Chapter 7
On Controlling Elasticity of Cloud Applications in CELAR

Georgiana Copil
Vienna University of Technology, Austria

Daniel Moldovan
Vienna University of Technology, Austria

Hung Duc Le
Vienna University of Technology, Austria

Hong-Linh Truong
Vienna University of Technology, Austria

Schahram Dustdar
Vienna University of Technology, Austria

Chrystalla Sofokleous
University of Cyprus, Cyprus

Nicholas Loulloudes
University of Cyprus, Cyprus

Demetris Trihinas
University of Cyprus, Cyprus

George Pallis
University of Cyprus, Cyprus

Marios D. Dikaiakos
University of Cyprus, Cyprus

Ioannis Giannakopoulos
National University of Athens, Greece

Nikolaos Papailiou
National University of Athens, Greece

Ioannis Konstantinou
National University of Athens, Greece

Craig Sheridan
FLEXIANT, UK

Christos K. K. Loverdos
Greek Research and Technology Network, Greece

Evangelos Floros
Greek Research and Technology Network, Greece

DOI: 10.4018/978-1-4666-8213-9.ch007
ABSTRACT

Today’s complex cloud applications are composed of multiple components executed in multi-cloud environments. For such applications, the possibility to manage and control their cost, quality, and resource elasticity is of paramount importance. However, given that the cost of different services offered by cloud providers can vary a lot with their quality/performance, elasticity controllers must consider not only complex, multi-dimensional preferences and provisioning capabilities from stakeholders but also various runtime information regarding cloud applications and their execution environments. In this chapter, the authors present the elasticity control approach of the EU CELAR Project, which deals with multi-dimensional elasticity requirements and ensures multi-level elasticity control for fulfilling user requirements. They show the elasticity control mechanisms of the CELAR project, from application description to multi-level elasticity control. The authors highlight the usefulness of CELAR’s mechanisms for users, who can use an intuitive, user-friendly interface to describe and then to follow their application elasticity behavior controlled by CELAR.

1. INTRODUCTION

With the popularity and diversity of cloud-based solutions from cloud providers and application providers/developers, there is a considerable need to customize these solutions and to provide cloud users with fine-grained mechanisms of controlling their cloud applications.

Many existing frameworks allow the specification of various cloud application-related information, like the cloud application complex structure (e.g., Di Nitto et al. (2013)) and functional requirements (e.g., Di Cosmo et al. (2013)) when deploying the cloud application on the cloud. Moreover, many tools are capable of describing and deploying cloud applications (e.g., Binz et al. (2013)) on different cloud infrastructures. The requirements of the cloud application stakeholders differ and depend on a number of variables, e.g., the cost of the cloud application reported to the number of clients, or the various cloud application quality parameters (e.g., a banking cloud application differs greatly in requirements from a scientific cloud application). However, current state-of-the-art on elasticity control techniques require the specification of low-level, detailed information. For instance, Auto Scale applications provided by Amazon, Rackspace, Azure or RightScale enable users to specify, for each Virtual Machine they are using, scaling policies, depending on IaaS-level metrics. Proposed frameworks take into consideration cloud application level metrics, e.g., response time, but do not allow users to specify their requirements, the optimization factor being defined in an ad-hoc manner (e.g., equilibrium between the cost and response time) (e.g., Serrano et al. (2013), Simjanoska et al. (2013)).

The concept of multi-dimensional elasticity, covering resources elasticity, cost elasticity and quality elasticity (see Dustdar et al. (2011)) and the relations among them, shows how complex the elasticity control of cloud applications actually is. Such a concept facilitates custom cloud application elasticity depending on what a cloud application stakeholder (e.g., service provider) actually needs. A visual representation of the elasticity dimensions is shown in Figure 1, each of the main dimensions, cost, resource and quality being further decomposed into storage cost and network cost, CPU and memory, and respectively quality of data and performance. Elasticity is defined as the relationship among these dimensions, in time, which change for fulfilling user’s elasticity requirements. Considering that distributed cloud applications have complex structures, each com-