UDC 681.3.093:044.3

Dos E.V.
DevOps EPAM

Zuev D.O.
International IT Expert, Independent Consultant

Kropachev A.V.
Bell Integrator

Babkin O.V.
Bell Integrator

MODELING AND MONITORING OF INFORMATION SYSTEM INFRASTRUCTURE

Development of power and temperature monitoring system is a key aspect of ensuring of data centers performance. It is important to build accurate model of the scalable server room power consumption system and cost-effective cooling facility. Systematic analysis demonstrates that the servers' power consumption is always correlated with key workload parameters of shared storage, memory, computational capability and network bandwidth. It was considered that for management of temporal and spatial temperature variations of servers' stability it is necessary to develop accurate temperature map modeling algorithm. It was shown that computational fluid dynamics simulation is most effective instrument of analysis while it uses mathematical methods for development of precise fluid flow model. Though, it was proved to be a very complex model because it based on differential equation with no analytical solution so resource-intensive numerical procedures have to be used in this case. It was proposed to use algorithm which allows decreasing complexity of computational fluid dynamics simulation and building accurate temperature map.

Key words: *data center, temperature map, power consumption, cooling facility, servers' room, computational fluid dynamics simulation, RC network.*

Introduction. Modern scalable data centers platforms performance is one of the most important task of IT-area development. Building of power and temperature facilities models isproved to be effective instrument of data centers servers' room stability ensuring. Assigned task could be solved by development ofmathematical model of server room power consumption system.

To identify the main aspects of the problem, systematic analysis of recent studies and publications was done. There were analyzed aspects high-level power data center servers' models to estimate key workload parameters [1–3]. To solve problem of electrical cooling complex organization based of fan's system works [4; 14] which demonstrate that the system sets significant amount of data center infrastructure power utilization were studied. Computational fluid dynamics simulation as effective instrument of development of servers' thermal map [5; 14] was analyzed; as well as methods which allows decreasing complexity of this simulation [6–8]. Comparative analysis of cooling power as varying processor utilization process which leads to adjusting the server room temperature change [10–14] was also considered. Systematic analysis shows possibility to develop effective model based on heat recirculation scheme of datacenter platform.

Data center power system modeling. Development of efficient power and temperature monitoring system is a key aspect of ensuring of datacenters performance. It is necessary to build accurate model of the data center server room power consumption and cooling facility and thenwork on scalable and cost-effective power with temperature monitoring systems.

Most accurate power models usuallysimulate and analyze individual components of servers, but for large-scale data centers these algorithmswould be resource-intensive and speed of such a simulation proves to be low enough. Thereby our goal is to simulate the large clusters of servers in datacenters' infrastructure network (Figure 1).

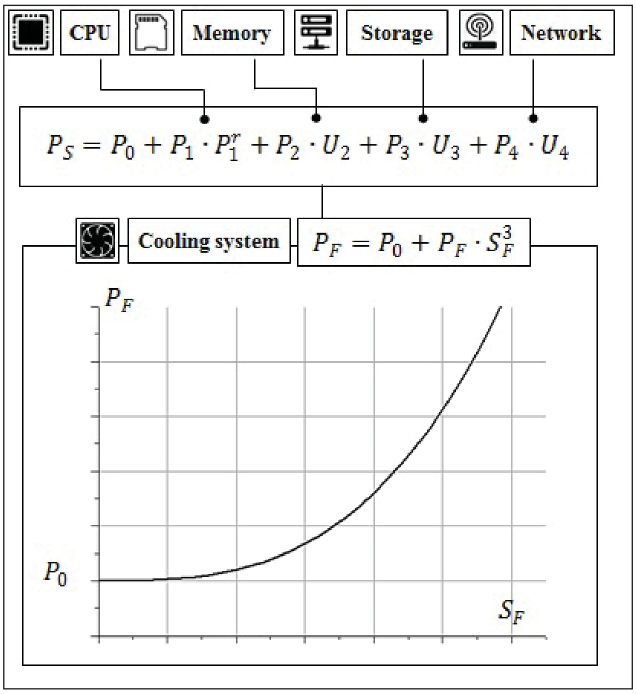

Fig. 1. Data center power system modeling scheme

It was demonstrated [1–3], that power models are widely used to monitor and estimate the power consumption of servers, analysis shows that the power consumption for a given server is always correlated with key workload parameters: shared storage, memory (RAM and cash-memory), computational capability (CPU) and network bandwidth (Figure 1).

