

УДК 004.383.3

В.А.Романчук

ОПТИМИЗАЦИЯ ПРОГРАММНЫХ И АППАРАТНЫХ СРЕДСТВ СЛОЖНЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ НА БАЗЕ НЕЙРОПРОЦЕССОРОВ

Рязанский государственный университет имени С.А.Есенина

Рассматриваются вопросы анализа и оптимизации вычислительных систем на базе нейропроцессоров с целью увеличения их производительности для заданного класса алгоритмов. Приведена адаптация общей методики анализа и оптимизации многопроцессорных устройств для нейропроцессоров семейства NM640x. Предложены структуры нейропроцессорных систем и получены оценки производительности для каждой структуры.

Ключевые слова: нейропроцессорная система, оптимизация, нейропроцессор.

Введение

В настоящее время для процессоров наступил так называемый "технологический предел", означающий что они достигли максимального уровня повышения быстродействия. Все разработки в данное время направлены на повышение числа процессоров на кристалле. Одним из выходов из данной ситуации является новая элементная база, например использование нейрокомпьютеров. Но для развития области нейропроцессорных технологий существует ряд проблем [1]:

1. Невысокая производительность нейропроцессорных устройств в связи с небольшой частотой нейрочипов (30-320 МГц).
2. Малое количество программного обеспечения для нейропроцессоров по сравнению с общераспространенными процессорами.
3. Секретность информационных материалов в данной области.
4. Слишком большая цена перехода от существующих процессоров к нейропроцессорам (изменение не только аппаратных, но и программных средств).

Одним из способов решения первой проблемы является организация многопроцессорных систем. Но имеются ряд проблем, мешающих созданию эффективных мультимикропроцессорных структур на базе нейропроцессоров, одной из которых является отсутствие средств оптимизации таких структур с целью увеличения их производительности.

Целью работы является исследование работы многопроцессорных систем на базе нейропроцессоров и разработка средств оптимизации программных и аппаратных средств вычислительных систем на базе нейропроцессоров с целью увеличения их производительности.

Теоретические исследования

За основу методики оптимизации вычислительных систем на базе нейропроцессоров – нейропроцессорных систем (НПС) была взята общая методика анализа и оптимизации многопроцессорных систем для заданного класса алгоритмов. В дальнейшем она была адаптирована для нейропроцессорных устройств и систем обработки информации. Поэтапная схема анализа и последующей оптимизации НПС приведена на рисунке 1.

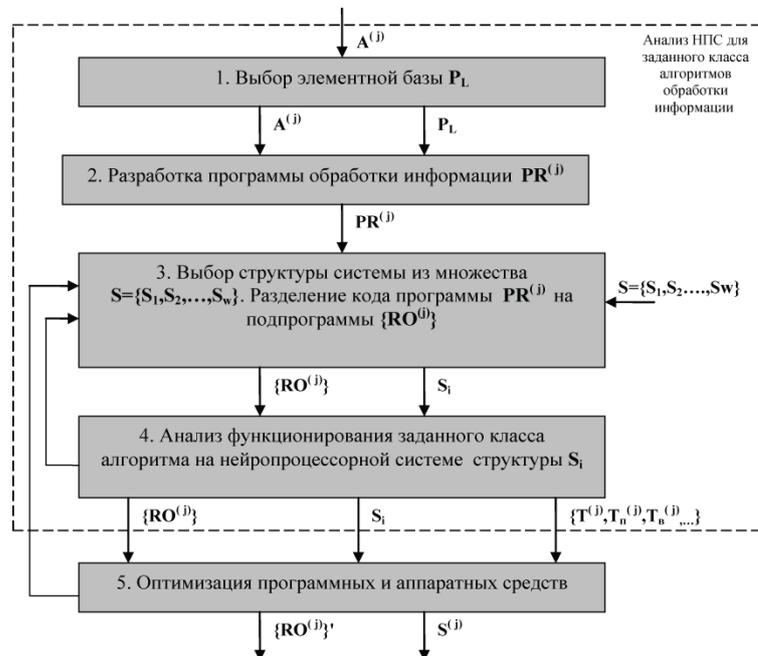


Рисунок 1 – Схема анализа и последующей оптимизации нейропроцессорных систем

Первым этапом является выбор элементной базы. Будем считать, что в результате исследования предметной области и представленного класса алгоритмов наиболее рациональным выбором является семейство процессоров NM640x.

