

Best Practices For Hp Vertica Oem Customers

Get command of your organizational Big Data using the power of data science and analytics Key Features A perfect companion to boost your Big Data storing, processing, analyzing skills to help you take informed business decisions Work with the best tools such as Apache Hadoop, R, Python, and Spark for NoSQL platforms to perform massive online analyses Get expert tips on statistical inference, machine learning, mathematical modeling, and data visualization for Big Data Book Description Big Data analytics relates to the strategies used by organizations to collect, organize and analyze large amounts of data to uncover valuable business insights that otherwise cannot be analyzed through traditional systems. Crafting an enterprise-scale cost-efficient Big Data and machine learning solution to uncover insights and value from your organization's data is a challenge. Today, with hundreds of new Big Data systems, machine learning packages and BI Tools, selecting the right combination of technologies is an even greater challenge. This book will help you do that. With the help of this guide, you will be able to bridge the gap between the theoretical world of technology with the practical ground reality of building corporate Big Data and data science platforms. You will get hands-on exposure to Hadoop and Spark, build machine learning dashboards using R and R Shiny, create web-based apps using NoSQL databases such as MongoDB and even learn how to write R code for neural networks. By the end of the book, you will have a very clear and concrete understanding of what Big Data analytics means, how it drives revenues for organizations, and how you can develop your own Big Data analytics solution using different tools and methods articulated in this book. What you will learn - Get a 360-degree view into the world of Big Data, data science and machine learning - Broad range of technical and business Big Data analytics topics that caters to the interests of the technical experts as well as corporate IT executives - Get hands-on experience with industry-standard Big Data and machine learning tools such as Hadoop, Spark, MongoDB, KDB+ and R - Create production-grade machine learning BI Dashboards using R and R Shiny with step-by-step instructions - Learn how to combine open-source Big Data, machine learning and BI Tools to create low-cost business analytics applications - Understand corporate strategies for successful Big Data and data science projects - Go beyond general-purpose analytics to develop cutting-edge Big Data applications using emerging technologies Who this book is for The book is intended for existing and aspiring Big Data professionals who wish to become the go-to person in their organization when it comes to Big Data architecture, analytics, and governance. While no prior knowledge of Big Data or related technologies is assumed, it will be helpful to have some programming experience.

This book constitutes the reviewed proceedings of the first Conference on Performance Evaluation and Benchmarking, TPCTC 2009, held in Lyon, France, August 24-28,2009. The 16 full papers and two keynote papers were carefully selected from 34 submissions. This book considers issues such as appliance, business intelligence, cloud computing, complex event processing, database performance optimizations, green computing, data compression, disaster tolerance and recovery, energy and space efficiency, hardware innovations, high speed data generation, hybrid workloads or operational data warehousing, unstructured data management, software management and maintenance, virtualization and very large memory systems This is a step-by-step guide with an easy-to-follow approach. A variety of diagrams and images support the concepts explored in the book. This book also contains important tips and tricks, thus making it a really useful reference for Vertica administration activities. If you are a Vertica user or DBA who wants to perform basic administration and fine tuning, this book is for you. Prior knowledge of Vertica will help you to understand the concepts better, but is not mandatory. Some intermediate knowledge of RDBMS, SQL, Linux, and scripting languages such as Perl or Python will be helpful.

English Mechanic and Mirror of Science and Art

Practical Big Data Analytics

Summary of World Broadcasts

The Municipal Journal

Readings in Database Systems

From Strategic Planning to Enterprise Integration with Tools, Techniques, NoSQL, and Graph

Get the most out of your data by getting more out of Tableau Tableau Your Data! shows you how to build dynamic, best of breed visualizations using the Tableau Software toolset. This comprehensive guide covers the core feature set for data analytics, and provides clear step-by-step guidance toward best practices and advanced techniques that go way beyond the user manual. You'll learn how Tableau is different from traditional business information analysis tools, and how to navigate your way around the Tableau 9.0 desktop before delving into functions and calculations, as well as sharing with the Tableau Server. Coverage includes settings customization, data security, scaling, syntax, and more, with plenty of examples that simplify advanced techniques. Use cases demonstrate how Tableau is applied throughout the enterprise, so you can utilize these analysis tools across sales, marketing, operations, financials, and much more. The companion website features actual working models of the book's visualizations, plus a host of useful links to web-based resources that can help you customize your Tableau experience. Tableau is designed specifically to provide fast and easy visual analytics. The intuitive drag-and-drop interface helps you create interactive reports, dashboards, and visualizations, all without any special or advanced training. This book is your Tableau companion, helping you get the most out of this invaluable business toolset. Analyze data more effectively with Tableau Desktop Deploy visualizations to consumers throughout the enterprise Understand Tableau functions and calculations Leverage Tableau across every link in the value chain You need to make sense of your data before you can use it effectively to make good business decisions. Tableau helps you unlock the stories within the numbers, and Tableau Your Data! puts the software's full functionality right at your fingertips.

