Introduction à la théorie des files d’attente

C. Rigault (COFORTIC)
Ancien directeur d'études à Telecom-ParisTech
claude.rigault@cofortic.com
http://www.cofortic.com

Sommaire

- Prologue
- Qu’est ce que le trafic ?
- Le nombre de clients dans un système
- Petits rappels sur les probabilités
- Les processus des clients
- Les processus à perte
- Les processus à attente
- Les réseaux de files d’attente
- Simulation

Prologue

Aucune mathématique difficile

- Ce cours est destiné à donner à des praticiens des outils pour dimensionner les systèmes qu'ils conçoivent.
- Ce n'est pas un cours de mathématiques, même s'il y a des équations à toutes les pages, ces équations sont élémentaires.
- Elles ne servent qu'à comprendre les résultats.
- L'accent est mis sur les concepts et leurs significations

A peu près tout ce qu'il faut savoir en mathématiques (1)

- Une somme infinie de termes positifs non nuls peut être finie.
- C'est ce que n'avait pas pu admettre le grand philosophe Parménide dans le fameux paradoxe d'Achille et de la Tortue.
- En particulier, vous n'aurez aucun mal à montrer que :

$$\left(1 + x + x^2 + x^3 + \ldots + x^N\right)(1 - x) = 1 - x^{N+1}$$

- Donc $$\sum_{i=0}^{N} x^i = \frac{1 - x^N}{1 - x}$$

- Et si $x<1$, lorsque N tend vers l'infini : $\boxed{\sum_{i=0}^{\infty} x^i = \frac{1}{1 - x}}$ C'est probablement la série la plus célèbre.
- En annexe on trouvera d'autres séries déduites de celle là.

A peu près tout ce qu'il faut savoir en mathématiques (2)

- Une autre série est à connaître :

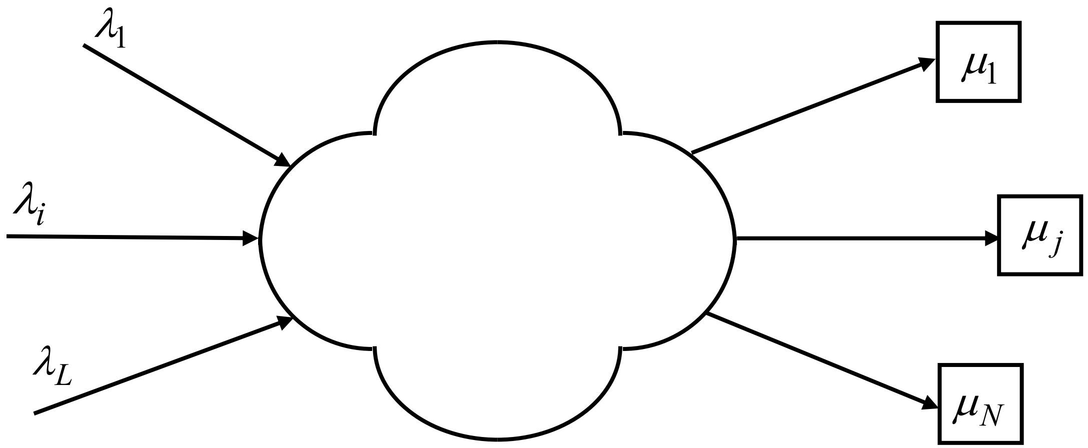
$$\sum_{i=0}^{\infty} \frac{x^i}{i!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + ... = e^x$$

- On obtient ce résultat en effectuant le développement de Taylor de e^x autour de $x=0$

Qu’est ce que le trafic ?

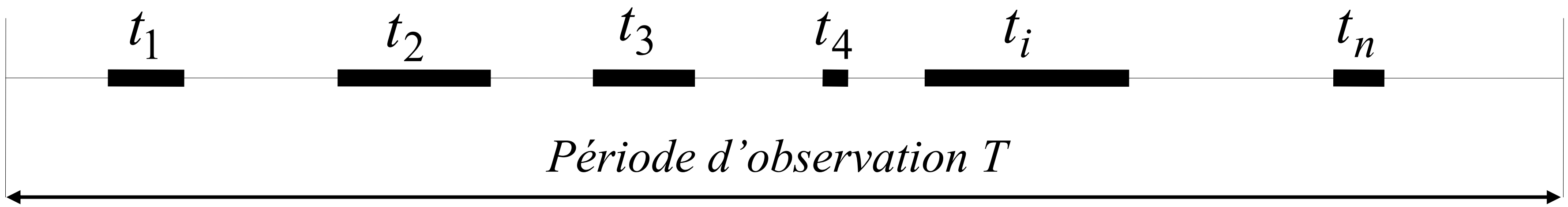
Clients et serveurs

- L clients peuvent être servis par N serveurs. Le but du jeu est que N<L
- Les clients « prennent » (seizure) des serveurs.
- Les « clients » peuvent être des humains, des appels ou des paquets
- Les serveurs peuvent être des guichets, des canaux de transmission ou des récepteurs (modems, machines de traitement), ou … des humains.



Le trafic d’une machine (1)

- Nous appelons ***« machine »*** tout organe ou tout intervenant qui a la propriété d’être soit libre (disponible) soit occupé (indisponible).
- Le trafic de la machine (ou sa charge) ρ est la mesure de son occupation : ρ est la fraction du temps où la machine est occupée. C’est donc la probabilité temporelle d’occupation de cette machine.



$$\rho = \frac{t_{cum}}{T} = \frac{\sum t_i}{T}$$

Le trafic d'une machine (2)

- Si en **« moyenne »** on a ***n*** « prises » de la machine par heure d'une durée **« moyenne »** $\boldsymbol{\tau}$: $t_{cum} = n\tau$

- Le trafic est : $\boxed{\rho = \frac{n\tau}{T}}$

- On définit le taux de prises de machines (nombre de prises par unité de temps) $\lambda = \frac{n}{T}$

- Le taux de service de la machine (nombre maximal de services par unité de temps) $\mu = \frac{1}{\tau}$

- Le trafic est donc également : $\boxed{\rho = \frac{\lambda}{\mu}}$ ou encore : $\boxed{\rho = \lambda\tau}$

Unités de trafic

- L'unité de trafic correspond à 3600 secondes d'occupation cumulée par heure.
  Dans ce cas : $\rho = \frac{t_{cum}}{T} = \frac{3600}{3600} = 1$

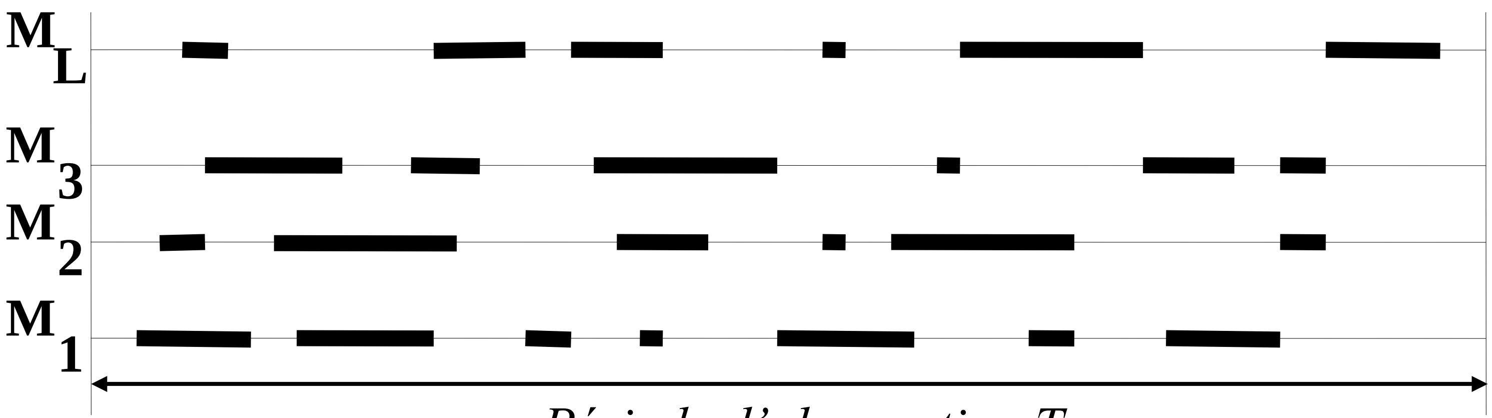
- L'occupation permanente représente l'unité de trafic. Elle est appelée 1 Erlang, en honneur de Agner Krarup Erlang (1878, 1929), un ingénieur danois pionnier du calcul de trafic.

- Les Américains utilisent la CCS (Cent Call Second) :
  1 CCS = 100 seconds d'occupation cumulée par heure
- 1 Erlang = 36 CCS

Le trafic d'un groupe de machines

- Le trafic du groupe est donné par : $A = \sum_M \left(\frac{t_{cum}}{T} \right) = \sum_M \left(\sum_i \frac{t_{cum_i}}{T} \right) = \sum_M \rho_M$
- Ou : $\boxed{A = \frac{n\tau}{T}}$ ou $\boxed{A = \lambda\tau}$ ou $\boxed{A = \frac{\lambda}{\mu}}$

- Dans ces formules λ est la somme des taux d'arrivée de toutes les machines et μ est le taux de service pour **une** machine



Période d'observation T

Remarques sur le trafic :

- Le trafic est aussi le nombre d'arrivées pendant le temps d'un service .
- En effet, par unité de temps le nombre de prises est : $\lambda = \dfrac{n}{T}$
- Pendant le temps τ d'un service le nombre d'arrivées est : $\lambda\tau = \dfrac{n\tau}{T} = A$
- Le trafic d'un groupe de machines n'est pas une probabilité
- Pour N machines, on a : $A \leq N$

COFORTIC

Probabilités temporelles et Probabilités d'ensemble

- On peut dire qu'une probabilité est une fraction du poids d'un ensemble.
- La notion de probabilité est donc relative à un ensemble.
- Si l'ensemble est le temps, la probabilité est dite ***Temporelle***
- Par exemple la probabilité temporelle de l'état x est $P_T(x) = \frac{t_x}{T}$ où t_x est le temps cumulé d'existence de x pendant la période d'observation T
- Si l'ensemble est la totalité des N choses qui peuvent avoir l'état x en question, la probabilité est dite d'***Ensemble***
- Par exemple la probabilité d'ensemble de l'état x est $P_E(x) = \frac{n_x}{N}$ où n_x est le nombre des choses ayant l'état x par rapport au nombre total N de choses

Ergodicité

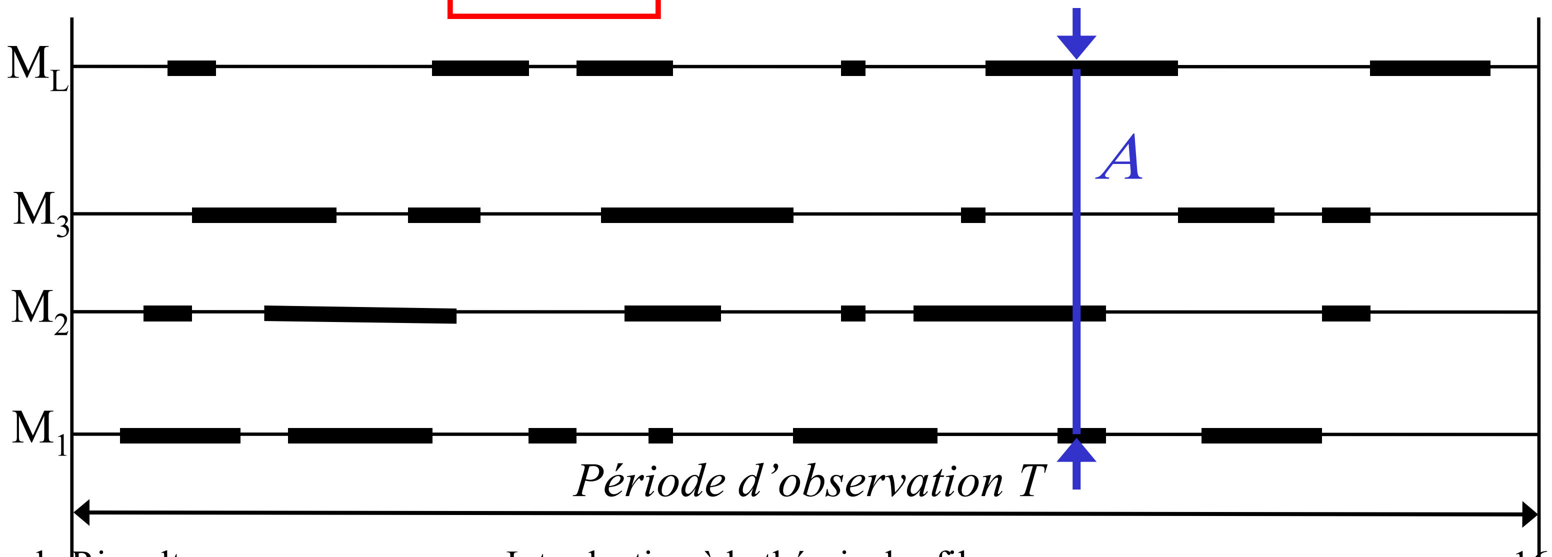
- On ne peut pas comparer des choses de natures différentes (réverbères et torchons) mais ont peut établir des relations (la quantité Q de torchons pour nettoyer N réverbères est $Q=5N$ par exemple)

- On dit qu'un système est ***ergodique*** quand les probabilités *temporelles* sont égales aux probabilités d'*ensemble*

- L'ergodicité a quelque chose à voir avec la stationnarité mais les deux notions sont différentes :
  * Un système ergodique est stationnaire
  * L'inverse n'est pas toujours vrai : un système stationnaire peut ne pas être ergodique

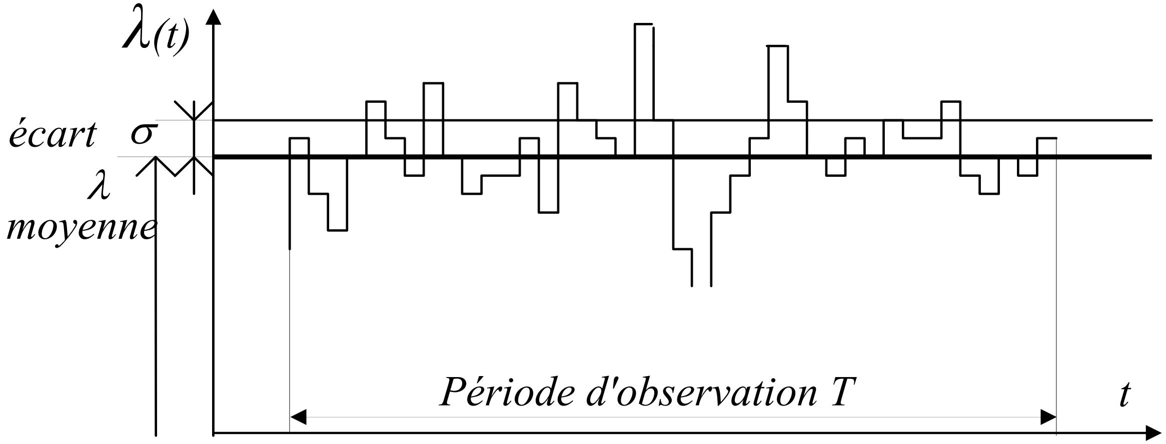
Trafic Ergodique

- Si le trafic est ***ergodique*** les probabilités *temporelles* sont égales aux probabilités d'*ensemble* $P_T = P_E$
- Dans ce cas : $P_T = \rho = P_E = \dfrac{\bar{x}}{L} \Rightarrow \bar{x} = L\rho = A$
- Alors ***le nombre moyen de machines simultanément occupées est égal au trafic*** : $\boxed{\bar{x} = A}$



Arrivées aléatoires des paquets (1)

- $\lambda(t)$ est le nombre de paquets arrivant par seconde entre *t et t+dt*
- τ est le temps moyen de traitement (émission) d'un paquet : $\tau = \dfrac{1}{\mu}$



Arrivées aléatoires des paquets (2)

- S'il arrive pendant une courte période plus de μ unités de données par secondes, il faut pouvoir retenir les unités de données de la pointe au-dessus de μ jusqu'au prochain creux pour pouvoir les écouler.
- Le nombre maximum d'unités de données à mémoriser dans les tampons est donc lié à la taille des pointes de la fonction aléatoire ***x(t)***.

- Normalement la probabilité d'avoir des pointes supérieures à ***k*** fois l'écart type σ est inférieure à un nombre petit ***ε(k)*** dépendant de ***k***.
- Inégalité de Bienaymé Tchebichev :

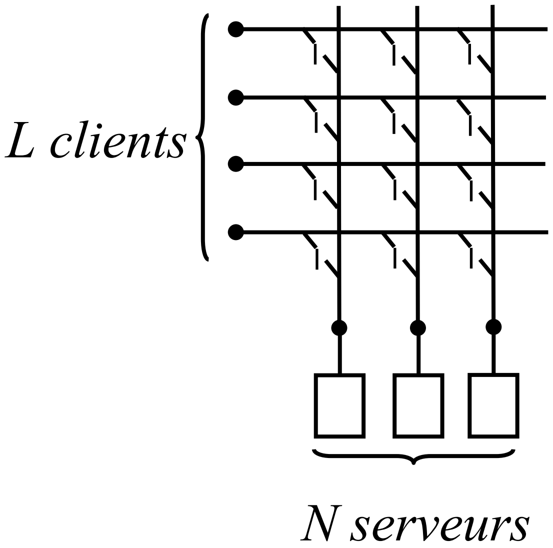
$$P\left(x > \bar{x} + k\sigma\right) \leq \frac{1}{k^2}$$

Arrivées aléatoires des paquets (3)

- Si nous dimensionnons la mémoire $\mathbf{M_i}$ à une taille $||M_i||$ égale à $k\sigma$ unités de données, il existe une probabilité B (inférieure à $\varepsilon(k)$) pour qu'une pointe de trafic amène dans la mémoire tampon un nombre d'unités de données supérieur à $||M_i||$.
- Dans ce cas, toutes les unités de données ne pourront pas être mémorisées et certaines unités de données seront perdues. Cette probabilité est la probabilité de perte.
- ***Les réseaux paquets perdent des paquets***

Arrivées aléatoires des appels

- L clients (abonnés) peuvent être servis par N récepteurs ou N canaux. Le but du jeu est que $N<L$.
- Les clients « prennent » (seizure) des serveurs.

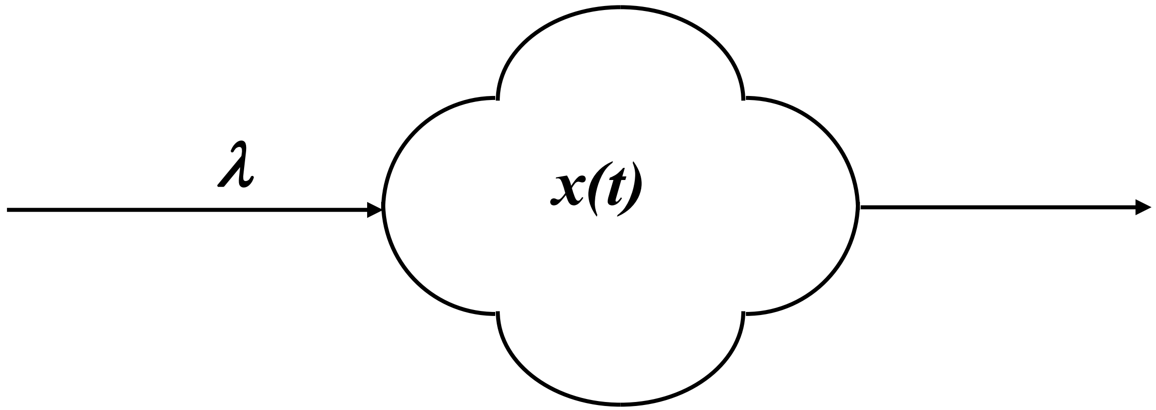


- ***Les réseaux à commutation de circuits perdent des appels***.

