

# Dimensionnement et Évaluation de Performances des réseaux

## Chapitre 1 : Introduction

Florence Perronnin

Université Joseph Fourier, LIG Laboratory

January 18, 2017

## Plan

- 1 Organisation pédagogique
- 2 Introduction
  - But
  - Problématiques
- 3 Méthodologie
  - Méthodologie
- 4 Qualité de service
  - QoS et SLA
  - Métriques
  - Choix d'une métrique
  - Caractérisation

## Compétences visées (et évaluées...)

- **Modéliser** un problème d'évaluation de performance
- **Concevoir et développer** un outil qui produit des indicateurs de performance
- **Analyser** les résultats obtenus
- Savoir **critiquer** une étude de performance

## Organisation pédagogique

### Team

Cours : Florence Perronnin

TD : Arnaud Legrand

- 1 Introduction et rappels
- 2 Expérimentation
- 3 Modélisation
- 4 Simulation numérique
- 5 Chaînes de Markov à temps discret
- 6 Modèles de trafic
- 7 Chaînes de Markov à temps continu
- 8 Files d'attente classiques
- 9 Réseaux de files d'attente

## Ce qui est attendu de vous

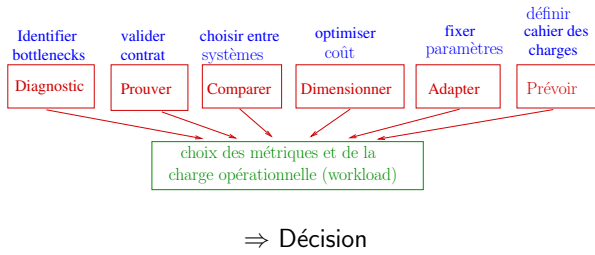
- Pré-requis: Probabilités et Simulation
- Attitude :
  - ▶ Poser des questions (que vous ayez compris ou non...)
  - ▶ **Curiosité**
  - ▶ Esprit critique
- Évaluation :
  - ▶ CC = Quicks + Projet (15-20h min) : 50%
  - ▶ Examen : 50% (*coefficients à confirmer*)

## L'évaluation de performances, pourquoi faire?

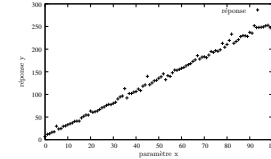
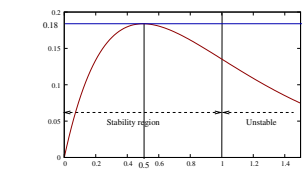
### Exemples

- 1 De quel débit fixe a-t-on besoin sur un réseau étendu pour garantir un temps de réponse inférieur à 100ms?
- 2 Prouvez que votre algorithme est plus efficace que celui actuellement utilisé.
- 3 Quels sont les goulots d'étranglement des réseaux 802.11p?
- 4 Une application parallèle prend 1h sur un cluster de 10 machines. Pour diviser ce temps par deux, faut-il doubler le nombre de machines?

## Questions typiques



## Buts possibles d'une étude

- Tracer une **courbe de réponse**: 
- Détecter un **effondrement de congestion**: 
- Identifier des **goulots d'étranglement** (bottlenecks)
- Comparer deux produits
- Savoir où placer ses ressources (**paradoxe de Braess**)

Parmi d'autres!

## Point de vue

### Quantifier la performance

Choix d'un critère à quantifier : **métrique** ... et d'une représentation statistique pertinente (moyenne, pire cas...)

### Exemple

Une compagnie de bus fait partir 3 bus par heure. Quel est le temps moyen d'attente?

- Pour la compagnie : 10mn
- Pour le passager : dépend des heures de départ

Les bus partent à :  $h$ ,  $h + 5mn$ ,  $h + 20mn$ . **Temps moyen d'attente** :

$$\mathbb{E}[W] = \frac{5}{60} \cdot 2.5' + \frac{15}{60} \cdot 7.5' + \frac{40}{60} \cdot 20' = 15'25$$

Compromis performance / efficacité

## Charge opérationnelle

### Problème

Dimensionner le buffer d'un serveur:

- débit serveur  $\mu$
- taux d'arrivée des requêtes  $\lambda$



On suppose que la charge  $\rho = \lambda/\mu = 1$ . Quelle doit être la taille  $K$  du buffer?

- arrivées et départs déterministes:  $K = 1$  suffit.
- arrivées et départs exponentiels:  $K = +\infty!$
- probabilité de perte maximum?

## Facteurs externes

### Problème

Mesurer l'efficacité d'un nouvel algorithme  $A$  et le comparer à l'ancien algorithme  $B$ .

