UDC 004.2

A.S. LYAPANDRA
Ternopil National Economic University

LOAD BALANCING OF RECONFIGURABLE CLUSTER SYSTEM

Abstract - In of the article a model of load balancing between the branches of the reconfigurable cluster system is developed on purpose to productivity increase. The of high productivity of such systems is conditioned their structure represents the structure of algorithm of task decision. For of the further increase of the productivity it is suggested to minimize duration of algorithm implementation by diminishing of information duration of communication channels transfer. For of this purpose program algorithm and the computer system are presented as complete graphs, and for vertices and edges weighing coefficients are intended. For of vertices are the weighing coefficients of calculable complexity $\mathrm{C_A}(i)$, *productivity* $\omega(n)$, *and for edges are weighing coefficients of volumes given* $\mathrm{d}(i, j)$, *mean value of sizes of packets* $\mathrm{m}(i, j)$, *latence* $\mathrm{t_{sd}}(n_n, n_o, m)$ *and carrying capacity* $\mathrm{r}(n_n, n_o, m)$.

The of task of load balancing as task finding of such injective reflection function to the algorithm graph on graph of the distributed computer system enables to minimize duration of operation on one compute node at implementation of algorithm. Duration of calculations is determined with the resource of computation time and productivity; duration of interconnect operations on the basis of the basis of model of LogP/LogGP.

It of enabled to develop a model of load balancing for the increase of efficiency the use of compute node of the reconfigurable cluster system.

As of a result of the conducted work analytical dependence enables to balance loading by the increase of using efficiency of reconfigurable cluster system's compute node is built.

Keywords: load balancing, communication part, reconfigurable cluster system, duration of implementation of algorithm, graph of algorithm of the program, graph of cluster system.

А.С. ЛЯПАНДРА
Тернопільський національний економічний університет

БАЛАНСУВАННЯ НАВАНТАЖЕННЯ РЕКОНФІГУРОВАНОЇ КЛАСТЕРНОЇ СИСТЕМИ

В статті розроблено модель розподілу навантаження між гілками реконфігурованої кластерної системи з метою підвищення її продуктивності. Висока продуктивність таких систем обумовлена тим, що їх структура відображає структуру алгоритму розв'язання задачі. Для подальшого підвищення продуктивності запропоновано мінімізувати тривалість виконання алгоритму шляхом зменшення тривалості передачі інформації комунікаційними каналами. Для цього алгоритм програми та розподілена обчислювальна система представлені у вигляді повних графів, а для вершин та ребер призначені вагові коефіцієнти. Для вершин це вагові коефіцієнти обчислювальної складності $\mathrm{C_A}(i)$, *продуктивності* $\omega(n)$, *а для ребер – вагові коефіцієнти обсягів даних* $\mathrm{d}(i, j)$, *усереднених розмірів повідомлень* $\mathrm{m}(i, j)$, *латентності* $\mathrm{t_{sd}}(n_n, n_o, m)$ *і пропускної здатності* $\mathrm{r}(n_n, n_o, m)$.

Визначено задачу балансування як задачу знаходження такої одно-однозначної функції відображення графу алгоритму на граф розподіленої обчислювальної системи, яка дає можливість мінімізувати тривалість операції на одному обчислювальному вузлі при виконанні алгоритму. Тривалість обчислень визначено за часовою обчислювальною складністю і продуктивністю, а тривалість комутаційних операцій – на основі моделі LogP/LogGP, що дало можливість розробити модель балансування навантаження шляхом підвищення ефективності використання обчислювального вузла реконфігурованої кластерної системи.

У результаті проведеної роботи побудовано аналітичні залежності збалансування навантаження, що дає можливість підвищити ефективність використання ресурсів обчислювального вузла реконфігурованої кластерної системи.

Ключові слова: комунікаційна частина, реконфігурована кластерна система, тривалість виконання алгоритму, граф алгоритму програми, граф розподіленої обчислювальної системи.

