Introduction aux systèmes répartis

Module 1

INF8480 Systèmes répartis et infonuagique

Michel Dagenais

École Polytechnique de Montréal
Département de génie informatique et génie logiciel

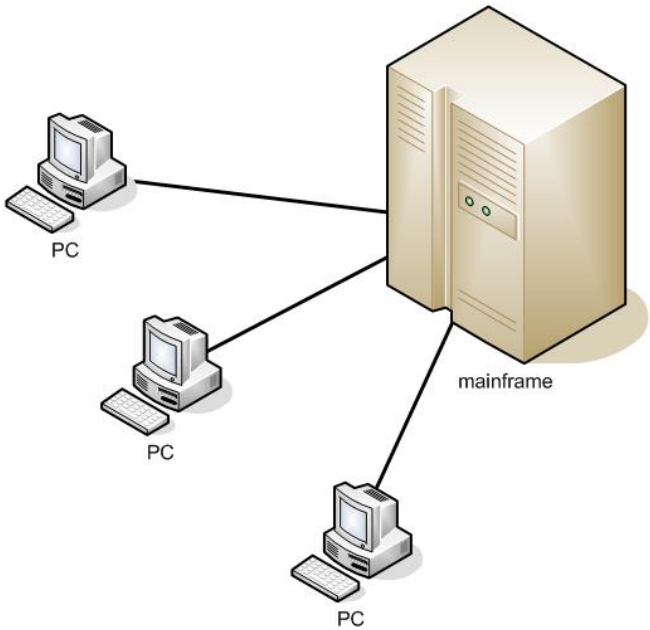
Sommaire

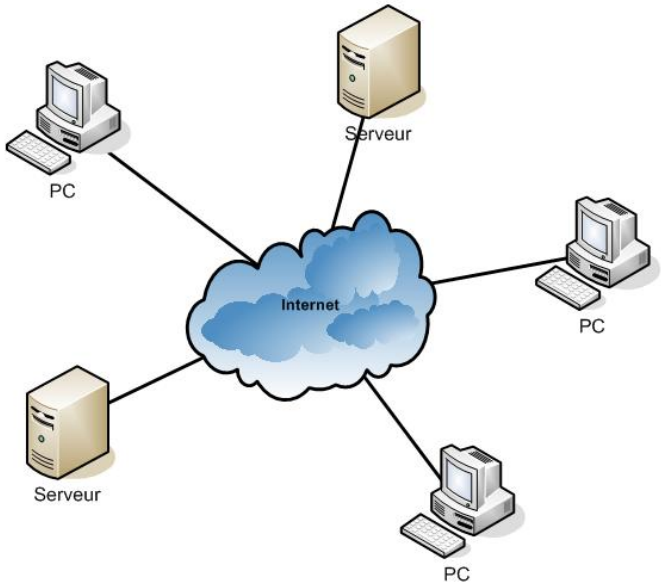
Introduction aux systèmes répartis

Évolution des organisations informatiques

- Serveur centralisé avec terminaux
  - Serveur coûteux;
  - Engorgement au serveur.
- Ordinateurs personnels
  - Faible coût d'achat;
  - Grand choix d'applications;
  - Autonomie mais manque de service et de coordination.
- Systèmes répartis
  - Le réseau partout et en continu;
  - Matériel et logiciels modulaires à faible coût;
  - Environnement hétérogène mais protocoles normalisés;
  - Redécoupage des responsabilités client et serveur;
  - Systèmes de plus en plus complexes.

