



Note sur la file d'attente avec bouclage

Philippe Robert

► **To cite this version:**

| Philippe Robert. Note sur la file d'attente avec bouclage. RR-0324, INRIA. 1984. <inria-00076233>

HAL Id: inria-00076233

<https://hal.inria.fr/inria-00076233>

Submitted on 24 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

IRIA

CENTRE DE ROCQUENCOURT

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105

78153 Le Chesnay Cedex
France

Tél. (3) 954 90 20

Rapports de Recherche

N° 324

NOTE SUR LA FILE D'ATTENTE AVEC BOUCLAGE

Philippe ROBERT

Août 1984

NOTE SUR LA FILE D'ATTENTE AVEC BOUCLAGE

Philippe ROBERT
INRIA
Domaine de Voluceau
Rocquencourt
BP 105
78153 Le Chesnay Cedex
France

ABSTRACT

We study the stability of a single server queue with a loop. Relations between the waiting times are established and in the case of a Poisson arrival, their means are computed.

RESUME

Nous étudions l'état stationnaire d'une file d'attente avec bouclage. Des relations entre les temps d'attente sont établies et dans le cas d'une arrivée Poissonnienne nous calculons leur moyenne.



La file d'attente avec bouclage

Nous étudions dans cette note un modèle de file d'attente s'apparentant au modèle "round robin" : Les requêtes arrivent au processeur avec deux opérations à effectuer, la discipline de la file d'attente est premier arrivé premier servi (P.A.P.S), la première opération terminée la requête retourne immédiatement en fin de file pour effectuer la deuxième.

Nous donnons les conditions de stabilité de cette file d'attente en utilisant le cadre des processus ponctuels. Cela nous permet d'écrire simplement les variables d'attente relatives à ce système et d'en calculer les moyennes dans le cas d'une arrivée poissonnienne de requêtes. Nous terminons en comparant cette file d'attente avec une file d'attente P.A.P.S (celle où les deux opérations sont menées consécutivement).

Ce type de file d'attente a jusqu'à présent été peu étudié. Les modèles étudiés (cf. Disney R.L. pour un état de l'art) supposent habituellement une homogénéité des opérations à un même processeur (i.e. les opérations effectuées par celui-ci ont même loi), ce qui n'est pas le cas ici, les lois des deux opérations pouvant être différentes.

I Notations et Rappels sur la file d'attente P.A.P.S

Les variables utilisées seront définies sur l'espace de probabilité $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$, on notera $(\theta_t)_{t \in \mathbb{R}}$ le flot des translations dans le temps, nous supposons que ce flot préserve la probabilité $\bar{\mathbb{P}}$ ($\int f(\theta_t(\omega)) d\bar{\mathbb{P}}(\omega) = \int f(\omega) d\bar{\mathbb{P}}(\omega)$ pour toute fonction f $\bar{\mathcal{A}}$ -mesurable positive et tout $t \in \mathbb{R}$).

Les arrivées des requêtes au processeur seront représentées par une distribution aléatoire de points sur la droite réelle $N(\omega, ds) = \sum_{n \in \mathbb{Z}} \varepsilon_{t_n}(\omega)$ où ε_a est la masse de Dirac en a et les $(t_n(\omega))_n$ sont les instants d'arrivée des requêtes numérotés par ordre croissant :

$$t_n(\omega) < t_{n+1}(\omega) \text{ et } t_0(\omega) \leq 0 < t_1(\omega).$$

Nous supposons le processus ponctuel N stationnaire :

$$N(\theta_t(\omega), ds) = \sum_{n \in \mathbb{Z}} \varepsilon_{t_n(\omega) - t}, \text{ d'intensité } 1 : \bar{E}(N[0, 1[) = \int \int_0^1 N(\omega, ds) d\bar{P}(\omega) = 1,$$

ce qui à une constante multiplicative près ne restreint pas la généralité de cette note.

À N on associe le sous ensemble de $\bar{\Omega}, \Omega = \{t_0=0\} = \{\text{une requête arrive à l'instant } t=0\}$ et la mesure de probabilité P sur Ω telle que $\int f(\theta_{t_1}(\omega)) dP(\omega) = \int f(\omega) dP(\omega)$, l'opérateur de $\Omega \longrightarrow \Omega \omega \longrightarrow \theta_{t_1}(\omega)$ sera noté θ , ainsi θ préserve P sur Ω . (Ω, P, θ) est l'espace de Palm associé à N.

