TECHNISCHE
UNIVERSITÄT
DRESDEN

Fakultät Informatik Institute für Systemarchitektur, Professur Rechnernetze

MODELLING THE LIVE MIGRATION TIME OF VIRTUAL MACHINES

M.C.S. Kateryna Rybina
Born on: 10th July 1986 in Shytomyr

DISSERTATION SUMMARY

to achieve the academic degree

DOKTOR-INGENIEUR (DR.-ING.)

Referee
Prof. Dr. rer. nat. habil. Dr. h. c. Alexander Schill

Submitted on: 4th December 2015

1 INTRODUCTION AND MOTIVATION

The resource usage of the *IT infrastructure* fluctuates over the course of the day, month, and year. Often, cloud service providers over-provision available hardware resources in order to meet the resource demands during infrequent short workload peaks, whereas most of the time the system is underutilised. Current research work [1], [2], [3] revealed that servers in server clusters of Google, Twitter, or Amazon are running underutilised. Reiss et al. [1] analysed the resource usage of Google cluster and showed that most of the time the hourly average CPU utilisation is between 25% and 40%, whereas memory utilisation is about 50%. Delimitrou and Kozyrakis [2] showed that in the production cluster of Twitter the CPU utilisation is below 20% and the memory utilisation is between 40% and 50%. Liu [3] analysed the CPU utilisation of servers of Amazon EC2 and revealed the average values of CPU utilisation below 20%. Another research work has shown that servers are consuming between 50% and 60% of their peak power consumption when running idle or underutilised [4], [5]. Such inefficient usage of resources results in a high energy cost with a negative environmental impact.

All the above mentioned statistics, however, show a huge potential for optimising the way Internet servers are utilised. In periods of low workloads, the services could be consolidated from a bigger set of underutilised servers to a smaller set of optimally loaded servers with the aim to power off idle servers, and thus, save energy [6]. Service consolidation can be realised in a way transparent to the end-user via *live migration* of virtual machines (VMs). During live migration the VM, together with its encapsulated services, is executed interruption free while being moved between the source and the destination servers [7], [8], [9]. The merits of workload consolidation via live VM's migration are manifold:

1. Idle servers can be switched off, after migrating their virtual machines to other servers, thereby saving energy [10], [11];

2. Resource consumption of overloaded servers can be reduced by moving some of the VMs from it, thereby applying load-balancing to the system [12], [13];

3. Servers can be shut down for maintenance, after migrating their VMs, thereby accommodating transparent *IT* maintenance.

However, beside the merits of migration, the process has also drawbacks which will be referred to as *migration costs*. *Migration costs* can be defined as penalties associated with the VM's migration process. These penalties include: 1) Degradation of the quality of services (QoS), executing inside the VMs, 2) increased power consumption of the servers due to migration process, 3) the energy overhead of migration; and 4) migration time itself [10], [11]. The longer it takes to migrate the VM, the stronger is the negative impact of migration on the applications' performance. The quantitative knowledge of VM's migration costs is vital in order to realise sophisticated system adaptation and reconfiguration actions effectively.

2 RESEARCH GOAL AND QUESTIONS

Migration costs are influenced by many parameters. Some of them are application specific, for example, the type of workload executing within the VM, the degree to which this workload utilises the main memory, etc. Other parameters are VM specific, for instance, the memory size of the VM and its CPU utilisation. Another type of parameters is migration specific, for example, network bandwidth available for migration, etc.

The goal of this PhD thesis is to investigate and classify which parameters influence the VM's migration time and the energy consumption of the servers during migration most intensively. This influence will be quantified. VM's migration time and the energy consumption of the servers during migration are the VM's migration costs considered in this thesis. Another goal is to develop mathematical models which best represent the relationships between the selected parameters and the VM's migration costs.

Thus, the research questions, addressed in this PhD thesis, are formulated as follows:

1. Which parameters influence the VM's migration time and energy overhead of migration?

2. Which models best represent the relationships between the parameters mentioned in the first research question, on the one hand, and the migration time and the energy consumption of the servers during migration, on the other hand?

