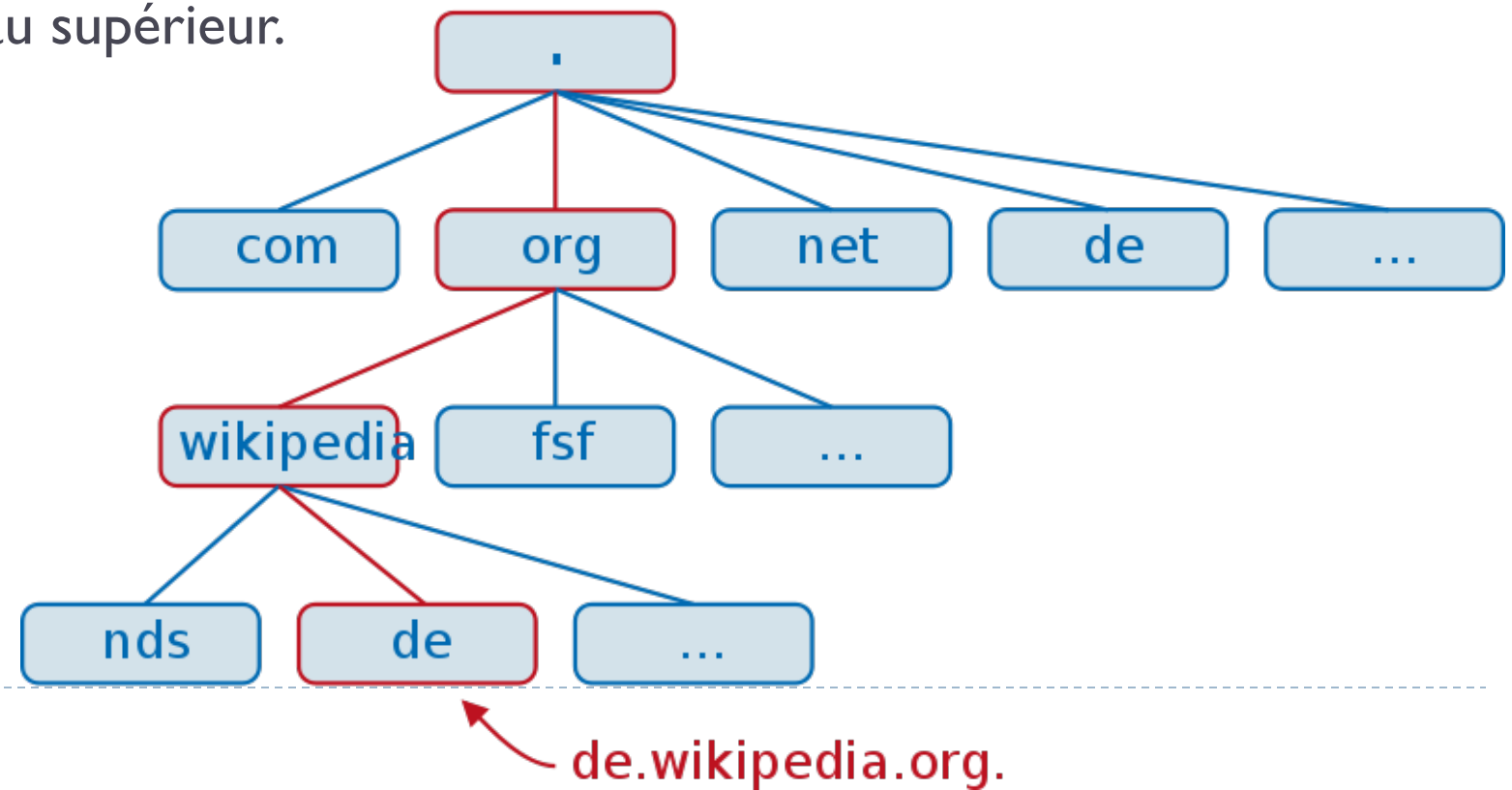
- L'URL répond à trois questions :
  - Comment ?
  - où ?
  - quoi ?

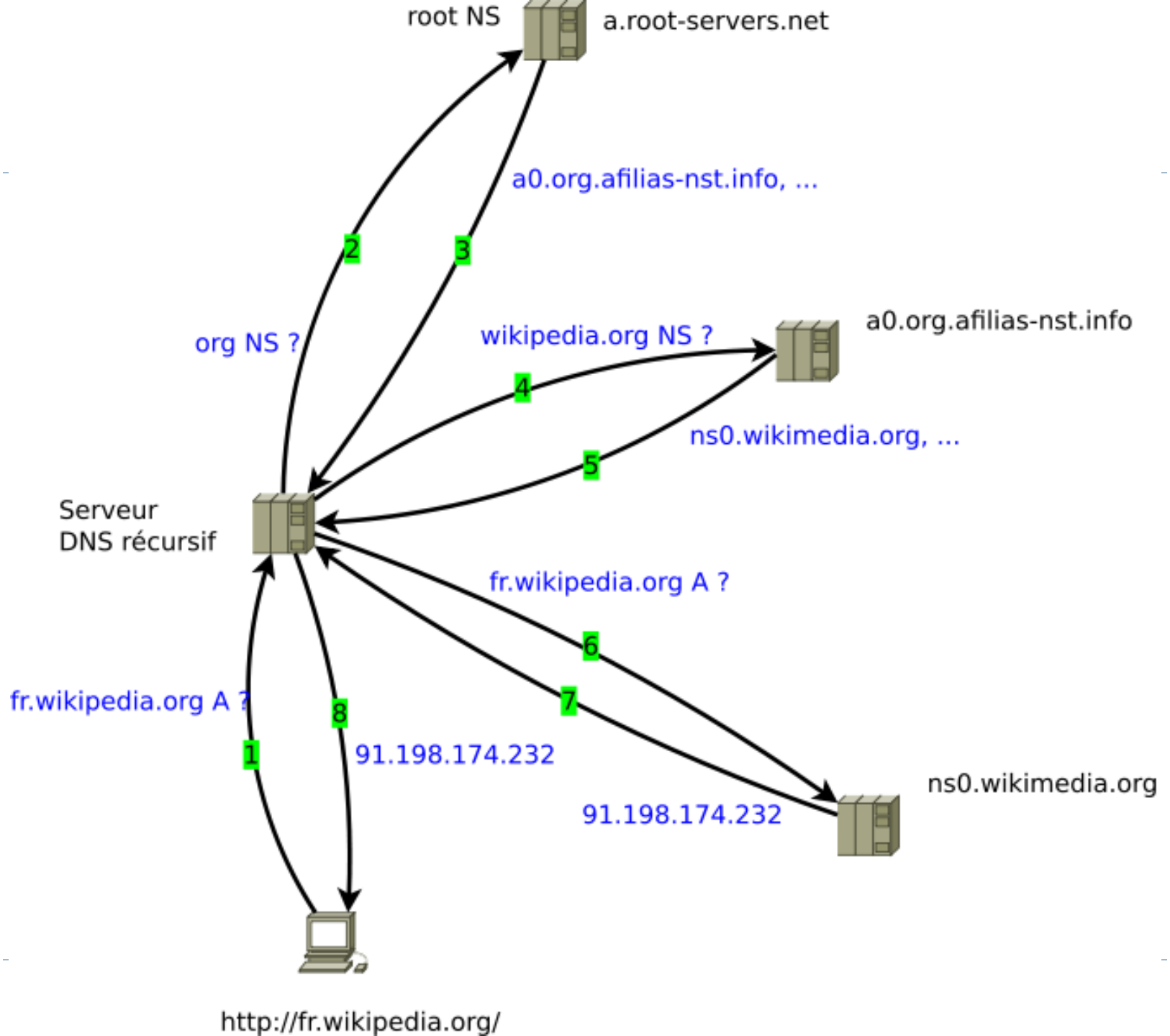
Exemple 1: `http://www.sebsauvage.net/comprendre/index.html`		
comment ?	`http://`	On utilise le protocole *http*.
où ?	`www.sebsauvage.net`	On va chercher l'information sur le serveur *www* du domaine *sebsauvage.net*.
	`/comprendre/`	dans le répertoire */comprendre/*
quoi ?	`index.html`	On va chercher le fichier *index.html*.

La composition d'une adresse URL

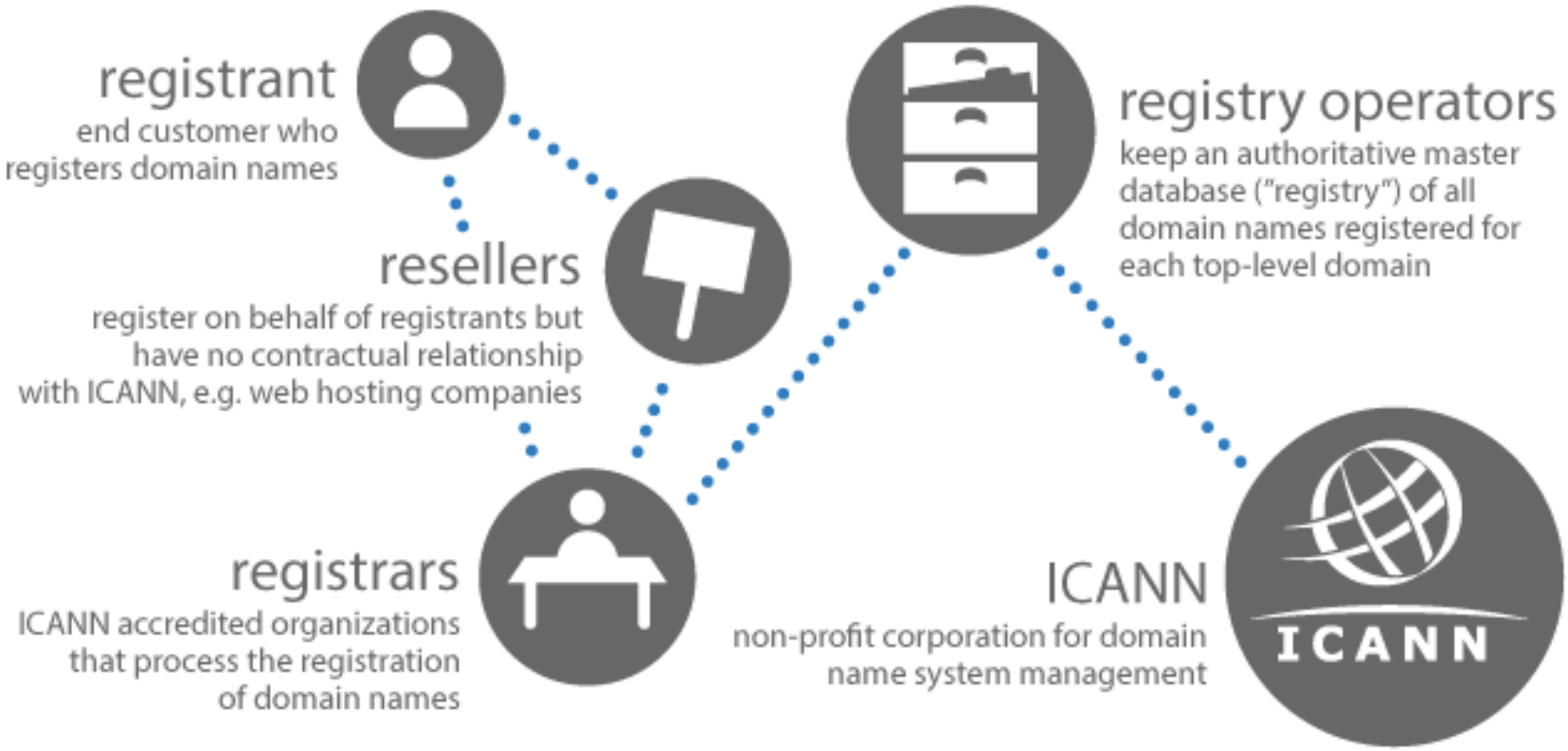


- Le Domain Name System (DNS) : une succession de délégations
  - Pour qu'il fonctionne normalement, un nom de domaine doit avoir fait l'objet d'une délégation correcte dans le domaine de niveau supérieur.




root NS
a.root-servers.net
a0.org.afilias-nst.info, ...
2
3
org NS ?
wikipedia.org NS ?
a0.org.afilias-nst.info
4
ns0.wikimedia.org, ...
5
Serveur
DNS récursif
fr.wikipedia.org A ?
6
7
fr.wikipedia.org A ?
8
1
91.198.174.232
ns0.wikimedia.org
91.198.174.232
http://fr.wikipedia.org/

Le Domain Name System (DNS)



domain registry process

Domain Name System (DNS)

- le domaine .fr représente l'ensemble des ordinateurs hébergeant des activités pour des personnes ou des organisations qui se sont enregistrées auprès de l' Association française pour le nommage Internet en coopération (Afnic) qui est le registre responsable du domaine de premier niveau .fr

- Sur le schéma précédent, les infos détenues par l'Afnic se trouve sur la deuxième catégorie de serveurs, dont l'adresse est donné par le serveur des noms racine (13 dans le monde), et qui donnera l'adresse du serveur ut-capitole.fr, qui lui donnera les infos sur l'adresse IP de ut-capitole.fr

Les systèmes alternatifs du DNS

Le projet « P2P DNS » (également appelé Dot-P2P9) de Peter Sunde, cofondateur de The Pirate Bay, ambitionne la création d'un système de noms de domaine « alternatif » et décentralisé basé sur la technologie P2P L'idée est de mettre en place un nouveau serveur DNS racine alternatif puis un système DNS distribué par l'intermédiaire de la technologie pair-à-pair. Le système serait distribué par les utilisateurs via l'installation d'une application basée sur le protocole de communication BitTorrent et où la transmission des données serait signée. Le projet P2P DNS vise à maintenir un Internet non censurable.

