

Big Data Benchmarks Performance Optimization And Emerging Hardware 4th And 5th Workshops Bpoe 2014 Salt Lake City Usa March 1 2014 And Papers Lecture Notes In Computer Science

Thank you for downloading **Big Data Benchmarks Performance Optimization And Emerging Hardware 4th And 5th Workshops Bpoe 2014 Salt Lake City Usa March 1 2014 And Papers Lecture Notes In Computer Science** . Maybe you have knowledge that, people have look hundreds times for their favorite novels like this Big Data Benchmarks Performance Optimization And Emerging Hardware 4th And 5th Workshops Bpoe 2014 Salt Lake City Usa March 1 2014 And Papers Lecture Notes In Computer Science , but end up in harmful downloads. Rather than reading a good book with a cup of coffee in the afternoon, instead they cope with some infectious bugs inside their computer.

Big Data Benchmarks Performance Optimization And Emerging Hardware 4th And 5th Workshops

Bpoe 2014 Salt Lake City Usa March 1 2014 And Papers Lecture Notes In Computer Science is available in our digital library an online access to it is set as public so you can get it instantly. Our digital library hosts in multiple countries, allowing you to get the most less latency time to download any of our books like this one.

Merely said, the Big Data Benchmarks Performance Optimization And Emerging Hardware 4th And 5th Workshops Bpoe 2014 Salt Lake City Usa March 1 2014 And Papers Lecture Notes In Computer Science is universally compatible with any devices to read

Industry 4.0 - Aydin Azizi 2022-12-07

This book presents Industry 4.0 enabler technologies and tools. It also highlights some of the existing empirical applications in the context of manufacturing. The book elucidates innovative thematic concepts of Industry 4.0 and its perspectives. It establishes routes to empirically utilize Industry 4.0 standards for manufacturing companies. The book can be used as a reference for professionals/engineers, researchers, and students.

Communication Optimization for Customizable Domain-specific Computing -

Bingjun Xiao 2015

This dissertation investigates the communication optimization for customizable domain-specific computing at different levels in a customizable heterogeneous platform (CHP) to improve the system performance and energy efficiency. Fabric-level optimization driven by emerging devices. Programmable fabrics (e.g., FPGAs) can be used to improve domain-specific computing by >10x in energy efficiency over CPUs since FPGAs can be customized to the application kernels in the target domain. But the programmable interconnects inside FPGAs

occupy >50% of the FPGA area, delay and power. We propose a novel architecture of programmable interconnects based on resistive RAM (RRAM), a type of emerging device with high density and low power. We optimize the layout and the programming circuit of the new architecture. We also extend RRAM benefits to routing buffers. We observe the high defect rate in the emerging RRAM manufacturing and further develop a defect-aware communication mechanism. Conventional defect avoidance leaves a large portion of the chip in the new architecture unusable. So we propose new defect utilization methodologies by treating stuck-closed defects as shorting constraints in the routing of signals. We develop a scalable algorithm to perform timing-driven routing under these extra constraints and successfully suppress the impact of defects. Chip-level optimization driven by accelerator-centric architectures. A chip can also be customized to an application domain by integrating a sea of

accelerators designed for the frequently used kernels in the domain. The design of interconnects among customized accelerators and shared resources (e.g., shared memories) is a serious challenge in chip design. Accelerators run 100x faster than CPUs and post a high data demand on the communication infrastructure. To address this challenge, we develop a novel design of interconnects between accelerators and shared memories and exploit several optimization opportunities that emerge in accelerator-rich computing platforms. Experiments show that our design outperforms prior work that was optimized for CPU cores or signal routing. Another design challenge lies in the data reuse optimization within an accelerator to minimize its off-chip accesses and on-chip buffer usage. Since the fully pipelined computation kernel consumes large amounts of data every clock cycle, and the data access pattern is the major difference among applications, existing accelerators use ad hoc

data reuse schemes that are carefully tuned per application to fit the data demand. To reduce the engineering cost of accelerator-rich architectures, we develop a data reuse infrastructure that is generalized for the stencil computation domain and can be instantiated to the optimal design for any application in the domain. We demonstrate the robustness of our method over a set of real-life benchmarks. Server-level and cluster-level optimization driven by big data. In the era of big data, workloads can no longer fit into a single chip. Most data are stored in disks, and we can only load a small part of it into main memories during computation. Due to the low access speed of disks, our primary design goal becomes minimization of the data transfer between disks and the main memory. We select a popular big data application, convolutional neural network (CNN), as a case study. We analyze the linear algebraic properties of CNN, and propose algorithmic modifications to reduce the total

computational workload and the disk access. Furthermore, when the application data become even larger, it needs to be distributed among a cluster of server nodes. This motivates us to develop an accelerator-centric computing cluster. We test two machine learning applications, logistic regression and artificial neural network (ANN), on our prototyping cluster and try to minimize the total data transfer incurred during the computation in this cluster. We select the distributed stochastic gradient descent (dSGD) as our training algorithm to eliminate the inter-node communication within a training iteration. We also deploy an in-memory cluster computing infrastructure, Spark, to eliminate the inter-node communication across training iterations. The baseline Spark only supports CPUs, and we develop a software layer to allow Spark tasks to offload their major computation to accelerators which are equipped by each server node. During the computation offloading, we group multiple

tasks into a batch and transfer it to the target accelerator in one transaction to minimize the setup overhead of the data transfer between accelerators and host servers. We further realize accelerator data caching to eliminate the unnecessary data transfer of training data based on the properties of iterative machine learning applications.