Système centralisé versus réparti

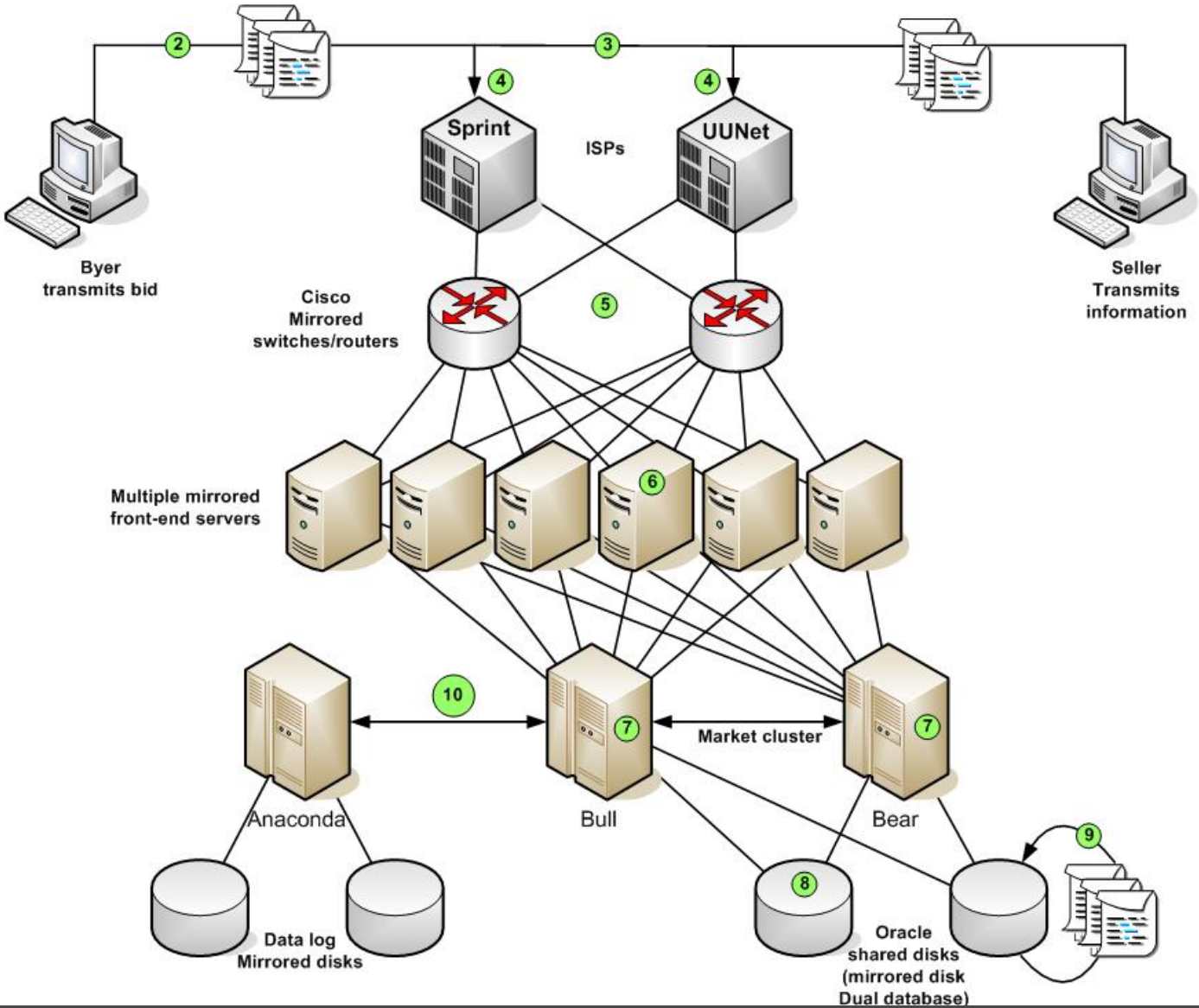




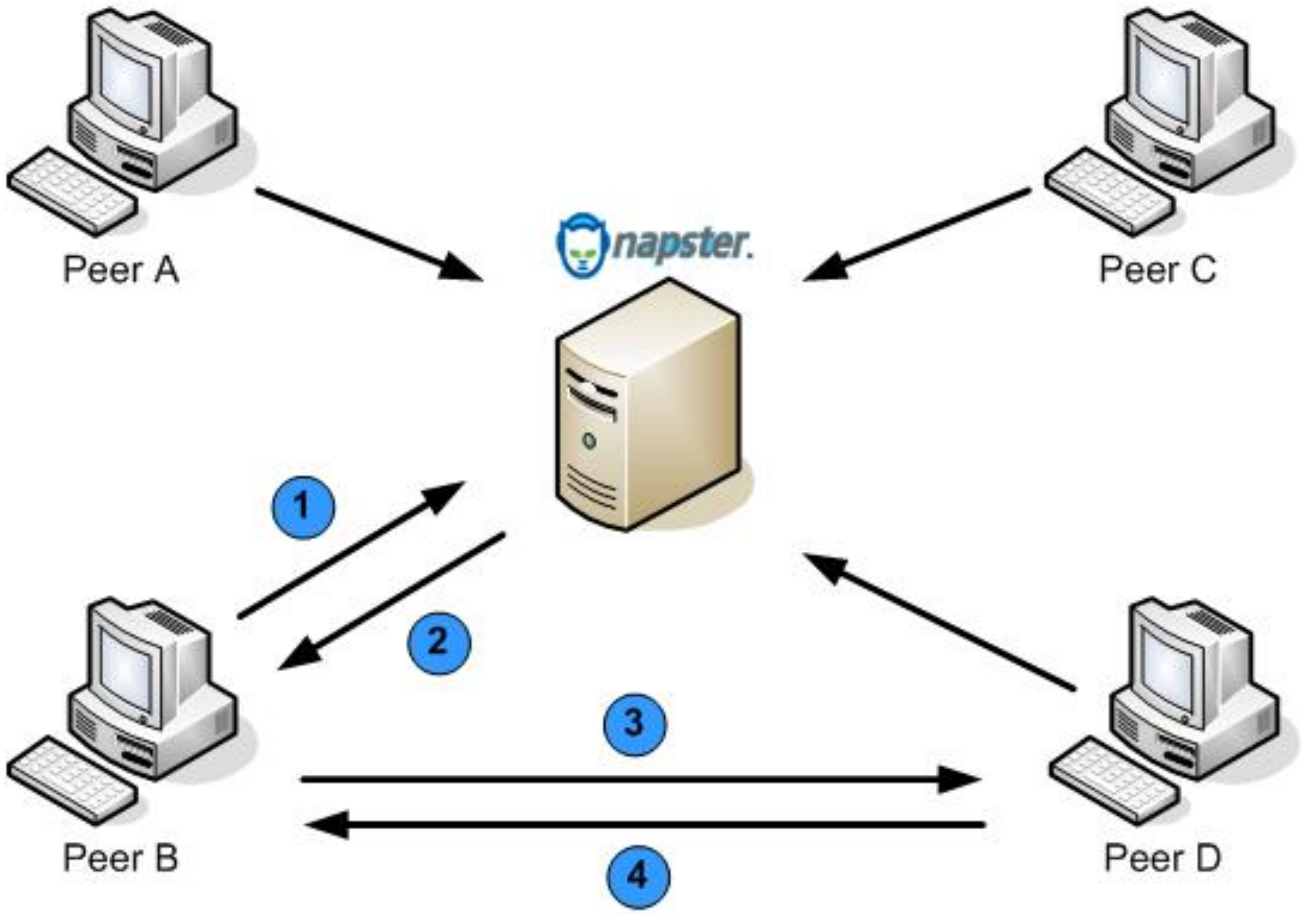
Système réparti

- Système dont les composantes sont réparties sur plusieurs ordinateurs en réseau et qui communiquent entre eux et coordonnent leurs actions uniquement par transmission de messages.
- Un ensemble d'ordinateurs indépendants qui, du point de vue de l'usager, apparaissent comme un système unique et cohérent.
- Une définition alternative par Leslie Lamport (1987):
  - "You know you have one when the crash of a computer you've never heard of stops you from getting any work done."

