Chapter 14
Cross-Layer Monitoring in Cloud Computing

Jose M. Alcaraz Calero
Hewlett-Packard Research Laboratories, UK

Benjamin König
Hewlett-Packard Research Laboratories, UK

Johannes Kirschnick
Hewlett-Packard Research Laboratories, UK

ABSTRACT

This book chapter describes the different aspects related to designing a suitable monitoring architecture for Cloud Computing, aiming to support cross-layer monitoring across all layers available in the Cloud stack. For this purpose, the importance of monitoring services in Cloud scenarios is outlined, followed by a comprehensible analysis of a wide set of distributed monitoring solutions. After that, the particular requirements related to cross-layer monitoring for Cloud Computing architectures are identified and explained. Then, diverse aspects which may fit a monitoring architecture for fulfilling such requirements are explained. Finally, some future research directions and conclusions are highlighted.

INTRODUCTION

Traditionally, businesses had to invest in the acquisition of their own dedicated IT infrastructure to satisfy their computational requirements. This led to a planning problem, as requirements for IT infrastructure in businesses are constantly changing according to current workload and business strategy. As a consequence, traditional businesses have to make a significant up-front investment in IT infrastructure to provide enough computational power and storage capacity to handle the highest peaks of workload and successfully face future challenges.

Cloud Computing (Armbrust, et al., 2010) is an upcoming computing technology with the potential to change how businesses implement their services. The idea behind Cloud Computing is to use virtualized resources which are provided by a third-party. These resources are acquired on-demand and typically paid for in a pay-per-usage model (Buyyaa, Shin Yeo, Venugopal, Broberg,

DOI: 10.4018/978-1-4666-0960-0.ch014

& Brandic, 2009). This paradigm enables Cloud customers to adapt the allocated IT infrastructure dynamically, based on changing business objectives and user demands. This leads to a significant reduction in investment needed to acquire IT infrastructure and creates a paradigm, which leads to a balance between own dedicated IT computation resources and IT resources rented from third parties for dealing with peaks of workload.

Cloud Computing provides services on a stack composed of three service layers (Hurwitz, Bloor, Kaufman, & Halper, 2009): *Infrastructure as a Service* (*IaaS*), *Platform as a Service* (*PaaS*), and *Software as a Service* (*SaaS*).

The *IaaS* layer exposes services for which the users can deploy virtual resources (e.g. virtual machines) on demand. This *IaaS* layer uses the underlying physical resources (e.g. physical machines, network, and storage components) to achieve the functionality. The *PaaS* layer relies on these virtualized resources to provide more complex middleware services. These middleware services are intended for being used as components and functionalities for other applications and services. Finally, the *SaaS* layer provides consumable high-level services to users (i.e. entire enterprise applications or web services). To provide these services, the *SaaS* layer uses the other layers of the cloud stack. Each layer available in the cloud stack is not restricted to end users but target different groups of users, i.e. developers and operations staff illustrates the types of services provided by the *IaaS*, *PaaS,* and *SaaS* layers and their primary user groups.

Each service layer may be provided by a different entity (*service provider*). Likewise, there may be different consumers for Cloud services (*service consumer*). In this context, a *private* Cloud is a usage scenario in which service provider and consumer are of the same party, whereas a *public* Cloud refers to a scenario in which provider and consumer are different parties and have to be from different administrative domains. Note that a private cloud can be considered as a software stack for internally governing the computational resources available in a company whereas a public cloud involves different providers of services.

Currently, almost all research works are focused on public Cloud scenarios, which are the predominant option in Cloud Computing due to imminent advantages for the service consumer, such as, a reduction of acquisition cost in IT infrastructure and outsourcing of IT maintenance tasks to the service provider. This book chapter is focused on a public Cloud scenario in which a service provider offers Cloud services, generally by means of the Internet, to a variable number of service consumers.

The design of a stack for a public Cloud scenario is generally challenging as the architecture has to reflect the individual requirements of different service consumers and service providers. Many challenges associated to the design of the service layers are still open issues nowadays and they are regarded as drawbacks for which some businesses are reluctant to move to the Cloud. One of the main concerns reported by Cloud customers (Fujitsu, 2010) is how consumers get full control over virtualized resources rented on-demand.

Monitoring enable service providers a way to increase the control that customers have over computational resources by providing monitored information about the infrastructure and the services deployed therein.

Monitoring systems in cloud computing entail several problems, which are not correctly addressed in the current monitoring solutions. Firstly, as shown in Figure 1, monitoring services for the Cloud have to interact with all service layers available in the cloud stack. This requires an architecture for performing a monitoring across all the layers available in the cloud stack, i.e. a cross-layer monitoring rather than traditional monitoring solutions focused on monitoring homogeneous resources, i.e. a single layer. The reason for which cross-layer monitoring is critical is due to the high level of demanded control over the rented computational architecture. The final user demands