

УДК 681.3

<https://doi.org/10.32362/2500-316X-2024-12-6-20-25>

EDN PWDKPB



НАУЧНАЯ СТАТЬЯ

Анализ и выбор структуры многопроцессорной вычислительной системы по критерию быстродействия

Г.В. Петушков[@],
А.С. Сигов

МИРЭА – Российский технологический университет, Москва, 119454 Россия

[@] Автор для переписки, e-mail: petushkov@mirea.ru

Резюме

Цели. Анализ различных архитектур вычислительных систем (ВС), использовавшихся в последние десятилетия, позволил выделить наиболее распространенные структуры. Одной из ключевых особенностей является использование серийно производимого оборудования для создания подсистем обработки данных (например, многоядерные процессоры и полупроводниковая память большой емкости) и сетевого оборудования для построения коммуникационных подсистем. Это снижает затраты на оборудование и позволяет создавать типовые или кластерные конфигурации, что особенно важно для дорогостоящих ВС. Стремление достичь высокой вычислительной скорости и производительности в таких ВС требует минимизации времени на выполнение задачи и балансировки временных задержек как в подсистемах обработки данных, так и в коммуникационной подсистеме, обеспечивающей передачу данных внутри ВС. Целью работы является анализ вычислительных модулей (ВМ) и структур, на основе которых проводится построение кластерных ВС.

Методы. Основные результаты работы получены с использованием методов математического анализа и моделирования.

Результаты. Рассмотрена структура современных многоядерных микропроцессоров (МП), являющихся основой построения ВМ кластерных ВС. По мере увеличения числа ядер в структуре МП усложняется коммуникационная сеть, объединяющая их в единую структуру. Показано, что в новых разработках МП коммуникация между ядрами выполняется в виде сети, а сами МП представляют собой MIMD-структуры (множественный поток команд, множественный поток данных) в соответствии с известной классификацией Флинна.

Выводы. Предложенная методика выбора эффективной структуры ВС позволяет получить оптимальную структуру ВС по критерию быстродействия.

Ключевые слова: сеть InfiniBand, быстродействие, микропроцессоры, вычислительные модули, метрики Холстеда, анализ

• Поступила: 18.10.2024 • Доработана: 26.10.2024 • Принята к опубликованию: 05.11.2024

Для цитирования: Петушков Г.В., Сигов А.С. Анализ и выбор структуры многопроцессорной вычислительной системы по критерию быстродействия. *Russ. Technol. J.* 2024;12(6):20–25. <https://doi.org/10.32362/2500-316X-2024-12-6-20-25>

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

RESEARCH ARTICLE

Analysis and selection of the structure of a multiprocessor computing system according to the performance criterion

Grigory V. Petushkov[@],
Alexander S. Sigov

MIREA – Russian Technological University, Moscow, 119454 Russia

[@] Corresponding author, e-mail: petushkov@mirea.ru

Abstract

Objectives. Analysis of the various architectures of computing systems (CSs) used in recent decades has allowed us to identify the most common structures. One of the key features is the use of mass-produced equipment to create data processing subsystems (for example, multicore processors and high-capacity semiconductor memory), as well as network equipment to build communication subsystems. This reduces hardware costs and allows typical or cluster configurations to be created, which is especially important for expensive CSs. The desire to achieve high computational speed and performance in such CSs requires minimizing the time to complete the task and balancing time delays both in data processing subsystems and in the communication subsystem which provides data transmission inside the CS. The aim of this work is to analyze computing modules (CMs) and structures on the basis of which the construction of cluster CSs is carried out.

Methods. The main results of the work were obtained using methods of mathematical analysis and modeling.

Results. The study considers the structure of modern multicore microprocessors as the basis for building CMs of cluster CSs. As the number of cores in the microprocessor structure increases, the communication network which unites them into a single structure becomes more complicated. It has been shown that in new developments of microprocessors, communication between cores is performed in the form of a network. The microprocessors themselves are MIMD structures in accordance with the well-known Flynn classification.

Conclusions. The proposed method of selecting an effective structure of a CS allows us to obtain the optimal structure of a CS according to the criterion of performance.