Le système DNS alternatif chinois lancé le 1er septembre 2006 utilise son propre DNS racine et ne passe plus par les serveurs de noms de domaines de l'ICANN. Par exemple, les domaines .com.cn et .net.cn apparaissent aux résidents chinois sous la forme .com et .net. ; cela créé alors dans la pratique deux réseaux Internet distincts : l'un est chinois, l'autre est mondial. Quant à un site chinois, il ne pourra être accessible du reste du monde que s'il en fait la demande auprès des autorités chinoises, qui publieront alors son nom dans les DNS officiels chinois qui eux sont reliés au DNS racine de l'ICANN

Les Common Gateway Interface

Dans certaines URL, on voit des ?p=home&action=search... ? C'est quoi ?

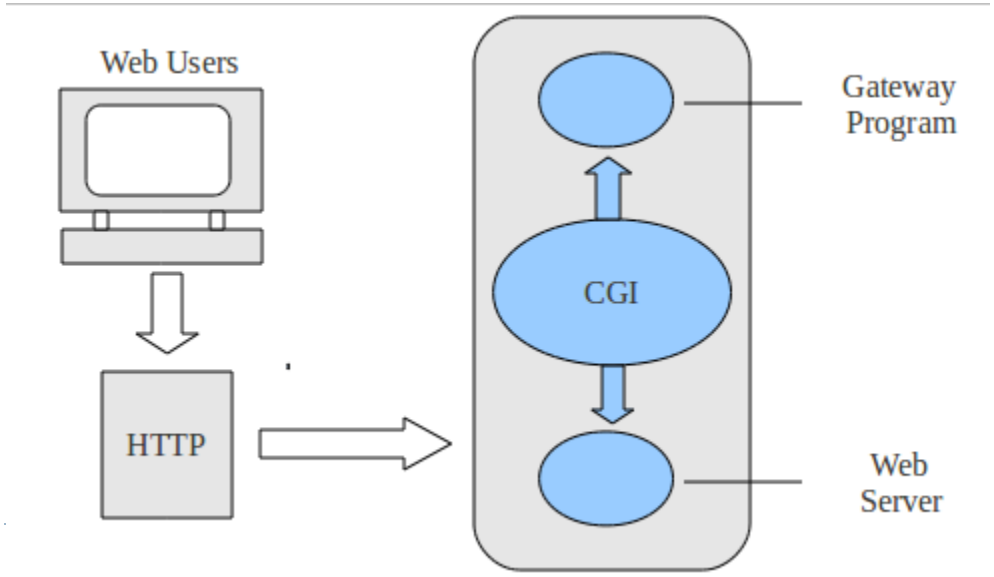
Une URL peut indiquer un fichier html, une image... et aussi un programme.

Quand vous voyez une URL contenant un point d'interrogation (?), tout ce qui se trouve après le point d'interrogation est donné au programme.

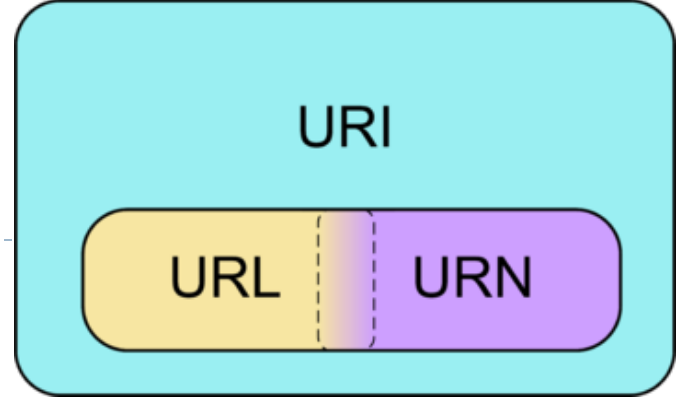
Par exemple: **`http://www.google.com/search?q=gateau+pommes&hl=fr`**

Dans ce cas, on va donner au programme *search* situé sur le serveur *www.google.com* les informations suivantes:

- `q=gateau+pommes` : nous recherchons les pages qui contiennent les mots *gateau* et *pommes*.
- `hl=fr` : nous voulons que l'interface de Google soit présentée en français.



URI ? URN ?



- ***Uniform Resource Identifier***,
  - Courte chaîne de caractères identifiant une ressource sur un réseau (par exemple une ressource Web) physique ou abstraite, et dont la syntaxe respecte une norme d'Internet mise en place pour le World Wide Web (voir IETF : RFC 3986).
- Deux types d'URI :
  - Les URL
  - Les URN
    - Un Uniform Resource Name (URN) est un URI qui identifie une ressource par son nom dans un espace de noms. Un URN peut être employé pour parler d'une ressource sans que cela préjuge de son emplacement ou de la manière de la référencer. Par exemple, l'URN urn:isbn:0-395-36341-1 est un URI qui, étant un numéro de l'International Standard Book Number (ISBN), permet de faire référence à un livre, mais ne suggère ni où, ni comment en obtenir une copie réelle.

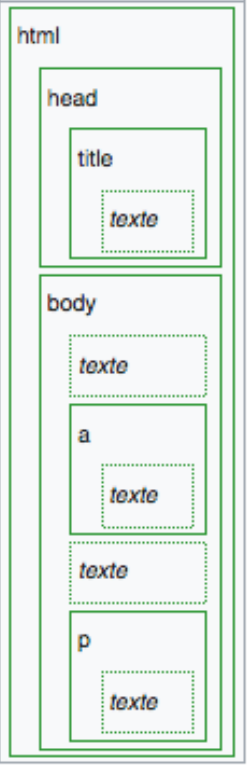
2.4 L'HyperText Markup Language (HTML)

- Permet de décrire les documents par zones balisées et d'effectuer les liens hypertextes

Source HTML

```
<!DOCTYPE html>
<html>
 <head>
  <title>
   Exemple de HTML
  </title>
 </head>
 <body>
  Ceci est une phrase avec un <a href="cible.html">hyperlien</a>.
  <p>
   Ceci est un paragraphe où il n'y a pas d'hyperlien.
  </p>
 </body>
</html>
```

Modèle du document



2.4 L'HyperText Markup Language (HTML)

- <!DOCTYPE html> définit ce document pour être HTML5
- <html> est l'élément racine d'une page HTML
- <head> contient des méta - informations sur le document
- <title> spécifie un titre pour le document
- <body> contient le contenu visible de la page
- <h1> définit un grand titre
- <p> définit un paragraphe

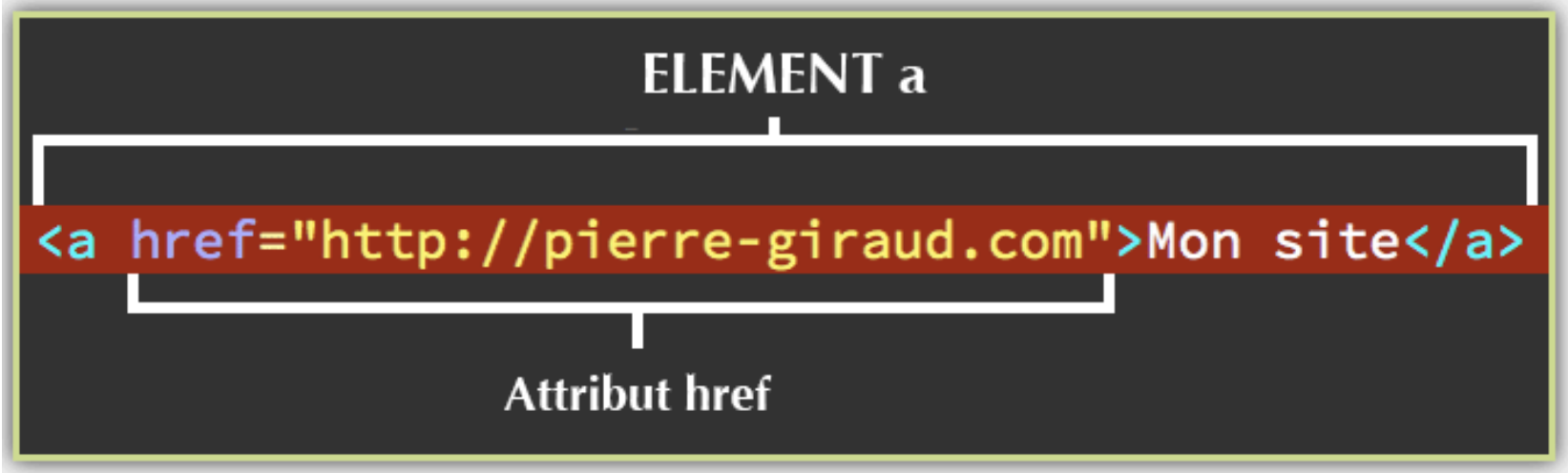
HTML définit également des éléments pour présenter le texte avec une signification spéciale .

- <B> - Le texte en gras
- <Strong> - texte Important
- <I> - Le texte en italique
- <Mark> - Marqué texte
- <Small> - Petit texte
- <Del> - texte supprimé
- <Sub> - texte Subscript
- <Sup> - texte Exposant
- <code> - code source

2.4 L'HyperText Markup Language (HTML)

- Les attributs HTML

Par exemple, l'élément `a` (pour "anchor") servant à créer des liens vers d'autres sites ou d'autres pages, va avoir besoin d'un attribut `href` ("hypertexte reference") qui va prendre comme valeur l'adresse (relative ou absolue) de la page vers laquelle on souhaite faire un lien.



Source : http://pierre-giraud.com/html-css/cours-complet/elements-balises-attributs-html.php

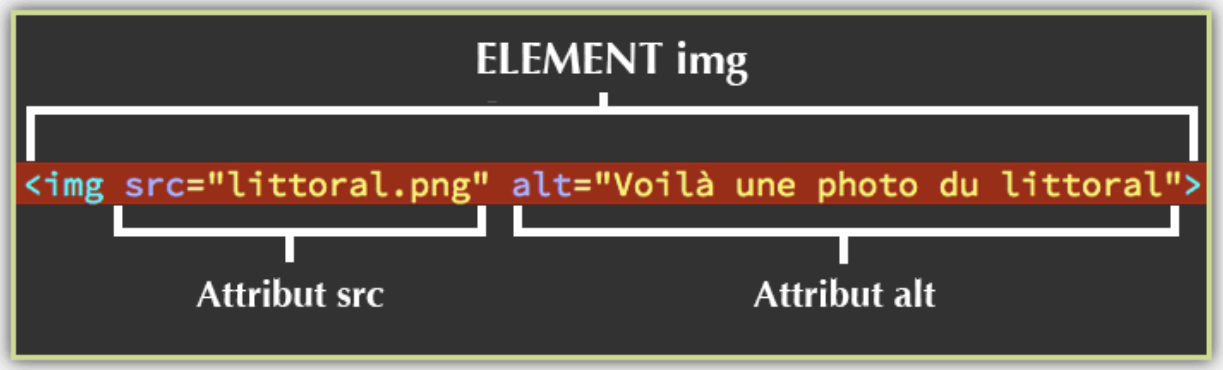
2.4 L'HyperText Markup Language (HTML)

- Les attributs HTML

L'élément `img`, servant à insérer une image dans une page HTML, va lui demander deux attributs : `src` et `alt`.

L'attribut `src` va prendre comme valeur le nom et l'emplacement de l'image tandis que l'attribut `alt` va afficher un texte alternatif dans le cas où l'image ne serait pas disponible (pour les non voyants par exemple).

Notez que cet élément n'est constitué que d'une seule balise orpheline, tout comme l'élément `br`.



