

# On Transactional Concurrency Control

# Synthesis Lectures on Data Management

## Editor

H.V. Jagadish, *University of Michigan*

## Founding Editor

M. Tamer Özsu, *University of Waterloo*

*Synthesis Lectures on Data Management* is edited by H.V. Jagadish of the University of Michigan. The series publishes 80–150 page publications on topics pertaining to data management. Topics include query languages, database system architectures, transaction management, data warehousing, XML and databases, data stream systems, wide scale data distribution, multimedia data management, data mining, and related subjects.

### On Transactional Concurrency Control

Goetz Graefe  
2019

### Answering Queries Using Views, Second Edition

Foto Afrati and Rada Chirkova  
2019

### Transaction Processing on Modern Hardware

Mohammad Sadoghi and Spyros Blanas  
2019

### Data Management in Machine Learning Systems

Matthias Boehm, Arun Kumar, and Jun Yang  
2019

### Non-Volatile Memory Database Management Systems

Joy Arulraj and Andrew Pavlo  
2019

### Scalable Processing of Spatial-Keyword Queries

Ahmed R. Mahmood and Walid G. Aref  
2019

## Data Exploration Using Example-Based Methods

Matteo Lissandrini, Davide Mottin, Themis Palpanas, and Yannis Velegrakis  
2018

## Data Profiling

Ziawasch Abedjan, Lukasz Golab, Felix Naumann, and Thorsten Papenbrock  
2018

## Querying Graphs

Angela Bonifati, George Fletcher, Hannes Voigt, and Nikolay Yakovets  
2018

## Query Processing over Incomplete Databases

Yunjun Gao and Xiaoye Miao  
2018

## Natural Language Data Management and Interfaces

Yunyao Li and Davood Rafiei  
2018

## Human Interaction with Graphs: A Visual Querying Perspective

Sourav S. Bhowmick, Byron Choi, and Chengkai Li  
2018

## On Uncertain Graphs

Arijit Khan, Yuan Ye, and Lei Chen  
2018

## Answering Queries Using Views

Foto Afrati and Rada Chirkova  
2017

## Databases on Modern Hardware: How to Stop Underutilization and Love Multicores

Anastasia Ailamaki, Erieta Liarou, Pınar Tözün, Danica Porobic, and Iraklis Psaroudakis  
2017

## Instant Recovery with Write-Ahead Logging: Page Repair, System Restart, Media Restore, and System Failover, Second Edition

Goetz Graefe, Wey Guy, and Caetano Sauer  
2016

## Generating Plans from Proofs: The Interpolation-based Approach to Query Reformulation

Michael Benedikt, Julien Leblay, Balder ten Cate, and Efthymia Tsamoura  
2016

**Veracity of Data: From Truth Discovery Computation Algorithms to Models of Misinformation Dynamics**

Laure Berti-Équille and Javier Borge-Holthoefer  
2015

**Datalog and Logic Databases**

Sergio Greco and Cristina Molinaro  
2015

**Big Data Integration**

Xin Luna Dong and Divesh Srivastava  
2015

**Instant Recovery with Write-Ahead Logging: Page Repair, System Restart, and Media Restore**

Goetz Graefe, Wey Guy, and Caetano Sauer  
2014

**Similarity Joins in Relational Database Systems**

Nikolaus Augsten and Michael H. Böhlen  
2013

**Information and Influence Propagation in Social Networks**

Wei Chen, Laks V.S. Lakshmanan, and Carlos Castillo  
2013

**Data Cleaning: A Practical Perspective**

Venkatesh Ganti and Anish Das Sarma  
2013

**Data Processing on FPGAs**

Jens Teubner and Louis Woods  
2013

**Perspectives on Business Intelligence**

Raymond T. Ng, Patricia C. Arocena, Denilson Barbosa, Giuseppe Carenini, Luiz Gomes, Jr., Stephan Jou, Rock Anthony Leung, Evangelos Milios, Renée J. Miller, John Mylopoulos, Rachel A. Pottinger, Frank Tompa, and Eric Yu  
2013

**Semantics Empowered Web 3.0: Managing Enterprise, Social, Sensor, and Cloud-based Data and Services for Advanced Applications**

