A Survey of CPU Process Surrogate Placement Methods and Techniques in Cloud Computing

Komal Rane [1], Gagan Sharma [2], Rajesh Sharma [3],
[1,2,3]Department of Computer Science & Engineering,
RKDF University, Bhopal, India

Abstract: In cloud computing network, the efficient process management and power saving have been a challenging task for researchers. The proper process migrations and consolidations are required for energy efficiency in cloud network. The running CPU processes under a server are migrated into another server if overloading occurs and processes are consolidated from different server into single server if loads are fewer for saving energy. This survey provides an overview of different CPU process surrogate methods and techniques used in cloud computing network. Various process surrogate placement methods present their benefits and conditions as a prominent approach toward cloud computing environment.

Keywords: Cloud Computing, Process Surrogation, Migration, Consolidation.

I Introduction

The convenience of high-speed Internet computing and high scale computational ability at processor level has lead to work a completely innovative model of computation as cloud computing. The resources such as high speed processors, huge storage, memory and network etc. are utilized on rental basis and payment is made in pay-per-use form in cloud computing [1, 2]. The computing resources into servers, storage, applications, networks and services are available as on-demand to the users. These resources are provided to demanding users and cost to the users is calculated on the basis of the resource-usage [3 – 5]. Further, the provisioning is elastic i.e. resources may be added if required and removed when not required.

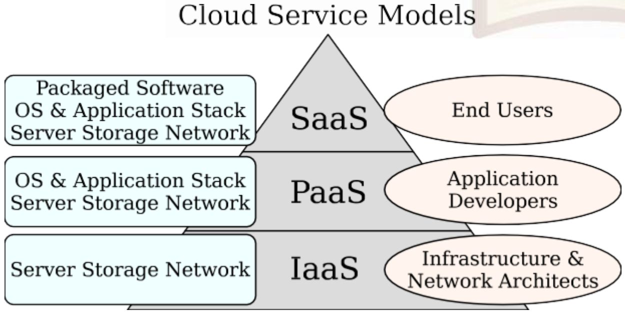

Figure 1: Cloud Service Model

Additionally, based on the type of services offered, there are classified three service models Infrastructure-as-a-Service (IaaS) [6], Platform-as-a-Service (PaaS) [7] and Software-as-a-Service (SaaS) [8] as presented in Figure 1. Based on territorial deployment concerns, there are four deployment models as Private, Public, Hybrid and Community model. Traditionally, the concept of Cloud Computing is implemented by a group of data centers managed and operated by Cloud service providers. These data centers are equipped with large number of processing elements and huge amount of other resources such as memory, storage, and bandwidth etc. The power consumed by these computational resources is significant and has attracted the attention of many researchers.

II Virtualization and Process Surrogation

The fundamental technology which makes the idea of Cloud computing practically possible is virtualization. It creates virtual instances of a physical server and these instances are offered as a service to the user on a shared basis. These instances are often known as Virtual Machines (VMs). Physically, every host (sometimes known as server or node) consists of many VMs. And many such hosts comprise a data center. To address the issue of energy consumption, it is suggested to keep minimal number of host live (active or running) at any given point of time. To achieve this, one may need to migrate few VMs from one host to another based on certain criteria. Hence, VM migration can be used to address the issue of energy consumption [9, 10].

In this paper, we aim to focus on the issue of reducing energy consumption by considering the communication cost between user and service provider using geographical location as one of the factors. Moreover, after selection of data center with lowest communication cost, we intend to reduce number of running hosts by applying existing VM consolidation and migration techniques [11, 12].

III Related Work

In increasing demand for Internet based services, in that large amount of process as like computational data, resource management and network based communications that are significantly contribute to energy consumption. And Cloud computing is a multi-directional solution to make process and network communication easier.

Chatziprimou *et al.* [13] presented a runtime optimization framework for cloud configuration management. The suggested framework may operate with two fitness functions, the first applies a simulation-guided fitness evaluation model. The second applies lightweight surrogate fitness evaluation models based on statistical regression techniques. Their evaluation results provide evidence that the use of surrogate models can efficiently guide the search for solution within seconds. This is the first work o introduce the use of fitness approximation models for tackling the challenges of optimality versus timeliness in the problem of loud configuration optimizations.