На втором этапе требуется алгоритму обработки информации поставить в соответствие некоторую программу, написанную на внутреннем языке выбранного процессора, т.е. процессора семейства NM640x. То есть определить некоторое отображение φ , представленное в виде: $\varphi: A^{(j)} \rightarrow PR^{(j)}, j=1, N$,

где $A^{(j)}$ - некоторый j -й алгоритм обработки информации;

$PR^{(j)}$ - некоторая j -я программа для однопроцессорного варианта.

На третьем этапе необходимо рассмотреть множество возможных структур $S_w \in S; w=1, W$, позволяющих некоторой программе поставить в соответствие множество подпрограмм.

На четвертом этапе в результате анализа каждой структуры S_w

определяются основные выходные характеристики, необходимые для выбора наиболее рациональной структуры:

- время выполнения программы T ;
- время выигрыша T_e - это время, представляющее собой разницу времени обработки программы на нейропроцессорной системе относительно однопроцессорного варианта;
- время проигрыша системы T_n - это суммарное для всех процессорных модулей время, в течение которого i -й процессорный блок не был занят обработкой информации;
- время простоев в процессоре T_n' - суммарное время, в течение которого некоторые параллельные устройства процессора не были заняты обработкой информации.

Основными проблемами на данном этапе являются отсутствие единого стандарта нейропроцессорных архитектур и отсутствие критериев оценки эффективности рассматриваемых архитектур.

В связи с этим предлагается следующая методика:

1. Анализ программы $PR^{(j)}$.

Исследование программы $PR^{(j)}$ основано на понятии равенства подпрограмм $RO_j = RO_k$ (под которым понимаются равенство длин подпрограмм $|RO_j| = |RO_k|$ и полное совпадение подпрограмм с точностью до микрокоманды $MK_l^{(j)} = MK_l^{(k)}, \forall l = 1, |RO_j|$). Такие подпрограммы можно объединить в отдельный класс, называемый классом эквивалентности, а затем можно определить вид параллельной системы (конвейерная, векторная, конвейерно-векторная, векторно-конвейерная) обработки исходя из числа найденных классов $L^{(j)}$ и порядка каждого класса $|a_l|$.

2. Выбор архитектуры в соответствии с числом классов эквивалентности $L^{(j)}$ и порядком класса $|a_l|$.

Рассмотрим следующие архитектуры и оценки эффективности:

а) если число классов эквивалентности равно L и порядок каждого класса равен единице $|a_l|=1$, тогда рациональным выбором будет конвейерная структура обработки информации.

Время простоя в данном случае будет равно:

$$T_{np}^{(j)} = \sum_{k=1}^{q-1} (k * (T_c^{(j)} - TO_{q-k+1}^{(j)})) = \sum_{k=1}^{q-1} (k * (\max_{l \in L} TO_l^{(j)} - TO_{q-k+1}^{(j)})), \forall l = \overline{1, L},$$

где $TO_i^{(j)}$ - время выполнения i -й подпрограммы.

$$\text{Время выигрыша: } T_e^{(j)} = \sum_{k=1}^q (k * TO_{q-k+1}^{(j)}) - \sum_{l=1}^q TO_l^{(j)} = \sum_{k=1}^q ((k-1) * TO_{q-k+1}^{(j)});$$

б) если число классов эквивалентности L равно единице и порядок класса $|a_q|$ равен q , то выбирается векторная структура.

Время проигрыша: $T_n^{(j)} = \sum_{i=1}^q (\max_{l \in L} TO_l^{(j)} - TO_i^{(j)})$.

Время выигрыша: $T_g^{(j)} = \sum_{l=1}^L TO_l^{(j)} - \max_{l \in L} TO_l^{(j)}$.

в) если число классов эквивалентности равно L и порядок каждого l -го класса равен $|a_l|$ и в случае, когда информация не требуется одновременно для всех представителей классов, тогда рационально использовать так называемую конвейерно-векторную структуру обработки информации.

Время проигрыша: $T_n^{(j)} = K * \max_{l \in L} TO_l^{(j)} - \sum_{l=1}^q TO_l^{(j)}, \forall l = \overline{1, L}$.

