

Chapter 8.7

Service Level Agreements for Real-Time Service- Oriented Infrastructures

Roland Kübert

University of Stuttgart, Germany

Theodora Varvarigou

National Technical University of Athens, Greece

Georgina Gallizo

University of Stuttgart, Germany

Eduardo Oliveros

Telefónica Investigación y Desarrollo, Spain

Thodoris Polychniatis

National Technical University of Athens, Greece

Stephen C. Phillips

*University of Southampton IT Innovation Centre,
UK*

Karsten Oberle

Alcatel Lucent Bell Labs, Germany

ABSTRACT

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.

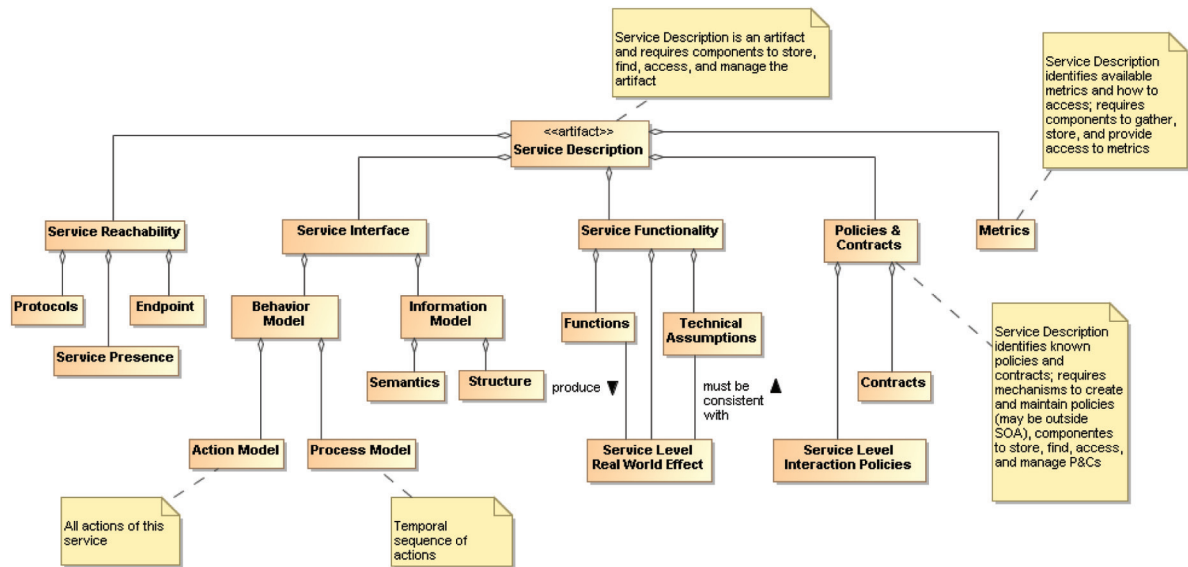
INTRODUCTION

In the context of service oriented architectures, a service is a mechanism to enable access to one or more capabilities, where the access is provided using a prescribed interface and is exercised consistent with constraints and policies as specified by the service description (OASIS).

An SLA is an agreement between the provider of the service and the consumer of the service that specifies the function performed by the service, the obligations on both the provider and consumer of the service, the agreed bounds of performance (Quality of Service, QoS) for the service, and how deviations are handled (exceptions and compensation). In this sense, an SLA is a contract between the participants in the service, which is typically

DOI: 10.4018/978-1-4666-0879-5.ch8.7

Figure 1. OASIS SOA reference model for service description



the provider and the consumer of the service, but can also include mediators or other actors that are stakeholders in the service.

SLAs form part of wider frameworks used to describe, access, use and govern services (policies and contracts) for example as shown in Figure 1.

An SLA is made in some business context, which may include decisions made by each party leading up to the agreement, the presence of an endorser for the agreement and simply some prior conditions that make the terms of the agreement acceptable to both sides. An SLA is typically established before deploying a service and covers the whole lifecycle including execution and monitoring through to decommissioning. However, it is also possible to form an SLA with an existing service, e.g. through a federation process orchestrated by an existing consumer that produces new interactions with other consumers. SLAs therefore have a huge influence on all aspects of the service, from as early as design time, to the infrastructure the service is deployed and executed on and the monitoring components that will be required for the provider to offer a service successfully in a cloud computing environment.

While many aspects of context following the creation of an SLA can be shared or agreed (e.g. roots of trust, expected QoS, etc), some should not (e.g. no commercial service provider is likely to reveal their resource plan to a consumer).

On a more practical level, some examples of the use of SLAs in real systems are given. Within the academic Grid community SLAs, e.g. as described in (Buyya, Abramson, & Venugopal, 2005),(Czajkowski, Foster, & Kesselman, 2005),(Yeo & Buyya, 2005), tend to focus on resources (e.g. computers, network bandwidth, storage devices) rather than services. This is appropriate for a community of experienced users running often experimental codes. However, in order to address the needs of business users (as opposed to technical users) the value of the service must be articulated at the appropriate business level rather than the resource level. The customer should not be concerned with the resources required to provide the service but just that the service exists and provides a clear business benefit. It is also important to consider SLA information when composing services to provide an integrated business solution. The NextGRID

25 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/service-level-agreements-real-time/64570

Related Content

A Novel Load Balancing Technique for Smart Application in a Fog Computing Environment

Mandeep Kaur and Rajni Aron (2022). *International Journal of Grid and High Performance Computing* (pp. 1-19).

www.irma-international.org/article/a-novel-load-balancing-technique-for-smart-application-in-a-fog-computing-environment/301583

Flow-Based Networking Architectures under Mixed Internet and Grid Traffic

César Cárdenas, Maurice Gagnaire and Víctor López (2012). *Computational and Data Grids: Principles, Applications and Design* (pp. 343-362).

www.irma-international.org/chapter/flow-based-networking-architectures-under/58753

A Dynamic Load Balancing Strategy with Adaptive Thresholds (DLBAT) for Parallel Computing System

Taj Alam and Zahid Raza (2014). *International Journal of Distributed Systems and Technologies* (pp. 54-69).

www.irma-international.org/article/a-dynamic-load-balancing-strategy-with-adaptive-thresholds-dlbat-for-parallel-computing-system/104764

Adaptive Routing Strategy for Large Scale Rearrangeable Symmetric Networks

Amitabha Chakrabarty, Martin Collier and Sourav Mukhopadhyay (2012). *Evolving Developments in Grid and Cloud Computing: Advancing Research* (pp. 212-222).

www.irma-international.org/chapter/adaptive-routing-strategy-large-scale/61993

Resource Scheduling for Energy-Aware Reconfigurable Internet Data Centers

Mohammad Shojafar, Nicola Cordeschi and Enzo Baccarelli (2016). *Innovative Research and Applications in Next-Generation High Performance Computing* (pp. 21-46).

www.irma-international.org/chapter/resource-scheduling-for-energy-aware-reconfigurable-internet-data-centers/159038