Le nombre de clients dans un système

Nombre de clients dans un système

- Taux d'arrivée des clients : λ
- Nombre de clients dans le système au moment t : $x(t)$
- Nombre de clients arrivés entre 0 et T : $n(T)$

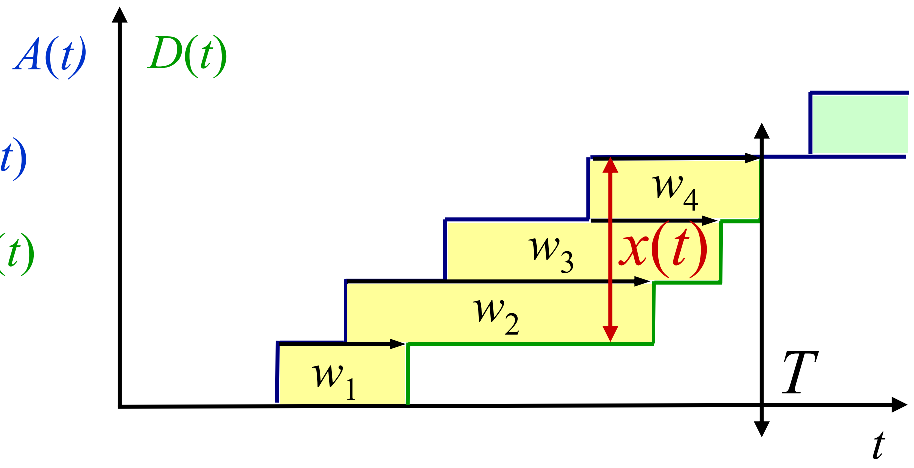


Arrivées et départs

- Le nombre x(t) de clients dans le système est la différence entre le processus d'arrivée et le processus de départ

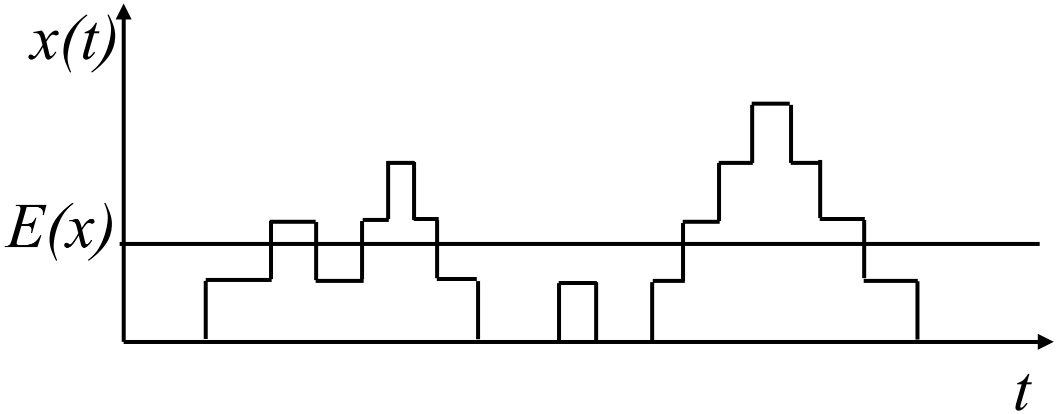
Processus d'arrivée $A(t)$

Processus de départ $D(t)$



Le trafic arrive par bouffées

- Le nombre de clients *x(t)* dans le système est une variable aléatoire. Sa valeur instantanée est très différente de sa valeur moyenne (son espérance *E(x)*). En particulier ce nombre retombe de temps en temps à zéro.
- Il en résulte que le raisonnement sur les moyennes est très délicat. Il induit à raisonner comme si les temps de service étaient constants (on dit Déterministes) ce qui conduit à des résultats faux



Moyennes (Espérances)

- Le nombre moyen de clients dans le système :

$$E(x) = \lim_{T \to \infty} \frac{1}{T} \int_0^T x(t)dt$$

- Le taux moyen d'arrivée est : $\lambda = \lim_{T \to \infty} \frac{n(T)}{T}$

- Le temps moyen dans le système est : $E(w) = \lim_{T \to \infty} \frac{\sum_{i=0}^{n(T)} w_i}{n(T)}$

Des probabilités de natures différentes

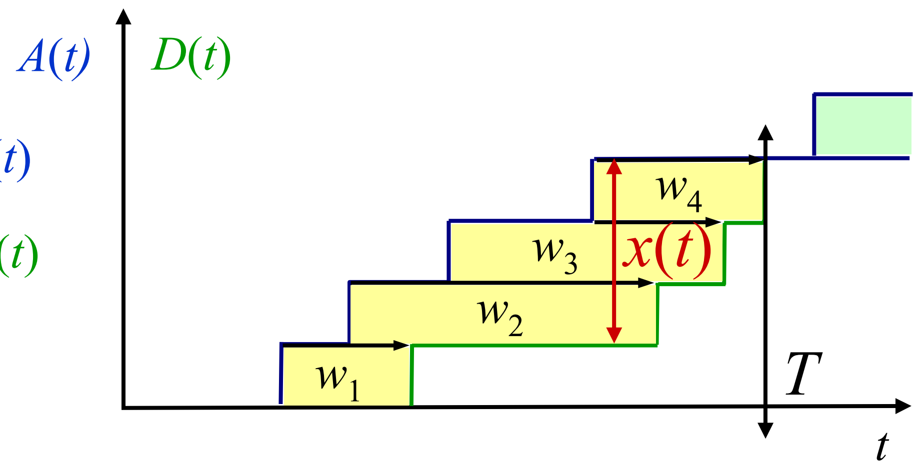
- *E(x)* est une probabilité temporelle
- *E(w)* est une probabilité d'ensemble
- La formule de John Dutton Conant Little (professeur au MIT) ou formule de Little établit un lien entre ces deux quantités de natures différentes

La somme des attentes

- La surface ombrée S est, si l'on s'arrête à un endroit où $x(t) = 0$ (entre 2 bouffées de trafic) : $$S = \sum_i w_i$$
- La surface ombrée S est aussi : $$S = \int_0^T x(t)dt = E(x)t$$

Processus d'arrivée $A(t)$

Processus de départ $D(t)$



Formule de Little

- On calcule la surface ombrée de 2 façons :

$$S = \sum_i w_i = \int_0^T x(t)dt \qquad \text{et} \qquad \frac{S}{T} = \frac{\sum_i w_i}{T} = \frac{\int_0^T x(t)dt}{T}$$

- Ou

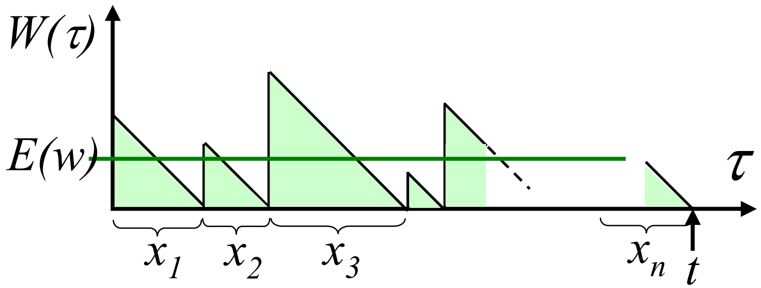
$$\frac{\int_0^T x(t)dt}{T} = \frac{n_C(t)}{T} \frac{\sum_i w_i}{n_C(t)}$$

- En prenant la limite quand $T \to \infty$

$$\boxed{\text{Formule de Little} \quad E(x) = \lambda E(w)}$$

Travail du système (1)

- Au fil du temps des gens arrivent et attendent un temps $w_i = x_i$ pour être servis



- L'espérance temporelle du temps d'attente est appelée le travail du système

$$E(w) = \frac{1}{t}\int_0^t w(\tau)d\tau = \frac{1}{t}\sum_1^n \frac{1}{2}x_i^2$$

- (L'intégrale est la somme des surfaces des triangles)
- Si $t \to \infty$, $t \to nE(x)$ où $E(x)$ est la moyenne d'*ensemble* des durées d'attente, alors que $E(w)$ est la moyenne temporelle des durées d'attente

Travail du système (2)

$$E(w) = \frac{1}{t}\sum_{1}^{n}\frac{1}{2}x_i^2 = \frac{1}{nE(x)}\frac{1}{2}\sum_{1}^{n}x_i^2 = \frac{1}{2E(x)}\frac{1}{n}\sum_{1}^{n}x_i^2$$

$$\boxed{E(w) = \frac{1}{2}\frac{E\left(x^2\right)}{E(x)}}$$

- Mais $\sigma^2 = E\left(x^2\right) - (E(x))^2 \Rightarrow E\left(x^2\right) = \sigma^2 + (E(x))^2$

$$\boxed{E(w) = \frac{1}{2}\frac{\left(\sigma^2 + (E(x))^2\right)}{E(x)}}$$

Trafic offert, Trafic perdu, Trafic écoulé ou Throughput

- Nombre de clients arrivant pendant T : $n = n_C + n_B$

  n : arrivées offertes → Trafic offert $A = \frac{n\tau}{T}$

  n_B : arrivées perdues → Trafic perdu $A_B = \frac{n_B \tau}{T}$

  n_C : arrivées écoulées → Trafic écoulé $A_C = \frac{n_C \tau}{T}$

- Fraction d'arrivées perdues ou probabilité de perte : $B = \frac{n_B}{n}$

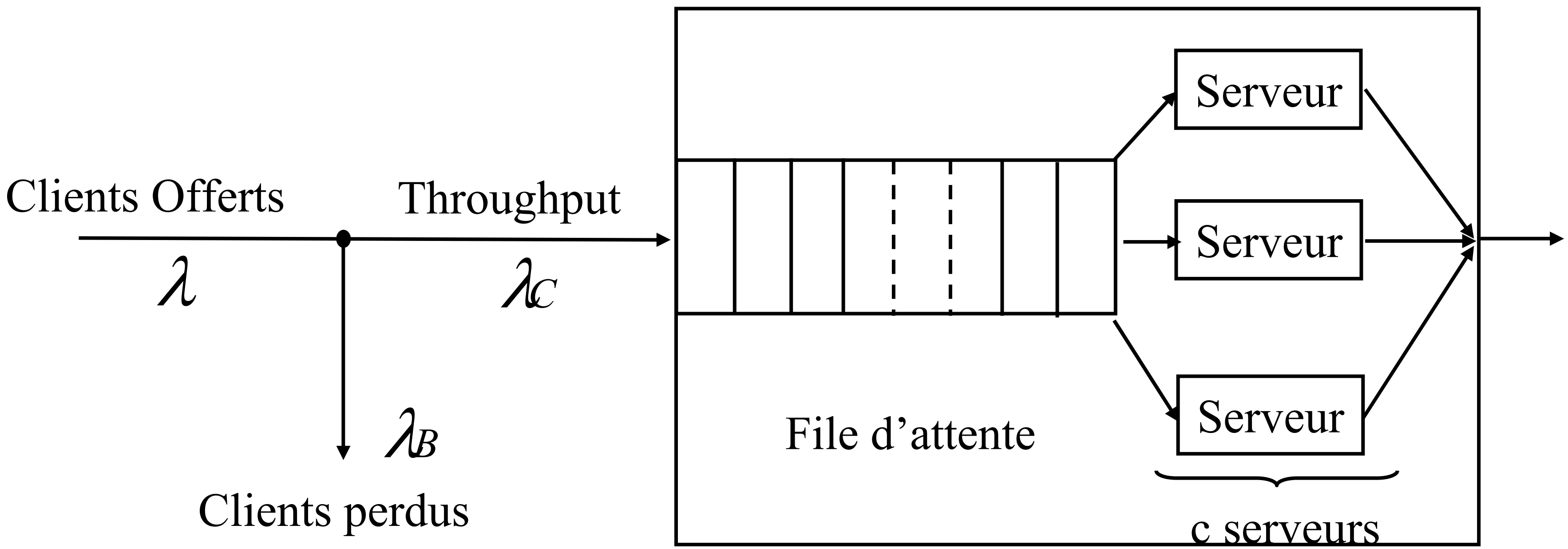
$$n_B = nB \text{ et } n_c = n - n_B = n(1-B)$$

- Le trafic perdu est : $A_B = A \times B$

- Le trafic écoulé aussi appelé ***"Throughput"*** est : $A_C = A \times (1-B)$

File d'attente

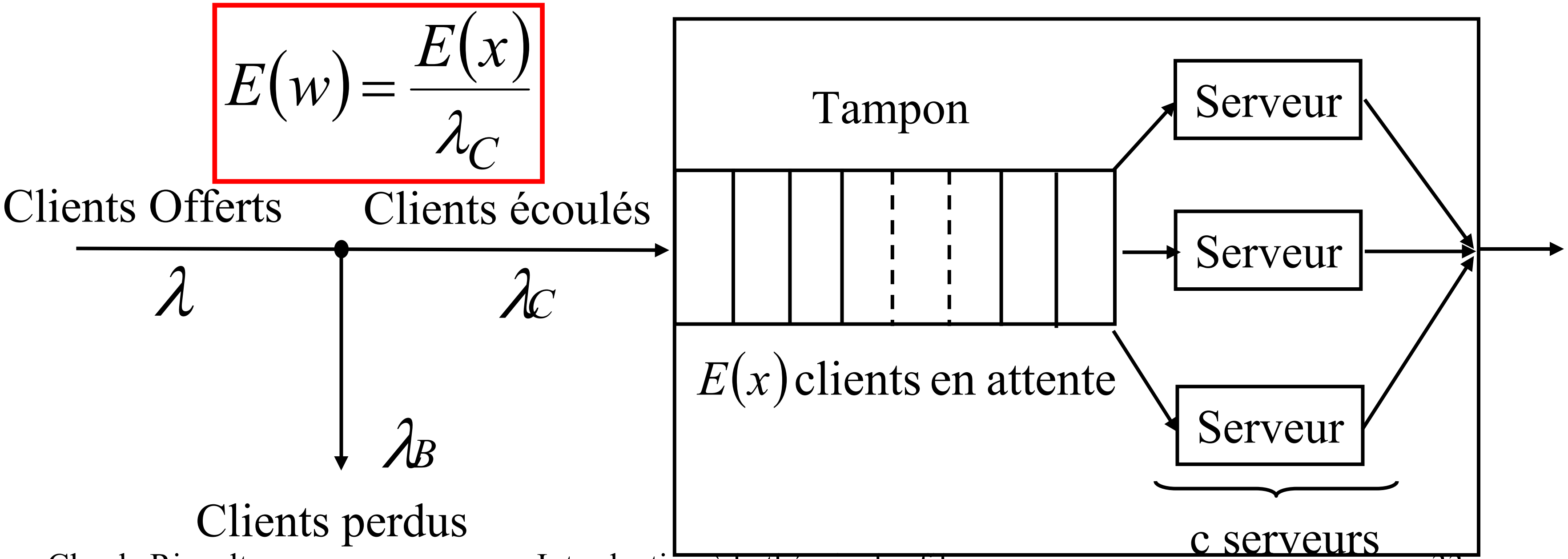
- Espérance du temps passé dans la file : $E(\boldsymbol{w})$
- Nombre de clients dans la file au moment t : $x(t)$
- Nombre de clients arrivés entre 0 et T : $n(T)$



Espérance du temps d'attente

- Espérance du temps d'attente : temps moyen passé dans la file, ***avant*** d'être servi : *E(**w**)*
- Espérance du nombre de clients dans la file d'attente : *E(x)*
- La formule de Little donne :

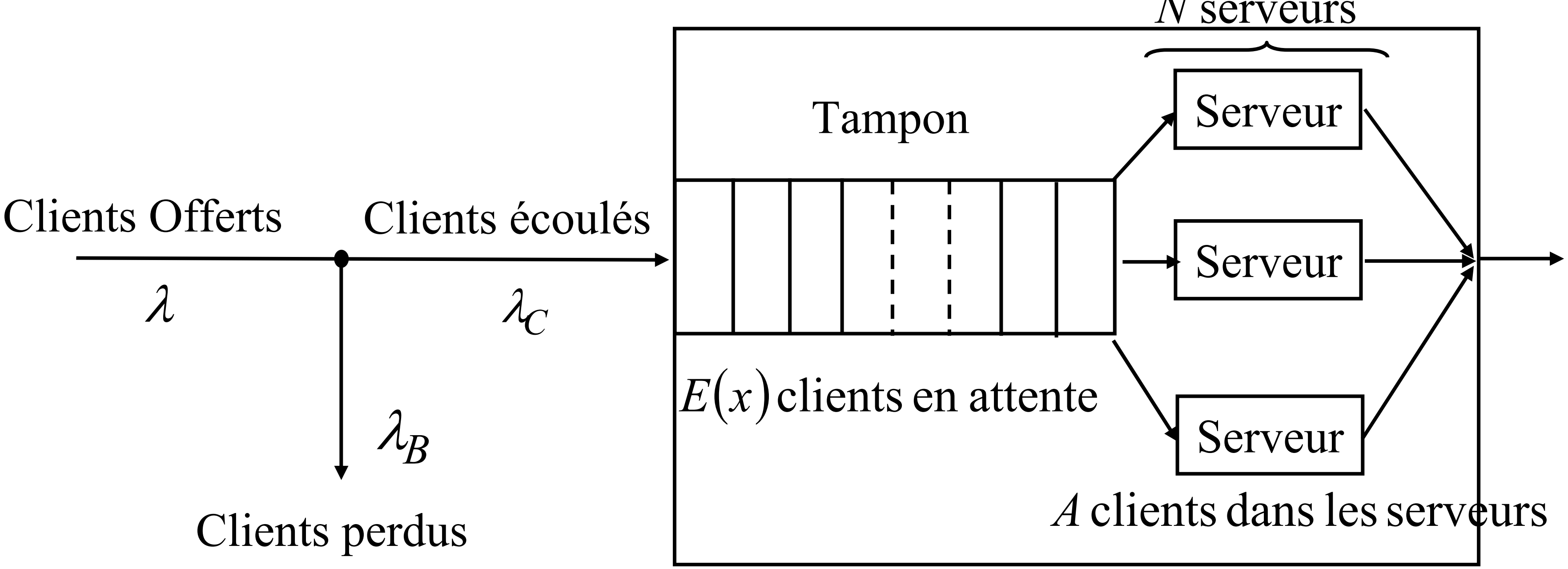
$$\boxed{E(w) = \frac{E(x)}{\lambda_C}}$$



Espérance du temps de séjour

- **Temps de séjour : temps d'attente + temps de service :** $\boxed{E(s) = E(w) + \tau}$
- Nombre de clients dans le système :

$$\boxed{E(c) = \lambda E(s) = \lambda E(w) + \lambda\tau = E(x) + A}$$



Ce qu'il nous reste à déterminer

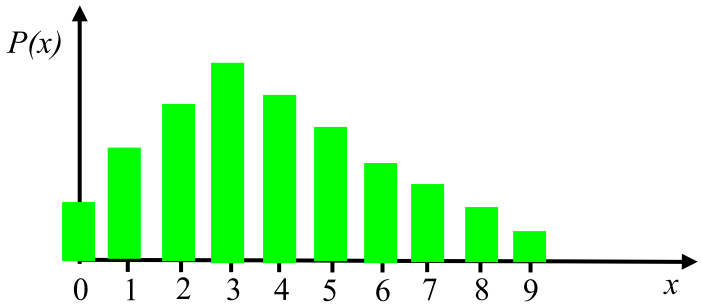
- Nous avons découvert une relation très importante entre le nombre de clients dans le système et leur temps de séjour.
- Ce qui nous intéresse vraiment est le temps de séjour ou la probabilité de perte
- Pour l'évaluer il faut d'abord arriver à déterminer le nombre de clients dans le système

Diapositive volontairement vide

Quelques petits rappels sur les probabilités

Processus à temps discret

- Loi de probabilité (ou fonction (ou loi) de distribution des probabilités) : diagramme en bâtons : ***P(x$_i$) =*** Probabilité que ***x=x$_i$***



- Moyenne ou Espérance $E(x)=\sum_{x=0}^{x_{\max}} xP(x)$

- Variance (moyenne du carré des écarts) :

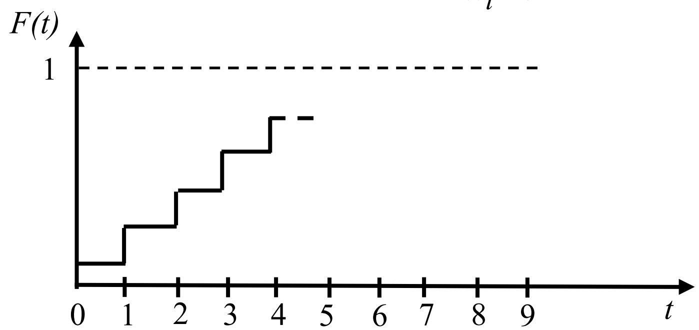
$$Var(x)=\sum_{x=0}^{x_{\max}}(x-E(x))^2P(x)=\sum_{x=0}^{x_{\max}}x^2P(x)-2E(x)\sum_{x=0}^{x_{\max}}xP(x)+E(x)^2\sum_{x=0}^{x_{\max}}P(x)=E\left(x^2\right)-(E(x))^2$$

- Ecart type : $\sigma=\sqrt{Var(x)}$

Fonction de répartition, cas discret

- La fonction de répartition (cumulative distribution function cdf) de ***x*** est la probabilité que la valeur de ***x*** soit $\leq$ ***t***

$$F_X(t) = P(x \leq t) = \sum_{x_i \leq t} P(x_i)$$



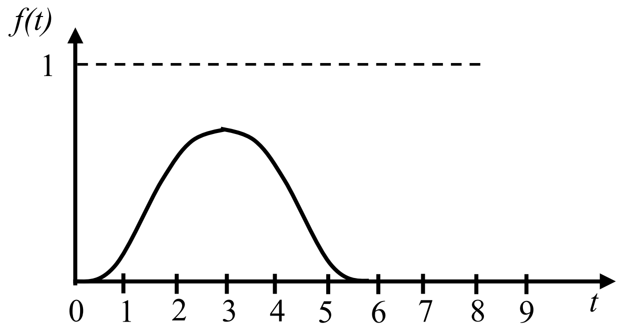
- Probabilité que ***x*** soit dans un intervalle [a,b] : $P(a < x \leq b) = F_X(b) - F_X(a)$

Processus à temps continu

- La loi de probabilité devient une fonction de densité de probabilité (probability density function pdf)
- La probabilité que ***x*** soit comprise entre ***t*** et ***t+dt*** est $\boldsymbol{f_X(t)dt}$

$$P(t < x \le t + dt) = f_X(t)dt$$

$$P(a < x \le b) = \int_a^b f_X(t)dt$$

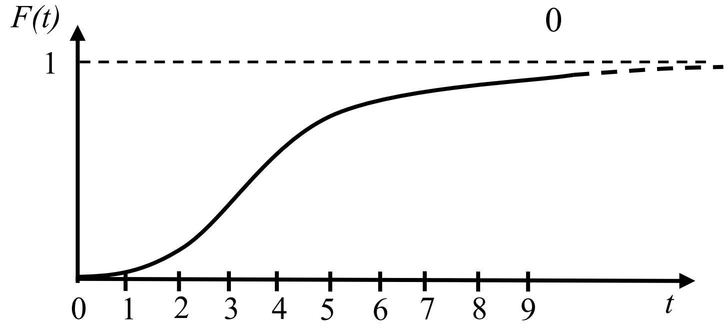


- Espérance : $E(x) = \int_0^\infty x f_X(x)dx$
- Variance : $Var(x) = \int_0^\infty (x - \bar{x})^2 f_X(x)dx = \left(\int_0^\infty x^2 f_X(x)dx\right) - \bar{x}^2$
- Ecart type : $\sigma = \sqrt{Var(x)}$

Fonction de répartition, cas continu

- Fonction de répartition (cumulative distribution function cdf) de $\boldsymbol{x}$: probabilité que la valeur de $\boldsymbol{x}$ soit $\leq \boldsymbol{t}$

$$F_X(t) = P(0 < x \leq t) = \int_0^t f_X(t)\,dt$$



- Probabilité que $\boldsymbol{x}$ soit dans un intervalle [a,b] : $P(a < x \leq b) = F_X(b) - F_X(a)$

