

УДК 539.374

Doi: 10.31772/2712-8970-2021-22-2-218-226

Для цитирования: Постановка задачи оптимизации структуры аппаратно-программного комплекса системы управления реального времени / С. Н. Ефимов, В. А. Терсков, О. Ю. Серикова, А. В. Попова // Сибирский аэрокосмический журнал. 2021. Т. 22, № 2. С. 218–226. Doi: 10.31772/2712-8970-2021-22-2-218-226.

For citation: Efimov S. N., Terskov V. A., Serikova O. Y., Popova A. V. Statement of the problem of optimization of the structure information processing computer appliances for real-time control systems. *Siberian Aerospace Journal*. 2021, Vol. 22, No. 2, P. 218–226. Doi: 10.31772/2712-8970-2021-22-2-218-226.

Постановка задачи оптимизации структуры аппаратно-программного комплекса системы управления реального времени

С. Н. Ефимов^{1*}, В. А. Терсков¹, О. Ю. Серикова², А. В. Попова¹

¹Сибирский государственный университет науки и технологий имени академика М. Ф. Решетнева
Российская Федерация, 660037, г. Красноярск, просп. им. газ. «Красноярский рабочий», 31

²Красноярский институт железнодорожного транспорта –
филиал Иркутского государственного университета путей сообщения
Российская Федерация, 660028, г. Красноярск, ул. Новая Заря, 2

*E-mail: efimov@bk.ru

В статье приведена постановка задачи оптимизации структуры аппаратно-программных комплексов, предназначенных для систем управления реального времени, применяемых, в том числе, в ракетно-космической отрасли. Кроме того, изучены особенности данной задачи, влияющие на выбор методов оптимизации. Делается вывод, что данная задача может быть эффективно решена с использованием эволюционных методов оптимизации.

Существующие модели производительности позволяют определять минимальную аппаратную конфигурацию многопроцессорного вычислительного комплекса. Предложенный в данной статье подход позволяет находить конфигурации, обладающие аппаратной избыточностью (по сравнению с минимальной конфигурацией), но, за счёт этого, имеющие большую вероятность нахождения в состояниях, обеспечивающих производительность, достаточную для достижения целей функционирования проектируемой системы управления реального времени. Описанный подход является более гибким, чем простое дублирование всех аппаратных компонентов минимальной конфигурации, что может быть использовано для уменьшения затрат на создание и эксплуатацию проектируемой системы управления.

Предложенная модель может быть использована для оптимизации производительности многопроцессорных аппаратно-программных комплексов систем управления реального времени. При этом нужно учитывать, что ресурсы, выделенные на создание и эксплуатацию аппаратно-программного комплекса, всегда ограничены. Поэтому целесообразно рассматривать задачу оптимизации производительности как многокритериальную: одним критерием будет производительность, а другим – затраты на создание аппаратно-программного комплекса.

Ключевые слова: аппаратно-программный комплекс, модель, производительность, системы управления реального времени, теория массового обслуживания.

Statement of the problem of optimization of the structure information processing computer appliances for real-time control systems

S. N. Efimov^{1*}, V. A. Terskov¹, O. Y. Serikova², A. V. Popova¹

¹Reshetnev Siberian State University of Science and Technology
31, Krasnoyarskii rabochii prospekt, Krasnoyarsk, 660037, Russian Federation

²Krasnoyarsk Institute of Railway Transport, branch of the Irkutsk State University of Communications
2, Novaya Zarya St., Krasnoyarsk, 660028, Russian Federation

*E-mail: efimov@bk.ru

The article presents the problem of optimizing the structure of information processing computer appliances for real-time control systems used, among other things, in the rocket and space industry. In addition, the features of this problem that affect the choice of optimization methods are studied. It's concluded that this problem can be effectively solved using evolutionary optimization methods.

Existing performance models allow you to determine the minimum hardware configuration of a multiprocessor computing system. The approach proposed in this article allows us to find configurations that have hardware redundancy (compared to the minimum configuration), but, due to this, have a greater probability of being in states that provide performance sufficient to achieve the goals of functioning of the designed real-time control system. The described approach is more flexible than simply duplicating all hardware components of the minimum configuration, which can be used to reduce the cost of creating and operating the designed control system.

