POWER MANAGEMENT FOR SERVER CLUSTERS HARDWARE

UDC 681.3.093:044.3

ZUEV D. O.
International IT Expert Independent Consultant USA, Colorado
e-mail: root@dzuev.pro

DOS E.V.
Lead DevOps Architect EPAM Belarus, Minsk
e-mail: bokarevdos@gmail.com

KROPACHEV A. V.
Automation Solution Department Manager Bell Integrator, USA, Colorado
e-mail: beckett@protonmail.ch

BABKIN O. V.
Bell Integrator Cloud Infrastructure Department Manager Moscow, Russia
e-mail: o.v.babkin@gmail.com

VARLAMOV A. A.
SharxDC LLC CTO Moscow, Russia
e-mail: varlamov.it@gmail.com

INTRODUCTION

Low performance of data center infrastructure work is usually associated with disproportion of servers' utilization. Statistically less than 30% of data servers are under-utilized more than 90 % of the total time while other servers normally cause idle power consumption which leads to the 50% of power loss and inefficient work of data center during peak period [1-5]. Overconsumption problem could be solved by minimization of the active servers' number within the bounds of the server consolidation procedure. Server consolidation is server virtualization technique, which allows workloads encapsulating as virtual machines (VMs) and, thus, run multiple VMs at single server with the aid of hypervisor block (Figure 1).

Main task of data center server consolidation implementation is maintenance of the prior performance level of the servers room infrastructure work. It leads to necessity of the data center peak utilization regime analysis, usually at the 90%, 95% and 99% of maximal recorded value threshold level. Threshold level should be based on the recorded sample in order to get a compromise between stable data center work and opportunity for power savings which is associated with skipping of rare cases of servers' peak load [6, 7]. Thereby, server consolidation may cause performance degradation due to the conflict of using shared resources by VMs [8, 9], specifically last level cache (LLC). The results of cache co-located VMs usage analysis show that sharing LLC between two copies of VMs leads to 20%-30% performance decreasing. The amount of interference could be characterized with a set of parameters, such as effective number of used sets [10]. Allocation of VMs, thus, can be realized by accounting for the amount of the interference and its minimization through the estimation of the required performance requirement.

