



(19) **RU** ⁽¹¹⁾ **2 084 953** ⁽¹³⁾ **C1**
 (51) МПК⁶ **G 06 F 15/16**

РОССИЙСКОЕ АГЕНТСТВО
 ПО ПАТЕНТАМ И ТОВАРНЫМ ЗНАКАМ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ РОССИЙСКОЙ ФЕДЕРАЦИИ

(21), (22) Заявка: 5010148/09, 12.11.1991

(30) Приоритет: 13.11.1990 US 611594

(46) Дата публикации: 20.07.1997

(56) Ссылки: Патент Японии N 60-144867, кл. G 06 F 15/16, 1985. Патент Великобритании N 2223867, кл. G 06 F 13/12, 1990.

(71) Заявитель:

Интернэшнл Бизнес Машинз Корпорейшн (US)

(72) Изобретатель: Джеймс Уоррен

Диффендерфер[US],

Питер Майкл Когге[US], Пол Амба

Уилкинсон[US], Николас Джером Шуновер[US]

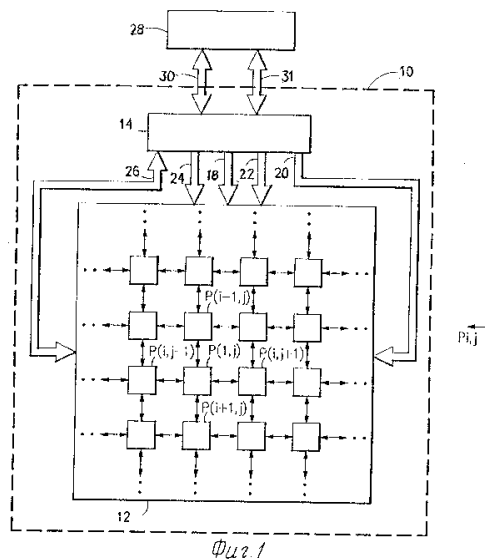
(73) Патентообладатель:

Интернэшнл Бизнес Машинз Корпорейшн (US)

(54) ПАРАЛЛЕЛЬНАЯ ПРОЦЕССОРНАЯ СИСТЕМА

(57) Реферат:

Параллельная процессорная система используется для параллельной матричной обработки с упрощенной архитектурой, приспособленной для реализации на чипах и работы в среде с воздушным охлаждением. Система представляет собой N размерную матрицу из байтовой ширины процессорных устройств, каждое из которых сопряжено с адекватным сегментом байтовой ширины, памятью и управляющей логикой. Разделяемая секция матрицы, содержащая несколько процессорных устройств, размещена на кремниевом чипе, оснащенном пикетами, причем элемент процессорной матрицы предпочтительно состоит из комбинации процессорного элемента и локальной памяти для обработки байтов параллельных битов информации за тактовый цикл. 5 з.п. ф-лы, 7 ил.





(19) **RU** ⁽¹¹⁾ **2 084 953** ⁽¹³⁾ **C1**
 (51) Int. Cl.⁶ **G 06 F 15/16**

RUSSIAN AGENCY
 FOR PATENTS AND TRADEMARKS

(12) **ABSTRACT OF INVENTION**

(21), (22) Application: 5010148/09, 12.11.1991

(30) Priority: 13.11.1990 US 611594

(46) Date of publication: 20.07.1997

(71) Applicant:
 Internehshnl Biznes Mashinz Korporejshn (US)

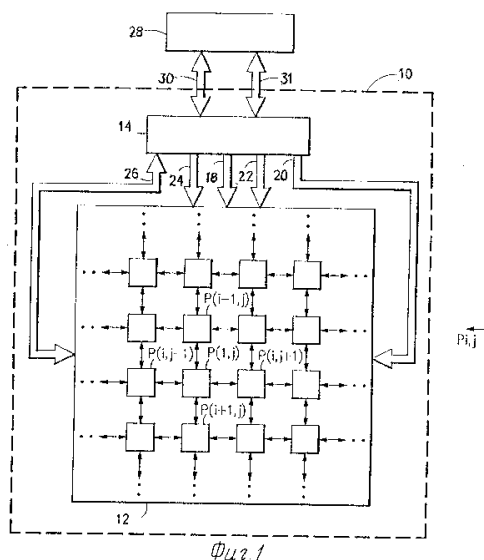
(72) Inventor: Dzhejms Uorren Diffenderfer[US],
 Piter Majkl Kogge[US], Pol Amba
 Uilkinson[US], Nikolas Dzherom Shunover[US]

(73) Proprietor:
 Internehshnl Biznes Mashinz Korporejshn (US)

(54) **PARALLEL PROCESSOR SYSTEM**

(57) Abstract:

FIELD: computer engineering, in particular, parallel array processing.
 SUBSTANCE: device is designed as N-dimensional matrix of byte-length processing units each of which is connected to corresponding segment of byte-length memory unit and controller. Shared part of matrix has several process units and is located on silicon chip which has pickets. Each element of process matrix is designed as processor element and local memory for processing bytes with parallel information bits for single clock pulse. EFFECT: simplified design, possibility of chip implementation, possibility of operations in air-cooled environment. 6 cl, 7 dwg



RU 2 084 953 C1

RU 2 084 953 C1

Изобретение относится в основном к мультипроцессорным параллельным вычислительным системам и, в частности, к параллельной архитектуре на базе байтовых последовательных SIMD процессоров и к системе и методам параллельной матричной обработки с упрощенной архитектурой, предназначенной для чиповой реализации в среде с воздушным охлаждением.

Стало нормой описывать большинство современных SIMD машин как синхронные битовые последовательные процессоры, организованные в виде N и X матрицы, проиллюстрированной примерной мультипроцессорной системой, описанной в работе Г.Ли и М. Мареска, озаглавленной "Полиморфично-торная сеть", Труды Международной конференции по проблемам параллельной обработки, с. 411-414. Способы матричного векторного умножения для такой в основном параллельной архитектуры, описанной Г.Ли и др. из NxH процессоров, соединенных физически сотовой технологией, причем сотовая сеть перекрыта другой переключательной сетью для целей реконфигурации, была подобно описана в Бюллетене технических открытий IBM, т. 32, N 3A, август 1989, для повышения скорости выполнения умножения разрешенной матрицы на векторе.

Имеется публикация 1990 г, которая показывает, что уделяется внимание созданию SIMD и SIMD/MIMD машин, которые работают с многострочными процессорами, в которых все процессоры в данной строке выполняют одну и ту же команду, но каждая строка программируется различно. См. например, приведенный Г. Ли, Б.Дом и Г. Джаффом в Бюллетене технических открытий, т. 3, 8В, январь 1990, в котором описана архитектура с конфигурацией, известной под названием "параллельный локальный операторный двигатель".

Заявка 07/519.332 Джеймса Л.Тейлора, зарегистрированная 4 мая 1990, продолжающая заявку США N 07/250.595, зарегистрированную 27 сентября 1988 под названием "SIMD матричный процессор", описывает многомерную матрицу из процессорных элементов, которая обладает повышенной степенью гибкости для обеспечения возможности лучшего использования параллельной обработки без использования дорогого и сложного MIMD процессора. Описанная в ней система посылает глобальную команду для локального битового последовательного исполнения по управляющей соединительной шине логики на различные параллельные процессорные элементы и программно модифицирует выбранные биты глобальной команды для использования на локальных битовых линиях, где модифицированные биты декодируются.

Известно устройство "Проектирование SIMD микропроцессорной матрицы", авторы К. Р. Джесшор, Р.О.Герман и др. Труды IEEE, т. 136, май 1989. В работе описаны результаты исследований SIMD архитектур, которые авторы провели на факультете Электроники и вычислительной техники в университете, Сауфамптон, Великобритания. В данной работе описан процессор, обладающий битовой ширины SIMD архитектурой. Можно заметить, что

предложенная структура не обеспечивает байтовой ширины соседской связи, что обеспечивает настоящее изобретение.

