Hbase The Definitive Guide Random Access To Your Planet Size Data

Presto: The Definitive Guide
Cloud Computing for Science and Engineering
Programming Pig
A 360-degree Approach
Building Effective Algorithms and Analytics for Hadoop and Other Systems
Mathematics of Big Data
Web Engineering
First BenchCouncil International Symposium, Bench 2018, Seattle, WA, USA, December 10-13, 2018, Revised Selected Papers
Distributed Applications and Interoperable Systems
Cloud Computing in Remote Sensing
Architecting HBase Applications
Intelligent Systems
Random Access to Your Planet-Size Data
Distributed Data at Web Scale
13th Conference, ACA 2020, Kunming, China, August 13–15, 2020, Proceedings
First International Conference, BigSDM 2018, Beijing, China, November 30 – December 1, 2018, Revised Selected Papers
Advanced Computer Architecture
Hadoop in Action
Hadoop Application Architectures
A Guidebook for Successful Development and Design
HBase: The Definitive Guide
18th International Conference, ICA3PP 2018, Guangzhou, China, November 15-17, 2018, Proceedings, Part I
Practical Hadoop Ecosystem

Handbook of Research on Big Data Storage and Visualization Techniques
Smart Technologies, Systems and Applications
Fundamentals, Techniques, and Applications
The Definitive Guide
Random Access to Your Planet-Size Data
Volume 5: Advanced Intelligent Systems for Computing Sciences
HBase in Action
Big Data Processing Made Simple
Seven Databases in Seven Weeks
A Guide to Modern Databases and the NoSQL Movement
Big Scientific Data Management
SQL at Any Scale, on Any Storage, in Any Environment
The Definitive Guide
13th Italian Research Conference on Digital Libraries, IRCDL 2017, Modena, Italy, January 26-27, 2017, Revised Selected Papers
20th International Conference, ICWE 2020, Helsinki, Finland, June 9–12, 2020, Proceedings
First International Conference, SmartTech-IC 2019, Quito, Ecuador, December 2-4, 2019, Proceedings
Big Data Systems

CARLA RAMOS

Presto: The Definitive Guide "O'Reilly Media, Inc."
The digital age has presented an exponential growth in the amount of data available to individuals looking to draw conclusions based on given or collected information across industries. Challenges associated with the analysis, security, sharing, storage, and visualization of large and complex data sets continue to plague data scientists and analysts alike as traditional data processing applications struggle to adequately manage big data. The Handbook of Research on Big Data Storage and Visualization Techniques is a critical scholarly resource that explores big data analytics and technologies and their role in developing a broad understanding of issues pertaining to the use of big data in multidisciplinary fields. Featuring coverage on a broad range of topics, such as architecture patterns, programing systems, and computational energy, this publication is geared towards professionals, researchers, and students seeking current research and application topics on the subject.
Cloud Computing for Science and

Engineering "O'Reilly Media, Inc."
There’s a lot of information about big data technologies, but splicing these technologies into an end-to-end enterprise data platform is a daunting task not widely covered. With this practical book, you’ll learn how to build big data infrastructure both on-premises and in the cloud and successfully architect a modern data platform. Ideal for enterprise architects, IT managers, application architects, and data engineers, this book shows you how to overcome the many challenges that emerge during Hadoop projects. You’ll explore the vast landscape of tools available in the Hadoop and big data realm in a thorough technical primer before diving into: Infrastructure: Look at all component layers in a modern data platform, from the server to the data center, to establish a solid foundation for data in your enterprise Platform: Understand aspects of deployment, operation, security, high availability, and disaster recovery, along with everything you need to know to integrate your platform with the rest of your enterprise IT Taking Hadoop to the cloud: Learn the important architectural aspects of running a big data platform in the cloud while maintaining enterprise security and high availability