The proposed model can be used to optimize the performance of multiprocessor hardware and software complexes of real-time control systems. At the same time, it should be taken into account that the resources allocated for the creation and operation of the hardware and software complex are always limited. Therefore, it is advisable to consider the problem of performance optimization as a multi-criterion: one criterion will be performance, and the other—the cost of creating a hardware and software complex.

Keywords: Computer appliance, model, performance, real-time system, queuing theory.

Introduction

A real-time system (RTS) is a hardware and software complex (HSC) that solves the problems of controlling various processes in conditions of time constraints.

Many modern control systems are real-time systems for which performance is a critical parameter: the control action must be developed in the required time, otherwise it becomes useless. This class of control systems includes, for example, control systems used in the rocket and space industry, air traffic control systems or technological process control systems. [1; 2].

Such control systems are hardware and software complexes that are a set of hardware and software that work together to accomplish a given task.

Requirements for the performance of computing systems used in real-time control systems are constantly increasing due to the increasing complexity of control objects.

Increasing the speed of computing technology has traditionally gone in two ways: increasing the clock frequency of processors and developing multiprocessor systems. Today, we can state that the possibilities of increasing the clock frequency have been exhausted, which is due to physical limitations [3]. This means that real-time control systems will inevitably be created on the basis of multiprocessor computing systems.

It is important to understand that the hardware performance requirements for real-time control systems are determined by the software that is used to generate the control action. Special requirements

are also imposed on the software of real-time control systems related to the need to ensure that the correct control action is obtained in a strictly defined time. Therefore it is advisable to study the performance of multiprocessor computing systems in close connection with the functioning of software.

For the design of multiprocessor hardware and software systems, a model of their performance is needed, which would make it possible to determine the speed of architecture options without experimentation, which can be extremely time-consuming and require significant costs.

The existing models of performance of multiprocessor computing systems [4–6] do not take into account the possibility of hardware failures and its recovery. In practice, when designing hardware and software complexes of control systems real-time, this aspect cannot be ignored, since a decrease in performance due to the failure of one of the processors can lead to the impossibility of generating a control action in the required time, which is unacceptable for real-time systems.

Performance model and formulation of the optimization problem

We consider a more general performance model, which includes additional states in which not all processors and buses are healthy, as well as transitions between states corresponding to processor and bus failures, as well as their recovery. The computing system is considered as a queuing system (QS).

The investigated HSC consists of N types of processors containing by M_i ($i = 1, 2, \dots, N$) processors of each type with the average execution time of one instruction T_{0i} . Processors are combined with RAM via N_1 buses. The service time of a request from a processor of the i type is τ_i . It is assumed that the time interval between any two adjacent claims obeys the Poisson distribution law with the parameter ν_i . The total flow of failures from processors of all types and interface buses also obeys the Poisson distribution law with the parameter λ_i . In addition, when evaluating the performance of a computing system, it is assumed that the time interval between two adjacent services obeys an exponential distribution law with the parameter μ_i , and the recovery time of buses and processors of the i type obeys the exponential law with the parameter ξ_i .

The states in which the considered system can be will be denoted as $a_{n, m_1, m_2, \dots, m_N, j_1, j_2, \dots, j_N}^{k, l}$. In this case, $(N_1 - n)$ interface buses are in good order and participate in the computational process, and n are faulty and are restored, $(M_1 - m_1)$ processors of the first type are serviceable and participate in the computational process, while m_1 are defective and are being restored, $(M_2 - m_2)$ processors of the second type are serviceable and participate in the computational process, and m_2 are faulty and restored, ..., $(M_N - m_N)$ processors of the N type are in good order and participate in the computational process, and m_N are faulty and restored. The system contains j_1 requests from processors of the first type, j_2 requests from processors of the second type, ..., j_N requests from processors of N type, k buses are busy with servicing, and l requests are in queues for servicing.

Due to the ordinariness of the streams of memory accesses, memory servicing of processors, failures and recovery of hardware components, transitions are possible only between states that differ in the value of only one index, and this index can either increase or decrease by one.

