

**Journées Francophones sur la Cohérence des Données en Univers Réparti
CDUR'2005**

**Une famille de protocoles de diffusion
préservant la cohérence**

Laurent GEORGE

ECE
53 rue de Grenelle
Paris

Pascale MINET

INRIA Rocquencourt
Domaine de Voluceau
Le Chesnay

Steven MARTIN

LRI
Univ. Paris-Sud 11
Orsay

Introduction

Modèles

Classification

Solution proposée

Exemple

Conclusion



PLAN

- ✓ Introduction
- ✓ Modèles
- ✓ Classification des protocoles de diffusion
- ✓ Solution proposée
- ✓ Exemple
- ✓ Conclusion

INTRODUCTION

- ✓ De nombreuses applications temps réel font appel à un service de diffusion dans un groupe.
- ✓ Diffusion fiable
 - Tous les membres corrects du groupe destinataire sont d'accord pour remettre le message ou pour le rejeter
 - Insuffisant pour assurer la cohérence.

Dans groupe= $\{N_1, N_2\}$ N_1 remet m_1 puis m_2 tandis que N_2 remet m_2 puis m_1

- ✓ Diffusion atomique
 - Tous les membres corrects remettent les messages dans le même ordre

MODELE RESEAU

- ✓ Nœuds interconnectés par des liens FIFO.
 - Nœuds forment des groupes
 - Groupes sont supposés déjà établis.

- ✓ Délais réseau dans $[L_{\min}, L_{\max}]$

- ✓ Horloges locales ε -synchronisées.

MODELE DE DEFAILLANCE

- ✓ Défaillance des nœuds par arrêt.
 - Tolérer jusqu'à f_c défaillances de noeud

- ✓ Défaillance du réseau par omission
 - Tolérer jusqu'à f_o omissions du réseau

- ✓ Horloges locales des nœuds non défaillants sont monotones croissantes et ont une dérive bornée.

MODELE DE TRAFIC

- ✓ Chaque flux est supposé suivre une séquence **fixe** de nœuds.

Par exemple, MPLS peut être utilisé pour fixer le chemin suivi par le flux.

- ✓ Les flux sont caractérisés par des arrivées **sporadiques**.

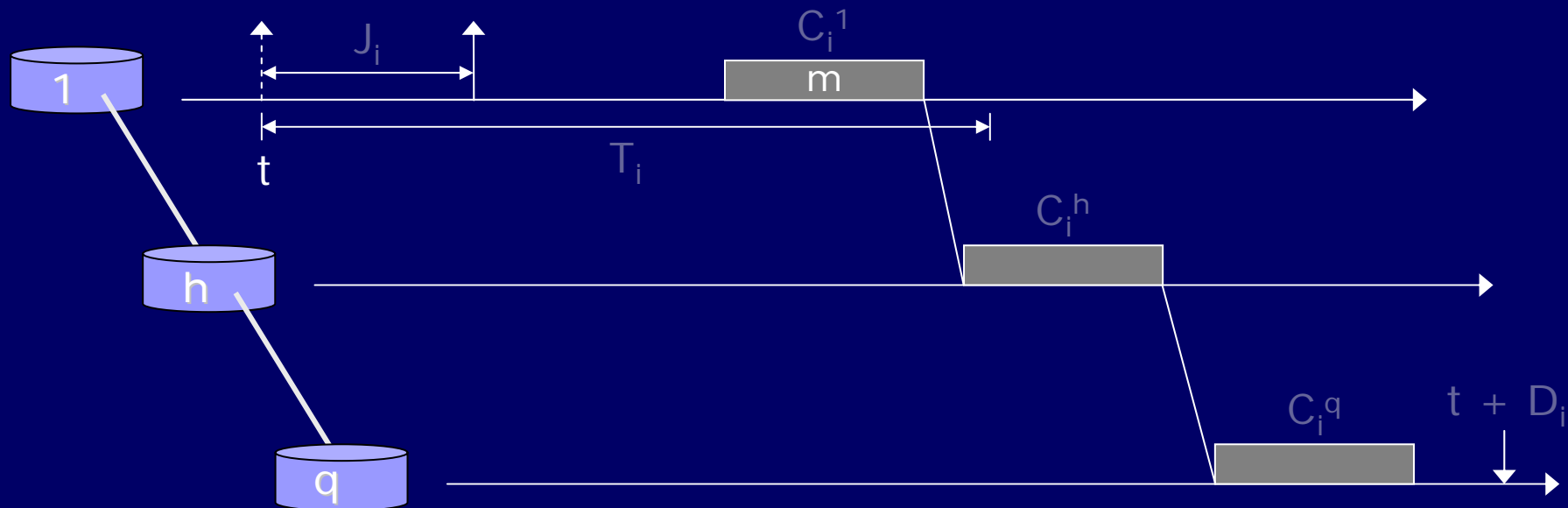
Chaque flux τ_i est défini par : T_i, C_i^h, J_i, D_i

