Preface

BACKGROUND

Infrastructure as a Service (IaaS) cloud systems allow the dynamic creation, destruction and management of Virtual Machines (VM) on virtualized clusters. IaaS clouds provide a high-level of abstraction to the end user that allows the creation of on-demand services through a pay-as-you-go infrastructure combined with elasticity. The increasing range of choices and availability of IaaS toolkits has also allowed creation of cloud solutions and frameworks suitable for private deployment and practical use even on smaller scales. As a result, many academic infrastructure service providers have started transitions to add cloud resources to their previously existing campus and shared grid deployments. To complete such solutions, they should also support the unification of multiple cloud and/or cloud and grid solutions in a seamless, preferably interoperable way. Hybrid, community or multi-clouds may utilize more than one cloud systems, which are also called cloud federations. The management of such federations raises several challenges and opens issues that require significant research work to be done in this area.

The Positioning of the Book

The more widespread cloud computing technologies are, the more likely people will face the interoperability issues when several cloud infrastructures must be used in parallel. This book is well suited to the practitioners who utilize cloud infrastructures and would like to avoid lock in issues or to increase the reliability of their virtual infrastructures. On the other hand, with the rise of private cloud infrastructures, hybrid clouds, cloud bursting technologies and partial outsourcing, the solutions offered in this book aid the private infrastructure providers to help efficiently deal with temporal under-provisioning situations.

Whom Is This Book Intended For?

This book provides a dedicated forum for sharing the latest results, exchanging ideas and experiences, presenting new research, development and deployment efforts in developing and running interoperable, federated IaaS cloud systems. It is aimed at defining the current state, determining further goals and presenting architectures and service frameworks to achieve highly interoperable federated cloud infrastructures. It presents solutions to interoperability and efficient management challenges faced by current and future infrastructure clouds, with a specific focus on cloud federation aimed at readers with specific interest in topics not addressed elsewhere such as federation policies, energy awareness, federation use cases, legal aspects of federation, scheduling and interoperability.
OVERVIEW OF THE SUBMISSIONS AND THE ACCEPTED CHAPTERS

The ‘Call for chapter proposals’ for the book was launched in winter 2014/15 and resulted 22 submissions, which were of good quality and generally relevant to the theme of the special issue. The chapter proposals, submitted chapters and their subsequent revisions were reviewed both by the experts in the field and by the guest editors - this procedure guaranteed that each paper received at least 2-3 reviews. After the initial selection process, the chapters were sent for one further revision round. Finally, we have selected the 8 strongest chapters (giving us a 36% acceptance rate). These chapters represent three major research areas, requirements and solutions for creating federated cloud infrastructures. These chapters are organized into two major sections. The first section overviews the theoretical background and provides an analysis of current research on cloud federations and multi-clouds, in general. The second section is more technical and provides in-depth details of particular topics. To strengthen the sections, we have also invited 3 additional papers. One on the legal aspects of the topic, how European data protection laws affect cloud federations. The second additional paper presents a H2020 funded European research project (called ENTICE) and presents how it envisions altering the current federated cloud landscape with a novel view on virtual machine image delivery techniques. Finally, the third additional paper presents the aspect of volunteer clouds to conclude the book with an interesting outlook for potential future practices. In the followings we provide a short overview of the selected papers.

Chapter 1, is titled “Cloud Federations: Requirements, Properties, and Architectures”, and presents a generic view on Cloud Federations. It summarizes the inherent characteristics of this paradigm and the limitations many cloud providers face. The chapter argues that multiple cloud organizations open the opportunity for the providers to utilize resources with more attractive prices, increase the resilience and monetize their own idle resources. When considering customers, problems such as interruption of services, lack of interoperability that lead to lock-in and loss of quality of services due to locality are presented as limiting factors to the adoption of Cloud Computing. The chapter presents an introduction to the conceptual characterization of Cloud Federation solutions, and highlights the challenges of implementing federated architectures, the requirements for the development of this type of organization and the relevant proposed architectures.

