Учреждение образования

«БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ

ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ»

**А. С. Кобайло**

**КОМПЬЮТЕРНЫЕ СИСТЕМЫ И СЕТИ**

Конспект лекций по одноименной дисциплине

для студентов IT-специальностей

Минск 2017

УДК 004.2.031(075.8)

ББК 32.97.я73

К55

Рассмотрено и рекомендовано редакционно-издательским советом Белорусского государственного технологического университета

Рецензенты:

кандидат технических наук, доцент,

доцент кафедры программного обеспечения информационных технологий Белорусского

государственного университета информатики и радиоэлектроники

*П. Ю. Бранцевич;*

Заведующий кафедрой информационных технологий и экологического менеджмента ГУО «Институт бизнеса и менеджмента технологий» БГУ

*Б.В. Лесун*

**Кобайло, А. С.**

К55 Арифметико-логические основы цифровых вычислительных машин и архитектура компьютеров: конспект лекций по одноименной дисциплине для студентов IT-специальностей / А. С. Кобайло. – Минск : БГТУ, 2015. – 79 с.

Конспект лекций содержит основные сведения в области схемотехники ЦВМ, архитектурной организации микропроцессоров, основ программирования на языках низкого уровня. Рассмотрены вопросы организации обмена данными в микропроцессорных системах, а также структуры конкретных микропроцессоров.

Издание предназначено для студентов, изучающих дисциплину «Арифметико-логические основы цифровых вычислительных машин и архитектура компьютеров». Пособие также может быть полезно аспирантам, изучающим избранные аспекты вышеуказанных проблем.

**УДК 004.2.031(075.8)**

**ББК 32.97.я73**

© УО «Белорусский государственный

технологический университет», 2017

© Кобайло А. С., 2017

**ПЕРЕЧЕНЬ РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ**

**a) основная**

1. Урбанович, П. П. Компьютерные сети / П. П. Урбанович, Д. М. Романенко, Е. В. Кабак. – Минск: БГТУ, 2011.
2. Романенко, Д. М. Компьютерные сети. Лабораторный практикум / Д. М. Романенко, Н. В. Пацей. – Минск: БГТУ, 2011. – 133 c.
3. Олифер, В. Г. Компьютерные сети. Принципы, технологии, протоколы. – 2-е изд. / В. Г. Олифер, Н. А. Олифер. – СПб.: Питер, 2004.
4. Танненбаум, Э. Компьютерные сети / Э. Танненбаум. – СПб.: Питер, 2002.
5. Танненбаум, Э. Архитектура компьютера / Э. Танненбаум. – СПб.: Питер, 2010.

**б) дополнительная**

1. Рошан, П. Основы построения беспроводных локальных сетей стандарта 802.11. / П. Рошан, Л. Джонатан . – М.: Издательский дом «Вильямс», 2004. – 304 с.

2. Закер, К. Компьютерные сети: модернизация, поиск неисправностей / К. Закер. – СПб.: БВХ-Санктпетербург, 2001. – 1001 с.

3. Кульгин, М.В. Компьютерные сети. Практика построения для профессионалов. 2-е изд. / М.В. Кульгин. – СПб.:Питер, 2003. – 462 с.

4. Гейер, Д. Беспроводные сети. Первый шаг / Д. Гейер. – М.: Вильямс, 2005.

5. Кобайло, А. С. Теория синтеза вычислительных систем реального времени / А.С. Кобайло. – Минск: БГТУ, 2010. – 256 с.

* Компьютерная [система](http://official.academic.ru/23018/%D0%A1%D0%B8%D1%81%D1%82%D0%B5%D0%BC%D0%B0)" означает любое устройство или группувзаимосвязанных или смежных устройств, одно или более из которых, действуя в соответствии с программой, осуществляет автоматизированную обработку данных..."
* — Компьютерная система – Набор аппаратных средств и периферийных устройств, составляющих в совокупности вычислительную машину. Основными частями стандартной компьютерной системы являются процессорный блок с дисководами для жестких и гибких дисков, монитор, клавиатура,…
* [**компьютерная система**](http://normative_reference_dictionary.academic.ru/27549/%D0%BA%D0%BE%D0%BC%D0%BF%D1%8C%D1%8E%D1%82%D0%B5%D1%80%D0%BD%D0%B0%D1%8F_%D1%81%D0%B8%D1%81%D1%82%D0%B5%D0%BC%D0%B0) компьютерная система (computer system): Совокупность аппаратных средств, управляемых программным обеспечением (операционной системой) как единый модуль. Компьютерная система может также предоставлять общие услуги, такие как управление, …

**Микропроцессорная система** может рассматриваться как частный случай электронной системы, предназначенной для обработки входных сигналов и выдачи выходных сигналов ([рис. 4.1](http://www.intuit.ru/department/hardware/mpbasics/1/#image.1.1)). В качестве входных и выходных сигналов при этом могут использоваться аналоговые сигналы, одиночные цифровые сигналы, цифровые коды, последовательности цифровых кодов. Внутри системы может производиться хранение, накопление сигналов (или информации), но суть от этого не меняется. Если система цифровая (а *микропроцессорные системы* относятся к разряду цифровых), то входные аналоговые сигналы преобразуются в последовательности кодов выборок с помощью *АЦП*, а выходные аналоговые сигналы формируются из последовательности кодов выборок с помощью *ЦАП*. Обработка и хранение информации производятся в цифровом виде.

Характерная особенность традиционной цифровой системы состоит в том, что алгоритмы обработки и хранения информации в ней *жестко связаны* со схемотехникой системы. То есть изменение этих алгоритмов возможно только путем изменения структуры системы, замены электронных узлов, входящих в систему, и/или связей между ними. Например, если нам нужна дополнительная операция суммирования, то необходимо добавить в структуру системы лишний *сумматор*. Или если нужна дополнительная функция хранения кода в течение одного такта, то мы должны добавить в структуру еще один регистр. Естественно, это практически невозможно сделать в процессе эксплуатации, обязательно нужен новый производственный цикл проектирования, изготовления, отладки всей системы. Именно поэтому традиционная цифровая система часто называется системой на «жесткой логике».

Аналоговые

сигналы

Цифровые коды

Цифровые

сигналы

Электронная система (обработка, хранение)

Аналоговые

сигналы

Цифровые

сигналы

Цифровые коды

Рис. 4.1.  Электронная система.

Любая система на «жесткой логике» обязательно представляет собой специализированную систему, настроенную исключительно на одну задачу или (реже) на несколько близких, заранее известных задач. Это имеет свои бесспорные преимущества.

Во-первых, специализированная система (в отличие от универсальной) никогда не имеет аппаратурной избыточности, то есть каждый ее элемент обязательно работает в полную силу (конечно, если эта система грамотно спроектирована).

Во-вторых, именно специализированная система может обеспечить максимально высокое быстродействие, так как скорость выполнения алгоритмов обработки информации определяется в ней только быстродействием отдельных логических элементов и выбранной схемой путей прохождения информации. А именно логические элементы всегда обладают максимальным на данный момент быстродействием.

Но в то же время большим недостатком цифровой системы на "жесткой логике" является то, что для каждой новой задачи ее надо проектировать и изготавливать заново. Это процесс длительный, дорогостоящий, требующий высокой квалификации исполнителей. А если решаемая задача вдруг изменяется, то вся аппаратура должна быть полностью заменена. В нашем быстро меняющемся мире это довольно расточительно.

Путь преодоления этого недостатка довольно очевиден: надо построить такую систему, которая могла бы легко адаптироваться под любую задачу, перестраиваться с одного алгоритма работы на другой без изменения аппаратуры. И задавать тот или иной алгоритм мы тогда могли бы путем ввода в систему некой дополнительной управляющей информации, **программы** работы системы ([рис. 4.2](http://www.intuit.ru/department/hardware/mpbasics/1/#image.1.2)). Тогда система станет универсальной, или **программируемой**, не жесткой, а гибкой. Именно это и обеспечивает *микропроцессорная система*.

Аналоговые

сигналы

Цифровые коды

Цифровые

сигналы

Аналоговые

сигналы

Цифровые

сигналы

Цифровые коды

Управляющая

информация

(программа)

Рис. 4.2.  Программируемая (она же универсальная) электронная система

Но любая универсальность обязательно приводит к избыточности. Ведь решение максимально трудной задачи требует гораздо больше средств, чем решение максимально простой задачи. Поэтому сложность универсальной системы должна быть такой, чтобы обеспечивать решение самой трудной задачи, а при решении простой задачи система будет работать далеко не в полную силу, будет использовать не все свои ресурсы. И чем проще решаемая задача, тем больше избыточность, и тем менее оправданной становится универсальность. Избыточность ведет к увеличению стоимости системы, снижению ее надежности, увеличению потребляемой мощности и т.д.

Кроме того, универсальность, как правило, приводит к существенному снижению быстродействия. Оптимизировать универсальную систему так, чтобы каждая новая задача решалась максимально быстро, попросту невозможно. Общее правило таково: чем больше универсальность, гибкость, тем меньше быстродействие. Более того, для универсальных систем не существует таких задач (пусть даже и самых простых), которые бы они решали с максимально возможным быстродействием. За все приходится платить.

Таким образом, можно сделать следующий вывод. Системы на «жесткой логике» хороши там, где решаемая задача не меняется длительное время, где требуется самое высокое быстродействие, где алгоритмы обработки информации предельно просты. А универсальные, программируемые системы хороши там, где часто меняются решаемые задачи, где высокое быстродействие не слишком важно, где алгоритмы обработки информации сложные. То есть любая система хороша на своем месте.

Однако за последние десятилетия быстродействие универсальных (*микропроцессорных*) систем сильно выросло (на несколько порядков). К тому же большой объем выпуска микросхем для этих систем привел к резкому снижению их стоимости. В результате область применения систем на «жесткой логике» резко сузилась. Более того, высокими темпами развиваются сейчас программируемые системы, предназначенные для решения одной задачи или нескольких близких задач. Они удачно совмещают в себе как достоинства систем на «жесткой логике», так и программируемых систем, обеспечивая сочетание достаточно высокого быстродействия и необходимой гибкости. Так что *вытеснение* «жесткой логики» продолжается.

4.1.2. Что такое микропроцессор

Ядром любой *микропроцессорной системы* является *микропроцессор* или просто процессор (от англ.processor). Перевести на русский язык это слово правильнее всего как "обработчик", так как именно ***микропроцессор***– это тот узел, блок, который производит всю обработку информации внутри *микропроцессорной системы*. Остальные узлы выполняют всего лишь *вспомогательные функции*: хранение информации (в том числе и управляющей информации, то есть программы), связи с внешними устройствами, связи с пользователем и т.д. Процессор заменяет практически всю «жесткую логику», которая понадобилась бы в случае традиционной цифровой системы. Он выполняет *арифметические функции* (сложение, умножение и т.д.), логические функции (сдвиг, сравнение, маскирование кодов и т.д.), временное хранение кодов (во внутренних регистрах), пересылку кодов между узлами *микропроцессорной системы* и многое другое. Количество таких элементарных операций, выполняемых процессором, может достигать нескольких сотен. Процессор можно сравнить с мозгом системы.

Но при этом надо учитывать, что все свои операции процессор выполняет **последовательно**, то есть одну за другой, по очереди. Конечно, существуют процессоры с параллельным выполнением некоторых операций, встречаются также *микропроцессорные системы*, в которых несколько процессоров работают над одной задачей параллельно, но это редкие исключения. С одной стороны, последовательное выполнение операций – несомненное достоинство, так как позволяет с помощью всего одного процессора выполнять любые, самые сложные алгоритмы обработки информации. Но, с другой стороны, последовательное выполнение операций приводит к тому, что время выполнения алгоритма зависит от его сложности. Простые алгоритмы выполняются быстрее сложных. То есть *микропроцессорная система* способна сделать все, но работает

Процессор

Управляющая

информационная

(программа)

Рис. 4.3.  Информационные потоки в микропроцессорной системе

Итак, *микропроцессор* способен выполнять множество операций. Но откуда он узнает, какую операцию ему надо выполнять в данный момент? Именно это определяется управляющей информацией, программой. Программа представляет собой набор **команд (инструкций)**, то есть цифровых кодов, расшифровав которые, процессор узнает, что ему надо делать. Программа от начала и до конца составляется человеком, программистом, а процессор выступает в роли послушного исполнителя этой программы, никакой инициативы он не проявляет (если, конечно, исправен). Поэтому сравнение процессора с мозгом не слишком корректно. Он всего лишь исполнитель того алгоритма, который заранее составил для него человек. Любое отклонение от этого алгоритма может быть вызвано только неисправностью процессора или каких-нибудь других узлов *микропроцессорной системы*.

Все команды, выполняемые процессором, образуют **систему команд** процессора. Структура и объем системы команд процессора определяют его быстродействие, гибкость, *удобство использования*. Всего команд у процессора может быть от нескольких десятков до нескольких сотен. Система команд может быть рассчитана на узкий круг решаемых задач (у специализированных процессоров) или на максимально широкий круг задач (у универсальных процессоров). Коды команд могут иметь различное количество разрядов (занимать от одного до нескольких байт). Каждая команда имеет свое время выполнения, поэтому время выполнения всей программы зависит не только от количества команд в программе, но и от того, какие именно команды используются.

По числу больших интегральных схем (БИС) в микропроцессорном комплекте различают микропроцессоры однокристальные, многокристальные и многокристальные секционные.

Процессоры даже самых простых ЭВМ имеют сложную функциональную структуру, содержат большое количество электронных элементов и множество разветвленных связей. Изменять структуру процессора необходимо так, чтобы полная принципиальная схема или ее части имели количество элементов и связей, совместимое с возможностями БИС. При этом микропроцессоры приобретают внутреннюю магистральную архитектуру, т. е. в них к единой внутренней информационной магистрали подключаются все основные функциональные блоки (арифметико-логический, рабочих регистров, стека, прерываний, интерфейса, управления и синхронизации и др.).

Для обоснования классификации микропроцессоров по числу БИС надо распределить все аппаратные блоки процессора между основными тремя функциональными частями: операционной, управляющей и интерфейсной. Сложность операционной и управляющей частей процессора определяется их разрядностью, системой команд и требованиями к системе прерываний; сложность интерфейсной части разрядностью и возможностями подключения других устройств ЭВМ (памяти, внешних устройств, датчиков и исполнительных механизмов и др.). Интерфейс процессора содержит несколько десятков информационных шин данных (ШД), адресов (ША) и управления (ШУ).

Однокристальные микропроцессоры получаются при реализации всех аппаратных средств процессора в виде одной БИС или СБИС (сверхбольшой интегральной схемы). По мере увеличения степени интеграции элементов в кристалле и числа выводов корпуса параметры однокристальных микропроцессоров улучшаются. Однако возможности однокристальных микропроцессоров ограничены аппаратными ресурсами кристалла и корпуса. Для получения многокристального микропроцессора необходимо провести разбиение его логической структуры на функционально законченные части и реализовать их в виде БИС (СБИС). Функциональная законченность БИС многокристального микропроцессора означает, что его части выполняют заранее определенные функции и могут работать автономно.

На рис. 4.4, *а* показано функциональное разбиение структуры процессора при создании трехкристального микропроцессора (пунктирные линии), содержащего БИС операционного (ОП), управляющего (УП) и интерфейсного (ИП) процессоров.

УП

ОП

ИП

УП

ОП

ИП

МА

МД

МУ

*а* *б*

Рис. 4.4. Функциональная структура процессора (*а*) и ее разбиение для реализации процессора в виде комплекта секционных БИС (*б*)

Операционный процессор служит для обработки данных, управляющий процессор выполняет функции выборки, декодирования и вычисления адресов операндов и также генерирует последовательности микрокоманд. Автономность работы и большое быстродействие БИС УП позволяет выбирать команды из памяти с большей скоростью, чем скорость их исполнения БИС ОП. При этом в УП образуется очередь еще не исполненных команд, а также заранее подготавливаются те данные, которые потребуются ОП в следующих циклах работы. Такая опережающая выборка команд экономит время ОП на ожидание операндов, необходимых для выполнения команд программ. Интерфейсный процессор позволяет подключить память и периферийные средства к микропроцессору; он, по существу, является сложным контроллером для устройств ввода/вывода информации. БИС ИП выполняет также функции канала прямого доступа к памяти.

Выбираемые из памяти команды распознаются и выполняются каждой частью микропроцессора автономно и поэтому может быть обеспечен режим одновременной работы всех БИС МП, т. е. конвейерный поточный режим исполнения последовательности команд программы (выполнение последовательности с небольшим временным сдвигом). Такой режим работы значительно повышает производительность микропроцессора.

Многокристальные секционные микропроцессоры получаются в том случае, когда в виде БИС реализуются части (секции) логической структуры процессора при функциональном разбиении ее вертикальными плоскостями (рис. 4.4, *б*). Для построения многоразрядных микропроцессоров при параллельном включении секций БИС в них добавляются средства «стыковки».

Для создания высокопроизводительных многоразрядных микропроцессоров требуется столь много аппаратных средств, не реализуемых в доступных БИС, что может возникнуть необходимость еще и в функциональном разбиении структуры микропроцессора горизонтальными плоскостями. В результате рассмотренного функционального разделения структуры микропроцессора на функционально и конструктивно законченные части создаются условия реализации каждой из них в виде БИС. Все они образуют комплект секционных БИС МП.

Таким образом, микропроцессорная секция – это БИС, предназначенная для обработки нескольких разрядов данных или выполнения определенных управляющих операций. Секционность БИС МП – определяет возможность «наращивания» разрядности обрабатываемых данных или усложнения устройств управления микропроцессора при «параллельном» включении большего числа БИС.

Однокристальные и трехкристальные БИС МП, как правило, изготовляют на основе микроэлектронных технологий униполярных полупроводниковых приборов, а многокристальные секционные БИС МП на основе технологии биполярных полупроводниковых приборов. Использование многокристальных микропроцессорных высокоскоростных биполярных БИС, имеющих функциональную законченность при малой физической разрядности обрабатываемых данных и монтируемых в корпус с большим числом выводов, позволяет организовать разветвление связи в процессоре, а также осуществить конвейерные принципы обработки информации для повышения его производительности.

По назначению различают универсальные и специализированные микропроцессоры.

Универсальные микропроцессоры могут быть применены для решения широкого круга разнообразных задач. При этом их эффективная производительность слабо зависит от проблемной специфики решаемых задач. Специализация МП, т. е. его проблемная ориентация на ускоренное выполнение определенных функций позволяет резко увеличить эффективную производительность при решении только определенных задач.

Среди специализированных микропроцессоров можно выделить различные микроконтроллеры, ориентированные на выполнение сложных последовательностей логических операций, математические МП, предназначенные для повышения производительности при выполнении арифметических операций за счет, например, матричных методов их выполнения, МП для обработки данных в различных областях применений и т. д. С помощью специализированных МП можно эффективно решать новые сложные задачи параллельной обработки данных. Например, конволюция позволяет осуществить более сложную математическую обработку сигналов, чем широко используемые методы корреляции. Последние в основном сводятся к сравнению и определению подобия всего двух серий данных: входных, передаваемых формой сигнала, и фиксированных опорных. Конволюция дает возможность в реальном масштабе времени находить соответствие для сигналов изменяющейся формы путем сравнения их с различными эталонными сигналами, что, например, может позволить эффективно выделить полезный сигнал на фоне шума.

Разработанные однокристальные конвольверы используются в устройствах опознавания образов в тех случаях, когда возможности сбора данных превосходят способности системы обрабатывать эти данные.

По виду обрабатываемых входных сигналов различают цифровые и аналоговые микропроцессоры. Сами микропроцессоры – цифровые устройства, однако могут иметь встроенные аналого-цифровые и цифро-аналоговые преобразователи. Поэтому входные аналоговые сигналы передаются в МП через преобразователь в цифровой форме, обрабатываются и после обратного преобразования в аналоговую форму поступают на выход. С архитектурной точки зрения такие микропроцессоры представляют собой аналоговые функциональные преобразователи сигналов и называются аналоговыми микропроцессорами. Они выполняют функции любой аналоговой схемы (например, производят генерацию колебаний, модуляцию, смещение, фильтрацию, кодирование и декодирование сигналов в реальном масштабе времени и т. д., заменяя сложные схемы, состоящие из операционных усилителей, катушек индуктивности, конденсаторов и т. д.). При этом применение аналогового микропроцессора значительно повышает точность обработки аналоговых сигналов и их воспроизводимость, а также расширяет функциональные возможности за счет программной «настройки» цифровой части микропроцессора на различные алгоритмы обработки сигналов.

Обычно в составе однокристальных аналоговых МП имеется несколько каналов аналого-цифрового и цифро-аналогового преобразования. В аналоговом микропроцессоре разрядность обрабатываемых данных достигает 24 бит и более, большое значение уделяется увеличению скорости выполнения арифметических операций.

Отличительная черта аналоговых микропроцессоров – способность к переработке большого объема числовых данных, т. е. к выполнению операций сложения и умножения с большой скоростью при необходимости даже за счет отказа от операций прерываний и переходов. Аналоговый сигнал, преобразованный в цифровую форму, обрабатывается в реальном масштабе времени и передается на выход обычно в аналоговой форме через цифро-аналоговый преобразователь. При этом согласно теореме Котельникова частота квантования аналогового сигнала должна вдвое превышать верхнюю частоту сигнала.

Сравнение цифровых микропроцессоров производится сопоставлением времени выполнения ими списков операций. Сравнение же аналоговых микропроцессоров производится по количеству эквивалентных звеньев аналого-цифровых фильтров рекурсивных фильтров второго порядка. Производительность аналогового микропроцессора определяется его способностью быстро выполнять операции умножения: чем быстрее осуществляется умножение, тем больше эквивалентное количество звеньев фильтра в аналоговом преобразователе и тем более сложный алгоритм преобразования цифровых сигналов можно задавать в микропроцессоре.

Одним из направлений дальнейшего совершенствования аналоговых микропроцессоров является повышение их универсальности и гибкости. Поэтому вместе с повышением скорости обработки большого объема цифровых данных будут развиваться средства обеспечения развитых вычислительных процессов обработки цифровой информации за счет реализации аппаратных блоков прерывания программ и программных переходов.

По характеру временной организации работы микропроцессоры делят на синхронные и асинхронные.

Синхронные микропроцессоры – микропроцессоры, в которых начало и конец выполнения операций задаются устройством управления (время выполнения операций в этом случае не зависит от вида выполняемых команд и величин операндов).

Асинхронные микропроцессоры позволяют начало выполнения каждой следующей операции определить по сигналу фактического окончания выполнения предыдущей операции. Для более эффективного использования каждого устройства микропроцессорной системы в состав асинхронно работающих устройств вводят электронные цепи, обеспечивающие автономное функционирование устройств. Закончив работу над какой-либо операцией, устройство вырабатывает сигнал запроса, означающий его готовность к выполнению следующей операции. При этом роль естественного распределителя работ принимает на себя память, которая в соответствии с заранее установленным приоритетом выполняет запросы остальных устройств по обеспечению их командной информацией и данными.

По организации структуры микропроцессорных систем различают микроЭВМ одно- и многомагистральные.

В одномагистральных микроЭВМ все устройства имеют одинаковый интерфейс и подключены к единой информационной магистрали, по которой передаются коды данных, адресов и управляющих сигналов.

В многомагистральныхмикроЭВМ устройства группами подключаются к своей информационной магистрали. Это позволяет осуществить одновременную передачу информационных сигналов по нескольким (или всем) магистралям. Такая организация систем усложняет их конструкцию, однако увеличивает производительность.

По количеству выполняемых программ различают одно- и многопрограммные микропроцессоры.

В однопрограммных микропроцессорах выполняется только одна программа. Переход к выполнению другой программы происходит после завершения текущей программы.

В много- или мультипрограммных микропроцессорах одновременно выполняется несколько (обычно несколько десятков) программ. Организация мультипрограммной работы микропроцессорных управляющих систем позволяет осуществить контроль за состоянием и управлением большим числом источников или приемников информации.

**4.1.3. Структура типового микропроцессора**

Архитектура типичной небольшой вычислительной системы на основе микроЭВМ показана на рис. 1.2. Такая микроЭВМ содержит все пять основных блоков цифровой машины: устройство ввода информации, управляющее устройство (УУ), арифметико-логическое устройство (АЛУ) (входящее в состав микропроцессора), запоминающие устройства (ЗУ) и устройство вывода информации.

Порты

Входы

Память данных ЗУПВ

Память программ ПЗУ

Микропроцессор

Выходы

Порты

От периферийных устройств

Шина

управления

Шина

адреса

16 линий

Шина

данных

(8 линий)

Рис. 4.5. Архитектура типового микропроцессора.

Микропроцессор координирует работу всех устройств цифровой системы с помощью шины управления (ШУ). Помимо ШУ имеется 16-разрядная адресная шина (ША), которая служит для выбора определенной ячейки памяти, порта ввода или порта вывода. По 8-разрядной информационной шине или шине данных (ШД) осуществляется двунаправленная пересылка данных к микропроцессору и от микропроцессора. Важно отметить, что МП может посылать информацию в память микроЭВМ или к одному из портов вывода, а также получать информацию из памяти или от одного из портов ввода.

Постоянное запоминающее устройство (ПЗУ) в микроЭВМ содержит некоторую программу (на практике программу инициализации ЭВМ). Программы могут быть загружены в запоминающее устройство с произвольной выборкой (ЗУПВ) и из внешнего запоминающего устройства (ВЗУ). Это программы пользователя.

В качестве примера, иллюстрирующего работу микроЭВМ, рассмотрим процедуру, для реализации которой нужно выполнить следующую последовательность элементарных операций:

1. Нажать клавишу с буквой «А» на клавиатуре.

2. Поместить букву «А» в память микроЭВМ.

3. Вывести букву «А» на экран дисплея.

Это типичная процедура ввода-запоминания-вывода, рассмотрение которой дает возможность пояснить принципы использования некоторых устройств, входящих в микроЭВМ.

На рис. 1.3 приведена подробная диаграмма выполнения процедуры ввода-запоминания-вывода. Обратите внимание, что команды уже загружены в первые шесть ячеек памяти. Хранимая программа содержит следующую цепочку команд:

1. Ввести данные из порта ввода 1.

2. Запомнить данные в ячейке памяти 200.

3. Переслать данные в порт вывода 10.

В данной программе всего три команды, хотя на рис. 1.2 может показаться, что в памяти программ записано шесть команд. Это связано с тем, что команда обычно разбивается на части. Первая часть команды 1 в приведенной выше программе – команда ввода данных. Во второй части команды 1 указывается, откуда нужно ввести данные (из порта 1). Первая часть команды, предписывающая конкретное действие, называется кодом операции (КОП), а вторая часть – операндом. Коды операции и операнд размещаются в отдельных ячейках памяти программ. На рис. 1.3 КОП хранится в ячейке 100, а код операнда – в ячейке 101 (порт 1); последний указывает, откуда нужно взять информацию.

