

Cours n°8

TCP: Transmission de Données sur Internet

Mercredi 29 novembre 2006

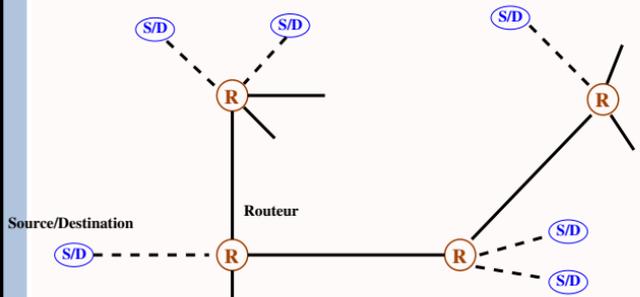
<http://www.cmap.polytechnique.fr/~probert>

Table des matières

1	Introduction	3
2	Algorithmique	20
3	Modélisation	30
4	Quand le taux de perte tend vers 0	38

1. Introduction

Internet : Une vue simplifiée



De Palaiseau à Stanford :

```
1 cmap_router
2 * * *
3 orsay-g5-0-10.cssi.renater.fr
4 nri-c-pos4-0.cssi.renater.fr
5 nri-b-g14-0-0-101.cssi.renater.fr
6 renater.rt1.par.fr.geant2.net
7 so-5-0-0.rt1.lon.uk.geant2.net
8 so-2-0-0.rt1.ams.nl.geant2.net
9 so-7-0-0.rt1.nyc.us.geant2.net
10 198.32.11.50
11 chinnng-nycmng.abilene.ucaid.edu
12 iplsng-chinnng.abilene.ucaid.edu
13 kscynng-iplsng.abilene.ucaid.edu
14 dnrng-kscynng.abilene.ucaid.edu
15 snvang-dnrng.abilene.ucaid.edu
16 losang-snvng.abilene.ucaid.edu
17 hpr-lax-gsr1-abilene-LA-10ge.cenic.net
18 svl-hpr-lax-hpr-10ge.cenic.net
19 hpr-stan-ge-svl-hpr.cenic.net
20 bbr2-rtr.Stanford.EDU
21 www5.Stanford.EDU
```

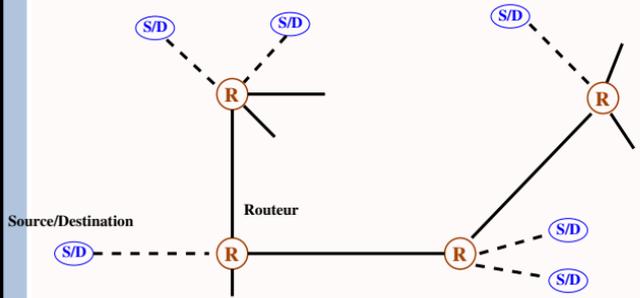
<http://www.cmapx.polytechnique.fr/~probert>

Cours de Majeure 2006

École Polytechnique

Philippe Robert

Internet : Une vue simplifiée



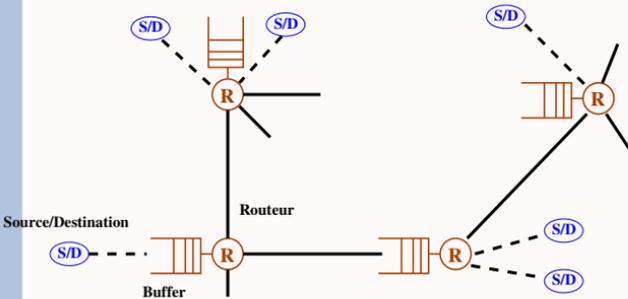
Cours de Majeure 2006

École Polytechnique

Philippe Robert

<http://www.cmapx.polytechnique.fr/~probert>

Internet : Une vue simplifiée



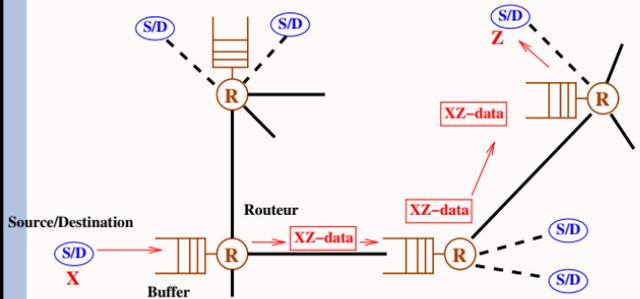
<http://www.cmapx.polytechnique.fr/~probert>

