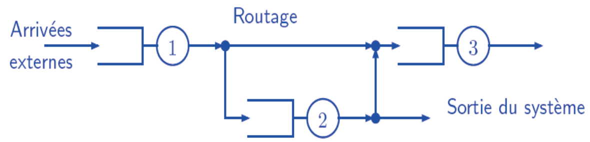
LES RESEAUX DE FILES D'ATTENTES

Réseaux de files d'attente

Un réseau de files d'attente est simplement un système composé d'**une ou plusieurs files** d'attente reliées entre elles.

Les clients (dans les cas les plus simples, tous « identiques »), une fois leur service terminé dans une **station** (file), se déplacent vers une autre station ou quittent le système selon des règles de **routage** (déterministes ou stochastiques).



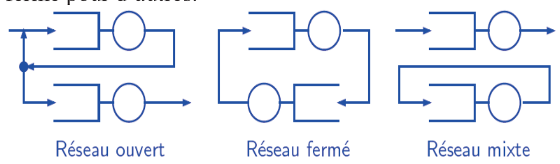
LES RESEAUX DE FILES D'ATTENTES

Réseaux ouverts, fermés et mixtes

Un réseau est **ouvert** si tout client présent ou entrant dans le système peut le quitter.

Un réseau est **fermé** si les clients ne peuvent le quitter. Dans un réseau fermé, le nombre de client est fixe et ces derniers sont présents dans le système dès le début de son évolution.

Finalement, un réseau est **mixte** s'il est ouvert pour certains clients et fermé pour d'autres.



LES RESEAUX DE FILES D'ATTENTES

Réseaux à formes produits

La définition la plus simple de l'état d'un réseau consiste à définir l'état *x(t)* du système au temps t comme le vecteur ($x_1(t)$, …, $x_n(t)$) où $x_j(t)$ est le nombre de clients présents à l'instant t dans la file j.

Sous certaines conditions, un réseau stable possède une distribution stationnaire π^* de la forme :

$$\pi^*(x) = \pi^*(x_1 \ldots x_n) = \prod_{j=1}^{n} \pi_j^*(x_j)$$

Un tel réseau est dit **à forme produit** et se comporte comme autant *n* files **indépendantes** de distributions stationnaires respectives π^*_j, j = 1, …, n.

LES RESEAUX DE FILES D'ATTENTES

Les réseaux de Jackson

Un réseau de Jackson est composé de **n** files **exponentielles** c'est-à-dire de files d'attente

- ➢ comportant chacun un ou plusieurs serveurs identiques (m_j pour la file j),
- ➢ fournissant des services de durée exponentielle (le taux de service de la file j est noté μ_j),
- ➢ de capacité infinie,
- ➢ utilisant une discipline de service FIFO.

Les clients (appartenant tous à la même classe) arrivent dans le système selon des processus de Poisson indépendants, le **taux d'arrivée externe** dans la file j étant égal à γ_j .

LES RESEAUX DE FILES D'ATTENTES

Les réseaux de Jackson : règles de routage

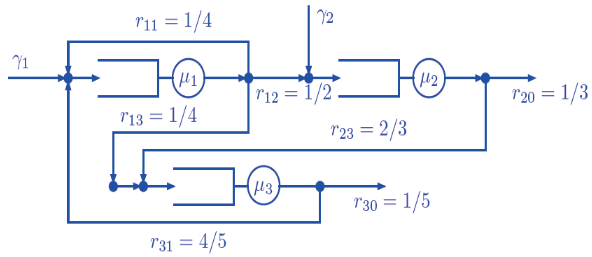
Après avoir terminé son service à la station *j*, un client se déplace à la station *k* avec probabilité r_{jk} et quitte le système avec une probabilité r_{j0} où

$$r_{j0} = 1 - \sum_{k=1}^{n} r_{jk}$$

De telles règles définissent un **routage markovien**.