Известна "Мультипроцессорная система передачи данных", описанная в патенте GB N 2.223.867, кл. G 06 F 13/12, опублик. 1988, содержащая матрицу процессорных элементов. Данная система является наиболее близким аналогом предлагаемой.

Настоящая заявка предполагает развитие систем с воздушным охлаждением, которые могут быть изготовлены по существующей технологии, которая может выполняться на чиповом наборе с заново разработанной архитектурой задачи, которые ранее предназначались для SIMD машин, а также, которые ранее предполагалось решать на фон-неймановских машинах.

Для облегчения понимания описания ниже даны общие, неограничивающие определения новой терминологии, принятой в настоящей области техники:

пикет элемент процессорной матрицы, предпочтительно состоящий из обрабатывающего (процессорного) элемента, скомбинированного с локальной памятью, для обработки байтов информации из параллельных байтов в тактовом цикле. Предпочтительный вариант осуществления состоит из процессора байтовой ширины потока данных, 32 К байтов или более памяти, примитивных органов управления, связей для коммуникации с другими пикетами.

пикетный чип содержит множество пикетов на одиночном кремниевом чипе.

пикетная процессорная система (или подсистема) полная система, состоящая из матрицы пикетов, связанной сети, системы ввода-вывода и SIMD контроллера, состоящего из микропроцессора, процессора с стандартным программным обеспечением и микроконтроллера, который управляет этой матрицей.

пикетная архитектура предпочтительный вариант осуществления SIMD архитектуры с признаками, которые удовлетворяют нескольким разнообразным проблемам, включающим:

ассоциативную обработку наборов, интенсивную параллельную числовую обработку, обработку физических матриц, подобных изобретениям.

Пикетная матрица совокупность пикетов, организованных в некотором геометрическом порядке.2 Архитектура, которая будет пояснена ниже, может быть названа "пикетной" архитектурой. Она включает в себя определенные структуры из процессорных элементов, в некотором смысле подобные известным SIMD процессорным элементам (PE), которые при необходимости работают совместно как пикетная линия, однако достаточно усовершенствованы и способны выполнять независимые фон-неймановские операции. Такая архитектура была создана для использования N-размерной матрицы из процессорных устройств, организованных для параллельной обработки и обладающих управляющей логикой (внутри каждого процессорного устройства, которое мы называем пакетом).

Предпочтительная система способна вести вычисления как SIMD (один поток команд несколько потоков данных) потоковая

машина, однако исполнение ее улучшено за счет большей степени параллелизма процессорных элементов. Проблемы, порождаемые зависимостью данных, были устранены. В ходе операций SIMD машины ни один процессор или функция не может иметь такую зависимость от данных, которая заставляла бы один процессорный элемент требовать некоторое отличное число циклов.

Архитектурные усовершенствования, проделанные нами, привели к появлению системы с множеством пикетных устройств или пикетов, причем локальная память соединена с процессорным элементом для параллельной обработки информации ассоциативным способом, где каждый пикет приспособлен для выполнения одного элемента ассоциативного процессора. Заявитель предусмотрел путь для горизонтальной ассоциации с каждым пикетом. Память пикетных устройств организована в виде матрицы. Матрица пикетов, таким образом организованная, включает в себя наборную ассоциативную память.

Изобретение наборной ассоциативной параллельно обрабатывающей системы на одиночном чипе позволяет выбирать из большего выбора, хранящегося в памяти, меньший выбор данных туда, где над ним может быть выполнена некоторая ассоциативная операция. Такая ассоциативная операция, которой обычно бывает сравнение на совпадение, исполняется над целым набором данных параллельно с использованием памяти и исполнительного устройства.

В пикетной матрице каждый пикет получает порцию данных из большего набора. Дополнительно каждый пикет выбирает одну часть данных из своей порции. Таким образом одна часть данных в каждом наборе пикетов представляет собой набор данных, над которыми любая ассоциативная операция выполняется всеми пикетами параллельно.

Эта конструкция может быть реализована в настоящее время с 50 К вентилями управляемой потоком данных логики и с многомегабитной DRAM произвольного доступа памятью, на одном чипе. Каждый из этих чипов сконфигурирован так, что содержит некоторое множество пикетов или процессорных устройств. В предпочтительном варианте осуществления, предназначенном для обработки текстов, а также пригодном для работы с графикой, имеется 16 пикетов, причем каждый пикет обладает 32 килобайтами DRAM произвольного доступа памяти, и система содержит матрицу из 64 таких пикетных чипов, образуя матрицу из 1024 процессорных элементов. Такая пикетная архитектура может быть изготовлена по CMOS технологии, которая позволяет разместить 4 миллиона битов DRAM произвольного доступа памяти повторяющимися структурами на одном чипе, и оставшаяся площадь поверхности чипа заполняется стандартными логическими ячейками, которые используют плотность до 50К логических элементов и которые могут образовывать управляемую потоком данных логику, чтобы образовывать пикеты на чипе. Заявитель так организовал систему, что пикеты могут обрабатывать данные в режиме локальной автономии, и предусмотрел

"проход" между пикетами.

Пикетная технология допускает расширение, и при наличии 128К байтов DRAM памяти в каждом отдельном пикете (16 Мбитовой DRAM произвольной памяти чип) пикетная архитектура способна обрабатывать 24-битовые цветные графические изображения точно так, как текст и 8-битовые цветные или полутоновые графические изображения обрабатываются предпочтительным вариантом осуществления настоящего изобретения. Экспериментальная производственная технология показывает, что такая плотность станет достижимой в ближайшем будущем в виде изготовленного как единое целое продукта, способного работать в среде с воздушным охлаждением. Что касается цветной графики, заявленная предпочтительная пикетная архитектура повысит количество DRAM на чипе до 128 килобайтов на пикете, однако при сохранении 16 пикетов на чипе. Альтернативно, 24 пикетных устройства на пикетный чип при 96 килобайтной памяти можно будет использовать для цветных графических процессоров.

На фиг.1 приведена схематическая иллюстрация известного SIMD процессора, который описан в одновременно рассматриваемой заявке, которая может рассматриваться как отражающая известный уровень техники; на фиг.2 пара базовых пикетных устройств, которые сконфигурированы на кремниевой базе с процессором памятью, управляющей логикой и ассоциативной памятью с байтовой связью и другими пикетами матрицы; на фиг.3 показана обработка ассоциативной памятью; на фиг.4 базовая 16 (n) пикетная конфигурация для SIMD подсистемы, которая использует микропроцессорный контроллер, аппаратный последовательный контроллер для программы стандартного обеспечения и пикетную матрицу и которая образует базовую параллельную пикетную процессорную систему, которая также может быть самостоятельным устройством; на фиг.5 - мультипикетная процессорная система, которая включает в себя некоторое множество пикетных процессоров, представленных на фиг.4; на фиг.6 диаграмма подсистемного функционального блока; на фиг.7 схема подсистемного контроллера с картами, представленными на фиг.5.

Фиг.1 представляет типичную известную SIMD систему, тип которой описан в поданной Тейлором на IBM европейской патентной заявке N88307855/88-A и в заявке Великобритании UK-A-1.445.714. В подобных известных устройствах SIMD представлял собой "одна команда множество данных" компьютер, имеющий параллельный матричный процессор, содержащий некоторое множество параллельно соединенных последовательных процессоров, каждый из которых ассоциирован с одним из некоторого множества SIMD запоминающих устройств. Система ввода-вывода действует как каскадная система относительно SIMD устройства и содержит временную память для двунаправленного двухразмерного обмена данных между главным компьютером (которым может быть большая ЭВМ или микропроцессор) и SIMD компьютером. Система ввода-вывода включает в себя

обработывающее ввод-вывод устройство для управления потоком данных между главным компьютером и временного хранения устройством и для управления потоком данных между временным хранилищем и некоторым множеством SIMD запоминающих устройств, которые обычно являются буферными секциями или частями некоторой большей памяти. Таким образом, операция ввода системы ввода-вывода включает в себя посылку данных от памяти главного компьютера на временное хранилище и от временного хранилища на SIMD запоминающее устройства на втором шаге, и для вывода также определен двухшаговый процесс, который пересылает по двухразмерной шине данные между главным компьютером и SIMD компьютером. Система ввода-вывода для передач ввода-вывода может представлять собой отдельное устройство, подустройство в главном компьютере или некоторое устройство внутри SIMD компьютера, в котором SIMD контроллер осуществляет управление буферным временным хранилищем (накопителем) ввода-вывода.