High Performance Computing - Rio Yokota
2019-01-24

This book constitutes the refereed post-conference proceedings of 13 workshops held at the 33rd International ISC High Performance 2018 Conference, in Frankfurt, Germany, in June 2018: HPC I/O in the Data Center, HPC-IODC 2018; Workshop on Performance and Scalability of Storage Systems, WOPSSS 2018; 13th Workshop on Virtualization in High-Performance Cloud Computing, VHPC 2018; Third International Workshop on In Situ Visualization, WOIV 2018; 4th International Workshop on Communication Architectures for HPC, Big Data,

Deep Learning and Clouds at Extreme Scale, ExaComm 2018; International Workshop on OpenPOWER for HPC, IWOPH 2018; IXPUG Workshop: Many-Core Computing on Intel Processors; Workshop on Sustainable Ultrascale Computing Systems; Approximate and Transprecision Computing on Emerging Technologies, ATCET 2018; First Workshop on the Convergence of Large-Scale Simulation and Artificial Intelligence; Third Workshop for Open Source Supercomputing, OpenSuCo 2018; First Workshop on Interactive High-Performance Computing; Workshop on Performance Portable Programming Models for Accelerators, P³MA 2018. The 53 full papers included in this volume were carefully reviewed and selected from 80 submissions. They cover all aspects of research, development, and application of large-scale, high performance experimental and commercial systems. Topics include HPC computer architecture and hardware; programming models, system software, and applications;

solutions for heterogeneity, reliability, power efficiency of systems; virtualization and containerized environments; big data and cloud computing; and artificial intelligence.

Big Data Benchmarks, Performance Optimization, and Emerging Hardware - Jianfeng Zhan 2016-01-28

This book constitutes the thoroughly revised selected papers of the 6th workshop on Big Data Benchmarks, Performance Optimization, and Emerging Hardware, BPOE 2015, held in Kohala Coast, HI, USA, in August/September 2015 as satellite event of VLDB 2015, the 41st International Conference on Very Large Data Bases. The 8 papers presented were carefully reviewed and selected from 10 submissions. The workshop focuses on architecture and system support for big data systems, aiming at bringing researchers and practitioners from data management, architecture, and systems research communities together to discuss the research issues at the intersection of these

areas. This book also invites three papers from several industrial partners, including two papers describing tools used in system benchmarking and monitoring and one paper discussing principles and methodologies in existing big data benchmarks.

Interdisciplinary Approaches to Spatial Optimization Issues - Faiz, Sami 2021-01-15

As metropolises continue to see a growth in population, planners are continually searching for trending methods for utilizing space and seeking the best geographical arrangements for these cities. Professionals have continually used geographic information systems (GIS) to solve these issues; however, limitations in this technology remain prevalent. Integrating multiple-criteria decision analysis and evolutionary computing tools with GIS has created an array of robust solutions for spatial optimization problems in densely populated areas. *Interdisciplinary Approaches to Spatial Optimization Issues* is a pivotal reference source

that provides vital research on advancements within the field of GIS and evolutionary solutions for spatial optimization issues. While highlighting topics such as computing machinery, vehicular routing, and operational research, this publication is ideally designed for practitioners, technicians, developers, academicians, students, government officials, planners, and researchers seeking current research on applications and improvements within spatial optimization and GIS.

Algorithms and Architectures for Parallel Processing - Sheng Wen 2020-01-21

The two-volume set LNCS 11944-11945 constitutes the proceedings of the 19th International Conference on Algorithms and Architectures for Parallel Processing, ICA3PP 2019, held in Melbourne, Australia, in December 2019. The 73 full and 29 short papers presented were carefully reviewed and selected from 251 submissions. The papers are organized in topical sections on: Parallel and Distributed

Architectures, Software Systems and Programming Models, Distributed and Parallel and Network-based Computing, Big Data and its Applications, Distributed and Parallel Algorithms, Applications of Distributed and Parallel Computing, Service Dependability and Security, IoT and CPS Computing, Performance Modelling and Evaluation.

Big Data 2.0 Processing Systems - Sherif Sakr
2016-08-24

This book provides readers the “big picture” and a comprehensive survey of the domain of big data processing systems. For the past decade, the Hadoop framework has dominated the world of big data processing, yet recently academia and industry have started to recognize its limitations in several application domains and big data processing scenarios such as the large-scale processing of structured data, graph data and streaming data. Thus, it is now gradually being replaced by a collection of engines that are dedicated to specific verticals (e.g.

structured data, graph data, and streaming data). The book explores this new wave of systems, which it refers to as Big Data 2.0 processing systems. After Chapter 1 presents the general background of the big data phenomena, Chapter 2 provides an overview of various general-purpose big data processing systems that allow their users to develop various big data processing jobs for different application domains. In turn, Chapter 3 examines various systems that have been introduced to support the SQL flavor on top of the Hadoop infrastructure and provide competing and scalable performance in the processing of large-scale structured data. Chapter 4 discusses several systems that have been designed to tackle the problem of large-scale graph processing, while the main focus of Chapter 5 is on several systems that have been designed to provide scalable solutions for processing big data streams, and on other sets of systems that have been introduced to support the

development of data pipelines between various types of big data processing jobs and systems. Lastly, Chapter 6 shares conclusions and an outlook on future research challenges. Overall, the book offers a valuable reference guide for students, researchers and professionals in the domain of big data processing systems. Further, its comprehensive content will hopefully encourage readers to pursue further research on the subject.