Architecture répartie de eBay

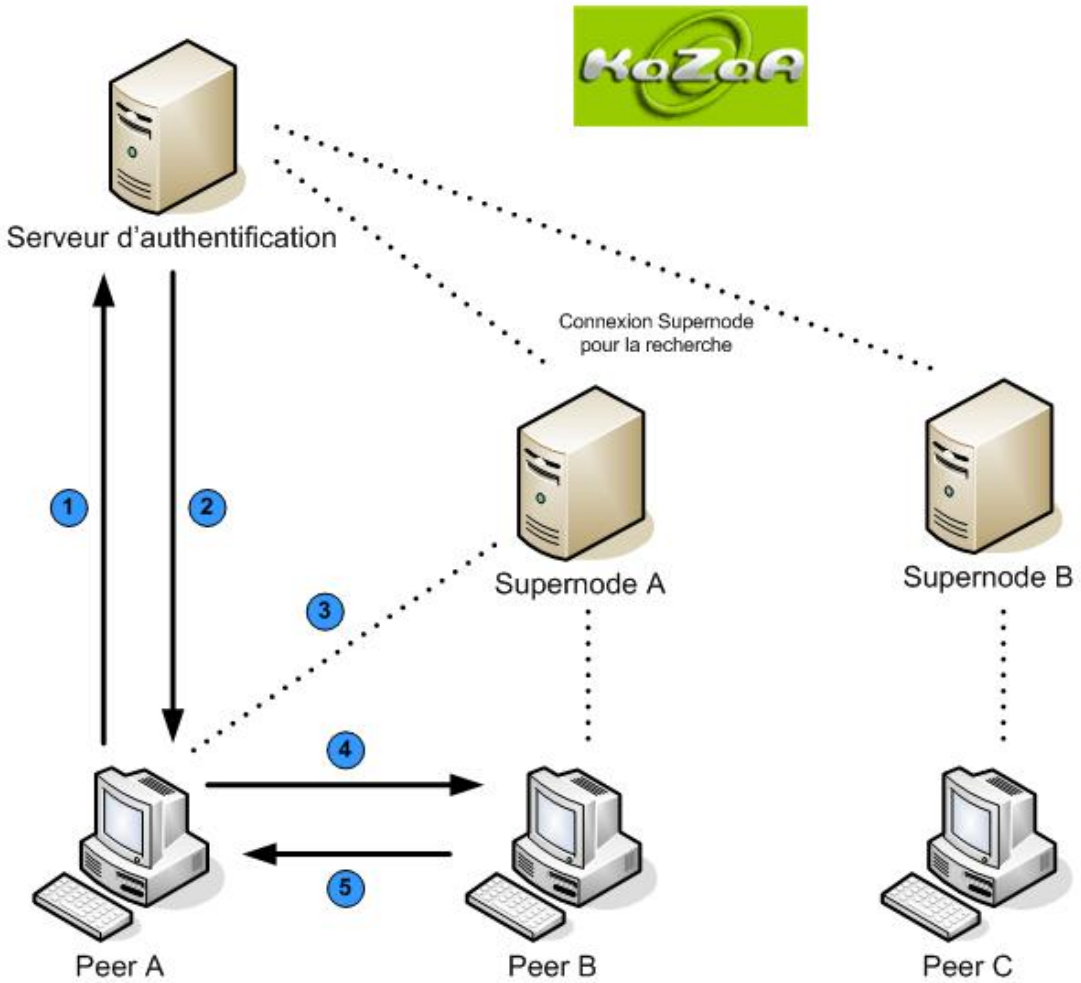


Réseau P2P centralisé



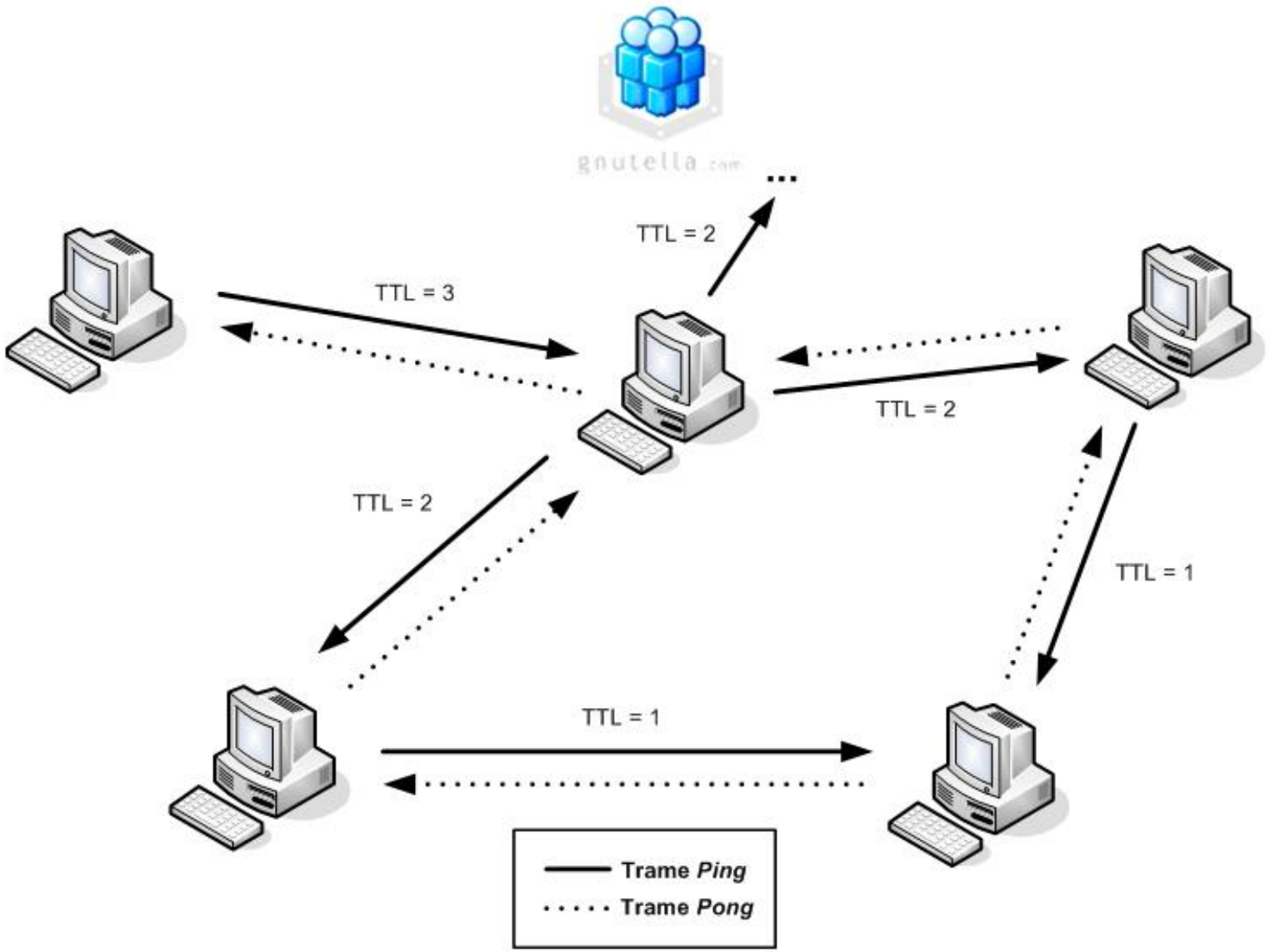
1. Recherche un fichier sur le serveur central
2. Serveur renvoit la liste des *peer* qui possèdent le fichier
3. *Peer B* contacte *Peer D* pour obtenir le fichier
4. Transfert du fichier

