Chapitre 3
Les réseaux de files d'attente

I Généralités

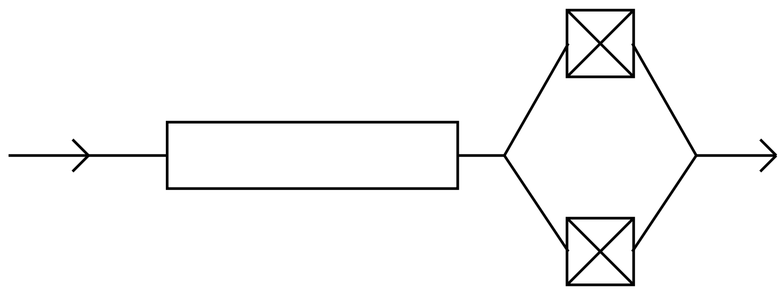
1 Introduction

L'étude théorique des files d'attente a pour origine les problèmes de liaisons téléphoniques que voulait résoudre Erlang en 1906. Il étudiait le dimensionnement des standards téléphoniques et, pour cela, l'effet des fluctuations des demandes de service sur l'utilisation des circuits lorsque le nombre d'usagers, composant un numéro, varie.

L'un des résultats essentiels a été obtenu par Baskett, Chandy, Muntz et Palacios (**BCMP**) à partir du **théorème de Jackson**. Il concerne les réseaux ouverts, fermés ou mixtes avec plusieurs classes de clients. D'autres études ont été faites depuis : méthodes de diffusion, méthodes itératives et méthodes numériques.

2 Définitions

Une station de service est un système aléatoire composé d'une file d'attente et d'un ou plusieurs serveurs exécutant des tâches identiques et généralement de façon identique.



Station de service

Exemples :

station	service	client
circuit téléphonique	mise en liaison de deux abonnés	abonné
feu de signalisation	passage d'un carrefour	véhicule
secrétaire	dactylographie	lettre
mécanicien	réparation	machine à réparer
dock	déchargement	bateau

Une station de service est entièrement définie par la donnée du processus des arrivées, des temps de service, de la discipline, de la capacité de la file d'attente et du nombre de serveurs.
La capacité peut être infinie, ainsi que le nombre de serveurs.

Notation de Kendall A / B / s / d / e

- **A** indique la loi d'entrée,
- **B** la loi de service,
- **s** le nombre de serveurs,
- **d** le nombre maximum de clients dans la station,
- **e** la discipline de la file d'attente.

Symboles utilisés pour les lois

- **M** loi exponentielle,
- **H** loi hyper-exponentielle,
- **E** loi d'Erlang hypo-exponentielle,
- **D** loi déterministe (constante),
- **G** loi générale (caractérisée par son espérance mathématique et sa variance),
- **K** loi à transformée de Laplace rationnelle.

Lorsque la capacité est infinie et la discipline Premier Arrivé Premier Servi **d** et **e** sont facultatifs.

Le processus des arrivées est défini par la loi des inter-arrivées. Ce processus est poissonnien si la loi est exponentielle.

Les disciplines les plus utilisées :

- **PAPS** Premier Arrivé, Premier Servi (ou First In First Out),
- **DAPS** Dernier Arrivé, Premier Servi (ou Last In First Out),
- **DAPP** Dernier Arrivé, Premier servi avec Préemption : lorsqu'un client arrive, il est pris immédiatement en charge ; si un client est en cours de service son service est interrompu et il rejoint la tête de la file d'attente.
- **PS** Processeur Partagé ; chaque client reçoit une fraction du temps de service par unité de temps égale à $1/n$ s'il y a n clients dans le système est un serveur.

La loi de service la plus utilisée est la loi exponentielle. Si X est la variable aléatoire représentant la durée de service et μ le taux de service, on a

$$P[t < X \leq t+dt] = \mu \, e^{-\mu t} \, dt$$
$$E[X] = 1/\mu$$
$$Var[X] = 1/\mu^2$$