Сам SIMD компьютер содержит процессорную матрицу, имеющую некоторое множество процессорных устройств и сеть, которая соединяет отдельные процессорные устройства и некоторое множество известных отдельных SIMD запоминающих устройств. SIMD процессорное устройство представляет собой параллельный матричный процессор, имеющий большое число отдельных процессорных элементов, соединенных и работающих параллельно. SIMD компьютер включает в себя управляющее устройство, которое генерирует поток команд для процессорных элементов и также необходимые для компьютера синхронизирующие сигналы. Сеть, объединяющая разнообразные процессорные элементы, включает в себя соединительную сеть некоторого вида для отдельных процессорных устройств, имеет вид сетки, полиморфичного тора или гиперкуба. Множество запоминающих устройств служит для временного хранения битовых данных для отдельных процессорных устройств, и здесь существует взаимоднозначное соответствие между числом процессорных элементов и числом запоминающих устройств, которые могут быть названной выше буферной частью памяти большего объема.

Например (см. фиг.1), имеется главный процессор 28. Этот процессор используется для загрузки микрокодовых программ в матричный контроллер 14 (который имеет буфер временного хранения) для обмена с ним данными и для слежения за его состоянием по шине данных "главный контроллер" 30 и шине "адрес и управление" 31. В этом примере главным компьютером может быть любого типа компьютер общего назначения, подобный большой ЭВМ или персональному компьютеру. В примере, относящемся к известному уровню техники, матрица процессоров представлена двухразмерной (2-D), однако матрицу можно организовать по-иному, по 3-D или 4-D кластерной схеме. SIMD матричное устройство 10 имеет матрицу 12 из процессорных элементов $P(i,j)$ и матричный

контроллер 14 для выдачи потока глобальных команд на процессорные элементы $P(i,j)$. Хотя на Фиг.1 не показано, также имеются процессорные элементы, которые обрабатывают каждый один бит в текущий момент, и имеется блок для хранения, который является частью памяти, ассоциированной с процессорным элементом. Процессорные элементы соединены так называемой NEWS (N север, E восток, W запад, S юг) сетью с их соответствующими соседями двунаправленными битовыми линиями. Таким образом, процессорный элемент $P(i,j)$ соединен с процессорными элементами $P(i-1, j)$, $P(i, j+1)$, $P(i, j-1)$ и $P(i+1, j)$ в северном, восточном, западном и южном направлениях соответственно. В этом типичном примере NEWS сеть тороидально соединена на своих концах так, что северный и южный концы взаимосоединены двунаправлено и западный и восточный концы соединены аналогичным образом. Чтобы обеспечить ввод/вывод в/из матрицы процессоров, шина данных "контроллер-матрица" 26 соединена с NEWS сетью. Как видно, она соединена с восточно-западной границей матрицы. Она может быть соединена вместо или дополнительно с северо-южной границей посредством двунаправленных трехуровневых драйверов, которые соединены с тороидальным восточно-западным-NEWS соединением. Как и в предпочтительном варианте осуществления, который будет описан ниже, число процессорных элементов в известном устройстве достигнет 1024, если вместо проиллюстрированного количества 16×16 процессорных элементов будет использовано 32×32 процессорных элементов. На иллюстрации однониточная линия обозначает однобитовую линию, а двуниточная линия, соединяющая функциональные элементы, применяется для обозначения некоторого множества соединительных линий или шины.

В примере известного уровня техники матричный контроллер выдает сигналы выбора строки и выбора столбца по линиям "выбор строки" 20 и линиям "выбор столбца" 22 соответственно. Эти команды заставляют процессорные элементы загружать данные из хранилища и обрабатывать эти данные, а затем записывать данные вновь в хранилище. Для этой цели каждый процессорный элемент имеет доступ к битовой части (секции или буферу) главной памяти. Поэтому логично, что главная память матричного процессора разделена на 1024 порционных части для матрицы из 1024 процессорных элементов. Это значит, что вплоть до 32 битовых слов можно одновременно передавать на или выбирать из хранилища за один шаг передачи. Для выполнения операции считывания или записывания память адресуется индексным адресом, который подается на линии адреса памяти по адресной шине 24 и команда считывания или записывания подается параллельно на каждый из процессорных элементов. Во время операции считывания сигналы выбора строки и столбца на линиях выбора строк и столбцов идентифицируют тот процессорный элемент, который будет выполнять эту операцию. Таким образом, в описанном примере можно считать одно 32-битовое

слово из памяти в тридцать два процессорных элемента в выбранной строке, когда матрица 32x32. Процессорный элемент ассоциирован с частью или блоком памяти (i,j) памяти шириной в один бит. Хотя часть или блок памяти логически ассоциирован взаимнооднозначным образом с ассоциированным отдельным процессорным элементом, на самом деле он может быть и обычно является физически удаленным на другой чип. Мы не знаем, как при этой известной архитектуре описанный матричный процессор можно изготовить, в то время как наш пикет можно изготовить с матрицей процессоров и достаточной памятью на одиночном чипе описываемого ниже типа.

Процессорные элементы P (i,j) предыдущего примера включают в себя арифметико-логическое устройство с входным и выходным регистрами, которые имеют разряд переноса, каждый из которых способен на один бит информации. Имеется мультиплексор, соединенный с входом и выходом арифметико-логического устройства, а также двунаправленным портом данных части памяти (i,j), ассоциированной с отдельным процессорным элементом P (i,j).

Имеются отдельные шины команд и данных, и матричный контроллер имеет микрокодвое хранилище, в которое микрокод, определяющий вид обработки, подлежащий выполнению матрицей, загружается главным компьютером 28 посредством шины данных 30 и шины адресов и управления 31. После инициализирования главным компьютером 28 работы матричного контроллера последовательность микрокода находится под управлением управляющего микрокодом устройства, которое соединено с хранилищем микрокодов внутри матричного контроллера 14. Арифметико-логическое устройство и регистровый банк матричного контроллера используются для генерирования адресов матричной памяти, подсчета циклов, вычислений адресов переходов и в операциях регистров общего назначения, которые выдаются на адресную шину матричного контроллера. Матричный контроллер также имеет масочные регистры для декодирования строчных и столбцовых масочных кодов, и коды специальных операций проходят на процессорные элементы по шине 18 команд. В этом примере матричный контроллер мог бы иметь буфер данных внутри контроллера, однако функционально между шиной данных "главный-контроллер" и шиной данных "контроллер-матрица". С этой целью буфер располагается в матричном контроллере в виде двунаправленного буфера типа "первым вошел первым уйдет". Более подробно с подобными известными системами по приведенным выше публикациям, в частности по патентной заявке США 07/519.332 "SIMD матричный процессор", поданной Тейлором.

Обзор названных выше известных попыток можно сравнить с предпочтительными вариантами осуществления изобретений, описываемых ниже. Фиг.2 представляет базовое пикетное устройство 100, содержащее комбинацию процессорного элемента арифметико-логическое устройство 101 с локальной памятью 102, соединенной с процессорным элементом, для обработки однобайтовой информации в тактовый цикл.

