Review of the book
”Cloud Computing and Software Services - Theory and Techniques”
edited by Syed A. Ahson and Mohammad Ilyas
CRC Press, Taylor & Francis Group, 2011

Dominik Birk

Horst Goertz Institute for IT Security - Ruhr-University Bochum

1   Summary of the review

The book ”Cloud Computing and Software Services” tries to serve as a single comprehensive source of information on the emerging notion of cloud computing. In 17 chapters, different aspects and technologies in the field of cloud computing are discussed by various authors from industry and academia. It has to be emphasized that this book is primarily not security related, however provides interesting background information on cloud technologies. This information is of high importance also for researchers working in the security field as cloud technologies have to be completely understood in order to make distributed systems and infrastructure as secure as possible.

2   Summary of the book

- **Chapter 1: Understanding the Cloud Computing Landscape**

  Within this chapter, the SPI (SaaS, PaaS, IaaS) cloud classification is discussed and different ontologies for clouds are introduced. Therefore, this chapter can be considered as an introduction to cloud technologies in general. Albeit the intention of this chapter is straightforward, it is questionable why the SPI model is only briefly discussed related to the other models. To the best of the reviewer's knowledge, SPI is the most often used service model for cloud computing offers so far.

- **Chapter 2: Science Gateways: Harnessing Clouds and Software Services for Science**

  Chapter 2 is about cloud computing technologies in the field of science. At first, the authors motivate cloud technology for the scientific community. Afterwards, various service models (e.g. IaaS, SaaS) and concrete system environments for scientific purposes are discussed.

- **Chapter 3: Enterprise Knowledge Clouds: Next Generation Knowledge Management Systems**

  Although Enterprise Knowledge Management (EKM) Systems do not seem to be a substantial part of current cloud computing technology on the first view, the authors showed in chapter 3 that in fact they are. They state that knowledge management is and will be a major issue especially in the enterprise field. Cloud technologies could be used to improve this situation and get the overwhelming amount of data and information managed.

- **Chapter 4: Real Cases and Applications of Cloud Computing**

  At the very first view, chapter 4 provides an interesting topic of clouds: Real world cases and applications of cloud computing. However, already the outline seems to be quite confusing. In section 4.2 the authors focus on cloud computing security on almost three pages which looks like the security topic wasn't worth more. In section 4.3 the different strategical phases are discussed

which are needed in case you want to move your data to the cloud. This was interesting to read albeit pretty straightforward. In the following sections, use cases of clouds were discussed unfortunately without digging deeper into technical issues.

- **Chapter 5: Large Scale Data Processing**

  Readers who are not familiar with the notion of "MapReduce" and "GridBatch" should not skip this chapter. The authors did a good job in explaining the MapReduce Framework and the implementation on a Cloud OS. The chapter is well written and relatively easy to understand.

- **Chapter 6: Toward a Reliable Cloud Computing Service**

  One of the challenges current CSP have to face is provisioning of services. How many resources should be provisioned to reliably service a workload? And what is the probability of blocking new requests at a specific resource provisioning level? These questions are discussed in chapter 6 with the help of a mathematical model.

- **Chapter 7: Abstractions for Cloud Computing with Condor**

  Generally, the MapReduce abstraction allows the user to specify two functions that transform and summarize data. If the desired computation can be expressed using this abstraction, then the computation can be scaled up to thousands of nodes. The authors of chapter 7 bring up different abstractions and discuss the issue coming up if the MapReduce abstraction cannot be applied but computations still should be scalable.

- **Chapter 8: Exploiting the Cloud of Computing Environments: An Application's Perspective**

  Within chapter 8, different types of computing environments are examined and their characteristics and example deployments are introduced. The authors also describe different classes of scientific applications that are being used across the domains of science today.