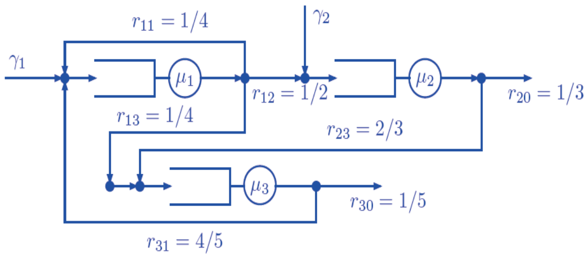
LES RESEAUX DE FILES D'ATTENTES

Exemple : chaque file n'a qu'un seul serveur



LES RESEAUX DE FILES D'ATTENTES

Exemple : chaque file n'a qu'un seul serveur



$$R = \begin{pmatrix} 1/4 & 1/2 & 1/4 \\ 0 & 0 & 2/3 \\ 4/5 & 0 & 0 \end{pmatrix} \qquad r_{.0} = \begin{pmatrix} 0 \\ 1/3 \\ 1/5 \end{pmatrix}$$

LES RESEAUX DE FILES D'ATTENTES

Conditions de stabilité

Sans surprise :

Un réseau est stable <=> chacune de ses files l'est.

Notant λ_j le **taux effectif** (ou taux moyen) d'arrivée dans la file j, on peut écrire les **équations de conservation :**

$$\lambda_j = \gamma_j + \sum_{i=1}^{n} \lambda_i r_{ij}$$

Sous forme matricielle, ces équations deviennent

$$\lambda = \gamma + \lambda R \iff \lambda(I - R) = \gamma$$

LES RESEAUX DE FILES D'ATTENTES

Conditions de stabilité

Pour un réseau **ouvert**, on a $\lim_{n\to\infty} R^n = 0$ et l'unique solution du système précédent est

$$\lambda = \gamma(I - R)^{-1}$$

Connaissant les taux effectifs d'arrivée λ, la file j est stable si et seulement si

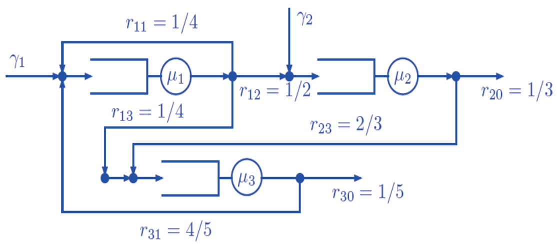
$$\rho_j = \frac{\lambda_j}{m_j \mu_j} < 1$$

et le réseau est stable si et seulement si chacune des files le composant l'est.

REMARQUE. Un réseau fermé est toujours stable !

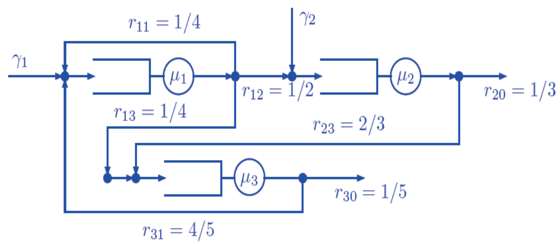
LES RESEAUX DE FILES D'ATTENTES

Exemple



LES RESEAUX DE FILES D'ATTENTES

Exemple



Les équations de conservation du flot dans le réseau sont

$$\begin{cases} \lambda_1 = \gamma_1 + \lambda_1 r_{11} + \lambda_3 r_{31} \\ \lambda_2 = \gamma_2 + \lambda_1 r_{12} \\ \lambda_3 = \lambda_1 r_{13} + \lambda_2 r_{23} \end{cases}$$

LES RESEAUX DE FILES D'ATTENTES

Exemple

$$\iff \begin{cases} \lambda_1 = \gamma_1 + \lambda_1/4 + 4\lambda_3/5 \\ \lambda_2 = \gamma_2 + \lambda_1/2 \\ \lambda_3 = \lambda_1/4 + 2\lambda_2/3 \end{cases}$$

$$\iff \begin{cases} \lambda_1 = \frac{60}{17}\gamma_1 + \frac{32}{17}\gamma_2 \\ \lambda_2 = \frac{30}{17}\gamma_1 + \frac{33}{17}\gamma_2 \\ \lambda_3 = \frac{35}{17}\gamma_1 + \frac{30}{17}\gamma_2 \end{cases}$$

Le réseau est donc stable si γ_1, γ_2, μ_1, μ_2 et μ_3 sont tels que :

$$\rho_1 = \frac{\lambda_1}{\mu_1} < 1 \qquad \rho_2 = \frac{\lambda_2}{\mu_2} < 1 \qquad \rho_3 = \frac{\lambda_3}{\mu_3} < 1$$

LES RESEAUX DE FILES D'ATTENTES

Distribution stationnaire

Les réseaux de Jackson sont des réseaux à forme produit.

