

Conception des systèmes répartis

3ième Année Informatique et Mathématiques Appliquées, Mastères SLCP et RTSA
Durée : 2 heures — Documents autorisés : Précis et notes de cours

Janvier 2005

1 Protocoles de diffusion (multicast)

Dans les intergiciels de communication offrant une pile de protocoles de diffusion (par exemple Java-Groups), il existe en particulier le protocole fifo de diffusion et le protocole causal de diffusion. Dans ce qui suit, on considère, pour simplifier, uniquement des diffusions intra-groupe : seuls les membres du groupe diffusent des messages au(x) groupe(s) au(x)quel(s) il(s) apparten(en)t.

Notation Dans ce qui suit, $e_x(m)$ désigne l'événement de diffusion par x du message m et $d_x(m)$ désigne l'événement de délivrance du message m au destinataire x .

Protocole fifo Le protocole fifo de diffusion vers un groupe $G = \{p_i\}_{0 \leq i < n}$ de cardinalité n assure la propriété suivante : si 2 messages sont diffusés par un même membre, ceux-ci sont délivrés aux autres membres du groupe G dans le même ordre que leur ordre chronologique de diffusion.

$$\forall p \in G : \forall m, m' \in Messages, r \in G : e_p(m) \prec e_p(m') \Rightarrow d_r(m) \prec d_r(m')$$

Protocole causal Le protocole causal de diffusion vers un groupe $G = \{p_i\}_{0 \leq i < n}$ de cardinalité n assure la propriété suivante : si 2 messages sont diffusés par des membres du groupe, ceux-ci sont délivrés aux autres membres du groupe G dans le même ordre que leur ordre causal de diffusion. S'il n'existe pas de précédence causale entre les deux messages, aucun ordonnancement de la délivrance des messages aux sites destinataires n'est garanti.

$$\forall p, p' \in G : \forall m, m' \in Messages, r \in G : e_p(m) \prec e_{p'}(m') \Rightarrow d_r(m) \prec d_r(m')$$

Questions (2 points par question)

1. Dessiner un chronogramme le plus simple possible (groupe de 4 processus) montrant que, dans le cas du protocole fifo, les messages diffusés ne sont pas forcément délivrés dans le même ordre aux différents destinataires des mêmes messages?
2. On considère deux possibilités pour assurer une diffusion fifo entre n processus :
 - soit, on crée n groupes fifo G_i contenant les mêmes membres mais dans lesquels, pour le groupe G_i , seul le processus p_i utilise le groupe pour diffuser des messages aux membres.
 - soit on crée un seul groupe G et chaque processus diffuse des messages aux membres via ce groupe unique.Montrer que les deux approches sont équivalentes : pour cela, montrer que les propriétés d'ordonnement des messages assurées par les deux approches sont les mêmes.
3. Montrer que le protocole causal délivre les messages dans un ordre respectant le protocole fifo.
4. Montrer qu'il est impossible, avec le protocole causal de diffusion, qu'un processus p , membre d'un groupe G comportant n membres, reçoive plus de n messages consécutifs non causalement liés au niveau de leur diffusion.

2 Prise de cliché global

On considère le chronogramme de la figure (1) représentant un calcul réparti entre 3 processus A,B et C.

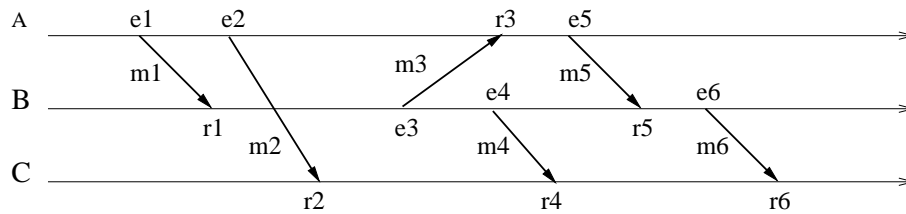


FIG. 1 – Calcul réparti

Questions (2 points par question)

- Donner, sous forme de l'ensemble des événements qui les compose, un exemple de coupure incohérente, un exemple de coupure cohérente sans message en transit et un exemple de coupure cohérente avec 2 messages en transit. Placer sur le chronogramme les événements de prise de cliché local associés à chaque coupure en notant par c_i , ce type d'événement.
- Montrer que l'intersection (ensembliste) de 2 coupures cohérentes est encore une coupure cohérente.

Algorithme de prise de cliché utilisant un jeton circulant et un observateur

On se propose d'étudier un algorithme de prise de cliché global utilisant un jeton circulant et un processus observateur qui collecte les clichés partiels évalués et transmis par les processus applicatifs. On suppose que le processus observateur fait partie de l'anneau virtuel sur lequel circule le jeton et qu'il est en particulier à l'origine de la création et propagation du jeton. Le jeton exécute des tours d'anneau numérotés à partir de 1. Par convention, on suppose que les processus applicatifs sont ordonnés sur l'anneau selon leur indice et que le processus observateur est prédécesseur de P_1 et successeur de P_N .

Les communications sont supposées fiables, sans erreur ni duplication. Le réseau est connexe et les liaisons point à point sont **fifo** : tous les messages émis par un processus A vers un autre processus B sont reçus par B dans l'ordre chronologique de leur envoi par A.