Liste complète des attributs :
https://developer.mozilla.org/fr/docs/Web/HTML/Attributs

2.4 L'HyperText Markup Language (HTML)

- Les microdonnées / microformats / Rdfa / Schema.org
  - Un certain type d'attribut, par exemple dans les balises structurantes span :

```
Les oiseaux se sont perchés à
    <span class="geo">
      <span class="latitude">52.48</span>,
      <span class="longitude">-1.89</span>
    </span>
```

2.4 L'HyperText Markup Language (HTML)

- Les microdonnées / microformats / Rdfa / Schema.org

```
<div itemscope itemtype="http://schema.org/Movie">
  <h1 itemprop="name">Avatar</h1>
  <div itemprop="director" itemscope itemtype="http://schema.org/Person">
  Réalisateur : <span itemprop="name">James Cameron</span>
(né le <time itemprop="birthDate" datetime="1954-08-16">16 août 1954</time>)
  </div>
  <span itemprop="genre">Science-fiction</span>
  <a href="../movies/avatar-theatrical-trailer.html" itemprop="trailer">Bande-annonce</a>
</div>
```

2.5 Les autres formats

- Les langages du web :
  - HTML 5 : HyperText Markup Language 5 (format ouvert)
  - Javascript : petits programmes avec interfaces (les Applets) (ouvert)
  - Active X : semblable à Java, développé par Microsoft (propriétaire)
  - Flash : permet de créer des animations (format propriétaire – Adobe, nécessite un plug-in Flash Player)

2.5 Les autres formats

- Les formats

  - Un format est ce qui permet à un logiciel d'interpréter (i.e. de traduire) les données brutes d'un fichier. C'est, autrement dit, le mode de représentation de ces données.

  - Les formats des fichiers sont indiqués par leur extension : le suffixe (souvent trois lettres) attaché au nom de fichier.

  - En plus du .html, il existe des formats pour les images (par exemple JPEG, PNG, GIF, TIF, BMP), pour les pages prêtes à l'impression (PDF, PS) etc.

2.5 Les autres formats

- Les formats ouverts
  - On dira que le format d'un fichier est ouvert si le mode de représentation de ses données est transparent et/ou sa spécification appartient au domaine public. Il s'agit principalement de standards établis par des autorités publiques ou des institutions internationales, dont le but est de fixer des normes assurant l'interopérabilité entre logiciels. Il existe pourtant plusieurs cas de formats ouverts promus par des entreprises ayant décidé de rendre publique la spécification de leurs formats.

- Les formats propriétaires (ou fermés)
  - On dira qu'un format est propriétaire si le mode de représentation de ses données est opaque et sa spécification n'appartient pas au domaine public. Il s'agit de formats développés par des entreprises afin de coder les données d'une application qu'elle produit : seuls les logiciels produits par cette entreprise seront capables de lire correctement et complètement les données contenues dans un fichier à format propriétaire. Les formats propriétaires peuvent être légalement protégés par des brevets et imposer le versement de royalties à tous ceux qui les utilisent.

2.6 Les principaux acteurs

- Le développeur

2.6 Les principaux acteurs

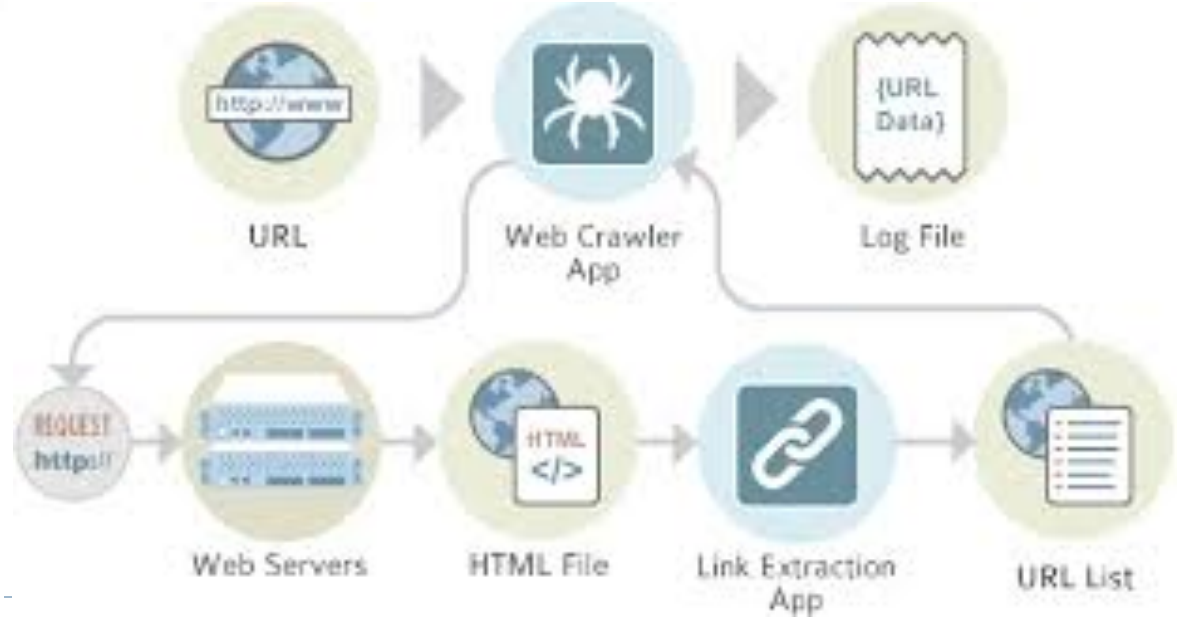
- Le navigateur

L'éditeur peut choisir de s'adresser différemment à différents navigateurs :

```
1 <!--[if !IE]><!-->
2 Code HTML pour les navigateurs différents d'IE
3 <!--<![endif]-->
```

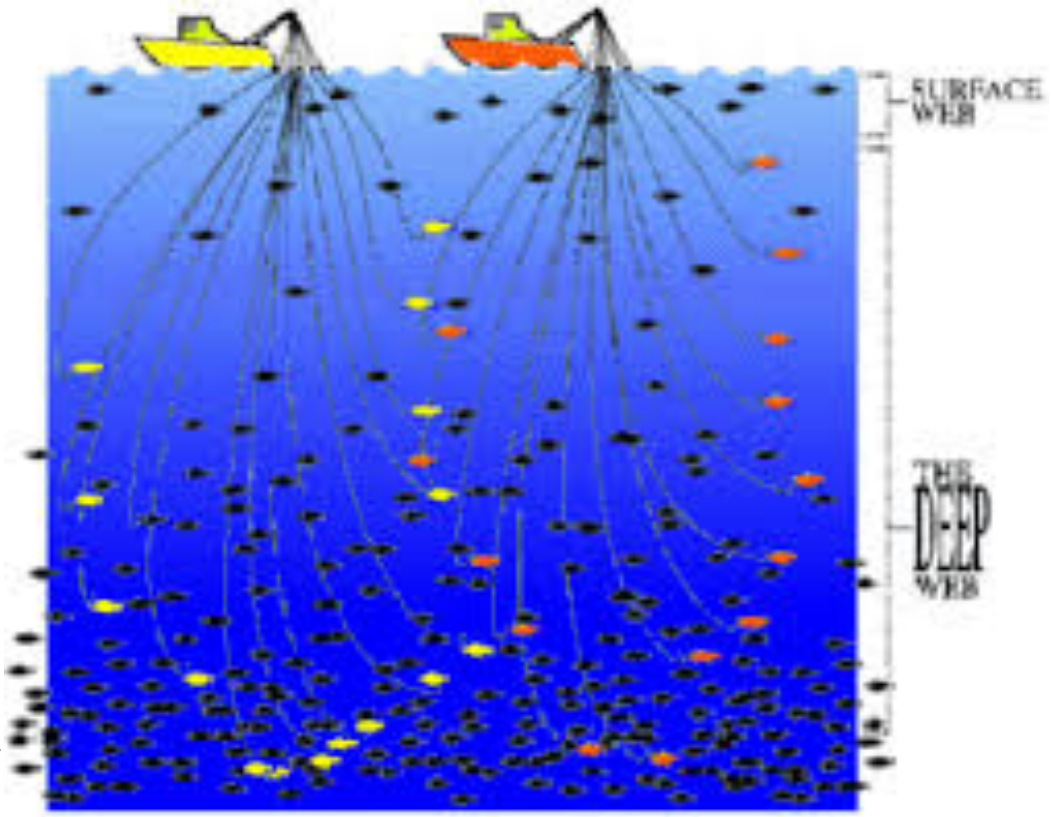
- Le crawler

XML
XML
XML
XML
Google


URL
Web Crawler App
{URL Data}
Log File
REQUEST http://
Web Servers
HTML
HTML File
Link Extraction App
URL List

2.7 Les profondeurs du web

- Le web opaque : indexable, non profond, mais non indexé (car le moteur n'a pas eu le temps d'y passer)



2.7 Les profondeurs du web

- Le web privé : l'éditeur a exclu du champ d'indexation son contenu, grâce au protocole robots.txt (pour le site) ou à une meta-tag : <noindex> (pour une page)

2.7 Les profondeurs du web

- Le web profond (web invisible ou encore web caché (en anglais deep web) : non indexé par les moteurs
  - contenu dynamique ;
  - contenu non lié ;
  - contenu à accès limité ;
  - Les pages liées par JavaScript (incompris par les robots)
  - formats non indexables.

  - En 2008, le web dit « invisible » non référencé par les moteurs de recherche représente 70 à 75 % de l'ensemble, soit environ un trilliard de pages non indexées

  - En 2014, selon NewScientist : only 0.03% of the web is searchable, hence leaving 99.97% of all data being dark Internet

2.7 Les profondeurs du web

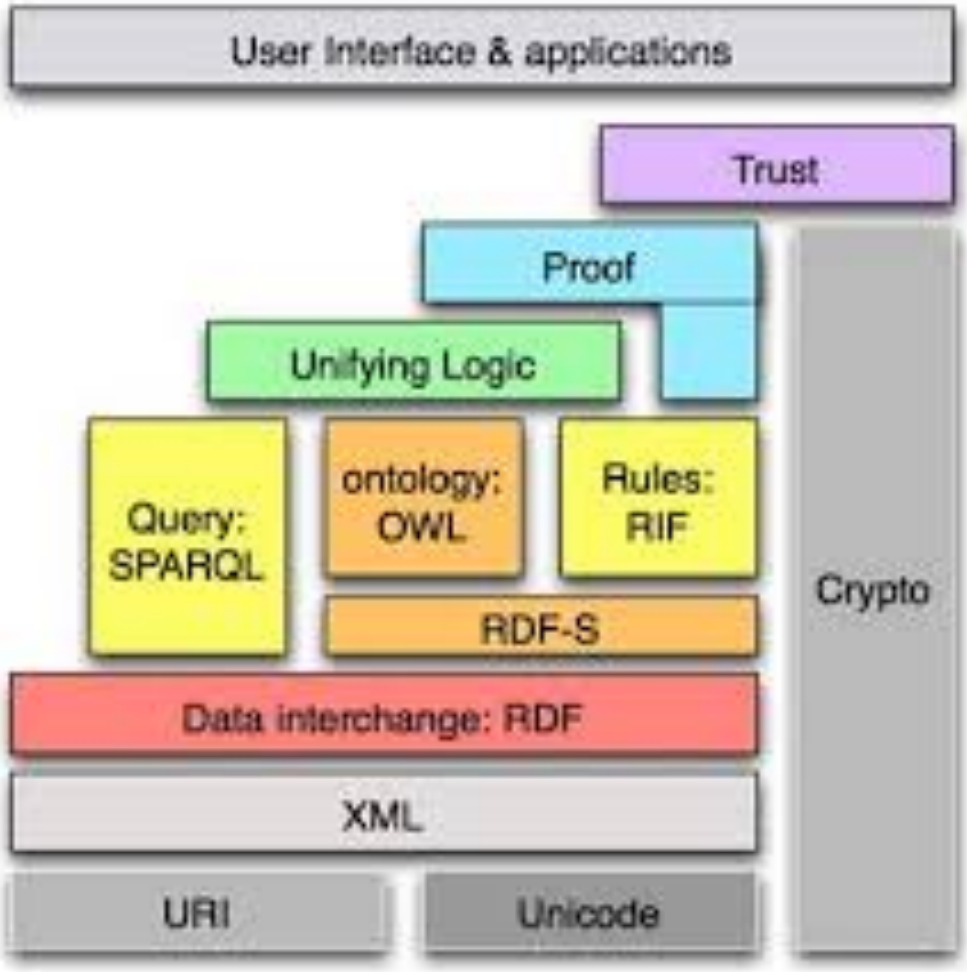
2.8 Le web sémantique

Objectif : ajouter des attributs spécifiques afin et de donner du « sens » aux données

Problème : les moteurs et les navigateurs ne « comprennent pas » les documents

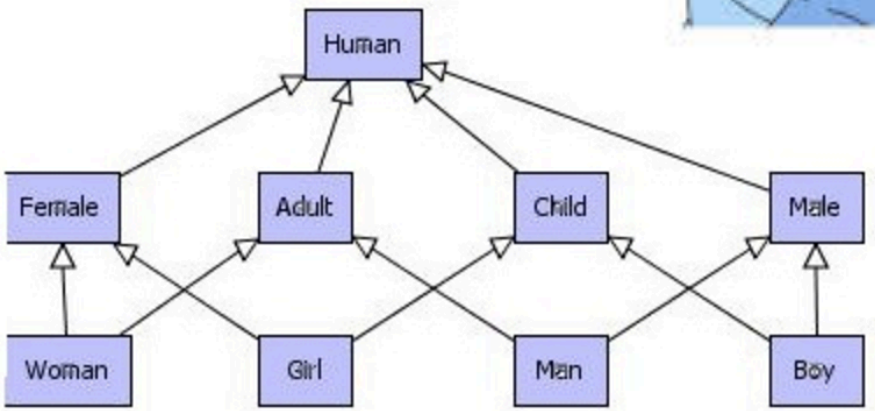
Ex : « ordinateur portable » et « laptop » ne donnent pas le même résultat sur Google alors qu'ils signifient la même chose