Conquering Big Data with High Performance Computing - Ritu Arora 2016-09-16

This book provides an overview of the resources and research projects that are bringing Big Data and High Performance Computing (HPC) on converging tracks. It demystifies Big Data and HPC for the reader by covering the primary resources, middleware, applications, and tools that enable the usage of HPC platforms for Big Data management and processing. Through interesting use-cases from traditional and non-traditional HPC domains, the book highlights the

most critical challenges related to Big Data processing and management, and shows ways to mitigate them using HPC resources. Unlike most books on Big Data, it covers a variety of alternatives to Hadoop, and explains the differences between HPC platforms and Hadoop. Written by professionals and researchers in a range of departments and fields, this book is designed for anyone studying Big Data and its future directions. Those studying HPC will also find the content valuable.

Network Data Analytics - K. G. Srinivasa
2018-04-26

In order to carry out data analytics, we need powerful and flexible computing software. However the software available for data analytics is often proprietary and can be expensive. This book reviews Apache tools, which are open source and easy to use. After providing an overview of the background of data analytics, covering the different types of analysis and the basics of using Hadoop as a tool, it

focuses on different Hadoop ecosystem tools, like Apache Flume, Apache Spark, Apache Storm, Apache Hive, R, and Python, which can be used for different types of analysis. It then examines the different machine learning techniques that are useful for data analytics, and how to visualize data with different graphs and charts. Presenting data analytics from a practice-oriented viewpoint, the book discusses useful tools and approaches for data analytics, supported by concrete code examples. The book is a valuable reference resource for graduate students and professionals in related fields, and is also of interest to general readers with an understanding of data analytics.

High-Performance Big Data Computing -
Dhabaleswar K. Panda 2022-08-02

An in-depth overview of an emerging field that brings together high-performance computing, big data processing, and deep learning. Over the last decade, the exponential explosion of data known as big data has changed the way we

understand and harness the power of data. The emerging field of high-performance big data computing, which brings together high-performance computing (HPC), big data processing, and deep learning, aims to meet the challenges posed by large-scale data processing. This book offers an in-depth overview of high-performance big data computing and the associated technical issues, approaches, and solutions. The book covers basic concepts and necessary background knowledge, including data processing frameworks, storage systems, and hardware capabilities; offers a detailed discussion of technical issues in accelerating big data computing in terms of computation, communication, memory and storage, codesign, workload characterization and benchmarking, and system deployment and management; and surveys benchmarks and workloads for evaluating big data middleware systems. It presents a detailed discussion of big data computing systems and applications with high-

performance networking, computing, and storage technologies, including state-of-the-art designs for data processing and storage systems. Finally, the book considers some advanced research topics in high-performance big data computing, including designing high-performance deep learning over big data (DLoBD) stacks and HPC cloud technologies. *Advanced Methodologies and Technologies in Network Architecture, Mobile Computing, and Data Analytics* - Khosrow-Pour, D.B.A., Mehdi 2018-10-19

From cloud computing to data analytics, society stores vast supplies of information through wireless networks and mobile computing. As organizations are becoming increasingly more wireless, ensuring the security and seamless function of electronic gadgets while creating a strong network is imperative. *Advanced Methodologies and Technologies in Network Architecture, Mobile Computing, and Data Analytics* highlights the challenges associated

with creating a strong network architecture in a perpetually online society. Readers will learn various methods in building a seamless mobile computing option and the most effective means of analyzing big data. This book is an important resource for information technology professionals, software developers, data analysts, graduate-level students, researchers, computer engineers, and IT specialists seeking modern information on emerging methods in data mining, information technology, and wireless networks.

Big Data Benchmarking - Tilmann Rabl 2015
This book constitutes the thoroughly refereed post-workshop proceedings of the 5th International Workshop on Big Data Benchmarking, WBDB 2014, held in Potsdam, Germany, in August 2014. The 13 papers presented in this book were carefully reviewed and selected from numerous submissions and cover topics such as benchmarks specifications and proposals, Hadoop and MapReduce - in the

different context such as virtualization and cloud - as well as in-memory, data generation, and graphs.

Advances in Databases and Information Systems - Morzy Tadeusz 2015-08-14

This book constitutes the thoroughly refereed proceedings of the 19th East European Conference on Advances in Databases and Information Systems, ADBIS 2015, held in Poitiers, France, in September 2015. The 31 full papers and 18 short papers presented were carefully selected and reviewed from 135 submissions. The papers are organized in topical sections such as database theory and access methods; user requirements and database evolution; multidimensional modeling and OLAP; ETL; transformation, extraction and archiving; modeling and ontologies; time series processing; performance and tuning; advanced query processing; approximation and skyline; confidentiality and trust.

