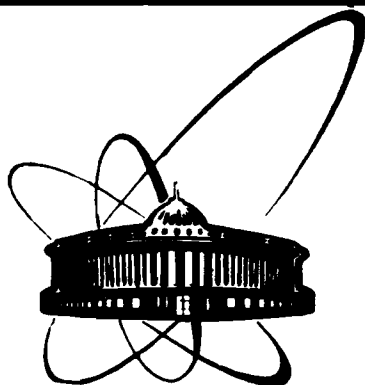


89-691



ОБЪЕДИНЕННЫЙ
ИНСТИТУТ
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ
ДУБНА

К 615

P11-89-691

И. Ф. Колпаков

МНОГОПРОЦЕССОРНЫЕ СИСТЕМЫ
ОТБОРА, РЕГИСТРАЦИИ,
ОБРАБОТКИ И АНАЛИЗА ДАННЫХ
В ФИЗИКЕ ЭЛЕМЕНТАРНЫХ ЧАСТИЦ

Направлено в Оргкомитет V Международного симпозиума
"Автоматизация и научное приборостроение (АНП-89)",
Варна, НРБ, октябрь 1989 г.

1989

1. Введение

Для обработки и регистрации экспериментальных данных в физике элементарных частиц используются все известные современные параллельные методы: конвейерная обработка, эмуляция и многопроцессорные системы. По принципу конвейера построена многоуровневая система триггеров в системах регистрации данных, эмуляторы используются на этапах отбора и обработки данных, многопроцессорные системы применяются на стадиях регистрации, обработки, моделирования и анализа данных. Далее будут рассмотрены способы построения и применения многопроцессорных систем, в которых, в основном, используются шина и модули VME. Остальные методы параллельной обработки будут продемонстрированы иллюстративно, как части систем.

2. Структура многопроцессорных систем

Элементами многопроцессорной системы являются узлы N и связи между ними. Под узлом понимается процессорный элемент, память или их комбинация. Оптимальная структура многопроцессорных систем обеспечивает минимальные задержки взаимодействия между узлами при минимальном числе связей L . Задача выбора структурной схемы сводится к нахождению оптимума структурной функции $S_N(L, T)$, а точнее, к поиску структуры, обеспечивающей приближение к этому оптимуму. Структурная функция достигает максимума (то есть выбранная структурная схема является оптимальной), чем меньше число используемых связей и короче задержка сообщения.

$$S_{N_{\max}} = \frac{1}{L \cdot T} \quad (1)$$

Можно условно разделить возможные структуры на три основных типа: пространственные, временные и пространственно-временные, или коммутационные. Примеры этих структур показаны на рис. 1.