### Idée

Faire l'expérience:

- même problème (mêmes paramètres d'entrée)
- même machine

Lancer l'algorithme  $A$  et mesurer le temps d'exécution  $\tau_A$ . Puis lancer l'algorithme  $B$  et mesurer  $\tau_B$ .

### Facteurs externes

- charge du processeur variable
- occupation carte réseau

## 3 méthodes

- Mesure expérimentale
- Simulation
- Analyse théorique

## Caractéristiques de la première méthode

### Mesure de systèmes réels

- générateur de charge (ou sélection)
- instrumentation parfois coûteuse
- immobilisation du système
- plan d'expériences
- interprétation
- crédibilité

## Caractéristiques de la deuxième méthode

### Simulation

- Durée de simulation
- fiabilité du programme (bugs)
- degré de précision
- pertinence du modèle simulé
- qualité du générateur aléatoire!
- analyse des résultats

## Caractéristiques de la troisième méthode

### Analyse théorique

- Abstraction des comportements représentatifs (modèle simplifié)
- Résultats ciblés
- Exemple:
  - ▶ temps moyen de réponse
  - ▶ temps d'inoccupation
  - ▶ probabilité d'attendre plus de 10s...
- Coût minime!

## Comparaison des 3 méthodes

Critère	Analyse	Simulation	Mesure
étape de développement	∇	∇	après proto
durée	faible	moyenne	longue
outils	cerveau	programmation	instrumentation
précision	faible	moyenne	variable
comparaison	facile	moyenne	difficile
tradeoff	facile	moyen	difficile
coût	faible	moyen	élevé
crédibilité	faible	moyenne	élevée

### Validation

Quelle que soit la méthode choisie, ses résultats doivent être validés!

## Diagnostic

### Problème [Le Boudec]

On considère un e-kiosque vendant des vidéos à la demande à des PDA sur un réseau WIFI. Le gérant veut vérifier le passage à l'échelle de son installation.

- 1 Première expérience: mesure du débit en fonction de la charge.
- 2 Deuxième expérience: ajouter une station de base et refaire les mesures.
- 3 Troisième expérience: mesure du nombre de collisions
- 4 Quatrième expérience: augmenter la mémoire du serveur et refaire les mesures.

## Présentation de résultats

On considère les débits mesurés suivants:

Système	expérience 1	expérience 2
A	20	10
B	10	20

Quel est le meilleur système?

### Égalité

Débit moyen = 15

### B gagne

ratio par rapport à A :

Sys	exp 1	exp 2	moyenne
A	1	1	1
B	0.5	2	1.25

## 10 steps [Jain 1991]

- 1 State the goals of the study and define system boundaries.
- 2 List system services and all possible outcomes.
- 3 Carefully choose performance metrics.
- 4 List system and workload parameters.
- 5 Select factors and their values.
- 6 Select evaluation techniques.
- 7 Select the workload.
- 8 Design the experiments.
- 9 Analyze and interpret the data.
- 10 Present the results. Start over, if necessary.

## Common Mistakes [Jain 1991]

- 1 No goals: know what you're looking for
- 2 Biased goal: show my algorithm is better than my challenger's
- 3 unsystematic approach: vary parameters randomly, on the fly
- 4 wrong or biased workload: consider real-life conditions
- 5 wrong technique: overly complex model, costly or marginal experiments, unreliable simulation
- 6 Overlooking factors: some hidden parameters may tamper the results
- 7 ignoring parameters: e.g. quantify memory for cpu-bound problem
- 8 Improper experimental design: use a methodology
- 9 Don't, or wrongly, analyze results: e.g. tests over 13 subjects
- 10 No confidence intervals
- 11 Ignoring outliers
- 12 Ignoring variability don't dimension data centers with mean load
- 13 Bad presentation of results see checklist for good graphics
- 14 Omitting/hiding assumptions: just don't!

## Quelle métrique?

### Exemple

- 1 Un fournisseur d'accès met en place une hotline facturée à la minute. Les métriques qui l'intéressent sont:
  - ▶ la durée moyenne d'attente (élevée)
  - ▶ la durée moyenne de service (faible)
  - ▶ le nombre moyen de clients
 Les métriques qui intéressent le client sont:
  - ▶ le temps de réponse (résolution du problème)
- 2 Un serveur de e-commerce est mis en ligne. Le commerçant regarde:
  - ▶ le nombre de transactions par jour (débit)
  - ▶ la disponibilité du service
  - ▶ le taux d'abandon
 Le client regarde:
  - ▶ le temps de réponse

Ces critères peuvent parfois être **contradictoires**.