La requête arrivée à l'instant $t = t_n(\omega)$ requiert le service de durée $\sigma(\theta_{t_n}(\omega)) = \sigma \circ \theta^n(\omega)$ où $\sigma : \Omega \longrightarrow \mathbb{R}^+$ est une variable aléatoire : le service du client arrivé à l'instant $t = 0$. Le processus d'arrivée de cette file d'attente sera noté $\sum_{n \in \mathbb{Z}} \varepsilon_{(t_n, \sigma \circ \theta^n)}(\omega)$ Rappelons pour terminer le résultat classique sur la file d'attente P.A.P.S de processus d'arrivée $\sum_{n \in \mathbb{Z}} \varepsilon_{(t_n, \sigma \circ \theta^n)}(\omega)$:

Proposition 1

Sous la condition $E(\sigma) = \int \sigma(\omega) dP(\omega) < 1$, il existe une unique variable temps d'attente de la file P.A.P.S., i.e. une unique variable aléatoire $W : \Omega \longrightarrow \mathbb{R}^+$ vérifiant

$$W \circ \theta = (W + \sigma - t_1)^+ \text{ (avec } a^+ = \text{Max}\{a, 0\})$$

De plus $P(W=0) > 0$ et si $v = \sup \{k \leq 0 / W \circ \theta^k = 0\}$

$$\text{alors } W = \sum_{k=v}^{-1} (\sigma - t_1) \circ \theta^k(\omega).$$

II La file d'attente avec bouclage : construction d'un modèle stationnaire

Les requêtes arrivent au processeur suivant le processus ponctuel stationnaire $N(\omega, ds) = \sum_{n \in \mathbb{Z}} \varepsilon_{t_n}(\omega)$. Le serveur sert dans l'ordre des arrivées, à leur entrée dans la file d'attente les clients demandent $\sigma_1 : \Omega \longrightarrow \mathbb{R}^+$ mesurable, celui ci effectué ils se placent en fin de file pour requérir $\sigma_2 : \Omega \longrightarrow \mathbb{R}^+$ ensuite ils quittent la file d'attente. Si lors du bouclage un client revient en fin de file en même temps qu'un client qui arrive pour la première fois dans la file, on conviendra de placer le nouvel arrivant après celui qui vient de boucler.

Nous noterons (F_1) ce modèle de file d'attente, définissons ce que l'on entend par modèle stationnaire de (F_1) :

Stationnarité de la file d'attente avec bouclage

Un état stationnaire de la file d'attente (F_1) sera la donnée de deux variables $W_1, W_2 : \Omega \longrightarrow \mathbb{R}$ mesurables représentant respectivement le temps d'attente pour effectuer σ_1 (resp σ_2).

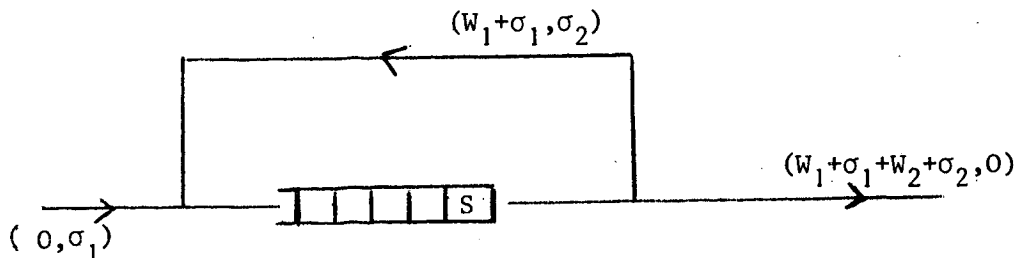
Plus précisément si

$$N_0 = \sum_{n \in \mathbb{Z}} \varepsilon_{(t_n, \sigma_1 \circ \theta^n)} \text{ et } N_1 = \sum_{n \in \mathbb{Z}} \varepsilon_{(t_n + (W_1 + \sigma_1) \circ \theta^n, \sigma_2 \circ \theta^n)}$$

alors $N = N_0 + N_1$ est le processus d'arrivée d'une file d'attente P.A.P.S stable dont la charge $C : \bar{\Omega} \longrightarrow \mathbb{R}^+$ vérifie

$$C \circ \theta_{t_0} = W_1 \circ \theta_{t_0} \text{ et } C \circ \theta_{t_0 + (W_1 + \sigma_1) \circ \theta_{t_0}} = W_2 \circ \theta_{t_0}$$

C est le temps d'un client qui arrive à l'instant $t=0$ (comme convenu, si un client boucle et revient en fin de file à $t=0$, celui ci est prioritaire sur le nouvel arrivant).