Pour arranger cela il ne faudrait plus que Google se réfère aux mots mais aux « classes » (i.e. le concept d'ordinateur portable, la chose en soi…)



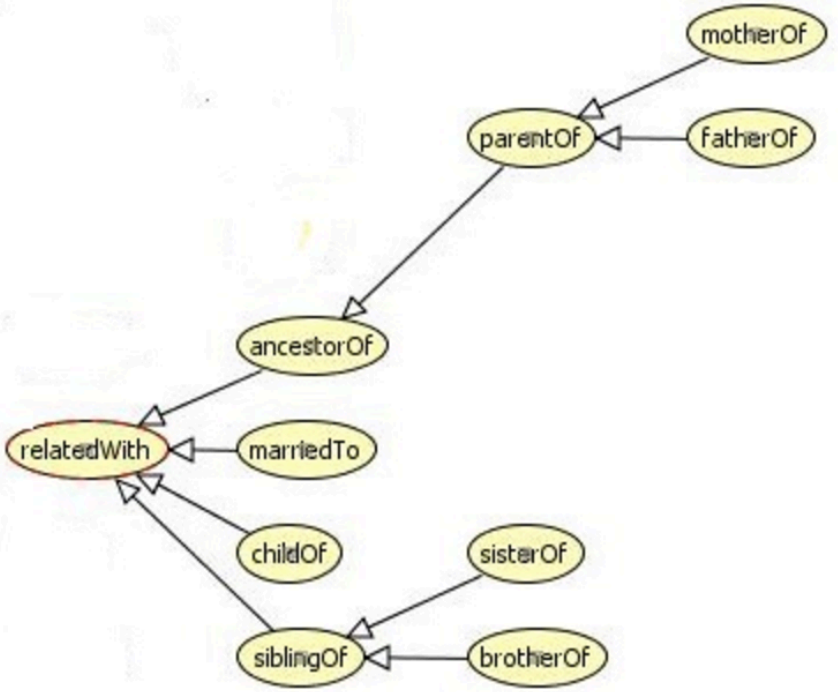
2.8 Le web sémantique

Classes :



2.8 Le web sémantique

Propriétés :



2.8 Le web sémantique

Maintenant, admettons quelques règles, logiques pour un humain sur ces classes et propriétés :

- règle 1 : si une classe « Male » appartient à une propriété « brotherOf » avec une classe « Female » alors la classe « Female » appartient à la propriété « sisterOf » avec la classe « Male » ;
- règle 2 : réciproque de la règle 1 ;
- règle 3 : si une classe « Adult » appartient à une propriété « parentOf » avec une classe « Human » alors la classe « Human » appartient à la propriété « childOf » avec la classe « Adult » ;
- règle 4 : réciproque de la règle 3.

De cette manière, on peut facilement imaginer toutes les règles possibles entre ces propriétés et classes. Pour finir, prenons quelques instances : {Bart, Lisa, Homer}. On dit ensuite que l'instance « Bart » appartient à la classe « Boy », « Lisa » appartient à la classe « Girl » et « Homer » appartient à la classe « Man ». On dit ensuite que « Bart » est le frère de « Lisa » et que « Bart » et « Lisa » sont les enfants de « Homer » avec nos propriétés.

Notre ontologie ainsi construite, il nous est maintenant facile d'effectuer les requêtes suivantes dessus :

- est-ce que Lisa est la sœur de Bart ?
- est-ce que Homer est un parent de Lisa et Bart ?

Bien entendu cette ontologie est ridiculement petite et facile comparée aux ontologies de domaines plus complexes. Mais, maintenant que vous savez ce qu'est une ontologie, vous pouvez vous amuser à continuer l'ontologie de la famille Simpson en ajoutant des propriétés, des classes, des règles et des instances.

Source : http://jplu.developpez.com/tutoriels/web-semantique/introduction/

2.8 Le web sémantique

- La sémantique des données est décrite grâce à des langages prévus pour fournir une description formelle de classes, relations et règles. Ces langages sont :
  - RDFS (Ressource Description Framework Schema)
  - OWL (Web Ontology Language).

- Il existe des langages de description des données structurées dans le HTML afin que des outils comme les moteurs de recherche puissent effectuer un traitement automatique de ces différentes données. Ces langages sont :
  - RDFa
  - Microformat
  - Microdata

Microformats

RDFa

Microdata

RDFa (W3C) : RDF intégré au HTML

- Exemple avec RDFa (Albuquerque, Nouveau Mexique) :

```
<span typeof="v:Address">
<span property="v:locality">Albuquerque</span>
<span property="v:region">NM</span>
</span>
</div>
```

https://www.slideshare.net/shamod/rdfa-what-why-and-how-by-mike-hewett-and-shamod-lacoul

Microdata (Whatwg)

Boston Cream Pie | Recipe - Allrecipes.com

allrecipes.com/recipe/boston-cream-pie-i/

★★★½☆ 45 reviews - 325 cal

A rich creamy dessert also known as **Boston Cream** Cake. It is a wonderful combination of yellow cake, custard and chocolate glaze.

↳ Boston Cream Pie II - Boston Cream Pie III - Recipes Like - Reviewer

Boston Cream Pie Recipes

www.foodnetwork.com/recipe.../boston-cream-pie/index.html

Don't wait to indulge in a slice of **Boston cream pie**, a homey treat that calls to mind Mom's best baking. Try one of these, our most popular **Boston cream pies**, ...

Properties from Recipe		
cookTime	Duration	The time it takes to actually cook the dish, in ISO 8601 duration format.
cookingMethod	Text	The method of cooking, such as Frying, Steaming, ...
ingredients	Text	An ingredient used in the recipe.
nutrition	NutritionInformation	Nutrition information about the recipe.
prepTime	Duration	The length of time it takes to prepare the recipe, in ISO 8601 duration format.

```
<div itemscope itemtype="http://schema.org/Recipe">
  <span itemprop="name">Mom's World Famous Banana Bread</span>
  By <span itemprop="author">John Smith</span>,
  <meta itemprop="datePublished" content="2009-05-08">May 8, 2009
  <img itemprop="image" src="bananabread.jpg" />

  <span itemprop="description">This classic banana bread recipe comes
  from my mom -- the walnuts add a nice texture and flavor to the banana
  bread.</span>

  Prep Time: <meta itemprop="prepTime" content="PT15M">15 minutes
  Cook time: <meta itemprop="cookTime" content="PT1H">1 hour
  Yield: <span itemprop="recipeYield">1 loaf</span>

  <div itemprop="nutrition"
    itemscope itemtype="http://schema.org/NutritionInformation">
    Nutrition facts:
    <span itemprop="calories">240 calories</span>,
    <span itemprop="fatContent">9 grams fat</span>
  </div>

  Ingredients:
  - <span itemprop="ingredients">3 or 4 ripe bananas, smashed</span>
  - <span itemprop="ingredients">1 egg</span>
  - <span itemprop="ingredients">3/4 cup of sugar</span>
  ...
```

```HTML
<article itemscope itemtype="http://schema.org/Article">
  <header>
    <h1 itemprop="name">Microformats versus Microdatas</h1>
    <ul>
      <li>Le <time itemprop="dateCreated" datetime="2012-11-20T20:00">20 novembre
2012</time> par <span itemprop="creator">Jojaba</span></li>
      <li>Permalien : <a href="http://lien_vers_article.html"
itemprop="url">Microformats versus Microdatas</a></li>
    <ul>
  </header>
  <section itemprop="text">
    <p>
    Les Microdatas sont relativement récents, puisque apparus avec HTML5 alors que les
```

Microformats (CommerceNet, non standardisés)

- Wiki ouvert / liste de diffusion
- http://microformats.org/

```
<div class="location vcard">
 <span class="fn org">Shakespeare's Pub </span>
 <div class="adr">
  <div class="street-address">314 East 6th Street</div>
  <div class="locality">Austin, TX 78701</div>
 </div>
</div>
```

Microformats (CommerceNet, non standardisés)

Considérez l'information de contact :

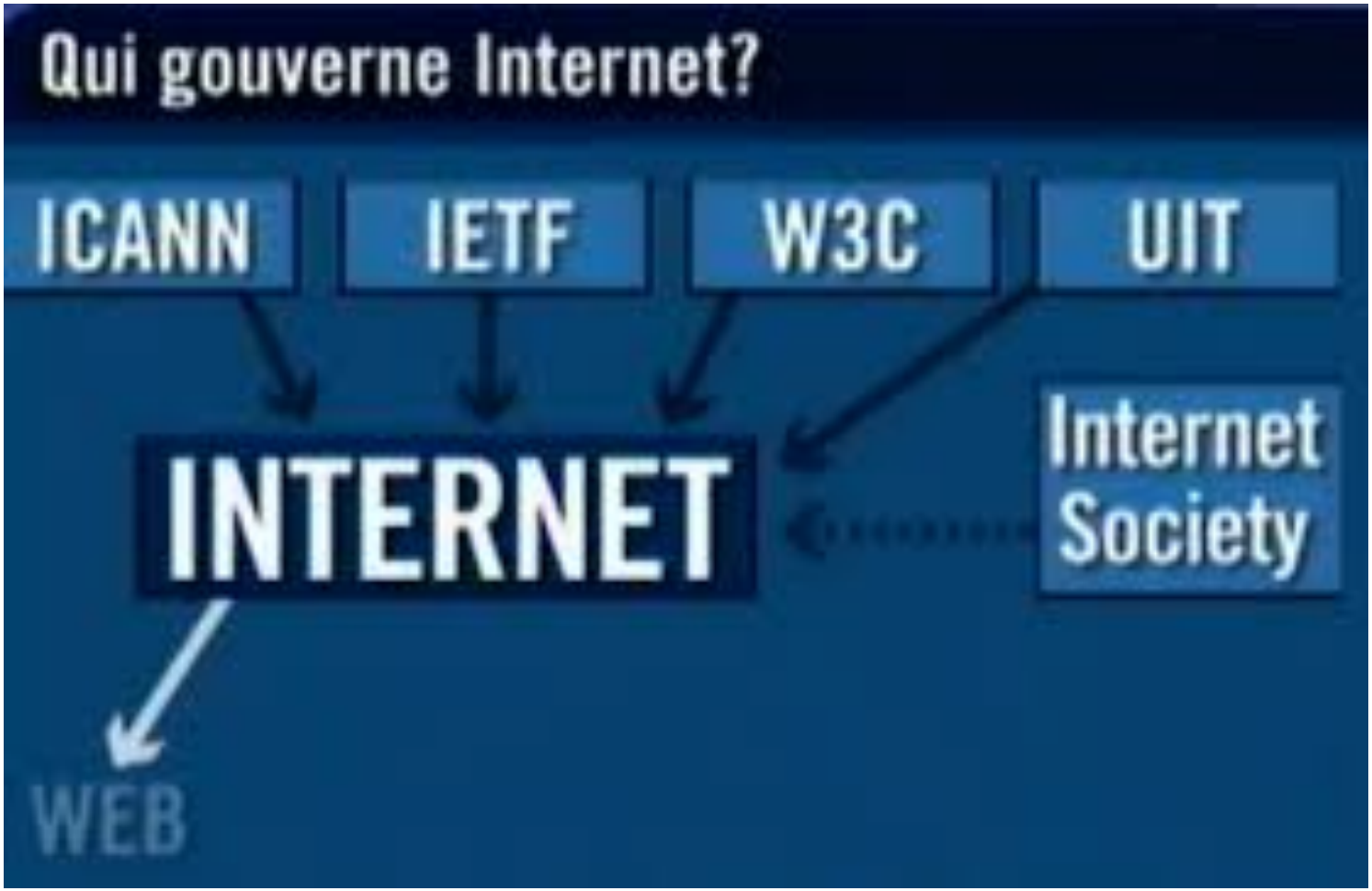
```
<address>
 <p>
  Jean Bout<br/>
  Société Exemple<br/>
  604-555-1234<br/>
  <a href="http://exemple.com/">http://exemple.com/</a>
 </p>
</address>
```

Avec le marquage microformat hCard, ceci devient :

```
<address class="vcard">
 <p>
  <span class="fn">Jean Bout</span><br/>
  <span class="org">Société Exemple</span><br/>
  <span class="tel">604-555-1234</span><br/>
  <a class="url" href="http://exemple.com/">http://exemple.com/</a>
 </p>
</address>
```

3. Gouvernance et gouvernementalité

3.1 Les principaux acteurs de la régulation technique



Internet Corporation for Assigned Names and Numbers (ICANN)

