TRANSPARENT REPLICATION FOR EFFICIENT DISTRIBUTED TRANSACTIONS

By
REKHA SINGHAL
Department of Computer Science & Engineering

Submitted
in fulfillment of the requirements of the degree of
DOCTOR OF PHILOSOPHY

to the

INDIAN INSTITUTE OF TECHNOLOGY, DELHI

FEBRUARY 1998

Certificate

This is to certify that the thesis titled "Transparent Replication for Efficient Distributed Transactions" being submitted by Rekha Singhal to the Indian Institute of Technology, Delhi, for the award of the Degree of Doctor of Philosophy, is a record of bonafide research carried out by her under my supervision. The results obtained in this thesis have not been submitted to any other University or Institute for the award of any other degree or diploma.

Gautam Shroff

Acknowledgments

I express my sincere thanks to Gautam Shroff, my thesis advisor, for all his encouragement and constant support over the past four years. The genesis of this thesis is the course on 'Parallel Computing' taken by Gautam shroff at the Institute. He has a good insight power which initiated me to relate the popular message ordering protocols with transaction management protocols. His methodical approach to problems in search for efficient, simple and elegant solutions with their potential applicability in practice has left a deep impact on me. I am thankful to him for being so generous with his time and for being patient with me.

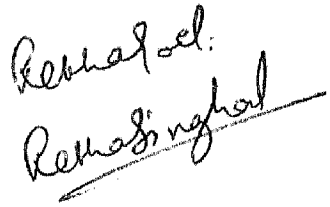
I would like to express my thanks and gratitude to Prof S.N. Maheshwari without whom this thesis would not have even started. He motivated me to pursue research as a Phd student of the Institute. His course on 'Concurrency Control' and frequent discussions with him clarified my basics on performance issues which contributed to the thesis.

I am thankful to the research organisation, Tata Research Development and Design Center (TRDDC), for its consistent support to provide platforms for applying this thesis ideas to industrial products, as well as sponsoring me to attend international conferences held in India.

I would like to thank Prof B.N.Jain for his generosity in letting me to use the department as well as ERNET lab resources to carry out the research. I would like to acknowledge the co-operation and timely help extended to me by the inmates of the ERNET lab and other staff in the department, especially Mani, Nandini, Raja Kumar and Jaswant. I am very fortunate to have been here with great set of people who were always ready to share their ideas with me. Thanks to R.P. Rustagi for the rigorous discussions I had with him on concurrency control techniques, R.N. Moorthy and Naveen for being helpful for any discussion at any time.

Many of those who do not have anything directly to do with this thesis or my research, but who have made my stay here lively with numerous interesting discussions. These include Lalita, Deepa and others at the hostel, and Neena, Neelima, Jadon, Sumantra and Shoma at the department. I thank all of them for their companionship.

The one quarter that I have always received constant support and encouragement is my family: my parents and my brothers, Rakesh and Vijay. This thesis is dedicated to them.



Abstract

This thesis proposes the parallel replication system (PRS), a transaction processing middleware designed specifically for object oriented applications. The goal is to enable an object oriented transaction processing application to be deployed in parallel at multiple physical locations, each potentially supporting different sets of data objects, and yet having transparent access to all application data objects. PRS can work with standard OLTP monitors; alternatively it can provide some of these services independently. PRS can support transaction processing in world wide web as well as in mobile computing environments.

PRS exploits replication as well as weak consistency to provide efficient transaction processing while preserving transaction transparency, distribution transparency and object storage transparency; PRS utilizes only application specific object level interfaces to access the persistent store. PRS is an object aware system which exploits application semantics to efficiently achieve any required level of data transparency in a wide area distributed scenario.

PRS uses application semantics to classify *modify* operations into *write*, *update* and *commutative update* operations, and read operations into *browse_read* and *critical_reads*. This categorization helps in defining an intermediate level of transaction isolation, Consistent Committed Read (degree 2.5), between Repeatable Read (degree 3) and Read Committed (degree 2). PRS uses this classification to provide transaction transparency by guaranteeing at least degree 2 isolation (i.e., degree 2, degree 2.5 or degree 3 as required) to users while ensuring correctness.

PRS employs a selective mix of eager and lazy propagation of data modifications to ensure mutual consistency: Efficient message ordering protocols are used to ensure mutually exclusive access to all replicas of an object whenever required. Due to lazy propagation, concurrent access to replicas of the same object may introduce conflicts potentially violating transaction transparency; PRS uses time vectors to correctly resolve such conflicts.

PRS supports all possible distribution of data objects including full distribution, full replication, partial replication and large objects. To summarize, PRS is a generic approach for efficient and transparent access to any distribution of data objects accessed by application replicated across wide area network.

Contents