ПОРТ1

ВХОД

Аккумулятор

МП

Регистр команд

Адрес

200

202

203

Содержимое

201

“А”

11

16

ВЫХОД

Порт 10

Адрес

100

101

102

103

104

105

106

Ввести данные

Из порта 1

Запомнить данные

В ячейки памяти 200

Вывести данные

В порт 10

Содержимое

1

3

6

8

12

14

16

2

4

7

9

13

15

5

16

5

Код символа “А” с

клавиатуры

Код символа “А” к монитору

Память программ

ЗУПВ

10

Рис. 4.6. Диаграмма выполнения процедуры ввода-запоминания-вывода

В МП на рис. 4.6 выделены еще два новых блока – регистры: аккумулятор и регистр команд.

Рассмотрим прохождение команд и данных внутри микроЭВМ с помощью занумерованных кружков на диаграмме. Напомним, что микропроцессор – это центральный узел, управляющий перемещением всех данных и выполнением операций.

Итак, при выполнении типичной процедуры ввода-запоминания-вывода в микроЭВМ происходит следующая последовательность действий:

1. МП выдает адрес 100 на шину адреса. По шине управления поступает сигнал, устанавливающий память программ (конкретную микросхему) в режим считывания.

2. ЗУ программ пересылает первую команду («Ввести данные») по шине данных, и МП получает это закодированное сообщение. Команда помещается в регистр команд. МП декодирует (интерпретирует) полученную команду и определяет, что для команды нужен операнд.

3. МП выдает адрес 101 на ША; ШУ используется для перевода памяти программ в режим считывания.

4. Из памяти программ на ШД пересылается операнд «Из порта 1». Этот операнд находится в программной памяти в ячейке 101. Код операнда (содержащий адрес порта 1) передается по ШД к МП и направляется в регистр команд. МП теперь декодирует полную команду («Ввести данные из порта 1»).

5. МП, используя ША и ШУ, связывающие его с устройством ввода, открывает порт 1. Цифровой код буквы «А» передается в аккумулятор внутри МП и запоминается.Важно отметить, что при обработке каждой программной команды МП действует согласно микропроцедуре выборки-декодирования-исполнения.

6. МП обращается к ячейке 102 по ША. ШУ используется для перевода памяти программ в режим считывания.

7. Код команды «Запомнить данные» подается на ШД и пересылается в МП, где помещается в регистр команд.

8. МП дешифрирует эту команду и определяет, что для нее нужен операнд. МП обращается к ячейке памяти 103 и приводит в активное состояние вход считывания микросхем памяти программ.

9. Из памяти программ на ШД пересылается код сообщения «В ячейке памяти 200». МП воспринимает этот операнд и помещает его в регистр команд. Полная команда «Запомнить данные в ячейке памяти 200» выбрана из памяти программ и декодирована.

10. Теперь начинается процесс выполнения команды. МП пересылает адрес 200 на ША и активизирует вход записи, относящийся к памяти данных.

11. МП направляет хранящуюся в аккумуляторе информацию в память данных. Код буквы «А» передается по ШД и записывается в ячейку 200 этой памяти. Выполнена вторая команда. Процесс запоминания не разрушает содержимого аккумулятора. В нем по-прежнему находится код буквы «А».

12. МП обращается к ячейке памяти 104 для выбора очередной команды и переводит память программ в режим считывания.

13. Код команды вывода данных пересылается по ШД к МП, который помещает ее в регистр команд, дешифрирует и определяет, что нужен операнд.

14. МП выдает адрес 105 на ША и устанавливает память программ в режим считывания.

15. Из памяти программ по ШД к МП поступает код операнда «В порт 10», который далее помещается в регистр команд.

16. МП дешифрирует полную команду «Вывести данные в порт 10». С помощью ША и ШУ, связывающих его с устройством вывода, МП открывает порт 10, пересылает код буквы «А» (все еще находящийся в аккумуляторе) по ШД. Буква «А» выводится через порт 10 на экран дисплея.

В большинстве микропроцессорных систем (МПС) передача информации осуществляется способом, аналогичным рассмотренному выше. Наиболее существенные различия возможны в блоках ввода и вывода информации.

Подчеркнем еще раз, что именно микропроцессор является ядром системы и осуществляет управление всеми операциями. Его работа представляет последовательную реализацию микропроцедур выборки-дешифрации-исполнения. Однако фактическая последовательность операций в МПС определяется командами, записанными в памяти программ.

Таким образом, в МПС микропроцессор выполняет следующие функции:

- выборку команд программы из основной памяти;

- дешифрацию команд;

- выполнение арифметических, логических и других операций, закодированных в командах;

- управление пересылкой информации между регистрами и основной памятью, между устройствами ввода/вывода;

- отработку сигналов от устройств ввода/вывода, в том числе реализацию прерываний с этих устройств;

- управление и координацию работы основных узлов МП.

**4.1.4. Логическая структура микропроцессора**

Логическая структура микропроцессора, т. е. конфигурация составляющих микропроцессор логических схем и связей между ними, определяется функциональным назначением. Именно структура задает состав логических блоков микропроцессора и то, как эти блоки должны быть связаны между собой, чтобы полностью отвечать архитектурным требованиям. Срабатывание электронных блоков микропроцессора в определенной последовательности приводит к выполнению заданных архитектурой микропроцессора функций, т. е. к реализации вычислительных алгоритмов. Одни и те же функции можно выполнить в микропроцессорах со структурой, отличающейся набором, количеством и порядком срабатывания логических блоков. Различные структуры микропроцессоров, как правило, обеспечивают их различные возможности, в том числе и различную скорость обработки данных. Логические блоки микропроцессора с развитой архитектурой показаны на рис. 1.3.

БС

БУВВ

БУВП

БУПК

БУФКА

БУПРПр

БУВОп

БЗП

РгСОЗУ

АЛБ

БДА1

БДАn

II

I

Рис. 4.7. Общая логическая структура микропроцессора:

*I – управляющая часть, II – операционная часть; БУПК – блок управления последовательно-стью команд; БУВОп – блок управления выполнением операций; БУФКА ­ блок управления формированием кодов адресов; БУВП – блок управления виртуальной памятью; БЗП – блок защиты памяти; БУПРПр – блок управления прерыванием работы процессора; БУВВ ­ блок управления вводом/выводом; РгСОЗУ­ регистровое сверхоперативное запоминающее устройство; АЛБ – арифметико-логический блок; БДА – блок дополнительной арифметики; БС – блок синхронизации*

При проектировании логической структуры микропроцессоров необходимо рассмотреть:

1) номенклатуру электронных блоков, необходимую и достаточную для реализации архитектурных требований;

2) способы и средства реализации связей между электронными блоками;   
3) методы отбора наиболее рациональных вариантов логических структур из возможного числа структур с отличающимся составом блоков и конфигурацией связей между ними.

При проектировании микропроцессора приводятся в соответствие внутренняя сложность кристалла и количество выводов корпуса. Относительный рост числа элементов по мере развития микроэлектронной технологии во много раз превышает относительное увеличение числа выводов корпуса, поэтому проектирование БИС в виде конечного автомата, а не в виде набора схем, реализующих некоторый набор логических переключательных функций и схем памяти, дает возможность получить функционально законченные блоки и устройства ЭВМ.

Использование микропроцессорных комплектов БИС позволяет создать микроЭВМ для широких областей применения вследствие программной адаптации микропроцессора к конкретной области применения: изменяя программу работы микропроцессора, изменяют функции информационно-управляющей системы. Поэтому за счет составления программы работы микропроцессоров в конкретных условиях работы определенной системы можно получить оптимальные характеристики последней.

Если уровень только программной «настройки» микропроцессоров не позволит получить эффективную систему, доступен следующий уровень проектирования – микропрограммный. За счет изменения содержимого ПЗУ или программируемой логической матрицы (ПЛМ) можно «настроиться» на более специфичные черты системы обработки информации. В этом случае частично за счет изменения микропрограмм затрагивается аппаратный уровень системы. Технико-экономические последствия здесь связаны лишь с ограниченным вмешательством в технологию изготовления управляющих блоков микроЭВМ.

Изменение аппаратного уровня информационно-управляющей микропроцессорной системы, включающего в себя функциональные БИС комплекта, одновременно с конкретизацией микропрограммного и программного уровней позволяет наилучшим образом удовлетворить требованиям, предъявляемым к системе.

Решение задач управления в конкретной системе чисто аппаратными средствами (аппаратная логика) дает выигрыш в быстродействии, однако приводит к сложностям при модификации системы. Микропроцессорное решение (программная логика) является более медленным, но более гибким решением, позволяющим развивать и модифицировать систему. Изменение технических требований к информационно-управляющей микропроцессорной системе ведет лишь к необходимости перепрограммирования работы микропроцессора. Именно это качество обеспечивает высокую логическую гибкость микропроцессоров, определяет возможность их широкого использования, а значит и крупносерийного производства.

**4.1.5. Типы архитектур**

Существует несколько подходов к классификации микропроцессоров по типу архитектуры. Так, выделяют МП с CISC (CompleteInstructionSetComputer) архитектурой, характеризуемой полным набором команд, и RISC (ReduceInstructionSetComputer) архитектурой, которая определяет систему с сокращенным набором команд одинакового формата, выполняемых за один такт МП.

Определяя в качестве основной характеристики МП разрядность, выделяют следующие типы МП архитектуры:

* с фиксированной разрядностью и списком команд (однокристальные);
* с наращиваемой разрядностью (секционные) и микропрограммным управлением.

Анализируя адресные пространства программ и данных, определяют МП с архитектурой фон Неймана (память программ и память данных находятся в едином пространстве и нет никаких признаков, указывающих на тип информации в ячейке памяти) и МП с архитектурой Гарвардской лаборатории (память программ и память данных разделены, имеют свои адресные пространства и способы доступа к ним).

Для выполнения команд в структуру процессора входят внутренние регистры, *арифметико-логическое устройство* (*АЛУ*, *ALU* – *ArithmeticLogicUnit*), мультиплексоры, буферы, регистры и другие узлы. Работа всех узлов синхронизируется общим внешним тактовым сигналом процессора. То есть процессор представляет собой довольно сложное цифровое устройство ([рис. 4.](http://www.intuit.ru/department/hardware/mpbasics/1/#image.1.4)8).

Микропроцессор

Арифметическо-логическое устройство(АЛУ)

Схема управления выборкой команд

Схема управления прерываниями

RG

RG

**.**

**.**

**.**

Регистр признаков

Схема управления ПДП

Логика

управления

RG

RG

Рис. 4.8.  Пример структуры простейшего процессора

Впрочем, для разработчика *микропроцессорных систем* информация о тонкостях внутренней структуры процессора не слишком важна. Разработчик должен рассматривать процессор как «*черный ящик*», который в ответ на входные и *управляющие коды* производит ту или иную операцию и выдает выходные сигналы. Разработчику необходимо знать систему команд, режимы работы процессора, а также правила взаимодействия процессора с внешним миром или, как их еще называют, протоколы обмена информацией. О внутренней структуре процессора надо знать только то, что необходимо для выбора той или иной команды, того или иного режима работы.

**4.2. Обмен данными в микропроцессорной системе**

**4.2.1. Шинная структура связей**

Для достижения максимальной универсальности и упрощения *протоколов обмена* информацией в *микропроцессорных системах* применяется так называемая шинная структура связей между отдельными устройствами, входящими в систему. Суть шинной структуры связей сводится к следующему.

При классической структуре связей ([рис. 4.9](http://www.intuit.ru/department/hardware/mpbasics/1/2.html#image.1.5)) все сигналы и коды между устройствами передаются по отдельным линиям связи. Каждое устройство, входящее в систему, передает свои сигналы и коды независимо от других устройств. При этом в системе получается очень много линий связи и разных *протоколов обмена* информацией.

Устройство 1

Устройство 3

Устройство 2

Устройство 4

Рис. 4.10.  Классическая структура связей

При шинной структуре связей ([рис. 4.10](http://www.intuit.ru/department/hardware/mpbasics/1/2.html#image.1.6)) все сигналы между устройствами передаются по одним и тем же линиям связи, но в разное время (это называется мультиплексированной передачей). Причем передача по всем линиям связи может осуществляться в обоих направлениях (так называемая двунаправленная передача). В результате количество линий связи существенно сокращается, а правила обмена (протоколы) упрощаются. Группа линий связи, по которым передаются сигналы или коды, как раз и называется ***шиной*** (англ. bus).

Устройство 1

Устройство 2

Устройство 3

Устройство 4

Рис. 4.11.  Шинная структура связей

Понятно, что при шинной структуре связей легко осуществляется пересылка всех информационных потоков в нужном направлении, например, их можно пропустить через один процессор, что очень важно для *микропроцессорной системы*. Однако при шинной структуре связей вся информация передается по линиям связи последовательно во времени, по очереди, что снижает быстродействие системы по сравнению с классической структурой связей.

Большое достоинство шинной структуры связей состоит в том, что все устройства, подключенные к *шине*, должны принимать и передавать информацию по одним и тем же правилам (протоколам обмена информацией по *шине*). Соответственно, все узлы, отвечающие за обмен с *шиной* в этих устройствах, должны быть единообразны, унифицированы.

Существенный недостаток шинной структуры связан с тем, что все устройства подключаются к каждой линии связи параллельно. Поэтому любая неисправность любого устройства может вывести из строя всю систему, если она портит линию связи. По этой же причине отладка системы с шинной структурой связей довольно сложна и обычно требует специального оборудования.

В системах с шинной структурой связей применяют все три существующие разновидности выходных каскадов цифровых микросхем:

* стандартный выход или выход с двумя состояниями (обозначается 2С, 2S, реже ТТЛ, *TTL*);
* выход с открытым коллектором (обозначается ОК, *OC*);
* выход с тремя состояниями или (что-то же самое) с возможностью отключения (обозначается 3С, 3S).

Упрощенно эти три типа выходных каскадов могут быть представлены в виде схем на [рис. 4.](http://www.intuit.ru/department/hardware/mpbasics/1/2.html#image.1.7)12.

Выход 2С

Выход 3С

Рис. 4.12.  Три типа выходов цифровых микросхем

У выхода 2С два ключа замыкаются по очереди, что соответствует уровням логической единицы (верхний ключ замкнут) и логического нуля (нижний ключ замкнут). У выхода ОК замкнутый ключ формирует уровень логического нуля, разомкнутый – логической единицы. У выхода 3С ключи могут замыкаться по очереди (как в случае 2С), а могут размыкаться одновременно, образуя третье, высокоимпедансное состояние. Переход в третье состояние (Z-состояние) управляется сигналом на специальном входе EZ.

Выходные каскады типов 3С и ОК позволяют объединять несколько выходов микросхем для получения мультиплексированных ([рис. 4.13](http://www.intuit.ru/department/hardware/mpbasics/1/2.html#image.1.8)) или двунаправленных ([рис. 4.14](http://www.intuit.ru/department/hardware/mpbasics/1/2.html#image.1.9)) линий.

◊

Вых.

Вх.

◊

Вых.

◊

Вых.

Рис. 4.13.  Мультиплексированная линия

◊

Вх. Вых.

Вх.

◊

Вых.

◊

Вх. Вых.

Рис. 4.14.  Двунаправленная линия

При этом в случае выходов 3С необходимо обеспечить, чтобы на линии всегда работал только один активный выход, а все остальные выходы находились бы в это время в третьем состоянии, иначе возможны конфликты. Объединенные выходы ОК могут работать все одновременно, без всяких конфликтов.

Типичная структура *микропроцессорной системы* приведена на [рис. 2.10](http://www.intuit.ru/department/hardware/mpbasics/1/2.html#image.1.10). Она включает в себя три основных типа устройств:

* процессор;
* *память*, включающую оперативную ***память*** (ОЗУ, RAM –*RandomAccess*Memory) и постоянную *память*( ПЗУ, ROM –ReadOnlyMemory), которая служит для хранения данных и программ;
* ***устройства ввода/вывода*** (УВВ, I/O–Input/*OutputDevices*), служащие для связи *микропроцессорной системы* с внешними устройствами, для приема (ввода, чтения, Read) входных сигналов и выдачи (вывода, записи, Write) выходных сигналов.

Выходные

сигналы

Входные

сигналы

Устройство ввода/вывода

Память

Процессор

Шина

питания

Системная шина

(магистраль)

Шина данных

Шина

управления

Шина адреса

Рис. 4.15.  Структура микропроцессорной системы

Все устройства *микропроцессорной системы* объединяются общей системной ***шиной*** (она же называется еще **системной магистралью** или **каналом**). Системная *магистраль* включает в себя четыре основные *шины* нижнего уровня:

* *шина* адреса (*AddressBus*);
* *шина* данных (*DataBus*);
* *шина* управления (*ControlBus*);
* *шина* питания (*PowerBus*).

*Шина* адреса служит для определения адреса (номера) устройства, с которым процессор обменивается информацией в данный момент. Каждому устройству (кроме процессора), каждой ячейке *памяти* в *микропроцессорной системе* присваивается собственный адрес. Когда код какого-то адреса выставляется процессором на *шине* адреса, устройство, которому этот адрес приписан, понимает, что ему предстоит обмен информацией. *Шина* адреса может быть однонаправленной или двунаправленной.

*Шина* данных – это основная *шина*, которая используется для передачи информационных кодов между всеми устройствами *микропроцессорной системы*. Обычно в пересылке информации участвует процессор, который передает код данных в какое-то устройство или в ячейку *памяти* или же принимает код данных из какого-то устройства или из ячейки *памяти*. Но возможна также и передача информации между устройствами без участия процессора. *Шина* данных всегда двунаправленная.

*Шина* управления в отличие от *шины* адреса и *шины* данных состоит из отдельных управляющих сигналов. Каждый из этих сигналов во время обмена информацией имеет свою функцию. Некоторые сигналы служат для стробирования передаваемых или принимаемых данных (то есть определяют моменты времени, когда информационный код выставлен на *шину* данных). Другие управляющие сигналы могут использоваться для подтверждения приема данных, для сброса всех устройств в исходное состояние, для тактирования всех устройств и т.д. Линии *шины* управления могут быть однонаправленными или двунаправленными.

Наконец, *шина* питания предназначена не для пересылки информационных сигналов, а для питания системы. Она состоит из линий питания и общего провода. В *микропроцессорной системе* может быть один источник питания (чаще +5 В) или несколько источников питания (обычно еще –5 В, +12 В и –12 В). Каждому напряжению питания соответствует своя линия связи. Все устройства подключены к этим линиям параллельно.

Если в *микропроцессорную систему* надо ввести входной код (или входной сигнал), то процессор по *шине* адреса обращается к нужному *устройству ввода/вывода* и принимает по *шине* данных *входную информацию*. Если из *микропроцессорной системы* надо вывести выходной код (или выходной сигнал), то процессор обращается по *шине* адреса к нужному *устройству ввода/вывода* и передает ему по *шине* данных *выходную информацию*.

Если информация должна пройти сложную многоступенчатую обработку, то процессор может хранить промежуточные результаты в системной оперативной *памяти*. Для обращения к любой ячейке *памяти* процессор выставляет ее адрес на *шину* адреса и передает в нее информационный код по *шине* данных или же принимает из нее информационный код по *шине* данных. В *памяти* (оперативной и постоянной) находятся также и *управляющие коды* (команды выполняемой процессором программы), которые процессор также читает по *шине* данных с адресацией по *шине* адреса. Постоянная *память* используется в основном для хранения программы начального пуска *микропроцессорной системы*, которая выполняется каждый раз после включения питания. Информация в нее заносится изготовителем раз и навсегда.

Таким образом, в *микропроцессорной системе* все информационные коды и коды команд передаются по *шинам* последовательно, по очереди. Это определяет сравнительно невысокое быстродействие *микропроцессорной системы*. Оно ограничено обычно даже не быстродействием процессора (которое тоже очень важно) и не скоростью обмена по системной *шине* (магистрали), а именно последовательным характером передачи информации по системной *шине* (магистрали).

Важно учитывать, что *устройства ввода/вывода* чаще всего представляют собой устройства на «жесткой логике». На них может быть возложена часть функций, выполняемых *микропроцессорной системой*. Поэтому у разработчика всегда имеется возможность перераспределять функции системы между аппаратной и программной реализациями оптимальным образом. Аппаратная реализация ускоряет выполнение функции, но имеет недостаточную гибкость. Программная реализация значительно медленнее, но обеспечивает высокую гибкость. Аппаратная реализация функций увеличивает стоимость системы и ее *энергопотребление*, программная –не увеличивает. Чаще всего применяется комбинирование аппаратных и программных функций.

Иногда *устройства ввода/вывода* имеют в своем составе процессор, то есть представляют собой небольшую специализированную *микропроцессорную систему*. Это позволяет переложить часть программных функций на *устройства ввода/вывода*, разгрузив центральный процессор системы.

**4.2.2. Режимы работы микропроцессорной системы.** Как уже отмечалось, *микропроцессорная система* обеспечивает большую гибкость работы, она способна настраиваться на любую задачу. Гибкость эта обусловлена прежде всего тем, что функции, выполняемые системой, определяются программой (программным обеспечением, software), которую выполняет процессор. Аппаратура (аппаратное обеспечение, hardware) остается неизменной при любой задаче. Записывая в *память* системы программу, можно заставить *микропроцессорную систему* выполнять любую задачу, поддерживаемую данной аппаратурой. К тому же шинная *организация связеймикропроцессорной системы* позволяет довольно легко заменять аппаратные модули, например, заменять *память* на новую большего объема или более высокого быстродействия, добавлять или модернизировать *устройства ввода/вывода*, наконец, заменять процессор на более мощный. Это также позволяет увеличить гибкость системы, продлить ее жизнь при любом изменении требований к ней.

Но гибкость *микропроцессорной системы* определяется не только этим. Настраиваться на задачу помогает еще и выбор режима работы системы, то есть режима обмена информацией по системной магистрали ( *шине* ).

Практически любая развитая *микропроцессорная система* (в том числе и компьютер) поддерживает три основных режима обмена по магистрали:

* программный обмен информацией;
* обмен с использованием прерываний (*Interrupts*);
* обмен с использованием прямого доступа к *памяти* (*ПДП*, *DMA* — *DirectMemoryAccess*).

**Программный обмен информацией** является основным в любой *микропроцессорной системе*. Он предусмотрен всегда, без него невозможны другие режимы обмена. В этом режиме процессор является единоличным хозяином (или задатчиком, Master) системной магистрали. Все операции (циклы) обмена информацией в данном случае инициируются только процессором, все они выполняются строго в порядке, предписанном исполняемой программой.

Процессор читает (выбирает) из *памяти* коды команд и исполняет их, читая данные из *памяти* или из *устройства ввода/вывода*, обрабатывая их, записывая данные в *память* или передавая их в *устройство ввода/вывода*. Путь процессора по программе может быть линейным, циклическим, может содержать переходы (прыжки), но он всегда непрерывен и полностью находится под контролем процессора. Ни на какие внешние события, не связанные с программой, процессор не реагирует (рис. 4.16). Все сигналы на магистрали в данном случае контролируются процессором.

Команда 1

Команда 2

Команда 3

●

●

●

Команда N

Команда N+1

Команда N+2

Команда N+3

Команда N+4

●

●

●

Выполнение команд

Программа

Последовательно

Цикл

Прыжок

**Рис. 4.16.**  Программный обмен информацией.

**Обмен по прерываниям** используется тогда, когда необходима реакция *микропроцессорной системы* на какое-то внешнее событие, на приход внешнего сигнала. В случае компьютера внешним событием может быть, например, нажатие на клавишу клавиатуры или приход по локальной сети пакета данных. Компьютер должен реагировать на это, соответственно, выводом символа на экран или же чтением и обработкой принятого по сети пакета.

В общем случае организовать реакцию на внешнее событие можно тремя различными путями:

* с помощью постоянного программного контроля факта наступления события (так называемый метод опроса флага или *polling*);
* с помощью прерывания, то есть насильственного перевода процессора с выполнения текущей программы на выполнение экстренно необходимой программы;
* с помощью прямого доступа к *памяти*, то есть без участия процессора при его отключении от системной магистрали.

Проиллюстрировать эти три способа можно следующим простым примером. Допустим, вы готовите себе завтрак, поставив на плиту кипятиться молоко. Естественно, на закипание молока надо реагировать, причем срочно. Как это организовать? Первый путь — постоянно следить за молоком, но тогда вы ничего другого не сможете делать. Правильнее будет регулярно поглядывать на молоко, делая одновременно что-то другое. Это программный режим с опросом флага. Второй путь — установить на кастрюлю с молоком датчик, который подаст звуковой сигнал при закипании молока, и спокойно заниматься другими делами. Услышав сигнал, вы выключите молоко. Правда, возможно, вам придется сначала закончить то, что вы начали делать, так что ваша реакция будет медленнее, чем в первом случае. Наконец, третий путь состоит в том, чтобы соединить датчик на кастрюле с управлением плитой так, чтобы при закипании молока горелка была выключена без вашего участия (правда, аналогия с *ПДП* здесь не очень точная, так как в данном случае на момент выполнения действия вас не отвлекают от работы).

Первый случай с опросом флага реализуется в *микропроцессорной системе* постоянным чтением информации процессором из *устройства ввода/вывода*, связанного с тем внешним устройством, на поведение которого необходимо срочно реагировать.

Во втором случае в режиме прерывания процессор, получив запрос прерывания от внешнего устройства (часто называемый *IRQ* — *InterruptReQuest*), заканчивает выполнение текущей команды и переходит к программе обработки прерывания. Закончив выполнение программы обработки прерывания, он возвращается к прерванной программе с той точки, где его прервали ([рис. 4.1](http://www.intuit.ru/department/hardware/mpbasics/1/3.html#image.1.12)7).

Здесь важно то, что вся работа, как и в случае программного режима, осуществляется самим процессором, внешнее событие просто временно отвлекает его. Реакция на внешнее событие по прерыванию в общем случае медленнее, чем при программном режиме. Как и в случае программного обмена, здесь все сигналы на магистрали выставляются процессором, то есть он полностью контролирует *магистраль*. Для обслуживания прерываний в систему иногда вводится специальный модуль *контроллера прерываний*, но он в обмене информацией не участвует. Его задача состоит в том, чтобы упростить работу процессора с внешними *запросами прерываний*. Этот контроллер обычно программно управляется процессором по системной магистрали.

Команда 1

Команда 2

Команда 3

●

●

●

Команда N

Команда N+1

Команда N+2

●

●

●

Команда 1

Команда 2

Команда 3

●

●

●

Команда M

Программа обработки прерывания

Запрос прерывания

Выполнение команд

Выполнение команд

Основная программа

**Рис. 4.17.**  Обслуживание прерывания.

Естественно, никакого ускорения работы системы прерывание не дает. Его применение позволяет только отказаться от постоянного опроса флага внешнего события и временно, до наступления внешнего события, занять процессор выполнением каких-то других задач.

**Прямой доступ к памяти (ПДП, DMA)** — это режим, принципиально отличающийся от двух ранее рассмотренных режимов тем, что обмен по системной *шине* идет без участия процессора. Внешнее устройство, требующее обслуживания, сигнализирует процессору, что режим *ПДП* необходим, в ответ на это процессор заканчивает выполнение текущей команды и отключается от всех *шин*, сигнализируя запросившему устройству, что обмен в режиме *ПДП* можно начинать.