Пикетное устройство сформировано на чипе с кремниевой подложкой или пикетном чипе с линейной матрицей пикетов с соседями по сторонам (на чертеже слева и справа) таким образом, что на кремниевой подложке чипа сформирована пикетная процессорная матрица с множеством локальных памяти, по одной на каждый, и байтовой ширины поток обрабатываемых данных, расположенных в логическую строку или линейную матрицу с шинами связи с соседями для двунаправленного пропуска данных влево и вправо. Набор пикетов на пикетном чипе размещен в геометрическом порядке, предпочтительно горизонтально на этом чипе. Фиг.2 показывает типовое осуществление двух пикетов пикетной матрицы на пикетном чипе с множеством запоминающих устройств и связанных путей для потоков данных между каждым процессорным элементом пикета и памятью. В нашем предпочтительном варианте осуществления пути передачи данных между взаимно однозначно сопоставленными памятью и процессорными элементами матрицы имеют ширину в один байт поперек, влево или вправо от соседа или "с проходом" для связи с более дальними пикетными процессорами.

"Проход" можно определить как средство для передачи информации за один цикл на несоседнюю позицию через пикетную адресную ячейку, которая нормально была бы способна принять эту информацию, если бы не была прозрачна к передаваемому сообщению, до тех пор пока оно не прибывает и не будет принято на ближайшем активном соседе, который примет ее. Таким образом, проход функционирует путем посылки информации на несоседнюю позицию через "отключенные" пикеты. Например, пикет "А" желает передать информацию на отдаленный пикет "Q". Перед таким циклом промежуточные пикеты делаются прозрачными, для чего пикеты с "В" по "F" отключаются. Затем в следующем одиночном цикле "А" пикет посылает свое сообщение направо и, выполняя это, проходит через пикеты "В" по "F", которые прозрачны, поскольку они отключены, и пикет "С", принимает сообщение, поскольку он включен. При нормальном использовании "проходная" информация передается линейно через решетку, однако "проходной" проход может также работать с двуразмерными решетками или в многомерной матрице.

Доступ процессорных элементов в нашем предпочтительном варианте осуществления является не битовой последовательностной операцией, а байтовой последовательностной. Каждый процессор имеет доступ к сопряженной с ним памяти, а не доступ лишь к блоку локальной памяти или некоторой ассоциированной части или страницы памяти. Шина имеет ширину не один бит, а ширину символа или нескольких символов. Не один бит, а байт (или в будущих системах, предназначенных повысить коэффициент полезности символического байта, несколько байтов) информации обрабатывается за один тактовый цикл. Поэтому 8,16 или 32 бита могут двигаться между отдельными пикетами процессорного элемента, отвечая ширине ассоциированной памяти. В нашем предпочтительном варианте осуществления каждый пикетный чип имеет

память шириной 8 (9) битов объемом 32 Кбайта, предпочтительно 16 пикетов с этими 32 Кбайтами каждое хранилище на пикетный узел линейной матрицы. В нашем варианте осуществления каждая ассоциированная память встроена в CMOS в виде DRAM произвольного доступа к памяти и символичный байт составляют 9 байтов (который функционирует как 8- битовой символ с самопроверкой).

Параллельный поток данных по шине шириной в байт между пикетами и между процессорным элементом и его памятью обеспечивает существенные преимущества в сравнении с битовой последовательностной структурой известных систем, однако также следует признать, если уже считать это усовершенствованием, что повышение степени параллелизма порождает дополнительные проблемы, которые требуют решений для достижения понимания свойств вновь созданной архитектуры.

Признак, который надо понять, состоит в том, что в дополнение к передачам левому и правому соседям и проходному механизму, который был описан со ссылками на приложенные чертежи, мы дополнительно предусмотрели "широковещательную" шину шириной два байта, чтобы все пикеты могли знакомиться с одними и теми же данными одновременно. Пикетное управление и адресное распространение также происходит по этой широковещательной шине. Именно эта шина передает сравниваемые данные, когда происходит исполнение операции ассоциации наборов и другие операции сравнения или синхронных математических операций.

Задачи, обладающие высокой степенью параллельности структур данных, которые пригодны для обработки пикетными обрабатывающими элементами под управлением одного потока команд, включают приложения, касающиеся сравнения образов искусственным интеллектом, совмещения датчиков и дорожек при оптимизации присвоения нескольких датчиков, контекстном поиске и обработке изображений. Однако большинство из таких приложений, в настоящее время весьма полезных, не были реализованы SIMD процессорами по причине последовательной побитовой обработки при одноконтантной синхронизации. Например, известный последовательностный обрабатывающий (процессорный) элемент SIMD машины работает над одним битом операции И в каждом процессорном цикле, в то время как 32-битовая параллельная машина может работать на 32 бита операции И на одном цикле.

Конфигурация 32 Кбайтного процессорного элемента предоставляет больше локальной памяти каждому процессорному элементу, чем предоставляет известна SIMD машина.

Количество контактов чипа сведено до минимума, поскольку сведено до минимума количество данных, поступающих и из чипа. DRAM память представляет собой известную CMOS матрицу, которая обеспечивает "строчно-столбцовый" доступ исключением столбцового демультиплексирования на основе запоминающей памяти и выполнением строчной адресации, при которой считывание строки запоминающей матрицы для

организации потоков данных происходит параллельно.

Память, помимо данных, содержит "трехбиты" или триты, чтобы логика распознавала три состояния, а не традиционные двоичные числа, т.е. логическую "1", логический "0" или "ничего". "Ничего" в математическом смысле совпадает с логической единицей, либо с логическим нулем. Трит содержится в последовательных запоминающих ячейках запоминающей матрицы. Маски являются другой формой данных, хранящихся в памяти, которые направляются на масочный регистр пикетного процессорного элемента.

Поскольку запоминающая матрица может содержать команды, это позволяет одному пикету выполнять операцию, отличающуюся от операции, выполняемой другим пикетом. Чиповое управление отдельными пикетами в ходе операций, осуществляемое большинством пикетов, но не обязательно всеми, позволяет реализации, которые уникальны для SIMD операций. Предусмотрена одна простая функция, которая осуществляет ждущие операции в любом пикете, чем статусный выход отвечает специальному условию. Поэтому ненулевое условие может означать "дремоту". "Дремота" представляет собой условие, которое предотвращает выполнение операций и делает пикет неактивным, но чувствительным. Другой предусмотренной командой является запрет или разрешение записи в память с учетом состояний пикета или с учетом команды, подаваемой на шину до начала проходческой операции.

Если создавать на пикетном чипе 16 мощных пикетов, каждый с памятью 32 Кбайта, то лишь 64 чипа предоставят 1024 процессора и память 32768 Кбайтов. Матрица пикетов включает в себя память ассоциативных наборов. Изобретение также полезно для цифрового интенсивного анализа обрабатываемых изображений, а также обработки векторов. Эту мощную пикетную процессорную матрицу можно уже на известном уровне техники упаковать всего лишь на двух небольших платах. Можно видеть, что тысячи пикетов можно должным образом упаковать в более портативный маломощный пикет, что позволит выполнять задачи, касающиеся обработки изображений с минимальной задержкой или в течение видеорамки, например во время полета самолета, без существенных финансовых затрат.

Мощность пикетов позволяет использовать большие ассоциативной памяти системы, упакованные в плотные пакеты, и позволяет свое использование при обработке разнообразных прикладных задач после того, как проектировщик научится применять эту новую систему.

Фиг. 3 иллюстрирует то, что можно назвать полностью ассоциативной памятью, поскольку когда запрашивается ассоциация, то сравниваемое значение предъявляется всем запоминающим ячейкам и все запоминающие ячейки одновременно отвечают по своим линиям сравнения. Ассоциативная память известна на текущем уровне техники. В описываемой системе, использующей параллельные пикеты памяти и обрабатывающих элементов, которые делают

байтовые передачи при выполнении поиска, имеется выход данных и маска для поиска, чтобы локализовать слово К среди К слов, находящихся в памяти. Все сравнивающие пикеты поднимают линии состояния и затем отдельная операция считывает или выбирает первое совпадение К. Эта операция, которую принято называть "установить ассоциацию", может быть повторена для последующих слов, хранящихся в пикетной памяти. Аналогично записывание осуществляется посредством операции "широкого вещания", при которой возбужденная линия выбора указывает участие, и широковещательные данные копируются во все выбранные пикеты.

