Journal of King Saud University – Computer and Information Sciences (2017) 29, 365–377

King Saud University

Journal of King Saud University – Computer and Information Sciences

www.ksu.edu.sa
www.sciencedirect.com

A Variable Service Broker Routing Policy for data center selection in cloud analyst

Ahmad M. Manasrah [a,*], **Tariq Smadi** [a], **Ammar ALmomani** [b]

[a] *Yarmouk University, Faculty of Information Technology and Computer Science, Jordan*
[b] *Al-Balqa Applied University, Dept. of Information Technology, Jordan*

Received 2 October 2015; revised 12 December 2015; accepted 12 December 2015
Available online 28 March 2016

KEYWORDS

Cloud-Analyst;
Cost effective analysis;
Data centers;
Service broker policy;
Simulation and modeling

Abstract Cloud computing depends on sharing distributed computing resources to handle different services such as servers, storage and applications. The applications and infrastructures are provided as pay per use services through data center to the end user. The data centers are located at different geographic locations. However, these data centers can get overloaded with the increase number of client applications being serviced at the same time and location; this will degrade the overall QoS of the distributed services. Since different user applications may require different configuration and requirements, measuring the user applications performance of various resources is challenging. The service provider cannot make decisions for the right level of resources. Therefore, we propose a Variable Service Broker Routing Policy – VSBRP, which is a heuristic-based technique that aims to achieve minimum response time through considering the communication channel bandwidth, latency and the size of the job. The proposed service broker policy will also reduce the overloading of the data centers by redirecting the user requests to the next data center that yields better response and processing time. The simulation shows promising results in terms of response and processing time compared to other known broker policies from the literature.

1. Introduction and background

Cloud computing is a model for facilitating on-demand network access to shared and configurable computing resources such as servers (IaaS), operating systems (PaaS), applications and services (PaaS) that can be made available and released with less administration efforts or service provider involvements. In a short time, cloud computing has been applied widely in many applications, it became an essential part of the next generation of computing infrastructure at low cost, that enables users to utilize their resources as a pay-per-use as portrayed in Fig. 1.

The main facet of cloud computing is the adoption of virtualization, in which virtual machines (VM) are running on top of the available hardware to satisfy the users need and demand (Kremer, 2013; Armbrust et al., 2010). Therefore, managing VMs is an important aspect to be considered to keep the whole

* Corresponding author.
E-mail addresses: ahmad.a@yu.edu.jo (A.M. Manasrah), tariksmadi@gmail.com (T. Smadi), ammarnav6@gmail.com (A. ALmomani).
Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.jksuci.2015.12.006
1319-1578 © 2016 The Authors. Production and hosting by Elsevier B.V. on behalf of King Saud University.

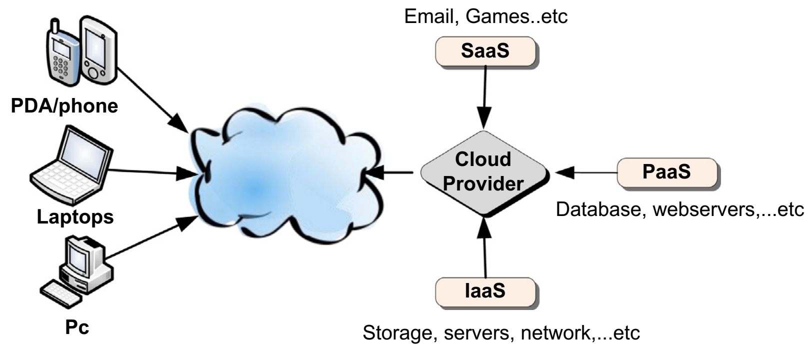

Figure 1 Cloud computing overview.

