There is an easier way to build Hadoop applications. With this hands-on book, you'll learn how to use Cascading, the open source abstraction framework for Hadoop that lets you easily create and manage powerful enterprise-grade data processing applications—without having to learn the intricacies of MapReduce. Working with sample apps based on Java and other JVM languages, you'll quickly learn Cascading's streamlined approach to data processing, data filtering, and workflow optimization. This book demonstrates how this framework can help your business extract meaningful information from large amounts of distributed data. Start working on Cascading example projects right away. Model and analyze unstructured data in any format, from any source. Build and test applications with familiar constructs and reusable components. Work with the Scalding and Cascalog Domain-Specific Languages. Easily deploy applications to Hadoop, regardless of cluster location or data size. Build workflows that integrate several big data frameworks and processes. Explore common use cases for Cascading, including features and tools that support them. Examine a case study that uses a dataset from the Open Data Initiative.

The two-volume set LNCS 10271 and 10272 constitutes the refereed proceedings of the 19th International Conference on Human-Computer Interaction, HCII 2017, held in Vancouver, BC, Canada, in July 2017. The total of 1228 papers presented at the 15 colocated HCII 2017 conferences was carefully reviewed and selected from 4340 submissions. The papers address the latest research and development efforts and highlight the human aspects of design and use of computing systems. They cover the entire field of Human-Computer Interaction, addressing major advances in knowledge and effective use of computers in a variety of application areas. The papers included in this volume cover the following topics: HCI theory and education; HCI, innovation and technology acceptance; interaction design and evaluation methods; user interface development; methods, tools, and architectures; multimodal interaction; and emotions in HCI.

Although you don't need a large computing infrastructure to process massive amounts of data with Apache Hadoop, it can still be difficult to get started. This practical guide shows you how to quickly launch data analysis projects in the cloud by using Amazon Elastic MapReduce (EMR), the hosted Hadoop framework in Amazon Web Services (AWS). Authors Kevin Schmidt and Christopher Phillips demonstrate best practices for using EMR and various AWS and Apache technologies by walking you...
Read Free Programming Elastic Mapreduce Using Aws Services To Build An End To End Application

through the construction of a sample MapReduce log analysis application. Using code samples and example configurations, you'll learn how to assemble the building blocks necessary to solve your biggest data analysis problems. Get an overview of the AWS and Apache software tools used in large-scale data analysis Go through the process of executing a Job Flow with a simple log analyzer Discover useful MapReduce patterns for filtering and analyzing data sets Use Apache Hive and Pig instead of Java to build a MapReduce Job Flow Learn the basics for using Amazon EMR to run machine learning algorithms Develop a project cost model for using Amazon EMR and other AWS tools

A practical guide to using modern software effectively in quantitative research in the social and natural sciences. This book offers a practical guide to the computational methods at the heart of most modern quantitative research. It will be essential reading for research assistants needing hands-on experience; students entering PhD programs in business, economics, and other social or natural sciences; and those seeking quantitative jobs in industry. No background in computer science is assumed; a learner need only have a computer with access to the Internet. Using the example as its principal pedagogical device, the book offers tried-and-true prototypes that illustrate many important computational tasks required in quantitative research. The best way to use the book is to read it at the computer keyboard and learn by doing. The book begins by introducing basic skills: how to use the operating system, how to organize data, and how to complete simple programming tasks. For its demonstrations, the book uses a UNIX-based operating system and a set of free software tools: the scripting language Python for programming tasks; the database management system SQLite; and the freely available R for statistical computing and graphics. The book goes on to describe particular tasks: analyzing data, implementing commonly used numerical and simulation methods, and creating extensions to Python to reduce cycle time. Finally, the book describes the use of LaTeX, a document markup language and preparation system.