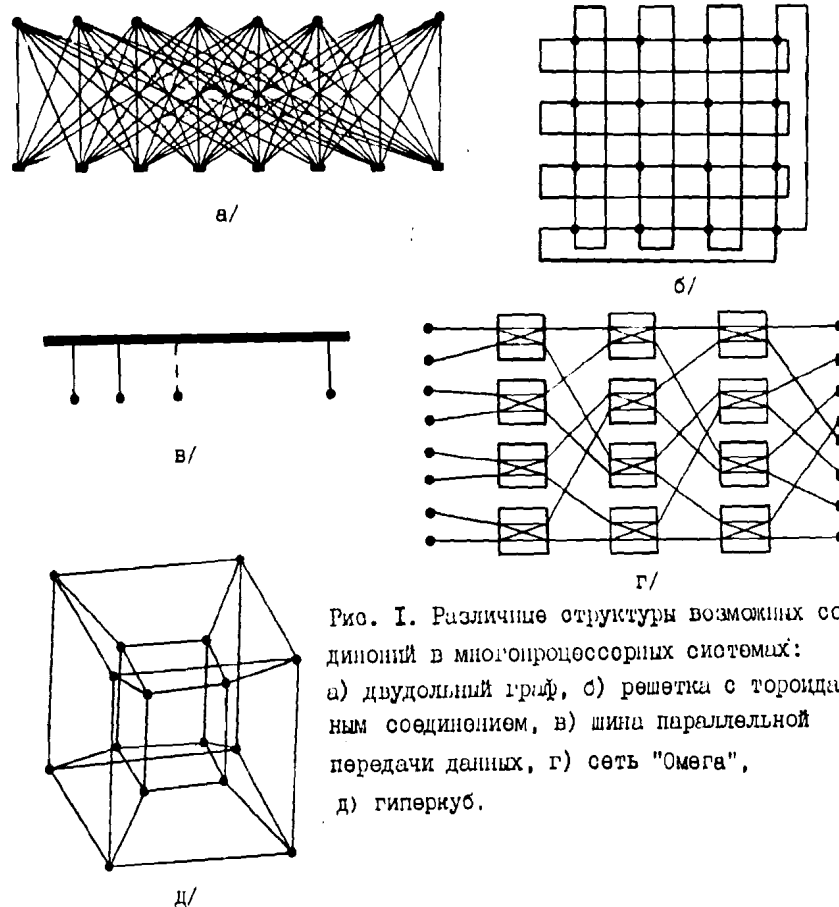


Рис. 1. Различные структуры возможных соединений в многопроцессорных системах: а) двудольный граф, б) решетки с торoidalным соединением, в) шина параллельной передачи данных, г) сеть "Омега", д) гиперкуб.

Структура, изображенная на рис. 1а, — так называемый двудольный граф, обеспечивает двустороннюю связь каждого из узлов с каждым с минимальной временной задержкой, фактически единичной, однако число связей $\frac{L \cdot N^2}{2}$ и в многопроцессорной системе является трудно реализуемым практически. Например, для двудольного графа в некотором нулевом приближении можно считать, что

$$S_N \leq \frac{1}{N^2} \quad (2)$$

Каждая связь требует определенных затрат на подсоединение интерфейса и т.д. Стремление уменьшить число связей неизбежно

приводит к увеличению задержки обращения к узлам. Задача оптимизации структурной схемы сводится, таким образом, к поиску структур, уходящих от зависимости числа связей N^2 , как в самом неэкономичном двудольном графе, и в то же время не очень увеличивающей задержку обращения. Простейшей временной структурой является кольцо, из-за очевидности не сопровождаемое рисунком. В кольце из N узлов требуются только $L=N$ связей, однако для связи двух произвольных узлов может понадобиться до $\frac{N}{2}$ временных циклов:

$$T \leq \frac{N}{2}. \quad (3)$$

Структурная функция такой системы будет описываться выражением:

$$S_N = \frac{2}{N^2}. \quad (4)$$

В структурной схеме типа тороидальной решетки (см. рис.1б) для обращения к любому узлу необходимо пройти два кольца, объединяющих \sqrt{N} узлов, и максимальная временная задержка составит $2\sqrt{N}$, поэтому эта схема нашла широкое применение в системах типа ММД (множество команд - множество данных). В структурную схему решетки хорошо вписываются многоходовые процессорные элементы, в частности, типа транспьютеров.

Типичной и наиболее широко используемой является структура типа шины (см. рис.1в). Топологически это соединение описывается как звезда. Оно требует минимального числа связей, но для обращения к произвольному узлу в общем случае необходимо N временных циклов. Однако для схем обработки данных типа SIMD (одна команда - множество данных) в случаях, когда на установление связей между узлами (засылку команд) требуется время, малое по сравнению с временем, затрачиваемым на исполнение самих команд, такая структура является оптимальной.