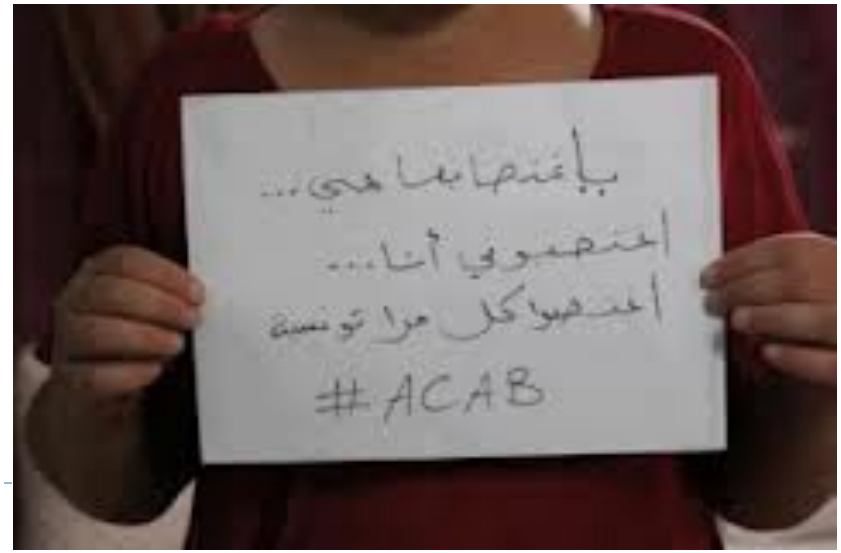
- Créée en 1998 sous l'impulsion du vice-président Al Gore
- Contrôle l'affectation des noms de domaines de premier niveau, l'ICANN délivre en pratique un droit de délégation sur la vente des noms de domaines à différentes organisations, comme VeriSign pour les domaines .com et .net ou l'AFNIC pour le domaine .fr.
- Sa compétence est mondiale et ses décisions s'imposent de fait aux États, alors même qu'elle est de droit californien, se trouve soumise de ce fait au procureur général de Californie et relève en dernière instance du département du Commerce des États-Unis.

Internet Corporation for Assigned Names and Numbers (ICANN)

- ICANN
  - Le rôle de l'ICANN est régulièrement remis en question, notamment à cause de ses liens avec le gouvernement américain.
  - De nombreux pays aimeraient que la fonction dévolue actuellement à l'ICANN soit prise en charge par un organisme dirigé par l'ONU.
  - L'ONU elle-même a indiqué avoir eu des projets dans ce domaine, mais le gouvernement Bush s'y était fortement opposé.
  - D'autres personnes préféreraient une autre organisation sous forme coopérative qui ne soit ni l'ICANN ni l'ONU, qui est inter-gouvernementale et où la société civile n'est pas représentée.

Internet Corporation for Assigned Names and Numbers (ICANN)

- ICANN
  - Le 30 octobre 2009, l'ICANN vote la fin de l'exclusivité de l'alphabet latin pour la rédaction des noms de domaine Internet.
  - Depuis le 16 novembre 2009, peuvent être enregistrées des adresses web rédigées avec des caractères arabes, chinois, coréens ou japonais



Icann : pourquoi l'ONU ne récupère pas le contrôle d'Internet

Dès samedi 1er octobre, la Maison-Blanche transfère son pouvoir sur les noms de domaine à l'Icann, une organisation à but non lucratif.

PAR GUERRIC PONCET

Publié le 30/09/2016 à 17:06 | Le Point.fr

Depuis des années, des voix s'élevaient pour réclamer la fin de la tutelle américaine sur l'organisme qui gère les noms de domaine sur Internet, l'Icann (Internet Corporation for Assigned Names and Numbers). Le système des noms de domaine permet à un ordinateur d'accéder à un site .com, .fr ou .org, par exemple. C'est désormais chose faite : le département américain du Commerce, qui contrôlait jusqu'alors l'Icann, va lui donner sa liberté samedi 1er octobre. Même si Washington n'exerçait plus qu'une autorité symbolique, ce changement est un symbole fort.

L'Icann se définit sur son site web comme un « organisme à but non lucratif responsable de la sécurité, de la stabilité et de la coordination mondiale du système d'identificateurs uniques de l'Internet ». Il a été choisi par les États-Unis pour continuer à assurer le bon fonctionnement du réseau, mais cette solution ne satisfait personne à part l'Icann. Concrètement, ce sont les membres de l'organisation, c'est-à-dire des représentants des États, des entreprises et de la société civile originaires du monde entier, qui prennent les décisions. Désormais officiellement libérés du contrôle de Washington, ils décideront de l'avenir du Net selon des règles de codécision censées prévenir toute dérive. Ce modèle autorégulé, qui transfère une partie du pouvoir hors des États-Unis, a été préféré à d'autres, et notamment à la piste onusienne.

L'ONU, une piste dangereuse

Un transfert à l'ONU et à son agence spécialisée l'ITU (Union internationale des télécommunications) avait été réclamé à Barack Obama par de nombreux États, qui disposent d'un pouvoir plus important dans l'organisation internationale qu'à l'Icann. Mais ce scénario aurait impliqué de donner beaucoup de pouvoir à des gouvernements potentiellement totalitaires. La Maison-Blanche a préféré éviter que le scénario du Conseil des droits de l'homme de l'ONU ne se reproduise : des pays piétinant régulièrement les droits de l'homme, comme l'Arabie saoudite, y ont obtenu leur siège... De là à imaginer que des chefs d'État totalitaire puissent s'allier pour présider aux destinées d'Internet en lui imposant une censure et une surveillance encore plus fortes qu'aujourd'hui, il n'y a qu'un pas.

De plus, compte tenu de la lourdeur bureaucratique des Nations unies, il est peut-être préférable qu'un média aussi instantané qu'Internet soit gouverné ailleurs, dans des cercles plus réactifs et moins vulnérables aux manoeuvres de politique internationale. L'Icann doit désormais faire bonne figure et assurer son rôle avec le plus de transparence possible.

Internet Engineering Task Force (IETF)

- Groupe informel, international, ouvert à tout individu/organisation, qui participe à l'élaboration de standards Internet.
- Le but du groupe est la rédaction d'un ou plusieurs Request for comments (RFC), nom donné aux documents de spécification à la base d'Internet.
  - Les groupes de travail sont répartis dans une dizaine de domaines d'intérêt, chaque domaine ayant un ou deux directeurs. Ces directeurs de domaine nomment le ou les directeurs de chaque nouveau groupe de travail.
  - Les directeurs de domaine font partie de l'Internet Engineering Steering Group (IESG). L'Internet Architecture Board (IAB) se charge de l'orientation à long terme d'Internet, et donc des activités données à l'IETF. L'IESG et l'IAB sont chapeautés par l'Internet Society (ISoc).

Internet Society (ISOC)

- Association américaine créée en 1992 par les « pionniers »
- Autorité morale et technique sans doute la plus influente dans l'univers du réseau Internet.
- Apporte à l'IETF un support organisationnel et financier
- 28 000 membres (personnes physiques) venant de 170 pays, ainsi que plus de 80 organisations et 90 filiales réparties dans le monde entier.
- Chaque discussion concernant Internet passe par eux…

Internet Society France™

Internet Society (ISOC)

- **L'Isoc dans l'actualité**
  - Défense de l'Internet libre et ouvert (idéologie des pionniers)
  - Critique de iTunes, Facebook : jardin clos…
  - Très virulent vis-à-vis des applications Apple…
    - « Vous êtes un utilisateur d'Apple, moi de Microsoft, et nous ne parvenons plus à communiquer ensemble parce que nous sommes sur deux protocoles différents. Vous êtes sur votre terrain de jeu, moi sur le mien. La fragmentation peut venir aussi au niveau de la gestion des câbles. Les fournisseurs sont en train de revendiquer l'idée de faire payer certains distributeurs de contenus. Si ce genre de chose arrivait, pourquoi Google livrerait-il encore ses services en Afrique ? Le troisième risque, ce sont les grandes discussions avec les Etats. Il ne faut pas regarder si loin que la Russie ou la Chine : même chez vous, l'Etat commence à filtrer Internet pour de fausses bonnes raisons, par exemple le « cyberterrorisme ». (Frédéric Donck, patron européen de l'Internet Society)

World Wide Web Consortium (W3C)

- Organisme de normalisation à but non lucratif
- Fondé par Tim-Berners Lee
- Géré par des universités :
  - MIT (USA),
  - ERCIM (UE) European Research Consortium for Informatics
  - Université de Keio (Japon)
- Ouvert à tous (prix d'entrée variable, qui peut aller jusqu'à 59 000 euros par an pour une grande entreprise).
- 65 employés / 463 membres en 2017 : industriels, éditeurs, navigateurs, moteurs de recherche, fournisseurs d'accès
- Partisan des formats ouverts (HTML) et des normes stables
- Recommandations concernant le web en particulier (l'IETF se chargeant d'Internet en général)

World Wide Web Consortium (W3C)

- Politique des brevets libres de droits (royalty-free)
  - Les membres du W3C qui détiennent un brevet susceptible d'être inclus dans une recommandation du consortium sont sommés d'abandonner leurs droits
  - Il n'y a pas de « licence W3C », seulement des conditions à remplir pour qu'une technologie propriétaire, quelle qu'elle soit, puisse être intégrée à une REC du W3C

Web Hypertext Application Technology Working Group (Whatwg)

- Créé en 2004 par Ian Hickson (« Hicksie »)
- Auteur du HMTL5 (contre le XHTML 2.0)
- Procédure beaucoup plus souple
- Membres : tout le monde peut contribuer en théorie, mais ce sont essentiellement les concepteurs des principaux navigateurs qui oeuvrent
- Partisan du Living Standard

« Le débat Living Standard / anti-Living Standard est stupide : les moyens techniques dont nous disposons permettent de travailler en continu sur une version « vivante » et à intervalles réguliers de faire un « snapshot » qui est tout simplement une version du même document purgée de ce qui n'est pas stable ou réaliste. (...) Le problème principal est que même si la situation entre le W3C et le Whatwg est à relativement normaliser (disons, comme entre l'Inde et le Pakistan), personne ne veut travailler à franchir les dernières marches. J'avais proposé un modus vivendi *(l'accord « Kadesh ») fondé sur l'idée qu'il était possible de produire des documents Living Standard publiés régulièrement en Royalty-Free. Mais le Whatwg l'a malheureusement rejeté après de longues négociations. »*

Robin Berjon, 2016

International Telecommunication Union ou (ITU)

- L'organe des Nations Unis pour essayer d'harmoniser les normes relatives aux TIC

- Du point de vue d'Internet, cela sert essentiellement de théâtre à la contestation de la domination américaine

Electronic Frontier Fondation (EFF)

- ONG fondée en 1990 (entre autre par John Perry Barlow, connu pour sa Déclaration d'indépendance du Cyberespace)
- Farouche défenseur de la liberté d'expression
- S'oppose aux brevets qu'elle estime nuisible au développement et à l'usage public

En France : la Quadrature du Net

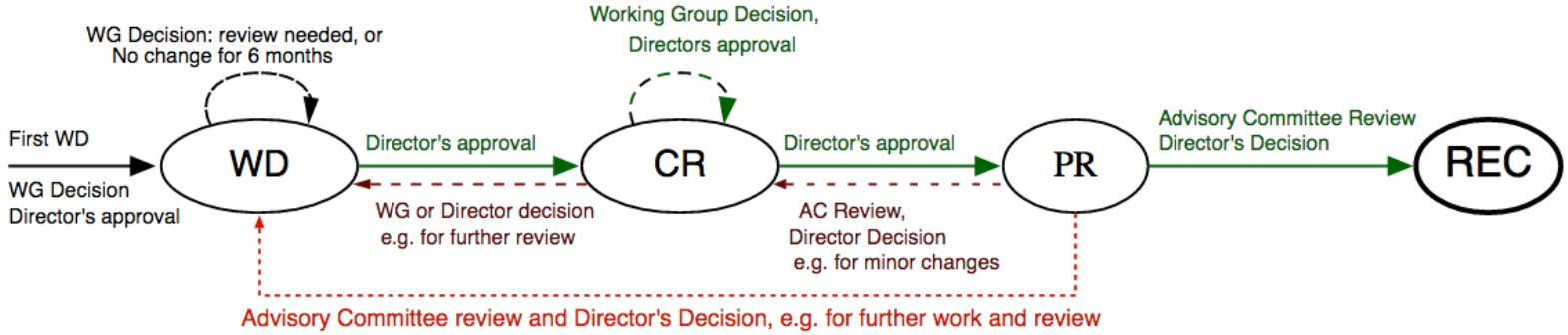
Qu'est-ce que La Quadrature?

La Quadrature du Net est une **association de défense des droits et libertés des citoyens sur Internet**. Elle promeut une adaptation de la législation française et européenne qui soit fidèle aux valeurs qui ont présidé au développement d'Internet, notamment la libre circulation de la connaissance.