This fully revised, self-paced learning tool lays out all the necessary steps to quickly and easily start writing SQL programs Thoroughly updated to reflect the most recent ANSI/ISO standard, SQL: A Beginner's Guide, Fourth Edition will get you up-and-running with SQL

programming right away. Clear tutorials, annotated code, and proven instructional tools guide you to easily performing queries and modifications, building databases, creating and reviewing embedded statements, troubleshooting system- and data-related problems, and much more. You will learn how to retrieve, insert, update, and delete database data, and perform management and administrative functions. The book also covers new features, including SQL/XML and the long-awaited temporal support. Code examples are provided throughout along with notes on using them with the latest RDBMS software versions such as MySQL 5.7, SQL Server 2014, and Oracle Database 12c. Platform-neutral coverage; all skills can be applied to any database product, and any SQL version Features hands-on exercises and self-tests that reinforce basic knowledge "Ask the Expert" sections throughout are filled with bonus information and useful tips

Until recently, Hadoop deployments existed on hardware owned and run by organizations. Now, of course, you can acquire the computing resources and network connectivity to run Hadoop clusters in the cloud. But there's a lot more to deploying Hadoop to the public cloud than simply renting machines. This hands-on guide shows developers and systems administrators familiar with Hadoop how to install, use, and manage cloud-born clusters efficiently. You'll learn how to architect clusters that work with cloud-provider features—not just to avoid pitfalls, but also to take full advantage of these services. You'll also compare the Amazon, Google, and Microsoft clouds, and learn how to set up clusters in each of them. Learn how Hadoop clusters run in the cloud, the problems they can help you solve, and their potential drawbacks Examine the common concepts of cloud providers, including compute capabilities, networking and security, and storage Build a functional Hadoop cluster on cloud infrastructure, and learn what the major providers require Explore use cases for high availability, relational data with Hive, and complex analytics with Spark Get patterns and practices for running cloud clusters, from designing for price and security to dealing with maintenance

Performance Evaluation and Benchmarking

Engineering News

Colliery Guardian

Colliery Guardian, and Journal of the Coal and Iron Trades

The Journal of the Industrial Designers Society of America

Cloud Computing

The big data era is upon us: data are being generated, analyzed, and used at an unprecedented scale, and data-driven decision making is sweeping through all aspects of society. Since the value of data explodes when it can be linked and fused with other data, addressing the big data integration (BDI) challenge is critical to realizing the promise of big data. BDI differs from traditional data integration along the dimensions of volume, velocity, variety, and veracity. First, not only can data sources contain a huge volume of data, but also the number of data sources is now in the millions. Second, because of the rate at which newly collected data are made available, many of the data sources are very dynamic, and the number of data sources is also rapidly exploding. Third, data sources are extremely heterogeneous in their structure and content, exhibiting considerable variety even for substantially similar entities. Fourth, the data sources are of widely differing qualities, with significant differences in the coverage, accuracy and timeliness of data provided. This book explores the progress that has been made by the data integration community on the topics of schema alignment, record linkage and data fusion in addressing these novel challenges faced by big data integration. Each of these topics is covered in a systematic way: first starting with a quick tour of the topic in the context of traditional data integration, followed by a detailed, example-driven exposition of recent innovative techniques that have been proposed to address the BDI challenges of volume, velocity, variety, and veracity. Finally, it presents merging topics and opportunities that are specific to BDI, identifying promising directions for the data integration community.

One issue each year consists of an annual conference review.

The latest edition of a popular text and reference on database research, with substantial new material and revision; covers classical literature and recent hot topics. Lessons from database research have been applied in academic fields ranging from bioinformatics to next-generation Internet architecture and in industrial uses including Web-based e-commerce and search engines. The core ideas in the field have become increasingly influential. This text provides both students and professionals with a grounding in database research and a technical context for understanding recent innovations in the field. The readings included treat the most important issues in the database area--the basic material for any DBMS professional. This fourth edition has been substantially updated and revised, with 21 of the 48 papers new to the edition, four of them published for the first time. Many of the sections have been newly organized, and each section includes a new or substantially revised introduction that discusses the context, motivation, and controversies in a particular area, placing it in the broader perspective of database research. Two introductory articles, never before published, provide an organized, current introduction to basic knowledge of the field; one discusses the history of data models and query languages and the other offers an architectural overview of a database system. The remaining articles range from the classical literature on database research to treatments of current hot topics, including a paper on search engine architecture and a paper on application servers, both written expressly for this edition. The result is a collection of papers that are seminal and also accessible to a reader who has a basic familiarity with database systems.