Composing the system of Kolmogorov – Chapman equations [7] according to the general rules for queuing systems, we obtain a system of linear differential equations for the probabilities of states in which the system can be.

Equating the derivatives to zero in this system, we obtain a system of linear algebraic equations for the probabilities of states in a stationary mode.

Solving the system with one of the numerical methods of linear algebra, we obtain the values of the probabilities of various states that can be used to determine any performance characteristics of the analyzed system [8].

In order for the failures of software elements to be considered statistically independent, like failures of various pieces of equipment, these elements must be developed independently [9]. This approach to critical software development is called multiversion programming (N -version programming) [10]. It is easy to understand that the performance of software developed using this approach increases with an increase in the number of different versions and an increase in the performance of their runtime environment [11].

Obviously by increasing the number of redundant hardware and software components, the system performance can be brought to any a given level [12]. However, such systems can be too expensive to develop and / or operate. Therefore performance models must be complemented by cost models. The cost of building hardware comes down to adding up the costs of the components. Models for estimating the costs of creating fault-tolerant software take into account the costs of developing multi-version software, labor costs of personnel employed at different stages of the software life cycle, etc. [13; 14].

The constructed models allow one to proceed to the formalization of the problem of choosing the optimal architecture options for multiprocessor hardware and software complexes of real-time control systems. In this case, two groups of criteria are obvious:

- performance criteria that must be maximized (the probability of being in a state in which performance is sufficient to generate a control action, etc.);
- cost criteria to be minimized (system cost, system development cost, operating cost, repair cost, etc.).

In this case, constraints will be imposed on the variable tasks, for example, in terms of energy consumption, speed, etc. To simplify the task, the cost criteria can be translated into constraints, since for all cost characteristics of the system, as a rule, there are upper bounds set by the customer of the system management. Having singled out the leader among the performance criteria, we get a one-criterion conditional optimization problem with a set of essential constraints, into which the rest of the criteria will go. In addition, there will be a set of natural constraints (for example, the number of hardware components is an integer and positive number).

We consider the type of variables in our optimization problem. In this case, we will assume given the maximum number of processor types N and software versions K , the maximum and minimum possible number of processors of each type and buses (for processors m_i^+ and m_i^- , respectively, $i=1, \dots, N$, and for buses n^+ and n^-). Let us denote by m_i the number of processors of the i type included in the structure of the hardware-software complex ($i = 1, \dots, N$), by n - the number of buses, and by k - the number of software versions. It is not difficult to see that the variables of our optimization problem (k, m_i, n) are integer, that is, we have a discrete optimization problem.

Let us give a formal record of the posed problem of optimizing the structure of a hardware-software complex with multiversion software for real-time control systems:

$$\begin{aligned}
 & R_0(m_1, \dots, m_N, n, k) \rightarrow \max, \\
 \text{under conditions} \quad & R_l(m_1, \dots, m_N, n, k) \geq R_l^0, l = 1, \dots, L_R, \\
 & C_l(m_1, \dots, m_N, n, k) \leq C_l^0, l=1, \dots, L_C, \\
 & m_i^- \leq m_i \leq m_i^+, i = 1, \dots, N, \\
 & n^- \leq n \leq n^+, \\
 & 1 \leq k \leq K.
 \end{aligned}$$

In this problem, the following designations are adopted: R_0 is the leading criterion for evaluating performance; $R_l, l = 1, \dots, L_R$, are secondary performance evaluation criteria; $C_l, l = 1, \dots, L_C$, - cost estimation criteria; R_l^0, C_l^0 - maximum permissible levels of criteria converted to restrictions.

