Achieving Real-Time in Distributed Computing: From Grids to Clouds

Dimosthenis P. Kyriazis  
*National Technical University of Athens, Greece*

Theodora A. Varvarigou  
*National Technical University of Athens, Greece*

Kleopatra G. Konstanteli  
*National Technical University of Athens, Greece*
Chapter 1
Modelling and Analysing QoS for Real-Time Interactive Applications on the Cloud

Matthew Addis, University of Southampton IT Innovation Centre, UK
Michael Boniface, University of Southampton IT Innovation Centre, UK
Juri Papay, University of Southampton IT Innovation Centre, UK
Arturo Servin, University of Southampton IT Innovation Centre, UK
Zlatko Zlatev, University of Southampton IT Innovation Centre, UK
George Kousiouris, National Technical University Athens, Greece

The complexity of determining resource provisioning policies for applications in such complex environments introduces significant inefficiencies in the cloud. Novel approaches are needed to efficiently model and analyse QoS for such applications, especially those with real-time constraints. This chapter investigates some of the techniques that can be used to explore these trade-offs and to find business models where value can be provided at all stages in the value-chain. For example, how can an application user identify a set of service providers that in combination provide the best solution to their workflow? How much risk is involved, e.g. what is the probability and impact of a failure of a service provider to deliver the QoS they promise? How can Service Level Agreements be specified that provide the flexibility to accommodate variability in service use, yet don’t result in unnecessarily high cost to the consumer due to the service provider having to use massive over-provisioning of resources to ensure they can meet times of peak-load?
Model-driven software engineering (MDE) has the basic assumption that the development of software systems from high-level abstractions along with the generation of low-level implementation code can improve the quality of the systems and at the same time reduce costs and improve time to market. This chapter provides an overview of MDE, state of the art approaches, standards, resources, and tools that support different aspects of model-driven software engineering: language development, modeling services, and real-time applications. The chapter concludes with a reflection over the main challenges faced by projects using the current MDE technologies, pointing out some promising directions for future developments.

Chapter 3
Programming Interfaces for Realtime and Cloud-Based Computing

Gregory Katsaros, National Technical University of Athens, Greece
Tommaso Cucinotta, Scuola Superiore Sant'Anna, Italy

Research in the fields of Grid Computing, Service Oriented Architectures (SOA), as well as virtualization technologies has driven the emergence of Cloud service models such as Software-as-a-Service (SaaS), Platform-as-a-Service (PaaS), and Infrastructure-as-a-Service (IaaS). The appearance of different business roles according to this classification, potentially with differing interests, introduces new challenges with regard to the tools and mechanisms put in place in order to enable the efficient provisioning of services. Security, QoS assurance, and real-time capabilities are just a few issues that the providers are trying to tackle and integrate within the new products and services that they offer. In this chapter, we make an overview of the approaches that aim to APIs for real-time computing. In the first part of this chapter, several real-time application interfaces will be presented and compared. After that, we will document the state-of-the-art regarding the Cloud APIs available and analyze the architecture and the technologies that they support.

Chapter 4
Service Quality Model Evaluation

Fotis Aisopos, National Technical University of Athens, Greece
Magdalini Kardara, National Technical University of Athens, Greece
Vrettos Moulos, National Technical University of Athens, Greece
Athanasios Papaoikonomou, National Technical University of Athens, Greece
Konstantinos Tserpes, National Technical University of Athens, Greece
Theodora A. Varvarigou, National Technical University of Athens, Greece

In this chapter, we present the current state-of-the-art technology and methodologies, regarding the evaluation of the provided QoS in service oriented environment. With the emergence of service provisioning infrastructures and the adoption of Service Level Agreements acting as electronic contracts between service providers and customers, the need to control and validate the offered quality has appeared throughout the service lifecycle. This monitoring is performed either in the client side, using the customer’s quality of experience and employing trust and reputation mechanisms for the service selection and evaluation phase, or in the provider side, dynamically reconfiguring the service and allocating resources accordingly, in order to optimize the quality metrics guaranteed. The latter, of course, initially requires mapping of the high-level quality parameters, which are closer to the customer perception, to low-level computing terms related to the resource management process. Dynamic resource allocation based on monitoring and evaluation can lead to optimizing resource utilization and provider’s profits.
Section 2
Platform as a Service