If you plan to use Amazon Web Services to run applications in the cloud, the end-to-end approach in this book will save you needless trial and error. You'll find practical guidelines for designing and building applications with Amazon Elastic Compute Cloud (EC2) and a host of supporting AWS tools, with a focus on critical issues such as load balancing, monitoring, and automation. How do you move an existing application to AWS, or design your application so that it scales effectively? How much storage will you require? Programming Amazon EC2 not only helps you get started, it will also keep you going once you're successfully positioned in the cloud. This book is a must-read for application architects, developers, and administrators. Determine your application's lifecycle and identify the AWS tools you need Learn how to build and run your application as part of the development process Migrate simple web applications to the cloud with EC2, Amazon Simple Storage Service, and CloudFront content delivery Meet traffic demand with EC2's Auto Scaling and Elastic Load Balancing Decouple your application using Simple Queue Service, Simple Notification Service, and other tools Use the right tools to minimize downtime, improve uptime, and manage your decoupled system "Jurg and Flavia have done a great job in this book building a practical guide on how to build real systems using AWS." --Werner Vogels, VP & CTO at Amazon.com

Cloud Computing

The Azure Services Platform is a cloud-computing technology from Microsoft. It is composed of four core components—Windows Azure, .NET Services, SQL Services, and Live Services—each with a unique role in the functioning of your cloud service. It is the goal of this book to show you how to use these components, both separately and together, to build flawless cloud services. At its heart, Windows Azure Platform is a down-to-earth, code-centric book. This book aims to show you precisely how the components are employed and to demonstrate the techniques and best practices you need to know to use them to best effect. That said, author Tejaswi Redkar regularly takes time out to provide a thorough overview of the architectural concepts that underpin Windows Azure. Without this understanding, you will find it hard to use the platform to its full potential. By the time you've read this book, you will be comfortable building high-quality end-to-end Azure services of your own.

Although you don't need a large computing infrastructure to process massive amounts of data with Apache Hadoop, it can still be difficult to get started. This practical guide shows you how to quickly
Read Free Programming Elastic Mapreduce Using Aws Services To Build An End To End Application

launch data analysis projects in the cloud by using Amazon Elastic MapReduce (EMR), the hosted Hadoop framework in Amazon Web Services (AWS). Authors Kevin Schmidt and Christopher Phillips demonstrate best practices for using EMR and various AWS and Apache technologies by walking you through the construction of a sample MapReduce log analysis application. Using code samples and example configurations, you’ll learn how to assemble the building blocks necessary to solve your biggest data analysis problems. Get an overview of the AWS and Apache software tools used in large-scale data analysis Go through the process of executing a Job Flow with a simple log analyzer Discover useful MapReduce patterns for filtering and analyzing data sets Use Apache Hive and Pig instead of Java to build a MapReduce Job Flow Learn the basics for using Amazon EMR to run machine learning algorithms Develop a project cost model for using Amazon EMR and other AWS tools

About the Book Recent industry surveys expect the cloud computing services market to be in excess of $20 billion and cloud computing jobs to be in excess of 10 million worldwide in 2014 alone. In addition, since a majority of existing information technology (IT) jobs is focused on maintaining legacy in-house systems, the demand for these kinds of jobs is likely to drop rapidly if cloud computing continues to take hold of the industry. However, there are very few educational options available in the area of cloud computing beyond vendor-specific training by cloud providers themselves. Cloud computing courses have not found their way (yet) into mainstream college curricula. This book is written as a textbook on cloud computing for educational programs at colleges. It can also be used by cloud service providers who may be interested in offering a broader perspective of cloud computing to accompany their own customer and employee training programs. The typical reader is expected to have completed a couple of courses in programming using traditional high-level languages at the college-level, and is either a senior or a beginning graduate student in one of the science, technology, engineering or mathematics (STEM) fields. We have tried to write a comprehensive book that transfers knowledge through an immersive “hands-on approach”, where the reader is provided the necessary guidance and knowledge to develop working code for real-world cloud applications. Additional support is available at the book’s website: www.cloudcomputingbook.info Organization The book is organized into three main parts. Part I covers technologies that form the foundations of cloud computing. These include topics such as virtualization, load balancing, scalability & elasticity, deployment, and replication. Part II introduces the reader to the design & programming aspects of cloud computing. Case studies on design and implementation of several cloud applications in the areas such as image processing, live streaming and social networks analytics are provided. Part III introduces the reader to specialized aspects of cloud computing including cloud application benchmarking, cloud security, multimedia applications and big data analytics. Case studies in areas such as IT, healthcare, transportation, networking and education are provided.