Amit Sheth and Krishnaprasad Thirunarayan  
2012

## Data Management in the Cloud: Challenges and Opportunities

Divyakant Agrawal, Sudipto Das, and Amr El Abbadi

2012

## Query Processing over Uncertain Databases

Lei Chen and Xiang Lian

2012

## Foundations of Data Quality Management

Wenfei Fan and Floris Geerts

2012

## Incomplete Data and Data Dependencies in Relational Databases

Sergio Greco, Cristian Molinaro, and Francesca Spezzano

2012

## Business Processes: A Database Perspective

Daniel Deutch and Tova Milo

2012

## Data Protection from Insider Threats

Elisa Bertino

2012

## Deep Web Query Interface Understanding and Integration

Eduard C. Dragut, Weiyi Meng, and Clement T. Yu

2012

## P2P Techniques for Decentralized Applications

Esther Pacitti, Reza Akbarinia, and Manal El-Dick

2012

## Query Answer Authentication

HweeHwa Pang and Kian-Lee Tan

2012

## Declarative Networking

Boon Thau Loo and Wenchoao Zhou

2012

## Full-Text (Substring) Indexes in External Memory

Marina Barsky, Ulrike Stege, and Alex Thomo

2011

## Spatial Data Management

Nikos Mamoulis

2011

**Database Repairing and Consistent Query Answering**

Leopoldo Bertossi

2011

**Managing Event Information: Modeling, Retrieval, and Applications**

Amarnath Gupta and Ramesh Jain

2011

**Fundamentals of Physical Design and Query Compilation**

David Toman and Grant Weddell

2011

**Methods for Mining and Summarizing Text Conversations**

Giuseppe Carenini, Gabriel Murray, and Raymond Ng

2011

**Probabilistic Databases**

Dan Suciu, Dan Olteanu, Christopher Ré, and Christoph Koch

2011

**Peer-to-Peer Data Management**

Karl Aberer

2011

**Probabilistic Ranking Techniques in Relational Databases**

Ihab F. Ilyas and Mohamed A. Soliman

2011

**Uncertain Schema Matching**

Avigdor Gal

2011

**Fundamentals of Object Databases: Object-Oriented and Object-Relational Design**

Suzanne W. Dietrich and Susan D. Urban

2010

**Advanced Metasearch Engine Technology**

Weiyi Meng and Clement T. Yu

2010

**Web Page Recommendation Models: Theory and Algorithms**

Sule Gündüz-Öğüdücü

2010

**Multidimensional Databases and Data Warehousing**

Christian S. Jensen, Torben Bach Pedersen, and Christian Thomsen

2010

**Database Replication**

Bettina Kemme, Ricardo Jimenez-Peris, and Marta Patino-Martinez  
2010

**Relational and XML Data Exchange**

Marcelo Arenas, Pablo Barcelo, Leonid Libkin, and Filip Murlak  
2010

**User-Centered Data Management**

Tiziana Catarcì, Alan Dix, Stephen Kimani, and Giuseppe Santucci  
2010

**Data Stream Management**

Lukasz Golab and M. Tamer Özsu  
2010

**Access Control in Data Management Systems**

Elena Ferrari  
2010

**An Introduction to Duplicate Detection**

Felix Naumann and Melanie Herschel  
2010

**Privacy-Preserving Data Publishing: An Overview**

Raymond Chi-Wing Wong and Ada Wai-Chee Fu  
2010

**Keyword Search in Databases**

Jeffrey Xu Yu, Lu Qin, and Lijun Chang  
2009

© Springer Nature Switzerland AG 2022

Reprint of original edition © Morgan & Claypool 2019

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means—electronic, mechanical, photocopy, recording, or any other except for brief quotations in printed reviews, without the prior permission of the publisher.

On Transactional Concurrency Control

Goetz Graefe

ISBN: 978-3-031-00745-3      paperback

ISBN: 978-3-031-01873-2      ebook

ISBN: 978-3-031-00100-0      hardcover

DOI 10.1007/978-3-031-01873-2

A Publication in the Springer series

*SYNTHESIS LECTURES ON DATA MANAGEMENT*

Lecture #59

Series Editor: H.V. Jagadish, *University of Michigan*

Founding Editor: M. Tamer Özsü, *University of Waterloo*

Series ISSN

Print 2153-5418   Electronic 2153-5426

# On Transactional Concurrency Control

Goetz Graefe

*SYNTHESIS LECTURES ON DATA MANAGEMENT #59*

## ABSTRACT

This book contains a number of chapters on transactional database concurrency control. A two-sentence summary of the volume's entire sequence of chapters is this: traditional locking techniques can be improved in multiple dimensions, notably in lock scopes (sizes), lock modes (increment, decrement, and more), lock durations (late acquisition, early release), and lock acquisition sequence (to avoid deadlocks). Even if some of these improvements can be transferred to optimistic concurrency control, notably a fine granularity of concurrency control with serializable transaction isolation including phantom protection, pessimistic concurrency control is categorically superior to optimistic concurrency control, i.e., independent of application, workload, deployment, hardware, and software implementation.

