

A FORMALIZATION OF THE EVALUATION PROCESS AT THE COMPUTER SYSTEMS INVESTIGATION

(1)Radi ROMANSKY

(1)Department of Computer Systems, Technical University – Sofia
e-mail address: rrom@tu-sofia.bg
Bulgaria

Abstract: The computer systems investigation is a process that requires precise organization of experiments for evaluation of different performance indexes. These experiments could be realized on the base of correct designed plan that describes the full sequence of steps for successful investigation. Each experiment gives some data that should be used for evaluation of systems characteristics and performance. The process of system evaluation and assessments defining should be mathematically formalized to secure correct carry out of the experiment and adequate experimental results. A formal procedure for system investigation and evaluation process using modeling is proposed in this paper. It has a hierarchical structure and formal description of the main problems of the evaluation process is given.

Key words: Computer systems investigation; Performance evaluation; Formal procedure.

1. INTRODUCTION

The computer systems investigation is a process that requires precise organization of experiments for evaluation of different performance indexes [1, 2]. This investigation could be made by two basic methods – monitoring [3, 4] or modeling [5, 6], but each experiment could be realized on the base of correct designed plan that describes the full sequence of steps for successful investigation. A plan for carry out of the experiments could be designed. This plan is very specific for each experiment and depends on investigated object and purpose of the assessments [7, 8], but in general the process of system evaluation and assessments defining should be mathematically formalized to secure correct carry out of the experiment and adequate experimental results.

A mathematical formalization of the evaluation process at the computer system investigation is proposed in this paper and sequence of formal steps is defined. A hierarchical procedure for system investigation using computer modeling is presented and some problems of this investigation are discussed.

2. PROBLEMS OF THE COMPUTER SYSTEM INVESTIGATION

It is known that the computer system consists of two parts defined its static discrete structure and dynamic stochastic workload. This defines the special features of the investigation process and problems of the experiments for system parameters evaluation. In general each computer processing should be described as a mathematical relation $f:D \rightarrow R$ for transformation of data set D in results set R , that is realized by computer units as equivalent relation $f^*:X \rightarrow Y$ using additional functions $\phi_1:D \rightarrow X$ (input function) and $\phi_2:Y \rightarrow R$ (output function). The realization of this general functional model is based on two sub-systems and each of them is object of investigation.

The first part of the problems is connected with *workload investigation*. The workload is formed by all program and information resources processed by computer units. These resources define a dynamical flow passed between hardware system components and changes the performance indexes. The servicing of this flow generates

a sequence of processes and each process should be described as a triple $I = \langle t, A, T \rangle$, where t is initial time for a process generation, A presents a set of defining parameters for profile characterisation should be presented as a time for task execution, times for separate events realization, initial times for events, effective usage of the system resources, number of accesses to the each resources, etc. As a general the input flow of tasks is described as a stochastic flow with known distribution. The units of this flow describe the tasks for servicing in the discrete static structure of the computer system. The parameters of this servicing have a stochastic nature too.

The second part of the problems is the *system resource investigation*. The system resource $S = \{S_1, S_2, \dots, S_n\}$ is formed on the base

system work load will be $R = \sum_{j=1}^n R_j$.

◆ Response time u_i is the time from input the current task for servicing to the time for result obtaining. The average response time

$u_{av} = \frac{1}{N} \sum_{i=1}^N u_i$ is calculated for N tasks.

◆ Reliability and fault tolerance are stochastic parameters estimated by number of failures in time period.

◆ Cost of the processing $C = \sum_{i=1}^n k_i \cdot W_i$ is

an assessment of the resources used for each task execution in the system, where W_i is the used volume of the resource S_i and k_i is a weight coefficient.

The performance is a general conception for computer work and includes different performance indexes characterised the system workload and system resources. These indexes could be evaluated by experiments based on the modelling or monitoring. An adequate experiment plan must be designed for correct realization and valid assessments obtaining. A general formalization of the evaluation procedure is proposed in the next section.

3. A PROCEDURE FOR SYSTEM INVESTIGATION AND EVALUATION

The evaluation process is based on

attributes and $T = \langle e_1, e_2, \dots, e_n \rangle$ is the process trace that defines the process profile (sequence of events for process realization).

