Future Generation Computer Systems 75 (2017) 1–13

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

QoS-aware genetic Cloud Brokering

Gaetano F. Anastasi, Emanuele Carlini *, Massimo Coppola, Patrizio Dazzi

Information Science and Technologies Institute, National Research Council (ISTI-CNR), Pisa, Italy

HIGHLIGHTS

- Cloud brokering based on a genetic algorithm, finding near-optimal solution according to customers preferences.
- Support to qos-aware allocation for local (e.g. cpu, ram) and global constraints (network), and different cost models.
- Scalability up to hundreds of providers, suitable for scientific application workflows.

ARTICLE INFO

Article history:
Received 19 June 2015
Received in revised form
17 January 2017
Accepted 15 April 2017
Available online 25 April 2017

Keywords:
Cloud computing
Cloud federations
Multi-cloud
Genetic algorithms
Resource management
Cloud brokering
Workflow scheduling
Distributed computing
Quality of service

ABSTRACT

The broad diffusion of Cloud Computing has fostered the proliferation of a large number of cloud computing providers. The need of Cloud Brokers arises for helping consumers in discovering, considering and comparing services with different capabilities and offered by different providers. Moreover, consuming services exposed by different providers may alleviate the vendor lock-in issue. While it can be straightforward to choose the best provider when deploying small and homogeneous applications, things get more challenging with large and complex applications. In this paper we propose QBROKAGE, a genetic approach for Cloud Brokering, aiming at finding Infrastructure-as-a-Service (IaaS) resources for satisfying Quality of Service (QoS) requirements of cloud applications. Our approach is capable of evaluating such requirements both for the single application service and for the application as whole. We performed a set of experiments with an implementation of such broker, by considering three-tier applications and scientific application workflows. Results show that our broker can find near-optimal solutions even when dealing with hundreds of providers, providing optimized deployment solutions that includes data transferring cost across multiple clouds.

1. Introduction

Cloud Computing is nowadays one of the most popular computational paradigms. It has been adopted by many companies and considered by many more others for the unquestionable benefits offered, such as potential cost reductions offered by the pay-per-use model, flexibility and scalability, fault-tolerance and increased availability due to the geographic distribution of resources.

Many Cloud providers implicitly force their customers to use proprietary interfaces, virtualization technologies, and communication protocols, so that the cost of switching from that provider's technology to another one would be too high and the migration become materially unfeasible for the customer (i.e. the so called vendor lock-in). For protecting themselves against vendor lock-in, some small to mid-sized businesses (SMBs) may decide to under-invest or simply hesitate to adopt Cloud Computing. Recent surveys [1] also point out that some SMBs are forgoing Cloud Computing because of security and trust reasons, being afraid of losing control on their data, worrying about reliability, integrity and compliance with data privacy laws.

Recently, the adoption of multiple clouds for running cloud-based applications and services has been considered as a mitigation factor towards the vendor lock-in issue. In addition, a multi-cloud environment may be beneficial to cloud-based applications in many other ways. For example, some application services may have special functional and/or non-functional demands that cannot be fulfilled by a single target cloud. In this case, considering a multi-cloud scenario is simply mandatory. Moreover, the multi-cloud scenario can show its advantage in terms of cost-saving for the users: since different services may have different requirements, simply choosing the cheapest provider by considering a single resource may not be cost-effective.

Two orthogonal approaches are commonly exploited for addressing deployments across multiple clouds: Cloud Brokering and

* Corresponding author.
E-mail addresses: g.anastasi@isti.cnr.it (G.F. Anastasi), emanuele.carlini@isti.cnr.it (E. Carlini), m.coppola@isti.cnr.it (M. Coppola), p.dazzi@isti.cnr.it (P. Dazzi).

http://dx.doi.org/10.1016/j.future.2017.04.026

Cloud Federation [2]. Cloud Brokers can leverage abstraction APIs, such as Apache Libcloud[1] or Delta Cloud[2] for allowing users to exploit different providers at the same time whereas Cloud Federations provide common platforms providers must be compliant with. Even if Cloud Federation may subsume the Cloud Brokering approach, they can be considered orthogonal from the viewpoint of the goals they pursue. In fact, if on the one hand a Cloud Broker should always consider user profits neglecting provider ones, on the other hands the Cloud Federation must operate a trade-off between these two apparent discording objectives, for example ensuring fairness in exploiting resources belonging to the federated providers.