Big Scientific Data Benchmarks,

Architecture, and Systems - Rui Ren

2019-01-11

This book constitutes the refereed proceedings of the First Workshop on Big Scientific Data Benchmarks, Architecture, and Systems, SDBA 2018, held in Beijing, China, in June 2018. The 10 revised full papers presented were carefully reviewed and selected from 22 submissions. The papers are organized in topical sections on benchmarking; performance optimization; algorithms; big science data framework.

Programming Hive - Edward Capriolo

2012-09-26

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

Big Data Analytics - 2015-08-04

While the term Big Data is open to varying interpretation, it is quite clear that the Volume, Velocity, and Variety (3Vs) of data have impacted every aspect of computational science and its applications. The volume of data is increasing at a phenomenal rate and a majority

of it is unstructured. With big data, the volume is so large that processing it using traditional database and software techniques is difficult, if not impossible. The drivers are the ubiquitous sensors, devices, social networks and the all-pervasive web. Scientists are increasingly looking to derive insights from the massive quantity of data to create new knowledge. In common usage, Big Data has come to refer simply to the use of predictive analytics or other certain advanced methods to extract value from data, without any required magnitude thereon. Challenges include analysis, capture, curation, search, sharing, storage, transfer, visualization, and information privacy. While there are challenges, there are huge opportunities emerging in the fields of Machine Learning, Data Mining, Statistics, Human-Computer Interfaces and Distributed Systems to address ways to analyze and reason with this data. The edited volume focuses on the challenges and opportunities posed by "Big Data" in a variety of

domains and how statistical techniques and innovative algorithms can help glean insights and accelerate discovery. Big data has the potential to help companies improve operations and make faster, more intelligent decisions. Review of big data research challenges from diverse areas of scientific endeavor Rich perspective on a range of data science issues from leading researchers Insight into the mathematical and statistical theory underlying the computational methods used to address big data analytics problems in a variety of domains *Mobility Patterns, Big Data and Transport Analytics* - Constantinos Antoniou 2018-11-27 *Mobility Patterns, Big Data and Transport Analytics* provides a guide to the new analytical framework and its relation to big data, focusing on capturing, predicting, visualizing and controlling mobility patterns - a key aspect of transportation modeling. The book features prominent international experts who provide overviews on new analytical frameworks,

applications and concepts in mobility analysis and transportation systems. Users will find a detailed, mobility 'structural' analysis and a look at the extensive behavioral characteristics of transport, observability requirements and limitations for realistic transportation applications and transportation systems analysis that are related to complex processes and phenomena. This book bridges the gap between big data, data science, and transportation systems analysis with a study of big data's impact on mobility and an introduction to the tools necessary to apply new techniques. The book covers in detail, mobility 'structural' analysis (and its dynamics), the extensive behavioral characteristics of transport, observability requirements and limitations for realistic transportation applications, and transportation systems analysis related to complex processes and phenomena. The book bridges the gap between big data, data science, and Transportation Systems Analysis with a

study of big data's impact on mobility, and an introduction to the tools necessary to apply new techniques. Guides readers through the paradigm-shifting opportunities and challenges of handling Big Data in transportation modeling and analytics Covers current analytical innovations focused on capturing, predicting, visualizing, and controlling mobility patterns, while discussing future trends Delivers an introduction to transportation-related information advances, providing a benchmark reference by world-leading experts in the field Captures and manages mobility patterns, covering multiple purposes and alternative transport modes, in a multi-disciplinary approach Companion website features videos showing the analyses performed, as well as test codes and data-sets, allowing readers to recreate the presented analyses and apply the highlighted techniques to their own data

Computer-Assisted Learning for Engaging Varying Aptitudes: From Theory to Practice -

Dhaya, R. 2022-09-30

Computer-assisted learning has completely modernized the way that students learn both in the average classroom as well as in language learning contexts. Through its ability to provide interactive and engaging learning resources, computer-assisted learning is a useful tool for engaging all learners. It is essential that educators stay current with the emerging learning technologies so that they can create more dynamic and engaging classrooms and pique the interest of even the most apathetic students. Computer-Assisted Learning for Engaging Varying Aptitudes: From Theory to Practice is an essential reference source that provides insights on the practical applications of technology-based learning and its measurement and explains the applicability of this method in various classrooms. Covering topics in facial recognition technology, big data technology, and learning challenges, this premier reference source is a dynamic resource for faculty and

administrators of both K-12 and higher education, pre-service teachers, IT consultants, educational software developers, government officials, superintendents, researchers, and academicians.

