Research Report

How to Reliably Verify Compliance with Cloud SLAs from a Consumer's Perspective?

TODAY'S CLOUD CONSUMERS HAVE VERY LIMITED MEANS FOR VERIFYING THE COMPLIANCE TO PRIOR NEGOTIATED SLAS INDEPENDENTLY FROM A CLOUD PROVIDER. IN ORDER TO OBTAIN RELIABLE MEANS FOR SERVICE MONITORING, WE PROPOSE AN INTERMEDIARY-BASED SOLUTION IN CONJUNCTION WITH DIFFERENT STRATEGIES FOR ROBUST CLOUD MONITOR PLACEMENT.

Melanie Holloway

Ralf Steinmetz

Introduction

Nowadays, the use of IT has become a necessity in a variety of industrial sectors. Even in the financial industry, the provisioning of financial services highly depends on IT. Due to the highly competitive environment in the financial services industry, lowering IT-related costs constitutes a major requirement. Furthermore, the productivity and quality of the financial products and services is increasingly dependent on the IT. Thus, the requirements on system availability and IT service quality are also increasing.

In terms of cost and flexibility, cloud computing can provide competitive advantages. Basically, cloud computing enables the dynamic provisioning of IT resources. The functional properties of these resources (e.g., CPU or RAM) can be configured by consumers. Besides this high level of functional flexibility, cloud services also permit cost savings due to, e.g., the consolidation of IT resources in large data centers (Creeger, 2009).

In general, the needs of cloud consumers not only differ from a functional perspective, but also include – depending on the particular business requirements – different levels of service quality. However, when using services from the cloud, consumers face a loss of control over service quality due to a shift of responsibility to the cloud provider. In order to address that issue, today's cloud providers offer some quality guarantees in the form of Service Level Agreements (SLAs). However, cloud providers often put the burden of reporting SLA violations on their customers and even provide own monitoring solutions for detecting such violations. Since data obtained in this way cannot be regarded as an independent base of evidence, we explore how to verify the compliance of cloud providers with the negotiated SLAs in a reliable manner from a consumer's perspective in our research work. In detail, we address the following two research questions:

1. How to provide reliable means for monitoring cloud SLAs from a consumer's perspective?
2. How to determine reliable locations for placing monitoring components?

Solution: SLA Management-as-a-Service

In order to address the issues mentioned above, we have designed an intermediary-based solution. The intermediary, in the form of a cloud broker, is acting as an independent, trusted third party which performs SLA management tasks, such as SLA monitoring, on behalf of cloud consumers. Besides the monitor design, the access to the monitoring components during monitoring must also be taken into account. Since failures of IT resources, such as outages of virtual machines (VMs) or network components, may prevent the monitoring components from delivering any data, an optimization of the distribution of the monitoring components is required. The resulting research problem is denoted as Robust Cloud Monitor Placement Problem. A tangible example is provided in Figure 1.

Given a number of cloud services (e.g., S1) to be monitored residing in one or more data centers of a cloud provider (e.g., PD1 and PD2), locations for the monitoring components should be chosen in such a manner that maximizes the reliability of the whole monitoring system. In this respect, we use redundant monitor components for each cloud service in order to increase the fault tolerance of the monitoring system and, thus, to increase its overall reliability. We further assume that the cloud broker is also running several private data centers (e.g., BD1 and BD2) at different geographical locations in order to additionally assess the quality of a service from a consumer's perspective. In doing so, monitoring components can be placed on VMs in a data center on provider- and broker-side. Depending on the placement of a monitoring component, the underlying VM as well as the network connection in between the monitoring component and the cloud service to be