In order to answer the first research question we carried out an extensive set of practical experiments and conducted the following detailed investigations:

1. Experimentally investigate the influence of the following parameters: VM's memory size, available network bandwidth, workload type on the total migration time of the VMs and the energy overhead of migration.

2. Investigate in detail the influence of the VM's size, available network bandwidth, workload type on the power consumption of both source and destination servers.

3. Experimentally investigate the influence of such parameters as last level cache line misses, total number of CPU instructions retired, "dirty" memory pages of the source server, ratio of active memory to the network bandwidth available for migration, and CPU utilisation of the source server on the total migration time of the VMs.

4. Learn the influence of co-located VMs on the total migration time and investigate whether the order in which the multiple VMs are migrated matters or not.

In order to answer the second research question we conducted the following extensive investigations and evaluations:

1. Develop models which estimate the migration time of the VMs and the energy consumption of servers during migration based on the above mentioned parameters.

2. Evaluate developed models which can estimate the live VM's migration time and the energy consumption of the source server during migration.

Moreover, in this PhD thesis we will make an additional investigation on the performance degradation of the applications caused by the VM's migration process. We will show the trade-off between the saved energy and the service degradation due to the workload consolidation.

3 CONTRIBUTIONS OF THE THESIS

The practical implications of the PhD thesis are as follows:

1. *Parameters:* The analysed and classified parameters can be considered by researchers when developing further algorithms for predicting migration time and energy overhead of migration prior to making migration decisions. The parameters which were classified by us as the most significant ones can be used as input to new migration models and those parameters which are insignificant can be directly eliminated, thus saving time and effort.

2. *Models:* The models developed in this PhD thesis which can estimate the VM's migration time and the energy consumption of the servers during migration can be applied in further sophisticated system optimisation decisions. The system can be optimised in order to minimise its energy utilisation or the QoS degradation. The quantitative costs of migration derived with the help of our models could constitute a subset of input parameters for the future optimisation algorithms. The optimisation algorithms may address the following questions: 1) When should the VM's migration take place; 2) which VM has to be migrated; 3) what is the destination server; 4) how frequent can the VM's migration take place [14]. Moreover, the migration costs can be considered by system administrators when accounting for the risks (penalties) associated with the VM's migration.

3. *Techniques:* We are the first who applied multiple linear regression (MLR) techniques to build models for predicting the migration time of the VMs and the energy consumption of the servers during migration. So far, only simple linear regression (SLR) techniques were used to model migration time in related work. With our evaluation of results we show that the models built using MLR techniques outperform those models built using SLR techniques both in terms of prediction accuracy as well as expressive power. Thus, MLR techniques can be used in the future research work in order to model the live VMs' migration time.

4 RELATED WORK

Strunk [11] as well as Xu et al. [10] tried to systematise in their survey work VM's migration costs considered in the literature as well as the parameters that influence them. We extend the taxonomy of Strunk by two additional types of the migration costs considered in the related work: *Power consumption* and *Network traffic* [8], [15] which is created due to the VM's migration process.

Several approaches to modelling VM's migration time were presented in related work. Strunk [16] applied SLR to model migration time of the idle VMs and linear regression with two parameters for modelling energy overhead of migration. The limitation of these models is that they were built for a VM which was executing no workload and it was running in isolation. The average errors of the models which predict the migration time of the idle VM and energy overhead of migration were 3% and 10%, respectively. In this PhD thesis we identify multiple parameters, that significantly influence the migration time of active VMs and the energy consumption of the servers during migration. Based on the selected parameters we build our models using MLR techniques. The VMs which are migrated are continuously executing CPU and memory intensive applications.

Wu et al. [17] examine the relationship between the total migration time and the amount of CPU resources available for migration. The limitation of the approach is that a separate model for predicting migration time has to be built for each type of application and it requires several workload pre runs. We address this issue by building MLR models which can predict migration time for combined (CPU and memory intensive) workloads.

