

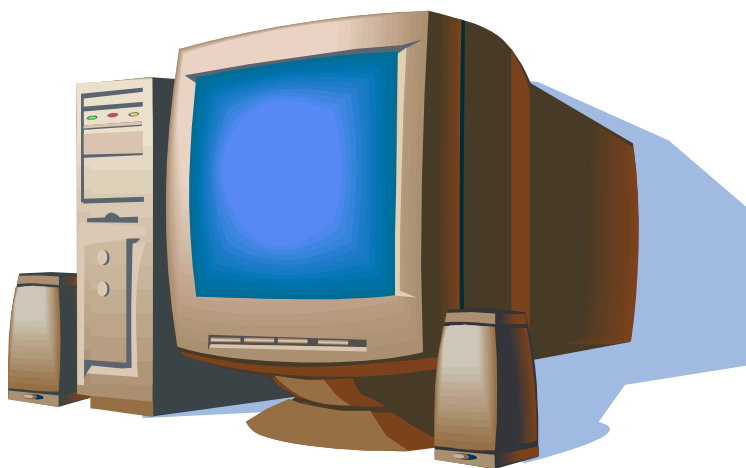


*Томский межвузовский центр
дистанционного образования*

Б.В. Илюхин

АППАРАТНЫЕ СРЕДСТВА И СЕТИ ЭВМ

Учебное пособие



ТОМСК – 2005

Рецензент: кандидат технических наук Пушкарев В.П.

Корректор: Осипова Е.А.

Илюхин Б.В.

Аппаратные средства и сети ЭВМ: Учебное пособие. — Томск: Томский межвузовский центр дистанционного образования, 2005. — 226 с.

Изложены основы аппаратного обеспечения персональных ЭВМ. Рассмотрены вопросы взаимодействия вычислительных систем с окружающей средой, принципы построения и функционирования различных компонентов персональных электронно-вычислительных машин. Показано единство подходов к проектированию вычислительных систем и сетей, исследованы взаимные связи между ними.

© Илюхин Б.В., 2005
© Томский межвузовский центр
дистанционного образования, 2005

ОГЛАВЛЕНИЕ

Введение.....	8
1 АППАРАТНЫЕ СРЕДСТВА ЭВМ	10
1.1 Общие сведения о работе вычислительных устройств	10
1.2 Важные определения	12
1.3 Микропроцессоры	13
1.3.1 Архитектура и команды микропроцессора	13
1.3.2 Семейство процессоров INTEL	15
1.3.3 Процессоры с умножением частоты	20
1.3.4 Процессоры-клоны	23
1.3.5 Теплообмен	24
1.3.6 Сопроцессоры	25
1.3.7 Что означает термин «сверхмасштабный»	26
1.3.8 Тенденции развития микропроцессоров семейств INTEL-AMD	27
1.4 Память.....	30
1.4.1 Классификация оперативной памяти	31
1.4.2 Корпуса и маркировка	33
1.4.3 Режимы работы памяти	36
1.4.4 Новые типы динамической памяти	38
1.4.5 Видеопамять	42
1.4.6 Кэш-память	42
1.5 BIOS и CMOS RAM.....	45
1.6 Вспомогательные микросхемы	49
1.6.1 Тактовый генератор	49
1.6.2 Контроллер прерываний	50
1.6.3 Контроллер прямого доступа в память	51
1.6.4 Другие вспомогательные микросхемы	52
1.6.5 Асинхронные и синхронные компьютеры	52
1.6.6 Часы внутри CPU	53
1.7 Набор микросхем, или chipset	54
1.8 Системные и локальные шины	55
1.8.1 Шина ISA	56
1.8.2 Шина EISA	56
1.8.3 Локальные шины	58
1.8.4 Шины VL-bus и PCI	59

