

EXPERIMENTAL STUDY OF AN EXPONENTIAL NEIGHBORHOOD FOR THE SERVERS LOAD BALANCING PROBLEM

I. A. Davydov

*Sobolev Institute of Mathematics,
Novosibirsk State University, Novosibirsk
e-mail: vann.davydov@gmail.com*

We consider the servers load balancing problem which is originated from optimal web hosting in cloud computing. We are given a set of servers. Each server contains the set of disks (images of disks). Internet sites with heterogeneous information are distributed among the disks. A lot of users visit these sites and generate the load for the servers. The load is changed during time interval and can be characterized by some parameters: CPU, RAM and others. We assume that the activities of users are known for each site and as result for each disk and each server. If the load of the server does not exceed a given threshold, we say that the server is working in regular mode. Otherwise, it works in heavy mode. To avoid the heavy mode we can move the disks from one server to another. This moving requires some computational efforts. We will call them as overhead expenses. We assume that the expenses are known for each parameter if a disk is extracted from one server and moving to another one. Initial allocation of disks among the servers is given. The problem is to find a new allocation of the disks to minimize the total exceeding over the thresholds during the time interval subject to overhead expenses for each server.

We develop a new exponential neighborhood in order to tackle this problem. Elements of this neighborhood are optimal solutions of the assignment problem, which is constructed as follows: we eject one disc from each server and redistribute these discs optimally. We study different disk ejection strategies and exploit the behavior of the local search algorithms based on the randomized version of this neighborhood. Computational results for 20 servers and 200 disks are discussed. Comparison with results from [1] is presented.

REFERENCES

1. Yu. A. Kochetov, N. A. Kochetova. The servers load balancing problem // Vestnik of Novosibirsk State University. Series Information Technology. 2013. Vol. 11, issue 4. P. 71-76. (in Russian)