- ✓ Pour tout message m_i d'un flux τ_i
 - $\text{src}(m_i)$ désigne le nœud source de m_i
 - $\text{dest}(m_i)$ désigne les nœuds destination de m_i

MODELE DE TRAFIC

- ✓ Les flux sont caractérisés par des arrivées **sporadiques**.

Chaque flux τ_i est défini par : T_i, C_i^h, J_i, D_i



DIFFUSION FIABLE UNIFORME

- ✓ **(P1) Validité** : Si un noeud correct diffuse un message m , alors il existe un noeud correct de $dest(m)$ qui remet le message m .
- ✓ **(P2) Accord Uniforme** : Si un noeud, correct ou non, remet un message m , alors tous les noeuds corrects de $dest(m)$ finissent par remettre le message m .
- ✓ **(P3) Intégrité Uniforme** : Pour tout message m , chaque noeud j , correct ou non, délivre m au plus une fois et uniquement si j appartient à $dest(m)$ et m a été préalablement diffusé par $src(m)$.

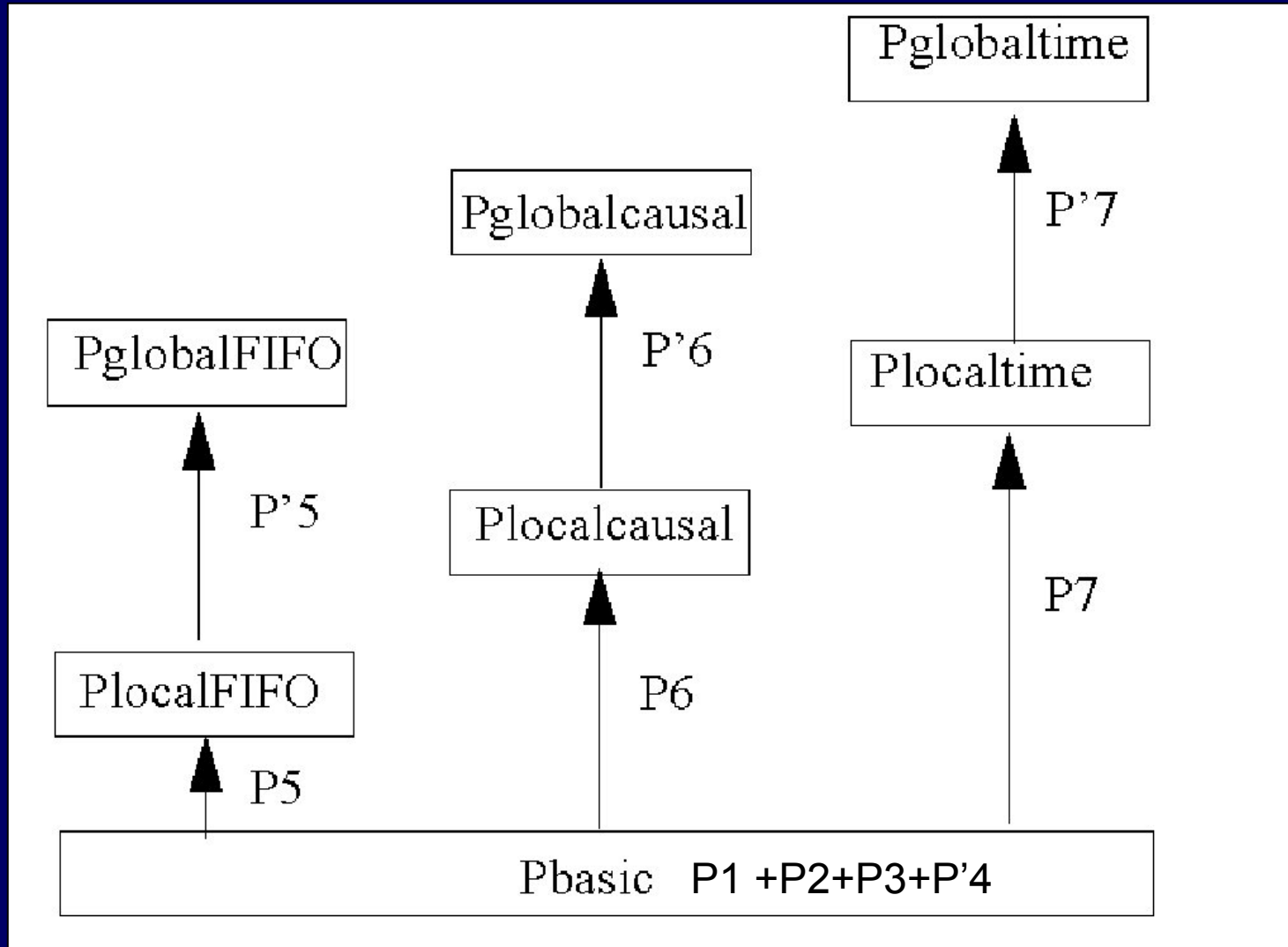
ORDRE LOCAL UNIFORME

- ✓ **(P4) Ordre total local uniforme** : Si les noeuds j et k , corrects ou non, remettent les messages m_1 et m_2 avec $dest(m_1) = dest(m_2)$, alors j remet m_1 avant m_2 ssi k remet m_1 avant m_2 .
- ✓ **(P5) Ordre FIFO local uniforme** : Si sur un noeud la demande de diffusion de m_1 précède celle de m_2 et $dest(m_1) = dest(m_2)$ alors aucun noeud, correct ou non, ne remet m_2 s'il n'a pas au préalable remis m_1 .
- ✓ **(P6) Ordre causal local uniforme** : Si la demande de diffusion de m_1 précède causalement celle de m_2 dans $dest(m_1)$ et $dest(m_1) = dest(m_2)$ alors aucun noeud, correct ou non, ne remet m_2 s'il n'a pas au préalable remis m_1 .