Fonction de densité de probabilité

- La fonction de densité est la dérivée de la fonction de répartition :

$$P(x < X \leq x + dx) = \frac{dF_X(x)}{dx} dx$$

- En effet $P(x < X \leq x + dx) = F(x + dx) - F(x) = f(x)dx$

- et $f_X(x) = \frac{dF_X(x)}{dx}$

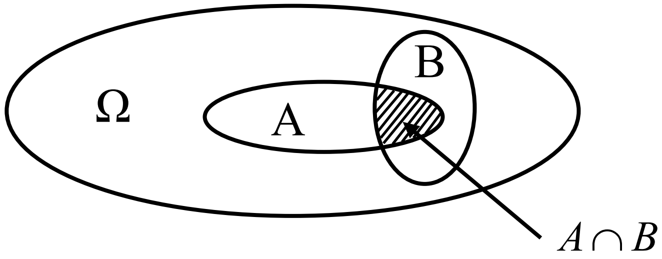
- La fonction de densité de probabilité est :

$$P(x < X \leq x + dx) = f_X(x)dx$$

Probabilité conditionnelle (sachant que…)

- Probabilité de l'événement A ***sachant que*** l'événement B s'est déjà réalisé :

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$



- En général, une probabilité conditionnelle (sachant que) est plus grande qu'une probabilité ne sachant rien.
- (Sachant qu'un homme a déjà atteint 60 ans, son espérance de vie (82 ans en France) est plus grande que celle de la population totale des hommes (78 ans))

PDF de la somme de 2 variables aléatoires indépendantes

- La fonction de densité de probabilité de la somme de deux variables aléatoires indépendantes est le ***produit de convolution*** des fonctions de densité de chacune des variables :
- Si $Z=X+Y$ et si X et Y ***sont indépendantes***, alors :

$$f_Z(z)=\int_{-\infty}^{\infty} f_X(x')f_Y(z-x')dx'$$

Espérance et variance d'une somme de variables aléatoires

- Qu'elles soient indépendantes ou non, l'espérance de la somme de 2 variables aléatoires est la somme des espérances :

$$E(x+y)=E(x)+E(y)$$

- En général la variance de la somme de 2 variables aléatoires ***n'est pas*** la somme des variances, c'est la somme des variance plus 2 fois la ***covariance*** :

$$Var(x+y)=Var(x)+Var(y)+2Cov(x,y)$$

- ***Si x et y sont indépendantes***, alors $Cov(x,y)=0$ et

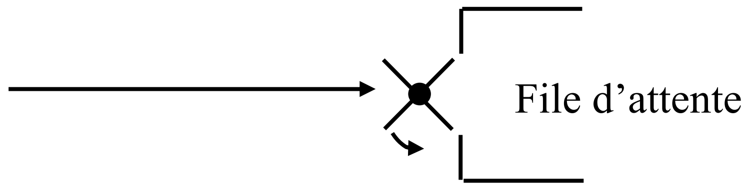
$$Var(x+y)=Var(x)+Var(y)$$

Diapositive volontairement vide

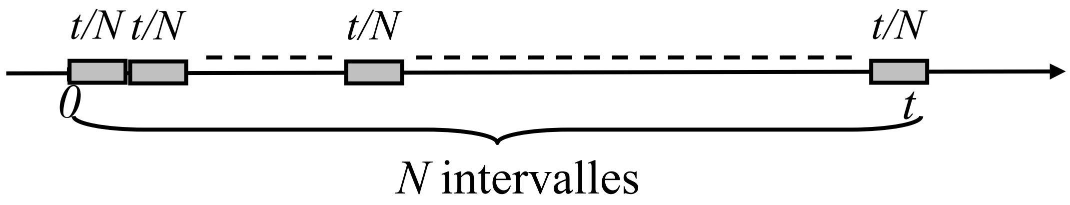
Les processus des clients

Arrivées à des temps discrets

- L'équivalent d'une porte à tourniquet tourne à vitesse constante et divise le temps en temps élémentaires Δt où ne peut entrer qu'un client à la fois



- L'intervalle $[0,t]$ se trouve divisé en N intervalles de durée t/N



Arrivées à des temps discrets

- On suppose que la probabilité d'arrivée dans un intervalle est :

$$P(\Delta t) = \lambda \Delta t$$

- La probabilité de x arrivées sur $[0, t=N\Delta t]$ est :

$$P(x,t) = C_N^x (\lambda \Delta t)^x (1 - \lambda \Delta t)^{N-x}$$

- Il s'agit de la loi de probabilité Binomiale ou
- de ***Bernoulli (Jacques)***
- Nous pouvons poser $\lambda\Delta t = p$ et $(1- \lambda\Delta t) = q$

$$P(x,t) = C_N^x p^x q^{N-x}$$

La loi de Bernoulli est une loi de probabilité

- Nous devons avoir :

$$\sum_{x=0}^{N} P(x) = \sum_{x=0}^{N} C_N^x p^x (1-p)^{N-x} = 1$$

- Or : $\sum_{x=0}^{N} C_N^x p^x (1-p)^{N-x}$ est le développement du binôme $(p+(1-p))^N$

- Ce binôme est égal à 1. CQFD

Espérance de la loi de Bernoulli

- Nous pouvons calculer l'espérance de la loi de Bernouilli par la méthode de la force brute :

$$E(x) = \sum_{x=0}^{N} x p^x (1-p)^{N-x}$$

- Nous faisons le calcul en annexe et nous trouvons : $\boxed{E(x) = Np}$

- Une méthode plus fine est de dire que nous cherchons l'espérance d'un ensemble de *N* variables *indépendantes identiquement distribuées* (iid) dont la valeur peut être *1* avec probabilité *p* et *0* avec probabilité *(1-p)*
- L'espérance d'une de ces variables est : $(1 \times p) + 0 \times (1 - p) = p$
- L'espérance de l'ensemble des *N* variables est *Np*

Variance de la loi de Bernoulli

- Nous pouvons également calculer la variance par la méthode de la force brute :

$$Var(x) = \left(\sum_{x=0}^{N} x^2 p^x (1-p)^{N-x} \right) - (Np)^2$$

- Nous faisons ce calcul en annexe et nous trouvons : $\boxed{Var(x) = Np(1-p)}$

- Nous pouvons aussi utiliser la méthode de l'iid comme nous l'avons fait pour l'espérance :

$$Var(x) = N\left[1^2 \times p + 0^2(1-p) - p^2\right] = N\left[p - p^2\right] = Np(1-p)$$

Temps inter-arrivées dans le cas discret

- On appelle ***x*** la variable aléatoire donnant le ***rang*** de la première arrivée
- La probabilité qu'il n'y ait aucune arrivée dans [*0, (N-1) Δt*] est :

$$P(0, t-1) = \left(1 - \lambda \frac{t}{N}\right)^{N-1} = q^{N-1}$$

- La probabilité que la première arrivée soit au temps $t = N\Delta t$ est :

$$\boxed{P(x=t) = \lambda\Delta t(1-\lambda\Delta t)^{N-1} = pq^{N-1}}$$

- Il s'agit d'une loi de probabilité ***géométrique du deuxième type***
- La fonction de répartition du temps inter-arrivées est la probabilité pour que le temps inter-arrivées soit $\leq \boldsymbol{t}$:

$$F(t) = \sum_{i=1}^{N} P(x_i) = \sum_{i=1}^{N} pq^{i-1} = p\sum_{i=1}^{N} q^{i-1} = p\frac{1-q^N}{1-q} = 1 - q^N$$

$$\boxed{F(t) = 1 - q^N = 1 - (1 - \lambda \frac{t}{N})^N}$$

- *La loi géométrique est la seule fonction de probabilité discrète à avoir une fonction de répartition simple*
- La probabilité que la première arrivée se produise à un temps > ***t*** est :

$$P(x > t) = 1 - P(x \leq t) = 1 - \left(1 - q^N\right) = q^N$$

Il y a 2 types de lois géométriques

- Premier type : $p(x)=p(1-p)^x$
- Alors :

$$E(x)=\frac{1-p}{p} \text{ et } Var(x)=\frac{1-p}{p^2}$$

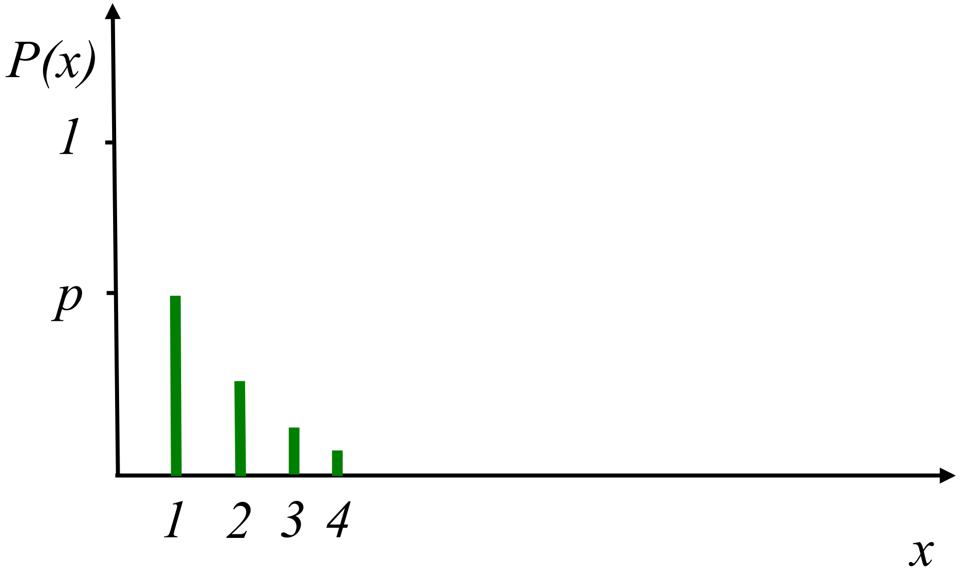
- Deuxième type : $p(x)=p(1-p)^{x-1}$
- Alors :

$$E(x)=\frac{1}{p} \text{ et } Var(x)=\frac{1-p}{p^2}$$

- La loi géométrique du premier type donne les probabilités d'occupation des files d'attente M/M/1
- La loi géométrique du deuxième type donne les temps inter-arrivées dans le cas discret

La loi géométrique de deuxième type

- La loi géométrique $P(x) = p(1-p)^{x-1}$



La loi géométrique est une loi de probabilité

- Nous supposons que nous observons les inter-arrivées pendant un temps infiniment long : *x* varie de 1 à l'infini
- Nous devons avoir : $\sum_{x=1}^{\infty} P(x) = \sum_{x=1}^{\infty} p(1-p)^{x-1} = 1$
- mais $\sum_{x=1}^{\infty} p(1-p)^{x-1} = p \sum_{x=0}^{N} (1-p)^{x-1}$
- En posant *y=x-1* et *(1-p)=q*

$$\sum_{x=1}^{\infty} p(1-p)^{x-1} = p \sum_{y=0}^{\infty} q^{y} = \frac{p}{1-q} = \frac{p}{p} = 1$$

Espérance de la loi géométrique

- Le temps inter-arrivée peut être aussi grand que l'on veut :

$$P(x) = p(1-p)^{x-1} = pq^{x-1} \quad \text{où} \quad p = \lambda \Delta t \;\; \text{et} \;\; q = (1-p)$$

- Nous montrons en annexe que l'espérance de x est : $E(x) = \frac{1}{p} = \frac{1}{\lambda \Delta t}$

- L'espérance du temps entre 2 arrivées est : $E(t) = E(x)\Delta t = \frac{\Delta t}{\lambda \Delta t} = \frac{1}{\lambda}$

$$\boxed{E(t) = \frac{1}{\lambda}}$$

- La variance et l'écart type sont :

$$\boxed{\begin{aligned} Var(x) &= \frac{1-p}{p^2} \\ \sigma &= \frac{\sqrt{1-p}}{p} \end{aligned}}$$

La propriété de Markov

- La propriété de Markov est la propriété sans mémoire : la probabilité conditionnelle d'un état futur sachant l'état présent ne dépend que de cet état présent et pas des états antérieurs.

- A la roulette, la probabilité d'avoir un noir sachant que l'on a déjà eu 8 rouges est la même que la probabilité d'avoir un noir au premier lancer (la boule n'a pas de mémoire).
- La probabilité d'attendre 2 minutes sachant que l'on a déjà attendu 8 minutes est la même que la probabilité d'attendre 2 minutes sachant que l'on a pas encore attendu

Andreï A. Markov

La loi géométrique possède la propriété de Markov

- Probabilité d'un temps inter-arrivées $\leq$ ***s*** sachant qu'un temps ***t*** s'est déjà écoulé sans arrivées

$$P(x \leq t+s \mid x>t) = \frac{P((x \leq t+s) \cap (x>t))}{P(x>t)} = \frac{P(t<x \leq t+s)}{P(x>t)}$$

$$P(x \leq t+s \mid x>t) = \frac{1-q^{N+M}-(1-q^{N})}{q^{N}} = \frac{q^{N}-q^{N+M}}{q^{N}} = 1-q^{M} = P(x \leq s)$$

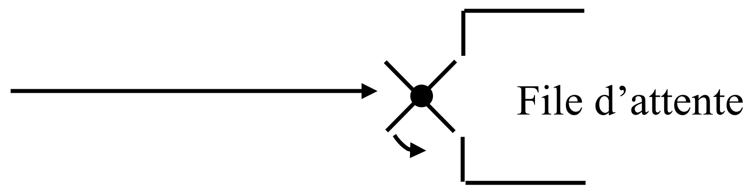
- La loi géométrique est sans mémoire (propriété de Markov)
- ***C'est la seule loi discrète à avoir cette propriété***

$$\boxed{P(x \leq t+s \mid x>t) = P(x \leq s)}$$

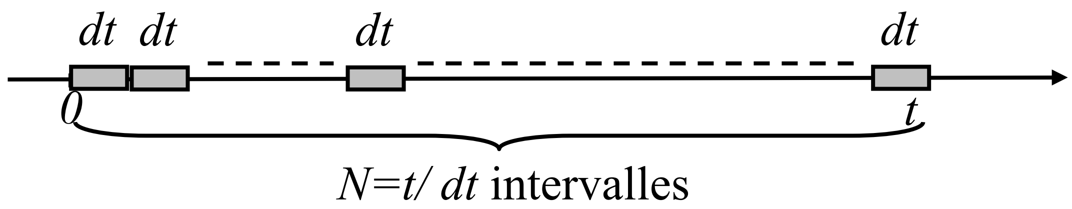
Diapositive volontairement vide

Arrivées à des temps continus

- La porte à tourniquet tourne maintenant à vitesse non constante aussi rapide que l'on veut (elle n'a pas de frein).
- Elle divise le temps en temps élémentaires *dt* où ne peut entrer qu'un client à la fois *(dt→0)*



- L'intervalle [*0,t*] se trouve divisé en *N* intervalles de durée *dt*

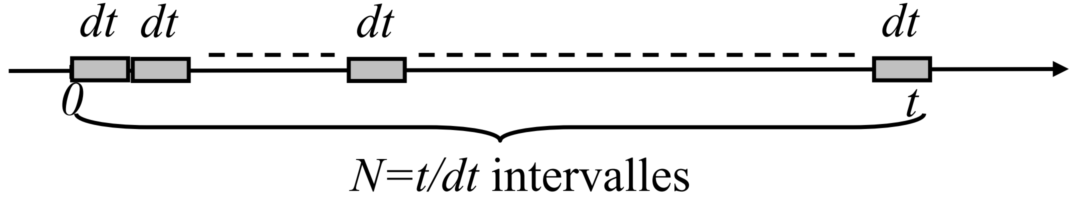


Proba temporelles, arrivées Poissoniennes

- Divisons l'intervalle [$0,t$] en $N=t/dt$ intervalles de durée dt
- La probabilité d'arrivée d'un client sur cet intervalle est : $P(dt) = \lambda dt$
- La probabilité de x arrivées sur [$0,t$] est : $P(x,t) = C_N^x (\lambda dt)^x (1-\lambda dt)^{N-x}$
- Lorsque $dt \to 0$, $N \to \infty$ la loi de Bernoulli tend vers la loi de Poisson

$$P(x,t) \to \frac{(\lambda t)^x}{x!} e^{-\lambda t}$$

- *(voir démonstration sur la diapo suivante)*



Bernoulli → Poisson

- $N=t/dt$, lorsque $dt \to 0$, $N \to \infty$

$$P(x,t) = C_N^x (\lambda dt)^x (1-\lambda dt)^{N-x}$$

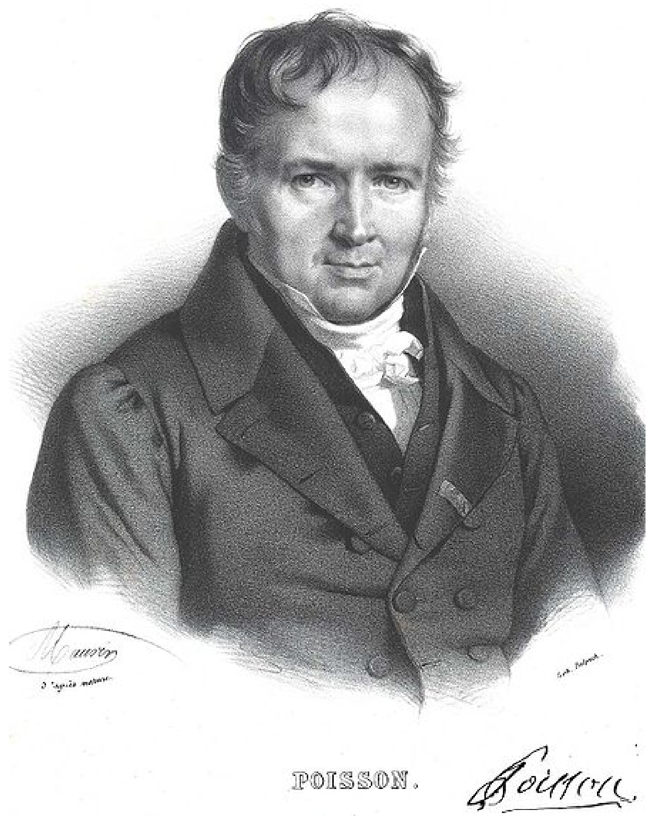
$$P(x,t) = \frac{N!}{x!(N-x)!} (\lambda dt)^x (1-\lambda dt)^{N-x}$$

$$P(x,t) = \frac{N(N-1)\ldots(N-x+1)}{x!} (\lambda dt)^x (1-\lambda dt)^{N-x}$$

$$N(N-1)\ldots(N-x+1) \to N^x = \left(\frac{t}{dt}\right)^x$$

$$(1-\lambda dt)^{\frac{t}{dt}-x} \to e^{-\lambda t}$$

$$P(x,t) \to \frac{\left(\frac{t}{dt}\right)^x}{x!} (\lambda dt)^x e^{-\lambda t} = \frac{(\lambda t)^x}{x!} e^{-\lambda t}$$



Siméon Denis Poisson

La loi de Poisson est une loi de probabilité

- Nous devons avoir : $$\sum_{x=0}^{N} P(x) = \sum_{x=0}^{N} \frac{A^x}{x!} e^{-A} = 1$$

$$\sum_{x=0}^{N} P(x) = e^{-A} \sum_{x=0}^{N} \frac{A^x}{x!}$$

- Mais : $$\sum_{x=0}^{N} \frac{A^x}{x!} = e^A$$

- Donc : $$\sum_{x=0}^{N} P(x) = e^{-A} e^{A} = 1$$

Espérance de la loi de Poisson

- Nous utilisons la méthode de la force brute :

$$E(x) = \sum_{x=0}^{\infty} x \frac{A^x}{x!} e^{-A} = \sum_{x=1}^{\infty} A \frac{A^{(x-1)}}{(x-1)!} e^{-A}$$

- Nous posons *y=x-1* (ou *x=y+1*) : $E(x) = A \sum_{y=0}^{\infty} \frac{A^y}{y!} e^{-A}$ mais $\sum_{y=0}^{\infty} \frac{A^y}{y!} e^{-A} = 1$

- Donc : $\boxed{E(x) = A}$

Variance de la loi de Poisson

- Nous utilisons la méthode de la force brute :

$$Var(x) = \left(\sum_{x=0}^{\infty} x^2 \frac{A^x}{x!} e^{-A} \right) - A^2 = \left(\sum_{x=1}^{\infty} Ax \frac{A^{(x-1)}}{(x-1)!} e^{-A} \right) - A^2$$

- Nous posons $y=x-1$ (ou $x=y+1$) :

$$Var(x) = \left(A \sum_{y=0}^{\infty} (y+1) \frac{A^y}{y!} e^{-A} \right) - A^2 = \left(A \sum_{y=0}^{\infty} y \frac{A^y}{y!} e^{-A} + A \sum_{y=0}^{\infty} \frac{A^y}{y!} e^{-A} \right) - A^2$$

- ou $$Var(x) = A^2 + \mathrm{A} - A^2$$
- Donc : $Var(x) = A$ et $\sigma = \sqrt{A}$

Pour la loi de Poisson, l'écart type est la <u>racine carrée de la moyenne</u>

Temps inter-arrivées

- Soit *0* l'instant de l'arrivée *i* et *t* l'instant de l'arrivée *i+1*
- Divisons l'intervalle [*0, t*] en *N=t/dt* intervalles
- La probabilité qu'il n'y ait aucune arrivée dans [*0, t*] est : $P(0,t) = (1 - \lambda dt)^N$
- Lorsque $dt \to 0$, $N \to \infty$ $P(0,t) \to e^{-\lambda t}$
- Ceci représente la probabilité pour que le temps entre 2 arrivées soit **>*t***
- Appelons *x* le temps entre 2 arrivées.

$$P(x > t) = e^{-\lambda t}$$

- La fonction de répartition du temps inter-arrivées est la probabilité pour que le temps inter-arrivées soit $\boldsymbol{\leq t}$

$$\boxed{P(x \leq t) = F(t) = 1 - e^{-\lambda t}}$$

Fonction de densité exponentielle (1)