Cours de Majeure 2006

École Polytechnique

Philippe Robert

Internet : Une vue simplifiée



<http://www.cmapx.polytechnique.fr/~probert>

Cours de Majeure 2006

École Polytechnique

Philippe Robert

Entête de paquet TCP/IP

clockstamp				Header
clockstamp		FIFO depth		
ATM header				
LLC/SNAP				IP
LLC/SNAP				
Version IHL	Type of Service	Total Length		
Identification		Flags	Fragment Offset	
Time to Live	Protocol	Header Checksum		
Source Address				TCP
Destination Address				
Source Port		Destination Port		
Sequence Number				
Acknowledgement Number				
Data	Reserved	Window		

<http://www.cam.ac.uk/tech/qedr/~proubet>

Mesures du trafic sur Internet

<http://www.cam.ac.uk/tech/qedr/~proubet>

Cours de Majeure 2006

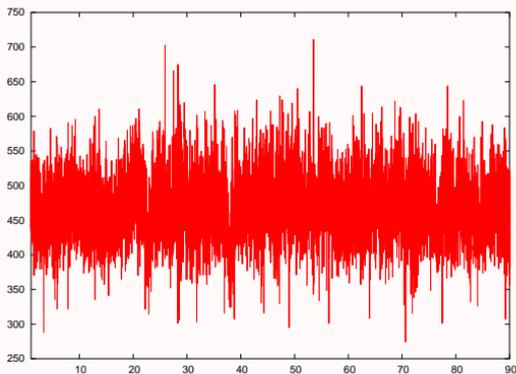
École Polytechnique

Philippe Robert

Cours de Majeure 2006

École Polytechnique

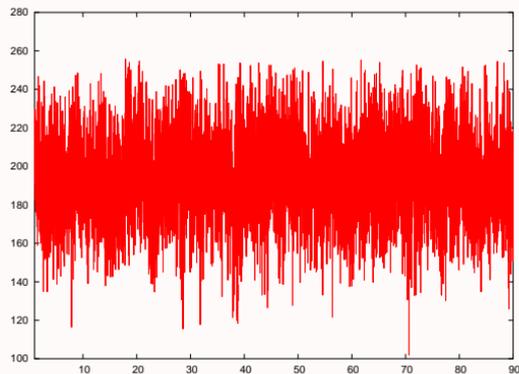
Philippe Robert



Trafic : Nombre de paquets IP par 1/100s

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

<http://www.cam.ac.uk/tech/qedr/~proubet>



Débit : Mb/s

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

<http://www.cam.ac.uk/tech/qedr/~proubet>

Cours de Majeure 2006

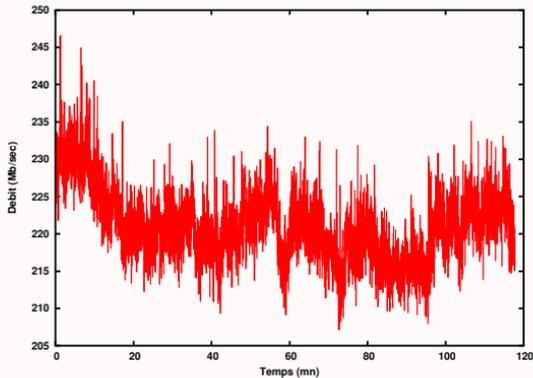
École Polytechnique

Philippe Robert

Cours de Majeure 2006

École Polytechnique

Philippe Robert



Débit d'un trafic ADSL

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

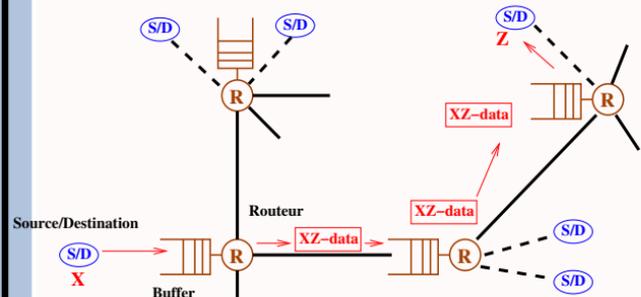
Colorado State Univ. 03/11/2003 à 21:51:26

