

# Notions d'Internet

Daniel POULIN

DRT-3808 (version 2008)

Faculté de droit, Université de Montréal

# Sommaire

- Les notions de base doivent être maîtrisée : protocoles, types de protocole, systèmes analogiques et numériques, les types d'adresses, les différences et relation entre elles.
- Sans devoir tout mémoriser, il faut avoir une compréhension générales des protocoles TCP, IP et DNS

# Notions

- Qu'est-ce qu'Internet? (GDT)
  - Le réseau informatique mondial constitué d'un ensemble de réseaux nationaux, régionaux et privés, qui sont reliés par le protocole de communication TCP-IP. »
  - En d'autres termes, **Internet est le réseau informatique résultant de la combinaison des réseaux capables de communiquer au moyen des protocoles TCP-IP** (nous y reviendrons)

# Internet est plus qu'un réseau

- Métaphoriquement, c'est un lieu
- 1 milliards d'Internautes
- Au Canada (StatCan, 2003)
  - 64 % des ménages, soit 7,9 M des ménages ont accès à la haute vitesse
- C'est aussi de plus en plus le lieu où trouver l'information de tout genre ainsi que tout type d'expression culturelle du monde entier (du moins du monde technologiquement et économiquement développé)

# Volumes d'information

1 page de texte	3 kilo octets
1 monographie (un livre de 300 pages)	1 Mo
Une pièce musicale en MP3	1-3 Mo
Une pièce musicale sur cédérom	30 Mo
Un cédérom	700 Mo
1 000 monographies ou 1,5 cédérom	1 Go
1 000 000 monographies 1 500 cédéroms, ou 1 000 Go	1 Téraoctet

# Internet en chiffres

Le Web de surface	167 T
Offert sur KaZaA en 2003	5 000 T ou 10 M de cédéroms
Le Web profond (tout ce qui est accessible par le Web dans les bases de données accessibles)	91 850 T
Le courriel	440 606 T
La messagerie instantanée	274 T

# Internet suite...le courriel

•Selon HMI 2003 :	•1/3 de la population en dispose aux ÉU.
•Temps consacré	•Forrester 35
•Nombre quotidien de messages	•60% (10 ou moins) •23% (plus de 20)
•Nombre quotidien	•60 milliards de messages (est. 2006)
•Pourriel ou spam	•200 milliards/an (Forrester, 2004) •AOL et Yahoo en intercepteraient chacun 1 milliards par jour, soit 28 par abonné.

# Systemes analogiques et numériques

- Un systeme *numerique* est un systeme qui utilise les nombres, souvent binaires, afin d'acquérir, de traiter, de transmettre, de stocker ou d'afficher des informations (ou donnees)
- Un systeme analogique fait usage d'un spectre continu de valeurs. (Wikipedia/numerique)
- Avantages du numerique: verification, compression et chiffrement
- Avec le numerique, la convergence devient possible

# Autres distinctions

- Signal v. message
- Réseau physique v. réseau logique

# Les réseaux téléphoniques ("circuit switching")

- Conçu il y a plus de cent ans, le réseau téléphonique (Public Switched Telephone Network, PSTN) offre une communication de point à point entre deux téléphones.
- En principe, **un circuit s'établit entre les deux postes téléphoniques pour la durée de la communication** et les infrastructures utilisées ne peuvent être utilisées à d'autres fins. La « bande passante » est réservée.