Large Scale and Big Data: Processing and Management provides readers with a central source of reference on the data management techniques currently available for large-scale data processing. Presenting chapters written by leading researchers, academics, and practitioners, it addresses the fundamental challenges associated with Big Data processing t

Data is arriving faster than you can process it and the overall volumes keep growing at a rate that keeps you awake at night. Hadoop can help you tame the data beast. Effective use of Hadoop however requires a mixture of programming, design, and system administration skills. "Hadoop Beginner's Guide" removes the mystery from Hadoop, presenting Hadoop and related technologies with a focus on building working systems and getting the job done, using cloud services to do so when it makes sense. From basic concepts and initial setup through developing applications and keeping the system running as the data grows, the book gives the understanding needed to effectively use Hadoop to solve real world problems. Starting with the basics of installing and configuring Hadoop, the book explains how to develop applications, maintain the system, and how to use additional products to integrate with other systems. While learning different ways to develop applications to run on Hadoop the book also covers tools such as Hive, Sqoop, and Flume that show how Hadoop can be integrated with relational databases and log collection. In addition to examples on Hadoop clusters on Ubuntu uses of cloud services such as Amazon, EC2 and Elastic MapReduce are covered.

The growth of Internet use and technologies has increased exponentially within the business sector. When utilized properly, these applications can enhance business functions and make them easier to
Read Free Programming Elastic Mapreduce Using Aws Services To Build An End To End Application

Exploring the Convergence of Big Data and the Internet of Things is a pivotal reference source featuring the latest empirical research on the business use of computing devices to send and receive data in conjunction with analytic applications to reduce maintenance costs, avoid equipment failures, and improve business operations. Including research on a broad range of topics such as supply chain, aquaculture, and speech recognition systems, this book is ideally designed for researchers, academicians, and practitioners seeking current research on various technology uses in business.

Cyber security has become a topic of concern over the past decade as private industry, public administration, commerce, and communication have gained a greater online presence. As many individual and organizational activities continue to evolve in the digital sphere, new vulnerabilities arise. Cyber Security and Threats: Concepts, Methodologies, Tools, and Applications contains a compendium of the latest academic material on new methodologies and applications in the areas of digital security and threats. Including innovative studies on cloud security, online threat protection, and cryptography, this multi-volume book is an ideal source for IT specialists, administrators, researchers, and students interested in uncovering new ways to thwart cyber breaches and protect sensitive digital information.

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

Web browsing would not be what it is today without the use of Service-Oriented Architecture (SOA). Although much has been written about SOA methodology, this emerging platform is continuously under development. Exploring Enterprise Service Bus in the Service-Oriented Architecture Paradigm is a detailed reference source that examines current aspects and research methodologies that enable enterprise service bus to unify and connect services efficiently on a common platform. Featuring relevant topics such as SOA reference architecture, grid computing applications, complex event computing, and java business integration, this is an ideal resource for all practitioners, academicians, graduate students, and researchers interested in the discoveries on the relationship that Service-Oriented architecture and enterprise service bus share.

For many organizations, Hadoop is the first step for dealing with massive amounts of data. The next step? Processing and analyzing datasets with the Apache Pig scripting platform. With Pig, you can batch-process data without having to create a full-fledged application, making it easy to experiment with new datasets. Updated with use cases and programming examples, this second edition is the ideal learning tool for new and experienced users alike. You'll find comprehensive coverage on key features such as the Pig Latin scripting language and the Grunt shell. When you need to analyze terabytes of data, this book shows you how to do it efficiently with Pig. Delve into Pig’s data model, including scalar and complex data types Write Pig Latin scripts to sort, group, join, project, and filter your data Use Grunt to work with the Hadoop Distributed File System (HDFS) Build complex data processing pipelines with Pig’s macros and modularity features Embed Pig Latin in Python for iterative processing and other advanced tasks Use Pig with Apache Tez to build high-performance batch and interactive data processing applications Create your own load and store functions to handle data formats and storage mechanisms

Fog computing is rapidly expanding in its applications and capabilities through various parts of society. Utilizing different types of virtualization technologies can push this branch of computing to even greater heights. Fog Computing: Breakthroughs in Research and Practice contains a compendium of the latest academic material on the evolving theory and practice related to fog computing. Including innovative studies on distributed fog computing environments, programming models, and access control mechanisms, this publication is an ideal source for programmers, IT professionals, students, researchers, and engineers.