To estimate this connection various experimental studies of high-level power data center servers' models were to be done [1–3]. Basically model should use simplified simulation equation of linear or nonlinear regression power model which estimates the server power consumption up to service resource occupancy level:

$$\begin{cases} P_S = P_0 + \sum_i P_i \cdot U_i^{r_i} \\ r_i \geq 1; U_i \in [0..100\%] \end{cases} \quad (1),$$

where P_S is server power consumption, U_i is physical resources utilization level and P_i is a set of fitting parameters, which varies according to the physical resource's type of analyzed data center server system.

Evaluations for developing the high-level server power model could be conducted by comparing different forms of power models which refers to different values of i and r_i. Most simplified model one could set $i = 1$ and analyzes only computational capability of data center's servers ($r_i = 1$ is used for linear model and $r_i > 1$ is for nonlinear one). For accurate simulation it is better to set $i = 1$ and analyze all servers' physical resources occupancy (CPU, RAM and storage workload intensity, as well as network bandwidth).

Electrical cooling complex based of fan's systemstands significant amount of data center infrastructure power utilization. Fan power consumption has a cubic relationship with fan speed [4], as follows:

$$P_F = P_0 + P_F \cdot S_F^3 \quad (2),$$

where P_0 and P_F are fitting parameters and S_F is a fan speed. Thus, lowering of the fan speed lets us to significantly reduce power consumption (Figure 1).

Data center temperature control system modeling. To manage temporal and spatial temperature variations stability it is necessary to develop accurate temperature model. It allows to significantly save expenses on placing of thermal sensors a high area data center server room and prevent problems caused by its' frequent failures. Computational fluid dynamics (CFD) simulation is proved to be effective instrument of development of servers' thermal map. It uses mathematical methods and algorithms for precise analysis of fluid flow model. CFD-based thermal modeling [5] is based following equation:

$$\frac{\partial(\rho\varphi)}{\partial t} + \frac{\partial(\bar{V}\rho\varphi)}{\partial x \partial y \partial z} = \frac{\partial\left(\frac{D\cdot\partial\varphi}{\partial x \partial y \partial z}\right)}{\partial x \partial y \partial z} + S(\varphi), \quad (3)$$

where ρ is a air fluid density, x, y, z are coordinates, $\bar{V}$ is velocity for each of x, y, z direction, S is the source for each φ variable and φ is a variable that can be used for following properties:

– mass;
– velocity;
– temperature;
– turbulence.

D is the diffusion coefficient which could be estimated as

$$D = \frac{A \cdot T^{\frac{3}{2}} \sqrt{\frac{1}{M_1} + \frac{1}{M_2}}}{\Omega \cdot \rho_A \sigma} \quad (4),$$

where A is coefficient, M_N are molar masses of molecules in the gaseous mixture, T is the absolute temperature, ρ_A is the pressure, σ is the average collision diameter, Ω is a temperature-dependent collision integral.

It should be noticed that four components in Eq. (3) refers to main parts of air fluid transport process model:

– transient: $\frac{\partial(\rho\varphi)}{\partial t}$;

– convection: $\frac{\partial(\bar{V}\rho\varphi)}{\partial x \partial y \partial z}$;

– diffusion: $\frac{\partial(\frac{D\cdot\partial\varphi}{\partial x\partial y\partial z})}{\partial x\partial y\partial z}$;

– source: $S(\varphi)$.