À ce titre, La Quadrature du Net intervient notamment dans les débats concernant la liberté d'expression, le droit d'auteur, la régulation du secteur des télécommunications ou encore le respect de la vie privée.

3.2 La fabrique des normes au W3C

Il y a cinq étapes à franchir pour chaque modification significative : First Working Draft (First WD), Working Draft (WD), Candidate Recommendation (CR), Proposed Recommendation (PR) et Recommendation (REC). Seule la REC peut être considérée comme une norme stable.



Pour franchir chacune des étapes, il existe des conditions qui doivent obligatoirement être remplies (*must*) ainsi que certaines conditions facultatives (*may* et *should*). S'il est jugé par le Comité consultatif (*Advisory Committee*), composé d'un représentant pour chaque membre du W3C, ou par Tim Berners-Lee directement, que les conditions techniques ne sont pas rencontrées, le texte est renvoyé à l'étape en cours, ou bien à l'étape précédente

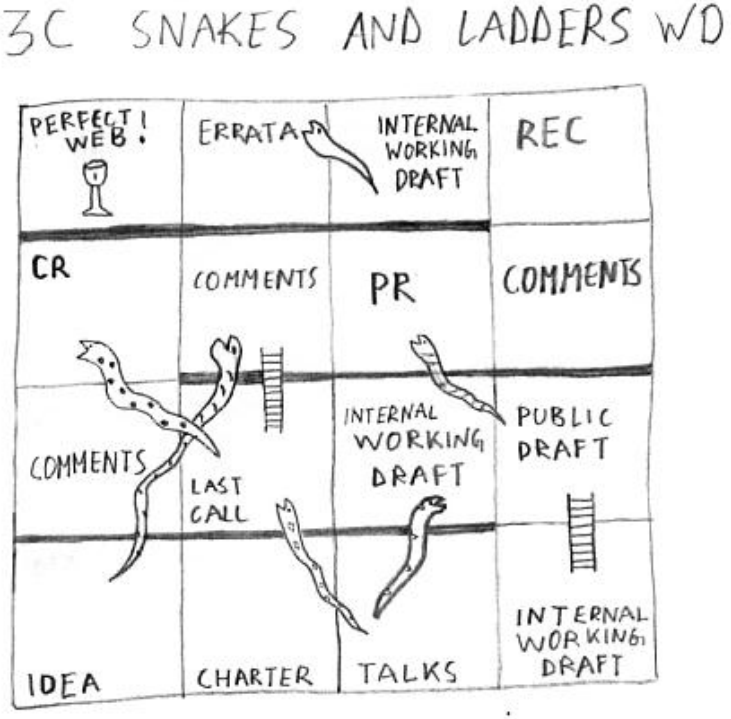
3.2 La fabrique des normes au W3C

- Une négociation ouverte à tous les membres et des conversations publiques
  - Chaque groupe a sa propre charte, comportant la description du sujet, une durée envisagée (variant entre six mois et deux ans), les objectifs et les règles de la discussion.
  - C'est l'annonce d'un « *Call for participation* » qui fait office de création d'un groupe de travail.
  - Les discussions ont ensuite lieu par liste mail, vidéoconférence, téléphone (*distributed meetings*) et lors de colloques et de séminaires (*face-to-face meetings*).

	2012	2013	2014	2015	2016
HTML5.0	CR start	...CR, LC	Rec	...	...
HTML5.1	FPWD	---	LC + CR	...CR	Rec

3.2 La fabrique des normes au W3C

- Une méthode par consensus (et non par vote)
- Des recommandations qui n'ont rien d'obligatoire (et qui ont donc des concurrents)

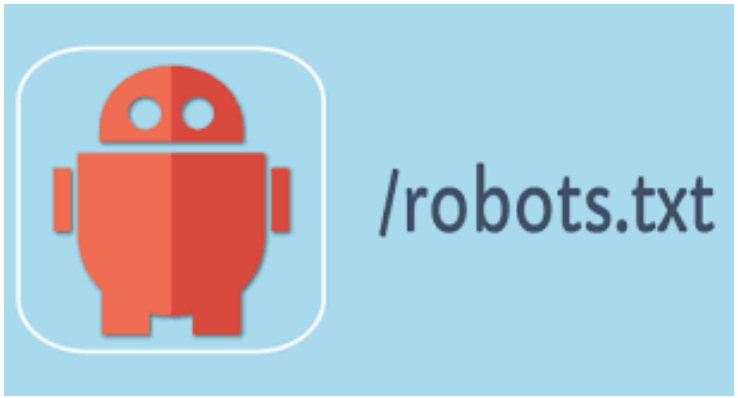


3.2 La fabrique des normes au W3C

- **Le pouvoir du directeur (Tim Berners-Lee)**
  - C'est lui, ou celui à qui il en a délégué le pouvoir, qui décide du franchissement de chaque étape
  - « dictateur à perpétuité » (Malcolm, 2008, p. 57).
  - Il peut s'opposer au Comité consultatif
  - La règle des 5% (jamais invoquée… jusqu'en 2017)

3.3 Les stratégies de contournement

- **1. Contourner les deux arènes : exemple du robots.txt et d'ACAP (2009)**



3.3 Les stratégies de contournement

- **1. Contourner les deux arènes : exemple d'ACAP**

  - Robots.txt :

Exemple n° 1 : l'éditeur autorise tous les *crawlers* à scanner toutes les pages
User-agent: *
Disallow:

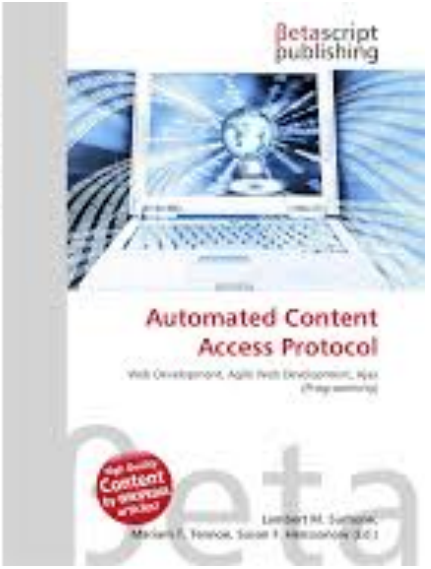
Exemple n° 2 : l'éditeur interdit tous les *crawlers* de scanner toutes les pages
User-agent: *
Disallow: /

Exemple n° 3 : l'éditeur exclut les pages du répertoire « /monde » du champ d'indexation de Google
User-agent: Googlebot
Disallow: /monde

- On utilise en plus des meta-tags pour les « nofollow », « noindex »…

3.3 Les stratégies de contournement

- **1. Contourner les deux arènes : exemple d'ACAP**



3.3 Les stratégies de contournement

- **1. Contourner les deux arènes : exemple d'ACAP**

  - ACAP :

Exemple 1

```
ACAP-crawler: *
ACAP-disallow-crawl: /
ACAP-allow-crawl: /$
ACAP-allow-crawl: /public/
ACAP-disallow-follow: /public/comments/
```

Exemple 2

```
ACAP-crawler: *
ACAP-disallow-crawl: /
ACAP-allow-crawl: /$
ACAP-allow-crawl: /public/
ACAP-allow-index: /public/
ACAP-disallow-other: /public/
```

3.3 Les stratégies de contournement

- 2. SI le Whatwg décide de n'en faire qu'à sa tête
  - 2005-2008
  - 2012-…

3.4 La controverse encrypted media extensions (EME)

- En février 2012 Google, Microsoft et Netflix proposèrent, via la *liste de diffusion* du W3C, d'intégrer au standard HTML5 une API (*Application Programming Interface*) nommée « *Encrypted Media Extensions* » (EME).
  - Réaction immédiate de Hicksie sur la mailing list : « Je crois que cette proposition n'est pas éthique et que la discussion à son sujet devrait s'arrêt là.»

3.4 La controverse encrypted media extensions (EME)

- Dans le but d'empêcher le W3C d'intégrer officiellement l'EME à la charte du HTML5 (pour des raisons que nous détaillerons ci-après), l'Electronic Frontier Foundation (EFF) et la Free Software Foundation (FSF) lancèrent une pétition avec pour objectif d'atteindre les 50 000 signataires avant le 3 mai 2013.

- Une lettre ouverte fut également adressée le 24 avril 2013 par 27 organisations à Tim Berners-Lee, pour « supplier le comité du *World Wide Web* ainsi que ses organisations participantes de rejeter la proposition EME ».
  - Malgré cette lettre et les 27 500 signatures reçues par la pétition, le 9 mai 2013, Tim Berners-Lee accepta de publier l'EME sous forme de FPWD et de l'inscrire sur la charte du HTML5.

3.4 La controverse encrypted media extensions (EME)

- Selon l'EFF :
  - « C'est exactement l'opposé de la raison pour laquelle le World Wide Web Consortium existe à l'origine »
  - Ces propositions *« sont une tentative d'apaiser Hollywood qui est mécontenté par Internet depuis presque aussi longtemps qu'existe le Web et qui a toujours demandé à ce qu'on lui fournisse des infrastructures techniques évoluées pour contrôler le fonctionnement des ordinateurs de son public »*,

3.4 La controverse encrypted media extensions (EME)

- Début 2014, la *Motion Picture Association of America* (MPAA), qui jusque là avait été soupçonnée de négocier avec Google, Netflix et Microsoft pour qu'ils proposent eux-mêmes et défendent l'EME au sein du W3C, annonça qu'elle rejoignait le W3C pour discuter du standard HTML5, et intégrer en particulier le groupe de travail où il était question de l'EME.
  - La fondation Mozilla est opposée aux DRM et à tout ce qui contribuait à contrôler ou à restreindre la circulation de l'information. Quand, finalement, en mai 2014, la fondation opta « pour le réalisme, quitte à égratigner ses principes »

« *La décision de poursuivre le projet de standardisation des DRM pour le web est venue de Tim Berners-Lee lui-même, qui semble avoir cru au mensonge selon lequel les producteurs d'Hollywood allaient abandonner le web et s'intéresser à d'autres médias (AOL ?) dans le cas où ils n'obtiendraient pas que l'Internet ouvert soit re-programmé de manière à correspondre à leurs projets de maximisation des profits.* »

Cory Doctorow (EFF) — octobre 2013

3.4 La controverse encrypted media extensions (EME)

- L'EME est passé du stade WD au stade CR, avec l'accord de Tim Berners-Lee le 5 juillet 2016.
- L'EFF se met alors à envisager le recours à la règle des 5%, jamais utilisée dans l'histoire du W3C, au cas où Tim Berners-Lee déciderait de stabiliser le format

3.4 La controverse encrypted media extensions (EME)

- La question de l'observabilité et du DMCA (section 1201)

« Nous ne pouvons pas accepter que la recherche concernant la sécurité de navigation sur le web soit pénalisée. Les navigateurs sont continuellement attaqués. Tous ceux qui utilisent le web utilisent un navigateur et deviendront vulnérables si l'on arrête la recherche en matière de sécurité de navigation. Dès lors que l'EME introduit un DRM dans les navigateurs, cela introduit un risque. »

Hickson (Google et Whatwg), Pieters (Opera) et van Kesteren (Mozilla) — 21 septembre 2016[13]

« Je suis en discussion avec plusieurs autres acteurs pour essayer de sauver la mise en place d'un « Open Covenant *» dans lequel les éditeurs utilisant EME s'engageraient à ne pas poursuivre les chercheurs en sécurité qui regardent sous le capot. »*

Robin Berjon

3.4 La controverse encrypted media extensions (EME)

- En janvier 2016, l'EFF propose que soit mis en place d'un accord formel impliquant que les ayants droit des vidéos protégées par l'EME s'engageraient à ne pas poursuivre en justice les personnes qui auraient procédé à des tests dans la mesure où celles-ci auraient ensuite publié les anomalies identifiées
  - Malgré la signature de 198 chercheurs en informatique favorables à cette proposition et le soutien de l'Open Source Initiative (OSI), la proposition de l'EFF a été refusée par les parties prenantes de la discussion concernant l'EME au W3C.

3.4 La controverse encrypted media extensions (EME)

- En mars 2017, l'Unesco a rejoint officiellement les rangs des opposants à l'EME

- Frank La Rue, sous-directeur général pour la communication et l'information, adresse une lettre publique à Tim Berners-Lee destinée à lui faire savoir
  - qu'une des valeurs fondamentales de l'Unesco était « la libre circulation des idées et de l'information »
  - et à le prévenir que l'EME « pourrait avoir un impact sur les navigateurs au point de rendre impossible l'exercice des utilisateurs de leur droit légal d'une utilisation équitable des vidéos sous copyright ».

