

Netezza System Administrators Guide

This book constitutes the refereed proceedings of the 15th International Conference on Advances in Databases and Information Systems, ADBIS 2011, held in Vienna, Austria, in September 2011. The 30 revised full papers presented together with 2 full length invited talks were carefully reviewed and selected from 105 submissions. They are organized in topical sections on query processing; data warehousing; DB systems; spatial data; information systems; physical DB design; evolution, integrity, security; and data semantics.

Mastering material for dealing with DBA certification exams Key Features Prepare yourself for the IBM C2090-600 certification exam Cover over 50 Db2 procedures including database design, performance, and security Work through over 150 Q&As to gain confidence on each topic Book Description IBM Db2 is a relational database management system (RDBMS) that helps you store, analyze, and retrieve data efficiently. This comprehensive book is designed to help you master all aspects of IBM Db2 database administration and prepare you to take and pass IBM's Certification Exams C2090-600. Building on years of extensive experience, the authors take you through all areas covered by the test. The book delves deep into each certification topic: Db2 server management, physical design, business rules implementation, activity monitoring, utilities, high availability, and security. IBM Db2 11.1 Certification Guide provides you with more than 150 practice questions and answers, simulating real certification examination questions. Each chapter includes an extensive set of practice questions along with carefully explained answers. This book will not just prepare you for the C2090-600 exam but also help you troubleshoot day-to-day database administration challenges. What you will learn Configure and manage Db2 servers, instances, and databases Implement Db2 BLU Acceleration and a DB2 pureScale environment Create, manage, and alter Db2 database objects Use the partitioning capabilities available within Db2 Enforce constraint checking with the SET INTEGRITY command Utilize the Db2 problem determination (db2pd) and dsmtop tools Configure and manage HADR Understand how to encrypt data in transit and at rest Who this book is for The IBM Db2 11.1 Certification Guide is an excellent choice for database administrators, architects, and application developers who are keen to obtain certification in Db2. Basic understanding of Db2 is expected in order to get the most out of this guide.

The big tech companies are increasingly relying on the database management systems to store and maintain the massive volume of data generated by our digital lives. The Relational Database Management System (RDBMS) is extensively used by these tech giants to not only store the large volume of data but as an advanced tool to gain insight from massive volume of data generated by our increasingly digital lives. The Structured Query Language (SQL) is the language of choice to define, manipulate, control and query the data within a RDBMS. This book is written to serve as your personal guide so you can efficiently and effectively learn and write SQL statements or queries to retrieve from and update data on relational databases such as MySQL. You will be able to install the free and open MySQL user interface with the instructions provided in this book. This will allow you to get hands-on practice utilizing a variety of exercises included in this book, so you will be able to create not only correct but efficient SQL queries to succeed at work and ace those job interview questions. Some of the highlights of this book are: - Foundational concepts of SQL language as well as 5 fundamental types of SQL queries namely - Learn the thumb rules for building SQL syntax or query - A variety of SQL data types that are a pre-requisite for learning SQL - Overview of a wide range of user interfaces available with MySQL servers - Learn how to create an effective database on the MySQL server - Learn the concept of temporary tables, derived tables and how you can create a new table from an existing one - Learn how to create new user accounts, update the user password as needed, grant and revoke access privileges - Learn CREATE VIEW, MERGE, TEMPTABLE, UNDEFINED, Updatable SQL Views and ALTER VIEW - The properties of SQL transactions as well as various SQL transaction statements with controlling clauses Don't miss the opportunity to quickly learn a programming language like SQL. Don't you think it can be that easy? If you really want to have proof of all this, don't waste any more time! Grab your copy now!

The Installation and Configuration Guide includes information to install and configure MicroStrategy products on Windows, UNIX, Linux, and HP platforms, as well as basic maintenance guidelines.

IBM® InfoSphere® Guardium® provides the simplest, most robust solution for data security and data privacy by assuring the integrity of trusted information in your data center. InfoSphere Guardium helps you reduce support costs by automating the entire compliance auditing process across heterogeneous environments. InfoSphere Guardium offers a flexible and scalable solution to support varying customer architecture requirements. This IBM Redbooks® publication provides a guide for deploying the Guardium solutions. This book also provides a roadmap process for implementing an InfoSphere Guardium solution that is based on years of experience and best practices that were collected from various Guardium experts. We describe planning, installation, configuration, monitoring, and administering an InfoSphere Guardium environment. We also describe use cases and how InfoSphere Guardium integrates with other IBM products. The guidance can help you successfully deploy and manage an IBM InfoSphere Guardium system. This book is intended for the system administrators and support staff who are responsible for deploying or supporting an InfoSphere Guardium environment.

