Journal of Network and Computer Applications 39 (2014) 1–16

Contents lists available at ScienceDirect

Journal of Network and Computer Applications

journal homepage: www.elsevier.com/locate/jnca

Automatic services instantiation based on a process specification

Hélder Pereira Borges [a,b,*], José Neuman de Souza [b], Bruno Schulze [c], Antonio Roberto Mury [c]

[a] Computer Science, Federal Institute of Maranhao—IFMA, Av. Getulio Vargas, 04-Monte Castelo, 65.025-001 Sao Luis, Maranhao, Brazil
[b] Computer Science, Federal University of Ceara (UFC), Campus do Pici-Bloco 910, 60.455-760 Fortaleza, Ceara, Brazil
[c] Computer Science, National Laboratory for Scientific Computing (LNCC), Av. Getulio Vargas, 333-Quitandinha, 25.651-075 Petropolis, Rio de Janeiro, Brazil

ARTICLE INFO

Article history:
Received 31 October 2012
Received in revised form
20 February 2013
Accepted 1 April 2013
Available online 18 April 2013

Keywords:
Cloud computing
Models
Standardization
SLA
QoS

ABSTRACT

A relevant challenge for cloud computing is related to quality control of provided services. Normally, cloud providers deliver services, but do not clearly guarantee their quality level. In addition, each provider uses a particular process to provide service, considering only its own business goals. This paper presents an approach based in an environment model for clouds and aims to propose a solution for these cloud challenges. It intends to define a service procurement process taking into account users requirements, and also considers service quality maintenance. This process uses an environment model containing all relevant information to create a virtual workspace (from just an individual virtualized platform to platforms hosted in clouds), considering hardware and software requirements and quality parameters, all of which are specified by users, to precisely offer the expected service. From this model, it will be possible to automatically provide infrastructure and platform as a service. The agreement negotiation happens during the service acquisition process, using automated agents to create services and to monitor their quality, generating an environment less error-prone and increasing the customer satisfaction level.

1. Introduction

A fact commonly observed in cloud computing context is the lack of a standard process for services request. Each provider according to its own convenience defines how services requisition should happen, not necessarily considering users requirements. The National Science Foundation (NSF) confirms this reality through the program Strategic Technology for Cyber Infrastructure (STCI) (NFS, 2012), which calls for proposals to standardize and automate access to cloud services.

From this perspective, this study defines a process for service procurement and presents a system associated with clouds, intending to automatically and dynamically create personalized services, complying with the user's needs. This process also has the objective to ensure the quality of services levels, proposing some quality parameters adequate for clouds. The main artifact in this structure is an environment model, containing all information required to make a service available. After the model creation, three perspectives will be evaluated.

1 A cloud provider develops a mechanism that performs the mapping of the environment model to the virtual image structure adopted. This is the least realistic prospect, given the large external dependency.
2 Addition of a mapping mechanism, to one or more specific kinds of virtual images used by cloud providers to the proposed system. This one represents a more real approach, with a small level of dependence, being necessary only the establishment of communication mechanisms using the APIs given by the cloud provider to create the virtualized environment. In this perspective, the provider only needs to instantiate a virtual environment.
3 From the model, the proposed system should automatically and dynamically create and deliver services to users. This is an ideal perspective, because it does not create dependence on third parties to provide the service. The system should perform the whole process of service provision and then be able to be used from standalone virtual environment to private clouds (this research is mostly focused on this aspect).

The purpose of this paper is the creation of a flexible mechanism for specification of cloud services and its quality parameters, enabling the real-time creation of services from the architecture developed for the proposed system without the use of prefabricated virtual images. This article is organized as follows: Section 2 presents the main topics related to this study, describing some related works. In Section 3 we depict the GerNU and its proposal to define a requirement process, also describing how it is possible to create a personalized cloud service. Section 3.8 presents three case studies. Finally, in Section 4, we have closing remarks and a research agenda.

* Corresponding author at: Computer Science, Federal Institute of Maranhao—IFMA, Av. Getulio Vargas, 04-Monte Castelo, 65.025-001 Sao Luis, Maranhao, Brazil. Tel.: +55 2499098731.
E-mail addresses: helderpb@hotmail.com, helder@ifma.edu.br (H.P. Borges), neuman.souza@gmail.com (J.N. de Souza), bruno.schulze@gmail.com (B. Schulze), a.roberto.m@gmail.com (A.R. Mury).

1084-8045/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jnca.2013.04.005

2. Related works

As far as it was possible to research, we could not identify a similar study to the general context of this proposal. Thus we described research results related to some challenges in the cloud context that encouraged the policies adopted in this proposal as Quality of Service (QoS) parameters, Service Level Agreement (SLA) as well as models.

