Transactions Concurrence d'accès

HLIN511

Pascal Poncelet

LIRMM

Pascal.Poncelet@lirmm.fr

http://www.lirmm.fr/~poncelet



• Dans votre session:

```
LIRE (A)
A := A+100
ECRIRE (A)
LIRE (B)
B := B +100
ECRIRE (B)
```





ECRIRE (B)



 Nécessité d'exécuter les programmes et requêtes de façon concurrente

- Garantir que l'exécution simultanée de plusieurs programmes se fait correctement
 - Cohérence de la base de données
 - Exécution correcte des programmes



- Régler les conflits lecture / écriture
- Garder de très bonnes performances
- Eviter les blocages

- De nombreux contextes
 - Annuaire téléphonique (accès en lecture)
 - Systèmes de réservation de billets (accès en lecture et mise à jour)



Les problèmes de concurrence

Perte d'opérations

```
T_1: A:=LIRE(A);
T_2: B:=LIRE(A);
T_2: B:=B+1;
T_2: A:=ECRIRE(B);
T_1: A:=A*2;
T_1: ECRIRE(A);
Valeur de A pour T_1?
```

Cette anomalie résulte de la dépendance ECRITURE-ECRITURE entre actions de plusieurs transactions qui s'interfèrent



Les problèmes de concurrence

Introduction d'incohérence

```
A = B (A DOIT TOUJOURS ETRE EGAL A B)
```

```
T_1 : A := A * 2;
                               T_2 : A := A + 1;
                               T_2 : B := B + 1;
                    T_1 : B := B * 2;
Test : A=5, B=5
T_1: A:=A*2; // A = 10
    T_2 : A := A + 1; // A = 11
    T_2: B:=B+1; //B=6
```

T₁: B:=B*2; // B=12 et A=11 !!!

Les problèmes de concurrence

Vision SGBD

```
T_1: A:=A*2;
T_2: A:=A+1;
T_2: B:=B+1;
T_1: B:=B*2;
T_1 L(A), E(A)
L(B), E(B)
L(A), E(A), L(B), E(B)
```

Séquence = $<T_1L(A), T_1E(A), T_2L(A), T_2E(A), T_2L(B), T_2E(B), T_1L(B), T_1E(B)>$



- Transaction : unité de programme exécutée sur un SGBD
- Une transaction débute souvent par un programme d'application
 - DEBUT TRANSACTION
 - Accès à la base de données (lectures, écritures)
 - Calculs en mémoire centrale
 - FIN TRANSACTION
 - COMMIT (si la transaction est bonne) ou ROLLBACK (s'il y a une erreur)



- Une transaction = ensemble d'instructions séparant un COMMIT ou un ROLLBACK du COMMIT ou du ROLLBACK suivant
- Une transaction T_i fait l'état de la base d'un état consistant à un autre état consistant
 - A noter que pendant l'exécution de T_i soit la transaction dans son intégralité fonctionne soit les opérations doivent être annulées (ROLLBACK)
- Chaque transaction T_i est composée d'une séquence d'actions <a₁, a₂, ..., a_{ni}>



- Une exécution simultanée (Histoire) des transactions {T₁, T₂, T_n} est une séquence d'actions
 - $-H = \langle ai_1j_1, ai_2j_2 ai_kj_k \rangle$ telle que $a_{ij} \langle a_{ij+1}$ pour tout i et tout j et quelque soit a_{ij} de T_1 ,... T_n , a_{ij} est dans H
 - C'est une séquence d'actions complète respectant l'ordre des actions des transactions

– Exemple :



$$H = \langle T_1L(A), T_1E(A), T_2L(A), T_2E(A), T_2L(B), T_2E(B), T_1L(B), T_1E(B) \rangle$$

 Quand une transaction est validée (par commit) toutes les opérations sont validées et on ne peut plus en annuler aucune : les mises à jour sont définitives

 Quand une transaction est annulée (par ROLLBACK ou par une panne), on annule toutes les opérations depuis le dernier commit ou ROLLBACK ou depuis le premier ordre SQL s'il n'y a eu ni COMMIT ni ROLLBACK