La distribution stationnaire d'un réseau **ouvert et stable** est

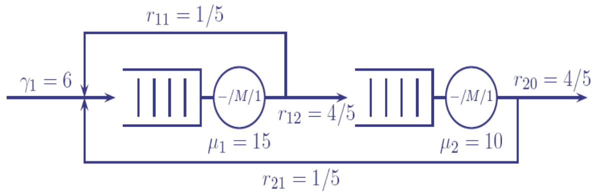
$$\pi^*(x) = \pi^*(x_1 \dots x_n) = \prod_{j=1}^{n} \pi_j^*(x_j).$$

où $\pi_j^*(x_j)$ est égal à la probabilité stationnaire d'observer x_j clients dans une file $M/M/m_j$.

Un tel système se comporte donc comme n files $M/M/m_j$ indépendantes d'intensité respective $\rho_1 = \frac{\lambda_1}{m_1\mu_1}, \dots \rho_n = \frac{\lambda_n}{m_n\mu_n}$

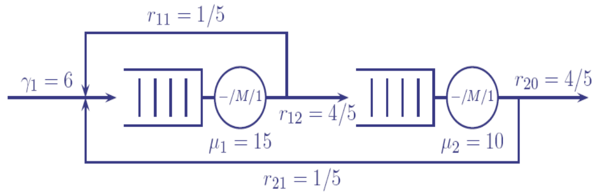
LES RESEAUX DE FILES D'ATTENTES

Réseau de Jackson ouvert



LES RESEAUX DE FILES D'ATTENTES

Réseau de Jackson ouvert



Équations de conservation du flot :

$$\begin{cases} \lambda_1 = \gamma_1 + \lambda_1 r_{11} + \lambda_2 r_{21} \\ \lambda_2 = \gamma_2 + \lambda_1 r_{12} + \lambda_2 r_{22} \end{cases} \Leftrightarrow \begin{cases} \lambda_1 = 6_1 + 1/5\,\lambda_1 + 1/5\,\lambda_2 \\ \lambda_2 = 4/5\,\lambda_1 \end{cases}$$

Solution unique : $\lambda_1 = \frac{75}{8}, \lambda_2 = \frac{15}{2}.$

LES RESEAUX DE FILES D'ATTENTES

Stabilité du réseau

- Première file stable $\Leftrightarrow$
$$\rho_1 = \frac{\lambda_1}{m_1\mu_1} < 1 \Leftrightarrow \rho_1 = \frac{75/8}{1.15} = \frac{75}{120} = \frac{5}{8} < 1 \qquad \checkmark$$
- Deuxième file stable $\Leftrightarrow$
$$\rho_2 = \frac{\lambda_2}{m_2\mu_2} < 1 \Leftrightarrow \rho_2 = \frac{15/2}{1.10} = \frac{15}{20} = \frac{3}{4} < 1 \qquad \checkmark$$

$\Rightarrow$le réseau est stable et possède une distribution stationnaire unique et à forme produit :

$$\pi^*(x_1,x_2) = (1-\rho_1)\rho_1^{x_1}(1-\rho_2)\rho_2^{x_2} = \frac{3}{8}\left(\frac{5}{8}\right)^{x_1}\frac{1}{4}\left(\frac{3}{4}\right)^{x_2} = \frac{3}{32}\left(\frac{5}{8}\right)^{x_1}\left(\frac{3}{4}\right)^{x_2}$$

où $x_1 \geq 0$ et $x_2 \geq 0$ représentent, respectivement, le nombre de clients dans la première et la deuxième file.