1.8.5 Шина PCI как ось «Север — Юг» в PC	61
1.8.6 Шина L2: обходной путь для повышения производительности центрального процессора	64
1.8.7 Последовательные шины (USB и IEEE 1394)	66
1.8.8 Стандарт PCMCIA	68
1.9 Выбор системной платы	70
1.10 Накопители	71
1.10.1 Винчестеры	73
1.10.2 Как устроен винчестер	74
1.10.3 Методы хранения информации	76
1.10.4 Важные параметры	78
1.11 Интерфейсы накопителей	78
1.11.1 Интерфейс ST-506/412	79
1.11.2 Интерфейс ESDI	81
1.11.3 Интерфейс SCSI	81
1.11.4 Интерфейс ATA	88
1.11.5 Интерфейс Enhanced IDE и другие	91
1.11.6 Кэширование диска	93
1.11.7 Выбор интерфейса винчестера	93
1.12 Другие виды накопителей	95
1.12.1 Флоппи-диски	95
1.12.2 Стримеры	98
1.12.3 Накопители, использующие эффект Бернулли	98
1.12.4 Накопители на компакт-дисках	99
1.12.5 Магнитооптические накопители	101
1.13 Видеоподсистемы	103
1.13.1 Принцип работы мониторов	104
1.13.2 Важные параметры	105
1.13.3 «Зеленые» мониторы	107
1.13.4 Плоские экраны	109
1.13.5 Жидкокристаллические экраны	109
1.14 Видеоадаптеры	112
1.14.1 Видеоадаптеры SVGA	116
1.14.2 Проблемы цветопередачи	118
1.14.3 Графические и мультимедиа-акселераторы	121
1.14.4 Видеокабели	123
1.14.5 Видеокарты	125
1.14.6 TV-тюнеры	125

1.14.7 Фрейм-грабберы.....	125
1.14.8 Преобразователи VGA-TV	126
1.14.9 MPEG-плейеры.....	127
1.15 Порты ввода-вывода	127
1.15.1 Параллельный порт.....	128
1.15.2 Последовательный порт	129
1.15.3 USB (Universal Serial Bus)	131
1.15.4 Инфракрасный порт.....	132
1.15.5 Игровой порт	133
1.16 Устройства ввода	133
1.16.1 Клавиатуры.....	133
1.16.2 Мыши и трекболы.....	136
1.16.3 Сканеры.....	139
1.16.4 Джойстики	139
1.17 Устройства вывода	139
1.17.1 Принтеры	139
1.17.2 Классификация принтеров	140
1.17.3 Матричные печатающие устройства.....	141
1.17.4 Струйные принтеры.....	143
1.17.5 Лазерные и LEty-принтеры	144
1.17.6 Цветная печать	146
1.17.7 Модемы.....	147
1.17.8 Факс-модемы.....	149
1.17.9 Звуковые карты	150
1.17.10 Акустические системы.....	152
2 ФИЗИЧЕСКОЕ ОКРУЖЕНИЕ ВЫЧИСЛИТЕЛЬНОЙ	
ТЕХНИКИ.....	154
2.1 Климатические условия	154
2.2 Проблемы электропитания	155
2.2.1 Классификация устройств защиты.....	157
2.2.2 Выбор подходящего ИБП	162
2.3 Электромагнитные помехи	163
2.4 Электростатические разряды (Electro Static Discharge)	164
3 КОМПЬЮТЕРНЫЕ СЕТИ.....	166
3.1 Понятие компьютерной сети	166
3.1.1 Локальные вычислительные сети.....	167
3.1.2 Расширение компьютерных сетей.....	167

3.1.3	Приложения.....	168
3.1.4	Одноранговые сети.....	170
3.1.5	Сети на основе сервера.....	170
3.1.6	Специализированные серверы.....	170
3.1.7	Комбинированные сети.....	172
3.2	Топология сети.....	172
3.2.1	Базовые топологии.....	173
3.2.2	Шина.....	174
3.2.3	Звезда.....	176
3.2.4	Кольцо.....	176
3.2.5	Комбинированные топологии.....	177
3.2.6	Выбор топологии.....	179
3.3	Кабельные каналы связи.....	179
3.3.1	Основные группы кабелей.....	179
3.3.2	Коаксиальный кабель.....	180
3.3.3	Витая пара.....	184
3.3.4	Оптоволоконный кабель.....	188
3.4	Беспроводная среда передачи данных.....	189
3.4.1	Типы беспроводных сетей.....	190
3.4.2	Инфракрасные беспроводные сети.....	191
3.4.3	Лазерные технологии организации сетей.....	192
3.4.4	Радиопередача в узком диапазоне (одночастотная передача).....	192
3.4.5	Радиопередача в рассеянном спектре.....	192
3.4.6	Передача «точка-точка».....	193
3.4.7	Многоточечное беспроводное соединение.....	193
3.5	Назначение платы сетевого адаптера.....	194
3.5.1	Подготовка данных.....	194
3.5.2	Передача и управление данными.....	196
3.6	Работа сети.....	196
3.6.1	Эталонная модель взаимодействия открытых систем.....	197
3.6.2	Понятие открытой системы.....	198
3.6.3	Процессы.....	199
3.6.4	Модель OSI.....	201
3.6.5	Взаимодействие уровней модели OSI.....	202
3.6.6	Услуги различных уровней модели OSI.....	208
3.7	Доступ абонентских систем к моноканалу.....	209