Keywords: InfiniBand network, performance, microprocessors, computing modules, Halstead metrics, analysis

• Submitted: 18.10.2024 • Revised: 26.10.2024 • Accepted: 05.11.2024

For citation: Petushkov G.V., Sigov A.S. Analysis and selection of the structure of a multiprocessor computing system according to the performance criterion. *Russ. Technol. J.* 2024;12(6):20–25. <https://doi.org/10.32362/2500-316X-2024-12-6-20-25>

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

ВВЕДЕНИЕ

Основой построения вычислительных модулей (ВМ) кластерных вычислительных систем (ВС) являются современные многоядерные микропроцессоры. По мере увеличения числа ядер в структуре микропроцессора усложняется коммуникационная сеть, объединяющая их в единую структуру. В новых разработках микропроцессоров коммуникация между ядрами выполняется в виде сети, а сами микропроцессоры представляют собой MIMD-структуры¹ в соответствии с известной классификацией Флинна [1–3].

На следующем уровне на основе одного или нескольких микропроцессоров выполняется ВМ – ячейка кластерной ВС. В ВМ микропроцессоры объединяются с модулями оперативной памяти либо стандартными интерфейсами класса PCI² – Express 3.0, либо через коммутатор, обеспечивающий связь всех микропроцессоров со всеми модулями памяти [4].

На уровне системы в кластерных ВС сравнительно большое число ВМ объединяется сетевыми средствами. Как правило, для этого требуется несколько сетей [5]:

- сеть, обеспечивающая передачу данных между ВМ в процессе решения задачи;
- сеть, обеспечивающая связь отдельных ВМ с хранилищем данных, используемым как для начальной загрузки данных задачи, так и для хранения результатов;
- служебная сеть, связанная с управлением ВС, по которой циркулирует информация о мониторинге работоспособности ВМ и ВС в целом.

Самой быстрой из перечисленных сетей должна быть первая сеть, называемая также сетью передачи данных и выполняемая как сеть InfiniBand (IBA) [6]. Эта сеть поддерживает наиболее плотный трафик передач данных в процессе решения задач в ВС. По мере увеличения числа ВМ в ВС время обработки данных для задачи снижается за счет роста числа обрабатывающих устройств, а «накладные» расходы времени на передачу данных между ВМ будут расти [7].

Возникает вопрос об оптимальном числе ВМ в ВС, обеспечивающем минимальное время выполнения задачи при известных характеристиках ВМ и сети передачи данных.

Эта задача сформулирована в работе как задача выбора структуры кластерной ВС [8].

МЕТОДИКА ВЫБОРА ЭФФЕКТИВНОЙ СТРУКТУРЫ ВС, ОПТИМАЛЬНОЙ ПО КРИТЕРИЮ БЫСТРОДЕЙСТВИЯ

Оценим производительность ВС, использующих сеть IBA, при увеличении числа микропроцессоров и ВМ [9]. Рост производительности таких систем с увеличением числа ВМ происходит нелинейно, т.к. увеличение количества ВМ приводит к увеличению «накладных расходов», связанных с временем, необходимым для обмена данными между модулями. Время, затрачиваемое на обмен данными, особенно увеличивается при взаимодействии ВМ через сеть, где возникают задержки и увеличивается объем трафика.

Чтобы проанализировать влияние этих «накладных расходов» [10], рассмотрим идеализированный случай выполнения хорошо распараллеливаемой программы, которая состоит из N параллельно выполняемых фрагментов, распределенных по K ВМ, где $N > K$.

Время выполнения программы в рассматриваемом случае можно оценить следующим образом:

$$T_{\text{пр}} = T_{\text{выч}} + T_{\text{обм}}, \quad (1)$$

где $T_{\text{выч}}$ – время, затрачиваемое на вычисления в ВМ; $T_{\text{обм}}$ – время, затрачиваемое на обмен данными между ВМ.

Величину $T_{\text{выч}}$ можно оценить с учетом $P_{\text{ВМ}}$ – производительности одного ВМ и K – числа ВМ по формуле:

$$T_{\text{выч}} = \frac{G}{KP_{\text{ВМ}}}, \quad (2)$$

где G – оценка числа операций в программе.

Значение G можно получить при анализе алгоритма программы, используя, например, метрики Холстеда.

Затраты времени на обмен пакетом данных $T_{\text{п}}$ в сети передачи данных определяются следующим образом:

$$T_{\text{п}} = T_{\text{н}} + T_{\text{з}} + \frac{Q}{V}, \quad (3)$$

где $T_{\text{н}}$ – задержки формирования пакета передачи данных в адаптере сети; $T_{\text{з}}$ – задержка передачи пакета в сети, связанная с задержками в коммутаторе; Q – объем передаваемых данных в пакете данных; V – скорость передачи данных в сети.