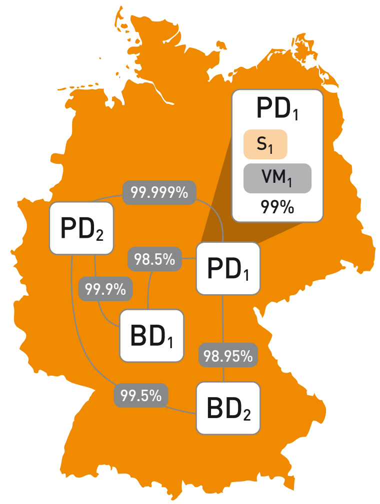

Figure 1: Simplified Example of a Robust Cloud Monitor Placement Problem

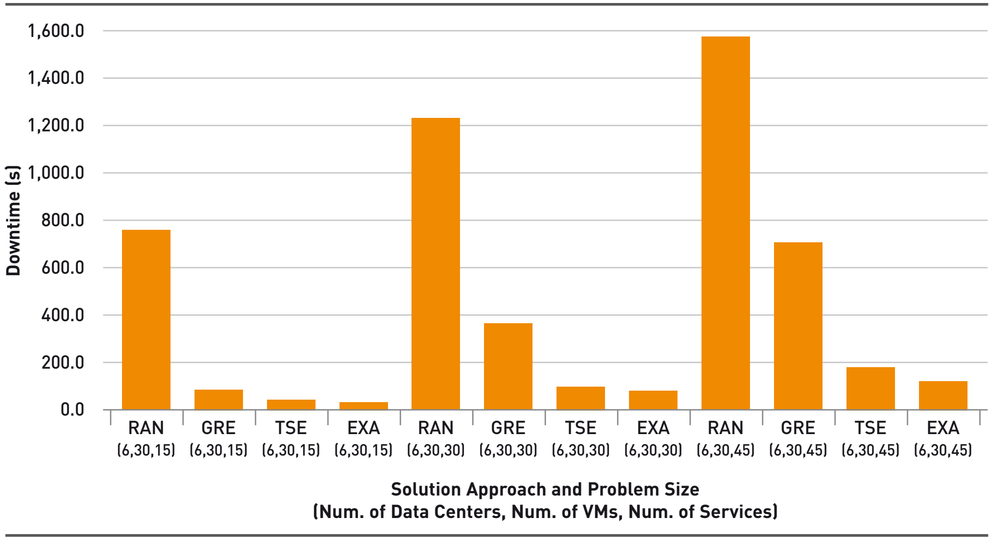

Figure 2: Mean Reliability in Terms of Probable Downtime

monitored may be more stable or unstable. In addition, resource constraints must be taken into account. The monitoring components have a certain resource demand depending on the cloud service to be monitored and the VM candidates for monitor placement only offer a limited resource supply. For the complete optimization model, the interested reader is referred to our publication (Siebenhaar et al., 2014).

To minimize the probability of failures and thus, to maximize the reliability of the monitoring tasks, we deal in our current work with the development of appropriate procedures for the optimal distribution of the monitoring components. In order to obtain an exact solution of the underlying optimization problem, techniques from the field of Operations Research are applied.

Furthermore, we have developed different heuristic approaches that trade solution quality against computation time.

Empirical Findings

To evaluate the practical applicability of our proposed solution approaches, we have performed a quantitative evaluation. The objective of this evaluation was to examine the computation time and the solution quality. For this purpose, we have generated 14 test cases, each incorporating a different number of data center sites, VMs on each site, services running on each VM, and redundant monitor components to be placed for each service. For each of these test cases, we have randomly generated 100 different problem instances using realistic data concerning availability guarantees and VM capacities from well-known cloud providers, such as Amazon, and statistics from the PingER project (PingER, 2016) in order to model the reliability of the network connections. Each problem instance was solved using our exact solution approach (EXA), our two heuristic approaches (GRE and TSE), and by using a random placement (RAN). Selected results of the evaluation are depicted in Figure 2.

The evaluation revealed that the optimization problem can only be solved exactly for a small number of VMs and cloud services. This is due to the fact that the computation time grows exponentially with an increasing problem size. For example, it already takes around 2 min in case of 30 VMs and 45 services. Hence, the practical applicability of the exact solution approach with regard to real data centers is very limited. In contrast, our heuristic approaches are able to solve similar problem instances in a few seconds. In terms of solution quality, the results of a random placement also emphasize the need for heuristic solutions. By applying a random-based procedure to the problem instance mentioned before, the monitoring components would still exhibit a level of unreliability that equals a downtime of 25 min (on a yearly basis). In contrast, the monitoring components will probably only experience a few seconds of downtime when applying a heuristic solution approach for their placement.

Conclusion

Cloud computing promises high functional flexibility and cost-savings, but consumers also face a loss of control over service quality. Therefore, cloud providers usually offer quality guarantees in the form of SLAs. However, cloud consumers have very limited means for verifying the compliance to prior negotiated SLAs in a reliable manner.

In this research report, we have presented an intermediary-based solution for SLA management, in which a trusted cloud broker conducts monitoring tasks on behalf of consumers. In this respect, the reliability of the monitoring components not only depends on the monitoring itself, but is also affected by the reliability of the monitor locations. Therefore, we have developed different optimization approaches for robust monitor placement. The evaluation revealed that an exact solution approach is not applicable in practice due to its high computational complexity. In contrast, our heuristic solution approaches permit a tremendous reduction of computation time, while their solution quality is very close to the exact approach.

References

Creeger, M.:
Cloud Computing: An Overview.
In: ACM Queue, 7 (2009) 5, pp. 1-5.

Siebenhaar, M.; Schuller, D.; Wenge, O.; Steinmetz, R.:
Heuristic Approaches for Robust Cloud Monitor Placement.
In: Proceedings of the 12th International Conference on Service Oriented Computing, Paris, France, 2014.

PingER:
Standford University. http://www-iepm.slac.stanford.edu/pinger, 2016.