- Problèmes à éviter :
 - Ecriture impropre (dirty write)
 - Lecture impropre (dirty read)
 - Lecture non répétable (non repeatable read)
 - Données fantômes (phantom read)



Ecriture impropre

- La mise à jour d'une transaction est écrasée par une autre transaction
 - T_1 crédite le compte A de 100 euros. T_2 débite le compte A de 50 euros. Au départ A = 500. Si T_1 et T_2 sont appliquées correctement A = 550.

```
Au départ A = 500. Si T_1 et T_2 sont appliquées correctement A = 550 T_1 (crédit) T_2 (débit) Lire (A); {A = 500} T_2 Lire (A); {A=500} T_3 A := A+100; {A=600} T_3 A:=A-50; {A=450} T_3 Ecrire (A); {A=600} T_3 Ecrire (A); {A=450}
```

La valeur finale de A = 450. Le crédit de T₁ n'est pas pris en compte



Lecture Impropre

Pour le vol 2, U1 ajoute 10 places et annule sa transaction, U2 souhaite 7 places si elles sont disponibles

```
\mathsf{T}_1
                                      T_2
                                     BEGIN TRANSACTION
BEGIN TRANSACTION
UPDATE T VOL
                                     IF (SELECT VOL PLACES LIBRES FROM T VOL
SET VOL_PLACES_LIBRES=
                                                   WHERE VOL_ID=2)>= 7
VOL_PLACES_LIBRES+10
WHERE VOL_ID = 2
                                     THEN
                                     BEGIN
                                          UPDATE T VOL
                                          SET VOL PLACES LIBRES= VOL PLACES LIBRES-7
                                          WHERE VOL_ID= 2
                                          INSERT INTO T_CLIENT_VOL VALUES (77, 2, 7)
ROLLBACK
                                          COMMIT
                                     END
                                     ELSE
```



Lecture Impropre

Tem ps	T_Vol Vol_id	Vol_ref	Vol_places_libres	T_Clien Cli_id		Vol_pla	aces_prises	Transactions
t1	2	AF 121	6					Début transaction pour u2
t2	2	AF 121	6					Début transaction pour u1
t3	2	AF 121	16					(+10) PAR U1
t4	2	AF 121	16					
t5	2	AF 121	6					Fin transaction rollback u1
t6	2	AF 121	-1					(-7) par U2
t7	2	AF 121	-1	77	2	7	7	
t8	2	AF 121	-1	77	2	•	7	Fin transaction commit u2

Le temps d'un update abandonné, la transaction a lu des informations qu'elle ne devait pas voir et en a déduit qu'elle pouvait réserver surbooking

Lecture Non Répétable

U2 souhaite toutes les places d'un vol s'il y en a plus de 4

 T_1

BEGIN TRANSACTION
UPDATE T_VOL
SET VOL_PLACES_LIBRES=
VOL_PLACES_LIBRES-4
WHERE VOL_ID = 2
COMMIT

```
T_2
BEGIN TRANSACTION
IF (SELECT VOL PLACES LIBRES FROM T VOL
WHERE VOL ID=2)>= 4
THEN
BEGIN
     UPDATE T VOL
      SET VOL PLACES LIBRES= VOL PLACES LIBRES-
       (SELECT VOL PLACES LIBRES FROM T VOL
       WHERE VOL ID= 2)
     WHERE ID=2
     INSERT INTO T CLIENT VOL VALUES (77, 2,
           (SELECT VOL PLACES LIBRES
           FOM T VOL WHERE VOL ID = 2))
     COMMIT
END
ELSE
     ROLLBACK
```



Lecture Non Répétable

Tem ps	T_Vol Vol_id	Vol_ref	Vol_places_libres	T_Clien Cli_id		Vol_pla	ces_prises	Transactions
t1	2	AF 121	6					Début transaction pour u2
t2	2	AF 121	6					Début transaction pour u1
t3	2	AF 121	2					(-4) par U1
t4	2	AF 121	2					
t5	2	AF 121	2					Fin transaction commit u1
t6	2	AF 121	0					(-2 toutes les places libres) par U2
t7	2	AF 121	0	77	2	2	!	
t8	2	AF 121	0	77	2	2		Fin transaction commit u2