Операция *ПДП* сводится к пересылке информации из *устройства ввода/вывода* в *память* или же из *памяти* в *устройство ввода/вывода*. Когда пересылка информации будет закончена, процессор вновь возвращается к прерванной программе, продолжая ее с той точки, где его прервали ([рис. 1.13](http://www.intuit.ru/department/hardware/mpbasics/1/3.html#image.1.13)). Это похоже на режим обслуживания прерываний, но в данном случае процессор не участвует в обмене. Как и в случае прерываний, реакция на внешнее событие при *ПДП* существенно медленнее, чем при программном режиме.

Понятно, что в этом случае требуется введение в систему дополнительного устройства (контроллера *ПДП*), которое будет осуществлять полноценный обмен по системной магистрали без всякого участия процессора. Причем процессор предварительно должен сообщить этому контроллеру *ПДП*, откуда ему следует брать информацию и/или куда ее следует помещать. Контроллер *ПДП* может считаться специализированным процессором, который отличается тем, что сам не участвует в обмене, не принимает в себя информацию и не выдает ее ([рис. 4.1](http://www.intuit.ru/department/hardware/mpbasics/1/3.html#image.1.14)8).

**Рис. 4.18.**  Обслуживание ПДП.

Команда 1

Команда 2

Команда 3

●

●

●

Команда N

Команда N+1

Команда N+2

Команда N+3

Команда N+4

●

●

●

Выполнение команд

Запрос ПДП

Точка остановки на время ПДП

Продолжение

выполнения команд

Программа

Процессор

Контроллер

ПДП

Память

Устройство

Ввода/выводы

Системная магистраль

**Рис. 4.19.**  Информационные потоки в режиме ПДП.

В принципе контроллер *ПДП* может входить в состав *устройства ввода/вывода*, которому необходим режим *ПДП* или даже в состав нескольких *устройств ввода/вывода*. Теоретически обмен с помощью прямого доступа к *памяти* может обеспечить более высокую скорость передачи информации, чем программный обмен, так как процессор передает данные медленнее, чем специализированный контроллер *ПДП*. Однако на практике это преимущество реализуется далеко не всегда. Скорость обмена в режиме *ПДП* обычно ограничена возможностями магистрали. К тому же необходимость программного задания режимов контроллера *ПДП* может свести на нет выигрыш от более высокой скорости пересылки данных в режиме *ПДП*. Поэтому режим *ПДП* применяется редко.

Если в системе уже имеется самостоятельный контроллер *ПДП*, то это может в ряде случаев существенно упростить аппаратуру *устройств ввода/вывода*, работающих в режиме *ПДП*. В этом, пожалуй, состоит единственное бесспорное преимущество режима *ПДП*.

**1.4. Архитектура микропроцессорных систем.** До сих пор мы рассматривали только один тип *архитектуры*   *микропроцессорных систем* — *архитектуру* с общей, единой *шиной* для данных и команд (одношинную, или **принстонскую**, **фон-неймановскую**   *архитектуру* ). Соответственно, в составе системы в этом случае присутствует одна общая *память*, как для данных, так и для команд ([рис. 4.](http://www.intuit.ru/department/hardware/mpbasics/1/4.html#image.1.15)20).

Процессор

Устройства

Ввода/вывода

Память

Команды

Данные

Общая шина данных и команд

**Рис. 4.20.**  Архитектура с общей шиной данных и команд.

Но существует также и альтернативный тип *архитектурымикропроцессорной системы* — это *архитектура* с раздельными *шинами* данных и команд (двухшинная, или гарвардская, *архитектура* ). Эта *архитектура* предполагает наличие в системе отдельной *памяти* для данных и отдельной *памяти* для команд ([рис. 1.16](http://www.intuit.ru/department/hardware/mpbasics/1/4.html#image.1.16)). Обмен процессора с каждым из двух типов *памяти* происходит по своей *шине*.

*Архитектура* с общей *шиной* распространена гораздо больше, она применяется, например, в персональных компьютерах и в сложных *микрокомпьютерах*. *Архитектура* с раздельными *шинами* применяется в основном в однокристальных *микроконтроллерах*.

Рассмотрим некоторые достоинства и недостатки обоих архитектурных решений.

*Архитектура* с общей *шиной* (принстонская, фон-неймановская) проще, она не требует от процессора одновременного обслуживания двух *шин*, контроля обмена по двум *шинам* сразу. Наличие единой *памяти* данных и команд позволяет гибко распределять ее объем между кодами данных и команд. Например, в некоторых случаях нужна большая и сложная программа, а данных в *памяти* надо хранить не слишком много. В других случаях, наоборот, программа требуется простая, но необходимы большие объемы хранимых данных. Перераспределение *памяти* не вызывает никаких проблем, главное — чтобы программа и данные вместе помещались в *памяти* системы. Как правило, в системах с такой *архитектурой память* бывает довольно большого объема (до десятков и сотен мегабайт). Это позволяет решать самые сложные задачи.

Шина команд

Память команд

Процессор

Устройства

ввода/вывода

Память данных

Шина данных

**Рис. 4.21.**  Архитектура с раздельными шинами данных и команд.

*Архитектура* с раздельными *шинами* данных и команд сложнее, она заставляет процессор работать одновременно с двумя потоками кодов, обслуживать обмен по двум *шинам* одновременно. Программа может размещаться только в *памяти* команд, данные — только в *памяти* данных. Такая узкая специализация ограничивает круг задач, решаемых системой, так как не дает возможности гибкого перераспределения *памяти*. *Память* данных и *память* команд в этом случае имеют не слишком большой объем, поэтому применение систем с данной *архитектурой* ограничивается обычно не слишком сложными задачами.

В чем же преимущество *архитектуры* с двумя *шинами* (гарвардской)? В первую очередь, в быстродействии.

Дело в том, что при единственной *шине* команд и данных процессор вынужден по одной этой *шине* принимать данные (из *памяти* или *устройства ввода/вывода* ) и передавать данные (в *память* или в *устройство ввода/вывода* ), а также читать команды из *памяти*. Естественно, одновременно эти пересылки кодов по магистрали происходить не могут, они должны производиться по очереди. Современные процессоры способны совместить во времени выполнение команд и проведение циклов обмена по системной *шине*. Использование конвейерных технологий и быстрой кэш-памяти позволяет им ускорить процесс взаимодействия со сравнительно медленной системной *памятью*. Повышение тактовой частоты и совершенствование структуры процессоров дают возможность сократить время выполнения команд. Но дальнейшее увеличение быстродействия системы возможно только при совмещении пересылки данных и чтения команд, то есть при переходе к *архитектуре* с двумя *шинами*.

В случае двухшинной *архитектуры* обмен по обеим *шинам* может быть независимым, параллельным во времени. Соответственно, структуры *шин* (количество разрядов кода адреса и кода данных, порядок и скорость обмена информацией и т.д.) могут быть выбраны оптимально для той задачи, которая решается каждой *шиной*. Поэтому при прочих равных условиях переход на двухшинную *архитектуру* ускоряет работу *микропроцессорной системы*, хотя и требует дополнительных затрат на аппаратуру, усложнения структуры процессора. *Память* данных в этом случае имеет свое распределение адресов, а *память* команд — свое.

Проще всего преимущества двухшинной *архитектуры* реализуются внутри одной микросхемы. В этом случае можно также существенно уменьшить влияние недостатков этой *архитектуры*. Поэтому основное ее применение — в *микроконтроллерах*, от которых не требуется решения слишком сложных задач, но зато необходимо максимальное быстродействие при заданной тактовой частоте.

**4. Архитектура 16-тиразрядного микропроцессора**

Мы рассмотрим более подробно основные типы архитектурных решений, выделяя связь со способами адресации памяти. 1. Регистровая архитектура определяется наличием достаточно большого регистрового файла внутри МП. Команды получают возможность обратиться к операндам, расположенным в одной из двух запоминающих сред: оперативной памяти или регистрах. Размер регистра обычно фиксирован и совпадает с размером слова, физически реализованного в оперативной памяти. К любому регистру можно обратиться непосредственно, поскольку регистры представлены в виде массива запоминающих элементов – регистрового файла. Типичным является выполнение арифметических операций только в регистре, при этом команда содержит два операнда (оба операнда в регистре или один операнд в регистре, а второй в оперативной памяти).

**4.1. Структура микропроцессора Intel8080**

Intel8080 является однокристальным микропроцессором, работающим с 8-разрядной шиной данных и 16-разрядной шиной адреса. Управляющие сигналы передаются по шине управления. Шины отделены друг от друга. Структура Intel 8080 приведена на рис. 4.22.

Буфер данных

Аккумулятор

Буфер аккумулятора

Арифметико-

Логическое устройство

Регистр

флагов

Схема инкремента

JP(16)

SP(16)

B(8) C(8)

D(8) E(8)

H(8) L(8)

W(8) Z(8)

Буфер адреса

Дешифратор выбора регистров

Устройство

управления

Intel 8080

Дешифратор

команд

Регистр

команд

16

8

Внутренняя шина данных

8

Схема десятичной коррекции

8

RC

Управление обменом информацией

TR

Рис. 4.22. Внутренняя структура микропроцессораIntel8080

Микропроцессор содержит внутреннюю шину данных, посредством которой происходит обмен информацией между внутренними регистрами, арифметико-логическим устройством, обрабатывающим 8-разрядные данные и передающим их через буфер на внешнюю шину данных. Кроме того, в состав 8080 входит устройство управления, буфер адресной шины, связанный с регистром команд, и блок регистров. В общих чертах работа микропроцессора выглядит следующим образом: в регистре, называемом программным счетчиком, хранится адрес следующей команды, которую необходимо выполнить. Устройство управления подключает этот регистр к шине адреса (конечно, через буфер) и выдает управляющие сигналы, необходимые для чтения кода команды из памяти. На этом завершается первый такт.

В следующем такте микропроцессор проверяет состояние сигналов на входе готовности и запрос останова. При их наличии микропроцессор переходит в соответствующее состояние. В противном случае, после появления на шине управления сигналов, подтверждающих выдачу кода команды на шину данных, устройство управления подключает к ней регистр команды и записывает в нее полученный код. Это требуется потому, что команда передается только в первом машинном цикле, а сохранить ее нужно на все время выполнения команды. Из регистра команды ее код поступает в дешифратор команды и затем в устройство управления, которое в зависимости от поступившей команды либо сразу переходит к ее выполнению, либо считывает данные или адрес, расположенные сразу после кода команды и необходимые для ее выполнения. На это тратится третий такт и, если это необходимо, четвертый и пятый такты. Таким образом, вся команда выполняется за 3–5 тактов. При тактовой частоте 2 МГц это составляет 1,5 – 2,5 мкс.

Перед выполнением команды проверяется состояние сигнала на входе захвата шины HLD (этот сигнал отключает микропроцессор от шины, давая внешним устройствам возможность прямого доступа в память по общей шине). При его обнаружении микропроцессор переходит в состояние захвата и вырабатывает сигнал подтверждения захвата шины. После снятия сигнала HLD процессор продолжает выполнение команды. В конце машинного цикла вновь анализируется состояние входа захвата, а затем проверяется, завершено ли выполнение команды. Если нет, то микропроцессор переходит к выполнению следующего цикла команды. Это может быть цикл работы с оперативной памятью или с внешним устройством.

После выполнения каждой команды проверяется состояние входа запроса прерывания. Если этот сигнал присутствует, то текущая программа приостанавливается и на шину данных выдается сигнал «подтверждение прерывания». Затем внешний контроллер прерываний передает по шине данных команду и адрес перехода к подпрограмме обработки прерывания. По окончании обработки прерывания происходит возврат к выполнению прерванной программы.

Сигнал готовности позволяет синхронизировать работу микропроцессора с работой более медленных устройств, и, кроме того, используется для пошагового выполнения программ.

Следующим важным узлом микропроцессора является блок регистров. Он включает в себя 16-разрядный регистр для временного хранения данных WZ, шесть 8-разрядных регистров общего назначения B, C, D, E, H, L, которые могут использоваться парами в качестве 16-разрядных – BC, DE, HL (это сделано прежде всего для удобной работы с адресами). Кроме того, блок регистров содержит 16-разрядный регистр адреса команды IP (программный счетчик), 16-разрядный регистр указателя стека SP, а также 16-разрядную схему инкремента-декремента. С помощью последней изменяется, например, состояние программного счетчика после выполнения каждой следующей команды.

Еще один важный узел – регистр результата (аккумулятор), связанный с АЛУ и используемый для хранения одного из исходных операндов или результата выполнения команды.

Последний регистр –это регистр флажков. В нем записан байт, каждый бит которого содержит информацию о результате выполнения последней команды.

**4.2. Архитектура микропроцессора 8086**

Микропроцессор Intel 8086 существенно отличается от i8080. В нем применена новая значительно более мощная и гибкая система команд, есть возможность адресации 1 Мбайта памяти, обращения к 65536 устройствам ввода и такому же количеству устройств вывода информации.

В i8086 имеется возможность изменения внутренней аппаратной конфигурации с помощью специального управляющего сигнала. В более простом режиме 8086 ориентирован на использование в простых вычислительных и управляющих устройствах. При этом микропроцессор сам вырабатывает сигналы управления шиной и обеспечивает прямой доступ к ней посредством контроллера Intel 8257. В режиме полной конфигурации обеспечивается работа с контроллером шины 8288, который декодирует три сигнала состояния процессора и в зависимости от них выдает семь сигналов управления шиной. Такой режим используется в мультипроцессорных системах и в сложных вычислительных устройствах, в частности, в компьютере IBM PC/XT.

Интересно организована память: хранение 16-разрядных слов осуществляется в виде отдельных байтов, причем байты, передающиеся по восьми младшим линиям шины данных (D7-D0), собраны в банк 0, а передаваемые по восьми старшим линиям – в банк 1. Объем каждого банка составляет 512 Кбайт. Таким образом, нечетные байты хранятся в банке 1, а четные– в банке 0. Выбор банка осуществляется с помощью младшего адреса и сигнала управления старшими разрядами шины данных.

Еще одна важная особенность –возможность обработки 256 типов прерываний (от 0 до 255), в том числе есть прерывания, определяемые пользователем, и пошаговые прерывания.

Микропроцессор Intel 8086 приспособлен для работы с несколькими процессорами в одной системе, причем возможно использование как независимых процессоров, так и сопроцессоров. Отличие заключается в том, что независимый процессор выполняет свою собственную последовательность команд, а сопроцессор следит за потоком команд центрального процессора и выделяет из него «свои» команды, расширяя набор команд основного процессора и улучшая таким образом характеристики системы. Для поддержки этих режимов используются команды ESC, LOCK и XCHG, а также специальные управляющие сигналы, позволяющие разрешать конфликты доступа к общим ресурсам.

Внешние шины адреса и данных в 8086 объединены, и поэтому наличие на шине в данный момент времени информации или адреса определяется порядковым номером такта внутри цикла. Процессор ориентирован на параллельное выполнение команды и выборки следующей команды. В целом выполнение команды происходит примерно так же, как и в 8080. Команда выбирается из памяти и принимается микропроцессором в свободный регистр очереди команд, причем в то же самое время выполняется предыдущая команда. Конвейеризация команд позволяет значительно повысить быстродействие системы. При выполнении команд проверяются состояния входов запросов прерываний и захвата шины, и при необходимости выполняются соответствующие действия.

Микропроцессор i8086 состоит из трех основных частей: устройства сопряжения шины, устройства обработки и устройства управления и синхронизации.

Устройство сопряжения шины состоит из шести 8-разрядных регистров очереди команд, четырех 16-разрядных регистров адреса команды, 16-разрядного регистра команды и 16-разрядного сумматора адреса (см. рис. 2.12). Оно выполняет следующие функции: выбирает команды из памяти и записывает их в регистр очереди команд, вычисляет и формирует физический адрес, читает операнды из памяти или из регистров и записывает результат выполнения команд в память или в регистры.

А0/D0-A15/D15

16

А16-А19

4

Устройство управления и синхронизации

Буфер шины адреса/данных

Внутренняя шина команд

Внутренняя шина

Регистры очереди команд (8-разрядные)

1

2

3

4

5

6

16

Сумматор адреса

CS

DS

SS

ES

IP

Регистр обмена

Внутренняя шина данных

16

Микропрограммное устройство управления

Арифметико-

логическое устройство

Регистр флагов

AH AL

BH BL

CH CL

DH DL

SP

BP

SI

DI

A

B

C

D

Рис. 4.23. Внутренняя структура микропроцессора 8086

Устройство обработки преобразует данные. Команда из очереди команд по запросу устройства обработки поступает на внутреннюю шину команд, а с нее на микропрограммное устройство управления, декодирующее ее и генерирующее соответствующие последовательности микрокоманд, необходимые для выполнения текущей операции. В отличие от первых микропроцессоров, устройство обработки в 8086 не связано с внешней шиной, а обменивается с ней информацией через регистр обмена устройства сопряжения шины.

Устройство обработки содержит 16-разрядное арифметико-логическое устройство, восемь 16-разрядных регистров общего назначения и 16-разрядный регистр флагов. Регистры могут использоваться как 16-разрядные или как пары 8-разрядных (при этом их количество удваивается).

2.6. Архитектура микропроцессора Intel80286

Кристалл 80286 представляет для читателя интерес прежде всего потому, что является, пожалуй, наиболее распространенным микропроцессором из применяющихся в персональных компьютерах. Как и его предшественник – 8086 – он имеет 16-разрядные шины данных и адреса и самым характерным его отличием можно считать, помимо большей тактовой частоты, возможность работы в режиме виртуальной адресации (адресация памяти объемом более 1 Мбайта), речь о котором пойдет ниже.

**Регистры**

Как и любой процессор, Intel-80286 содержит некоторое количество ячеек памяти быстрого доступа, называемых регистрами. В состав i802286 входят три набора по четыре регистра и один специальный регистр – указатель команды.

***Регистры общего назначения***

Первый набор включает в себя регистры общего назначения или РОН, необходимые для временного хранения тех операндов и результатов вычислений, доступ к которым постоянно повторяется в процессе выполнения программы. Использование РОН в подобных случаях существенно ускоряет работу системы за счет сокращения времени чтения/записи и пересылки данных из ОЗУ. Всего регистров общего назначения четыре, они, разумеется, 16-разрядные, но могут использоваться и как 8-разрядные (однобайтные), при этом их количество удваивается.

Функции всех РОН, в основном, идентичны, но в некоторых случаях архитектура предполагает их строгую специализацию. Например, при выполнении команд обработки строк и циклов, в одном из регистров должно храниться число, равное количеству итераций. Этот регистр выполняет роль счетчика (counter) и носит название CX. Остальные регистры выполняют функции аккумулятора (AX), базы (BX) и ячейки временного хранения данных (DX). Как мы уже знаем, каждый регистр из числа РОН может быть разделен на два однобайтных, один из которых (0-7) называется младшим (Low), а другой (7-15) – старшим (High). В соответствии с этим, каждый 8-разрядный регистр получил свое название: младшие именуются AL, BL, CL, DL, а старшие – AН, BН, CН и DH (рис. 2.13).

AH AL

BH BL

CH CL

DH DL

AX

BX

CX

DX

Аккумулятор

База

Счетчик

Данные

7

7

0

0

Рис. 4.24 Регистры общего назначения

Перед тем как познакомиться с назначением и функциями остальных наборов регистров, разберемся, каким образом процессору с 16-разрядной шиной адреса удается работать с памятью объемом в 1 Мбайт.

***Режим реального адреса***. Адресная шина процессора 80286 имеет ширину 16 бит, к тому же известно, что максимальное двоичное число длиной в два байта равно 216 или 64 Кбайт и, если адрес задается таким числом, то, вроде бы, пространство ОЗУ, с которым может работать процессор, не должно превышать 64 Кбайт. С другой стороны, 1 Мбайт памяти можно адресовать с помощью двоичного числа длиной 20 бит (220). Как быть?

Можно, например, воспользоваться двойным адресом, ведь в повседневной жизни нам приходится постоянно сталкиваться с многоступенчатыми адресами: мы пишем на почтовом конверте сначала название города, потом – улицы, дома и т.д. (Предположим на мгновенье, что все квартиры в СССР пронумерованы последовательно, каково придется почте в такой ситуации?) Разработчики 80286 решили проблему подобным же образом: полный адрес ячейки памяти состоит из комбинации двух 16-разрядных чисел, причем одно из них предназначили для адресации внутри некоторой области ОЗУ размером 64 Кбайта, а второе – для локализации этой области во всем пространстве ОЗУ. Область, внутри которой происходит адресация, называется сегментом, а адрес внутри сегмента – внутрисегментным смещением. Адрес, локализующий положение сегмента в оперативной памяти, содержится в одном из специальных сегментных регистров процессора, но он тоже 16-разрядный. Для того, чтобы при помощи этого адреса можно было перекрыть все пространство ОЗУ, со стороны младшего байта его дополняют четырьмя нулями. Например, если содержимое сегментного регистра: 0001.1101.1000.1111 (или 1D8F16), то адрес начала соответствующего сегмента будет равен: 0001.1101.1000.1111.0000 (или 1D8F016). Таким образом можно искусственно разделить всю память на сегменты, начинающиеся по адресам, кратным 1610. Предположим, что внутрисегментное смещение нашей ячейки задано числом 1001.1011.0010.0101 или 9В2516, в этом случае ее реальный адрес будет равен сумме адреса сегмента и внутрисегментного смещения: 1D8F016+9B2516 = 2701516. Схема такого формированияреального адреса будет рассмотрена в следующей лекции.

Выполняемая программа может обращаться к любому из четырех сегментов, именуемых: текущий сегмент кода (то есть программы), текущий сегмент данных, текущий сегмент стека и текущий дополнительный сегмент.

Теперь вернемся к регистрам.

***Указательные и индексные регистры***

Второй набор состоит из четырех 16-разрядных регистров, которые, в основном, используются для хранения внутрисегментных смещений. При выполнении многих команд функции каждого из данных регистров строго определены.

Регистры SI (индекс источника) и DI (индекс приемника) называются указательными и содержат смещения в текущем сегменте данных. Регистры SP (указатель стека) и ВР (указатель базы) называются базовыми и содержат смещения в текущем сегменте стека. Для тех, кто забыл или не знает, следует напомнить, что стек – это способ организации работы с оперативной памятью по принципу LIFO (LastIn-FirstOut), что в переводе означает: «последним вошел – первым вышел», то есть слово данных, помещенное в стек последним, будет извлечено оттуда в первую очередь. Само слово стек произошло от английского stack–скирда. И действительно, область памяти, организованную в виде стека, можно сравнить со скирдой сена: последние, уложенные сверху снопы, будут использованы первыми – никому не придет в голову выгребать сено из середины. Для работы со стеком необходимо знать две величины: адрес дна и адрес вершины стека. Если адрес дна – число фиксированное, то адрес вершины зависит от того, сколько байтов данных содержится в стеке. В нашем случае адрес вершины находится в регистре SP.

Регистры SP и BP содержат указатели стека и базы соответственно, регистры SI и DI – смещения, соответственно, источника и приемника при выполнении команд обработки строк MOVSB, MOVSW, LODSB, LODSW, STOSB, STOSW и команды LOOP.

***Сегментные регистры***

Если мы вспомним, что программа в любой момент может обратиться к одному из четырех сегментов: к текущему сегменту кода, данных, стека или к дополнительному (сегменту данных), то нас вряд ли удивит, что в состав процессора входят четыре 16-разрядных регистра, являющихся указателями адресов текущих Сегментов. Их функции строго дифференцированы, а потому каждый регистр имеет свою «профессию»: CS определяет сегмент кода, DS – сегмент данных, SS – сегмент стека и ES – дополнительный сегмент.

Теперь для того, чтобы, к примеру, произвести выборку слова данных из стека, программе достаточно обратиться к регистрам SS и SP, сложить находящиеся в них числа по уже известному нам правилу и в качестве результата получить реальный адрес вершины стека.

***Флажки***

Совокупность значений флажков и определяет состояние процессора во время его работы. В самом общем случае слово состояния – это двоичное число, каждый бит которого отражает строго определенный параметр состояния устройства. Что касается 80286, то здесь биты слова состояния называются флажками, всего их девять, причем шесть из них регистрируют состояние процессора, а три применяются для управления его работой.

C

T

S

NT

IOPL

O

D

I

Z

A

P

IP

Флажки

Указатель команды

15

0

Рис.4.25 Указатель команды и флажки

Верхний регистр IP является указателем команды, нижний – регистром флажков. К флажкам состояния относятся: флажок переноса CF (имеет значение равное 1 при переносе из старшего бита), флажок вспомогательного переноса AF (индицирует перенос из младших 4-х бит), флажок переполнения OF (устанавливается равным единице при выходе знакового результата за границу диапазона), флажок нуля ZF (фиксирует нулевой результат выполнения команды), флажок знака SF (фиксирует отрицательный результат выполнения команды) флажок четности PF (фиксирует четное число единиц в последнем байте, полученном в результате выполнения команды).

К флажкам управления относятся: флажок направления DF (указывает направление прохождения строк в строковых командах), флажок разрешения прерывания IF (разрешает или запрещает прерывание по входу INTR), флажок трассировки TF (переводит процессор в пошаговый режим).

Размещение флажков в разрядах регистра слова состояния будет также рассмотрено в следующей лекции.

5. УРОВЕНЬ АРХИТЕКТУРЫ НАБОРА КОМАНД

5.1. Система команд

5.1.1. Классификация команд

Проектирование системы команд оказывает влияние на структуру ЭВМ. Оптимальную систему команд иногда определяют как совокупность команд, которая удовлетворяет требованиям проблемно-ориентированных применений таким образом, что избыточность аппаратных и аппаратно-программных средств на реализацию редко используемых команд оказывается минимальной. В различных программах ЭВМ частота появления команд различна; например, по данным фирмы DEC в программах для ЭВМ семейства PDP-11 наиболее часто встречается команда передачи MOV(B), на ее долю приходится приблизительно 32% всех команд в типичных программах. Систему команд следует выбирать таким образом, чтобы затраты на редко используемые команды были минимальными.

При наличии статистических данных можно разработать (выбрать) ЭВМ с эффективной системой команд. Одним из подходов к достижению данной цели является разработка команд длиной в одно слово и кодирование их таким образом, чтобы разряды таких коротких команд использовать оптимально, что позволит сократить время реализации программы и ее длину.

Другим подходом к оптимизации системы команд является использование микроинструкций. В этом случае отдельные биты или группы бит команды используются для кодирования нескольких элементарных операций, которые выполняются в одном командном цикле. Эти элементарные операции не требуют обращения к памяти, а последовательность их реализации определяется аппаратной логикой.

Сокращение времени выполнения программ и емкости памяти достигается за счет увеличения сложности логики управления.

Важной характеристикой команды является ее формат, определяющий структурные элементы команды, каждый из которых интерпретируется определенным образом при ее выполнении. Среди таких элементов (полей) команды выделяют следующие: код операции, определяющий выполняемое действие; адрес ячейки памяти, регистра процессора, внешнего устройства; режим адресации; операнд при использовании непосредственной адресации; код анализируемых признаков для команд условного перехода.