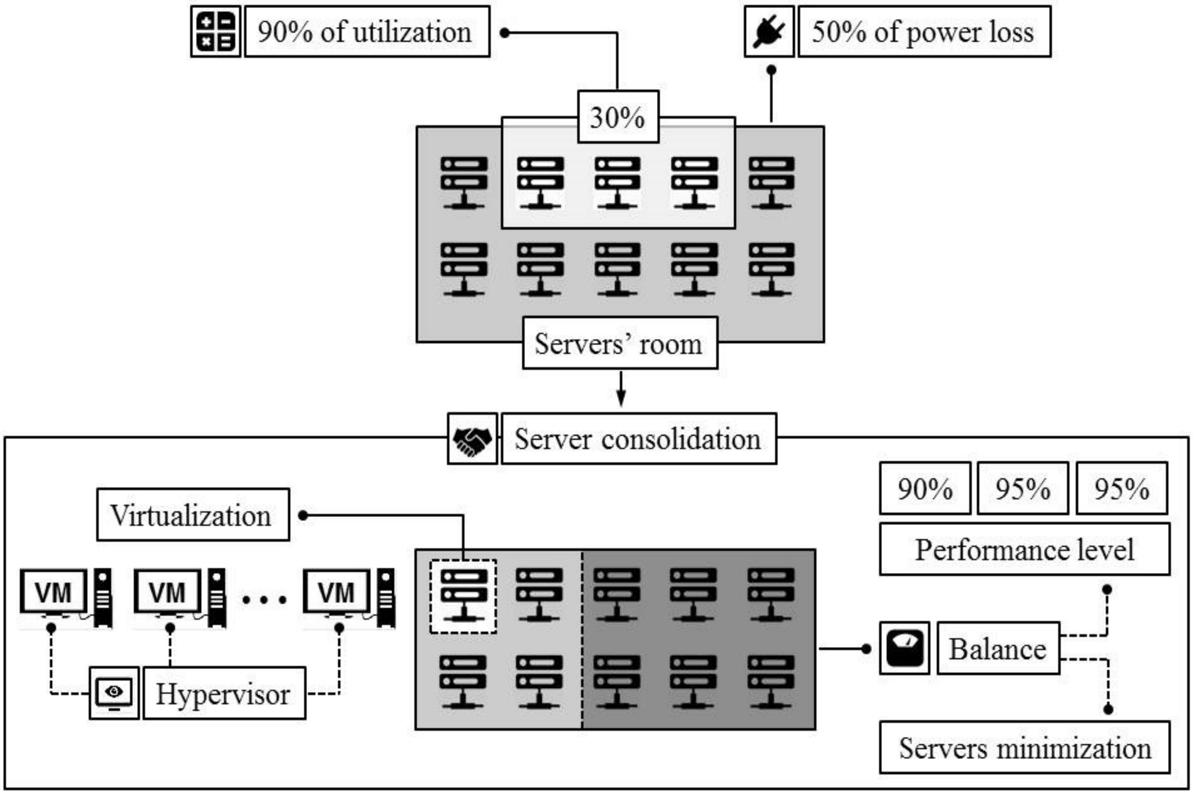

Fig. 1. Data center server consolidation scheme

To develop data center server consolidation methodology there were analyzed recent studies and publications. Statistics of modern data servers' physical resources utilization rate and proportions was considered [1-5]. In order to overcome restrictions and optimize power consumption models that based on the resources utilization threshold value rather than the peak value were discussed [6-7]. Sharing of the servers' resources among co-located VMs, especially LLC problem were studied [8, 9]. To develop the methodology co-located VMs interference with a set of parameters, particularly effective number of used sets was analyzed [10], as well as correlations among VMs' workload [11-14]. Finally it was studied power management solution for data centers scale-out application and targeting distinctive workload characteristics of scale-out applications [15, 16].

KEY ASPECTS OF THE CORRELATION-AWARE POWER MANAGEMENT

As it was mentioned above, server consolidation could be achieved by considering correlation among workload variation. Basic scheme of clustering-based correlation-aware VM development [11, 16] solution includes:

- trace data center servers' physical resources utilization level;
- transform utilization traces into binary sequence up to the utilization threshold value;
- clustering of VMs up to the binary sequence in order to maintain not overlapping of different clusters;
- VMs allocation at physical servers in order to minimize the possibility of the service performance degradation at peak period.

Typical engineering solution is pairing of two uncorrelated VMs into super-VM. Maintaining of the super-VMs can be done by predicting of the aggregated workloads. Once two uncorrelated VMs are paired correlations of them within a same super-VM have not be considered, and possibility of further power consumption decrease will be lost. If servers' utilization is perfectly known this scheme could be extended by utilization of multiple VMs workload of such that for VM placement. To overcome those drawbacks it was proposed [15, 16] to develop power management procedure for cloud services that includes:

- user-interactive and fast changing service;
- maintaining of the minimal performance degradation caused by physical resources sharing conflict;
- high correlation level of VMs.

While the scale-out applications usually operate as highly parallel processes, it is advisable to assign the right number of CPU cores for each VM. At Figure 2 are demonstrated generalized results of recent studies [6, 17] of response time of a websearch cluster with respect to the number of queries for 90% threshold value. The number of allocated cores varied from 4 to 16. It should be noticed that resource utilization level depends on time and usually is lower than the available amount of resources, though dynamic power gating cannot be applicable to this type of applications due to the performance degradation caused by the unapropriable transition delay of power modes switching. Thereby it should be noted necessity of allocating the right number of cores for each VM according to its peak and off-peak resources utilization demands. This procedure has to be implemented at the stage of scaling voltage/frequency level (V/F level).