Si on définit le carré du coefficient de variation d'une variable aléatoire par

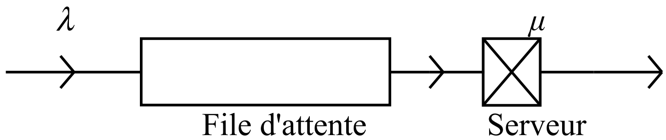
$$CV^2[X] = \frac{Var[X]}{(E[X])^2}$$

On a évidemment, si X suit une loi exponentielle $CV^2[X] = 1$

Si $CV^2[X]$ est nul, le phénomène est régulier et X prend une valeur constante.
Si $CV^2[X]$ est supérieur à 1, le phénomène est hyper-exponentiel ou par grappes, comme un système de communication avec beaucoup de messages courts et de messages longs et très peu entre les deux.
Si $CV^2[X]$ est compris entre 0 et 1, le phénomène est plus ou moins régulier. Le phénomène est hypo-exponentiel et peut être représenté par une loi d'Erlang ; pour servir un client il y a k opérations consécutives, chacune suivant une loi exponentielle de même taux.

3 La file M/M/1

3.1 Généralités



Les clients arrivent de l'extérieur suivant une loi de Poisson de débit λ et viennent se placer dans la file d'attente. Ils sont servis suivant une discipline PAPS. Les temps de service sont des variables aléatoires indépendantes identiquement distribuées selon une loi exponentielle de moyenne $1/\mu$. Lorsque le serveur a terminé un service, le client servi quitte le système. La capacité de la file est supposée infinie.

Soit $p(n,t)$ la probabilité d'avoir n clients dans la file (y compris celui recevant le service) à l'instant t et soit δt un intervalle de temps. On pose o(δt), pour toute fonction de δt , telle que

$$\lim_{\delta t \to 0} \frac{o(\delta t)}{\delta t} = 0$$

Les équations du système à l'instant $t+\delta t$ sont données à partir de son état à l'instant t

$$\begin{aligned} p(n,t+\delta t) = {} & p(n,t)\,(1-(\lambda+\mu)\delta t) + o(\delta t) \\ & + p(n-1,t)\lambda\delta t + o(\delta t) \\ & + p(n+1,t)\mu\delta t + o(\delta t) \end{aligned}$$

En faisant tendre δt vers 0, on obtient les équations différentielles aux différences caractérisant l'évolution du système.

$$\frac{d}{dt} p(n,t) = -(\lambda+\mu)\, p(n,t) + \lambda\, p(n-1,t) + \mu\, p(n+1,t) \quad \text{pour } n > 0$$
$$\frac{d}{dt} p(0,t) = -\lambda\, p(0,t) + \mu\, p(1,t)$$

3.2 Solution stationnaire

L'état d'équilibre est défini par

$$p(n) = \lim_{t \to \infty} p(n,t)$$

où la limite doit être indépendante des conditions initiales $p(n,0)$.
On montre que si cette limite existe, on a

$$\lim_{t \to \infty} p(n,t) = 0$$

Les équations d'équilibre du système s'écrivent donc

(1) $\lambda\, p(n-1) - (\lambda+\mu)\, p(n) + \mu\, p(n+1) = 0$ pour $n > 0$
(2) $-\lambda\, p(0) + \mu\, p(1) = 0$

Ce sont des équations aux différences finies, homogènes, linéaires, à coefficients constants et du premier degré.
L'équation caractéristique du système s'écrit :

$\lambda - (\lambda+\mu)\, r + \mu\, r^2 = 0$

Si $\lambda \neq \mu$, l'équation a deux racines distinctes :

$r_1 = 1$ et $r_2 = \lambda / \mu$

La solution générale est donc :

$p(n) = C_1 + C_2\, r_2^n$ où C_1 et C_2 sont des constantes

L'équation (2) est satisfaite par la solution $C_2\, r_2^n$ donc $C_1 = 0$.
D'autre part, on doit avoir $\lim_{n \to \infty} p(n) = 0$ pour que $p(n)$ soit une distribution de probabilité.
Il est donc nécessaire que $r_2 < 1$, c'est-à-dire $\lambda < \mu$.
C_2 est obtenu en exprimant

$$\sum_{n=0}^{\infty} p(n) = 1$$

On obtient finalement en posant $\rho = r_2 = \lambda / \mu$

$$p(n)=(1-\frac{\lambda}{\mu})(\frac{\lambda}{\mu})^n=(1-\rho)\rho^n \quad \text{pour n} \geq 0$$

La condition $\lambda < \mu$ signifie que le serveur est en mesure d'écouler le trafic.
Le serveur est oisif avec probabilité 1 - ρ et son taux d'utilisation est ρ.
Si on suppose que les clients ayant terminé un service ont une probabilité a de revenir dans la file d'attente et une probabilité $q = 1 - a$ de quitter le système, les équations précédentes donnant $\{p(n)\}_n$ restent valables. Il suffit de remplacer μ par $q\,\mu$.

