



## Internet: Passé, Présent et Futurs

Philippe Robert

### ► To cite this version:

Philippe Robert. Internet: Passé, Présent et Futurs. Conférence au lycée Jeanne d'Albret en 2009, 2009, Paris, France. hal-00816136

**HAL Id: hal-00816136**

**<https://inria.hal.science/hal-00816136>**

Submitted on 19 Apr 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Internet: Passé, Présent et Futurs

Philippe Robert  
INRIA Paris-Rocquencourt

Lundi 27 avril 2008

# Table des matières

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Introduction</b>                      | <b>5</b>  |
|          | <b>Comment transmettre des données ?</b> | <b>16</b> |
| <b>3</b> | <b>Internet du Futur</b>                 | <b>57</b> |

## Présentation

- Directeur de recherche à l'INRIA  
Institut de recherche en Informatique  
et en Mathématiques Appliquées.  
Responsable de l'équipe de recherche  
“Réseaux, Algorithmes et Probabilités”  
  
Mathématicien  
Spécialité : Probabilités.

## Présentation

- Directeur de recherche à l'INRIA  
Institut de recherche en Informatique  
et en Mathématiques Appliquées.  
Responsable de l'équipe de recherche  
“Réseaux, Algorithmes et Probabilités”  
  
Mathématicien  
Spécialité : Probabilités.
- Professeur Chargé de Cours  
à l'École Polytechnique.

## Plan de la Conférence

- Introduction.
- Transmission de données dans les réseaux.
- Internet aujourd'hui : TCP et Google.
- Le(s) Futur(s).

# 1. Introduction

# Problématique générale

La raison d'être d'un réseau :

- Diffuser
- Rechercher

**l'information.**



## Un bref aperçu historique

- $\simeq$  1900 : Réseaux téléphoniques.

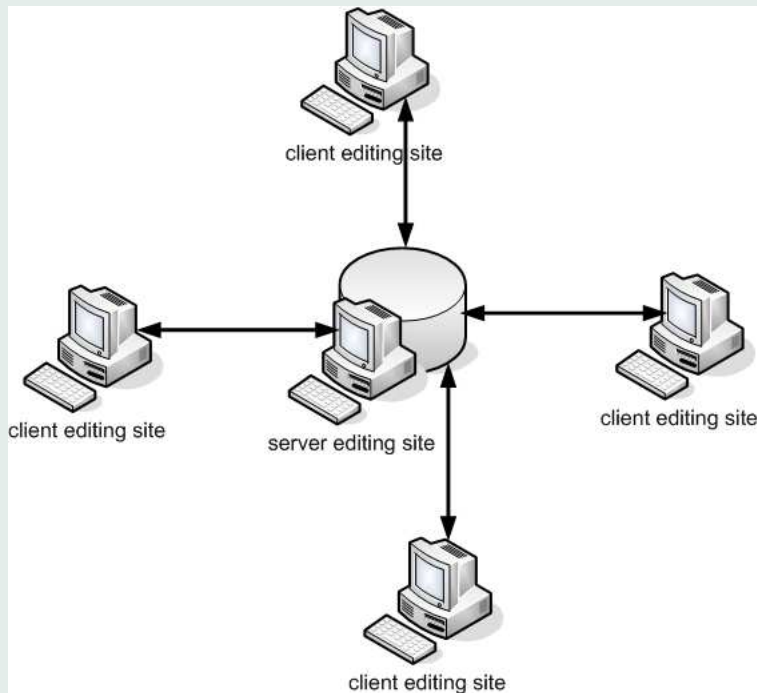
## Un bref aperçu historique

- $\simeq$  1900 : Réseaux téléphoniques.
- 1960 : Réseaux informatiques.
- Serveur central.

## IBM System/360



# Le Modèle du Serveur Central



## Un bref aperçu

- 1909 : Réseaux téléphoniques.
- 1960 : Réseaux informatiques.
- Serveur central.

## Un bref aperçu

- 1909 : Réseaux téléphoniques.
- 1960 : Réseaux informatiques.
  - Serveur central.
- 1980 : Systèmes distribués.
  - Réseaux Locaux.
  - Internet.
  - Réseaux Mobiles.

## Système Distribué

- Ensemble de Machines communicantes.
- Pas de Contrôle Central.
- Chacune agit de façon autonome.

## Système Distribué

- Ensemble de Machines communicantes.
- Pas de Contrôle Central.
- Chacune agit de façon **autonome**.

### Exemples

- Internet
- Réseaux Pair à Pair
- ...



# Problématique

## Fonctionnement d'un Système Distribué

**Pb :** Trouver une procédure, **un Algorithme** tel que

- Chaque machine utilise la même procédure.
- Globalement le réseau effectue la fonction demandée.

# Problématique

## Fonctionnement d'un Système Distribué

**Pb** : Trouver une procédure, **un Algorithme** tel que

- Chaque machine utilise la même procédure.
- Globalement le réseau effectue la fonction demandée.

## Exemples

- **Internet** : Transmission des données.
- **Réseaux Pair à Pair**  
Recherche et Stockage des données  
Téléphone (**Skype**). ...

# Les Systèmes Distribués dans la Nature

- Bacs de Poisson.
- Fourmilière, Essaims.

# Les Systèmes Distribués dans la Nature

- Bacs de Poisson.
- Fourmilière, Essaims.
- Cerveau.

# Les Systèmes Distribués dans la Nature

- Bacs de Poisson.
- Fourmilière, Essaims.
- Cerveau.
- ...

# Les Domaines Scientifiques pour les Réseaux

- Électronique.
- Informatique.
- Mathématiques. Modèles de la Physique.
- . . .

# La Recherche

## Universités :

États-Unis, France, Grande Bretagne, Hollande, Allemagne, ...

## Centres de Recherche :

### Privés :

IBM, Microsoft, Google, Alcatel-Lucent, Cisco,

### Publics :

CNRS, CWI, INRIA, INRS, ...

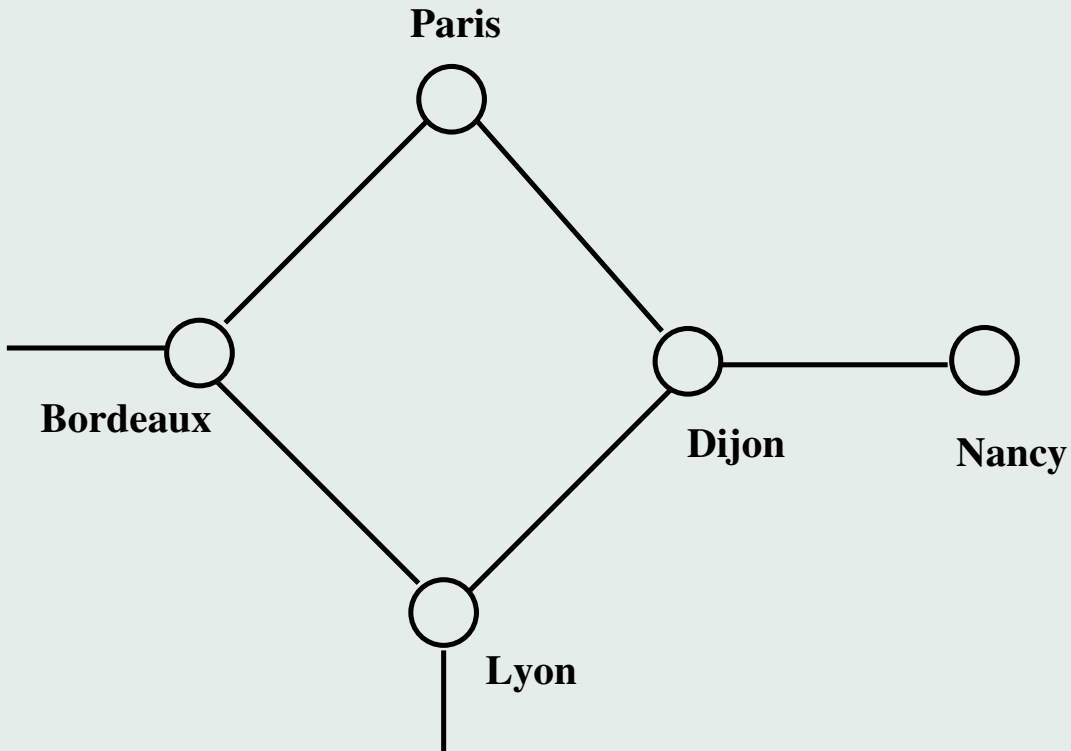
## 2. Comment transmettre des données ?

