

104-7114

2

**AN XML-BASED MIDDLEWARE INTERFACE INTEGRATING
WITH HETEROGENEOUS RELATIONAL DATABASES**

KUM LEE CHING (WGCY 01011)

**DISSERTATION SUBMITTED IN FULFILLMENT
OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SOFTWARE ENGINEERING**

**FACULTY OF COMPUTER SCIENCE AND
INFORMATION TECHNOLOGY
UNIVERSITY OF MALAYA
KUALA LUMPUR**

DECEMBER 2003



ABSTRACT

The increasing growth of XML as the undisputable *lingua franca* standard for data interchange and content management has changed the prospect of how to store XML data effectively into relational databases. Most organisations still rely heavily on existing relational database management systems (RDBMS) to store and manage daily business transactions. However, due to the underlying intrinsic difference principles between XML structures and relational databases structures, there is a growing need for tools that can perform data mappings and transfer of XML data to relational databases. While most existing XML middleware provide highly valuable XML functionalities, they still remain as static pieces of APIs or classes, which are programmatic-oriented approach that requires some programming skills. This research is designed to provide a generic yet extensive and graphical user-friendly XML-based middleware interface framework, which is known as JXDB, that allows a user to use XML for dealing with semi-structured data for creating, accessing, updating and transferring of XML data to multiple relational databases that store structured data and vice versa. Additionally, it aims to incorporate XQuery language to manipulate XML data from multiple XML sources. JXDB prototype can serve as a basis for future work to extend its functionalities, such as to provide the ability to integrate XML with other types of databases besides relational databases and web-enabled core services as web services so that users can access JXDB across the Internet. Hence, it will assist XML to fulfill its potential and gain wide acceptance in data interchange and integration with heterogeneous relational databases.

DECLARATION

I certify that this dissertation submitted for the degree of Master of Software Engineering is the result of my own research, except where otherwise acknowledged, and that this dissertation or any part of it has not been submitted in support of an application for higher degree or qualification to any other university or institution.

Kum Lee Ching.

Kum Lee Ching (WGCY01011)

ACKNOWLEDGEMENTS

First of all, I would like to dedicate this dissertation to my beloved parents, Dato' Kum Boo and Datin Mah Gaik Hong for giving me love, moral support and strength to enable me to complete this project. Secondly, I would like to thank all the professors and lecturers for their advice and guidance during my years of Master Degree study in the Faculty of Computer Science and Information Technology, University of Malaya.

I would specially wish to acknowledge the invaluable guidance and constructive comments received from my supervisor, Associate Professor Dr. Lee Sai Peck from whom this project has greatly benefited. I hope that I have satisfactorily accommodated all her suggestions and concerns. I would like to extend my appreciation to my bosses from Heitech Padu Berhad, Encik Hasnan Hashim and Encik Khairol Amin for their support. Last but not least, I would also like to express my sincere gratitude to Christina, my sister, and Eric who has continuously encouraged me to pursue and complete my Master Degree course.

TABLE OF CONTENTS

CHAPTER 1	INTRODUCTION.....	1
1.1	Problem Statement.....	2
1.2	Research Objectives.....	5
1.3	Scope of Research	7
1.4	Significance of Research.....	9
1.5	Dissertation Outline.....	10
CHAPTER 2	LITERATURE REVIEW	12
2.1	XML and Databases	12
2.1.1	XML.....	13
2.1.2	Relational Database Management Systems (RDBMS)	14
2.2	XML-Relational Mapping Technologies.....	15
2.2.1	XML-Relational Mapping Features	16
2.2.2	XML-Relational Mapping Requirements	17
2.2.2.1	Mapping Definition	18
2.2.2.2	Mapping Tools	20
2.2.3	Integration of Query, Selection and Transformation XML Processors of XML Views over Relational Databases	21
2.2.4	XML-Relational Mapping Advantages	22
2.3	XML versus Relational Data Structure	24
2.3.1	XML Data Structure.....	24
2.3.1.1	Document Type Definition (DTD)	25
2.3.1.2	XML Schema	25
2.3.2	Relational Database Structure	26