Классификация команд по основным признакам представлена на рис. 2.1. Важнейшим структурным элементом формата любой команды является код операции (КОП), определяющийдействие, которое должно быть выполнено. Большое число КОП в процессоре очень важно, так как аппаратная реализация команд экономит память и время. Но при выборе ЭВМ необходимо концентрировать внимание на полноте операций с конкретными типами данных, а не только на числе команд, на доступных режимах адресации. Число бит, отводимое подКОП, является функцией полного набора реализуемых команд.

Команды

По функциональному числу

По числу

адресов

По способу кодирования

команд

По длине команды

По способу

адресации

Передача данных

Обработка данных

Передача управления

Дополнительные

Безадресные

(нуль адресные)

Одноадресные

Двухадресные

С фиксированным полем КОП

С расширяющимся полем КОП

Длиной в одно слово(байт)

Длиной в два слова(байта)

Длиной в три слова(байта)

Рис. 5.1. Классификация команд.

При использовании фиксированного числа бит под КОП для кодирования всех *m* команд необходимо в поле КОП выделитьn двоичных разрядов, чтобы 2*n*=*m*. Однако, учитывая ограниченную длину слова мини- и микроЭВМ, различное функциональное назначение команд, источники и приемники результатов операций, а также то, что не все команды содержат адресную часть для обращения к памяти и периферийным устройствам, в малых ЭВМ для кодирования команд широко используется принцип кодирования с переменным числом бит под поле КОП для различных групп команд.

В некоторых командах необходим только один операнд и они называются однооперандными (или одноадресными) командами в отличие от двухоперандных (или двухадресных), в которых требуются два операнда. При наличии двух операндов командой обычно изменяется только один из них. Так как информация берется только из одной ячейки, эту ячейку называют источником; ячейку, содержимое которой изменяется, называют приемником.

По форматам команд можно судить о возможностях ЭВМ.

5.1.2. Режимы адресации

Для взаимодействия с различными модулями в ЭВМ должны быть средства идентификации ячеек внешней памяти, ячеек внутренней памяти, регистров МП и регистров устройств ввода/вывода. Поэтому каждой из запоминающих ячеек присваивается адрес, т.е. однозначная комбинация бит. Количество бит определяет число идентифицируемых ячеек. Обычно ЭВМ имеет различные адресные пространства памяти и регистров МП, а иногда –отдельные адресные пространства регистров устройств ввода/вывода и внутренней памяти. Кроме того, память хранит как данные, так и команды. Поэтому для ЭВМ разработано множество способов обращения к памяти, называемых режимами адресации.

Режим адресации памяти – это процедура или схема преобразования адресной информации об операнде в его исполнительный адрес.

Все способы адресации памяти можно разделить на:

1) прямой, когда исполнительный адрес берется непосредственно из команды или вычисляется с использованием значения, указанного в команде, и содержимого какого-либо регистра (прямая адресация, регистровая, базовая, индексная и т.д.);   
2) косвенный, который предполагает, что в команде содержится значение косвенного адреса, т.е. адреса ячейки памяти, в которой находится окончательный исполнительный адрес (косвенная адресация).

5.1.3. Адресация операндов

Большая часть команд процессора работает с кодами данных (операндами). Одни команды требуют входных операндов (одного или двух), другие выдают выходные операнды (чаще один операнд). Входные операнды называются еще операндами-источниками, а выходные называются операндами-приемниками. Все эти коды операндов (входные и выходные) должны где-то располагаться. Они могут находиться во внутренних регистрах процессора (наиболее удобный и быстрый вариант). Они могут располагаться в системной памяти (самый распространенный вариант). Наконец, они могут находиться в устройствах ввода/вывода (наиболее редкий случай). Определение места положения операндов производится кодом команды. Причем существуют разные методы, с помощью которых код команды может определить, откуда брать входной операнд и куда помещать выходной операнд. Эти методы называются **методами адресации**. Эффективность выбранных *методов адресации* во многом определяет эффективность работы всего процессора в целом.

Методы адресации

Количество *методов адресации* в различных процессорах может быть от 4 до 16. Рассмотрим несколько типичных *методов адресации операндов*, используемых сейчас в большинстве микропроцессоров.

***Непосредственная адресация*** (рис. 5.2.) предполагает, что *операнд* (входной) находится в памяти непосредственно за кодом команды. *Операнд* обычно представляет собой константу, которую надо куда-то переслать, к чему-то прибавить и т.д. Например, команда может состоять в том, чтобы прибавить число 6 к содержимому какого-то внутреннего регистра процессора. Это число 6 будет располагаться в памяти, внутри программы в адресе, следующем за кодом данной команды сложения.

Программа

Команда

Операнд

Память

Рис.5.2. Непосредственная адресация

***Прямая (она же абсолютная) адресация*** (рис. 5.3.) предполагает, что *операнд* (входной или выходной) находится в памяти по адресу, код которого находится внутри программы сразу же за кодом команды. Например, команда может состоять в том, чтобы очистить (сделать нулевым) содержимое ячейки памяти с адресом 1000000. Код этого адреса 1000000 будет располагаться в памяти, внутри программы в следующем адресе за кодом данной команды очистки.

Память

Команда

Операнд

Программа

Операнд

Данные

Рис. 5.3.  Прямая адресация.

***Регистровая адресация*** (рис. 5.4) предполагает, что *операнд* (входной или выходной) находится во внутреннем регистре процессора. Например, команда может состоять в том, чтобы переслать число из нулевого регистра в первый. Номера обоих регистров (0 и 1) будут определяться кодом *команды пересылки*.

Косвенно-регистровая (она же косвенная) адресация предполагает, что во внутреннем регистре процессора находится не сам *операнд*, а его адрес в памяти (рис. 5.5). Например, команда может состоять в том, чтобы очистить ячейку памяти с адресом, находящимся в нулевом регистре. Номер этого регистра (0) будет определяться кодом команды очистки.

Память

Регистр

Команда

Программа

Операнд

Рис. 5.4.  Регистровая адресация.

Память

Регистр

Команда

Операнд

Программа

Операнд

Данные

Рис. 5.5.  Косвенная адресация.

Реже встречаются еще два *метода адресации*.

***Автоинкрементная адресация*** очень близка к косвенной адресации, но отличается от нее тем, что после выполнения команды содержимое используемого регистра увеличивается на единицу или на два. Этот *метод адресации* очень удобен, например, при последовательной обработке кодов из массива данных, находящегося в памяти. После обработки какого-то кода адрес в регистре будет указывать уже на следующий код из массива. При использовании косвенной адресации в данном случае пришлось бы увеличивать содержимое этого регистра отдельной командой.

***Автодекрементная адресация*** работает похоже на автоинкрементную, но только содержимое выбранного регистра уменьшается на единицу или на два перед выполнением команды. Эта адресация также удобна при обработке массивов данных. Совместное использование автоинкрементной и автодекрементной адресаций позволяет организовать память стекового типа.

Из других распространенных *методов адресации* можно упомянуть об индексных методах, которые предполагают для вычисления адреса *операнда* прибавление к содержимому регистра заданной константы (индекса). Код этой константы располагается в памяти непосредственно за кодом команды.

Отметим, что выбор того или иного *метода адресации* в значительной степени определяет время выполнения команды. Самая быстрая адресация – это регистровая, так как она не требует дополнительных циклов обмена по магистрали. Если же адресация требует обращения к памяти, то время выполнения команды будет увеличиваться за счет длительности необходимых циклов обращения к памяти. Понятно, что чем больше внутренних регистров у процессора, тем чаще и свободнее можно применять регистровую адресацию, и тем быстрее будет работать система в целом.

Сегментирование памяти.

Говоря об адресации, нельзя обойти вопрос о *сегментировании* памяти, применяемой в некоторых процессорах, например в процессорах IBM PC-совместимых персональных компьютеров.

В процессоре Intel 8086 *сегментирование* памяти организовано следующим образом.

Вся память системы представляется не в виде непрерывного пространства, а в виде нескольких кусков – сегментов заданного размера (по 64 Кбайта), положение которых в пространстве памяти можно изменять программным путем.

Для хранения кодов адресов памяти используются не отдельные регистры, а пары регистров:

* сегментный регистр определяет адрес начала сегмента (то есть положение сегмента в памяти);
* регистр указателя (регистр смещения) определяет положение рабочего адреса внутри сегмента.

При этом физический 20-разрядный адрес памяти, выставляемый на внешнюю *шину адреса*, образуется так, как показано на рис. 3.6, то есть путем сложения смещения и адреса сегмента со сдвигом на 4 бита. Положение этого адреса в памяти показано на рис. 3.7.

Сегмент может начинаться только на 16-байтной границе памяти (так как адрес начала сегмента, по сути, имеет четыре младших нулевых разряда, как видно из рис. 5.6), то есть с адреса, кратного 16.

Эти допустимые границы сегментов называются границами параграфов.

Отметим, что введение *сегментирования*, прежде всего, связано с тем, что внутренние регистры процессора 16-разрядные, а физический адрес памяти 20-разрядный (16-разрядный адрес позволяет использовать память только в 64 Кбайт, что явно недостаточно). В появившемся в то же время процессоре MC68000 фирмы Motorola внутренние регистры 32-разрядные, поэтому там проблемы *сегментирования* памяти не возникает.

16 бит

20 бит

4 бита

16 бит

Исполнительный адрес

(смещение)

Адрес сегмента

Физический адрес памяти

Рис. 5.6.  Формирование физического адреса памяти из адреса сегмента и смещения

15АВЕ

Смещение (ЗАВЕ)

12000

.

.

.

.

.

.

Память

Адрес сегмента

Размер сегмента

(10000 байт = 64 Кбайт)

Физический адрес

Адреса 00000

21FFF

22000

Рис. 5.7.  Физический адрес в сегменте (все коды – шестнадцатеричные)

Применяются и более сложные методы *сегментирования* памяти. Например, в процессоре *Intel 80286* в так называемом защищенном режиме адрес памяти вычисляется в соответствии с [рис. 3.](http://www.intuit.ru/department/hardware/mpbasics/4/#image.3.7)8.

В *сегментном регистре* в данном случае хранится не базовый (начальный) адрес сегментов, а коды селекторов, определяющие адреса в памяти, по которым хранятся дескрипторы (то есть описатели) сегментов. Область памяти с дескрипторами называется таблицей дескрипторов. Каждый дескриптор сегмента содержит базовый адрес сегмента, размер сегмента (от 1 до 64 Кбайт) и его атрибуты. Базовый адрес сегмента имеет разрядность 24 бит, что обеспечивает адресацию 16 Мбайт физической памяти.

Процессор

Сегментный регистр

15

0

Селектор

15

0

Смещение

Сумматор физического адреса

23

0

Базовый адрес сегмента

Память

Операнд

Операнд

Сегмент

Дескриптор

сегмента

Таблица

дескрипторов

сегмента

Рис. 5.8.  Адресация памяти в защищенном режиме процессора Intel 80286

Таким образом, на *сумматор*, вычисляющий физический адрес памяти, подается не содержимое сегментного регистра, как в предыдущем случае, а базовый адрес сегмента из таблицы дескрипторов.

Еще более сложный *метод адресации* памяти с *сегментированием* использован в процессоре *Intel 80386* и в более поздних моделях процессоров фирмы Intel. Этот метод иллюстрируется рис. 3.9.

Индекс

дескриптора

32

Эффективный

адрес

Вычисление эффективного адреса

INDEX

SCALE (1,2,4,8)

BASE

DISP

32

Линейный

адрес

32

Физический

адрес

Память

Блок страничной переадресации

Блок сегментации

Сегментный регистр

Рис. 5.9.  Формирование физического адреса памяти процессора 80386 в защищенном режиме.

Адрес памяти (физический адрес) вычисляется в три этапа. Сначала вычисляется так называемый **эффективный адрес** (32-разрядный) путем суммирования трех компонентов: базы, индекса и смещения (Base, Index, *Displacement*), причем возможно умножение индекса на масштаб (Scale). Эти компоненты имеют следующий смысл:

* смещение– это 8-, 16- или 32-разрядное число, включенное в команду;
* база – это содержимое базового регистра процессора. Обычно оно используется для указания на начало некоторого массива;
* индекс – это содержимое индексного регистра процессора. Обычно оно используется для выбора одного из элементов массива;
* масштаб – это множитель (он может быть равен 1, 2, 4 или 8), указанный в коде команды, на который перед суммированием с другими компонентами умножается индекс. Он используется для указания размера элемента массива.,

Затем специальный блок *сегментации* вычисляет 32-разрядный линейный адрес, который представляет собой сумму базового адреса сегмента из сегментного регистра с эффективным адресом. Наконец, физический 32-битный адрес памяти образуется путем преобразования линейного адреса блоком страничной переадресации, который осуществляет перевод линейного адреса в*физический страницами* по

4 Кбайта.

В любом случае *сегментирование* позволяет выделить в памяти один или несколько сегментов для данных и один или несколько сегментов для программ. Переход от одного сегмента к другому сводится всего лишь к изменению содержимого сегментного регистра. Иногда это бывает очень удобно. Но для программиста работать с сегментированной памятью обычно сложнее, чем с непрерывной, несегментированной памятью, так как приходится следить за границами сегментов, за их описанием, переключением и т.д.

Адресация байтов и слов.

Многие процессоры, имеющие разрядность 16 или 32, способны адресовать не только целое слово в памяти (16-разрядное или 32-разрядное), но и отдельные байты. Каждому байту в каждом слове при этом отводится свой адрес.

Так, в случае 16-разрядных процессоров все слова в памяти (16-разрядные) имеют четные адреса. А байты, входящие в эти слова, могут иметь как четные адреса, так и нечетные.

Например, пусть 16-разрядная ячейка памяти имеет адрес 23420, и в ней хранится код 2А5Е (рис. 3.10).

.

.

16 бит

Память

.

.

2A 5E

48 7F

Старший байт

(адрес 23421)

Старший байт

(адрес 23423)

Младший байт

(адрес 23420)

Младший байт

(адрес 23422)

Слово

(адрес 23422)

Слово

(адрес 23420)

Рис. 5.10. Адресация слов и байтов.

При обращении к целому слову (с содержимым 2А5Е) процессор выставляет адрес 23420. При обращении к младшему байту этой ячейки (с содержимым 5Е) процессор выставляет тот же самый адрес 23420, но использует команду, адресующую байт, а не слово. При обращении к старшему байту этой же ячейки (с содержимым 2А) процессор выставляет адрес 23421 и использует команду, адресующую байт. Следующая по порядку 16-разрядная ячейка памяти с содержимым 487F будет иметь адрес 23422, то есть опять же четный. Ее байты будут иметь адреса 23422 и 23423.

Для различия байтовых и словных циклов обмена на магистрали в шине управления предусматривается специальный сигнал байтового обмена. Для работы с байтами в систему команд процессора вводятся специальные команды или предусматриваются методы байтовой адресации.

5.1.4 Регистры процессора

Как уже упоминалось, внутренние регистры процессора представляют собой сверхоперативную память небольшого размера, которая предназначена для временного хранения служебной информации или данных. Количество регистров в разных процессорах может быть от 6—8 до нескольких десятков. Регистры могут быть универсальными и специализированными. Специализированные регистры, которые присутствуют в большинстве процессоров, – это *регистр*-*счетчик команд, регистр состояния* (**PSW**), *регистр указателястека*. Остальные регистры процессора могут быть как универсальными, так и специализированными.

Например, в 16-разрядном процессоре Т-11 фирмы DEC было 8 регистров общего назначения (РОН) и один *регистр состояния*. Все регистры имели по 16 разрядов. Из регистров общего назначения один отводился под *счетчик команд*, другой – под указатель стека. Все остальные регистры общего назначения полностью взаимозаменяемы, то есть имеют универсальное назначение, могут хранить как данные, так и адреса (указатели), индексы и т.д. Максимально допустимый объем памяти для данного процессора составлял 64 Кбайт (адрес памяти 16-разрядный).

В 16-разрядном процессоре MC68000 фирмы Motorola было 19 регистров: 16-разрядный *регистр состояния*, 32-разрядный регистр *счетчика команд*, 9 регистров адреса (32-разрядных) и 8 регистров данных (32-разрядных). Два регистра адреса отведены под указатели стека. Максимально допустимый объем адресуемой памяти – 16 Мбайт (внешняя шина адреса 24-разрядная). Все 8 регистров данных взаимозаменяемы, 7 регистров адреса – тоже взаимозаменяемы.

В 16-разрядном процессоре Intel 8086, который стал базовым в линии процессоров, используемых в персональных компьютерах, реализован принципиально другой подход. Каждый регистр этого процессора имеет свое особое назначение, и заменять друг друга регистры могут только частично или же не могут вообще. Остановимся на особенностях этого процессора подробнее.

Процессор 8086 имеет 14 регистров разрядностью по 16 бит. Из них четыре регистра (AX, BX, CX, DX) – это регистры данных, каждый из которых помимо хранения операндов и результатов операций имеет еще и свое специфическое назначение:

• регистр AX – умножение, деление, обмен с устройствами ввода/вывода (команды ввода и вывода);

• регистр BX – базовый регистр в вычислениях адреса;

• регистр CX – счетчик циклов;

• регистр DX –определение адреса ввода/вывода.

Для регистров данных существует возможность раздельного использования обоих байтов (например, для регистра AX они имеют обозначения AL – младший байт и AH — старший байт).

Следующие четыре внутренних регистра процессора — это сегментные регистры, каждый из которых определяет положение одного из рабочих сегментов (рис. 3.11):

• CS (CodeSegment) соответствует сегменту команд, исполняемых в данный момент;

• DS (DataSegment) соответствует сегменту данных, с которыми работает процессор;

• ES (ExtraSegment) соответствует дополнительному сегменту данных;

• SS (StackSegment) соответствует сегменту стека.

**.**

**.**

**.**

(CS)

Сегмент команд

**.**

**.**

**.**

(DS)

Сегмент данных

**.**

**.**

**.**

(ES)

Дополнительный сегмент данных

**.**

**.**

**.**

(SS)

Стек

**.**

**.**

**.**

Рис. 5.11. Сегменты команд, данных и стека в памяти.

В принципе, все эти сегменты могут и перекрываться для оптимального использования пространства памяти. Например, если программа занимает только часть сегмента, то сегмент данных может начинаться сразу после завершения работы программы (с точностью 16 байт), а не после окончания всего сегмента программы.

Следующие пять регистров процессора (SP –StackPointer, BP –BasePointer, SI –SourceIndex, DI –DestinationIndex, IP –InstructionPointer) служат указателями (то есть определяют смещение в пределах сегмента). Например, *счетчик команд* процессора образуется парой регистров CS и IP, а указатель стека – парой регистров SP и SS. Регистры SI, DI используются в строковых операциях, то есть при последовательной обработке нескольких ячеек памяти одной командой.

Последний регистр FLAGS – это регистр состояния процессора (PSW). Из его 16 разрядов используются только девять (рис. 3.12): CF (CarryFlag) – флаг пере-носа при арифметических операциях, PF (ParityFlag) – флаг четности результа-та, AF (AuxiliaryFlag) – флаг дополнительного переноса, ZF (ZeroFlag) — флаг нулевого результата, SF (SignFlag) – флаг знака (совпадает со старшим битом результата), TF (TrapFlag) – флаг пошагового режима (используется при отла-ке), IF (Interrupt-enableFlag) – флаг разрешения аппаратных прерываний, DF (DirectionFlag) – флаг направления при строковых операциях, OF (OverflowFlag) – флаг переполнения.

CF

1514 13 12 11 109 8 7 6 5 4 3 2 1 0

OF DF IF TF SF ZF AF PF CFCFАCF

Рис. 5.12. Регистр состояния процессора 8086.

Биты регистра состояния устанавливаются или очищаются в зависимости от результата исполнения предыдущей команды и используются некоторыми командами процессора. Биты регистра состояния могут также устанавливаться и очищаться специальными командами процессора (о системе команд процессора будет рассказано в следующем разделе).

Во многих процессорах выделяется специальный регистр, называемый **аккумулятором** (то есть накопителем). При этом, как правило, только этот регистр-аккумулятор может участвовать во всех операциях, только через него может производиться взаимодействие с устройствами ввода/вывода. Иногда в него же помещается результат любой выполненной команды (в этом случае говорят даже об «аккумуляторной» архитектуре процессора). Например, в процессоре 8086 регистр данных АХ можно считать своеобразным аккумулятором, так как именно он обязательно участвует в командах умножения и деления, а также только через него можно пересылать данные в устройство ввода/вывода и из устройства ввода/вывода. Выделение специального регистра-аккумулятора упрощает структуру процессора и ускоряет пересылки кодов внутри процессора, но в некоторых случаях замедляет работу системы в целом, так как весь поток информации должен пройти через один регистр-аккумулятор. В случае, когда несколько регистров процессора полностью взаимозаменяемы, таких проблем не возникает.

К данному типу архитектуры относится микропроцессор фирмы Zilog. Процессор Z80 – детище фирмы Zilog помимо расширенной системы команд, одного номинала питания и способности исполнять программы, написанные для i8080, имел архитектурные «изюминки».

Шина адреса (16 бит)

Шина данных (8бит)

Управление (13 линий)

Регистр команды

8 бит

Индексный регистр 1Х

Индексный регистр 1Y

Указатель стека SP

Программный

счетчик PC

16 бит

Основной набор регистров

Аккумулятор A

B

D

H

S

Z

H

PV

N

C

C

E

L

Аккумулятор A

B`

D`

H`

S

Z

H

PV

N

C

C`

E`

L`

АЛУ

8 бит

8 бит

8 бит

8 бит

PSW

Дополнительный набор регистров

Управление шиной

Вектор прерывания 1

Регенерация памяти - R

8 бит

8 бит

Рис. 5.13. Микропроцессор Z80 фирмы Zilog

В дополнение к основному набору РОН, в кристалле был реализован второй комплект аналогичных регистров. Это значительно упрощало работу при вызове подпрограмм или процедур обслуживания прерываний, поскольку программист мог использовать для них альтернативный набор регистров, избегая сохранения в стеке содержимого РОНов для основной программы с помощью операций PUSH. Кроме того, в систему команд был включен ряд специальных инструкций, ориентированных на обработку отдельных битов, а для поддержки регенерации динамической памяти в схему процессора введены соответствующие аппаратные средства. Z80 применялся в машинахSinclairZX, SinclairSpectrum, TandyTRS80.

Предельный вариант – архитектура с адресацией посредством аккумуляторов (меньший набор команд).

МП фирмы Motorola имел ряд существенных преимуществ. Прежде всего, кристалл МС6800 требовал для работы одного номинала питания, а система команд оказалась весьма прозрачной для программиста. Архитектура МП также имела ряд особенностей.

Адрес(16 бит)

Данные (8бит)

Управление (9 линий)

8 бит

АЛУ

Управление шиной

16 бит

16 бит

16 бит

Программный счетчик

Указатель стека

Индексный регистр

H

I

N

Z

V

C

Регистр команды

Состояние процессора

А (8 бит)

А (8 бит)

Аккумуляторы

Рис 5.14. Микропроцессор МС6800 фирмы Motorola.

Микропроцессор МС 6800 содержал два аккумулятора, и результат операции АЛУ мог быть помещен в любой из них. Но самым ценным качеством структуры МС 6800 было автоматическое сохранение в стеке содержимого всех регистров процессора при обработке прерываний (Z80 требовалось для этого несколько команд PUSH). Процедура восстановления РОН из стека тоже выполнялась аппаратно.

1. Стековая архитектура дает возможность создать поле памяти с упорядоченной последовательностью записи и выборки информации.
2. В общем случае команды неявно адресуются к элементу стека, расположенному на его вершине, или к двум верхним элементам стека.
3. Архитектура МП, ориентированная на оперативную память (типа «память-память»), обеспечивает высокую скорость работы и большую информационную емкость рабочих регистров и стека при их организации в оперативной памяти.

Архитектура этого типа не предполагает явного определения аккумулятора, регистров общего назначения или стека; все операнды команд адресуются к области основной памяти.

С точки зрения важности для пользователя-программиста под архитектурой в общем случае понимают совокупность следующих компонентов и характеристик:

* разрядности адресов и данных;
* состава, имен и назначения программно-доступных регистров;
* форматов и системы команд;
* режимов адресации памяти;
* способов машинного представления данных разного типа;
* структуры адресного пространства;
* способа адресации внешних устройств и средств выполнения операций ввода/вывода;
* классов прерываний, особенностей инициирования и обработки прерываний.

5.1.5. Система команд процессора

В общем случае система команд процессора включает в себя следующие четыре основные группы команд:

• команды пересылки данных;

• арифметические;

• логические;

• команды переходов.

**Команды пересылки** данных не требуют выполнения никаких операций над операндами. Операнды просто пересылаются (точнее, копируются) из источника (Source) в приемник (Destination). Источником и приемником могут быть внутренние регистры процессора, ячейки памяти или устройства ввода/вывода. АЛУ в данном случае не используется.

**Арифметические команды** выполняют операции сложения, вычитания, умножения, деления, увеличения на единицу (инкрементирования), уменьшения на единицу (декрементирования) и т.д. Этим командам требуется один или два входных операнда. Формируют команды один выходной операнд.

**Логические команды** производят над операндами логические операции, например, логическое И, *логическое* ИЛИ, исключающее ИЛИ, очистку, инверсию, разнообразные сдвиги (вправо, влево, арифметический сдвиг, циклический сдвиг). Этим командам, как и *арифметическим*, требуется один или два входных операнда, и формируют они один выходной операнд.

Наконец, **команды переходов** предназначены для изменения обычного порядка последовательного выполнения команд. С их помощью организуются переходы на подпрограммы и возвраты из них, всевозможные циклы, ветвления программ, пропуски фрагментов программ и т.д. Команды переходов всегда меняют содержимое счетчика команд. Переходы могут быть условными и безусловными. Именно эти команды позволяют строить сложные алгоритмы обработки информации.

В соответствии с результатом каждой выполненной команды устанавливаются или очищаются биты регистра состояния процессора (PSW). Но надо помнить, что не все команды изменяют все имеющиеся в PSW флаги. Это определяется особенностями каждого конкретного процессора.

У разных процессоров системы команд существенно различаются, но в основе своей они очень похожи. Количество команд у процессоров также различно. Например, у упоминавшегося уже процессора МС68000 всего 61 команда, а у процессора 8086 – 133 команды. У современных мощных процессоров количество команд достигает нескольких сотен. В то же время существуют процессоры с *сокращенным набором команд* (так называемые RISC-процессоры), в которых за счет максимального сокращения количества команд достигается увеличение эффективности и скорости их выполнения.

Рассмотрим теперь особенности четырех выделенных групп команд процессора более подробно.

Команды пересылки данных.

*Команды пересылки* данных занимают очень важное место в системе команд любого процессора. Они выполняют следующие важнейшие функции:

• загрузка (запись) содержимого во внутренние регистры процессора;

• сохранение в памяти содержимого внутренних регистров процессора;

• копирование содержимого из одной области памяти в другую;

• запись в устройства ввода/вывода и чтение из устройств ввода/вывода.

В некоторых процессорах (например, Т-11) все эти функции выполняются одной единственной командой MOV (для байтовых пересылок – MOVB), но с различными *методами адресации* операндов.