Другой вариант осуществления, который не является предпочтительным, ограничивает объем DRAM произвольного доступа к памяти, доступной для каждого пикета, что обеспечивало включение секции полностью ассоциативной памяти того типа, который проиллюстрирован на фиг.3. Если 512 байтов полностью ассоциативной памяти будут включены, то каждый пикет сможет содержать набор поисковых индексов и за одну операцию будет выполняться 512×1024 пикетов, т. е. 512 К сравнений за операцию или 512 Гиг сравнений в секунду при длительности операции одна микросекунда. При расширении эта концепция может быть перенесена в диапазон многих тера сравнений. Этот вариант осуществления делает решаемыми задачи, касающиеся ассоциации, которые предполагают расширяемый поиск информации, предоставляя возможности, которые превосходят известные вычислительные способности.

Эта ассоциативная операция, использующая память и байтовой ширины сопряжения процессорные элементы, как показано на фиг.2, помимо приложений к отдельным алгоритмам или операциям, искусственному интеллекту и параллельному программированию, относящихся в SIMD (одна команда множество потоков данных) ситуациям, открывает возможность реализации большого числа дополнительных приложений для машины с конфигурацией описанного чипа, в SIMD окружающей среде, среди которых:

простые, допускающие распараллеливание арифметические задачи, включающие матричное умножение, и другие задачи, которые можно выполнять в специализированных запоминающих машинах;

задачи, сравнения образцов и обработки образцов, которые можно решать с помощью фон-неймановских машин, но которые значительно быстрее решаемы с приложениями, приспособленными к чрезвычайно высокому параллелизму, например к шаблонному сравнению трехразмерного изображения; функции запроса по данным; образцовое сравнение в области искусственного интеллекта; сетевое управление мостами с целью быстрой идентификации сообщений, которые идут к пользователю по другой стороне моста в сети;

симуляция шлюзовых уровней, контроллеры основных VLSI правил.

Процессорные задачи, успешно решаемые с помощью банка запоминающих устройств и

ассоциативных обрабатывающих элементов, помогают прикладному программисту, когда он попытается воспользоваться мощностью этой новой системной архитектуры.

Процесс сохранения описания цифровой системы может быть улучшен путем использования матрицы для одного вентиля или логического элемента на пикет 100. В такой системе процесс может начинаться с присвоения каждому вентилю описания в виде списка сигналов, которые этот вентиль воспринимает как входы, и наименования сигнала, который он генерирует. Требуется, чтобы при каждом изменении сигнала его имя передается по широковещательной шине 103 на все пикеты и сравнивается параллельно с именами ожидаемых входных сигналов. Если будет обнаружено совпадение, производится запись в пикет нового значения сигнала в бит регистра потока данных. Когда будут записаны все сигнальные изменения, все пикеты будут считать параллельно управляющее слово, которое сообщает их потоку данных как использовать текущий набор входов для вычисления выхода. Эти вычисления обязательно выполняются параллельно, причем результаты сравниваются со старым значением, хранящимся в локальном вентиле. Выполняется запись в бит состояния потока данных всех тех вентилях пикетов, чьи выходы (не сигналы) изменились. Заставьте внешний контроллер обратиться ко всем пикетам и запросить вентиль, который изменился следующим. Затем широковещательно передайте подходящие сигнальные имя и значение с этого пикета на все другие пикеты, как описано выше, и производите этот цикл, пока не прекратятся сигнальные изменения или не остановится процесс.

Другим процессом может быть поиск словарного имени. Имена хранятся в пикетной памяти 102 таким образом, что первая буква всех имен может быть сравнена с желаемым широковещательным именем на широковещательной шине адресов данных 103. Все пикеты без совпадения отключаются по управляющему признаку, нами предусмотренному. Затем происходит сравнение второй буквы, и процедура сравнения и отключения воспроизводится для последующих букв (символов) до тех пор, пока не останется активных пикетных устройств или пока не будет достигнут конец этого слова. В этот момент запрашиваются остальные пикетные устройства, и индекс желаемых данных считывается последовательным устройством.

Фиг. 4 иллюстрирует базовую пикетную конфигурацию некоторого множества параллельных процессоров и памяти, пикетных устройств, организованных в строку на одиночном кремниевом чипе как часть некоторой параллельной матрицы, которую можно конфигурировать в виде SIMD подсистемы, иллюстрирующей управляющую структуру такой системы. Она также иллюстрирует управляющий процессор и супервизорный микропроцессор. На фиг.4 показаны память и логическая схема из параллельных процессорных элементов, размещенных на одном чипе, которые изображены в секции, обозначенной надписью "Матрица пикетов". Каждая память

имеет ширину в n битов, предпочтительно ширину символа 8 (9) битов, как уже было сказано, но концептуально имеет ширину слова памяти, ширина слов которой измеряется несколькими байтами. Таким образом, секция памяти параллельного пикетного процессорного элемента будет иметь ширину предпочтительно 8 (9) битов или альтернативно 16 или 32 бита. При использовании CMOS технологии мы предпочитаем оснащать каждый пикетный процессорный элемент 8-битовой или символьной ширины ассоциативной памятью (т.е. 9-битовой ширины байт с самоконтролем). Запоминающие устройства непосредственно ассоциированы однозначным образом с сопряженным процессорным элементом, который включает в себя арифметико-логическое устройство, масочные регистры (A и Q используются для маскировки) и вентиль 104 (обозначенный SR на Фиг.4), а также регистры состояний 107 и регистры потоков данных A 105 и Q 106 (обозначены DF на фиг. 4), которые подробно проиллюстрированы на пикетной диаграмме на фиг.2. DRAM произвольного доступа память и логика каждого пикетного процессора не несут какой-либо нагрузки, вызванной наличием связанной сети, поскольку существует прямая взаимодозначная ассоциация между многобитовой шириной произвольного доступа памяти DRAM и ее процессорным элементом на самом чипе.

На фиг. 4 можно заметить, что проходного В регистра вентиль (SR) 104 логически размещен между памятью и ассоциированной логикой арифметико-логического устройства процессорного элемента, и вентиль фактически является сопряженным портом для каждого процессорного элемента в пикетной матрице. Каждый пикетный чип содержит некоторое множество параллельных пикетных процессорных элементов, расположенных в линию (которая представлена в виде прямой шины) для связи с пикетным управлением. Векторных адресов шина является общей для памяти, и регистр векторных адресов данных управляет данными, проходящими через каждую память.

На фиг. 2 также изображена связь между основной или микропроцессорной платой MP, которой в предпочтительном варианте осуществления является микропроцессор 386, конфигурированный в виде PS /2 системы, с подсистемным контроллером, через который проходят глобальные команды на стандартно программно обеспеченный процессор CRP, предусмотренный для подачи команд на командный упорядочиватель 402. Управление исполнением 403 исполняет специфичный микрокод, вызванный командным упорядочивателем. Этот командный упорядочиватель аналогичен функционально контроллеру. Однако дополнительно внутри стандартно обеспеченного программами процессора CRP нами предусмотрено наличие локальных регистров 405, которые совместно с локальными регистрами ALV (не изображены) служат базой для всех операций адресации, которые широко вещаются на все пикеты внутри пикетной матрицы 406. Таким образом, адресные вычисления выполняются для всех пикетов в одном арифметико-логическом устройстве без использования пикетных

ресурсов или без использования пикетных исполнительных циклов. Это важное дополнение увеличивает гибкость пикетной матрицы, позволяя выполнять для специальных задач "дремоту", запреты и другие управляющие функции и оградить пикеты от любой широко вещаемой команды или исполнения функции над данными.