Chapter 5
Taxonomy and State of the Art of Service Discovery Mechanisms and Their Relation to the Cloud Computing Stack

George Kousiouris, National Technical University of Athens, Greece
Dimosthenis P. Kyriazis, National Technical University of Athens, Greece
Theodora A. Varvarigou, National Technical University of Athens, Greece
Eduardo Oliveros, Telefónica Investigación y Desarrollo, Spain
Patrick Mandic, Rechenzentrum Universität Stuttgart, Germany

Service discovery mechanisms are gaining interest in the last years due to the growing bulk of information available, especially to distributed computing infrastructures like Grids and Clouds. However, a vast number of characteristics of these implementations exist, each one suitable for a number of purposes. The aim of this chapter is to extract a taxonomy of these characteristics found in modern Service Discovery systems and produce a categorization of existing implementations in a grouped and comparative way, based on these features. Furthermore, the mapping of these characteristics to the Cloud business model is produced, in order to assist in selecting the suitable solutions for each provider based on his/her location in the value chain or identify gaps in the existing implementations.

Chapter 6
Monitoring and Metering in the Cloud

Eduardo Oliveros, Telefónica Investigación y Desarrollo, Spain
Tommaso Cucinotta, Scuola Superiore Sant'Anna, Italy
Stephen C. Phillips, University of Southampton IT Innovation Centre, UK
Xiaoyu Yang, University of Southampton IT Innovation Centre, UK
Stuart Middleton, University of Southampton IT Innovation Centre, UK
Thomas Voith, Alcatel-Lucent, Germany

Monitoring and Metering are essential activities for Service Oriented Infrastructures (SOI) and Cloud services. The information collected through monitoring is necessary to ensure the correct execution of the applications in the Cloud and the monitoring of the SLA compliance. This chapter will present the reasons and difficulties for monitoring and metering on Cloud infrastructures. The approaches for monitoring of the execution environment and the network on virtualised infrastructures will be described together with the existing monitoring tools present on different commercial and research platforms.

Chapter 7
Workflow Management Systems in Distributed Environments

Spyridon V. Gogouvitis, National Technical University of Athens, Greece
Kleopatra G. Konstanteli, National Technical University of Athens, Greece
Dimosthenis P. Kyriazis, National Technical University of Athens, Greece
Gregory Katsaros, National Technical University of Athens, Greece
Tommaso Cucinotta, Scuola Superiore Sant'Anna, Italy
Michael Boniface, University of Southampton IT Innovation Centre, UK
With the advent of Service Oriented Architectures, more applications are build in a distributed manner based on loose coupled services. In this context, workflow management systems play an important role, as they are the means to both define the processes that realize the application goals and implement the orchestration of the different services. The purpose of this chapter is to give an overview of various solutions regarding workflow semantics and languages as well as their enactment within the scope of distributed systems. To this end, major focus is given to solutions that are aimed at Grid environments. Scheduling algorithms and advance reservation techniques are also discussed as these are one of the hottest research topics in workflow management systems.

Chapter 8
Service Level Agreements for Real-Time Service-Oriented Infrastructures ........................................... 133
Roland Kübert, University of Stuttgart, Germany
Georgina Gallizo, University of Stuttgart, Germany
Theodoros Polychniatis, National Technical University of Athens, Greece
Theodora A. Varvarigou, National Technical University of Athens, Greece
Eduardo Oliveros, Telefónica Investigación y Desarrollo, Spain
Stephen C. Phillips, University of Southampton IT Innovation Centre, UK
Karsten Oberle, Alcatel Lucent Bell Labs, Germany

Service Level Agreements (SLAs) are nowadays used as a cornerstone for building service-oriented architectures. SLAs have been closely investigated in the scope of distributed and Grid computing and are now gaining uptake in cloud computing as well. However, most solutions have been developed for specific purposes and are not applicable generally, even though the most approaches propose a general usability. Only rarely have SLAs been applied to real-time systems. The purpose of this chapter is to analyze different fields where SLAs are used, examine the proposed solutions, and investigate how these can be improved in order to better support the creation of real-time service-oriented architectures.