Verma et al. [18] suggested an application aware model for predicting the total VM's migration time. It computes the VM's migration time by refining the calibrated migration time, obtained for this VM during the *calibration* phase at a fixed active memory, by additional time needed to transfer the current active memory used by the VM. In presence of heavy resource contention on the CPU resources the suggested migration model would not be sufficient. Besides, the model requires a workload pre-run for each type of application in the Cloud which does not scale well. The migration models we develop in this thesis can be build for combined workloads.

The models of Liu et al. [8] and Akoush et al. [19] which can predict migration time of VMs do not account for the CPU utilisation of the source server, though the migration time as well as the power utilised by the source server depend on its CPU utilisation [17]. Besides, the VMs are always running in isolation on the source server.

Liu et al. [8] assume that the energy overhead of migration of the source and the destination servers is the same, thus simplify their energy model accordingly. We made investigations of the energy overhead of both the source and the destination servers and found out that even though the servers are homogeneous, the overhead contributed to the energy overhead of migration by each of the servers differs not negligibly. We are going to address this issue in our research work and suggest to build MLR models which predict the energy consumption of each physical machine during migration separately based on its particular resource utilisation parameters.

In contrast to the related work, our models account for a broader spectrum of parameters, namely CPU related parameters such as CPU utilisation, total number of instructions retired, and last level cache line misses; application specific parameters, namely "dirty" memory pages, and combined parameters such as ratio of active memory to the network bandwidth (data transmission rate) available for migration. We show the impact of each individual parameter on the prediction of the migration time, analyse the impact of these parameters in combination and define a set of the most significant parameters. We explicitly point out the benefit of considering multiple parameters in combination.

5 RESULTS

The main contributions of the PhD thesis are as follows:

1. We extended a set of parameters that influence VM's migration costs based on our own investigations and practical experiments. In our work we considered twelve parameters (independent variables) that influence the VM's migration time. These parameters are listed in Table 1. We proved that migration time linearly depends on all of them by calculating the correlation coefficients between the migration time and each of the selected parameters. Moreover, we defined a set of the most significant parameters that influence the VM's migration time. It was realised by determining the relative importance of parameters based on the method of relative weights [20], [21] and based on the analysis of the standardised regression coefficients of the independent variables [20]. The results revealed that such parameters as total number of CPU instruction retired during migration $TotalINST\_Server$, last level cache line misses observed during migration $L3_{miss}$, "dirty" memory pages in the source server during migration $DirtyPages_{server}$ are the first, the second, and the third important independent variables, respectively. These parameters were not considered in related work. That is why we extend the taxonomy of parameters that influence the VM's migration time by including these three parameters.

Table 1: Resource utilisation parameters (independent variables) used in our experiments to determine total migration time t_{mig} of VMs.

Name of variable	Description
BW	Network bandwidth available for migration in MBps.
$L3_{miss}$ [22]	Total number of L3 cache line misses during VM's migration in millions
$INST$ [22]	Total number of instructions retired during migration in thousands
MAR [23]	Ratio of total number of L3 cache line misses to the total number of instructions retired during migration
CPU_{util_server}	Mean total CPU utilisation of the source server during the VM's migration in percentage
CPU_{util_vm}	Mean total CPU utilisation of the VM during the migration process in percentage
MEM_{server}	Mean active memory utilisation of the source server during the VM's migration in MB
MEM_{vm}	Mean active memory utilisation of the VM during the migration process in MB
$MEMtoBW_{server}$	Ratio of active memory utilised by the source server to the network bandwidth available for migration
$MEMtoBW_{vm}$	Ratio of active memory utilised by the VM to the network bandwidth available for migration
$DirtyPages_{server}$	Number of "dirty" pages observed in the source server during the migration process
$DirtyPages_{server\_per\_sec}$	Number of "dirty" pages observed in the source server per second during the migration process

2. We developed models which can predict the VM's migration time and the energy consumption of the servers during migration. These models were built using SLR and MLR techniques. We showed that MLR models outperform SLR models both in terms of prediction accuracy as well as expressive power. The summary of results of the best SLR and MLR models which can predict: 1) Migration time of VMs running combined workloads; 2) migration time of VMs running CPU intensive workloads; and 3) energy consumption of the source server during migration is given in Table 2, Table 3, and Table 4, respectively. Benchmarks from SPEC CPU2006[1] benchmark suite were used as workload in these experiments. The results pertaining to the selection of parameters that influence migration time of VMs as well as modelling migration time based on these parameters were published in the proceedings of the International Symposium on Secure Virtual Infrastructures

[1] https://www.spec.org/cpu2006/

(Cloud and Trusted Computing 2015) [24].