3.7.1 Методы доступа в сетях с шинной топологией	212
3.7.2 Методы доступа в кольцевых сетях	216
3.7.3 Модель IEEE Project 802	219
3.7.4 Категории стандартов ЛВС.....	221
3.7.5 Расширения модели OSI.....	223
Заключение	225
Список литературы.....	226

ВВЕДЕНИЕ

Из этого пособия вы сможете узнать об основных узлах и компонентах IBM PC-совместимого компьютера, о большинстве периферийных устройств, которые могут к нему подключаться, о факторах окружающей среды, влияющих на работу вычислительной техники, а также об основах построения компьютерных сетей. Изложение материала построено исходя из того, что «знание некоторых принципов заменяет знание многих фактов». Вот почему основной упор сделан на описание принципов функционирования и используемых технологий. Действительно, стремительные изменения в элементной базе и схемотехнике в ряде случаев исключают возможность рассказа о конкретных технических решениях. Ведь хорошо известно, что, например, новые варианты системных плат появляются почти через каждые два месяца. Именно поэтому основной упор сделан на описание принципов функционирования и используемых технологий.

В этой книге речь идет об аппаратном обеспечении IBM PC, а точнее, об IBM PC-совместимых (и не совсем совместимых) компьютерах. Можно сказать, что в 1981 году появление первого IBM PC (Personal Computer) произвело настоящую информационную революцию.

Безусловно, понятие ЭВМ охватывает гораздо более широкий круг устройств, нежели персональные компьютеры. Но в данном пособии мы будем говорить именно о персональных ЭВМ, более того, о персональных IBM PC-совместимых компьютерах, поскольку они являются на сегодняшний день наиболее распространенными.

IBM PC не был первым персональным компьютером (даже у IBM), однако по сравнению с уже существующими имел ряд существенных преимуществ. По сути, впервые была предложена действительно гибкая и легко расширяемая система, которая могла использовать аппаратные и программные дополнения других фирм-производителей. До появления IBM PC все подобные компьютеры для пользователя оставались «черным ящиком»: их модернизация оставалась уделом профессионалов, а самим пользователям приходилось довольствоваться только тем, что они приобретали. Принцип открытой (модульной) архитектуры,

заложенный в IBM PC, позволял легко менять конфигурацию, модернизировать блоки и узлы. Удачным в IBM PC было и конструктивное решение. Почти за 25 лет, естественно, многое изменилось, однако так или иначе миллионы пользователей во всем мире продолжают работать именно с IBM PC-совместимыми компьютерами.

1 АППАРАТНЫЕ СРЕДСТВА ЭВМ

1.1 Общие сведения о работе вычислительных устройств

Важнейший элемент любого компьютера — процессор. Данный элемент в большей степени определяет возможности вычислительной системы и, образно выражаясь, является ее сердцем. Вовсе не обязательно он представляет из себя одну микросхему. В больших вычислительных системах процессор (процессоры) могут занимать пространство размером не с один шкаф. Например, процессор «Эльбрус», разработанный в Зеленограде и признанный на сегодняшний день совершенством (с точки зрения архитектуры, то есть внутреннего строения), — представляет собой систему довольно больших геометрических размеров. Мы же будем вести речь о микропроцессорах, то есть об устройствах, которые используются, в основном, в персональных компьютерах. Микропроцессор, как правило, представляет собой сверхбольшую интегральную схему, реализованную в одном полупроводниковом кристалле и способную выполнять функции центрального процессора. Степень интеграции определяется размером кристалла и количеством реализованных в нем транзисторов. Зачастую интегральные микросхемы называют чипами (chips).

К обязательным компонентам микропроцессора относятся арифметико-логическое (исполнительное) устройство и блок управления. Они характеризуются скоростью (тактовой частотой), разрядностью или длиной слова (внутренней и внешней), архитектурой и набором команд. Архитектура процессора определяет необходимые регистры, стеки, систему адресации а также систему команд и типы обрабатываемых процессором данных. Обычно используются следующие типы данных: бит (один разряд), полубайт (4 бита), байт (8 бит), слово (16 бит), двойное слово (32 бита). Выполняемые процессором команды предусматривают, как правило, арифметические действия, логические операции, передачу управления (условную и безусловную) и перемещение данных (между регистрами, памятью и портами ввода-вывода).