Modern businesses depend on data for their very survival, creating a need for sophisticated databases and database technologies to help store, organise and transport their valuable data. This updated and expanded, easy-to-read textbook/reference presents a comprehensive introduction to databases, opening with a concise history of databases and of data as an organisational asset. As relational database management systems are no longer the only database solution, the book takes a wider view of database technology, encompassing big data, NoSQL, object and object-relational, and in-memory
Read Free Programming Elastic Mapreduce Using Aws Services To Build An End To End Application
databases. Presenting both theoretical and practical elements, the new edition also examines the
issues of scalability, availability, performance and security encountered when building and running a
database in the real world. Topics and features: Presents review and discussion questions at the end
of each chapter, in addition to skill-building, hands-on exercises Provides new material on database
adaptiveness, integration, and efficiency in relation to data growth Introduces a range of commercial
databases and encourages the reader to experiment with these in an associated learning environment
Reviews use of a variety of databases in business environments, including numerous examples
Discusses areas for further research within this fast-moving domain With its learning-by-doing
approach, supported by both theoretical and practical examples, this clearly-structured textbook will
be of great value to advanced undergraduate and postgraduate students of computer science,
software engineering, and information technology. Practising database professionals and application
developers will also find the book an ideal reference that addresses today's business needs.
Konstantinos Domdouzis is senior lecturer in the Communication and Computing Research Centre at
Sheffield Hallam University, UK. Peter Lake (now retired) was formerly course leader for the Oracle
IT&M MSc and the IT Professional MSc at Sheffield Hallam University. Paul Crowther (now retired)
was formerly head of postgraduate taught programmes in the Faculty of Arts, Computing, Engineering
and Sciences at Sheffield Hallam University.

Web service technologies are redefining the way that large and small companies are doing business
and exchanging information. Due to the critical need for furthering automation, engagement, and
efficiency, systems and workflows are becoming increasingly more web-based. Web Services:
Concepts, Methodologies, Tools, and Applications is an innovative reference source that examines
relevant theoretical frameworks, current practice guidelines, industry standards and standardization,
and the latest empirical research findings in web services. Highlighting a range of topics such as
cloud computing, quality of service, and semantic web, this multi-volume book is designed for
computer engineers, IT specialists, software designers, professionals, researchers, and upper-level
students interested in web services architecture, frameworks, and security.

Technology has become deeply integrated into modern society and various activities throughout
everyday life. However, this increases the risk of vulnerabilities, such as hacking or system errors,
among other online threats. Cybersecurity Breaches and Issues Surrounding Online Threat Protection
is an essential reference source for the latest scholarly research on the various types of unauthorized
access or damage to electronic data. Featuring extensive coverage across a range of relevant
perspectives and topics, such as robotics, cloud computing, and electronic data diffusion, this
publication is ideally designed for academicians, researchers, computer engineers, graduate students,
and practitioners seeking current research on the threats that exist in the world of technology.