U2 ne voulait que des places s'il y en avait plus de 4 et il se retrouve avec 2 places



Données Fantômes

U2 souhaite n'importe quel vol pas cher pour amener 11 personnes

ROLLBACK

 T_1

BEGIN TRANSACTION INSERT INTO T_VOL VALUES (5, 'AF 111', 125) COMMIT

```
T_2
BEGIN TRANSACTION
IF EXISTS (SELECT * FROM T VOL
       WHERE VOL PLACES LIBRES >=11)
THEN
    BEGIN
    UPDATE T VOL
      SET VOL PLACES LIBRES= VOL PLACES LIBRES-11
      WHERE VOL PLACES LIBRES >= 11
     INSERT INTO T CLIENT VOL VALUES (77, 4, 11)
    COMMIT
END
ELSE
```



Données Fantômes

Tem ps	T_Vol Vol_id	Vol_ref	Vol_places_libres	T_Clien Cli_id	_	Vol_pla	ces_prises	Transactions
t1	4	AF 325	258					Début transaction pour u2
t2	4	AF 325	258					Début transaction pour u1
t3	4 5	AF 325 AF 111	258 125					AF111 créé par U1
t4	4 5	AF 325 AF 111	258 125					
t5	4 5	AF 325 AF 111	247 114					Fin transaction commit u1 U2 enlève 11 places partout
t6	4 5	AF 325 AF 111	247 114					
t7	4 5	AF 325 AF 111	247 114	77	4	13	I	
t8	4 5	AF 325 AF 111	247 114	77	4	1	1	Fin transaction commit u2

Voilà comment un certain été des avions d'Air France volaient à vide Avec toutes les places réservées. Bug du service de réservation

- Pour éviter les problèmes :
 - Ecriture impropre (dirty write)
 - Lecture impropre (*dirty read*)
 - Lecture non répétable (non repeatable read)
 - Données fantômes (phantom read)
- Les transactions doivent suivre un certain nombre de propriétés : Propriétés d'une transaction (ACID)



Propriétés ACID

Atomicité

 Unité de cohérence : toutes les mises à jour doivent être effectuées ou aucune

Cohérence

 La transaction doit faire passer la base de données d'un état cohérent à un autre

Isolation

 Les résultats d'une transaction ne sont visibles aux autres transactions qu'une fois la transaction validée

Durabilité

 Les modifications d'une transaction validée ne seront jamais perdue



Avez vous remarqué?

- Séquence = $\langle T_1L(A), T_1E(A), T_2L(A), T_2E(A), T_2L(B), T_2E(B), T_1L(B), T_1E(B) \rangle$
- Où L = Lecture et E = Ecriture
- Les transactions n'affectent que les opérations du LMD et non pas le LDD

```
SQL>COMMIT;

SQL>CREATE TABLE R1 (A NUMBER(2));

SQL>CREATE TABLE R2 (B NUMBER(2));

SQL>INSERT INTO R1 VALUES (1);

SQL>INSERT INTO R2 VALUES (2);

SQL>ROLLBACK;
```

- Combien de tables créées ?
- Que donne la requête SELECT * FROM R1;

Avez vous remarqué?

- Séquence = $<T_1L(A), T_1E(A), T_2L(A), T_2E(A), T_2L(B), T_2E(B), T_1L(B), T_1E(B)>$
- Où L = Lecture et E = Ecriture
- Les transactions n'affectent que les opérations du LMD et non pas le LDD

```
SQL>COMMIT;

SQL>CREATE TABLE R1 (A NUMBER(2));

SQL>CREATE TABLE R2 (B NUMBER(2));

SQL>INSERT INTO R1 VALUES (1);

SQL>INSERT INTO R2 VALUES (2);

SQL>ROLLBACK;
```

- Combien de tables créées ? 2 R1 et R2 (CREATE TABLE : LDD)
- Que donne la requête SELECT * FROM R1; : aucun résultat car le ROLLBACK fait retourner la transaction dans l'état du dernier COMMIT