## KEYWORDS

concurrency, concurrency control, database, update, transaction, serializability, phantom protection, locking, index, b-tree, key-range locking, key-value locking, deadlock avoidance, lock acquisition sequence, pessimistic concurrency control, optimistic concurrency control, validation, timestamps, snapshot isolation, versioning, multi-version storage, write buffer, distributed systems, mirroring, replication, log shipping, two-phase commit, three-phase commit, controlled lock violation, deferred lock acquisition, deferred lock enforcement, weak lock enforcement, reserved locks, pending locks, orthogonal key-range locking, orthogonal key-value locking, partitioning

# Contents

<b>On Transactional Concurrency Control .....</b>	<b>1</b>
Lock Management .....	1
1 Published Papers .....	3
1.1 A Survey of B-Tree Locking Techniques (ACM TODS 2010) .....	3
1.2 Hierarchical Locking in B-Tree Indexes (BTW Keynote 2007) .....	3
1.3 Transaction Support for Indexed Views (ACM SIGMOD 2004) .....	3
1.4 Controlled Lock Violation (ACM SIGMOD 2013) .....	4
1.5 Orthogonal Key-Value Locking (BTW 2015 and Extended Draft 2019) .....	4
2 Optimistic Concurrency Control.....	4
2.1 Orthogonal Key-Value Validation .....	5
2.2 Serializable Timestamp Validation.....	5
2.3 Repairing Optimistic Concurrency Control .....	6
3 Locking.....	6
3.1 Avoiding Index Navigation Deadlocks.....	6
3.2 Two-Phase Commit .....	6
3.3 Deferred Lock Enforcement .....	7
4 The End of Optimistic Concurrency Control .....	8
Acknowledgements .....	9
References.....	9
<b>PART I Published Papers .....</b>	<b>15</b>
<b>1 A Survey of B-Tree Locking Techniques .....</b>	<b>17</b>
<i>Goetz Graefe</i>	
<i>Hewlett-Packard Laboratories</i>	
Abstract .....	17
1.1 Introduction .....	17
1.1.1 Historical Background .....	18

1.1.2	Overview . . . . .	20
1.2	Preliminaries . . . . .	20
1.3	Two Forms of B-Tree Locking . . . . .	21
1.3.1	Locks and Latches . . . . .	21
1.3.2	Recovery and B-Tree Locking . . . . .	23
1.3.3	Lock-Free B-Trees . . . . .	24
1.3.4	Summary . . . . .	24
1.4	Protecting a B-Tree's Physical Structure . . . . .	24
1.4.1	Issues . . . . .	25
1.4.2	Lock Coupling . . . . .	26
1.4.3	B <sup>link</sup> -Trees . . . . .	27
1.4.4	Load Balancing and Reorganization . . . . .	28
1.4.5	Summary . . . . .	29
1.5	Protecting a B-Tree's Logical Contents . . . . .	29
1.5.1	Key Range Locking . . . . .	30
1.5.2	Key Range Locking and Ghost Records . . . . .	31
1.5.3	Key Range Locking as Hierarchical Locking . . . . .	33
1.5.4	Locking in Nonunique Indexes . . . . .	36
1.5.5	Increment Lock Modes . . . . .	37
1.5.6	Summary . . . . .	39
1.6	Future Directions . . . . .	40
1.7	Summary and Conclusions . . . . .	40
	Acknowledgements . . . . .	41
1.8	References . . . . .	41
<b>2</b>	<b>Hierarchical Locking in B-Tree Indexes . . . . .</b>	<b>45</b>
	<i>Goetz Graefe</i>	
	<i>Hewlett-Packard Laboratories</i>	
	Abstract . . . . .	45
2.1	Introduction . . . . .	45
2.1.1	Problem Overview . . . . .	46
2.1.2	Solution Overview . . . . .	46
2.2	Related Work . . . . .	47
2.2.1	Related Work on Multi-Granularity Locking . . . . .	47
2.2.2	Related Work on Key Range Locking . . . . .	48
2.3	Assumptions . . . . .	51