Introduction

A necessity of far calculations induces to application of parallel architectures of the computer systems. So, for example, an important task is determination of amplitude-time parameters for authentication of diseases on the basis of harmonious analysis of chemiluminescence signals of bioassays, which is characterized by considerable calculable complexity [1, 2]. The use of the computer systems will enable to get the value of parameters quicker, that it is necessary at the production of biomedical implant [3].

Parallel computer systems passed arduous evolutive process and on the present stage of wide application the cluster systems purchased due to a high productivity at the insignificant charges of time and financial resources on the stages of planning and using of ones. Application of FPGA enables run-time reconfiguration of parallel computer system according to the task structure and can substantially decrease duration of decision of one. In addition, the use of the reconfigurable cluster systems in education process gives wide possibilities for research them and verification of the analytically calculated parameters of ones.

The tasks of balancing of calculable powers of the computer systems are devoted numerous works among which will mark works of Acharva A., Chanson S. Chen S., Chen Y., Cho Y., Downey A., Du X., Harchol-Balter M., Hui C., Isert C., Jiang H., Kunz T., Lee L., Qin X., Qu Y., S., Kuo J.L., Scheuermann P., Schwan K., Setia S., Swanson D.R., Vingralek R., Voelker G., Weikum G., Winslett M., Xiao L., Zabback P., Zhang X., Zhu Y.,

Shpakovskiy G.I., Gergel V.P. and Strogonin R.G. [4, 5].

Methods of compute nodes balancing with the criterion of resources efficiency using of central processing unit, main memory, combination of processor and memory, input-outputs, combinations of processor, memory and input-outputs are developed at works of the above-mentioned researchers. However, the question of consideration of the reconfigurable cluster systems features, namely possibilities of configuration of their structure according to the algorithm of task decision and data structure, is exposed not enough full [4, 5].

The goal of this work is a model of load balancing between the branches of the reconfigurable cluster system for the purpose its productivity increase. In future this model will be used for design of the reconfigurable computer cluster systems nodes.

In the article a research object is a calculable process of the reconfigurable cluster system, and its object is a model of load balancing between the compute nodes of the reconfigurable cluster system.

Load balancing model of the reconfigurable cluster system

For the design of industrial constructions or crash-tests it is necessary to execute 10^{15} operations, for the design of flowing around of airplane - 10^{16} operations, atmosphere or nuclear explosion - 10^{17} operations, and that is why an increase of the productivity of the computer systems is an actual task. Diminishing of duration of algorithms implementation are got with using both technological and structural methods. The increase of frequency of system work is limited to the physical phenomena. More potential is characterized structural methods those enable parallel task decision in time and/or space. Basic implementation of calculable powers parallel were founded in D825 in 1962 years that possibility of multiprocessor machines construction was well-proven; array processor ILLIAC IV in 1972 that SIMD structure was implemented; vector-pipeline CRAY in 1976 that pipeline calculations was created; cluster Beowulf was collected with serial equipments; multi-core processors with shared and private memory; quantum computer Orion, on which it is attained exponential speed due to superposition [5]. Development of parallel architectures allows to conduct a calculation in parallel. At the decision of many tasks of digital treatment of signals, design, optimization there is a requirement in the use of base algorithms of these areas. Basic descriptions of algorithms is instrument space, time and counting complexity. Algorithm time complexity of list ranging or finding an Eulerian path tree presents $O(\log n)$; finding a minimum spanning tree - $O(\log^2 n)$, and depth-first search - $O(\log^3 n)$ for the parallel algorithms in tasks using graph theory. Considerable part of algorithms in algebra, such as a decision of the triangular matrix, inverse three-cornered matrix, decision of System of linear equations, using Gauss elimination method for solving linear equations, calculation of rank of matrix, matrix similarity are characterized $O(\log^2 n)$ time complexity. Thus, perspective direction is the concerted development of parallel architectures of the computer systems and parallel algorithms [5].