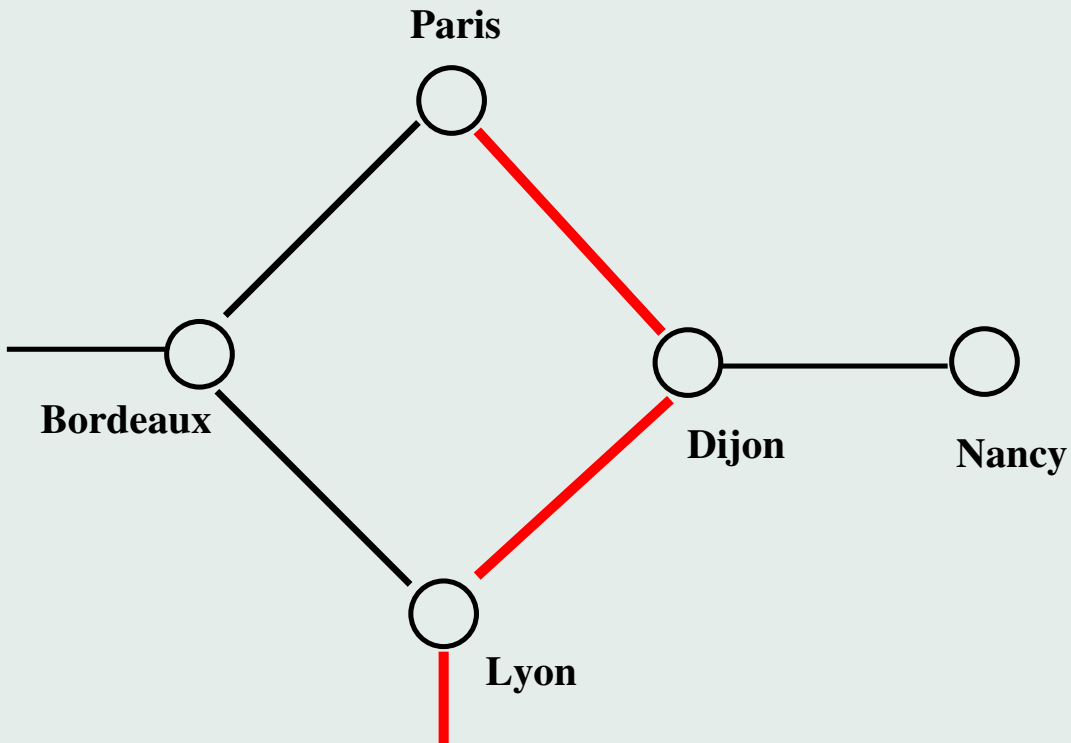


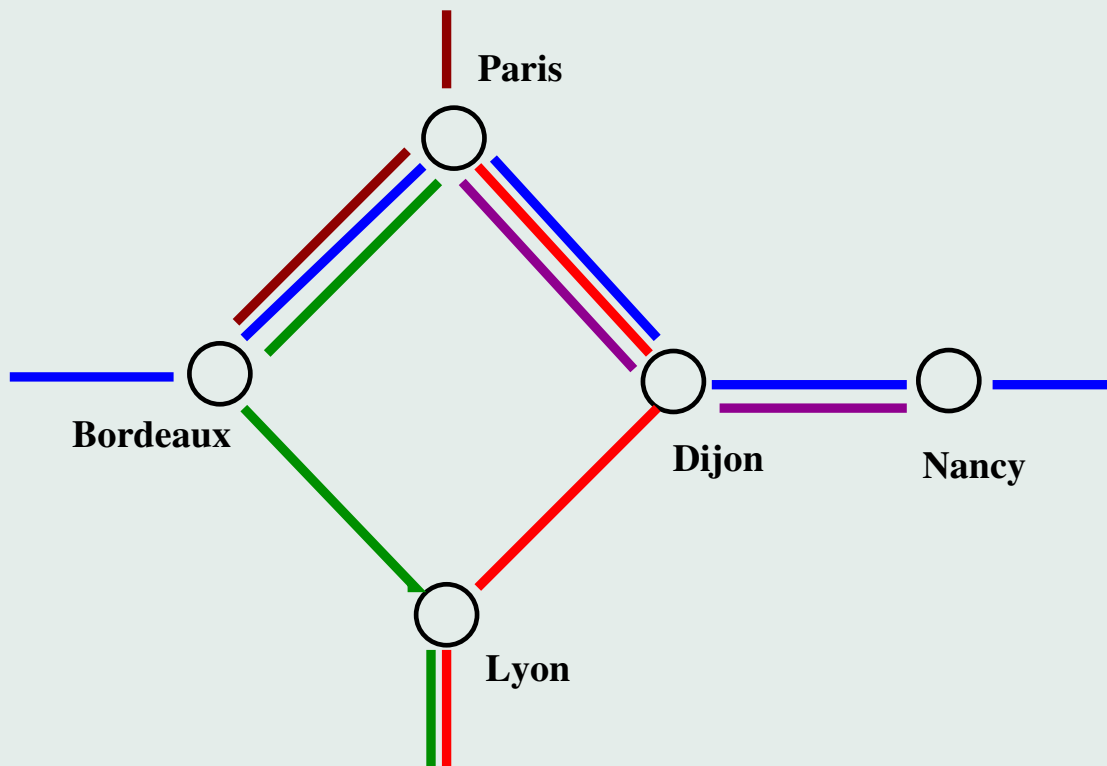
## Deux Modèles “Type” de Réseaux

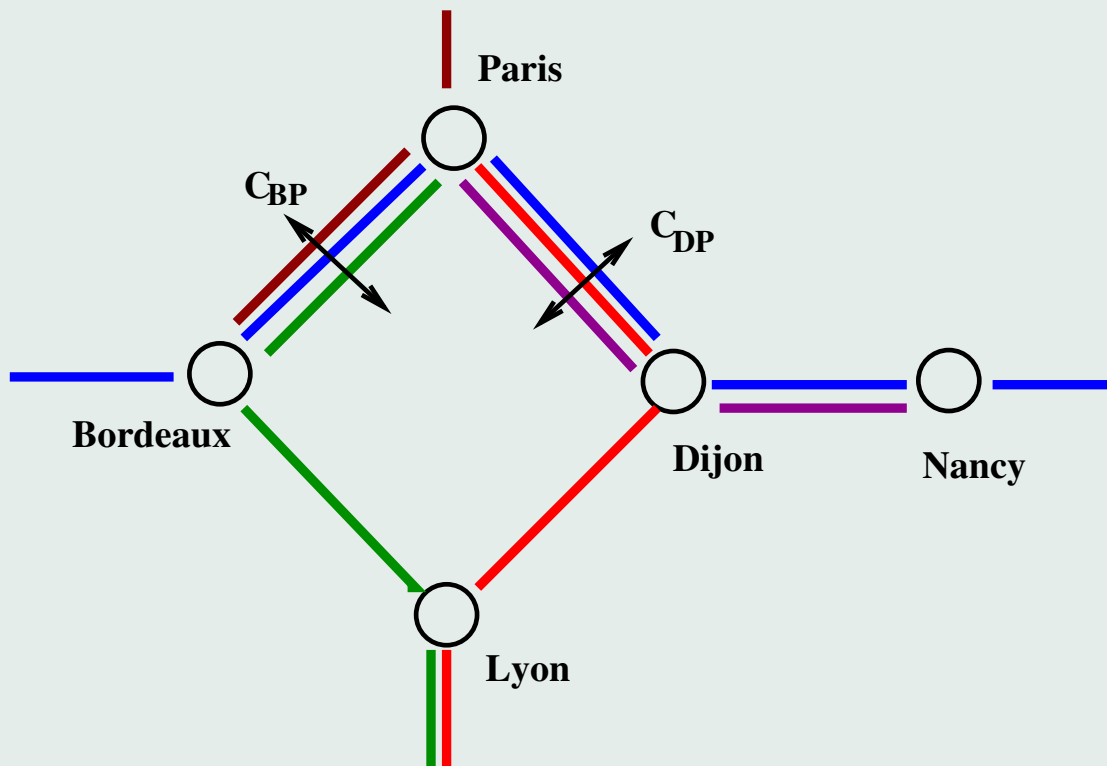
- Réseaux Téléphoniques
- Internet

# Les Réseaux Téléphoniques









- Le réseau téléphonique :  
Réseau à commutation de circuits.

- Le réseau téléphonique :  
Réseau à commutation de circuits.
- $\neq$  Réseaux à commutation de paquets.  
(Internet).



## Le réseau téléphonique

Un réseau à commutation de circuits :

- **Avantages**
  - Réservation de ressources.
  - Garantie de service.

# Le réseau téléphonique

Un réseau à commutation de circuits :

- **Avantages**

- Réserve de ressources.
- Garantie de service.

- **Problèmes**

- Système de type un peu centralisé.
- Reconfiguration complexe à mettre en œuvre.

Internet

Un peu d'histoire

## Débuts de l'Internet

- **1957.** Lancement du satellite Spoutnik par l'URSS

## Débuts de l'Internet

- 1957. Lancement du satellite Spoutnik par l'URSS  
⇒ création de l'ARPA aux États-Unis  
— Advanced Research Projects Agency—

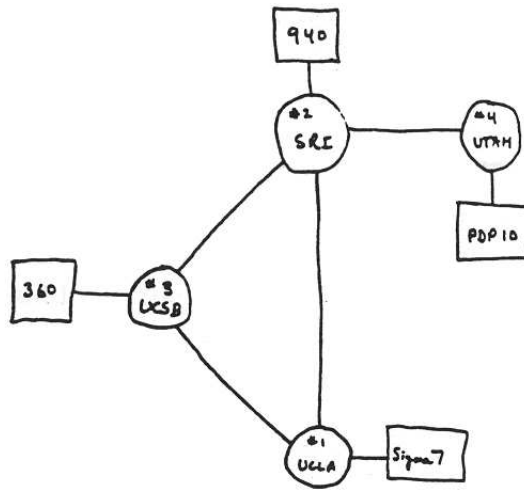
## Débuts de l'Internet

- **1957.** Lancement du satellite Spoutnik par l'URSS  
⇒ création de l'ARPA aux États-Unis  
— Advanced Research Projects Agency—
- **1957-1969.** Recherches Scientifiques sur la conception d'un réseau “robuste”.

## Débuts de l'Internet

- **1957.** Lancement du satellite Spoutnik par l'URSS  
⇒ création de l'ARPA aux États-Unis  
— Advanced Research Projects Agency—
- **1957-1969.** Recherches Scientifiques sur la conception d'un réseau “robuste”.
- **1969.** Réseau à 4 machines.
  1. University of California Los Angeles
  2. Stanford Research Institute
  3. University of California Santa Barbara
  4. University of Utah.





THE ARPA NETWORK

DEC 1969

4 NODES

FIGURE 6.2 Drawing of 4 Node Network  
(Courtesy of Alex McKenzie)

## La croissance du nombre de machines sur Internet

- 1970 : 15
- 1984 : 1000
- 1987 : 10 000
- 1989 : 100 000
- 1992 : 1 000 000
- 2004 : 290 000 000
- 2007 : 500 000 000

## La croissance du nombre de machines sur Internet

- 1970 : 15
- 1984 : 1000
- 1987 : 10 000
- 1989 : 100 000
- 1992 : 1 000 000
- 2004 : 290 000 000
- 2007 : 500 000 000

2007 :

Nb de PC : 1 000 000 000

Nb de téléphones portables : 3 000 000 000.

## Quelques Innovations liées à l'Internet

- 1960 : Réseau avec des paquets. Kleinrock.
- 1973 : Transmission Fiable des Données : TCP. Cerf et Kahn.
- 1989 : WWW. Berners-Lee.
- 1997 : Recherche sur le web. Brin et Page.

## Quelques Innovations liées à l'Internet

- 1960 : Réseau avec des paquets. Kleinrock.
- 1973 : Transmission Fiable des Données : TCP. Cerf et Kahn.
- 1989 : WWW. Berners-Lee.
- 1997 : Recherche sur le web. Brin et Page.



# Innovations : Évolutions

## Progrès technologiques :

- Vitesse des processeurs, des composants.
- Nouveaux matériaux, ...

# Innovations : Évolutions

## Progrès technologiques :

- Vitesse des processeurs, des composants.
- Nouveaux matériaux, ...

**Important mais n'est plus la source dominante d'innovation.**

# Innovations : Évolutions

## Progrès technologiques :

- Vitesse des processeurs, des composants.
- Nouveaux matériaux, ...

**Important mais n'est plus la source dominante d'innovation.**

## Conception d'Algorithmes

- Langues/Programmes Informatiques.
- Modélisation mathématique.
- ...

**Importance croissante.**



Un peu de détail

## La lettre de l'alphabet électronique : l'octet

- L'ordinateur ne traite que des **0** et des **1**

## La lettre de l'alphabet électronique : l'octet

- L'ordinateur ne traite que des 0 et des 1
- Unité minimale : un octet  $O$   
 $O = o_1o_2o_3o_4o_5o_6o_7o_8$  avec  $o_i = 0$  ou 1 pour  $1 \leq i \leq 8$ .

## La lettre de l'alphabet électronique : l'octet

- L'ordinateur ne traite que des 0 et des 1
- Unité minimale : un octet  $O$   
 $O = o_1o_2o_3o_4o_5o_6o_7o_8$  avec  $o_i = 0$  ou 1 pour  $1 \leq i \leq 8$ .

Exemples :

a  $\rightarrow$  01100001    b  $\rightarrow$  01101100

10  $\rightarrow$  00001010

## La lettre de l'alphabet électronique : l'octet

- L'ordinateur ne traite que des **0** et des **1**
- **Unité minimale** : un octet ***O***  
***O*** =  $o_1 o_2 o_3 o_4 o_5 o_6 o_7 o_8$  avec  $o_i = 0$  ou  $1$  pour  $1 \leq i \leq 8$ .

Exemples :

**a**  $\rightarrow$  01100001    **b**  $\rightarrow$  01101100

**10**  $\rightarrow$  00001010

- **Nombre entier** est codé avec 4 octets.
- **Nombre à virgule** est codé avec 8 octets.

## Ordres de grandeur

- 10 octets  
un mot
- 2 Kilo-octets :  $2 \cdot 10^3$  octets  
une page de texte
- 50 Méga-octets :  $50 \cdot 10^6$  octets  
le contenu d'un album de musique mp3
- 5 Giga-octets :  $5 \cdot 10^9$  octets  
Un film
- 25 Giga-octets  
le contenu complet d'une bibliothèque publique

## L'unité d'information de l'Internet : LE PAQUET

Un paquet :  $\simeq 15\text{Ko}$   
un entête + les données

- L'entête : 64 octets  
contient entre autres l'adresse de la machine qui doit recevoir le paquet.
- Les données : le contenu du fichier transféré.