The basic of different system components S_j that is loaded during processes realization. The system profile presents the level of system workload and should be described as a vector $V(t) = V_t = \langle v_1, v_2, \dots, v_n \rangle$ with elements define free ($v_j=0$) or busy ($v_j=1$) unit $S_j \in S$ (for $j=1, 2, \dots, n$) for sequence of times $t=1, 2, \dots$. Some of the main system characteristics for evaluation are listed below:

◆ Workload of the unit $S_j \in S$ during time interval $\tau_{ij} = t_{i+1} - t_i$ ($i=1, 2, \dots, k$) for the time period

$T = \langle t_1, t_2, \dots, t_{k+1} \rangle$: $R_j = \frac{1}{T} \sum_{i=1}^k \tau_{ij} \leq 1$. If

the processes are more than one the assessments of

the experiments for system investigation and experimental results obtaining. One of the methods for investigation of computer systems is the computer modelling. A general procedure for model investigation is shown in fig. 1. The planning of an experiment for computer modelling and investigation consists of the following base phases:

a) Empirical data collection (current values) for base systems parameters defined as performance indexes.

b) Structuring and initial processing of the collected empirical data and design of function scheme of the model.

c) Defining of 'a priori' information and spaces of definition for work parameters used at the model design.

d) Carry out of model experiments, collection of model information (experimental results) and analysis of obtained data for performance indexes evaluation.

The main steps of the procedure (marked on the fig. 1) are follows:

(1) Formulation of the main conception for model designing by system decomposition and building of a formal frame.

(2) Mathematical formalization of the system structure and relations by suitable formal system.

(3) Mathematical description of the system functionality and building of a functional model.

(4) Mathematical model designing on the base of built formal and functional models.

(5) A program model realization on the base of suitable program language.

(6) Carry out experiments by program model and analysis of the obtained experimental results.

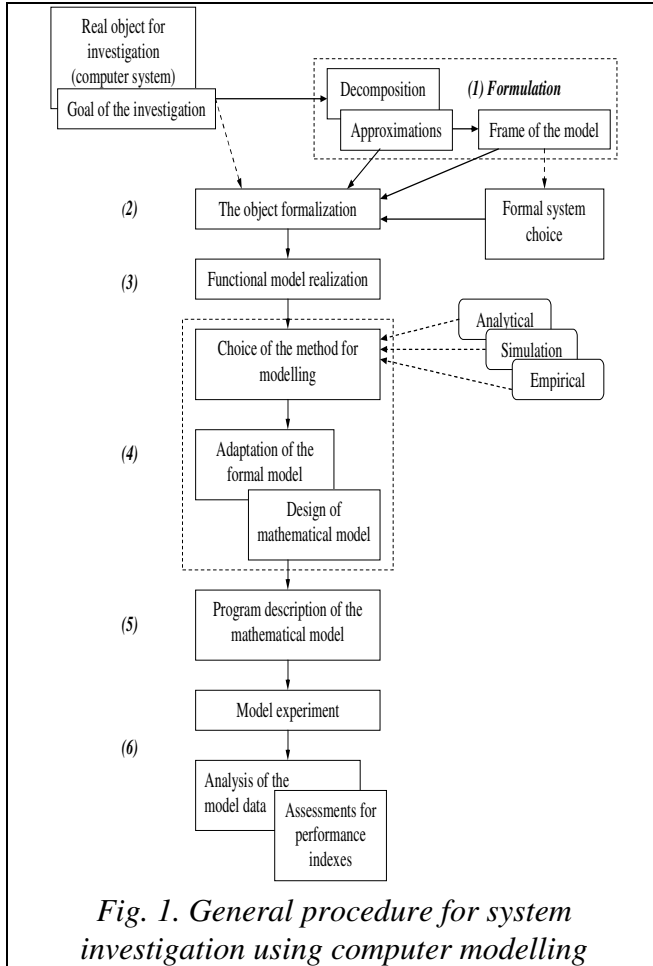


Fig. 1. General procedure for system investigation using computer modelling

4. A FORMALIZATION OF THE EVALUATION PROCESS

To decide each individual task for performance investigation and evaluation of computer system characteristics an experimental plan designing is needed. The main component of this plan must be the process connected to data collection, data analysis and data processing for performance indexes evaluating. In this reason proposed general procedure defines some basic problems connected to experiment planning as a first phase of the investigation and evaluation process.

1. *Precise definition and analysis of the task for investigation.* The task definition is very important for precise determining of the problems to be decided. The task analysis will define the size of the experiment and the need of task decomposition to the sub-tasks. This phase must include building of conceptual model based on the actions of the step (1) “Formulation” of the procedure presented in fig. 1.