We defined the best MLR model with respect to the adjusted R-square metric R^2_{Adj} by applying method of all subsets regression [20]. The closer R^2_{Adj} is to 1 the better the model fits the measured data. The best MLR model which can determine the live migration time of the VMs running combined (CPU and memory intensive) workloads includes five parameters. These are: $MEMtoBW_{server}$, $L3_{miss}$, $INST$, CPU_{util_server}, and $DirtyPages_{server}$. Its R^2_{adj} equals to 0.946. The mean absolute percentage error of the model $Perc_m$ is equal to 10.14%. Its residual standard error on the 654 degrees of freedom $Res_{st.err}$ is 7.8 seconds. Model's standard error of estimate on the test data $Pred_{st.err}$ is 5.4 seconds. The acceptable error rates show that the best MLR model can be applied in order to estimate the migration time of VMs which execute combined CPU and memory intensive workloads (see Table 2).

Table 2: The best SLR and MLR models for predicting live migration time of VMs running combined (CPU and memory intensive) workloads.

The best SLR and MLR models: $lm\,(t_{mig} \sim Predictors)$	R^2	R^2_{Adj}	$Res_{st.err}$	$Pred_{st.err}$	$Perc_m$
$INST$	0.7405	0.7401	17.05	15.35	30.79
$INST + L3_{miss}$	**0.9251**	**0.9248**	**9.171**	**8.02**	**15.18**
$L3_{miss} + INST + DirtyPages_{server}$	0.936	0.9356	8.487	7.028	13.39
$L3_{miss} + INST + CPU_{util\_server} + DirtyPages_{server}$	0.943	0.943	7.984	5.811	12.12
$L3_{miss} + INST + DirtyPages_{server} + CPU_{util\_server} + MEMtoBW_{server}$	**0.946**	**0.9456**	**7.802**	**5.379**	**10.14**

Table 3: The best SLR and MLR models with one, two, three, four, and five independent variables for predicting live migration time of VMs running CPU intensive workloads.

The best SLR and MLR models: $lm\,(\sqrt{t_{mig}} \sim Predictors)$	R^2	R^2_{Adj}	$Res_{st.err}$	$Pred_{st.err}$	$Perc_m$
$MEMtoBW_{server}$	0.839	0.838	0.23	0.16	1.7
$MEMtoBW_{server} + INST$	0.857	0.856	0.21	0.18	2.22
$MEMtoBW_{server} + L3_{miss} + INST$	0.929	0.928	0.15	0.197	1.87
$MEMtoBW_{server} + L3_{miss} + INST + CPU_{util\_server}$	**0.952**	**0.951**	**0.124**	**0.12**	**1.55**
$MEMtoBW_{server} + L3_{miss} + INST + CPU_{util\_server} + DirtyPages_{server}$	0.952	0.950	0.124	0.121	1.56

Moreover, we defined which parameters out of the five used in the model are the most important by calculating their relative weights [20], [21] in contributing to the model's R-square. $INST$, $L3_{miss}$, and $DirtyPages_{server}$ are the most important parameters as they contribute 37.8%, 31.4%, and 27.6% to the total model's R-square of 0.946, respectively.

This model is generalisable. We tested it by carrying out the 10-fold cross-validation of the R^2 metric. The results are as follows: original R-square = 0.946, 10-fold cross-validated R-square = 0.937. The difference is very small (0.009), and thus, the model is performing well on the test data. The model satisfies *linearity*, and *normality* assumptions. But due to the high variances in migration time introduced mainly by *mcf* and *astar* memory intensive benchmarks from SPEC CPU2006 benchmark suite the model faces the problem of non-constant variance which is often the case in practice. Nevertheless, the non-constant variance is not substantial in this case because the model's prediction error is still acceptable ($Perc_m$ = 10.14%). In order to satisfy additionally constant variance assumption we designed models specifically for CPU intensive workloads, as those do not reveal such high variances in migration time.