LES RESEAUX DE FILES D'ATTENTES

Mesure de performance

Le nombre moyen de clients présents dans la première file est :

$$\overline{N}_1 = \sum_{x_1=0}^{\infty}\sum_{x_2=0}^{\infty} x_1\pi^*(x_1,x_2) = \sum_{x_1=0}^{\infty}\sum_{x_2=0}^{\infty} x_1(1-\rho_1)\rho_1^{x_1}(1-\rho_2)\rho_2^{x_2}$$

$$= \left(\sum_{x_1=0}^{\infty} x_1(1-\rho_1)\rho_1^{x_1}\right)\underbrace{\left(\sum_{x_2=0}^{\infty}(1-\rho_2)\rho_2^{x_2}\right)}_{1}$$

$$= \frac{\rho_1}{(1-\rho_1)} = \frac{5/8}{1-5/8} = \frac{5}{3}.$$

De même, le nombre moyen de clients dans la deuxième file est

$$\overline{N}_2 = \frac{\rho_2}{(1-\rho_2)} = \frac{3/4}{1-3/4} = 3.$$

LES RESEAUX DE FILES D'ATTENTES

Mesure de performance

Le temps moyen de séjour d'un client dans le système est :

$$\overline{T} = \frac{\overline{N}_{sys}}{\lambda_{sys}} = \frac{\overline{N}_1 + \overline{N}_2}{\gamma_1} = \frac{5/3+3}{6} = \frac{7}{9}[h]$$

où λ_{sys} est le taux global d'arrivée dans le système (ici, simplement γ_1).

Remarque. Une erreur à ne pas commettre est de calculer $\mathcal{T}$ à l'aide de

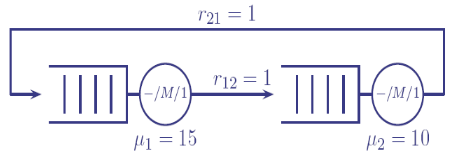
$$\overline{T} = \overline{T}_1 + \overline{T}_2$$

avec

$$\overline{T}_1 = \frac{\overline{N}_1}{\lambda_1} = \frac{5/3}{75/8} = \frac{8}{45} \quad \text{et} \quad \overline{T}_2 = \frac{\overline{N}_2}{\lambda_2} = \frac{3}{15/2} = \frac{2}{5}$$

LES RESEAUX DE FILES D'ATTENTES

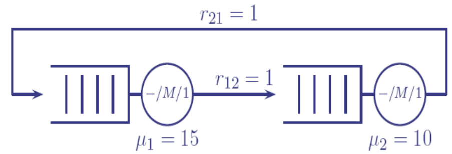
Réseau de Jackson fermé : exemple



Le système contient à tout instant K = 10 clients.

LES RESEAUX DE FILES D'ATTENTES

Réseau de Jackson fermé : exemple



Le système contient à tout instant K = 10 clients.

Équations de conservation du flot : $\begin{cases} \lambda_1 = \lambda_2 \\ \lambda_2 = \lambda_1 \end{cases}$

$\Rightarrow$ le système possède une infinité de solutions !

Le système est toujours stable car le nombre de clients est fixe !

LES RESEAUX DE FILES D'ATTENTES

Distribution stationnaire

Les probabilités stationnaires $\pi(x_1, x_2)$ d'avoir x_1 clients dans la première file et x_2 clients dans la seconde ne sont définies que pour les valeurs entières et non négatives de x_1 et x_2 vérifiant :

$$x_1 + x_2 = 10.$$

Il n'y a donc que 11 couples différents pour lesquelles elles sont définies.

La distribution stationnaire du système est, ici aussi, à forme produit et est donnée par :

$$\pi^*(x) = \pi^*(x_1, x_2) = \frac{1}{G(10)}^{a} F_1(x_1)F_2(x_2) = \frac{1}{G(10)}\left(\frac{\lambda_1}{\mu_1}\right)^{x_1}\left(\frac{\lambda_2}{\mu_2}\right)^{x_2}$$