# L'unité d'information de l'Internet : LE PAQUET

Un paquet :  $\simeq 15\text{Ko}$   
un entête + les données

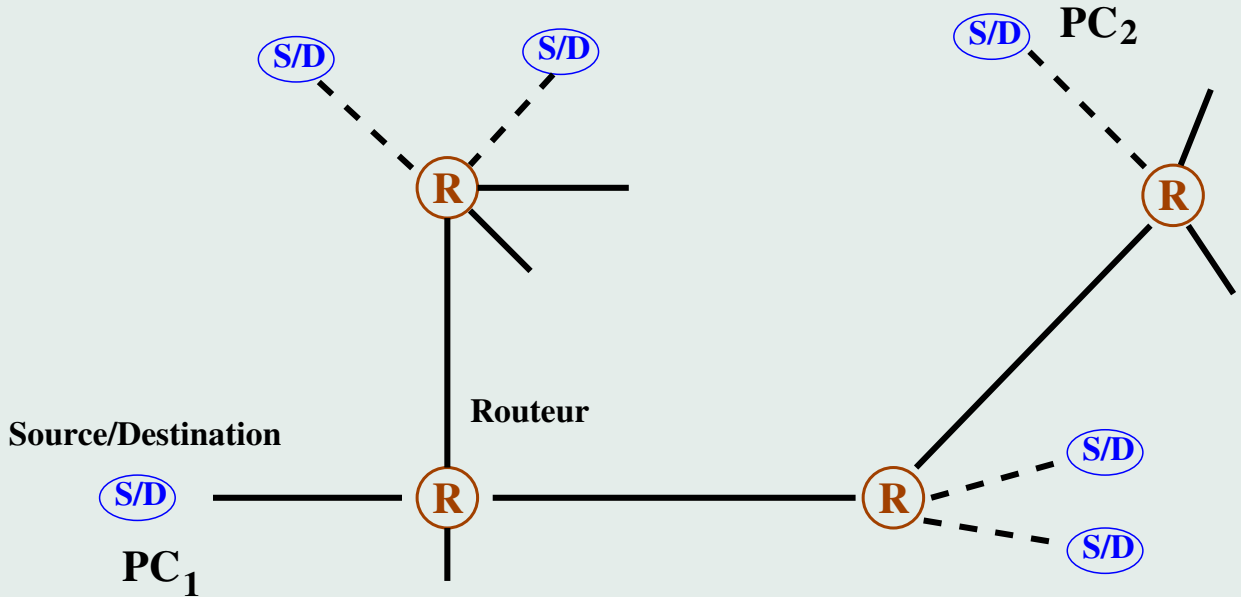
- L'entête : 64 octets  
contient entre autres l'adresse de la machine qui doit recevoir le paquet.
- Les données : le contenu du fichier transféré.

## Exemples

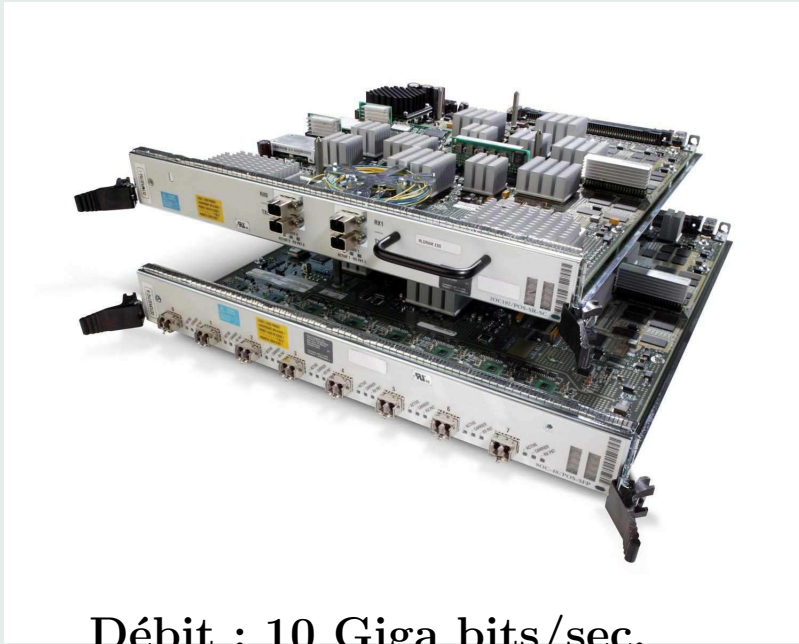
- Un CD mp3 : 400 000 paquets.
- Un film : 4 000 000 paquets.



## Internet : Une vue simplifiée



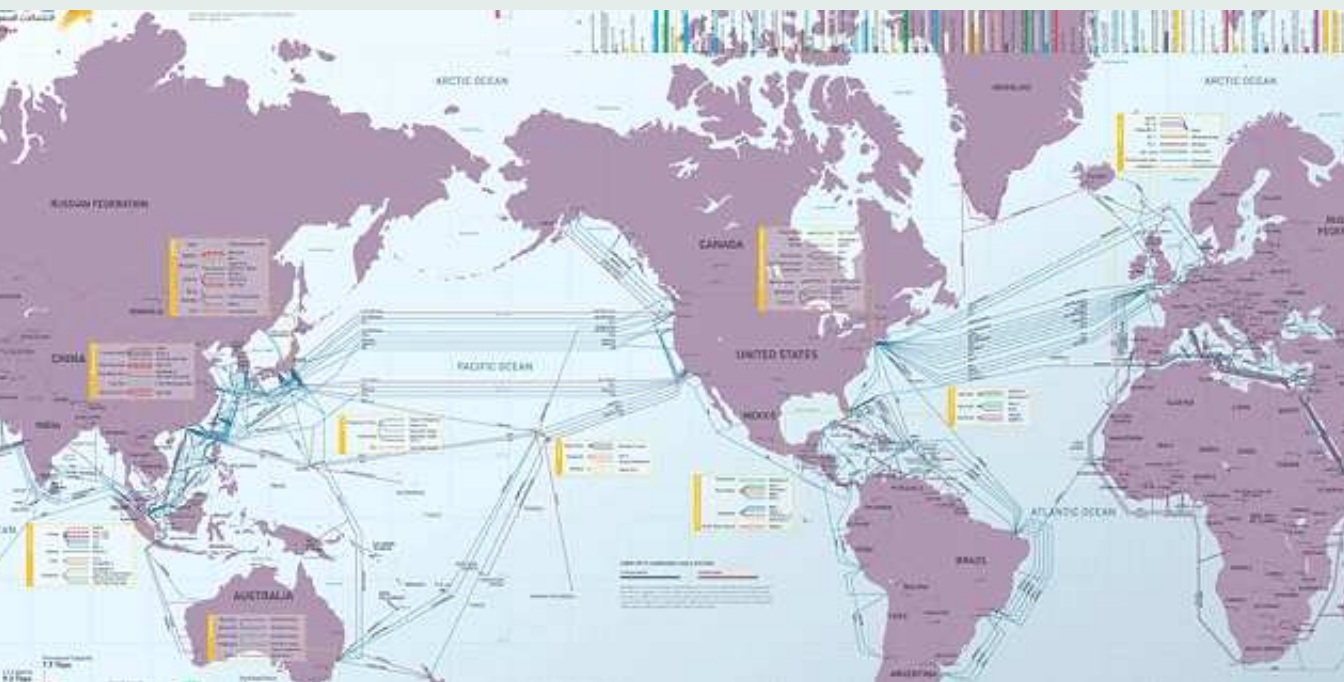
## Un Routeur : Cisco OC192



Débit : 10 Giga bits/sec.

## Voyage d'un paquet de Paris à Stanford (Californie)

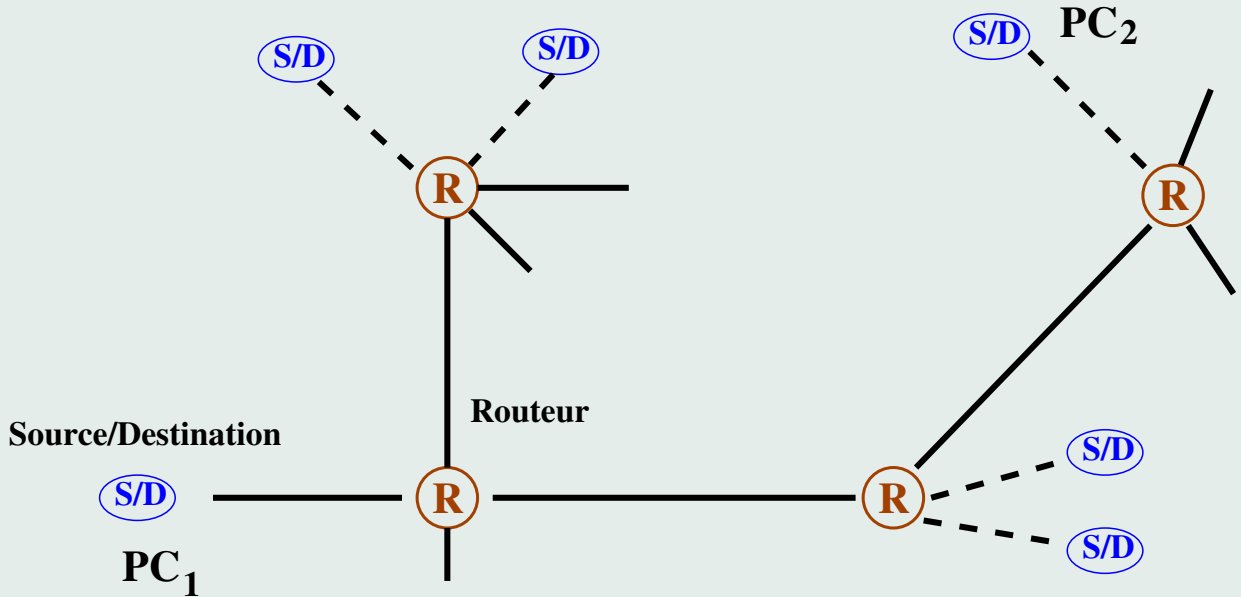
|    |   |             |
|----|---|-------------|
| 1  | @work   | France      |
| 2  | rocq-renater-gw.inria.fr                      |             |
| 3  | gi2-0-inria.cssi.renater.fr                   |             |
| 4  | gi8-10-paris1-rtr-021.cssi.renater.fr         |             |
| 5  | vl89-te0-0-0-3-paris1-rtr-001.cssi.renater.fr |             |
| 6  | renater.rt1.par.fr.geant2.net                 |             |
| 7  | so-3-0-0.rt1.lon.uk.geant2.net                | UK          |
| 8  | so-2-0-0.rt1.ams.nl.geant2.net                | Hollande    |
| 9  | 198.32.11.50                                  | États-Unis  |
| 10 | so-0-0-0.0.rtr.wash.net.internet2.edu         | Washington  |
| 11 | so-0-0-0.0.rtr.atla.net.internet2.edu         | Atlanta     |
| 12 | so-3-2-0.0.rtr.hous.net.internet2.edu         | Houston     |
| 13 | so-3-0-0.0.rtr.losa.net.internet2.edu         | Los Angeles |
| 14 | hpr-lax-hpr-i2-newnet.cenic.net               |             |
| 15 | svl-hpr-lax-hpr-10ge.cenic.net                |             |
| 16 | oak-hpr-svl-hpr-10ge.cenic.net                |             |
| 17 | hpr-stan-ge-oak-hpr.cenic.net                 |             |
| 18 | bbrb-i2.Stanford.EDU                          | Stanford    |



