ARTICLE IN PRESS

Future Generation Computer Systems ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at SciVerse ScienceDirect

Future Generation Computer Systems

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

Business-driven management of infrastructure-level risks in Cloud providers

J. Oriol Fitó *, Jordi Guitart

Barcelona Supercomputing Center and Technical University of Catalonia, Barcelona, E-08034, Spain

ARTICLE INFO

Article history:
Received 28 October 2011
Received in revised form
6 April 2012
Accepted 17 May 2012
Available online xxxx

Keywords:
Business-driven risk management
Self-managed Cloud providers
Private Cloud provisioning
Resource failures

ABSTRACT

Cloud computing is an innovative and promising paradigm that is leading to remarkable changes in the way in which hardware and software are designed and purchased, as well as how IT systems are managed. However, the Cloud is a risky paradigm. For instance, the use of Cloud services, which usually are external assets to their consumers, implies unprecedented risks that must be taken into account.

In this paper, we propose the involvement of the risk management discipline into the Cloud computing realm. We present a risk management approach led by business-level objectives (BLOs) of Cloud organizations. Its main goal is to assist in business-driven self-managed Cloud providers, by facing uncertainties always present in their internal decision-making processes. Our Cloud-aware risk management method includes a SEmi-quantitative BLO-driven Cloud Risk Assessment (SEBCRA) as the core subprocess. Its aim is to constantly rank and prioritize risks affecting the governing business-level goals.

In addition, we present, as a use case, a PaaS provider that incorporates our risk management approach to enhance the achievement of two BLOs, i.e. maximization of profit and customer satisfaction. In particular, it can manage – identify, assess, and treat – the most critical Cloud infrastructure-level risks, i.e. provisioning its private Cloud, either under- or over-provisioning, as well as resource failures. We present some risk treatment responses to face these risks and we evaluate their impact on the above-mentioned BLOs. Our results show that the best responses to address risks may change over time depending on the current provider's status. As a result, an adaptive management of risks should be considered as a mandatory process for Cloud providers to ensure their success in the ever-growing worldwide ecosystem of Clouds.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Nowadays, Cloud computing is widely recognized as the most promising computing paradigm of the last several years [1]. A recent Gartner report [2] identifies Cloud computing as the most strategic technology and trend for the majority of today's companies, basically because the use of Cloud systems leads to promising business models. The benefits for both stakeholders, i.e. providers and end-users, are actually very clear [3].

Already, several Cloud computing models have emerged: *SaaS providers*, e.g. Salesforce CRM [4], which deliver software over the Internet; *PaaS providers*, e.g. Google App Engine [5], which mainly offer virtualized execution environments to host Cloud services; and *IaaS providers*, e.g. Amazon EC2 [6], which provide virtualized computing resources as a service and, thus, serve as the foundation layer for Cloud systems. Each model offers different features and/or services, at different degrees of flexibility, and involves distinct risks. This includes new risks to be determined (due to the usage of new technologies such as virtualization), as well as well-known risks to be re-evaluated within the Cloud domain.

The current trend involves several of those providers interoperating in Cloud ecosystems. These multi-Cloud scenarios (e.g. Cloud federations) comprise for instance PaaS providers outsourcing resources from public Clouds of third-party IaaS providers when their customers demand overcomes their private Cloud capacity. As physical Cloud infrastructures are, of course, provisioned in a static way, Cloud providers have to accurately size their private Cloud capacity in order to tighten capital expenditures (CapEx). In fact, Spellmann et al. [7] state that those initial investments are much more difficult to secure than operating expenses (OpEx), e.g. electric bills, since they are usually budgeted for technology refresh every 3–5 years. Moreover, providers are exposed to physical or virtual resource failures, which may represent significant losses for them.

All in all, Cloud providers are constantly subjected to uncertainties during their operation, which may greatly impact their business expectations. These uncertainties can represent threats for their success which, therefore, can greatly reduce the well-known benefits of using Cloud systems. In this case, the involvement of risk-aware decision-making processes into self-managed

* Corresponding author. Tel.: +34 934054282.
E-mail addresses: josep.oriol@bsc.es (J.O. Fitó), jordi.guitart@bsc.es (J. Guitart).

0167-739X/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.future.2012.05.008

Please cite this article in press as: J.O. Fitó, J. Guitart, Business-driven management of infrastructure-level risks in Cloud providers, Future Generation Computer Systems (2012), doi:10.1016/j.future.2012.05.008

Cloud providers is clearly a need in order to minimize undesirable expenditures. On the other hand, uncertainties may result in opportunities or positive impacts for Cloud providers. In fact, there are events and management actions that may produce either positive or negative results for the business. For instance, overbooking resources in a Cloud provider may have a twofold impact: it increases the probability of violating service-level agreements (SLAs), and thus it implies revenue loss for the provider if violations occur, but conversely it can lead to obtaining more profit because the provider is serving more customers. Therefore, a remarkable tradeoff appears when considering the best risk-aware management action(s) to carry out in order to face risks. In fact, the best decision will highly depend on the provider's business-level objectives (BLOs) (e.g. profit and ecological efficiency maximization), which should be used to drive the whole risk management process.

1.1. Contributions

This paper, which extends our previous work [8], contributes to the inclusion of the risk management discipline into the Cloud computing paradigm.