Learn all you need to know about seven key innovations disrupting business analytics today. These innovations—the open source business model, cloud analytics, the Hadoop ecosystem, Spark and in-memory analytics, streaming analytics, Deep Learning, and self-service analytics—are radically changing how businesses use data for competitive advantage. Taken together, they are disrupting the business analytics value chain, creating new opportunities. Enterprises who seize the opportunity will thrive and prosper, while others struggle and decline: disrupt or be disrupted. Disruptive Business Analytics provides strategies to profit from disruption. It shows you how to organize for insight, build and provision an

open source stack, how to practice lean data warehousing, and how to assimilate disruptive innovations into an organization. Through a short history of business analytics and a detailed survey of products and services, analytics authority Thomas W. Dinsmore provides a practical explanation of the most compelling innovations available today. What You'll Learn Discover how the open source business model works and how to make it work for you See how cloud computing completely changes the economics of analytics Harness the power of Hadoop and its ecosystem Find out why Apache Spark is everywhere Discover the potential of streaming and real-time analytics Learn what Deep Learning can do and why it matters See how self-service analytics can change the way organizations do business Who This Book Is For Corporate actors at all levels of responsibility for analytics: analysts, CIOs, CTOs, strategic decision makers, managers, systems architects, technical marketers, product developers, IT personnel, and consultants.

Payment fraud can be defined as an intentional deception or misrepresentation that is designed to result in an unauthorized benefit. Fraud schemes are becoming more complex and difficult to identify. It is estimated that industries lose nearly \$1 trillion USD annually because of fraud. The ideal solution is where you avoid making fraudulent payments without slowing down legitimate payments. This solution requires that you adopt a comprehensive fraud business architecture that applies predictive analytics. This IBM® Redbooks® publication begins with the business process flows of several industries, such as banking, property/casualty insurance, and tax revenue, where payment fraud is a significant problem. This book then shows how to incorporate technological advancements that help you move from a post-payment to pre-payment fraud detection architecture. Subsequent chapters describe a solution that is specific to the banking industry that can be easily extrapolated to other industries. This book describes the benefits of doing fraud detection on IBM System z®. This book is intended for financial decisionmakers, consultants, and architects, in addition to IT administrators.

IBM® Watson™ Content Analytics (Content Analytics) Version 3.0 (formerly known as IBM Content Analytics with Enterprise Search (ICAwES)) helps you to unlock the value of unstructured content to gain new actionable business insight and provides the enterprise search capability all in one product. Content Analytics comes with a set of tools and a robust user interface to empower you to better identify new revenue opportunities, improve customer satisfaction, detect problems early, and improve products, services, and offerings. To help you gain the most benefits from your unstructured content, this IBM Redbooks® publication provides in-depth information about the features and capabilities of Content Analytics, how the content analytics works, and how to perform effective and efficient content analytics on your content to discover actionable business insights. This book covers key concepts in content analytics, such as facets, frequency, deviation, correlation, trend, and sentimental analysis. It describes the content analytics miner, and guides you on performing content analytics using views, dictionary lookup, and customization. The book also covers using IBM Content Analytics Studio for domain-specific content analytics, integrating with IBM Content Classification to get categories and new metadata, and interfacing with IBM Cognos® Business Intelligence (BI) to add values in BI reporting and analysis, and customizing the content analytics miner with APIs. In addition, the book describes how to use the enterprise search capability for the discovery and retrieval of documents using various query and visual navigation techniques, and customization of crawling, parsing, indexing, and runtime search to improve search results. The target audience of this book is decision makers, business users, and IT architects and specialists who want to understand and analyze their enterprise content to improve and enhance their business operations. It is also intended as a technical how-to guide for use with the online IBM Knowledge Center for configuring and performing content analytics and enterprise search with Content Analytics.