SERIALISABILITE

HYPOTHESES

- Exécution d'une transaction individuelle est correcte
- Exécution de transactions en série (les unes derrière les autres) est correcte
- PRINCIPE:

SE RAMENER A UNE EXECUTION DES TRANSACTIONS EN SERIE

Concept de **SERIALISABILITE**



Sérialisabilité

- Critère permettant de dire que l'exécution d'un ensemble de transactions (schedule) est correct
 - Une exécution est sérielle si toutes les actions des transactions ne sont pas entremêlées. Elle est donc de la forme : <Tp(1), Tp(2), ...Tp(n)> où p est une permutation de 1, 2, ... n.
- L'exécution des mêmes transactions en série : mêmes valeurs de départ et mêmes valeurs finales

Sérialisabilité : exemple

T ₁	T ₂
LIRE (A)	LIRE (A)
A := A+100	A := A*2
ECRIRE (A)	ECRIRE (A)
LIRE (B)	LIRE (B)
B := B +100	B := B*2
ECRIRE (B)	ECRIRE (B)

Au début : A = B = 25

 T_1 puis T_2 : A = B = 250

 T_2 puis T_1 : A = B = 150



Sérialisabilité: exemple

t0 : A = B = 25t1 t2 t3 t4 t5 t6 t7 t8 t9 t10 t11 t12

<T1 T2 T1 T2> T_1 T_2 LIRE (A) A := A+100ECRIRE (A) LIRE (A) A := A*2ECRIRE (A) LIRE (B) B := B + 100ECRIRE (B) LIRE (B) B := B*2ECRIRE (B) A=250, B=250 **SERIALISABLE**



Sérialisabilité: exemple

t0 : A = B = 25t1 t2 t3 t4 t5 t6 t7 t8 t9 t10 t11 t12

<T1 T2 T2 T1> T_1 T_2 LIRE (A) A := A+100ECRIRE (A) LIRE (A) A := A*2ECRIRE (A) LIRE (B) B := B*2ECRIRE (B) LIRE (B) B := B + 100ECRIRE (B) A=250, B=150 **NON SERIALISABLE**



- Précédences
 - Techniques basées sur la seule sémantique des opérations de lecture / écriture - Recherche des conflits

```
T_i(L) avant T_j(L): pas de conflit T_i(L) avant T_j(E): conflit T_i(E) avant T_j(L): conflit T_i(E) avant T_j(E): conflit
```

- T_i lit O avant T_j écrit => lien T_i précède T_j
- $-T_i$ écrit O avant T_j écrit/lit => lien T_i précède T_j
- Condition de sérialisabilité
 - Le graphe de précédence doit rester sans circuit



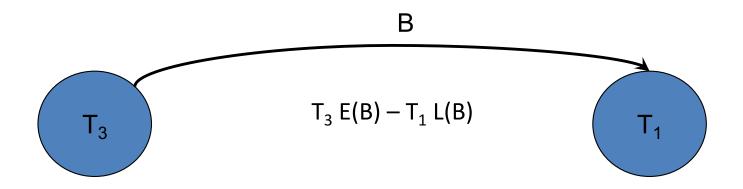
```
T<sub>1</sub>: LIRE (A)
T<sub>2</sub>: LIRE (A)
T<sub>3</sub>: ECRIRE (B)
T<sub>1</sub>: LIRE (B)
T<sub>2</sub>: ECRIRE(A)
T<sub>3</sub>: LIRE (B)
```

$$T_1$$
 L(A) L(B) T_2 L(A) E(A) $E(A)$ $E(B)$



$$T_1$$
 L(A) L(B)
 T_2 L(A) E(A)
 T_3 E(B) L(B)