2.4	Java Technology	27
2.5	Conclusion	28
CHAPTER 3 RESEARCH METHODOLOGY		29
3.1	Project Framework	29
3.2	Feasibility Review	30
3.2.1	Review of XML Middleware	31
3.2.1.1	Allora	32
3.2.1.2	XML-DBMS.....	33
3.2.1.3	Comparison of Existing XML Middleware and the Proposed System .	34
3.2.2	Review of XML-Enabled Databases	38
3.2.2.1	Oracle	39
3.2.2.2	Microsoft SQL	40
3.2.2.3	Comparison of XML-Enabled Database and the Proposed System.....	43
3.3	System Architectural View: 4+1	44
3.3.1	Layered System Architectural View.....	46
3.3.2	System Architectural Design	47
3.3.3	Object-Oriented Software Development Life Cycle	50
3.3.4	Object-Oriented Analysis: Use Case Driven.....	51
3.3.5	Object-Oriented Design	54
3.4	System Implementation	55
3.5	System Testing and Evaluation.....	57
3.6	Selection and Familiarisation of Development Tools	60
3.7	Conclusion	62

CHAPTER 4 OBJECT-ORIENTED ANALYSIS AND DESIGN..... 63

4.1 Object-oriented Analysis 63

4.1.1 Analysis Model..... 66

4.1.2 Load XML Documents Use Case Scenario 67

4.1.3 Insert XML Data Use Case Scenario 68

4.1.4 Update XML Data Use Case Scenario..... 70

4.1.5 Delete XML Data Use Case Scenario 71

4.1.6 Read XML Data Use Case Scenario 72

4.2 JXDB System Architecture 73

4.3 Object-oriented Design 77

4.3.1 Design Model 77

4.3.2 Sequence Diagrams 78

4.3.2.1 Load XML Documents Sequence Diagram 79

4.3.2.2 Insert XML Data Sequence Diagram..... 80

4.3.2.3 Update XML Data Sequence Diagram..... 82

4.3.2.4 Delete XML Data Sequence Diagram..... 84

4.3.2.5 Read XML Data Sequence Diagram 85

4.4 Class Diagram 87

4.5 Summary..... 90

CHAPTER 5 JXDB IMPLEMENTATION AND TESTING 91

5.1 System Development Environment 91

5.1.1 Software Development Tools..... 92

5.1.2 Development Environment..... 92

5.2 JXDB XML-based Interfaces Application System 93

5.3 JXDB Component Systems 96

5.3.1 Implementation of Database Connection Component Systems 97

5.3.2 Implementation of XQuery Wizard Component Systems..... 99

5.3.2.1 FLWOR Expressions 103

5.3.3 Implementation of Data Transfer Component Systems 106

5.4 JXDB Testing Strategies..... 110

5.5 JXDB Evaluation 113

5.5.1 JXDB 114

5.5.2 XML-DBMS..... 116

5.5.3 Comparison between JXDB and XML-DBMS..... 119

5.6 Summary..... 120

CHAPTER 6 FUTURE WORK AND CONCLUSION 122

6.1 Significance of Research 122

6.2 Critical Remarks..... 125

6.3 Future Work..... 128

6.4 Conclusion 131

APPENDIX A 135

APPENDIX B 136

APPENDIX C 141

APPENDIX D 147

REFERENCES..... 150

SUPPLEMENT 154

LIST OF FIGURES

Figure 2-1: Generic Overview of Mapping and Transformation of XML Data.....	16
Figure 3-1: Generic Graphical View of Middleware Architecture.....	32
Figure 3-2: Generic Graphical View of XML-Enabled Database Architecture.....	38
Figure 3-3: The 4+1 View of Architecture (Quatrani, 2000).....	45
Figure 3-4: A Typical Layered System Architectural View	47
Figure 3-5: System Architecture Design.....	49
Figure 3-6: Object-oriented Software Development Life Cycle.....	50
Figure 3-7: XML-based Middleware Use Case Packages.....	53