Рассмотрим сеть передачи данных в составе ВС [11], поскольку объем трафика в ней много больше, чем в служебной сети. Анализируя время обмена данными как «накладные расходы» в процессе вычислений, в первом приближении $T_{\text{обм}}$ можно

¹ Multiple instruction, multiple data – множественный поток команд, множественный поток данных.

² Peripheral component interconnect – взаимосвязь периферийных компонентов.

оценить с учетом ограниченной пропускной способности сети как

$$T_{\text{обм}} = K \frac{Q}{V}. \quad (4)$$

Поясним более подробно получение выражения (4) на примере системы, состоящей из K ВМ. Поскольку объем трафика каждого ВМ пропорционален числу ВМ, предположим, что каждый ВМ пересылает пакеты другим ВМ по завершении выполнения своего фрагмента программы. При объединении трафика всех ВМ [12], пересылаемого по сети, можно оценить общее время обмена данными. Это время и отображено в виде (4).

Объем данных в байтах пакета обмена связан с числом операций в выполняемой программе G (вычислительной сложностью алгоритма) как

$$Q = CGL, \quad (5)$$

где C – коэффициент, характеризующий класс исполняемых алгоритмов в отношении связности по данным (алгоритмы с большей связностью по данным характеризуются большей интенсивностью обменов между ВМ и большим значением коэффициента C), а L – доля расчетных операций в выполняемом фрагменте программы ($L \approx 1$).

С учетом выражений (2)–(5), выражение (1) можно переписать в виде:

$$T_{\text{пр}} = \frac{G}{KP_{\text{ВМ}}} + \frac{CGL}{V}K. \quad (6)$$

Как следует из приведенного выражения (6), с ростом K снижается время, затрачиваемое на вычисления, и растет длительность времени на обмен данными в ВС. Таким образом, характеристики сети передачи данных будут сильно влиять на работу ВС.

Из выражения (6) следует, что можно провести оценку оптимального значения числа ВМ $K_{\text{оп}}$, обеспечивающего минимальное значение времени выполнения программы $T_{\text{пр}}$. Значение коэффициента k , определяющего число ВМ в полной ВС, определяется из условия $dT_{\text{пр}}/dk = 0$ и оценивается следующим образом:

$$K_{\text{оп}} = \sqrt{\frac{V}{CP_{\text{ВМ}}L}}. \quad (7)$$

Для современных ВС [13] значения коэффициента находятся на уровне единицы для пропускной способности сети $V = 10$ Гбайт/с и скорости обработки данных, составляющей 10 млрд операций в секунду. Однако для значений коэффициента $C = 0.01$ байт/операция, характеризующих алгоритмы со слабой связностью по данным, он может составлять порядка 10.

При увеличении пропускной способности сети передачи данных число ВМ в системе можно увеличивать без значительного снижения производительности из-за времени, затраченного на обмен данными. Использование оптической шины со скоростью передачи информации $V = 1$ Тбайт/с для сети передачи данных может увеличить оптимальное число ВМ более чем в 10 раз (до 100 ВМ), что значительно повышает эффективность работы многопроцессорной и многомодульной ВС.

Выражение (7) для коэффициента будем называть правилом «выбора эффективной структуры ВС» и отметим его важное практическое значение при проектировании ВС. Методика выбора структуры для кластерной вычислительной системы включает несколько этапов [14, 15]:

1. Анализ производительности ВМ.
2. Оценка скорости передачи данных в самой быстрой сети передачи данных ВС.
3. Анализ алгоритма для определения коэффициента связности по данным C .
4. Определение коэффициента в соответствии с выражением (7).

В реальных ВС число ВМ, как правило, превышает значение 10. Это связано с разработкой эффективных конструкций ВС. Однако с ростом скоростей передачи данных в сети расчетное и реальное количество ВМ будут все лучше согласовываться.

ВЫВОДЫ

1. Наиболее эффективные по соотношению «стоимость/производительность» структуры кластерных ВС строятся на основе стандартного оборудования сетей IBA и Ethernet.
2. На основе анализа временных затрат на обработку и передачу данных в ВС предложена методика выбора эффективной структуры ВС, что позволяет получить оптимальную структуру ВС по критерию быстродействия. Методика включает в себя следующие шаги:
 - анализ особенностей обмена данными между ВМ в зависимости от алгоритма решения задачи;
 - анализ производительности ВМ в зависимости от используемого оборудования;
 - анализ пропускной способности сети передачи данных в ВС;
 - выбор структуры ВС как выбор оптимального числа ВМ в соответствии с выражением (7).