When designing the optimal structure of the hardware-software complex one cannot focus on the maximum performance of special processors, but it is necessary to choose it so as to ensure the maximum performance of the entire hardware-software complex as a whole. For the formal statement of the problem, this means that the values of the average execution time of one instruction T_{0i} by processors of the i type cannot be constant, but must also be included in the number of optimization variables. Moreover, the parameters of the system v_i and μ_i become functions of T_{0i} , that is, $v_i = v_i(T_{0i})$, $\mu_i = \mu_i(T_{0i})$. This leads to a significant complication of the optimization problem, turning it into a two-level first hierarchical task:

$$(R_0^*(T_{01}, \dots, T_{0i}, T_{0N}), R_l^*(T_{01}, \dots, T_{0i}, T_{0N}), C_l^*(T_{01}, \dots, T_{0i}, T_{0N})) \rightarrow \text{extr},$$

где R_0^* , R_l^* и C_l^* is optimization problem solution.

First of all, it should be noted that the space of possible solutions is discrete, since the configuration of the hardware-software complex is determined by the number of processors of various types and RAM buses, which can only be integers. At the same time, the power of the search space grows rapidly with the increase in the number of processor types.

If we roughly estimate the power of the optimization space, then we get the total number of possible configurations more than $1,6 \cdot 10^{20}$. At the same time, significant restrictions will not significantly reduce the number of search points.

A significant problem for solving the resulting optimization problem is created by the method of calculating objective functions (criteria), which are mostly given algorithmically.

There are all the signs of a complex optimization problem: algorithmically specified functions, different types of problem variables, a variable number of sought variables, a large search area for an optimal solution.

When solving such optimization problems, evolutionary optimization algorithms have proven themselves well [15–18]. Therefore, the study of the effectiveness of evolutionary algorithms when optimizing the structure of hardware and software complexes of real-time control systems can be indicated as a possible direction for further research.

Conclusion

The existing performance models make it possible to determine the minimum hardware configuration of a multiprocessor computing complex. The approach proposed in this article makes it possible to find configurations that have hardware redundancy (compared to the minimum configuration), but, due to this, have a high probability of being in states that provide performance sufficient to achieve the goals of functioning of the designed real-time control system. The described approach is more flexible than simple duplication of all hardware components of the minimum configuration, which can be used to reduce the cost of creating and operating the designed control system.

The proposed model can be used to optimize the performance of multiprocessor hardware and software complexes of real-time control systems. It should be borne in mind that the resources allocated for the creation and operation of the hardware and software complex are always limited. Therefore, it is advisable to consider the problem of performance optimization as a multi-criteria one: one criterion will be performance, and the other will be the cost of creating a hardware-software complex.

Thus, this article presents the formulation of the problem of optimizing the structure of hardware and software complexes with multiversion software designed for real-time control systems. In the

future, it is proposed to investigate the effectiveness of using evolutionary optimization methods to solve this problem.

