



Analysis and Strategy for the Performance Testing in Cloud Computing

By Eljona Proko & Ilia Ninka

Vlora University, Albania

Abstract - The aim of this study is the analysis and presentation of some ideas on performance testing in Cloud Computing. Performance is an important factor in testing a web application. Performance testing in cloud computing is different from that of traditional applications. Our research methodology in this article includes an overview of existing works on testing performance in Cloud Computing, focusing on discussion that the traditional benchmarks are not sufficient to analyze performance testing in Cloud Computing. In this study we are focused mainly on analysis performance metrics in Cloud Computing, based on their characteristics such as elasticity, scalability, pay-per-use and fault tolerance, and then we discuss why needed new strategies for performance testing in Cloud Computing and creation of new benchmarks. From this study we conclude that the performance testing and evaluation should be performed using new models testing, which are created according to Cloud Computing characteristics and metrics.

Keywords : *Cloud computing, characteristics, performance, testing, benchmarks, strategy.*

GJCST-B Classification: *C.2.1*



Strictly as per the compliance and regulations of:



Analysis and Strategy for the Performance Testing in Cloud Computing

Eljona Proko^α & Ilia Ninka^σ

Abstract - The aim of this study is the analysis and presentation of some ideas on performance testing in Cloud Computing. Performance is an important factor in testing a web application. Performance testing in cloud computing is different from that of traditional applications. Our research methodology in this article includes an overview of existing works on testing performance in Cloud Computing, focusing on discussion that the traditional benchmarks are not sufficient to analyze performance testing in Cloud Computing. In this study we are focused mainly on analysis performance metrics in Cloud Computing, based on their characteristics such as elasticity, scalability, pay-per-use and fault tolerance, and then we discuss why needed new strategies for performance testing in Cloud Computing and creation of new benchmarks. From this study we conclude that the performance testing and evaluation should be performed using new models testing, which are created according to Cloud Computing characteristics and metrics.

Keywords : *Cloud computing, characteristics, performance, testing, benchmarks, strategy.*

I. INTRODUCTION

Modern computer system is becoming more complex and this depends on the network technologies on the internet. Performance testing [1] intended to measure system throughput and latency with varying number of concurrent users, over extended periods of times, and with different load profiles. Performance testing in cloud computing is different from that of traditional applications. The traditional performance testing focused on the performance metrics for applications that are under a particular workload for a fixed configuration. Cloud test need to measure the performance metrics related to the workloads that run in a distributed fashion on multiple virtual and real machines. The growth of cloud computing created a demand for new strategy that can measure the performance characteristics of cloud applications.

This paper begins by describing Cloud Computing definition. Section III describes cloud computing characteristics. Section IV discusses traditional benchmarks problems regarding performance testing in Cloud Computing. Section V discusses ideas for new strategies and creation of new

Author α : Eljona Proko Computer Science Department, Vloja University, Albania. E-mail : elzavalani@gmail.com

Author σ : Ilia Ninka Informatics Head Department, Tirana University, Albania. E-mail : ilia.ninka@unitir.edu.al

models in testing cloud computing. In Section V we conclude this study.

II. CLOUD COMPUTING

Cloud Computing is a model that offers the vision of a virtually infinite pool of computing, storage and networking resources where applications can be scalable deployed [2]. Fig.1 illustrates cloud computing scheme. This cloud model promotes availability and is composed of five essential characteristics, four deployment models, and three service models [3].

a) Essential Characteristics

i. On-demand self-service based usage model

A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service's provider.

ii. Multi Tenancy with resource pooling

The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter). Examples of resources include storage, processing, memory, network bandwidth, and virtual machines.

iii. Broad network access for distributed resources

Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, laptops, and PDAs).

iv. Elasticity to provision capabilities quickly

Capabilities can be rapidly and elastically provisioned, in some cases automatically, to quickly scale out and rapidly released to quickly scale in. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be purchased in any quantity at any time.

v. Measured Service

Cloud systems automatically control and optimize resource use by leveraging a metering

capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts).

Resource usage can be monitored, controlled, and reported providing transparency for both the provider and consumer of the utilized service.

b) *Deployment Models*

i. *Private cloud*

The cloud infrastructure is operated solely for an organization. It may be managed by the organization or a third party and may exist on premise or off premise.

ii. *Public cloud*

The cloud infrastructure is made available to the general public or a large industry group and is owned by an organization selling cloud services.

iii. *Community cloud*

The cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be managed by the organizations or a third party and may exist on premise or off premise.

iv. *Hybrid cloud*

The cloud infrastructure is a composition of two or more clouds (private, community, or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load-balancing between clouds)

c) *Service Models*

i. *Cloud Software as a Service (SaaS)*

The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through a thin client interface such as a web browser (e.g., web-based email). The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

ii. *Cloud Platform as a Service (PaaS)*

The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly application hosting environment configurations.

iii. *Cloud Infrastructure as a Service (IaaS)*

The capability provided to the consumer is to provide processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating system; storage, deployed applications, and possibly limited control of select networking components (e.g., host firewalls).