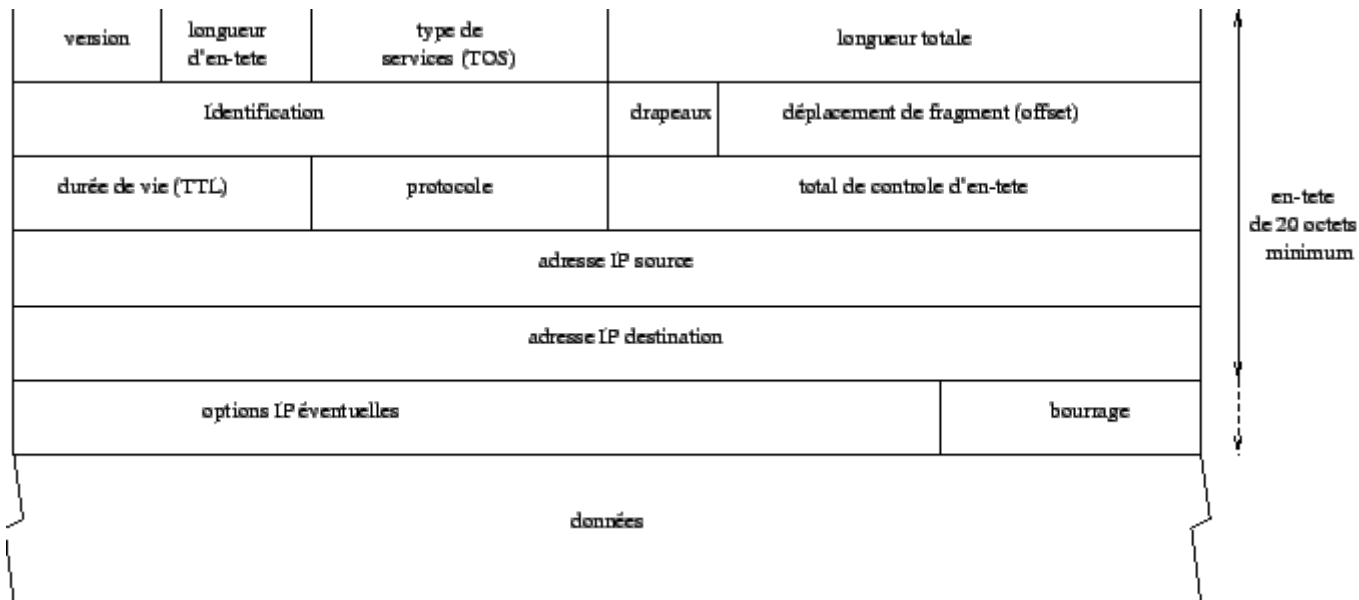
# Internet et réseaux antérieurs

- Réseaux informatiques antérieurs
  - Moins polyvalents, confinés à un type particulier d'ordinateur : AppleTalk, Novell
  - Réseaux X.25 : Datapac, Bloomberg, CompuServe
- Internet, un réseau logique
  - Les protocoles TCP/IP font abstraction du support physique et offrent des services réseau à des applications non définies
  - Le réseau est hétérogène

# Technologies de bases d'Internet

- Commutation par paquet (datagrammes)
- Paquet d'information
  - Méta-informations : adresse de la source, de la destination, TTL, taille
  - Contenu: information ou charge utile, généralement 1500 caractères ou moins
- Ils se meuvent de façon indépendante par des routeurs, pas de route pré-établie

# Datagramme IP



Source : Pascal NICOLAS Université d'Angers, <http://www.info.univ-angers.fr/pub/pn/poly/node39.html>

# Les protocoles

- La notion de protocole est familière : téléphone, adressage
- Un protocole: un ensemble de règles décrivant en termes techniques comment une action doit être exécutée
- Dans un réseau, les protocoles ont généralement pour objet de définir les modalités de connexion et d'échange des données
- Exemple: Le protocole décrit dans RFC2822 précise l'adresse d'un courriel
- Il s'agit donc d'une forme de normativité technique

# Protocoles propriétaires et protocoles ouverts

- Protocole privé ou propriétaire
- Protocole ouvert
  - Disponible, documentation publiée
  - Gratuité (pas de brevet)
  - Non discriminatoire, parfois pourvu de licences contre la prédation
  - Favorise le choix
- Les protocoles de la suite TCP-IP sont des protocoles ouverts
- Ils peuvent donc être implémentés sans qu'il soit nécessaire de demander la permission et d'acquitter des frais de licence

# Sources et types des protocoles

- Source
  - Les forums publics : XHTML (de W3C)
  - L'industrie
    - SSL, les cookies (de Netscape)
    - PDF (de Adobe)
- Type
  - De facto : QWERTY
  - De jure : IPv6

# Les architectures client-serveur

- Une approche selon laquelle le service est centralisé sur un serveur et est rendu accessible par un logiciel-client opérant sur un autre ordinateur
- Le client et le serveur communiquent par un protocole déterminé
- Exemple: un logiciel serveur Apache sert des pages web via le protocole HTTP
- Ces pages sont accessibles par le biais d'un logiciel client compatible, comme Internet Explorer
- Contre-exemple
  - Le Peer2Peer : Napster, KazAa, Bittorent

