

Борзов Д.Б., канд. техн. наук,
Картамышев А.В., аспирант,
Ткачев П.Ю., аспирант

Юго-Западный государственный университет

АППАРАТНЫЕ СРЕДСТВА ПЛАНИРОВАНИЯ РАЗМЕЩЕНИЯ ПАРАЛЛЕЛЬНЫХ ПРОГРАММ В МУЛЬТИПРОЦЕССОРНЫХ СИСТЕМАХ

borzovdb@kursknet.ru

В статье описывается проблема размещения параллельных программ в мультипроцессорных системах. Делается вывод о невозможности использования программных средств планирования размещения при использовании критических систем (системы наблюдения, слежения, наведения, атомные объекты и т.д.) и предлагаются альтернативные аппаратные средства решения проблемы.

Ключевые слова: процессор, программа, отказ, критическая система, резерв, переразмещение.

В настоящее время мультипроцессоры являются наиболее распространенным классом вычислительной техники. Они, как правило, являются системами высокого параллелизма и включают в своем составе сотни процессорных модулей, которые обычно имеют матричную топологическую организацию [1]. В случае использования мультипроцессорных систем (МС), особенно в объектах критического назначения (системы слежения, наблюдения, управления, атомные объекты различного назначения и т.п.), требования к производительности, отказоустойчивости и быстродействию возрастают.

В МС могут возникать отказы и сбои внутренних процессорных модулей и/или каналов связи. В этом случае уменьшается внутренняя коммуникационная задержка, производительность и время реакции системы. Одним из вариантов решения данной проблемы может быть внутреннее переразмещение отказавших процессорных модулей и/или каналов связи [2–4].

Работа является продолжением исследований, начатых в [5–6]. Как показал анализ [2–6], в случае отказов в критических МС использование программных средств неприменимо из-за длительного времени решения. В связи с этим целесообразным является применение специализированных аппаратных средств планирования размещения и/или переразмещения.

Предлагается структурная схема мультипроцессорной хост-системы, представленная на рисунке 1. На нем каждый из процессорных ядер системы (ПЯ1, ПЯ2, ..., ПЯN) включает в себя устройство управления и арифметико-логическое устройство, ПЗУ, подключаемые к общей шине МС. Предлагаемые специализированные аппаратные средства планирования размещения программ предлагается использовать как периферийные устройства, также подключенные к общей шине, к ним относятся также внешняя память и устройства ввода-вывода.



Рис. 1. Структурная схема мультипроцессорной хост-системы

В случае выполнения задачи размещения предложен метод, методика и алгоритм [3, 4].

Пакет программ в МС представляется графом взаимодействия задач вида:

$$G = \langle X, E \rangle, \quad (1)$$

где $X = \{x_1, x_2, \dots, x_g, \dots, x_N\}$ – множество вершин графа G , $x_g \in X$ соответствуют про-

граммам, а дуги $e_{ij} \in E$ – связям между ними и представляют объемы данных m_{ij} , передаваемые между задачами, задаваемая матрицей обмена информацией (МОИ) $M = \|m_{ij}\|_{N \times N}$, где $N = n^2 = |X|$, $i, j = \overline{1, N}$.

Матричный блок МС представляется топологической моделью в виде графа $H = \langle P, V \rangle$, где $P = \{p_1, p_2, \dots, p_n\}$ – множество идентификаторов процессорных модулей блока, организованных в матрицу $|P|_{n \times n}$, где $|P| = N = n^2$ – процессоры; V – множество связей, описываемых матрицей смежности $\|W\|_{N \times N}$ размером $n^2 \times n^2$.

Тогда размещение пакета программ графа G в МС представляется отображением

$$\beta_s = \begin{pmatrix} x_1^{(s)} & x_2^{(s)} & \dots & x_N^{(s)} \\ p_1 & p_2 & \dots & p_N \end{pmatrix}, \quad (2)$$

где S – номер очередной перестановки, соответствующей s -му варианту размещения.