Библиографические ссылки

1. Васильев В. А., Легков К. Е., Левко И. В. Системы реального времени и области их применения // *Информация и космос*. 2016. № 3. С. 68–70.
2. Buttazzo G. *Hard Real-Time Computing Systems: Predictable Scheduling Algorithms and Applications*. New York, NY, Springer. 2011.
3. Sutter H. The free lunch is over: A fundamental turn toward concurrency in software // *Dr. Dobbs's Journal*. 2005. 30(3) [Электронный ресурс]. URL: <http://www.gotw.ca/publications/concurrency-ddj.htm> (дата обращения: 11.03.2021).
4. Liu Wang, Xiao Li, Shanghong Li Research on the Performance of Robot Multiprocessor Control System Based on BS Structure Digital Media // *Microprocessors and Microsystems*. 2020. Vol. 75. P. 103067.
5. Ефимов С. Н., Терсков В. А. *Реконфигурируемые вычислительные системы обработки информации и управления*. Красноярск : КРИЖТ ИрГУПС, 2013. 249 с.
6. Костров Б. В., Мартышкин А. И. Исследование структурной организации и оценка производительности многопроцессорных вычислительных систем с общей шиной // *Известия Тульского гос. ун-та. Технич. науки*. 2018. Вып. 2. С. 152–162.
7. Вентцель А. Д. *Курс теории случайных процессов*. М. : Наука, 1996. 400 с.
8. Бахвалов Н. С., Жидков Н. П., Кобельков Г. М. *Численные методы*. М. : БИНОМ. Лаборатория знаний, 2004. 636 с.
9. Липаев В. В. *Экономика производства программных продуктов*. М. : СИНТЕГ, 2011. 358 с.
10. Использование метода роя частиц для формирования состава мультиверсионного программного обеспечения / И. В. Ковалев, Е. В. Соловьев, Д. И. Ковалев и др. // *Приборы и системы. Управление, контроль, диагностика*. 2013. № 3. С. 1–6.
11. К вопросу реализации мультиверсионной среды исполнения бортового программного обеспечения автономных беспилотных объектов средствами операционной системы реального времени / И. В. Ковалев, В. В. Лосев, М. В. Сарамуд и др. // *Вестник СибГАУ*. 2017. Т. 18, № 1. С. 58–61.
12. Efimov S. N., Tyapkin V. N., Dmitriev D. D., Terskov V. A. Methods of Assessing the Characteristics of the Multiprocessor Computer System Adaptation Unit // *Journal of Siberian Federal University. Mathematics & Physics*. 2016. No. 9 (3). P. 288–295.
13. Глазова М. А. Модель СОСОМО II: анализ и пути совершенствования // *Экономика, статистика и информатика*. 2008. № 3. С. 63–67.
14. Шеенок Д. А., Кукарцев В. В. Прогнозирование стоимости разработки систем с программной избыточностью // *Известия Волгоградского гос. технич. ун-та*. 2013. № 14 (117). С. 101–105.
15. Тархов Д. А., Радченко Д. С. Распределенные алгоритмы оптимизации // *Современные информационные технологии и ИТ-образование*. 2015. Том 11, № 2. С. 404–408.
16. Semenkina O. E., Popov E. A., Seminkin E. S. Cooperative self-configuring nature-inspired algorithm for a scheduling problem // *IOP Conference Series: Materials Science and Engineering*. 2021. P. 12080.
17. Goldberg D. E. *Genetic algorithms in search, optimization, and machine learning*, Reading, MA: Addison-Wesley Professional. 1989.

18. Полякова А. С., Липинский Л. В., Семенкин Е. С. Автоматизированная система формирования состава коллектива многокритериальным генетическим алгоритмом. М. : Роспатент. 2020. № государственной регистрации программы для ЭВМ RU 2020663770 от 22.10.2020.

References

1. Vasil'ev V. A., Legkov K. E., Levko I. V. [The real-time systems and applications]. *Informaciya i kosmos*. 2016, No. 3, P. 68–70. (In Russ.)
2. Buttazzo G. *Hard Real-Time Computing Systems: Predictable Scheduling Algorithms and Applications*. New York, NY, Springer. 2011.
3. Sutter H. The free lunch is over: A fundamental turn toward concurrency in software // *Dr. Dobbs's Journal*. 2005, No. 30 (3). Available at: <http://www.gotw.ca/publications/concurrency-ddj.htm> (accessed: 11.03.2021).
4. Liu Wang, Xiao Li, Shanghong Li Research on the Performance of Robot Multiprocessor Control System Based on BS Structure Digital Media. *Microprocessors and Microsystems*. 2020, Vol. 75, P. 103067.
5. Efimov S. N., Terskov V. A. *Rekonfiguriruyemye vychislitel'nye sistemy obrabotki informacii i upravleniya* [Reconfigurable computing systems for information processing and management]. Krasnoyarsk, KRIZHT IrGUPS Publ., 2013, 249 p.
6. Kostrov B. V., Martyshkin A. I. [Investigation of the structural organization and performance evaluation of multiprocessor computing systems with a common bus]. *Izvestiya Tul'skogo gosudarstvennogo universiteta. Tekhnicheskie nauki*. 2018, Vol. 2, P. 152–162. (In Russ.)
7. Wentzel A. D. *Kurs teorii sluchajnyh processov* [Course of the theory of random processes]. Moscow, Nauka Publ., 1996, 400 p.
8. Bakhvalov N. S., Zhidkov N. P., Kobelkov G. M. *Chislennyye metody* [Numerical methods]. Moscow, BINOM. Laboratoriya znaniy Publ., 2004, 636 p.
9. Lipaev V. V. *Ekonomika proizvodstva programmnyh produktov* [The economics of the software engineering]. Moscow, SINTEG Publ., 2011, 358 p.
10. Kovalev I. V., Solov'ev E. V., Kovalev D. I. et al. [Application of particle swarm optimization to design of N-version software composition]. *Pribory i sistemy. Upravlenie, kontrol', diagnostika*. 2013, No. 3, P. 1–6. (In Russ.)
11. Kovalev I. V., Losev V. V., Saramud M. V. et al. [On the issue of implementing a multiversion execution environment for on-board software of autonomous unmanned objects by means of a real-time operating system]. *Vestnik SibGAU*. 2017, Vol. 18, No. 1, P. 58–61. (In Russ.)
12. Efimov S. N., Tyapkin V. N., Dmitriev D. D., Terskov V. A. Methods of Assessing the Characteristics of the Multiprocessor Computer System Adaptation Unit. *Journal of Siberian Federal University. Mathematics & Physics*. 2016, No. 9 (3), P. 288–295.
13. Glazova M. A. [COCOMO II Model: Analysis and Improvement]. *Ekonomika, statistika i informatika*. 2013, No. 14 (117), P. 101–105. (In Russ.)
14. Sheenok D. A., Kukarcev V. V. [Forecasting the cost of developing systems with software redundancy]. *Izvestiya Volgogradskogo gosudarstvennogo tekhnicheskogo universiteta*. 2018, Vol. 2, P. 152–162. (In Russ.)
15. Tarhov D. A., Radchenko D. S. [Distributed optimization algorithms]. *Sovremennyye informacionnyye tekhnologii i IT-obrazovanie*. 2015, Vol. 11, No. 2, P. 404–408. (In Russ.)