Chapter 9
Securing Real-Time Interactive Applications in Federated Clouds ...................................................... 160
Michael Boniface, University of Southampton IT Innovation Centre, UK
Bassem Nasser, University of Southampton IT Innovation Centre, UK
Mike Surridge, University of Southampton IT Innovation Centre, UK
Eduardo Oliveros, Telefónica Investigación y Desarrollo, Spain

Enterprise adoption of cloud computing for real-time interactive applications processes is limited by their ability to meet inter-enterprise security requirements. Although some clouds offerings comply with security standards, no solution today allows businesses to assess security compliance of applications at the business level and dynamically link to security countermeasures on-demand. In this chapter we examine cloud security, privacy, and trust issues from three levels; business, jurisdiction, and technical. Firstly, we look at the business level to identify issues arising from the motivations and concerns of business stakeholders. Secondly, we explore jurisdictional level to identify risks that arise from legislation, gaps in legislation, or conflicts between legislation in different jurisdictions related to a cloud deployment, given the concerns of stakeholders. Finally, we examine the technical level to identify issues that arise from technical causes such as ICT vulnerabilities, and/or require technical solutions, such as data confidentiality and integrity protection.
Chapter 10
Web Service Specifications Relevant for Service Oriented Infrastructures

Eduardo Oliveros, Telefónica Investigación y Desarrollo, Spain
José Movilla, Telefónica Investigación y Desarrollo, Spain
Andreas Menychtas, National Technical University of Athens, Greece
Roland Kuebert, University of Stuttgart, Germany
Michael Braitmaier, University of Stuttgart, Germany
Stuart Middleton, University of Southampton IT Innovation Centre, UK
Stephen C. Phillips, University of Southampton IT Innovation Centre, UK
Michel Boniface, University of Southampton IT Innovation Centre, UK
Bassem Nasser, University of Southampton IT Innovation Centre, UK

Service Oriented Infrastructures (SOIs) have recently seen increased use, mainly thanks to technologies for data centre virtualization and the emergence and increasing commercial offering of Cloud solutions. Web Services have been seen as a tool to implement SOI solutions thanks to their versatility and interoperability, but at the same time, Web services have been considered not suitable for providing interactive real-time solutions. In this chapter the state of the art of Web service technology will be analysed, and their different communication mechanisms and the existing implementations will be compared. Firstly, the different standardisation bodies working on Web service specifications relevant to SOI will be introduced. The various approaches to implement Web services will be described followed by the Web service specifications and the middleware that make use of those specifications, including the description of the commercial interfaces and development tools to create services for the cloud. In the last part of the chapter, the interoperability problems present on the different frameworks and the existing solutions to minimize those interoperability problems will be explained.

Section 3
Infrastructure as a Service

Chapter 11
Execution and Resource Management in QoS-Aware Virtualized Infrastructures

Dominik Lamp, University of Stuttgart, Germany
Sören Berger, University of Stuttgart, Germany
Manuel Stein, Alcatel-Lucent Bell Labs, Germany
Thomas Voith, Alcatel-Lucent Bell Labs, Germany
Tommaso Cucinotta, Scuola Superiore Sant’Anna, Italy
Marko Bertogna, Scuola Superiore Sant’Anna, Italy

Both real-time systems and virtualization have been important research topics for quite some time now. Having competing goals, research the correlation of these topics has started only recently. This chapter overviews recent results in the research literature on virtualized large-scale systems and soft real-time systems. These concepts constitute the fundamental background over which the execution environment of any large-scale service-oriented real-time architecture for highly interactive, distributed, and virtualized applications will be built in the future. While many aspects covered in this chapter have already
been adopted in commercial products, others are still under intensive investigation in research labs all over the world.