Autrement dit, W_1 vérifie une équation de point fixe : en superposant les clients $(t_n, \sigma_1 o\theta^n)_n$ et les clients décalés de $W_1 + \sigma_1$

$(t_n + (W_1 + \sigma_1) o\theta^n, \sigma_2 o\theta^n)$ alors les premiers attendent W_1 dans la file.

Proposition 2

Si $E(\sigma_1 + \sigma_2) < 1$ alors il existe un état stationnaire unique de la file d'attente avec bouclage.

Construction d'un modèle stationnaire

Si $E[\sigma_1 + \sigma_2] < 1$, alors d'après §I, la file d'attente P.A.P.S. avec le processus d'arrivée $\sum_{n \in \mathbb{Z}} \varepsilon_{t_n}(\omega)$ et le service $\sigma_1 + \sigma_2$ est stable : il existe une unique variable aléatoire $V : \Omega \longrightarrow \mathbb{R}^+$ telle que $V o\theta = (V + \sigma_1 + \sigma_2 - t_1)^+$ et si $v_0 = \sup \{k \leq 0 / V o\theta^k = 0\}$ alors $V = \sum_{k=v_0}^{-1} (\sigma_1 + \sigma_2 - t_1) o\theta^k = \sum_{k=v_0}^{-1} (\sigma_1 + \sigma_2) o\theta^k + t_{v_0}$, la file d'attente est constamment chargée entre t_{v_0} et 0 et le client arrivé à $t = t_{v_0}$ n'attend pas. Nous noterons (F_0) ce modèle de file d'attente de $C_0 : \Omega \longrightarrow \mathbb{R}^+$ sa charge.

Construisons $(W_1^k)_{v_0 \leq k \leq 0}$ sur Ω par récurrence en posant $W_1^{v_0} = 0$ et si $v_0 \leq i \leq i+1 \leq 0$

$$W_1^{i+1} = \sum_{k=v_0}^i \sigma_1 o\theta^k + \sigma_2 o\theta^k \cdot \mathbf{1}_{\{t_k + W_1^k + \sigma_1 o\theta^k \leq t_{i+1}\}} - (t_{i+1} - t_{v_0})$$

D'après la définition de v_0 , pour $v_0 \leq k \leq 0$, $v_0 o\theta^k = v_0 - k$ on en déduit $W_1^k = W_1^0 o\theta^k$, nous poserons $W_1 = W_1^0$,

$$W_1^0 = \sum_{k=v_0}^{-1} \sigma_1 o\theta^k + \sigma_2 o\theta^k \mathbf{1}_{\{t_k + (W_1^0 + \sigma_1) o\theta^k \leq 0\}} + t_{v_0} \leq \sum_{k=v_0}^{-1} (\sigma_1 + \sigma_2) o\theta^k + t_{v_0} = V$$

ainsi $W_1 \leq V$.

Considérons la file d'attente P.A.P.S. débutant à t_{v_0} avec les arrivées de clients $(t_k, \sigma_1 o \theta^k)_{k \geq v_0}$ et $(t_k + (W_1 + \sigma_1) o \theta^k, \sigma_2 o \theta^k)_{k \geq v_0}$ (ces derniers ayant priorité sur les premiers en cas d'arrivée simultanée). Si T_1 désigne la fin de la période d'activité commencée à t_{v_0} , la file fonctionnant de façon ininterrompue entre t_{v_0} et T_1 , sur l'intervalle $[t_{v_0}, T_1]$ la charge de cette file s'exprime comme

$$C(t) = \int_{t_{v_0}}^{t-} (\sigma_1 o \theta_s + \sigma_2 o \theta_s \cdot 1_{\{s + (W_1 + \sigma_1) o \theta s \leq t\}}) N(\omega, ds) - (t - t_{v_0})$$

$$C(t) \leq \int_{t_{v_0}}^{t-} (\sigma_1 + \sigma_2) o \theta_s N(\omega, ds) - (t - t_{v_0}) \leq C_0(t)$$