2. *Defining the primary and secondary factors.* The initial information is necessary to define the model structural parameters and conditions for servicing. It is necessary to define the type and size of these data, as well as to determine the main factors – primary and secondary. A vector of variables $\vec{x} = (x_1, x_2, \dots, x_n)$ that will be evaluated must be defined.

3. *Acceptance of hypotheses.* This problem should be discussed on the base of the information for investigated object. If this information is insufficient one or more hypothesis must be accepted for replacement of missed data. These hypotheses should be accepted or rejected during the experiment realization.

4. *Determining of the model main content.* This problem is connected to the used method of investigation and special features of the real object and processes. The modelled object is a set of structural units, operations, relations and parameters that by decomposition and approximation could define the main content of the designed model. The result of this step will be a conceptual model designed on the base of actions from all previous steps.

5. *Specifying the ranges of definition.* For each variable x_i ($i=1, \dots, n$) used in the evaluation process should be determined a range for values, registered on the base of time intervals t_j for the time period T . This process should be formally described as follows:

$$\text{for } \forall x_i \ (i = 1, \dots, n) \ \exists R_i \Rightarrow \forall \xi_{ij} \in R_i \ (i = 1 \div n; j = 1 \div N), \text{ where } \xi_{ij} = x_i(t_j); t_j \in T.$$

In this case is accepted that experimental data are obtained during set of virtual clocks $\Delta t = |t_j - t_{j-1}|$, where $t_j, t_{j-1} \in T$ and the number of virtual clocks should be calculated as $N = T/\Delta t$. The time-interval Δt presents the registration frequency during the experiment. On the base of processes nature in computer systems should be

accepted that the ranges R_i are sub-ranges of the set of rational numbers, i. e. $R_i \subset R$ ($i=1, \dots, n$), or should be modified as rational numbers by simple relation. An experimental data $x_i(t_j)=\xi_{ij}$ ($\xi_{ij} \in R_i$) for each model variable x_i from the vector should be collected and on this base should be determined the minimal ($\xi_{i \min}$) and maximal ($\xi_{i \max}$) values, i.e.

$\exists (\xi_{i \min}, \xi_{i \max}) \in R_i$, so as $\xi_{i \min} \leq x_i \leq \xi_{i \max}$,
or for $\forall \xi_{ij} (j=1 \div N) \Rightarrow \xi_{ij} \in [\xi_{i \min}, \xi_{i \max}] \Rightarrow R_i \equiv [\xi_{i \min}, \xi_{i \max}] \subset R$.

6. *Formulation of goal and purpose of the assessments.* This step must define the specificity of the investigation and evaluation process and global content of the obtained assessments. This problem reflects to the step (2) and corrects determining of the vector of variables (the factors for the analysis). Significant factors should be included in this vector for securing the high precision of the mathematical model. A complete definition for each model factor or parameter is needed including name, dimension, range of variation, type of the factor (observed or controlled), role in the model, etc. On this base should be defined the general goal of the global assessment and the main purpose of the evaluation.

7. *Specifying of individual criteria.* These criteria allow determining the individual assessments of all investigated factors. Independent on subjective factor of the evaluation process each criterion should be regarded as a unique relation $g_i: R_i \rightarrow I$ ($i=1 \div n$) that permits to calculate the individual assessment $E_i = g_i(x_i)$ for each factor. The initial information $\{\xi_{ij}\}$ from the experiment defines the individual values for all variables of the vector $\bar{x} = \{x_1, x_2, \dots, x_n\}$. The ranges of definition R_i should be determined by grouping of experimental data for each variable x_i . It is possible to define l_i groups of compatibly so as $x_{ik} = [\xi_{ij_k}, \xi_{ij_{(k+1)}})$, $k = 0, 1, 2, \dots, (l_i - 1)$ on the base of the following rules: for $k = 0 \Rightarrow \xi_{ij_k} = \xi_{ij_0} \equiv \xi_{i \min}$; for $k = (l_i - 1) \Rightarrow \xi_{ij_{(k+1)}} = \xi_{ij_{l_i}} \equiv \xi_{i \max}$. The groups must be formed on the base of the requirement compatibly, i.e.: if $x_i \geq \xi_{ij_k} \Rightarrow E_i = e_{ik}$, and because $x_i \geq \xi_{ij_{(k+1)}} \Rightarrow E_i = e_{i(k+1)}$, at that $\xi_{ij_k} \leq x_{ik} < \xi_{ij_{(k+1)}} \Rightarrow E_i = e_{ik}$. For suitable

presentation of the assessment it is recommended to define $I = [0, 1]$. Such the estimation coefficients $E_i \in I$ will be presented as a percentage values. In the reason of stochastic nature of the processes in the computer systems it supposes that experimental data structure contents the number of realization for each event named by variable x_i that is stochastic valuable and accepts different values for the time-intervals $\Delta t_j = t_j - t_{j-1} \in T = (t_N - t_0)$. In this case calculation of all assessments should

be made by $E_i = \frac{\sigma_i}{\bar{x}_i} \in I$ using the next formulas:

- expectation of x_i : $\bar{x}_i = E[x_i] = \frac{1}{N} \sum_{j=1}^N \xi_{ij}$;
- variation of x_i :

$$\sigma_i^2 = V[x_i] = \frac{1}{N-1} \sum_{j=1}^N (\xi_{ij} - \bar{x}_i)^2 = \frac{1}{N-1} \sum_{j=1}^N (\xi_{ij}^2 - \bar{x}_i^2)$$
 ;
- standard deviation of x_i : $\sigma_i = +\sqrt{\sigma_i^2}$.

8. *Logical completeness of the assessments and actualization.* The goal is to determine the global criteria for evaluation $G: R_1 \times R_2 \times \dots \times R_n \rightarrow I$ that permits to calculate general assessment of system performance $E = G(\bar{x}) = F(E_1, E_2, \dots, E_n) = F[g_1(x_1), \dots, g_n(x_n)]$. A possible estimation is the following:

$$E = \frac{1}{n} \sum_{i=1}^n \tau_i \cdot E_i ,$$

where τ_i ($i=1, 2, \dots, n$) is the weight for each variable $x \in \bar{x}$. If it is realized $\sum_{i=1}^n \tau_i = 1$ and

$E_i \in [0, 1]$ for the global assessment will be realized $E \in [0, 1]$.

An actualization of the calculated assessments is recommended as a last step of the evaluation process and it must be included in the experimental plan. The purpose is to actualize the all assessments by cost coefficients k_{ct} . This actualization should be presented by the following expression $E_A = k_{ct} \cdot E$, where $k_{ct} \in I$.

5. CONCLUSION

The main goal of the modeling organization is to secure an efficiency of the investigation and correct assessments for observed values. Each experiment must be realized by designed

experimental plan. The proposed formal procedure for computer systems investigation and evaluation using modeling permits to realize an effective experimental plan because it is independent on the type of model experiments and the set of investigated parameters. It should be used for organization of the evaluation process at different computer systems investigation. The presented formalization of the evaluation process gives a possibility for strong determination of the steps sequence for assessments calculation and adaptation to the goal of the investigation.

REFERENCES

- [1] AlWehaibi, A., A. Agarwal, S. AlWakeel, N.-E. Rikli, A.ElHakeem. Performance behavior evaluation of Internet congestion control policing mechanisms. IEEE Canadian Conference on Electrical and Computer Engineering, May 2003, vol. 2, pp. 1017-1020.
- [2] Cancian, R.L., L. F. Friedrich. Performance evaluation of real time schedulers for a multicomputer. Proceedings of the 6th IEEE International Workshop on Distributed Simulation and Real-Time Applications, 11-13 Oct. 2002, pp.71-76.
- [3] Fabre, E., V. Pigourier. Monitoring distributed systems with distributed algorithms. Proceedings of the 41st IEEE Conference on Decision and Control, 10-13 Dec. 2002, vol. 1, pp. 411-416.
- [4] Korres, G.N., G. C. Contaxis. External system monitoring based on a reduced state estimation model. Power Tech Proceedings, 2001 IEEE Porto, 10-13 Sept. 2001, vol. 3, 6 pp.
- [5] Goldsman, D., G. Tokol. Output analysis procedures for computer simulations. Proc. of the 32nd Winter Simulation Conference, 10-13 Dec. 2000, vol. 1, pp. 39-45.
- [6] de Carvalho Klingelfus, A.L., W. Jr. Godoy. Mathematical modeling, performance analysis and simulation of current Ethernet computer networks. Proc. of the 5th IEEE International Conference on High Speed Networks and Multimedia Communications, 3-5 July 2002, pp. 380-382.
- [7] Benjamin, P., D. Delen, R. Mayer, T. O'Brien. A model based approach for component simulation development. Proc. of the 32nd Winter Simulation Conference, Dec. 2000, pp. 1831-1839.
- [8] Nossal, R., R. Lang. Model-based system development. IEEE Micro, July-Aug. 2002, vol. 22, No 4, pp. 56-63.