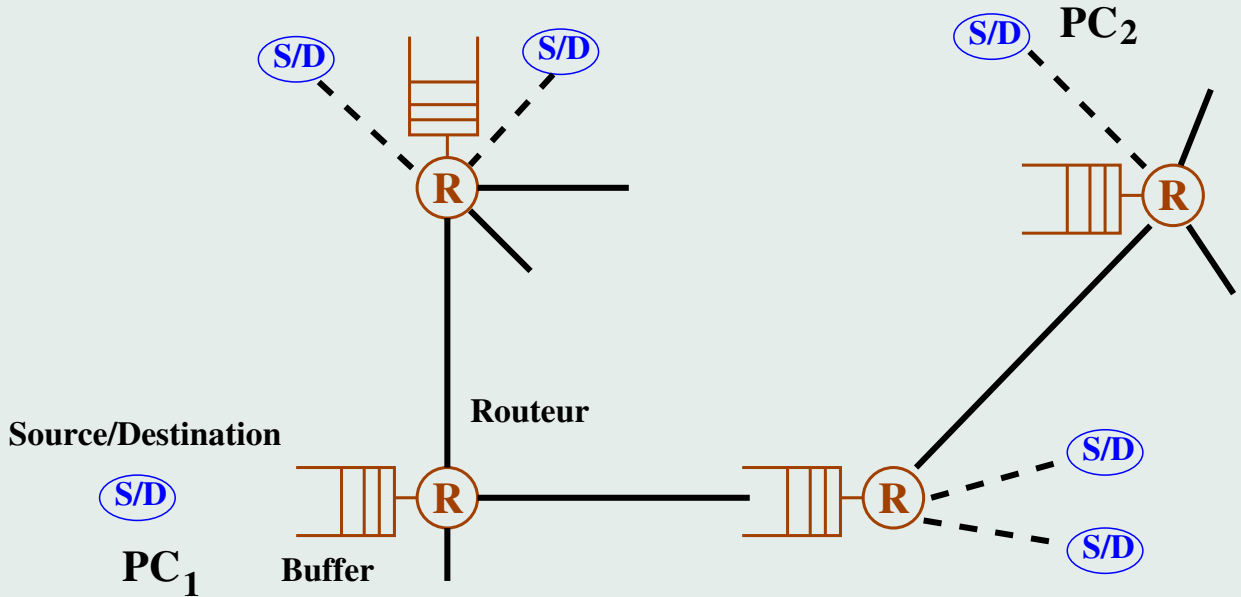
Source : TeleGeography Research

2009 PriMetrica, Inc. All ©

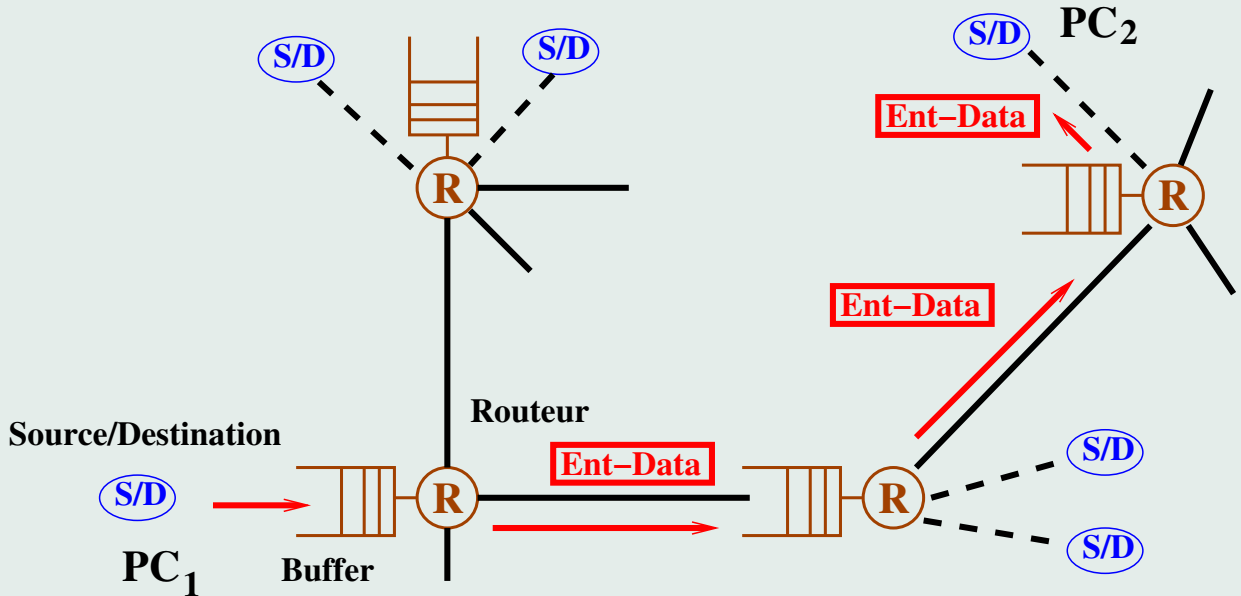
## Internet : Une vue simplifiée



## Internet : Une vue simplifiée



## Internet : Une vue simplifiée

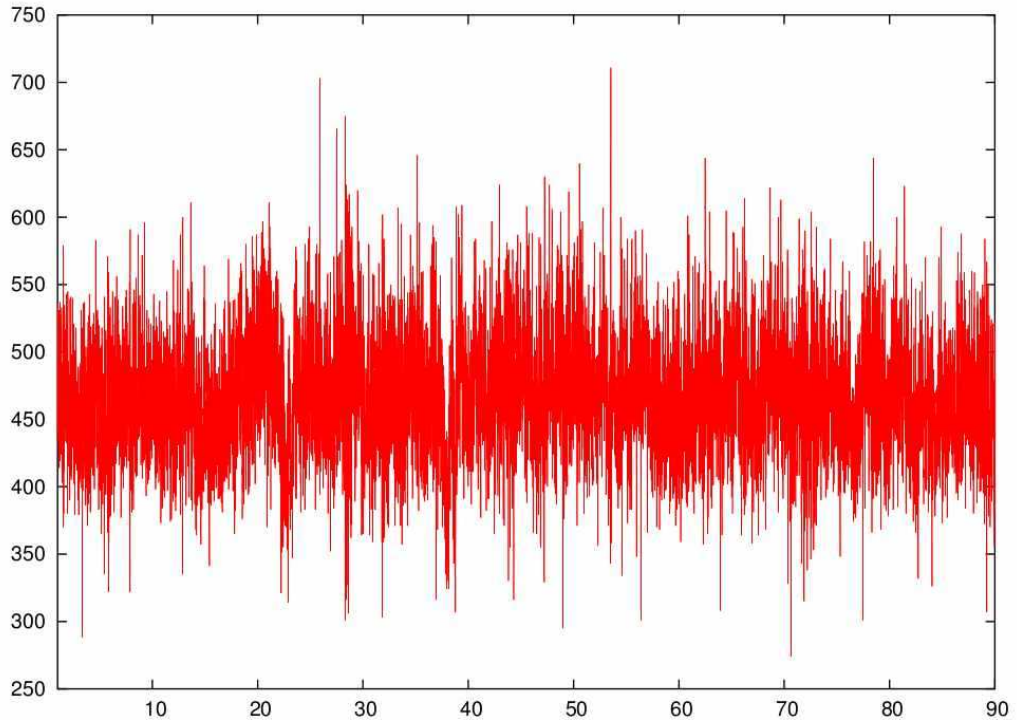


# Internet

## Un réseau à commutation de paquets

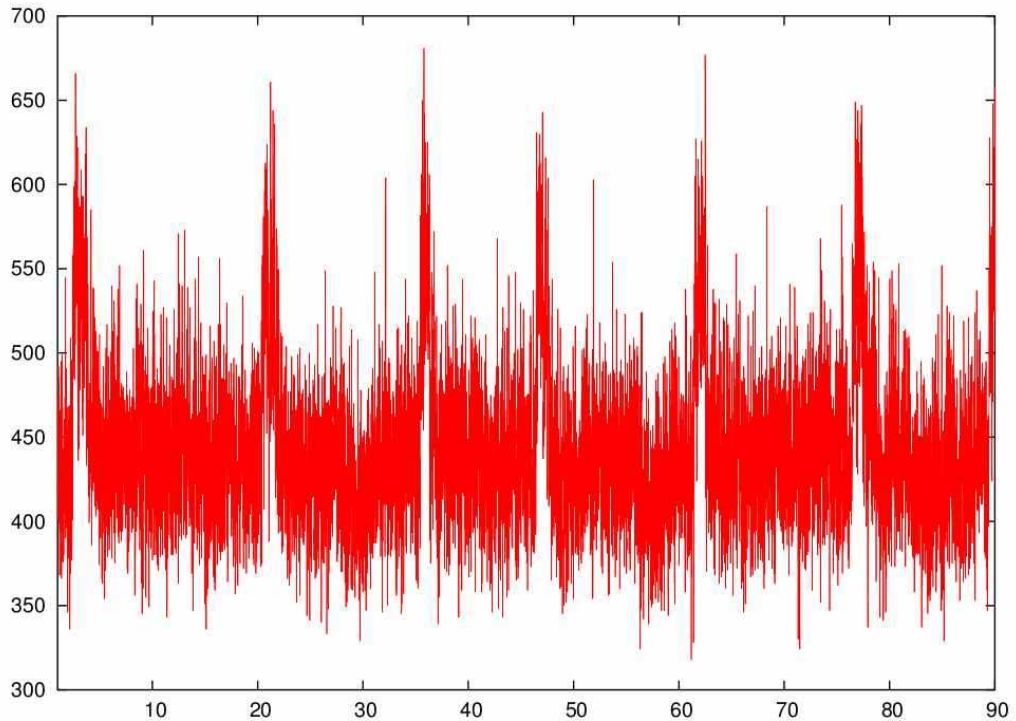
- Messages divisés en paquets.
- Paquets acheminés individuellement.
- **Avantages**
  - Système distribué.
  - Flexibilité : Évolution facile.





**Trafic : Nombre de paquets par 1/100s**

**Colorado State Univ. 03/10/2002 à 21 :00 :06**



**Trafic : Nombre de paquets par 1/100s**

Colorado State Univ. 01/09/2002 à 00 :56 :09

## Nature du Traffic Internet

- Haut débit.
- Imprévisible : événements aléatoires.  
⇒ débits variables.

## Nature du Traffic Internet

- Haut débit.
- Imprévisible : événements aléatoires.  
⇒ débits variables.

## Modèles Mathématiques pour décrire le trafic

- Composante Aléatoire Essentielle.
- Prévoir le “comportement moyen du réseau”.
- Prédire les performances de débit, délai, ...

## Composition du trafic

Colorado State Univ. 03/10/2002 à 21 :00 :06

| Application  | % nb. paquets |
|--------------|---------------|
| e-donkey     | 0.377860      |
| ms-streaming | 0.532267      |
| blubster     | 0.588439      |
| smtp         | 0.712183      |
| realtime     | 0.779447      |
| ftp-data     | 1.872506      |
| battlefield  | 2.011356      |
| nntp         | 2.512342      |
| gnutella-svc | 3.221723      |
| kazaa        | 8.200591      |
| www-http     | 10.292393     |

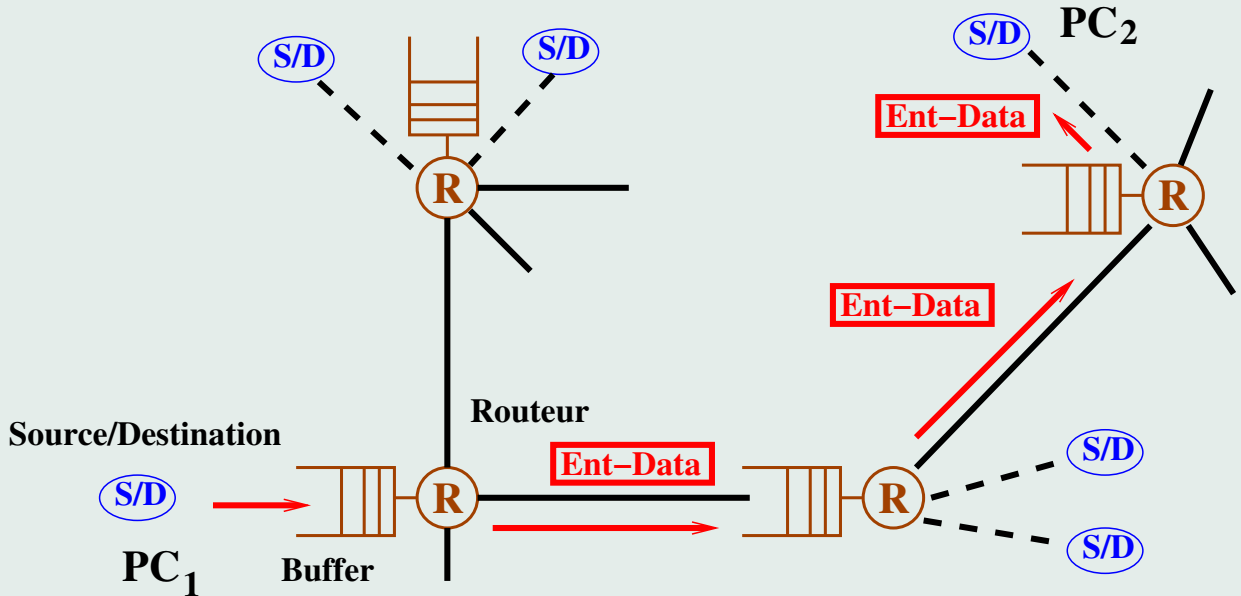
Durée : 90 sec.