 T_i lit O avant T_j écrit => T_i précède T_j
 T_i écrit O avant T_j écrit/lit => T_i précède T_j





```
T_1: LIRE (A)

T_2: ECRIRE (A)

T_2: LIRE (B)

T_2: ECRIRE (B)

T_3: LIRE (A)

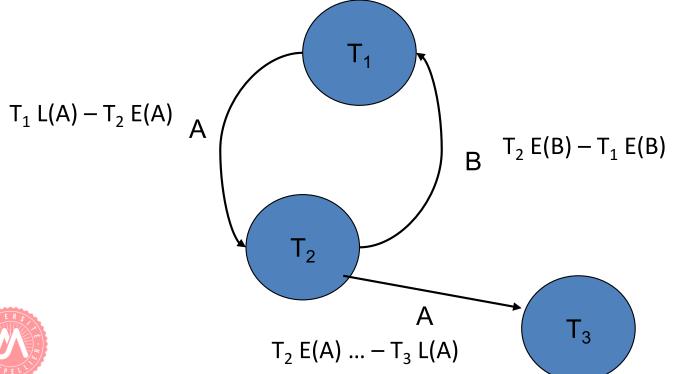
T_1: ECRIRE (B)
```

$$T_1$$
 L(A) E(B) T_2 E(A) L(B) E(B) T_3 L(A)



$$T_1$$
 L(A) E(B) T_2 E(A), L(B), E(B) T_3 L(A)

T_i lit O avant T_i écrit => T_i précède T_i T_i écrit O avant T_i écrit/lit => T_i précède T_i



Présence d'un cycle T₁ dépend $T_2 E(B) - T_1 E(B)$ de celui de T_2 et vice-versa



Ordonnancement des transactions

- Exécution en série : une transaction après l'autre
- Exécutions équivalentes :
 - quelque soit la base de données, la première exécution est identique à l'effet de la seconde exécution
 - Vérification : voir l'ordre des ordres de lecture/écriture conflictuelles

$$T_1 = L(A) A := A*5 E(A) C$$

 $T_2 = L(A) A := A+10 E(A) C$
 T_1T_2 non équivalente à T_2T_1

 Exécution sérialisable : équivalente à une exécution en série de toutes les transactions



35

Exécution sérialisable

Deux exécutions sont équivalentes si :

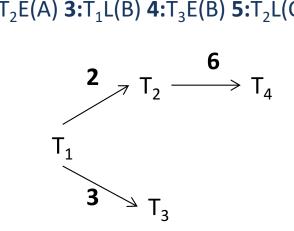
- Les mêmes actions (lecture/écriture) se retrouvent dans les mêmes transactions
- Chaque paire d'actions conflictuelles (lecture/écriture ; écriture/lecture ; écriture/écriture) sont ordonnées de la même façon dans les deux exécutions



Exemple

 $T_1L(A) T_2E(A) T_1L(B) T_3E(B)T_2L(C)T_4E(C)T_1CT_2CT_3CT_4C$

 $1:T_1L(A)$ $2:T_2E(A)$ $3:T_1L(B)$ $4:T_3E(B)$ $5:T_2L(C)$ $6:T_4E(C)T_1CT_2CT_3CT_4C$



• Les opérations sont sérialisables dans l'ordre suivant :

$$T_1 T_2 T_3 T_4$$
 $T_1 T_2 T_4 T_3$
 $T_1 T_3 T_2 T_4$



Approches de gestion de la concurrence

- Objectif : forcer la sérialisabilité, i.e. le graphe de précédence doit être sans cycle
- Problème: les mécanismes utilisant le graphe de sérialisabilité existent mais difficiles à mettre en œuvre car il faut construire tout le graphe en mémoire (coût trop élevé)
- Contrôle continu :
 - vérification au fur et à mesure de la sérialisabilité.
 - Pessimistes car considèrent que les conflits sont fréquents et qu'il faut les traiter le plus tôt possible
- Contrôle par certification :
 - Vérification de la sérialisabilité quand la transaction s'achève
 - Optimistes : considèrent conflits rares et que l'on peut réexécuter les transactions qui posent problèmes.