Application	% nb. paquets
bmc_patrolldb	0.516636
sm-pas-2	0.551968
tabula	0.620219
gnutella-svc	0.701901
ms-olap4	0.842090
e-donkey	0.938023
blubster	0.999990
realtime	1.392920
https	1.537388
nntp	1.579550
www-http	18.956738

Durée : 90 sec.

108865 Stations/2196322 Paquets/ 1.6 Go

Internet : Une vue simplifiée



Internet

Un réseau à commutation de paquets

- Messages divisés en paquets.
- Paquets acheminés individuellement.

Avantages

- Système distribué.
- Flexibilité : Évolution facile.

Problèmes

- Réserveation de ressources.
- Garantie de service. QoS.

<http://www.cmapx.polytechnique.fr/~proubet>

Réseau Téléphonique

Un réseau à commutation de circuits

Connexion de A à B :

- Réserveation de tous les liens reliant A à B .
- Occupation de tout le circuit / connexion.

Avantages

- Réserveation de ressources.
- Garantie de service. QoS.

Problème

- Peu flexible.

<http://www.cmapx.polytechnique.fr/~proubet>

Problématique : Le cadre

- Réseau non fiable.
 - Pas de contrôle global.
 - Système distribué.
-

Pb. : Transmission des données fiable ?

TCP : protocole de transmission de données.

Protocole : ensemble d'algorithmes.

<http://www.cmapx.polytechnique.fr/~proubet>

2. Algorithmique

<http://www.cmapx.polytechnique.fr/~proubet>

Les principes de base de TCP

Principes de Base Cerf and Kahn (1973)

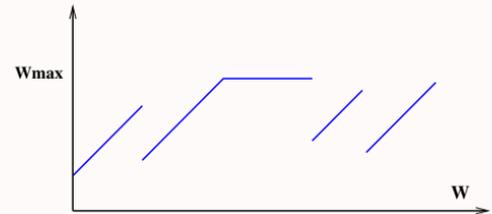
- Acquittements des paquets reçus.
 - Auto-limitation : une source a au plus W paquets dans le réseau.
- W : taille de la fenêtre de congestion.

Contrôle de congestion Jacobson (1987)

- Transmission de W paquets OK :
 $W \rightarrow W + 1$
- Un paquet est perdu : $W \rightarrow W/2$.

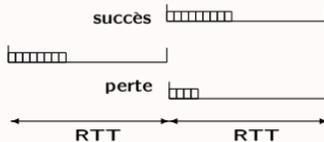
http://www.cmapx.polytechnique.fr/~robert

Évolution de la variable W



http://www.cmapx.polytechnique.fr/~robert

Évolution de la fenêtre de congestion (II)



RTT : Round Trip Time

http://www.cmapx.polytechnique.fr/~robert

Algorithme AIMD

$$W \rightarrow \begin{cases} W + \beta & \text{si pas de perte} \\ \max(\lfloor \delta W \rfloor, 1) & \text{sinon,} \end{cases}$$

$$\delta \in]0, 1[, \beta > 0.$$

Caractéristiques

- Accroissements additifs $\beta = 1$.
- Décroissance multiplicative : $\delta = 1/2$.

http://www.cmapx.polytechnique.fr/~robert

TCP (Transmission Control Protocol)

Quatre algorithmes :

- Slow Start.
- Fast Recovery.
- Fast Retransmit.
- Congestion Avoidance.

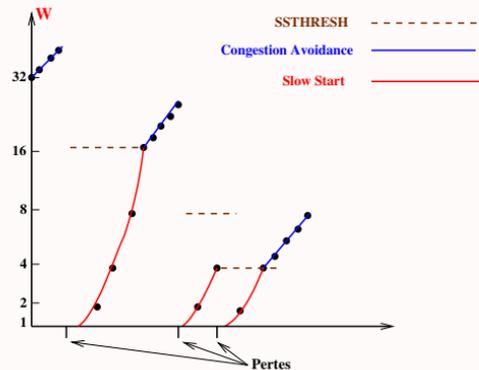
Beaucoup de versions :

Reno, Vegas, Tahoe, Sack, ...