Programming Pig Springer
Digital Technologies – An overview of concepts, tools and techniques associated with it Preface Digital technologies transformed the people’s day to day life and businesses organizations the way they work. It saved their time, accessibility and cost. It provided new opportunities, new markets, and new products to the organizations. Digital revolution started with the introduction of Computers, Internet, Mobile phones, Social media networks, Cloud computing, Big Data, Internet of Things, 3D Printing, Machine learning, Virtual reality, Natural Language Processing, Block Chain, Artificial Intelligence, Robotics and Quantum Computing. Digital technologies are so dynamic and it becomes difficult to cope up with the pace with which new technologies evolve over time. Currently digital technology is being used in the time of pandemic in many countries to contain the spread through detection, contact tracing and monitoring the people movement using Big Data, Machine learning, mobile applications and Artificial intelligence tools. This book provides an overview of digital technologies and concepts, tools and techniques associated with it and how organizations and business needs to adapt to these technologies and transform their business operations. Editor IJSMI International Journal of Statistics and Medical Informatics
www.ijsmi.com/book.php
https://www.amazon.com/dp/B08K38DBCQ
https://www.amazon.com/dp/B08K3YHW2M

A 360-degree Approach "O'Reilly Media, Inc."
A guide to cloud computing for students, scientists, and engineers, with advice and many hands-on examples. The emergence of powerful, always-on cloud utilities has transformed how consumers interact with information technology, enabling video streaming, intelligent personal assistants, and the sharing of content. Businesses, too, have benefited from the cloud, outsourcing much of their information technology to cloud services. Science, however, has not fully exploited the advantages of the cloud. Could scientific discovery be accelerated if mundane

chores were automated and outsourced to the cloud? Leading computer scientists Ian Foster and Dennis Gannon argue that it can, and in this book offer a guide to cloud computing for students, scientists, and engineers, with advice and many hands-on examples. The book surveys the technology that underpins the cloud, new approaches to technical problems enabled by the cloud, and the concepts required to integrate cloud services into scientific work. It covers managing data in the cloud, and how to program these services; computing in the cloud, from deploying single virtual machines or containers to supporting basic interactive science experiments to gathering clusters of machines to do data analytics; using the cloud as a platform for automating analysis procedures, machine learning, and analyzing streaming data; building your own cloud with open source software; and cloud security. The book is accompanied by a website, Cloud4SciEng.org, that provides a variety of supplementary material, including exercises, lecture slides, and other resources helpful to readers and instructors.

Building Effective Algorithms and Analytics for Hadoop and Other Systems "O'Reilly Media, Inc."

The four-volume set LNCS 11334-11337 constitutes the proceedings of the 18th International Conference on Algorithms and Architectures for Parallel Processing, ICA3PP 2018, held in Guangzhou, China, in November 2018. The 141 full and 50 short papers presented were carefully reviewed and selected from numerous submissions. The papers are organized in topical sections on Distributed and Parallel Computing; High Performance Computing; Big Data and Information Processing; Internet of Things and Cloud Computing; and Security and Privacy in Computing.

Mathematics of Big Data "O'Reilly Media, Inc."

A rigorous and comprehensive textbook covering the major approaches to knowledge graphs, an active and interdisciplinary area within artificial intelligence. The field of knowledge graphs, which allows us to model, process, and derive insights from complex real-world data, has emerged as an active and interdisciplinary area of artificial intelligence over the last decade, drawing on such fields as natural language processing, data mining, and the semantic web. Current projects involve predicting cyberattacks, recommending products, and even gleaning insights from thousands of papers on COVID-19. This textbook offers rigorous and comprehensive coverage of the field. It focuses systematically on the major approaches, both those that have stood the test of time and the latest deep learning methods.

Web Engineering Springer

This book constitutes the proceedings of the 20th International Conference on Web Engineering, ICWE 2020, which was planned to take place in Helsinki, Finland, during June 9-12, 2020. Due to the corona pandemic the conference changed to a virtual format. The total of 24 full and 10 short contributions presented in this volume were carefully reviewed and selected from 78 submissions. The book also contains 4 PhD and 7 demo papers. The papers were organized in topical sections named: User interface technologies; performance of Web technologies; machine learning; testing of Web applications; emotion detection;