В других процессорах помимо команды MOV имеется еще несколько команд для выполнения перечисленных функций. Например, для загрузки регистров могут использоваться команды загрузки, причем для разных регистров – разные команды (их обозначения обычно строятся с использованием слова LOAD – загрузка). Часто выделяются специальные команды для сохранения в стеке и для извлечения из стека (PUSH – сохранить в стеке, POP – извлечь из стека). Эти команды выполняют пересылку с автоинкрементной и с автодекрементной адресацией (даже если эти режимы адресации не предусмотрены в процессоре в явном виде).

Иногда в систему команд вводится специальная команда MOVS для строчной (или цепочечной) пересылки данных (например, в процессоре 8086). Эта команда пересылает не одно слово или байт, а заданное количество слов или байтов (MOVSB), то есть инициирует не один цикл обмена по магистрали, а несколько. При этом адрес памяти, с которым происходит взаимодействие, увеличивается на 1 или на 2 после каждого обращения или же уменьшается на 1 или на 2 после каждого обращения. То есть в неявном виде применяется автоинкрементная или автодекрементная адресация.

В некоторых процессорах (например, в процессоре 8086) специально выделяются функции обмена с устройствами ввода/вывода. Команда IN используется для ввода (чтения) информации из устройства ввода/вывода, а команда OUT используется для вывода (записи) в устройство ввода/вывода. Обмен информацией в этом случае производится между регистром-аккумулятором и устройством ввода/вывода. В более продвинутых процессорах этого же семейства (начиная с процессора 80286) добавлены команды строчного (цепочечного) ввода (команда INS) и строчного вывода (команда OUTS). Эти команды позволяют пересылать целый массив (строку) данных из памяти в устройство ввода/вывода (OUTS) или из устройства ввода/вывода в память (INS). Адрес памяти после каждого обращения увеличивается или уменьшается (как и в случае с командой MOVS).

Также к *командам пересылки* данных относятся команды обмена информацией (их обозначение строится на основе слова Exchange). Может быть предусмотрен обмен информацией между внутренними регистрами, между двумя половинами одного регистра (SWAP) или между регистром и ячейкой памяти.

Арифметические команды

*Арифметические команды* рассматривают коды операндов как числовые двоичные или двоично-десятичные коды. Эти команды могут быть разделены на пять основных групп:

• команды операций с *фиксированной запятой* (сложение, вычитание, умножение, деление);

• операций с *плавающей запятой* (сложение, вычитание, умножение, деление);

• очистки;

• инкремента и декремента;

• сравнения.

Команды операций с *фиксированной запятой* работают с кодами в регистрах процессора или в памяти как с обычными двоичными кодами. Команда сложения (ADD) вычисляет сумму двух кодов. Команда вычитания (SUB) вычисляет разность двух кодов. Команда умножения (MUL) вычисляет произведение двух кодов (разрядность результата вдвое больше разрядности сомножителей). Команда деления (DIV) вычисляет частное от деления одного кода на другой. Причем все эти команды могут работать как с числами со знаком, так и с числами без знака.

Команды операций с *плавающей запятой* (точкой) используют формат представления чисел с порядком и *мантиссой* (обычно эти числа занимают две последовательные ячейки памяти). В современных мощных процессорах набор команд с *плавающей запятой* не ограничивается только четырьмя арифметическими действиями, а содержит и множество других более сложных команд, например, вычисление тригонометрических функций, логарифмических функций, а также сложных функций, необходимых при обработке звука и изображения.

Команды очистки (CLR) предназначены для записи нулевого кода в регистр или ячейку памяти. Эти команды могут быть заменены *командами пересылки* нулевого кода, но специальные команды очистки обычно выполняются быстрее, чем *команды пересылки*. Команды очистки иногда относят к группе *логических команд*, но суть их от этого не меняется.

Команды инкремента (увеличения на единицу, INC) и декремента (уменьшения на единицу, DEC) также бывают очень удобны. Их можно в принципе заменить командами суммирования с единицей или вычитания единицы, но инкремент и декремент выполняются быстрее, чем суммирование и вычитание. Эти команды требуют одного входного операнда, который одновременно является и выходным операндом.

Наконец, команда сравнения (обозначается CMP) предназначена для сравнения двух входных операндов. По сути, она вычисляет разность этих двух операндов, но выходного операнда не формирует, а всего лишь изменяет биты в регистре состояния процессора (PSW) по результату этого вычитания. Следующая за командой сравнения команда (обычно это *команда перехода*) будет анализировать биты в регистре состояния процессора и выполнять действия в зависимости от их значений (о *командах перехода* речь идет в разделе 3.3.4). В некоторых процессорах предусмотрены команды цепочечного сравнения двух последовательностей операндов, находящихся в памяти.

Логические команды.

*Логические команды* выполняют над операндами логические (побитовые) операции, то есть они рассматривают коды операндов не как единое число, а как набор отдельных битов. Этим они отличаются от *арифметических команд*. *Логические команды* выполняют следующие основные операции:

• логическое И, логическое ИЛИ, сложение по модулю 2 (исключающее ИЛИ);

• логические, арифметические и циклические сдвиги;

• проверка битов и операндов;

• установка и очистка битов (флагов) регистра состояния процессора (PSW).

Команды логических операций позволяют побитно вычислять основные логические функции от двух входных операндов. Кроме того, операция И (AND) используется для принудительной очистки заданных битов (в качестве одного из операндов при этом используется код маски, в котором разряды, требующие очистки, установлены в нуль). Операция ИЛИ (OR) применяется для принудительной установки заданных битов (в качестве одного из операндов при этом используется код маски, в котором разряды, требующие установки в единицу, равны единице). Операция «исключающее ИЛИ» (XOR) используется для инверсии заданных битов (в качестве одного из операндов при этом применяется код маски, в котором биты, подлежащие инверсии, установлены в единицу). Команды требуют двух входных операндов и формируют один выходной операнд.

Команды сдвигов позволяют побитно сдвигать код операнда вправо (в сторону младших разрядов) или влево (в сторону старших разрядов). Тип сдвига (логический, арифметический или циклический) определяет, каково будет новое значение старшего бита (при сдвиге вправо) или младшего бита (при сдвиге влево), а также определяет, будет ли где-то сохранено прежнее значение старшего бита (при сдвиге влево) или младшего бита (при сдвиге вправо). Например, при логическом сдвиге вправо в старшем разряде кода операнда устанавливается нуль, а младший разряд записывается в качестве флага переноса в регистр состояния процессора. А при арифметическом сдвиге вправо значение старшего разряда сохраняется прежним (нулем или единицей), младший разряд также записывается в качестве флага переноса.

Циклические сдвиги позволяют сдвигать биты кода операнда по кругу (по часовой стрелке при сдвиге вправо или против часовой стрелки при сдвиге влево). При этом в кольцо сдвига может входить или не входить флаг переноса. В бит флага переноса (если он используется) записывается значение старшего бита при циклическом сдвиге влево и младшего бита при циклическом сдвиге вправо. Соответственно, значение бита флага переноса будет переписываться в младший разряд при циклическом сдвиге влево и в старший разряд при циклическом сдвиге вправо.

Для примера на рис. 5.13 показаны действия, выполняемые командами сдвигов вправо.

Команды проверки битов и операндов предназначены для установки или очистки битов *регистра состояния* процессора в зависимости от значения выбранных битов или всего операнда в целом. Выходного операнда команды не формируют. Команда проверки операнда (TST) проверяет весь код операнда в целом на равенство нулю и на знак (на значение старшего бита), она требует только одного входного операнда. Команда проверки бита (BIT) проверяет только отдельные биты, для выбора которых в качестве второго операнда используется код маски. В коде маски проверяемым битам основного операнда должны соответствовать единичные разряды.

Циклический сдвиг вправо

через перенос

Циклический сдвиг вправо

Арифметический сдвиг вправо

Логический сдвиг вправо

Флаг переноса

Мл.

Ст.

Операнд

Флаг переноса

Мл.

Ст.

Операнд

Флаг переноса

Мл.

Ст.

Операнд

Флаг переноса

Мл.

Ст.

Операнд

Рис. 5.13. Команды сдвигов вправо.

Наконец, команды установки и очистки битов *регистра состояния* процессора (то есть флагов) позволяют установить или очистить любой флаг, что бывает очень удобно. Каждому флагу обычно соответствуют две команды, одна из которых устанавливает его в единицу, а другая сбрасывает в нуль. Например, флагу переноса C (от Carry) будут соответствовать команды CLC (очистка) и SEC или STC (установка).

Команды переходов.

*Команды переходов* предназначены для организации всевозможных циклов, *ветвлений*, вызовов *подпрограмм* и т.д., то есть они нарушают последовательный ход выполнения программы. Эти команды записывают в регистр-*счетчик команд* новое значение и тем самым вызывают переход процессора не к следующей по порядку команде, а к любой другой команде в памяти программ. Некоторые *команды переходов* предусматривают в дальнейшем возврат назад, в точку, из которой был сделан переход, другие не предусматривают этого. Если возврат предусмотрен, то текущие параметры процессора сохраняются в стеке. Если возврат не предусмотрен, то текущие параметры процессора не сохраняются.

*Команды переходов* без возврата делятся на две группы:

• команды *безусловных переходов*;

• команды *условных переходов.*

В обозначениях этих команд используются слова Branch (ветвление) и Jump (прыжок).

Команды безусловных переходов вызывают переход в новый адрес независимо ни от чего. Они могут вызывать переход на указанную величину смещения (вперед или назад) или же на указанный адрес памяти. Величина смещения или новое значение адреса указываются в качестве входного операнда.

Команды условных переходов вызывают переход не всегда, а только при выполнении заданных условий. В качестве таких условий обычно выступают значения флагов в регистре состояния процессора (PSW). То есть условием перехода является результат предыдущей операции, меняющей значения флагов. Всего таких условий перехода может быть от 4 до 16. Несколько примеров команд условных переходов:

• переход, если равно нулю;

• если не равно нулю;

• если есть переполнение;

• если нет переполнения;

• если больше нуля;

• если меньше или равно нулю.

Если условие перехода выполняется, то производится загрузка в регистр-счетчик команд нового значения. Если же условие перехода не выполняется, счетчик команд просто наращивается, и процессор выбирает и выполняет следующую по порядку команду.

Специально для проверки условий перехода применяется команда сравнения (CMP), предшествующая команде условного перехода (или даже нескольким командам условных переходов). Но флаги могут устанавливаться и любой другой командой, например командой пересылки данных, любой арифметической или логической командой. Отметим, что сами команды переходов флаги не меняют, что как раз и позволяет ставить несколько команд переходов одну за другой.

Совместное использование нескольких команд условных и безусловных переходов позволяет процессору выполнять разветвленные алгоритмы любой сложности. Для примера на рис. 3.14 показано разветвление программы на две ветки с последующим соединением, а на рис. 3.15– разветвление на три ветки с последующим соединением.

Команды переходов с дальнейшим возвратом в точку, из которой был произведен переход, применяются для выполнения подпрограмм, то есть вспомогательных программ. Эти команды называются также командами вызова подпрограмм (распространенное название – CALL). Использование подпрограмм позволяет упростить структуру основной программы, сделать ее более логичной, гибкой, легкой для написания и отладки. В то же время надо учитывать, что широкое использование подпрограмм, как правило, увеличивает время выполнения программы.

нет

да

Программа 1

А + В А

Программа 3

Программа 2

А>0

Память

Сложение

Переход если больше 0

Начало программы 1

.

.

Безусловный переход

Начало программы 2

.

.

Конец программы 2

Начало программы 3

.

.

Рис. 5.14. Реализация разветвления на две ветки.

Память

Сложение

Переход, если равно 0

Переход если больше 0

Начало программы 3

.

.

Безусловный переход

Начало программы 1

.

.

Безусловный переход

Начало программы 2

.

.

Конец программы 2

Начало программы 4

нет

да

нет

Программа 1

А + В А

Программа 3

Программа 2

А=0

А>0

Программа 4

Рис. 5.15. Реализация разветвления на три ветки.

Все команды переходов с возвратом предполагают безусловный переход (они не проверяют никаких флагов). При этом они требуют одного входного операнда, который может указывать как абсолютное значение нового адреса, так и смещение, складываемое с текущим значением адреса. Текущее значение счетчика команд (текущий адрес) сохраняется перед выполнением перехода в стеке.

Для обратного возврата в точку вызова подпрограммы (точку перехода) используется специальная команда возврата (RET или RTS). Эта команда извлекает из стека значение адреса команды перехода и записывает его в регистр-счетчик команд.

Особое место среди команд перехода с возвратом занимают команды прерываний (распространенное название – INT). Эти команды в качестве входного операнда требуют номер прерывания (адрес вектора). Обслуживание таких переходов осуществляется точно так же, как и аппаратных прерываний. То есть для выполнения данного перехода процессор обращается к таблице векторов прерываний и получает из нее по номеру прерывания адрес памяти, в который ему необходимо перейти. Адрес вызова прерывания и содержимое регистра состояния процессора (PSW) сохраняются в стеке. Сохранение PSW – важное отличие команд прерывания от команд переходов с возвратом.

Команды прерываний во многих случаях оказываются удобнее, чем обычные команды переходов с возвратом. Сформировать таблицу векторов прерываний можно один раз, а потом уже обращаться к ней по мере необходимости. Номер прерывания соответствует номеру подпрограммы, то есть номеру функции, выполняемой подпрограммой. Поэтому команды прерывания гораздо чаще включаются в системы команд процессоров, чем обычные команды переходов с возвратом.

Для возврата из подпрограммы, вызванной командой прерывания, используется команда возврата из прерывания (IRET или RTI). Эта команда извлекает из стека сохраненное там значение счетчика команд и регистра состояния процессора (PSW).

Отметим, что у некоторых процессоров предусмотрены также команды условных прерываний, например, команда прерывания при переполнении.

Конечно, в данном разделе мы рассмотрели только основные команды, наиболее часто встречающиеся в процессорах. У конкретных процессоров могут быть и многие другие команды, не относящиеся к перечисленным группам команд. Но изучать их надо уже после того, как выбран тип процессора, подходящий для задачи, решаемой данной микропроцессорной системой.

5.2. Система команд микропроцессора Intel 8086

5.2.1. Способы адресации и форматы команд микропроцессора I 8086

Микропроцессор Intel8086(К1810ВМ80) имеет двухадресную систему команд. Ее особенностью является отсутствие команд, использующих оба операнда из оперативной памяти. Исключение составляют лишь команды пересылки и сравнения цепочек байт или слов, которые в данном пособии рассматриваться не будут. Таким образом, в командах допустимы следующие сочетания операндов: RR, RS, RI, SI. Здесь R обозначает операнд, находящийся в одном из регистров регистровой памяти микропроцессора, S– операнд, находящийся в оперативной памяти, адрес которого формируется по одному из допустимых способов адресации, I– непосредственный операнд, закодированный в адресном поле самой команды. Формат команды во многом определяется способом адресации операнда, находящего в оперативной памяти, длиной используемого непосредственного операнда, а также наличием и длиной смещения, используемого при относительных режимах адресации.

Микропроцессор имеет все режимы адресации, общая схема которых была рассмотрена выше. Естественно, они имеют определенные особенности, присущие данному процессору.

Непосредственная адресация предполагает, что операнд занимает одно из полей команды и, следовательно, выбирается из оперативной памяти одновременно с ней. В зависимости от форматов обрабатываемых процессором данных непосредственный операнд может иметь длину 8 или 16 бит, что в дальнейшем будем обозначать data8 и data16 соответственно.

Механизмы адресации операндов, находящихся в регистровой памяти и в оперативной памяти, существенно различаются. К регистровой памяти допускается лишь прямая регистровая адресация. При этом в команде указывается номер регистра, содержащего операнд. 16-разрядный операнд может находиться в регистрах AX, BX, CX, DX, DI, SI, SP, BP, а 8-разрядный – в регистрах AL, AH, BL, BH, CL, CH, DL, DH.

Адресация оперативной памяти имеет свои особенности, связанные с ее разбиением на сегменты и использованием сегментной группы регистров для указания начального адреса сегмента. 16-разрядный адрес, получаемый в блоке формирования адреса операнда на основе указанного режима адресации, называется эффективным адресом (ЭА). Иногда эффективный адрес обозначается как ЕА (effectiveaddress). 20-разрядный адрес, который получается сложением эффективного адреса и увеличенного в 16 раз значения соответствующего сегментного регистра, называется физическим адресом (ФА).

Именно физический адрес передается из микропроцессора по 20-ти адресным линиям, входящим в состав системной шины, в оперативную память и используется при обращении к ее ячейке на физическом уровне. При получении эффективного адреса могут использоваться все основные режимы адресации, рассмотренные выше, а также некоторые их комбинации.

Прямая адресация предполагает, что эффективный адрес является частью команды. Так как ЭА состоит из 16 разрядов, то и соответствующее поле команды должно иметь такую же длину.

При регистровой косвенной адресации эффективный адрес операнда находится в базовом регистре BX или одном из индексных регистров DI либо SI:

ЭА=

[BX]

[DI]

[SI]

**.**

Обозначение имени регистра в квадратных скобках указывает на содержимое соответствующего регистра. Фигурные скобки – символ выбора одной из нескольких возможных альтернатив.

При регистровой относительной адресации эффективный адрес равен сумме содержимого базового или индексного регистра и смещения:

ЭА=

[BX]

[BP]

[DI]

[SI]

+

disp8

disp16

**.**

Обозначения disp8 и disp16 здесь и далее указывают на 8- или 16-разрядное смещение соответственно.

Эффективный адрес при базово-индексной адресации равен сумме содержимого базового и индексного регистров, определяемых командой:

ЭА=

[BX]

[BP]

+

[DI]

[*SI]*

**.**

Наиболее сложен механизм относительной базово-индексной адресации. Эффективный адрес в этом случае равен сумме 8- или 16-разрядного смещения и базово-индексного адреса:

**.**

+

+

ЭА=

*[*BX]

[BP*]*

[DI]

[SI]

disp8

disp16

Форматы двухоперандных команд представлены на рис. 3.16. Пунктиром показаны поля, которые в зависимости от режима адресации могут отсутствовать в команде.

а) формат командтипа RR і RS

1 байт

1 байт

1 байт

1 байт

КАП *d w*

*mdreg r/m*

*disp L*

*disp H*

1 байт

1 байт

1 байт

1 байт

1 байт

1 байт

КАП *s w*

*md* КАП*r/m*

*disp L*

*disp H*

*data L*

*data H*

б) формат командснепосредственнымоперандом

Рис. 5.16. Форматы двухоперандных команд микропроцессора I808:

Поле КОП содержит код выполняемой операции. Признак w указывает на длину операндов. При *w* = 1 операция проводится над словами, а при *w* = 0 – над байтами. Признак *d* указывает положение приемника результата. Признак *d* = 1, если результат записывается на место операнда, закодированного в поле reg, и *d* = 0, если результат записывается по адресу, закодированному полями (md, r/m).

Второй байт команды, называемый постбайтом, определяет операнды, участвующие в операции. Поле reg указывает регистр регистровой памяти согласно табл. 5.1

|  |  |  |
| --- | --- | --- |
| Таблица 5.1 | | |
| **reg** | **Регистр** | |
| **w=1** | **w=0** |
| 000 | AX | AL |
| 001 | CX | CL |
| 010 | DX | DL |
| 011 | BX | BL |
| 100 | SP | AH |
| 101 | BP | CH |
| 110 | SI | DH |
| 111 | DI | BH |

Поля md и r/m задают режим адресации второго операнда согласно табл. 5.2.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Таблица 5.2 | | | | | | | |
| **r/m** | ***md*** | | | | | | |
| 00 | | 01 | 10 | | 11 | |
| ***w* =1** | ***w*=0** |
| 000 | (BX)+(SI)  (DS) | | (BX)+(SI)+disp 8  (DS) | (BX)+(SI)+disp 16  (DS) | | AX | AL |
| 001 | (BX)+(DI)  (DS) | | (BX)+(DI)+disp 8  (DS) | (BX)+(DI)+disp 16  (DS) | | CX | CL |
| 010 | (BP)+(SI)  (SS) | | (BP)+(SI)+disp 8  (SS) | (BP)+(SI)+disp 16  (SS) | | DX | DL |
| 011 | (BP)+(DI)  (SS) | | (BP)+(DI)+disp 8  (SS) | (BP)+(DI)+disp 16  (SS) | | BX | BL |
| 100 | (SI)  (DS) | | (SI)+disp 8  (DS) | (SI)+disp 16  (DS) | | SP | AH |
| 101 | (DI)  (DS) | | (DI)+disp 8  (DS) | (DI)+disp 16  (DS) | | BP | CH |
| 110 | disp16  (DS) | | (BP)+disp 8  (SS) | (BP)+disp 16  (SS) | | SI | DH |
| 111 | | (BX)  (DS) | (BX)+disp 8  (DS) | (BX)+disp 16  (DS) | DI | | BH |

В этой таблице помимо определения режима адресации оперативной памяти указан также сегментный регистр, используемый по умолчанию для получения физического адреса. Использование другого сегментного регистра возможно введением специального префикса (дополнительного байта, который записывается перед командой).

В командах, использующих непосредственный операнд, признак *s* вместе с признаком *w* определяет разрядность непосредственного операнда, записываемого в команде, и разрядность выполняемой операции согласно [табл. 5.3](http://www.intuit.ru/department/hardware/archhard2/6/2.html#table.6.3).

|  |  |  |  |
| --- | --- | --- | --- |
| Таблица 5.3  **Использование признака *s*** | | | |
| ***w*** | ***s*** | **Операция** | **Непосредственный операнд** |
| 0 | 0 | 8-разрядная | 8-разрядный |
| 0 | 1 | не используется | |
| 1 | 0 | 16-разрядная | 16-разрядный |
| 1 | 1 | 8-разрядный, расширяемый знаком до 16-ти разрядов при выполнении операции |

Изменение естественного порядка выполнения команд программы осуществляется с помощью команд передачи управления. К ним относятся команды переходов, циклов, вызова подпрограммы и возврата из нее, а также некоторые другие. Мы рассмотрим лишь первые две группы команд.

Классификация команд переходов в персональной ЭВМ представлена на рис. 5.17.

Команды переходов

Условные

Прямые

Косвенные

Внутрисегментные

Безусловные

Прямые

Косвенные

Межсегментные

Рис. 5.17. Классификация команд переходов IBM PC

Физический адрес выполняемой команды определяется содержимым указателя команд IP и сегментного регистра команд CS. Команды, меняющие значение обоих этих регистров, называются командами межсегментных переходов, а меняющие только значение IP, – командами внутрисегментных переходов.

Команды безусловных переходов производят модификацию регистра IP или регистров IP и CS без предварительного анализа ка-ких-либо условий. Существует пять команд безусловных переходов. Все они имеют одинаковую мнемонику JMP и содержат один операнд. Конкретный формат команды определяется соответствующим префиксом и приведен в общей таблице машинного представления команд (табл. 5.4).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Таблица 5.4  **Машинные коды некоторых команд** | | | | |
| **Команда** | **Байты кода команды** | | | **Схема операции** |
| **байт 1** | **байт 2** | **байты 3…6** |
| ADD | 000000dw | *md*reg r/m | (disp8/16) | r(r/m) = r + r/m |
| 100000sw | *md* 000 r/m | (disp8/disp16)d8/16 | r/m = r/m + d8/16 |
|  | 0000010w | data L | (data H) | ac = ac + d8/16 |
|  | | | | |
| SUB | 001010dw | *md*reg r/m | (disp8/16) | r(r/m)=r(r/m) - (r/m)r |
| 100000sw | *md* 101 r/m | (disp8/16)d8/16 | r/m = r/m - d8/16 |
|  | 0010110w | data L | (data H) | ac = ac - d8/16 |
| AND | 001000dw | *md*reg r/m | (disp8/16) | r(r/m) = r & r/m |
|  | 100000sw | *md* 100 r/m | (disp8/16)d8/16 | r/m = r/m & d8/16 |
|  | 0010010w | data L | (data H) | ac = ac& d8/16 |
| OR | 000010dw | *md*reg r/m | (disp8/16) | r(r/m) = r V r/m |
|  | 100000sw | *md* 001 r/m | (disp8/16)d8/16 | r/m = r/m V d8/16 |
|  | 0000110w | data L | (data H) | ac = ac V d8/16 |
| XOR | 001100dw | *md*reg r/m | (disp8/16) | r(r/m) = r + r/m |
|  | 100000sw | *md* 110 r/m | (disp8/16)d8/16 | r/m = r/m + d8/16 |
|  | 0011010w | data L | (data H) | ac = ac + d8/16 |
| MOV | 100010dw | *md*reg r/m | (disp8/16) | r = r/m, r/m = r |
| 1100011w | *md* 000 r/m | (disp8/16)d8/16 | r/m = d8/16 |
|  | 1011wreg | data L | (data H) | reg = d8/16 |
|  | 1010000w | disp L | disp H | аc=m; прямой адрес |
|  | 1010001w | disp L | disp H | m=ac; прямой адрес |
| CMP | 0011101w | *md*reg r/m | (disp8/16) | r - r/m |
|  | 0011100w | *md*reg r/m | (disp8/16) | r/m - r |
|  | 100000sw | *md* 111 r/m | (disp8/16)d8/16 | r/m - d8/16 |
|  | 0011110w | data L | (data H) | ac - d8/16 |
| INC | 1111111w | *md* 000 r/m | (disp8/16) | r/m = r/m+1 |
|  | 01000reg |  |  | reg = reg+1 |
| DEC | 1111111w | *md* 001 r/m | (disp8/16) | r/m = r/m-1 |
|  | 01001reg |  |  | reg = reg-1 |
| TEST | 1000010w | *md*regr/m | (disp8/16) | r & r/m |
|  | 1111011w | *md* 000 r/m | (disp8/16)d8/16 | r/m & d8/16 |
|  | 1010100w | data L | (data H) | ac& d8/16 |
| XCHG | 10010reg |  |  | reg↔AX |
|  | 1000011w | *md*reg r/m | (disp8/16) | reg↔r/m |
| JMP short | 11101011 | disp L |  | IP=IP+dispL |
| nearptr | 11101001 | disp L | disp H | IP=IP+dispH,L |
| wordptr | 11111111 | *md* 100 r/m | (disp8/16) | IP=(*EA*) |
| farptr | 11101010 | IP-L | IP-H,CS-L,CS-H | IP=IPH,L, CS=*CSH*,L |
| dwordptr | 11111111 | *md* 101 r/m | (disp8/16) | IP=(*EA*), CS=(*EA*+2) |
| *Условный переход* | \_\_\_\_\_\_ | \_\_\_\_\_\_ | «нет данных» | IP=IP+dispL, если условие выполнено, иначе к следующей команде |
| JZ (JE) | 01110100 | disp L | «нет данных» | ноль (равно) |
| JNZ (JNE) | 01110101 | disp L | «нет данных» | не ноль (не равно) |
| JS | 01111000 | disp L | «нет данных» | минус |
| JNS | 01111001 | disp L | «нет данных» | плюс |
| JO | 01110000 | disp L | «нет данных» | переполнение |
| JNO | 01110001 | disp L | «нет данных» | нет переполнения |
| JL (JNGE) | 01111100 | disp L | «нет данных» | меньше для чисел |
| JNL (JGE) | 01111101 | disp L | «нет данных» | не меньше для чисел |
| JG (JNLE) | 01111100 | disp L | «нет данных» | больше для чисел |
| JNG (JLE) | 01111101 | disp L | «нет данных» | не больше для чисел |
| JB (JNAE,JC) | 01110010 | disp L | «нет данных» | меньше для кодов |
| JNB (*JAE*, JNC) | 01110011 | disp L | «нет данных» | не меньше для кодов |
| JA (JNBE) | 01110010 | disp L | «нет данных» | больше для кодов |
| JNA (JBE) | 01110011 | disp L | «нет данных» | не больше для кодов |
| JP (JPE) | 01111010 | disp L | «нет данных» | четное число "1" |
| JNP (JPO) | 01111011 | disp L | «нет данных» | нечетное число "1" |

*Примечание:* в столбце «Схема операции»ac означает регистр-аккумулятор, в качестве которого используется регистр AX при *w*=1 и регистр AL при *w*=0.