Today's global organizations depend on being able to unlock business insights from massive volumes of data. Now, with IBM® FlashSystem 900, powered by IBM FlashCore™ technology, they can make faster decisions based on real-time insights and unleash the power of the most demanding applications, including online transaction processing (OLTP) and analytics databases, virtual desktop infrastructures (VDIs), technical computing applications, and cloud environments. This IBM Redbooks® publication introduces clients to the IBM FlashSystem® 900. It provides in-depth knowledge of the product architecture, software and hardware, implementation, and hints and tips. Also illustrated are use cases that show real-world solutions for tiering, flash-only, and preferred-read, and also examples of the benefits gained by integrating the FlashSystem storage into business environments. This book is intended for pre-sales and post-sales technical support professionals and storage administrators, and for anyone who wants to understand how to implement this new and exciting technology. This book describes the following offerings of the IBM Spectrum™ Storage family: IBM Spectrum Storage™ IBM Spectrum Control™ IBM Spectrum Virtualize™ IBM Spectrum Scale™ IBM Spectrum Accelerate™

The IBM® DB2® Analytics Accelerator for IBM z/OS® is a high-performance appliance that integrates the IBM zEnterprise® infrastructure with IBM PureData™ for Analytics, powered by IBM Netezza® technology. With this integration, you can accelerate data-intensive and complex queries in a DB2 for z/OS highly secure and available environment. DB2 and the Analytics Accelerator appliance form a self-managing hybrid environment running online transaction processing and online transactional analytical processing concurrently and efficiently. These online transactions run together with business intelligence and online analytic processing workloads. DB2 Analytics Accelerator V4.1 expands the value of high-performance analytics. DB2 Analytics Accelerator V4.1 opens to static Structured Query Language (SQL) applications and row set processing, minimizes data movement, reduces latency, and improves availability. This IBM Redbooks® publication provides technical decision-makers with an understanding of the benefits of version 4.1 of the Analytics Accelerator with DB2 11 for z/OS. It describes the installation of the new functions, and the advantages to existing analytical processes as measured in our test environment. This book also introduces the DB2 Analytics Accelerator Loader V1.1, a tool that facilitates the data population of the DB2 Analytics Accelerator. This supplement describes the SAS/ACCESS features and options that are available for Netezza. When you use this supplement along with SAS/ACCESS Interface to Relational Databases: Reference (available separately), you have a comprehensive guide to SAS/ACCESS software for your database management system (DBMS). Also available online, this supplement explains how SAS/ACCESS software is customized to accommodate the particular requirements and capabilities of your DBMS. It is designed

for users of SAS/ACCESS software who need specific information about Netezza.

Big Data in a nutshell: It is the ability to retain, process, and understand data like never before. It can mean more data than what you are using today; but it can also mean different kinds of data, a venture into the unstructured world where most of today's data resides. In this book you will learn how cognitive computing systems, like IBM Watson, fit into the Big Data world. Learn about the concept of data-in-motion and InfoSphere Streams, the world's fastest and most flexible platform for streaming data. Capturing, storing, refining, transforming, governing, securing, and analyzing data are important topics also covered in this book.

Explore the modern market of data analytics platforms and the benefits of using Snowflake computing, the data warehouse built for the cloud. With the rise of cloud technologies, organizations prefer to deploy their analytics using cloud providers such as Amazon Web Services (AWS), Microsoft Azure, or Google Cloud Platform. Cloud vendors are offering modern data platforms for building cloud analytics solutions to collect data and consolidate into single storage solutions that provide insights for business users. The core of any analytics framework is the data warehouse, and previously customers did not have many choices of platform to use. Snowflake was built specifically for the cloud and it is a true game changer for the analytics market. This book will help onboard you to Snowflake, present best practices to deploy, and use the Snowflake data warehouse. In addition, it covers modern analytics architecture and use cases. It provides use cases of integration with leading analytics software such as Matillion ETL, Tableau, and Databricks. Finally, it covers migration scenarios for on-premise legacy data warehouses. **What You Will Learn** Know the key functionalities of Snowflake Set up security and access with cluster Bulk load data into Snowflake using the COPY command Migrate from a legacy data warehouse to Snowflake integrate the Snowflake data platform with modern business intelligence (BI) and data integration tools Who This Book Is For Those working with data warehouse and business intelligence (BI) technologies, and existing and potential Snowflake users