Под конвейерным режимом понимают такой вид обработки,

при которой интервал времени, необходимый для выполнения процесса в функциональном узле (например, в арифметико-логическом устройстве) микропроцессора, продолжительнее, чем интервалы, через которые данные могут вводиться в этот узел. Предполагается, что функциональный узел выполняет процесс в несколько этапов, то есть когда первый этап завершается, результаты передаются на второй этап, на котором используются другие аппаратные средства. Разумеется, что устройство, используемое на первом этапе, оказывается свободным для начала новой обработки данных. Можно выделить четыре конкретных этапа обработки команды микропроцессора: выборка, декодирование, выполнение и запись результата. Иными словами, в ряде случаев пока команда выполняется, вторая может декодироваться, а третья выбираться.

С внешними устройствами микропроцессор может общаться «благодаря» шинам адреса, данных и управления, выведенным на специальные контакты корпуса микросхемы. Стоит отметить, что разрядность внутренних регистров микропроцессора может не совпадать с количеством внешних выводов для линий данных. Иначе говоря, микропроцессор с 32-разрядными регистрами может иметь только 16 внешних линий данных. Объем физически адресуемой процессором памяти однозначно определяется разрядностью внешней шины адреса как 2^N , где N — количество адресных линий.

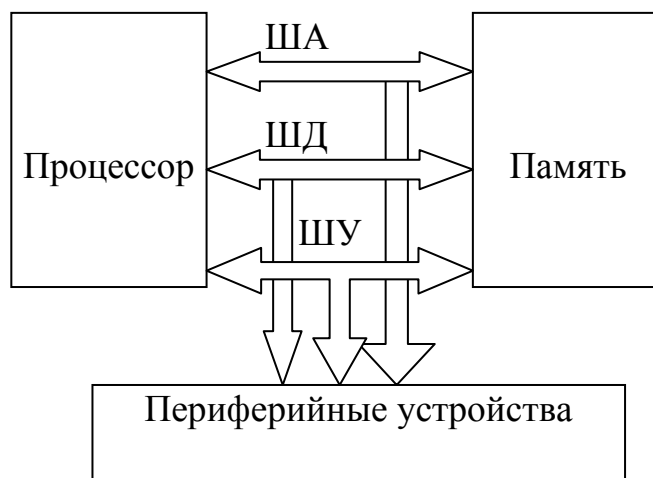


Рис. 1 — Структурная схема вычислительного устройства

Структурная схема устройства, выполняющего, например, операции сложения будет выглядеть, как показано на рисунке 1, и состоять из процессора, памяти, шин адреса, данных и управления. При этом возникает необходимость ввода информации и вывода результатов. Это достигается использованием внешних устройств для ввода-вывода информации (например, клавиатуры, монитора).

1.2 Важные определения

Прежде чем далее говорить о микропроцессорах, необходимо запомнить некоторые важные определения, которые пригодятся в дальнейшем:

Периферийным устройством называется любое внешнее устройство, совершающее по отношению к микропроцессору операции ввода-вывода.

Регистр — совокупность бистабильных устройств (то есть имеющих два устойчивых состояния), предназначенных для хранения информации и быстрого доступа к ней. В качестве таких устройств в интегральных схемах используют триггеры. Триггер в свою очередь выполнен на транзисторных переключателях (электронных ключах). В регистре из N триггеров можно запомнить слово из N бит информации.

Порт — это некая схема сопряжения, обычно включающая в себя один или несколько регистров ввода-вывода и позволяющая подключить, например, периферийное устройство к внешним шинам микропроцессора. Практически каждая микросхема использует для различных целей несколько портов ввода-вывода. Каждый порт в персональном компьютере имеет свой уникальный номер. Номер порта — это, по сути, адрес регистра ввода-вывода, причем адресные пространства основной памяти и портов ввода-вывода не пересекаются.