При безусловном прямом внутрисегментном переходе новое значение указателя команд IP равно сумме смещения, закодированного в соответствующем поле команды, и текущего значения IP, в качестве которого используется адрес команды, записанной вслед за командой перехода. Команды прямых межсегментных переходов содержат в сeбе помимо нового значения IP и новое значение сегментного регистра CS.

Команды косвенных переходов (внутрисегментных и межсегментных) передают управление на команду, адрес которой определяется содержимым регистра или ячеек оперативной памяти, на которые указывает закодированный в команде перехода постбайт.

Команды условных переходов являются только внутрисегментными. По своему формату и способу формирования нового значения IP они полностью аналогичны команде внутрисегментного прямого перехода с 8-разрядным смещением. Отличие их заключается в том, что в командах условного перехода механизм формирования нового значения IP включается лишь при выполнении определенных условий, а именно, при определенном состоянии регистра флагов. При невыполнении проверяемого условия в IP остается его текущее значение, то есть адрес команды, следующей за командой условного перехода.

Ниже приведены примеры команд переходов различных типов.

1) Команды условного перехода

Формат:

КОП

disp L

IP = IP + 2, если условие не выполнено;

IP = IP +2 + disp L, если условие выполнено;

Пример:

JZ MARK; переход на метку MARK, если ZF = 1.

2) Команды прямого внутрисегментного перехода

Формат:

КОП

disp L

disp

IP=IP+∆+disp, где ∆ - длина команды перехода (2 или 3 в зависимости от длины смещения).

Примеры:

JMP shortptr MARK; переход на метку MARK, с использованием 8-разрядного смещения;

JMP nearptr MARK; переход на метку MARK, с использованием 16-разрядного смещения.

3) Команды прямого межсегментного перехода

Формат:

КОП

IP\_L

IP\_H

CS\_H

CS\_H

IP = IP\_H, IP\_L,

CS = CS\_H, CS\_L.

Пример:

JMP farptr MARK; переход на метку MARK к команде, находящейся в другом сегменте.

4) Команды косвенного внутрисегментного перехода.

Формат:

КОП

md

disp

disp

IP = [*EA* + 1, *EA*]; или

IP = <регистр>, если в постбайте задано обращение к регистровой памяти;

Пример:

JMP wordptr [BX + SI]; новое значение IP берется из двух последовательных байт памяти, эффективный адрес первого из которых определяется суммой регистров BX и SI.

5) Команды косвенного межсегментного перехода

Формат:

КОП

md

dispL

disp

IP = [*EA* + 1, *EA*],

CS = [*EA* + 3, *EA* + 2].

Пример:

JMP dwordptr [BX + SI]; сумма регистров BX и SI определяет эффективный адрес области памяти, первые два байта которой содержат новое значение IP, а следующие два байта–новое значение CS.

Команды циклов идентичны по формату и очень близки по вы-полняемым действиям командам условных переходов. Однако по сравнению с последними они имеют ряд особенностей, позволяющих эффективно использовать их при программировании циклических участков алгоритмов.

Один из наиболее распространенных видов циклического участка программы представлен на рис. 5.18.

да

нет

СЧ=0

СЧ=<число повторений>

тело цикла

СЧ=СЧ-1

Рис. 5.18. Структура счетного цикла с постпроверкой.

Команды циклов предназначены для упрощения действий декремента (уменьшения на 1) счетчика цикла, проверки условия выхода из цикла и перехода.

Некоторые команды цикла реализуют выход из цикла не только по значению счетчика, но и при выполнении некоторых других условий.

Описание команд цикла сведено в табл. 5.5. За исключением команды JCXZ, которая не изменяет значения регистра CX, при выполнении команд циклов производятся следующие действия:

CX=(CX)–1. Затем, если проверяемое условие выполнено, то IP=(IP)+disp8 с расширением смещения знаком до 16 разрядов, в противном случае IP не изменяется, и программа продолжает выполнение в естественном порядке.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Таблица 5.5  **Команды циклов** | | | | |
| **Название** | **Мнемоника** | **Альтернативная мнемоника** | **КОП** | **Проверяемое условие** |
| Зациклить | LOOP |  | 11100010 | (CX)=0 |
| Зациклить пока ноль (равно) | LOOPZ | LOOPE | 11100001 | (ZF=1)&(CX)=0) |
| Зациклить пока не ноль (неравно) | LOOPNZ | LOOPNE | 11100000 | (ZF=0)&((CX)=0) |
| Перейти по (CX) | JCXZ |  | 11100011 | (CX)=0 |

**6. Архитектура параллельных компьютерных систем**

Развитие элементно-конструкторской базы, основанное на физических принципах электроники, всегда (и это диалектически правильно) отстает от требований конкретно решаемых задач. Быстродействия одного компьютера недостаточно для, например, массового обслуживания группы перехватчиков, стартующих против налета баллистических целей. Своевременно пришло понимание того, что только *структурными* методами можно добиться необходимой производительности вычислительных средств. Таким основным структурным методом является *распараллеливание вычислений*.

Сначала стало ясно, что вместо одной ЭВМ в систему управления необходимо встраивать несколько, решающих единый набор функционально связанных задач. Такой коллектив ЭВМ получил название вычислительного комплекса (ВК). Информационная взаимосвязанность задач потребовала применения средств оперативного обмена данными внутри ВК и средств синхронизации. Однако временные затраты на организацию взаимодействия ЭВМ в ВК (накладные расходы) оказались весьма высоки. Возникла идея использования взамен *разделенной* оперативной памяти (ОП), реализованной в ВК, *общей (разделяемой)* ОП, которая связывает несколько центральных процессоров. Проблема затрат времени на обмен оказалась решенной. Так зародилось понятие *параллельной вычислительной системы* (ВС).

Впоследствии принципы распараллеливания стали применяться на всех уровнях разработки ВС: и отдельных устройств, и процессоров. Отметим, что такой способ распараллеливания, как конвейерный, использовался на раннем этапе создания ЭВМ и используется сейчас.

Приоритет в создании ВС на общей памяти принадлежит С. Крею, признанному «отцу супер-ЭВМ», главному конструктору системы CDC 6600, разработанной в 1963 году. Однако условия соперничества и сокрытия информации определили независимость отечественных разработок.

Необходимо отметить тех выдающихся ученых, чьими усилиями осуществлялась не только разработка, но и обоснование эффективности и целесообразности пути создания многопроцессорных ВС, ибо своевременно принятая техническая политика определяет эффективное развитие государства. (К сожалению, известна и негативная политика в области вычислительной техники, отвлекшая значительные технологические ресурсы.) Такими учеными являются: В. М. Бахарев, настойчиво пропагандировавший, продвигавший и руководивший исследованиями в области ВС на общей памяти; В. С. Бурцев и Б. А. Бабаян — конструкторы и ведущие разработчики семейства «Эльбрус», продолжающие борьбу за честь Отечества и поныне; М. А. Карцев — теоретик ВС и разработчик семейства, по его словам, «универсальных спецпроцессоров», объединивших в одном вычислительном устройстве скалярные и векторные операции; Б. А. Головкин, исследовавший эффективность внедрения ВС в сложные управляющие системы.

Какова же цель распараллеливания?

Универсальным критерием, используемым при решении задач распараллеливания, является минимум времени выполнения совокупности работ, распределяемых между процессорами.

Если несколько процессоров составляют ВС, то важной характеристикой ее эффективности (основные составляющие эффективности — производительность, надежность, стоимость) при специализированном использовании (например, в составе автоматизированной системы управления, АСУ) является коэффициент полезной загрузки *k*3. Для его определения находят коэффициенты загрузки процессоров:

**,

где *Ti*, *i* = 1, ..., *n*, — суммарное время занятости каждого процессора решением задачи на всем отрезке полного решения задачи, длиной *T*реш([рис.В.1](http://www.intuit.ru/department/hardware/paralltech/0/#image.B.1)). Тогда



Если *P0* — производительность одного процессора, то реальная производительность *PBC* ВС, состоящей из n процессоров, при решении данной задачи (!) находится:



где *P*0 определяется классом решаемых задач.

Т3

Т2

Т1

**6.1 Два уровня распараллеливания**

Развитие вычислительной техники характеризуется тем, что на каждом этапе новых разработок требования к производительности значительно превышают возможности элементной базы.

Это обусловлено задачами сложных систем управления в реальном времени, централизованным решением задач в сетях, имитационным моделированием сложных процессов (например, в ядерной физике), оперативным планированием и управлением и решением других задач исследования операций, преодолевающих «проклятие размерности». Такие задачи требуют концентрации вычислительных мощностей, постоянно поддерживая высокую актуальность проблемы создания супер-ЭВМ.

Уже давно стало ясно, что только структурными методами можно уравнять возможности вычислительных средств и требуемые скорости решения на них задач. Под структурными понимают методы распараллеливания работ. К распараллеливанию прибегают при проектировании отдельных устройств ЭВМ — устройств управления, буферов команд, каналов обращения к памяти и модулей памяти, многофункциональных арифметическо-логических устройств (АЛУ), повсеместно применяемых конвейеров и т. д. Но к распараллеливанию же прибегают и в проектировании совместной работы многих процессоров при параллельной или распределенной обработке информации, вводя в обращение термин "вычислительная система (ВС)".

Технический прогресс, несомненно, сказывается на росте частоты работы элементной (элементно-конструкторской) базы, на повышении степени интеграции, но благодаря ему появляются все новые задачи, требующие еще более значительного роста производительности вычислительных средств. Это можно считать законом, приводящим к новым уловкам при совмещении работы устройств ВС, при увеличении их количества в системе, при увеличении их эффективности в процессе решения задач.

Под ***эффективностью работы устройства*** в составе ВС понимают степень его участия в общей работе ВС при решении конкретной задачи — коэффициент загрузки устройства. Распараллеливание работ оправдано, если приводит к существенному росту усредненного по всем устройствам коэффициента загрузки оборудования при решении задач. Это непосредственно сказывается на времени решения. Сегодня говорят не о специальном классе задач, а о задачах, ориентирующих ВС на универсальность, что обусловлено современными областями применения.

Важным революционизирующим моментом стал переход на микропроцессорную элементно-конструкторскую базу, обусловившую построение мультимикропроцессорных ВС.

Сложилось представление о двух основных уровнях, на которых в ВС применяются практические методы распараллеливания:

* на уровне программ, процессов, процедур (первый уровень распараллеливания);
* на уровне команд и операций (второй уровень распараллеливания).

Эти уровни обусловили уровни структуризации ВС на пути превращения ее в супер-ЭВМ. Современным практическим воплощением первого уровня структуризации являются однородные многопроцессорные ВС на общей (разделяемой) оперативной памяти. Они получили название симметричных ВС за обеспечение «равноправия» составляющих модулей. Окончательное признание симметричных ВС положило конец поиску «экзотических» архитектур, эффективных лишь при решении определенных классов задач. Универсальность симметричных ВС, возможность реализации на них любых вычислительных процессов с высокой эффективностью оборудования иллюстрируются многими применениями и анализируются ниже.

Уровень команд и операций наиболее ярко представлен многофункциональными АЛУ и их обобщением — решающими полями, представляющими разделяемый вычислительный ресурс многих процессоров. Некоторые современные проекты ВС в разной степени предполагают такой ресурс. Здесь основная проблема — полная загрузка отдельных исполнительных устройств (ИУ) в процессе выполнения программы.

Различают два способа реализации такой загрузки: динамический и статический.

Динамическая загрузка осуществляется аппаратурой в процессе выполнения программы. Она использует упрощенные алгоритмы распараллеливания.

Статическая загрузка предусматривается при трансляции программы. Транслятор оптимизирует использование оборудования, также решая задачи распараллеливания. Это выражается в формировании «длинных» командных слов, задающих работы устройствам АЛУ в каждом машинном такте.

Основная сложность распараллеливания заключается в соблюдении частичной упорядоченности распределяемых работ. Поэтому решение задач синхронизации параллельного вычислительного процесса сопровождает все усилия по организации совместной работы устройств. Это сказывается и при решении всех задач эффективного использования расслоенной многоуровневой памяти ВС.

Отечественный опыт создания семейства МВК (многопроцессорных вычислительных комплексов) «Эльбрус», модели которого относятся к симметричным ВС, и последующее проектирование позволили проанализировать, разработать и применить ряд существенно новых, важных и перспективных решений в распараллеливании как самого вычислительного процесса, так и работы отдельных устройств. Разработка пронизана такими решениями, они ложатся в основу проектирования развития семейства, являются основой обобщений и дальнейшего исследования возможности применения.

**6.2 Классификация параллельных ВС.**

**6.2.1 Потоки команд и потоки данных**

Общепринятую удачную классификацию ВС, которую предложил в 1970 г. Г. Флин (США). Основным определяющим архитектурным параметром он выбрал взаимодействие *потока команд* и *потока данных* (операндов и результатов).

В ЭВМ классической архитектуры ведется последовательная обработка команд и данных. Команды поступают одна за другой (за исключением точек ветвления программы), и для них из ОЗУ или регистров так же последовательно поступают операнды. Одной команде (операции) соответствует один необходимый ей набор операндов (как правило, два для бинарных операций). Этот тип архитектуры – «***один поток команд – один поток данных», ОКОД (SISD  – "SingleInstruction, SingleData")*** условно изображен на [рис. 6.1](http://www.intuit.ru/department/hardware/paralltech/1/#image.1.1).

Память

Поток данных после обработки

Поток команд

Поток данных для обработки

Процессор

Рис. x.1.  ВС типа ОКОД (SISD)

Тип ***ОКМД – «один поток команд – много потоков данных» (SIMD – "SingleInstruction – MultiplеData")*** охватывает ВС, в которых одной командой обрабатывается набор данных, множество данных, вектор, и вырабатывается множество результатов. Это векторные и матричные системы, в которых по одной команде выполняется одна и та же операция над всеми элементами массива – вектора или матрицы, распределенными между *процессорными (обрабатывающими) элементами ПЭ или процессорами.* Принцип обработки показан на [рис. 6.2](http://www.intuit.ru/department/hardware/paralltech/1/#image.1.2).

ПЭ1

ПЭ2

ПЭ*n*

Память данных

Память программ

Устройство управления

Поток данных для обработки

Поток данных после обработки

Поток команд

……………

Рис. 6.2.  ВС типа ОКМД (SIMD)

Отечественные векторные ВС – ПС-2000, ПС-2100. Допускают организацию матричной обработки. Классический пример матричной архитектуры – ILLIAC-1V (США).

К типу ***МКОД – «много потоков команд — один поток данных» (MISD – "MultipleInstruction – SingleData")*** принято относить векторный конвейер (обычно в составе ВС, чтобы подчеркнуть основной используемый принцип вычислений), например, в составе ВС *Crey*-1, «Электроника ССБИС». На векторном конвейере производится последовательная обработка одного потока данных многими обрабатывающими устройствами (ступенями, станциями) конвейера.

К такому же типу относится ВС, реализующая *макроконвейер* (ВС «Украина»). В ней задача, решаемая циклически, «разрезается» на последовательные этапы, закрепляемые за отдельными процессорами. Запускается конвейер многократного выполнения цикла, составляющего задачу.

Принцип обработки показан на [рис. x.3](http://www.intuit.ru/department/hardware/paralltech/1/#image.1.3)*.*



Память

Потоки команд

Поток данных для обработки

Поток данных после обработки

УУ

1

2

*n*

……………

Рис.6.3.  ВС типа МКОД (MISD)

Тип ***МКМД – «много потоков команд – много потоков данных» (MIMD – "MultipleInstruction – MultipleData")***cоответствует более полному и независимому распараллеливанию. К этому типу относятся, например, все многопроцессорные вычислительные комплексы (МВК) семейства «Эльбрус».

**6.1.2** «**Фон-Неймановские» и «не-Фон-Неймановские» архитектуры**

**Немного истории.** Первую ЭВМ создал в 1939 г. в США профессор Джон Атанасов, болгарин, со своим аспирантом К. Берри. Две малые ЭВМ, созданные ими в период 1937 – 1942 гг., были прототипами большой ЭВМ АВС для решения систем линейных уравнений, которая в 1942 г. доводилась по устройствам ввода-вывода и должна была войти в строй в 1943 г., но призыв Атанасова в армию в 1942 г. воспрепятствовал этому. Проект электронной ЭВМ Эниак (ElectronicsNumericalIntegratorandComputer) был сделан в 1942 г. Д. Моучли и Д. Эккертом и осуществлен в 1945 г. в Муровской электротехнической лаборатории Пенсильванского университета. В 1946 г. Эниак был публично продемонстрирован в работе. В нем впервые были применены триггеры. Рождение Эниак считают началом компьютерной эры, посвящая ему научные симпозиумы и другие торжественные мероприятия. (Международный симпозиум, посвященный 50-летию первой ЭВМ, был проведен и в Москве в июне 1996 г.)

Однако еще в начале 40-х г. XX века Атанасов поделился с Моучли информацией о принципах, заложенных в ЭВМ АВС. Хотя Моучли впоследствии утверждал, что он не воспользовался этой информацией в патенте на Эниак, суд не согласился с этим. Вернувшись из армии после войны, Атанасов узнал, что более мощная ЭВМ Эниак уже создана, и потерял интерес к этой теме, не поинтересовавшись, насколько Эниак похож на его ЭВМ АВС.

Известный английский математик Алан Тьюринг был не только теоретиком по информации и теории алгоритмов, автором теоретического автомата «машины Тьюринга», но и талантливым инженером, создавшим в начале 1940-х г. первую работающую специализированную ЭВМ. Эта ЭВМ под названием «Колосс» была сконструирована и сделана им совместно с Х. А. Ньюменом в Блетчи (Англия) и начала работать в 1943 г. Сообщения о ней своевременно не публиковались, т. к. она использовалась для расшифровки секретных германских кодов во время войны.

Основные архитектурно-функциональные принципы построения ЦВМ были разработаны и опубликованы в 1946 г. венгерским математиком и физиком Джоном фон Нейманом и его коллегами Г. Голдстайном и А. Берксом в ставшем классическим отчете «Предварительное обсуждение логического конструирования электронного вычислительного устройства».Основополагающими принципами ЭВМ на основании этого отчета являются: 1) принцип программного управления выполнением программы, и 2) принцип хранимой в памяти программы. Они легли в основу понятия ***фон-Неймановской архитектуры***, широко использующей *счетчик команд*.

Вернемся к настоящему. Счетчик команд отражает «узкое горло», которое ограничивает поток команд, поступающих на исполнение, их последовательным анализом.

Альтернативной архитектурой является «***не-фон-Неймановская» архитектура***, допускающая одновременный анализ более одной команды. Поиски ее обусловлены необходимостью распараллеливания выполнения программы между несколькими исполнительными устройствами – процессорами. Счетчик команд при этом не нужен. Порядок выполнения команд определяется наличием исходной информации для выполнения каждой из них. Если несколько команд готовы к выполнению, то принципиально возможно их назначение для выполнения таким же количеством свободных процессоров. Говорят, что такие ВС управляются *потоком данных (dataflow).*

В командах проверки условия возможно альтернативное задание адреса результата (ИЛИ – ИЛИ).

Адреса результатов являются адресами ПК, т. е. результаты выполнения одних команд в качестве операндов могут поступать в текст других команд. Команда не готова к выполнению, если в ее тексте отсутствует хотя бы один операнд. Ячейка, обладающая полным набором операндов, переходит в возбужденное состояние и передает в селекторную сеть информационный пакет (токен), содержащий код операции и необходимую числовую и связную информацию. Он поступает по сети на одно из исполнительных устройств.

Там же операция выполняется, и в распределительную сеть выдается результирующий пакет, содержащий результат вычислений и адреса назначения в ПК (возможно, за счет выбора альтернативы, т. е. такой выбор – тоже результат). По этим адресам в ПК результат и поступает, создавая возможность активизации новых ячеек. После выдачи токена в селекторную сеть операнды в тексте команды уничтожаются, что обеспечивает повторное выполнение команды в цикле, если это необходимо.

Общая схема потоковых ВС представлена на [рис. 6.4](http://www.intuit.ru/department/hardware/paralltech/1/2.html#image.1.4).

ИСПОЛНИТЕЛЬНЫЕ

УСТРОЙСТВА

**. . . . . . . . . . . . .**

КОММУНИКАЦИОННАЯ

СЕТЬ

Распределительная

сеть

Селекторная сеть

ПАМЯТЬ КОМАНД (ПК)

ТОКЕНЫ

**. . . . . . . . . . . . .**

Рис. 6.4. «Идеальная» потоковая ВС

Программа или ее часть (сегмент) размещается в памяти команд ПК, состоящей из ячеек команд. Команды имеют следующую структуру.

{код операции, операнд 1, ..., операнд L,

адрес результата 1, ..., адрес результата M}

Селекторная и распределительная сети образуют ***коммуникационную сеть*** ВС.

Ожидаемая сверхвысокая производительность такой системы может быть достигнута за счет одновременной и независимой активизации большого числа готовых команд, проблематичном допущении о бесконфликтной передаче пакетов по сетям и параллельной работы многих исполнительных устройств.

Существует ряд трудностей, в силу которых «не-фон-Неймановские» архитектуры не обрели технического воплощения для массового применения в «классическом», отраженном выше, исполнении. Однако многие устройства используют данный принцип, но чаще всего взаимодействие процессоров при совместном решении общей задачи и их синхронизация при использовании общих данных основаны на анализе готовности данных для их обработки. Это дает основание многим конструкторам заявлять, что в своих моделях они реализовали принцип *dataflow*.

**6.1.3. Системы с общей и распределенной памятью**

***Системы с общей (разделяемой) оперативной памятью*** образуют современный класс ВС – многопроцессорных супер-ЭВМ. Одинаковый доступ всех процессоров к программам и данным представляет широкие возможности организации параллельного вычислительного процесса *(параллельных вычислений).* Отсутствуют потери реальной производительности на межпроцессорный (между задачами, процессами и т. д.) обмен данными ([рис. 6.5, *a*](http://www.intuit.ru/department/hardware/paralltech/1/2.html#image.1.5)).

***Системы с распределенной памятью*** образуют вычислительные комплексы (ВК) – коллективы ЭВМ с межмашинным обменом для совместного решения задач ([рис. 6.5, *б*](http://www.intuit.ru/department/hardware/paralltech/1/2.html#image.1.5)).

П

С

*n*

1

. . . . . . . .

ЛОП1

Система обмена

1

*n*

Процессоры

1

ОП1

ОП1

ОП*n*

*а*

*б*

Рис. 6.5.  ВС с общей (*а*) и распределённой (*б*) памятью

В ВК объединяются вычислительные средства систем управления, решающие специальные наборы задач, взаимосвязанных по данным. Принято говорить, что такие ВК выполняют *распределенные вычисления*, а сами ВК называют *распределенными ВК.*

Другое, противоположное воплощение принципа МИМД – *масспроцессорные* или высокопараллельные архитектуры, объединяющие сотни – тысячи – десятки тысяч процессоров.

В современных супер-ЭВМ наметилась тенденция объединения двух принципов: общей (распределяемой) и распределенной (локальной) оперативной памяти (ЛОП). Такая структура используется в проекте МВК «Эльбрус-3» и «Эльбрус-3М» ([рис. 6.6](http://www.intuit.ru/department/hardware/paralltech/1/2.html#image.1.6)).

1

*n*

**. . . . . . .**

Коммутатор

**. . . . . . . . . . . . . . .** .

Общая ОП

ЛОП*1*

ЛОП*n*

М1

М*n*

Рис. 6.6.  Схема ВС с модулями локальной памяти

**6.2. Способы межмодульного соединения (комплексирования)**

Различают два противоположных способа комплексирования: *с общей шиной (шинная архитектура) и с перекрестной (матричной) коммутацией* модулей ВС (процессоров, модулей памяти, периферии).

На [рис. х.7](http://www.intuit.ru/department/hardware/paralltech/1/2.html#image.1.7) представлена система с общей шиной. Шина состоит из линий, по которым передаются информационные и управляющие сигналы.

1

Общая шина

**. . . . . . . . . . . . . . . . . . . . .**

Периферийные устройства

Процессоры

Модули памяти

*n*

1

*m*

**Рис. 6.7.**  Схема ВС с общей шиной

Шина используется в режиме разделения времени, при котором лишь один модуль в данный момент работает на передачу. Принимать принципиально могут все модули, хотя преимущественно информация при выдаче в нее адресуется. Применяется в микро- и мини-ЭВМ при сравнительно небольшом числе модулей. Практически производится разделение шины на управляющую, адресную и шину данных.

В высокопроизводительных ВС для возможности одновременного обмена многими парами абонентов используется перекрестная или матричная коммутация.

***Матричный коммутатор*** можно представить (прямоугольной) сеткой шин. К одному концу каждой подсоединен источник-потребитель информации ([рис. 6.8](http://www.intuit.ru/department/hardware/paralltech/1/2.html#image.1.8)). Точки пересечения — *узлы* этой сетки — представляют собой *управляющие ключи*, которые соединяют или разъединяют соответствующие шины, устанавливая или прекращая связь между модулями. Реализуется связь «каждый с каждым». Одновременно могут связываться многие (до *n*/2) пары модулей.

1

2

*n*

Коммутатор

Управляющие

сигналы

.

.

Процессоры

.

. . .

. . .