High-Performance Modelling and Simulation for Big Data Applications -

Joanna Kołodziej 2019-03-25

This open access book was prepared as a Final Publication of the COST Action IC1406 “High-Performance Modelling and Simulation for Big Data Applications (cHiPSet)” project. Long considered important pillars of the scientific method, Modelling and Simulation have evolved from traditional discrete numerical methods to complex data-intensive continuous analytical optimisations. Resolution, scale, and accuracy have become essential to predict and analyse natural and complex systems in science and engineering. When their level of abstraction raises to have a better discernment of the domain at hand, their representation gets

increasingly demanding for computational and data resources. On the other hand, High Performance Computing typically entails the effective use of parallel and distributed processing units coupled with efficient storage, communication and visualisation systems to underpin complex data-intensive applications in distinct scientific and technical domains. It is then arguably required to have a seamless interaction of High Performance Computing with Modelling and Simulation in order to store, compute, analyse, and visualise large data sets in science and engineering. Funded by the European Commission, cHiPSet has provided a dynamic trans-European forum for their members and distinguished guests to openly discuss novel perspectives and topics of interests for these two communities. This cHiPSet compendium presents a set of selected case studies related to healthcare, biological data, computational advertising, multimedia, finance, bioinformatics, and telecommunications.

Performance and Capacity Implications for Big Data - Dave Jewell 2014-02-07

Big data solutions enable us to change how we do business by exploiting previously unused sources of information in ways that were not possible just a few years ago. In IBM® Smarter Planet® terms, big data helps us to change the way that the world works. The purpose of this IBM Redpaper™ publication is to consider the performance and capacity implications of big data solutions, which must be taken into account for them to be viable. This paper describes the benefits that big data approaches can provide. We then cover performance and capacity considerations for creating big data solutions. We conclude with what this means for big data solutions, both now and in the future. Intended readers for this paper include decision-makers, consultants, and IT architects.

Building Big Data and Analytics Solutions in the Cloud - Wei-Dong Zhu 2014-12-08

Big data is currently one of the most critical

emerging technologies. Organizations around the world are looking to exploit the explosive growth of data to unlock previously hidden insights in the hope of creating new revenue streams, gaining operational efficiencies, and obtaining greater understanding of customer needs. It is important to think of big data and analytics together. Big data is the term used to describe the recent explosion of different types of data from disparate sources. Analytics is about examining data to derive interesting and relevant trends and patterns, which can be used to inform decisions, optimize processes, and even drive new business models. With today's deluge of data comes the problems of processing that data, obtaining the correct skills to manage and analyze that data, and establishing rules to govern the data's use and distribution. The big data technology stack is ever growing and sometimes confusing, even more so when we add the complexities of setting up big data environments with large up-front investments.

Cloud computing seems to be a perfect vehicle for hosting big data workloads. However, working on big data in the cloud brings its own challenge of reconciling two contradictory design principles. Cloud computing is based on the concepts of consolidation and resource pooling, but big data systems (such as Hadoop) are built on the shared nothing principle, where each node is independent and self-sufficient. A solution architecture that can allow these mutually exclusive principles to coexist is required to truly exploit the elasticity and ease-of-use of cloud computing for big data environments. This IBM® Redpaper™ publication is aimed at chief architects, line-of-business executives, and CIOs to provide an understanding of the cloud-related challenges they face and give prescriptive guidance for how to realize the benefits of big data solutions quickly and cost-effectively.

Big Data Management, Technologies, and Applications - Hu, Wen-Chen 2013-10-31

"This book discusses the exponential growth of information size and the innovative methods for data capture, storage, sharing, and analysis for big data"--Provided by publisher.

Research and Practical Issues of Enterprise Information Systems - A Min Tjoa 2016-11-17

This book constitutes the proceedings of the 10th International IFIP WG 8.9 Working Conference on Research and Practical Issues of Enterprise Information Systems, CONFENIS 2016, held in Vienna, Austria, in December 2016. The conference provided an international forum for the broader IFIP community to discuss the latest research findings in the area of EIS and specifically aimed at facilitating the exchange of ideas and advances on all aspects and developments of EIS. The 25 papers presented in this volume were carefully reviewed and selected from 63 submissions. They were organized in topical sections on: semantic concepts and open data; customer relationship management; security and privacy

issues; advanced manufacturing and management aspects; business intelligence and big data; decision support in EIS; and EIS-practices.

Performance Evaluation and Benchmarking for the Era of Artificial Intelligence -

Raghunath Nambiar 2019-01-29

This book constitutes the thoroughly refereed post-conference proceedings of the 10th TPC Technology Conference on Performance Evaluation and Benchmarking, TPCTC 2018, held in conjunction with the 44th International Conference on Very Large Databases (VLDB 2018) in August 2018. The 10 papers presented were carefully reviewed and selected from numerous submissions. The TPC encourages researchers and industry experts to present and debate novel ideas and methodologies in performance evaluation, measurement, and characterization.

Performance Evaluation and Benchmarking

- Raghunath Nambiar 2021-08-03

This book constitutes the refereed post-conference proceedings of the 12th TPC Technology Conference on Performance Evaluation and Benchmarking, TPCTC 2020, held in August 2020. The 8 papers presented were carefully reviewed and cover the following topics: testing ACID compliance in the LDBC social network benchmark; experimental performance evaluation of stream processing engines made easy; revisiting issues in benchmarking metric selection; performance evaluation for digital transformation; experimental comparison of relational and NoSQL document systems; a framework for supporting repetition and evaluation in the process of cloud-based DBMS performance benchmarking; benchmarking AI inference; a domain independent benchmark evolution model for the transaction processing performance council.