Réseau P2P hybride

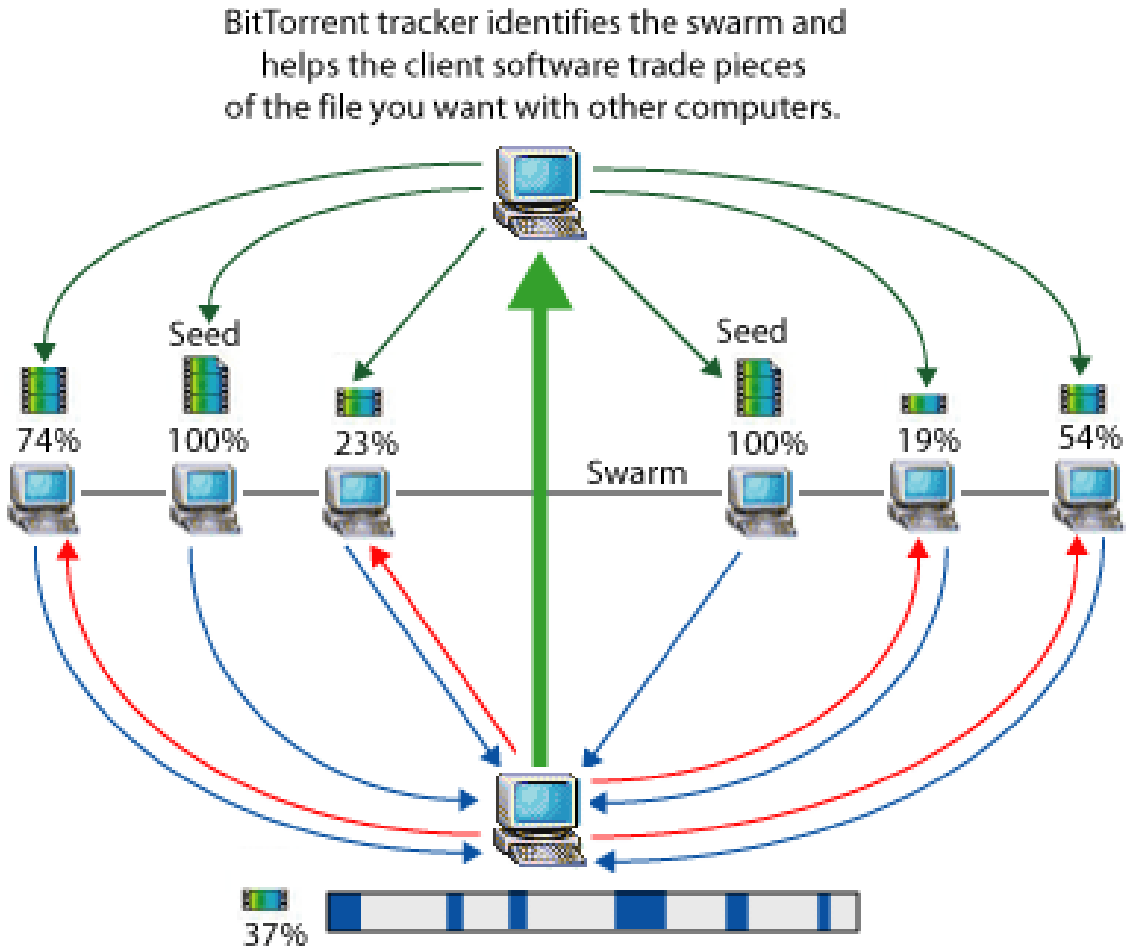


1. Authentification auprès du serveur
2. Serveur renvoie le *Supernode* le plus près
3. Effectue la recherche du fichier sur le *Supernode*
4. Contacte le peer possédant le fichier
5. Transfère le fichier

Réseau P2P décentralisé



Réseau P2P autre



Autres exemples

- L'Internet et tous ses protocoles, DNS, SMTP, BGP;
- L'intranet d'une entreprise;
- Le système Interac;
- Google, Facebook...

Pourquoi les systèmes répartis

- Partage des ressources (données, périphériques...);
- Accès à des ressources distantes;
- Augmentation modulaire de la capacité du système;
- Possibilité de tolérance aux pannes.

Inconvénients des systèmes répartis

- Plusieurs points de défaillance;
- Sécurité;
- Difficulté pour le système d'avoir un état global;
- Complexité accrue.

Caractéristiques des systèmes répartis

- Les composantes du système:
  - Sont réparties matériellement et/ou géographiquement;
  - Sont autonomes;
  - Sont concurrentes;
  - Peuvent défaillir indépendamment;
  - Possèdent des horloges asynchrones;
  - Communiquent par envoi de message sur le réseau.

Conséquences

- Nombreux points de défaillance possibles;
- Décalage de temps entre les horloges de chaque système;
- Pas d'état global;
- Pas de garantie que les messages sont reçus;
- Messages peuvent être interceptés, modifiés, ajoutés.