First, we present an overview of risk management and assessment methods, as well as of the most important Cloud-related risks to be addressed.

Second, we propose the adoption by Cloud providers of a risk management process led by business interests, such as BLOs. We also propose a SEmi-quantitative BLO-driven Cloud Risk Assessment (SEBCRA) as its core subprocess. Its main goal is to constantly evaluate the impact of Cloud-specific risks, either with positive or negative consequences, on business objectives. Based on that, any Cloud organization is able to be aware, at any moment, of how to efficiently tackle uncertainties and their related risks, thus aligning its low-level management decisions with those high-level (business) objectives.

Third, we suggest several risk mitigation and adoption responses to deal with the negative and positive impacts, respectively, of the most critical infrastructure-level risks identified in the assessment step, namely the risk of provisioning a private Cloud, i.e. under- and over-provisioning, and the risk of physical and virtual resource failures.

Last, we demonstrate the applicability of the proposed BLO-driven risk management approach in a PaaS provider, which offers execution environments to host Web-based services. We experiment with different utilizations of its private Cloud (number of hosted services) and time-varying workloads for these services and we evaluate the impact of applying different risk treatment responses with respect to the achievement of two important BLOs for current Cloud providers: the maximization of profit and customer satisfaction. Our results demonstrate that a dynamic risk treatment strategy, which is able to apply the most appropriate risk response according to the provider's status to fulfill its BLOs, is needed to deal with the aforementioned Cloud infrastructure-related risks.

The remainder of the paper is organized as follows. Section 2 exposes useful background around risk management and assessment methods, as well as identifying the most important Cloud-related risks and re-evaluating traditional ones. Section 3 presents some related work on risk management and assessment. Section 4 details the business-driven risk management and assessment procedures. Section 5 presents a use case of how a PaaS provider is able to deal with critical Cloud infrastructure-related risks. Section 6 details the experimental environment and the evaluation of the presented risk management approach and of several risk treatment responses to face those risks. Finally, Section 7 draws our conclusions and exposes future work.

2. Background

2.1. Risk management and assessment

There are many different definitions of risk which have been developed and adopted by several disparate organizations over recent years. After considering dozens of them, ISO 31000:2009 [9], together with ISO/IEC Guide 73 [10], defines *risk* as the "effect of uncertainty on objectives". It also states that risk is the consequence of an organization setting and pursuing objectives against an uncertain environment. The uncertainty comes from both internal and external events which may or not happen. In general, they can represent opportunities for benefit or threats to success, i.e. positive and negative impacts of risks on an organization's objectives. Thus, and in contrast to traditional risk avoidance (mitigation) strategies, adopting positive risks may lead to obtaining significant benefits from the business point of view. Actually, managing risks can be seen as a process of optimization that causes organizations to minimize uncertainties in achieving their objectives.

Risk management is governed by generic guidelines and principles established in the widely accepted ISO 31000:2009. By definition, it is the process whereby organizations treat, in a methodical way, risks related with their activities. Its main goal is to obtain benefits and sustainable values for the business in each of its activities and across all of them. For this reason, it should be a fundamental part of any organization's strategic management.

Risk assessment is a core subprocess of any risk management strategy. It consists of the determination of a quantitative or qualitative value of each risk, also known as risk-level estimation, related to a particular situation and a recognized event (representing either a threat, an opportunity, or both). There are three primary methods to assess risks according to [11]: *qualitative*, which roughly categorizes risks and thus does not need to determine the numerical value of all assets at risk and frequencies; *quantitative*, which assigns numerical values to both the impact and the likelihood of risks; and *semi-quantitative (or hybrid)*, which is less numerically intensive than quantitative methods and classifies (prioritizes) risks according to consequences and foreseen probabilities.

Quantitative risk assessments have been criticized for being overly reductive and diverting attention from preventive actions. In addition, they ignore qualitative differences among risks. Although the calculations involved are tedious and include a strong element of arbitrariness, their main advantage is that they provide accurate measurements of the magnitude of the impacts. However, those quantitative impacts may be unclear, thus requiring to be somehow interpreted in a qualitative way. In contrast, the main advantages of qualitative assessments are the prioritization of risks and the identification of the most important areas for improvement. Even so, they do not provide enough quantifiable measurements concerning the probabilities and impacts of risks. As a result, semi-quantitative methods basically take advantage of both these aforesaid aspects and, therefore provide risk prioritizations and useful (semi-)quantifiable impact analyses.

2.2. Is risk management needed in Cloud organizations?

New risks have appeared together with the evolution of the Cloud computing paradigm. Within them we find specific issues imposed by law or regulations, as well as operational risks inherent to the use of Cloud systems, either local or external assets. These risks can have a great impact on the operation of Cloud providers, making it inconsistent with their respective business strategies, represented by means of business objectives (BLOs) and/or constraints. As suggested by ISO 31000:2009, proper

Please cite this article in press as: J.O. Fitó, J. Guitart, Business-driven management of infrastructure-level risks in Cloud providers, Future Generation Computer Systems (2012), doi:10.1016/j.future.2012.05.008

Download English Version:

https://daneshyari.com/en/article/6873576

Download Persian Version:

https://daneshyari.com/article/6873576

Daneshyari.com