Chapter 12
Network Management in Virtualized Infrastructures

Manuel Stein, Alcatel-Lucent Bell Labs, Germany
Karsten Oberle, Alcatel-Lucent Bell Labs, Germany
Thomas Voith, Alcatel-Lucent Bell Labs, Germany
Dominik Lamp, University of Stuttgart, Germany
Sören Berger, University of Stuttgart, Germany

Service Oriented Infrastructures (SOI) build upon previous advancements in distributed systems, Grid computing, Cloud computing, virtualization, SOA, and technologies alike. Capabilities merged under the banner of SOI offer a solution serving long-standing business needs, but also meet increasing demand for infrastructures enabling the fast and flexible deployment of new services. However, typical current SOI realizations, e.g., Grid or Cloud solutions, do not take the network infrastructure, necessary for flawless service interaction, sufficiently into consideration. In most cases, those frameworks focus on providing huge and extremely divisible applications with hardware resources possibly distributed over several provider domains. They manage just computing related resources like CPU and RAM or Storage, but network connectivity is typically taken for granted, while network QoS aspects (e.g., jitter, delay) of the data exchange is usually not considered. Consequently, the data exchange between changeably deployed components cannot be comprehensively treated. This chapter provides an overview on related state of the art technologies regarding topics such as QoS provisioning, virtualization, and network resource management. This background is enriched with latest research results on future trends and advances in state of the art in Network Management.

Chapter 13
Data Storage in Cloud Based Real-Time Environments

Sai Narasimhamurthy, Xyratex, UK
Malcolm Muggeridge, Xyratex, UK
Stefan Waldschmidt, Digital Film Technology, Germany
Fabio Checconi, Scuola Superiore Sant'Anna, Italy
Tommaso Cucinotta, Scuola Superiore Sant'Anna, Italy

The service oriented infrastructures for real-time applications ("real-time clouds") pose certain unique challenges for the data storage subsystem, which indeed is the "last mile" for all data accesses. Data storage subsystems typically used in regular enterprise environments have many limitations which impede direct applicability for such clouds, particularly in the ability to provide QoS for applications. Provision of QoS within storage is possible through a deeper understanding of the behaviour of the storage system under a variety of conditions dictated by the application and the network infrastructure. We intend to arrive at a QoS mechanism for data storage keeping in view the important parameters that come into play for the storage subsystem in a soft real-time cloud environment.
Chapter 14
Fault Detection and Recovery Mechanisms and Techniques for Service Oriented Infrastructures

Andreas Menychtas, National Technical University of Athens, Greece
Kleopatra G. Konstanteli, National Technical University of Athens, Greece

The need for guaranteed QoS and efficient management in Service Oriented Infrastructures is an essential requirement for the deployment, execution, and management of modern business applications. In that frame, the capabilities for fault detection and recovery in all layers of a Service Oriented Infrastructure are essential for the smooth operation of the business applications and the wide adoption of these solutions in the global market. In this chapter, we present the concepts of fault detection and recovery, including terminology, classification of faults, and analysis of the key processes taking place in a system in order to diagnose and recover from failures. The state of the art mechanisms and techniques for fault detection and recovery are also analyzed while recommendations for applying them in Service Oriented Infrastructure are presented.

Chapter 15
Real-Time Attributes in Operating Systems

Tommaso Cucinotta, Scuola Superiore Sant'Anna, Italy
Spyridon V. Gogouvitis, National Technical University of Athens, Greece

General-Purpose Operating Systems (GPOSes) are being used more and more extensively to support interactive, real-time, and distributed applications, such as found in the multimedia domain. In fact, the wide availability of supported multimedia devices and protocols, together with the wide availability of libraries and tools for handling multimedia contents, make them an almost ideal platform for the development of this kind of complex applications. However, contrarily to Real-Time Operating Systems, General-Purpose ones used to lack some important functionality that are needed for providing proper scheduling guarantees to application processes. Recently, the increasing use of GPOSes for multimedia applications is gradually pushing OS developers towards enriching the kernel of a GPOS so as to provide more and more real-time functionality, thus enhancing the performance and responsiveness of hosted time-sensitive applications. In this chapter, an overview is performed on the efforts done in the direction of enriching GPOSes with real-time capabilities, with a particular focus on the Linux OS. Due to its open-source nature and wide diffusion and availability, Linux is one of the most widely used OSes for such experimentations.

Compilation of References

About the Contributors

Index