location-aware applications; sentiment analysis; open data; liquid Web applications; Web-based learning; PhD symposium; demos and posters.
First BenchCouncil International Symposium, Bench 2018, Seattle, WA, USA, December 10-13, 2018, Revised Selected Papers "O'Reilly Media, Inc."
This book constitutes the refereed proceedings of the 13th Conference on Advanced Computer Architecture, ACA 2020, held in Kunming, China, in August 2020. Due to the COVID-19 pandemic the conference was held online. The 24 revised full papers presented were carefully reviewed and selected from 105 submissions. The papers of this volume are organized in topical sections on: interconnection network, router and network interface architecture; accelerator-based, application-specific and reconfigurable architecture; processor, memory, and storage systems architecture; model, simulation and evaluation of architecture; new trends of technologies and applications.
Distributed Applications and Interoperable Systems Springer
Until now, design patterns for the MapReduce framework have been scattered among various research papers, blogs, and books. This handy guide brings together a unique collection of valuable MapReduce patterns that will save you time and effort regardless of the domain, language, or development framework you’re using. Each pattern is explained in context, with pitfalls and caveats clearly identified to help you avoid common design mistakes when modeling your big data architecture. This book also provides a complete overview of MapReduce that explains its origins and implementations, and why design patterns are so important. All code examples are written for Hadoop. Summarization patterns: get a top-level view by summarizing and grouping data Filtering patterns: view data subsets such as records generated from one user Data organization patterns: reorganize data to work with other systems, or to make MapReduce analysis easier Join patterns: analyze different datasets together to discover interesting relationships Metapatterns: piece together several patterns to solve multi-stage problems, or to perform several analytics in the same job Input and output patterns: customize the way you use Hadoop to load or store data "A clear exposition of MapReduce programs for common data processing patterns—this book is indespensible for anyone using Hadoop." --Tom White, author of Hadoop: The Definitive Guide
Cloud Computing in Remote Sensing IGI Global
Summary Hadoop in Practice, Second Edition provides over 100 tested, instantly useful techniques that will help you conquer big data, using Hadoop. This revised new edition covers changes and new features in the Hadoop core architecture, including MapReduce 2. Brand new chapters cover YARN and integrating Kafka, Impala, and Spark SQL with Hadoop. You'll also get new and updated techniques for Flume, Sqoop, and Mahout, all of which have seen major new versions recently. In short, this is the most practical, up-to-date coverage of Hadoop available anywhere. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Book It's always a good time to upgrade your Hadoop skills! Hadoop in Practice, Second Edition provides a collection of 104 tested,

instantly useful techniques for analyzing real-time streams, moving data securely, machine learning, managing large-scale clusters, and taming big data using Hadoop. This completely revised edition covers changes and new features in Hadoop core, including MapReduce 2 and YARN. You'll pick up hands-on best practices for integrating Spark, Kafka, and Impala with Hadoop, and get new and updated techniques for the latest versions of Flume, Sqoop, and Mahout. In short, this is the most practical, up-to-date coverage of Hadoop available. Readers need to know a programming language like Java and have basic familiarity with Hadoop. What's Inside Thoroughly updated for Hadoop 2 How to write YARN applications Integrate real-time technologies like Storm, Impala, and Spark Predictive analytics using Mahout and RR Readers need to know a programming language like Java and have basic familiarity with Hadoop. About the Author Alex Holmes works on tough big-data problems. He is a software engineer, author, speaker, and blogger specializing in large-scale Hadoop projects. Table of Contents PART 1 BACKGROUND AND FUNDAMENTALS Hadoop in a heartbeat Introduction to YARN PART 2 DATA LOGISTICS Data serialization—working with text and beyond Organizing and optimizing data in HDFS Moving data into and out of Hadoop PART 3 BIG DATA PATTERNS Applying MapReduce patterns to big data Utilizing data structures and algorithms at scale Tuning, debugging, and testing PART 4 BEYOND MAPREDUCE SQL on Hadoop Writing a YARN application

Architecting HBase Applications HBase - The Definitive GuideRandom Access to Your Planet-Size Data

This book constitutes the proceedings of the 17th IFIP International Conference on Distributed Applications and Interoperable Systems, DAIS 2017, held in Neuchâtel, Switzerland, in June 2017. The 11 papers presented together with 4 short papers in this volume were carefully reviewed and selected from 23 submissions. The papers are organized in topical sections on running system efficiently, storing data smartly, roaming in graph, building collaborative services, and making things safe.

Intelligent Systems "O'Reilly Media, Inc."