2.4	Traditional Locking Hierarchies . . . . .	53
2.4.1	Locks on Keys and Ranges . . . . .	53
2.4.2	Key Insertion Using Ghost Records . . . . .	56
2.4.3	Splitting and Merging Key Ranges . . . . .	57
2.4.4	Summary . . . . .	58
2.5	Locks on Separator Keys . . . . .	58
2.5.1	Techniques . . . . .	59
2.5.2	Advantages . . . . .	63
2.5.3	Disadvantages . . . . .	63
2.5.4	Opportunities . . . . .	63
2.5.5	Summary . . . . .	64
2.6	Locks on Key Prefixes . . . . .	64
2.6.1	Techniques . . . . .	65
2.6.2	Advantages . . . . .	67
2.6.3	Disadvantages . . . . .	68
2.6.4	Opportunities . . . . .	68
2.6.5	Summary . . . . .	69
2.7	Summary and Conclusions . . . . .	69
	Acknowledgments . . . . .	71
2.8	References . . . . .	71

### **3 Concurrent Queries and Updates in Summary Views and Their Indexes . . . . . 75**

	<i>Goetz Graefe</i>	
	Abstract . . . . .	75
3.1	Introduction . . . . .	75
3.1.1	Serializable View Maintenance . . . . .	77
3.1.2	Contributions . . . . .	78
3.1.3	Overview . . . . .	79
3.2	Prior Work . . . . .	80
3.2.1	Deltas for Materialized Views . . . . .	80
3.2.2	Concurrency Control for Materialized Views . . . . .	80
3.2.3	Multi-Level Transactions and Aries . . . . .	81
3.2.4	Concurrent Update Techniques . . . . .	82
3.2.5	Snapshot Isolation and Multi-Version Concurrency Control . . . . .	84
3.2.6	Derived Lock Modes . . . . .	85
3.2.7	Lock Escalation and De-Escalation . . . . .	86
3.2.8	Transaction Save Points . . . . .	87

3.2.9	Local Commit Processing . . . . .	88
3.2.10	Two-Phase Commit . . . . .	89
3.2.11	Managing Individual Records and Keys . . . . .	90
3.2.12	Online Index Operations . . . . .	91
3.2.13	Duplicate Records and Keys . . . . .	92
3.2.14	Incremental View Maintenance . . . . .	93
3.2.15	Expanded “Group by” Clauses . . . . .	94
3.3	Multi-Version Snapshot Isolation . . . . .	95
3.4	Concurrent Updates and Linear Version History . . . . .	97
3.4.1	Multiple Records per Summary Row . . . . .	98
3.4.2	Multiple Delta Records per Summary Row . . . . .	100
3.4.3	Deferred Actions and Virtual Log Records . . . . .	100
3.4.4	Commit-Time-Only Locks . . . . .	101
3.4.5	Commit Processing . . . . .	103
3.4.6	Transaction Abort . . . . .	104
3.4.7	Transaction Save Points . . . . .	104
3.4.8	Two-Phase Commit . . . . .	105
3.5	Logging and Recovery . . . . .	105
3.5.1	Transaction Abort . . . . .	106
3.5.2	Partial Rollback for Save Points . . . . .	106
3.5.3	Transaction Commit . . . . .	107
3.5.4	System and Media Recovery . . . . .	108
3.5.5	Failures and Transaction Abort During Commit Processing . . . . .	109
3.5.6	Two-Phase Commit . . . . .	109
3.5.7	Optimistic Concurrency Control . . . . .	110
3.6	Multi-Granularity Locking . . . . .	112
3.6.1	Derived Lock Types . . . . .	112
3.6.2	Interpretation of the Lock Table . . . . .	114
3.6.3	Commit Processing with Intention Locks . . . . .	114
3.7	Update and Upgrade Locks . . . . .	115
3.8	Insert and Delete . . . . .	116
3.8.1	Escrow Locks and Key Range Locking . . . . .	117
3.8.2	Record Creation by System Transactions . . . . .	118
3.8.3	Locking and Logging in System Transactions . . . . .	118
3.9	Online Index Operations . . . . .	119
3.10	Correctness . . . . .	121