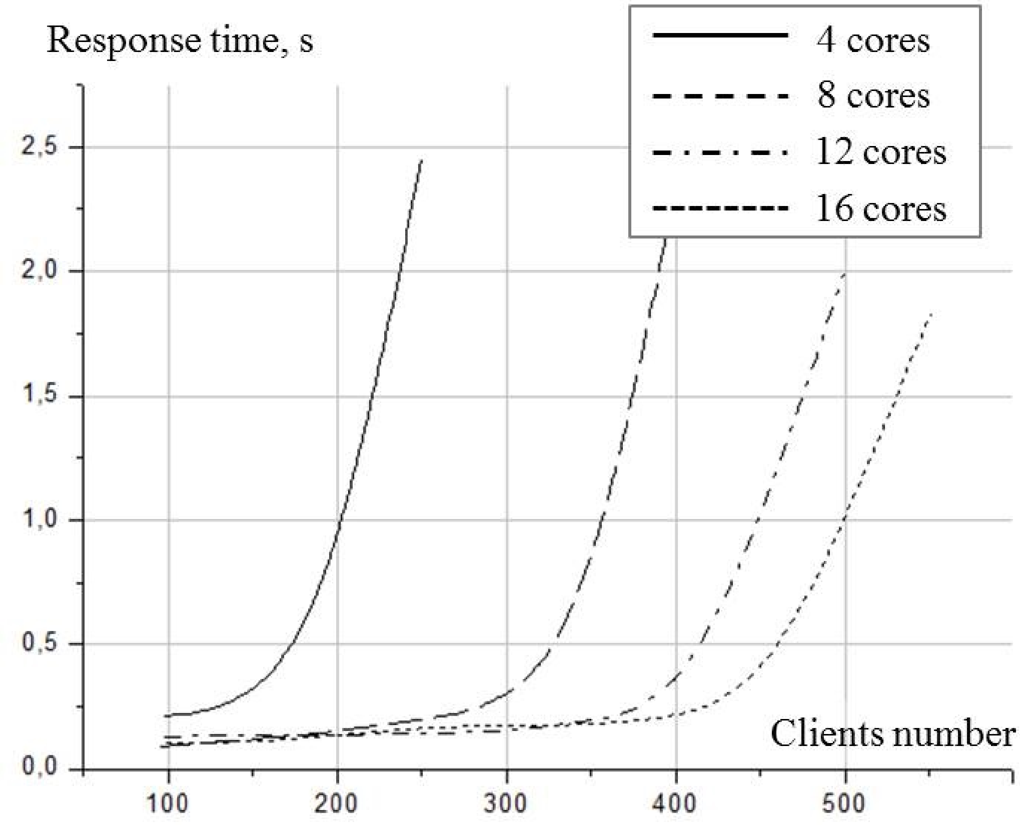

Fig. 2. Response time of cluster with respect to the number of queries for 90% threshold value

Dynamics of the websearch cluster's CPU utilization level is shown at Figure 3. CPU utilization level was traced for 2 VMs with respects to the number of clients' queries. CPU utilizations of both VMs are synchronized with the variation of the number of queries and it could be seen that loads between VMs are not perfectly balanced. Therefore it should be mentioned that resource utilization efficiency has to be improved by sharing cores among multiple VM.

This procedure will provide more flexible use of the core cores up to the resource demands as a real-time function.