Вклад авторов

Все авторы в равной степени внесли свой вклад в исследовательскую работу.

Authors' contribution

All authors equally contributed to the research work.

СПИСОК ЛИТЕРАТУРЫ

1. Воеводин В.В., Воеводин Вл.В. *Параллельные вычисления*. СПб.: БХВ-Петербург; 2002. 608 с.
2. Орлов С.А., Цилькер Б.Я. *Организация ЭВМ и систем*. СПб.: Питер; 2016. 688 с.
3. Хенесси Д.Л., Паттерсон Д.А. *Компьютерная архитектура. Количественный подход*. 5-е изд. М.: Техносфера; 2016. 936 с.
4. Кирк Б., Шарад С., Стенли В. Как преуспеть в условиях экономики изобилия памяти? *Открытые системы. СУБД*. 2016;2:25–32.
5. Сугак Е.В. *Прикладная теория надежности. Практикум*. М.: Лань; 2023. 312 с. ISBN 978-5-507-47014-3
6. Коваленко С.М. Оценка надежности информационно-управляющих систем на основе непрерывных моделей. *Вопросы радиоэлектроники*. 2005;4(2):143–146.
7. Половко А.М., Гуров С.В. *Основы теории надежности*. СПб.: БХВ-Петербург; 2006. 702 с. ISBN 5-94157-541-6
8. Черкесов Г.Н. *Надежность аппаратно-программных комплексов*. СПб.: Питер; 2005. 479 с. ISBN 5-469-00102-4
9. Коваленко С.М., Платонова О.В. Анализ задачи эффективной эксплуатации комплексов систем автоматизации и расчеты надежности на основе непрерывных моделей. *Известия вузов. Машиностроение*. 2014;8(653):75–89. <http://doi.org/10.18698/0536-1044-2014-8-75-79>
10. Podgorny Y.V., Antonovich A.N., Petrushin A.A., Sigov A.S., Vorotilov K.A. Effect of metal electrodes on the steady-state leakage current in PZT thin film capacitors. *J. Electroceram.* 2022;49:15–21. <https://doi.org/10.1007/s10832-022-00288-5>
11. Абдуллаев Д.А., Милованов Р.А., Волков Р.Л., Боргардт Н.И., Ланцев А.Н., Воротилов К.А., Сигов А.С. Сегнетоэлектрическая память: современное производство и исследования. *Russian Technological Journal*. 2020;8(5):44–67. <https://doi.org/10.32362/2500-316X-2020-8-5-44-67>
12. Конюхова О.В., Кравцова Э.А., Лукьянов П.В. *Техническое и программное обеспечение вычислительных машин и систем*. Инфра-Инженерия; 2023. 200 с. ISBN 978-5-9729-1186-8
13. Журавлев А.А. *Организация и архитектура ЭВМ. Вычислительные системы*. Лань; 2022. 144 с. ISBN 978-5-507-48089-0
14. Гельбух С.А. *Сети ЭВМ и телекоммуникации. Архитектура и организация*. Лань; 2019. 208 с. ISBN 978-5-8114-3474-9
15. Андреев А.М., Можаров Г.П., Сюзев В.В. *Многопроцессорные вычислительные системы. Теоретический анализ, математические модели и применение*. Издательство МГТУ им. Н.Э. Баумана; 2011. 336 с. ISBN 978-5-7038-3439-6