Additionally, such approaches can help to overcome the trust problem that limits the adoption of Cloud Computing, for instance by selecting time by time providers that are most suitable to fit the security needs of the users. As an example, the user may want to choose a particular provider location when submitting applications for ensuring law compliance in data management. Recent advances in this research field, designed and developed in the Contrail approach to Cloud Federation [3–5], treat security needs by explicitly addressing Quality of Protection (QoP) terms as a special case of Quality of Service (QoS).

One of the most relevant research challenges focuses on the problem of scheduling complex applications by respecting user constraints, that have to match the providers' offer. The related aspect to consider is the number of worldwide providers. While it can be considered acceptable to manually search for resources on handful of providers, this task becomes unfeasible when the number of providers grows up to hundreds.

To address this issue we conceived, designed and developed QBROKAGE, a Cloud Brokering approach that provides an optimized deployment solution for a cloud-based application across multiple clouds. QBROKAGE exploits only the information that commercial providers are likely to made available for customers, such as Virtual Machine (VM) costs and their features in term of storage, memory, etc. Let us consider a scenario in which customers submit their applications to QBROKAGE requesting for a deployment configuration that meets QoS requirements, that could be formally expressed by Service Level Agreements (SLAs). Such requirements may involve both non-functional aspects, such as security capabilities of providers, and functional aspects as coming from other specification formats, such as the Open Virtualization Format (OVF [6]). For example, application requirements may specify that VMs require at least a certain amount of memory, and a minimum number of physical CPUs, along with the exact match of geographic location where to place specific parts of the application. Such requirements are used as constraints by QBROKAGE for choosing a set of Cloud providers that can host the services (appliances) and at the same time guaranteeing the respect of the QoS negotiated for the whole application.

In such context, QBROKAGE advocates the exploitation of a Genetic Algorithm (GA) to match services and Cloud resources. GA is a well-known heuristic approach that permits to iteratively find near-optimal solutions for NP-hard problems in large search spaces. Being an heuristic approach, it usually has a computational advantage w.r.t. optimal algorithms and thus it is suitable for being leveraged in an interactive service like our broker. Moreover, our work leverages the GA approach because its model is flexible enough to support multiple constraints at the same time and the injection of additional constraints in the future with minimal interventions on the algorithm. Clearly, this is a crucial aspect for software reuse in the context of Cloud Computing, where QoS models are continuously enriched as providers support QoS guarantees previously not addressed, such as soft real-time guarantees for virtualized services [7] or multi-user virtual environments [8].

1.1. Paper contributions

The main contribution of this paper is the design and implementation of a generic framework supporting cloud brokering. Such framework embodies a genetic algorithm driving the allocation of applications, which is designed and implemented by considering the following software requirements:

- meeting the heterogeneous QoS requirements of applications;
- finding near-optimal solution according to customers preferences trying at the same time to mitigate vendor lock-in;
- supporting providers with different cost models;
- scaling up with hundreds of providers, while maintaining interactivity.

Several capabilities of QBROKAGE has been already presented in a previous paper [9]. With respect to that paper, we advanced with the implementation of our prototype and we extend our work with the following new contributions:

- we extend our conceptual framework by introducing QoS constraints with *global scope* i.e., constraints that cannot be evaluated by considering VMs in isolation;
- we add the capability of considering network characteristics by implementing two types of QoS constraints with *global scope*, i.e., cost and bandwidth;
- we evaluate our approach in terms of network awareness by setting up some experiments targeting well-known scientific application workflows;
- we further study the scalability of QBROKAGE, presenting an additional experiment to this purpose.

To foster further research in this field and to make our results reproducible, we made the source code and dataset publicly available [10].

1.2. Paper outline

In Section 2, we present our work with respect to the state of the art. The model proposed in this paper is presented in Section 3. The reference architecture for QBROKAGE and an insight on the algorithm are given respectively in Section 3.3 and Section 4.1. This paper also provides an experimental evaluation of QBROKAGE by means of simulations (Section 5), including a comparison with a state of the art approach, the tuning of the genetic algorithm, scalability performances, the capability of mitigating vendor lock-in and the QoS evaluation of a scientific workflow application when mapping it to multiple clouds.

2. Related work

2.1. Cloud Brokering

In the research community there is a wide consensus on the importance that brokers can have on Cloud environments for helping consumers in discovering, considering and comparing services with different capabilities as offered by different providers [11]. The need of brokering mechanism particularly arises in Cloud Federation architectures, such as Intercloud [2], the first approach going in the direction of building a unified platform composed by federated providers that can exchange information through super-entities (e.g. the Contrail approach) or as peers (e.g. the Sky [12] approach).

[1] http://libcloud.apache.org/.
[2] https://deltacloud.apache.org/.