Many companies move workloads to the cloud only to encounter issues with legacy processes and organizational structures. How do you design new operating models for this environment? This practical book shows IT managers, CIOs, and CTOs how to address the hardest part of any cloud transformation: the people and the processes. Author Mike Kavis (Architecting the Cloud) explores lessons learned from enterprises in the midst of cloud transformations. You'll learn how to rethink your approach from a technology, process, and organizational standpoint to realize the promise of cost optimization, agility, and innovation that public cloud platforms provide. Learn the difference between working in a data center and operating in the cloud Explore patterns and anti-patterns for organizing cloud operating models Get best practices for making the organizational change required for a move to the cloud Understand why site reliability engineering is essential for cloud operations Improve organizational performance through value stream mapping

If you've been asked to maintain large and complex Hadoop clusters, this book is a must. Demand for operations-specific material has skyrocketed now that Hadoop is becoming the de facto standard for truly large-scale data processing in the data center. Eric Sammer, Principal Solution Architect at Cloudera, shows you the particulars of running Hadoop in production, from planning, installing, and configuring the system to providing ongoing maintenance. Rather than run through all possible scenarios, this pragmatic operations guide calls out what works, as demonstrated in critical deployments. Get a high-level overview of HDFS and MapReduce: why they exist and how they work Plan a Hadoop deployment, from hardware and OS selection to network requirements Learn setup and configuration details with a list of critical properties Manage resources by sharing a cluster across multiple groups Get a runbook of the most common cluster maintenance tasks Monitor Hadoop clusters—and learn troubleshooting with the help of real-world war stories Use basic tools and techniques to handle backup and catastrophic failure

This IBM® Redbooks® publication describes how the IBM Big Data Platform provides the integrated capabilities that are required for the adoption of Information Governance in the big data landscape. As organizations embark on new use cases, such as Big Data Exploration, an enhanced 360 view of customers, or Data Warehouse modernization, and absorb ever growing volumes and variety of data with accelerating velocity, the principles and practices of Information Governance become ever more critical to ensure trust in data and help organizations overcome the inherent risks and achieve the wanted value. The introduction of big data changes the information landscape. Data arrives faster than humans can react to it, and issues can quickly escalate into significant events. The variety of data now poses new privacy and security risks. The high volume of information in all places makes it harder to find where these issues, risks, and even useful information to drive new value and revenue are. Information Governance provides an organization with a framework that can align their wanted outcomes with their strategic management principles, the people who can implement those principles, and the architecture and platform that are needed to support the big data use cases. The IBM Big Data Platform, coupled with a framework for Information Governance, provides an approach to build, manage, and gain significant value from the big data landscape.

The Implementing MicroStrategy: Development and Deployment course provides an overview of the stages involved in developing, implementing, and maintaining a business intelligence project. You will first get an intensive, yet high-level overview of the project design and report creation processes, followed by the document and dashboard creation basics. The course also covers deployment to MicroStrategy Web™ and MicroStrategy Mobile™, as well as administration and maintenance of MicroStrategy environment.

The MicroStrategy Analytics Desktop User Guide describes the steps for a business analyst to execute and analyze a Visual Insight dashboard in MicroStrategy Analytics Desktop. It provides the information necessary for a business analyst to import data from a data source, and use that data to create and modify dashboards.

For more than 40 years, Computerworld has been the leading source of technology news and information for IT influencers worldwide. Computerworld's award-winning Web site (Computerworld.com), twice-monthly publication, focused conference series and custom research form the hub of the world's largest global IT media network.

Provides a point of entry for understanding the basics of the SAS Intelligence Platform. It discusses the benefits of the SAS Intelligence Platform to businesses, describes the architecture, and provides an overview of each software component. This document can be used alone or as an introduction to the deployment and administration guides.

Geared to IT professionals eager to get into the all-important field of data warehousing, this book explores all topics needed by those who design and implement data warehouses. Readers will learn about planning requirements, architecture, infrastructure, data preparation, information delivery, implementation, and maintenance. They'll also find a wealth of industry examples garnered from the author's 25 years of experience in designing and implementing databases and data warehouse applications for major corporations. Market: IT Professionals, Consultants.

Typical deployment architectures introduce challenges to fully using the shared metadata platform across products, environments, and servers. Data privacy and information security requirements add even more levels of complexity. IBM® InfoSphere® Information Server provides a comprehensive, metadata-driven platform for delivering trusted information across heterogeneous systems. This IBM Redbooks® publication presents guidelines and criteria for the successful deployment of InfoSphere Information Server components in typical logical infrastructure topologies that use shared metadata capabilities of the platform, and support development lifecycle, data privacy, information security, high availability, and performance requirements. This book can help you evaluate information requirements to determine an appropriate deployment architecture, based on guidelines that are presented here, and that can fulfill specific use cases. It can also help you effectively use the functionality of your Information Server product modules and components to successfully achieve your business goals. This book is for IT architects, information management and integration specialists, and system administrators who are responsible for delivering the full suite of information integration capabilities of InfoSphere Information Server.