Мощность множества $\psi = \{\beta_s\}$ всевозможных отображений (2) равна числу всевозможных перестановок задач $x_N^{(s)}$ в матрице $X: |\psi| = N!$. Множество длин d_{ij} кратчайших маршрутов передачи данных в пределах блока описывается матрицей минимальных расстояний (ММР) $D = \|d_{ij}\|_{N \times N}$, $N = n^2 = |P|$, которую можно построить по матрице смежности.

Задача планирования размещения сводится к поиску максимальной коммуникационной задержки при передаче данных между процессорными модулями $p_{a,b}$ и $p_{x,y}$, соответствующей отображению $\beta^* \in \psi$ и вычислению показателя $A = \max_{\beta_s \in \psi} \{T_{\beta_s}(p_{a,b}, p_{x,y})\}$ с последующей его минимизацией. В результате

$$T_{\beta^*} = \min_{\psi} \{A\}. \quad (3)$$

Отображение β_s вычисляется как

$$T_{\beta_s}(p_{a,b}, p_{x,y}) = d_{ij} \cdot m_{ij}, \quad (4)$$

где $i = (a-1) \cdot n + b$ и $j = (x-1) \cdot n + y$,

$$a \max_{\beta_s \in \psi} \{T_{\beta_s}(p_{a,b}, p_{x,y})\} = \max \{d_{ij} m_{ij}\}. \quad (5)$$

Поиск наилучшего варианта размещения по критерию (3) является сложной переборной задачей. Вариантом ее ускорения является применение целенаправленных поисковых перестановок строк и столбцов МОИ с выбором в ней α -го места перестановки элемента $m_{\alpha\beta}$ по критериям (3, 4).

$$d_{\alpha k} < d_{\alpha\beta}, \quad (6)$$

где $d_{\alpha k}, d_{\alpha\beta}$ – одноименные элементы матрицы ММР; $m_{\alpha\beta}$ – элемент МОИ, которому соответствует $\max \{m_{ij} | d_{ij}\}$, найденный в предыдущем шаге перестановок.

Число требуемых перестановок уменьшается, если не учитывать явно нецелесообразные из них, разрешая очередную перестановку по следующим дополнительным критериям [8, 9]:

$$m_{\alpha k} \cdot d_{\alpha k} < m_{\alpha\beta} \cdot d_{\alpha\beta}, \quad (7)$$

$$m_{\alpha k} < m_{\alpha\beta}. \quad (8)$$

Для ускорения поиска разработана методика ускоренного выполнения процедур планирования размещения программ [3, 4]:

1. Составляются две матрицы: обмена информацией (МОИ) и кратчайших маршрутов (ММР) между процессорами и коммуникационной средой процессорного блока.

2. Вычисляются минимум коммуникационной задержки T_{inf} и коэффициент эффективности исходного размещения задач η_H . При нахождении T_{inf} веса дуг графа G назначаются на смежные процессорные модули без учета реальной топологии.

3. По порогу эффективности $\eta_H > 2$ принимается решение о целесообразности инициализации процедуры поиска размещения. Коэффициент эффективности перестановок $\eta = T / T_{inf}$ определяется как отношение полученной величины задержки (3) к T_{inf} .

4. Выполняются шаги целенаправленных перестановок столбцов и строк матрицы обмена информацией. Далее находится максимальное значение коммуникационной задержки (2) по предыдущему варианту перестановок программ.

5. Находится минимум (3) из максимумов по всем вариантам перестановок и вычисляется коэффициент эффективности η .

6. Если η оказывается менее установленного порога эффективности $\eta \leq 2$, шаги поиска прекращаются и найденный вариант матрицы

обмена информацией считается соответствующим полученному варианту размещения.

В результате был разработан аппаратно-ориентированный алгоритм ускоренного планирования размещения программ в МС [2].

На основе предложенного метода, методики и алгоритма планирования размещения программ в МС, разработана структурная схема устройства планирования размещения программ, представленная на рисунке 2 [7, 8, 9].