On déduit de cette dernière inégalité que la fin de la période d'activité de la file (F_0) est supérieure à T_1 . Si nous montrons que $C(t) > 0$ sur $\{C_0(s) > 0, \forall s \in [t_{v_0}, t[]\}$, on en déduira que ces deux files d'attente ont les mêmes périodes d'activité (i.e. $\{t/C(t) = 0\} = \{t/C_0(t) = 0\}$). En particulier la file d'attente avec les arrivées

$$(t_k, \sigma_1 o \theta^k)_{k \in \mathbb{Z}} \text{ et } (t_k + (W_1 + \sigma_1) o \theta^k, \sigma_2 o \theta^k)_{k \in \mathbb{Z}}$$

est stable. D'autre part sa charge C vérifie

$$Co\theta_{t_0} = \int_{t_{v_0}}^{t_0-} (\sigma_1 o \theta_s + \sigma_2 o \theta_s \cdot 1_{\{s + (W_1 + \sigma_1) o \theta s \leq t_0\}}) N(\omega, ds) - (t_0 - t_{v_0})$$

d'après la définition de W_1 on en déduit

$$Co\theta_{t_0} = W_1 o \theta t_0$$

on posant

$$W_2 o \theta t_0 = Co\theta_{t_0} + (W_1 + \sigma_1) o \theta t_0$$

(W_1, W_2) constitue un état stationnaire de la file d'attente (F_1)

Il nous reste à montrer que $C(t) > 0$ sur $\{C_0(s) > 0, t_{v_0} < s \leq t\}$
ou encore que la quantité

$$D(t) = \int_{t_{v_0}}^{t-} (\sigma_1 \theta_s + \sigma_2 \theta_s \mathbb{1}_{\{s + (W_1 + \sigma_1) \theta_s \leq t\}}) N(\omega, ds) - (t - t_{v_0})$$

est > 0 sur $\{C_0(s) > 0 / \forall s \in]t_{v_0}, t]\}$, le lemme suivant achève donc la construction d'un état stationnaire de cette file :

lemme 2 :

$$D(t) = \int_{t_{v_0}}^{t-} (\sigma_1 \theta_s + \sigma_2 \theta_s \mathbb{1}_{\{s + (W_1 + \sigma_1) \theta_s \leq t\}}) N(\omega, ds) - (t - t_{v_0}) \text{ est } > 0 \text{ sur } \{\omega / C_0(\theta_s) > 0 / \forall s \in]t_{v_0}, t]\}$$

□ Il suffit de la montrer pour $t = 0$, comme $W_1 \leq V$ sur 0

$$D(0) \geq \int_{t_{v_0}}^{0-} (\sigma_1 \theta_s + \sigma_2 \theta_s \mathbb{1}_{\{s + (V + \sigma_1) \theta_s \leq 0\}}) N(\omega, ds) + t_{v_0}$$

Si pour tout $s \in [t_{v_0}, 0[$ tel que $N(\omega, \{s\}) > 0$, $\underline{s + (V + \sigma_1) \theta_s} \leq 0$ alors

$$D(0) \geq \int_{t_{v_0}}^{0-} (\sigma_1 + \sigma_2) \theta_s N(\omega, ds) + t_{v_0} = C_0(\omega) > 0$$

Sinon il existe $s_0 \in [t_{v_0}, 0[$, $N(\omega, \{s_0\}) > 0$ tel que $\underline{s_0 + (V + \sigma_1) \theta_{s_0}} > 0$

et $s + (V + \sigma_1) \theta_s \leq 0$ pour $s \in [t_{v_0}, s_0[$ tel que $N(\omega, \{s\}) > 0$. Ainsi

$$D(0) \geq \int_{t_{v_0}}^{s_0} (\sigma_1 \theta_s + \sigma_2 \theta_s \mathbb{1}_{\{s + (V + \sigma_1) \theta_s \leq 0\}}) N(\omega, ds) + t_{v_0}$$