Время выигрыша: $T_g^{(j)} = \sum_{l=1}^K ((K-l) * \sum_{i=lM}^{iN} TO_i^{(j)})$.

г) если информация требуется одновременно для всех представителей классов. Тогда, если подпрограммы внутри l -го класса обмениваются последовательно, назначим число процессорных модулей No для случая $|a_l| = q, \forall l = 1, L$, равное $No = L * q$, и для общего случая $No = L * \max_{l \in L} |a_l|$. В противном случае рационально использовать конвейерно-векторную структуру.

Получаемая структура обработки является векторно-конвейерной структурой обработки информации.

Время выигрыша: $T_g^{(j)} = \sum_{i=1}^K \sum_{k=iM}^{iN} (k * TO_{iN-k+1}^{(j)}) - \sum_{l=1}^q TO_l^{(j)}$.

Время проигрыша:

$T_n^{(j)} = \max_{G_i \in G} |G_i| * \max_{l \in L} TO_l^{(j)} - \sum_{l=1}^q TO_l^{(j)}, i = \overline{1, V}, \forall l = \overline{1, L}$.

Особенностью нейропроцессора является то, что он сам является векторным устройством, способным обрабатывать параллельно несколько потоков данных. Скалярную команду процессора NM6403 можно представить в виде совокупности двух операций (рисунок 2).

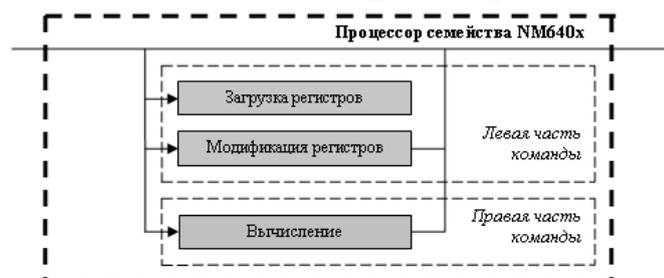


Рисунок 2 - Схема выполнения скалярной команды на процессоре семейства NM640x

Следовательно, если рассматривать процессор, обрабатывающий скалярную команду, то можно сказать, что процессор представляет собой

систему, параллельно обрабатывающую два потока данных.

Тогда время проигрыша будет равно: $T_{ns}^{(k)} = T_{os}^{(k)} - \sum_{l=1}^3 TO_{ls}^{(k)}$,

время простоя равно: $T_{nps}^{(k)} = \sum_{l=1}^3 (T_{os}^{(k)} - TO_{ls}^{(k)}) = 3 * T_{os}^{(k)} - \sum_{l=1}^3 TO_{ls}^{(k)}$.

Векторная команда также состоит из левой и правой части, но в правой части из-за переменной разрядности операндов возможно выполнение до 64 операций одновременно (рисунок 3).

Время простоя в этом случае равно:

$$T_{npv}^{(k)} = 20 * (N - 1) * T_o^{(k)} + (N - 1) * \sum_{l=5}^9 TO_l^{(k)} - \sum_{l=1}^4 TO_l^{(k)}.$$

Время проигрыша: $T_{nv}^{(k)} = N * (T_o^{(k)} - \sum_{l=1}^5 TO_{lvr}^{(k)}) - \sum_{l=1}^5 TO_{lvr}^{(k)} - \sum_{l=1}^4 TO_{lvt}^{(k)}$.