59919 Stations/4210294 Paquets/2.2 Go

## Composition du trafic

- **ADSL** : 80% Pair à pair.
- Montée en puissance des transferts vidéo :  
**Youtube, Dailymotion, ...**

## Internet : Une vue simplifiée



Transfert du fichier  $F$   
de la machine  $PC_1$  vers la machine  $PC_2$

- Sur chaque machine : un programme contrôle l'échange.



Transfert du fichier  $F$   
de la machine  $PC_1$  vers la machine  $PC_2$

- Sur chaque machine : un programme contrôle l'échange.
- $PC_1$  segmente en  $n$  paquets une copie de  $F$  et envoie le numéro  $1$ , puis  $2, \dots, n$  à  $PC_2$ .

Transfert du fichier  $F$   
de la machine  $PC_1$  vers la machine  $PC_2$

- Sur chaque machine : un programme contrôle l'échange.
- $PC_1$  segmente en  $n$  paquets une copie de  $F$  et envoie le numéro  $1$ , puis  $2$ ,  $\dots$   $n$  à  $PC_2$ .

Mémoire des routeurs finie : en cas de congestion  
 $\Rightarrow$  Le réseau perd des paquets.

## Transfert du fichier $F$ de la machine $PC_1$ vers la machine $PC_2$

- Sur chaque machine : un programme contrôle l'échange.
- $PC_1$  segmente en  $n$  paquets une copie de  $F$  et envoie le numéro  $1$ , puis  $2$ ,  $\dots$   $n$  à  $PC_2$ .

Mémoire des routeurs finie : en cas de congestion  
 $\Rightarrow$  Le réseau perd des paquets.

Problème : Comment transmettre de façon fiable dans un réseau qui ne l'est pas ?

# Transmission de Données sur Internet

**TCP** : Transmission Control Protocol.

- Protocole de transmission de données.  
(Protocole : ensemble de procédures, d'algorithmes).
- **> 95%** du trafic Internet contrôlé par TCP.

# Les principes de base TCP

Cerf and Kahn (1973)

- Accusé de réception des messages.
  - Régulation des envois : à un instant une source a au plus  $W$  paquets en circulation dans le réseau.
- $W$  : Taille de la fenêtre de congestion.

## Les principes de base TCP (II)

### Contrôle de la congestion Jacobson (1987)

- Transmission de  $W$  paquets OK :

$$W \rightarrow W + 1$$

- Un paquet est perdu :

$$W \rightarrow W/2.$$

## Exemple simplifié

Transfert de  $F$  entre les machines  $PC_1$  et  $PC_2$  :

- $PC_2$  : envoie un paquet pour demander  $F$ .
- $PC_1$  envoie les paquets  $n^{\circ} 1, 2, \dots W$  à  $PC_2$

## Exemple simplifié

Transfert de  $F$  entre les machines  $PC_1$  et  $PC_2$  :

- $PC_2$  : envoie un paquet pour demander  $F$ .
- $PC_1$  envoie les paquets  $n^{\circ} 1, 2, \dots W$  à  $PC_2$
- $PC_2$  reçoit le  $n^{\circ}1 \Rightarrow$  envoie paquet  $OK-1$  à  $PC_1$   
“j’ai bien reçu 1”...



## Exemple simplifié

Transfert de  $F$  entre les machines  $PC_1$  et  $PC_2$  :

- $PC_2$  : envoie un paquet pour demander  $F$ .
- $PC_1$  envoie les paquets  $n^{\circ} 1, 2, \dots, W$  à  $PC_2$
- $PC_2$  reçoit le  $n^{\circ} 1 \Rightarrow$  envoie paquet  $OK-1$  à  $PC_1$   
“j’ai bien reçu 1”...
- ...
- Si tout va bien :  
 $PC_1$  a reçu  $OK-1, \dots, OK-W$   
 $\Rightarrow PC_1$  Envoie les  $W + 1$  paquets suivants  
 $n^{\circ} W + 1, \dots, W + W + 1 = 2W + 1$ .

## Exemple simplifié

- Une perte :

le paquet  $n^{\circ}i$  a été perdu

$\Rightarrow PC_1$  renvoie le paquet  $n^{\circ}i$

$PC_1$  envoie les  $\lfloor W/2 \rfloor$  paquets

$n^{\circ} W + 1, \dots W + \lfloor W/2 \rfloor$ .

- etc...

## Commutation de paquets

+++ Adaptation aux conditions de trafic.

-- Pas de garantie de débit, d'accès, ...

## Commutation de paquets

+++ Adaptation aux conditions de trafic.

-- Pas de garantie de débit, d'accès, ...