Le membre de droite de cette inégalité est égal à

$$s_0 + \int_{t_{v_0}}^{s_0-} (\sigma_1 \theta_s + \sigma_2 \theta_s) N(\omega, ds) - (s_0 - t_{v_0}) + \sigma_1 \theta_{s_0}$$

et ce terme est $s_0 + (V + \sigma_1) \theta_{s_0}$, on en déduit

$$D(0) \geq s_0 + (V + \sigma_1) \theta_{s_0} > 0$$

De ce lemme on déduit que $D(t) = C(t)$ sur $\{\omega/C_0 \circ \theta_s^0 > 0, \forall s \in [t_{v_0}, t]\}$
 donc que $C(t) > 0$ sur $\{\omega/C_0(\theta_s^0) > 0, \forall s \in [t_{v_0}, t]\}$. □

Unicité de (W_1, W_2)

Si (W'_1, W'_2) est un modèle stationnaire de (F_1) la file d'attente

$$\sum_{n \in \mathbb{Z}} \varepsilon_{(t_n, \sigma_1 \circ \theta^n)} + \sum_{n \in \mathbb{Z}} \varepsilon_{(t_n + (W'_1 + \sigma_1) \circ \theta^n, \sigma_2 \circ \theta^n)}$$

est stable, elle a donc des périodes de liberté. Une période d'activité ne peut commencer à un instant $t_n + (W'_1 + \sigma_1) \circ \theta^n$, en effet par hypothèse sur W'_1 , le temps d'attente du client arrivé à t_n est $W'_1 \circ \theta^n$, entre t_n et $t_n + (W'_1 + \sigma_1) \circ \theta^n$ la file d'attente est donc constamment chargée, $t_n + (W'_1 + \sigma_1) \circ \theta^n$ ne peut être un début de période d'activité. Les débuts de période d'activité de cette file forment donc un sous processus

$\sum_{n \in \mathbb{Z}} \varepsilon_{S'_n}$, de $\sum_{n \in \mathbb{Z}} \varepsilon_{t_n}$. Si $S'_0 = t_{v_0}$, avec $S_0 \leq 0 < S_1$ alors W'_1 est W_1^0 construit précédemment relativement à t_{v_0} , on a vu que la période d'activité de

cette file est la même que celle de la file d'attente P.A.P.S de processus d'entrée $\sum_{n \in \mathbb{Z}} \varepsilon_{t_n}$ de service $\sigma_1 + \sigma_2$ débutant à t_{v_0} . Ainsi $\sum_{n \in \mathbb{Z}} \varepsilon_{S'_n}$ est un processus de débuts de période d'occupation de la file d'attente P.A.P.S, de service $\sigma_1 + \sigma_2$ par unicité du modèle stationnaire d'une telle file on en déduit

$$\sum_{n \in \mathbb{Z}} \varepsilon_{S'_n} = \sum_{p \in \mathbb{Z}} \varepsilon_{t_{v_p}} \text{ et donc } W'_1 = W_1$$

ainsi

$$(W'_1, W'_2) = (W_1, W_2)$$

ce qui achève la démonstration de la proposition 2.

La construction précédente nous donne directement :

Proposition 3

W_1 et W_2 vérifient

$$1) \quad W_1 = V - \int_{-\infty}^{0-} \sigma_2 \circ \theta_s^{-1} \mathbb{1}_{\{s+(W_1+\sigma_1) \circ \theta_s > 0\}} N(\omega, ds)$$

$$2) \quad W_2 = \int_0^{(W_1+\sigma_1)^-} \sigma_1 \circ \theta_s^{-1} N(\omega, ds) + \int_{-\infty}^{0-} \sigma_2 \circ \theta_s^{-1} \mathbb{1}_{\{s+(W_1+\sigma_1) \circ \theta_s > 0\}} N(\omega, ds)$$

Remarque

Des relations (1) et (2) on déduit les inégalités intuitives $W_1 \leq V$ et $W_1 + W_2 \geq V$, la façon de répartir la charge est défavorable, toutefois les clients rentrent plus rapidement en service. L'"excès" d'attente $W_1 + W_2 - V$ est

$$\int_0^{(W_1+\sigma_1)^-} \sigma_1 \circ \theta_s^{-1} N(\omega, ds)$$

la charge due aux clients arrivés pendant le premier séjour dans la file.

III Cas d'une arrivée de poisson : calcul des moyennes de W_1 et W_2

Nous supposons :

$\cdot N = \sum_{n \in \mathbb{Z}} \varepsilon_{t_n}$ est un processus de poisson de paramètre 1

\cdot Les familles de variables aléatoires i.i.d.