The comparative metrics of ***goodness-of-fit*** and ***predictive ability*** is evaluated of surrogate models to examine their accuracy. The ***goodness-of-fit*** measures how well the surrogate reflects the system behavior. To evaluate it, the coefficient of determination (R^2) [14] is measured as:

$$R^2 = 1 - \frac{\sum \left(Y_i - \hat{Y}_i \right)^2}{\sum \left(Y_i - \frac{1}{N}\sum Y_i \right)^2} \leq 1 \qquad (1)$$

To evaluate the predictive ability of the surrogates, the mean absolute percent error (MAPE) [15, 16] is calculated. the mean absolute percent error (MAPE) offers a weighted error measure of the predictive ability of the surrogate model and it is defined as:

$$\text{MAPE} = \frac{1}{N} \sum^{N} \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \qquad (2)$$

where $\hat{Y}_i$ is the predictive value for the sample Y_i and N the number of observations in the dataset.

Salahuddin *et al.* [17] designed a push-based content placement (CP) algorithm that pre-emptively stores content and pulls popular and correlated content within QoS distance. Content placement (CP) algorithm leverages the induced Online Social Networking (OSN) – based relationships and their affect on correlation and popularity of videos. We show that popularity and correlation aware placement with popular and correlated content within QoS strikes a balance between cost and latency. Our future research directions include designing dynamic push-based CP algorithms that can adapt to changing end-user requests and using reinforcement learning strategies to determine when the content placement should be updated.

Karthikeyan *et al.* [18] demonstrated Naive Bayes with hybrid optimization to limit the energy consumption for virtual machine (VM) migrations in distributed computing. From this exploration, possibility of prediction of virtual machine (VM) failure is exhibited along with the outcomes that utilize cloudsim. The outcomes reveal that the proposed approach yielded better outcomes in prediction of virtual machine (VM) failure in cloud server farms. In the analysis of the quantity of failures in virtual machine (VM) migration occasions, the Artificial Bee Colony Optimization, Bat Algorithm (ABC – BA) approach has less number of failures in virtual machine (VM) migration occasions than that of random migration and ideal migration. It has been moved from servers that neglect to fulfill the heap adjust condition to the goal servers and the load mindful migration algorithm was utilized. The execution of the proposed framework is compared with previous studies and contracted with the current strategies using achievement and failure rate, and energy consumption as measures.

IV Proposed Approach

CPU process surrogation in cloud computing is a virtual machine (VM) Migration process which is a $NP-$ Hard problem and this problem can be solved in less time using some meta-heuristic algorithm. All such implementations can be effectively simulated using a tool called **CloudSim**. Many $NP-$Hard problem can effectively solved by Any Colny Optimization (ACO) method which has many variants; different variants may be applied on VM migration problem.

So various Ant Colny Optimization (ACO) can be applied and performance of all these variants is compared with each other in terms of number. of VM Migrations.

V Conclusion

In this paper, a survey of CPU process surrogate placement methods and techniques in cloud computing system are presented. We also investigate the relevance methods and techniques for the purpose of efficient process surrogation. The various researchers presented exploratory study of process surrogate placement methods and analysis represented with their proposed methods and approaches. This survey provides an fundamental overview of popular methods applied for process surrogation in cloud computing environment.

REFERENCES

[1] H. Jin, S. Ibrahim, T. Bell, L. Qi, H. Cao, S. Wu, and X. Shi, *Tools and Technologies for Building Clouds*. London: Springer London, 2010, pp. 3–20. [Online]. Available: http://dx.doi.org/10.1007/978-1-84996-241-4 1

[2] B. Xu, Z. Peng, W. Ke, M. Zhong, and A. M. Gates, “Deployment method of vm cluster based on graph theory for cloud resource management,” *IET Communications*, vol. 11, no. 5, pp. 622–627, 2017. [Online]. Available: https://doi.org/10.1049/iet-com.2016.0071

[3] S. Chaisiri, B. Lee, and D. Niyato, “Optimization of resource provisioning cost in cloud computing,” *IEEE Transactions on Services Computing*, vol. 5, no. 2, pp. 164–177, April 2012. [Online]. Available: https://doi.org/10.1109/TSC.2011.7