Командный упорядочиватель 402, загруженный микрокодом 407, широко вещает на матрицу пикетов для исполнения в соответствии с последовательностью команд SIMD, определяемой основным программным микропроцессором MP и стандартными программами процессора CRP прогона временной библиотеки 408, чтобы дать возможность SIMD обработки данных, содержащихся в матрице пикетов.

Команды, поступающие на микропроцессор MP через подсистемный интерфейс, являются процессорными командами высокого уровня, которые могут включать в себя "Начать процесс", "Записать наблюдения" и "Считать результат", которые проходят на микропроцессор через подсистемный контроллер микропроцессора MP. Микропроцессор можно рассматривать как основной системный или управляющий процессор в подсистемной схеме (см. фиг. 4-70). Должно быть ясно, что это устройство может быть самостоятельным устройством с добавлением периферийного вводного устройства (не показано), подобного клавиатуре и дисплейному устройству. В такой самостоятельной конфигурации система может быть в виде коммерческой PS/2 системы, в которую плату упорядочивателя (которая конфигурирует стандартно обеспеченный программами процессор) и платы процессорной матрицы вставлены вдоль линий, изображенных на фиг. 7. Стандартных программ библиотека 411 содержит стандартные последовательности для общего управления процессом, подобным, например, CALL (.), Kalman Convolve и Nav Update. Выбор из этих стандартных программ происходит через посредство программы пользователя, и поэтому общая обработка может происходить под управлением внешнего основного компьютера или под управлением программы пользователя 412, размещенной в микропроцессоре MP. В памяти микропроцессора MP предусмотрен буфер данных 413 для передачи данных внутрь и из параллельной пикетной процессорной системы. Командный упорядочиватель 402 конфигурирован так, чтобы выполнять поток управления от микропроцессора MP, а также стандартные программы, постоянно хранящиеся во временной библиотечной памяти 408 прогона стандартных программ. Среди таких стандартных программ имеются CALL (Вызвать) Load Blok (загрузить блок), Siu (синус), Cos (косинус), (найти), Min (минимум), Pange Comp. (сравнение диапазонов) и Matrix Multiply, которые содержатся в библиотеке стандартных программ 408 исполнительных систем.

Внутри CRP также имеется микрокод 407 для исполнительного управления функций низкого уровня, подобных "Загрузить", "Считать", "Сложить", "Умножить" и "Сравнить".

Заявитель предпочел внешнее управление FOR/NEXT (вперед/следующий) для каждого процессорного устройства. Он также предусмотрел детерминистическую реализацию нормализации байта с плавающей запятой.

Использование детерминистического подхода к макроразвитию системы позволяет группировать пикеты и осуществлять групповое управление. Функция локальной "дремоты" предусмотрена для учета особенностей обработки, выполняемой каждым отдельным пикетом.

Если программа по пользователю нуждается в исполнении с помощью матрицы процессоров, примитивные команды, адреса и широковещательные данные подаются на матрицу пикетных процессоров. Конкретная функция, которую реализует каждая часть системы, определяется задачей, подлежащей исполнению, и присваивается во время компилирования программы пользователя.

На фиг. 5 представлена многопикетная параллельная процессорная система 510, которая содержит некоторое множество параллельных пикетных процессоров. Для приложений, подобных слежению за многочисленными целями, сопоставлению датчиков и данных, обработке сигналов, искусственному интеллекту, обработке спутниковых изображений, распознаванию образца/цели, операциям кодирования/декодирования "Флейта Соломона", мы создали систему, которая может быть сконфигурирована в некоторой предпочтительный вариант в виде SIMD системы с 1024 параллельными процессорами с числом SEM E плат 511 от 2 до 4 (здесь представленными как четыре платы на систему) для каждого из 1024 процессоров. Индивидуальные платы 512 вставляются в ячейку 513 стоечной системной конструкции с парообразными направляющими 514, и платы снабжены рычагами "установки/извлечения" 516, для того чтобы закрытая крыша 517 надежно герметизировала смонтированную систему в стойке с хранилищем объемом 32 или 64 Мбайта, способную работать со скоростью приблизительно 2 миллиарда операций в секунду. Эта система компактна, и матрица из некоторого множества пикетов вставлена в заднюю панель 518, на которой есть логическая схема, обеспечивающая связь некоторого множества плат. Процессор 32 М объемом памяти 32 М формируется из 4 EM E плат, и система весит всего около 30 фунтов. Мощность подается источниками мощности 519. Мощность, потребляемая таким процессором, охлаждаемым воздухом, оценивается около 280 Вт. Каждая SIMD система имеет два входных-выходных порта 520 для связи канальных адаптеров с ассоциированными главными ЭВМ или, другими словами, с остальным миром. При использовании многопикетного параллельного процессора, состоящего из 4 логических станиц и использующего стандартные модульные электронные блоки и шинную структуру для присоединения к внешней памяти (например, PI, TM и IEEE 488 шины), такой процессор можно присоединить через порт ввода-вывода к шине памяти целевого процессора и его можно рассматривать как продолжение

запоминающего пространства целевого процессора.

В проиллюстрированном многопикетном параллельном процессоре, составленном из 1024 параллельных процессорных элементов, каждый процессор обладает локальной памятью 32 килобайта, а ассоциированный тракт на пикетный параллельный процессор имеет в ширину 8 битов или символов (9 битов) в параллель.

Процессоры внутри каждого пикета обмениваются данными с другими соседними процессорами и между страницами через связную сеть задней платы, предпочтительно кросс-шину, но, альтернативно: скользящую кросс-шину, тасующую сеть, N-куб по основанию 3 или N-куб по основанию 8.

Индивидуальные пикетные процессоры системы размещены на двухплатном модуле из четырех плат и PS/2 микропроцессор на одной плате, а упорядочиватель процессора с стандартным программным обеспечением размещен на другой плате из четырех плат, которые составляют систему, схематично изображенную на фиг. 6 и 7. Индивидуальные пикеты 100 или платы с пикетами 512 могут конфигурировать по ходу работы стандартно обеспеченными программами процессором CRP, осуществляющим вход, или могут быть выведены из операций в зависимости от состояния данных под управлением вентильной 104 архитектуры и локальных регистров, которые сопряжены с управлением исполнительной платы упорядочивателя CRP 703. Поэтому пикетные процессоры могут выполнять независимо операции выравнивания и нормализации, которые ассоциируются с операциями над числами с плавающей запятой.

Процессоры управляют параллельно общим упорядочивателем. Плата упорядочивателя 703 содержит контроллер CRP пикетного процессора и может заставить пикетный процессор выполнить одну последовательность команд, закодированную для исполнения на матрицу из SIMD процессоров побайтовым образом, подобным классической побитовой последовательной обработке. Контроллер имеет три слоя. Микроконтроль для пикетов микрозакодирован аналогично современным процессорам и передается параллельно на все пикеты. Микроуправление и пикеты синхронизируются с одной и той же тактирующей системой CLK, чтобы функции, управляемые упорядочивателем, могли быть исполнены за одно и то же тактовое время. Подача команд на микроуправляемый упорядочиватель является функцией стандартно обеспеченного программами процессора. Плата упорядочивателя 703 представляет собой аппаратный контроллер, который во время большинства функций выполняет команды управления циклами и рекурсивно начинает новые последовательности микроуправления. Этот контроллер со своей стандартной программной библиотекой 408 и своей функцией организации циклов обеспечивает хорошую подачу (команд) на пикеты и не понуждается командами. Контроллер стандартно обеспеченного программами процессора CRP содержит большое число микрокодов, которые вызываются основной системой, который в подсистеме действуют

как первичный супервизорный пикетный контроллер. Он является главной управляющей системой пикетной матрицы. Именно процессор 386-й управляет деятельностью матрицы из пикетов. В каждый данный момент все пикеты матрицы могут исполнять одну и ту же команду, хотя некоторый поднабор процессоров может реагировать индивидуально на поток управляющих команд.