According to rating of the most powerful computers in the world clusters present 86%. Such their considerable part is conditioned by the relative cheapness of ones at considerable calculable power. The top cluster system Tianhe-2 achieves productivity 33,86 PFlops [6]. However the acceleration of the parallel system largely depends on efficiency the run-time load balancing between compute nodes. Application of different methods enables to take into consideration critical factors in different task.

Divide of the reconfigurable cluster system (RClS): managing server MS (Management Server), compute node CN (Compute Node), communication system CmS (Communication System).

Basis of managing server MS of the reconfigurable cluster system RClS is the programmable logical integrated circuit (FPGA). One of variants its choice is application of morphological method of great number synthesis of alternative components structures on the basis of the generalized evaluation of resources of FPGA [7]. The choice of hardware tools at the conceptual design level realize with cross arrangement of critical to time structural blocks of FPGA and exchange speed of accessible part of communication matrix for the increase of efficiency of RClS. The managing server MS carries out planning and distribution of tasks, forming of result. The increases of time informative FPGA signals descriptions achieve by the use the specialized hardware resources. Ones is chosen according to goal function [8].

The load balancing methods with the efficiency criterion of central processing unit resources are used for the central processing units of compute nodes systems considerable loading. They consist in comparing of set threshold value to the value of loading Z of pipeline p_{pl} of core p_{cr} of processor p_{pr} of node p_{cn} of the reconfigurable heterogeneous cluster system $RClS$ defined analytically on the basis of tasks amount w_{RClS} core $p_{RClS-cn-pr-pl}$ execution, as a relation of amount of operations $\mathrm{w}_{RClS-cn}$ on a node cn to computational capacity $\mathrm{q}_{RClS-cn}$ of the same node:

$$Z_{RClS-cn} = \frac{\mathrm{w}_{RClS-cn}}{\mathrm{q}_{RClS-cn}}. \quad (1)$$

At $Z_{\text{threshold}} < Z_{RClS-1}$ the program of balancing determines loading of other node. If $Z_{\text{threshold}} \geq Z_{RClS-2}$, then

this processor can be high-usage. In case $Z_{RClS-2} > Z_{RClS-3} > Z_{RClS-4}$ there is loading of processor of node p_{RClS-4}.

However an finished goal is duration diminishing of implementation of calculations according to the concurrent algorithm. Therefore will consider the task of time minimization of program's implementation with optimization of load partition between the processor nodes of the cluster system.

The algorithm of the concurrent program are displayed a complete graph $\mathrm{AL} = (I, L)$. It's built by connection of instructions of parallel algorithm $\mathrm{I} = \{1,2,..., N_A\}$ finite set and the set of transitions $\mathrm{L} \subseteq I \times I$. The weights of calculable complexity $\mathrm{C_A}(i)$ of arithmetic $\mathrm{C_{Aa}}(i)$ and logical $\mathrm{C_{Al}}(i)$, $i \in I$ operations are assigned vertices I. The weights of volumes given $\mathrm{d}(\mathrm{i}, j)$ and middle sizes of reports $\mathrm{m}(i, j)$ are assigned edges L [9].

The computer architecture hardware is set a complete graph $\mathrm{HW} = (GNPC, L)$. One is created by the cluster compute nodes $\mathrm{N} = \{1,2,..., N_D\}$ finite set and interconnect connections $\mathrm{L} \subseteq I \times I$ set. Cluster compute nodes N are described by the weights of the productivity $\omega(n)$, and interconnects are described by the weight functions of latence $\mathrm{t_{sd}}(n_n, n_o, m)$ and channel capacity $\mathrm{r}(n_n, n_o, m)$.

Data processing duration t_{pa} of parallel algorithm AL is determined as a sum of global operations duration t_G and local operations duration t_l on a compute node n_n.

The injective function $f : AL \to HW$, $f(i) = n$, $x_{ni} = 1$, and in another cases $x_{ni} = 0$ on condition of minimization t_{pa}, is the decision of the task.