Introduction aux systèmes répartis

Historique

- Xerox DFS, Xerox PARC, Xerox Alto, Ethernet, 1977;
- Cambridge Distributed Computing Services (DCS), M68000, Cambridge ring, 1979;
- Système d'exploitation Locus de UCLA, VAX, Ethernet, 1980;
- Apollo Domain, Token Ring, 1980;
- Grapevine, Xerox PARC, Xerox Alto, Ethernet, 1981 (replicated distributed application-oriented database service);
- Cedar, Xerox PARC, Xerox Dorado, Ethernet, 1982, development environment for office and personnal systems;
- Amoeba, Vrije University, VAX/M68000..., Ethernet, 1984, distributed system based on capabilities;
- Unix BSD 4.2 + SUN RPC/NFS, Vax/SUN, Ethernet, 1985;
- Mach, CMU, VAX/SUN, Ethernet, 1986, système d'exploitation réparti basé sur un micro-noyau.

Historique (suite)

- World Wide Web, HTTP sur TCP/IP, 1992;
- Groupe OMG, CORBA (Common Object Request Broker Architecture) 1992;
- Langage Java, RMI (Remote Method Invocation), 1995;
- Google, 1998;
- VMWare, 1998;
- Langage C#, Remoting, 2001;
- Facebook, 2004;
- Amazon EC2, 2006;
- iPhone, Android, 2007, 2008;
- OpenStack, 2010;
- Docker, 2013;
- Kubernetes et la Cloud Native Computing Foundation, 2015;

Introduction aux systèmes répartis

Les principaux problèmes à résoudre

- Répartition de l'application;
- Hétérogénéité des équipements et technologies, besoin d'interopérabilité;
- Ouverture de système;
- Sécurité;
- Évolutivité et mise à l'échelle;
- Tolérance aux fautes et la fiabilité/ Détection et isolation des fautes/défaillances;
- Concurrence, Synchronisation et Interblocage;
- Transparence;
- Validation et tests;

Répartition de l'application

- Partitionnement de l'appplication en différents composants;
- Equilibrer la charge de l'application à travers différents composants répartis (client, noeuds de la grappe), statiquement ou dynamiquement;
- Architecture simple, propice à l'évolutivité et au maintien de la sécurité;

Hétérogénéité

- Réseaux et protocoles utilisés;
- Matériel;
- Systèmes d'exploitation;
- Langages de programmation;
- Implémentations;
- Représentations internes.

Solutions

- Protocoles et formats de stockage normalisés;
- Intergiciels d'adaptation (e.g. gRPC, CORBA, Java RMI, .NET).

Systèmes ouverts

- Possibilité d'évoluer, de re-développer le système en tout ou en partie;
- Interopérabilité avec des systèmes complémentaires;
- Portabilité vers du nouveau matériel;
- Services développés selon des règles normalisées, formalisées à l'intérieur de protocoles, formats de stockage et interfaces de programmation.

Evolution vers les systèmes ouverts

- Système unique, homogène;
- Développement interne, en plein contrôle;
- Applications commerciales prêtes à utiliser, plus performantes, moins chères;
- Fournisseur unique, perte de contrôle sur le prix et le cycle de mise à jour;
- Systèmes ouverts, interface de programmation (CORBA), protocoles (IIOP) et formats de stockage normalisés, code source ouvert, implémentation de référence libre (Orbit);
- Mélange de logiciels internes, logiciels libres et logiciels commerciaux

Sécurité

- Transmettre des informations sensibles sur un lien de communication non sécuritaire et non fiable de manière sécuritaire;
- Confidentialité, intégrité, disponibilité.

Évolutivité et mise à l'échelle

- Taille du système (nombre d'utilisateurs, de requêtes, de ressources);
- Etendue géographique (avec les latences associées);
- Structure administrative (décentralisée, sécuritaire);
- Architecture du logiciel réparti, séparer les politiques des mécanismes;

Tolérance aux fautes et fiabilité :

- Les fautes et les défaillances sont plus courantes que dans les systèmes centralisés;
- Les défaillances sont habituellement indépendantes;
- Détection des fautes/défaillances;
- Masquage ou tolérance des fautes/défaillances;
- Redondance et réplication;

Concurrence

- Permettre au système de traiter simultanément plusieurs requêtes à une même ressource;
- Les opérations doivent être sérialisées ou donner un résultat cohérent équivalent.

Transparence

- Masquer à l'utilisateur tous les aspects reliés à la répartition du système;
- Accès, localisation, concurrence, réplication, défaillance, mobilité, performance, évolutivité.

Validation et tests

- Comment tester le système complet? Chaque composante?
- Les fautes lors des tests pourraient être masquées par la tolérance aux pannes?
- Validation formelle de certaines portions.
- SPIN, modelchecker développé par Bell Labs, http://spinroot.com/spin/whatispin.html.
- UPPAAL est un modelchecker développé par l'Université Uppsala, en Suède et l'Université d'Aalborg en Danemark, http://www.uppaal.com/. Standard opensource pour la fiabilité et interopérabilité:
- Service Availability Forum (SAF) et Availability Management Framework (AMF), http://www.saforum.org/.

Introduction aux systèmes répartis

Modèles de systèmes

- Client-serveur (multiples, imbriqués, micro-service...)
- Client-proxy-serveur.
- Collègues (peer to peer).
- Client + code mobile - serveur.
- Agents mobiles.
- Ordinateur réseau ou client minimal (X, VNC, Citrix).
- Réseaux spontanés (découverte de ressources DHCP, réseaux infra-rouge, bluetooth).

Modèles de pannes

- Auto-détection;
- Omission;
- Mauvaise réponse plus ou moins aléatoire;
- Erreur byzantine;
- Erreur de synchronisme.

Exemple: World Wide Web

- Format HTML, XML, CSS, XSLT, XSL-FO, Javascript;
- Réseau TCP/IP;
- Convention pour les adresses (URL), et protocole pour les requêtes (HTTP);
- Client qui exécute un fureteur: envoi de requêtes par HTTP, affichage du résultat en XML/CSS, exécution d'applet;
- Serveur: sert des requêtes HTTP à partir de fichiers HTML, de fichiers de script (CGI, PHP, jsp), ou de modules spéciaux (XML, XSLT);
- Cette plate-forme est utilisée pour accéder de l'information, rechercher des documents, consulter des annuaires, faire du courriel, interagir avec des groupes de discussion...