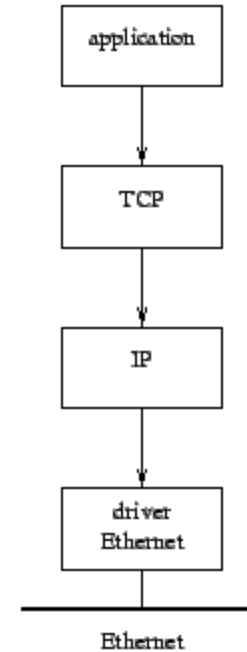
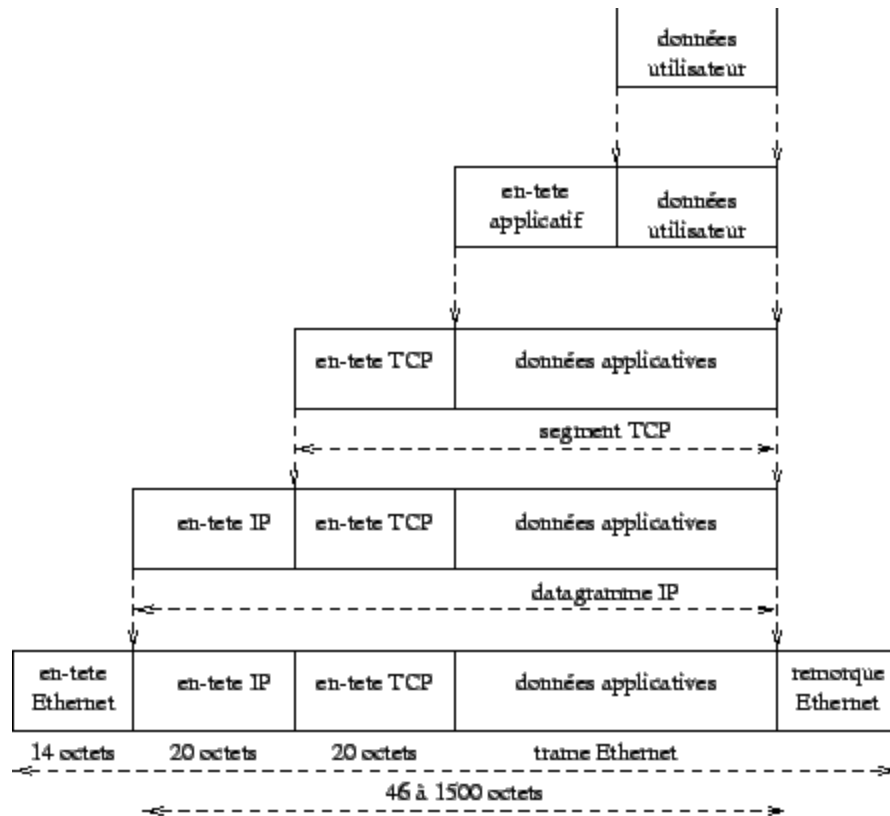
# Le protocole IP (Internet Protocol)

- Le protocole IP est relatif à l'adressage, au morcellement des paquets et à leur acheminement d'une adresse IP à une autre.
- Circulation des paquets
  - Sans connexion, chaque paquet circule de façon indépendante
  - De la source, le paquet est acheminé vers une passerelle qui l'achemine à premier routeur qui le réacheminera à un autre jusqu'à destination
  - Les routeurs ont des tables leur indiquant vers lequel des routeurs auxquels ils sont attachés acheminer un paquet
- Le protocole IP ne suffit pas

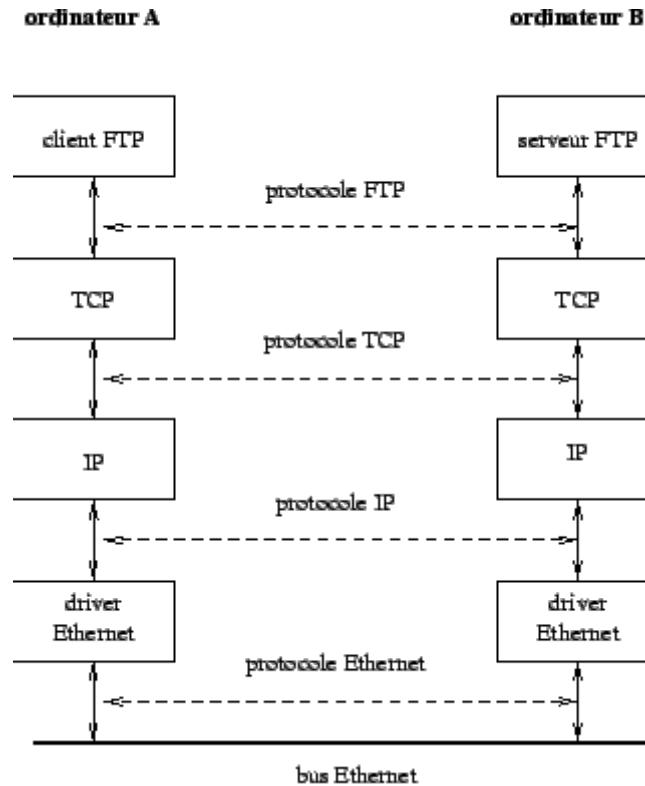
# Le protocole TCP (Transmission Control Protocol)