Intelligent Data Engineering and Automated Learning -- IDEAL 2013 - Hujun Yin 2013-10-16

This book constitutes the refereed proceedings of the 14th International Conference on Intelligent Data Engineering and Automated Learning, IDEAL 2013, held in Hefei, China, in October 2013. The 76 revised full papers presented were carefully reviewed and selected from more than 130 submissions. These papers provided a valuable collection of latest research outcomes in data engineering and automated learning, from methodologies, frameworks and techniques to applications. In addition to various topics such as evolutionary algorithms, neural networks, probabilistic modelling, swarm intelligent, multi-objective optimisation, and practical applications in regression, classification, clustering, biological data processing, text processing, video analysis, including a number of special sessions on emerging topics such as adaptation and learning multi-agent systems, big data, swarm intelligence and data mining, and combining learning and optimisation in intelligent data

engineering.

New Horizons for a Data-Driven Economy - José María Cavanillas 2016-04-04

In this book readers will find technological discussions on the existing and emerging technologies across the different stages of the big data value chain. They will learn about legal aspects of big data, the social impact, and about education needs and requirements. And they will discover the business perspective and how big data technology can be exploited to deliver value within different sectors of the economy. The book is structured in four parts: Part I “The Big Data Opportunity” explores the value potential of big data with a particular focus on the European context. It also describes the legal, business and social dimensions that need to be addressed, and briefly introduces the European Commission’s BIG project. Part II “The Big Data Value Chain” details the complete big data lifecycle from a technical point of view, ranging from data acquisition, analysis, curation and

storage, to data usage and exploitation. Next, Part III “Usage and Exploitation of Big Data” illustrates the value creation possibilities of big data applications in various sectors, including industry, healthcare, finance, energy, media and public services. Finally, Part IV “A Roadmap for Big Data Research” identifies and prioritizes the cross-sectorial requirements for big data research, and outlines the most urgent and challenging technological, economic, political and societal issues for big data in Europe. This compendium summarizes more than two years of work performed by a leading group of major European research centers and industries in the context of the BIG project. It brings together research findings, forecasts and estimates related to this challenging technological context that is becoming the major axis of the new digitally transformed business environment.
Advancing Big Data Benchmarks - Tilmann Rabl
2014-10-08

This book constitutes the thoroughly refereed

joint proceedings of the Third and Fourth Workshop on Big Data Benchmarking. The third WBDB was held in Xi'an, China, in July 2013 and the Fourth WBDB was held in San José, CA, USA, in October, 2013. The 15 papers presented in this book were carefully reviewed and selected from 33 presentations. They focus on big data benchmarks; applications and scenarios; tools, systems and surveys.
Supercomputing Frontiers - Rio Yokota
2018-03-20

It constitutes the refereed proceedings of the 4th Asian Supercomputing Conference, SCFA 2018, held in Singapore in March 2018. Supercomputing Frontiers will be rebranded as Supercomputing Frontiers Asia (SCFA), which serves as the technical programme for SCA18. The technical programme for SCA18 consists of four tracks: Application, Algorithms & Libraries Programming System Software Architecture, Network/Communications & Management Data, Storage & Visualisation The 20 papers presented

in this volume were carefully reviewed and selected from 60 submissions.

The Internet of Things and Big Data Analytics - Pethuru Raj 2020-06-07

This book comprehensively conveys the theoretical and practical aspects of IoT and big data analytics with the solid contributions from practitioners as well as academicians. This book examines and expounds the unique capabilities of the big data analytics platforms in capturing, cleansing and crunching IoT device/sensor data in order to extricate actionable insights. A number of experimental case studies and real-world scenarios are incorporated in this book in order to instigate our book readers. This book Analyzes current research and development in the domains of IoT and big data analytics Gives an overview of latest trends and transitions happening in the IoT data analytics space Illustrates the various platforms, processes, patterns, and practices for simplifying and streamlining IoT data analytics The Internet of

Things and Big Data Analytics: Integrated Platforms and Industry Use Cases examines and accentuates how the multiple challenges at the cusp of IoT and big data can be fully met. The device ecosystem is growing steadily. It is forecast that there will be billions of connected devices in the years to come. When these IoT devices, resource-constrained as well as resource-intensive, interact with one another locally and remotely, the amount of multi-structured data generated, collected, and stored is bound to grow exponentially. Another prominent trend is the integration of IoT devices with cloud-based applications, services, infrastructures, middleware solutions, and databases. This book examines the pioneering technologies and tools emerging and evolving in order to collect, pre-process, store, process and analyze data heaps in order to disentangle actionable insights.