Introduction aux systèmes répartis

Réseaux logiques

- Réseau logique bâti par-dessus un réseau physique (overlay network);
- Ethernet VLAN;
- Réseau défini par logiciel, par exemple avec OpenFlow;
- Réseau de machines virtuelles en infonuagique;

Architectures de réseau

- PSTN
- Internet
- Multiprotocol Layer Switching (MPLS)
- Cellulaires (GSM, GPRS, EDGE, UMTS: 3G, LTE et WiMax Mobile :4G)
- Wi-Fi
- WiMAX
- Bluetooth
- Réseaux ad hoc

Circuits commutés

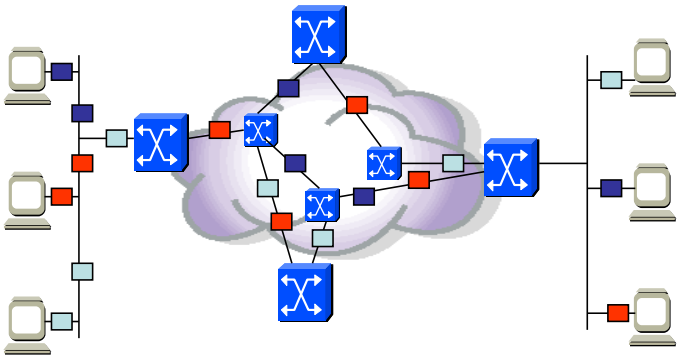
Internet

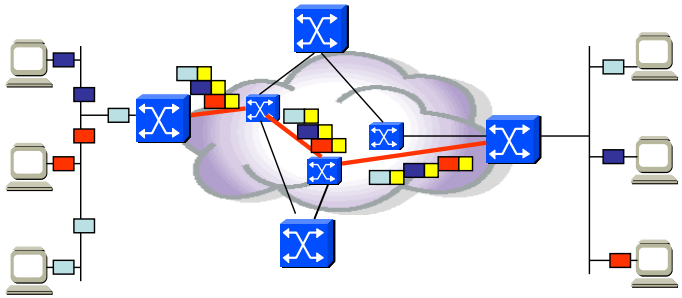
- Réseau mondial basé sur l'envoi de paquets;
- Familles de protocoles IP, UDP, TCP, FTP, SMTP, HTTP...
- Les aspects techniques et architecturaux sont régis par l'Internet Engineering Task Force (IETF);
- Les protocoles sont documentés dans les Request For Comments (RFC);
- Réseaux locaux avec routage statique et réseau global avec routage dynamique;
- Initialement, rien n'était prévu pour assurer la qualité de service, par exemple afin de transmettre la voix ou le vidéo en temps réel.

Multiprotocol Label Switching (MPLS)

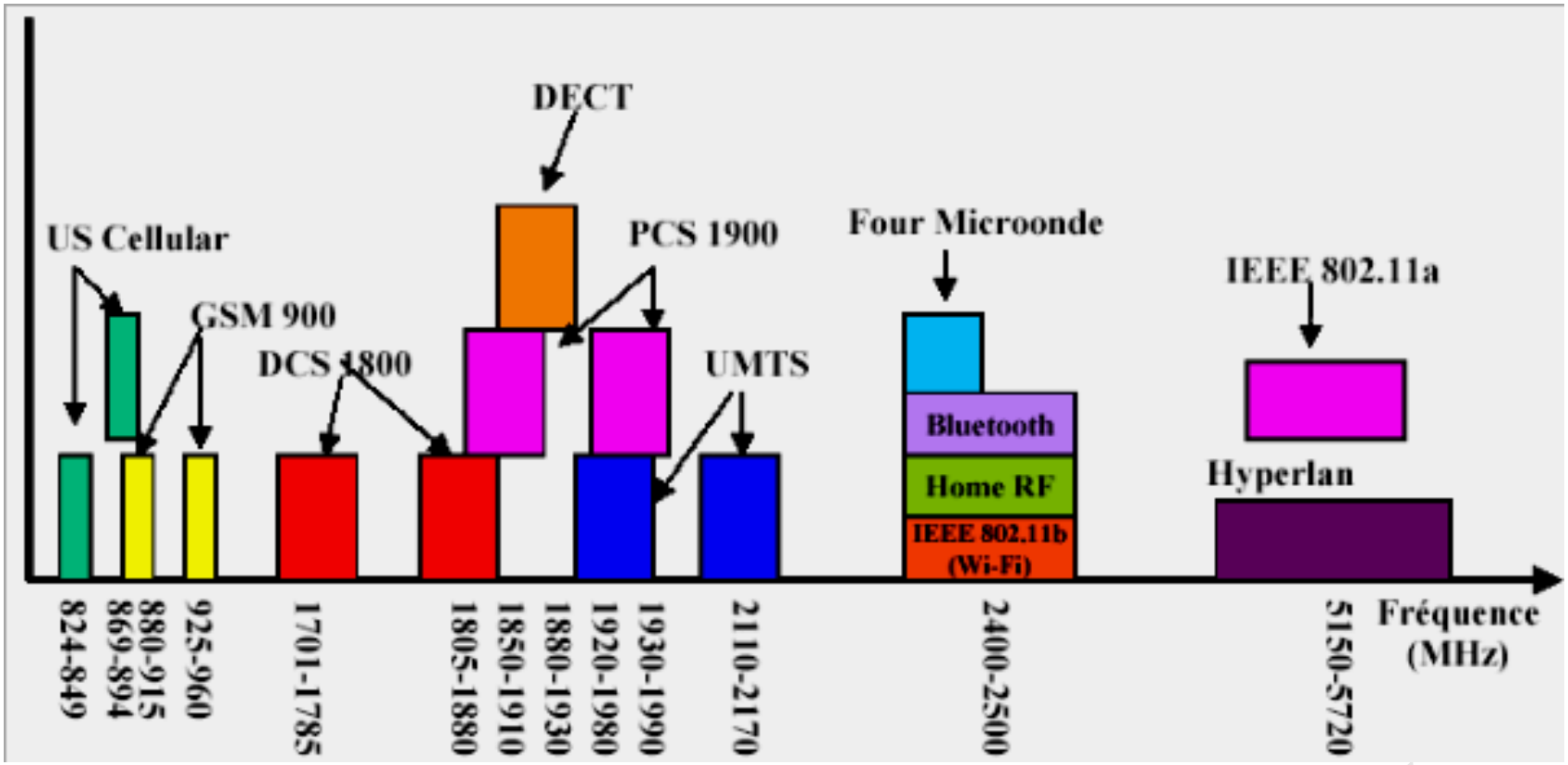
- Le principe du MPLS consiste à générer une étiquette courte, d'une longueur fixe, correspondant à un bref résumé de tout l'en-tête du datagramme IP.
- Le premier routeur MPLS rencontré apposera une telle étiquette et le datagramme pourra être envoyé très rapidement dans le réseau MPLS en fonction de cette étiquette.
- De l'autre côté du réseau, le datagramme IP sera de nouveau déballé et acheminé de la manière classique.
- L'étiquette n'est pas seulement créée en fonction de l'adresse de destination, mais aussi à partir de caractéristiques comme la qualité de service.
- Cette méthode peut être comparée à celle utilisée par la Poste. En mettant un code postal sur une lettre, il n'est pas nécessaire d'interpréter toute l'adresse avant d'arriver près de la destination.