- La fonction de répartition du temps inter-arrivées est :

$$\boxed{P(x \le t) = F(t) = 1 - e^{-\lambda t}}$$

- La fonction de densité de probabilité est :

$$\boxed{f(t) = \frac{dF(t)}{dt} = \lambda e^{-\lambda t}}$$

- C'est la loi de densité de probabilité ***exponentielle***

$$P(t < x < t + dt) = \lambda e^{-\lambda t} dt$$

$$P(a < x \le b) = \int_a^b \lambda e^{-\lambda t} dt = -\int_{t=a}^{t=b} e^{-\lambda t} d(-\lambda t) = -\left[e^{-\lambda t}\right]_{t=a}^{t=b} = \left[e^{-\lambda t}\right]_{t=b}^{t=a} = e^{-\lambda a} - e^{-\lambda b}$$

La loi exponentielle est une loi de densité de probabilité

- Nous devons avoir : $\int_0^\infty f(x)dx = \int_0^\infty \lambda e^{-\lambda t} dt = 1$

$$\int_0^\infty \lambda e^{-\lambda t} dt = \int_0^\infty e^{-\lambda t} d\lambda t$$

- Nous posons $y=\lambda t$:

$$\int_0^\infty e^{-\lambda t} d\lambda t = \int_0^\infty e^{-y} dy = \left[-e^{-y}\right]_0^\infty = e^{-0} - e^{-\infty} = 1$$

Espérance du temps inter-arrivées

- Espérance : $E(t)=\int_0^\infty tf(t)dt=\int_0^\infty t\lambda e^{-\lambda t}dt=\frac{1}{\lambda}\int_0^\infty \lambda te^{-\lambda t}d\lambda t$
- On pose : $x=\lambda t \Rightarrow E(t)=\frac{1}{\lambda}\int_0^\infty xe^{-x}dx$
- On intègre par parties : $\int_0^\infty xe^{-x}dx=\left[-xe^{-x}\right]_0^\infty+\int_0^\infty e^{-x}dx=\left[-xe^{-x}\right]_0^\infty+\left[e^{-x}\right]_0^\infty=1$
- Donc $\boxed{E(t)=\frac{1}{\lambda}}$

Variance du temps inter-arrivées

$$Var(t) = \left(\int_0^\infty t^2 \lambda e^{-\lambda t} dt \right) - \frac{1}{\lambda^2}$$

Calculons l'intégrale : $\int_0^\infty t^2 \lambda e^{-\lambda t} dt = \frac{1}{\lambda^2} \int_0^\infty (\lambda t)^2 e^{-\lambda t} d\lambda t = \frac{1}{\lambda^2} \int_0^\infty x^2 e^{-x} dx$

$$\int_0^\infty x^2 e^{-x} dx = \left[-xe^{-x} \right]_0^\infty + 2\int_0^\infty xe^{-x} dx = 2$$

On obtient : $Var(t) = \left(\frac{2}{\lambda^2} \right) - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$ $\boxed{Var(t) = \frac{1}{\lambda^2} \Rightarrow \sigma = \frac{1}{\lambda}}$

Pour la loi exponentielle, l'écart type est <u>égal à la moyenne</u>

La loi exponentielle possède la propriété de Markov (1)

- Probabilité d'un temps inter-arrivées $\leq \boldsymbol{s}$ sachant qu'un temps $\boldsymbol{t}$ s'est déjà écoulé sans arrivées

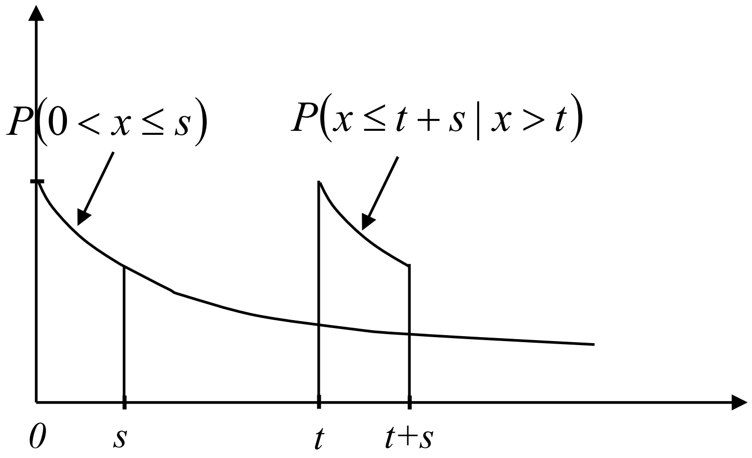
$$P(x \leq t+s \mid x>t) = \frac{P((x \leq t+s) \cap (x>t))}{P(x>t)} = \frac{P(t \leq x \leq t+s)}{P(x>t)}$$

$$P(x \leq t+s \mid x>t) = \frac{F(t+s)-F(t)}{P(x>t)} = \frac{\left(1-e^{-\lambda(t+s)}\right)-\left(1-e^{-\lambda t}\right)}{e^{-\lambda t}} = \frac{e^{-\lambda t}-e^{-\lambda(t+s)}}{e^{-\lambda t}} = 1-e^{-\lambda s} = P(x \leq s)$$

- *La loi exponentielle est sans mémoire. C'est la seule loi de probabilité continue qui a cette propriété.*

$$\boxed{P(x \leq t+s \mid x>t) = P(x \leq s)}$$

La loi exponentielle possède la propriété de Markov (2)



$$\boxed{P(x \leq t + s \mid x > t) = P(x \leq s)}$$

La propriété PASTA

- PASTA : Poisson Arrivals See Time Average
- Nous nous souvenons que le travail est : $E(w)=\frac{1}{2}\frac{\left(\sigma^2+\bar{t}^2\right)}{\bar{t}}$

- Dans le cas d'arrivées Poissoniennes , les inter-arrivées suivent une loi exponentielle : $E(w)=\frac{1}{\lambda}\ ,\ \sigma=\frac{1}{\lambda}$

- On trouve que l'espérance du temps d'attente d'un client dans la file (s'il n'y a personne avant lui) est :

$$E(w)=\frac{1}{2}\frac{\left(\frac{1}{\lambda^2}+\frac{1}{\lambda^2}\right)}{\frac{1}{\lambda}}=\frac{1}{\lambda}$$

Le paradoxe de l'auto-stoppeur

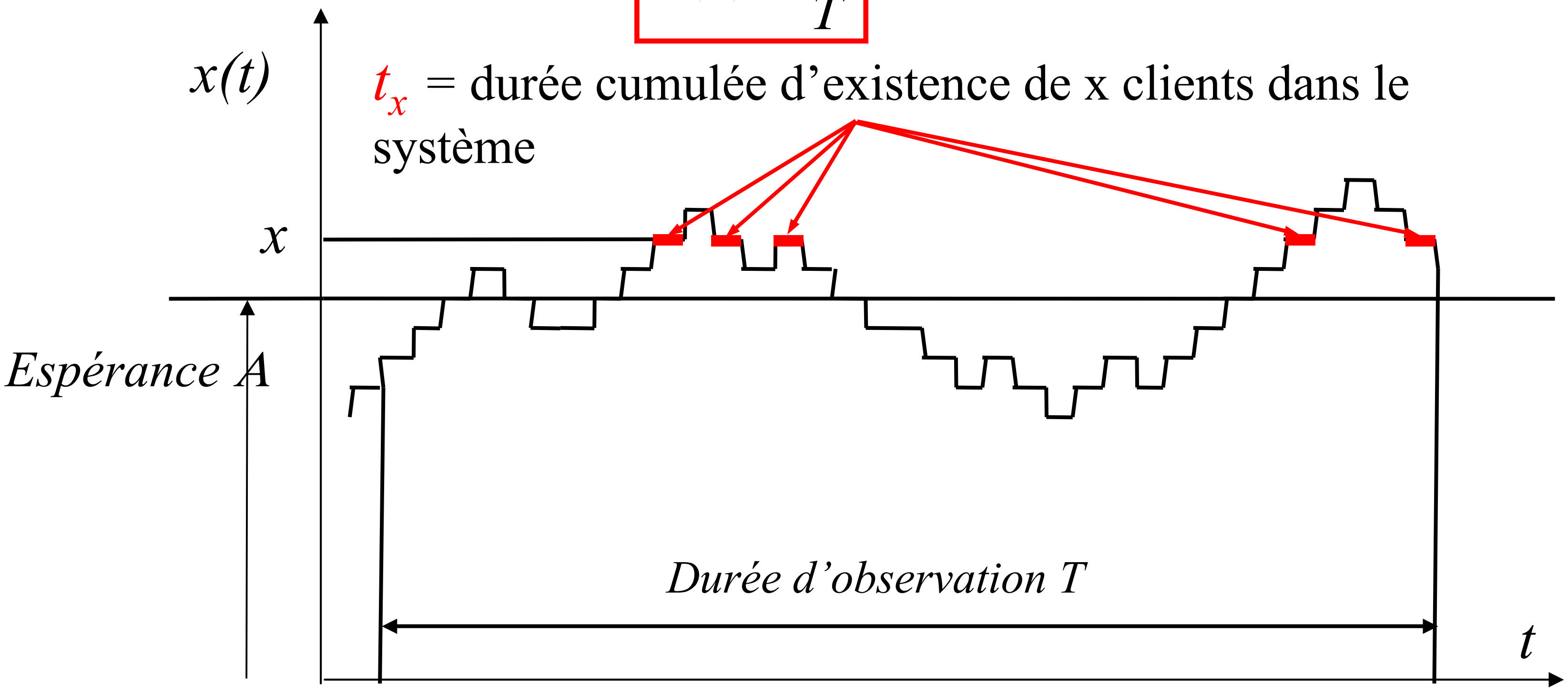
- Un auto-stoppeur arrive au bord d'une route. Les voitures passent selon une loi exponentielle de taux 10 mn
- Combien de temps, en moyenne l'auto-stoppeur doit il attendre?
- On s'attend à une réponse du genre 5 mn (qui serait vraie si les arrivées des voitures étaient déterministes)
- La propriété PASTA nous dit que la réponse vraie est 10 mn. (Plus de gens attendent des temps long que des temps courts)

Diapositive volontairement vide

Probabilité de x clients simultanés

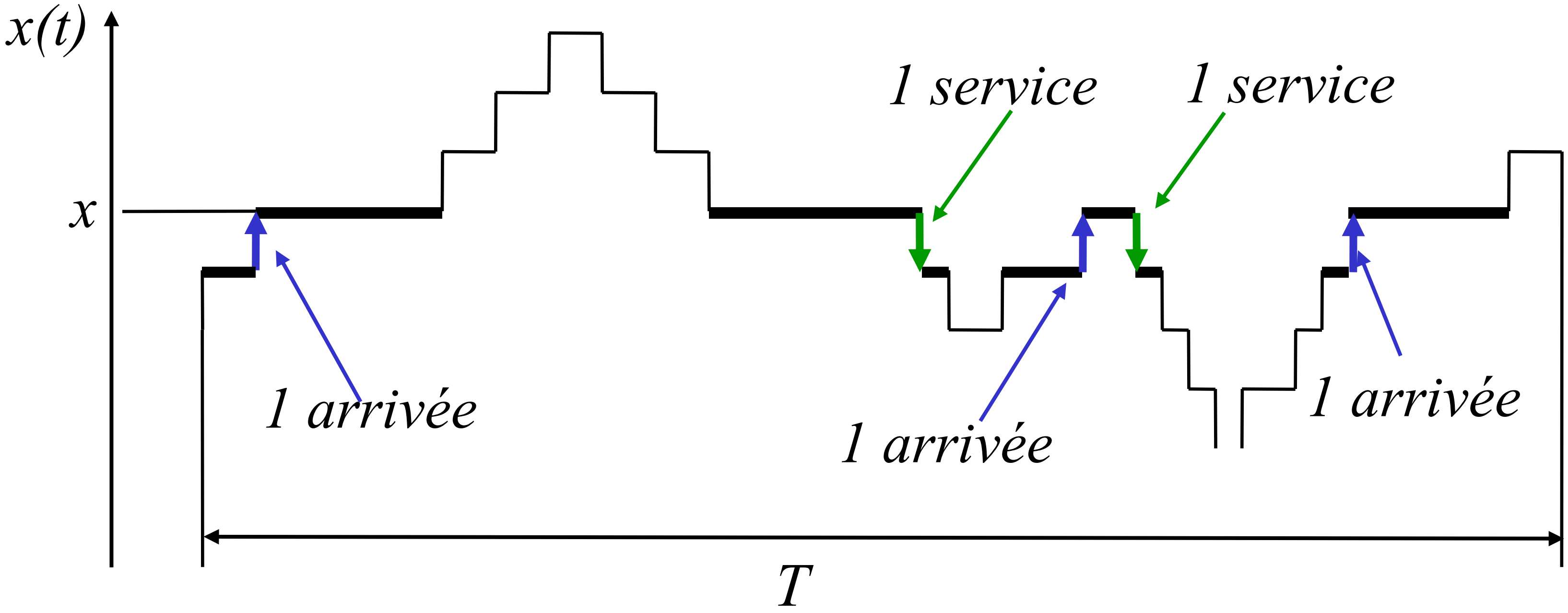
- La probabilité temporelle d'avoir simultanément *x* clients dans le système est :

$$\boxed{\pi(x) = \frac{t_x}{T}}$$



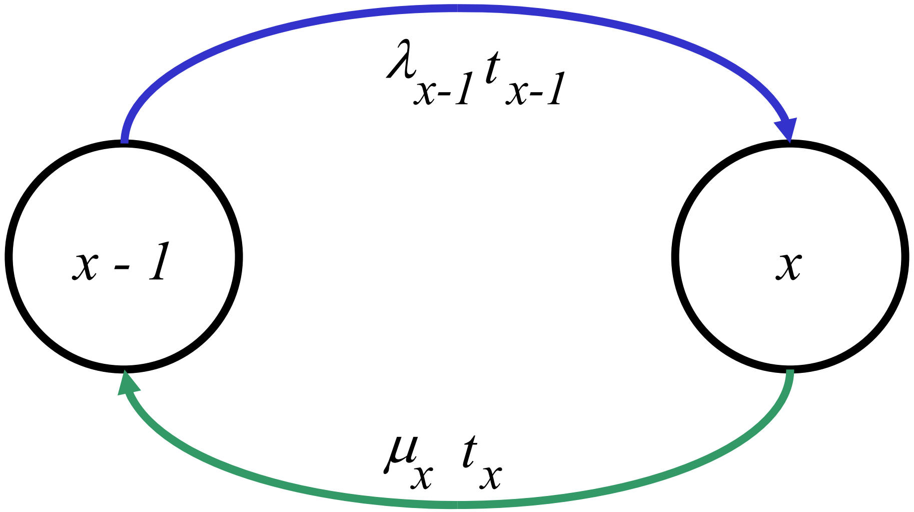
Equilibre statistique

- Pour T très grand, le nombre de transitions de x à x-1est égal au nombre de transitions de x-1 à x (l'effet de bord devient négligeable)



L'équation d'équilibre

- A l'équilibre : $\lambda_{x-1} t_{x-1} = \mu_x t_x$
- Si $\boldsymbol{x} \le \boldsymbol{N}$ où $\boldsymbol{N}$ est le nombre de serveurs (pas d'attente) , alors $\mu_x = \dfrac{x}{\tau}$
  où $\boldsymbol{\tau}$ est l'espérance du temps de service



Cas d'un grand nombre de clients

- λ_x est indépendant de x : $\lambda = \frac{n}{T}$ Quand c'est le cas, on parle d'un processus d'arrivées Markovien (le passé, n'a pas d'influence).
- S'il n'y a pas d'attente μ_x est : $\mu_x = \frac{x}{\tau}$

$$\lambda_{x-1} t_{x-1} = \mu_x t_x$$

- L'équation d'équilibre est : $\frac{n}{T} t_{x-1} = \frac{x}{\tau} t_x$
-
- Elle devient : $\frac{n\tau}{T} \frac{t_{x-1}}{T} = x \frac{t_x}{T}$
- Ou : $\boxed{\pi(x) = \frac{A}{x} \pi(x-1)}$

La distribution de Poisson

- Pour un grand nombre de clients : $\pi(x) = \dfrac{A^x}{x!}\pi(0)$
- On calcule $P(0)$ par : $\sum_{x=0}^{\infty} \pi(x) = \pi(0) \sum_{x=0}^{\infty} \dfrac{A^x}{x!} = 1 \quad \Rightarrow \quad \pi(0) = e^{-A}$
- ou : $\boxed{\pi(x) = \dfrac{A^x}{x!} e^{-A}}$
- On obtient (voir annexe) $E(x) = A$ et $\sigma = \sqrt{A}$

Conséquences du trafic Poissonien

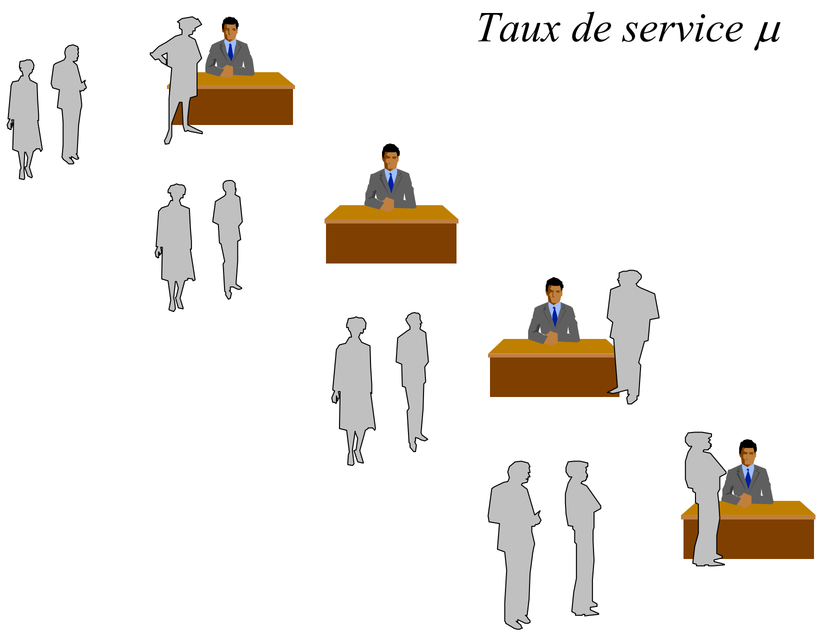
- La fluctuation mesurée par l'écart type ne varie pas linéairement avec le trafic :

$$\boxed{\sigma = \sqrt{A}}$$

- Un réseau n'est pas dimensionné pour le nombre moyen de clients dans le système mais pour ce nombre plus la fluctuation.
- La fluctuation varie comme la racine carrée du trafic $\Rightarrow$ Plus le trafic est grand, moins la fluctuation est grande (en valeur relative)
- ***Le dimensionnement de réseau n'est pas linéaire :***

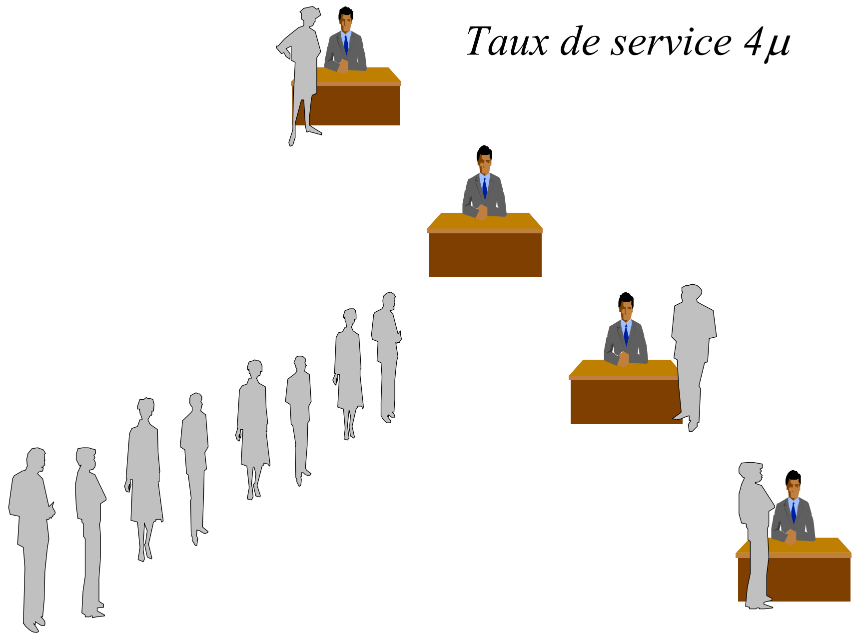
$$\boxed{N \cong A + k\sqrt{A}}$$

La règle de l’accessibilité totale



Grandes fluctuations ressenties par le client !

Mauvaise solution



Petites fluctuations ressenties par le client !