Critères de performance

Notons N la variable aléatoire nombre de clients dans la station, l'espérance mathématique de N est le nombre moyen de clients dans la station

$$L=E(N)=\sum_{n=1}^{\infty} np(n)=\rho(1-\rho)\sum_{n=1}^{\infty} n\rho^{n-1}=\frac{\rho}{1-\rho}=\frac{\lambda}{\mu-\lambda} \quad \text{en utilisant}$$

$$\sum_{n=0}^{\infty} x^n=\frac{1}{1-x} \quad \text{et} \quad \frac{d}{dx}\sum_{n=0}^{\infty} x^n=\sum_{n=0}^{\infty} n\,x^{n-1}=\frac{1}{(1-x)^2} \quad \text{pour } |x| < 1.$$

Nombre moyen de clients dans la file : $L_q=\frac{\lambda^2}{\mu(\mu-\lambda)}$

Nombre moyen de clients dans la station : $L=L_q+\frac{\lambda}{\mu}=\frac{\lambda}{\mu-\lambda}$

Temps moyen d'attente : $W_q=\frac{L_q}{\lambda}=\frac{\lambda}{\mu(\mu-\lambda)}$

Temps moyen de séjour (réponse) : $W=\frac{L}{\lambda}=\frac{1}{\mu-\lambda}$

Loi de **Little** : $L = \lambda\, W$

3.3 Solution transitoire

Une solution transitoire (dépendant du temps) peut être obtenue par la méthode des fonctions génératrices.

4 Extension de la file M/M/1

4.1 File M/M/1 avec dépendance d'état

Cette file est caractérisée par un processus d'arrivée et une distribution des temps de service qui dépendent du nombre n de clients se trouvant dans la station.

Le **processus d'arrivée est poissonnien** de paramètre $\lambda(n)$ et le **temps de service est exponentiel** de taux $\mu(n)$.

La **loi exponentielle** est "**sans mémoire**" donc il est possible de recommencer ou de continuer le service, lorsqu'un client entre, avec un nouveau paramètre $\mu(n+1)$.
Si l'on n'avait pas deux lois exponentielles il faudrait fixer le temps de service au moment où le client entre dans le guichet de service.

Définition Un processus de naissance et de mort est un processus de Markov pour lequel deux transitions seulement sont possibles

le processus décroît de 1 : **mort**
le processus croît de 1 : **naissance**

Soit N_t le nombre de clients dans la station au temps t.

Propriété : N_t est un processus de naissance et de mort.

Indication pour la démonstration :
Soient V_t date d'arrivée et W_t date de départ

$$P(V_t > t \ \ et \ \ W_t > t \mid N_s = i \ \ et \ \ s \leq t) = e^{-(\lambda(i)+\mu(i))t}$$

c'est une loi exponentielle donc il n'y a pas simultanément arrivée et départ.
On utilise l'indépendance des lois exponentielles d'arrivée et de service.
Probabilité d'une arrivée avant un départ :

$$P(V_t \leq W_t \mid N_s = i \ \ et \ \ s \leq t) = \frac{\lambda(i)}{\lambda(i)+\mu(i)}$$

N_t est un processus de Markov.

Transitions du processus

0 vers 0	$1-\lambda(0)\Delta t$
0 vers 1	$\lambda(0)\Delta t$
i vers i-1	$\mu(i)\Delta t$
i vers i	$1-(\lambda(i)+\mu(i))\Delta t$
i vers i+1	$\lambda(i)\Delta t$

Exercice : Construire le graphe et la matrice P des transitions

On pose $P = I + Q\ \Delta t$

$$Q = \begin{pmatrix} -\lambda(0) & \lambda(0) & 0 & 0 & \dots \\ \mu(1) & -(\mu(1)+\lambda(1)) & \lambda(1) & 0 & \dots \\ 0 & \mu(2) & -(\mu(2)+\lambda(2)) & \lambda(2) & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

Recherche d'une suite de probabilités stationnaire v

v doit vérifier $\begin{bmatrix} vQ = 0 \\ \sum_{i=0}^{\infty} v_i = 1 \end{bmatrix}$