CFD-simulation shows high accuracy, but this kind of simulation is a very complex one because there is noanalytical solution for differential equation so it has to be solved by numerical procedures which prove to be resource-intensive.

At this study is presented solution based on the works [6; 8; 14] which allows to decrease complexity of CFD-simulation and build accurate temperature map. The algorithm is based on building of heat- and air-flow graphs. Simplified temperature model for servers is oriented on the CPU- and RAM-blocks of servers, as well as on heat removal capability referring to the fan speed changes. Model includes building of thermal RCnetwork scheme of the system (Figure 2) based on connection between thermal and electrical losses [7].

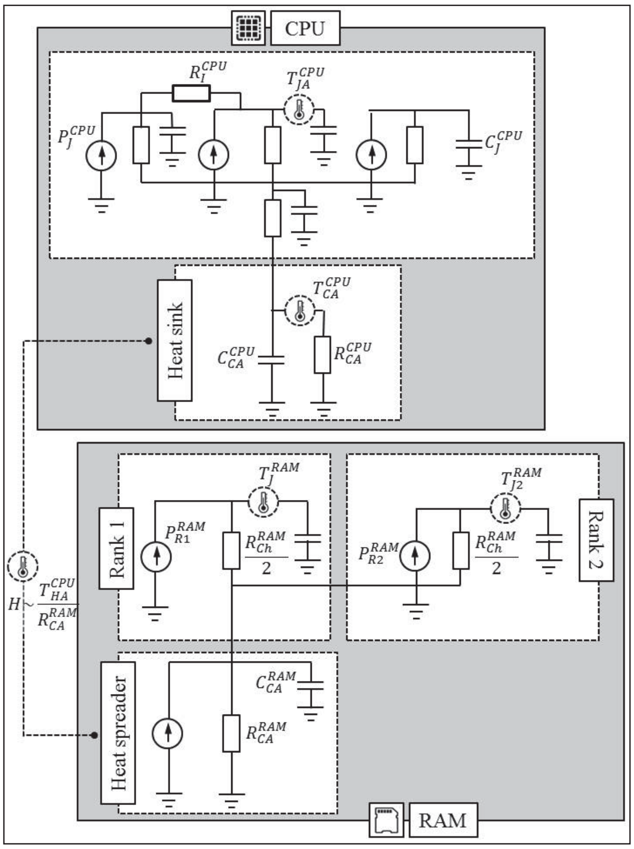

Fig. 2. RC network based temperature model which includes CPU and memory sockets

Figure 2 demonstrates that CPU socket RC network scheme includes:

– power consumption of each core in a socket P_J^{CPU};

– lateral thermal resistance R_l^{CPU};

– vertical thermal resistance R_v^{CPU};

– thermal resistance of heat spreader R_S^{CPU};

– case-to-ambient thermal resistance of heat sink R_{CA}^{CPU};

– thermal capacitances of die C_J^{CPU};

– thermal capacitances of heat spreader C_S^{CPU};

– thermal capacitances of heat sink C_{CA}^{CPU};

– junction temperature T_{JA}^{CPU}.

It has to be noticed that R_v^{CPU} is usually neglected while $R_v^{CPU} \ll R_l^{CPU}$ and R_{CA}^{CPU} could be obtained as a sum of the thermal resistances of heat sink R_{HS}^{CPU} and convective resistance R_{Conv}^{CPU} as function of the fan speed S_F

$$\begin{cases} R_{CA}^{CPU} = R_{HS}^{CPU} + R_{Conv}^{CPU}(S_F) \\ R_{Conv}^{CPU} \sim (EA \cdot S_F^{\alpha})^{-1} \\ \alpha \in [80..100\%] \end{cases} \quad (5),$$

where R_{Conv}^{CPU} estimation is based on parameters of effective area EA and factor α.