Прерывание — сигнал, по которому процессор узнает о совершении некоего асинхронного события. При этом исполнение текущей последовательности команд приостанавливается (прерывается), а вместо нее начинает выполняться другая последовательность, соответствующая данному прерыванию. Прерывания можно классифицировать как:

- *аппаратные,*
- *логические,*
- *программные.*

Аппаратные прерывания обычно связаны с запросами от периферийных устройств (например, нажатие клавиши на клавиатуре), логические возникают при работе самого микропроцессора (например, деление на ноль), а программные инициируются выполняемой программой и используются для вызова специальных подпрограмм. Кроме того, прерывания могут быть:

маскируемыми — при определенных условиях (например, запрете на определенные прерывания) микропроцессор не обращает на них внимания;

немаскируемыми — не игнорируемыми ни при каких условиях. В этом случае, как правило, должны обрабатываться почти катастрофические события (падение напряжения питания или ошибка памяти).

В режиме прямого доступа (*DMA, Direct Memory Access*) периферийное устройство связано с оперативной памятью непосредственно, минуя внутренние регистры микропроцессора. Наиболее эффективна такая передача данных в ситуациях, когда требуется высокая скорость обмена при передаче большого количества информации (например, при загрузке данных в память с внешнего накопителя).

Довольно часто для адресов, номеров портов, прерываний и т.п. используется шестнадцатеричная система счисления. В этом случае после соответствующего числа стоит буква «h» (hexadecimal).

1.3 Микропроцессоры

1.3.1 Архитектура и команды микропроцессора

Исторически сложилось так, что набор инструкций, «понимаемых» (т.е. выполняемых) процессорами очень велик. Всего существует более 100 различных типов инструкций, причем многие из них принимают модификаторы, по существу превращающие один тип инструкций в целое семейство типов. В общем, процес-

сор готов исполнить несколько сотен приказов программиста. Некоторые из них чрезвычайно просты, другие изрядно сложны. (Но не сложнее даже очевидной фразы на практически любом умеренно сложном человеческом языке.)

Длина инструкций процессоров, например x86, варьируется от одного — до десятка байтов. Короткими инструкциями кодируются простые действия. Сложные — длинными, причем некоторые могут достигать со всеми модификаторами до 20 байт длины.

Все эти характеристики типичны для так называемого *CISC-компьютера* (Complex Instruction Set Computer, CISC — компьютер со сложным набором инструкций). С каждым новым поколением процессоров x86 CISC-компьютеры становятся все сложнее. Забавно, но одновременно они все больше напоминают *RISC-машины* (Reduced Instruction Set Computer, RISC — компьютер с уменьшенным набором инструкций). RISC-машины характеризуются относительно скудным набором инструкций одинаковой длины.

В научных кругах уже много лет ведутся дискуссии о сравнительных достоинствах подходов CISC и RISC к проектированию. Проблема выбора уступила место идее объединения двух подходов.

По мнению адвокатов RISC, хотя такой тип машин и не может выполнить много действий, реализуя одну инструкцию, зато каждая инструкция быстрее декодируется в набор операций для логического устройства процессора в силу их единого размера. В среднем, для выполнения RISC-инструкции требуется меньше тактов, чем для выполнения средней CISC-инструкции. Более того, обычно RISC-машинам под силу исполнение сразу нескольких инструкций за такт а на CISC-машинах есть инструкции, для выполнения требующие более 100 тактов.

На эти выпады защитники CISC спокойно отвечают, что любимый ими подход имеет перед RISC два больших преимущества. Во-первых, заявляют они, программировать легче и естественнее под CISC-процессор, потому что он «знает» больше операций, элементарных с точки зрения программиста. А во-вторых, грамотно сделанный CISC-процессор может работать даже быстрее соответствующей RISC-машины, потому что сложные инструкции делают программы короче, а значит, и быстрее. Банально, но истина опять где-то посередине. Всего 10 или 20 лет назад все

высокопроизводительные рабочие станции, используемые инженерами, программистами, компьютерными художниками, были построены на базе RISC-процессора (или процессоров). Теперь же уровня RISC-станции можно достичь, купив CISC-персональный компьютер.

Просто создатели CISC-процессоров, наконец, позаимствовали некоторые идеи RISC-подхода. В некоторых случаях, примером чему AMD, процессор научился разбивать сложные инструкции x86, которые программист считает неделимыми, на множество меньших RISC-инструкций, и уже они выполняются логическим устройством процессора. По сути дела, имеет место эмуляция работы CISC-процессора RISC-процессором.

Это, в свою очередь очень неплохо описывает работу процессоров Intel Pentium Pro, Pentium II и последующих. Теперь даже Intel, до этого главный сторонник CISC-подхода, поумнела и на микропрограммном уровне использует RISC подход, создавая уже RISC-, но на программном уровне выглядящие как очень быстрые CISC-, машины.