3.4 La controverse encrypted media extensions (EME)

- En juillet 2017, on annonça que l'EME était sur le point de devenir une REC.
- L'EFF, par l'intermédiaire de son représentant au W3C Cory Doctorow, réunit les 5% des signatures
  - Sur les 463 membres du W3C, 108 ont voté oui, 57 non, 20 ont voté blanc, et le reste n'a pas participé
  - L'EME gagne avec 58,4% des voix et passe au stade REC le 18 septembre 2017

3.4 La controverse encrypted media extensions (EME)

- Un EME pour les formats textuels ?
  - En février 2017 le W3C et l'International Digital Publishing Forum (IDPF), au sein duquel a été développé la norme EPUB, ont fusionné et qu'il sera possible à partir de 2018 pour les membres du W3C de participer à un groupe appelé « Publishing business »

3.5 Les « faire-faire » du web

- La gouvernementalité
  - Conduire les conduites
  - Borner les libertés
  - Orienter les décisions
  - Inciter à certaines actions
  - Diminuer / augmenter les probabilités de
  - Contraindre, négocier…
  - « Soft power »
  - Organiser des « indicateurs statistiques », une « surveillance douce »…

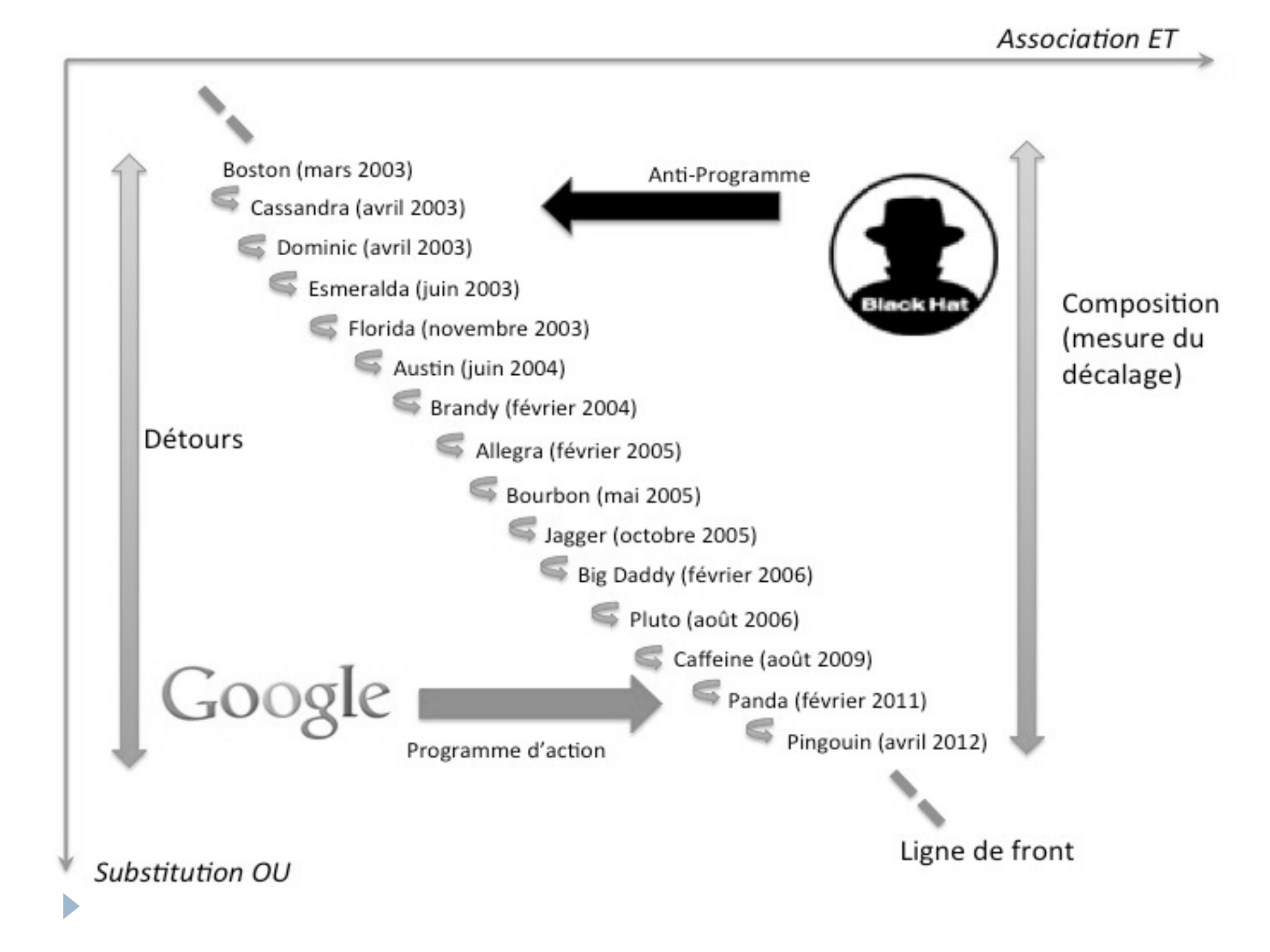
Certains acteurs ont le pouvoir « de disposer des choses, c'est-à-dire d'utiliser plutôt des tactiques que des lois, ou d'utiliser au maximum des lois comme des tactiques ; faire en sorte, par un certain nombre de moyens, que telle ou telle fin puisse être atteinte » (Foucault, 2004, p. 103).

Les incitations : l'exemple du SEO

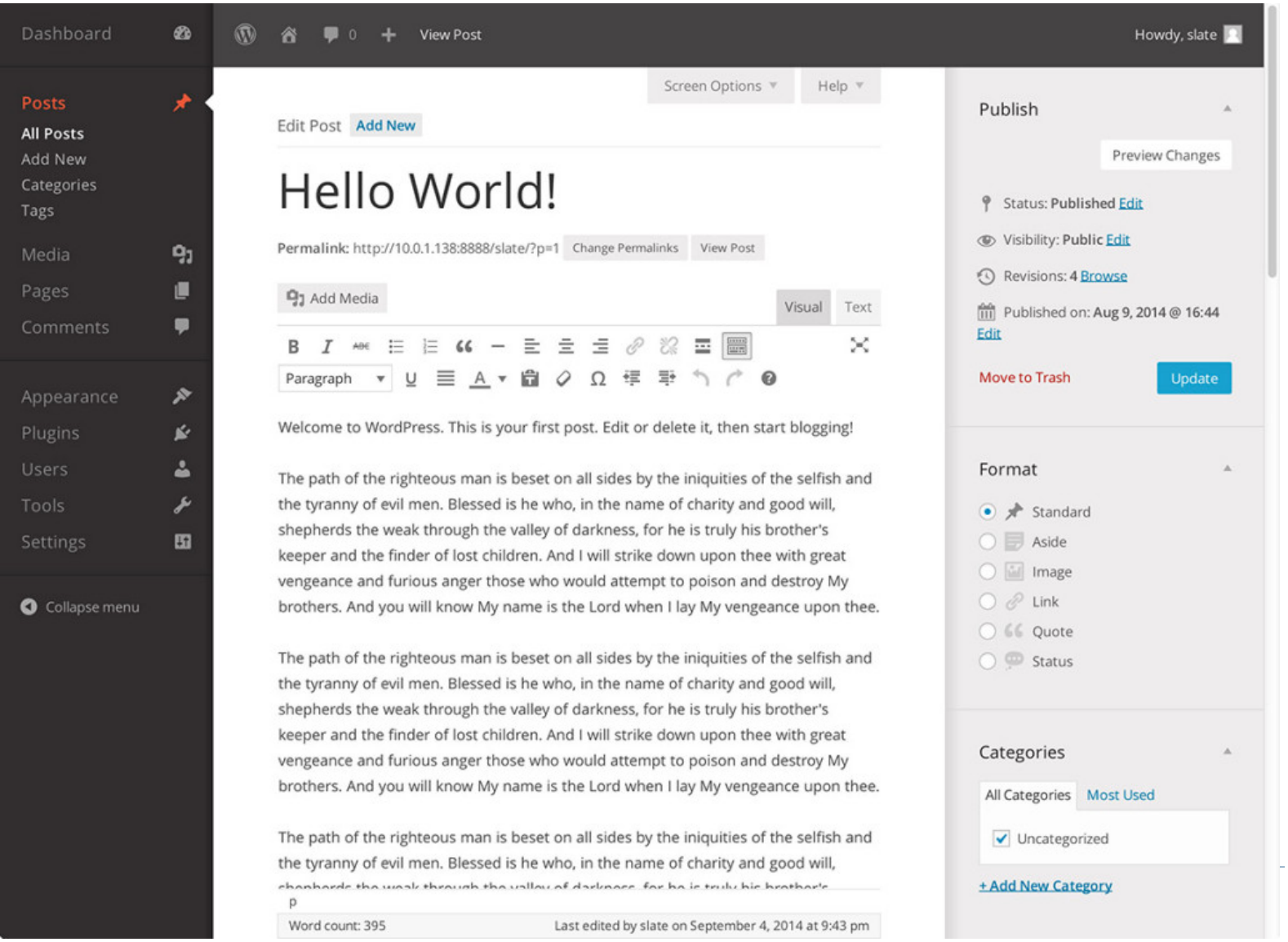
- Exemple de gouvernementalité par incitation : Google et les techniques de Search Engine Optimization

- « Les webmasters sont encouragés à adapter leur contenu d'une manière avantageuse pour Google. De plus, les webmasters sont invités à indiquer les sites qui ne respectent pas les règles définies par Google » (Röhle, 2009).

- Mais les éditeurs peuvent regimber : black hat SEO
- Google modifie son algorithme : négociations perpétuelles


Association ET
Boston (mars 2003)
Cassandra (avril 2003)
Dominic (avril 2003)
Esmeralda (juin 2003)
Florida (novembre 2003)
Austin (juin 2004)
Brandy (février 2004)
Allegra (février 2005)
Bourbon (mai 2005)
Jagger (octobre 2005)
Big Daddy (février 2006)
Pluto (août 2006)
Caffeine (août 2009)
Panda (février 2011)
Pingouin (avril 2012)
Anti-Programme
Black Hat
Composition (mesure du décalage)
Détours
Google
Programme d'action
Ligne de front
Substitution OU

Les contraintes : l'exemple du CMS



Les contraintes : l'exemple du CMS

« Au-delà de la simple question de la facilité de la publication, ce sont les possibilités de contrôle sur les écrits à venir à partir du CMS qui sont concernées : selon que les créateurs d'un CMS envisagent ou projettent des enjeux différents dans leur propre projet de communication, ils vont distribuer cette liste de rôles selon des logiques sociales diverses, positionnant dès l'amont de la conception les sites selon des degrés de médiations » (Jeanne-Perrier, 2005, p. 77).

Le pouvoir de validation : l'exemple d'Apple

Pour développer une application pour iPhone :

- Souscrire au *Mac Developper Program* (99$ par an)
- Respecter les exigences de l'*iOS Developper Library*
- Se soumettre au processus d'évaluation
- Corriger si besoin est (courber l'échine !)
- Accepter qu'Apple ponctionne 30% des revenus

HOW LONG IS THE LINE?
REVIEW FACILITY
ONLY THE APPLE KNOWS...
GAME MARKETEERS
WWW.GAMEMARKETEERS.COM

4. Controverses et perspectives

4.1 Le pouvoir des Etats-Unis

- L'influence sur les instituts de normalisation
- Le contrôle des infrastructures type DNS
- Le Patriot Act et les injonctions

4.2 Le pouvoir des navigateurs et des logiciels

- En matière de norme et de gouvernance, le pouvoir est dans les mains de ceux qui exécutent

4.3 La neutralité du réseau

4.3 La neutralité du réseau

- Les 4 atteintes à la neutralité selon Tim Whu (2003) :
  - 1) Blocking : Bloquer du contenu
  - 2) Termination Monopoly Pricing : les FAI peuvent faire payer ceux qui veulent accéder à leurs usagers
  - 3) "Playing Favorites" ou MFN (Most Favored Network) violations : favoriser le passage d'une information, d'un service, plutôt qu'une autre
  - 4) Transparency Failures : pour l'opérateur, ne pas donner à ceux qui le souhaitent les informations techniques à propos des capacités du réseau (ou bien mentir)

4.3 La neutralité du réseau

Les 4 principes essentiels de Benjamin Bayart (*French Data Network*) :

- Transmission des données par les opérateurs sans en examiner le contenu ;
- Transmission des données sans prise en compte de la source ou de la destination des données ;
- Transmission des données sans privilégier un protocole de communication ;
- Transmission des données sans en altérer le contenu.

4.3 La neutralité du réseau

- L'héritage des pionniers

« Que les auteurs du réseau [Internet] aient eu conscience ou non de ce qui naîtrait de leur création, ils l'ont bâtie en fonction d'une certaine philosophie : en un mot, l'idée selon laquelle le réseau lui-même ne serait pas en mesure de réguler son mode de croissance. Ce sont les applications qui le feraient. Tel était l'enjeu d'une structure end-to-end. »

(Lawrence Lessig, *L'Avenir des idées, 2005)*

La Neutralité : comment techniquement est-elle menacée ?

- Le protocole IP permet de signaler différents niveaux de priorité des données qui composent le trafic Internet

- Les FAI ont mis en place des équipements leur permettant de trier les types de données. C'est ce qu'on nomme « l'examen approfondi du paquet » (deep packet inspection).

La Neutralité : qui la menace ?