ORDRE GLOBAL UNIFORME

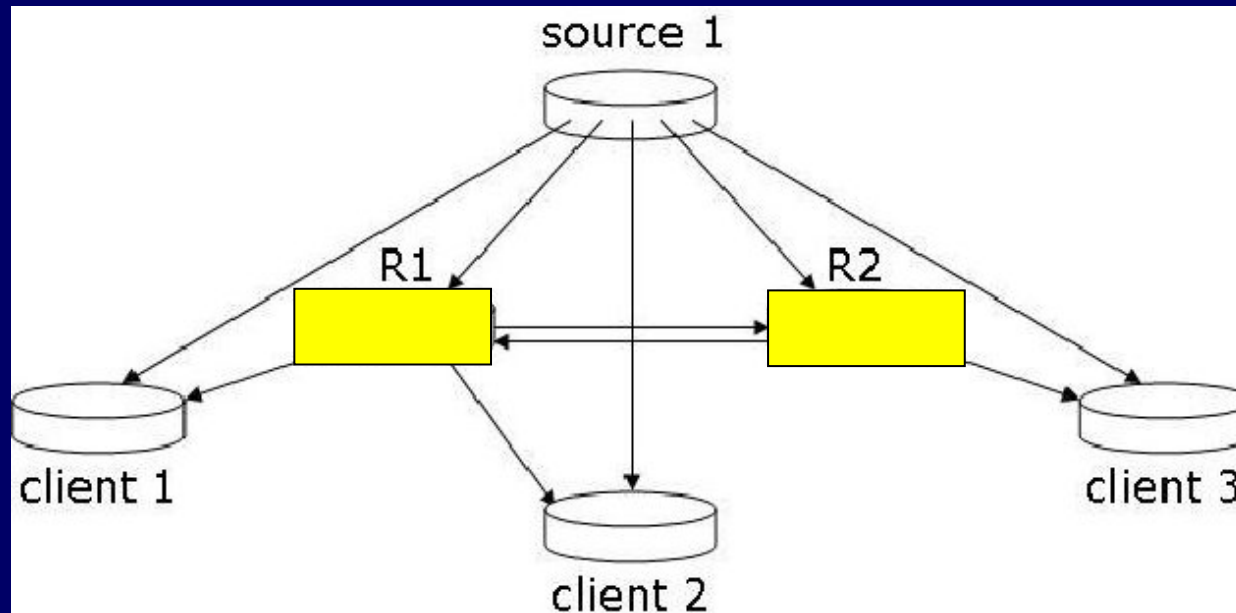
- ✓ **(P'4) Ordre total global uniforme** : La relation 'est remis avant' est acyclique.
Un message est dit remis avant un autre ssi ceci est vrai sur au moins un noeud, correct ou non. .
- ✓ **(P'5) Ordre FIFO global uniforme** : Si sur un noeud la demande de diffusion de m_1 précède celle de m_2 alors aucun noeud dans $dest(m_1)$, correct ou non, ne remet m_2 s'il n'a pas au préalable remis m_1 .
- ✓ **(P'6) Ordre causal global uniforme** : Si la demande de diffusion de m_1 précède causalement celle de m_2 alors aucun noeud dans $dest(m_1)$, correct ou non, ne remet m_2 s'il n'a pas au préalable remis m_1 .

FAMILLE DE PROTOCOLES DE DIFFUSION



DIFFUSION FIABLE

- ✓ Basée sur le concept de relais.
- Chaque source j a f_c+f_o+1 relais différents de j dans chaque groupe G où elle peut demander une diffusion fiable.



- ✓ Condition de faisabilité $\text{card}(G-\{j\}) \geq f_c+f_o+1$.

DIFFUSION ATOMIQUE

✓ Basée sur la diffusion fiable précédente

- Un message m dont la diffusion atomique est demandée à l'instant t est estampillé avec t
- Est diffusé avec la diffusion fiable précédente
- Attend l'instant $t+F_i+\varepsilon$ pour remettre le message.

✓ Calcul de F_i dépend du protocole de diffusion retenu.

- Basé sur l'approche par trajectoire pour calculer le pire temps de réponse de bout-en-bout

- Pour PBasic: F_i est le temps de réponse de bout en bout maximum pour $\text{dest}(m_i)$
- Pour PlocalFIFO: F_i est le temps de réponse de bout en bout maximum pour l'ensemble des clients d'une même source
- Pour PglobalFIFO: F_i est le temps de réponse de bout en bout maximum étendu à l'ensemble des groupes destination

MODELE D'ORDONNANCEMENT

✓ Ordonnancement FP/FIFO* non-préemptif

- FP

- Les messages sont ordonnancés en fonction de leur priorité fixe.
- La priorité fixe d'un message reflète le degré d'importance du flux.