Имеется несколько вариантов индивидуальных реакций благодаря наличию байтовой управляющей функции для каждого пикета ("дремота", запрет, т.д.), что обеспечивает локальную автономию, программирование которой дает преимущества для адресации, которая может быть произведена под управлением системы во время компилирования программы.

Дополнительно, как сказано выше, предусмотрена автономность адресации локальной памяти. Упорядочиватель SIMD контроллера подает общий адрес на все пикеты. Каждый пикет может увеличивать этот адрес локально, чтобы повысить свою способность выполнять независимые от данных доступы к памяти.

Дополнительно пикет может принимать или не принимать участие в действиях матрицы в зависимости от местных условий.

При наличии такой характеристики теперь можно ввести концепцию групп в SIMD обработку, предусмотрев устройство, позволяющее каждому пикету приписывать себя к одной или более группам, и обработка может продолжаться посредством этих групп, в которых конфигурация может существенно меняться в ходе работы. В одном варианте только одна группа или комбинация групп может быть активной одновременно и каждая выполняет один и тот же поток команд SIMD. Некоторые операции требуют работы лишь подгруппы или группы пикетов. При программировании можно использовать преимущества, вытекающие из такой способности. Автономия локального участия позволяет работать в этом направлении. Очевидно, что чем больше пикетов, занимающихся вычислением, тем лучше.

В предпочтительном варианте система будет выглядеть так, как она изображена схематично на функциональной блок-схеме подсистемы параллельного процессора на фиг.6. Через порты ввода-вывода системы, управляемой управлением интерфейса с главной ЭВМ 413, упорядочиватель контроллера функционирует аналогично SIMD программе с высокого уровня функциональными макрокодами, управляющими функциями процессорных элементов. Адресация памяти обеспечивает поток данных шириной 8 битов, байтовой ширины, и арифметическая логика с основанием 8 используется для функций (логических, сложения, умножения и деления). Предусмотрены формат с плавающей запятой и автономная работа пикетов в режимах сна и дремоты по отдельности и раздельная адресация.

Схема подсистемного контроллера

изображена на фиг.7. Каждая из плат процессорной матрицы 512 (которых 4 на этой иллюстрации подсистемы, но их число можно сократить до 2 SEM E плат) соединена с упорядочивателем CRP 703, который сопряжен с подсистемным контроллером 702, который соединен портами с основной запоминающей системой или с другой подсистемой в конфигурации через интерфейс чипа 705 с ассоциированной микроканальной шиной 706. В предпочтительном варианте подсистемный контроллер представляет собой общего назначения микропроцессорное устройство IBM Корпорейшенс PS/2 (торговый знак IBM) и использует Интел 386 процессорный чип и память 4 Мбайта. Микропроцессор персонального компьютера MP 702 сопряжен с платой упорядочивателя через микроканального типа шину 705, 706 внутри подсистемы.

Формула изобретения:

1. Параллельная процессорная система, содержащая N процессорных устройств, причем входы-выходы "Адрес-данные" всех процессорных устройств соединены через шину "Адрес-данные", входы-выходы "Управление" всех процессорных устройств соединены через шину "Управление", отличающаяся тем, что каждое процессорное устройство содержит локальную память, процессорный элемент, три регистра, причем первый вход-выход локальной памяти соединен через шину "Адрес-данные" с входом-выходом "Адрес-данные" системы, управляющие входы-выходы которой соединены через шину "Управление" с управляющими входами первого регистра, выход которого соединен с первым входом-выходом процессорного элемента, второй вход-выход которого соединен с первым входом второго регистра, выход которого соединен с входом третьего регистра, выход которого соединен с третьим входом-выходом процессорного элемента, с первым входом первого регистра, с вторым входом-выходом локальной памяти, выход которой соединен с вторым входом первого регистра, второй вход второго регистра соединен через шину "Адрес-данные" с входом-выходом "Адрес-данные" системы.

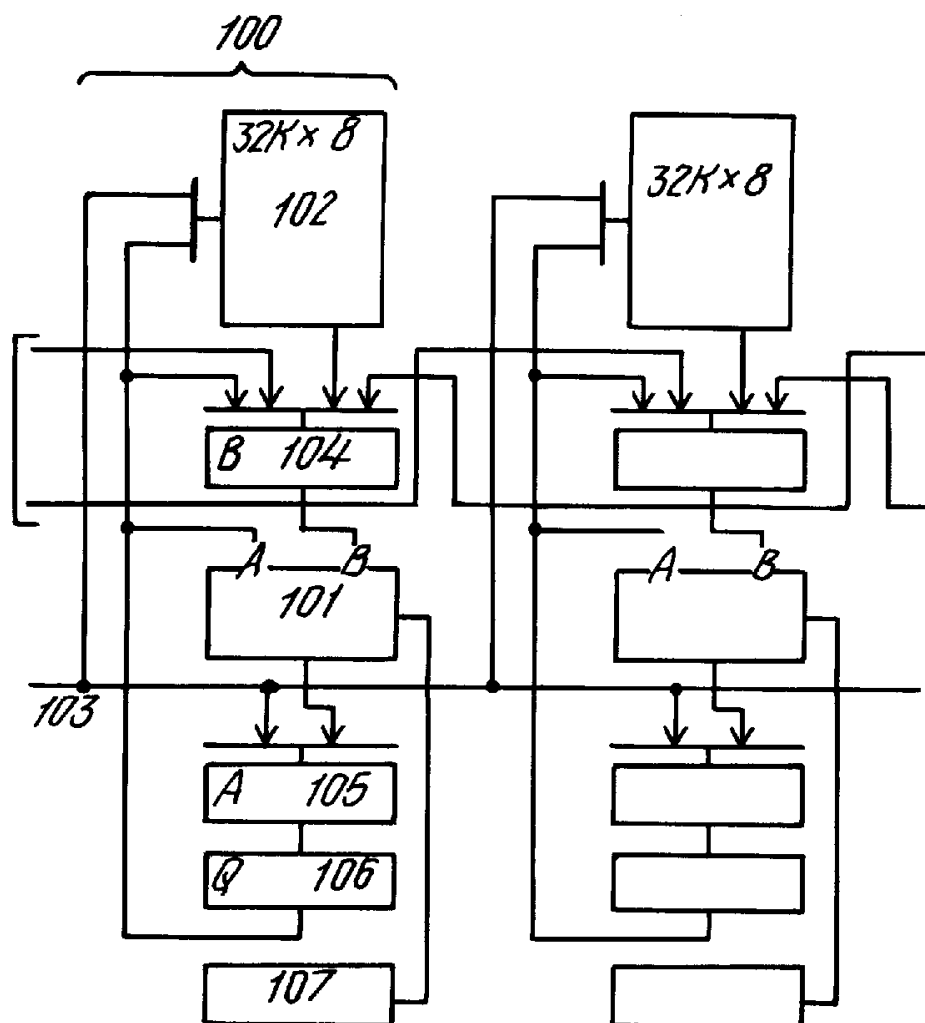
2. Система по п.1, отличающаяся тем, что шина "Адрес-данные" выполнена в виде шины широковещательной рассылки адресов-данных.

3. Система по п.1, отличающаяся тем, что шина "Управление" выполнена в виде двунаправленной шины.

