

Preface

Data centers play an important role in modern IT infrastructures. A data center is a home to computational power, storage, and applications necessary to support an enterprise business. Data centers process billions of Internet transactions every day. Nowadays, data centers are the beating hearts of the companies they serve. Large data centers with thousands of servers have been deployed by renowned ICT organizations, like IBM, Microsoft, Amazon, Yahoo, E-Bay, and Google, to provide cloud computing services. The phenomenal increase in the size and number of data centers and resultant increase in operational cost has stimulated the research in various domains of data centers, such as energy efficiency, resource management, networking, and security. This detailed *Handbook on Data Centers* covers in a succinct and orderly manner all aspects pertaining to data centers technologies.

Given the fast growing expansion of data centers, it is not surprising that a variety of methods are now available to researchers and practitioners striving to improve data center performance. This handbook aims to organize all major concepts, theories, methodologies, trends, challenges and applications of data centers into a coherent and unified repository. Moreover, the handbook provides researchers, scholars, students and professionals with a comprehensive, yet concise source of reference to data centers design, energy efficiency, resource management, and scalability issues. The handbook consists of nine parts, where each part consists of several chapters. The following topics spanning the data center technologies are covered in detail: (a) energy efficiency, (b) networking, (c) cloud computing, (d) hardware, (e) modeling and simulation, (f) security, (g) data services, (h) monitoring, and (i) resource management. Each part describes the state of the art methods, as well as the extensions and novel methodologies developed recently for data centers.

The first part describes energy efficiency methods for data centers and addresses the designing of energy efficient scheduling mechanisms for high performance computing environments. The authors have discussed energy-aware algorithms for task graph scheduling, replica placement, check-pointing, numerical processing, and to estimate energy consumption for the given workload. Also, this part addresses the problem of minimizing energy consumption with schedule length constraint on multicore processors. The following chapters covers the energy efficiency aspects of data centers presented in the handbook: (a) *Energy-Efficient and High-Performance*

Processing of Large-Scale Parallel Applications in Data Centers, (b) Energy-aware algorithms for task graph scheduling, replica placement and checkpoint strategies, (c) Energy efficiency in HPC Data Centers: Latest Advances to Build the Path to Exascale, (d) Techniques to achieve energy proportionality in data centers: a survey, (e) A Power-Aware Autonomic Approach for Performance Management of Scientific Applications in a Data Center Environment, (f) CoolEmAll: Models and Tools for Planning and Operating Energy Efficient Data Centres, (g) Smart Data center, (h) Power and Thermal Efficient Numerical Processing, and (k) Providing Green Services in HPC Data Centers: A Methodology based on Energy Estimation.

The second part of the handbook provides a study of various communication and networking methodologies for data centers. The network virtualization concepts, the architecture of optical networks for data centers, and network scalability issues are discussed in detail. Moreover, an emphasis is drawn over packet classification in multicore platforms. The routing techniques for data center networks are discussed in detail along with study on TCP congestion control in data center networks. The following chapters provide a detailed overview in terms of networking technologies for data centers are included in this part: *(a) Network Virtualization in Data Centers: A Data Plane Perspective, (b) Optical data center networks: Architecture, performance, and energy efficiency, (c) Scalable Network Communication using Unreliable RDMA, (d) Packet Classification on Multi-core Platforms, (e) Optical Interconnects for Data Center Networks, (f) TCP Congestion Control in Data Center Networks, and (g) Routing Techniques in Data Center Networks.*

The third part of the handbook discusses the role of data centers in cloud computing and highlights various challenges faced in ensuring the data integrity, reliability, and privacy in cloud computing environments. The role of trusted third parties in performing monitoring and auditing of service level agreements is illustrated, as well as integrity of big data in cloud computing is discussed. The data intensive applications in cloud are discussed along with storage challenges. This part includes the following chapters: *(a) Auditing for Data Integrity and Reliability in Cloud Storage, (b) I/O and File Systems for Data-Intensive Applications, (c) Cloud resource pricing under tenant rationality, (d) Online Resource Management for Carbon-Neutral Cloud Computing, (e) A Big Picture of Integrity Verification of Big Data in Cloud Computing, (f) An Out-of-Core Task-based Middleware for Data-Intensive Scientific Computing, (g) Building Scalable Software for Data Centers: An Approach to Distributed Computing at Enterprise Level, and (h) Cloud Storage over Multiple Data Centers.*