-FIFO*

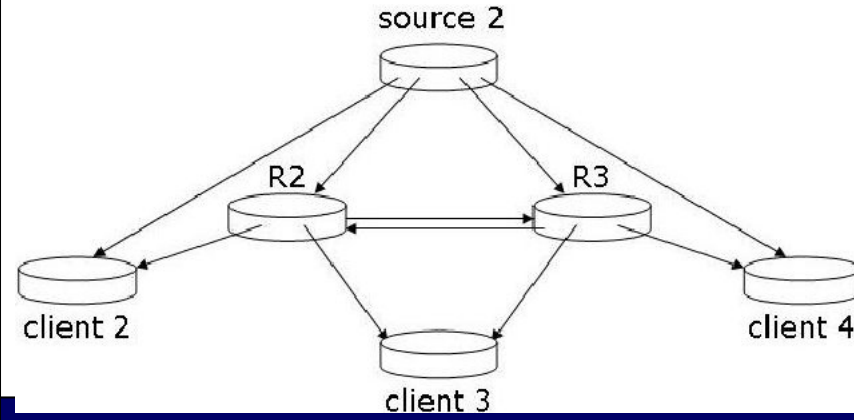
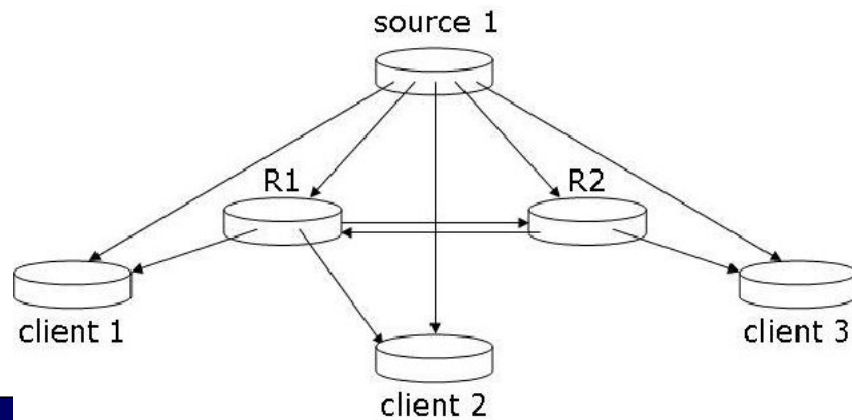
- Si plusieurs messages ont la même priorité fixe, ils sont départagés par leur instant d'arrivée dans le réseau.

APPROCHE PAR TRAJECTOIRE

- ✓ Calcul du temps de réponse pire cas de bout-en-bout pour les flux les plus prioritaires

$$R \max_i = \sum_{h \in P_i} \left(\sum_{first_{j,i}=h, j=1..n} \left(1 + \frac{S \max_i^h + J_j^h}{T_j} \right) \cdot C_j^{slow_{j,i}} \right) \\ + \sum_{\substack{h \in P_i \\ h \neq slow_i}} \max \{ C_{\max}^h \} + \delta_i + (|P_i| - 1) \cdot L_{\max}$$

EXAMPLE



	C_i	T_i	J_i	D_i	P_i
τ_1	2	10	0	25	1
τ_2	3	10	0	30	1

R_i	Client 1	Client 2	Client 3	Client 4
τ_1	19	21	21	-
τ_2	-	28	28	23

CONCLUSION

- ✓ Nous avons présenté une famille de protocoles de diffusion préservant la cohérence.



Racine = Diffusion fiable uniforme tolérant :
fc crashes de noeuds
et fo omissions réseau

- ✓ Définition d'un ordre total
 - **Dépend des propriétés requises par l'application**
 - local/global
 - FIFO/causal/....
 - Est obtenu à partir du temps de réponse pire cas**
- ✓ Calcul du temps de réponse pire cas de bout-en-bout selon l'approche par trajectoire.