

O'REILLY®

Database Internals

A Deep-Dive into How Distributed Data
Systems Work



Alex Petrov

Database Internals

When it comes to choosing, using, and maintaining a database, understanding its internals is essential. But with so many distributed databases and tools available today, it's often difficult to learn what each one offers and how they differ. With this practical guide, Alex Petrov guides developers through the concepts behind modern database and storage engine internals.

Throughout the book, you'll explore relevant material gleaned from numerous books, papers, blog posts, and the source code of several open source databases. You'll discover that the most significant distinctions among many modern databases reside in subsystems that determine how storage is organized and how data is distributed.

You'll delve into:

- **Storage engines:** Learn storage classification and taxonomy and examine B-tree-based and immutable log-structured storage engines
- **Storage building blocks:** Understand how database files are organized to build efficient storage using auxiliary data structures such as page cache and buffer pools
- **Distributed systems:** Learn step-by-step how nodes and processes connect and build complex communication patterns
- **Database clusters:** Explore which consistency models are commonly used by modern databases and how distributed storage systems achieve consistency

"In order to pick the right tool for the job, we need to understand the ideas and algorithms behind their design. This book is a beginner-friendly introduction to a number of relevant topics written by an industry practitioner."

—Michael Klishin

along time RabbitMQ contributor

"This is a must-have book for anyone working with any sort of database technology, particularly when in the position of having to make decisions about which systems to use."

—Nate McCall

Apache Cassandra Committer
and PMC Chair

Alex Petrov is a data infrastructure engineer, database and storage systems enthusiast, Apache Cassandra committer, and PMC member. His expertise is in storage, distributed systems, and algorithms.

DATABASES

US \$59.99 CAN \$79.99

ISBN: 978-1-492-04034-7



9 781492 040347



Twitter: @oreillymedia
facebook.com/oreilly

Database Internals

*A Deep Dive into How
Distributed Data Systems Work*

Alex Petrov

Beijing • Boston • Farnham • Sebastopol • Tokyo

O'REILLY®