cloud running efficiently, which is carried out by the hypervisor (Kremer, 2013; Rimal et al., 2009). The selection of the VM for a particular workload is done by the load balancer. The load balancer distributes the load in a way that ensures no VM is swamped with requests at one time (Mell and Grance, 2009; Rajesh and Sreenivasulu, 2014). Above this level, another abstraction level called the service broker, acts as an intermediary between the users and the cloud service providers. The service broker utilizes existing service broker policies to route user requests to the most appropriate data center (Houidi et al., 2011; Limbani and Oza, 2012; Kapgate, 2014; Mishra et al., 2014). Therefore, the optimal response time of a particular request and the efficient utilization of the datacenters are governed through a proper data center selection policy. Especially that, data centers are under the service providers control at diverse locations, which can be configured with different types of hardware based on utilization and clients' demands (Houidi et al., 2011; Limbani and Oza, 2012; Mishra et al., 2014; Sharma, 2014; Mishra and Bhukya, 2014; Rekha and Dakshayini, 2014). The existing broker policies of data center selection are based on the location of the data centers, response time or current execution load, and the cost of the data center usage (Rimal et al., 2009; Dinh et al., 2013). Therefore, the objective of this paper is to illustrate an enhanced proximity service broker policy that selects a data center based on the network latency and bandwidth to ensure efficient and reliable request execution over data centers (i.e. minimized response and execution time).

Cloud computing is a very complex process; it depends on uncontrollable factors like network congestion and servers varying workloads. However, measuring the performance of internet based applications using real cloud platform is difficult (Armbrust et al., 2010; Iosup et al., 2011; Dillon et al., 2010). Therefore, simulation-based approaches are provided to solve such issue virtually and free of charge under stable and controllable environment (Dinh et al., 2013; Wickremasinghe et al., 2010). Calheiros et al. (2011), proposes an extensible toolkit for modeling and simulating cloud computing systems called CloudSim. The CloudSim provides a set of components that provide the base for cloud computing, including Virtual Machines (VM), Cloudlets (Jobs and user request will be used interchangeably), datacenters (DC), Service broker and hosts. Each of them has its own characteristics and functionality. DCs consist of a number of physical hosts, each of which manages a number of allocated VMs. There is a policy to maintain the efficiency of the VM allocation. CloudSim offers a straight forward policy, which is a first-come-first-serve (FCFS) policy. The Datacenter within the CloudSim has its own characteristics like architecture, OS, list of Virtual Machines, allocation policy (time-shared or space-shared), time zone and the cost of the provided services. Virtual machine parameters are: Size, Ram, MIPS (Million Instructions per Second) and bandwidth. Cloudlet (job) parameters are: length (Number of Instructions), user, input file size and output file size (Calheiros et al., 2011). Users within the user base generate their own requests and send them as a job to the cloud through the Service Broker. The service broker selects an appropriate data center according to the service broker policy. The result is returned back to the user through the service broker in a reversed order after the job is finished (Limbani and Oza, 2012; Wickremasinghe et al., 2010). Therefore, this paper proposes a service broker policy for datacenter selection with the best possible response time, delay and bandwidth (i.e. availability) to serve the user's requests. The algorithm balances between the delay, bandwidth and the request size in selecting the most suitable data center. Few scenarios were conducted to introduce heavy and light loads on the datacenters. The proposed algorithm shows an enhancement in response and processing time compared to other algorithms but almost similar in terms of the overall cost. In this paper we present some enhancement over the service proximity broker policy through taking the communication channel bandwidth, latency and the size of the job into account in an attempt to come up with a new service broker policy that achieves minimum response and processing time.

The rest of the paper is organized as follows: Section 2 discusses the existing service broker policies and the datacenter selection algorithms, hence, the problem formulation. Section 3 demonstrates the proposed service broker policy. Section 4 discusses the simulation environment setup and description as well as discussing the results of the simulation using the new proposed policy. Section 5 concludes the topic and Section 6 provides an idea on the future scope.

2. Data center selection problem

Since the main goal of the service brokers is to direct the user requests to the best DC with optimal performance, the service broker policy has to efficiently select the best data center for the job considering many factors such as time, cost, and