

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

DISTRIBUTION DESIGN  
IN  
OBJECT ORIENTED DATABASES

A thesis presented in partial fulfilment of the requirements for the degree of

MASTER OF INFORMATION SCIENCE  
IN  
INFORMATION SYSTEMS

at Massey University, Palmerston North,  
New Zealand

Hui Ma  
2003

# Abstract

The advanced development of object oriented database systems has attracted much research. However, very few of them contribute to the distribution design of object oriented databases. The main tasks of distribution design are fragmenting the database schema and allocating the fragments to different sites of a network. The aim of *fragmentation* and *allocation* is to improve the performance and increase the availability of a database system. Even though much research has been done on distributed databases, the research almost always refers to the relational data model (RDM). Very few efforts provide distribution design techniques for distributed object oriented databases.

The aim of this work is to generalise distribution design techniques from relational databases for object oriented databases. First, the characteristics of distributed databases in general and the techniques used for fragmentation and allocation for the RDM are reviewed. Then, fragmentation operations for a rather generic object oriented data model (OODM) are developed. As with the RDM, these operations include horizontal and vertical fragmentation. A third operation named *splitting* is also introduced for OODM. Finally, normal predicates are introduced for OODM. A heuristic procedure for horizontal fragmenting of OODBs is also presented. The adaption of horizontal fragmentation techniques for relational databases to object oriented databases is the main result of this work.

# Acknowledgements

I would like to thank Professor Klaus-Dieter Schewe, my supervisor, for his patience, guidance, suggestions and constant support during this research. I am also thankful to Markus Kirchberg for his encouragement and guidance through the early stage of chaos and confusion. A special thanks goes to Madre Chrystal for her kindly devoting valuable time to proof read my draft.

The *Massey Masterate Scholarship*, which was awarded to me for the period February 2002 – February 2003 for graduate studies, was crucial to the successful completion of this project. Finally, I am grateful to my husband and my parents for their patience and *love*. Without them this work would never have come into existence (literally).

Hui Ma  
March 31, 2003

# Table of Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	What is a Distributed Database? . . . . .	2
1.2	Why Distribution? . . . . .	2
1.3	Distribution Design: Fragmentation and Allocation . . . . .	4
1.4	Objective: Fragmentation in Object Oriented Databases . . . . .	5
1.5	The Outline of the Thesis . . . . .	6
<b>2</b>	<b>Distributed Databases</b>	<b>7</b>
2.1	Characteristics . . . . .	7
2.1.1	Data Independence . . . . .	8
2.1.2	Network Transparency (Distribution Transparency) . . . . .	9
2.1.3	Replication Transparency . . . . .	9
2.1.4	Fragmentation Transparency . . . . .	10
2.2	Key Concepts . . . . .	10
2.2.1	Heterogeneity . . . . .	10
2.2.2	Autonomy . . . . .	11
2.2.3	Distribution . . . . .	11
2.2.4	Classification of Distributed DBMS . . . . .	12
2.3	A Framework for Distributed Databases . . . . .	13
2.3.1	The Objective of the Design of Data Distribution . . . . .	14
2.3.2	The Reasons for Fragmentation . . . . .	15
2.3.3	Alternative Design Strategies . . . . .	16

<b>3</b>	<b>Distribution Design for Relational Databases</b>	<b>19</b>
3.1	The Relational Data Model . . . . .	20
3.2	Characteristics . . . . .	21
3.3	Horizontal Fragmentation . . . . .	22
3.4	Vertical Fragmentation . . . . .	26
3.5	Mixed Fragmentation . . . . .	29
3.6	Allocation . . . . .	29
3.7	Related Work . . . . .	30
3.7.1	Horizontal Fragmentation . . . . .	30
3.7.2	Vertical Fragmentation . . . . .	35
<b>4</b>	<b>Object Oriented Databases</b>	<b>37</b>
4.1	Fundamentals of the OODM . . . . .	38
4.1.1	Type Definitions . . . . .	38
4.1.2	Class Definitions . . . . .	42
4.1.3	Schema Definition . . . . .	42
4.2	An Example of Object Oriented Database Schema . . . . .	43
4.3	Queries . . . . .	51
4.3.1	Path Expressions . . . . .	51
4.3.2	Queries . . . . .	54
<b>5</b>	<b>Fragmentation Operations in Object Oriented Databases</b>	<b>61</b>
5.1	Split Fragmentation . . . . .	62
5.2	Horizontal Fragmentation . . . . .	64
5.2.1	Horizontal Fragmentation on Class Level . . . . .	64
5.2.2	Horizontal Fragmentation on Type Level . . . . .	66
5.3	Vertical Fragmentation for Object Oriented Databases . . . . .	68
5.3.1	Vertical Fragmentation on Class Level . . . . .	68
5.3.2	Vertical Fragmentation on Type Level . . . . .	70
5.4	Fragmentation Strategies . . . . .	71
5.5	Related Work . . . . .	72

<b>6</b>	<b>A Method for Horizontal Fragmentation in Object Oriented Databases</b>	<b>75</b>
6.1	Simple Predicates . . . . .	75
6.2	Normal Predicates . . . . .	78
6.3	The Heuristic Fragmentation Process . . . . .	80
6.4	A Cost Model . . . . .	82
6.4.1	Query-Trees . . . . .	82
6.4.2	Calculation of Size for Classes . . . . .	83
6.4.3	Calculation of Size for Fragments or Intermediate Nodes . . . . .	88
6.4.4	Allocate Intermediate Nodes to Sites . . . . .	89
6.4.5	Calculation of Query Costs . . . . .	90
6.5	A Heuristic Procedure for Horizontal Fragmentation . . . . .	92
6.6	Summary . . . . .	97
<b>7</b>	<b>Conclusion and Possible Extension</b>	<b>99</b>
7.1	Summary . . . . .	99
7.2	Future Work . . . . .	100
	<b>Bibliography</b>	<b>103</b>