Résolution par récurrence

$$v_i = \frac{\lambda(0)\lambda(1)\dots\lambda(i-1)}{\mu(1)\mu(2)\dots\mu(i)} v_0$$

CNS d'existence :

$$v_0 \neq 0 \ \ et \ \ \sum_{i=0}^{\infty} v_i = 1$$

avec $i = n+1$ ou $i-1 = n$, pour avoir deux "belles formules"

$$v_0 \left(1 + \sum_{n=0}^{\infty} \frac{\lambda(0)\lambda(1)\dots\lambda(n)}{\mu(1)\mu(2)\dots\mu(n+1)}\right) = 1$$

La condition d'ergodicité du processus devient

$$1 + \sum_{n=0}^{\infty} \frac{\lambda(0)\lambda(1)\dots\lambda(n)}{\mu(1)\mu(2)\dots\mu(n+1)} < \infty$$

La distribution limite v renommée p (et $n = i$) devient

$$p(n) = \frac{\lambda(0)\lambda(1)\dots\lambda(n-1)}{\mu(1)\mu(2)\dots\mu(n)} p(0)$$

et

$$p(0) = \left[1 + \sum_{n=1}^{\infty} \frac{\lambda(0)\lambda(1)\dots\lambda(n-1)}{\mu(1)\mu(2)\dots\mu(n)}\right]^{-1}$$

On retrouve les probabilités d'état de la M/M/1 en posant $\lambda(n) = \lambda$ et $\mu(n) = \mu$.

Exercice :
1 - Etudier la file M/M/c à c serveurs. Comparer la M/M/1 et la M/M/2 avec deux serveurs deux fois plus lents.
2 - Comment prendre en compte la capacité limitée à k places ? Etudier la file M/M/1/k.

4.2 File M/M/∞

La file possède un nombre suffisant de serveurs pour que les clients entrant dans la station trouvent toujours un serveur disponible. Il n'y a jamais d'attente.

Exemple : Les terminaux d'une salle de TP.

Le processus d'arrivée est poissonnien de paramètre λ et le temps de service est exponentiel de taux μ par serveur.
N_t est un processus de naissance et de mort de taux :

$$\lambda(i)=\lambda \quad \forall i\geq 0$$

$$\mu(i)=i\mu \quad \forall i\geq 0$$

D'après le résultat obtenu en 4.1

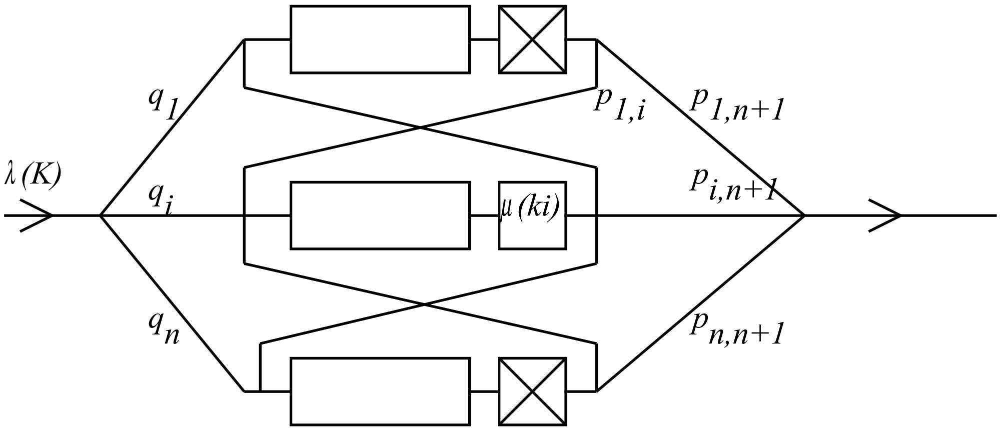
$$p(n)=p(0)\frac{\lambda^n}{n!\mu^n} \quad et \quad 1=p(0)\sum_{n=0}^{\infty}\frac{\lambda^n}{n!\mu^n}$$

d'où $\quad p(0)=e^{-\lambda/\mu} \quad et \quad p(n)=e^{-\lambda/\mu}\frac{\lambda^n}{n!\mu^n}$ distribution de Poisson