*а*

1

2

…

n

1 2 *m*

Процессоры

Коммутатор

Управляющие

сигналы

Модули памяти

*б*

Рис. 6.8.  Матричные коммутаторы: *а*– перекрёстная коммутация процессоров, *б*– коммутация процессоров и модулей памяти

На [рис. 6.8, *а*](http://www.intuit.ru/department/hardware/paralltech/1/2.html#image.1.8)– перекрестная связь между процессорами в ВС с распределенной памятью, на [рис. 6.8, *б*](http://www.intuit.ru/department/hardware/paralltech/1/2.html#image.1.8)– между n процессорами и m модулями ОП.

|  |  |  |
| --- | --- | --- |
| mhtml:file://F:\INTUIT_ru%20Курс%20Архитектура%20__%20Лекция%20№2%20Микропроцессорные%20системы%20и%20способы%20распараллеливания2.mht!http://www.intuit.ru/img/empty.gif | **6.3. МИКРОПРОЦЕССОРНЫЕ СИСТЕМЫ И СПОСОБЫ РАСПАРАЛЛЕЛИВАНИЯ**   |  | | --- | | **6.3.1. Мультимикропроцессорные вычислительные системы**  В настоящее время выбор сделан в пользу многопроцессорных симметричных ВС типа *MIMD*, обеспечивающих виртуализацию вычислительных ресурсов. Основу такой ВС составляет *суперскалер*, сосредоточивший в себе все способы достижения максимального быстродействия при выполнении одиночной программы. Векторные и векторно-конвейерные процессоры и системы получили своё место. Их эффективность как самостоятельных установок могла быть достаточно высокой только при решении специальных задач и тестов. Поэтому достаточно быстро выяснилось, что эти установки могут выполнять функции интеллектуальных терминалов при решении основной задачи на другом универсальном вычислительном средстве и выполнять лишь отдельные его заявки. Сегодня стало окончательно ясно, что первые эффективны лишь в роли специализированных вычислительных устройств для решения специальных задач. Вторые твердо заняли место в составе многофункциональных арифметическо-логических устройств (АЛУ) суперскалеров, ибо без конвейеров мы не мыслим себе выполнение всех операций ВС.  Складывается и структура памяти ВС, которая может совмещать в одной установке все способы доступа: от разделяемой (общей) до распределенной оперативной памяти. Однако ограниченные возможности эффективной работы с общей памятью часто диктуют иерархическую структуру ВС, где уровни иерархии (кластеры) отличаются или способом доступа к оперативной памяти, или тем, что каждый кластер имеет свою собственную физическую память в общем адресном пространстве. При этом принцип буферизации, основанный на многоуровневой по быстродействию (и, конечно, — различной по технологии) памяти, на активном использовании Кэш-памяти, продолжает развиваться. Кэш-память, как память самого высокого уровня, претерпевает функциональное разбиение в зависимости от типа данных, для хранения которых она предназначена, либо, в зависимости от вида обработки, — программ или данных.  Все сказанное выше подтверждает перспективность структурных решений при проектировании многопроцессорного комплекса «Эльбрус-3» и его микропроцессорного развития «Эльбрус-3М», «Эльбрус-2К». Таким образом, структура «длинного командного слова» (архитектура *VLIW*, лежащая в основе *EPIC*) попадает в разряд классических.  Сейчас микропроцессор, сконцентрировавший все достижения микроэлектроники, является основной составляющей элементно-конструкторской базы ВС. Поэтому понятие «мультимикропроцессорные ВС» пришло на смену понятию «микропроцессорные ВС».  Анализ современных мультимикропроцессорных ВС позволяет выделить те развиваемые характерные решения, которые в условиях микроминиатюризации и снижения энергоемкости, «экономного» логического развития обеспечивают необходимые свойства универсального применения.  Такими решениями являются следующие.  **6.3.2. Многопроцессорные кристаллы**.   1. Воспроизведение многопроцессорной ВС на одном кристалле в значительной степени характерно для сигнальных вычислительных средств, специализирующихся на обработке двух- и трехмерных изображений, которые применяются в цифровом телевидении и радиовещании, при передаче изображений по каналам связи и др. Такие средства эффективно используются в качестве нейрокомпьютеров.   Например, на одном кристалле *MVP (MultimediaVideoProcessor)* семейства *TMS 320 C80* (фирма *TexasInstrument*) расположены 4 32-разрядных цифровых сигнальных процессора (*DSP*– DigitalSignalProcessor) с фиксированной запятой (*ADSP-0 – ADSP-3*). Их особенность — высокая степень конвейеризации и до 64 бит длина командного слова для параллельного выполнения нескольких операций. Система команд содержит команды над битовыми полями и структурами данных, несущими графическую информацию. Такая специализация обусловила понятие — *DSP-архитектура*.  Процессоры работают независимо. Т. е. ВС – типа *MIMD* – (Multiple-Instruction, Multiple-Data). Программируются отдельно на ассемблере или ЯВУ. Данными обмениваются через общую внутрикристальную память.  Каждый из *ADSP* содержит КЭШ-память команд (2 Кбайта), и через матричный коммутатор *Crossbar* получает доступ к 32 из имеющихся 50 Кбайт быстродействующей статической внутренней памяти. Память расслоенная — поделена на сегменты. Если два и более процессора в одном цикле попытаются обратиться к одному сегменту, аппаратная система управления доступом с циклическим изменением приоритета (*roundrobinprioritization*) позволит сделать это только одному процессору.  32-разрядное АЛУ *ADSP* может работать как два 16-разрядных или четыре 8-разрядных АЛУ. Этого достаточно для обработки видеоизображений. Специальные блоки ускоряют обработку графики. Блоки генерации адресов формируют кольцевые (бесконечные) буферы. Аппаратно поддержаны три вложенных цикла.  *RISC*-процессор управляет четырьмя *ADSP* с помощью диспетчера. Диспетчер и планировщик заданий тесно взаимодействуют с контроллером пересылок. Кроме того, управляющий процессор самостоятельно выполняет вычисления и обеспечивает обмен с внешними устройствами. Содержит встроенный блок плавающей арифметики и набор векторных операций с плавающей запятой, оптимизированных для обработки изображений, звука и трехмерной графики.  **6.3.3. Транспьютерная технология.**   1. Представленная выше архитектура обладает такой конструктивной законченностью, которая позволяет как встраивать ее в некоторую систему, так и организовать взаимодействие нескольких кристаллов. Это обеспечивается развитыми средствами связи и обмена данными.   Возможность комплексирования привлекла внимание еще на раннем этапе развития микропроцессоров (в середине 1980-х годов) и привела к построению *транспьютеров* — микропроцессоров, снабженных развитыми средствами комплексирования. Таким образом, создавались "кирпичики", на основе которых можно было создавать сложные структуры. Эта тенденция не только сохранилась, но является необходимым средством построения мультимикропроцессорных ВС.  Преследуя многофункциональность средств обмена, не обязательно требовать их размещения на одном кристалле с центральным процессором. Так, фирма *AnalogDevices* предлагает микропроцессоры *ADSP-21060/62 SHARC* ("АКУЛА") для цифровой обработки сигналов, специально предназначенные для комплексирования.  Средства комплексирования «АКУЛЫ»:   * + магистраль для подключения 6 «АКУЛ» и одного ХОСТ-процессора (управляющего, с привилегированным доступом к магистрали, а также к памяти каждого процессора — через специальный порт);   + сигнальные регистры в составе каждого процессора, непосредственно связанные (одной ножкой) с каждым из других процессоров — для контроля их состояния;   + ЛИНКи — каждый процессор имеет 6 выходов (ЛИНКов) для непосредственной связи «процессор – процессор».  1. ***Общее адресное пространство***комплексируемых микропроцессоров «АКУЛА» обеспечивает псевдообщую память и исключает необходимость программной организации обмена данными. Если адрес физически принадлежит ОП другого процессора, то обмен организуется автоматически, без вмешательства пользователя (т. е. программно не предусматривается). 2. ***Межпроцессорный (магистральный) обмен*** инициируется в том случае, если адрес считывания или записи принадлежит адресному пространству другого процессора (единичный обмен). Аналогично возникают групповые пересылки данных с использованием "чужого" адресного пространства.   Пользователь не составляет программу обмена, даже для контроллера обмена данных. Достаточно указать «чужие» адреса.  Процессоры обмениваются сигналами состояния. Поэтому каждый процессор знает, кто является «хозяином» магистрали, т.е. ведет обмен, и свой приоритет в очереди к магистрали. По завершении каждого обмена производится циклическая смена приоритетов процессоров, которым нужна магистраль. Процессор с максимальным приоритетом становится «хозяином». Обмен может прерываться только ХОСТ-процессором.  Микропроцессор утверждается в роли основы элементно-конструкторской базы ВС, и это поняли ведущие разработчики.  В этом смысле привлекает внимание трансформация интересов "отца суперкомпьютеров" С.Крея, который признал определяющую роль принципа *MIMD* при построении ВС *CraySuperserver 6400 System* (*CS640*), выпускаемой корпорацией *CrayResearch* в сотрудничестве с компанией *SUN Microsystems* (сотрудничество с фирмой *SUN* ныне характерно и для ведущих российских разработчиков).  Система предполагает наращиваемую конфигурацию от 4 до 64 процессоров *SuperSPARC*. Принято компромиссное решение на основе классической схемы разделения (общей) ОП при многопроцессорной обработке и распределенной памяти при параллельной обработке массивов. Чтобы работать с частично распределенной памятью в ОЗУ, ВС имеет в любой конфигурации 4 шины. Шина использует сетевую технологию «коммутации пакетов». Это позволяет находить путь обмена единицами информации в соответствии с занятостью или освобождением шин.  В целом, архитектуру следует считать шинной, хотя наличие нескольких шин делает ее промежуточной между шинной и использующей матричный коммутатор.  **6.4. Направление «мини-супер» призвано поддержать персональный компьютер**  То, что говорилось выше, «по умолчанию» соответствует разработке супер-ЭВМ, предназначенных для решения особо сложных задач в составе систем управления в реальном времени, моделирования сложнейших физических процессов, решения задач исследования операций, задач искусственного интеллекта, выполнения роли майнфреймов и серверов в локальных, корпоративных и глобальных сетях.  Супер-ЭВМ уникальна, мало тиражируема, цена ее высока.  С другой стороны, ничто уже не может остановить "победного шествия" персональных компьютеров. Область применения их стала всеобъемлющей. Они используются и там, где могут справиться с задачами, и там, где уже не справляются, несмотря на применение современныхсуперскалеров.  Тогда целесообразно поставить следующую проблему.  Введем в состав персонального компьютера (*РС*), как его внешнее устройство, мультимикропроцессорную систему (мультипроцессор), использование которого в монопольном и однозадачном режиме может обеспечить успешное решение задач повышенной сложности.  Действительно, разрешение этой проблемы позволило бы заполнить определенную нишу между супер-ЭВМ и *PC*, вывести персональный компьютер на уровень мини-супер-ЭВМ. Применение мультипроцессора *РС* в однопрограммном режиме, при жестком распределении памяти, использование (см. далее) прогрессивной технологии «одна программа — много потоков данных» позволяют существенно снизить издержки производительности на работу ОС, легко «врезать» их в современные операционные системы компьютеров. Сборка такой системы должна производиться на основе существующей микропроцессорной элементно-конструкторской базы, с минимальным использованием вновь разрабатываемых элементов.  Здесь воспроизводится упомянутая выше идея о наличии мониторной системы, на которой решается основная задача, и о наличии интеллектуального терминала, который берет на себя функции, обеспечивающие общую эффективность системы.  Общая схема такой установки показана на [рис. 6.9](http://www.intuit.ru/department/hardware/paralltech/2/2.html#image.2.1). Выбраны конкретные значения параметров.  Мониторная  система — PC  Память команд  Память команд  ЛОП  ЛОП  КЭШ команд  КЭШ команд  Процессор 2  Процессор 1  Общая шина или коммутатор  Средства синхронизации (семафоры, закрытые адреса, предикаты)  Синхронизатор  Общая разделяемая память  Модуль 1  Модуль 16  Рис. 6.9.  Схема ВС для персонального компьютера  Мультимикропроцессорную приставку к персональному компьютеру целесообразно разработать на основе исследования принципов построения локально-асинхронной архитектуры (*SPMD*-технологии). Важным достоинством архитектуры является сведение традиционных функций ОС на уровень команд. Т.е. система команд мультипроцессора такова, что позволяет реализовать функции управления параллельным процессом, не требуя запуска процедур ОС. Способствует простоте управления параллельным процессом также монопольный и однозадачный режим использования мультипроцессора. Ниже мы подробнее остановимся на принципах *SPMD*-технологии. Предполагая первоначальное знакомство с этими принципами, отметим следующее.  Известно (см. далее), что семафоры — универсальное средство синхронизации. Однако семафоры традиционно используют ОС. Чтобы этого избежать, семафоры следует реализовать с помощью *предикатного механизма*, т.е. с использованием памяти предикатов.  Семафорный механизм может быть эффективно реализован с помощью *механизма закрытия адресов* (памяти закрытых адресов).  Тогда, в общем случае применения семафоров, должны быть введены команды следующего вида.  ***Считать по семафору (Сч(С) А***). Считывание по адресу производится в случае, если указанный семафор (реализованный в памяти регистрового типа, наряду с индексными и базовыми регистрами) открыт. Если семафор закрыт, реализуется ожидание данного ПЭ без прерывания (т.е. в данном применении пользователь может быть допущен к операциям над семафорами типа «жужжать»).  ***Записать по семафору (Зап(С) А***). Запись по адресу производится аналогично предыдущей команде.  При использовании памяти закрытых адресов необходима лишь команда*Закрыть адрес*. Любое последующее считывание по этому адресу циклически возобновляется (в режиме "жужжания") до тех пор, пока по этому же адресу другой процессор не произведет запись.  В случае использования механизма предикатов адрес некоторой булевой переменной записывается в специальные разряды командного слова. Команда, для которой указанный в ней предикат имеет значение 0, выполняется, в соответствии с кодом операции, в *спекулятивном режиме* в двух модификациях:   * ожидается присвоение данному предикату значения 1 (в режиме «жужжания»); * пропускается выполнение данной команды.   ПЭ реализует идею *RISC*-архитектуры и представляет собой функционально законченное устройство, состоящее из микропроцессора, схем обрамления и локальной оперативной памяти (ЛОП). Локальная память процессора содержит область для хранения стеков вычислительного процесса, в том числе — стеков подпрограмм и вложенных циклов. В других областях этой памяти хранятся модификаторы, дескрипторы массивов и локальные величины. Здесь же находятся микропрограммы, реализующие систему команд ВС.  Общая (разделяемая) память (ОП) содержит M модулей с общим адресным пространством и реализует принцип ***интерливинга***, предполагающий, что смежные ячейки памяти находятся в разных модулях.  ***Синхронизатор*** предназначен для обеспечения одновременного пуска программ или их модулей.  Возможно использование простейших коммутаторов для обмена ПЭ с модулями памяти. | |

**6.5. Распределенный и разделяемый вычислительный ресурс второго уровня. Решающие поля**

Произведем некоторые обобщения.

Итак, второй уровень распараллеливания предполагает распределение команд, инструкций, операций, элементарных функций и других несложных процедур — для выполнения исполнительными устройствами процессоров или в общем вычислительном ресурсе симметричной ВС. Здесь существуют свои проблемы, связанные с "элементарным" характером операций, небольшим объемом содержащихся в них работ, с их еще большей критичностью по отношению к "накладным расходам" на организацию и синхронизацию. Мы предполагаем, что исполнительные устройства ВС образуют вычислительный ресурс второго уровня (распараллеливания).

Сложились традиции построения этого ресурса, где основное внимание уделяется построению многофункциональных АЛУ. Однако в ряде архитектур пока еще робко пробивает себе дорогу объединение АЛУ в единый разделяемый ресурс системы — построение решающих полей.

Проработка этой идеи проводилась неоднократно в отечественной практике разработки ВС. Она проявлялась во включении в состав ВС специализированных процессоров на правах интеллектуальных терминалов для эффективного выполнения определенных операций. Это были векторные процессоры с доступом от нескольких ЦП (ПС-3100), векторный сопроцессор в ПС-2100. Эта же идея фактически воплощена в семействе «Эльбрус», допускающем включение в свой состав спецпроцессоров — эмуляции других систем, векторно-конвейерных модулей и др.

При реализации идеи решающего поля проблема выбора и развития вычислительного ресурса неотделима от выбора вариантов архитектуры системы вообще. Единственным средством обоснования и исследования этого выбора является моделирование. Построение детерминированных имитаторов позволяет с любой детализацией выявить целесообразные технические решения и обосновать язык системы. Целью стохастических моделей является оценка эффективности различных архитектур на основе полноты полезной загрузки оборудования.

На основе традиций разработки многопроцессорных симметричных вычислительных систем можно сделать вывод о практике и тенденции развития вычислительного ресурса второго уровня.

Все модели семейства МВК «Эльбрус» предполагают наличие в составе АЛУ процессора нескольких исполнительных устройств (ИУ), специализированных по типам операций. Тогда в целом для ВС можно сказать, что вычислительный ресурс второго уровня является *распределенным и неоднородным* ([рис.6.10](http://www.intuit.ru/department/hardware/paralltech/2/3.html#image.2.2)).

УУ

ИУ1

ЦП 1

. . . . .

ЦП N

УУ

ИУ*n*

ИУ1

ИУ*n*

Рис.6.10. ВС с распределённым решающим полем

Выше говорилось, что использование ресурса второго уровня неотделимо от общих идей функционирования процессоров ВС — от их архитектуры и от архитектуры ВС в целом. Поэтому сказанного о распределенном ресурсе недостаточно, надо говорить и о способе его использования.

Так, в МВК «Эльбрус-2» применена динамическая загрузка ИУ в процессе выполнения последовательности безадресных команд программы, которую подробно рассмотрим в лекции 3. Обобщенный алгоритм такой загрузки основан на промежуточном переводе безадресных команд в трехадресные. Появление адресов аргументов и результатов в каждой команде позволяет на основе совместного анализа нескольких команд, представленных в "окне просмотра", выявлять их частичную упорядоченность и выделять независимые команды для одновременного выполнения. Это — адресный метод распараллеливания, который рассматривался в архитектуре *dataflow*. Команды проходят стадии обработки, как показано на [рис.Y.3](http://www.intuit.ru/department/hardware/paralltech/2/3.html#image.2.3).

Однако проект МВК «Эльбрус-3», породивший микропроцессорное воплощение — МВК «Эльбрус-3М», — основан на использовании идеи «длинного» командного слова и управления каждым тактом системы. Динамическое распределение работ между ИУ заменено статическим — предписанием каждому ИУ, что он должен начать делать в данном такте. В «длинном» командном слове, в соответствующих позициях, записаны инструкции каждому ИУ.

Дополнение снизу «окна просмотра» новой командой

Исключение из «окна просмотра» выполненных команд

Формирование для новой команды признака возможности выполнения на основе совместного анализа адресов команды и адресов выше находящихся команд

Выборка из «окна просмотра» команд, готовых к выполнению, и назначение их на свободные ИУ

Рис. 6.11. Схема оптимизатора-компоновщика «длинных» командных слов

Это означает, что решение проблем оптимального использования ИУ, их синхронизации при выполнении данного алгоритма возлагаются на оптимизирующий транслятор. Он фактически производит диспетчирование, оптимальное планирование параллельного вычислительного процесса на одном процессоре.

**6.6. Способы распараллеливания**

Различают два основных способа распараллеливания: *по управлению* и *по информации*.

**Первый способ** — представление алгоритма задачи в виде частично-упорядоченной последовательности выполняемых работ. Затем в результате диспетчирования реализуется оптимальный план выполнения работ в ВС при ограничениях на время выполнения всего алгоритма или за минимальное время.

Основой является представление алгоритма граф-схемой G, отражающей информационные связи между работами (задачами, процессами, процедурами, операторами, макрокомандами и т.д.), на которые разбит алгоритм. Граф G — взвешенный, ориентированный, без контуров.

Для исследования графа и диспетчирования используют матрицы следования S; их дополняют столбцом T весов — получают расширенные матрицы следования S\* ([рис. 6.12](http://www.intuit.ru/department/hardware/paralltech/2/4.html#image.2.4)).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 | T |
| 1 |  |  |  |  |  |  | 2 |
| 2 |  |  |  |  |  |  | 1 |
| 3 |  |  |  |  |  |  | 2 |
| 4 |  |  |  |  |  |  | 3 |
| 5 |  |  |  |  |  |  | 2 |
| 6 | 1 |  | 1 |  |  |  | 2 |

Рис. 6.12.  Исходная информация для распараллеливания

Здесь предполагаем, что ВС — однородная, с общей (разделяемой) памятью, т.е. потерями времени на обмен между работами можно пренебречь.

Пусть ВС содержит два процессора (*n* = 2). Тогда в результате оптимального распределения получим план ([рис. 6.13](http://www.intuit.ru/department/hardware/paralltech/2/4.html#image.2.5)).

0

2

5

6

1

1

3

4

2

1

2

3

4

5

6

7

*t*

Рис.6.13.  Временная диаграмма параллельного выполнения работ

План действительно совпадает с оптимальным, т.к. длина расписания T = 7, что совпадает с длиной критического пути в графе, Tкр = 7 (путь 1 mhtml:file://F:\INTUIT_ru%20Курс%20Архитектура%20__%20Лекция%20№2%20Микропроцессорные%20системы%20и%20способы%20распараллеливания4.mht!http://www.intuit.ru/img/symbols/srarr.gif3 mhtml:file://F:\INTUIT_ru%20Курс%20Архитектура%20__%20Лекция%20№2%20Микропроцессорные%20системы%20и%20способы%20распараллеливания4.mht!http://www.intuit.ru/img/symbols/srarr.gif4).

В общей схеме организации параллельного вычислительного процесса мы не полностью раскрыли содержание блока 3 — интерпретации потока макроинструкций в виде, удобном для работы диспетчера. Сейчас мы определили, что такой вид — это матрица следования. Значит, в случае необходимости автоматического формирования матрицы следования надо определять информационную взаимосвязь макроинструкций в пределах видимости, т. е. в «окне просмотра». Таким образом, по текущему содержимому «окна просмотра» надо формировать текущий вид матрицы следования.

Вспомним, что мы уже в упрощенном виде решали подобную задачу, например, когда по формируемому потоку трехадресных команд определяли их информационную взаимосвязь и определяли возможность одновременного выполнения этих команд.

Обобщим эту задачу.

Возвращаясь к названной схеме, представим себе, что поток макроинструкций (блок 2) следует через «окно просмотра» так, что для планирования оптимальной загрузки процессоров диспетчер может анализировать некоторое множество этих макроинструкций и из них выбирать вариант назначения их на процессоры для выполнения. Каждая макроинструкция может интерпретироваться и как процедура, где можно выделить имя θμ множество {mhtml:file://F:\INTUIT_ru%20Курс%20Архитектура%20__%20Лекция%20№2%20Микропроцессорные%20системы%20и%20способы%20распараллеливания4.mht!http://www.intuit.ru/img/symbols/alpha.gifμ} входных параметров, множество {βμ} выходных параметров. На [рис. 6.14](http://www.intuit.ru/department/hardware/paralltech/2/4.html#image.2.6) отображено «окно просмотра», через которое следует поток макроинструкций.

m



«окно просмотра»

Рис. 6.14.  Обработка «окна просмотра»

Составим по его содержимому соответствующую матрицу следования размерности *m* × *m*:

По матрице следования *S* диспетчер производит назначение.

После выполнения макроинструкций они исключаются из "окна прросмотра", оставшиеся макроинструкции уплотняются вверх, а снизу "окно просмотра" пополняется новыми макроинструкциями. С учетом вновь поступившихмакроинструкций уточняется текущий вид матрицы следования *S* и процесс диспетчирования продолжается.

По такой же схеме, а именно, на основе первого способа распараллеливания — по управлению — решается другая важная задача распараллеливания: *компоновки длинных командных слов в оптимизирующем трансляторе*. Назначение работы на ИУ осуществляется здесь в виде записи соответствующей инструкции в позицию длинного командного слова, соответствующую ИУ. т.е. план параллельного выполнения работ (команд, операций) фиксируется в длинных командных словах, в которых предусмотрены инструкции каждому ИУ, которые они должны начать выполнять с данного такта.

**Второй способ распараллеливания** — по информации — используется тогда, когда можно распределить обрабатываемую информацию между процессорами для обработки по идентичным алгоритмам (по одному алгоритму)

1. **Рассмотрим задачу умножения матриц:**

Развернем матрицу — результат *С* — в линейный (одномерный) массив, переименуем ее элементы и заменим два индекса на один:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | … |  |  | … |  |  | … |  |
|  |  | … |  |  | … |  |  | … |  |

Пусть ВС содержит *n* процессоров. Выберем следующий план счета элементов матрицы *C*:

процессор 1 считает элементы *d*1, *d*1+*n*, *d*1+2*n*, ...

процессор 2 считает элементы *d*2, *d*2+*n*, *d*2+2*n*, ...

…………………………………………………………..

процессор *n* считает элементы *dn*, *d*2*n*, *d*3*n*, ...

По-видимому, все они будут выполнять одну и ту же программу, но обрабатывать разные наборы данных. (Мы снова столкнулись с целесообразностью *SPMD*-технологии.)

Здесь не потребовалась какая-либо синхронизация параллельного вычислительного процесса.

**2. РАССМОТРИМ ЗАДАЧУ СЧЕТА СПОСОБОМ «ПИРАМИДЫ».**

Эту задачу мы исследовали при рассмотрении ВС типа *SPMD*. Посмотрим еще раз, какая синхронизация нам здесь потребуется.

Пусть необходимо перемножить все элементы некоторого массива {*a*1,*a*2,... , *a*10}. Каждый элемент занимает одну ячейку памяти. Пусть число процессоров в ВС *n*=4. Чтобы распараллелить этот процесс, примем схему счета «пирамидой» ([рис. 6.15](http://www.intuit.ru/department/hardware/paralltech/2/5.html#image.2.7)).

Рис. 6.15.  Граф-схема выполнения операции «свёртки»

Количество уровней операций в ней ]log2 *m*[=]log210[=4 (]*x*[ — ближайшее целое, не меньшее *x*).

Расширим массив, дополнив его ячейками, в которых будем хранить промежуточные частные произведения. Тогда весь план счета примем таким, как показано на [рис. 6.16](http://www.intuit.ru/department/hardware/paralltech/2/5.html#image.2.8). Отмечены процессоры, выполняющие указанную операцию.

Рис. 6.16.  Схема выполнения операции свёртки четырьмя процессорами

Следовательно, надо так написать программу, одну для всех процессоров, предусмотрев необходимую переадресацию для выборки и вычисления «своих» данных, чтобы по ней выбирались два соседних элемента этого удлиненного массива, а результат их умножения отправлялся в очередную ячейку этого «удлинения».

Возникает только одна трудность: для первых пяти произведений данные есть, а вот последующие произведения должны выполняться тогда, когда для них будут найдены исходные данные.

Значит, процессоры, которым выпало произвести такие умножения, должны «уметь» обнаруживать отсутствие данных и дожидаться их появления. Т. е. требуется синхронизация процессоров по использованию общих данных.

Здесь распараллеливание по данным смыкается с распараллеливанием по управлению.