$\cdot (\sigma_1 \circ \theta^n)_{n \in \mathbb{Z}}, (\sigma_2 \circ \theta^n)_{n \in \mathbb{Z}}, (t_1 \circ \theta^n)_{n \in \mathbb{Z}}$ sont indépendantes entre elles.

$\cdot \sigma_1$ et σ_2 sont de carré intégrable et $E[\sigma_1 + \sigma_2] < 1$

On obtient la proposition suivante :

Proposition 4

$$(a) \quad E(W_1) = \frac{E(\sigma_1^2) + E(\sigma_2^2) + 2E(\sigma_1) E(\sigma_2) \cdot E(\sigma_1 + \sigma_2)}{2(1 - E(\sigma_1 + \sigma_2)) (1 + E(\sigma_2))}$$

$$(b) \quad E(W_2) = \frac{E((\sigma_1 + \sigma_2)^2) + E(\sigma_1)^2 - E(W_1) (1 - E(\sigma_1))}{2(1 - E(\sigma_1 + \sigma_2))}$$

□ Puisque $\sigma_1 + \sigma_2$ est de carré intégrable alors V est intégrable (cf[K]), en intégrant la relation (1) de la proposition 3, par invariance de la mesure $N(\omega, ds)dP$ sur $\Omega \times \mathbb{R}$ pour l'application $(\omega, s) \longrightarrow (\theta_s(\omega), -s)$ (cf[N]p.325) on obtient :

$$E(W_1) = E(v) + E\left(\sigma_2 \int_{0^+}^{+\infty} 1_{\{s < W_1 + \sigma_1\}} N(\omega, ds)\right)$$

$$E(W_1) = E(V) - E(\sigma_2 \cdot N]0, W_1 + \sigma_1[)$$

W_1 étant d'après la construction précédente mesurable par rapport à la tribu engendrée par les variables $(\sigma_1 \circ \theta^n)_{n < 0}$, $(\sigma_2 \circ \theta^n)_{n < 0}$, $(t_1 \circ \theta^n)_{n < 0}$, σ_2 , est indépendant de $N]0, W_1 + \sigma_1[$.

$$E(W_1) = E(V) - E(\sigma_2) E(N]0, W_1 + \sigma_1[)$$

D'après une propriété bien connue des processus de poisson (cf.[N]p. 369 ex.3a)) :

$$E(N]0, W_1 + \sigma_1[) = \int N]0, W_1 + \sigma_1[\bar{dP} = \bar{E}(N]0, W_1 + \sigma_1[)$$

par indépendance de $W_1 + \sigma_1$ et de N restreint à \mathbb{R}^{**+}

$$E[N]0, W_1 + \sigma_1[) = \bar{E}(W_1 + \sigma_1) = E(W_1 + \sigma_1)$$

cette relation et la formule de Pollaczek-Khintchine :

$$E(v) = \frac{1}{2} \frac{E((\sigma_1 + \sigma_2)^2)}{1 - E(\sigma_1 + \sigma_2)}$$

nous donnent

$$E(W_1) = \frac{1}{2} \frac{E((\sigma_1 + \sigma_2)^2)}{1 - E(\sigma_1 + \sigma_2)} - E(W_1 + \sigma_1) E(\sigma_2)$$

la première relation est établie.

Suivant la même méthode avec W_2 en intégrant la relation (2) de la proposition (3)

$$E(W_1) + E(W_2) = E(V) + E(\sigma_1) E(W_1 + \sigma_1)$$

$$E(W_2) = E(V) + E(\sigma_1^2) - E(W_1) (1 - E(\sigma_1)).$$

□

Remarque

On déduit de cette proposition la moyenne de l'excès d'attente par rapport à la discipline P.A.P.S.

$$\underline{E(W_1 + W_2 - V) = E(\sigma_1^2) + E(W_1) E(\sigma_1)}$$

BIBLIOGRAPHIE

Disney R.L, Wyszweanski R.J. [D]

"Feedback queues in the modelling of computer systems : a survey"

Preprint

Neveu J. [N]

" sur les processus pontuels"

école d'été de Probabilités de Saint Flour 1976

L.N. in Math. n°598

Kingman J.F.C. [K]

"Some inequalities for the queue GI/G/1"

Biométrie - 1962. 49. p315-324.

Imprimé en France

par

l'Institut National de Recherche en Informatique et en Automatique