The IBM® DB2® Analytics Accelerator Version 2.1 for IBM z/OS® (also called DB2 Analytics Accelerator or Query Accelerator in this book and in DB2 for z/OS documentation) is a marriage of the IBM System z® Quality of Service and Netezza® technology to accelerate complex queries in a DB2 for z/OS highly secure and available environment. Superior performance and scalability with rapid appliance deployment provide an ideal solution for complex analysis. This IBM Redbooks® publication provides technical decision-makers with a broad understanding of the IBM DB2 Analytics Accelerator architecture and its exploitation by documenting the steps for the installation of this solution in an existing DB2 10 for z/OS environment. In this book we define a business analytics scenario, evaluate the potential benefits of the DB2 Analytics Accelerator appliance, describe the installation and integration steps with the DB2 environment, evaluate performance, and show the advantages to existing business intelligence processes.

This IBM® Redpaper™ publication provides suggestions, hints and tips, directions, installation steps, checklists of prerequisites, and configuration information collected from several IBM InfoSphere® Information Server experts. It is intended to minimize the time required to successfully install and configure InfoSphere Information Server. The information in this document is based on field experiences of experts who have implemented InfoSphere Information Server. As such, it is intended to supplement, and not replace, the product documentation. Discover the proven choices and combinations for installing InfoSphere Information Server that have been the most successful for the IBM InfoSphere Center Of Excellence. This paper includes a broad range of customer needs and experiences, with a focus on the following areas: InfoSphere Information Server architecture Checklists Prerequisites Configuration choices that work well together This paper is based on thousands of hours of production systems experience, from which you can now reap significant benefits.

This IBM® Redbooks® publication describes IBM PureApplication™ System preferred practices that are based on IBM client and Business Partner experience. It explains how PureApplication System enables industries to consolidate workloads, increase efficiency, automate routine processes, reduce costs, and become more agile to respond to continually changing business needs. This book is particularly useful to solution specialists, system or software architects, and the IT teams who implement PureApplication System cloud services.

This IBM® Redbooks® publication explains how IBM Cognos® Business Intelligence (BI) administrators, authors, modelers, and power users can use the dynamic query layer effectively. It provides guidance for determining which technology within the dynamic query layer can best satisfy your business requirements. Administrators can learn how to tune the query service effectively and preferred practices for managing their business intelligence content. This book includes information about metadata modeling of relational data sources with IBM Cognos Framework Manager. It includes considerations that can help you author high-performing applications that satisfy analytical requirements of users. This book provides guidance for troubleshooting issues related to the dynamic query layer of Cognos BI. Related documents: Solution Guide : Big Data Analytics with IBM Cognos BI Dynamic Query Blog post : IBM Cognos Dynamic Query Extensibility

In today's infrastructure, it is common to build networks based on 10 Gb Ethernet technology. The IBM® portfolio of 10 Gb systems networking products includes Top-of-Rack switches, and the embedded switches in the IBM BladeCenter® family. In 2010, IBM formed the IBM System Networking business (by acquiring BLADE Network Technologies), which is now focused on driving data center networking by using the latest Ethernet technologies. The main focus of this IBM Redbooks® publication is on the IBM System Networking 10Gb Switch Modules, which include both embedded and Top-of-Rack (TOR) models. After reading this book, you can perform basic to advanced configurations of IBM System Networking 10Gb Switch Modules. In this publication, we introduce the various 10 Gb switch models that are available today and then describe in detail the features that are applicable to these switches. We then present two architectures that use these 10 Gb switches, which are used throughout this book. These designs are based on preferred practices and the experience of authors of this book. Our intention is to show the configuration of the different features that are available with IBM System Networking 10Gb Switch Modules. We follow the three-tier Data Center design, focusing on the Access and Aggregation Layers, because those layers are the layers that IBM System Networking Switches use. Regarding online transaction processing (OLTP) workloads, IBM® z Systems™ platform, with IBM DB2®, data sharing, Workload Manager (WLM), geoplex, and other high-end features, is the widely acknowledged leader. Most customers now integrate business analytics with OLTP by running, for example, scoring functions from transactional context for real-time analytics or by applying machine-learning algorithms on enterprise data that is kept on the mainframe. As a result, IBM adds investment so clients can keep the complete lifecycle for data analysis, modeling, and scoring on z Systems control in a cost-efficient way, keeping the qualities of services in availability, security, reliability that z Systems solutions offer. Because of the changed architecture and tighter integration, IBM has shown, in a customer proof-of-concept, that a particular client was able to achieve an orders-of-magnitude improvement in performance, allowing that client's data scientist to investigate the data in a more interactive process. Open technologies, such as Predictive Model Markup Language (PMML) can help customers update single components instead of being forced to replace everything at once. As a result, you have the possibility to combine your preferred tool for model generation (such as SAS Enterprise Miner