Routage classique IP versus MPLS

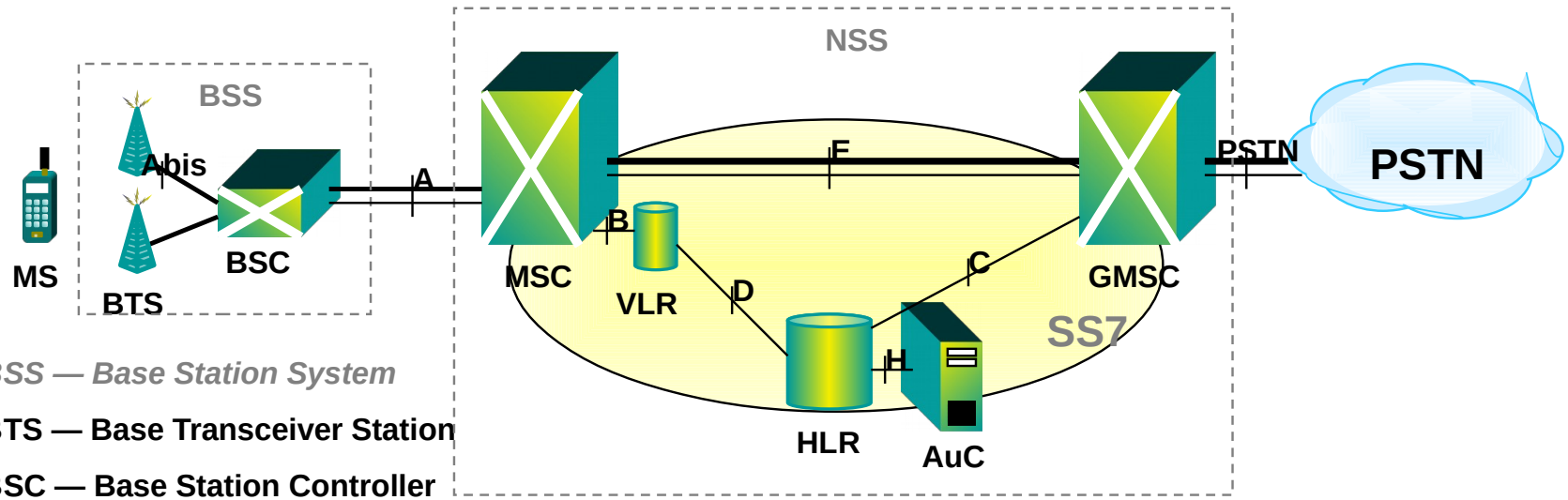




Les réseaux sans fil, fixes ou mobiles



Architecture GSM



BSS — Base Station System

BTS — Base Transceiver Station

BSC — Base Station Controller

MS — Mobile Station

NSS — Network Sub-System

MSC — Mobile-service Switching Controller

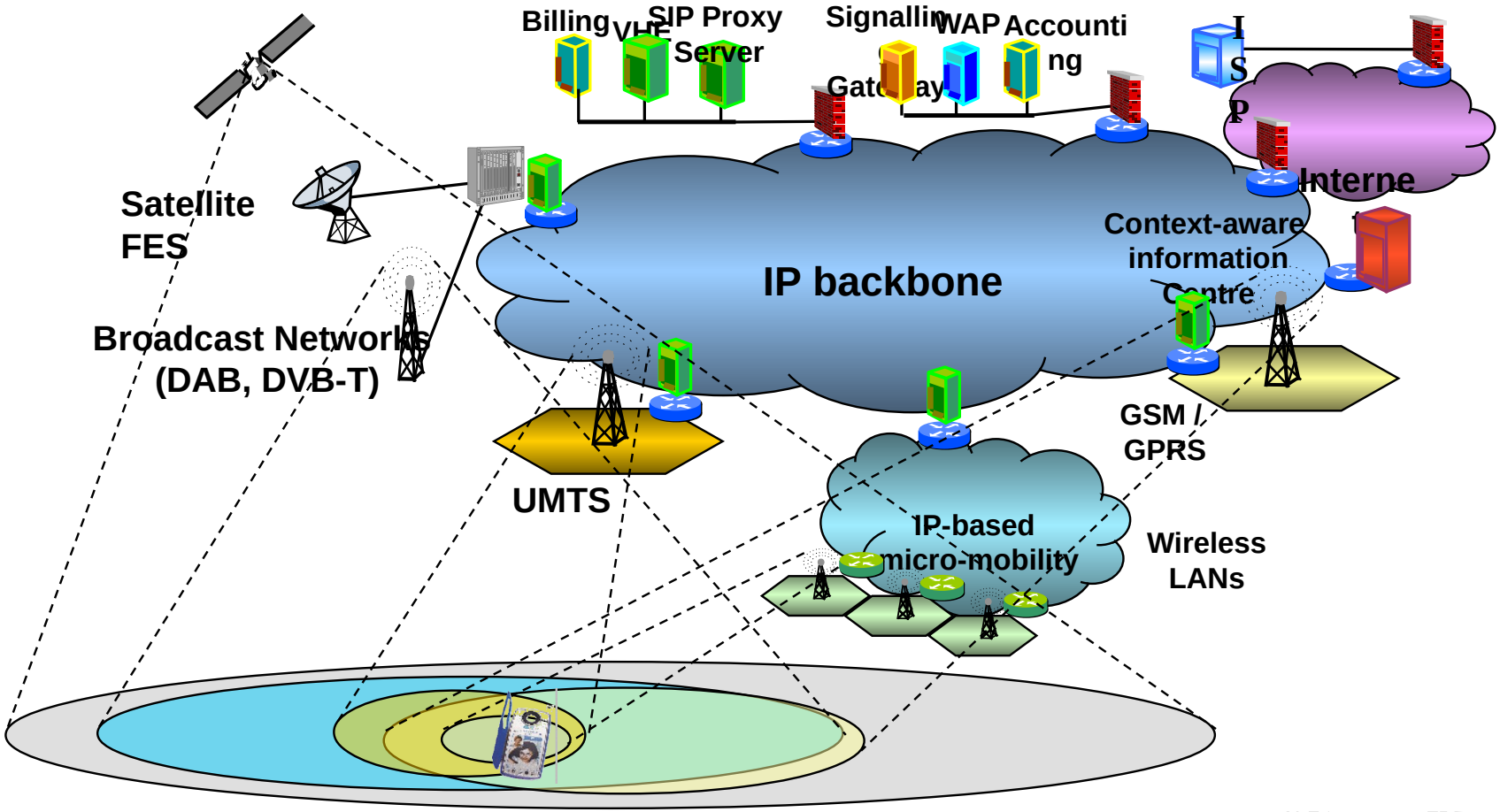
VLR — Visitor Location Register

HLR — Home Location Register

AuC — Authentication Server

GMSC — Gateway MSC

Vision tout-IP



Norme IEEE 802.11 (Wi-Fi)

- Wi-Fi : Wireless Fidelity
- Originellement IEEE 802.11b (11 Mbps) mais il y a eu aussi 802.11a (54 Mbps).
- Maintenant 802.11g (54 Mbps) et 802.11n (600 Mbps) sont très répandus et 802.11ac (6930 Mbps) est aussi disponible.
- Rayon de 50 mètres approximativement pour chaque point d'accès selon les obstacles.

WiMax

- WiMax (Worldwide Interoperability for Microwave Access) est une une famille de normes techniques permettant de livrer une connectivité haute vitesse sur le dernier kilomètre;
- Haut débit;
- Grande couverture;
- Alternative à ADSL.

Équipements WiMax

Bluetooth

- Technologie ayant évolué des principes de conception des réseaux cellulaires (basé sur 802.11 en mode ad hoc)
- Norme de communication de courte portée (jusqu'à 10 m mais peut être étendue à 100 m)
- Fonctionne à 3.4 GHz, près de la fréquence micro-onde, dans la part de la bande de fréquence qui ne requiert pas de licence d'opération (ISM - Industrial, Scientific and Medical)
- Effectue des sauts de fréquence rapides (1600 sauts/seconde) entre 79 fréquences de manière à éviter les interférences
- Technologie full-duplex (canal de communication dans les deux sens) en utilisant le TDD (Time Division Duplex)

Réseaux ad hoc

- La topologie change fréquemment, nœuds entrent, sortent, bougent;
- Découverte des voisins par diffusion de message;
- Capacités réduites en mémoire, calcul et puissance
- Pas d'identificateur global
- Déployés en grand nombre (10^3... 10^6)
- Réseaux de capteurs
- Exemple: Zigbee et Z-Wave pour la domotique

La sécurité des réseaux et applications

- Trois volets:
  - Intégrité de l'information;
  - Confidentialité de l'information;
  - Disponibilité du service, coût, réputation;