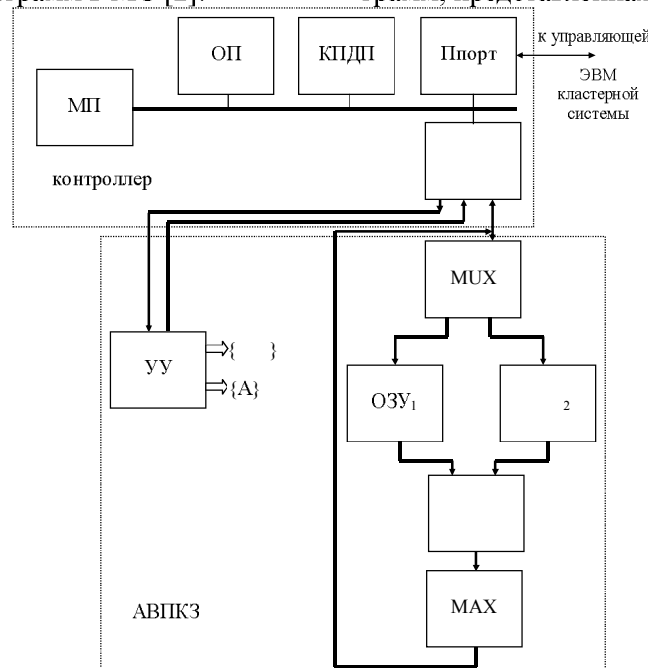


Рис. 2. Устройство планирования размещения программ

Операцию вычисления максимальной коммуникационной задержки, как одну из наиболее трудоемких, предполагается выполнять в аппаратном ускорителе: акселераторе вычисления показателя коммуникационной задержки (АВПКЗ). В блоке АВПКЗ применены конвейерный и матричный подходы для поэлементного перемножения векторов с одновременным подсчетом максимального произведения. Данные с параллельного порта поступают в блок специализированного мультиплексора (MUX). В зависимости от режима работы MUX загружает одно из ОЗУ данными соответствующей разрядности. Если идет загрузка в ОЗУ₂, из 8-и разрядного порта за один цикл приема MUX принимает два 4-х разрядных слова. Если же идет загрузка в ОЗУ₁, то MUX принимает два байта 16-ти разрядного слова и после их склеивания помещает целое слово в ОЗУ₁. Блок умножения матрично-конвейерный (БУМК) осуществляет перемножение синхронно считанных из ОЗУ₁ и ОЗУ₂ двух слов и выдает результат в блок нахождения максимума (МАХ). Умножение происходит конвейерно за один такт. Блок МАХ находит максимум и по сигналу об окончании расчета выдает результат за три цикла вывода в порт контроллера. Устройство управления (УУ) выдает управляющие сигналы, обозначенные на рис. 2 как множество микроопераций (МО), а

также значения адресов для ОЗУ₁ и ОЗУ₂ в режимах загрузки и вычисления.

В случае отказа процессора и/или межпроцессорной связи примем, множество программ описывается графом взаимодействия задач $\Phi = \langle X, E \rangle$, где $X = \{x_i\}$ – множество вершин, соответствующих программам, $E = \{e_{ij}\}$ – множество дуг, отражающих связи программам. Граф Φ задается матрицей смежности: $AM = \|m_{ij}\|_{N \times N}$, где $N = |X|$. Топология мультипроцессора представляется графом $H = \langle P_1, V \rangle$, где P_1 – процессоры, а множество V – межмодульные связи. Множество $P_1 = P \cup L$, где $P = \{p_{ij}\}$ – основные процессоры, $L = \{l_{ij}\}$ – резервные процессоры, причем $|P| = |L| = n^2$, $n = 2, 3, 4, \dots$. Множества процессоров P и L задаются матрицами $\|p_{ij}\|_{n \times n}$ и $\|l_{ij}\|_{n \times n}$ соответственно. Множество P_1 представляется объединением указанных матриц. Размещение задается отображением [5, 6].

$$\beta: X \rightarrow P, \tag{9}$$

ставящее в соответствие программам один из процессоров системы.

Тогда задача размещения программ в мультикомпьютере заключается в выборе такого

отображения $\beta_s \in Y$, которое соответствует критерию [10].

$$\zeta_s = \min_{\beta \in Y} \left\{ \max_{i=1, N, j=1, N} \{d_{ij} \cdot m_{ij}\} \right\}, \quad (10)$$

где максимум в фигурных скобках представляет собой наибольшую частную коммуникационную задержку для заданного отображения β .

