

УДК 004.75, 004.272.45

В.Н. Ручкин, В.А. Романчук, В.В. Лукашенко

**ОБОБЩЕННАЯ МОДЕЛЬ ВЫЧИСЛЕНИЙ КЛАСТЕРА
НЕЙРОКОМПЬЮТЕРОВ**

В статье рассматривается обобщенная модель вычислений кластера, рабочие узлы (машины) которого работают на базе нейропроцессоров. Данная работа выполнена в рамках гранта РФФИ 14-07-00261 «Кластеризация и организация облачных и распределенных вычислительных систем на базе нейропроцессоров».

Кластеризация, нейропроцессорные технологии, распределенные вычисления.

В общем виде любая задача, поступившая для выполнения в кластер, может быть представлена в виде алгоритма — последовательности машинных команд и данных [4]. Известно, что нейрокомпьютерные системы, осуществляя обработку информации, разбивают ее на фрагменты, которые можно сгруппировать по классам эквивалентности и в зависимости от решения этой задачи формируют оптимальную структуру обработки [2]. Данный принцип был перенесен на распределенную структуру типа кластер, вычислительными узлами которого являются нейрокомпьютерные модули. В предлагаемой нами модели распределенный вычислительный кластер виртуализован до абстрактной вычислительной машины с пулом ресурсов, как описано в работе [1].

В основе предлагаемой модели вычислений кластера нейрокомпьютеров лежит функциональный принцип, согласно которому главную роль играет множество операций, выполняемых на процессорных модулях распределенной вычислительной системы нейрокомпьютеров.

Операции (O) — это последовательность действий в алгоритме:

$$O = \{O_1, O_2, \dots, O_i, \dots, O_m\}.$$

В результате анализа алгоритма, поступившего в кластер, необходимо выбрать оптимальную структуру обработки. Множество возможных структур представим как $S = \{S_1, S_2, \dots, S_i, \dots\}$, где S_i — конвейер, вектор или их сочетание в виде комбинаций конвейерно-векторной, или векторно-конвейерной, или матричной и других структур.

Проектируемая структура должна выполнять заданный список операций под управлением алгоритма.

Алгоритм (A) — упорядоченная последовательность введенных операций или действий. Такая последовательность называется кортежем $A = \{O_1, O_2, \dots, O_N\}$, под

которым понимается конечный ряд операций O_l , определяемых согласно следующему выражению:

$$A^{(j)} = \{O_1, O_2, \dots, O_L, O_2, O_1, \dots, O_r\} \text{ длиной } |A^{(j)}|.$$

Структура распределенного нейрокомпьютерного подкластера рассматривается на уровне устройств. Поэтому на первом этапе определяется элементная база. Решение такой задачи математически сводится к нахождению отображения

$$O_l \rightarrow \{MK_m\}; \forall m = \overline{1, K_m}; \forall l = \overline{1, L}, \quad (1)$$

где под множеством команд $MK = \{MK_1, MK_2, \dots, MK_m\}$ конкретного комплекта СБИС подразумевается внутренний язык вычислительной системы; MK_m — минимальное количество команд, реализующее операцию O_l .

Вопрос выбора элементной базы вычислительных машин в виде конечных узлов вычислительного кластера рассмотрен выше.

На следующем этапе в зависимости от решения задачи проектирования (1) каждому j -му алгоритму обработки ставится в соответствие программа обработки $PR^{(j)}$ посредством определяемого отображения:

$$A^{(j)} \rightarrow PR^{(j)}, j = \overline{1, N}. \quad (2)$$

Под программой обработки информации $PR^{(j)}$ понимается кортеж команд $PR^{(j)} = \{MK_1, MK_2, \dots, MK_i, \dots, MK_M\}$. При этом характеристиками рассматриваемой программы $PR^{(j)}$

$$PR^{(j)} \Rightarrow (PR^{(j)}, T^{(j)}, X_m^{(j)}) \quad (3)$$

являются длина программы $|PR^{(j)}|$, определяемая как число макрокоманд, входящих в программу, частота повторения макрокоманд $X_m^{(j)}$, $m = 1, M$, и время выполнения программы $T = \sum_{m=1} X_m^{(j)} t_m$, где t_m — время выполнения m -й команды.

В результате решения задачи (2) определяется программа обработки однопроцессорного варианта нейрокомпьютера, что соответствует централизованной структуре обработки информации. Для решения задачи в распределенной вычислительной системе введем понятие структуры S , которую будем понимать как отношение параллельности выполнения подпрограмм RO_l, RO_k двумя различными

процессорными модулями: $RO_l S_w RO_k; RO_l, RO_k \in PR$. Отношение параллельности понимается как выполнение одновременно двух и более подпрограмм на разных вычислительных машинах кластера.