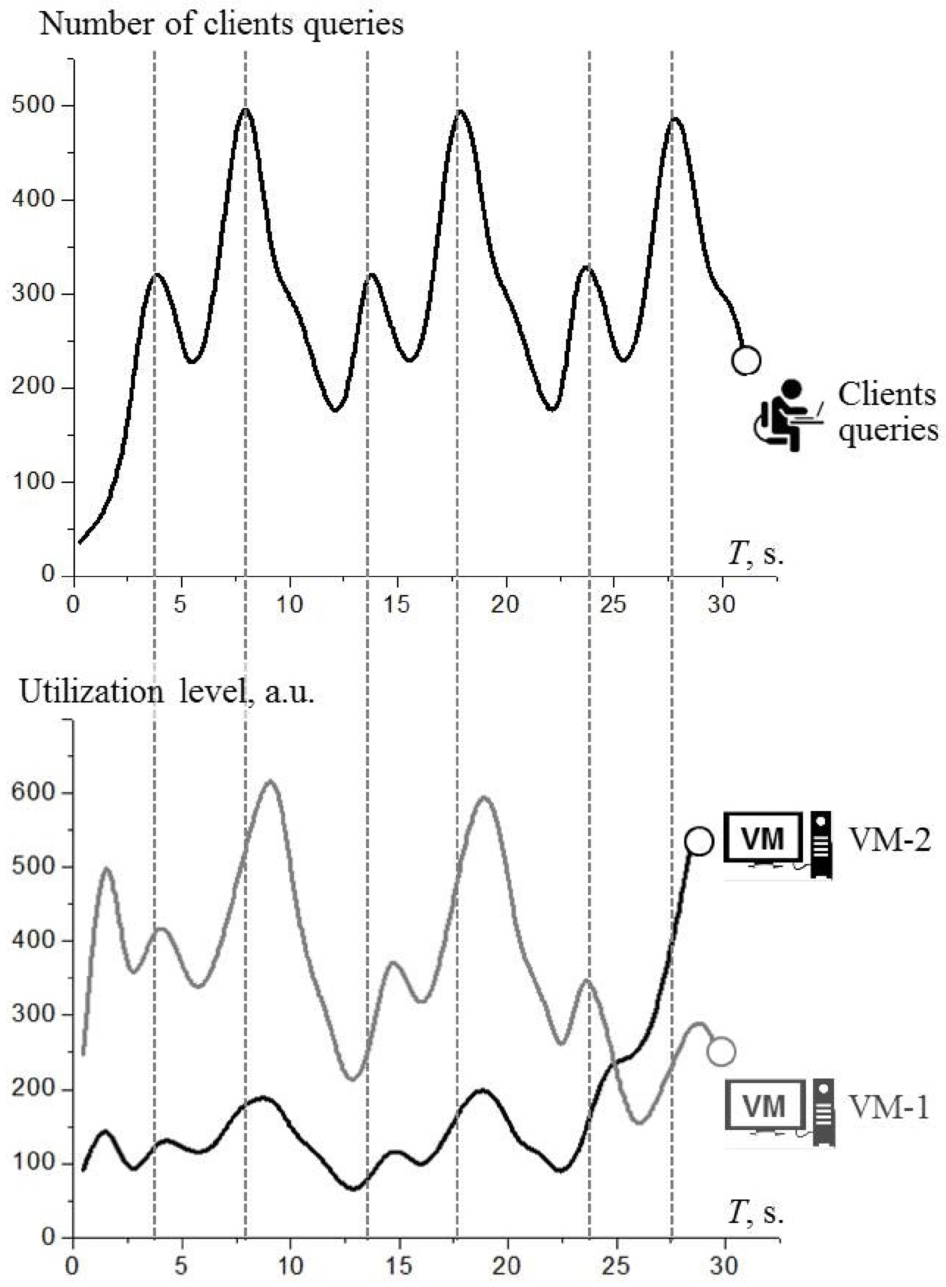

Fig. 3. CPU utilization level of two VMs with respect the number of clients queries

PROPOSED SERVER CONSOLIDATION METHOD

As it was mentioned before clients queries are distributed between multiple VMs and of every cluster and workloads of VM within a same cluster are highly correlated in comparison of correlation of different clusters VMs. In Figures 4-6 is shown intra-cluster correlation of 2 VMs. It is can be seen that VMs resources utilization are strongly synchronized. Proposed method includes analysis of VMs pervasive correlation within a cluster and among clusters. The Figures demonstrates the effectiveness of the correlation-aware VM maintaining of 2 servers which possess 2x8 cores. Servers virtualization produces 4 VMs: VM-1, VM-2, VM-3, and VM-4 where VM-1 is co-located with VM-2 and VM-3 is co-located with VM-4 (Figure 4).