or IBM SPSS® Modeler) with a different technology for model scoring (such as Zementis, a company focused on PMML scoring). IBM SPSS Modeler is a leading data mining workbench that can apply various algorithms in data preparation, cleansing, statistics, visualization, machine learning, and predictive analytics. It has over 20 years of experience and continued development, and is integrated with z Systems. With IBM DB2 Analytics Accelerator 5.1 and SPSS Modeler 17.1, the possibility exists to do the complete predictive model creation including data transformation within DB2 Analytics Accelerator. So, instead of moving the data to a distributed environment, algorithms can be pushed to the data, using cost-efficient DB2 Accelerator for the required resource-intensive operations. This IBM Redbooks® publication explains the overall z Systems architecture, how the components can be installed and customized, how the new IBM DB2 Analytics Accelerator loader can help efficient data loading for z Systems data and external data, how in-database transformation, in-database modeling, and in-transactional real-time scoring can be used, and what other related technologies are available. This book is intended for technical specialists and architects, and data scientists who want to use the technology on the z Systems platform. Most of the technologies described in this book require IBM DB2 for z/OS®. For acceleration of the data investigation, data transformation, and data modeling process, DB2 Analytics Accelerator is required. Most value can be achieved if most of the data already resides on z Systems platforms, although adding external data (like from social sources) poses no problem at all.

Thinking of migrating to PostgreSQL? This updated guide helps you quickly understand and use the 9.3 release of this open source database system. You'll not only learn about its unique enterprise-class features, but also discover that PostgreSQL is more than just a database system—it's also an impressive application platform. Using numerous examples, this book shows you how to achieve tasks that are difficult or impossible in other databases. The second edition covers LATERAL queries, augmented JSON support, materialized views, and other key topics. If you're an existing PostgreSQL user, you'll pick up gems you may have missed along the way. Learn basic administration tasks, such as role management, database creation, backup, and restore Apply the psql command-line utility and the pgAdmin graphical administration tool Explore PostgreSQL tables, constraints, and indexes Learn powerful SQL constructs not generally found in other databases Use several different languages to write database functions Tune your queries to run as fast as your hardware will allow Query external and variegated data sources with Foreign Data Wrappers Learn how to replicate data, using built-in replication features

The IBM® DB2® Analytics Accelerator Version 3.1 for IBM z/OS® (simply called Accelerator in this book) is a union of the IBM System z® quality of service and IBM Netezza® technology to accelerate complex queries in a DB2 for z/OS highly secure and available environment. Superior performance and scalability with rapid appliance deployment provide an ideal solution for complex analysis. In this IBM Redbooks® publication, we provide technical decision-makers with a broad understanding of the benefits of Version 3.1 of the Accelerator's major new functions. We describe their installation and the advantages to existing analytical processes as measured in our test environment. We also describe the IBM zEnterprise® Analytics System 9700, a hybrid System z solution offering that is surrounded by a complete set of optional packs to enable customers to custom tailor the system to their unique needs..

The MicroStrategy Project Design guide gives you information you need to create and modify MicroStrategy projects, and understand facts, attributes, hierarchies, transformations, advanced schemas, and project optimization.