- Méchanismes:
  - Authentification: garantie de l'identité du correspondant;
  - Somme de contrôle cryptographique: intégrité et non répudiation
  - Encryption: confidentialité;
  - Contrôle d'accès: usager, groupe, administrateur, rôle, délégation...
- Attaques en déni de service, virus, cheval de troie, exploitation de vulnérabilité réseau, clé USB, accès physique...

La sécurité avec un réseau non fiable

- Un message peut être vu, intercepté, modifié, ajouté, retardé ou rejoué;
- Systèmes de clés publiques pour initier une connexion; des clés symétriques peuvent être communiquées et utilisées par la suite.
  - Paire de clés, une publique, l'autre privée;
  - Chaque utilisateur a deux paires, l'une avec chiffrement publique (déchiffrement secret) permettant à chacun d'écrire un message que seul cet utilisateur peut lire, l'autre avec déchiffrement publique (chiffrement secret) permettant à cet utilisateur d'envoyer un message que lui seul peut avoir envoyé mais tous peuvent lire.
- Avoir un message avec temps, date, numéro de séquence chiffré avec la clé publique du destinataire et la clé privée de l'envoyeur.

La sécurité d'un système

- Sécurité physique: accès au serveur, à l'ordinateur utilisé par l'administrateur de système, au courrier contenant les logiciels à installer...
- Sécurité humaine: persuader un employé de donner un accès...
- Les vulnérabilités dans les logiciels existent;
- Mise à jour de sécurité fréquentes;
- Multiples lignes de défense, pare-feu, détection d'intrusion, vérification d'intégrité, monitoring réseau, contrôle fin des accès, vérification des log...
- Analyse des risques, plan de contingence.

Les files d'attente (pour les paquets, requêtes...)

- On suppose que la queue a une capacité très grande et que le taux d'arrivée des requêtes n'est pas influencé par l'attente;
- Taux d'arrivée des requêtes, les requêtes se présentent aléatoirement selon un processus de Poisson: λ;
- Capacité de traitement des requêtes: μ requêtes par seconde;

Utilisation U d'un service est la fraction de temps occupé

$$U = \frac{\lambda}{\mu}$$

Nombre moyen de requêtes dans le système

$$\overline{N} = \frac{U}{1 - U}$$

Résumé