[a] G(10) est une constante de normalisation

LES RESEAUX DE FILES D'ATTENTES

Distribution stationnaire

Dans l'expression :

$$\pi^*(x_1, x_2) = \frac{1}{G(10)}\left(\frac{\lambda_1}{\mu_1}\right)^{x_1}\left(\frac{\lambda_2}{\mu_2}\right)^{x_2}$$

qui n'est applicable que pour

$$(x_1, x_2) \in S = \{(x1, x2) \in IN^2 \mid x_1 + x_2 = 10\}$$

les « taux » λ_1 et λ_2 correspondent à une solution quelconque des équations de conservation du flot. Fixant, arbitrairement, $\lambda_1 = 1$, on obtient dans notre cas $\lambda_2 = \lambda_1 = 1$ et

$$\pi^*(x_1, x_2) = \frac{1}{G(10)}\left(\frac{1}{15}\right)^{x_1}\left(\frac{1}{10}\right)^{x_2} \qquad (x1, x2) \in S$$

LES RESEAUX DE FILES D'ATTENTES

Constante de normalisation

La constante *G(10)* est une **constante de normalisation** assurant que la somme des probabilités stationnaires est égale à 1 :

$$G(10) = \sum_{x_1=0}^{10}\left(\frac{1}{15}\right)^{x_1}\left(\frac{1}{10}\right)^{10-x_1} = \left(\frac{1}{10}\right)^{10}\sum_{x_1=0}^{10}\left(\frac{10}{15}\right)^{x_1}$$

$$= \left(\frac{1}{10}\right)^{10}\sum_{x_1=0}^{10}\left(\frac{2}{3}\right)^{x_1} = \left(\frac{1}{10}\right)^{10}\left(\frac{1-\left(\frac{2}{3}\right)^{11}}{1-\frac{2}{3}}\right).$$

LES RESEAUX DE FILES D'ATTENTES

Constante de normalisation : cas général

Pour un réseau de Jackson fermé comptant n files exponentielles et K clients, la distribution stationnaire n'est définie que pour les états $x = (x_1, \ldots, x_n)$ vérifiant :

$$x = (x_1, \ldots, x_n) \in S_K = \left\{(x_1, \ldots, x_n) \in IN^n \middle| \sum_{j=1}^{n} x_j = K\right\}.$$

Même pour des valeurs modérées de n et K, l'ensemble S_K est énorme :

$$|S_K| = \binom{n+K-1}{K} = \frac{(n+K-1)!}{K!(n-1)!}$$

$\Rightarrow$ **le calcul de la constante *G(K)* constitue l'une des principales difficultés lors de l'étude d'un réseau de Jackson fermé.**

LES RESEAUX DE FILES D'ATTENTES

Réseaux BCMP

Les réseaux de Jackson ne sont pas les seuls réseaux possédant une distribution stationnaire à forme produit. Une classe importante de tels réseaux, englobant ceux de Jackson, sont les réseaux BCMP dont le nom fait référence à leurs "inventeurs " : *Baskett, Chandy, Muntz et Palacios*.

Ces réseaux généralisent ceux de Jackson et sont formés

a) de files exponentielles *-/M/m - FIFO*
b) de centre de délai *-/M/∞* ou *-/G/∞*
c) de files *-/G/1 - LIFO - RR*
d) de files *-/G=1 - PS*

La forme produit est cependant loin d'être la règle !

LES RESEAUX DE FILES D'ATTENTES

Réseaux BCMP

➢ Round Robin (RR) : on définit un "quantum de temps de service" : le serveur traite à tour de rôle chaque client pour une portion du temps égale au quantum ; il accomplit ainsi une sorte de cycle sur les clients en attente. Dès qu'un client a reçu, en plusieurs fois, le service qu'il réclame, il quitte le serveur.

➢ Processor Sharing (PS): version asymptotique du "round robin": le quantum tend vers 0 ; si n clients sont présents entre t et $t+\Delta t$, chacun d'eux reçoit un service élémentaire égal à $\Delta t/n$ pendant ce temps (avec $\Delta t \to 0$).

LES RESEAUX DE FILES D'ATTENTES

Exercice

Des pièces arrivent selon un processus de Poisson de paramètre $\gamma = 1$ dans un atelier composé de quatre machines $M_1, \ldots, M_4$ dont les temps de traitement obéissent a des lois exponentielles indépendantes de paramètres respectifs $\mu_1 = 5/2$, $\mu_2 = 2$, $\mu_3 = 3/2$, $\mu_4 = 5/2$. Sachant que :

_ les pièces pénètrent dans l'atelier par M_1 et le quittent par M_4 ;
_ la probabilité qu'une pièce soit traitée incorrectement par une machine et doive donc être retraitée (par cette machine) est de 1/5 pour M_1 et M_2, de 1/10 pour M_3, et de 0 pour M_4 ;
_ les pièces correctement traitées par M_1 sont acheminées avec probabilités égales vers M_2 ou M_3, alors que celles traitées par M_2 ou M_3 sont directement envoyées vers M_4.

