UDC 621.391

AN APPROACH TO DATA PROCESSING OPTIMIZATION CONSIDERING ENERGY EFFICIENCY AND PERFORMANCE PARAMETERS

Gvozdetska N.A., Globa L.S.

Institute of Telecommunications of National Technical University of Ukraine Igor Sikorsky KPI, Ukraine

E-mail: n.gvozdetska@gmail.com

Підхід до оптимізації обробки даних з урахуванням параметрів енергоефективності та продуктивності

Запропоновано підхід до оптимізації обчислень у центрі обробки даних (ЦОД) за критеріями енергоефективності та продуктивності обчислень. Ефективність запропонованого підходу було перевірено за допомогою натурного експерименту та імітаційного моделювання у середовищі MATLAB.

New services that are going to be provided with the 4G and 5G technologies development require higher rates of data processing and its transmitting. Such services as IoT (Internet of Things) and M2M (Machine to Machine communication) produce big amounts of data that need to be processed as more efficiently as possible. Thus, data centers are consuming more and more power every year.

Servers' sources are used very inefficiently. Some statistics got by Google in 2013 shows that the servers' load is very unsteady. The average load ranges between 15% and 60% for one server and between 40% and 85% for another [1]. It means that the amount of the consumed power is much bigger, than it could be. It causes the challenge to develop a strategy of load balancing within the server cluster that would perform high processing rates and reduce power consumption at the same time [2].

In this paper, an approach to data processing optimization considering energy efficiency and performance parameters is proposed. It is aimed to increase server cluster's performance increasing at the same time its energy efficiency.

According to the proposed approach each node of the server cluster must be preliminary certified to determine its function $P_j = f_j(CPU_j)$. This information should be included into descriptive table of each unit with its RAM volume value, number of processing cores and performance (measured in number of floating point operations per time unit (FLOPS)).

The algorithm that presents an approach consists of 7 steps. It is described by fig.1.

Step 1. Evaluate the state of the cluster at the moment τ_{k-1}

Step 2. Exclude all the impropriate nodes (by RAM and cores available)

Step 3. Find the set of summary consumptions $\boldsymbol{P_\Sigma} = \{\boldsymbol{P_{\Sigma_j}}\}$

Step 4. Sort the set $\boldsymbol{N_j}$by theoretically inserted gain of the consumed power: $\boldsymbol{N_P} = \{\boldsymbol{N_{P_j}}\}$

Step 5. Sort the set of leftover nodes by performance (FLOPS): $\boldsymbol{N_{FLOPS}} = \{\boldsymbol{N_{FLOPS_j}}\}$

Step 6. Grant each node in sets $\boldsymbol{N_P}$and $\boldsymbol{N_{FLOPS}}$ with the marks according to its position in these sets

Step 7. Allocate the task to the node with maximal total mark

Fig. 1. General description of proposed approach

To evaluate the proposed approach performance in real environment an experiment was carried out. For comparison Round Robin algorithm was chosen as simple and widely used algorithm.

For the experiment five nodes were used. Server cluster topology that was used in experiment is presented on the fig.2.

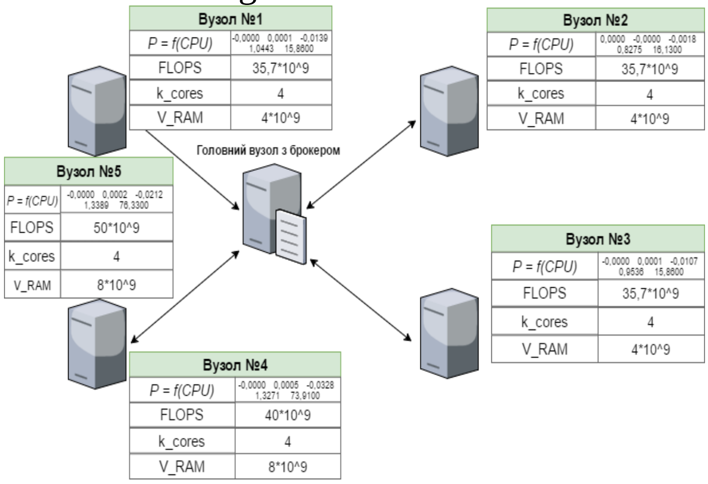

Fig. 2. Topology of cluster evaluated in experiment

On fig.2 each server has its description that must be formed according to the proposed approach within preliminary cluster analysis process. The first row of each descriptive table contains coefficients of polynomials that present $P_j = f_j(CPU_j)$ functions. These functions were got with the help of wattmeter that is presented on the fig.3.1. $P_j = f_j(CPU_j)$ functions in graphical form are presented on the fig.3.2.