3.11	Performance .....	122
3.12	Summary and Conclusions .....	123
3.13	References.....	124
<b>4</b>	<b>Controlled Lock Violation .....</b>	<b>129</b>
<i>Goetz Graefe, Mark Lillibridge, Harumi Kuno, Joseph Tucek, and Alistair Veitch</i>		
<i>Hewlett-Packard Laboratories</i>		
	Abstract .....	129
4.1	Introduction .....	129
4.2	Related Prior Work .....	132
4.2.1	Main Memory Databases .....	133
4.2.2	Early Lock Release in Shore-MT .....	133
4.2.3	Early Lock Release in Foster B-Trees .....	135
4.2.4	Other Cases of Early Lock Release .....	137
4.2.5	Distributed Commit .....	138
4.3	Controlled Lock Violation.....	139
4.3.1	Principles .....	140
4.3.2	Approach.....	141
4.3.3	Implementation Techniques .....	142
4.3.4	Comparison with Early Lock Release .....	143
4.3.5	Combined Locks.....	144
4.3.6	Weak Transaction Isolation Levels.....	144
4.3.7	Summary .....	145
4.4	Distributed Transactions .....	146
4.4.1	Implementation of Commit Dependencies .....	147
4.4.2	Performance Effects .....	147
4.4.3	Summary .....	148
4.5	Canned Transactions .....	148
4.6	Performance Evaluation .....	150
4.6.1	Implementation .....	150
4.6.2	Results .....	151
4.7	Discussion .....	152
4.8	Summary and Conclusions .....	154
4.9	Acknowledgements .....	155
4.10	References .....	156

<b>5</b>	<b>Orthogonal Key-Value Locking . . . . .</b>	<b>159</b>
<i>Goetz Graefe</i>		
5.1	Introduction . . . . .	160
5.1.1	Motivation . . . . .	160
5.1.2	Concurrency Control and Lock Scopes . . . . .	162
5.1.3	Design Goals . . . . .	162
5.1.4	Outline . . . . .	163
5.2	Prior Designs . . . . .	163
5.2.1	History of Locking in B-Trees . . . . .	163
5.2.2	ARIES/KVL “Key-Value Locking” . . . . .	167
5.2.3	ARIES/IM “Index Management” . . . . .	168
5.2.4	SQL Server Key-Range Locking . . . . .	169
5.2.5	Orthogonal Key-Range Locking . . . . .	170
5.2.6	Summary of Prior Designs . . . . .	171
5.3	Orthogonal Key-Value Locking . . . . .	172
5.3.1	Design Goals . . . . .	172
5.3.2	Orthogonality . . . . .	172
5.3.3	Partitioned Lists . . . . .	173
5.3.4	Partitioned Gaps . . . . .	175
5.3.5	Lock Manager Implementation . . . . .	178
5.3.6	Unique Secondary Indexes . . . . .	180
5.3.7	Primary Indexes . . . . .	181
5.3.8	Master-Detail Clustering . . . . .	182
5.3.9	Summary of Orthogonal Key-Value Locking . . . . .	183
5.4	Case Studies . . . . .	184
5.4.1	Empty Queries—Phantom Protection . . . . .	184
5.4.2	Successful Equality Queries . . . . .	187
5.4.3	Phantom Protection with Ghost Records . . . . .	189
5.4.4	Range Queries . . . . .	190
5.4.5	Non-Key Updates . . . . .	192
5.4.6	Deletions . . . . .	193
5.4.7	Insertions . . . . .	195
5.4.8	Key Updates . . . . .	196
5.4.9	Non-Unique Primary Indexes . . . . .	196
5.4.10	Summary of the Case Studies . . . . .	199
5.5	Future Opportunities . . . . .	199