Figure 3-8: Abstract Class Diagram.....	55
Figure 3-9: Schematic of evolutionary prototyping (Smith, 1991)	56
Figure 3-10: Software Testing Stages (Burch, 1992).....	59
Figure 4-1: Use Case Model.....	65
Figure 4-2: Load XML Documents Collaboration Diagram.....	68
Figure 4-3: Insert XML Data Collaboration Diagram	69
Figure 4-4: Update XML Data Collaboration Diagram	71
Figure 4-5: Delete XML Data Collaboration Diagram.....	72
Figure 4-6: Read XML Data Collaboration Diagram.....	73
Figure 4-7: Layered Architecture of the JXDB XML-based System	76
Figure 4-8: Load XML Documents Sequence Diagram.....	80
Figure 4-9: Insert XML Data Sequence Diagram	82
Figure 4-10: Update XML Data Sequence Diagram.....	83
Figure 4-11: Delete XML Data Sequence Diagram.....	85
Figure 4-12: Read XML Data Sequence Diagram	86
Figure 4-13: Class Diagram.....	89
Figure 5-1: JXDB Main Screen	94
Figure 5-2: Database Connection Wizard Dialog Box.....	98
Figure 5-3: Execute SQL Statement Screen.....	99
Figure 5-4: JXDB XQuery Wizard Dialog	101
Figure 5-5: Building XQuery Expressions Using XQuery Wizard Screen	102
Figure 5-6: Execute XQuery Expressions Screen.....	103
Figure 5-7: JXDB's Data Transfer Strategy	109
Figure 5-8: JXDB Middleware Approach	116
Figure 5-9: XML-DBMS Middleware Approach	118
Figure C-1: Displaying Data on Data Grid Control.....	141
Figure C-2: Displaying Data in Tree Format	142
Figure C-3: Connecting to Oracle 9i Database Dialog.....	143
Figure C-4: XQuery Wizard.....	146

LIST OF TABLES

Table 3-1: Project Framework Objectives	30
Table 3-2: Comparison between XML Middleware and Proposed System	37
Table 3-3: Comparison between XML-Enabled Database and Proposed System	43
Table 3-4: Selection of Development Tools	61
Table 5-1: Comparison between JXDB and XML-DBMS	119
Table A-1: JXDB Packages (xdb)	135

LIST OF ACRONYMS AND ABBREVIATIONS

API: Application Programming Interface

B2B: Business to Business

B2C: Business to Customer

DBMS: DataBase Management System

DOM: Document Object Model

DTD: Document Type Definition

EAI: Enterprise Application Integration

GUI: Graphical User Interface

HTTP: HyperText Transfer Protocol

HTML: HyperText Markup Language

IT: Information Technology

IDE: Integrated Development Environment

JAXB: Java Architecture for XML Binding

JAXP: Java API for XML Processing

JDBC: Java DataBase Connectivity

JDOM: Java DOM

JDK: Java Development Kits

ODBC: Open DataBase Connectivity

OLE DB: OLE DataBase

RDBMS: Relational DataBase Management System

RUP: Rational Unified Process

SDLC: Software Development Life Cycle

SGML: Standard Generalized Markup Language

SOAP: Simple Object Access Protocol

SQL: Structured Query Language

TCP/IP: Transmission Control Protocol / Internet Protocol

UML: Unified Modeling Language

URI: Uniform Resource Identifier

W3C: World Wide Web Consortium

XML: EXtensible Markup Language

XPath: XML Path Language

XSLT: EXtensible Stylesheet Language Transformation

XQuery: XML Query Language

XSU: XML-SQL Utility (XML API introduced by Oracle)