Remarquable propriétés d'auto-stabilisation

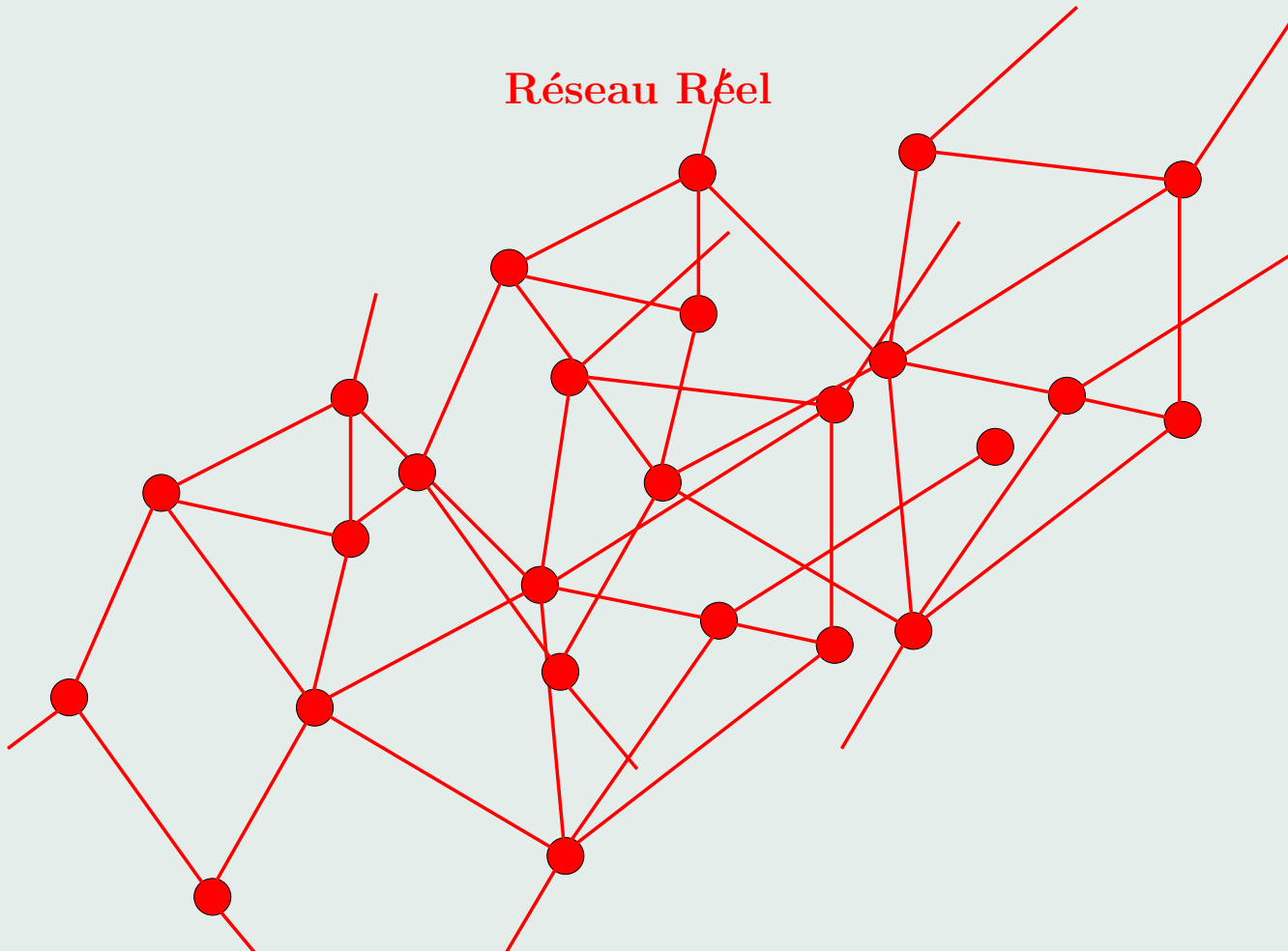
### 3. Internet du Futur

## Vue d'Ensemble

- Hauts Débits
- Mobilité
- Réseaux Virtuels
- Sécurité

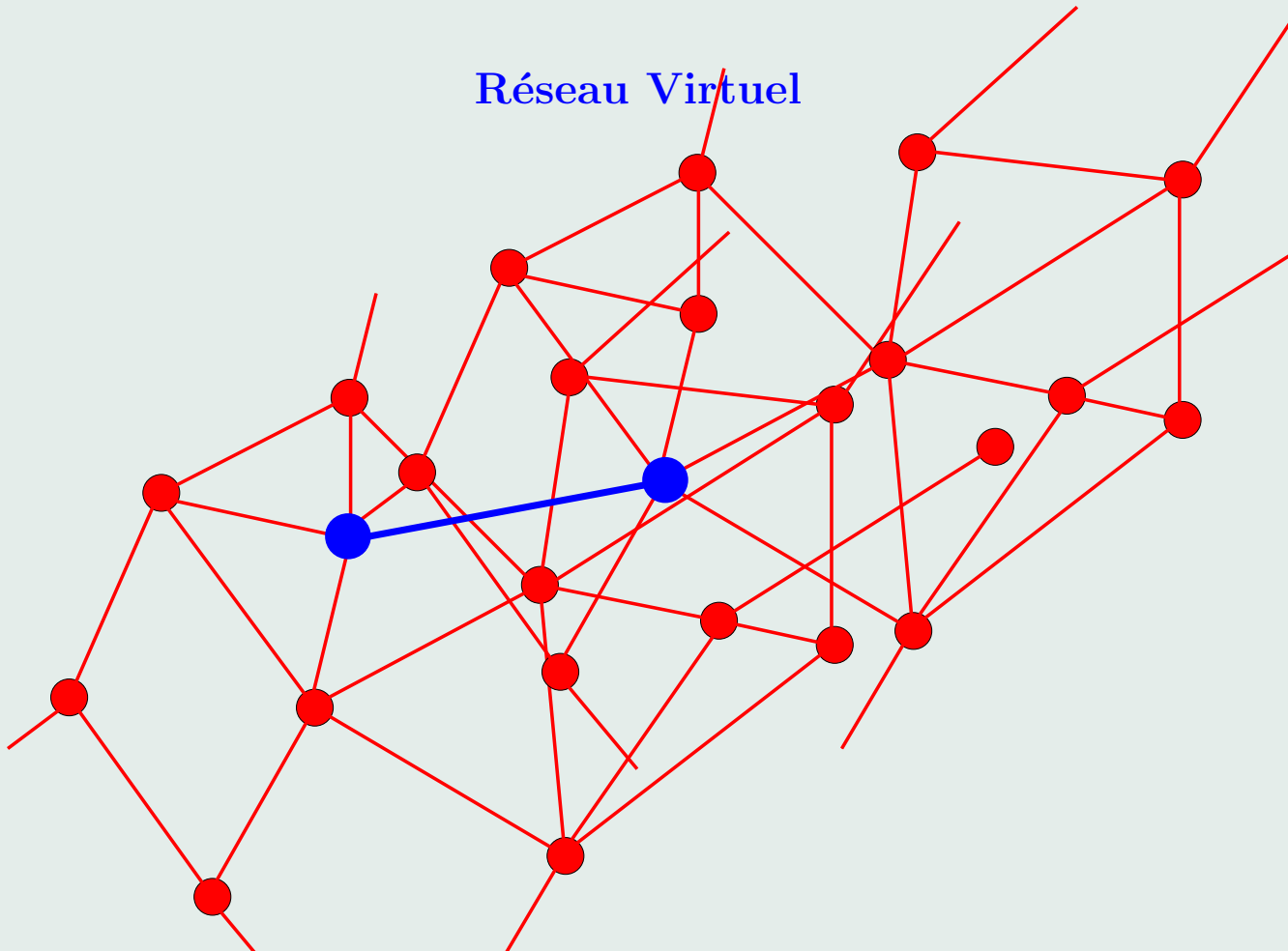
# Réseaux Virtuels

# Réseau Réel

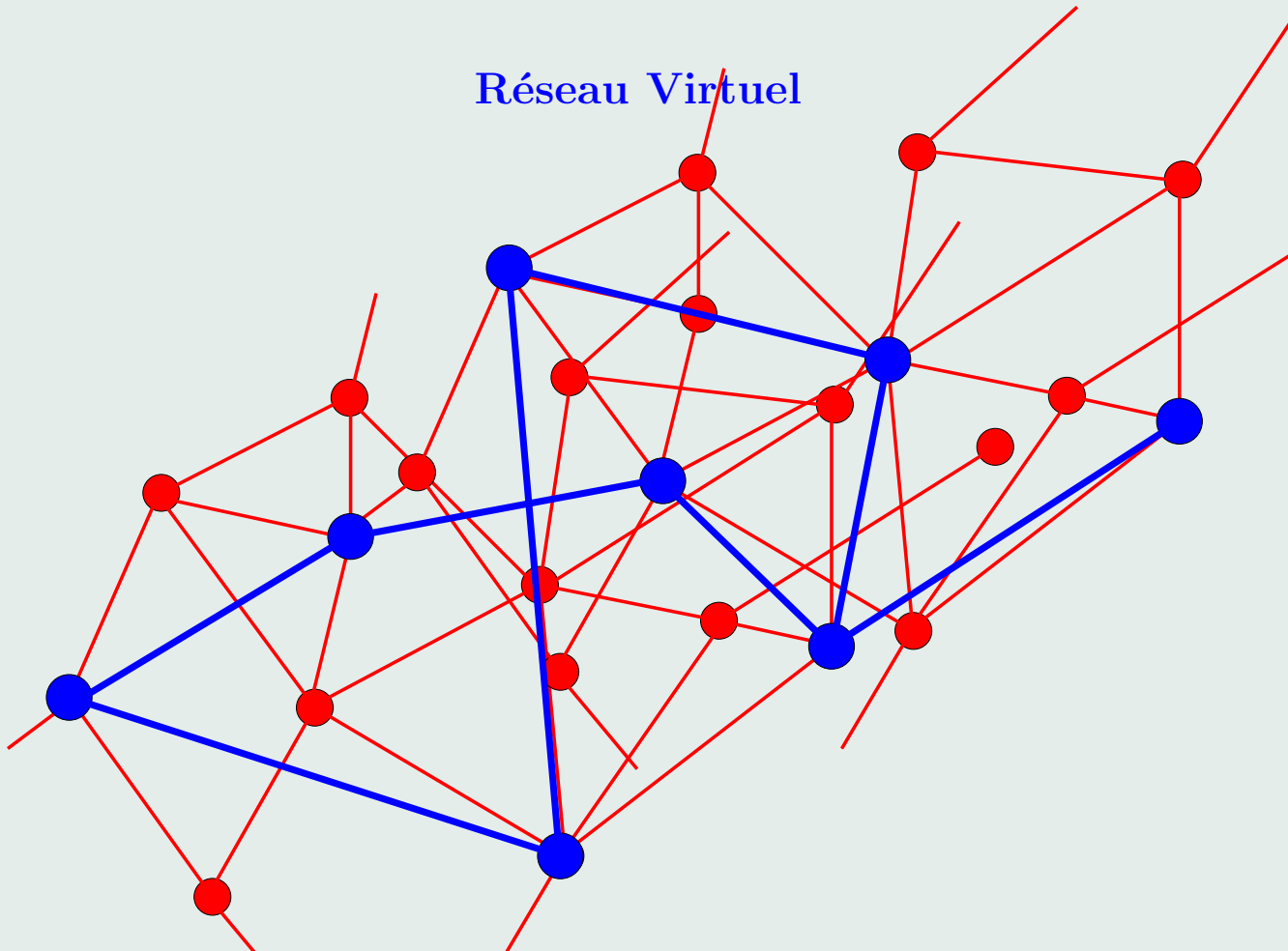




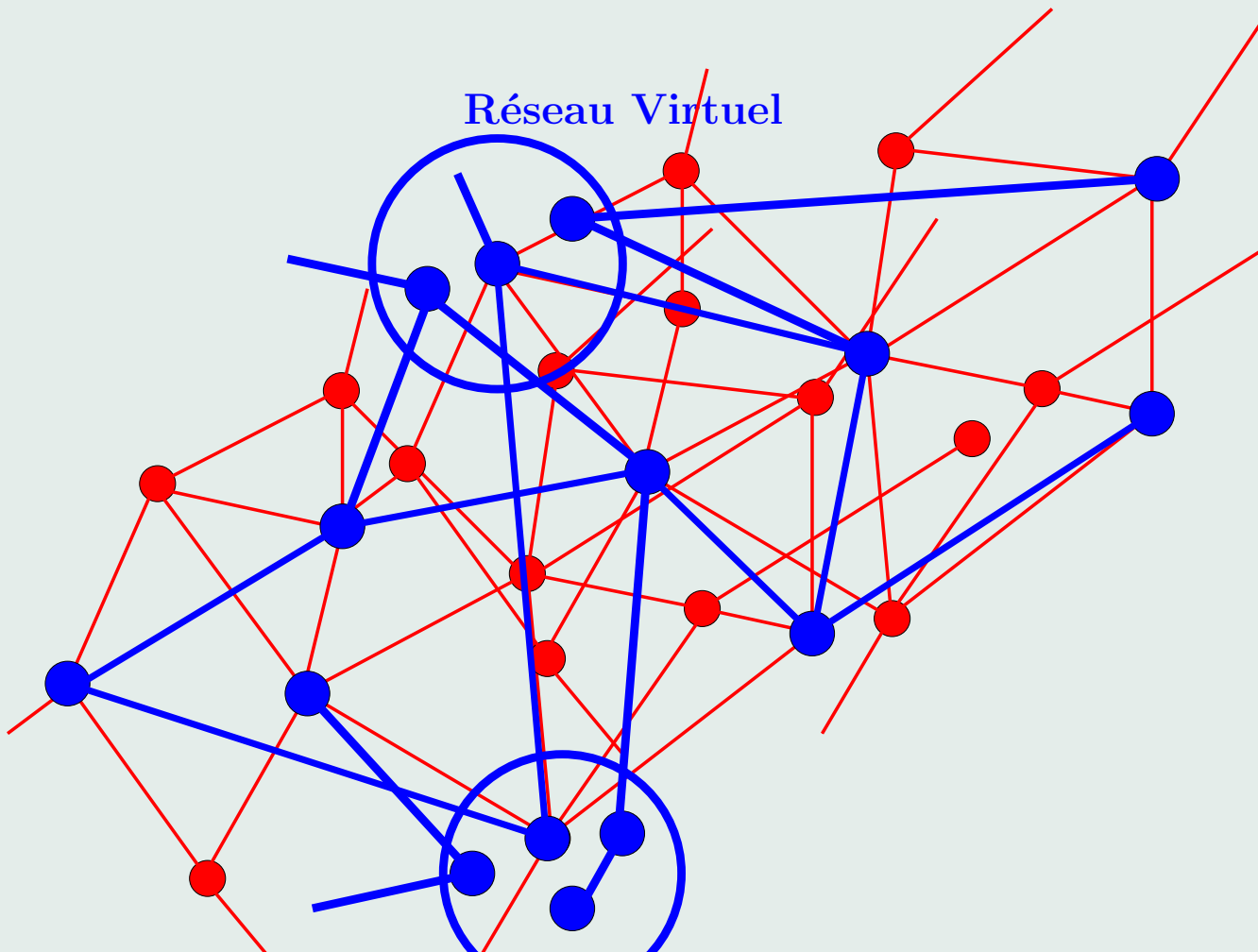
# Réseau Virtuel



# Réseau Virtuel



# Réseau Virtuel



## Exemple : Les Réseaux Pair à Pair

- Ensemble  **$\mathcal{S}$**  de nœuds ayant un thème commun.
  - **Très grand nombre** de nœuds.
  - Chacun a un ensemble de fichiers.
- Organisation en serveurs de fichiers.

## Exemple : Les Réseaux Pair à Pair

- Ensemble **S** de nœuds ayant un thème commun.
- **Très grand nombre** de nœuds.
- Chacun a un ensemble de fichiers.  
Organisation en serveurs de fichiers.
- **Deux nœuds sont voisins si** (par exemple)
  - Peu de sauts entre celles-ci.
  - Grand débit offert par un des nœuds.

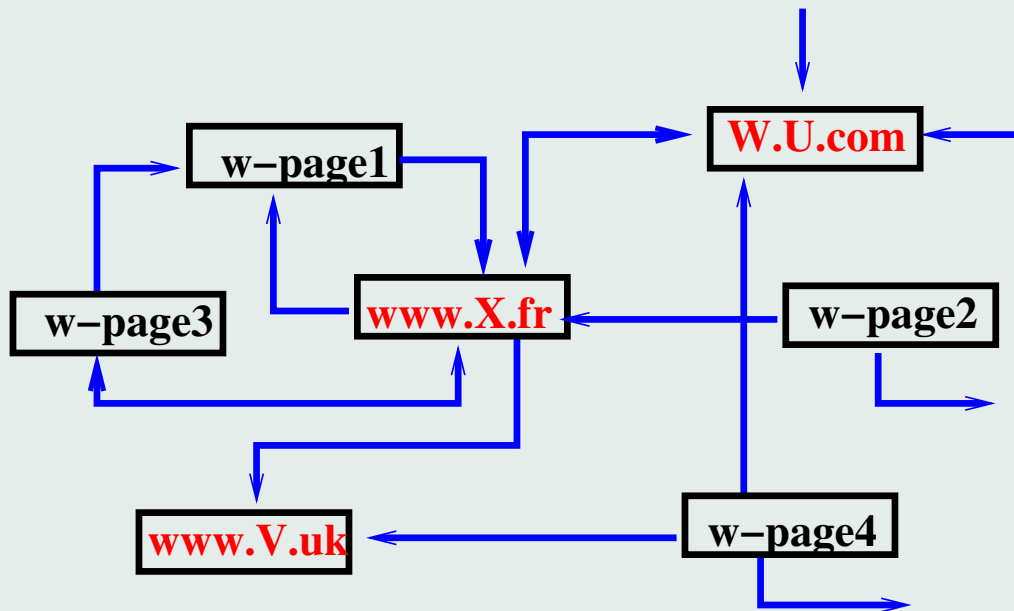
## Exemple : Les Réseaux Pair à Pair

- Ensemble **S** de nœuds ayant un thème commun.
- **Très grand nombre** de nœuds.
- Chacun a un ensemble de fichiers.  
Organisation en serveurs de fichiers.
- **Deux nœuds sont voisins si** (par exemple)
  - Peu de sauts entre celles-ci.
  - Grand débit offert par un des nœuds.