The best MLR model which can predict the VM's migration time (for VMs running CPU benchmarks) includes four independent variables (see Table 3). These are: $MEMtoBW_{server}$, $L3_{miss}$, $INST$, CPU_{util_server}. Its R^2_{adj} is equal to 0.951 and it has the highest prediction accuracy. The mean absolute percentage error of the model is 1.55% only and its residual and prediction errors are low (0.124 seconds and 0.12 seconds, respectively). Moreover, we calculated the relative importance of the parameters in contributing to the model's R^2. $MEMtoBW_{server}$, $INST$, and $L3_{miss}$ are the first, the second, and the third important variables as they contribute 48%, 34.9%, and 15.7% to the model's R-square of 0.952, respectively. The independent variables applied in all models are significant with p-value lower than the smallest significance level 0.001. This model is generalisable and it satisfies all main assumptions of the regression analysis, namely: *linearity*, *normality*, and *constant-variance*.

Last but not least, we trained and tested MLR models which can estimate the energy consumption of the source server during the VM migration process. The best MLR model is built upon four parameters: $MEMtoBW_{server}$, $L3_{miss}$, $INST$, and CPU_{util_server}. Its R^2_{adj} equals 0.94 and its mean absolute percentage error is equal to 4.6% (see Table 4). The four independent variables applied in this model are significant with p-value lower than the smallest significance level 0.001. This model is generalisable and satisfies *normality*, *linearity*, and *constant-variance* assumptions. The multicollinearity among the independent variables is not present in all models.

Table 4: The best SLR and MLR models for predicting energy consumption of the source server during migration of VMs running CPU workloads.

The best SLR and MLR models: $lm\,(E_{s-during} \sim Predictors)$	R^2	R^2_{Adj}	$Res_{st.err}$	$Pred_{st.err}$	$Perc_m$
$MEMtoBW_{server}$	0.766	0.765	0.059	0.039	5.44
$L3_{miss} + INST$	0.891	0.89	0.04	0.052	6.76
$MEMtoBW_{server} + L3_{miss} + INST$	0.927	0.9268	0.032	0.038	5.42
$MEMtoBW_{server} + L3_{miss} + INST + CPU_{util\_server}$	**0.936**	**0.9353**	**0.031**	**0.03**	**4.6**
$MEMtoBW_{server} + L3_{miss} + INST + CPU_{util\_server} + DirtyPages_{server}$	0.936	0.9352	0.031	0.03	4.7

One of the *limitations* of the developed models is that the models' parameters (coefficients) have to be retrained for each hardware platform. Another limitation is that the models do not account for the co-located VMs. In our experiments the VM was migrated in isolation. This allowed us to control the server-level as well as the VM-level parameters. The derived MLR models can be applied for different data transmission rates, though, due to constraints of our physical infrastructure, we realised the VM's migrations at maximum data transmission rate of 1 Gbps. Thus, conducting further migration experiments at higher data transmission rates would be also of high practical interest.

3. We investigated the energy overhead of the VM's migration process and parameters that influence it. The VM's migration process creates a not negligible energy overhead which exists regardless of the type of workload the VM was hosting. The energy overhead of migration increases with an increment in the RAM utilised by the VMs. The variation in network bandwidth does not make a considerable impact on the energy overhead. The overhead contributed to the energy overhead by the destination server was higher than the overhead contributed by the source server. Moreover, we found out that migration time and the energy overhead of migration were significantly high when the source server was overloaded. The results pertaining to these investigations were published in the following conferences: 1) The third IFIP Conference on Sustainable Internet and ICT for Sustainability [25], 2) the 23rd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing [26].