Bonne solution

Cas d'un petit nombre de clients

- λ_x dépend de x : s'il y a en tout L clients et *x-1* sont déjà là, alors seulement $L-(x-1)=L-x+1$ clients peuvent produire des arrivées
- Un client a le trafic $\rho=\frac{\nu\tau}{T}$
- Un client libre produit $\lambda_f=\frac{\nu}{T-\nu\tau}$ arrivées par unité de temps
- L'équation d'équilibre devient : $(L-x+1)\frac{\nu}{T-\nu\tau}t_{x-1}=\frac{x}{\tau}t_x$
- Ou $\pi(x)=\frac{(L-x+1)}{x}\left(\frac{\rho}{1-\rho}\right)\pi(x-1)$
- et : $\pi(x)=C_L^x\left(\frac{\rho}{1-\rho}\right)^x\pi(0)$

La distribution de Bernoulli

- On calcule $\pi(0)$ par : $\sum_{x=0}^{\infty} \pi(x) = \pi(0) \sum_{x=0}^{L} C_L^x \left(\frac{\rho}{1-\rho} \right)^x = 1$
- Ou : $\pi(0)\left(1+\frac{\rho}{1-\rho}\right)^L = 1$
- et : $\pi(0) = (1-\rho)^L$
- donc : $\boxed{\pi(x) = C_L^x \rho^x (1-\rho)^{L-x}}$
- On obtient : $\overline{X} = L\rho = A$ et : $\sigma = \sqrt{A(1-\rho)}$

Comparaison de Bernoulli et de Poisson

- Pour une moyenne donnée, la loi de Poisson a la plus grande entropie $\Rightarrow$ Poisson est un modèle pessimiste (worst case)
- La fluctuation de Bernoulli est un peu plus petite que celle de Poisson
- $\Rightarrow$ Le modèle de Bernoulli est plus précis pour un petit nombre de clients
- Quand $L \to \infty$ Bernoulli $\to$ Poisson

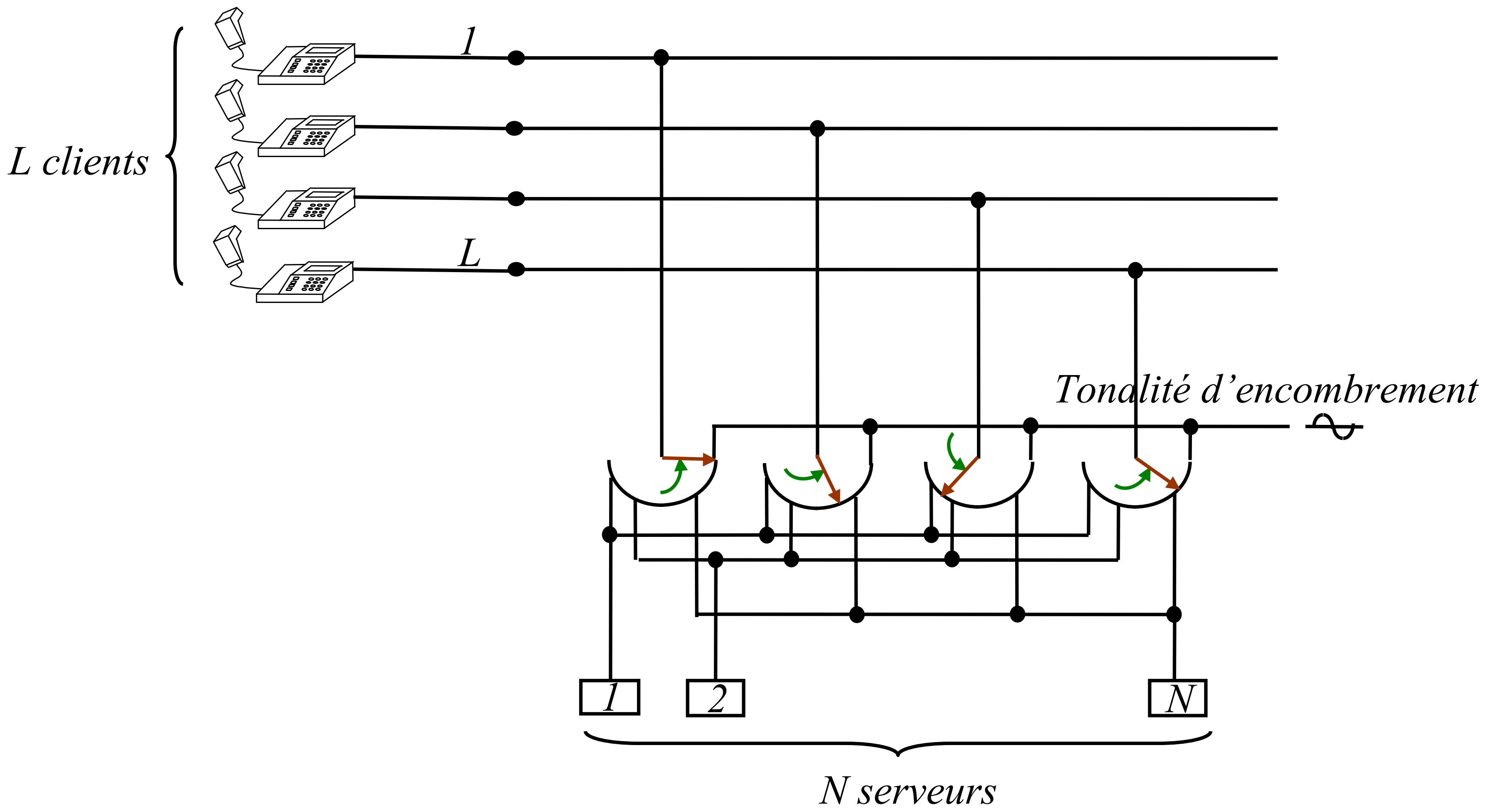
$$\pi(x) = \frac{L!}{x!(L-x)!}\rho^x(1-\rho)^{L-x} = \frac{\rho^x}{x!}\frac{L!}{(L-x)!}\left(1-\frac{A}{L}\right)^{L-x}$$

- et $\frac{L!}{(L-x)!} \to L^x\ ,\ \left(1-\frac{A}{L}\right)^{L-x} \to e^{-A}$

$$\pi(x) \to \frac{(L\rho)^x}{x!}e^{-A} = \frac{A^x}{x!}e^{-A}$$

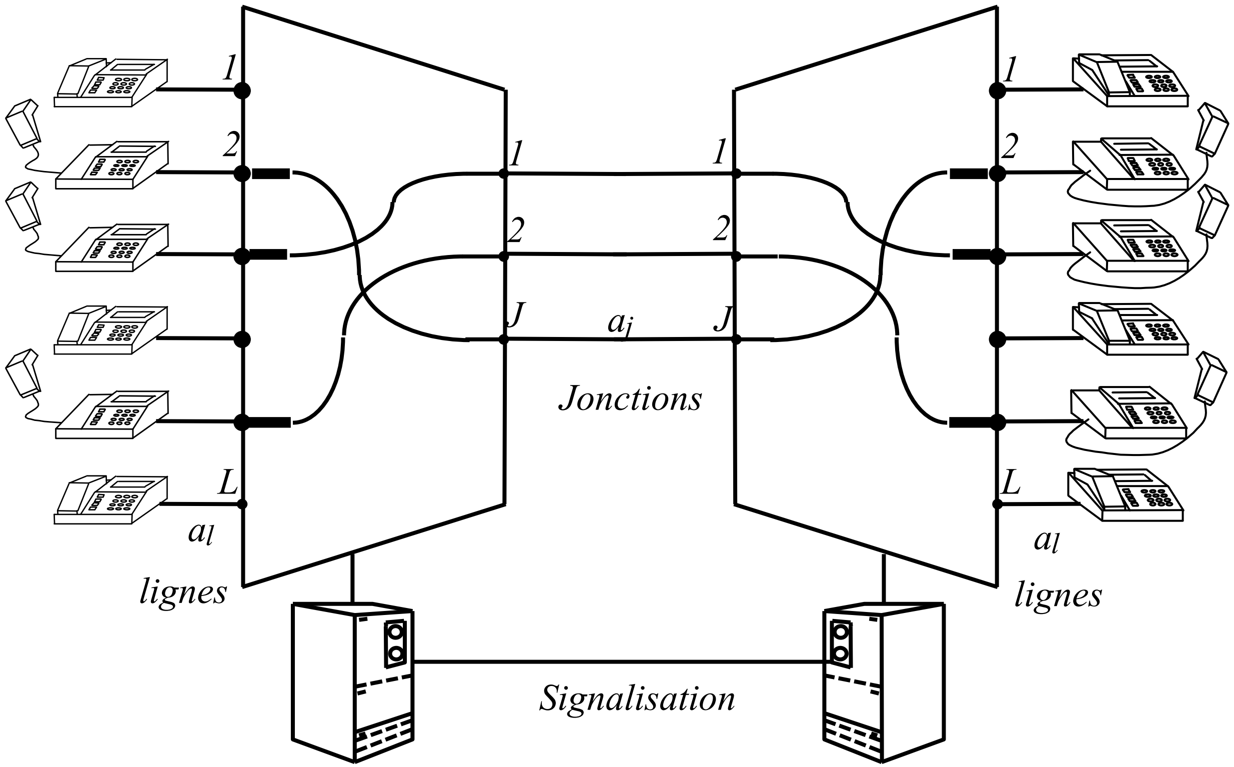
Les processus à perte

Processus à pertes

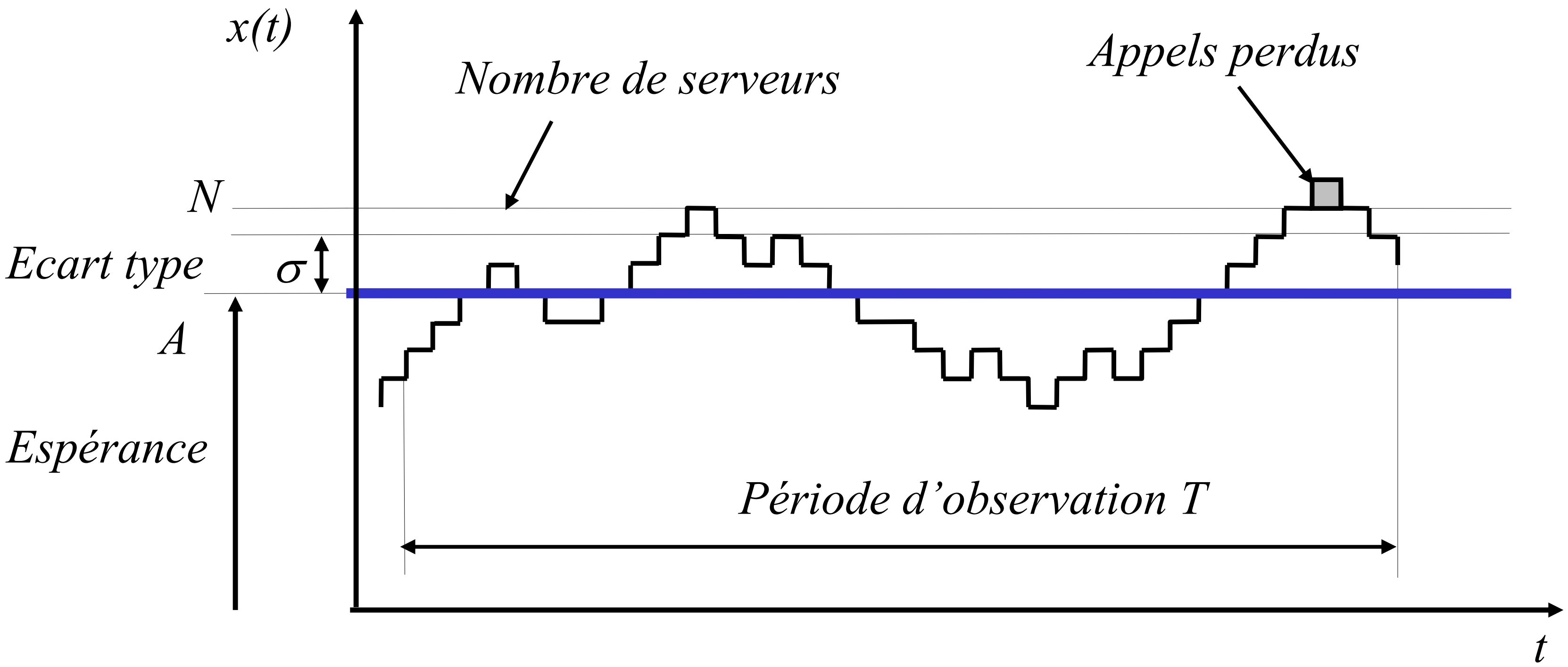


Connexions

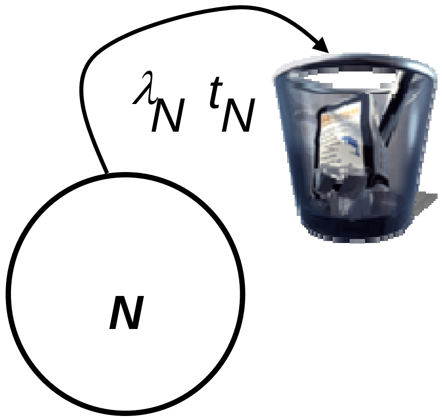
- Les serveurs sont les jonctions (trunks). Combien en faut-il ?



Combien de serveurs ?



Probabilité de perte et congestion temporelle



$$B = \frac{n_B}{n} \qquad n_B = (L-N)\frac{\nu}{T-\nu_C\tau}t_N \,,\; \nu = \frac{n}{L} \,,\; \nu_C = \frac{n_C}{L}$$

$$\frac{n_B}{n} = \frac{(L-N)}{L}\frac{1}{\left(1-\frac{n_C}{L}\frac{\tau}{T}\right)}\frac{t_N}{T}$$

$$\boxed{B = \frac{L-N}{L-A_C}\pi(N)}$$

La probabilité de perte est égale à la congestion temporelle dans le cas des clients Poissoniens

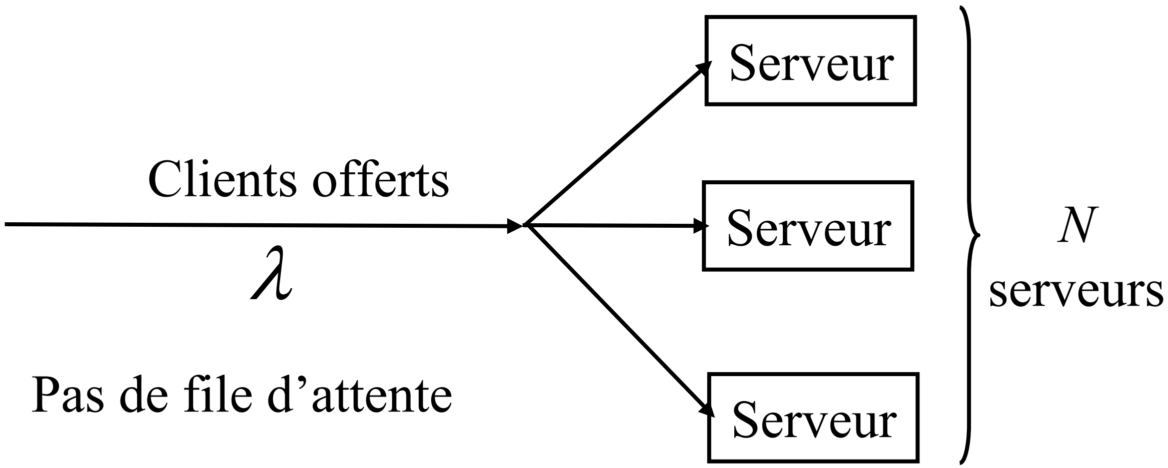
$$B = \frac{L - N}{L - A_C} \pi(N)$$

- Clients sont Poissoniens $\Rightarrow L$ est infini et : $B \to \pi(N)$
- Une autre façon de le démontrer : $\lambda = \frac{n}{T}$ est indépendant de N

$$n_B = \lambda t_N = \frac{n}{T} t_N \;\Rightarrow\; B = \frac{n_B}{n} = \frac{t_N}{T} = \pi(N)$$

Le système M/M/N/N

- Il y a N serveurs et on ne peut pas avoir plus de N clients dans le système (pas de file d'attente)
- Les clients sont Poissoniens $\Rightarrow \lambda$ est indépendant de x
- Il y a N serveurs de capacité μ indépendante de x



Probabilité d'avoir x clients

- Les clients sont Poissonniens $\Rightarrow$ λ est indépendant de x $\Rightarrow$ $\pi(x) = \frac{A^x}{x!}\pi(0)$
- Il ne peut pas y avoir plus de N clients dans le système (pas de file d'attente)

$$\pi(0)\sum_{i=0}^{N}\frac{A^i}{i!} = 1 \Rightarrow \pi(0) = \frac{1}{\sum_{i=0}^{N}\frac{A^i}{i!}}$$

$$\boxed{\pi(x) = \frac{\frac{A^x}{x!}}{\sum_{i=0}^{N}\frac{A^i}{i!}}}$$

La première loi d'Erlang ou loi d'Erlang "B"

- Les clients sont Poissoniens $\Rightarrow$ L est infini $\Rightarrow$ $B = \pi(N)$

- On obtient : $$\boxed{B = \mathrm{E}_{1,N}(A) = \frac{\dfrac{A^N}{N!}}{\displaystyle\sum_{x=0}^{N} \frac{A^x}{x!}}}$$

- Les Européens appellent cette loi "la première loi d'Erlang", les Américains l'appellent "loi d'Erlang B"

Inversion de la loi d'Erlang B

- On peut écrire une petite boucle de programme pour tirer profit de la relation de récurrence :

$$\boxed{\frac{1}{E_{1,N}(A)} = 1 + \frac{N}{A} \times \frac{1}{E_{1,N-1}(A)}}$$

- Sachant que : $E_{1,0}(A) = 1$

La règle de Rigault

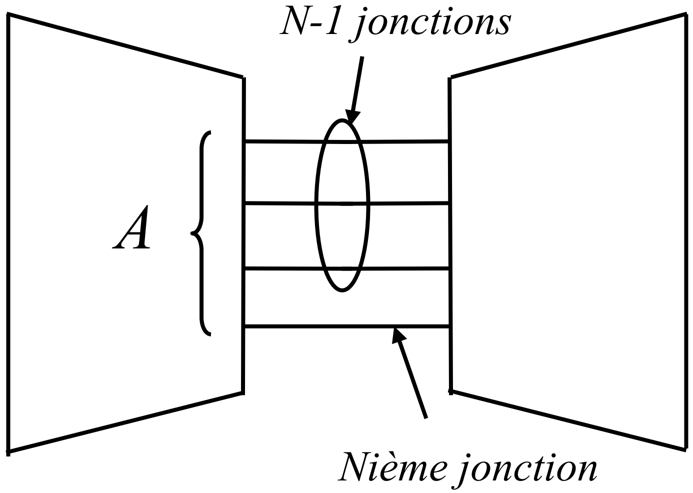
$$B = E_{1,N}(A) = \frac{\frac{A^N}{N!}}{\sum_{i=0}^{N} \frac{A^i}{i!}}$$

$$\boxed{\text{Si } B = 10^{-k} \quad \text{alors} \quad N \cong A + k\sqrt{A}}$$

Excellente approximation à quelques pour cents près (par excès)

Trafic écoulé par la *Nième* jonction

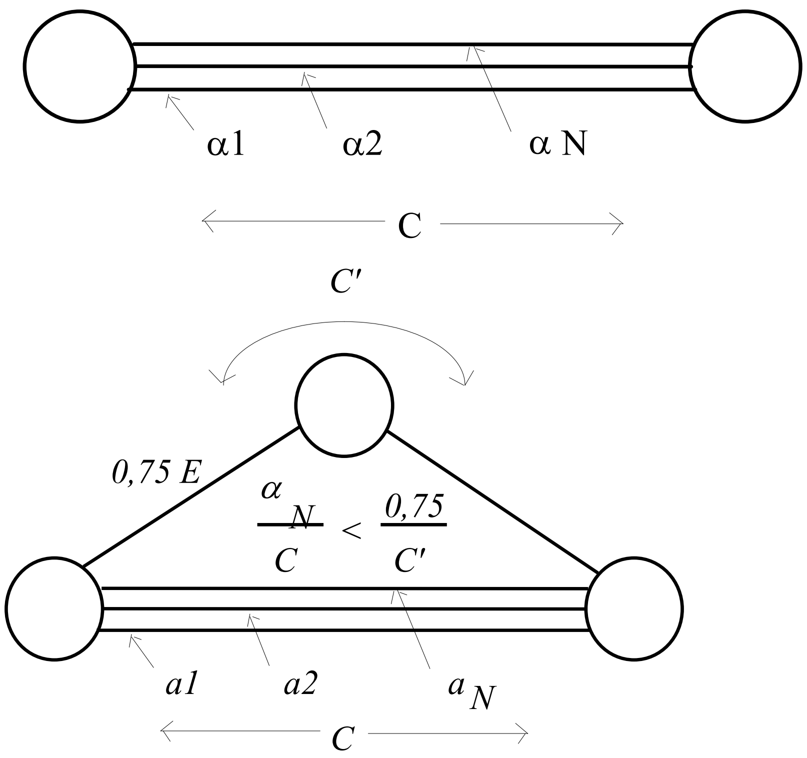
- La *Nième* jonction écoule le trafic perdu par le groupe des *N-1* premières jonctions et perd le trafic perdu par le groupe des *N* jonctions



$$\rho_N = A\left[E_{N-1}(A) - E_N(A)\right]$$

La règle de l'utilité marginale (principe de Moe)