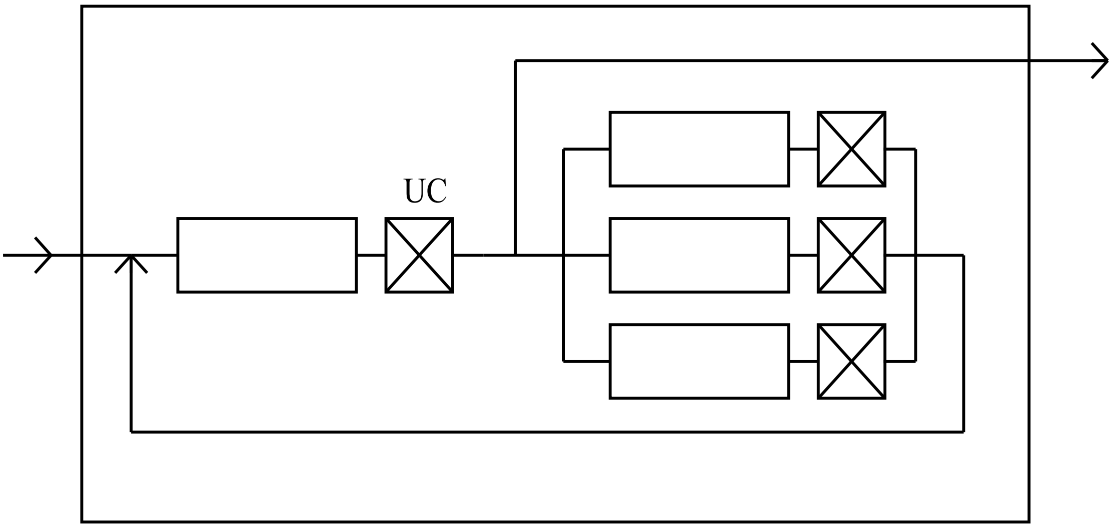
Remarques : Idem pour M/G/∞ car p ne dépend que des valeurs moyennes ; $E(N) = \rho = \lambda / \mu$ nombre moyen de serveurs occupés.

II Les théorèmes de Jackson

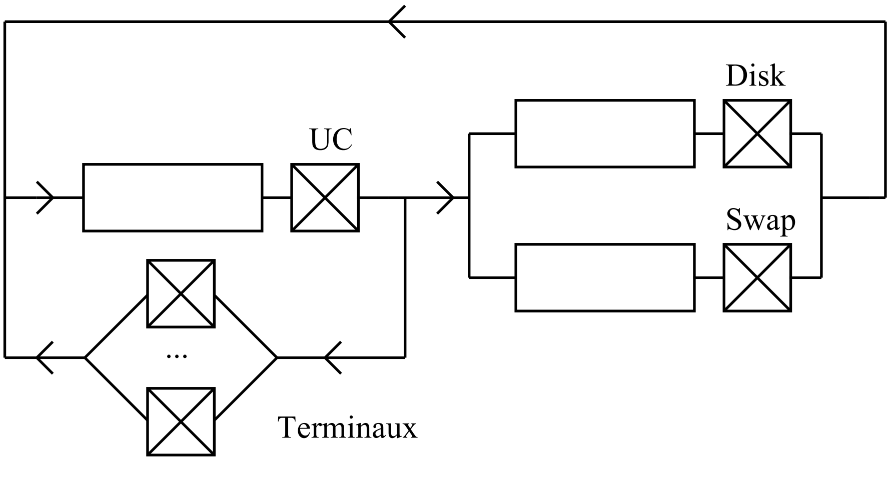
1 Le réseau général de Jackson



Exemple : Réseau ouvert à serveur central ; un central téléphonique



Exemple : Réseau fermé ; un système informatique



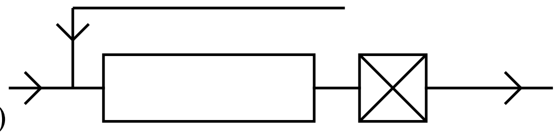
Dans un réseau fermé, le nombre N de clients est fixé.