Пространственно-временные системы представляют собой некоторый компромисс с целью оптимизации структурной схемы, позволяя

снизить число связей между узлами по сравнению с двудольным графом, показанным на рис. 1а, при сравнительно небольшом увеличении задержек. Типичным примером является сеть "Омега", изображенная на рис.1г. Каждый узел в такой сети связывается с любым другим узлом напрямую через коммутирующий элемент, содержащий по два входа и выхода. Топологически соединенные через коммутирующие на два выхода элементы эквивалентны известной схеме дерева. Коммутирующий элемент можно представить как многоходовый процессорный узел, в частности, например, 4-канальный транспьютер. В целом значительное сокращение числа связей здесь достигается некоторым допустимым увеличением числа узлов, задержка растет как

$$T = \ell + 1, \quad (5)$$

здесь ℓ - число слоёв, равное $\log_2 \frac{n}{2}$, где n - число соединяемых узлов. Полное число узлов N в системе с учетом находящихся в промежуточных слоях:

$$N = n + \sqrt{\frac{n}{2}} (\log_2 n - 1). \quad (6)$$

Структура гиперкуба (см. рис.1д) позволяет размещать процессорные элементы с таким же числом связей, как и сеть "Омега", и с минимальной задержкой. Поэтому она получила относительно широкое распространение в многопроцессорных компьютерах. Гиперкуб представляет собой многомерный куб, в вершинах которого размещаются узлы, а грани являются связями. Число узлов, которое может содержать p -мерный гиперкуб, равно:

$$N = 2^p, \quad (7)$$

а максимальная задержка сообщения составляет

$$T = \log_2 p. \quad (8)$$

Показанные выше структуры (рис. 1а+1д) в настоящее время нашли применение в зависимости от типа решаемых задач в области физики элементарных частиц. В перечень этих задач входят предварительный отбор событий в спектрометрах с помощью так называемых

Примером использования новых принципов построения систем отбора является триггер эксперимента ZEUS /3/, где применена сеть "Омега" (см. рис. 1г), выполненная на транспьютерах.

4. Многопроцессорные системы для обработки данных и задач моделирования

Системы обработки данных в настоящее время развиваются с целью достижения производительности в единицы Гоп/с и могут содержать до 1000 процессорных узлов, каждый из которых эквивалентен VAX II/780 при стоимости одного узла около 300 долларов^{/4/}. Первая система обработки данных АСР объединяла несколько десятков процессорных плат VME, выполненных на основе микропроцессора MC 68020, с помощью шин крейтов VME и шины специально разработанной 32-разрядной ветви со скоростью обмена 20 Мбайт/с. Структура объединяла равноправные узлы, работающие под управлением одного ведомого - хосткомпьютера. В системе АСР второго поколения используется более высокопроизводительный процессор и новая структурная организация. В качестве процессорного узла применяется модуль VME на основе микропроцессора R2000 CPU фирмы MIPS с тактовой частотой 16 МГц и сопроцессором с плавающей запятой R2010 и динамической памятью 8 Мбайт. Модуль поддерживается операционной системой UNIX. Структурная схема многопроцессорной системы основана на оригинальном коммутаторе ветвей (см. рис. 4). В коммутаторе используется конструктив крейта VME и донная плата с 16 разъемами

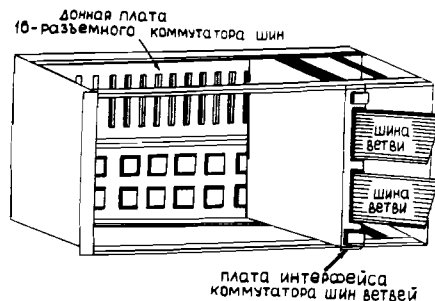


Рис. 4. Коммутатор ветвей системы АСР.

для переключения ветвей. Коммутатор обеспечивает кроссовые соединения любых ветвей числом до 16. Например, любые 8 ветвей могут быть связаны с 8 другими, при этом скорости передачи суммируются, и обеспечивается пропускная способность 8x20 Мбайт/с, то есть 160 Мбайт/с.

Система имеет однородный характер, то есть любой узел системы АСР может быть ведущим шины ветви и читать или записывать в любой адрес шины ветви. Это позволяет, в отличие от системы АСР I-го поколения, связываться любому процессору с любым другим без обращения к хосту.

Коммутатор выполнен на кроссовой микросхеме TI 74AS8840, представляющей собой 4-разрядный переключатель на 16x16 направлений. На донной плате размещено 14 таких микросхем.