Assured Cloud Computing - Roy H. Campbell 2018-08-06

Explores key challenges and solutions to assured cloud computing today and provides a provocative look at the face of cloud computing tomorrow. This book offers readers a comprehensive suite of solutions for resolving many of the key challenges to achieving high levels of assurance in cloud computing. The distillation of critical research findings generated by the Assured Cloud Computing Center of Excellence (ACC-UCoE) of the University of Illinois, Urbana-Champaign, it provides unique insights into the current and future shape of robust, dependable, and secure cloud-based computing and data cyberinfrastructures. A survivable and distributed cloud-computing-based infrastructure can enable the configuration of any dynamic systems-of-systems that contain both trusted and partially trusted resources and services sourced from multiple organizations. To assure mission-critical computations and workflows that rely on such systems-of-systems

it is necessary to ensure that a given configuration does not violate any security or reliability requirements. Furthermore, it is necessary to model the trustworthiness of a workflow or computation fulfillment to a high level of assurance. In presenting the substance of the work done by the ACC-UCoE, this book provides a vision for assured cloud computing illustrating how individual research contributions relate to each other and to the big picture of assured cloud computing. In addition, the book: Explores dominant themes in cloud-based systems, including design correctness, support for big data and analytics, monitoring and detection, network considerations, and performance. Synthesizes heavily cited earlier work on topics such as DARE, trust mechanisms, and elastic graphs, as well as newer research findings on topics, including R-Storm, and RAMP transactions. Addresses assured cloud computing concerns such as game theory, stream processing, storage, algorithms, workflow,

scheduling, access control, formal analysis of safety, and streaming Bringing together the freshest thinking and applications in one of today's most important topics, Assured Cloud Computing is a must-read for researchers and professionals in the fields of computer science and engineering, especially those working within industrial, military, and governmental contexts. It is also a valuable reference for advanced students of computer science.

Big Data, Analytics, and the Future of Marketing and Sales - Mckinsey Chief Marketing & Sales Officer Forum 2014-08-02
Big Data is the biggest game-changing opportunity for marketing and sales since the Internet went mainstream almost 20 years ago. The data big bang has unleashed torrents of terabytes about everything from customer behaviors to weather patterns to demographic consumer shifts in emerging markets. This collection of articles, videos, interviews, and slideshares highlights the most important

lessons for companies looking to turn data into above-market growth: Using analytics to identify valuable business opportunities from the data to drive decisions and improve marketing return on investment (MROI) Turning those insights into well-designed products and offers that delight customers Delivering those products and offers effectively to the marketplace. The goldmine of data represents a pivot-point moment for marketing and sales leaders. Companies that inject big data and analytics into their operations show productivity rates and profitability that are 5 percent to 6 percent higher than those of their peers. That's an advantage no company can afford to ignore. Encyclopedia of Information Science and Technology, Fourth Edition - Khosrow-Pour, D.B.A., Mehdi 2017-06-20

In recent years, our world has experienced a profound shift and progression in available computing and knowledge sharing innovations. These emerging advancements have developed

at a rapid pace, disseminating into and affecting numerous aspects of contemporary society. This has created a pivotal need for an innovative compendium encompassing the latest trends, concepts, and issues surrounding this relevant discipline area. During the past 15 years, the Encyclopedia of Information Science and Technology has become recognized as one of the landmark sources of the latest knowledge and discoveries in this discipline. The Encyclopedia of Information Science and Technology, Fourth Edition is a 10-volume set which includes 705 original and previously unpublished research articles covering a full range of perspectives, applications, and techniques contributed by thousands of experts and researchers from around the globe. This authoritative encyclopedia is an all-encompassing, well-established reference source that is ideally designed to disseminate the most forward-thinking and diverse research findings. With critical perspectives on the impact of

information science management and new technologies in modern settings, including but not limited to computer science, education, healthcare, government, engineering, business, and natural and physical sciences, it is a pivotal and relevant source of knowledge that will benefit every professional within the field of information science and technology and is an invaluable addition to every academic and corporate library.

Incorporating Nature-Inspired Paradigms in Computational Applications - Khosrow-Pour, Mehdi 2018-04-06

Many techniques have been developed to control the variety of dynamic systems. To develop those control techniques, it is fundamental to know the mathematical relations between the system inputs and outputs. *Incorporating Nature-Inspired Paradigms in Computational Applications* is a critical scholarly resource that examines the application of nature-inspired paradigms on system identification. Featuring

coverage on a broad range of topics such as biogeographic computation, evolutionary control systems, and natural computing, this book is geared towards IT professionals, engineers, computer scientists, academicians, researchers, and graduate-level students seeking current research on the application of nature-inspired paradigms.

Big Data Benchmarks, Performance Optimization, and Emerging Hardware -

Jianfeng Zhan 2014-11-10

This book constitutes the thoroughly revised selected papers of the 4th and 5th workshops on Big Data Benchmarks, Performance Optimization, and Emerging Hardware, BPOE 4 and BPOE 5, held respectively in Salt Lake City, in March 2014, and in Hangzhou, in September 2014. The 16 papers presented were carefully reviewed and selected from 30 submissions. Both workshops focus on architecture and system support for big data systems, such as benchmarking; workload characterization;

performance optimization and evaluation; emerging hardware.