Hypothèses et notations

Le réseau a n stations.
Les clients sont statistiquement identiques (il y a une seule classe de clients).
La discipline de service est PAPS.
Les clients arrivent de l'extérieur à partir d'une source infinie de débit $\lambda(K)$ où K est le nombre de clients dans le réseau.
La loi d'arrivée est une loi de Poisson.
Un client arrive de l'extérieur dans la station i avec une probabilité qi.
Le temps de service de la station i suit une loi exponentielle de moyenne $1/\mu_i(k_i)$ où k_i est le nombre de clients dans la station i (définition du serveur jacksonnien).
On note p_{ij} la probabilité constante qu'un client, ayant terminé son service dans la station i, aille dans la station j ou sorte du réseau ($j = n+1$).
Les stations sont interconnectées de manière quelconque.
On note $k = (k_1,k_2, \dots ,k_n)$ et $q = (q_1,q_2, \dots ,q_n)$, on a $K = k_1+k_2+ \dots +k_n$.

Remarques :

- Si $\lambda(K) = \lambda$ constant, on a des files M/M/1, M/M/c ou M/M/∞.
- Caractéristiques de la station i :



2 Réseau ouvert

Théorème

Soient $a(k,i)$ [resp. $b(k,i)$] un vecteur identique à k à l'exception de sa composante k_i qui est remplacée par k_i+1 [resp. k_i-1].
Soit $c(k,i,j)$ un vecteur identique à k à l'exception des composantes k_i et k_j qui sont remplacées par k_i+1 et k_j-1.
Les équations du système s'écrivent alors :

$$\frac{d}{dt}p(k,t) = -[\lambda(K)+\sum_{i=1}^{n}\mu_i(k_i)(1-p_{ii})]\,p(k,t)+\sum_{i=1}^{n}\lambda(K-1)q_i\,p(b(k,i),t)+$$
$$\sum_{i=1}^{n}\mu(k_i+1)\,p_{i,n+1}\,p(a(k,i),t)+\sum_{i,j=1|i\neq j}^{n}\mu(k_i+1)\,p_{ij}\,p(c(k,i,j),t)$$

Les termes correspondent aux cas : état inchangé, arrivée, départ, transfert.
Les notations ont été simplifiées en posant

$\mu_i(k_i) = 0$ si $k_i = 0$ et

$p(k,t) = 0$ si le vecteur k contient une composante négative.

Le système obtenu est un système infini d'équations différentielles.
La solution stationnaire est la distribution de probabilité $p(k)$ indépendante des conditions initiales

$$p(k) = \lim_{t\to\infty} p(k,t)$$

Cette limite peut ne pas exister. Si elle existe elle est obtenue en posant

$$\frac{d}{dt}p(k,t) = 0$$

On note

P la matrice des transitions de serveur à serveur $P = (p_{ij})$ $1 \le i \le n$, $1 \le j \le n$

q le vecteur ligne $q = (q_1, \dots ,q_n)$

e le vecteur ligne $e = (e_1, \dots, e_n)$ tel que $e = q + e\,P$

$$v(k) = \prod_{i=1}^{n} \prod_{m=1}^{k_i} \frac{e_i}{\mu_i(m)}$$

$$w(K) = \prod_{m=0}^{K-1} \lambda(m) \quad \text{pour } K \geq 0$$

$$T(K') = \sum_{k \mid K = K'} v(k) \quad \text{pour } K' \geq 0$$

$$C = [\sum_{K'=0}^{K'=\infty} w(K')T(K')]^{-1} \quad C \text{ est nulle si la somme n'existe pas.}$$

Théorème

Si (1) la solution de $e = q + e\,P$ est unique et tous ses éléments sont non négatifs

et si (2) $C > 0$

alors il existe une solution stationnaire unique $\{p(k)\}$ définie par

$p(k) = C.v(k).w(K)$

Remarques :

On a une FORME MULTIPLICATIVE de $p(k)$.

$e = q + e\,P$	nombres moyens de passages par les stations
$\lambda\,e = \lambda\,q + \lambda\,e\,P$	conservation du flux
$\lambda_i = \lambda\,e_i$	taux d'arrivée à la station i

Exercice : Calculer les critères de performance d'un réseau ouvert de stations M/M/1 PAPS, lorsque $\lambda(K) = \lambda$ et $\mu_i(k_i) = \mu_i$.

Montrer d'abord que $C = \prod_{i=1}^{n} (1 - \frac{\lambda\,e_i}{\mu_i})$ si $\rho_i = \lambda\,e_i / \mu_i < 1$.

En déduire que $p(k)$ est un produit de distributions de files M/M/1 et donc que les stations se comportent indépendamment les unes des autres.