This book constitutes the refereed proceedings of the First International Conference on Big Scientific Data Management, BigSDM 2018, held in Beijing, Greece, in November/December 2018. The 24 full papers presented together with 7 short papers were carefully reviewed and selected from 86 submissions. The topics involved application cases in the big scientific data management, paradigms for enhancing scientific discovery through big data, data management challenges posed by big scientific data, machine learning methods to facilitate scientific discovery, science platforms and storage systems for large scale scientific applications, data cleansing and quality assurance of science data, and data policies.

Random Access to Your Planet-Size Data "O'Reilly Media, Inc."

Data is getting bigger and more complex by the day, and so are your choices in handling it. Explore some of the most cutting-edge databases available - from a traditional relational database to newer NoSQL approaches - and make informed decisions about challenging data storage problems. This is the only comprehensive

guide to the world of NoSQL databases, with in-depth practical and conceptual introductions to seven different technologies: Redis, Neo4J, CouchDB, MongoDB, HBase, Postgres, and DynamoDB. This second edition includes a new chapter on DynamoDB and updated content for each chapter. While relational databases such as MySQL remain as relevant as ever, the alternative, NoSQL paradigm has opened up new horizons in performance and scalability and changed the way we approach data-centric problems. This book presents the essential concepts behind each database alongside hands-on examples that make each technology come alive. With each database, tackle a real-world problem that highlights the concepts and features that make it shine. Along the way, explore five database models - relational, key/value, columnar, document, and graph - from the perspective of challenges faced by real applications. Learn how MongoDB and CouchDB are strikingly different, make your applications faster with Redis and more connected with Neo4J, build a cluster of HBase servers using cloud services such as Amazon's Elastic MapReduce, and more. This new edition brings a brand new chapter on DynamoDB, updated code samples and exercises, and a more up-to-date account of each database's feature set. Whether you're a programmer building the next big thing, a data scientist seeking solutions to thorny problems, or a technology enthusiast venturing into new territory, you will find something to inspire you in this book. What You Need: You'll need a *nix shell (Mac OS or Linux preferred, Windows users will need Cygwin), Java 6 (or greater), and Ruby 1.8.7 (or greater). Each chapter will list the downloads required for that database.

Distributed Data at Web Scale Apress

Learn how to use, deploy, and maintain Apache Spark with this comprehensive guide, written by the creators of the open-source cluster-computing framework. With an emphasis on improvements and new features in Spark 2.0, authors Bill Chambers and Matei Zaharia break down Spark topics into distinct sections, each with unique goals. You'll explore the basic operations and common functions of Spark's structured APIs, as well as Structured Streaming, a new high-level API for building end-to-end streaming applications. Developers and system administrators will learn the fundamentals of monitoring, tuning, and debugging Spark, and explore machine learning techniques and scenarios for employing MLlib, Spark's scalable machine-learning library. Get a gentle overview of big data and Spark Learn about DataFrames, SQL, and Datasets—Spark's core APIs—through worked examples Dive into Spark's low-level APIs, RDDs, and execution of SQL and DataFrames Understand how Spark runs on a cluster Debug, monitor, and tune Spark clusters and applications Learn the power of Structured Streaming, Spark's stream-processing engine Learn how you can apply MLlib to a variety of problems, including classification or recommendation

13th Conference, ACA 2020, Kunming, China, August 13–15, 2020, Proceedings "O'Reilly Media, Inc."

If you're looking for a scalable storage solution to accommodate a virtually endless amount of data, this book shows you how Apache HBase can fulfill your needs. As the open source implementation of Google's BigTable architecture, HBase scales to billions of rows and millions of

columns, while ensuring that write and read performance remain constant. Many IT executives are asking pointed questions about HBase. This book provides meaningful answers, whether you're evaluating this non-relational database or planning to put it into practice right away. Discover how tight integration with Hadoop makes scalability with HBase easier Distribute large datasets across an inexpensive cluster of commodity servers Access HBase with native Java clients, or with gateway servers providing REST, Avro, or Thrift APIs Get details on HBase's architecture, including the storage format, write-ahead log, background processes, and more Integrate HBase with Hadoop's MapReduce framework for massively parallelized data processing jobs Learn how to tune clusters, design schemas, copy tables, import bulk data, decommission nodes, and many other tasks

First International Conference, BigSDM 2018, Beijing, China, November 30 – December 1, 2018, Revised Selected Papers "O'Reilly Media, Inc."