Databases A Beginner's Guide

Fast and Easy Visual Analysis with Tableau Software

HP Vertica Essentials

The Pragmatic Wisdom of Michael Stonebraker

Engineering and Mining Journal

Big Data Beyond the Hype

Make your data work for you! Tableau For Dummies brings order to the chaotic world of data. Understanding your data and organizing it into formats and visualizations that make sense to you are crucial to making a real impact on your business with the information that's already at your fingertips. This easy-to-use reference explores the user interface, and guides you through the process of connecting your data sources to the software. Additionally, this approachable, yet comprehensive text shows you how to use graphs, charts, and other images to bring visual interest to your data, how to create dashboards from multiple data sources, and how to export the visualizations that you have developed into multiple formats that translate into positive change for your business. The mission of Tableau Software is to grant you access to data that, when put into action, will help you build your company. Learning to use the data available to you helps you make informed, grounded business decisions that can spell success for your company. Navigate the user interface to efficiently access the features you need Connect to various spreadsheets, databases, and other data sources to create a multi-dimensional snapshot of your business Develop visualizations with easy to use drag and drop features

Start building your data with templates and sample workbooks to spark your creativity and help you organize your information Tableau For Dummies is a step-by-step resource that helps you make sense of the data landscape—and put your data to work in support of your business.

Build your own cloud based Database as a Service using OpenStack Trove About This Book Familiarize yourself with the concept of Database as a Service and make your existing system scalable and efficient with OpenStack Trove Minimize the administrative tasks and complexities of managing your cloud infrastructure This is a fast-paced guide to datastore management on the OpenStack platform using OpenStack Trove Who This Book Is For If you are a DBA / system administrator / architect, or a student who wants to build a Database as a Service based on OpenStack, this book is for you. You should have a basic knowledge of OpenStack components, RDBMS/NoSQL, IaaS, and cloud computing. What You Will Learn Get to grips with the basics of OpenStack and the prerequisites to install Trove Understand the expectations of DBaaS and how Trove can help you achieve them Set up a basic installation of DevStack (Development Stack) in a virtual box Install Trove and utilize its configuration groups to manage and tune databases Use Image builder to create guest images for Trove Utilize Trove to provision your first database instance Back up and restore your databases with the help of Trove In Detail OpenStack has become an extremely popular solution to build public and private clouds with. Database as a Service (DBaaS) enables the delivery of more agile database services at lower costs. Some other benefits of DBaaS are secure database deployments and compliance to standards and best practices. Trove is a DBaaS built on OpenStack and is becoming more popular by the day. Since Trove is one of the most recent projects of OpenStack, DBAs and system administrators can find it difficult to set up and run a DBaaS using OpenStack Trove. This book helps DBAs make that step. We start by introducing you to the concepts of DBaaS and how is it implemented using OpenStack Trove. Following this, we look at implementing OpenStack and deploying Trove. Moving on, you will learn to create guest images to be used with Trove. We then look at how to provision databases in self-service mode, and how to perform administration tasks such as backup and recovery, and fine-tuning databases. At the end of the book, we will examine some advanced features of Trove such as replication. Style and approach This fast-paced, step-by-step guide introduces you to DBaaS, OpenStack Trove, and its components, leading you through building your own Cloud-based DBaaS. Using the DevStack deployment method, you will spend less time on installing OpenStack so you can devote more time to learning how to provision and manage databases in a DBaaS environment.

This practical book covers both strategies and tactics around managing a data governance initiative to help make the most of your data.

Tableau Your Data!

Making Databases Work

Hands-on techniques to implement enterprise analytics and machine learning using Hadoop, Spark, NoSQL and R Power

The Electrical Journal

The Aeroplane

Big Data Analytics will assist managers in providing an overview of the drivers for introducing big data technology into the organization and for understanding the types of business problems best suited to big data analytics solutions, understanding the value drivers and benefits, strategic planning, developing a pilot, and eventually planning to integrate back into production within the enterprise. Guides the reader in assessing the opportunities and value proposition Overview of big data hardware and software architectures Presents a variety of technologies and how they fit into the big data ecosystem

