Les files d'attente
(dans les équipements de réseau)

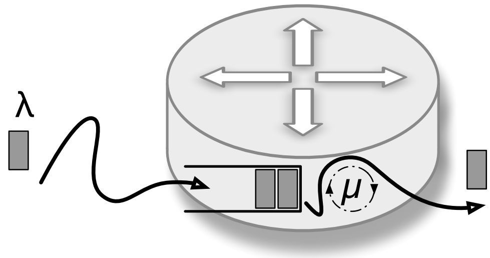

Figure 1 – Le taux moyen d'arrivée des paquets est λ, le taux moyen des départs est μ.

On considère un routeur (fig. 1) (ou commutateur) auquel, en moyenne, λ paquets par seconde se présentent avec toujours la même interface de sortie. On suppose que les instants d'arrivée n'ont aucun lien entre eux — leur processus de génération est sans mémoire ; de plus, ils ne sont jamais superposés. C'est un processus de Poisson dont vous ré-entendrez parler dans d'autres cours. Sur cette interface, on suppose que le routeur transmet un paquet toutes les s secondes en moyenne, et donc se débarrasse de $\mu = \frac{1}{s}$ paquets par seconde (typiquement, s est lié à la taille des paquets — elle peut être aléatoire ou fixe). λ et μ sont les taux (ou « *intensités* ») d'arrivée et de départ des paquets.

Comme les arrivées et les départs sont à des instants aléatoires, il faut un tampon pour absorber les irrégularités. On suppose que ce tampon est de taille infinie et on reviendra sur ce point plus bas. Si le tampon est vide quand un paquet arrive, alors il sera immédiatement transmis ; sinon, il devra attendre que les j paquets présents dans le tampon à son arrivée aient été mis en ligne, ce qui prendra en moyenne $js = \frac{j}{\mu}$ secondes.

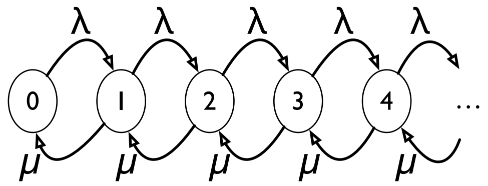

Figure 2 – Chaque état i représente i paquets dans la file d'attente. On change d'état quand un paquet arrive ou part.

$\boxed{\text{Notre problème est de déterminer la durée d'attente moyenne en fonction de } \lambda \text{ et } \mu.}$

Le sort d'un paquet est déterminé par le remplissage (l'état) de la file d'attente à son arrivée. On notera N le nombre de paquets déjà présents à l'arrivée d'un nouveau paquet et $P_i = P[N = i]$, la probabilité qu'il trouve i paquets dans la file d'attente. Au cours du temps, i augmente de 1 à chaque fois qu'un paquet arrive et diminue de 1 (si non nul) quand un paquet est transmis, comme en figure 2.

Q. 1 – Si on observe la file d'attente en continu, on est témoin —dans la représentation de la figure 2— de λ transitions vers la droite par seconde, et μ vers la gauche (à moins que la file d'attente ne soit pas vide, bien sûr). Si on se restreint aux transitions à partir de l'état i et vers la droite (la transition de i à $i+1$), combien en observe-t-on par seconde ? (Ça n'est pas compliqué)

Q. 2 – Comment décrire ce qu'il se passerait si pour un état $i > 0$ on avait $\lambda P_{i-1} - \mu P_i + \mu P_{i+1} - \lambda P_i \neq 0$? Dans la suite, on considère que les probabilités P_i sont stabilisées, c'est à dire qu'elles ne varient pas au cours du temps.

Q. 3 – Il n'y a aucune transition à gauche de l'état 0, donc quelle expression lie P_i et P_{i+1} (par récurrence) ? De même, quelle est l'expression de P_i en fonction de P_0 ?

On pose $\rho = \frac{\lambda}{\mu}$ pour simplifier. Comme, par définition,

$$\sum_{i=0}^{\infty} P_i = 1$$

Q. 4 – en déduire P_i en fonction de ρ ?
Préciser aussi :
- Pourquoi cette somme vaut 1 ?
- À quoi correspond ρ ?

Bien sûr, il est préférable d'avoir $\rho < 1$. Sinon le tampon croît indéfiniment…

Q. 5 – Quelle est la valeur moyenne de N, qu'on pourra noter $\bar{N}$?

Q. 6 – Tracez $\bar{N}$ en fonction de ρ.

Q. 7 – Pour $\rho = 90\%$, quelle fraction du temps le routeur passe-t-il à ne rien faire ? Quel est le remplissage moyen de la file d'attente ? Quelle est l'expression du temps d'attente moyen $\bar{W}$?

Q. 8 – Quel est l'état q en deçà duquel on a une probabilité Q de se trouver ?

Si le tampon est de taille finie (en nombre de paquets, comme dans tous les vrais routeurs d'ailleurs), la probabilité de dépasser cette taille est une bonne indication du taux de perte (même si les choses changent quand des paquets sont rejetés avant d'être mis en file d'attente)(Et ça n'est pas tellement plus compliqué à appréhender d'ailleurs).

Q. 9 – Quelle est la valeur q_{99} qu'on ne dépasse que 1% du temps pour $\rho = 0.9$?

Q. 10 – On compare le cas d'une paire de routeurs opérant en parallèle pour lesquels le trafic et la capacité de traitement sont les mêmes (λ, μ), à celui d'un routeur plus gros, pour lequel $\lambda' = 2\lambda$ et $\mu' = 2\mu$. Quelle est l'attente dans ce 2e cas ?

Q. 11 – En régime à l'équilibre (avec $\lambda < \mu$), quelle est l'intensité de sortie des paquets du routeur ?