Chapter 2, “Regulating European Clouds - The New European Data Protection Framework” is one of the additional papers that introduces legal aspects of federated cloud computing. In the supply of any goods and services, the law gives certain rights that protect the consumer and provider, and this also applies for Cloud Computing services. This new technology moves functions and responsibilities away from local ownership and management to a service provided by third-parties and raises several legal issues, such as data protection, which require the service to comply with certain regulation. This chapter overviews and analyses the revised legislation of the European Union resulting in the General Data Protection Regulation, which will be used to set up the new European Data Protection Framework. The author gathers and summarizes the most relevant changes this regulation brings to the field of Clouds and relate them to the antecedent legislation called the Data Protection Directive currently in force.

Following, Chapter 3, “Identity and Access Management in the Cloud Computing Environments” explores the issue of identity and access management that became an important research topic in cloud computing recently. Since cloud infrastructures and even more federated cloud infrastructures may collect and provide an enormous quantity and quality of resources, their access must be controlled and effective authentication and authorization are essential to make sure that unauthorized users cannot access the resources, thereby ensuring the validity, confidentiality, integrity, and availability of information.
Preface

hosted in the cloud environment. This chapter discusses the identity and access management by analyzing the work carried out in this area. Also, various issues in the current identity and access management practices in cloud computing, such as authentication, authorization, access control models, identity lifecycle management, cloud identity-as-a-service, federated identity management and also, the identity and access management in the inter-cloud environment are presented.

Chapter 4, “Characterizing PaaS Solutions enabling Cloud Federations” gives an outlook at realizing interoperable cloud services at a higher level above IaaS. Recently many businesses migrate their IT applications and data to the Cloud to take advantage of the flexible resource provisioning that can bring benefits to businesses by responding quickly to new demands from customers. Recent solutions hide the diversity of multiple utilized clouds and form a unified federation on top of them. Many approaches follow recent trends in cloud application development, and offer federation capabilities at the platform level thus, creating Platform-as-a-Service solutions. This chapter overviews the capabilities of these approaches: what levels of developer experience they offer, how they follow recent trends in cloud application development, what types of APIs, developer tools they support and what web GUIs they provide.

Subsequently, Chapter 5, “Highly Available Fault-Tolerant Cloud Database Services” presents a different aspect of utilizing the virtual machines in the cloud. Legacy database systems manage transactions under a concurrency control and a recovery protocol. The underlying operating system creates a transaction execution platform and the database executes transactions concurrently. The availability of virtual machines in clouds however, makes it possible to eliminate the effect of system or transaction failure by always taking the database to the next consistent state. This chapter presents a novel scheme of eliminating such failures by applying transaction “roll-forward” which resumes the execution from the point of failure.

Chapter 6, “Distributed Multi-Cloud Based Building Data Analytics” sets the focus on data storage and data movement in a scenario where computation is dispersed and also defines a hierarchical view of the processing capabilities. The efficient computation of data intensive applications requires understanding how to store, process, and analyze large volumes of data in a timely manner. Many smart city applications, for instance, identify how data from building sensors can be combined together to support applications such as emergency response, energy management, etc. Enabling sensor data to be transmitted to a cloud environment for processing provides a number of benefits, such as scalability and on-demand provisioning of computational resources. This chapter proposes the use of a multi-layer cloud infrastructure that distributes processing over sensing nodes, multiple intermediate / gateway nodes, and large data centres. The presented solution aims at utilizing the pervasive computational capabilities located at the edge of the infrastructure and along the data path to reduce data movement to large data centres located deep into the infrastructure and perform a more efficient use of computing and network resources.

Next, Chapter 7, “dEcentralised repositories for traNsparent and efficienT vRvirtual maChine oPperations: Architecture of the ENTICE Project “ is another additional paper that presents the aims and approaches of a major project on federated clouds. ENTICE is an H2020 European project aimed at creating a novel Virtual Machine (VM) repository and operational environment for federated Cloud infrastructures. The operational environment in focus is envisioned to simplify the creation of lightweight and highly optimized VM images tuned for functional descriptions of applications. Furthermore, it also automatically decomposes and distributes VM images based on multi-objective optimization to meet application runtime requirements, and enables the elastic scaling of applications on Cloud resources, based on their fluctuating load, by optimized VM interoperability across Cloud infrastructures. This chapter gives an inside view into the ENTICE project architecture. Based on stakeholders that interact
with ENTICE, it also describes the different functionalities of the various components and services and how they interact with each other.