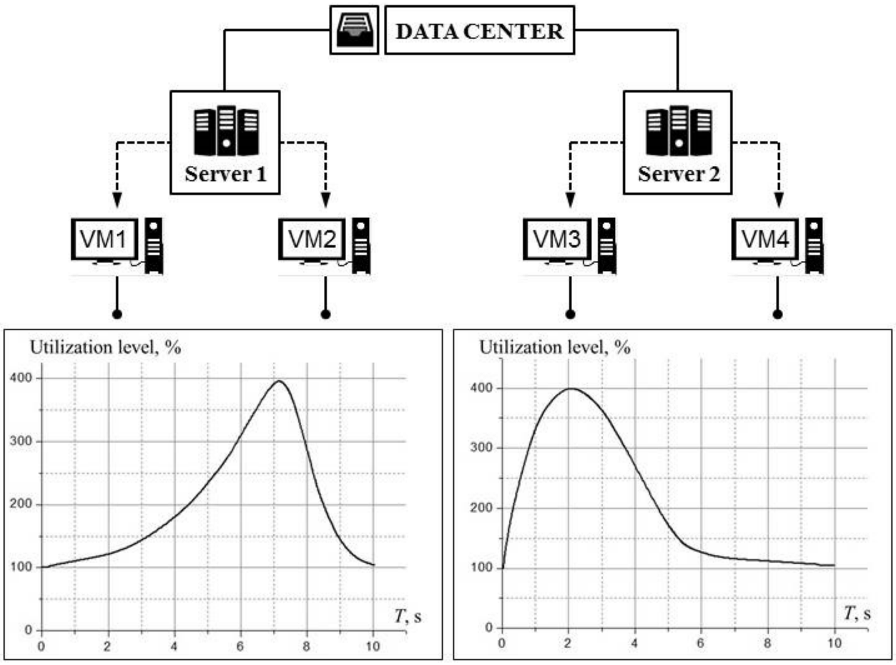

Fig. 4. Data center servers' virtualization procedure simulation

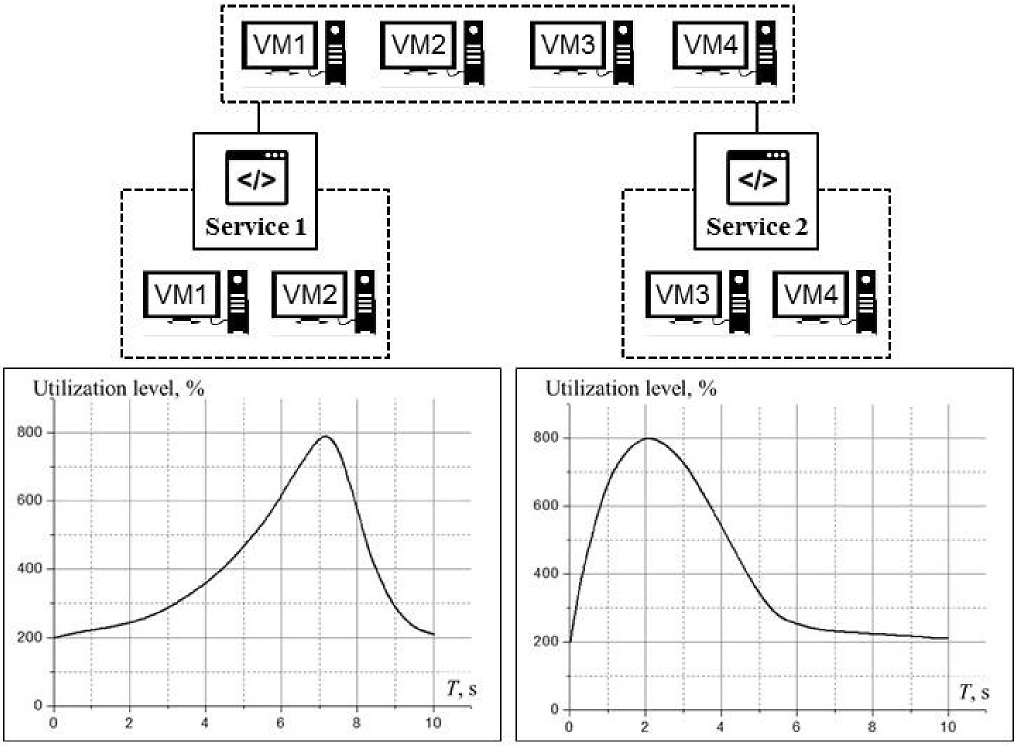

Fig. 5. Drawbacks of VMs allocations without considering correlation value

VMs have the same tail distribution of CPU computational resource utilization and co-located ones are highly correlated. If one will not take into account the correlation (Figure 5), services 1 will allocate sets of VM-1 and VM-2, while service 2 will allocate sets of VM-3 and VM-4. In this case, extremum value of CPU utilization will attain 8x100% of core of each server (active state of all cores). In other hand, if one will pair services [VM-1; VM-3] and [VM-2; VM-

4], extremum value of CPU utilization for each server cores may be lowered down to 6x100% (Figure 6), which allow to lower v/f level without services performance degradation.