This book constitutes the refereed proceedings of the First International Symposium on Benchmarking, Measuring, and Optimization, Bench 2018, held in Seattle, WA, USA, in December 2018. The 20 full papers presented were carefully reviewed and selected from 51 submissions. The papers are organized in topical sections named: AI Benchmarking; Cloud; Big Data; Modelling and Prediction; and Algorithm and Implementations.

Advanced Computer Architecture Springer

The era of rapidly progressing technology we live in generates vast amounts of data; however, the challenge exists in understanding how to aggressively monitor and make sense of this data. Without a better understanding of how to collect and manage such large data sets, it becomes increasingly difficult to successfully utilize them. Managing Big Data Integration in the Public Sector is a pivotal reference source for the latest scholarly research on the application of big data analytics in government contexts and identifies various strategies in which big data platforms can generate improvements within that sector. Highlighting issues surrounding data management, current models, and real-world applications, this book is ideally designed for professionals, government agencies, researchers, and non-profit organizations interested in the benefits of big data analytics applied in the public sphere.

Hadoop in Action IGI Global

Describes the features and functions of Apache Hive, the data infrastructure for Hadoop.

Hadoop Application Architectures "O'Reilly Media, Inc."

Hadoop: The Definitive Guide helps you harness the power of your data. Ideal for processing large datasets, the Apache Hadoop framework is an open source implementation of the MapReduce algorithm on which Google built its empire. This comprehensive resource demonstrates how to use Hadoop to build reliable, scalable, distributed systems: programmers will find details for analyzing large datasets, and administrators will learn how to set up and run Hadoop clusters. Complete with case studies that illustrate how Hadoop solves specific problems, this book helps you: Use the Hadoop Distributed File System (HDFS) for storing large datasets, and run distributed

computations over those datasets using MapReduce Become familiar with Hadoop's data and I/O building blocks for compression, data integrity, serialization, and persistence Discover common pitfalls and advanced features for writing real-world MapReduce programs Design, build, and administer a dedicated Hadoop cluster, or run Hadoop in the cloud Use Pig, a high-level query language for large-scale data processing Take advantage of HBase, Hadoop's database for structured and semi-structured data Learn ZooKeeper, a toolkit of coordination primitives for building distributed systems If you have lots of data -- whether it's gigabytes or petabytes -- Hadoop is the perfect solution. Hadoop: The Definitive Guide is the most thorough book available on the subject. "Now you have the opportunity to learn about Hadoop from a master-not only of the technology, but also of common sense and plain talk."--Doug Cutting, Hadoop Founder, Yahoo!

A Guidebook for Successful Development and Design Springer Nature

The first book to present the common mathematical foundations of big data analysis across a range of applications and technologies. Today, the volume, velocity, and variety of data are increasing rapidly across a range of fields, including Internet search, healthcare, finance, social media, wireless devices, and cybersecurity. Indeed, these data are growing at a rate beyond our capacity to analyze them. The tools—including spreadsheets, databases, matrices, and graphs—developed to address this challenge all reflect the need to store and operate on data as whole sets rather than as individual elements. This book presents the common mathematical foundations of these data sets that apply across many applications and technologies. Associative arrays unify and simplify data, allowing readers to look past the differences among the various tools and leverage their mathematical similarities in order to solve the hardest big data challenges. The book first introduces the concept of the associative array in practical terms, presents the associative array manipulation system D4M (Dynamic Distributed Dimensional Data Model), and describes the application of associative arrays to graph analysis and machine learning. It provides a mathematically rigorous definition of associative arrays and describes the properties of associative arrays that arise from this definition. Finally, the book shows how concepts of linearity can be extended to encompass associative arrays. Mathematics of Big Data can be used as a textbook or reference by engineers, scientists, mathematicians, computer scientists, and software engineers who analyze big data.

Related with Hbase The Definitive Guide Random Access To Your Planet Size Data:
• Magna Wave Therapy For Horses : click here