- Le protocole TCP assure un acheminement ordonné, de la source à la destination, de l'ensemble des paquets d'un message
- Travail de TCP
  - Construit des paquets (source, destination, information de contrôle et de séquence et le contenu lui-même)
  - Établit une connexion et la gère
  - Vérifie que tous les paquets sont reçus et au besoin les réachemine
  - Fournit l'information au logiciel qui l'a demandée
- UDP est un autre protocole de transport, mais sans connexion et sans vérification d'erreur

# L'encapsulation des données



# Les piles des protocoles



# Autres protocoles TCP-IP

- HTTP
- FTP
- Telnet, SSH
- POP, IMAP

# Illustrations

- Traceroute/DNSLookup/
  - [www.austlii.org](http://www.austlii.org)
  - [www.canlii.org](http://www.canlii.org)
  - [www.juriburkina.bf](http://www.juriburkina.bf)
- Utiliser <http://www.iptools.com> ou <http://xwhois.com>

# Les adresses d'Internet

- Adresses IP (IPv4, IPv6, NAT)
- Noms de domaine
- URL
- Adresses de courriel

# L'adresse IP (traditionnelle)

- Numérique, unique, nécessaire pour utiliser Internet
- Formée de quatre octets, des nombres binaires de huit bits, séparés par des points
  - Exemple : 132.204.34.18

# Nombre binaire, bits et octets

Nombres décimaux	Nombres binaires
0	0
1	1
2	10
3	11
4	100
5	101
6	110
7	111
8	1000
9	1001

# L'adresse IP

- Un nombre binaire de 32 bits représenté comme 4 octets
- La valeur de chaque élément ne peut dépasser 255, i.e. 11111111
- L'adresse comporte deux éléments, un désigne le réseau, l'autre un hôte (ordinateur) particulier

# Les classes de réseaux IP

Cl	de	à	Pour le réseau	# adr. de réseau	# adr. pour les hôtes
A	0.1.0.0	126.0.0.0	1er octet	128	16 777 214
B	128.0.0.0	191.255.0.0	Le deux 1er octets	16 384	65 000
C	192.0.1.0	223.255.255.0	Trois 1er octets	2,1 M	253

**L'Université de Montréal possède un réseau de classe B: 132.204.0.0**

# Les adresses IP et leurs propriétés

- Les nombres formant l'adresse sont en quelque sorte hiérarchisés (132.204.0.0)
- L'administration des sous-réseaux et des plages d'adresses est décentralisée
- Traditionnellement, les adresses étaient stables
- L'adressage dynamique
  - Au bureau
  - À la maison

# L'adressage IP, les Firewalls et NAT (Network Address Translation)

- Au départ, 1 ordinateur = 1 adresse
- Cela a changé
  - Les adresses manquent
  - Des ordinateurs, les Firewalls (coupe-feu), se sont insérés entre le réseau interne et Internet
- Le firewall convertit l'adresse interne (ex. : 192.168.0.0) en adresse externe par (ex. : 132.204.132.24:3434)
- Enjeux

# Les noms de domaine

- Caractéristiques
  - Hiérarchiques
  - Correspondent plus ou moins à une adresse IP
  - Jamais utilisés par les ordinateurs
- Deux sens ou acceptions
  - Domaine totalement qualifié : `www.ibm.com`
  - Domaine de second niveau : `ibm.com`
- Avantages
  - Mnémonique, configurable

# Les Domain Name Servers (DNS)

- L'utilisation de noms de domaine est possible grâce aux Domain Name Servers (DNS)
- Pour utiliser Internet, tous les ordinateurs doivent disposer d'un accès à un DNS
- Le problème et sa solution
  - Des milliards d'adresses, des millions d'utilisateurs faisant des centaines de requêtes chaque jour
  - Une base de données distribuée
  - Organisée de façon hiérarchique
  - L'utilisation de caches