Duration of execution local operation t_l on one compute node n_n, $n_n \in GNPC$, is calculated as the sum of duration calculations t_c and communications between nearby compute nodes t_s with the communication system CmS:

$$t_l = t_c + t_s . \tag{2}$$

Duration of implementation of arithmetic and logical operations is determined a processor core t_c:

$$t_c = \sum_{n=1}^{N_N} x_{ni} \frac{C_A(i)}{\omega(n)}, \tag{3}$$

$C_A(n)$ is time calculable complexity of algorithm of A.

Duration of communications between nearby compute nodes t_s defines with LogP/LogGP model [4]. It describes a computer network in terms of delay t_{sd} and overhead costs t_{sl}:

$$t_{\mathrm{s_{mean}}} = 2(t_{sd} + 2t_{sl}) . \tag{4}$$

Thus, load balancing of branches in accordance with the function $f : AL \to HW$, at $\left(2(t_{sd} + 2t_{sl}) + \sum_{n=1}^{N_N} x_{ni} \frac{C_A(i)}{\omega(n)}\right) \to \min$ is one of terms of development of the cluster system high performance. The compare of results of load balancing reconfigurable cluster system taking into account the heterogeneity (LBRCSH) and without this accounting (LBRCS) are displayed in the table ($t_G = 1.2 \times 10^{-6}$ s):

Table 1

Results of modelling

Method	n	$\mathrm{C_A}(i)$, op	t_{sl}, 10^{-8} s	$\mathrm{t_{sd}}(n_n, n_o, m)$, 10^{-6} s	$\omega(n)$, 10^6 op/s	t_l, s	t_{pa}, s
	n_1	562	5.0	9.8	5.0	1.32	2.52
	n_2	456	5.3	19.6	6.7	1.07	2.27
	n_3	786	4.7	14.7	4.0	2.26	3.46
LBRCSH							8.25
LBRCS	n	601	1.5	13.4	5.1		8.68
Gain							3.00

Summary

In the article the model of load partition between the branches of the reconfigurable cluster system on purpose to increase its productivity is developed. The ones' high productivity of such systems is conditioned with

their structure reflects the structure of task decision algorithm. It is suggested to minimize duration of algorithm implementation by diminishing of information transfer duration of communication channels for the further increase of the productivity. For this purpose the program algorithm and the computer system are presented as complete graphs, and for vertices and edges weights are intended. For vertices are the weights of calculable complexity $C_A(i)$, productivity $\omega(n)$, and for edges are weight of volumes given $d(i, j)$, middle sizes of reports $m(i, j)$, latence $t_{sd}(n_n, n_o, m)$ and channel capacity $r(n_n, n_o, m)$.

The task of balancing is determined as the task of finding the injective function of reflect the algorithm graph to the distributed computer system graph for minimize duration of operation algorithm execution of one compute node. Calculations duration is determined with time and productivity complexity, and duration of interconnect operations accordingly to the LogP/LogGP model.

It enabled to develop the load balancing model and to increase using efficiency of the reconfigurable cluster system compute node.

As a result of the conducted work is analytical dependences of load balance that enables to increase efficiency resources using of compute node of the reconfigurable cluster system.

References

1. Lyapandra A.S. Amplitudno-chasovi parametry bioelektromagnitnykh sygnaliv fazovogo rivnya biotekhnichnykh system. Herald of Khmelnytsky National University. Technical sciences. Khmelnytsky. 2011. Volume 176. Issue 3. Part 1. pp.171–175.

2. Lyapandra A.S. Identyifikatsiya zakhvoryuvannya na osnovi garmoniynogo analizu khemilyuminestsentnykh sygnaliv bioprob. Herald of Khmelnytsky National University. Technical sciences. Khmelnytsky. 2007. Volume 100. Issue 6. Part 1. pp.137–144.