16. Semenkina O. E., Popov E. A., Semenkin E. S. Cooperative self-configuring nature-inspired algorithm for a scheduling problem. *IOP Conference Series: Materials Science and Engineering*. 2021, P. 12080.

17. Goldberg D. E. Genetic algorithms in search, optimization, and machine learning, Reading, MA: Addison-Wesley Professional. 1989.

18. Polyakova A. S., Lipinskij L. V., Semenkin E. S. *Avtomatizirovannaya sistema formirovaniya sostava kollektiva mnogokriterial'nym geneticheskim algoritmom* [An automated system for forming the composition of a team using a multicriteria genetic algorithm]. Moscow, Rospatent, 2020, No. *gosudarstvennoj registracii programmy dlya EVM* [state registration of a computer program] RU 2020663770. (In Russ.)

© Ефимов С. Н., Терсков В. А., Серикова О. Ю., Попова А. В., 2021

Ефимов Сергей Николаевич – кандидат технических наук, доцент, доцент кафедры информационно-управляющих систем; Сибирский государственный университет науки и технологий имени академика М. Ф. Решетнева. E-mail: efimov@bk.ru.

Терсков Виталий Анатольевич – доктор технических наук, профессор, профессор кафедры информационно-управляющих систем; Сибирский государственный университет науки и технологий имени академика М. Ф. Решетнева. E-mail: terskovva@mail.ru.

Серикова Олеся Юрьевна – аспирант; Красноярский институт железнодорожного транспорта – филиал Иркутского государственного университета путей сообщения. E-mail: olesyaserik@mail.ru.

Попова Анастасия Валерьевна – магистрант; Сибирский государственный университет науки и технологий имени академика М. Ф. Решетнева. E-mail: nasty.popowa@yandex.ru.

Efimov Sergei Nikolaevich – Cand. Sc., assistant professor, department of informational and control systems; Reshetnev Siberian State University of Science and Technology. E-mail: efimov@bk.ru.

Terskov Vitalii Anatolyevich – Dr. Sc., professor, department of informational and control systems; Reshetnev Siberian State University of Science and Technology. E-mail: terskovva@mail.ru.

Serikova Olesya Yuryevna – graduate student; Krasnoyarsk Institute of Railway Transport, branch of the Irkutsk State University of Communications. E-mail: olesyaserik@mail.ru.

Popova Anastasiya Valer'evna – master's degree student; Reshetnev Siberian State University of Science and Technology. E-mail: nasty.popowa@yandex.ru.