- L'approche la plus fréquente : contrôle continu
- Verrouillage: on bloque l'accès à une donnée dès qu'elle est lue ou écrite par une transaction (« pose de verrou ») et on libère cet accès quand la transaction termine par COMMIT ou ROLLBACK (« libération du verrou »)
- Permet à une transaction de se réserver l'usage exclusif d'une donnée aussi longtemps que nécessaire
 - Avec un mécanisme de verrouillage, l'accès à une donnée E est encadrée par la paire d'opérateurs :

LOCK E
UNLOCK E



- Le verrouillage définit le type d'opération : lecture, écriture
- Deux modes possibles :
 - Exclusif (eXclusive lock)
 - Si seule la transaction T peut mettre à jour la donnée E. Une opération d'écriture change la valeur d'une donnée. Une opération de lecture par T₁ et de lecture par T₂ ne sont pas permises
 - Partagé (shared lock)
 - Si aucune transaction T ne peut mettre à jour la donnée E mais si des transactions concurrentes peuvent accéder à E. Une opération Lecture ne bloque pas une transaction.



- Ces verrous sont posés de manière automatique par les SGBD
- Attention le verrouillage influence les performances d'une base de données soumise à un très grand nombre de transactions



- Prise en compte du temps
- Le verrou associé à la donnée E peut être de durée variable
 - Courte
 - S'il est maintenu pendant la durée d'une action de T (LIRE, ECRIRE)
 - Longue
 - S'il est maintenu pendant toute la durée de la transaction T



Cas de verrous exclusifs de longue durée

```
Begin T<sub>1</sub>
                            Begin T<sub>2</sub>
LOCK A
                            LOCK C
LIRE(A)
                            LIRE (C)
                           // blocage de T<sub>2</sub>
A := A + 10;
                           // jusqu'à libération
UNLOCK A
End T₁
                             LOCK A
                             LIRE(A)
                             A := A + C
                             DELETE (C)
                             UNLOCK A,C
                             End T<sub>2</sub>
```

La mise à jour de la transaction T_1 n'a pas été perdue. La valeur finale tient compte des modifications de T_1 .



Un autre exemple

T1 T2

LOCK A

LIRE (A)

A:=A+100

ECRIRE (A)

UNLOCK A

LOCK B

LIRE (B)

B := B*5

ECRIRE (B)

UNLOCK B

LOCK B

LIRE(B)

B:= B+10

ECRIRE (B)

UNLOCK B

LOCK A

LIRE A

A := A + 20

ECRIRE (A)

UNLOCK A



Ca marche toujours?

```
T1
                                  T2
LOCK A
LIRE (A) \{A = 10\}
A := A + 100
ECRIRE (A) (A=110)
UNLOCK A
                              LOCK B
                                        {B=20}
                             LIRE (B)
                              B := B*5
                             ECRIRE (B) {B=100}
                              UNLOCK B
LOCK B
           {B=100}
LIRE(B)
B := B + 10
            {B=110}
ECRIRE (B)
UNLOCK B
                              LOCK A
                             LIRE A
                                       {A=110}
                             A := A + 20
                             ECRIRE (A) {A=130}
                             UNLOCK A
```

Les valeurs finales sont A=130 et B=110. Ce n'est pas correct. Une exécution séquentielle aurait produit : A = 130 et B=150. Il manque quelque chose.

Protocole en deux phases

- Chaque transaction doit obtenir un verrou partagé P sur un granule avant de le lire, et un verrou exclusif X sur un granule avant d'écrire
- Tous les verrous émis par une transaction sont libérés à la terminaison
- Si une transaction émet un verrou X sur un granule, aucune transaction ne peut obtenir un verrou (P ou X) sur le même granule
- Une transaction ne peut émettre de verrou dès qu'elle commence à libérer ses verrous.



Gestion des verrous

- Les demandes de verrouillage/déverrouillage sont gérés par le gestionnaire de verrouillage
- Une table spécifique contenant le nombre de transactions avec verrou, le type de verrou et un pointeur vers la file de demande de verrou
- Verrouillage et déverrouillage sont des opérations atomiques
- Une transaction peut demander à ce qu'un verrou partagé devienne un verrou exclusif



Améliorations du verrouillage

- Relâchement des verrous en lecture après opération
 - non garantie de la reproductibilité des lectures
 - + verrous conservés moins longtemps