The fourth part of the handbook shed a light on data centers emerging hardware technologies. The issues pertaining to the efficient data storage on redundant array of independent disks (RAID) technologies, the data synchronization challenges on many-cores, and hardware approaches to transactional memory on chip multiprocessors are detailed. The following chapters are included in this part: *(a) Realizing Accelerated Cost-Effective Distributed RAID, (b) Efficient Hardware-Supported Synchronization Mechanisms for Many-cores, and (c) Hardware Approaches to Transactional Memory in Chip Multiprocessors*

The fifth part of the handbook discusses modeling and simulation techniques for data centers and include following chapters: (a) *Data Center Modeling and Simulation Using OMNeT++*, (b) *Power-Thermal Modeling and Control of Energy-Efficient Servers and Datacenters*, (c) *Thermal modeling and management of storage systems in data centers*, and (d) *Modeling and Simulation of Data Center Networks*. This portion highlights various techniques to model power consumption, thermal response, storage systems, and communication in data center networks. The authors utilized discrete-time simulator OMNet++ to model the data center traffic.

The sixth part of the handbook provides a discussion on various security and privacy techniques for data centers. A model is proposed to detect and mitigate the covert channels in data centers. The privacy issues regarding data center outsourcing are discussed and a survey of various privacy attacks and their counter measures is presented. The following chapters are included in this portion: (a) *C2Hunter: Detection and Mitigation of Covert Channels in Data Centers*, (b) *Selective and Private Access to Outsourced Data Centers*, and (c) *Privacy in Data Centers: A Survey of Attacks and Countermeasures*.

The seventh part details the data services and their management in data center. The quality of service requirements for processing stream data for city applications is discussed. The data management and querying of big data is surveyed in detail. The authors also discussed the various methods of constructing on the fly data centers in wireless ad hoc network environments. The following chapters are included in this part: (a) *Quality-of-Service in Data Center Stream Processing for Smart City Applications*, (b) *Opportunistic Databank: A context aware on-the-fly data center for mobile networks*, (c) *Data Management: State-of-the-Practice at Open-Science Data Centers*, and (d) *Data Summarization Techniques for Big Data—A Survey*.

The eighth part of handbook illustrates various hardware and software-based monitoring solutions for data centers. The use of wireless sensor networks technology to sense the thermal activities and efficient circulation of cooling air is explored. Traffic monitoring is another vital topic discussed with a survey on network intrusion detection systems for data centers. The following chapters are covered in this portion: (a) *Central Management of Datacenters*, (b) *Monitoring of Data Centers using Wireless Sensor Networks*, (c) *Network Intrusion Detection Systems in Data Centers*, and (d) *Software Monitoring in Data Centers*.

Lastly, the ninth part of book chapter elaborates the resource management aspects of data centers and present a detailed case study on usage patterns in multi-tenant data centers. A discussion is also performed on scheduling of distributed transactional memory and the various tradeoffs and techniques in distributed memory management are elaborated in detail. Moreover, an emphasis is drawn on the resource scheduling in data-centric systems. This part includes the following chapters: (a) *Usage Patterns in Multi-tenant Data Centers: a Large-Case Field Study*, (b) *On Scheduling in Distributed Transactional Memory: Techniques and Tradeoffs*, (c) *Dependability-Oriented Resource Management Schemes for Cloud Computing Data Centers*, and (d) *Resource Scheduling in Data-Centric Systems*.

<http://www.springer.com/978-1-4939-2091-4>

Handbook on Data Centers

Khan, S.U.; Zomaya, A.Y. (Eds.)

2015, XIII, 1334 p. 439 illus., 283 illus. in color.,

Hardcover

ISBN: 978-1-4939-2091-4