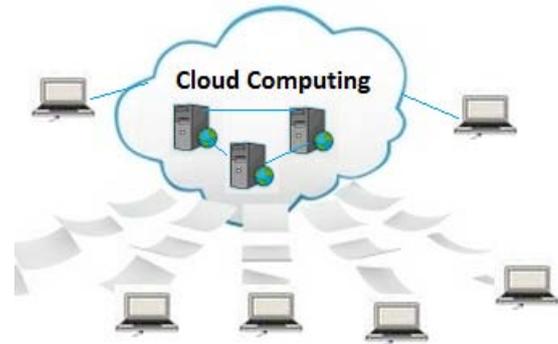


Fig. 1: Cloud Computing

III. CLOUD COMPUTING CHARACTERISTICS

A metric is used to measure and understand the behavior of software. Cloud metrics can be used to measure the behavior of cloud which utilizes the resources from the computers as a collective virtual computer, where the applications can run independently from particular computer or server configurations [10]. Cloud delivers its services through internet and provides the full user functionality of a software application by the web sites which provide Software as a Service. Dynamic web sites provide regularly changing information to users and utilize dynamically generated pages and maintain data for display in a database [9]. Cloud uses the dynamic web sites to deliver the web applications on demand. Cloud metrics should follow some characteristics which help to evaluate cloud on each and every parameter which is necessary for a good quality cloud, so that a client can rely on it to choose the best cloud.

The main advantages of cloud computing are scalability, pay-per-use and fault-tolerance [4].

a) *Elasticity* [8] is one of the major factors for the success of the cloud as an IT infrastructure. For a DBMS deployed on a pay-per-use cloud infrastructure, an added goal is to optimize the system's operating cost. Elasticity, i.e. the ability to deal with load variations by adding more resources during high load or consolidating the tenants to fewer nodes when the load decreases, all in a live system without service disruption, is therefore critical for these systems. Even though

elasticity is often associated with the scale of the system, a subtle difference exists between elasticity and scalability when used to express a system's behavior.

b) Scalability is a desirable property of a system, which indicates its ability to either handle growing amounts of work in a graceful manner or its ability to improve throughput when additional resources (typically hardware) are added. A system, whose performance improves after adding hardware, proportionally to the capacity added, is said to be a scalable system.

c) Reliability is the probability that a product or part will operate properly for a specified period of time (design life) under the design operating conditions (such as temperature, volt, etc.) without failure [6]. The outcome of the measurement process is reproducible that is similar to results over time for some different inputs and across many different situations. Cloud gets many requests simultaneously and will also give the similar results for some requests in a period of time so clouds have to be reliable.

d) Availability Cloud Services should be available maximum time [7]. The on demand, elastic, scalable, and customizable nature of the cloud must be considered when deploying cloud architectures. Many different clients might be accessing the same back-end applications, and many provider are providing the cloud services has the expectation that only their application will be properly delivered to users. In cloud computing it is essentially required to gather the information instantly without making a user to wait and the gathered information should be related to each other.

e) Cost Cloud Computing allows an organization to pay by the hour of computing resources, potentially leading to cost savings even if the hourly rate to rent a machine from a cloud provider is higher than the rate to own one. This is essentially preferable when demand for a service that varies over time.

f) Fault Tolerance is one of the key issues of cloud computing. There are many fault tolerance techniques in parallel computing [11]. Fault tolerance is concerned with all the techniques necessary to enable a system to tolerate software faults. These software faults may or may not Manifest themselves during systems

operations, but when they do, software fault tolerant techniques should provide the necessary mechanisms of the software system to prevent system failure occurrences.

IV. TRADITIONAL BENCHMARKS ARE NOT SUFFICIENT TO ANALYZE PERFORMANCE TESTING IN CLOUD COMPUTING

The goal of benchmarking a software system is to evaluate its average performance under a particular workload. TPC-W [14] has been designed for transactional database systems. Cloud systems usually do not offers strong consistency constraints because most web-based applications only require lower levels of consistency. As a consequence existing TPC-W implementations for the cloud are not conforming to the specification. The primary metric used by the TPC-W [13] is WIPS that the system under test can handle. By scaling the number of emulated browsers, the number of requests and the load on the system can be increased. WIPS is useful in the context of a static system it is not for adaptable and scalable systems. The second metric of the TPC-W is \$/WIPS, is based on the total cost of ownership of the system under test including software, hardware, maintenance and administration expenses. These overall costs are then divided by the maximum number of WIPS to calculate the \$/WIPS. In the context of cloud computing does not exists maximum number of WIPS. Thus, there exists no fixed load for which the overall cost can be calculated. TPC-W became outdated in front the evolution of web applications and does not reflect modern access-paths. TPC-W benchmark has not the adequate metrics for measuring the characteristics of cloud systems, such as pay-per use scalability and fault tolerance.