2.1. Cloud computing

Cloud computing has been established in recent years as an important research area that has many advantages, especially considering tasks such as obtaining, sharing, manipulating and exploring large amounts of data, which are quite common today. However its implementation requires many resources. The availability of these resources by cloud providers adds a number of advantages for organizations and users. Concerns about complex installations and infrastructure maintenance become responsibility of the service providers, allowing users to focus exclusively on their business rules.

One of the reasons for the cloud success is that its features can be used in both business and academic contexts, since its efficiency could be verified by applications in real problems such as high-performance computing applications, social networks and others.

The use of specialized companies, such as computing resources providers, corresponds to one basic proposition of cloud computing. In this context, the resource supply is abstracted from the user, and the management and maintenance become responsibility of the experts. Moreover, the resources provision must be done in several layers, which represent a specific kind of resource, provided as a service (Infrastructure—IaaS, Platform—PaaS or Software—SaaS). These abstractions hide the infrastructure complexity (Buyya et al., 2008).

Another important feature is the payment based on resource utilization (Vaquero et al., 2009). A cloud represents a scalability scenario giving its almost unlimited services, resources, processes and infrastructure without precedent, improving flexibility to the information technology structures, and furthermore decreasing the business cost by providing on-demand services.

Considering cloud computing, patterns are described in Fehling et al. (2012) as one way to document architectural principles and to create good solutions to recurring cloud challenges. Moreover, Binz et al. (2011) describe patterns taking into account an application moved to the cloud. It covers the best practices in how to handle differences of runtime environments in a standard format.

2.1.1. Infrastructure as a service—IaaS

An IaaS is the layer where the main objective is to make it easy and affordable the provision of resources, such as servers, networks, storage and others. The IaaS is very important for the construction of an on-demand environment, thereby providing the necessary infrastructure for PaaS and SaaS. This infrastructure is based on the virtualization of computing resources, being dynamically scalable to increase or decrease according to the demand in a transparent way.

Some advantages of working with IaaS are: reduction of investments in hardware as well as of concerns about its depreciation; elimination of safety and maintenance costs; performance optimization; release of physical space; flexibility to increase or decrease the processing power, storage and others.

2.1.2. Platform as a service—PaaS

The basic proposal of the PaaS model is the provision of services to enable the development, testing, deployment, hosting and application management, aiming to support the life cycle of the application development. These services can be configured as an integrated solution offered through the Internet without the costs and complexities related to the purchasing and management of hardwares and adjacent softwares.

Some advantages of using PaaS are described below: lower initial investment, representing the lowest risk business; no investment in infrastructure or software, and maintenance being the responsibility of the provider; prompt support and updates, being available immediately and transparently; focus applied in business; increased availability and data security.

2.1.3. Software as a service—SaaS

This layer is composed by applications that run in the cloud environment, and has been presented in recent years as a tendency for the software industry. It is a business model characterized by centralized access to information through the Internet, which means that from anywhere and with any computing device a user can access a service through an interface with a web browser. Thus, new features can be added to systems in a transparent way, therefore making the maintenance and evolution of systems a simpler task.

Some advantages of using a SaaS are listed below: lower initial cost because there are no license fees, being charged only for the use; transparent updates without files for download or installation; better management of the application given the centrality of the environment; high level of availability and infinite scalability of the infrastructure to meet the users demand.

2.1.4. Other services

There are many concepts derived from the as-a-service paradigm, normally used to distinguish a particular type of service, within which we can cite database as a service (DaaS), testing as a service (TaaS), security, simulation, communication, and others, all provided as a service.

2.2. Quality of service parameters—QoS

Normally, proposals for resource provisioning have been based on the best effort policies, making it possible to meet the user's request usually not taking into account the QoS. Distefano et al. (2011) depict this problem presenting a proposal focused on QoS aspects, describing how to deliver SLA-based guarantees for QoS.

A method for allocating resources for cloud services that considers processing, memory, network speed, reliability and throughput is shown in Prasad et al. (2010), where algorithms are presented for data partitioning, allowing the parallel data transfer.

Moreover, providers use virtualization to host independent services on their servers. This practice creates interference in the physical machine performance, which is reflected in the virtual machines, and may significantly affect the QoS requirements (Nathuji et al., 2010). Approaches evaluating and ranking components are important for the selection and identification of components with high/low performance (Zheng et al., 2010).

2.3. Service level agreement—SLA

Cloud computing uses the concept of SLA to control the use of computing resources arising from a provider. Normally, the SLA management strategies are based on an understanding of two distinct phases: one is related to contract negotiation and the