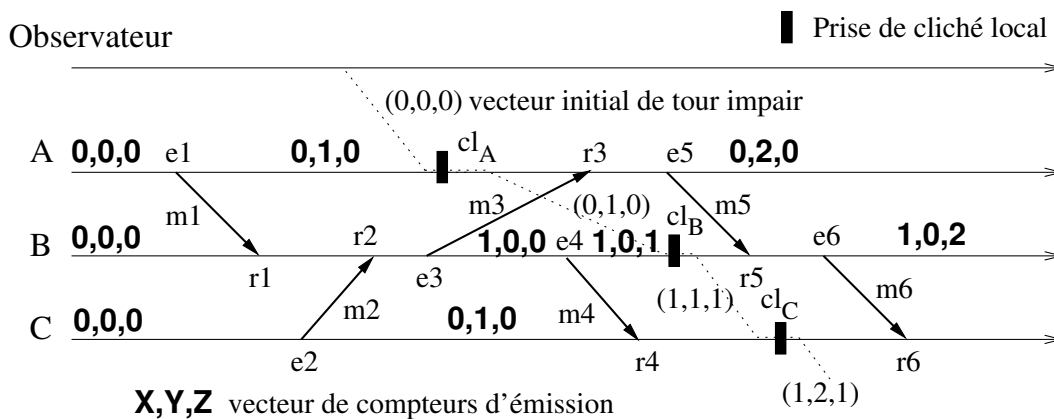


FIG. 2 – Circulation du jeton

Description de l'algorithme

Les principes de base de l'algorithme sont les suivants :

- le comportement de chaque processus est identique : un processus exécute son calcul applicatif et participe à une prise de cliché global lorsqu'il reçoit le jeton.
- chaque processus mémorise le nombre de passages du jeton de façon à détecter si la visite du jeton est associée à un tour pair ou impair afin d'exécuter un traitement différent selon cette parité.
- les visites du jeton lors d'un tour impair permettent de fixer une frontière de coupure en prenant des cliqués locaux et en évaluant un vecteur comptabilisant le nombre de messages émis avant cette coupure.
- les visites du jeton lors d'un tour pair permettent de déterminer les messages en transit associés à la coupure précédente à partir du vecteur évalué au tour précédent.

Tour impair : La réception du jeton circulant par un processus P_i **lors d'un tour impair** provoque une prise de cliché de l'état local cl_i du processus. De plus, le jeton est valué et contient un vecteur d'entiers de dimension N égale au nombre de processus. Ce vecteur $E[N]$ a pour objectif de collecter, lors d'un tour impair, le nombre de messages émis vers chaque processus jusqu'à cette visite du jeton : autrement dit, après un tour complet impair, chaque élément $E[k]$ est égal au nombre de messages émis vers P_k avant la prise de chaque cliché local quel que soit le processus d'émission. Pour assurer la mise-à-jour correcte de ce vecteur, chaque processus P_i doit donc lui-même gérer un vecteur $ve_i[N]$ de dimension N qui comptabilise le nombre de messages émis par ce processus vers chaque autre processus. Autrement dit, $ve_i[k]$ est égal au nombre de messages émis par P_i vers P_k ¹. La figure (2) montre un exemple de parcours impair du jeton entre 3 sites, ainsi que la valeur du vecteur E transporté par le jeton jusqu'à sa valeur $(1, 2, 1)$ à la fin du tour. Lorsque le processus observateur lance un tour impair, le vecteur E est réinitialisé à zéro.

Tour pair ; La réception du jeton circulant **lors d'un tour pair** fournit au processus le vecteur final évalué durant le tour précédent (impair). Grâce à ce vecteur, chaque processus peut déterminer quels messages reçus ou à recevoir doivent être considérés comme des messages en transit vis-à-vis des clichés locaux mémorisés lors du tour précédent du jeton. Une fois collectés, ces messages seront transmis au processus observateur.

Questions (2 points par question)

7. Compte tenu du placement des processus sur l'anneau et du déroulement des prises de clichés locaux lors d'un tour impair du jeton, montrer que le cliché global final ne peut pas contenir de message en transit émis d'un processus P_i , avec $1 \leq i < N$ vers un processus P_{i+1} . Quelle propriété similaire serait vérifiée si le protocole de communication était causalement ordonné (et non plus fifo) ?
8. Lors d'un tour impair, décrire l'opération de mise-à-jour du vecteur E en fonction du vecteur ve_i du processus visité P_i .
9. Lorsqu'un processus P_i reçoit le vecteur E lors d'un tour pair, donner, en le justifiant, le prédicat qui permet de décider si un message reçu m , émis par un processus P_k , doit être considéré comme en transit dans le cliché global. Pour que ce test soit possible, quel variable d'état doit être gérée par chaque processus ?
10. Pour pouvoir déterminer les messages en transit, que doit faire chaque processus P_i
 - lorsqu'il reçoit un message applicatif ?
 - lorsqu'il reçoit le jeton d'un tour impair ?
 - lorsqu'il reçoit le jeton d'un tour pair ?Préciser les variables qui lui sont nécessaires et leur mise-à-jour.

¹Pour tous ces vecteurs ve_i , on a naturellement **invariant** $\forall i : ve_i[i] = 0$.