Тогда на третьем этапе проектирования определяется множество всевозможных структур $S_w \in S$, позволяющих некоторой j -й программе обработки PR поставить в соответствие множество подпрограмм RO_i :

$$S_w \in S : PR \Rightarrow \{RO_i\}; \forall i = \overline{1, L}; \forall w = \overline{1, W}. \quad (4)$$

В результате решение задачи (4) определяет технические затраты $Q^{(jw)}$, такие как требуемое число нейрокомпьютеров кластера, программные затраты $R^{(jw)}$, суммарный набор макрокоманд для каждой программы:

$$R^{(s_w)} = \sum_{j=1}^N \sum_{l=1}^L MK_l^{(j)}$$

На четвертом этапе осуществляется оптимальный выбор вычислительной структуры для последующей реструктуризации вычислительного кластера или его части из пула виртуализованных ресурсов. Формирование оптимальной структуры также осложнено возможной стратегией проектирования, включающей в себя минимум аппаратных и программных средств и удовлетворяющей требованиям по производительности. Алгоритм разбиения поступившей для выполнения в вычислительный кластер задачи на подзадачи, а также задачи выбора оптимальной структуры кластера по заданной стратегии в рамках данной работы не рассматривается.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Лукашенко, В.В. Разработка математической модели реструктуризуемого подклассы задач виртуализуемого кластера вычислительной grid-системы на базе нейропроцессоров [Текст] / В.В. Лукашенко, В.А. Романчук // Вестник Рязанского государственного университета имени С.А. Есенина : науч. журнал. — 2014. — № 1/42. — С. 176–181.
2. Романчук, В.А. Разработка модели сложной нейропроцессорной системы [Текст] / В.А. Романчук, В.Н. Ручкин, В.А. Фулин // Цифровая обработка сигналов. — Рязань : Информационные технологии, 2012. — № 4. — С. 70–74.
3. Ручкин, В.Н. Когнитология и искусственный интеллект [Текст] / В.Н. Ручкин, В.А. Романчук, В.А. Фулин. — Рязань : Узорочь, 2012. — 260 с.
4. Ручкин, В.Н. Архитектура компьютерных сетей [Текст] / В.Н. Ручкин, В.А. Фулин // Диалог-МИФИ. — М., 2008. — С. 238.
5. Ручкин, В.Н. Проектирование и выбор специализированных средств обработки информации [Текст] / Московский государственный открытый университет. — М., 1997. — 120 с.
6. Foster, I. The anatomy of the GRID [Text] / I. Foster, C. Kesselman, S. Tuecke // International Journal of High Performance Computing Applications. — 2001. — № 3/15. — P. 200–222.

REFERENCES

1. Foster, I. The anatomy of the GRID [Text] / I. Foster, C. Kesselman, S. Tuecke // International Journal of High Performance Computing Applications. — 2001. — № 3/15. — P. 200–222.
2. Lukashenko, V.V. Razrabotka matematicheskoy modeli restrukturizuyemogo pod klassy zadach virtualizuyemogo klastera vychislitel'noy grid-sistemy na baze neyroprotsessorov [Development of Mathematical Model of restructured under the classes of problems virtualizuemogo cluster computing grid-based systems neuroprocessors] / V.V. Lukashenko, V.A. Romanchuk // Bulletin of the Ryazan State University named after S. A. Yessenin : science magazine. — 2014. — № 1/42. — P. 176–181.
3. Romanchuk, V.A. Razrabotka modeli slozhnoy neyroprotsessornoy sistemy [The development model of a complex system neuroprocessor] / V.A. Romanchuk, V.N. Ruchkin, V.A. Fulin // Digital signal processing. - Ryazan: Information Technology, 2012. — № 4. — P. 70–74.
4. Ruchkin, V.N. Kognitologiya i iskusstvennyy intellekt [Cognitive science and artificial intelligence] / V.N. Ruchkin, V.A. Romanchuk, V.A. Fulin. — Ryazan : Uzorochye, 2012. — 260 p.
5. Ruchkin, V.N. Arkhitektura komp'yuternykh setey [Architecture of computer networks] / V.N. Ruchkin, V.A. Fulin // Dialog-MIFI. — Moscow, 2008. — P. 238.
6. Ruchkin, V.N. Proyektirovaniye i vybor spetsializirovannykh sredstv obrabotki informatsii [Design and selection of specialized information processing facilities] / Moscow State Open University. — Moscow, 1997. — 120 p.

V.N. Ruchkin, V.A. Romanchuk, V.V. Lukashenko

A GENERALIZED MODEL OF NEUROCOMPUTER CLUSTERS

The paper presents a generalized model for computing clusters, the production nodes of which function on the basis of neuroprocessors. The research is supported by RFFI grant 14-07-00261 “The Formation of Cloud System and Distributed Computer System Based on Neuroprocessor”.

Cluster formation, neuroprocessor technologies, distributed computing.