

Resource Prediction for Inter-cloud Broker

Mohammad Aazam, Eui-Nam Huh

► **To cite this version:**

Mohammad Aazam, Eui-Nam Huh. Resource Prediction for Inter-cloud Broker. 11th IFIP International Conference on Network and Parallel Computing (NPC), Sep 2014, Ilan, Taiwan. pp.559-562, 10.1007/978-3-662-44917-2_52 . hal-01403142

HAL Id: hal-01403142

<https://hal.inria.fr/hal-01403142>

Submitted on 25 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Resource Prediction for Inter-Cloud Broker

Mohammad Aazam¹, Eui-Nam Huh¹

¹ Department of Computer Engineering, Kyung Hee University,
Suwon, Republic of Korea
aazam@ieee.org, johnhuh@khu.ac.kr

Abstract. Media content over the Internet has massively been increasing, resulting in popularity gain of cloud computing. Cloud computing is the only solution in hand to handle rapidly increasing digital media content. Through cloud computing, digital media can be manipulated, stored, and communicated in a much better and easier way. But due to increase in user's demands and diversity of applications, it is, at times, not possible for a single cloud to fulfill all the requests. At that point, multiple clouds have to communicate and share resources through an intermediary, called cloud broker. To handle requests properly, broker has to predict the amount of resources required when a service is requested. This paper focuses on this particular issue. We present resource prediction part of our model here, along with its implementation and evaluation using CloudSim toolkit.

1. Introduction

The rapidly increasing digital media content has already surpassed traditional media, as a result of which long-term and vast changes are required for the contents shared over the Internet. In 2010, Internet video traffic had surpassed global peer-to-peer (P2P) traffic [1]. Excluding the amount of video exchanged through P2P file sharing, at the time being, Internet video is 40 percent of consumer Internet traffic. Since 2012, it has become over 50 percent and will reach 62 percent by the end of 2015. Counting all form of videos, the number will be approximately 90 percent by 2015 [2]. This media revolution not only brings great opportunities, but also bears some challenges. To meet those challenges, much better infrastructure, sophisticated technologies, and powerful capabilities are required to be incorporated.

Cloud computing still faces some open challenges, but to provide better reliability, availability, cost-efficiency, and QoS, inter-cloud computing has already been envisioned. Research on inter-cloud computing is still in its start, but its effectiveness cannot be denied by any means [3]. Cloud Service Providers (CSPs) have their customers dispersed all around the globe. To serve them optimally, CSPs have to setup many of their data centers at different geographical locations. Existing systems are not capable enough to coordinate dynamically the load distribution among data centers, to determine optimal location for hosting services to achieve

desired performance. Furthermore, users' geographical distribution cannot be predicted as well. Thus, load coordination and service distribution has to be done automatically. Inter-cloud computing is meant to counter this problem. It provides scalable provisioning of services with consistent performance, under variable workload and dynamically changing requirements. It supports dynamic expansion and contraction of resources, to handle abrupt variations in service demands [4]. In inter-cloud computing, an intermediary, broker, is responsible to identify appropriate CSP, according to the needs of its customer, through cloud exchange. Broker negotiates with the gateway to allocate resources, according to user and service requirements [4]. Resource management is a key attribute of broker. So far, the available literature addresses resource management issue in a trivial way.

In this paper, we present a part of our resource management model, in which we focus mainly on advanced reservation of resources, according to the type of customer and service. Implementation and simulation of our model was performed on CloudSim 3.0.3 toolkit.

2. Resource Prediction and Premium Amount Calculation

When a cloud service customer (CSC) [5] requests the broker [6] for a particular service, broker has to further contact cloud service provider (CSP) and negotiate the contract, including service level agreement (SLA) [7]. For some services, broker performs ad hoc allocation of resources, while for others; resources are required to be predicted. Based on the type of service, broker's resource prediction is formulated as under:

$$R_{res} = \sum_{i=0}^n \left\{ \begin{array}{l} (C_{m_i} * n * P_i(L|H)) * (1 - P_i(L|H)) * (1 - \mu_i) * D_s \\ 0 \end{array} \right. \quad (1)$$

Where R_{res} is required service, C_m is the maximum cost a particular user can afford or willing to pay, n is the number of cloud customers, $P(L|H)$ is the probability of a particular customer of giving up the resource. For simplicity, we have categorized it into two, as low (L) or high (H) probability.

$$0 > L \leq 0.5, 0.5 > H \leq 1 \quad (2)$$

' μ ' is User Characteristic, which represents the characteristic of the requesting user, based on its history. This value is assigned by the broker, on the basis of previous resource consumption log-file of a particular user. For an altogether new user, this characteristic is set to maximum positive value for the first time. After predicting the resources, broker asks its users to pay a particular premium amount, based on following formulation:

$$\rho_{prem} = C_m * P_L * \mu_L \quad (3)$$

Where, ρ_{prem} determines premium amount, P_L represents probability as “low”. Actual and final price is paid once the resources are started to be consumed. Therefore, in equation 3, relinquish probability is set as low, instead of both low and high. Same is the case with user characteristic, which is set as positive for low relinquish probability customers.