V. IDEA FOR TESTING PERFORMANCE IN CLOUD COMPUTING

Performance is generally tied to an application's capabilities within the cloud infrastructure itself. Testing is a periodic activity and requires new environments to be set up for each project [12]. Web applications must be tested for multiple operating systems and updates, multiple browser platforms and versions, different types of hardware and a large number of concurrent users to understand their performance in real-time [5]. Cloud Computing is growing at a rapid pace. With the advent of this technology, there is bound to be an increase in demand for Cloud Testing. New cloud test should be based on an e-commerce scenario (i.e., a web-shop) and define web interactions as test drivers. Thus, the test should allow the evaluation of the complete application stack. A new cloud test should analyze the ability of a dynamic system to adapt to a changing load

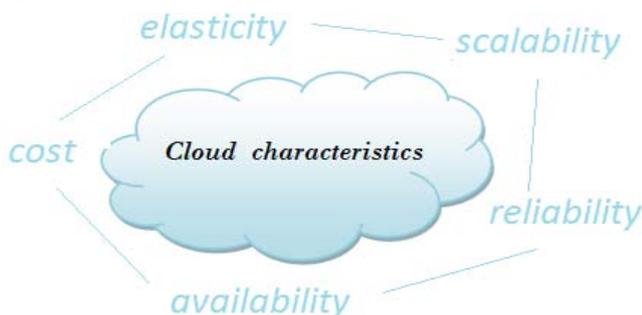


Fig. 2: Some of the cloud computing characteristics

(including peaks) in terms of scalability and costs. Moreover, another goal is to test to the assumption of infinite scalability of an application in the cloud. Cloud providers often replicate data over different data centers for availability but also performance reasons. In order to get a fair comparison of the test results, the emulated browsers should run in different locations (worldwide). By doing this, we can achieve that the test results are not biased due to the location where the test driver is running. A solution to this problem is to run the test drivers on a cloud infrastructure of a provider which supports location based installations. A new test should comprise web interactions that resemble the access patterns of Web 2.0 like applications. One example is to add web interactions that allow users to write and read reviews of individual products or to add web interactions that allow user communities to exchange the latest news about certain products. Web 2.0 applications often include multimedia content (audio files, video files, pictures) which can be accessed by users. This content produces heavy load on the servers which host that content. Cloud makes it cost-effective for creating separate test regions for system testing. The test strategy should answer what is intended to be achieved by moving testing to the cloud, including cost savings, easy access to infrastructure, reduction in cycle times, etc. The strategy should define the type of tests to be performed in the cloud, the risks associated and the duration of the tests. We need to define the infrastructure requirements necessary for building a test environment by selecting the required testing tools and applications, hardware and software, bandwidth, etc. The next step is selection of a service provider for security, quality, reliability and any discrepancies in the terms and conditions. Executing the test is the critical phase where applications are tested according to the defined test strategy. Monitor and analyze test results is the last step. It is advised that test results be monitored in real-time to understand and react to capacity- or performance-related issues. Also, analyze cloud usage against chargeback costs to understand the financial performance of cloud services.

VI. CONCLUSION

The growth of cloud computing created a demand for benchmarks that can measure the performance characteristics of cloud applications.

The traditional performance benchmarking focused on the performance metrics for applications that run on single node systems. Cloud benchmarks need to measure the performance metrics related to the workloads that run in a distributed fashion on multiple virtual and real machines. Metrics are the necessary and important elements for evaluation the quality enabling the identification of a good Cloud Computing. The performance metrics for the distributed workloads need

to be defined based on the cloud application characteristics.

REFERENCES RÉFÉRENCES REFERENCIAS

1. J.D. Meier et al. Performance Testing Guidance for Web Applications. Microsoft Corporation, United States, 2007.
2. B. Hayes, "Cloud computing", Communications of the ACM, vol. 51, no. 7, pp. 9–11, Jul. 2008.
3. Peter Mell, Timothy Grance The NIST Definition of Cloud Computing, Special Publication 800-145
4. M. Armbrust et al. Above the clouds: A Berkeley view of cloud computing. Technical Report UCB/EECS-2009-28, 2009.
5. Ian Molyneux - The Art of Application Performance Testing.
6. Kareim M. Sobhe, Ahmed Sameh "Multi-Channel Clustered Web Application Server"
7. "Making Cloud Service Continuity a Reality" NetPrecept Software Ltd.
8. Team Sardes, Inria Rhône-Alpes, Elasticity in Cloud Computing, June 23, 2011
9. David Cleary "Web Based Development and Functional Size Measurement" IFPUG Annual conference.
10. Gurdev Singh, Shanu Sood, Amit Sharma CM-Measurement Facets for Cloud Performance. International Journal of Computer Applications (0975 – 8887) Volume 23– No.3, June 2011
11. Chunye Gong, Jie Liu, Qiang Zhang, Haitao Chen and Zhenghu Gong The Characteristics of Cloud Computin
12. Cognizant, Taking Testing to the Cloud.
13. TPC BENCHMARK™ W (Web Commerce) Specification Version 1.21.1 October 18 August 23, 2000
14. Márk Kaszó, Tihamér Levendovszky TPC Benchmark.