Dimensioning of multi-clouds with follow-the-renewable approaches for environmental impact minimization

Miguel Felipe Silva Vasconcelos[1,2], Daniel Cordeiro[2], Fanny Dufossé[1]

[1] Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LIG, 38000 Grenoble, France
{miguel.silva-vasconcelos, fanny.dufosse}@inria.fr

[2] School of Arts Science and Humanities, Univ. of São Paulo, São Paulo, Brazil
daniel.cordeiro@usp.br

Mots-clés : *Cloud computing, Scheduling, Follow-the-renewables.*

The growing appetite for Cloud services leads to an unprecedented increase in the electricity consumption of Data Centers. For instance, according to a recent report, in 2014, data centers in the U.S. consumed about 1.8% of total U.S. electricity [5]. This energy consumption is among the sources of pollution and global greenhouse gas emissions of Information and Communication Technologies (ICT), with the extraction of raw materials and end-of-life product recycling. To reduce the negative impacts of ICT on global warming, major Cloud actors like Amazon AWS, Apple or Microsoft are involved in projects to deploy solar power facilities [4].

This project is a follow-up of papers [1,2,3]. It focuses on a geographically distributed cloud. Virtual machines (VM) arrivals are considered unpredictable and no assumption will be made concerning future submissions. Each data center (DC) is supplied by a farm of photovoltaic panels (PV). The PV forecast model is based on a truncated normal law. For each data center and at each future time slot, an expected value of generation and a variance value compose the truncated normal law forecasting the PV generation. One challenge is to formulate a realistic communication network between data centers, including energy consumption. This energy consumption cannot be considered as an on-site consumption for the sending or receiving data center, and is thus not concerned by local PV production.

The first objective consists in developing a realistic model of Clouds supplied by green energy, and a simulation platform to compare scheduling algorithms. This model and this platform will then be used to examine the problem of cloud dimensioning, to minimize the ecological impact of data centers in terms of brown energy consumption and IT products manufacturing. The first objective of this project consists in developing an accurate model of geographically distributed clouds, including data centers associated with PV farms and batteries. The model will comprise a realistic battery model including charge and discharge rate, self-discharge and battery aging. The communication network model in each data center and between two data centers will consider the bandwidth constraints, and the energy consumption over time. The model will then be extended with a more realistic electric network model with each data center connected to a local network offering different sources of electricity with different Greenhouse gases (GHG) emissions, and different volatility levels. This platform modeling will be used to design scheduling heuristics targeting a reduction of the brown (nonrenewable) energy usage with low losses due to battery usage. Intuitively, the battery will mostly be discharged during the night, but can also be used during the day when workload requests exceed PV production. This work will give a detailed probability and expectation estimate to formulate the optimization criteria. Online scheduling requests short delays for scheduling decisions. At a given time slot, VMs waiting for scheduling imply to evaluate the distribution law of battery level and brown energy consumption for many future time slots, which are strongly correlated. However, the exact distribution law of the battery after few time slots does not follow a truncated normal law, due to the limit of capacity. An exact evaluation of expected battery level and brown energy consumption would thus imply complex formula that can not be computed within a short time. With these estimates, scheduling decisions need to determine VMs allocation based on multiple criteria, including battery storage and brown energy consumption, but also battery aging and battery level distribution in the different DCs. This requires trade-offs between these different parameters according to the available data and expectations.

The second objective corresponds to the dimensioning of PV and battery according to the DCs sizes and the amount of arriving VMs, for brown energy reduction but also for a purely financial criterion. This problem formulation will be based on the model detailed in the first objective. The idea is to determine the PV dimensioning for efficiently supplying the cloud and battery dimensioning, in particular for smoothing production during the day and for night computations. This problem is barely addressed in the literature, while green energy is ubiquitous in Clouds supply. This dimensioning problem will be formulated to include the environmental cost of manufacturing PV panels, battery and servers. If this project focuses on brown energy reduction, PV and servers manufacturing is a major cause of concern in the different sources of pollution related to cloud computing. A cloud dimensioning targeting the environmental concern cannot neglect this aspect of the issue. In particular, for both ecological and financial criteria, the purchase cost prohibits extensive battery usage. This problem implies taking into account multiple ecological and financial optimization criteria, but also the capacity of the system to process all VMs directly after submission, including during peaks of activity. This implies the knowledge of the maximum activity peaks workload. Another approach consists in delaying VM allocation during peaks under constraints of quality of service (VMs with deadline or penalty for delayed VMs). This implies new criteria of quality of service for the dimensioning problem. In any case, the problem implies a stochastic model for photovoltaic production overtime on a yearly basis, and a second yearly stochastic model for workload submission. The method will consist of evaluating, based on these distribution laws, the expected brown electricity consumption at any time slot on a yearly basis, the function of the number of servers and PV panels and the batteries sizes. The environmental cost of brown energy consumption per year will then be compared to the GHG emissions due to manufacturing. The considered scheduling heuristics do not make hypothesis on submission rates, but VMs characteristics will deeply impact the schedules. The most challenging part will however concern the combination of the yearly PV forecast with the distribution law computed by the scheduler to predict at any time future time slots production regarding actual and past ones. The accurate estimate of green energy production and energy consumption is however necessary to determine an optimal or approximate dimensioning of clouds.

Références

[1] B. Camus, A. Blavette, F. Dufossé, and A. Orgerie. Self-Consumption Optimization of Renewable Energy Production in Distributed Clouds. In *2018 IEEE International Conference on Cluster Computing (CLUSTER)*, pages 370–380, Belfast, UK, 2018. IEEE.

[2] B. Camus, F. Dufossé, A. Blavette, M. Quinson, and A. Orgerie. Network-Aware Energy-Efficient Virtual Machine Management in Distributed Cloud Infrastructures with On-Site Photovoltaic Production. In *2018 30th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD)*, pages 86–92, Lyon, France, 2018. IEEE.

[3] Benjamin Camus, Fanny Dufossé, and AnneCécile Orgerie. A stochastic approach for optimizing green energy consumption in distributed clouds. In *SMARTGREENS 2017 - International Conference on Smart Cities and Green ICT Systems*, Porto, Portugal, 2017. SMARTGREENS.

[4] Clicking green : who is winning the race to build a green internet. Greenpeace report, 2017.

[5] Arman Shehabi, Sarah Smith, Dale Sartor, Richard Brown, Magnus Herrlin, Jonathan Koomey, Eric Masanet, Nathaniel Horner, Inês Azevedo, and William Lintner. United states data center energy usage report, 2016.