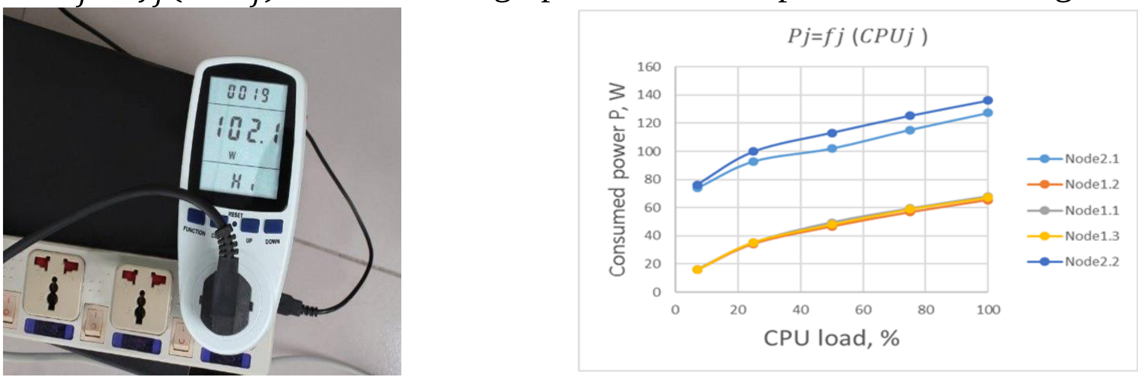

Fig. 3. Preliminary cluster evaluation: 1 – wattmeter for $P_j = f_j(CPU_j)$ defining;
2 - $P_j = f_j(CPU_j)$ functions got experimentally.

To compare the proposed approach with other algorithms, the most widely used and simple one was chosen. Round Robin schedules tasks to each server in the cluster one by one without any priorities. If there are no sources on the node chosen for current task processing, this task goes to a queue.

The scheduler coded in accordance to Round Robin was placed on the main node. The cluster was loaded with the π number to 1000, 5000, 10000, 25000, 50000, 75000, 100000 decimal places counting tasks.

To test the proposed approach the same set of tasks was used. To the main node new scheduler code was placed.

The results of experiment are presented in the table 1.

Table 1. Proposed approach experimental analysis results

Algorithm	Average time of 1 task processing (s)	Power consumed (W)	Gain (in comparison with Round Robin) (%)	
			in time	in power consumed
Round Robin	14.835	376.88	0%	0%
Proposed approach	13.324	365.8	**10,2%**	**3%**

It is possible to see from the table that proposed approach gives some gain in performance and power consumption as well. The gain is not high because of the small number of nodes in cluster and their homogeneity. To show approach's efficiency for bigger cluster MATLAB modeling was held.

The results of MATLAB modeling for cluster including 20 nodes processing typical day workload is presented in the table 2. The typical workload that was used is presented in [4].

Table 2. Modeling of 20 nodes in cluster working the whole day results

Algorithm	Average time of 1 task processing (s)	Power consum ed (W)	Gain (in comparison with Round Robin) (%)	
			in time	in power consumed
Round Robin	15.76	30578	0%	0%
Proposed approach	10.07	28568	**36,1%**	**8,3%**

The results of modeling showed that proposed approach has some benefits in power consumption and in performance as well. The gain of proposed approach reached 36,1% in average processing time and 8,3% in power consumption.

So, in this paper new approach to the energy efficient load balancing in data center was proposed. The proposed approach was evaluated experimentally and with the help of MATLAB modeling. The results showed that the proposed approach gives benefit up to 8,3% in energy efficiency and 36,1% in performance (in case of 20 nodes in cluster).

References

1. Natural Resources Defense Council (NRDS) "Data center efficiency assessment - scaling up energy efficiency across the data center industry: Evaluating key drivers and barriers", http://www.nrdc.org/energy/files/data-center-efficiencyassessment-IP.pdf, 2014.
2. Fermin Armenta-Cano, Andrei Tchernykh, Jorge M. Cortés-Mendoza, Ramin Yahyapour, Alexander Yu. Drozdov, Pascal Bouvry, Dzmitry Kliazovich, Arutyun Avetisyan "Min_c: Heterogeneous Job Consolidation for Power Aware Scheduling with Quality of Service", Russian Supercomputing Days, Moskow, 2015.
3. W. Dargie and A. Schill ''Analysis of the Power and Hardware Resource Consumption of Servers Under Different Load Balancing Policies,'' in Proc. IEEE 5th Int'l Conf. CLOUD, 2012, pp. 772- 778. [Online]. Available: http://dx.doi.org/10.1109/CLOUD. 2012.30.
4. http://www.osp.ru/os/2004/02/183912.