The definitive guide to successfully integrating social, mobile, Big-Data analytics, cloud and IoT
principles and technologies The main goal of this book is to spur the development of effective big-data
computing operations on smart clouds that are fully supported by IoT sensing, machine learning and
analytics systems. To that end, the authors draw upon their original research and proven track record
in the field to describe a practical approach integrating big-data theories, cloud design principles,
Internet of Things (IoT) sensing, machine learning, data analytics and Hadoop and Spark
programming. Part 1 focuses on data science, the roles of clouds and IoT devices and frameworks for
big-data computing. Big data analytics and cognitive machine learning, as well as cloud architecture,
IoT and cognitive systems are explored, and mobile cloud-IoT-interaction frameworks are illustrated
with concrete system design examples. Part 2 is devoted to the principles of and algorithms for
machine learning, data analytics and deep learning in big data applications. Part 3 concentrates on
cloud programming software libraries from MapReduce to Hadoop, Spark and TensorFlow and
describes business, educational, healthcare and social media applications for those tools. The first
book describing a practical approach to integrating social, mobile, analytics, cloud and IoT (SMACT)
principles and technologies Covers theory and computing techniques and technologies, making it
suitable for use in both computer science and electrical engineering programs Offers an extremely
well-informed vision of future intelligent and cognitive computing environments integrating SMACT
technologies Fully illustrated throughout with examples, figures and approximately 150 problems to
support and reinforce learning Features a companion website with an instructor manual and
PowerPoint slides www.wiley.com/go/hwangIOT Big-Data Analytics for Cloud, IoT and Cognitive
Computing satisfies the demand among university faculty and students for cutting-edge information
Read Free Programming Elastic Mapreduce Using Aws Services To Build An End To End Application

on emerging intelligent and cognitive computing systems and technologies. Professionals working in data science, cloud computing and IoT applications will also find this book to be an extremely useful working resource.

Presents a guide to the features of C♯, covering such topics as functions, generics, iterators, currying, caching, order functions, sequences, monads, and MapReduce.

Perform cloud-based machine learning and deep learning using Amazon Web Services such as SageMaker, Lex, Comprehend, Translate, and Polly Key Features Explore popular machine learning and deep learning services with their underlying algorithms Discover readily available artificial intelligence(AI) APIs on AWS like Vision and Language Services Design robust architectures to enable experimentation, extensibility, and maintainability of AI apps Book Description From data wrangling through to translating text, you can accomplish this more with the artificial intelligence and machine learning services available on AWS. With this book, you’ll work through hands-on exercises and learn to use these services to solve real-world problems. You’ll even design, develop, monitor, and maintain machine and deep learning models on AWS. The book starts with an introduction to AI and its applications in different industries, along with an overview of AWS artificial intelligence and machine learning services. You’ll then get to grips with detecting and translating text with Amazon Rekognition and Amazon Translate. The book will assist you in performing speech-to-text with Amazon Transcribe and Amazon Polly. Later, you’ll discover the use of Amazon Comprehend for extracting information from text, and Amazon Lex for building voice chatbots. You will also understand the key capabilities of Amazon SageMaker such as wrangling big data, discovering topics in text collections, and classifying images. Finally, you’ll cover sales forecasting with deep learning and autoregression, before exploring the importance of a feedback loop in machine learning. By the end of this book, you will have the skills you need to implement AI in AWS through hands-on exercises that cover all aspects of the ML model life cycle. What you will learn Gain useful insights into different machine and deep learning models Build and deploy robust deep learning systems to production Train machine and deep learning models with diverse infrastructure specifications Scale AI apps without dealing with the complexity of managing the underlying infrastructure Monitor and manage AI experiments efficiently Create AI apps using AWS pre-trained AI services Who this book is for This book is for data scientists, machine learning developers, deep learning researchers, and artificial intelligence enthusiasts who want to harness the power of AWS to implement powerful artificial intelligence solutions. A basic understanding of machine learning concepts is expected. While cloud computing continues to transform developments in information technology services, these advancements have contributed to a rise in cyber attacks; producing an urgent need to extend the applications of investigation processes. Cybercrime and Cloud Forensics: Applications for Investigation Processes presents a collection of research and case studies of applications for investigation processes in cloud computing environments. This reference source brings together the perspectives of cloud customers, security architects, and law enforcement agencies in the developing area of cloud forensics.

Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics combines elements of computer science, information technology, mathematics, statistics and biotechnology, providing the methodology and in silico solutions to mine biological data and processes. The book covers Theory, Topics and Applications, with a special focus on Integrative -omics and Systems Biology. The theoretical, methodological underpinnings of BCB, including phylogeny are covered, as are more current areas of focus, such as translational bioinformatics, cheminformatics, and environmental informatics. Finally, Applications provide guidance for commonly asked questions. This major reference work spans basic and cutting-edge methodologies authored by leaders in the field, providing an invaluable resource for students, scientists, professionals in research institutes, and a broad swath of researchers in biotechnology and the biomedical and pharmaceutical industries. Brings together information from computer science, information technology, mathematics, statistics and biotechnology Written and reviewed by leading experts in the field, providing a unique and authoritative resource Focuses on the main theoretical and methodological concepts before expanding on specific topics and applications Includes interactive images, multimedia tools and crosslinking to further resources and databases.
Distributed and Cloud Computing: From Parallel Processing to the Internet of Things offers complete coverage of modern distributed computing technology including clusters, the grid, service-oriented architecture, massively parallel processors, peer-to-peer networking, and cloud computing. It is the first modern, up-to-date distributed systems textbook; it explains how to create high-performance, scalable, reliable systems, exposing the design principles, architecture, and innovative applications of parallel, distributed, and cloud computing systems. Topics covered by this book include: facilitating management, debugging, migration, and disaster recovery through virtualization; clustered systems for research or ecommerce applications; designing systems as web services; and social networking systems using peer-to-peer computing. The principles of cloud computing are discussed using examples from open-source and commercial applications, along with case studies from the leading distributed computing vendors such as Amazon, Microsoft, and Google. Each chapter includes exercises and further reading, with lecture slides and more available online. This book will be ideal for students taking a distributed systems or distributed computing class, as well as for professional system designers and engineers looking for a reference to the latest distributed technologies including cloud, P2P and grid computing. Complete coverage of modern distributed computing technology including clusters, the grid, service-oriented architecture, massively parallel processors, peer-to-peer networking, and cloud computing Includes case studies from the leading distributed computing vendors: Amazon, Microsoft, Google, and more Explains how to use virtualization to facilitate management, debugging, migration, and disaster recovery Designed for undergraduate or graduate students taking a distributed systems course—each chapter includes exercises and further reading, with lecture slides and more available online

Product Description Amazon Cloud Computing With C#/.Net provides A comprehensive look at the emerging Amazon Web Services Platform and a peep into the emerging paradigm of cloud computing from the perspective of the leading Cloud vendors offerings. It will enable you to plan migration efforts from enterprise softwares to ones operating from the cloud. The book is technical in nature and walks the reader through development of tools and programs which work with AWS. The book is accompanied by the complete source code to the exercises covered in the book, which can be downloaded from the authors website. ‘A no fluff just stuff’ approach to utilizing AWS. Amazon Cloud Computing With C#/.Net covers - EC2 - CloudWatch - Elastic Load Balancing - AutoScaling - S3 - Virtual Private Cloud - SimpleDB - RDS - CloudFront - SQS - Elastic MapReduce - Agile Continuous Integration with AWS