- Les FAI : ils ne veulent pas que leurs réseaux explosent

- Les producteurs de contenus : ils ne veulent pas qu'on les pille ou les détourne

La Neutralité : que veulent faire ceux qui la menacent ?

- Contenir l'accroissement de la demande en bande passante en ralentissant certains usages intensifs (vidéo à la demande)

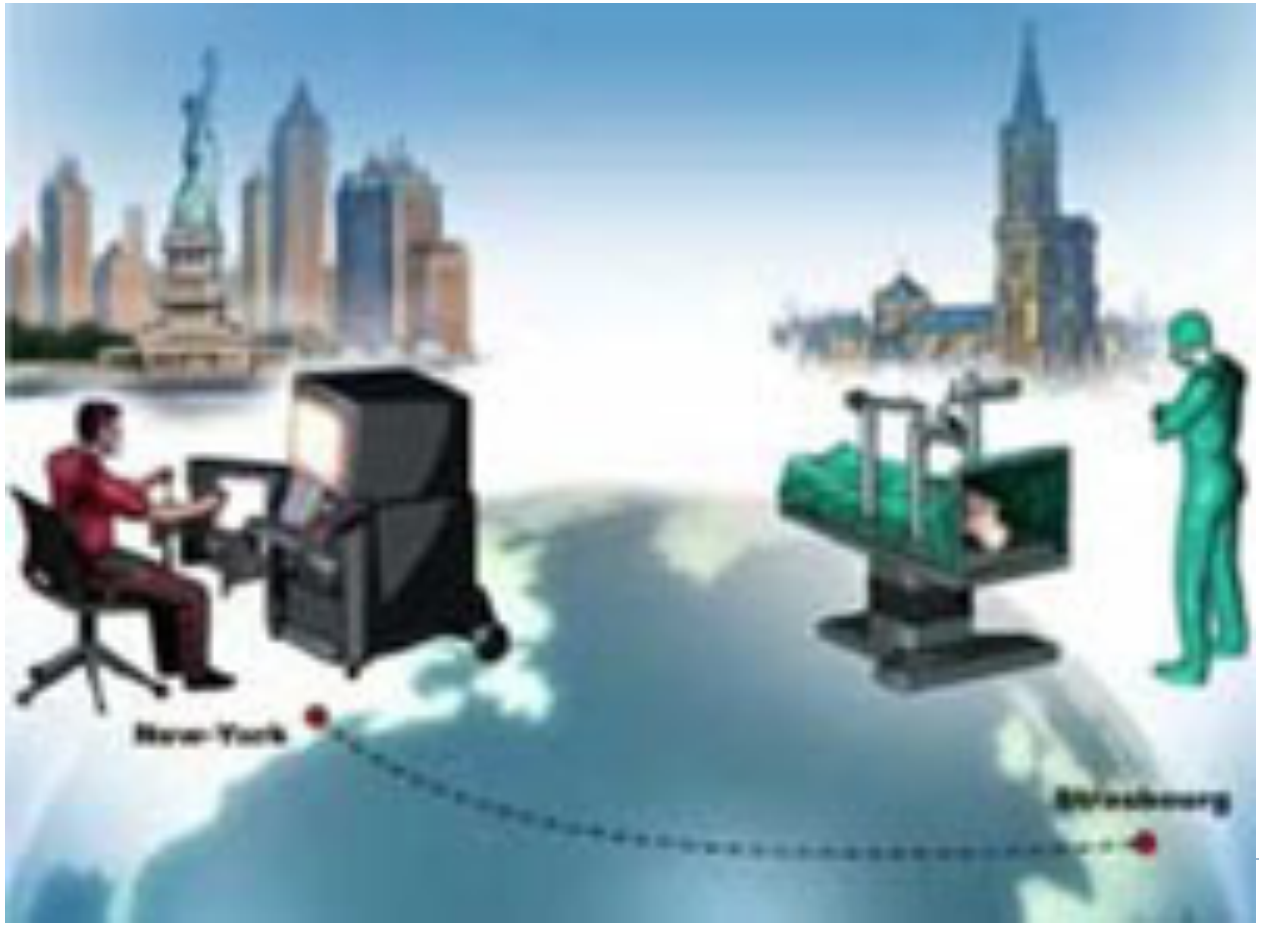
- Développer de nouveaux modèles économiques donnant naissance à un internet à plusieurs vitesses, en donnant une priorité de transport à certains services et applications contre rémunération. (Premium)

Neutralité

- Mais que fait la police
  - La violence, la pornographie illicite…
  - Le FBI, Interpol… doivent-ils surveiller ?
  - Au USA : le Patriot Act (oct. 2001) permet de violer la neutralité au nom de la lutte contre le terrorisme

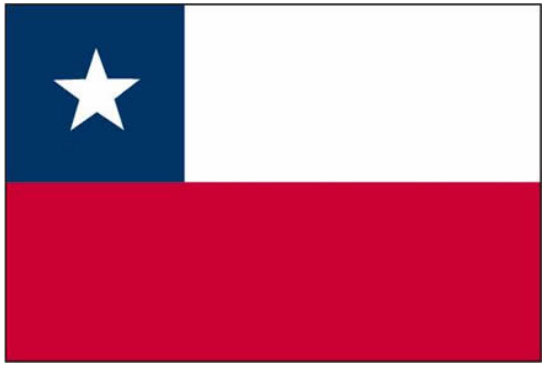
La Neutralité : les arguments contre

- L'exemple de la chirurgie à distance… et de la vidéo de chat !



Neutralité : les pays qui ont tranché le débat…

- La Neutralité protégée au Chili, au Pays-Bas et en Slovénie : interdiction de discriminer



Neutralité : France

- En France : le CNNum pour l'inscrire dans la Loi
- L'avis du CNNum sur la neutralité du Net est clair : il faut que le gouvernement inscrive le principe de neutralité des réseaux dans le marbre.
- Le CNNum recommande de modifier la loi « Léotard » de 1986 sur la liberté de communication, qui serait renommée pour l'occasion « loi relative à la liberté d'expression et de communication »

4.4 Extension du domaine d'Internet

- Internet des objets : Quelles normes pour connecter les objets et leur permettre de communiquer ?

4.5 L'avenir des protocoles du web par rapport aux autres

- Le web est-il mortel ?

(2010)

4.6 L'apprentissage du code

4.6 L'apprentissage du code

- *Le code : la langue maternelle du numérique…*

- *Ne pas connaître le code, c'est être analphabète*

4.6 L'apprentissage du code

- *« Oui »*, répondent 87 % des parents interrogés par BVA pour Syntec Numérique dans une consultation rendue publique mercredi 21 mai 2014.
- Le 14 mai 2014, un rapport d'information sur le « développement de l'économie numérique », signé par les députées Corinne Erhel (PS) et Laure de La Raudière (UMP), insistait sur la nécessité d'*« éveiller les élèves dès l'école primaire » et de créer « un enseignement dans le secondaire ».*

4.6 L'apprentissage du code

ès des lycées • Dates des vacances scolaires 2015-2016 • Orthographe révisée • Admission Post bac • Breve

"Un milliard d'euros sur trois ans" pour le numérique à l'école

Actualité / Education / Par LEXPRESS.fr avec AFP , publié le 07/05/2015 à 18:33 , mis à jour à 21:18

583 partages | Partager | Tweeter | Partager | 21 réactions

Axelle Lemaire
@axellelemaire

Suivre

"Au primaire tous les enfants seront sensibilisés à la programmation et à la culture numérique dès la rentrée 2016" #EcoleNumerique @Elysee

18:08 - 7 Mai 2015

201 97

4.6 L'apprentissage du code

L'école se prépare à enseigner le code informatique

A la rentrée, le code figurera dans les programmes de primaire et collège. L'éducation nationale poursuit sa mue numérique, même si l'informatique n'a pas encore trouvé sa place dans l'éventail des disciplines scolaires.

LE MONDE | 06.06.2016 à 06h41 • Mis à jour le 06.06.2016 à 10h19 |

Par **Aurélie Collas**

L'initiation commencera dès l'âge de 6 ans. Un peu en CP, CE1 et CE2, et surtout à partir du CM1. Les élèves apprendront à *« programmer les déplacements d'un robot ou d'un personnage sur écran »* ou à *« construire une figure simple »*. Au collège, le code devient un thème des programmes de mathématiques et de technologie. Le but est d'être capable d'*« écrire, mettre au point et exécuter un programme simple »*.

Au brevet 2017, l'épreuve de mathématiques et sciences comportera obligatoirement au moins un exercice d'algorithmique ou de programmation. *« L'idée n'est pas de former des spécialistes, mais d'apporter aux élèves des clefs de décryptage du monde numérique, de les amener à voir l'informatique autrement que comme une pensée magique à laquelle on n'aurait pas accès »,* explique Florence Robine, directrice générale de l'enseignement scolaire au ministère de l'éducation nationale.

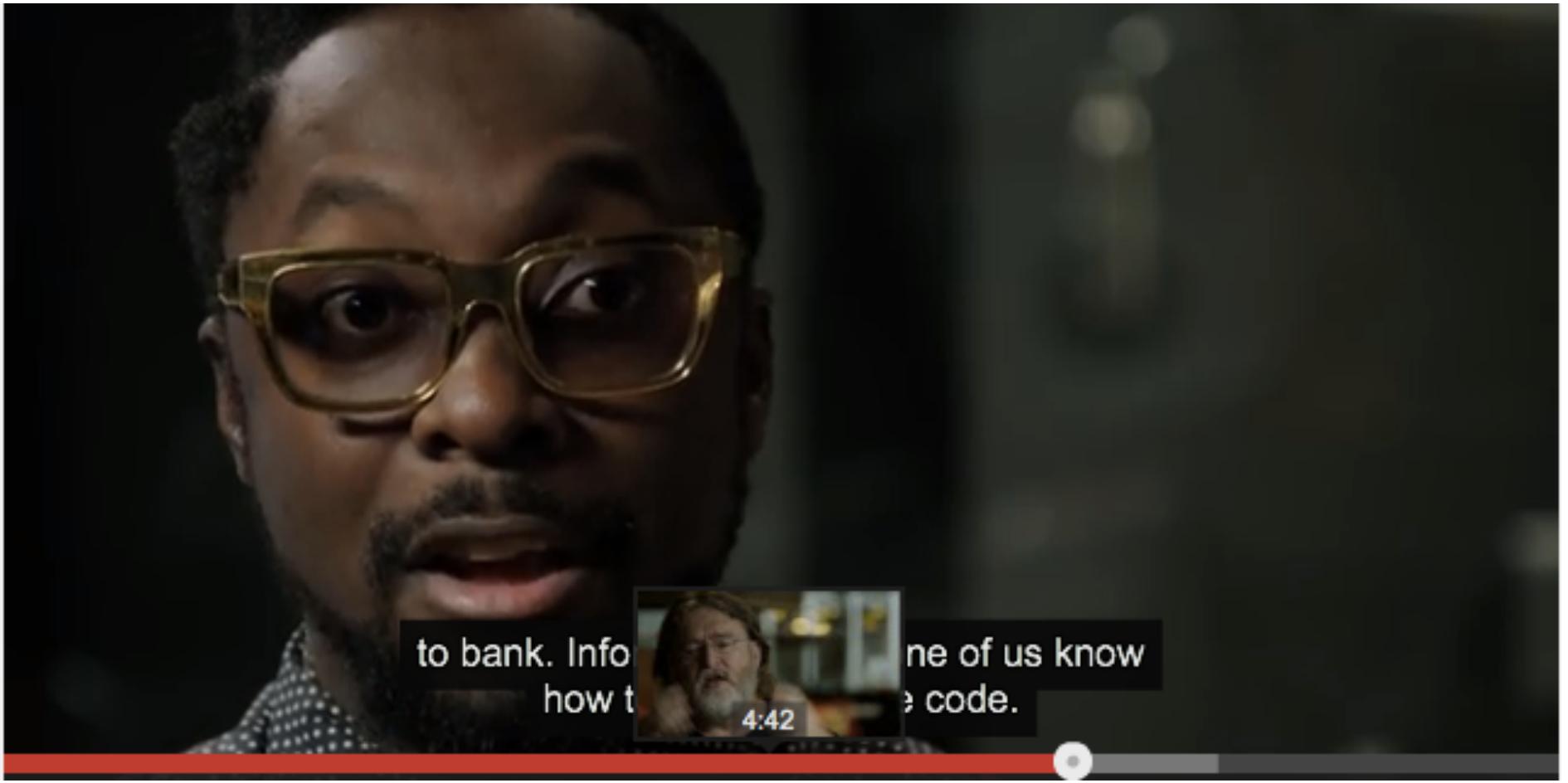
4.6 L'apprentissage du code

- *Apprendre soi-même…*

codecademy

Apprenez à coder de manière interactive, gratuitement.

4.6 L’apprentissage du code



Questions
Discussion

- Mail : guillaumesire@gmail.com
- Twitter : @guillaumesire

Université Fédérale Toulouse Midi-Pyrénées