В многопроцессорной системе обработки данных 2-го поколения, пример структурной схемы которой показан на рис. 5а, любой узел способен выполнять функции, выполняемые в системе I-го поколения хост-компьютером микро-VAX. В частности, любой узел может заслать, принять, сделать общее обращение и накапливать команды от любого узла и наоборот. Он может также установить все узлы в данном классе или ряду.

Распределенная схема арбитража позволяет иметь на одной ветви до 16 ведущих, что делает систему независимой от компьютера-хоста. Как видно из структурной схемы, каждый крейт имеет два модуля, организующие взаимодействие с ветвью - интерфейс ветви VME и контроллер ветви VME.

Структура процесса обработки данных показана на рис. 5б. Системная программа автоматически находит подходящий для пользователя узел, при этом пользователь не обязан знать особенности структурной схемы системы. Любой узел также имеет доступ к вводу-выводу, чтению или записи на магнитные ленты и к дисковым накопителям. Например, узлы I-го ряда могут считывать события с магнитных лент и посылать их к узлам класса 2 или 3, которые могут обрабатывать

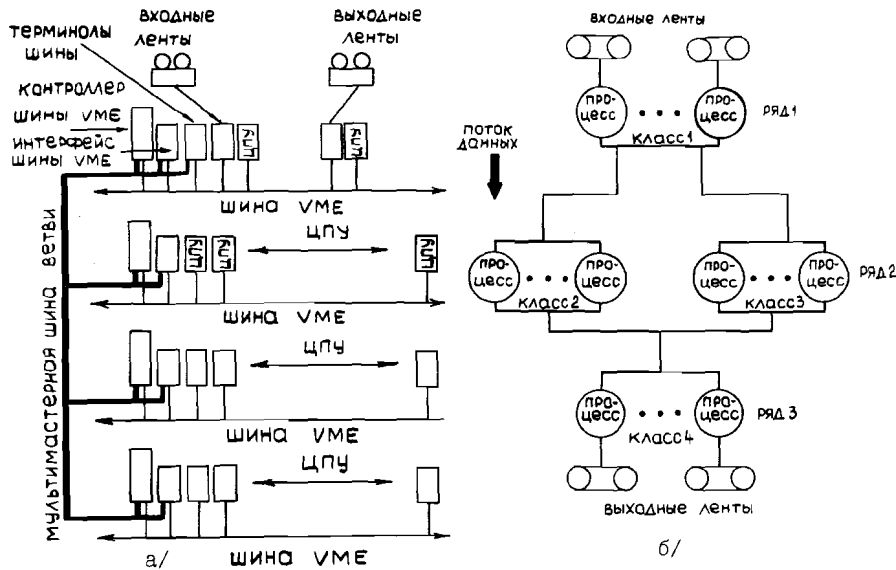


Рис. 5. Структурная схема многопроцессорной системы обработки данных 2-го поколения АСР: а) аппаратура, б) процесс обработки данных.

события от триггеров различного типа. Узлы ряда 3 могут собирать события от любых узлов в ряду 2, как в классе 2, так и в классе 3 для выдачи на магнитную ленту. Такая организация системы позволяет осуществлять множественный ввод/вывод информации параллельно на все узлы, что, в принципе, делает возможным анализ данных, в частности, в реальном времени, или производить одновременное считывание данных с множества магнитных лент. В целом применение подобных многопроцессорных систем позволяет практически полностью решить задачу реконструкции событий экспериментов.

Другое интересное направление в использовании многопроцессорных структур – анализ теоретических задач.