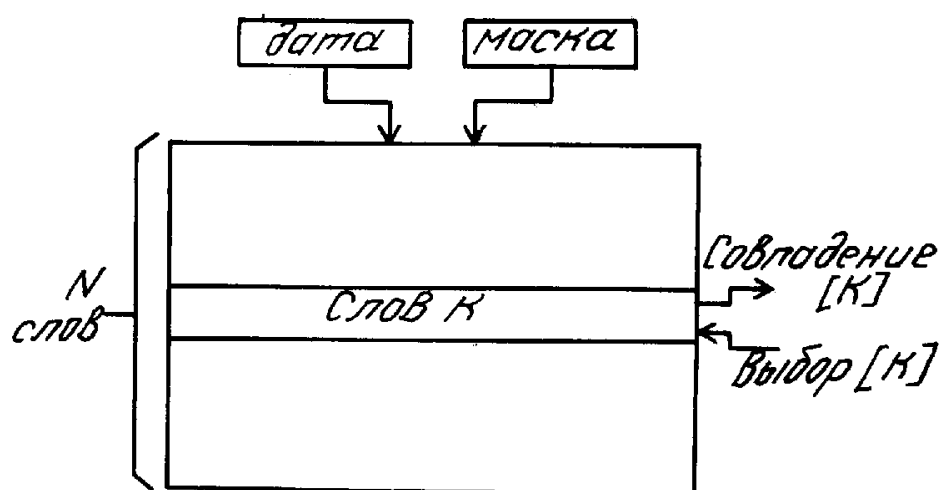
4. Система по п.1, отличающаяся тем, что каждое процессорное устройство содержит регистр команд-состояния, вход=выход которого соединен с четвертым входом-выходом процессорного элемента.

5. Система по п.1, отличающаяся тем, что процессорный элемент имеет ширину, по меньшей мере равную 8 бит.

6. Система по п.1, отличающаяся тем, что емкость локальной памяти составляет по меньшей мере 32 x 8 килобит.



Фиг. 2



Фиг. 3

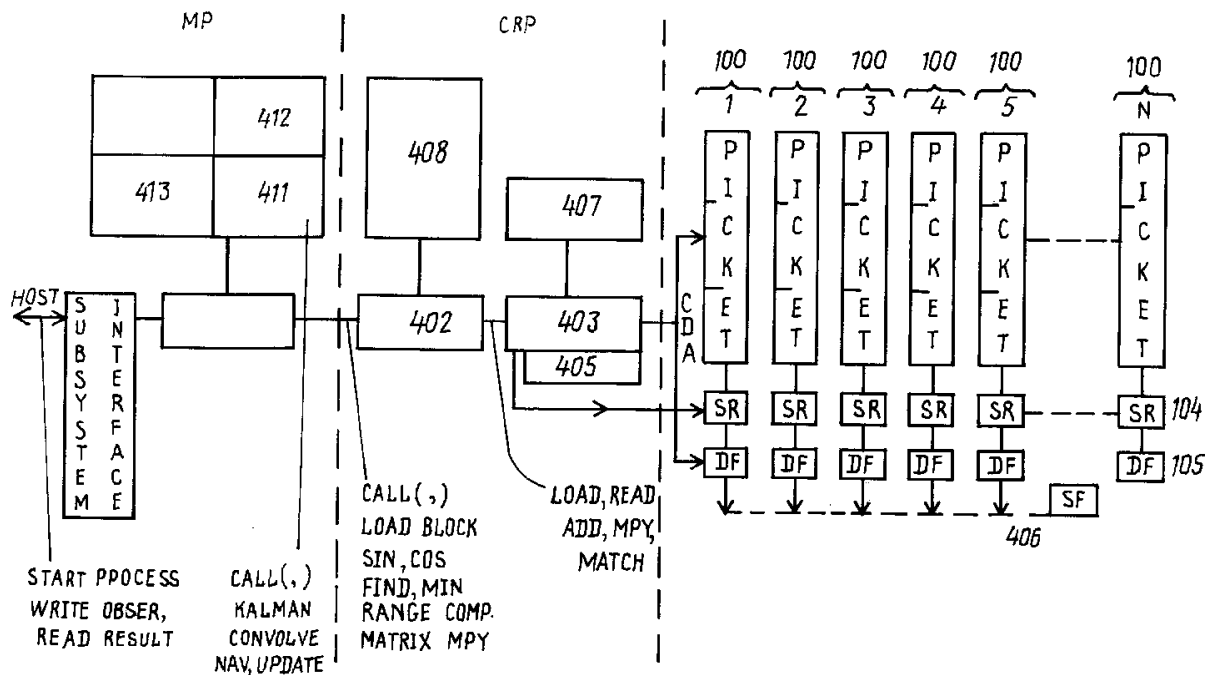


Fig. 4

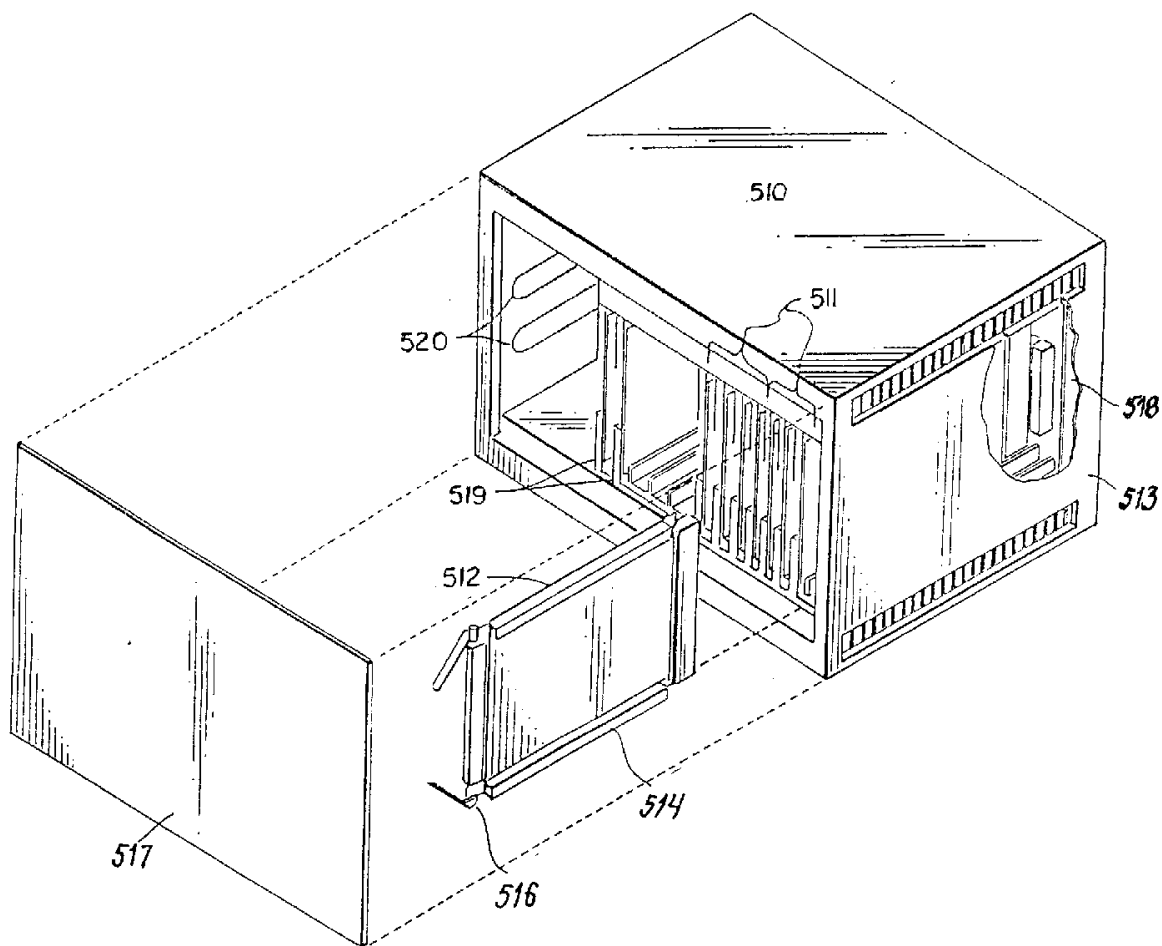


Fig. 5