The implementation and evaluation of this part of model was done on CloudSim 3.0.3 toolkit. Results are presented below.

Shown in figure 1, the unit is greater for L customers, having low relinquish probability, since they are more loyal, as compared to H customers. Prices for a total of 9 services have been shown, ranging from USD 100 to USD 500. On vertical axis, unit of resources to be reserved increase with the higher valued service. Better and costly service require more resources, hence, the unit of resources increase accordingly.

Figure 2 shows the premium amount to be paid, according to service type. Premium amount is shown on the top of the bars, in USD, while the values at the bottom represent the total price of services, in USD. Premium amount of each of the 9 services is shown.



Fig. 1. Resource prediction for different types of customers, according to different services.

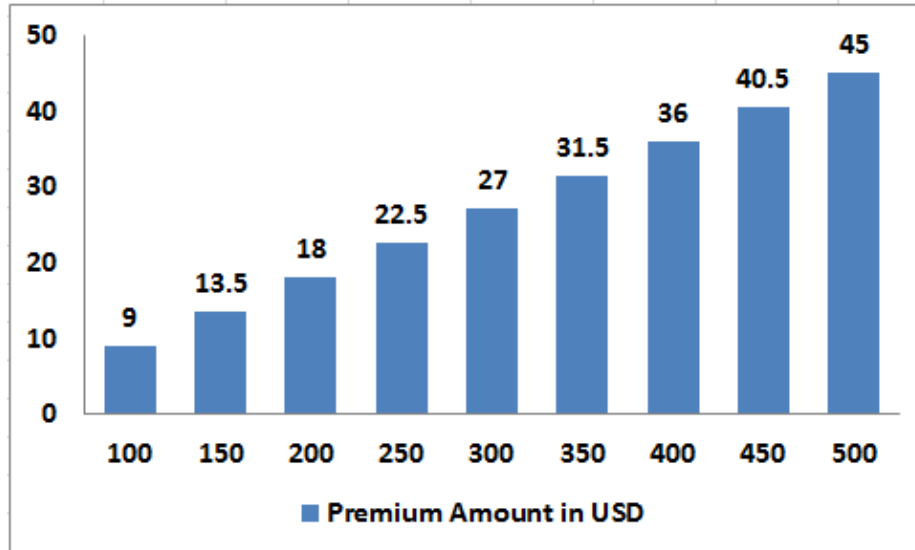


Fig. 2. Premium amount, according to service type.

Acknowledgment. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education(No.NRF-2013R1A1A2013620). The corresponding author is Prof. Eui-Nam Huh.

References

- [1] Mingfeng Tan, Xiao Su, "Media Cloud: When Media Revolution Meets Rise of Cloud Computing", Proceedings of The 6th IEEE International Symposium on Service Oriented System Engineering, Irvine, CA, USA, 12-14 December, 2011.
- [2] Cisco-White-Paper, "Cisco Visual Networking Index – Forecast and Methodology, 2010–2015," June 1, 2011.
- [3] Nikolay Grozev and Rajkumar Buyya, "Inter-Cloud Architectures and Application Brokering: Taxonomy and Survey", Wiley Software: Practice and Experience, 2012.
- [4] Buyya, Rajkumar, Rajiv Ranjan, and Rodrigo N. Calheiros. "Intercloud: Utility-oriented federation of cloud computing environments for scaling of application services." Algorithms and architectures for parallel processing. Springer Berlin Heidelberg, 13-31, 2010.
- [5] Z. Wenwu, L. Chong, W. Jianfeng, and L. Shipeng, "Multimedia Cloud Computing," Signal Processing Magazine, IEEE, vol. 28, pp. 59-69, 2011.
- [6] Mohammad Aazam, Eui-Nam Huh, "Inter-Cloud Architecture and Media Cloud Storage Design Considerations" in the proceedings of 7th IEEE CLOUD, Anchorage, Alaska, USA, 27 June – 02 July, 2014.
- [7] Daniel Díaz-Sánchez, Florina Almenarez, Andrés Marín, Davide Proserpio, and Patricia Arias Cabarcos, "Media Cloud: An Open Cloud Computing Middleware for Content Management", IEEE Transactions on Consumer Electronics, Vol. 57, No. 2, May 2011.