Exemple :

**Ancêtre** : Napster

Aujourd'hui : e-mule, bittorrent.



Un autre réseau virtuel : le web  
25 milliards de pages web indexées par Google

# Algorithmes sur les réseaux virtuels

## Sujets de recherche actuels :

- Recherche/Stockage de données sur Internet :  
Pair à Pair



# Algorithmes sur les réseaux virtuels

## Sujets de recherche actuels :

- Recherche/Stockage de données sur Internet :  
Pair à Pair
- Recherche d'informations sur le web :

## Algorithmes sur les réseaux virtuels

### Sujets de recherche actuels :

- Recherche/Stockage de données sur Internet :  
Pair à Pair
- Recherche d'informations sur le web :  
Google (le plus connu mais pas le seul).

## Comment Marche un Moteur de Recherche ?

**Problème :** Recherche d'un site web ayant une information sur le sujet “**XYZ**”

## Comment Marche un Moteur de Recherche ?

**Problème :** Recherche d'un site web ayant une information sur le sujet “**XYZ**”

– **Première étape (facile)**

Recherche de l'ensemble des sites web ayant ce mot “XYZ”



Lycee Jeanne d'Albret

Rechercher

[Recherche avancée](#)  
[Préférences](#)

☒ Rechercher sur le Web ☐ Rechercher les pages en français

Web

Résultats 1 - 10 sur un total d'environ 10 100 pour **Lycee Jeanne d'Albret** (0,21 secondes)

**[Lycée Jeanne d'Albret - Accueil](#)**

Lycée d'enseignement général et classes Prépas (CPGE) à Saint Germain en Laye.

[www.lyc-albret-st-germain-laye.ac-versailles.fr/](#) - 44k - [En cache](#) - [Pages similaires](#) -

[CPGE](#)

[Filière Commerciale Option ...](#)

[L'internat](#)

[Version imprimable](#)

[Mentions légales](#)

[Plan du site](#)

[Autres résultats, domaine ac-versailles.fr »](#)

**[L'E.P.S A JEANNE d'ALBRET - Accueil](#)**

L'E.P.S A JEANNE d'ALBRET. Vous êtes ici : Accueil du site ... 4 juges représenteront notre lycée pour la compétition départementale qui sera qualificative ...

[www.lyc-albret-st-germain-laye.ac-versailles.fr/EPS/](#) - 22k -

[En cache](#) - [Pages similaires](#) -

**[Collège Jeanne d'Albret - Pau 64](#)**

[www.ac-bordeaux.fr/Etablissement/CJAAlbretPau/](#) - 10k - [En cache](#) - [Pages similaires](#) -

Done

## Comment Marche un Moteur de Recherche ?

**Problème :** Recherche d'un site web ayant une information sur le sujet “**XYZ**”

- **Première étape (facile)**

Recherche de l'ensemble des sites web ayant ce mot “XYZ”

- **Deuxième étape**

Quel est le site web le plus pertinent ?

## Comment Marche un Moteur de Recherche ?

**Principe** : Il faut trouver une fonction  $f$  telle que :  
À une page web  $P$  on associe  $f(P) \in [0, 1]$ .

$P$  plus pertinent que  $Q$  si  $f(P) > f(S)$ .

## Comment Marche un Moteur de Recherche ?

**Principe :** Il faut trouver une fonction  $f$  telle que :  
À une page web  $P$  on associe  $f(P) \in [0, 1]$ .

$P$  plus pertinent que  $Q$  si  $f(P) > f(S)$ .

$$\mathcal{E}_{xyz} = \{P : P \text{ page web et } P \text{ contient "XYZ"}\}$$



## Comment Marche un Moteur de Recherche ?

**Principe :** Il faut trouver une fonction  $f$  telle que :  
À une page web  $P$  on associe  $f(P) \in [0, 1]$ .

$P$  plus pertinent que  $Q$  si  $f(P) > f(S)$ .

$$\mathcal{E}_{xyz} = \{P : P \text{ page web et } P \text{ contient "XYZ"}\}$$

**Action :** Afficher les adresses des pages  $Q_1, \dots, Q_{10}$   
ayant les 10 plus grandes valeurs pour  $f$  sur  $\mathcal{E}_{xyz}$ .

# La fonction $f$ pour Google

Brin et Page (1997)

Modèle mathématique : surfeur aléatoire  $S$

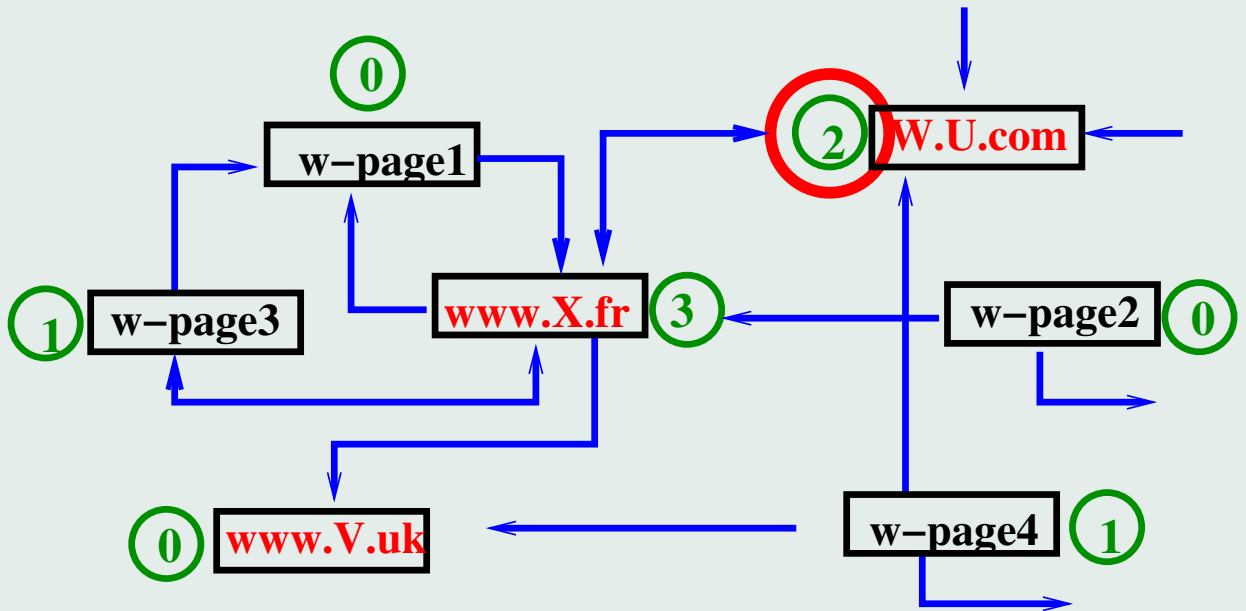
## La fonction $f$ pour Google

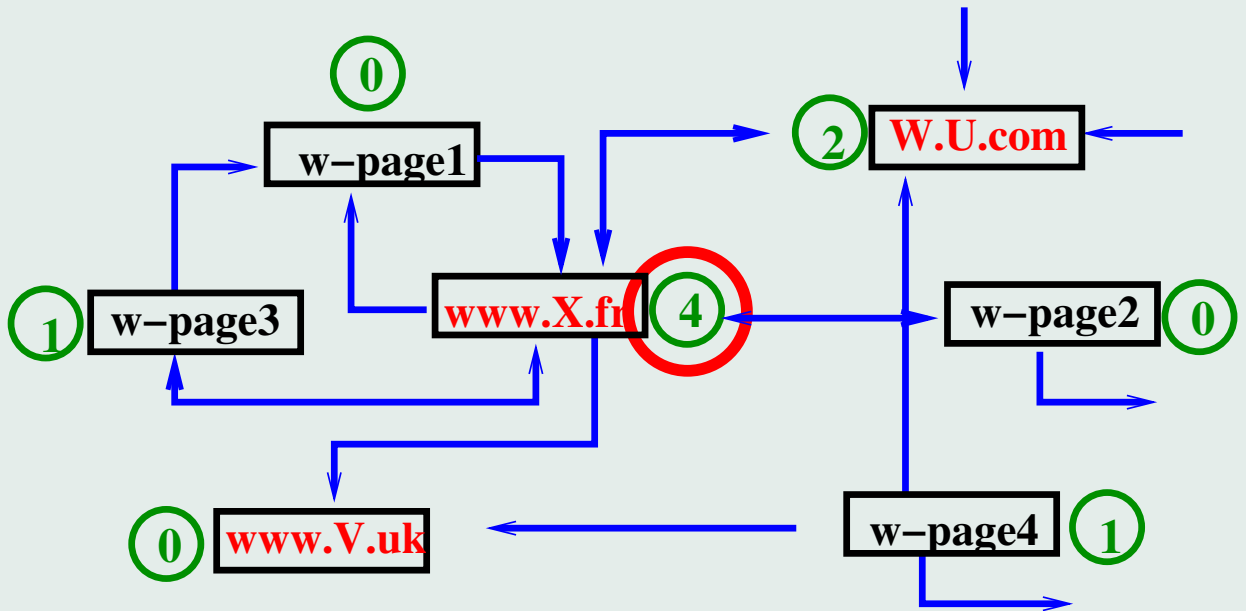
Brin et Page (1997)

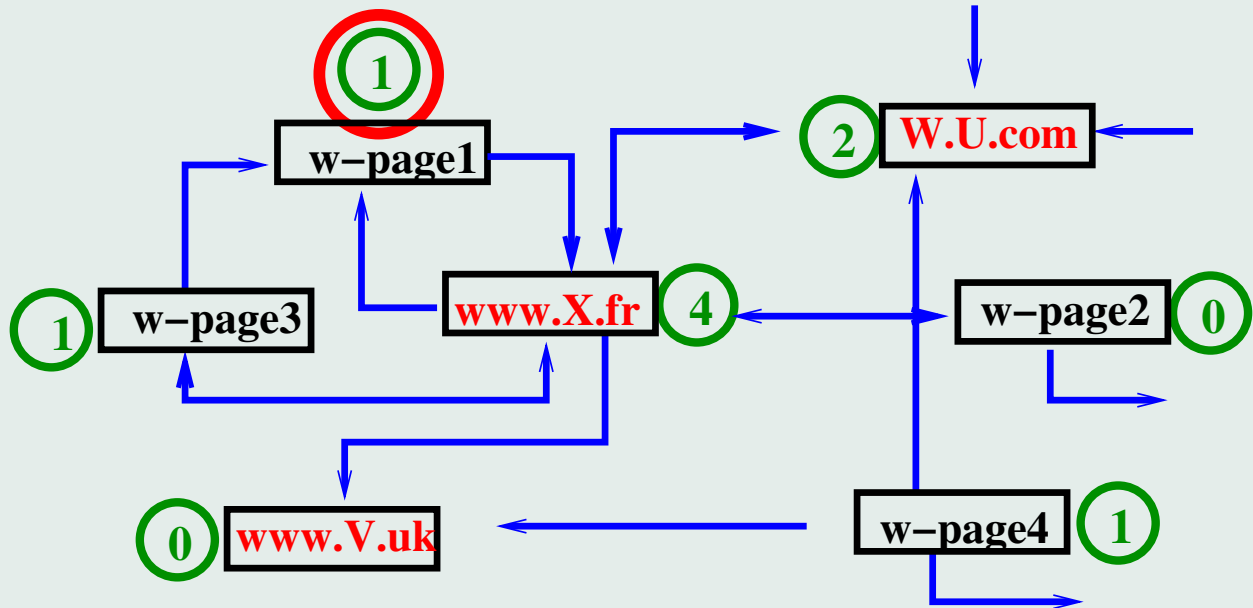
Modèle mathématique : surfeur aléatoire  $S$

–  $S$  navigue au hasard sur le web :

Si  $S$  sur une page web à l'instant  $t$   
à  $t + 1$ ,  $S$  choisit au hasard un lien de cette page  
et va sur la page web correspondante, etc...







