SOME STUDIES ON CONCURRENCY CONTROL AND RECOVERY IN DISTRIBUTED DATABASE SYSTEMS

By

RADHA RAMAN SINHA

Thesis submitted in fulfilment
of the requirements of the degree of
DOCTOR OF PHILOSOPHY

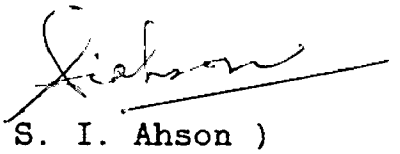
Department of Electrical Engineering
INDIAN INSTITUTE OF TECHNOLOGY, DELHI

1988

IN FOND MEMORY OF
MY FATHER WHO ALWAYS
ADVOCATED HONESTY AND
DETERMINATION FOR WORK

C E R T I F I C A T E

This is to certify that the thesis entitled "SOME STUDIES ON CONCURRENCY CONTROL AND RECOVERY IN DISTRIBUTED DATABASE SYSTEMS" being submitted by Radha Raman Sinha for the award of the degree of DOCTOR OF PHILOSOPHY to the Indian Institute of Technology, Delhi is a record of bonafide research work he has carried out under my supervision. The results contained in this thesis have not been submitted to any other University or Institute for award of a degree or diploma.



(S. I. Ahson)

Professor

Department of Electrical Engineering

Indian Institute of Technology, Delhi

NEW DELHI - 11ØØ16

A C K N O W L E D G E M E N T S

I express my deep sense of gratitude and appreciation to my supervisor Professor S.I.Ahson for his invaluable guidance, useful discussions and critical suggestions during the course of this research work.

I am thankful to Dr. Mukul K Sinha, Director, Expert Software Consultants Private Limited for his invaluable suggestions and help in shaping up this thesis.

Lastly thanks are due to Shri Arvind Kanaujia for typing the manuscript.

I.I.T. DELHI
JULY, 1988

RADHA RAMAN SINHA

ABSTRACT

A two-step commit algorithm which assures transaction update atomicity in a distributed database is modelled using Bipolar Synchronization schemes. The modelling helps in determining the numerous crash possibilities and in specifying recovery procedures and associated protocols. The use of unalterable logs is required and modelled accordingly.

A new scheme for maintaining multicopy data has been presented. Replicas of a data item, stored at various sites, are structured into one or more disjoint groups named representation trees. The scheme is more general and by restructuring the representation trees it is shown that : (1) the primary copy approach (2) the majority copies approach (3) the true copy token approach and (4) the quorum approach result as its special cases.

Some number of votes are assigned to each node_replica of the representation trees. Root_Replica represents the whole representation tree, and effectively has the cumulative sum of votes assigned to each node_replica (including the root) contained in the tree. A child has no right to vote till the Parent replica is active. A transaction scans only the root_replicas for collecting votes, resulting in less number of messages than that required in the basic quorum scheme.
Since the scheme presents a generalized and integrated view of multi_copy data, it provides a tool for comparing various schemes. By organizing the structure of the representation trees, by assigning different votes to replicas, and by tuning a read and write quorum the database administrator can control the reliability and performance characteristics most suited to the application. This scheme offers a performance improvement without degrading the resiliency in case of failures. It also provides increased concurrency as updates of non root_replicas are done in parallel, and are not involved in actual transactions.

CONTENTS PAGE NO