5. Многопроцессорные системы в задачах теоретической физики

В такой области физики, как квантовая хромодинамика, точнее в ее разделе, называемом "Калибровочные теории на решетке", требуются

многие тысячи счетных часов на самых современных суперкомпьютерах типа Cray. В расчетах, согласно калибровочным теориям, возможна такая дискретизация систем дифференциальных уравнений, которая позволяет применить параллельные многопроцессорные системы. Характер задачи позволяет приписать определенный набор точек в пространстве или пространстве-времени каждому узлу, который хранит данные для переменных поля, определенных в этих точках, в своей локальной памяти, и производит вычисления для этих точек. В лаборатории им. Э.Ферми создана высокопроизводительная система для расчетов по КХД, основанная на многопроцессорной параллельной системе, организованной структурно в виде гиперкуба^{5,6/}. Структурная схема системы показана на рис.6. Аппаратурная часть системы основана на коммутаторе шин ветвей, описанном ранее (см. рис.4), и высокопроизводительных процессорных узлах – модулях VME. Каждый узел представляет собой процессор с плавающей запятой на 32-рядном микропроцессоре конвейерного типа **Weitek XL** с производительностью 20 Мопз/с с памятью данных 8 Мбайт и памятью программы 2 Мбайт. Микропроцессор **XL** программируется на ФОРТРАН'e и СИ. Производительность процессора позволяет применить все существующие алгоритмы для моделирования в КХД, исключая внутренние фермионные петли. Несомненно, что и это станет возможным в ближайшем будущем. Узлы размещаются в сети коммутирующих крейтов со скоростью обмена 20 Мбайт/с в каждом соединении. Для создаваемого компьютера с 256 узлами это эквивалентно пропускной способности 2,56 Гбайт/с. В принципе структура позволяет использовать до 2048 узлов. В прототипе, показанном на рис.6, каждый из 2 примененных коммутирующих крейтов содержит 8 узлов и 8 интерфейсных связей, что как раз достаточно для образования гиперкуба 2^4 , где требуется 32 связи (см. рис.1д). Все узлы работают асинхронно. Используется структура команд типа MIMD, что допускает произвольные, независимые от аппаратуры, размер и форму решетки, а также позволяет системе работать, например, с меньшим числом узлов, независимо от программы.

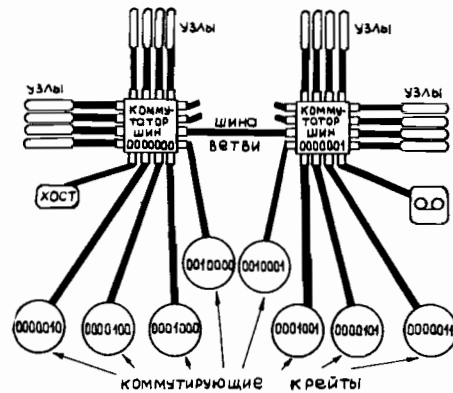


Рис. 6. Структурная схема многоматричной процессорной системы на 256 узлов, для задач КХД.

Для компьютера разработан специальный пакет программ **CANOPY**. Скелет всех программ пакета написан на ФОРТРАН'е или СИ, с использованием специальных вызовов подпрограмм, что позволяет их применять теоретикам, специализирующимся в КХД. Пакет автоматически привязывает точки решетки к узлам, и пользователю не требуется знать аппаратную часть системы.

Коммутирующий крейт организует доступ любого узла к памяти данных любого другого узла, без уточнения, где в сети расположен этот узел. Система с 256 узлами при стоимости около 750 тыс.долларов будет иметь производительность до 5 Гопз/с.

Имеется ряд других разработок параллельных многопроцессорных систем, ориентированных на задачи КХД. Одно из интересных решений – разработанный в центре Цукуба (Япония) для задач КХД компьютер **QSDPAK** /7/. Его структурная схема основана на решетке с тороидальным обращением (см. рис.10). Выбор такой структурной схемы, сложной аппаратно, позволяет сократить задержки обращения в системе, как говорилось выше. Каждый отдельный процессор имеет соединения с ближайшими четырьмя соседними (см.рис.7). Структура команд компьютера – МПМД.Обоснование предпочтительности выбора

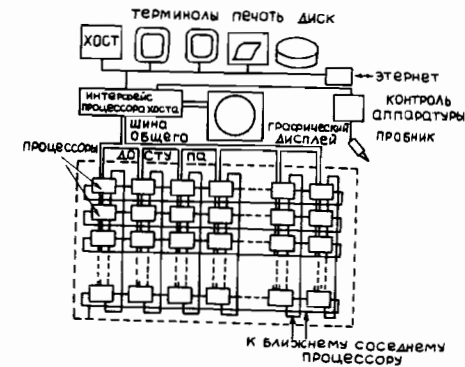


Рис. 7. Структурная схема с тороидальной решеткой параллельного компьютера для задач КХД.