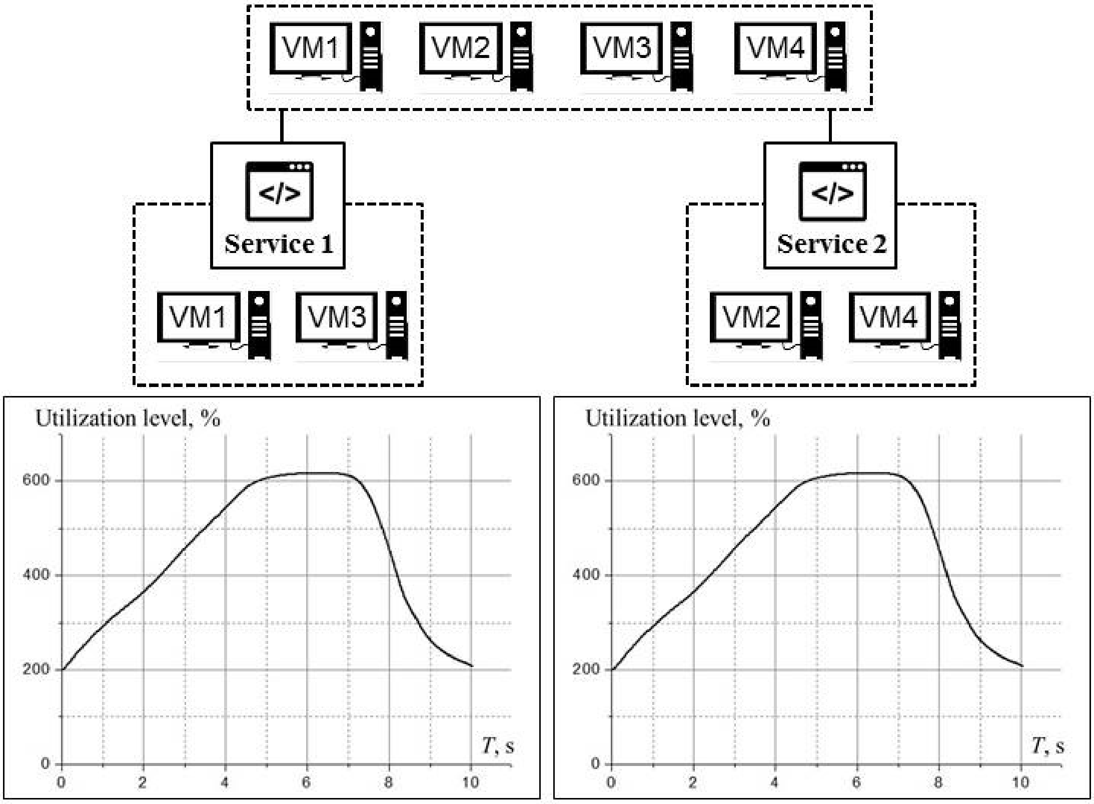

Fig. 6. Correlation-aware VMs allocations procedure

To develop an efficient mathematical model it is proposed to use Pearson's correlation [13] to quantify the correlation coefficient of used data center VMs CPU utilization. It can be calculated as the ratio of covariance of the two variables to the product of their standard deviations. However, Pearson's correlation could be inefficient for the task because this value refers to correlation throughout the corresponding time interval while only correlation at peak or threshold VMs utilization is required. Thereby, it was important to estimate proper measure to quantify the correlation coefficient between VMs that is able to overcome the inefficiency of the conventional correlation metric:

$$C_{i,j} = \left(1 - \frac{U_i^{CPU} + U_j^{CPU}}{U_{i,j}^{CPU}}\right) \cdot 100\%, \quad (1)$$

where $C_{i,j}$ is correlation measure of VM_i and VM_j, U_i^{CPU} is CPU utilization level of VM_i, of VM_j is CPU utilization level of VM_j and $U_{i,j}^{CPU}$ is aggregated actual peak utilization of co-located VM_i and VM_j. $C_{i,j} = 100\%$ refers to complete correlation, while $C_{i,j} = 0\%$ refers to no correlation.

It is important to note that values of each recorded period of utilization have to be updated. Correlation coefficients between all VMs have to be modeled by $C[i,j]$ matrix where each element corresponds to the $C_{i,j}$ measuring function. This model will allow storing all samples and evenly distributing computational utilization as well as correlation between the events in the bounds of certain time period.

CONCLUSIONS

It was shown that low performance of data center infrastructure work refers to disproportion of servers' utilization. Overconsumption problem could be solved by minimization of the active servers' number. In order to provide server consolidation implementation it is necessary to maintain acceptable performance level of the servers room infrastructure work. Server consolidation may cause performance degradation due to the conflict of using shared resources by virtual machines. Basic scheme of correlation-aware virtualization includes: tracing data center servers' physical resources utilization level, transforming of utilization traces into binary sequence up to the utilization threshold value, clustering of virtual machines up to the binary sequence and virtual machines allocation at physical servers. Power management procedure which consists from user-interactive and fast changing service, maintaining of the minimal performance degradation caused by physical resources sharing conflict, high correlation level of virtual machines was developed. Pearson's correlation coefficient was proved to be optimal instrument of the correlation of used data center virtual machines physical resources utilization quantifying. Developed model model allows storing all samples and evenly distributing computational utilization as well as correlation between the events in the bounds of certain time period.