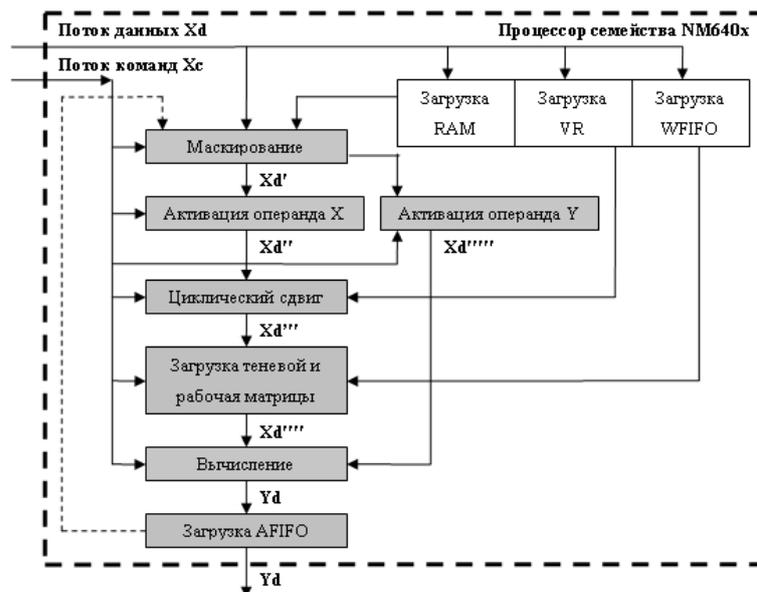


Рисунок 3 - Схема выполнения векторной команды на процессоре семейства NM640x

На пятом этапе необходимо минимизировать время простоев и время проигрыша $T_n^{(j)} \rightarrow \min$ и повысить время выигрыша $T_e^{(j)} \rightarrow \max$ исходя из полученных на четвертом этапе оценок эффективности НПС для реализации данного алгоритма.

Заключение

В результате работы были получены следующие результаты.

1. Предложена методика анализа и оптимизации средств обработки информации, адаптированная для нейропроцессоров семейства NM640x, позволяющая оценить эффективность НПС с учетом эффективности каждого процессорного модуля.
2. Предложена классификация многопроцессорных структур на базе нейропроцессоров, позволяющая определить аналитические вы-

ражения оценок эффективности для каждого типа структуры. Теоретико-множественный подход, использованный для классификации вычислительных структур, может являться основой для разработки алгоритмов распараллеливания программного кода в НПС.

3. Предложены аналитические выражения оценок эффективности НПС и отдельного нейропроцессора.

ЛИТЕРАТУРА

1. Галушкин А.И., Нейрокомпьютеры. Кн.3. М: ИПРЖР. 2000.
2. НТЦ «Модуль»: сайт НТЦ «Модуль», 2011: URL: <http://www.module.ru>.
3. Комарцова Л.Г., Максимов А.В. Нейрокомпьютеры: Учеб.пособие для вузов. М.: Изд-во МГТУ им. Н.Э.Баумана, 2004. 400 с.
4. Нейроинформатика / Горбань, А.Н. [и др.]. Новосибирск: Наука. Сибирское предприятие РАН, 1998. 296 с.
5. Ручкин В.Н., Романчук В.А., Фулин В.А. Когнитология и искусственный интеллект. – Рязань : Узорочье, 2012. – 260 с.
6. Vladimir Ruchkin, Vitaliy Romanchuk, Roman Sulitsa. Clustering, Restorability and Designing Of Embedded Computer System Based On Neuroprocessors // Proceedings of the 2nd Mediterranean Conference on Embedded Computing (MECO). - Budva, Montenegro, 2013. - С.58-62.
7. Романчук В.А., Ручкин В.Н. Оценка результатов моделирования вычислительных систем на базе нейропроцессоров // Известия тульского государственного университета. Технические науки. - Тула : Издательство ТулГУ, 2013. - Вып.9. - Ч.2. - С.194-203.
8. Романчук В.А., Ручкин В.Н. Алгоритмы анализа вычислительных структур на базе нейропроцессоров // Вестник РГРТУ. – Рязань : РГРТУ, 2012. – №2. – Вып.40. – С.60–66.
9. Романчук В.А. Моделирование нейропроцессорных систем // Отраслевые аспекты технических наук : научно-практический журнал. - Москва : ИНГН, 2013. - №10(34). - С.19-24.
10. Романчук В.А., Ручкин В.Н. Разработка программного комплекса для моделирования и анализа нейропроцессорных систем обработки изображений // Цифровая обработка сигналов. – Рязань : Информационные технологии, 2010. – №1. – С.53–58.

V.A.Romanchuk

**OPTIMIZATION OF SOFTWARE AND HARDWARE COMPLEX
COMPUTING SYSTEM BASED ON NEUROPROCESSORS**

Ryazan State University named for S. A. Yesenin

The problems of the analysis and optimization of computer systems based on neuroprocessors to increase their productivity for a given class of algorithms. Shows the adaptation of the general methods of analysis and optimization of multi-family units for neuroprocessors NM640x. The structures neuroprocessor system and obtained estimates of performance for each structure.

Keywords: neuroprocessor system optimization, neuroprocessor.