In other hand, memory socket RC network scheme also includes further components and definitions (Figure 2):

– power consumption of each RAM chip P_{Ch}^{RAM};

– thermal resistance of each RAM chip R_{Ch}^{RAM};

– thermal capacitance of each RAM chip C_{Ch}^{RAM};

– junction temperature of each RAM chip T_{Ch}^{RAM};

– thermal resistance of the case to ambient of the memory R_{CA}^{RAM};

– number of ranks of each RAM chip N.

Temperature of memory socket is correlated with the temperature of CPU socket due to air flows inside a server. Thereby, air absorbing heat in CPU socket affects to the temperature of RAM socket as it is equivalent to raising temperature at memory socket. Thermal coupling should be modeled as follows:

$$H \sim \frac{T_{HA}^{CPU}}{R_{CA}^{RAM}} \quad (6),$$

where H is the dependent coupling heat source of the memory; T_{HA}^{CPU} is CPU heat sink temperature, of the CPU.

Data center computing facility and cooling facility modeling. For precise estimation of data center servers temperature map, it is necessary to analyze account interactions of multiple servers' heat and hot air flows from bottom to top of the servers room. This procedure allows to develop heat recirculation scheme of datacenter. The model of recirculation can be built by a cross-interference matrix represented by

$$\varphi_{N\times N}:\begin{bmatrix}\varphi_{1-1} & \cdots & \varphi_{i-1}\\ \cdots & \varphi_{i-j} & \cdots\\ \varphi_{1-j} & \cdots & \varphi_{N-N}\end{bmatrix} \quad (7),$$

where φ_{i-j} parameter refers to the outlet heat rate of the i -th serverin the inlet heat rate of the j -th server of data center, N is the number of servers in a servers room.

Let us suppose that H_i^{out} is outlet heat of i -th server and H_j^{in} is inlet heat of j -th server. H_j^{in} can be calculated on server room environment heat H_{env}, power consumed by j -th server and H_i^{out} value (Figure 3):

$$H_j^{in} = \sum_{i=1}^{N} H_i^{out} \cdot \varphi_{i-j} + H_{env} + P_j. \quad (8),$$

Heat rate allows estimating the temperature at each server within a serverroom by temperature map models described at previous chapter.

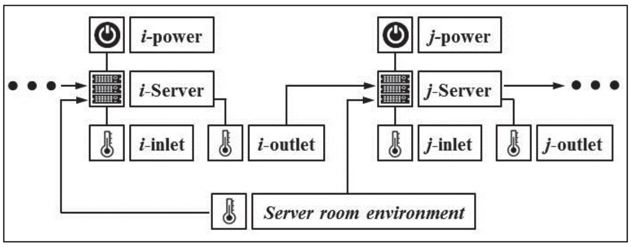

Fig. 3. Scheme of estimatio of data center server temperature level

Cooling facility model usually includes following functional components:

– cooling tower;
– chiller;
– server room air conditioning (SRAC);
– server room air handling (SRAH).

Thereby, the heat generated by data center's servers is absorbed by SRAC conditioned air provided from CRAH, and then it has to be drawn by SRAH system. SRAH exchanges the heat with cold air (or water) provided from a chiller based on refrigeration cycle. Comparative analysis of cooling power should be provided as varying processor utilization process which leads to adjusting the server room temperature change [10–14].

Up to this model power usage effectiveness (PUE) as a comparison of total power utilized by data center and power utilized by servers can be evaluated on server temperature set-point (Figure 4) which depends on CRAH efficiency [13; 14]:

$$E_{CRAH} = \frac{T_{SRAH}^{air} - T_{room}}{T_{SRAH}^{air} - T_{SRAH}^{water}} \quad (9),$$

where T_{SRAH}^{air} refers to the temperatures of air exhausted from server room and T_{SRAH}^{water} is the temperature of chilled water flowing into the SRAH.

