

SYNCHRONIZATION AND CRASH RECOVERY IN
DISTRIBUTED SYSTEMS

By
SUBHASH BHALLA
CENTRE FOR COMPUTER SCIENCE AND ENGINEERING

THESIS SUBMITTED IN FULFILMENT OF THE
REQUIREMENTS OF THE DEGREE OF
DOCTOR OF PHILOSOPHY

TO THE
INDIAN INSTITUTE OF TECHNOLOGY, DELHI
FEBURARY, 1983

DEDICATED TO MY PARENTS

CERTIFICATE

Certified that the thesis entitled ' SYNCHRONIZATION AND
CRASH RECOVERY IN DISTRIBUTED SYSTEMS '

being submitted by Mr Subhash Bhalla, is a record of work
carried out by him under my supervision. The matter
embodied in this thesis has not been submitted for the
award of any other degree.



(Dr PG Reddy)
Professor, Computer
Centre, IIT, New Delhi

ACKNOWLEDGEMENTS

I am grateful to my supervisor Professor P.G. Reddy for his personal interest. Without his constructive criticism, and constant encouragement this work would never have been a success.

My thanks are due to Professor M.K. Jain for his help and encouragement. I would like to acknowledge the contribution of faculty members and other staff of the computer centre, for their cooperation and help.

I would like to express my gratitude to Dr. B.E.Prasad, Dr. S.P. Rana, Arun Aggarwal, Nirmal Jain, Rekha Rani, Niraj Sharma, Amitab Sinha, Dhruv Nath, and all others who have helped me in several ways.

Finally I would like to thank Mr. D.R. Joshi for his neat typing.

(SUEHASH BHALLA)

ABSTRACT

In this dissertation, a new approach to the study of concurrency control techniques used in distributed databases has been adopted. A representative concurrency control technique has been chosen for the above purpose. A mathematical model of the above scheme, based on a stochastic process has been developed. Based on the inferences drawn from the above model a new concurrency control scheme has been proposed. The proposed scheme relies upon 'detection of non-serializable schedules at runtime, and correction by backup'.

In the later part, the case of network partitions and site crashes has been studied and an optimum corrective strategy has been proposed. Also, studied in the thesis is the case of priority response requirement of certain important applications.

A qualitative study of the costs incurred by the proposed scheme has been presented in order to have an idea of range of applicability of this scheme.

It has been pointed out that the proposed scheme has various advantages associated with it, in addition to simplicity and ease of implementation. In addition to offering flexibility the scheme is a unified approach where the same pool of information is used for various purposes, like concurrency control, recovery from site crashes and network partitions and for maintenance of accountbook-like logs.

C O N T E N T S

CHAPTER		PAGE
	ABSTRACT	
	NOTATIONS	
I	INTRODUCTION	1
	1.1 Introduction	1
	1.2 CC Problem	2
	1.3 Survey of Previous Work	8
	1.4 Thesis Objective and Outline	11
II	A MODEL OF CONCURRENCY CONTROL	16
	2.1 Introduction	16
	2.2 The Algorithm	18
	2.3 Model of Transaction Behaviour	22
	2.4 Discussion	27
	2.5 Application of the Model	28
	2.6 Summary	39
III	THE CONCURRENCY CONTROL SCHEME	40
	3.1 Introduction	40
	3.2 Review of Past Work	43
	3.3 Centralized vs. Distributed Control	51
	3.3.1 The Centralized Certifier	53
	3.3.2 Discussion	57
	3.3.3 Crash-Recovery	61
	3.3.4 Advantages of Distributed Systems	64
	3.3.5 Priority Response	67
	3.4 Proof of Consistency	70
	3.5 Performance Comparison	71
	3.6 Summary	73
IV	NETWORK PARTITIONING AND PRIORITY RESPONSE	75
	4.1 Introduction	75
	4.2 Previous Work on Partitioning	77
	4.2.1 Back-out of Transaction on Recovery	

4.3	Detection of Mutual Inconsistency in Case of Network Partitioning	82
4.4	Priority Response for Important Applications	87
4.4.1	Avoiding Delay due to Concurrency in a Restricted Set-up	89
4.5	Summary	91
V	PERFORMANCE OF THE CONCURRENCY CONTROL METHOD	92
5.1	Introduction	92
5.2	Cost Considerations	93
5.2.1	Communication Cost	94
5.2.2	Local Processing Overhead	99
5.2.3	Transaction Restarts	100
5.2.4	Transaction Blocking	102
5.3	Centralized vs. Decentralized Control	103
5.4	Transaction Processing, Integrity and Reliability	106
5.4.1	Update in Place	107
5.4.2	A Simple and Unified Approach	110
5.5	Summary	113
VI	CONCLUSIONS	115
	APPENDIX A	121
	APPENDIX B	124
	APPENDIX C	128
	REFERENCES	129