Возможная схема общей для всех процессоров программы — на [рис.6.17](http://www.intuit.ru/department/hardware/paralltech/2/5.html#image.2.9). Она реализована в примере для типа *SPMD.*

Общая синхронизация ВС

Запрет ображения по считыванию к ячейке, в которую будет помещен результат операции

Выполнение операции, запись результата в «закрытую» ячейку, снятие запрета на использование этой ячейки

Переадресация или выход

ВС

|  |
| --- |
| Рис. 6.17 Конвейеры операций  Выполнение любой операции складывается из нескольких последовательных этапов, каждый из которых может выполняться своим функциональным узлом. Это легко показать на операциях сложения и умножения. Выполнение деления мантисс (порядки вычитаются) чаще всего производятся с помощью вычитания из делимого делителя, сдвига влево полученного остатка, нового вычитания делителя из результата сдвига и т.д. В некоторых ВС находится обратная величина делителя с помощью аппроксимирующих полиномов. Затем делимое умножается на эту величину.  Пусть задана операция, выполнение которой разбито на n последовательных этапов. Пусть *ti* — время выполнения *i*-го этапа. При последовательном их выполнении операция выполняется за время  Выберем *время такта* — величину *t*T = maxt*i*} и потребуем при разбиении на этапы, чтобы для любого *i* = 1, ...,n выполнялось условие *ti* + *t*(*i*+1) mod *n*>*t*T. Т.е. чтобы никакие два последовательных этапа (включая конец и новое начало операции) не могли быть выполнены за время одного такта.  Функциональные узлы, выполняющие последовательные этапы одной операции, целесообразно выстроить в единую конвейерную линию, где устройство, выполняющее некоторый этап, закончив его для операции над одним набором данных, переходило бы в следующем такте к выполнению этого же этапа той же операции для другого набора исходных данных.  Например, на [рис. 6.1](http://www.intuit.ru/department/hardware/paralltech/3/#image.3.1)8 представлен конвейер выполнения операции сложения.  Рис. 6.18.  Выполнение операции сложения на конвейере  Загрузка операндов (этап 1)  Выравнивание порядков (этап 2)  Сложение мантисс (этап 3)  Загрузка операндов (этап 4)  Запись результата (этап 5)  Пусть реализуется поток команд одного процессора или существует доступ к этому устройству нескольких процессоров так, что в каждом такте возможно задание на выполнение сложения новой пары чисел. Тогда временная диаграмма работы конвейера может иметь вид, представленный на [рис. 6.1](http://www.intuit.ru/department/hardware/paralltech/3/#image.3.2)9.  2  1  3  4  5  1  2  3  1  2  3  4  уровень 1  уровень 2  уровень 3  уровень 4  уровень 5  0  *t*T  2*t*T  3*t*T  4*t*T  5*t*T  6*t*T  …  *t*  1  2  3  4  5  6  2  3  1  2  Рис. 6.19.  Схема заполнения конвейера  Максимальное быстродействие процессора при полной загрузке конвейера составляет  Число *n* — *количество уровней конвейера*, или *глубина перекрытия*, т. к. каждый такт на конвейере параллельно выполняются n операций. Чем больше число уровней (станций), тем больший выигрыш в быстродействии может быть получен.  В проекте МВК «Эльбрус-3» АЛУ его ЦП имеет конвейерные ИУ сложения (*n*=5), умножения (*n*=5), деления (*n*=8 для полусловного формата, — 32 разряда, *n*=16 для словного формата). Логические операции также выполняются на конвейере с *n*=2.  Известна оценка  т. е. выигрыш в быстродействии получается в *n*-раз.  Реальный выигрыш в быстродействии оказывается всегда меньше, чем указанный выше, поскольку:   1. некоторые операции, например, над целыми, могут выполняться за меньшее количество этапов, чем другие арифметические операции. Тогда отдельные станции конвейера будут простаивать. 2. при выполнении некоторых операций на определённых этапах могут требоваться результаты более поздних, ещё не выполненных этапов предыдущих операций. Приходится приостанавливать конвейер. 3. поток команд порождает недостаточное количество операций для полной загрузки конвейера.   **6.7. Векторные конвейеры. «Зацепление» векторов**  Наряду с использованием конвейеров для обработки единичных (скалярных) данных, используют так называемые *векторные конвейеры*, единичной информацией для которых являются вектора — массивы данных. Применение векторных конвейеров определило класс ВС — *векторно-конвейерных ВС*, сегодня ещё являющихся основой построения некоторых супер-ЭВМ — ВС сверхвысокой производительности.  Для эффективности векторно-конвейерных ВС (например, для подготовки алгоритмов решения задач на ВС «Электроника-ССБИС») необходима векторизация задач. Это — такое преобразование алгоритма, при котором максимально выделяются (если не вся задача сводится к этому) элементы обработки массивов данных одинаковыми операциями. Сюда входят все задачи, основанные на матричных преобразованиях, обработка изображений, сигналов, моделирование поведения среды и т. д.  В основе векторного конвейера лежит то же самое разбиение операции на уровни или этапы выполнения, но он дополняется средствами аппаратной поддержки, позволяющими по информации о векторах организовать последовательную загрузку конвейера элементами векторов, учитывая их длину.  Пусть необходимо выполнить операцию C = A × B, т. е. c*j* = *aj* + *bj*, *j* = 1, ...,N. Пусть на регистрах СОЗУ записаны вектора A и B. Группа регистров отведена для результатов C. Для управления этим процессом известны дескрипторы векторов DA ,DB ,DC, где Dα = {*a*α, *h*α, N}, α= A,B,C,hα — шаг переадресации. Если загрузка векторов производится всегда в одни и те же регистры АЛУ, то достаточно знать значение N. Может задаваться маска M длиной N, состоящая из нулей и единиц. Каждый элемент M соответствует элементу вектора-результата C. Если элемент *mj* = 1 (логическая переменная), то операция получения c*j* производится, в противном случае соответствующие элементы векторов пропускаются. Это применимо для альтернативного счёта в соответствии со значением логических переменных.  Пусть операция умножения выполняется за три этапа. Тогда можно представить временную диаграмму получения N результатов при предположении о назначении функциональных устройств ([рис. 6.20](http://www.intuit.ru/department/hardware/paralltech/3/#image.3.3))*.*  В составе АЛУ может быть два и более конвейерных устройств, специализированных каждое для выполнения некоторой операции. Тогда возможно и эффективно ***«зацепление» векторов***, иллюстрируемое примером на [рис. 6.](http://www.intuit.ru/department/hardware/paralltech/3/#image.3.4)20 для выполнения сложной операции над векторами: D=A× B+C.  Рис.6.20.  Умножение векторов на конвейере  *N*-3  *N*-2  *N*-1  *N*-2  *N*-1  *N*  *N*  1  2  3  Загрузка очередной пары (*a*,*b*), переадресация или остановка  Первый уровень сумматоров  Второй уровень сумматоров  Третий уровень сумматоров  Запись cj  1  2  3  4  5  6  1  2  3  4  5  1  2  3  4  1  2  1  2  3  4  5  6  …………  *N*  *N*+1  *N*+2  *N*+3  *N*+4  *N*  *N*-1  *N*  Рис. 6.21. «Зацепление» векторов  1  2  *n*x  …….  1  2  *n*x  …….  A  B  C  Здесь два конвейера образовали один, с глубиной перекрытия *n*  =  *n*× + *n*+. Очередной результат умножения немедленно направляется на конвейер сложения, куда параллельно направляется необходимый сомножитель. |
| mhtml:file://F:\INTUIT_ru%20Курс%20Архитектура%20__%20Лекция%20№3%20Распараллеливание%20в%20ВС%20на%20уровне%20исполнительных%20устройств1.mht!http://www.intuit.ru/img/empty.gif |
| **6.8. Выполнение операций на стеке**  В вычислительной технике повсеместно используется такая структура данных, как стек. Его применение при выполнении арифметических и логических операций в арифметическо-логическом устройстве (АЛУ) позволяет реализовать безадресную систему команд, что, в свою очередь, дает возможность минимизировать число обращений к оперативной памяти, осуществить буферизацию при многоуровневой памяти, кодировать большое число команд в одном слове.  Рассмотрим механизмы обработки информации на стеке и формирования безадресных команд.  Прежде всего - об идее применения безадресных команд.  Рассматривая структуру команд, мы можем отметить, что практически команды бывают одно-, двух-, трехадресные. Трехадресная команда, как правило, по законченности соответствует одному оператору, т.к. отображает действие над двумя операндами и результат. Двухадресная команда — это такая команда, где один из адресов подразумевается (например, аккумулятор или сумматор), или результат направляется по одному из указанных адресов. То же касательно одноадресной команды: по числу адресов - число обращений к памяти, не считая считывания самой команды.  Возникает вопрос: можно ли отделить загрузку регистров сверхоперативного запоминающего устройства (СОЗУ) от собственно выполнения команды, сделать эту загрузку опережающей, выполнение команды производить только с использованием подразумеваемых адресов СОЗУ и при этом в целом минимизировать количество обращений к ОП? Ответ приводит к структуре безадресных команд и к выполнению операций на стеке.  ***Стек*** представляет собой множество последовательно пронумерованных регистров СОЗУ или ячеек ОП, снабженное ***указателем вершины стека УВС***, в котором автоматически при записи и считывании устанавливается номер (адрес) последнего занятого регистра — ***вершины стека***. При записи в стек (загрузка стека) слово пишется в следующий по номеру регистр, а УВС увеличивается на единицу. При считывании извлекается слово с адресом, указанным в УВС. Затем УВС уменьшается на единицу. Таким образом, в стеке реализуется правило последний «пришел — первый ушел».  Адреса памяти фигурируют только в командах загрузки стека и записи из стека в память.  Значения выражений можно вычислять полностью безадресным способом.  Команда, состоящая только из кода операции, извлекает из стека один или два операнда, выполняет операцию и заносит результат в стек. На [рис. 6.](http://www.intuit.ru/department/hardware/paralltech/3/2.html#image.3.5)22 приведен пример преобразования стека при выполнении трех команд программы.    Рис. 6.22.  Выполнение команд на стеке  2  2  2  ……  17  3  25  3  ……  4  17  3  2  ……  21  3  25  1  ……  3  25  25  УВС:  УВС:  УВС:  УВС:  2  1  0  0  1  1  3  0  1  0  Команда: Загрузка стека изR, (R)=4  «+»  Записать в A  Вычисления с использованием стеков удобно задавать, описывать и программировать с помощью ***польской инверсной (бесскобочной) записи*** арифметических выражений ***ПОЛИЗ*** (предложил польский математик Лукашевич в 1929 г.). Запись производится по правилу "читаем арифметическое выражение слева направо и последовательно друг за другом выписываем встречающиеся операнды". Как только окажется, что все операнды некоторой операции выписаны, записываем знак этой операции, условно учитываем вместо операндов и знака операции потенциальное имя ее результата. Это может породить возможность записи новой операции; или же будем продолжать выписывать операнды, если условно учтенный результат не позволяет записать знак операции. Запись на ПОЛИЗ состоит из цепочек имен и цепочек операций.  Например,y := (*k* + (*l* - *mn*))(*a* - *b*) mhtml:file://F:\INTUIT_ru%20Курс%20Архитектура%20__%20Лекция%20№3%20Распараллеливание%20в%20ВС%20на%20уровне%20исполнительных%20устройств2.mht!http://www.intuit.ru/img/symbols/srarr.gify*klmn* × - + *ab* - ×  Последовательность символов в ПОЛИЗ может рассматриваться как программа вычисления значения данного выражения, если под именами понимать команды загрузки стека (кроме случая, когда имени следует знак операции :=, тогда это имя определяет запись из стека), а под знаками операций — безадресные команды, содержащие только коды операций. Удобно ввести команду вида Зп У — записи из стека по адресу У. Тогда мы получим окончательный вид программы  *klmn* × - + *ab* - × Зп Y (3.1)  Не будем отображать УВС, интерпретируя стек магазинной памятью. Тогда схема выполнения данной программы сводится к представленной на [рис. 6.](http://www.intuit.ru/department/hardware/paralltech/3/2.html#image.3.6)23.    Рис. 6.23.  Выполнение программы на стеке  *n*  *m*  *l*  *k*  *m*x*n*  *l*  *k*  ……..  *l*-*mn*  *k*  …….  *k*+(*l*-*mn*)  ……..  *b*  *a*  *k*+(*l*-*mn*)  ……….  *a*-*b*  *k*+(l-*mn*)  ………..  результат  ………..  A |
| mhtml:file://F:\INTUIT_ru%20Курс%20Архитектура%20__%20Лекция%20№3%20Распараллеливание%20в%20ВС%20на%20уровне%20исполнительных%20устройств2.mht!http://www.intuit.ru/img/empty.gif |
| mhtml:file://F:\INTUIT_ru%20Курс%20Архитектура%20__%20Лекция%20№3%20Распараллеливание%20в%20ВС%20на%20уровне%20исполнительных%20устройств2.mht!http://www.intuit.ru/img/empty.gif |
|  |
| **7. РАСПАРАЛЛЕЛИВАНИЕ В ВС НА УРОВНЕ ИСПОЛНИТЕЛЬНЫХ УСТРОЙСТВ** |

Очевидно, что одновременно или с перекрытием во времени могут выполняться такие две команды, каждая из которых не использует в качестве операнда результат выполнения другой. Если же команда меняет значение некоторой величины, то предшествующие команды в «окне просмотра» должны раньше выполнения данной команды успеть использовать предыдущее значение. При этом должна соблюдаться последовательность присваивания значений одной величине.

Пусть α*j* — логическая переменная (признак), соответствующая возможности назначения *j*-й команды для выполнения. Тогда в процессе назначения команд на ИУ *j*-я команда может быть назначена для выполнения в соответствии со значением истинности предикатов.

При наличии незанятых ИУ необходимой специализации или при наличии свободных регистров в буферах этих ИУ, команды с признаком α*j*=1 назначаются для выполнения. Так, в АЛУ процессора МВК «Эльбрус-2» каждое специализированное конвейерное ИУ имеет небольшой буфер, в которые записываются назначенные команды.

Назначенные команды снабжаются признаком назначения β*j* и не исключаются из «окна просмотра» до сигнала ИУ о завершении выполнения. Это необходимо для исключения преждевременного начала выполнения других команд, использующих результат данной.

Легко видеть, что в каждой из одновременно выполняющихся команд, на разных ИУ, в данном случае используются адреса СОЗУ, которых нет ни в одной из других команд. Т. е. одновременно работающие ИУ используют непересекающиеся множества адресов. Это исключает конфликты обращения к регистрам СОЗУ, ускоряет работу. В таком случае говорят, что ИУ *полностью сегментированы*. Хотя в общем случае работы на общих вычислительных ресурсах (решающих полях) это необязательно.

**Пример.** Пусть «окно просмотра» вместило в себя весь фрагмент ранее рассмотренной программы. АЛУ содержит необходимые устройства сложения, умножения и деления, а также два одновременно работающих канала обращения к ОП. Последовательное преобразование содержимого «окна просмотра» показано на [рис. 7.](http://www.intuit.ru/department/hardware/paralltech/3/4.html#image.3.9)1.

Рис. 7.1.  Распределение работ между исполнительными устройствами

+

Сч

Сч

×

:

3n

r1

<*m*>

<v>

r4

r3

r1

r2

r5

r6

r3

r4

r5

r6

r1

<*A*>

1

1

1

0

0

0

1

1

1

0

0

0

θ*j*

Сч

Сч

+

Сч

Сч

×

:

3n

A*j*1

<*a*>

<*b*>

r1

<m>

<*n*>

r4

r3

r1

A*j*2

r2

r5

r6

A*j*3

r1

r2

r3

r4

r5

r6

rl

3n

 α

1

1

0

1

1

0

0

0

β

1

1

0

0

0

0

0

0

x

:

3n

r4

r3

r1

r5

r6

r6

r1

<*A*>

1

0

0

1

0

0

:

3n

r3

r1

r6

Q

r1

<*A*>

1

0

1

0

Данный способ распределения команд, потребовавший промежуточного перевода безадресных команд в трехадресные, эффективен в случае многофункционального АЛУ. Универсальные исполнительные устройства, образующие АЛУ, — решающие поля, могут потребовать разработки других способов распределения работ.

Рассматривается проблема параллельного выполнения программы, которая представлена в безадресной системе команд, имитирующей выполнение операций на стеке. Параллельная обработка стека применима к программам счёта арифметических операторов, содержащих условия. Это позволяет значительно сократить количество условных переходов.

**8. ПАРАЛЛЕЛЬНАЯ ОБРАБОТКА СТЕКА И СТАТИЧЕСКОЕ РАСПАРАЛЛЕЛИВАНИЕ В РЕШАЮЩЕМ ПОЛЕ**

**8.1. Подстеки и их взаимодействие**

В процессорах супер-ЭВМ используются многофункциональные АЛУ, состоящие из специализированных по операциям исполнительных устройств. Широкое применение микропроцессоров позволяет реализовать в составе АЛУ *решающие поля* на основе универсальных исполнительных устройств, которые могут выполнять последовательности команд или целые процедуры. Рассмотрим возможность динамического распараллеливания в таких АЛУ при выполнении арифметических операторов программы в безадресной системе команд процессора, воспроизводящей выполнение работ на стеке.

Существуют пути обобщения такой структуры ВС на основе комплектации многопроцессорных ВС на общем вычислительном ресурсе, - решающем поле. Такая структура в наибольшей степени адекватна концепции двух основных уровней распараллеливания: уровня программ и уровня команд.

Процессор в обычном смысле выполняет функции устройства управления - управления выполнением программы. Исполнительные команды выполняются процессорными элементами (ПЭ) решающего поля РП, которые составляют общий вычислительный ресурс ВС. Т. е. РП выполняет заявки процессоров на выполнение операций. При этом порядок использования ПЭ определяется динамически в соответствии с потоком команд программы и с возможностями их параллельного выполнения.

Однако при такой структуре и универсальности ПЭ целесообразно распределять между ними не отдельные команды, а группы команд или законченные операторы. И принимать решение об этом можно на основе анализа бесскобочной записи (польской инверсной записи — ПОЛИЗ) арифметических операторов программы, или, что то же самое, — на основе потока безадресных команд программы. Выясняется, что производить распределение (распараллеливание) можно, минуя промежуточный перевод команд из *безадресной формы* в трехадресную, — непосредственно по тексту программы.

Тогда, если представить, что программа в ПОЛИЗ предполагает ее непосредственное выполнение на стеке, то, следовательно, производится распараллеливание обработки стека. Работа одного стека сводится к параллельной работе нескольких взаимосвязанных *подстеков* данного стека. Каждый подстек реализуется на стеке выделенного для этого ПЭ. Если предположить, что данные находятся в СОЗУ и не подлежат перемещению при организации совместной работы ПЭ, то для их обработки целесообразно организовать в локальной регистровой памяти каждого ПЭ адресные стеки, как это рассматривалось выше.

Процессорный элемент и его адресный стек со своим окружением показан на [рис 8.1](http://www.intuit.ru/department/hardware/paralltech/4/#image.4.1).

Рассмотрим арифметическое выражение

*A* :=*a* - (*b* × (*c* + *d*) - *e* : *f*) : (*g* × *h* × *i*)

Его бесскобочная запись

*Aabcd* + × *ef* : - *gh* × *i* × : - :=.

Указатель вершины стека

Ссылка налевый (или два левых) подстек

Цепочка (список) операций

Адресный стек

Рис. 8.1. Подстек

Запишем программу, произведя очевидное оптимизирующее преобразование, сокращающее количество цепочек имен и операций,

*abcd* + × *ef* : - *ghi* × × : - Зп*А*.

Составим информационный граф G, соответствующий порядку выполнения операций на стеке при счете значения этого выражения ([рис. 8.2](http://www.intuit.ru/department/hardware/paralltech/4/#image.4.2)).

Рис. 8.2.  Граф-схема счёта арифметического выражения

Строить этот граф будем в порядке выполнения операций. Сначала изобразим вершины *a,b,c,d* в соответствии с вызовом их в стек. Затем изобразим вершину +, соответствующую сложению *c* и *d*. Затем — вершину ×, соответствующую умножению *b* на *c*+*d*. Так как цепочка операций закончилась, изобразим вершины *e* и *f* и т.д. В результате последовательных действий развернется граф G, иллюстрирующий параллельную структуру алгоритма счета значения данного выражения. Выделены подструктуры, которые могут выполняться параллельно.

Важный вывод на основе анализа графа: можно независимо (и параллельно) извлекать из памяти величины по всем именам, составляющим цепочки имен в ПОЛИЗ. Т. е. можно их заблаговременно считывать в любой последовательности, в том числе и параллельно. Однако при возможности последовательного обращения к памяти предпочтительным является анализ цепочек имен слева направо, а каждой цепочки — справа налево, т. к. использование величин в операциях производится справа налево. (В нижеследующих примерах загрузка стеков производилась традиционно.)

Тогда разовьем стековый механизм, позволяющий производить распараллеливание вычислений.

По числу цепочек имен (и операций) в ПОЛИЗ сформируем несколько подстеков, каждый из которых имеет свою вершину. Подстеки, сформированные для счета одного выражения, упорядочены так, что каждый из них может обладать *ссылкой*на свой *правый*подстек (кроме последнего). Левый по отношению к данному подстек образуется цепочкой имен, находящейся в бесскобочной записи непосредственно левее цепочки имен, соответствующей данному подстеку. Каждому подстеку ставится в соответствие следующая за цепочкой имен цепочка операций и разрешается ее выполнение. На [рис 8.3, *а*](http://www.intuit.ru/department/hardware/paralltech/4/#image.4.3) показаны сформированные подстеки и их начальная загрузка (не по стековому принципу — последовательно через вершину, а по возможности параллельно или с перекрытием во времени).

Рис. 8.3.  Взаимодействие подстеков

*d*

*c*

*b*

*a*

Ссылка 2

*f*

*e*

Ссылка 3

*i*

*h*

*g*

Ссылки нет

*c*+*d*

*b*

*a*

Ссылка 2

*e*:*f*

Ссылка 3

*h*×*i*

*g*

Ссылки нет

*b*(*c*+*d*)

*a*

Ссылка 2

*e*:*f*

Ссылка 3

*g*×*h*×*i*

Ссылки нет

*b*(*c*+*d*)-*e*:*f*

*a*

Ссылка 3

*g*×*h*×*i*

Ссылки нет

*a*

*б*

*в*

*г*

1

2

3

1

2

3

1

2

3

1

3

+×

:-

× ×:-

×:-

-

×

:-

Определим возможный вариант реализации: подстеки могут заполняться не соответствующими величинами, а их адресами в СОЗУ. Т. е. традиционные приемы использования КЭШ-памяти должны быть применены здесь.

Взаимосвязь подстеков следует из неравенства длин цепочек имен и соответствующих им цепочек операций. Возможны ситуации:

1. *подстек выродился* в свою вершину, но цепочка операций оказалась не исчерпанной при том, что первая из этих операций - двуместная; такой подстек назовем *не полностью вырожденным;*
2. цепочка операций исчерпалась, но на подстеке содержится одна или более величин (или их имен, адресов), — такой подстек назовем *вырожденным;* если на подстеке находится единственная величина, составляющая вершину, то такой стек назовем *полностью вырожденным.*

Если данный подстек не полностью вырожденный, то последующая операция возможна над вершиной этого подстека и вершиной левого подстека, если тот является вырожденным или полностью вырожденным.

Если левый подстек также не полностью вырожденный или выполнение на нем возможных операций не закончилось, то данный подстек находится в состоянии ожидания. Для выполнения операции над вершинами двух подстеков (если правый подстек является вырожденным) правый подстек переводится наданныйподстек так, что его вершина становится новой вершиной данного левого подстека. Ссылка правого подстека заменяет собой ссылку левого подстека. Таким образом, на основе двух подстеков формируется один и появляется возможность дальнейшего выполнения операций из цепочки.

При правильной бесскобочной записи каждый не полностью вырожденный подстек дождется возможности продолжения выполнения операций за счет вырожденного или полностью вырожденного левого подстека. Он, в свою очередь, превратится в вырожденный или полностью вырожденный и пополнит левый не полностью вырожденный, если такой имеется. Окончательное значение данного выражения образуется в вершине единственного оставшегося полностью вырожденного подстека, не содержащего ссылку на правый подстек.

Продолжим рассмотрение примера, для простоты считая одинаковым время выполнения всех операций.

На первом шаге (в первом такте) выполняются операции на всех подстеках ([рис. 8.3, *б*](http://www.intuit.ru/department/hardware/paralltech/4/#image.4.3)). После этого подстек 2 оказывается не полностью вырожденным; в дальнейшем он будет ждать вырождения левого — подстека 1.

После выполнения операций на втором шаге ([рис. 8.3, *в*](http://www.intuit.ru/department/hardware/paralltech/4/#image.4.3)) подстек 1 оказывается вырожденным, его дополняет подстек 2 со своей ссылкой на третий подстек, который становится правым для первого. Подстек 3 оказывается не полностью вырожденным, он ждет вырождения своего левого подстека.

После выполнения операций на третьем шаге ([рис. 4.3, *г*](http://www.intuit.ru/department/hardware/paralltech/4/#image.4.3)) подстек 1 оказывается вырожденным; он дополняется подстеком 3.

После выполнения операций на подстеке 1 на четвертом и пятом шагах получается окончательный результат.

При произвольных временах выполнения операций динамическая картина меняется, но правила взаимодействия подстеков те же.

Таким образом, все подстеки образуют очередь заданий, которые распределяются между операционными или исполнительными универсальными устройствами — ПЭ для выполнения. ПЭ действительно образуют распределяемый ресурс системы, в общем случае — многопроцессорной, и назначаются динамически по необходимости. Реализуется концепция виртуальных исполнительных устройств или процессорных элементов решающего поля.

Усложним пример, введя в рассмотрение арифметические операторы, содержащие условие, а также образующие сложные конструкции с использованием условий.

Пусть необходимо распараллелить счет арифметического выражения

Y:=(*a*+*e*:*f*)× if *a*×lmhtml:file://F:\INTUIT_ru%20Курс%20Архитектура%20__%20Лекция%20№4%20Параллельная%20обработка%20стека%20и%20статическое%20распараллеливание%20в%20решающем%20поле1.mht!http://www.intuit.ru/img/symbols/le.gif*c* then if *l* <*g*×*h* then A else A+*l* else *c*×(*h*+16).

Предварительно необходимо распространить правила формирования ПОЛИЗ для отображения условий. Целесообразно оставлять на месте разграничителиif, then, else. Знак операции, в которой участвует условное выражение, необходимо предпосылать каждому альтернативному оператору.

Процессор, как указывалось ранее, выполняя функции лишь устройства управления, производит последовательный анализ символов программы. Обнаружив цепочку имен, он выделяет ПЭ и настраивает его на обработку этой цепочки и последующей цепочки операций. При нахождении каждой следующей цепочки имен, если анализ производится вне условного оператора, он формирует ссылку «левого» подстека на свой «правый». Однако после формирования последнего (а возможно — единственного) подстека оператора if он формирует ему ссылку на два подстека: первого подстека оператора then и первого подстека оператора else — после их анализа.

Учебное издание

**Кобайло** Александр Серафимович

**Архитектура компьютера**

Конспект лекций

Редактор Ю. Д. Нежикова

Компьтерная верстка

Корректор Ю. Д. Нежикова

Издатель:

УО «Белорусский государственный технологический университет».

Свидетельство о государственной регистрации издателя, изготовителя, распространителя печатных изданий

№1/227 от 20.03.2014

Ул. Свердлова, 13а, 220006, г.Минск