- Equation ECCS



$$\frac{1}{E_{1,N}(A)} = 1 + \frac{N}{A}\frac{1}{E_{1,N-1}(A)}$$

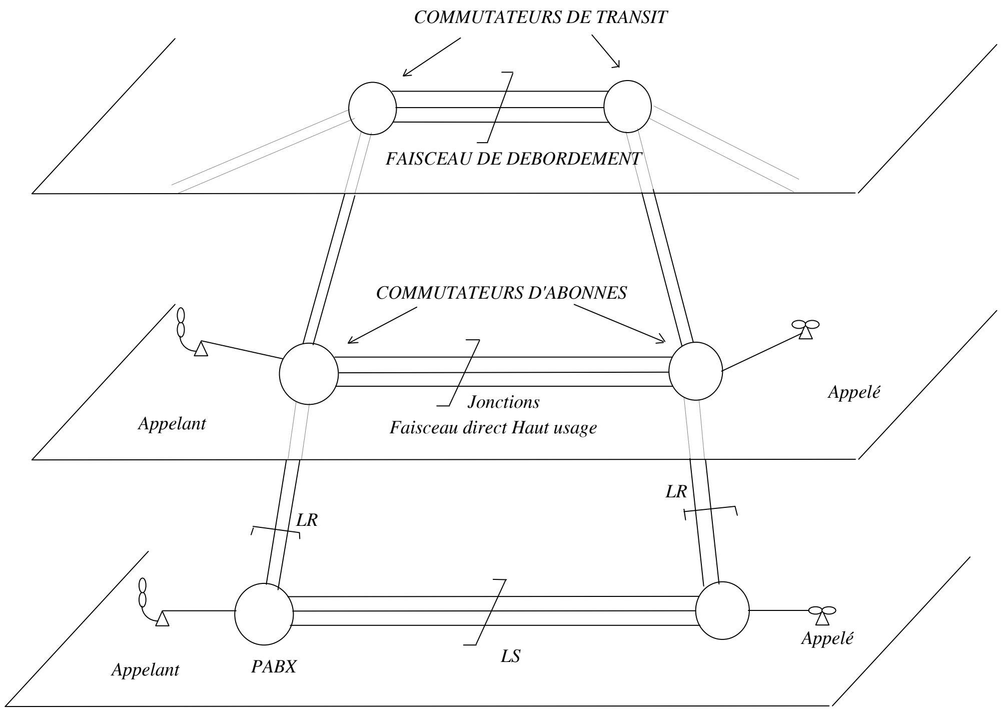
$$\alpha_N = \frac{E_{1,N}(A)}{1 - E_{1,N}(A)}\Big(N - A\big(1 - E_{1,N}(A)\big)\Big)$$

$$u_N = \frac{\alpha_N}{C} \cong \frac{(N - A)\,E_{1,N}(A)}{C}$$

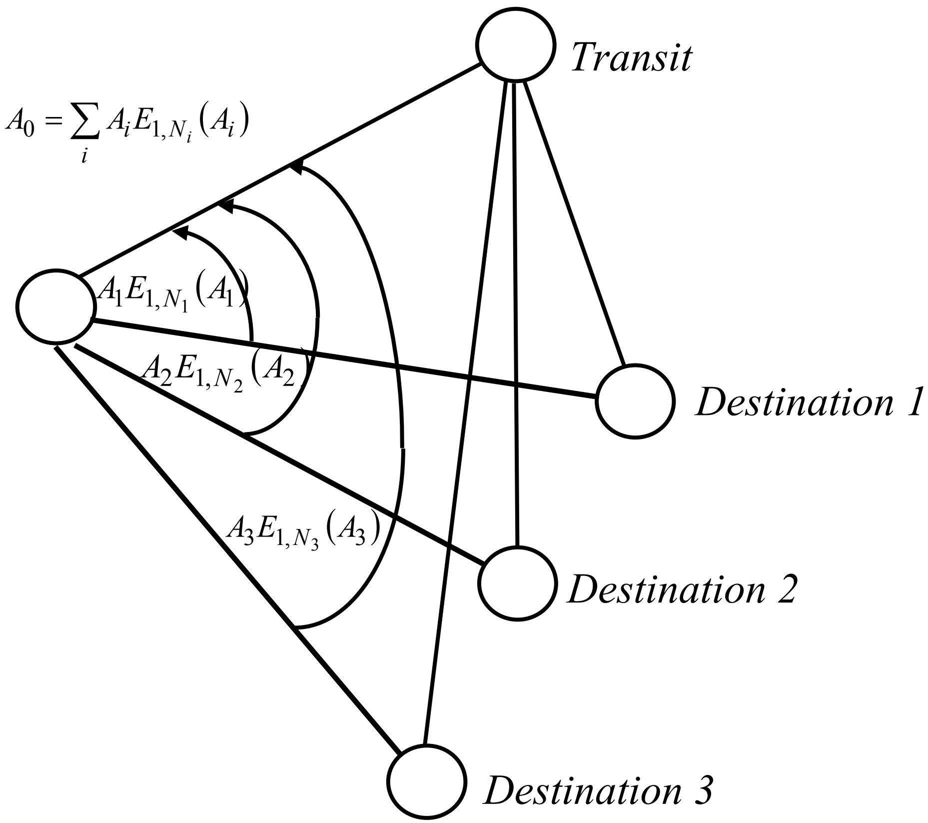
$$\boxed{u_N = \frac{\alpha_N}{C} \leq \frac{0{,}75}{C'}}$$

Construction du réseau : réseaux de transit

- La construction du réseau se fait en fonction des trafics à écouler



Trafic de débordement



Faisceau équivalent de Wilkinson (1)

- On cherche un faisceau unique qui aurait un trafic de débordement dont la moyenne et la variance serait la même que celles que nous obtenons (*o* pour Overflow) : $A_0 = \sum_i A_i E_{1,N_i}(A_i)$
- La variance du trafic perdu sur un faisceau où le trafic offert est A est :

$$\mathrm{var}_B(A) = A_B\left[1 - A_B + \frac{A}{1 + N + A_B - A}\right]$$

- La variance totale du trafic de débordement est donc (nous ne le démontrons pas) :

$$\mathrm{var}_0(A_0) = \sum_i A_{B_i}\left[1 - A_{B_i} + \frac{A}{1 + N + A_{B_i} - A}\right]$$

Faisceau équivalent de Wilkinson (2)

- Les formules de Rapp donnent le trafic et le nombre de jonctions d'un faisceau unique à trafic Poissonien qui aurait le même trafic de débordement et la même variance de débordement (*e* pour Équivalent):

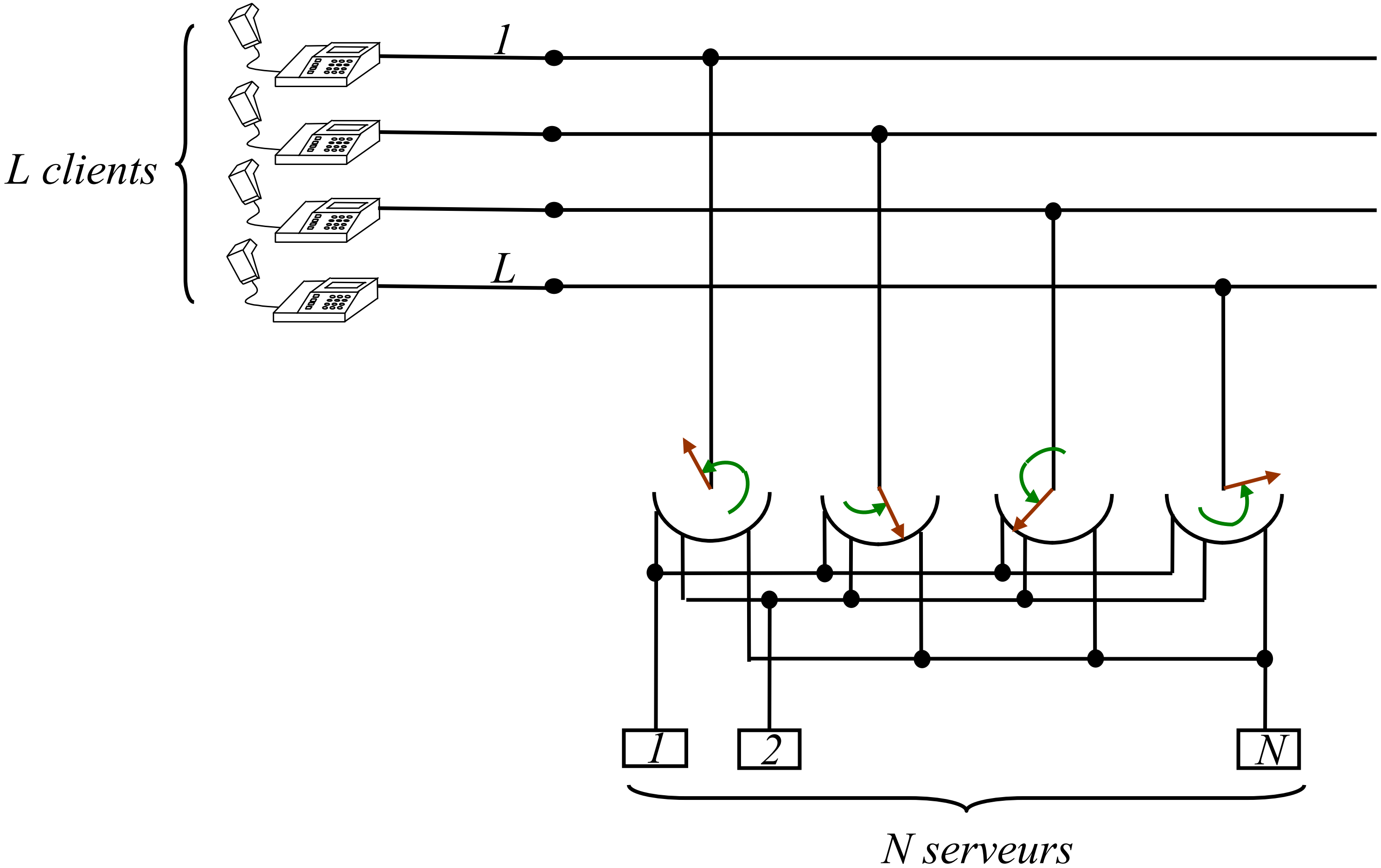
$$\left.\begin{aligned} A_e &= \mathrm{var}_0(A_0) + \left[\left(\frac{3\,\mathrm{var}_0(A_0)}{A_0}\right)\left(\frac{\mathrm{var}_0(A_0)}{A_0} - 1\right)\right] \\ N_e &= \frac{A_e}{1 - \dfrac{1}{A_0 + \dfrac{\mathrm{var}_0(A_0)}{A_0}}} - A_0 - 1 \end{aligned}\right\} \text{Faisceau équivalent}$$

- Le principe de la méthode de Wilkinson est de dire que si l'on équipe le faisceau de débordement de ***m*** jonctions la perte d'appel serait la même que la perte d'Erlang appliquée à un faisceau équivalent qui aurait $\boldsymbol{N_e + m}$ jonctions

Diapositive volontairement vide

Les processus à attente

Processus à attente



Paramètres intéressants

- Processus à perte :

  Probabilité de perte : $B = \frac{n_B}{n} < \varepsilon$

- Processus à attente :

  Probabilité qu'un certain pourcentage de clients dépasse un certain délai :

$$P\left(\frac{n(t > \tau)}{n}\right) < \varepsilon$$

Probabilité d'attente

- On appelle Probabilité d'attente (Delai) la fraction des arrivées qui doit attendre : $$D = \frac{n_D}{n}$$
- Les arrivées qui doivent attendre sont les arrivées qui ont lieu quand tous les serveurs ont occupés $$n_D = \lambda \sum_{x \geq N} t_x = \frac{n}{T} \sum_{x \geq N} t_x$$
- Ou : $n_D = n \sum_{x \geq N} \pi(x)$
- et : $$\boxed{D = \sum_{x \geq N} \pi(x)}$$

Deux espérances différentes pour le temps d'attente

- Soit ***E(w)*** l'espérance du temps d'attente (n'incluant pas le temps dans le serveur) moyennée sur toutes les ***n*** arrivées (même celles qui sont servies sans attente)
- Soit ***E(w$_D$)*** l'espérance du temps d'attente , moyennée seulement sur les ***n$_D$*** arrivées qui subissent de l'attente
- La somme des attentes est : $nE(w) = n_D E(w_D) \Rightarrow E(w) = \frac{n_D}{n} E(w_D)$
- et : $\boxed{E(w) = DE(w_D)}$

Classification des systèmes de files d'attente

- Ces systèmes sont désignés par la notation de Kendall* :
- Processus Clients/Processus Service /Nombre de Serveurs/Occupation maximale de la queue/taille de la population de clients/Discipline de service

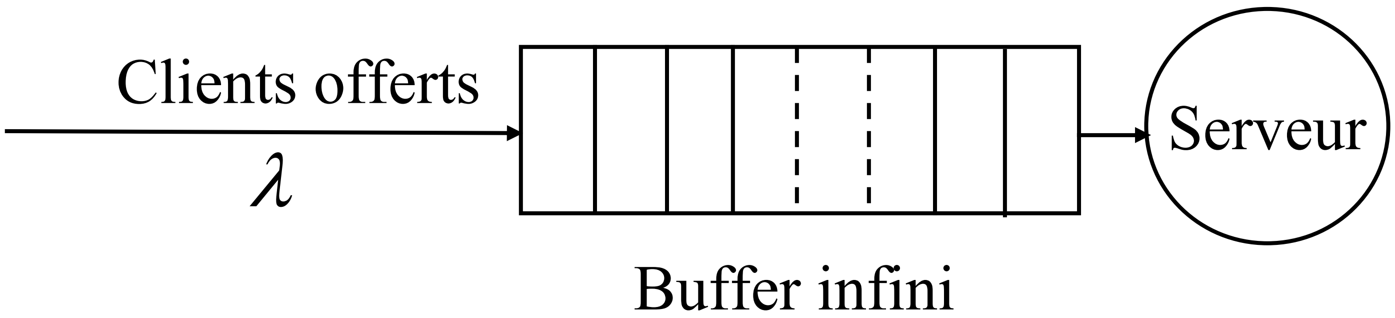
- M = exponentiel
- D = Déterministe
- G = Général

- Quand l'occupation maximale (et la population de clients) n'est pas mentionnée cela veut dire qu'elle est illimitée (le nombre de clients peut être infini). Si la discipline de service est FCFS (First Come First Serve) on ne met rien

❖ David George Kendall (1918 –2007) était un mathématicien et statisticien anglais de l'université de Cambridge

La file M/M/1

- Les clients sont Poissoniens $\Rightarrow \lambda$ est indépendant de x
- Il n'y a pas de perte $\Rightarrow \lambda_C = \lambda$
- Il n'y a qu'un serveur de capacité μ indépendante de x :
- (le serveur ne se dépêche pas plus si la file est longue)



Équation d'équilibre pour la file M/M/1

- L'équation d'équilibre devient : $\lambda t_{x-1} = \mu t_x$

  ou : $\pi(x) = \rho\, \pi(x-1)$

  et : $\pi(x) = \rho^x\, \pi(0)$

- On obtient *P(0)* par : $\sum_{x=0}^{\infty} \pi(x) = \pi(0) \sum_{x=0}^{\infty} \rho^x = 1$

- Donc Si $\rho < 1$, la somme $1 + \rho + \rho^2 + \rho^3 + \ldots$ vaut $\frac{1}{1-\rho}$
- Et $\pi(0) = (1 - \rho)$

$$\boxed{\pi(x) = \rho^x (1 - \rho)}$$

- Il s'agit d'une loi ***géométrique du premier type***

Espérance et variance du nombre de clients dans le système M/M/1 (file + serveur)

- Nous calculons en annexe l'espérance d'une loi géométrique de deuxième type : $P(x) = p^x(1-p) = qp^x$ où $q = (1-p)$

$$E(x) = \frac{p}{1-p}$$

- Ici $p=\rho$ donc l'espérance du nombre de clients dans le système est :

$$\boxed{E(x) = \frac{\rho}{1-\rho} = \frac{\lambda}{\mu - \lambda}}$$

- La variance et l'écart type sont :

$$Var(x) = \frac{\rho}{(1-\rho)^2}$$

$$\sigma = \frac{\sqrt{\rho}}{(1-\rho)}$$

Temps de séjour moyen dans la file M/M/1

- Selon la formule de Little, le temps de séjour moyen dans le système (temps dans la file + temps de service) est :

$$E(t) = \frac{E(x)}{\lambda} = \frac{1}{\lambda}\frac{\lambda}{\mu-\lambda} = \frac{1}{\mu - \lambda}$$

- $\boxed{E(t) = \frac{1}{\mu - \lambda}}$ et aussi : $\boxed{E(t) = \frac{1}{1-\rho}\tau}$

- L'attente moyenne (dans la file uniquement) est :

$$\boxed{E(w)_{M/M/1} = E(t) - \tau = \frac{\rho}{1-\rho}\tau}$$

Nombres moyens de clients dans la file et dans le serveur

- L'espérance du nombre de clients total dans le système (serveur + file) est :

$$\boxed{E(x) = \frac{\rho}{1-\rho}}$$

- L'espérance du nombre de clients dans la file est : $n_D = \lambda E(w) = \lambda \frac{\rho}{1-\rho} \tau$

- Ou : $\boxed{n_D = \frac{\rho^2}{1-\rho}}$

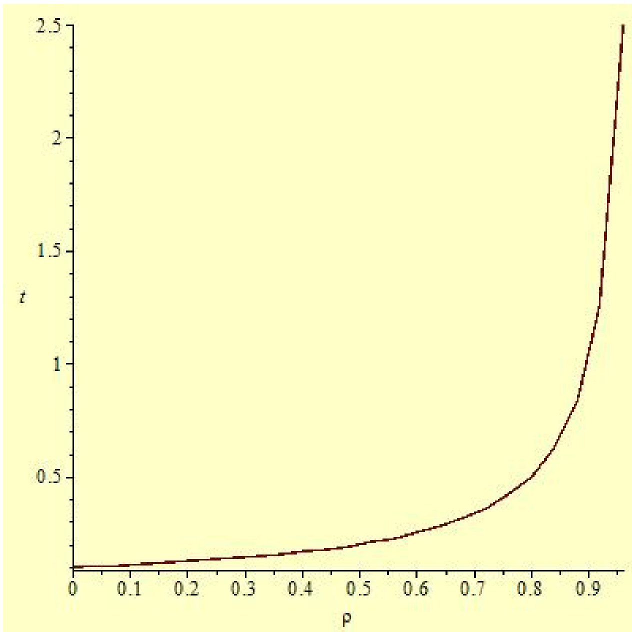
- L'espérance du nombre de clients dans le serveur est :

$$E(n_s) = E(x) - n_D = \frac{\rho}{1-\rho} - \frac{\rho^2}{1-\rho} = \frac{\rho(1-\rho)}{1-\rho} = \rho$$

- On aurait pu retrouver ce résultat par la formule de Little : $\boxed{E(n_s) = \lambda\tau = \rho}$

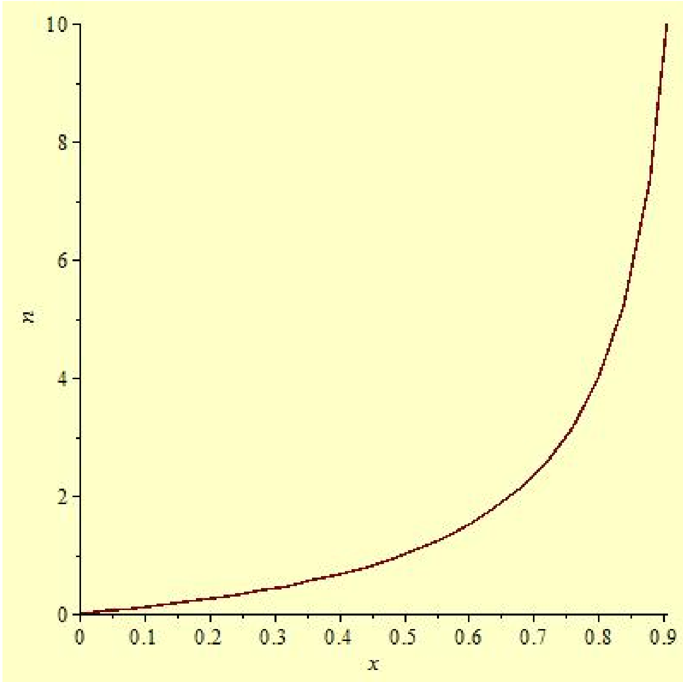
Sensibilité du temps de séjour au trafic

- Il est impossible d'avoir un trafic de 1.
- τ=0,1 : Une variation de charge de 0,7 à 0,9 Erlangs multiplie le délai par 3 !
- Une variation de charge de 0,1 à 0,3 Erlangs multiplie le délai par 1,2



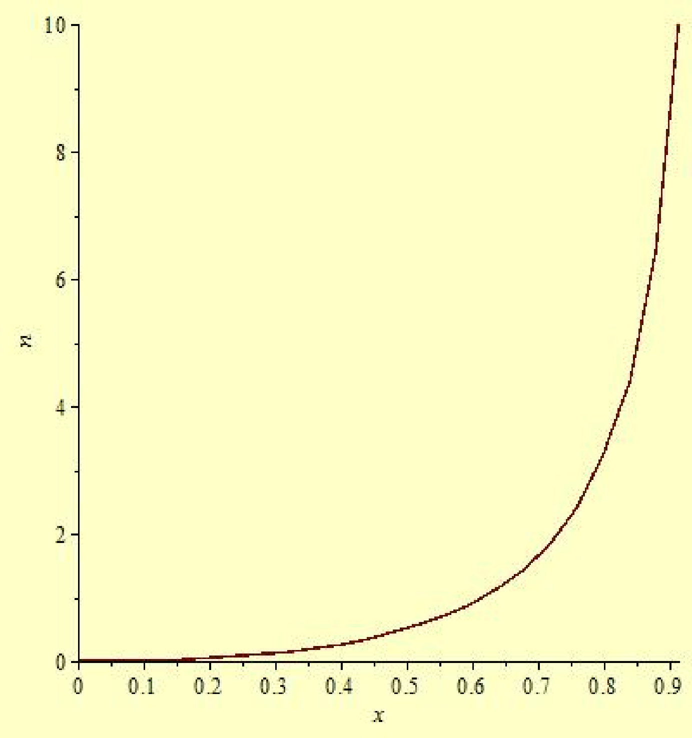
Sensibilité du nombre de clients total au trafic

- Le nombre total de clients (file + serveur) : $\boxed{c = E(x) = \frac{\rho}{1-\rho}}$



Sensibilité du nombre de clients en attente au trafic

- Le nombre de clients en attente (dans la file) : $\boxed{n_D = \dfrac{\rho^2}{1 - \rho}}$



Multiplexage statistique

- Le multiplexage statistique donne toute la capacité du canal à tout le monde. L'espérance du délai est :

$$E(t_S) = \frac{1}{\mu - \lambda}$$

- En TDM on divise la capacité en N intervalles de temps de capacité

$$\mu_T = \frac{\mu}{N}$$

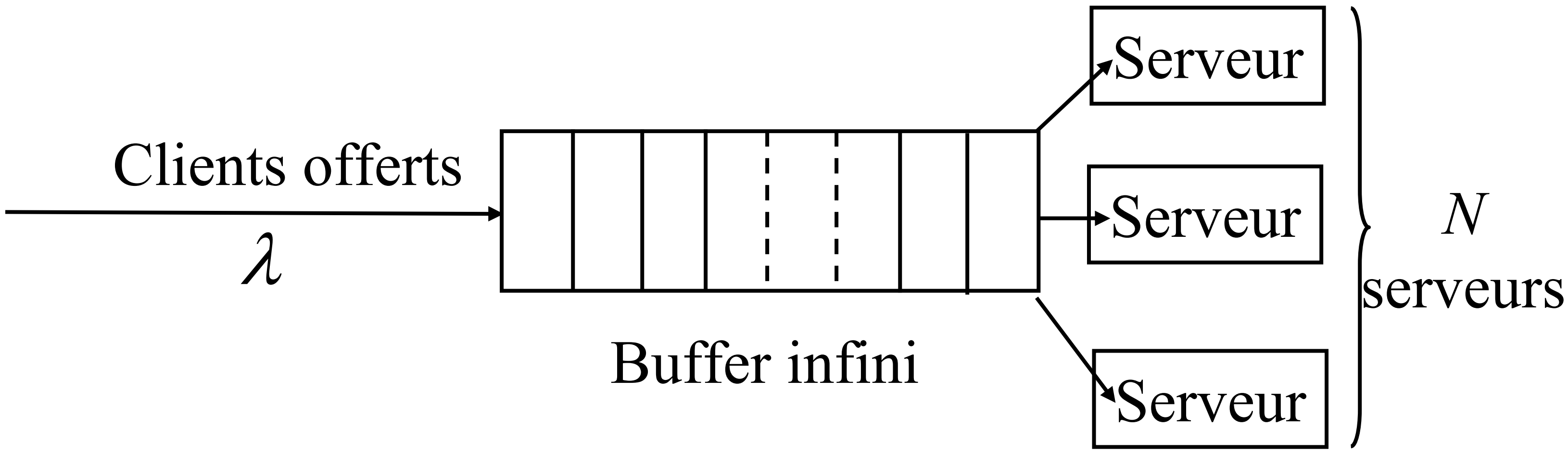
- Dans ces intervalles le taux d'arrivée est : $\lambda_T = \frac{\lambda}{N}$

- Le délai en TDM devient : $E(t_T) = \dfrac{1}{\dfrac{\mu}{N} - \dfrac{\lambda}{N}} = NE(t_S)$

- Le délai est N fois plus petit en multiplexage statistique !