В случае отказа отдельного канала связи между процессорами нарушаются маршруты транзитной передачи данных и необходимо найти новые кратчайшие маршруты обхода отказавшего канала. Для этого случая обоснована целесообразность использования известного алгоритма Дейкстры [11].

Для предложенного метода и аппаратно-ориентированного алгоритма отказоустойчивого перераспределения бала предложена соответствующая структурная схема микропроцессорного акселератора, представленная на рис. 3 [12].

На рис. 3 приняты следующие обозначения блоков и узлов: АППП – акселератор планирования перераспределения подпрограмм; ОП – оперативная память; КПДП – контроллер прямого доступа в память; Ппорт – последовательный порт; Прпорт – параллельный порт; УУ – устройство управления; БППОМ - блока перераспределения отказавших процессорных модулей; БПКМ - блока поиска кратчайшего маршрута; МО – микрооперации; А – адрес.

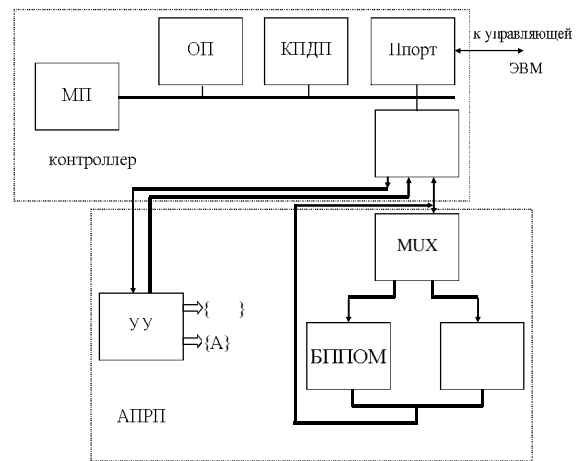


Рис. 3. Структурная организация акселератора отказоустойчивого перераспределения

МП работает по принципу, аналогичному описанной структурной схема выше. После замены процессорного модуля резервным либо после нахождения кратчайшего пути обхода найденный вариант передается через параллельный порт обратно в контроллер и далее через последовательный порт обратно в управляющую ЭВМ.

На основе представленной структурной схемы была разработана функциональная организация соответствующего устройства оперативной замены отказавшего процессорного модуля резервным, представленная на рисунке 4 [12].