## La fonction $f$ pour Google

Durée du surf  $T = 10^{12}$  (par exemple).

Si  $P$  est une page web,

$$f(P) = \frac{1}{T}N(P),$$

où  $N(P)$  : nb de passages à  $P$  entre  $0$  et  $T$ .

## La fonction $f$ pour Google

Durée du surf  $T = 10^{12}$  (par exemple).

Si  $P$  est une page web,

$$f(P) = \frac{1}{T}N(P),$$

où  $N(P)$  : nb de passages à  $P$  entre  $0$  et  $T$ .

$P_1$  plus pertinent que  $P_2$  si  $f(P_1) > f(P_2)$ .



## La fonction $f$ pour Google

Durée du surf  $T = 10^{12}$  (par exemple).

Si  $P$  est une page web,

$$f(P) = \frac{1}{T}N(P),$$

où  $N(P)$  : nb de passages à  $P$  entre  $0$  et  $T$ .

$P_1$  plus pertinent que  $P_2$  si  $f(P_1) > f(P_2)$ .

### Problèmes

– Dépend de  $T$  ?

## La fonction $f$ pour Google

Durée du surf  $T = 10^{12}$  (par exemple).

Si  $P$  est une page web,

$$f(P) = \frac{1}{T}N(P),$$

où  $N(P)$  : nb de passages à  $P$  entre  $0$  et  $T$ .

$P_1$  plus pertinent que  $P_2$  si  $f(P_1) > f(P_2)$ .

### Problèmes

- Dépend de  $T$  ?
- Dépend du chemin du surfeur ?

## Résultats de Maths :

La limite existe :

$$\pi(P) = \lim_{T \rightarrow +\infty} \frac{N_T(P)}{T}$$

$P_1$  plus pertinent que  $P_2$  si  $\pi(P_1) > \pi(P_2)$ .

## Résultats de Maths :

La limite existe :

$$\pi(P) = \lim_{T \rightarrow +\infty} \frac{N_T(P)}{T}$$

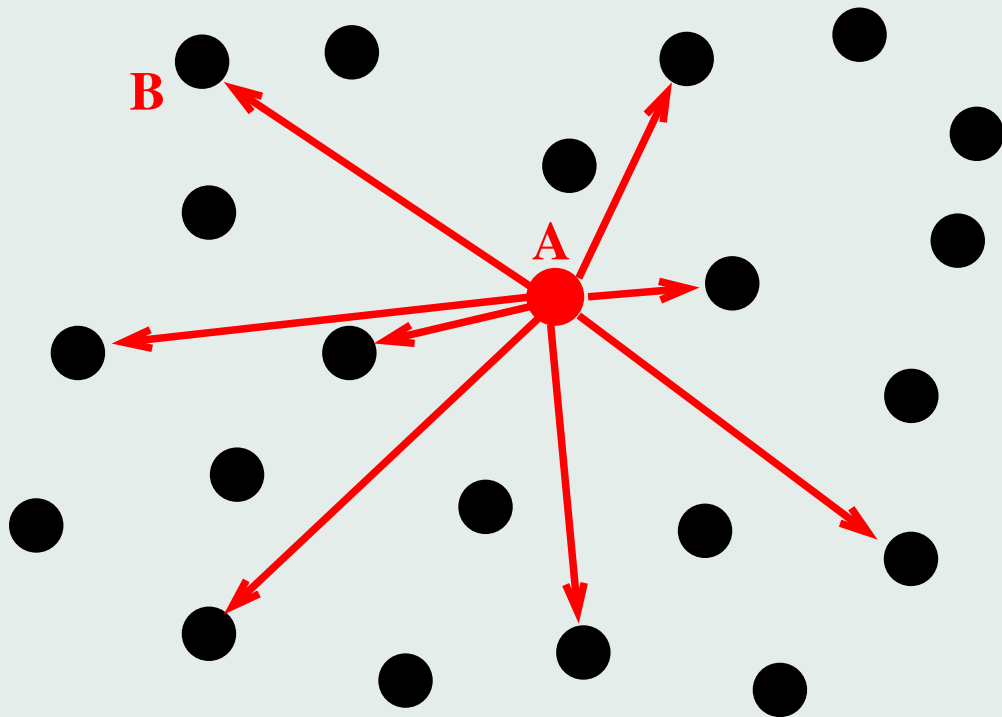
$P_1$  plus pertinent que  $P_2$  si  $\pi(P_1) > \pi(P_2)$ .

- $\pi(P)$  ne dépend pas de  $T$
- $\pi(P)$  ne dépend pas du chemin.
- $\pi(P)$  peut-être calculé par une procédure numérique.

# Réseaux Virtuels : les réseaux sociaux

## Réseaux Virtuels : les réseaux sociaux

- Réseaux par affinité : *A* ami de *B* :  
un lien de *A* vers *B*.



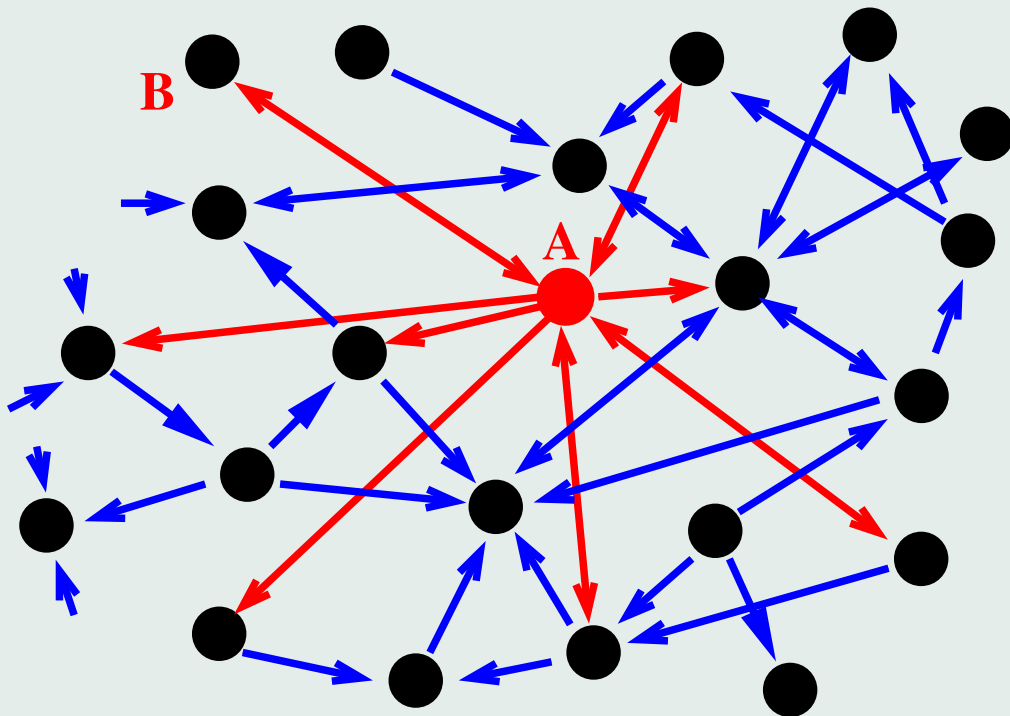
## Réseaux Virtuels : les réseaux sociaux

- Réseaux par affinité : *A* ami de *B* :  
un lien de *A* vers *B*.
- Très grand nombre de nœuds.

Facebook : 175 millions

Myspace : 50 millions.





# Réseaux Virtuels : les réseaux sociaux

Problème :

Comment extraire de l'information de ces réseaux ?

# Réseaux Virtuels : les réseaux sociaux

Problème :

Comment extraire de l'information de ces réseaux ?

Un domaine actif de recherche

- **Data Mining (Fouille de données)**

Algorithmes pour structurer les données.

- **Typologie des Réseaux Sociaux**

Caractérisation/Estimation des graphes

- ...

# Réseaux Virtuels : les réseaux sociaux

Problème :

Comment extraire de l'information de ces réseaux ?

Un domaine actif de recherche

- Data Mining (Fouille de données)

Algorithmes pour structurer les données.

- Typologie des Réseaux Sociaux

Caractérisation/Estimation des graphes

- ...

⇒ Enjeux Économiques



The New York Times

## Internet

WORLD

U.S.

N.Y. / REGION

BUSINESS

TECHNOLOGY

SCIENCE

HEALTH

SPORTS

OPINION

### Search Technology

Go

### Inside Technology

Internet

Start-Ups

Business Computing

Companies

## Facebook's Users Ask Who Owns Information

By BRIAN STELTER

Published: February 16, 2009

Reacting to an online swell of suspicion about changes to [Facebook's](#) terms of service, the company's chief executive moved to reassure users on Monday that the users, not the Web site, "own and control their information."

### Related

Facebook Withdraws

The online exchanges reflected the uneasy and evolving balance between sharing information and

COMMENTS (92)

E-MAIL

SEND TO PHONE

PRINT

REPRINTS

SHARE

ARTICLE TOOLS  
SPONSORED BY

# Les Modèles Économiques des Réseaux

1960 — 1985 IBM.

Machine+programme Informatique spécifique

# Les Modèles Économiques des Réseaux

1960 — 1985 IBM.

Machine+programme Informatique spécifique

1985 — 200? Microsoft.

Programme Informatique généraliste

# Les Modèles Économiques des Réseaux

1960 — 1985 IBM.

Machine+programme Informatique spécifique

1985 — 200? Microsoft.

Programme Informatique généraliste

2000 — ? Google.

Recherche sur Internet.



# Les Modèles Économiques des Réseaux

1960 — 1985 IBM.

Machine+programme Informatique spécifique

1985 — 200? Microsoft.

Programme Informatique généraliste

2000 — ? Google.

Recherche sur Internet.

? — Facebook, Myspace ?

???

## En guise de Conclusion : les tendances

Les informations ne sont plus dans les livres, les ordinateurs mais “sur le réseau”.

## En guise de Conclusion : les tendances

Les informations ne sont plus dans les livres, les ordinateurs mais “sur le réseau”.

Les moyens de calcul sont sur le réseau

## En guise de Conclusion : les tendances

Les informations ne sont plus dans les livres, les ordinateurs mais “sur le réseau”.

Les moyens de calcul sont sur le réseau

L'ordinateur ne sert qu'à avoir accès à Internet.

## En guise de Conclusion : les tendances

Les informations ne sont plus dans les livres, les ordinateurs mais “sur le réseau”.

Les moyens de calcul sont sur le réseau

L'ordinateur ne sert qu'à avoir accès à Internet.

Les données personnelles aussi sont sur le réseau.

## En guise de Conclusion : les tendances

Les informations ne sont plus dans les livres, les ordinateurs mais “sur le réseau”.

Les moyens de calcul sont sur le réseau

L'ordinateur ne sert qu'à avoir accès à Internet.

Les données personnelles aussi sont sur le réseau.

⇒ Enjeux de Société

**FIN**

**FIN**

**Questions ?**