Principes de base inchangés pour l'essentiel.

<http://www.comptel.ch/chaquefr/~probert>

Slow Start



Le problème des pertes

Destination :

Reçoit le paquet $n^o n \Rightarrow$ envoie entier $n + 1$
à la source.

Détection de perte : deux méthodes

A. Source :

réception de trois numéros identiques.

B. Dépassement de délai d'un paquet.

<http://www.comptel.ch/chaquefr/~probert>

Détection de perte

A. Source :

réception de trois numéros identiques.

B. Dépassement de délai d'un paquet.

Récupération d'une perte : $ssthresh \rightarrow W/2$

A. Source : envoi du paquet perdu.

$W \rightarrow W/2.$

B.

$W \rightarrow 1.$

<http://www.comptel.ch/chaquefr/~probert>

Fast Retransmit/Fast Recovery

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
S	E	1	2	3	4	5	6	7	8	4	9	10	11	12	
	R					2	3	4	4	4	4	4	4	9	10
D	R			1	2	3	∅	5	6	7	8	4	9	10	11
	E			2	3	4	4	4	4	4	4	9	10	11	

S : Source. D : Destination.
E : Émission. R : Réception.

3. Modélisation

Le cadre

- Longue connexion TCP.
Uniquement algorithmique
"Congestion Avoidance" étudié.
- Réseau
= boîte noire perdant certains paquets.
Modèle probabiliste des pertes de paquets.

Objectif : Estimation du débit de la connexion en fn du taux de perte.

Évolution asymptotique de la variable W .

Modèle du processus de perte de paquets

Hypothèses

- Un paquet est perdu avec proba $1 - e^{-\alpha}$.
- Les pertes des paquets sont indépendantes.

Discussion sur l'hypothèse d'indépendance.

Réaliste.

Le réseau évolue rapidement \Rightarrow oublié.

Irréaliste.

Les pertes sont groupées donc corrélées.

Solution : Modèle des pertes groupées.

Une chaîne de Markov

W_n^α : taille de la fenêtre de congestion au début du n -ième cycle.

$$W_{n+1}^\alpha = \begin{cases} W_n^\alpha + 1 & \text{pas de perte parmi les } W_n^\alpha \text{ paquets} \\ \lfloor \delta W_n^\alpha \rfloor & \text{sinon,} \end{cases}$$

$$\delta \in]0, 1[.$$

Indépendance des pertes

$\Rightarrow (W_n^\alpha)$ chaîne de Markov sur \mathbb{N} .

Une chaîne de Markov (II)

Transitions

$$\mathbb{P}(W_1^\alpha = x + 1 \mid W_0^\alpha = x) = e^{-\alpha x}$$

$$\mathbb{P}(W_1^\alpha = \lfloor \delta x \rfloor \mid W_0^\alpha = x) = 1 - e^{-\alpha x}$$

Irréductible :

Pour $x \in \mathbb{N}$, $\mathbb{P}(x \rightarrow 0) > 0$ et $\mathbb{P}(0 \rightarrow x) > 0$

Une chaîne de Markov (III)

(W_n^α) ergodique : Si $W_0 = x > 0$

$$\mathbb{E}(W_1^\alpha \mid W_0^\alpha = x) = (x + 1)e^{-\alpha x} + \lfloor \delta x \rfloor (1 - e^{-\alpha x})$$

$$\mathbb{E}_x(W_1^\alpha - W_0^\alpha) = e^{-\alpha x} + (\lfloor \delta x \rfloor - x)(1 - e^{-\alpha x})$$

$$\lim_{x \rightarrow +\infty} \frac{1}{x} \mathbb{E}_x(W_1^\alpha - W_0^\alpha) = \delta - 1 < 0$$

Donc $\exists x_0 : x \geq x_0 \Rightarrow \mathbb{E}_x(W_1^\alpha - W_0^\alpha) \leq -1$

Foster : (W_n^α) ergodique.