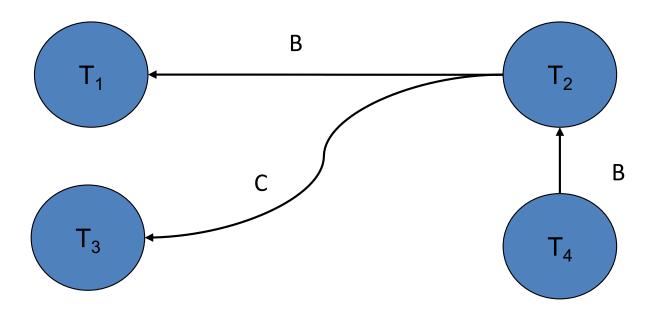
- Accès à la version précédente lors d'une lecture bloquante
 - nécessité de conserver une version (journaux)
 - + une lecture n'est jamais bloquante



Problèmes du Verrouillage

risque d'interblocage

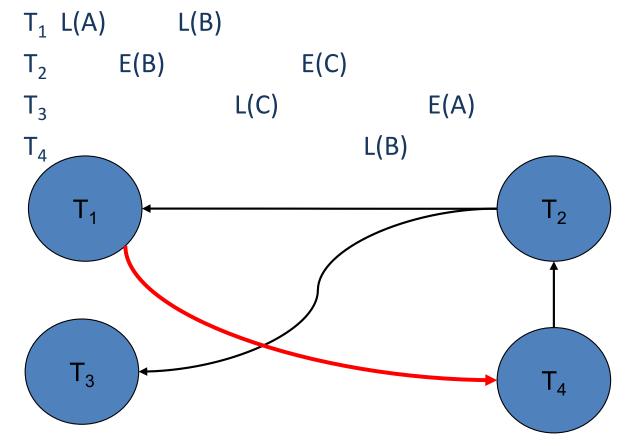
$$T_1$$
 L(A) L(B)
 T_2 E(B) E(C) E(C)
 T_3 L(C) L(B)





Problèmes du Verrouillage

- risque d'interblocage
 - des transactions en attente mutuelles



T1 doit précéder T4 Mais T1 doit être Précédé par T4 (via T2)



Résolution de l'interblocage

Prévention

- définir des critères de priorité de sorte à ce que le problème ne se pose pas
- exemple : priorité aux transactions les plus anciennes

Détection

- gérer le graphe des attentes : les nœuds représentent les transactions, il existe un lien entre T_i et T_j si T_i attend que T_j libère un verrou
- lancer régulièrement un algorithme de détection de circuits
- choisir une victime qui brise le circuit



- Une estampille est une étiquette de temps que l'on associe :
 - Aux données : valeur de l'horloge au moment de la dernière modification (=numéro de version)

Aux transactions : valeur de l'horloge courante

- Les conflits sont détectés au moment où ils se produisent par comparaison des estampilles
- En cas de conflit une transaction est annulée



 Principe: attribuer à chaque élément un readtimestamp (RTS) et un write-timestamp (WTS), donner à chaque nouvelle transaction une estampille (TS)

Si l'action a_i de la transaction T_i est en conflit avec l'action a_j de T_j et si $TS(T_i) < TS(T_j)$ alors a_i doit être réalisé avant a_j . Sinon relancer la transaction



Ce mécanisme garantit que :

- Pas de mise à jour perdue : Une transaction utilisant une valeur obsolète de la donnée pour effectuer une mise à jour est refusée
- Pas de lecture impropre : Si une transaction T lit une entité qui a un numéro de version supérieur à l'estampille de T, cette transaction est refusée
- Contrôle très simple d'ordonnancement des accès conformément à l'ordre de lancement des transactions.
- En cas de désordre : reprendre la transaction ayant créé le désordre



- Cependant
- Les estampilles remplacent en quelque sorte les verrous
- Pas d'attente car reprise des transactions en cas d'accès ne respectant pas l'ordre de lancement des transactions
- => beaucoup trop de reprises
- => difficile à mettre en œuvre dans de grosses applications très concurrentes



Conclusions

- Problème difficile
- Généralement dans les SGBD : verrou
- Les solutions « académiques »: in progress
- Prise en compte de la sémantique des application (opérations commutatives (e.g., ajouts d'informations))
- Quid des transactions longues
 - mise à jour d'objets complexes
- Quid du travail coopératif
 - modèles concurrents plutôt que séquentiels



• Des questions ?