Learn from the AWS subject-matter experts, apply real-world scenarios and clear the AWS Certified Solutions Architect -Associate exam Key Features Build highly reliable and scalable workloads on the AWS platform Pass the exam in less time and with confidence Get up and running with building and managing applications on the AWS platform Book Description Amazon Web Services (AWS) is currently the leader in the public cloud market. With an increasing global interest in leveraging cloud infrastructure, the AWS Cloud from Amazon offers a cutting-edge platform for architecting, building, and deploying web-scale cloud applications. As more the rate of cloud platform adoption increases, so does the need for cloud certification. The AWS Certified Solution Architect - Associate Guide is your one-stop solution to gaining certification. Once you have grasped what AWS and its prerequisites are, you will get insights into different types of AWS services such as Amazon S3, EC2, VPC, SNS, and more to get you prepared with core Amazon services. You will then move on to understanding how to design and deploy highly scalable applications. Finally, you will study security concepts along with the AWS best practices and mock papers to test your knowledge. By the end of this book, you will not only be fully prepared to pass the AWS Certified Solutions Architect - Associate exam but also capable of building secure and reliable applications. What you will learn Explore AWS terminology and identity and access management Acquaint yourself with important cloud services and features in categories such as compute, network, storage, and databases Define access control to secure AWS resources and set up efficient monitoring Back up your database and ensure high availability by understanding all of the database-related services in the AWS Cloud Integrate AWS with your applications to meet and exceed non-functional requirements Build and deploy cost-effective and highly available applications Who this book is for The AWS Certified Solutions Architect -Associate Guide is for you if you are an IT professional or Solutions Architect wanting to pass the AWS Certified Solution Architect - Associate 2018 exam. This book is also for developers looking to start building scalable applications on AWS
Cloud computing is a buzzword in today's information technology (IT) that nobody can escape. But what is really behind it? There are many interpretations of this term, but no standardized or even uniform definition. Instead, as a result of the multi-faceted viewpoints and the diverse interests expressed by the various stakeholders, cloud computing is perceived as a rather fuzzy concept. With this book, the authors deliver an overview of cloud computing architecture, services, and applications. Their aim is to bring readers up to date on this technology and thus to provide a common basis for discussion, new research, and novel application scenarios. They first introduce the foundation of cloud computing with its basic technologies, such as virtualization and Web services. After that they discuss the cloud architecture and its service modules. The following chapters then cover selected commercial cloud offerings (including Amazon Web Services and Google App Engine) and management tools, and present current related open-source developments (including Hadoop, Eucalyptus, and Open CirrusTM). Next, economic considerations (cost and business models) are discussed, and an evaluation of the cloud market situation is given. Finally, the appendix contains some practical examples of how to use cloud resources or cloud applications, and a glossary provides concise definitions of key terms. The authors' presentation does not require in-depth technical knowledge. It is equally intended as an introduction for students in software engineering, web technologies, or business development, for professional software developers or system architects, and for future-oriented decision-makers like top executives and managers.

This book is aimed at developers and system administrators who want to learn about Big Data analysis using Amazon Elastic MapReduce. Basic Java programming knowledge is required. You should be comfortable with using command-line tools. Prior knowledge of AWS, API, and CLI tools is not assumed. Also, no exposure to Hadoop and MapReduce is expected.

Hadoop in Action teaches readers how to use Hadoop and write MapReduce programs. The intended readers are programmers, architects, and project managers who have to process large amounts of data offline. Hadoop in Action will lead the reader from obtaining a copy of Hadoop to setting it up in a cluster and writing data analytic programs. The book begins by making the basic idea of Hadoop and MapReduce easier to grasp by applying the default Hadoop installation to a few easy-to-follow tasks, such as analyzing changes in word frequency across a body of documents. The book continues through the basic concepts of MapReduce applications developed using Hadoop, including a close look at framework components, use of Hadoop for a variety of data analysis tasks, and numerous examples of Hadoop in action. Hadoop in Action will explain how to use Hadoop and present design patterns and practices of programming MapReduce. MapReduce is a complex idea both conceptually and in its implementation, and Hadoop users are challenged to learn all the knobs and levers for running Hadoop. This book takes you beyond the mechanics of running Hadoop, teaching you to write meaningful programs in a MapReduce framework. This book assumes the reader will have a basic familiarity with Java, as most code examples will be written in Java. Familiarity with basic statistical concepts (e.g. histogram, correlation) will help the reader appreciate the more advanced data processing examples. Purchase of the print book comes with an offer of a free PDF, ePub, and Kindle eBook from Manning. Also available is all code from the book.

This illuminating text/reference surveys the state of the art in data science, and provides practical guidance on big data analytics. Expert perspectives are provided by authoritative researchers and practitioners from around the world, discussing research developments and emerging trends, presenting case studies on helpful frameworks and innovative methodologies, and suggesting best practices for efficient and effective data analytics. Features: reviews a framework for fast data applications, a technique for complex event processing, and agglomerative approaches for the partitioning of networks; introduces a unified approach to data modeling and management, and a distributed computing perspective on interfacing physical and cyber worlds; presents techniques for machine learning for big data, and identifying duplicate records in data repositories; examines enabling technologies and tools for data mining; proposes frameworks for data extraction, and adaptive decision making and social media analysis.

This book is for programmers and developers who want to improve the performance of their R programs by making them run faster with large data sets or who are trying to solve a pesky performance problem.
This book is aimed at developers and system administrators who want to learn about Big Data analysis using Amazon Elastic MapReduce. Basic Java programming knowledge is required. You should be comfortable with using command-line tools. Prior knowledge of AWS, API, and CLI tools is not assumed. Also, no exposure to Hadoop and MapReduce is expected.