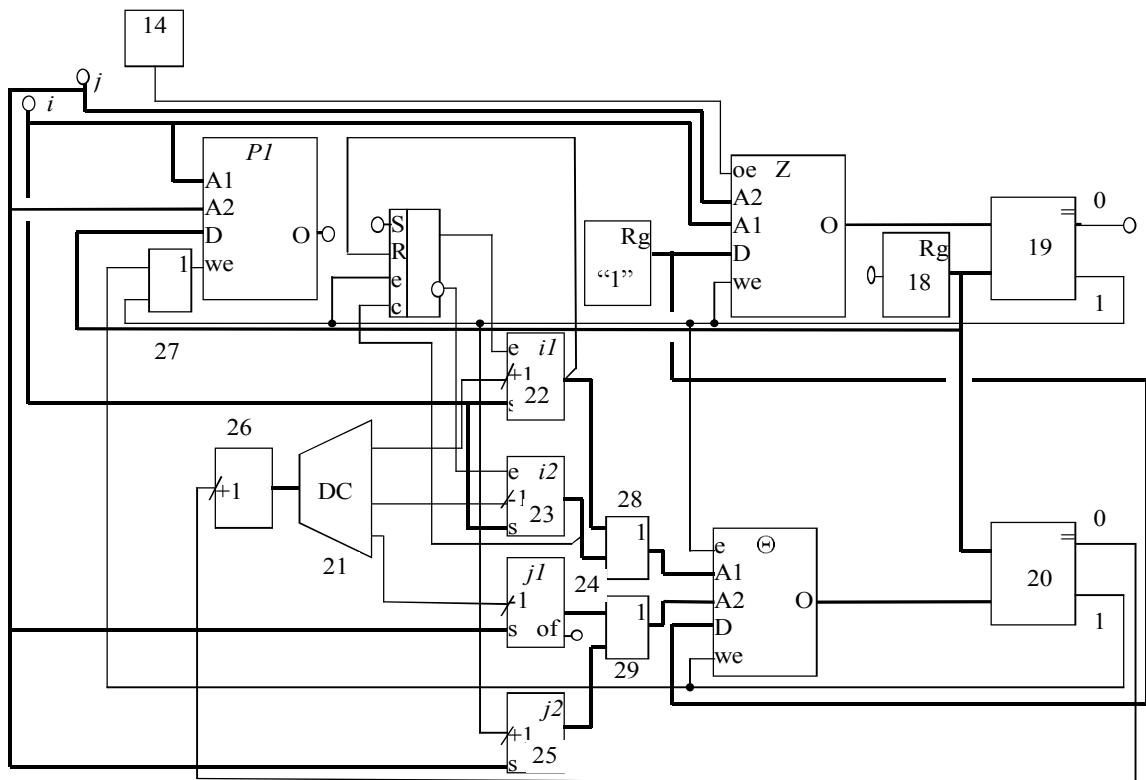


Рис. 4. Устройство замены отказавшего процессорного модуля резервным

На рис. 4 приняты следующие обозначения блоков и узлов: 14 – генератор импульсов, $P1$ – ОЗУ, Z – ОЗУ, Θ – ОЗУ, регистр 18, первый 19 и второй 20 элемент сравнения, дешифратор 21 выбора, первый 22 и второй 23 счетчик строки, первый 24 и второй 25 счетчик столбца, счетчик 26 номера предназначенный для подсчета номера выбираемого счетчика 22, 23, 24, 25, необходимого для проведения операций с заменой основного процессора резервным, первый 27, второй 28 и третий 29 Элементы ИЛИ.

Предложенное устройство функционирует в соответствии с предложенным алгоритмом отказоустойчивого перераспределения [13,14].

В начальном состоянии в ОЗУ 15 хранится матрица $P1$ процессорных модулей и матрица L резервных процессоров. В матрице Z хранятся нулевые коды, свидетельствующие о полной начальной работоспособности мультикомпьютера. В матрице Θ также хранятся коды нулей, свидетельствующие о наличии и полной начальной работоспособности резервных процессоров. В регистре 18 хранится код нуля («0...00»), в счетчике 26 содержится код нуля («0...00»), а значит, на выходе дешифратора 21 не появляется единичного импульса. Триггер 36 находится в высокоимпедансном состоянии.

В работе были предложены специализированные устройства ускоренного планирования размещения параллельных программ в МС, обеспечивающих сокращение времени формирования планов размещения программ, минимизирующих коммуникационную задержку МП при наличии отказов внутренних процессорных модулей и/или каналов связи.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Воеводин В.В., Воеводин Вл.В. Параллельные вычисления // БХВ– Петербург. Санкт-Петербург, 2002. 608 с.
2. Borzov D.B., J.A Azzeh, I.V. Zotov, D.E. Skopin, D. M.Al Hadidi. An Approach to Achieving Increased Fault-Tolerance and Availability of Multiprocessor-Based Computer Systems // Australian Journal of Basic and Applied Sciences. 8(6) April 2014. Pp. 512–522.
3. Борзов Д.Б., Титов В.С. Параллельные вычислительные системы (архитектура, принципы размещения задач): монография / Курск. гос. техн. ун-т. Курск, 2008. 156 с.
4. Борзов Д.Б., Титов В.С. Вопросы проектирования и динамической реконфигурации топологии систем логического управления в системах высокой готовности: монография / ЮЗГУ. Курск, 2015. 278 с.
5. Борзов Д.Б., Соколова Ю.В., Минайлов В.В. Перераспределение подпрограмм в отказо-

устойчивых мультипроцессорных системах / Известия вузов. Приборостроение. Санкт-Петербург. 2013. Т56. №6. С. 39–44

6. Борзов Д.Б., Борисенко Ю.В., Сизов А.С. Метод и аппаратно-ориентированный алгоритм перераспределения подпрограмм в мультикомпьютерах при отказе процессоров и связей между ними // Телекоммуникации. – Ежемесячный научно-технический, информационно-аналитический и учебно-методический журнал. 2013. №11. С. 45-48.