REFERENCES

1. Voevodin V.V., Voevodin V.I. *Parallel'nye vychisleniya (Parallel Computing)*. St. Petersburg: BHV-Petersburg; 2002. 608 p. (in Russ.).
2. Orlov S.A., Tsil'ker B.Ya. *Organizatsiya EVM i system (Organization of Computers and Systems)*. St. Petersburg: Piter; 2016. 688 p. (in Russ.).
3. Hennessey J.L., Patterson D.A. *Komp'yuternaya arkhitektura. Kolichestvennyi podkhod (Computer Architecture. A Quantitative Approach)*; transl. from Engl. 5th ed. Moscow: Tekhnosfera; 2016. 936 p. (in Russ.). [Hennessey J.L., Patterson D.A. *Computer Architecture. A Quantitative Approach*. 5th ed. Morgan Kaufmann; 2011. 856 p.]
4. Kirk B., Sharad S., Stanley V. How to succeed in an economy of abundance of memory? *Otkrytye sistemy = Open Systems. DBMS*. 2016;2:25–32 (in Russ.).
5. Sugak E.V. *Prikladnaya teoriya nadezhnosti. Praktikum (Applied Reliability Theory. The Workshop. Textbook)*. Moscow: Lan; 2023. 312 p. (in Russ.). ISBN 978-5-507-47014-3
6. Kovalenko S.M. Evaluation of the reliability of information management systems based on continuous models. *Voprosy radioelektroniki = Questions of Radio Electronics*. 2005;4(2):143–146 (in Russ.).
7. Polovko A.M., Gurov S.V. *Osnovy teorii nadezhnosti (Fundamentals of Reliability Theory)*. St. Petersburg: BHV-Petersburg; 2006. 702 p. (in Russ.). ISBN 5-94157-541-6
8. Cherkesov G.N. *Nadezhnost' apparatno-programmnykh kompleksov (Reliability of Hardware and Software Complexes)*. St. Petersburg: Piter; 2005. 479 p. (in Russ.). ISBN 5-469-00102-4
9. Kovalenko S.M., Platonova O.V. Analysis of the operational efficiency of complex automation systems and the calculation of their reliability on the basis of continuous models. *Izvestiya vysshikh uchebnykh zavedenii. Mashinostroyeniye = BMSTU J. Mechanical Engineering*. 2014;8(653):75–89 (in Russ.). <http://doi.org/10.18698/0536-1044-2014-8-75-79>
10. Podgorny Y.V., Antonovich A.N., Petrushin A.A., Sigov A.S., Vorotilov K.A. Effect of metal electrodes on the steady-state leakage current in PZT thin film capacitors. *J. Electroceram.* 2022;49:15–21. <https://doi.org/10.1007/s10832-022-00288-5>
11. Abdullaev D.A., Milovanov R.A., Volkov R.L., Borgardt N.I., Lantsev A.N., Vorotilov K.A., Sigov A.S. Ferroelectric memory: state-of-the-art manufacturing and research. *Russ. Technol. J.* 2020;8(5):44–67 (in Russ.). <https://doi.org/10.32362/2500-316X-2020-8-5-44-67>
12. Konyukhova O.V., Kravtsova E.A., Lukyanov P.V. *Tekhnicheskoe i programmnnoe obespechenie vychislitel'nykh mashin i system (Technical and Software Support of Computers and Systems)*. Infra-Inzheneriya; 2023. 200 p. (in Russ.). ISBN 978-5-9729-1186-8
13. Zhuravlev A.A. *Organizatsiya i arkhitektura EVM. Vychislitel'nye sistemy (Organization and Architecture of Computers. Computer Systems)*. Lan; 2022. 144 p. (in Russ.). ISBN 978-5-507-48089-0

14. Gelbukh C.A. *Seti EVM i telekommunikatsii. Arkhitektura i organizatsiya (Computer Networks and Telecommunications. Architecture and Organization)*. Lan; 2019. 208 p. (in Russ.). ISBN 978-5-8114-3474-9
15. Andreev A.M., Mozharov G.P., Syuzev V.V. *Mnogoprotsessornye vychislitel'nye sistemy. Teoreticheskii analiz, matematicheskie modeli i primeneniye (Multiprocessor Computing Systems. Theoretical Analysis, Mathematical Models and Applications)*. Moscow: Bauman Press; 2011. 336 p. (in Russ.). ISBN 978-5-7038-3439-6

Об авторах

Петушков Григорий Валерьевич, проректор, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: petushkov@mirea.ru. <https://orcid.org/0009-0006-0801-429X>

Сигов Александр Сергеевич, академик Российской академии наук, д.ф.-м.н., профессор, президент ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: sigov@mirea.ru. Scopus Author ID 35557510600, ResearcherID L-4103-2017, SPIN-код РИНЦ 2869-5663, https://www.researchgate.net/profile/A_Sigov

About the authors

Grigory V. Petushkov, Vice-Rector, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: petushkov@mirea.ru. <https://orcid.org/0009-0006-0801-429X>

Alexander S. Sigov, Academician at the Russian Academy of Sciences, Dr. Sci. (Phys.–Math.), Professor, President, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: sigov@mirea.ru. Scopus Author ID 35557510600, ResearcherID L-4103-2017, RSCI SPIN-code 2869-5663, https://www.researchgate.net/profile/A_Sigov