этой структуры по сравнению с **SIMD** приводилось выше. Авторы считают, что увеличение эффективности системы на 7%, достигаемое при использовании структуры гиперкуба, не компенсируется возрастанием аппаратной сложности.

В качестве процессорного узла используется микропроцессор **MC68020** с динамической памятью 8 Мбайт для команд и долго хранимых данных и 2 Мбайт – для данных, которые используются часто. Увеличение производительности узла достигнуто использованием защищенных программ для исполнения основных математических функций на полуказанных СБИС. Цель разработчиков системы – создание компьютера, содержащего 10^6 узлов с общей производительностью 10^6 Топз/с, и могущей решать практически все задачи КХД.

Заключение

Как правило, структура параллельных систем выбирается оптимальной для данного класса задач. Многопроцессорные параллельные системы позволят удовлетворить практически все потребности по обработке данных, моделированию и теоретическому анализу. Их применение в области триггерных систем и систем регистрации данных в случае поиска новых структурных решений также позволяет надеяться на быстрые изменения в этих областях в ближайшем будущем.

Л и т е р а т у р а

1. Джеффри К.Фокс, Пол К.Мессина. Архитектуры компьютеров. - В мире науки. - № 12, 1987, с.17-25.
2. Abolins M., Edmunds D., Linnemann J. DO Trigger Framework. DO Note 328, Michigan State University, Feb. 24, 1986.
3. Nash T. Computing Possibilities in the Mid 1990's. Fermilab-Conf-88/117, Batavia, FNAL, September 1988.
4. Gaines Irwin. Multi-Processor Developments in the United States for Future High Energy Physics Experiments and Accelerators. Fermilab-Conf-88/211, Batavia, March 1988.
5. Nash T., Areti H., Atac R. The Fermilab Advanced Computer Programm Multi-Array Processor System (ACMAPS) a Site Oriented Supercomputer for Theoretical Physics. FNAL, Batavia, FERMILAB-Conf-88/111, August 1988.
6. Fischler M., Atac R., Cook A. et al. The Fermilab Lattice Supercomputer Project, Fermilab-Conf-89/37, Batavia, February 1989.
7. Iwasaki Y. et al. QCDPAX: A Parallel Computer for Lattice QCD Simulation. Computer Physics Communications, v.49, 1988, p. 449-455.

Рукопись поступила в издательский отдел
29 сентября 1989 года.

Колпаков И.Ф. P11-89-691
Многопроцессорные системы отбора,
регистрации и анализа данных
в физике элементарных частиц

Рассматриваются структурные схемы организации многопроцессорных систем. Дан обзор основных применений этих систем для задач сбора, обработки и анализа данных в физике элементарных частиц. Как правило, выбираются специализированные структуры параллельных систем, оптимальные для данного класса задач. Приведены примеры современных многопроцессорных систем, используемых в этой области.

Работа выполнена в Лаборатории высоких энергий ОИЯИ.

Препринт Объединенного института ядерных исследований. Дубна 1989

Перевод Л.Н.Барабаш

Kolpakov I.F. P11-89-691
Multiprocessor Systems for Data Acquisition,
Handling and Analysis in Elementary
Particles Physics

Architecture of multiprocessor systems has been considered. A review of principle applications of these systems for acquisition, handling and analysis of data is presented. As a rule, specialized structures of parallel systems are chosen, which are optimal for the certain range of tasks. Examples of modern multiprocessor systems used in this field are presented.

The investigation has been performed at the Laboratory of High Energies, JINR.

Preprint of the Joint Institute for Nuclear Research. Dubna 1989