REFERENCES:

1. Harris, M., & Geng, H. (2015). Data center infrastructure management. *Data center handbook, 1st edn. Wiley, Hoboken, NJ*, 601-618.
2. Barroso, L. A., Clidaras, J., & Hölzle, U. (2013). The datacenter as a computer: An introduction to the design of warehouse-scale machines. *Synthesis lectures on computer architecture*, *8*(3), 1-154.
3. Ferdman, M., Adileh, A., Kocberber, O., Volos, S., Alisafaee, M., Jevdjic, D., ... & Falsafi, B. (2012, March). Clearing the clouds: a study of emerging scale-out workloads on modern hardware. In *ACM SIGPLAN Notices* (Vol. 47, No. 4, pp. 37-48). ACM.
4. Harris, M. (2014). Data Center Infrastructure Management. Data Center Handbook, 601-618. doi:10.1002/9781118937563.ch33.
5. Meisner, D., Sadler, C. M., Barroso, L. A., Weber, W. D., & Wenisch, T. F. (2011, June). Power management of online data-intensive services. In *ACM SIGARCH Computer Architecture News*(Vol. 39, No. 3, pp. 319-330). ACM.
6. Bobroff, N., Kochut, A., & Beaty, K. (2007, May). Dynamic placement of virtual machines for managing sla violations. In *Integrated Network Management, 2007. IM'07. 10th IFIP/IEEE International Symposium on* (pp. 119-128). IEEE.
7. Jaramillo, D., Furht, B., & Agarwal, A. (2014). Mobile virtualization technologies. In *Virtualization Techniques for Mobile Systems* (pp. 5-20). Springer, Cham. doi:10.1007/978-3-319-05741-5_2.
8. Tickoo, O., Iyer, R., Illikkal, R., & Newell, D. (2010). Modeling virtual machine performance: challenges and approaches. *ACM SIGMETRICS Performance Evaluation Review*, *37*(3), 55-60.
9. Govindan, S., Liu, J., Kansal, A., & Sivasubramaniam, A. (2011, October). Cuanta: quantifying effects of shared on-chip resource interference for consolidated virtual machines. In *Proceedings of the 2nd ACM Symposium on Cloud Computing* (p. 22). ACM.
10. Wang, L., & Lu, Y. (2010, December). Power-efficient workload distribution for virtualized server clusters. In *High Performance Computing (HiPC), 2010 International Conference on* (pp. 1-10). IEEE. doi: 10.1109/hipc.2010.5713178.
11. Meng, X., Isci, C., Kephart, J., Zhang, L., Bouillet, E., & Pendarakis, D. (2010, June). Efficient resource provisioning in compute clouds via vm multiplexing. In *Proceedings of the 7th international conference on Autonomic computing* (pp. 11-20). ACM.
12. Chen, M., Zhang, H., Su, Y. Y., Wang, X., Jiang, G., & Yoshihira, K. (2011, May). Effective VM sizing in virtualized data centers. In *Integrated Network Management (IM), 2011 IFIP/IEEE International Symposium on* (pp. 594-601). IEEE.
13. Halder, K., Bellur, U., & Kulkarni, P. (2012, June). Risk aware provisioning and resource aggregation based consolidation of virtual machines. In *Cloud Computing (CLOUD), 2012 IEEE 5th International Conference on* (pp. 598-605). IEEE.
14. Santos, J. R., Turner, Y. (2011). Virtual Machine Management. Mastering, 255-326. doi:10.1002/9781118257432.ch7.
15. Kim, J., Ruggiero, M., Atienza, D., & Lederberger, M. (2013, March). Correlation-aware virtual machine allocation for energy-efficient datacenters. In *Proceedings of the Conference on Design, Automation and Test in Europe* (pp. 1345-1350). EDA Consortium.
16. Khan, S. U. (2015). Handbook on data centers. Place of publication not identified: Springer.

Рецензент: д.т.н., проф. Марасанов В. В.
Херсонський національний технічний університет