In this IBM® Redbooks® publication, we present guidelines for the development of highly efficient and scalable information integration applications with InfoSphere™ DataStage® (DS) parallel jobs. InfoSphere DataStage is at the core of IBM Information Server, providing components that yield a high degree of freedom. For any particular problem there might be multiple solutions, which tend to be influenced by personal preferences, background, and previous experience. All too often, those solutions yield less than optimal, and non-scalable, implementations. This book includes a comprehensive detailed description of the components available, and descriptions on how to use them to obtain scalable and efficient solutions, for both batch and real-time scenarios. The advice provided in this document is the result of the combined proven experience from a number of expert practitioners in the field of high performance information integration, evolved over several years. This book is intended for IT architects, Information Management specialists, and Information Integration specialists responsible for delivering cost-effective IBM InfoSphere DataStage performance on all platforms.

Let Hadoop For Dummies help harness the power of your data and rein in the information overload Big data has become big business, and companies and organizations of all sizes are struggling to find ways to retrieve valuable information from their massive data sets with becoming overwhelmed. Enter Hadoop and this easy-to-understand For Dummies guide. Hadoop For Dummies helps readers understand the value of big data, make a business case for using Hadoop, navigate the Hadoop ecosystem, and build and manage Hadoop applications and clusters. Explains the origins of Hadoop, its economic benefits, and its functionality and practical applications Helps you find your way around the Hadoop ecosystem, program MapReduce, utilize design patterns, and get your Hadoop cluster up and running quickly and easily Details how to use Hadoop applications for data mining, web analytics and personalization, large-scale text processing, data science, and problem-solving Shows you how to improve the value of your Hadoop cluster, maximize your investment in Hadoop, and avoid common pitfalls when building your Hadoop cluster From programmers challenged with building and maintaining affordable, scaleable data systems to administrators who must deal with huge volumes of information effectively and efficiently, this how-to has something to help you with Hadoop. Using real-world examples and hands-on tasks, Oracle Data Guard 11gR2 Administration Beginner's Guide will give you a solid foundation in Oracle Data Guard. It has been designed to teach you everything you need to know to successfully create and operate Data Guard environments with maximum flexibility, compatibility, and effectiveness.If you are an Oracle database administrator who wants to configure and administer Data Guard configurations, then "Oracle Data Guard 11gR2 Administration Beginner's Guide" is for you. With a basic understanding of Oracle database administration, you'll be able to easily follow the book.

To make better informed business decisions, better serve clients, and increase operational efficiencies, you must be aware of changes to key data as they occur. In addition, you must enable the immediate delivery of this information to the people and processes that need to act upon it. This ability to sense and respond to data changes is fundamental to dynamic warehousing, master data management, and many other key

initiatives. A major challenge in providing this type of environment is determining how to tie all the independent systems together and process the immense data flow requirements. IBM® InfoSphere® Change Data Capture (InfoSphere CDC) can respond to that challenge, providing programming-free data integration, and eliminating redundant data transfer, to minimize the impact on production systems. In this IBM Redbooks® publication, we show you examples of how InfoSphere CDC can be used to implement integrated systems, to keep those systems updated immediately as changes occur, and to use your existing infrastructure and scale up as your workload grows. InfoSphere CDC can also enhance your investment in other software, such as IBM DataStage® and IBM QualityStage®, IBM InfoSphere Warehouse, and IBM InfoSphere Master Data Management Server, enabling real-time and event-driven processes. Enable the integration of your critical data and make it immediately available as your business needs it.

Perform fast interactive analytics against different data sources using the Trino high-performance distributed SQL query engine. With this practical guide, you'll learn how to conduct analytics on data where it lives, whether it's Hive, Cassandra, a relational database, or a proprietary data store. Analysts, software engineers, and production engineers will learn how to manage, use, and even develop with Trino. Initially developed by Facebook, open source Trino is now used by Netflix, Airbnb, LinkedIn, Twitter, Uber, and many other companies. Matt Fuller, Manfred Moser, and Martin Traverso show you how a single Trino query can combine data from multiple sources to allow for analytics across your entire organization. Get started: Explore Trino's use cases and learn about tools that will help you connect to Trino and query data Go deeper: Learn Trino's internal workings, including how to connect to and query data sources with support for SQL statements, operators, functions, and more Put Trino in production: Secure Trino, monitor workloads, tune queries, and connect more applications; learn how other organizations apply Trino

Hybrid Analytics Solution using IBM DB2 Analytics Accelerator for z/OS V3.1 IBM Redbooks

[Copyright: 49201b9afc5393ef94975db9bdc7d00e](https://www.ibm.com/redbooks/copyright/49201b9afc5393ef94975db9bdc7d00e)