## Qualité de Service

### Définition

Ensemble d'exigences de qualité sur le comportement global d'un système.

Un accord contractuel de QoS entre fournisseur et client s'appelle un **Service Level Agreement (SLA)**.

### Exemples

- Frame Relay : CIR et EIR
- ATM : CBR, VBR
- Diffserv and RED
- Intserv (RSVP)

Dans ce cas l'évaluation de performances intervient à différents niveaux:

- conformité du trafic (lissage, marquage)
- conformité du service

## Métriques classiques

### Temps de réponse (response time)

Temps écoulé entre l'émission de la demande de service et la terminaison de celui-ci. ex: temps de réponse au clic.

### Débit (throughput)

Quantité de travail fourni par le système en une unité de temps.

notion très variable selon le contexte

ex: débit d'une fibre optique (Mbps). Débit d'un péage (véhicule/mn). Débit utile. Débit crête.

## Métriques classiques (suite)

### Probabilité de perte/d'erreur (loss/error rate)

Probabilité, ou taux d'occurrence, d'une indisponibilité de service ou d'un service inapproprié. ex: probabilité de blocage d'un appel GSM (saturation). Probabilité d'erreur bit sur un réseau sans fil.

### Taux d'utilisation (utilization)

Pourcentage moyen de temps pendant lequel le système est actif. ex: taux d'utilisation de CPU. Taux d'utilisation des transports en commun. Taux d'utilisation d'un lien OC3.

### Gigue (jitter)

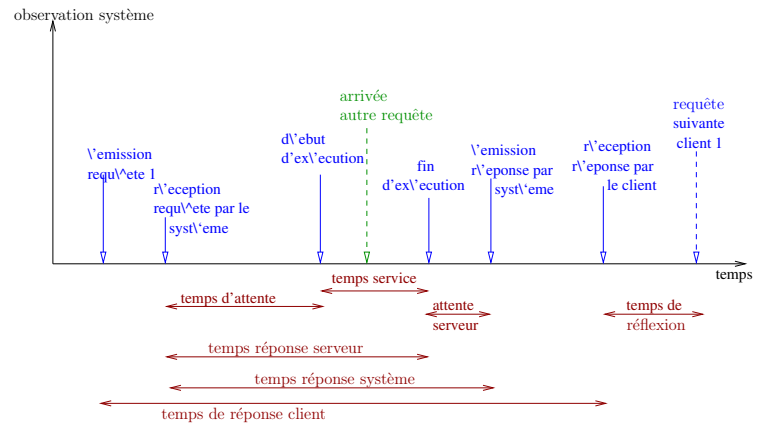
Variation des délais de transmissions. ex: bufferisation des flux temps-réel pour lisser les délais.

## Métriques classiques (suite)

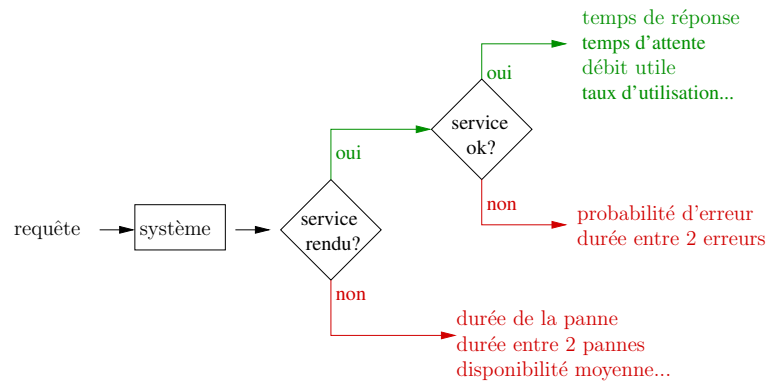
**Stabilité**  
 Un système est stable s'il peut atteindre un régime permanent (stationnaire). **Préalable à l'étude d'autres métriques!**

**Équité**  
 En cas d'utilisateurs concurrents d'un système, l'équité est une façon de vérifier qu'aucun utilisateur ne puisse être "jaloux" d'un autre. Il en existe plusieurs définitions : *max-min fairness*, *proportional fairness*, *alpha-fairness*...

## Temps de réponse



## Choix d'une métrique



## Description (statistique) d'une métrique

- Valeurs extrêmes (*pire cas; débit max...*)
- Condition de stabilité (*charge max*)
- Valeur moyenne (*pour une charge donnée*)
- Variabilité
- Distribution, courbe de réponse