

PROCEDURES DE REPRISE DANS LES
SYSTEMES DE GESTION DE BASE DE
DONNEES REPARTIES

P. BOUCHET

BIBLIOTHEQUE DU CERIST

IST

638

Procédures de reprise dans les bases de données réparties

c. 638

PROCEDURES DE REPRISE DANS LES SYSTEMES DE GESTION
DE BASE DE DONNEES REPARTIES *

P. BOUCHET **

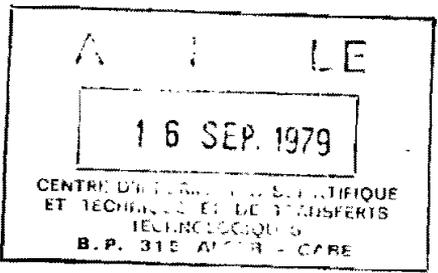


Recovery Techniques for Distributed Database Management System

Summary. In this paper, two recovery techniques used to restore the integrity of the distributed database management system operating in the presence of intermittent failures and with a checkpointing and roll-back recovery scheme are analysed and compared. A mathematical model of the distributed database management system is proposed to evaluate its disponibility at the updating and interrogating operations in the case of one, two, three and four distributed copies. Numerical illustrations are presented and it is shown that the recovery procedure based on the transfert of the valid copies accross the network is better in some given conditions of processing of the distributed database management system.

* Cette recherche a été effectuée dans le cadre de la convention de recherche IRIA-SIRIUS n° 77077

** Université Paris-Nord - I.U.T. de Villetaneuse - Département Informatique - 93400 Villetaneuse
Equipe de Recherche Associée au C.N.R.S. n° 452 "AL KHOWARIZMI"



1 - INTRODUCTION

L'automatisation des systèmes d'information impliquant des usagers largement répartis géographiquement ne se conçoit plus, depuis quelques années, de façon uniquement centralisée. Le développement des facilités de télécommunications, l'abaissement des coûts du matériel informatique, permet dans certaines conditions de distribuer les ressources en information, logiciel et matériel et de les utiliser avec un meilleur rapport coût-efficacité. C'est dans cette perspective que se situent les travaux de Chu [5,6] , Casey [3,4] Withney [16] , Levin, Morgan [12] sur l'implantation optimale des copies d'une base de données en différents noeuds d'un réseau d'ordinateurs et de Mahmoud, Riordon [13] , sur l'implantation optimale des copies et le choix des capacités de transmission. Les modèles mathématiques que ces auteurs ont étudiés permettent de déterminer le nombre de copies de la base de données et les noeuds du réseau auxquels elles sont affectées en tenant compte des paramètres caractéristiques du système d'information et du réseau, ainsi que d'un certain nombre de contraintes.

La fiabilité des systèmes informatique de gestion des copies est un paramètre important de la répartition qui n'a pas été pris en compte dans les modèles analysés dans [3, 4, 5, 6, 12, 16] . Le modèle analysé dans [13] suppose une fiabilité parfaite des systèmes informatiques et une fiabilité limitée des noeuds et des voies de communication qui intervient dans l'expression de la contrainte de disponibilité du système informatique.

L'influence de la multiplication des copies sur la fiabilité du système de gestion de la base de données est due au fait qu'elle accroît la probabilité de dégradation d'une copie ou d'immobilisation du système informatique permettant d'en assurer la gestion. La reprise de l'activité du système ou la remise en état de la copie dégradée passe par l'exécution éventuelle d'un processus de réparation et dans tous les cas par l'exécution de procédures de restauration des informations

BIBLIOTHEQUE DU CERIST

en mémoire centrale et auxiliaire et de reprise des traitements, que nous appellerons dans la suite plus brièvement procédures de reprise. Dans certaines conditions d'exploitation de la base de données, pour assurer la cohérence des copies, le traitement des transactions de mise-à-jour est interrompu pendant la durée d'exécution de la procédure de reprise. D'où l'influence de la multiplication des copies et de la nature de la procédure de reprise sur la disponibilité du système informatique pour traiter les transactions de mise-à-jour. Cette influence peut se traduire dans certains cas d'exploitation de la base de données par une diminution de la disponibilité du système au traitement de ces transactions.

Dans ce travail, deux procédures de reprise dans un système de gestion d'une base de données à copies réparties, sont analysées et comparées. Les systèmes informatiques de gestion des copies réparties dans le réseau d'ordinateurs, supposé parfaitement fiable, sont soumis à des défaillances intermittentes de deux types, nécessitant d'exécuter, selon que l'un ou l'autre type de défaillance se présente, une reprise "à froid" ou "à chaud". Dans l'une des deux procédures de reprise, la reprise à froid fait intervenir le transfert d'une des copies valides pour restaurer l'état des copies dégradées. Un modèle mathématique du système informatique réparti est élaboré pour analyser sa disponibilité au traitement des transactions de mise-à-jour d'une part et d'interrogation d'autre part. Dans certaines conditions d'exploitation du système informatique réparti, une meilleure disponibilité apparaît lorsque la procédure de reprise basée sur le transfert des copies est utilisée.

Cette étude se situe dans le prolongement des travaux qui ont été effectués par Gelenbe [2, 8, 9, 10, 11], sur l'analyse et l'optimisation des procédures de reprise dans un système de gestion d'une base de données centralisées et sur les travaux de Bouchet, Gelenbe, Tucci [1] sur l'étude du rapport coût -efficacité de différentes architectures permettant de restaurer l'intégrité d'une base de données et la reprise des traitements.