- **Chapter 9: Granules: A Lightweight Runtime for Scalable Computing with Support for Map-Reduce**

  Chapter 9 discusses the notion of "Granules", a lightweight streaming-based runtime for clouds computing environments. Granules allows processing tasks to be deployed on a single resource or a set of resources. Furthermore, besides the basic support for MapReduce, different variants of the MapReduce framework are also supported.

- **Chapter 10: Dynamic and Adaptive Rule-Based Workflow Engine for Scientific Problems in Distributed Environments**

  Within chapter 10, a brief overview on scientific workflows seen from a distributed system perspective is given. The authors argue that given the increased need of larger storage and computational resources, workflows have migrated toward a distributed approach by using web services.

- **Chapter 11: Transparent Cross-Plattform Access to Software Services Using Grid-Solve and GridRPC**

  The authors discuss "GridRPC", a model for remote procedure call (RPC) in the context of a computational grid or other loosely coupled distributed computing environment. Furthermore, "GridSolve", an implementation of the GridRPC model is introduced.

- **Chapter 12: High-Performance Parallel Computing with Cloud and Cloud Technologies**

  The authors of chapter 12 discuss limitations of current cloud technologies for parallel applications. Several questions are discussed: What applications are best handled by cloud technologies? What overheads do they introduce? Are there any alternatives? Moreover, the authors propose several different studies of clouds and cloud technologies on both real applications and standard benchmark.

- **Chapter 13: BioVLAB: Bioinformatics Data Analysis Using Cloud Computing and Graphical Workflow Composers**

  BioVLAB is a novel bio-informatics computing architecture using AWS and discussed in chapter 13. This chapter is a must-read for people working in computational medicine of bio-informatics in general as it contains a lot of information how clouds could be used in the context of bio-informatics.

- **Chapter 14: Scale-Out RDF Molecule Store for Efficient, Scalable Data Integration and Querying**

  Again, this chapter 14 also focuses on a bioinformatics - the authors try to apply semantic web technologies to molecular biology data.

- **Chapter 15: Enabling XML Capability for Hadoop and Its Applications in Healthcare**

  This chapter describes the architectural design and features of Xbase, the first healthcare-specific analytic engine to be built on top of existing cloud infrastructure Hadoop and RDBMS, to meet the requirements of advanced healthcare applications.

- **Chapter 16: Toward a QoS-Focused SaaS Evaluation Model**

  This chapter 16 discusses a model framework for evaluating SaaS applications based on QoS characteristics and forms the basis for a toolset to assist the IT planning process.

- **Chapter 17: Risk Evaluation-based Selection Apporach for Transactional Services Composition**

  According to the authors of chapter 17, todays QoS-driven service selection methods for web services composition (WSC) ignore the failure risk impact of each task, and few works focus on WSC with transactional properties. Hence, the authors present a risk-driven services selection method for WSC with transactional properties.

This book tries to bring clarity into the emerging field of cloud computing by presenting accurate, up-to-date information in a comprehensive manner - unfortunately it does not achieve this goal completely. After finishing the book, the reader still has several questions concerning "what" exactly is cloud computing and which technology is just ordinary and well researched for decades. This is due to the controversial discussion around the notion of cloud computing. Without the establishment of a concrete definition for cloud environments, the complete technology can hardly be understood by engineers. Additionally, without a concrete definition, security issues cannot be comprehensively solved.

3 What is the book like (style)?

The book is a collection of 17 articles on the subject of cloud computing and software services. Most of the articles are written by well known authors from academia working in the fields of computer science. Due to this structure, the book cannot be considered as an ordinary textbook but more as a book focusing on specific scenarios in the context of cloud computing. Buy this book if you like to further improve your knowledge about clouds.

4 Would you recommend this book?

The reviewed book is not written for beginners - basic knowledge of distributed systems and infrastructures should be known to the reader in order to fully understand the different articles. Hence, this book is recommended to readers interested in the deep technical basics of cloud infrastructures and systems.

The reviewer is a Ph.D. student at the Horst Goertz Institute for IT Security.