A contribution to scheduling of cluster networks with finite-source*

Attila Kuki[a], Tamás Bérczes[a], Ádám Tóth[a], János Sztrik[a]

[a]University of Debrecen, Hungary
[kuki.attila;berczes.tamas;toth.adam;sztrik.janos]@inf.unideb.hu

Introduction

This paper deals with scheduling jobs in heterogeneous resources of networks, like the computational grid. In the literature, various job allocation algorithms have been proposed to schedule arriving jobs in computational clusters [5], [4], [1]. In addition, some algorithms have been designed to consider knowledge about the characteristics of jobs; these algorithms may be classified as either clairvoyant or as non-clairvoyant [8], [10], [9].

Apart from the effective scheduling, the energy consumption of such grid systems turns into a really crucial requirement due to the rapid increase of the size of the grid and the goal of a green network. The most common techniques of reducing energy consumption are related to the dynamic power management used at runtime. It is therefore of interest to examine algorithms which offer the greatest performance while using an amount of energy that is as low as possible.

The Model

Do introduced a generalized infinite model for the performance evaluation of scheduling compute-intensive jobs with unknown service times in computational clusters [2]. In this paper we use a finite model instead of the infinite one [7] to make the queueing model more realistic and we introduce two new scheduling policies. In addition to the existing HP policy, new policies are introduced:

- *HP (High Performance priority)*: This policy chooses the shortest queue in the system. If there is more than one queue with this property, a queue whose server has the highest performance is chosen.
- *MRT (Mean Response Time priority)*: This policy first calculates the expected mean response time for every queue and then selects a queue where this value is the minimal.
- *MRTHP (Mean Response Time with High Performance priority)*: This policy is a combination of MRT and HP. If there is an idle server, it behaves like the HP policy; if all servers are busy, it behaves like MRT.

*The research work was supported by the construction EFOP - 3.6.3 - VEKOP - 16-2017-00002. The project was supported by the European Union, co-financed by the European Social Fund.

To calculate the performance, mean response times and energy consumption of a server, we consider every server of the cluster to be one of a specific type (class), which can be characterized by the following parameters: C_s - the throughput of the server; $P_{\text{ac},s}$ - the average active energy consumption of the server under full load; $P_{\text{id},s}$ - the power consumption of the server in the idle state. These measures are applied based on the SPECpower_ssj2008 benchmark [6]. According to these parameters the modeled servers were Intel Xeon E5-2670, Intel Xeon E5-2660, and Intel Xeon E5-4650L.

We investigate these policies with respect to three schemes of buffering the arriving jobs:

- *Separate Queue*: In this scheme every server has its own queue. Jobs are scheduled to the queue of a specific server according to the chosen policy.
- *Class Queue*: In this scheme a buffer is assigned to each class. Jobs are scheduled to the queue of a specific class according to the chosen policy.
- *Common Queue*: In this scheme only a single common buffer is available for all servers. If more than one server is idle, then the local scheduler chooses the server with the highest performance.

In this paper, we present a generalized finite source model for the performance evaluation of scheduling compute-intensive jobs with unknown service times in a computational cluster which is built from servers of different types. The state space of the describing Markov chain is extremely large. Therefore, to obtain the performance measures the SimPack, a collection of C/C++ libraries and executable programs for computer simulation is used [3].

In particular, we determine various performance measures for all combinations of three scheduling policies for assigning jobs to servers with three schemes for buffering arriving jobs; furthermore, we investigate the effect of switching off idle servers, thus the energy consumption of the system under these combinations of scheduling policies and buffering schemes can be estimated.

Large number of figures were generated to illustrate the effects of scheduling algorithms and buffering schemes for the performance measures and the energy consumption. Simulation results show that the choice of the scheduling policy and of the buffering scheme plays an important role in ensuring quality of service parameters such as the waiting time and the response time experienced by arriving jobs. The energy consumption, however, is only affected by the scheduling policy and the energy saving mode, while the buffering scheme does not have significant impact.

Keywords: performance evaluation, cluster network, finite-source queueing systems, buffering scheme

References

[1] M. Cankar, M. Artač, M. Šterk, U. Lotrič, B. Slivnik: *Co-Allocation with Collective Requests in Grid Systems*, Journal of Universal Computer Science 19.3 (2013), pp. 282–300, DOI: https://doi.org/10.3217/jucs-019-03-0282.

[2] T. V. Do, B. T. Vu, X. T. Tran, A. P. Nguyen: *A generalized model for investigating scheduling schemes in computational clusters*, Simulation Modelling Practice and Theory 37.0 (2013), pp. 30–42.

[3] P. A. Fishwick: *Simpack: Getting Started With Simulation Programming In C And C++*, in: WSC '92 Proceedings of the 24th Conference on Winter Simulation, ed. by J. S. et al., ACM, New York, 1992, pp. 154–162.

[4] A. Tchernykh, J. Ramírez, A. Avetisyan, N. Kuzjurin, D. Grushin, S. Zhuk: *Two level job-scheduling strategies for a computational grid*, in: Proceedings of the 6th International Conference on Parallel Processing and Applied Mathematics, 2006, pp. 774–781.

[5] G. Terzopoulos, H. D. Karatza: *Performance evaluation of a real-time grid system using power-saving capable processors*, The Journal of Supercomputing 61.3 (2012), pp. 1135–1153.

[6] The Standard Performance Evaluation Corporation: *SPECpower_ssj2008 Result File Fields*, Web Page: https://www.spec.org/power/docs/SPECpower_ssj2008-Result_File_Fields.html.

[7] Á. Tóth, T. Bérczes, A.Kuki, B. Almási, W. Schreiner, J. Wang, F. Wang: *Analysis of finite-source cluster networks*, Creative Mathematics and Informatics 2 (2016), pp. 223–235.

[8] S. Zikos, H. D. Karatza: *A clairvoyant site allocation policy based on service demands of jobs in a computational grid*, Simulation Modelling Practice and Theory 19.6 (2011), pp. 1465–1478.

[9] S. Zikos, H. D. Karatza: *Communication cost effective scheduling policies of nonclairvoyant jobs with load balancing in a grid,* Journal of Systems and Software 82.12 (2009), pp. 2103–2116.

[10] S. Zikos, H. D. Karatza: *The impact of service demand variability on resource allocation strategies in a grid system*, ACM Transactions on Modeling and Computer Simulation 20.19 (2010), pp. 1–29, doi: https://doi.org/10.1145/1842722.1842724.