Host Your Web Site On The Cloud is your step-by-step guide to this revolutionary approach to hosting and managing your web applications. Cloud computing gives you the tools you need to prepare and cope with a traffic onslaught. You'll have the confidence to withstand a traffic surge without melting your servers or sending you into bankruptcy. There are a number of ways to use the cloud to host existing applications, build creative new ones, and improve the cost-effectiveness and efficiency or organizations large and small. You'll learn how to: gain a thorough understanding of cloud computing master the fundamentals of Amazon Web Services install and configure visual and command line tools store, retrieve, and distribute data quickly and easily build applications that scale manage the monitoring, load balancing, and scaling capabilities of cloud computing As a developer, you need room & flexibility to be innovative. Why waste time worrying about the technical aspects of server capacity? AWS handles security, load balancing, and server resources virtually so you're not restricted to one physical server.

The efficient management of a consistent and integrated database is a central task in modern IT and highly relevant for science and industry. Hardly any critical enterprise solution comes without any functionality for managing data in its different forms. Web-Scale Data Management for the Cloud addresses fundamental challenges posed by the need and desire to provide database functionality in the context of the Database as a Service (DBaaS) paradigm for database outsourcing. This book also discusses the motivation of the new paradigm of cloud computing, and its impact to data outsourcing and service-oriented computing in data-intensive applications. Techniques with respect to the support in the current cloud environments, major challenges, and future trends are covered in the last section of this book. A survey addressing the techniques and special requirements for building database services are provided in this book as well.

The first textbook to teach students how to build data analytic solutions on large data sets using cloud-based technologies. This is the first textbook to teach students how to build data analytic solutions on large data sets (specifically in Internet of Things applications) using cloud-based technologies for data storage, transmission and mashup, and AI techniques to analyze this data. This textbook is designed to train college students to master modern cloud computing systems in operating principles, architecture design, machine learning algorithms, programming models and software tools for big data mining, analytics, and cognitive applications. The book will be suitable for use in one-semester computer science or electrical engineering courses on cloud computing, machine learning, cloud programming, cognitive computing, or big data science. The book will also be very useful as a reference for professionals who want to work in cloud computing and data science. Cloud and Cognitive Computing begins with two introductory chapters on fundamentals of cloud computing, data science, and adaptive computing that lay the foundation for the rest of the book. Subsequent chapters cover topics including cloud architecture, mashup services, virtual machines, Docker containers, mobile clouds, IoT and AI, inter-cloud mashups, and cloud performance and benchmarks, with a focus on Google’s Brain Project, DeepMind, and X-Lab programs, IBKai HwangM SyNapse, Bluemix programs, cognitive initiatives, and neurocomputers. The book then covers machine learning algorithms and cloud programming software tools and application development, applying the tools in machine learning, social media, deep learning, and cognitive applications. All cloud systems are illustrated with big data and cognitive application examples.

Computational methodologies and modeling play a growing role for investigating mechanisms, and for the diagnosis and therapy of human diseases. This progress gave rise to computational medicine, an interdisciplinary field at the interface of computer science and medicine. The main focus of computational medicine lies in the development of data analysis methods and mathematical modeling as well as computational simulation techniques specifically addressing medical problems. In this book, we present a number of computational medicine topics at several scales: from molecules to cells, organs, and organisms. At the molecular level, tools for the analysis of genome variations as well as cloud computing resources for medical genetics are reviewed. Then, an analysis of gene expression data and the application to the characterization of microbial communities are highlighted.
At the protein level, two types of analyses for mass spectrometry data are reviewed: labeled quantitative proteomics and lipidomics, followed by protein sequence analysis and a 3D structure and drug design chapter. Finally, three chapters on clinical applications focus on the integration of biomolecular and clinical data for cancer research, biomarker discovery, and network-based methods for computational diagnostics.

Copyright code: 278336fde3fa1a26ea1f81a8bc5b6c95