LES RESEAUX DE FILES D'ATTENTES

Exercice

a) représenter schématiquement le système ;

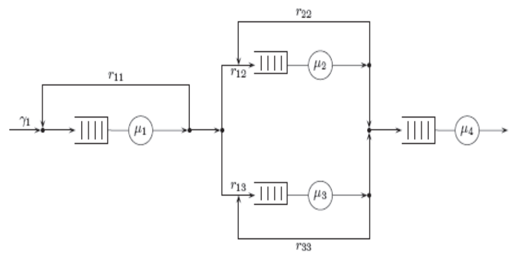
Puis déterminer :

b) si le système est stable ;
c) la loi de probabilités conjointes du nombre de requêtes dans le système ;
d) quelle est la machine critique de l'atelier ;
e) le temps moyen de réponse du système pour une requête quelconque.

LES RESEAUX DE FILES D'ATTENTES

Exercice / corrigé

a) L'atelier peut être assimile a un réseau de files d'attente schématisé par la figure suivante :



LES RESEAUX DE FILES D'ATTENTES

Exercice / corrigé

b) La matrice de routage R du réseau de Jackson s'écrit :

$$R = \begin{pmatrix} 1/5 & 2/5 & 2/5 & 0 \\ 0 & 1/5 & 0 & 4/5 \\ 0 & 0 & 1/10 & 9/10 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

et le vecteur λ des taux effectifs d'arrivée dans les files est donné par l'unique solution du système $\lambda = \gamma + \lambda R$ où $\gamma = (1\ 0\ 0\ 0)$ est le vecteur des taux d'arrivée externes.

Après calcul, on trouve :

$$\lambda = (5/4\ 5/8\ 5/9\ 1).$$

Exercice / corrigé

Le système est stable si et seulement si $\rho_i = \lambda_i/\mu_i < 1 \ \forall \ i = 1, \ldots, 4$, ce qui est le cas comme le montre le calcul suivant :

$$\rho_1 = \frac{5}{4}.\frac{2}{5} = \frac{1}{2} < 1$$

$$\rho_2 = \frac{5}{8}.\frac{1}{2} = \frac{5}{16} < 1$$

$$\rho_3 = \frac{5}{9}.\frac{2}{3} = \frac{10}{27} < 1$$

$$\rho_4 = 1.\frac{2}{5} = \frac{2}{5} < 1$$

Exercice / corrigé

c) Le système se comporte comme 4 files *M/M/1* indépendantes d'intensités ρ_1, ρ_2, ρ_3, ρ_4 et la distribution stationnaire du réseau vaut donc :

$$\pi^*(x_1, x_2, x_3, x_4) = \prod_{j=1}^{4}(1-\rho_j)\rho_j^{x_j}$$

où x_j dénote le nombre de pièces dans la machine j.

d) La machine critique est celle dont l'intensité du trafic ρ est maximale, c'est à dire la machine M_1.

Exercice / corrigé

e) Le nombre moyen de pièces $\mathcal{N}$ est la somme des $\mathcal{N}_i$, où $\mathcal{N}_i = \lambda_i/(\mu_i-\lambda_i)$ est le nombre moyen de pièces pour la machine *i*. Or après calcul, on trouve :

$$\overline{N}_1 = 1$$

$$\overline{N}_2 = \frac{5}{11}$$

$$\overline{N}_3 = \frac{10}{17}$$

$$\overline{N}_4 = \frac{2}{3}$$

Exercice / corrigé

En conséquence : $\overline{N} = \sum_{i=1}^{4} \overline{N}_i$

Pour calculer le temps moyen nécessaire au traitement d'une pièce, on utilise la formule de Little :

$$\overline{T} = \overline{N}/\gamma_1 = \frac{1520}{561}\,\mathrm{s}$$