3. Karpinskyi M.P., Lyapandra A.S. Osoblyvosti pobudovy kompiuternoi systemy dlya biokhemiliuminometrychnykh doslidzhen. Herald of Khmelnytsky National University. Technical sciences. Khmelnytsky. 2005. Volume 100. Issue 6. Part 2. pp.167–171.

4. Culler D., Karp R., Patterson D., Sahay A., Schauser K. E., Santos E., Subramonian R., and T. von Eicken. LogP: Towards a realistic model of parallel computation. In Proceedings of the fourth ACM SIGPLAN symposium on Principles and practice of parallel programming, pages 1–12. ACM Press, 1993.

5. Shpakovskiy G.I. Parallelnoe programmirovanie i apparatura. Minsk, BGU, 2012, 184 p.

6. Home | TOP500 Supercomputer Sites. URL: http://top500.org/

7. Lyapandra A.S. Pidkhid do uzagalnenogo otsinyuvannya resursiv programovanykh logichnykh integralnykh skhem. Elektrotekhnichni ta kompyuterni systemy. K.:"Tekhnika", 2012. Issue 7. pp.92–96.

8. Lyapandra A.S. Zmenshennya chasovikh zatrymok PLIS u rekonfigurovanykh kompyuternykh systemakh. Herald of Khmelnytsky National University. Technical sciences. Khmelnytsky. 2012. Volume 193. Issue 5. pp.147–151.

9. Lyapandra A.S. Vplyv parametriv komunikatsiinoi chastyny rekonfihurovanoi klasternoi systemy na tryvalist vykonannya alhorytmu. Vymiriuvalna ta obchysliuvalna tekhnika v tekhnolohichnykh protsesakh. 2014. № 4. pp.125–127.

Література

1. Ляпандра А.С. Амплитудно-часові параметри биоелектромагнитних сигналів фазового рівня биотехнічних систем / А.С. Ляпандра // Вісник Хмельницького національного університету. Технічні науки. –2011. – №3. – С. 171–175.

2. Ляпандра А.С. Идентыификатсия закхворювання на основи гармониыного анализу кхемилюминестсентныкх сыгналив биопроб / А.С. Ляпандра // Вісник Хмельницького національного університету. Технічні науки. – 2007. – № 6. – С. 137–144.

3. Карпінский М.П. Особливості побудови комп'ютерної системи для биохемилиуминометричних досліджень / М.П. Карпінський, А.С. Ляпандра // Вісник Хмельницького національного університету. Технічні науки. – 2005. – № 6. – Ч. 2. – С.167–171.

4. Culler D., Karp R., Patterson D., Sahay A., Schauser K. E., Santos E., Subramonian R., and T. von Eicken. LogP: Towards a realistic model of parallel computation. In Proceedings of the fourth ACM SIGPLAN symposium on Principles and practice of parallel programming, pages 1–12. ACM Press, 1993.

5. Шпаковский Г.И. Параллельное программирование и апаратура / Г.И. Шпаковский. – Минск, БГУ, 2012. – 184 с.

6. Home | TOP500 Supercomputer Sites. URL: http://top500.org/

7. 7. Ляпандра А.С. Підхід до узагальненого оцінювання ресурсів програмованих логічних інтегральних схем // Електротехнічні та комп'ютерні системи. – К.:"Техника", 2012. – № 7. – С.92–96.

8. 8. Ляпандра А.С. Зменшення часових затримок ПЛИС у реконфигурованих комп'ютерних системах / А.С. Ляпандра // Вісник Хмельницького національного університету. Технічні науки. – 2012. – № 5. – С.147–151.

9. Ляпандра А.С. Вплив параметрів комунікаційної частини реконфигурованої кластерної системи на тривалість виконання алгоритму / А.С. Ляпандра // Вимірювальна та обчислювальна техніка в технологічних процесах. – 2014. – № 4. – С.125–127.

Рецензія/Peer review : 5.1.2015 р. Надрукована/Printed :26.1.2015 р.
Рецензент: к.т.н., доц. В.В. Яцків.