Resilient service provisioning in cloud based data centers

Mahmoud Al-Ayyoub, Muneera Al-Quraan, Yaser Jararweh,
Elhadj Benkhelifa, Salim Hariri

PII: S0167-739X(17)31468-1
DOI: http://dx.doi.org/10.1016/j.future.2017.07.005
Reference: FUTURE 3535

To appear in: *Future Generation Computer Systems*

Received date : 16 November 2016
Revised date : 5 June 2017
Accepted date : 1 July 2017

Please cite this article as: M. Al-Ayyoub, M. Al-Quraan, Y. Jararweh, E. Benkhelifa, S. Hariri, Resilient service provisioning in cloud based data centers, *Future Generation Computer Systems* (2017), http://dx.doi.org/10.1016/j.future.2017.07.005

Resilient Service Provisioning in Cloud based Data Centers

Mahmoud Al-Ayyoub[a], Muneera Al-Quraan[a], Yaser Jararweh[a], Elhadj Benkhelifa[b], Salim Hariri[c]

[a]*Department of Computer Science, Jordan University of Science and Technology, Irbid 22110 Jordan, e-mail: maalshbool@just.edu.jo, yijararweh@just.edu.jo*
[b]*Staffordshire University, UK. e-mail: E.Benkhelifa@staffs.ac.uk*
[c]*University of Arizona, USA. e-mail: hariri@ece.arizona.edu*

Abstract

Cloud service providers usually have several geo-distributed data centers to meet these increasing demands. On peak times, these data centers receive unpredictable amount of workload that might go beyond their capacity and service providers need to distribute this workload with the objective of maximizing their revenue. This objective is hard to be guaranteed given the possible incidents that may cause cloud services interruption such as power outages, security breaches, natural disasters, etc. Since cloud providers are facing serious threats to the continuation of their services offering which will cause a serious decrease in their revenues, they must have proper strategies for handling such incidents. While a rich volume of recent research works focused on optimizing cloud data centers operations in order to reduce their operational cost (OPEX), little attention has been paid to such serious incidents and how they would affect service provisioning. In this study, we use mixed integer-linear programming to model the problem of maximizing revenue of a given workload by creating/expanding/renting data centers. The model takes into account a wide range of incidents affecting the users traffic and/or the data centers service capacities. Moreover, it handles many issues such as different types of costs (capital, operational and recovery), power consumption, carbon footprint, SLA related issues such as constraints (e.g., delay) and penalties, etc. To the best of our knowledge, this formation is unique among the body of existing works. Finally, to evaluate the proposed model and the impact of different parameters on its performance, several simulation experiments are conducted.

Keywords: cloud data centers, incident management, operational cost, workloads modeling, optimization problems, disaster recovery

Nomenclature

- T Number of time segments (i.e., the number of years).
- H Number of timeslots (i.e., the number of hours).
- U Set of user locations.
- S Set of data center locations.
- S^* Set of existing data center locations.
- x_s^t A binary variable representing whether a data center is built on location s at year t or not.
- X Set of x_s^t variables.
- $R_s^{t,h}$ The percentage of lost service capacity due to abnormal circumstances in data center s at hour h of year t.
- $K_{s,u}^{t,h}$ The percentage of failure over the link between u to s due to abnormal circumstances at hour h of year t.
- $K_s^{t,h}$ The total percentage of failures over all links to data center s due to abnormal circumstances at hour h of year t.
- $L_u^{t,h}$ The total number of service requests originating from user location u during hour h of year t.
- $\lambda_{s,u}^{t,h}$ A variable representing the portion of $L_u^{t,h}$ served under normal circumstances by the data center at location s.
- Λ Set of $\lambda_{s,u}^{t,h}$ variables.
- $y_{s,u}^{t,h}$ A binary variable representing the ability of data center s to handle service requests from user location u at hour h of year t under normal circumstances.
- Y Set of $y_{s,u}^{t,h}$ variables.