La file M/M/N

- Les clients sont Poissoniens $\Rightarrow \boldsymbol{\lambda}$ est indépendant de $\boldsymbol{x}$
- Il n'y a pas de perte $\Rightarrow \boldsymbol{\lambda_C} = \boldsymbol{\lambda}$
- Il y a N serveurs de taux $\boldsymbol{\mu}$ indépendant of x

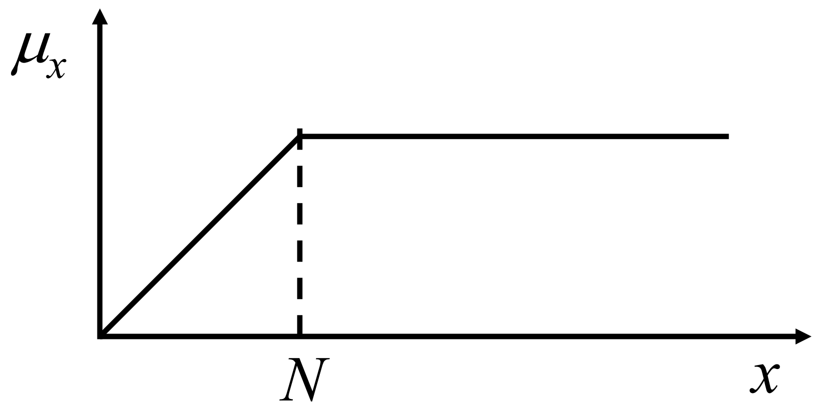


Taux de service dans la file M/M/N

- Il y a 2 cas :

1) Si $x \le N$ alors $\mu_x = \dfrac{x}{\tau} = x\mu$ où µ est le taux de service d'un serveur

2) Si $x > N$ alors $\mu_x = \dfrac{N}{\tau}$



Equations d'équilibre pour la file M/M/N

1) Si $x \leq N$ alors $\lambda\, t_{x-1} = x\mu\, t_x$ et $\pi(x) = \frac{A}{x}\pi(x-1)$

ou $\pi(x) = \frac{A^x}{x!}P(0)$

2) Si $x = N + j > N$ alors $\lambda_{x-1}\, t_{x-1} = N\mu\, t_x$ et

$$\pi(N+j) = \frac{A}{N}\pi(N+j-1)$$

ou $\pi(N+j) = \left(\frac{A}{N}\right)^{j} \pi(N)$

$\pi(0)$ dans la file M/M/N

- On obtient $\pi(0)$ par : $\sum_{x=0}^{\infty} \pi(x) = 1$
- Ce qui donne : $\sum_{x=0}^{\infty} \pi(x) = \pi(0) \sum_{\mathrm{x}=0}^{N} \frac{A^x}{x!} + \pi(N) \sum_{j=1}^{\infty} \left(\frac{A}{N} \right)^j = 1$

ou $\pi(0) \sum_{\mathrm{x}=0}^{N} \frac{A^x}{x!} + \pi(N) \left(\frac{A}{N-A} \right) = 1$

mais $\pi(N) = \frac{A^N}{N!} \pi(0)$ donc : $\pi(0) \left(\sum_{\mathrm{x}=0}^{N-1} \frac{A^x}{x!} + \frac{N}{N-A} \times \frac{A^N}{N!} \right) = 1$

et : $$\pi(0) = \frac{1}{\sum_{\mathrm{x}=0}^{N-1} \frac{A^x}{x!} + \frac{N}{N-A} \times \frac{A^N}{N!}}$$

Probabilités d'état dans la file M/M/N

1) Si $x \le N$

$$\pi(x) = \frac{\dfrac{A^x}{x!}}{\displaystyle\sum_{x=0}^{N-1} \frac{A^x}{x!} + \frac{N}{N-A} \times \frac{A^N}{N!}}$$

2) Si $x = N + j > N$

$$\pi(N+j) = \frac{A^N}{N!} \; \frac{\left(\dfrac{A}{N}\right)^{j}}{\displaystyle\sum_{x=0}^{N-1} \frac{A^x}{x!} + \frac{N}{N-A} \times \frac{A^N}{N!}}$$

Probabilité d'attente

- Nous avons défini la probabilité d'attente "Delay probability *D*" comme la fraction des arrivées qui subissent de l'attente :

$$D = \frac{n_D}{n}$$

- Les arrivées retardées sont celles qui se produisent quand tous les serveurs sont occupés :

$$n_D = \lambda \sum_{x \geq N} t_x = \frac{n}{T} \sum_{x \geq N} t_x$$

- ou :

$$n_D = n \sum_{x \geq N} \pi(x)$$

- soit :

$$\boxed{D = \sum_{x \geq N} \pi(x)}$$

La loi d'Erlang "C"

- La probabilité d'attente D est : $D = \sum_{x \geq N} \pi(x)$ et $\pi(N+j) = \pi(N)\left(\frac{A}{N}\right)^{\mathrm{j}}$
- On obtient : $D = \sum_{j=0}^{\infty} \pi(N+j) = \pi(N) \sum_{j=0}^{\infty} \left(\frac{A}{N}\right)^{j} = \frac{N}{N-A}\pi(N)$
- Ou : $$\boxed{D = \mathrm{E}_{2,N}(A) = \frac{\frac{N}{N-A} \times \frac{A^N}{N!}}{\sum_{\mathrm{x}=0}^{N-1} \frac{A^x}{x!} + \frac{N}{N-A} \times \frac{A^N}{N!}}}$$
- Aussi appelée "la deuxième loi Erlang"

Relation entre les 2 lois ‘Erlang

- Posons $S_{N-1} = \sum_{i=0}^{N-1} \frac{A^i}{i!}$ alors : $B = \dfrac{\frac{A^N}{N!}}{S_{N-1} + \frac{A^N}{N!}}$
- Et $S_{N-1} = \left(\frac{1-B}{B}\right)\frac{A^N}{N!}$
- En introduisant cette expression dans celle de D :

$$D = \frac{\frac{N}{N-A}\frac{A^N}{N!}}{\frac{1-B}{B}\frac{A^N}{N!} + \frac{N}{N-A}\frac{A^N}{N!}} = \frac{\frac{N}{N-A}}{\frac{1-B}{B} + \frac{N}{N-A}} = \frac{NB}{N-A+AB}$$

$$\boxed{D = \frac{NB}{N-A+AB}}$$

- Normalement AB est très inférieur à N-A donc : $\boxed{D \cong \frac{N}{N-A}B}$

- La probabilité d’attente est plus grande que la probabilité de perte
- En appliquant la règle de Rigault : $D \cong \left(1 + \frac{\sqrt{A}}{k}\right)10^{-k}$

Espérance du nombre de clients en attente dans la file M/M/N

- Le nombre moyen de clients en attente est :

$$E(j) = \sum_{j=1}^{\infty} j\pi(N+j) = \pi(N)\sum_{j=1}^{\infty} j\left(\frac{A}{N}\right)^j = \frac{\frac{A}{N}}{\left(1-\frac{A}{N}\right)^2}\pi(N)$$

- On a montré sur la diapos précédente que : $D = \dfrac{1}{\left(1-\frac{A}{N}\right)}\pi(N)$

- donc : $E(j) = \dfrac{\frac{A}{N}}{\left(1-\frac{A}{N}\right)}D$ et : $\boxed{E(j) = \dfrac{A}{N-A}D}$

Temps d'attente et temps de séjour dans la file M/M/N

- De la formule de Little : $E(j) = \lambda E(w)$

- avec : $E(j) = \dfrac{A}{N-A} D \quad E(w) = \dfrac{E(j)}{\lambda} = \dfrac{\tau}{N-A} D$

- soit : $\boxed{E(w) = \dfrac{\tau}{N-A} D}$

- Le temps de séjour est : $E(t) = E(w) + \tau = \dfrac{\tau}{N-A} D + \tau$

- Nous avons aussi : $\boxed{E(w_D) = \dfrac{\tau}{N-A}}$

Comparaison entre les files M/M/1 et M/M/N

- Pour M/M/N nous avons obtenu : $E(j) = \frac{A}{N-A} D$
- Si ρ est le trafic de chaque serveur $A = N\rho$
- Nous en déduisons : $E(j) = \frac{N\rho}{N - N\rho} D$ ou : $\boxed{E(j) = \frac{\rho}{1-\rho} D}$
- Calculons D si $N=1$. Dans ce cas $A = \rho$

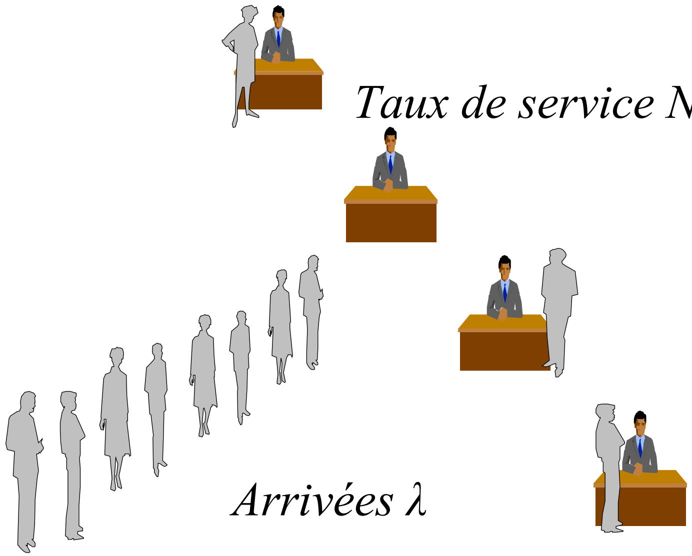
$$D = \frac{\frac{N}{N-A} \times \frac{A^N}{N!}}{\sum_{x=0}^{N-1} \frac{A^x}{x!} + \frac{N}{N-A} \times \frac{A^N}{N!}} = \frac{\frac{1}{1-\rho} \times \rho}{1 + \frac{1}{1-\rho} \times \rho} = \frac{\rho}{1 - \rho + \rho} = \rho$$

- Nous trouvons que, pour M/M/1, $D = \rho$ et comme prévu : $E(j) = n_D = \frac{\rho^2}{1-\rho}$

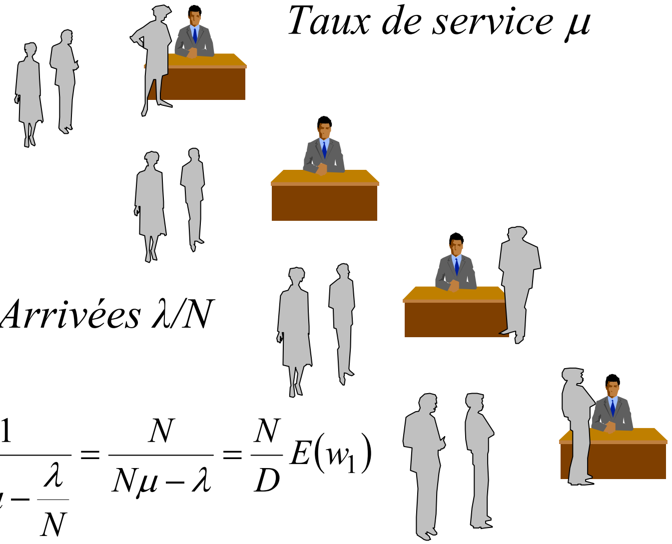
Comparaison entre les files M/M/1 et M/M/N (suite)

- Pour M/M/N nous avons obtenu : $E(w) = \dfrac{\tau}{N - A} D$
- Où $A=\lambda\tau$ avec $\tau=1/\mu$ étant le temps de service <u>*d'un serveur.*</u>
- Mais on a aussi $A=\lambda/\mu$, donc :
- $$E(w) = \frac{1}{\mu\left(N - \frac{\lambda}{\mu}\right)} D = \frac{D}{N\mu - \lambda}$$ ou : $$\boxed{E(w) = \frac{D}{N\mu - \lambda}}$$
- Dans le cas où $N=1$ (file M/M/1) nous avons vu que $D=\rho$
- Alors : $E(w) = \dfrac{\rho}{\mu - \lambda} = \dfrac{\rho}{\mu(1 - \rho)}$ ou : $E(w) = \dfrac{\rho}{(1 - \rho)} \tau$
- Ce qui est bien le résultat que nous avions obtenu pour la file M/M/1

Retour sur l'accessibilité totale



$$E(w_1) = \frac{D}{N\mu - \lambda}$$



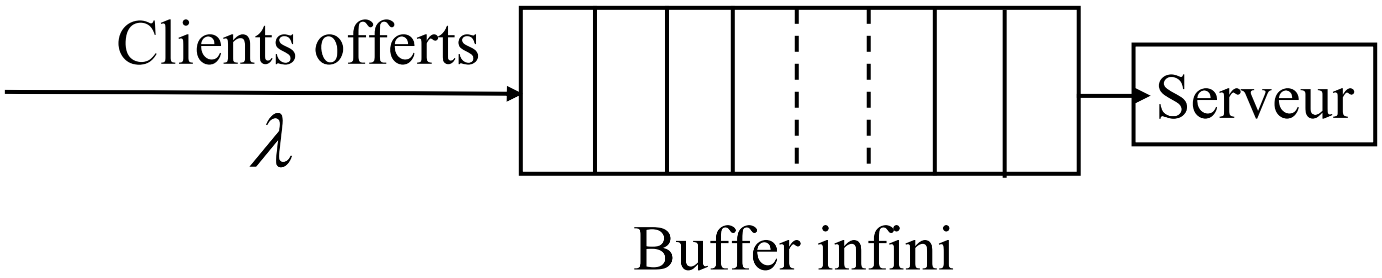
$$E(W_2) = \frac{1}{\frac{N}{N}\mu - \frac{\lambda}{N}} = \frac{N}{N\mu - \lambda} = \frac{N}{D}E(w_1)$$

Bonne solution

Vraiment très mauvaise solution

La file M/G/1

- Les clients sont Poissoniens $\Rightarrow \lambda$ est indépendant de x
- Il n'y a pas de perte $\Rightarrow \lambda_C = \lambda$
- Il n'y a qu'un serveur de capacité μ dont la distribution des temps de service n'est pas connue :

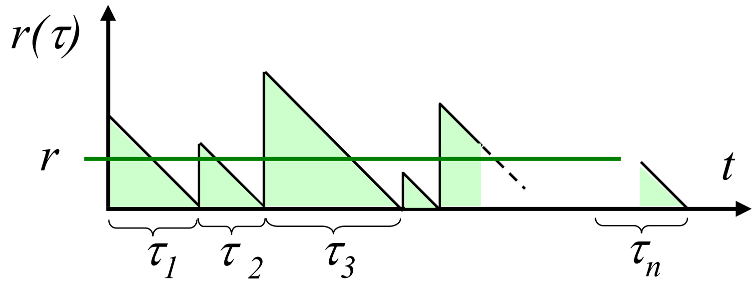


M/G/1 : temps d’attente

- Un client qui arrive attend en moyenne que le nombre moyen de gens dans la file avant lui soient servis ($n\tau$) plus la moyenne du temps résiduel de service r du client en train d’être servi
- $w=n\tau+r$
- Mais d’après la formule de Little $n=\lambda w$
- Donc $w=w\lambda\tau+r = w\rho+r$
- et $$w=\frac{r}{1-\rho}$$

M/G/1 : calcul du temps résiduel

- Calculons l’espérance des temps résiduels à chaque fois qu’un nouveau client arrive



- L’espérance des temps résiduels est :

$$r = \frac{1}{t}\int_0^t \tau(t)dt = \frac{1}{t}\sum_1^n \frac{1}{2}\tau_i^2$$

- (L’intégrale est la somme des surfaces des triangles)
- Le nombre ***n*** de triangles pendant ***t*** est ***n=λt***
- donc :

$$r = \frac{1}{t}\sum_1^n \frac{1}{2}\tau_i^2 = \frac{n}{t}\frac{1}{2}\frac{1}{n}\sum_1^n \tau_i^2$$

Formule de Pollaczek Khinchin

- Quand $t \to \infty$ $\quad \dfrac{n}{t} \to \lambda$
- Et la moyenne des temps résiduels est : $\quad r = \lambda \dfrac{1}{2}\dfrac{1}{n}\sum_{1}^{n} \tau_i^2 = \lambda \dfrac{E(\tau^2)}{2}$
- Le temps d'attente est : $\quad E(w) = \dfrac{r}{1-\rho} = \dfrac{\lambda}{1-\rho}\dfrac{E\left(\tau_i^2\right)}{2}$
- Mais d'après la formule de la variance : $\quad E\left(\tau^2\right) = \sigma^2 + \left(E(\tau)\right)^2$
- Donc : $E(w) = \dfrac{\lambda}{1-\rho}\dfrac{\left(\sigma^2+\tau^2\right)}{2} = \dfrac{\lambda}{1-\rho}\dfrac{\tau^2\left(1+\dfrac{\sigma^2}{\tau^2}\right)}{2} = \dfrac{\rho}{1-\rho}\dfrac{\left(1+\dfrac{\sigma^2}{\tau^2}\right)}{2}\tau$
- On pose $\boldsymbol{C_v = \sigma/\tau}$ (Coefficient de variation) : $\quad \boxed{E(w) = \dfrac{\rho}{1-\rho}\dfrac{\left(1+C_v^2\right)}{2}\tau}$

Nombre de clients en attente dans la file M/G/1

- On utilise la formule de Little $E(n) = \lambda E(w)$

$$E(n) = \lambda E(w) = \frac{\rho}{1-\rho} \frac{\left(1 + C_v^2\right)}{2} \lambda\tau$$

$$\boxed{E(n) = \frac{\rho^2}{1-\rho} \frac{\left(1 + C_v^2\right)}{2}}$$

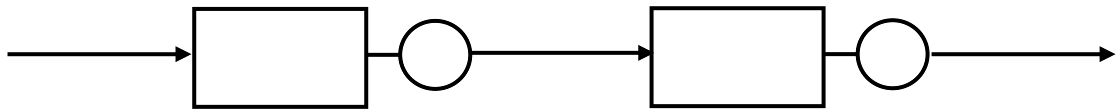
Applications de la formule de Pollaczek Kinchin

- Pour des temps de service Poissoniens (File M/M/1) : $\boldsymbol{\sigma = \tau}$
- On retrouve : $E(w)_{M/M/1} = \dfrac{\rho}{1-\rho}\dfrac{(1+1)}{2}\tau = \dfrac{\rho}{1-\rho}\tau$
- Pour des temps de service déterministes : $\boldsymbol{\sigma = 0}$
- On trouve : $E(w)_{M/D/1} = \dfrac{\rho}{1-\rho}\dfrac{(1+0)}{2}\tau = \dfrac{\rho}{1-\rho}\dfrac{\tau}{2}$
- *Pour un trafic donné le temps d'attente minimum est donné dans un réseau où la taille des paquets est fixe (ATM)*

Les réseaux de files d’attente

Le théorème de Burke

- Soient 1file M/M/1 en série avec une autre file ?/M/1

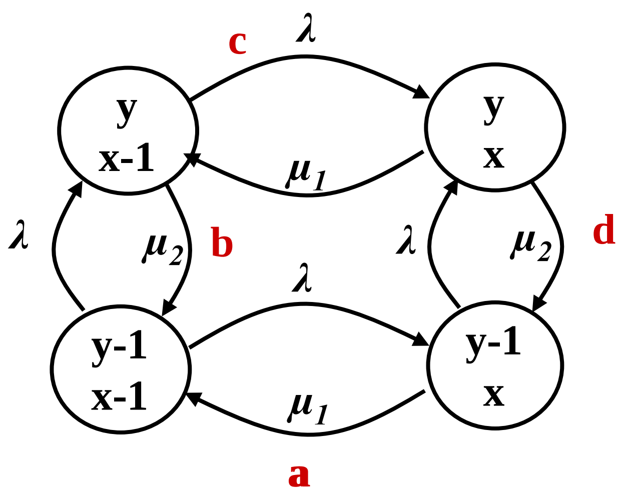


- La première file a un taux d'arrivée $\boldsymbol{\lambda}$ et un taux de service $\boldsymbol{\mu_1}$
- La deuxième file a un taux de service $\boldsymbol{\mu_2}$
- *A l'équilibre, il sort autant de client qu'il en entre par unité de temps. Les arrivée dans la deuxième file sont également Poissoniennes de même taux* $\boldsymbol{\lambda}$.
- Donc la deuxième file est également une file M/M/1 et avec le même taux d'arrivées que la première file.

$$\rho_1 = \frac{\lambda}{\mu_1} \qquad \rho_2 = \frac{\lambda}{\mu_2}$$

Equations d'équilibre

- x est le nombre de clients dans la première file, y est le nombre de clients dans la deuxième file
- Il y a 4 équations d'équilibre



a $\lambda\pi(x-1, y-1) = \mu_1\pi(x, y-1)$

b $\lambda\pi(x-1, y-1) = \mu_2\pi(x-1, y)$

c $\lambda\pi(x-1, y) = \mu_1\pi(x, y)$

d $\lambda\pi(x, y-1) = \mu_2\pi(x, y)$

Forme produit

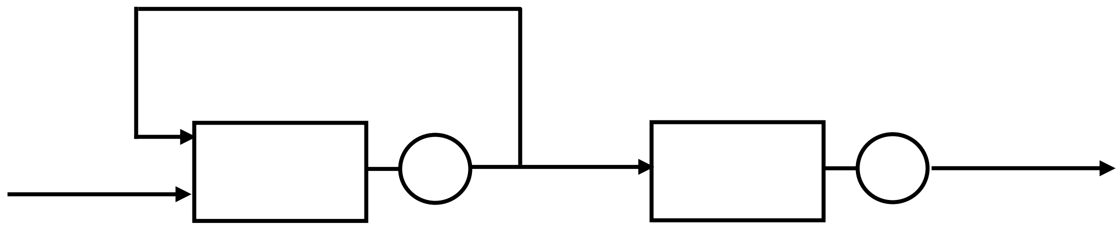
- On considère que les 2 processus sont indépendants, donc :

$$P(x, y) = \pi(x)\pi(y) = (1 - \rho_1)\rho_1^x (1 - \rho_2)\rho_2^y$$

- On constate que cette forme produit satisfait les 4 équations d'équilibre

Réseau de Jackson

- Lorsqu'un client peut visiter la même file d'attente plusieurs fois, le processus d'arrivée n'est plus Poissonien.