Asymptotique de la connexion TCP

(W_n^α) ergodique : $W_n^\alpha \xrightarrow{\mathcal{L}} W_\infty^\alpha$

Débit moyen : $\lim_{n \rightarrow +\infty} \mathbb{E}(W_n^\alpha) = \mathbb{E}(W_\infty^\alpha)$

(Convergence en loi des C.M.)

Sessions où le débit est $\geq d$: \mathbb{P} -p.s.

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^n \mathbf{1}_{\{W_k^\alpha \geq d\}} = \mathbb{P}(W_\infty^\alpha \geq d)$$

(Théorème ergodique pour les C.M.)

Le problème de la probabilité invariante

Si π^α proba. invariante de (W_n^α) alors

$$\pi_n^\alpha = e^{-\alpha(n-1)} \pi_{n-1}^\alpha + \sum_{k=\lceil n/\delta \rceil}^{\lceil (n+1)/\delta \rceil - 1} (1 - e^{-\alpha k}) \pi_k^\alpha$$

Pas de solution connue à cette équation.

<https://www.compari.fr/technique/fr/~probert>

4. Quand le taux de perte tend vers 0

<https://www.compari.fr/technique/fr/~probert>

La chaîne de Markov aux instants de perte

V_n^α : Taille de la fenêtre de congestion après la n -ième perte.

Si $V_0^\alpha = x$, alors

$$V_1^\alpha = \lfloor \delta (x + G_x^\alpha) \rfloor$$

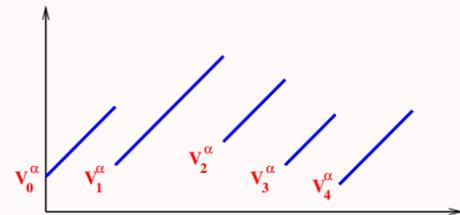
G_x^α nb de fenêtres transmises sans perte.

$$\mathbb{P}(G_x^\alpha \geq m) = \exp(-\alpha(mx + m(m-1)/2)).$$

(V_n^α) est une chaîne de Markov ergodique.

<https://www.compari.fr/technique/fr/~probert>

La chaîne de Markov incluse



<https://www.compari.fr/technique/fr/~probert>

Convergence en loi des transitions

Si $W_0 = x$:

G_x^α nb de fenêtres transmises sans perte.

Proposition. Qd $\alpha \rightarrow 0$

$$\sqrt{\alpha} G_{[x/\sqrt{\alpha}]}^\alpha \xrightarrow{\text{Loi}} \bar{G}_x$$

avec

$$\mathbb{P}(\bar{G}_x \geq y) = \exp(-xy - y^2/2),$$

Convergence en loi des transitions (II)

– $\bar{G}_0 \stackrel{\mathcal{L}}{=} \sqrt{2E_1}$ avec E_1 de loi exp. par. 1.

– \bar{G}_0 : loi de Rayleigh.

– $x + \bar{G}_x \stackrel{\mathcal{L}}{=} (\bar{G}_0 \mid \bar{G}_0 \geq x)$

Conséquence sur (V_n^α)

$$V_0^\alpha = [x/\sqrt{\alpha}]$$

$$\begin{aligned}\sqrt{\alpha} V_1^\alpha &= \sqrt{\alpha} \left[\delta \left([x/\sqrt{\alpha}] + G_{[x/\sqrt{\alpha}]}^\alpha \right) \right] \\ \sqrt{\alpha} V_1^\alpha &\sim \delta \left(\sqrt{\alpha} [x/\sqrt{\alpha}] + \sqrt{\alpha} G_{[x/\sqrt{\alpha}]}^\alpha \right) \\ &\sim \delta \left(x + \bar{G}_x \right)\end{aligned}$$

Si $\sqrt{\alpha} V_0^\alpha \sim x \stackrel{\text{def.}}{=} \bar{V}_0$ alors

$$\bar{V}_1 \stackrel{\text{def.}}{=} \sqrt{\alpha} V_1^\alpha \sim \delta(\bar{V}_0 + \bar{G}_{\bar{V}_0})$$