4. The summed migration time of multiple VMs is proportional to the summed amount of memory utilised by these VMs. The order of sequential migration of the VMs does not impact the summed migration time as long as the VMs do not compete for common resources (CPU caches, memory bus, and network bandwidth) and there exist enough free CPU resources to accommodate migration. When multiple VMs with resource intensive benchmarks are executed on the source server, interference effects occur, which significantly influence the migration time of these VMs. The results of these investigations were published in the 4th International Conference on Cloud Computing and Services Science [27].

5. We investigated the influence of the VM's migration on the transcoding time of three of the shelf transcoders: FFMPEG, MENCODER, and HANDBRAKE. The transcoding time of these transcoding applications increased during migration in both considered scenarios, namely when both servers were underutilised and when the source server was overloaded. But this increment should be understood in context. If the VM is migrated from the overloaded server to the server that has plenty of available resources (e.g. in order to realise load-balancing) the additional time added due to migration will be compensated by sufficient resources on the new server. Thus, this shows a trade-off between the energy consumption and the quality of services (in this example a transcoding time). Namely, to decrease the transcoding time we needed to turn on a new server and migrate the VM which run a transcoding application to it [26].

Moreover, we analysed the impact of frequent migrations of VMs on the quality of services running within the VMs. We first ran the benchmarks from SPEC CPU2006 benchmark suite until their completion within the VM without migration and measured their execution times. Then, we repeated the same experiment but during the benchmarks' normal execution the VM was 20 times migrated between two servers. The execution time of all benchmarks significantly increased due to frequent VM's migrations, whereas the execution time of memory-intensive benchmark *astar* even doubled.

6 SUMMARY AND OUTLOOK

We carried out extensive investigations, practical experiments, and detailed evaluations in order to answer the research questions defined in this PhD thesis. We analysed parameters that influence the total migration time of the VMs as well as energy overhead of migration. Moreover, we identified the most significant parameters based on which we built models which can predict VM's migration time and the energy consumption of the servers during migration.

The main contributions of the thesis are as follows: Firstly, based on our experimental results and evaluations we extended a set of parameters which can be used to model VM's migration time. We added the following parameters which were not considered in related work: $TotalINST\_Server$, $L3_{miss}$, and $DirtyPages_{server}$. Secondly, we identified the most significant parameters which can be used to model the migration time. These are: $TotalINST\_Server$, $L3_{miss}$, $MEMtoBW_{server}$, and $DirtyPages_{server}$. Thirdly, we are the first who applied MLR techniques in order to build models which can determine the live migration time of active VMs and the energy consumption of the servers during migration. The evaluations of the models showed that they are generalisable and can accurately estimate the VM's migration time and the energy consumption of the servers during migration. The mean absolute percentage errors on test data $Perc_m$ of the models which can predict VM's migration time are 10.14% and 1.55% for VMs running combined workloads (memory and CPU intensive) and CPU intensive workloads, respectively. $Perc_m$ of the best MLR model which can predict the energy consumption of the source server during migration is 4.6%.

These models can be used by other researches in order to realise system optimisation in effective way. Different optimisation criteria may be considered. For example, the system may be optimised to operate with minimal energy cost. Or it can be optimised for maximising the level of quality of provided services. The benefit of the models is that they can provide a quantitative energy cost of migration as well as define the time it takes to migrate a VM. These costs can be used as the input parameters of the sophisticated server consolidation algorithms. During migration the quality of provided services might be degraded, thus the VM migration candidate should be a VM which is migrated faster. Likewise, if the system is optimised for energy, the VM migration candidate should be the VM, which will create the lowest energy overhead. Upon reaching one of the critical conditions (system overload or underutilisation) such sophisticated algorithms may address the following aspects [14]:

1. Select the source servers;
2. Select the candidate VMs that have to be migrated;
3. Determine the destination servers.

Thus, the energy consumption of the servers during migration derived and quantified with our models may guide the choice of selection of the source servers and the destination servers. The models that predict the migration time of the VMs can directly influence the choice of the migration candidates.