- Il y a des arrivées extérieures et des rebouclages intérieurs

Réseau de Jackson

1) Toutes les arrivées extérieures doivent être Poissoniennes
2) Tous les temps de services doivent suivre une distribution exponentielle.
3) Toutes les files d'attente doivent être de capacité infinie.
4) Quand un client quitte un nœud, la probabilité qu'il aille à un autre nœud doit être indépendante de son histoire passée et indépendante de la localisation d'un autre client.

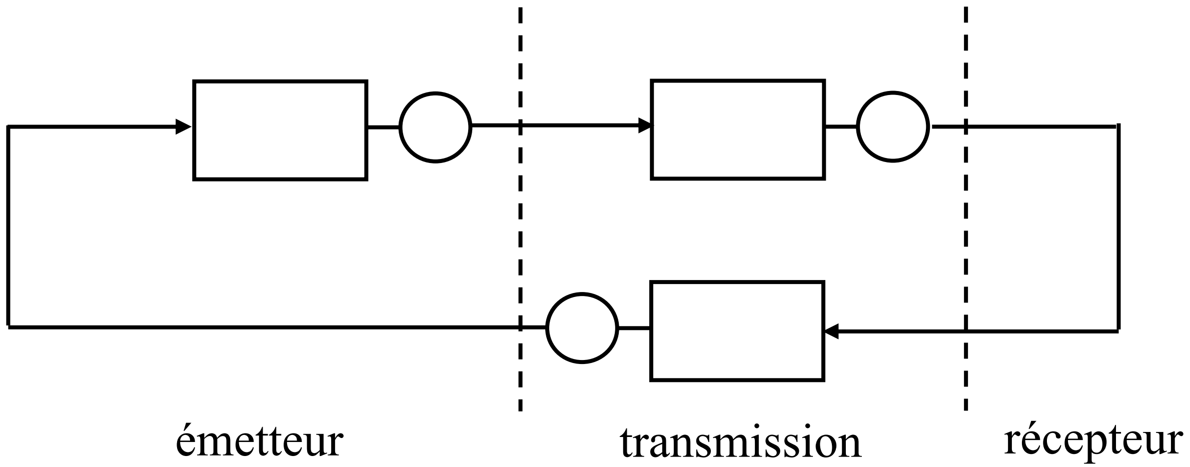
Le théorème de Jackson

1. Le nombre de clients dans la file d'attente de chaque nœud est indépendant des autres nœuds et les arrivées peuvent être considérées comme Poissoniennes.
2. Les probabilités stationnaires de chaque noeuds sont celles d'une file $M/M/1$ ou $M/M/s$.
3. Les délais moyens à chaque nœud peuvent être additionnés pour déterminer le délai total de traversée du réseau.

Dans ces conditions le réseau est à forme produit

Réseau de Jackson fermé

- Exemple d'un système de transmission



- Dans les réseaux de Jackson fermés le calcul de la constante de normalisation ($\pi(o)$) est difficile. On utilise la méthode MVA (Mean Value Analysis).

Simulation

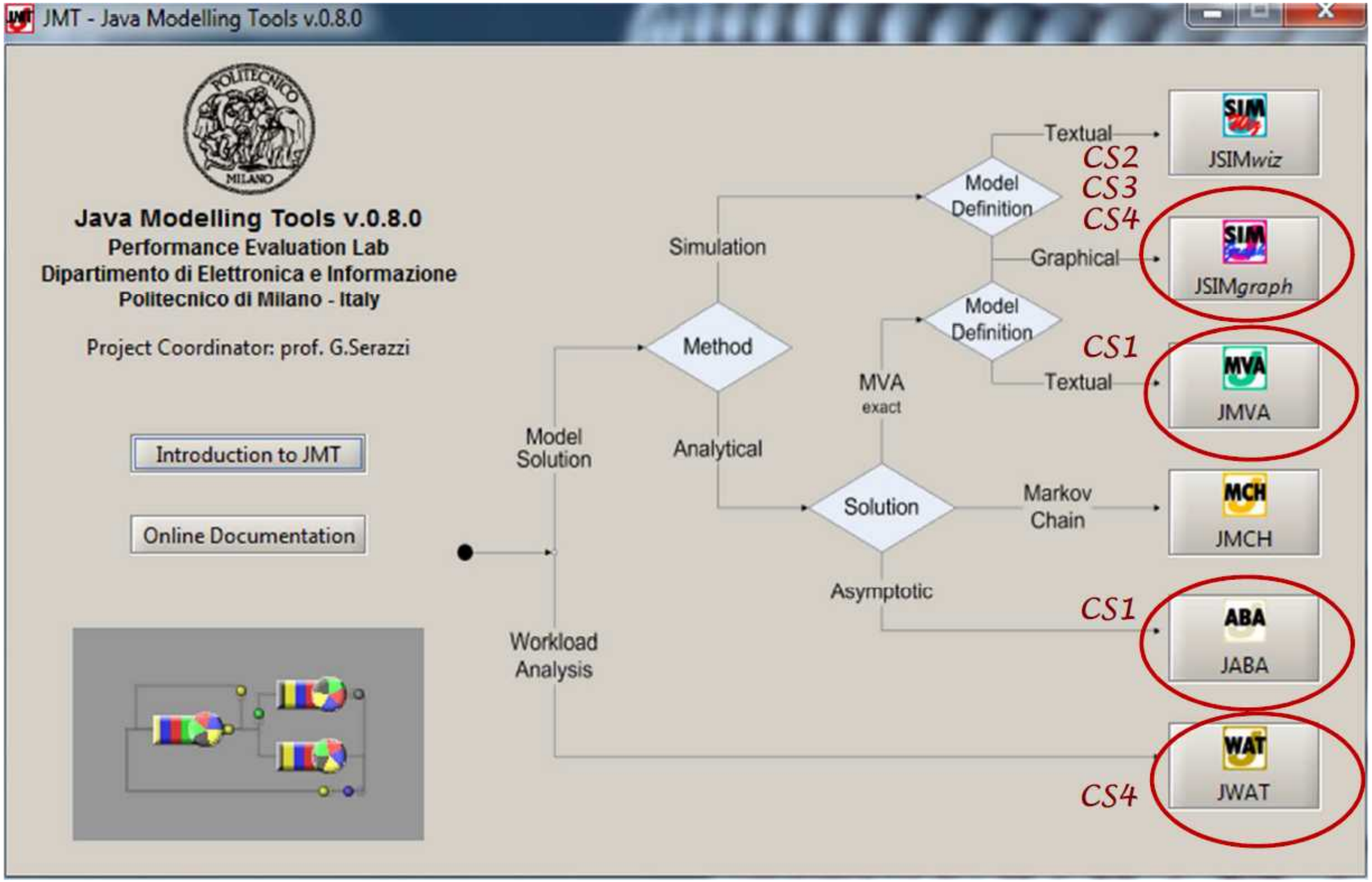
Principaux programmes de simulation (1)

- *Il est déconseillé d’écrire son propre simulateur. On en trouve de très bons et gratuits.*
- Opnet (coûteux), dispose de bibliothèques de modèles préconstruits. (www.opnet.com)
- NS : Network Simulator. Apprécié du monde académique et gratuit. Mais ne tourne que sur Linux. (www.isi.edu/nsnam)
- OMNeT++ : gratuit sur Windows, interface graphique. Semblable à Opnet mais pas de bibliothèques préconstruites. Plutôt orienté simulation de protocoles. (www.omnetpp.org)
- En France Modline/Qnap2 a été développé par l’INRIA et Bull dans les années 80. Qnap2 est le moteur de simulation et Modline est un interface graphique. Malheureusement il n’est plus maintenu.

JMT : Java Modelling Tools

- JMT (Java Modeling Tools) : un logiciel plus récent mais qui suit le même modèle que QNap2, développé par Polytechnique de Milan. (http://jmt.sourceforge.net/)
- Gratuit sur Windows, interface graphique. Orienté simulation de réseaux, bien adapté à la pédagogie et la recherche.
- Download installer :

  http://sourceforge.net/projects/jmt/files/jmt/JMT-0.8.0/JMT-installer-0.8.0.jar/download
- JMT user manual : http://jmt.sourceforge.net/Papers/JMT_users_Manual.pdf
- Une bonne petite vidéo pour démarrer : http://www.youtube.com/watch?v=YuJ3eDTYAp0


JMT - Java Modelling Tools v.0.8.0
Java Modelling Tools v.0.8.0
Performance Evaluation Lab
Dipartimento di Elettronica e Informazione
Politecnico di Milano - Italy
Project Coordinator: prof. G.Serazzi
Introduction to JMT
Online Documentation
Model Solution
Workload Analysis
Method
Simulation
Analytical
Model Definition
Textual
Graphical
MVA exact
Solution
Markov Chain
Asymptotic
CS2
CS3
CS4
CS1
CS1
CS4
JSIMwiz
JSIMgraph
JMVA
JMCH
JABA
JWAT

Installation de JMT

- Il faut avoir Java sur sa machine. L'installer d'abord.
-
- Ensuite télécharger JMT-installer-0.8.0.jar
- Le ranger dans le répertoire racine de la machine C:
- (pas dans program file)
-
- Passe en console de commande et taper :
-
- java -jar JMT-installer-0.8.0.jar
-

JMT: définir le système à simuler

- -Ouvrir JSIMGraph
- File New
- -Construire le système à simuler
- Par exemple 3 sources (Source 0, Source 1, Source 2, une file d'attente Queue Stat 0, 1 puit, Sink 0)
-

Définir les trafics

- -Définir les trafics : dans Class faire Add Class
- Class0 type : open puis cliquer sur edit : définir lamda=3 pour distribution exponentielle, cliquer sur mean qui va se mettre automatiquement à 1/3, puis OK, Reference Station : affecter à source 0.
- Refaire Add Class
- Class1 type : open puis cliquer sur edit : définir lamda=3 pour distribution exponentielle, cliquer sur mean qui va se mettre automatiquement à 1/3, puis OK, Reference Station : affecter à source 1 en cliquant sur le nom de la source
- Refaire Add Class   Class 2 type : open puis cliquer sur edit : définir lamda=3 pour distribution exponentielle, cliquer sur mean qui va se mettre automatiquement à 1/3, puis OK, Reference Station : affecter à source 1 en cliquant sur le nom de la source
- Finir la définition des trafics par Done

Définir ce qu'on veut observer

- Définir ce que vous voulez observer en cliquant sur Define Performance Indices :
- Pour Number of Customers choisir All Classes, Cliquer sur Station/Region sélectionner Queue Stat 0 pour que le nombre de clients observés soit celui dans la file, prendre un intervalle de confiance de 0.95 et une erreur max de 0.05
- Ajouter un autre paramètre à observer : sélectionner Queue Time puis clique sur Add Selected index, Station/Region Queue Stat 0, même intervalle de confiance et erreur
- Finir la définition de ce que vous voulez observer par Done

Définir les paramètres de simulation

- -Définir les paramètres de la simulation en cliquant sur Define Simulation Parameters
- Simulation random seed : definition du nombre servant à générer des nombres aléatoires. Cocher Random
- Maximum Duration : Décocher infinite, mettre 600
- Maximum number of Samples : décocher no automatic stop, mettre 1000000
- Cocher Animation. Mettre 2 secondes.
- Done

- - surtout de pas Cliquer sur What if . l'analyse What if enclanche N simulations successives avec à chaque fois une augmentation du trafic. Pour faire des courbes des paramètres observés en fonction du trafic.
- C'est intéressant mais cela ne nous intéressera que plus tard.

Définir les paramètres du modèle

- - Définir les paramètres du modèle. Cliquer sur Define Default's values of model parameters
- Station parameters : Vérifier que l'on a bien :
- Queue Capacity 1 (pas de file d'attente dans les sources)
- Queue Strategy FCFS (First Come First Serve)
- Drop Rule : Drop
- Simulation parameters : ne pas toucher, on retrouve ce qu'on a déjà défini
- Finite Capacity Region parameters : vérifier que le paramètre Drop est bien sur True
- Okay

Définir les paramètres de la file d’attente

- Maintenant sur le graphe Sélectionner la file d'attente et double cliquer dessus
- Aller à l'onglet service section pour définir le taux de service (mu) pour chaque trafic (classe)
- Pour chacune des classes que nous avons défini cliquer sur édit. le paramètre est noté lambda mais nous devons comprendre mu. Mettre la valeur 10.
- Nous aurons ainsi un trafic total du serveur de 3*3/10 = 0.9 Erlang
- Done

La simulation est prête

- On peut la lancer
- Quand elle se termine, vérifiez que l'on a bien une espérance du nombre de clients dans la file de 9 et une espérance du temps d'attente de 0,9 seconde conforme à la théorie

Annexes

Quelques sommes

- Rappelons que $\dfrac{1}{1-q} = \sum_{x=0}^{\infty} q^x$
- En dérivant une fois par rapport à *q :* $\dfrac{1}{(1-q)^2} = \sum_{x=1}^{\infty} x q^{x-1}$
- En dérivant encore une fois par rapport à *q :*

$$\frac{2(1-q)}{(1-q)^4} = \frac{2}{(1-q)^3} = \sum_{x=1}^{\infty} x(x-1) q^{x-2}$$

Espérance de la loi de Bernoulli

- Méthode de force brute :

$$E(x) = \sum_{x=0}^{N} x C_N^x p^x (1-p)^{N-x} = \sum_{x=1}^{N} x \frac{N!}{x!(N-x)!} p^x (1-p)^{N-x} = \sum_{x=1}^{N} \frac{N!}{(x-1)!(N-x)!} p^x (1-p)^{N-x}$$

- Posons *y=x-1*, *(x=y+1)*

$$E(x) = Np \sum_{y=0}^{N-1} \frac{(N-1)!}{(y)!(N-1-y)!} p^y (1-p)^{N-1-y}$$

- Or d'après la formule du binôme :

$$\sum_{y=0}^{N-1} \frac{(N-1)!}{(y)!(N-1-y)!} p^y (1-p)^{N-1-y} = (p + (1-p))^{N-1} = 1$$

- Donc :

$$\boxed{E(x) = Np}$$

Variance de la loi de Bernoulli

- Méthode de force brute, en posant comme pour l'espérance *y=x-1*, *(x=y+1)* :

$$Var(x)=\left(\sum_{x=0}^{N} x^2 C_N^x p^x (1-p)^{N-x}\right)-(Np)^2 = Np\sum_{y=1}^{N-1}(1+y)\frac{(N-1)!}{(y)!(N-1-y)!}p^y(1-p)^{N-1-y}$$

$$Var(x)=Np+Np(N-1)p-(Np)^2=Np-Np^2$$

- Donc :

$$\boxed{\begin{array}{l} Var(x)=Np(1-p) \\ \text{et} \\ \sigma=\sqrt{Np(1-p} \end{array}}$$

Espérance de loi géométrique du premier type

$$P(x) = p^x(1-p) = qp^x \quad \text{où} \;\; q = (1-p)$$

$$E(x) = \sum_{1}^{\infty} xqp^x = q\sum_{x=1}^{\infty} xp^x = qp\sum_{x=1}^{\infty} xp^{x-1}$$

- D'après la somme déjà calculée :

$$E(x) = qp\frac{1}{(1-p)^2} = \frac{p}{(1-p)}$$

$$\boxed{E(x) = \frac{p}{1-p}}$$

Variance de la loi géométrique du premier type

- Avec cette loi, nous avons : $Var(x) = \sum_{1}^{\infty} x^2 p^x q - \left(\frac{p}{1-p}\right)^2$

- D'après la somme déjà calculée :

$$\frac{2p^2q}{(1-p)^3} = \sum_{x=1}^{\infty} x(x-1)p^x q = \sum_{x=1}^{\infty} x^2 p^x q - \sum_{x=1}^{\infty} x p^x q = E\left(x^2\right) - \frac{p}{1-p}$$

- Donc : $E\left(x^2\right) = \frac{2p^2q}{q^3} + \frac{p}{q} = \frac{2p^2q + pq^2}{q^3} = \frac{2p^2 + pq}{q^2}$

- Et : $Var(x) = E\left(x^2\right) - \frac{p}{q^2} = \frac{2p^2 + pq - p^2}{q^2} = \frac{p^2 + p(1-p)}{q^2} = \frac{p}{q^2}$

$$\boxed{Var(x) = \frac{p}{(1-p)^2} \qquad \sigma = \frac{\sqrt{p}}{1-p}}$$

Espérance de loi géométrique du deuxième type

$$P(x) = p(1-p)^{x-1} = pq^{x-1} \quad \text{où } q = (1-p)$$

$$\bar{x} = \sum_{1}^{\infty} xpq^{x-1}$$

- D'après la somme déjà calculée :

$$\bar{x} = \sum_{1}^{\infty} xpq^{x-1} = p\sum_{1}^{\infty} xpq^{x-1} = p\frac{1}{(1-q)^2} = \frac{1}{p}$$

$$\boxed{\bar{x} = \frac{1}{p}}$$

Variance de la loi géométrique du deuxième type

- Rappelons nous que $$Var(x)=\sum_{1}^{\infty} x^2 pq^{x-1}-\frac{1}{p^2}$$

- D'après la somme déjà calculée :

$$\frac{2pq}{(1-q)^3}=pq\sum_{x=1}^{\infty} x(x-1)q^{x-2}=\sum_{x=1}^{\infty} x^2 pq^{x-1}-\sum_{x=1}^{\infty} xpq^{x-1}=E\left(x^2\right)-\frac{1}{p}$$

- Donc : $$E\left(x^2\right)=\frac{2pq}{p^3}+\frac{1}{p}=\frac{2pq+p^2}{p^3}=\frac{1-q+2q}{p^2}=\frac{1+q}{p^2}$$

- Et : $$Var(x)=E\left(x^2\right)-\frac{1}{p^2}=\frac{1+q-1}{p^2}=\frac{q}{p^2}$$

$$\boxed{\begin{aligned} Var(x)&=\frac{1-p}{p^2} \\ \sigma&=\frac{\sqrt{1-p}}{p}\end{aligned}}$$

Bibliographie

Bibliographie générale 1

- Toutes les références font appel à la théorie des chaines de Markov, et très souvent à la transformée en z et à la transformée de Laplace ce qui exige un niveau mathématique assez élevé
- Le grand classique : Leonard Kleinrock : Queueing Systems, Volume 1 :Theory.
- Une référence que j'ai trouvé bien faite : J. Virtamo. Queueing Course, from Finland. http://www.netlab.hut.fi/opetus/s383143/kalvot/english.shtml
- Un excellent livre de probabilités que j'avais ramené des USA et dont je me sers tout le temps : Alberto Leon Garcia : Probability and Random Processes for Electical Engineering. Addison-Wesley 1994

Bibliographie générale 2

- Pour aborder les chaines de Markov de manière très simple : Seymour Lipschutz : Theory and problems of Probability, Schaums's Outline Series, McGraw-Hill

- Un cours en français : Fabrice Valois, Modélisation et performance des réseaux, http://fvalois.insa-lyon.fr/courses/mot_1_Chaines_de_Markov_0304.pdf
http://fvalois.insa-lyon.fr/courses/mot_2_Formalisme_FAs_0304.pdf
http://fvalois.insa-lyon.fr/courses/mot_3_Perfs_FAs_0304.pdf
http://fvalois.insa-lyon.fr/courses/mot_4_FAs_simples_0304.pdf
http://fvalois.insa-lyon.fr/courses/mot_5_Reseaux_FA_0304.pdf

Bibliographie pour le calcul de trafic des réseaux à attente (réseaux paquets)

- Un livre qui couvre à la fois les performances des réseaux paquets et circuits : Misha Schwartz : Telecommunication Networks, protocols, Modeling and Analysis, Madison Wesley

- Un livre en français : Bonald Thomas, Feuillet Mathieu, Performances des réseaux et systèmes informatiques, Hermes Lavoisier 2011

- Un livre très complet et pas trop difficile : David McDysan, QoS & Traffic Management in IP & ATM Networks, McGrAw-Hill, 2000

Bibliographie pour le calcul de trafic des réseaux à perte (réseaux circuits)

- Dans votre bibliothèque : Rigault C. Principes de commutation numérique, Hermes, Paris, 1998

- D. Bear. Principles of telecommunications Traffic engineering, Peter Pergrinus Ltd

- Et aussi Mischa Schwartz déjà mentionné

Bibliographie pour la simulation

- Pour JMT : http://jmt.sourceforge.net/Documentation.html

- Pour Omnet : http://www.omnetpp.org/documentation

Pour aller plus loin

- Calculer les réseaux par des bornes supérieures : le Network calculus http://ica1www.epfl.ch/PS_files/netCalBookv4.pdf

- Un site qui donne une grande liste de références : http://web2.uwindsor.ca/math/hlynka/qonline.html