# Les Domain Name Servers (DNS)

- Si on cherche « www.ibm.com »
- Si le DNS de l'UdM ne l'a pas dans sa cache
  - Il s'adresse à un des serveurs « root »
    - De qui il obtient le serveur pour « .com »
      - De qui il obtient le serveur pour « ibm.com »
        - » De qui il obtient l'adresse IP de « www.ibm.com »
  - Il fournit l'adresse et l'inscrit dans sa cache pour l'avoir si un autre ordinateur le lui demande dans les prochains jours

# Les noms de domaine de tête (Top level domain, TLD)

- Structure géographique (code de pays ISO à deux lettres)
  - ca, fr, ch, be, de, us, bb et ainsi de suite
- Structure générique
  - Traditionnellement la plus recherchée
  - « .com », « .org », « .net »

# Les nouveaux noms de domaine de tête

.aero	Air-transport industry	Société Internationale de Télécommunications Aéronautiques (SITA).
.biz	Businesses	NeuLevel, Inc.
<b>.com</b>		VeriSign Global Registry Services.
.coop	Cooperative associations	Dot Cooperation LLC.
.info		Afilias Limited.
.jobs	Human resource managers	Employ Media LLC.
.museum	Museums	Museum Domain Management Association.
.name	Individuals	Global Name Registry.
<b>.net</b>		VeriSign Global Registry Services.
<b>.org</b>	Noncommercial community	Public Interest Registry
.pro	Credentialed professionals	RegistryPro.
.travel	Travel industry	Tralliance Corporation.
.int	Organizations established by international treaties	IANA .int Domain Registry
<b>.gov</b>	United States Government	US General Services Administration.
<b>.edu</b>	United States postsecondary institutions	Educause.
<b>.mil</b>	United States Military	DoD Network Information Center.

# La gestion des noms de domaine

- Les noms de domaines sont gérés hiérarchiquement
- Les serveurs DNS « root » sont sous le contrôle de l'Internet Corporation for Assigned Names and Numbers (ICANN)
- Le domaine « .ca » est sous le contrôle de l'Autorité canadienne pour les enregistrements Internet (ACEI)
- Le domaine de second niveau « umontreal.ca » est sous le contrôle de l'UdeM
- Le sous-domaine « lexum.umontreal.ca » est sous le contrôle du LexUM

# Observation sur les adresses et les noms

- Traditionnellement
  - 1 ordinateur = 1 adresse = 1 nom
- Maintenant
  - 1 adresse = plusieurs ordinateurs
  - 1 ordinateur = plusieurs adresses
  - 1 adresse = plusieurs noms
  - 1 nom = plusieurs adresses

# Enjeux des adresses IP et des noms de domaines

- Adresses IP
  - Répartition inégale
  - Décentralisation de leur attribution
  - Adresses dynamiques ou internes : NAT et FSI
- Noms de domaine
  - Gestion des TLD génériques « .com »
  - Gestion des ressources centrales « roots »
  - Décentralisation
  - Ils offrent un point de prise pour la régulation des États

# Adresse de courrier électronique

- [daniel@poste.lexum.umontreal.ca](mailto:daniel@poste.lexum.umontreal.ca)
- [daniel@lexum.umontreal.ca](mailto:daniel@lexum.umontreal.ca)
- [daniel.poulin@umontreal.ca](mailto:daniel.poulin@umontreal.ca)
- Daniel Poulin <daniel@lexum.ca>
- Madame la juge Tremblay  
<didi@aol.com>

# Les « Uniform Resource Locators » (URL)

- Exemple:
  - <http://www.canlil.org:80/index.html#top>
  - <http://132.204.136.40>
- Éléments
  - Le protocole (« HTTP »)
  - Le nom de domaine (ou l'adresse IP 132.204.136.40)
  - Le numéro de port (« :80 »)
  - Le nom de fichier (« /index.html »)
  - L'ancre (« #top »)
- Exemple : « Domain info » à <http://www.iptools.com> avec CanLII

# Conséquences juridiques

- En principe, sur Internet
  - L'absence de centre : pas de contrôle
  - Tout ordinateur peut communiquer avec tout autre
  - Les chemins empruntés peuvent varier
- Conséquences
  - Ce qui convient quelque part peut offenser ailleurs, mais être néanmoins accessible
  - Des choses imprévues surviennent : Napster, Skype
- En pratique, de plus en plus, des nuances sont nécessaires