5.5.1	Column-Level Locking . . . . .	199
5.5.2	Orthogonal Row Locking . . . . .	200
5.5.3	Orthogonal Locking in Tables and Indexes . . . . .	202
5.5.4	Orthogonal Schema Locking . . . . .	204
5.5.5	Summary of Future Opportunities . . . . .	206
5.6	Conclusions . . . . .	206
5.7	References . . . . .	208
<b>PART II   Optimistic Concurrency Control . . . . .</b>		<b>211</b>
<b>6</b>	<b>Orthogonal Key-Value Validation . . . . .</b>	<b>213</b>
<i>Goetz Graefe</i>		
6.1	Introduction . . . . .	214
6.2	Related Prior Work . . . . .	215
6.2.1	Two-Phase Locking . . . . .	215
6.2.2	Transaction Isolation Levels . . . . .	217
6.2.3	Ghost Records . . . . .	218
6.2.4	System Transactions . . . . .	219
6.2.5	Orthogonal Key-Value Locking . . . . .	220
6.2.6	A Brief Comparison of Locking Techniques . . . . .	222
6.2.7	Optimistic Concurrency Control . . . . .	224
6.2.8	Summary of Related Prior Work . . . . .	226
6.3	Orthogonal Key-Value Validation . . . . .	226
6.3.1	Read- and Write-Sets . . . . .	227
6.3.2	Ghost Records and System Transactions . . . . .	228
6.3.3	User Transactions . . . . .	229
6.3.4	Phantom Protection . . . . .	230
6.3.5	Integrity Constraints . . . . .	233
6.3.6	Transaction Isolation Levels . . . . .	233
6.3.7	Hierarchical and Multi-Granularity Concurrency Control . . . . .	234
6.3.8	Beyond B-Tree Indexes . . . . .	235
6.3.9	Summary of the Design . . . . .	236
6.4	Case Studies . . . . .	236
6.4.1	Queries with Empty Results—Phantom Protection . . . . .	237
6.4.2	Successful Equality Queries . . . . .	239

6.4.3	Range Queries . . . . .	240
6.4.4	Non-Key Updates . . . . .	241
6.4.5	Deletions . . . . .	241
6.4.6	Insertions . . . . .	242
6.4.7	Key Updates . . . . .	244
6.4.8	Summary of the Case Studies . . . . .	244
6.5	Alternative Approaches . . . . .	244
6.6	Conclusions . . . . .	246
6.7	References . . . . .	247
<b>7</b>	<b>Serializable Timestamp Validation . . . . .</b>	<b>251</b>
<i>Goetz Graefe</i>		
7.1	Introduction . . . . .	251
7.2	Related Prior Work . . . . .	253
7.2.1	Ghost Records and System Transactions . . . . .	253
7.2.2	Multi-Version Storage and Snapshot Isolation . . . . .	254
7.2.3	Traditional Locking . . . . .	254
7.2.4	Orthogonal Key-Value Locking . . . . .	255
7.2.5	A Quick Comparison . . . . .	257
7.2.6	Optimistic Concurrency Control . . . . .	259
7.2.7	Timestamp Validation . . . . .	260
7.2.8	Summary of Related Prior Work . . . . .	261
7.3	Phantom Protection with Timestamp Validation . . . . .	261
7.3.1	Basic Designs . . . . .	262
7.3.2	Technology Transfers from Traditional Locking Techniques . . . . .	264
7.3.3	Orthogonal Key-Value Timestamp Validation . . . . .	266
7.3.4	Summary of Timestamp Validation with Phantom Protection . . . . .	267
7.4	Management of Timestamps . . . . .	268
7.4.1	Compressing Timestamps . . . . .	269
7.4.2	Caching Timestamps . . . . .	270
7.4.3	Timestamp Validation in Databases Without Timestamps . . . . .	271
7.4.4	Client-Server Database Storage . . . . .	272
7.4.5	Summary of Timestamp Management . . . . .	273
7.5	Hierarchical Timestamps . . . . .	274
7.6	Conclusions . . . . .	276
7.7	References . . . . .	277

<b>8</b>	<b>Repairing Optimistic Concurrency Control . . . . .</b>	<b>279</b>
<i>Goetz Graefe</i>		
8.1	Introduction . . . . .	279
8.2	Related Prior Work . . . . .	281
8.2.1	Traditional Optimistic Concurrency Control . . . . .	281
8.2.2	Locking Techniques . . . . .	282
8.2.3	Two-Phase Commit . . . . .	284
8.2.4	Controlled Lock Violation . . . . .	285
8.2.5	Summary of Related Prior Work . . . . .	288
8.3	Concurrent Validation . . . . .	288
8.3.1	The Need for Concurrent Validation . . . . .	288
8.3.2	Simple and Efficient Concurrent Validation . . . . .	289
8.3.3	Summary of Concurrent Validation . . . . .	291
8.4	Premature Publication . . . . .	292
8.4.1	The Danger of Premature Publication . . . . .	292
8.4.2	Optimistic Transactions Without Premature Publication . . . . .	294
8.4.3	Summary of Premature Publication . . . . .	295
8.5	Distributed Transactions . . . . .	296
8.5.1	Two-Phase Commit . . . . .	296
8.5.2	Three-Phase Commit and Controlled Lock Violation . . . . .	297
8.5.3	Summary of Distributed Transactions . . . . .	298
8.6	Conclusions . . . . .	299
8.7	References . . . . .	299
<b>PART III Locking . . . . .</b>		<b>301</b>
<b>9</b>	<b>Avoiding Index-Navigation Deadlocks . . . . .</b>	<b>303</b>
<i>Goetz Graefe</i>		
9.1	Introduction . . . . .	303
9.2	Related Prior Work . . . . .	305
9.2.1	Ghost Records and System Transactions . . . . .	306
9.2.2	ARIES/KVL “Key-Value Locking” . . . . .	307
9.2.3	ARIES/IM “Index Management” . . . . .	307
9.2.4	Orthogonal Key-Value Locking . . . . .	309
9.2.5	A Brief Comparison of Locking Techniques . . . . .	311