Chapter 8, “EUBrazilCC Federated Cloud: A Transatlantic Multi-Cloud Infrastructure” is another overview of a major project on creating an intercontinental federated cloud. This chapter describes the architecture and the deployment of the EUBrazilCC federated e-infrastructure, a Research & Development project that aims at providing a user-centric test bench enabling European and Brazilian research communities to test the deployment and execution of scientific applications on a federated intercontinental e-infrastructure. This e-infrastructure exploits existing resources that consist of virtualized data centers, supercomputers, and even opportunistically exploited desktops spread over a transatlantic geographic area. These heterogeneous resources are federated with the aid of appropriate middleware that provides the necessary features to achieve the established challenges. In order to elicit the requirements and validate the resulting infrastructure, three complex scientific applications have been implemented, which are also presented.

Chapter 9, “Federated IaaS Resource Brokerage” is more towards the in-depth insights and analyses of technical realizations of Service Level Agreements. It presents the CloudAnchor brokerage platform for dealing with both individual providers as well as federated Infrastructure as a Service (IaaS) resources. The platform, which is a layered Multi-Agent System (MAS), provides multiple services including (consumer or provider) business registration and deregistration, provider coalition creation and termination, provider lookup and invitation and negotiation services regarding brokerage, coalitions and resources. Providers, consumers and virtual providers, representing provider coalitions, are modelled by dedicated agents within the platform. The main goal of the platform is to negotiate and establish Service Level Agreements (SLA). In particular, the platform supports multiple notions of SLA and contemplates the establishment of ‘brokerage SLA’ between the platform and each provider or consumer, ‘coalition SLA’ between the members of a coalition of providers and ‘resource SLA’ between a consumer and a provider. Federated resources are detained and negotiated by virtual providers on behalf of the corresponding coalitions of providers.

Chapter 10, “GPGPU as a Service: Providing GPU-Acceleration Services to Federated Cloud Systems” adds further technical challenges and solutions for enhancing cloud federations by advanced computing resources. Current data centers leverage virtual machines in order to efficiently use hardware resources. VMs allow reducing equipment acquisition costs as well as decreasing overall energy consumption. However, although VMs have noticeably evolved to make a smart use of the underlying hardware, the use of GPUs (Graphics Processing Units) for General Purpose computing (GPGPU) is still not efficiently supported. This concern might be addressed by remote GPU virtualization solutions, which may provide VMs with GPUs located in a remote node, detached from the host where the VMs are being executed. This chapter presents an in-depth analysis of providing GPU access to applications running inside VMs. This analysis is complemented with experimental results which show that the use of remote GPU virtualization is an effective mechanism to provide GPU access to applications with negligible overheads. Finally, the approach is presented in the context of cloud federations for providing GPGPU as a Service.

Chapter 11, “Volunteer Clouds: From Volunteer Computing to Interconnected Infrastructures” concludes the book and gives an outlook to potential future directions of realizing federated clouds. Volunteer Computing (VC) is an established and known paradigm that uses volatile, heterogeneous and unreliable resources. This chapter makes an attempt, starting from a definition for Cloud Computing, to identify the required steps and formulate a definition for what can be considered as the next evolutionary stage of Volunteer Computing: Volunteer Clouds (VCI). There are many idiosyncrasies of VC to overcome
Preface

(e.g., volatility, heterogeneity, reliability, responsiveness, scalability, etc.) for example, heterogeneity exists in VC at different levels whereas the vision of cloud computing promises to provide a homogeneous environment. The chapter identifies methods and proposes solutions that tackle the heterogeneity and thus, make a step towards Volunteer Clouds.