BIBLIOGRAPHY

[1] C. Reiss, A. Tumanov, G. R. Ganger, R. H. Katz, and M. A. Kozuch, "Heterogeneity and dynamicity of clouds at scale: Google trace analysis," in *Proceedings of the Third ACM Symposium on Cloud Computing*, ser. SoCC '12. New York, NY, USA: ACM, 2012, pp. 7:1–7:13. [Online]. Available: http://doi.acm.org/10.1145/2391229.2391236

[2] C. Delimitrou and C. Kozyrakis, "Quasar: Resource-efficient and qos-aware cluster management," in *Proceedings of the 19th International Conference on Architectural Support for Programming Languages and Operating Systems*, ser. ASPLOS '14. New York, NY, USA: ACM, 2014, pp. 127–144. [Online]. Available: http://doi.acm.org/10.1145/2541940.2541941

[3] L. Human, "Host server cpu utilization in amazon ec2 cloud," 2012. [Online]. Available: http://huanliu.wordpress.com/2012/02/17/host-server-cpu-utilization-in-amazon-ec2-cloud/

[4] X. Fan, W.-D. Weber, and L. A. Barroso, "Power provisioning for a warehouse-sized computer," in *Proceedings of the 34th Annual International Symposium on Computer Architecture*, ser. ISCA '07. New York, NY, USA: ACM, 2007, pp. 13–23. [Online]. Available: http://doi.acm.org/10.1145/1250662.1250665

[5] L. Barroso and U. Holzle, "The case for energy-proportional computing," *Computer*, vol. 40, no. 12, pp. 33–37, Dec 2007.

[6] A. Verma, P. Ahuja, and A. Neogi, "pmapper: Power and migration cost aware application placement in virtualized systems," in *Proceedings of the 9th ACM/IFIP/USENIX International Conference on Middleware*, ser. Middleware '08. New York, NY, USA: Springer-Verlag New York, Inc., 2008, pp. 243–264. [Online]. Available: http://dl.acm.org/citation.cfm?id=1496950.1496966

[7] W. Voorsluys, J. Broberg, S. Venugopal, and R. Buyya, "Cost of virtual machine live migration in clouds: A performance evaluation," in *Proceedings of the 1st International Conference on Cloud Computing*, ser. CloudCom '09. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 254–265.

[8] H. Liu, C.-Z. Xu, H. Jin, J. Gong, and X. Liao, "Performance and energy modeling for live migration of virtual machines," in *Proceedings of the 20th international symposium on High performance distributed computing*, ser. HPDC

'11. New York, NY, USA: ACM, 2011, pp. 171–182. [Online]. Available: http://doi.acm.org/10.1145/1996130.1996154

[9] C. Clark, K. Fraser, S. Hand, J. G. Hansen, E. Jul, C. Limpach, I. Pratt, and A. Warfield, "Live migration of virtual machines," in *Proceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation - Volume 2*, ser. NSDI'05. Berkeley, CA, USA: USENIX Association, 2005, pp. 273–286. [Online]. Available: http://dl.acm.org/citation.cfm?id=1251203.1251223

[10] F. Xu, F. Liu, H. Jin, and A. Vasilakos, "Managing performance overhead of virtual machines in cloud computing: A survey, state of the art, and future directions," *Proceedings of the IEEE*, vol. 102, no. 1, pp. 11–31, Jan 2014.

[11] A. Strunk, "Costs of virtual machine live migration: A survey," in *Services (SERVICES), 2012 IEEE Eighth World Congress on*, june 2012, pp. 323 –329.

[12] T. Wood, P. Shenoy, A. Venkataramani, and M. Yousif, "Black-box and gray-box strategies for virtual machine migration," in *Proceedings of the 4th USENIX conference on Networked systems design anbd implementation*, ser. NSDI'07. Berkeley, CA, USA: USENIX Association, 2007, pp. 17–17. [Online]. Available: http://dl.acm.org/citation.cfm?id=1973430.1973447

[13] B. Gerofi, H. Fujita, and Y. Ishikawa, "An efficient process live migration mechanism for load balanced distributed virtual environments," in *Cluster Computing (CLUSTER), 2010 IEEE International Conference on*, sept. 2010, pp. 197 –206.