9.2.6	Summary of Related Prior Work .....	313
9.3	Recommended Locking Sequences .....	313
9.3.1	New Techniques .....	313
9.3.2	Examples .....	315
9.3.3	Effects .....	316
9.4	Conclusions .....	317
9.5	References .....	318
<b>10</b>	<b>A Problem in Two-Phase Commit .....</b>	<b>321</b>
<i>Goetz Graefe</i>		
10.1	References .....	326
<b>11</b>	<b>Deferred Lock Enforcement .....</b>	<b>327</b>
<i>Goetz Graefe</i>		
11.1	Introduction .....	327
11.2	Related Prior Work .....	329
11.2.1	Multi-Version Storage .....	329
11.2.2	Transaction-Private Update Buffers .....	330
11.2.3	Update Locks and Intention Locks .....	331
11.2.4	Orthogonal Key-Value Locking .....	333
11.2.5	Early Lock Release and Controlled Lock Violation .....	335
11.2.6	Summary of Related Prior Work .....	336
11.3	Deferred Lock Acquisition .....	337
11.3.1	Transaction-Private Update Buffers .....	338
11.3.2	Lock Timing and Semantics .....	339
11.3.3	Integration with Multi-Version Storage .....	340
11.3.4	Summary of Deferred Lock Acquisition .....	341
11.4	Deferred Lock Enforcement .....	341
11.4.1	Lock Semantics and Transitions .....	342
11.4.2	Integration with Controlled Lock Violation .....	343
11.4.3	Integration with Multi-Version Storage .....	344
11.4.4	Transaction Commit Logic .....	345
11.4.5	Summary of Deferred Lock Enforcement .....	347
11.5	Deferred Lock Enforcement and Other Techniques .....	347
11.5.1	System Transactions .....	347
11.5.2	Weak Transaction Isolation Levels .....	348
11.5.3	Integration with Hierarchical Locking .....	350

11.5.4	Lock Escalation and De-Escalation . . . . .	350
11.5.5	Increment/Decrement Locks . . . . .	351
11.5.6	Deadlocks . . . . .	353
11.5.7	Transaction Rollback . . . . .	353
11.5.8	Instant Recovery . . . . .	354
11.5.9	Optimistic Concurrency Control . . . . .	355
11.5.10	Summary of Deferred Lock Enforcement with Other Techniques .	358
11.6	Distributed Operations . . . . .	359
11.6.1	Transaction Commit Logic . . . . .	359
11.6.2	Log Shipping and Replication . . . . .	361
11.6.3	Summary of Distributed Operations . . . . .	362
11.7	Summary and Conclusions . . . . .	362
11.8	References . . . . .	364
 <b>PART IV The End of Optimistic Concurrency Control 367</b>		
<b>12</b>	<b>The End of Optimistic Concurrency Control . . . . .</b>	<b>369</b>
<i>Goetz Graefe</i>		
12.1	Introduction . . . . .	369
12.2	Earlier Comparisons . . . . .	370
12.3	Common Misconceptions . . . . .	373
12.4	Locking in Optimistic Concurrency Control . . . . .	374
12.5	Advantage: Locking . . . . .	376
12.6	Another Quick Look at Optimistic Concurrency Control . . . . .	377
12.7	Recommendations . . . . .	379
12.8	Conclusions . . . . .	379
12.9	References . . . . .	380
<b>Author's Biography . . . . .</b>		<b>383</b>