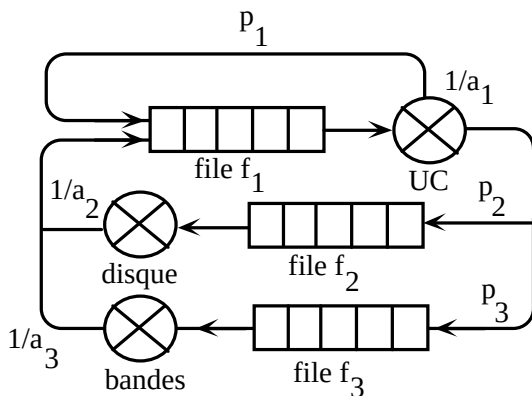


Projet de Recherche Opérationnelle

But : Simulation du comportement d'un système informatique pour améliorer ses performances.

Sujet : Modèle du serveur central (réseaux de files d'attente).

Considérons un système informatique composé d'un ordinateur disposant d'unités de disques et de bandes. Pour étudier a priori les performances (temps de réponse, débit, équilibrage de la configuration) d'un tel système, il est souvent intéressant de modéliser son comportement (ici, par un réseau de files d'attente).



Le modèle que nous utiliserons est constitué de trois files :

- la file f_1 , file d'attente de l'unité centrale ;
- la file f_2 , file d'attente disque ;
- la file f_3 , file d'attente du canal de gestion des unités de bandes magnétiques.

(Nous ne prenons pas en compte les sorties vers les périphériques lents qui sont faites en dehors des heures de pointe et à partir de fichiers disques.)

Le déroulement d'un travail s'effectue de la façon suivante : il entre dans la file d'attente de l'unité centrale et attend. Lorsqu'il obtient l'unité centrale, il exécute des instructions pendant un temps distribué exponentiellement de moyenne $1/a_1$. Cette séquence se termine à l'occurrence de l'un des trois événements suivants :

- a) le travail est terminé, cet événement se produit avec la probabilité p_1 ;
- b) le travail a demandé à lire ou à écrire sur le disque. Ceci se produit avec une probabilité égale à p_2 . Le travail sort alors de l'unité centrale et est placé dans la file d'attente du disque f_2 ;
- c) le travail a fait une demande d'accès à une unité de bandes magnétiques. Ceci se produit avec la probabilité p_3 ; le travail sort alors de l'unité centrale et est placé dans la file d'attente des bandes magnétiques f_3 .

Bien entendu, $p_1 + p_2 + p_3 = 1$.

Lorsqu'un travail est dans une file f_i ($i = 2, 3$) il attend son tour, puis reçoit un service distribué exponentiellement et de moyenne $1/a_i$ (c'est à dire que $1/a_2$ est le temps moyen pour écrire ou lire une page sur le disque et que $1/a_3$ est le temps nécessaire pour écrire ou lire un bloc sur une bande magnétique). Les demandes d'entrées/sorties sont uniformément réparties durant toute la durée du travail.

On considère que le système traite au maximum N travaux (le degré de multiprogrammation est N). L'état du système est alors décrit par un vecteur (n_1, n_2, n_3) où n_i représente le nombre de travaux en attente ou en cours de traitement dans la file f_i . On a, bien entendu $n_1 + n_2 + n_3 \leq N$. Au démarrage, les travaux entrent tous dans la file f_1 ; à la saturation, c'est à dire lorsque le degré de multiprogrammation

est maximal, dès qu'un travail se termine il est remplacé par un autre (qui provient, en pratique, d'une quatrième file d'attente en amont de la file f_1 pour éviter l'engorgement du système). Dans ces conditions, p_1 (probabilité pour qu'un travail quittant l'unité centrale soit terminé) est égale à la probabilité qu'un nouveau travail entre dans la file d'attente f_1 (ceci explique la boucle sur la file f_1 dans la figure). On parle, dans ce cas, de recyclage du travail.

On appellera :

- N_1, N_2, N_3 : les nombres de passage d'un travail par la file f_1 , la file f_2 et la file f_3 entre son entrée dans le système et sa fin d'exécution ;
- T_1, T_2, T_3 : les temps d'utilisation effective de l'unité centrale, du disque et de l'unité de bandes par un travail ;
- q_1, q_2, q_3 : les probabilités que les files f_1, f_2, f_3 ne soient pas vides, c'est à dire les taux d'utilisation des ressources ;
- ϕ_{11} : le débit du système, c'est à dire le nombre moyen de travaux traités, donc quittant le système, par unité de temps.

Les probabilités p_i et les durées moyennes $1/a_i$ sont des constantes qui ne sont a priori liées par aucune relation. Ces valeurs caractérisent le comportement statistique des travaux traités par le système. L'arrivée des travaux dans le système est supposée poissonnienne (indépendance des instants d'arrivée), λ désignant le nombre moyen de travaux arrivant par unité de temps.

Question 1 :

Calculer les nombres moyens $\bar{N}_1, \bar{N}_2, \bar{N}_3$ de passage d'un travail dans l'unité centrale, la file disque (f_2) et bande (f_3). Comparer aux résultats théoriques.

Question 2 :

Tracer les courbes $n_i(t)$; on mettra en évidence le régime transitoire, suivi du régime permanent. Donner les valeurs moyennes de n_i en régime permanent. Pouvez vous justifier ces résultats ? Tracer le régime transitoire de $n_1(t) + n_2(t) + n_3(t)$.

Question 3 :

Calculer, en régime permanent, les taux d'utilisation q_1^*, q_2^*, q_3^* des différentes ressources, le débit ϕ_{11} du système et le temps total de traitement de tous les travaux. Justifier ces résultats. Quel est le "goulet d'étranglement" du système ?

Question 4 :

Pour améliorer les performances du système, on ajoute un deuxième disque (D2), identique au premier (D1). Selon la nature du travail, le premier ou le second disque est sollicité. On notera d_1 le pourcentage des travaux accédant au disque D1 (et donc, $1 - d_1$ le pourcentage des travaux accédant au disque D2). Même étude qu'à la question 3.

Question 5 :

Comment peut-on améliorer encore les performances du système ? De combien ? La simulation des solutions proposées est facultative.

Remarques :

Les données du problème seront introduites dans un fichier de paramètres "para.p" dont le format est le suivant :

1/351	probabilité_p1
300/351	probabilité_p2
5/351	temps_moyen_UC(seconde)
0.05	temps_moyen_disque(seconde)
0.2	temps_moyen_bande(seconde)
0.25	lambda(nombre_de_travaux_par_seconde)
50	ordre_de_multiprogrammation
1000	nombre_de_travaux_a_traiter
0.5	taux_d'utilisation_du_disque_D1

Travail à réaliser :

Le choix du langage pour écrire le logiciel est laissé libre, l'exécution devra pouvoir tourner sur un PC standard.

Afin de pouvoir tester un autre fichier de paramètres, on devra pouvoir passer en premier argument, le nom du fichier de paramètres.

Les résultats sur ordinateur devront être présentés sous forme de tableaux.

Un exposé oral présentant la méthode et les résultats sera réalisé par chaque groupe. Le logiciel sera prêt pour que l'enseignant puisse le tester.

Travail à remettre :

Un rapport détaillé sur l'analyse et la conception de votre projet et un jeu de résultats sur papier (15 pages maximum).

Date de remise : le mardi 2 avril 2013 à 10 heures au plus tard.