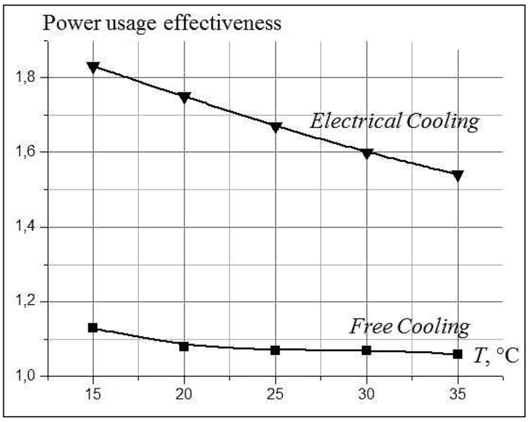

Fig. 4. Power usage effectiveness of data center server room in electrical and free cooling

Therefore, those parameters can be calculated by server power consumption, outside temperature estimation. While $E_{CRAH} < 1$ it should be noticed T_{room} has to be always higher than T_{SRAH}^{water}.

Conclusions. It was shown that development of power and temperature monitoring system is a key aspect of ensuring of data centers performance. Thereby it is important to build accurate model of the scalable server room power consumption system and cost-effective cooling facility. Analysis demonstrates that the servers' power consumption is always correlated with key workload parameters of shared storage, memory, computational capability and network bandwidth.

It was considered that for management of temporal and spatial temperature variations of servers' stability it is necessary to develop accurate temperature map simulation algorithm. Computational fluid dynamics simulation is proved to be effective instrument of analysis while it uses mathematical methods for development of precise fluid flow model. Though, it is a very complex model because it based on differential equation with no analytical solution so resource-intensive numerical procedures have to be used in this case. It was proposed to use algorithm which allows decreasing complexity of computational fluid dynamics simulation and building accurate temperature map. The algorithm is based on building of heat- and air-flow graphs. Simplified temperature model for servers is oriented on the computational and memory-sockets of servers, as well as on heat removal capability referring to the fan speed changes. Model includes building of thermal RC network scheme of the system which is based on connection between thermal and electrical losses

Cooling facility model included cooling tower, chiller, server room air conditioning and server room air handling. It was mentioned that for estimation of data center servers temperature map, it is necessary to analyze account interactions of multiple servers' heat and air flows within the bounds of the server's room. This procedure alloweddeveloping preciseheat recirculation scheme of datacenter.

References:

1. Kralicek E. Physical vs. Virtual Server Environments. The Accidental SysAdmin Handbook. 2016. P. 121–134.
2. Rivoire S., Ranganathan P., Kozyrakis C. A Comparison of High-Level Full-System Power Models. Hot Power. 2008. № 8 (2). P. 32–39.
3. Pedram M., Hwang I.. Power and performance modeling in a virtualized server system. Parallel Processing Workshops (ICPPW), 39th International Conference on. 2010, September. P. 520–526.
4. Megdiche M. Dependability Engineering for Data Center Infrastructures. Data Center Handbook. 2014. P. 275–305. DOI: 10.1002/9781118937563.ch15.
5. Choi J., Kim Y., Sivasubramaniam A., Srebric J., Wang Q., Lee J. A CFD-based tool for studying temperature in rack-mounted servers. IEEE transactions on computers. 2008. № 57 (8). P. 1129–1142.
6. Heath T., Centeno A.P., George P., Ramos L., Jaluria Y., Bianchini R. (2006, October). Mercury and freon: temperature emulation and management for server systems. In ACM SIGARCH Computer Architecture News. 2006. Vol. 34. № 5. P. 106–116.
7. IEEE Transactions on Very Large Scale Integration (VLSI) Systems publication information. IEEE Transactions on Very Large Scale Integration (VLSI) Systems. 2013. № 21 (2).
8. Ayoub R., Nath R., Rosing T. JETC: Joint energy thermal and cooling management for memory and CPU subsystems in servers. High Performance Computer Architecture (HPCA), IEEE 18th International Symposium on. 2012, February. P. 1–12.
9. Pakbaznia E., Pedram M. Minimizing datacenter cooling and server power costs. Proceedings of the 2009 ACM/IEEE international symposium on Low power electronics and design. 2009, August. P. 145–150.
10. Hwang D.C., Manno V.P., Hodes M., Chan G.J. Energy savings achievable through liquid cooling. Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm). 12th IEEE Intersociety Conference on. 2010, June. P. 1–9.
11. Breen T.J., Walsh E.J., Punch J., Shah A.J., Bash C.E. From chip to cooling tower datacenter modeling: Part I influence of server inlet temperature and temperature rise across cabinet. Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 12th IEEE Intersociety Conference on. 2010, June. P. 1–10.
12. Gao T., Samadiani E., Schmidt R., Sammakia B. Dynamic analysis of hybrid cooling datacenters subjects to the failure of CRAC units. ASME International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems. 2013, July. P. V002T09A015–V002T09A015.
13. Kim J., Ruggiero M., Atienza D. Free cooling-aware dynamic power management for green datacenters. High Performance Computing and Simulation (HPCS), International Conference on. 2012, July. P. 140–146.
14. Kim J., Sabry M.M., Ruggiero M., Atienza D. Power-Thermal Modeling and Control of Energy-Efficient Servers and Datacenters. Handbook on DataCenters. NewYork, 2015. P. 857–913.

