Database Replication
Synthesis Lectures on Data Management

**Editor**

**M. Tamer Özsu, University of Waterloo**

Synthesis Lectures on Data Management is edited by Tamer Özsu of the University of Waterloo. The series will publish 50- to 125 page publications on topics pertaining to data management. The scope will largely follow the purview of premier information and computer science conferences, such as ACM SIGMOD, VLDB, ICDE, PODS, ICDT, and ACM KDD. Potential topics include, but not are limited to: 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.

**Database Replication**
Bettina Kemme, Ricardo Jiménez Peris, and Marta Patiño-Martínez
2010

**User-Centered Data Management**
Tiziana Catarci, 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
Database Replication

Bettina Kemme
McGill University

Ricardo Jiménez Peris
Technical University of Madrid

Marta Patiño-Martínez
Technical University of Madrid

SYNTHESIS LECTURES ON DATA MANAGEMENT #7
ABSTRACT

Database replication is widely used for fault-tolerance, scalability and performance. The failure of one database replica does not stop the system from working as available replicas can take over the tasks of the failed replica. Scalability can be achieved by distributing the load across all replicas, and adding new replicas should the load increase. Finally, database replication can provide fast local access, even if clients are geographically distributed clients, if data copies are located close to clients.

Despite its advantages, replication is not a straightforward technique to apply, and there are many hurdles to overcome. At the forefront is replica control: assuring that data copies remain consistent when updates occur. There exist many alternatives in regard to where updates can occur and when changes are propagated to data copies, how changes are applied, where the replication tool is located, etc. A particular challenge is to combine replica control with transaction management as it requires several operations to be treated as a single logical unit, and it provides atomicity, consistency, isolation and durability across the replicated system. The book provides a categorization of replica control mechanisms, presents several replica and concurrency control mechanisms in detail, and discusses many of the issues that arise when such solutions need to be implemented within or on top of relational database systems.

Furthermore, the book presents the tasks that are needed to build a fault-tolerant replication solution, provides an overview of load-balancing strategies that allow load to be equally distributed across all replicas, and introduces the concept of self-provisioning that allows the replicated system to dynamically decide on the number of replicas that are needed to handle the current load. As performance evaluation is a crucial aspect when developing a replication tool, the book presents an analytical model of the scalability potential of various replication solution.

For readers that are only interested in getting a good overview of the challenges of database replication and the general mechanisms of how to implement replication solutions, we recommend to read Chapters 1 to 4. For readers that want to get a more complete picture and a discussion of advanced issues, we further recommend the Chapters 5, 8, 9 and 10. Finally, Chapters 6 and 7 are of interest for those who want get familiar with thorough algorithm design and correctness reasoning.

KEYWORDS
database replication, transactions, replica control, 1-copy-equivalence, consistency, scalability, fault-tolerance, performance, elasticity
To Maya and Sophia

Bettina

To my parents Ricardo and Maria Adelina, my siblings David and Ana, and my son Alejandro

Ricardo

To my parents Manuel and Cele, my brother Manolo, and my son Alejandro

Marta
## Contents

1 Overview

1.1 Motivation ............................................................ 1

1.2 Challenges ............................................................. 4
   1.2.1 Replica Control .................................................. 4
   1.2.2 Other Issues ..................................................... 6

2 1-Copy-Equivalence and Consistency .................................... 9

   2.1 Replication Model ...................................................... 9
   2.2 1-Copy-Isolation ...................................................... 10
   2.3 1-Copy-Atomicity ..................................................... 13
   2.4 1-Copy-Durability ..................................................... 14
   2.5 Relationship Between Isolation, Atomicity and Durability ............... 14
   2.6 1-Copy-Consistency ................................................... 15
   2.7 Session Consistency. ................................................... 15
   2.8 Eventual Consistency ................................................ 16

3 Basic Protocols .......................................................... 19

   3.1 Eager Protocols ....................................................... 19
      3.1.1 Protocol Description ............................................. 19
      3.1.2 Example Execution .............................................. 21
      3.1.3 Eager Properties ................................................ 22
      3.1.4 Primary Copy vs. Update Anywhere ......................... 23
   3.2 Lazy Protocols ........................................................ 25
      3.2.1 Protocol Description ............................................. 25
      3.2.2 Example Execution .............................................. 25
      3.2.3 Lazy vs. Eager Properties ....................................... 25
      3.2.4 Primary Copy vs. Update Anywhere ......................... 27
      3.2.5 Summary ....................................................... 29
# 1-Copy-Snapshot Isolation

7.1 1-Copy-Snapshot Isolation ............................................. 68  
7.1.1 Snapshot Isolation in a Non-Replicated System ...................... 68  
7.1.2 Snapshot Isolation in a Replicated System .......................... 70  

7.2 Primary Copy – Centralized Middleware .................................. 71  
7.2.1 Protocol ..................................................................... 72  
7.2.2 Example Execution .............................................. 72  
7.2.3 Algorithm Properties ............................................. 75  

7.3 Update Anywhere – Centralized Middleware ............................. 76  
7.3.1 Protocol ..................................................................... 77  
7.3.2 Example Execution .............................................. 79  
7.3.3 Algorithm Properties ............................................. 79  

7.4 Update-Anywhere – Decentralized Middleware ............................ 81  
7.4.1 Protocol Description ............................................. 82  
7.4.2 Example Execution .............................................. 83  
7.4.3 Algorithm Properties ............................................. 85  

7.5 Snapshot Isolation vs. Traditional Optimistic Concurrency Control ........... 86  
7.6 Related Work .................................................................. 86  

# Lazy Replication

8.1 Bounding the Staleness in Lazy Primary Copy ............................. 89  
8.1.1 Boundary Types .................................................. 89  
8.1.2 Basic Implementation ........................................... 90  
8.1.3 Push vs. Pull Based Refresh ..................................... 91  
8.1.4 Materialized Views ............................................... 92  
8.1.5 Transaction Propagation ......................................... 92  

8.2 Multiple Primaries ................................................................ 94  

8.3 Lazy Update Anywhere .................................................... 96  
8.3.1 Distributed vs. Central Conflict Management ..................... 96  
8.3.2 Conflict Detection ................................................ 97  
8.3.3 Conflict Resolution ................................................ 99  

8.4 Related Work .................................................................... 101  
8.4.1 Bounding Staleness .............................................. 101  
8.4.2 Replica Placement ............................................... 102  
8.4.3 Conflict Detection and Resolution ................................ 102
## 9 Self-Configuration and Elasticity

9.1 Self-healing
9.1.1 Fault-Tolerant Measures During Normal Processing
9.1.2 Failure Types
9.1.3 Failover: Client Side
9.1.4 Failover: Server Side
9.1.5 Recovery

9.2 Self-Optimization
9.2.1 Load-Balancing
9.2.2 Other Optimization Techniques

9.3 Elasticity: Self-Provisioning
9.3.1 System Reconfiguration
9.3.2 Deciding on the Right Number of Replicas

9.4 Related Work

## 10 Other Aspects of Replication

10.1 Multi-Tier Architectures
10.2 Quorums
10.3 Mobile and Peer-to-Peer Environments

## A Transactions and the ACID Properties

A.1 Atomicity
A.2 Consistency
A.3 Durability
A.4 Isolation
A.5 Distributed Transactions and 2-Phase Commit

## Bibliography

## Authors' Biographies