Conséquence sur (V_n^α) (II)

Proposition

Si $\lim_{\alpha \rightarrow 0} \sqrt{\alpha} V_0^\alpha = \bar{v}$, alors $(\sqrt{\alpha} V_n^\alpha)$ converge vers la chaîne de Markov (\bar{V}_n) , telle que $\bar{V}_0 = \bar{v}$ et

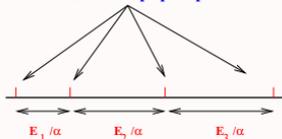
$$\bar{V}_{n+1} \stackrel{\mathcal{L}}{=} \delta \left(\bar{V}_n + \bar{G}_{\bar{V}_n} \right),$$

avec, pour $x \geq 0$,

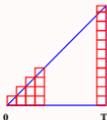
$$\mathbb{P}(\bar{G}_x \geq y) = e^{-(xy+y^2/2)}.$$

Pourquoi $1/\sqrt{\alpha}$?

Nombres des paquets perdus



E_1, E_2, \dots v.a. i.i.d. exp. de paramètre 1.



Première perte : $T^2/2 \sim E_1/\alpha$

$$T \sim \frac{\sqrt{2E_1} \frac{L}{\alpha}}{\sqrt{\alpha}} = \frac{\bar{G}_0}{\sqrt{\alpha}}$$

http://www.cmapx.jpl.technique.fr/~probert

Diagramme de convergence

$$\begin{array}{ccc} (\sqrt{\alpha} V_n^\alpha) & \xrightarrow{n \rightarrow +\infty} & \sqrt{\alpha} V_\infty^\alpha \\ \alpha \rightarrow 0 \downarrow & & \downarrow ? \\ (\bar{V}_n) & \rightarrow & \bar{V}(\infty) \end{array}$$

Théorème

Le diagramme commute.

Proba invariante de (\bar{V}_n)

\Rightarrow comportement de $(\sqrt{\alpha} V_\infty^\alpha)$

http://www.cmapx.jpl.technique.fr/~probert

Loi invariante asymptotique

loi de $\sqrt{\alpha} V_\infty^\alpha$: proba invariante de $(\sqrt{\alpha} V_n^\alpha)$.

$$\mathbb{E}(f(\sqrt{\alpha} V_\infty^\alpha)) = \frac{1}{\mathbb{E}_{\pi_K^\alpha}(\bar{T}_K)} \mathbb{E}_{\pi_K^\alpha} \left(\sum_{k=0}^{T_K^\alpha - 1} f(\sqrt{\alpha} V_k^\alpha) \right)$$

– $T_K^\alpha = \inf\{n > 0 : \sqrt{\alpha} V_n^\alpha \leq K\}$.

– π_K^α proba invariante des passages
de $(\sqrt{\alpha} V_n^\alpha)$ dans $[0, K]$.

http://www.cmapx.jpl.technique.fr/~probert

La chaîne asymptotique

$$\mathbb{E}(f(\bar{V}_\infty)) = \frac{1}{\mathbb{E}_{\pi_K}(\bar{T}_K)} \mathbb{E}_{\pi_K} \left(\sum_{k=0}^{\bar{T}_K - 1} f(\bar{V}_k) \right)$$

– $\bar{T}_K = \inf\{n > 0 : \bar{V}_n \leq K\}$.

– π_K proba invariante des passages
de (\bar{V}_n) dans $[0, K]$.

http://www.cmapx.jpl.technique.fr/~probert

Preuve de la convergence de V_∞^α vers \bar{V}_∞

1. $\pi_K^\alpha \xrightarrow{n \rightarrow +\infty} \bar{\pi}_K,$

2.

$$\lim_{\alpha \rightarrow 0} \mathbb{E}_{\pi_K^\alpha} \left(\sum_{k=0}^N f(V_k^\alpha) \right) = \mathbb{E}_{\bar{\pi}_K} \left(\sum_{k=0}^N f(\bar{V}_k) \right)$$

3.

$$\lim_{C \rightarrow +\infty} \sup_{\alpha > 0} \mathbb{E} \left(T_K^\alpha \mathbf{1}_{\{T_K^\alpha > C\}} \right) = 0$$