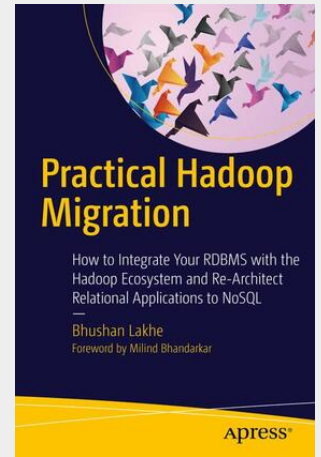


Lakhe

Practical Hadoop Migration

How to Integrate Your RDBMS with the Hadoop Ecosystem and Re-Architect Relational Applications to NoSQL

Re-architect relational applications to NoSQL, integrate relational database management systems with the Hadoop ecosystem, and transform and migrate relational data to and from Hadoop components. This book covers the best-practice design approaches to re-architecting your relational applications and transforming your relational data to optimize concurrency, security, denormalization, and performance. Winner of IBM's 2012 Gerstner Award for his implementation of big data and data warehouse initiatives and author of Practical Hadoop Security, author Bhushan Lakhe walks you through the entire transition process. First, he lays out the criteria for deciding what blend of re-architecting, migration, and integration between RDBMS and HDFS best meets your transition objectives. Then he demonstrates how to design your transition model. Lakhe proceeds to cover the selection criteria for ETL tools, the implementation steps for migration with SQOOP- and Flume-based data transfers, and transition optimization techniques for tuning partitions, scheduling aggregations, and redesigning ETL. Finally, he assesses the pros and cons of data lakes and Lambda architecture as integrative solutions and illustrates their implementation with real-world case studies. Hadoop/NoSQL solutions do not offer by default certain relational technology features such as role-based access control, locking for concurrent updates, and various tools for measuring and enhancing performance. Practical Hadoop Migration shows how to use open-source tools to emulate such relational functionalities in Hadoop ecosystem components. What You'll Learn - Decide whether you should migrate your relational applications to big data technologies or integrate them - Transition your relational applications to Hadoop/NoSQL platforms in terms of logical design and physical implementation - Discover RDBMS-to-HDFS integration, data transformation, and optimization techniques - Consider when to use Lambda architecture and data lake solutions - Select and implement Hadoop-based components and applications to speed transition, optimize integrated performance, and emulate relational functionalities Who This Book Is For Database developers, database administrators, enterprise architects, Hadoop/NoSQL developers, and IT leaders. Its secondary readership is project and program managers and advanced students of database and management information systems.



42,79 €

39,99 € (zzgl. MwSt.)

Lieferfrist: bis zu 10 Tage

Artikelnummer: 9781484212882

Medium: Buch

ISBN: 978-1-4842-1288-2

Verlag: Apress

Erscheinungstermin: 11.08.2016

Sprache(n): Englisch

Auflage: 1. Auflage 2016

Produktform: Kartoniert

Gewicht: 5037 g

Seiten: 305

Format (B x H): 155 x 235 mm