МОДЕЛЮВАННЯ ТА МОНІТОРИНГ ІНФРАСТРУКТУРИ ІНФОРМАЦІЙНИХ СИСТЕМ

Розробка системи контролю потужності та температури є ключовим аспектом забезпечення ефективності роботи центрів обробки даних. Важливо побудувати точну модель масштабованої серверної системи енергоспоживання й економічного охолодження. Систематичний аналіз демонструє, що енергоспоживання серверів завжди корелює з ключовими параметрами робочого навантаження для загального інформаційного сховища, пам'яті, обчислювальних потужностей і мережевої інфраструктури. Показано, що для управління температурними змінами, які впливають на стабільність роботи серверів, необхідно розробити точний алгоритм моделювання температурних карт. Було показано, що моделювання динаміки потоку є найбільш ефективним інструментом аналізу, оскільки використовуються точні математичні методи. Проте практика продемонструвала, що така модель є надто складною, адже вона заснована на диференціальних рівняннях, для яких не існує аналітичного рішення, тому необхідно використовувати ресурсомісткі чисельні методи. Було запропоновано розробити алгоритм, який дозволяє зменшити складність моделювання і побудувати точну карту температур.

Ключові слова: *центр обробки даних, карта температур, енергоспоживання, система охолодження, серверний зал, моделювання динаміки потоку, RC-контур.*

МОДЕЛИРОВАНИЕ И МОНИТОРИНГ ИНФРАСТРУКТУРЫ ИНФОРМАЦИОННЫХ СИСТЕМ

Разработка системы контроля мощности и температуры является ключевым аспектом обеспечения эффективности работы центров обработки данных. Важно построить точную модель масштабируемой серверной системы энергопотребления и экономичного охлаждения. Систематический анализ демонстрирует, что энергопотребление серверов всегда коррелирует с ключевыми параметрами рабочей нагрузки для общего хранилища, памяти, вычислительных мощностей и сетевой инфраструктуры. Показано, что для управления температурными изменениями, которые влияют на стабильность работы серверов, необходимо разработать точный алгоритм моделирования температурных карт. Было показано, что моделирование динамики потока является наиболее эффективным инструментом анализа, поскольку используются точные математические методы. Тем не менее, практика продемонстрировала, что данная модель является чрезмерно сложной, поскольку основана на дифференциальных уравнениях, для которых не существует аналитического решения, поэтому необходимо использовать ресурсоемкие численные методы. Было предложено разработать алгоритм, который позволяет уменьшить сложность моделирования и построить точную карту температур.

Ключевые слова: *центр обработки данных, карта температур, энергопотребление, охлаждающая установка, серверный зал, моделирование динамики потока, RC-контур.*