[14] A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing," *Future Gener. Comput. Syst.*, vol. 28, no. 5, pp. 755–768, May 2012. [Online]. Available: http://dx.doi.org/10.1016/j.future.2011.04.017

[15] U. Deshpande, X. Wang, and K. Gopalan, "Live gang migration of virtual machines," in *Proceedings of the 20th International Symposium on High Performance Distributed Computing*, ser. HPDC '11. New York, NY, USA: ACM, 2011, pp. 135–146. [Online]. Available: http://doi.acm.org/10.1145/1996130.1996151

[16] A. Strunk, "A lightweight model for estimating energy cost of live migration of virtual machines," in *2013 IEEE Sixth International Conference on Cloud Computing, Santa Clara, CA, USA, June 28 - July 3, 2013*, 2013, pp. 510–517. [Online]. Available: http://dx.doi.org/10.1109/CLOUD.2013.17

[17] Y. Wu and M. Zhao, "Performance modeling of virtual machine live migration," in *Cloud Computing (CLOUD), 2011 IEEE International Conference on*, july 2011, pp. 492 –499.

[18] A. Verma, G. Kumar, R. Koller, and A. Sen, "Cosmig: Modeling the impact of reconfiguration in a cloud," in *Modeling, Analysis Simulation of Computer and Telecommunication Systems (MASCOTS), 2011 IEEE 19th International Symposium on*, July 2011, pp. 3–11.

[19] S. Akoush, R. Sohan, A. Rice, A. Moore, and A. Hopper, "Predicting the performance of virtual machine migration," in *Modeling, Analysis Simulation of Computer and*

Telecommunication Systems (MASCOTS), 2010 IEEE International Symposium on, aug. 2010, pp. 37–46.

[20] R. I. Kabacoff, "R in action: Data analysis and graphics with r," 2011.

[21] J. W. Johnson, "A heuristic method for estimating the relative weight of predictor variables in multiple regression." *Multivariate Behavioral Research*, vol. 35, no. 1, pp. 1–19, 2000. [Online]. Available: http://dx.doi.org/10.1207/S15327906MBR3501_1

[22] T. W. (Intel), "Intel pcm column names decoder ring," *available at: https://software.intel.com/en-us/blogs/2014/07/18/intel-pcm-column-names-decoder-ring*, 2014.

[23] C. Poellabauer, L. Singleton, and K. Schwan, "Feedback-based dynamic voltage and frequency scaling for memory-bound real-time applications," in *Real Time and Embedded Technology and Applications Symposium, 2005. RTAS 2005. 11th IEEE*, March 2005, pp. 234–243.

[24] K. Rybina, W. Dargie, S. Umashankar, and A. Schill, "Modelling the live migration time of virtual machines," in *Cloud and Trusted Computing 2015, International Symposium on Secure Virtual Infrustructures, Rhodes, Greece, October 26-30*, 2015.

[25] K. Rybina, W. Dargie, A. Strunk, and A. Schill, "Investigation into the energy cost of live migration of virtual machines," in *Sustainable Internet and ICT for Sustainability, SustainIT 2013, Palermo, Italy, 30-31 October, 2013, Sponsored by the IFIP TC6 WG 6.3 "Performance of Communication Systems"*, 2013, pp. 1–8. [Online]. Available: http://dx.doi.org/10.1109/SustainIT.2013.6685192

[26] K. Rybina, W. Dargie, R. Schöne, and S. Malakuti, "Mutual influence of application- and platform-level adaptations on energy-efficient computing," in *IEEE Computer Society; 23rd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing, Turku, Finland, March 4-6*, 2015.

[27] K. Rybina, A. Patni, and A. Schill, "Analysing the migration time of live migration of multiple virtual machines," in *CLOSER 2014 - Proceedings of the 4th International Conference on Cloud Computing and Services Science, Barcelona, Spain, April 3-5, 2014.*, 2014, pp. 590–597. [Online]. Available: http://dx.doi.org/10.5220/0004951605900597