[4] W. J. Lloyd, S. Pallickara, O. David, M. Arabi, T. Wible, J. Ditty, and K. Rojas, "Demystifying the clouds: Harnessing resource utilization models for cost effective infrastructure alternatives," *IEEE Transactions on Cloud Computing*, vol. 5, no. 4, pp. 667–680, Oct 2017. [Online]. Available: https://doi.org/10.1109/TCC.2015.2430339

[5] B. Palanisamy, A. Singh, and L. Liu, "Cost-effective resource provisioning for mapreduce in a cloud," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 5, pp. 1265–1279, May 2015. [Online]. Available: https://doi.org/10.1109/TPDS.2014.2320498

[6] S. K. Panda and P. K. Jana, *An Efficient Resource Allocation Algorithm for IaaS Cloud*. Cham: Springer International Publishing, 2015, pp. 351–355. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-14977-637

[7] D. Beimborn, T. Miletzki, and S. Wenzel, "Platform as a service (paas)," *WIRTSCHAFTSINFORMATIK*, vol. 53, no. 6, pp. 371–375, 2011. [Online]. Available: http://dx.doi.org/10.1007/s11576-011-0294-y

[8] R. Sharma and M. Sood, *Cloud SaaS: Models and Transformation*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 305–314. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-24055-3 31

[9] G. Eslami and A. Toroghi Haghighat, "A new surrogate placement algorithm for cloud-based content delivery networks," *The Journal of Supercomputing*, vol. 73, no. 12, pp. 5310–5331, Dec 2017. [Online]. Available: https://doi.org/10.1007/s11227-017-2088-5

[10] P. Zoghi, M. Shtern, M. Litoiu, and H. Ghanbari, "Designing adaptive applications deployed on cloud environments," *ACM Trans. Auton. Adapt. Syst.*, vol. 10, no. 4, pp. 25:1–25:26, Jan. 2016. [Online]. Available: http://doi.acm.org/10.1145/2822896

[11] Y. Jin and Y. Wen, "When cloud media meet network function virtualization: Challenges and applications," *IEEE MultiMedia*, vol. 24, no. 3, pp. 72–82, 2017. [Online]. Available: https://doi.org/10.1109/MMUL.2017.3051519

[12] R. Ranjan, "The cloud interoperability challenge," *IEEE Cloud Computing*, vol. 1, no. 2, pp. 20–24, July 2014. [Online]. Available: https://doi.org/10.1109/MCC.2014.41

[13] K. Chatziprimou, K. Lano, and S. Zschaler, "Surrogate-assisted online optimisation of cloud iaas configurations," in *2014 IEEE 6th International Conference on Cloud Computing Technology and Science*, Dec 2014, pp. 138–145. [Online]. Available: https://doi.org/10.1109/CloudCom.2014.101

[14] *The coefficient of determination in multiple regression*. Boston, MA: Springer US, 1997, pp. 91–95. [Online]. Available: https://doi.org/10.1007/978-0-585-25657-3 19

[15] D. A. Swanson, J. Tayman, and T. M. Bryan, "Mape-r: a rescaled measure of accuracy for cross-sectional subnational population forecasts," *Journal of Population Research*, vol. 28, no. 2, pp. 225–243, Sep 2011. [Online]. Available: https://doi.org/10.1007/s12546-011-9054-5

[16] D. A. Ahlburg, "Error measures and the choice of a forecast method," *International Journal of Forecasting*, vol. 8, no. 1, pp. 99–100, 1992. [Online]. Available: https://doi.org/10.1016/0169-2070(92)90010-7

[17] M. A. Salahuddin, A. Mseddi, H. Elbiaze, and R. H. Glitho, "Popularity and correlation-aware content placement for hierarchical surrogates in cloud-based cdns," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, Dec 2017, pp. 1–6. [Online]. Available: https://doi.org/10.1109/GLOCOM.2017.8254119

[18] K. Karthikeyan, R. Sunder, K. Shankar, S. K. Lakshmanaprabu, V. Vijayakumar, M. Elhoseny, and G. Manogaran, "Energy consumption analysis of virtual machine migration in cloud using hybrid swarm optimization (abc–ba)," *The Journal of Supercomputing*, Sep 2018. [Online]. Available: https://doi.org/10.1007/s11227-018-2583-3