On déterminera les temps moyens de séjour et les nombres moyens de clients pour les stations et le système :

$$W = \sum_{i=1}^{n} e_i W_i \quad \text{et} \quad L = \sum_{i=1}^{n} L_i$$

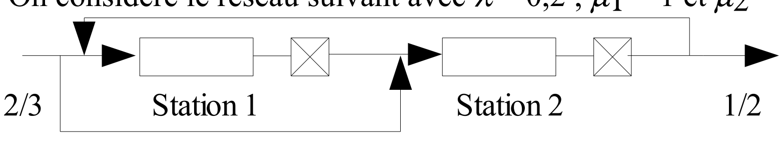
$W_i = L_i / \lambda_i$ temps moyen de séjour lors d'un passage par la station i,

$R_i = e_i W_i = L_i / \lambda$ temps moyen de séjour total dans la station i pour e_i passages.

Quelles sont les conditions d'existence du régime stationnaire (stabilité du réseau) ?

En déduire la valeur maximale de λ.

Application numérique : On considère le réseau suivant avec $\lambda = 0{,}2$; $\mu_1 = 1$ et $\mu_2 = 2$.



Quelles sont les stations prises en charge par le théorème de Jackson ?

Proposez un algorithme de principe d'un solveur de réseaux de file d'attente ouverts.

3 Réseau fermé

Les équations du système s'écrivent alors :

$$\frac{d}{dt}p(k,t)=-\sum_{i=1}^{n}\mu_i(k_i)(1-p_{ii})p(k,t)+\sum_{i,j=1|i\neq j}^{n}\mu(k_i+1)p_{ij}\,p(c(k,i,j),t)$$

Théorème : Il existe une solution stationnaire unique $\{p(k)$, k tel que $\sum_{i=1}^{n}k_i=K$ $\}$ définie par

$$p(k)=\frac{1}{G(K)}\prod_{i=1}^{n}\prod_{m=1}^{k_i}\frac{e_i}{\mu_i(m)}$$

$$G(K)=\sum_{k\,|\,K}\prod_{i=1}^{n}\prod_{m=1}^{k_i}\frac{e_i}{\mu_i(m)}$$

où $e=(e_1,\ldots,e_n)$ est une solution de $e = e\,P$, il faudra fixer une composante ($e_1 = 1$, par exemple).

Remarques : $G(K)$ est une somme finie donc pas de condition particulière d'existence. Elle dépend aussi du nombre n de stations et sera donc notée $G(K,n)$.

Exercices :

1- Quel est le nombre d'états (nombre de vecteurs k) pour un réseau fermé à n stations et K clients. On peut envisager de placer n-1 barrières (bit 1) parmi les K clients (bit 0) mis en rang, par exemple :
1 000 1 00 1 0 ; $n = 4$, $K = 6$; station n°1 vide, 3 clients dans la n°2, 2 dans la n°3 et 1 dans la n°4.
On choisit donc n-1 positions parmi K+n-1.

2- On considère un réseau fermé à n stations, K clients et $\mu_i(k_i) = \mu_i$ pour $1 \le i \le n$. Montrer que

$$p(k)=\frac{1}{G(K,n)}\prod_{i=1}^{n}\left(\frac{e_i}{\mu_i}\right)^{k_i}\quad\text{puis}\quad G(K,n)=\sum_{k|k_1+\ldots+k_n=K}\prod_{i=1}^{n}\rho_i^{k_i}\quad\text{où}\quad\rho_i=\frac{e_i}{\mu_i}\;.$$

On pose $g_i(z)=\sum_{k_i=0}^{\infty}(\rho_i z)_i^k=\frac{1}{1-\rho_i z}$ et $g(z)=\prod_{i=1}^{n}g_i(z)$ $G(K,n)$ est le coefficient de z^K dans $g(z)$.

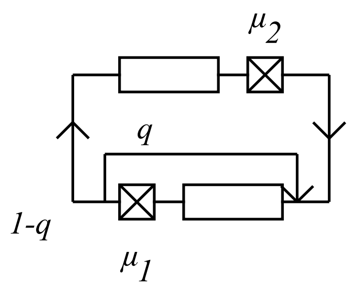
Posons $h_1(z)=g_1(z)$ et $h_i(z)=h_{i-1}(z)g_i(z)$ donc $h_i(z)=h_{i-1}(z)+\rho_i z\,h_i(z)$ et $G(j,i)$ est le coefficient de z^j dans $h_i(z)$. On a la relation de récurrence pour $1 \le j \le K$ et $1 \le i \le n$

$$G(j,i)=G(j,i-1)+\rho_i G(j-1,i)\quad\text{avec}\quad G(0,i)=1\quad\text{et}\quad G(j,1)=\rho_1^j\quad\text{ou}\quad G(j,0)=0\;.$$

Expliciter l'algorithme de BUZEN pour le calcul de $G(K,n)$.