7. Борзов Д.Б., Чеснокова Е.О. Устройство поиска нижней оценки размещения в полносвязных матричных системах при однонаправленной передаче информации / Патент РФ №2398270, заявл. 11.02.2009; опубл. 27.08.2010, БИ 24, 21 с. 2 ил.

8. Борзов Д.Б., Бобынцев Д.О. Устройство поиска нижней оценки размещения в системах с матричной организацией при направленной передаче информации / Патент РФ №2406135, заявл. 9.02.2009, опубл. 10.12.2010, БИ №34, 12 с, 2 ил.

9. Борзов Д.Б., Чеснокова Е.О., Марухленко А.Л., А-А Муджиб Мохаммед Яхья. Устройство поиска нижней оценки размещения в полносвязных матричных системах при двунаправленной передаче информации / Патент РФ №2421805, заявл. 24.11.2008, опубл. 27.06.2011, 17 с, 2 ил.

10. Борзов Д.Б., Зотов И.В., Титов В.С. О субоптимальном размещении процессов и данных в кольцевых сетях // Известия вузов. Приборостроение. Санкт-Петербург. 2003. Т46, №11. С. 48-54.

11. Морозов К.К., Одинокоев В.Г., Курейчик В.М. Автоматизированное проектирование конструкций радиоэлектронной аппаратуры: Учебное пособие для вузов. М.: «Радио и связь». 1983. 280 с.

12. Пат. 2447485 Российская Федерация, МПК G06F7/76, G06F17/10 Устройство поиска нижней оценки размещения в матричных системах при двунаправленной передаче информации / Борзов Д.Б., Соколова Ю.В., заявитель и патентообладатель Федеральное государственное бюджетное образовательное учреждение высшего образования "Юго-Западный государственный университет" (ЮЗГУ). - №2009134208/08, заявл. 11.09.2009; опубл. 10.04.2012, БИ №10. – 5 с.

13. Борзов Д.Б., Соколова Ю.В. Методика перераспределения подпрограмм в отказоустойчивых мультикомпьютерах // Сборник трудов XVIII Международной научно-технической конференции «Машиностроение и техносфера XXI века. Т1». Донецк, 2011. С. 86–89.

14. Борзов Д.Б., Соколова Ю.В., Масюков И.И. Алгоритм перераспределения подпрограмм в отказоустойчивых мультимедийных компьютерах // Сборник трудов XVIII Международной научно-

технической конференции «Машиностроение и техносфера XXI века. Т1». Донецк, 2012. С. 101–103.

Borzov D.B., Kartamyshev A.V., Tkachev P.Y.

HARDWARE FOR PARALLEL PROGRAM SCHEDULING IN MULTIPROCESSOR SYSTEMS

This article describes the placement of parallel programs in multiprocessor systems problem. Concluded the inability of using scheduling software when using critical systems (system monitoring, tracking, targeting, nuclear facilities, etc.) and suggested alternative hardware to solve the problem.

Key words: *processing, program, failure, critical system, reserve, accommodation.*

Борзов Дмитрий Борисович, кандидат технических наук, доцент, кафедра вычислительной техники.

Юго-Западный государственный университет.

Адрес: Россия, 305035, Курск, ул. Овечкина, д. 5, кв.69.

E-mail: borzovdb@kursknet.ru

Картамышев Александр Вячеславович, аспирант кафедры вычислительной техники.

Юго-Западный государственный университет.

Адрес: Россия, 305026, Курск, ПЛК, д. 54, кв.57.

E-mail: ak-1802@mail.ru

Ткачев Павел Юрьевич, аспирант кафедры вычислительной техники.

Юго-Западный государственный университет.

Адрес: Россия, 305003, Курск, Узенький пер., д. 48.

E-mail: amdathlon64@yandex.ru