HP Vertica EssentialsPackt Pub Limited

This book celebrates Michael Stonebraker's accomplishments that led to his 2014 ACM A.M. Turing Award "for fundamental contributions to the concepts and practices underlying modern database systems." The book describes, for the broad computing community, the unique nature, significance, and impact of Mike's achievements in advancing modern database systems over more than forty years. Today, data is considered the world's most valuable resource, whether it is in the tens of millions of databases used to manage the world's businesses and governments, in the billions of databases in our smartphones and watches, or residing elsewhere, as yet unmanaged, awaiting the elusive next generation of database systems. Every one of the millions or billions of databases includes features that are celebrated by the 2014 Turing Award and are described in this book. Why should I care about databases? What is a database? What is data management? What is a database management system (DBMS)? These are just some of the questions that this book answers, in describing the development of data management through the achievements of Mike Stonebraker and his over 200 collaborators. In reading the stories in this book, you will discover core data management concepts that were developed over the two greatest eras (so far) of data management technology. The book is a collection of 36 stories written by Mike and 38 of his collaborators: 23 world-leading database researchers, 11 world-class systems engineers, and 4 business partners. If you are an aspiring researcher, engineer, or entrepreneur you might read these stories to find these turning points as practice to tilt at your own computer-science windmills, to spur yourself to your next step of innovation and achievement.

The Electrician

The Far East. Weekly supplement

OpenStack Trove Essentials

Theory and Practice

Transit Journal

Big Data For Dummies

Popular Science gives our readers the information and tools to improve their technology and their world. The core belief that we and our readers share: The future is going to be better, and science and technology are the driving forces that will help make it. Big Data in a nutshell: It is the ability to retain, process, and understand data like never before. It can mean more data than we have today; but it can also mean different kinds of data, a venture into the unstructured world where most of today's data reside. You will learn how cognitive computing systems, like IBM Watson, fit into the Big Data world. Learn about the concept of data-in-motion. InfoSphere Streams, the world's fastest and most flexible platform for streaming data. Capturing, storing, refining, transforming, securing, and analyzing data are important topics also covered in this book.

Essential Database Skills--Made Easy! Learn standard database design and management techniques applicable to any type of database. Featuring clear examples using both Microsoft Access and Oracle, Databases: A Beginner's Guide begins by showing you how to use Structured Query Language (SQL) to create and access database objects. Then, you'll discover how to implement logical design

normalization, transform the logical design into a physical database, and handle data and process modeling. You'll also get details on security, online analytical processing (OLAP), connecting databases to applications, and integrating XML and object content into applications. Designed for Easy Learning Key Skills & Concepts--Chapter-opening lists of specific skills covered in the chapter Ask the Experts sections filled with bonus information and helpful tips Try This--Hands-on exercises that show you how to apply your skills Nuts and Bolts sections filled with information related to the topic being covered Self Tests--Chapter-ending quizzes to test your knowledge

The Data Governance Imperative
Popular Science

Tableau For Dummies

Transaction Processing Performance Council Technology Conference, TPCTC 2009, Lyon, France, August 24-28, 2009, Revised Edition
Papers

The Gas World

Find the right big data solution for your business or organization *Big data management is one of the major challenges facing business, industry, and not-for-profit organizations. Data sets such as customer transactions for a mega-retailer, weather patterns monitored by meteorologists, or social network activity can quickly outpace the capacity of traditional data management tools. If you need to develop or manage big data solutions, you'll appreciate how these four experts define, explain, and guide you through this new and often confusing concept. You'll learn what it is, why it matters, and how to choose and implement solutions that work. Effectively managing big data is an issue of growing importance to businesses, not-for-profit organizations, government, and IT professionals* **Authors are experts in information management, big data, and a variety of solutions** *Explains big data in detail and discusses how to select and implement a solution, security concerns to consider, data storage and presentation issues, analytics, and much more* **Provides essential information in a no-nonsense, easy-to-understand style that is empowering** **Big Data For Dummies cuts through the confusion and helps you take charge of big data solutions for your organization.**

The Design and Implementation of Modern Column-Oriented Database Systems *discusses modern column-stores, their architecture and evolution as well the benefits they can bring in data analytics.*

Cloud Computing: Theory and Practice *provides students and IT professionals with an in-depth analysis of the cloud from the ground up. Beginning with a discussion of parallel computing and architectures and distributed systems, the book turns to contemporary cloud infrastructures, how they are being deployed at leading companies such as Amazon, Google and Apple, and how they can be applied in fields such as healthcare, banking and science. The volume also examines how to successfully deploy a cloud application across the enterprise using virtualization, resource management and the right amount of networking support, including content delivery networks and storage area networks. Developers will find a complete introduction to application development provided on a variety of platforms. Learn about recent trends in cloud computing in critical areas such as: resource management, security, energy consumption, ethics, and complex systems* **Get a detailed hands-on set of practical recipes that help simplify the deployment of a cloud based system for practical use of computing clouds along with an in-depth discussion of several projects** **Understand the evolution of cloud computing and why the cloud computing paradigm has a better chance to succeed than previous efforts in large-scale distributed computing**

Innovation

Big Data Integration

SQL: A Beginner's Guide, Fourth Edition

With Catalogue of the Manufactures of the Babcock & Wilcox Co. ...

Scientific American

The Street Railway Review