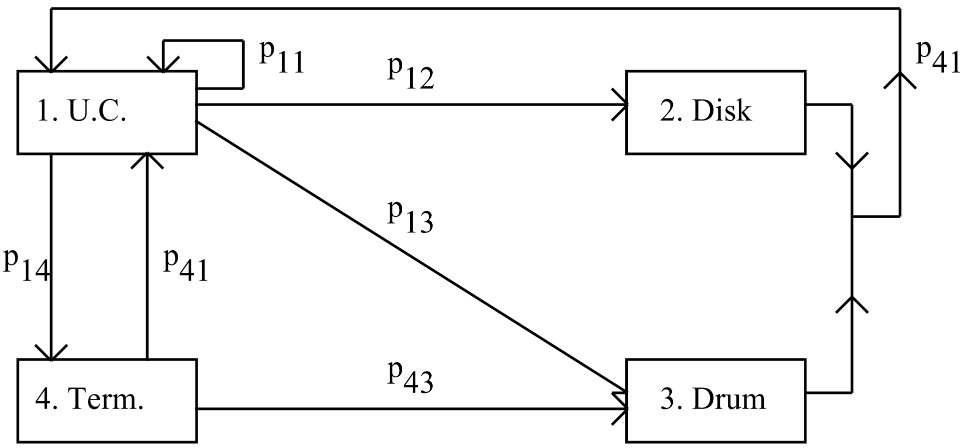
Montrer $U_i=1-p_i(0)=\rho_i G(K-1,n)/G(K,n)$ et $L_i=\rho_i\sum_{j=0}^{K-1}G(j,n)/G(K,n)$.

3- On considère le réseau fermé défini ci-dessous. A l'aide de l'algorithme de BUZEN, calculer $p(k)$ pour 3 clients, des services exponentiels, une discipline FIFO et $0 \le q < 1$.
Quel est le taux d'activité et le nombre moyen de clients de la station i ?



III Application PDP 10, Peter CHEN, 1975

1- Etude simplifiée



$$p_{11}+p_{12}+p_{13}+p_{14}=1$$
$$p_{41}+p_{43}=1$$
$p_{41}=1-p_{43}$ (p_{411} = (1 - p_{43}) / 2 par processeur pour une UC biprocesseur)

2- Etude avec une politique de swapping

Plan de l'article

- Présentation du système et des objectifs
- Le modèle
- La politique de swapping (la matrice des transitions P dépend de l'état du système)
- Proposition d'un algorithme itératif
  - C = { caractéristiques }, $\wp$ = { performances }
  - $\wp_i = \mathrm{f}(C_i)$ puis $C_{i+1}=\mathrm{g}(\wp_i)$ donc $\wp_{i+1}=\mathrm{F}(\wp_i)$ et recherche d'un point fixe.

Ncpu : Nombre moyen de requêtes CPU par interaction.

$$Ncpu=\sum_{k=1}^{\infty} k\, p_{14}(1-p_{14})^{k-1}=1/p_{14}$$

$$Tcpu=\frac{Ncpu}{\mu_1}=\frac{1}{\mu_1 p_{14}}$$

On a aussi $Ncpu=e_1$, première composante de e tel que $e=eP$ et $n_4=1$:

$$e=\left(\frac{1}{p_{14}},\frac{P_{12}}{p_{14}},\frac{P_{13}}{p_{14}}+p_{43},1\right)$$ nombres moyens de passages par CPU, DISK, DRUM et TERM.

Politique de swapping

$A=M/J$	nombre maximum de jobs disposant de mémoire
M	taille mémoire allouée aux utilisateurs
J	taille moyenne d'un job

Les probabilités de transition dépendent de n_4. Les nombres moyens n_i sont évalués à l'aide du théorème de Jackson puis utilisés pour calculer une nouvelle matrice de transition.
Le processus itératif s'arrête lorsque les performances se stabilisent.