



distributed data warehousing using web technology

how to build a more
cost-effective and
flexible warehouse

R. A. MOELLER

Contents

Preface	xi
Acknowledgments	xiii
1. Introduction	1
A Brief History of Data Warehousing	2
What Is a Data Warehouse?	5
Pressure to Distribute Data	8
Data Mart Issues	16
The Web and Other Emerging Technologies	22
In Brief	27
2. Data Warehouse Architecture	29
Role of a Data Warehouse	32
Data Warehouse Definition	34
Traditional Data Warehouse Architecture	45
Challenges of a Centralized Data Warehouse	58
Distributed Data Warehouse Architecture	63
Challenges of a Distributed Data Warehouse	75
What to Choose	80
3. Data Placement in Distributed Warehouses—Distribution and Transformation	83
The Data Placement Problem	83
A Practical Approach to the Placement Problem	88
Integration of Diverse Database Systems	98

The Global Data Model	102
What to Remember	109
4. Concurrency Control in Distributed Warehouses	115
Transactions	117
Interference between Concurrent Transactions	122
Schedules and Serialization	126
Concurrency Control Techniques	132
Special Case Concurrency Control	147
Data Warehousing and Concurrency Control	153
5. The Web and the Data Warehouse	159
Emerging Trends	161
Intranet Basics	167
The Advantages of Intranets and Data Warehousing	174
The Challenges of Intranets and Data Warehousing	176
Common Gateway Interfaces	180
The Future of Web-Enabled Data Warehousing	184
6. Data Marts	187
Data Mart Solutions	188
What Is a Data Mart?	189
Implementing a Data Mart	201
Data Mart Architecture	208
The Future of Data Marts	216
7. Object Technologies	219
Object-Oriented Technologies	220
Distributed Object Computing	226
ORB Architecture	230
Standards	232
Data Warehouses and Object Technologies	234
Extended Relational Database Management Systems	235
In Brief	236
8. Security Issues for the Distributed Data Warehouse	239
Key Concepts of Security Planning	242
Network Security Issues	248

Overview of Network Architecture	252
Common Gateway Interface (CGI)	259
Managing Security	262

9. Tools for Distributed Warehouse and Web Security **265**

Firewalls	266
Encryption	270
Authentication Services	279
Secure Sockets Layer/Transport Layer Security	284
Internet Protocol Security	287
Virtual Private Networks	289
Browser Security Options	293
Host Security	295
Final Thoughts	298

10. Implementing a Distributed Data Warehouse **301**

Overview	302
Data Marts	303
To Be Considered	306
Pitfalls	312
Business Factors	316
Implementation Plan	319
Final Considerations	324

11. Future Developments **327**

Evolution	328
Intelligent Distributed Systems	333
The Economic Reality of the Web	340
The New Paradigm	344
Conclusions	346

Glossary **349**

Index **375**