Big Data Benchmarking - Tilmann Rabl
2016-11-30

This book constitutes the thoroughly refereed post-workshop proceedings of the 6th International Workshop on Big Data Benchmarking, WBDB 2015, held in Toronto, ON, Canada, in June 2015 and the 7th International Workshop, WBDB 2015, held in New Delhi, India, in December 2015. The 8 full papers presented in this book were carefully reviewed and selected from 22 submissions. They deal with recent trends in big data and HPC convergence, new proposals for big data benchmarking, as well as tooling and performance results.

Role of Edge Analytics in Sustainable Smart City Development - G. R.

Kanagachidambaresan 2020-07-29

Efficient Single Board Computers (SBCs) and advanced VLSI systems have resulted in edge

analytics and faster decision making. The QoS parameters like energy, delay, reliability, security, and throughput should be improved on seeking better intelligent expert systems. The resource constraints in the Edge devices, challenges the researchers to meet the required QoS. Since these devices and components work in a remote unattended environment, an optimum methodology to improve its lifetime has become mandatory. Continuous monitoring of events is mandatory to avoid tragic situations; it can only be enabled by providing high QoS. The applications of IoT in digital twin development, health care, traffic analysis, home surveillance, intelligent agriculture monitoring, defense and all common day to day activities have resulted in pioneering embedded devices, which can offer high computational facility without much latency and delay. The book address industrial problems in designing expert system and IoT applications. It provides novel survey and case study report on recent industrial approach towards Smart

City development.

Hands-On Julia Programming - Sambit Kumar Dash 2021-10-21

Build production-ready machine learning and NLP systems using functional programming, development platforms, and cloud deployment.

KEY FEATURES ● In-depth explanation and code samples highlighting the features of the Julia language. ● Extensive coverage of the Julia development ecosystem, package management, DevOps environment integration, and performance management tools. ● Exposure to the most important Julia packages that aid in Data and Text Analytics and Deep Learning.

DESCRIPTION The Julia Programming language enables data scientists and programmers to create prototypes without sacrificing performance. Nonetheless, skeptics question its readiness for production deployments as a new platform with a 1.0 release in 2018. This book removes these doubts and offers a comprehensive glimpse at the language's use

throughout developing and deploying production-ready applications. The first part of the book teaches experienced programmers and scientists about the Julia language features in great detail. The second part consists of gaining hands-on experience with the development environment, debugging, programming guidelines, package management, and cloud deployment strategies. In the final section, readers are introduced to a variety of third-party packages available in the Julia ecosystem for Data Processing, Text Analytics, and developing Deep Learning models. This book provides an extensive overview of the programming language and broadens understanding of the Julia ecosystem. As a result, it assists programmers, scientists, and information architects in selecting Julia for their next production deployments. **WHAT YOU WILL LEARN** ● Get to know the complete fundamentals of Julia programming. ● Explore Julia development frameworks and how to work

with them. ● Dig deeper into the concepts and applications of functional programming. ● Uncover the Julia infrastructure for development, testing, and deployment. ● Learn to practice Julia libraries and the Julia package ecosystem. ● Processing Data, Deep Learning, and Natural Language Processing with Julia. **WHO THIS BOOK IS FOR** This book is for Data Scientists and application developers who want to learn about Julia application development. No prior Julia knowledge is required but knowing the basics of programming helps understand the objectives of this book. **TABLE OF CONTENTS** 1. Getting Started 2. Data Types 3. Conditions, Control Flow, and Iterations 4. Functions and Methods 5. Collections 6. Arrays 7. Strings 8. Metaprogramming 9. Standard Libraries Module 2. The Development Environment 10. Programming Guidelines in Julia 11. Performance Management 12. IDE and Debugging 13. Package Management 14. Deployment Module 3. Packages in Julia 15.

Data Transformations 16. Text Analytics 17.
Deep Learning
Encyclopedia of Big Data Technologies - Sherif
Sakr 2019-03-01

The Encyclopedia of Big Data Technologies provides researchers, educators, students and industry professionals with a comprehensive authority over the most relevant Big Data Technology concepts. With over 300 articles written by worldwide subject matter experts from both industry and academia, the encyclopedia covers topics such as big data storage systems, NoSQL database, cloud computing, distributed systems, data processing, data management, machine learning and social technologies, data science. Each peer-reviewed, highly structured entry provides the reader with basic terminology, subject overviews, key research results, application examples, future directions, cross references and a bibliography. The entries are expository and tutorial, making this reference a practical resource for students,

academics, or professionals. In addition, the distinguished, international editorial board of the encyclopedia consists of well-respected scholars, each developing topics based upon their expertise.

Database Systems for Advanced

Applications - Sourav S. Bhowmick 2014-04-16
These two volumes set LNCS 8421 and LNCS 8422 constitutes the refereed proceedings of the 19th International Conference on Database Systems for Advanced Applications, DASFAA 2014, held in Bali, Indonesia, in April 2014. The 62 revised full papers presented together with 1 extended abstract paper, 4 industrial papers, 6 demo presentations, 3 tutorials and 1 panel paper were carefully reviewed and selected from a total of 257 submissions. The papers cover the following topics: big data management, indexing and query processing, graph data management, spatio-temporal data management, database for emerging hardware, data mining, probabilistic and uncertain data management, web and social

data management, security, privacy and trust,

keyword search, data stream management and data quality.