Более простой пример заимствования — конвейерная обработка и продуманное исполнение, разработанные первоначально именно для RISC-машин.

Но нашлись и те, которые не спорили, а пошли в третьем направлении, названном *YLIW-обработкой* (Very Long Computer Word, VLIW — очень длинное машинное слово). Основная идея здесь состоит в том, чтобы поместить несколько относительно простых инструкций в одну длинную, а затем построить процессор, выполняющий их все за один раз. Если все делать правильно, то может получиться нечто, имеющее лучшие черты и RISC-, и CISC-подхода. Сейчас Hewlett-Packard и Intel сотрудничают в этой области.

1.3.2 Семейство процессоров INTEL

Первый микропроцессор — I4004 — был изготовлен в 1971 году, и с тех пор фирма Intel (INTEgrated ELectronics) прочно удерживает лидирующее положение на данном сегменте рынка. Стоит, пожалуй, напомнить, что максимальная тактовая частота этого прапрадедушки современных «числодробилок» составляла

всего 750 кГц.

Реализация ряда следующих проектов фирмы Intel по разработке однокристальных микропроцессоров (i4040, i8008) возвестила о наступлении эры персональных компьютеров. Наиболее успешным был, пожалуй, проект разработки микропроцессора i8080. Именно на этом микропроцессоре был основан компьютер «Альтаир», для которого молодой Билл Гейтс написал свой первый интерпретатор Бейсика. Он был выполнен по n-канальной МОП-технологии (n-MOS), а его тактовая частота не превышала 2 МГц. Не будет преувеличением сказать, что классическая архитектура i8080 оказала огромное влияние на дальнейшее развитие однокристальных микропроцессоров. Несмотря на заслуженный успех i8080, настоящим промышленным стандартом для персональных компьютеров стал другой микропроцессор фирмы Intel.

i8088

Микропроцессор i8088 был анонсирован Intel в июне 1979 года, а в 1981-м «Голубой Гигант», как иногда называют IBM, выбрал этот микропроцессор для своего первого персонального компьютера и, надо сказать, не ошибся. Новый чип содержал примерно 29 тысяч транзисторов. Одним из существенных достоинств микропроцессора i8088 была возможность (благодаря 20-адресным линиям) физически адресовать область памяти в 1 Мбайт. Здесь следует, правда, отметить, что для IBM PC в этом пространстве программам было отведено всего лишь 640 Кбайт. Хотя с внешними периферийными устройствами (дисками, видео) i8088 был связан через свою внешнюю 8-разрядную шину данных, его внутренняя структура (адресуемые регистры) позволяла работать с 16-разрядными словами. Как известно, на системной шине IBM PC для передачи данных было отведено 8 линий (1 байт). Первоначально микропроцессор i8088 работал на частоте 4,77 МГц и имел быстродействие около 0,33 MIPS (Million Instruction Per Second), однако впоследствии были разработаны его модификации, рассчитанные на более высокую тактовую частоту (например, 8 МГц).

Но ровно на год раньше своего счастливого последователя появился процессор 8086 (в июне 1978 года), ставший популярным в основном благодаря компьютеру Compaq DeskPro. Про-

граммная модель (доступные регистры) этого микропроцессора полностью совпадала с моделью i8088. Основное отличие данных микропроцессоров состоит в различной разрядности внешней шины данных: 8 разрядов у i8088 и 16 разрядов у i8086. Понятно, что более высокой производительности с новым микропроцессором можно было достичь только при использовании компьютера, на системной шине которого под данные предусмотрено 16 линий. Адресная шина микропроцессора i8086 по-прежнему позволяла адресовать 1 Мбайт памяти.

Опираясь на архитектуру i8086 и учитывая запросы рынка, в феврале 1982 года фирма Intel выпустила свой новый микропроцессор — i80286. На кристалле было реализовано около 130 тысяч транзисторов. Надо сказать, что этот чип появился практически одновременно с новым компьютером фирмы IBM — PC/AT. Наряду с увеличением производительности этот микропроцессор (i80286) мог теперь работать в двух режимах — реальном и защищенном. Если первый режим был (за рядом исключений) похож на обычный режим работы i8088/86, то второй использовал более изощренную технику управления памятью. В частности, защищенный режим работы позволял программным продуктам, например Windows 3.0 и OS/2, работать с оперативной памятью выше 1 Мбайта. Благодаря 16 разрядам данных на новой системной шине, которая была впервые использована в IBM PC/ AT286, микропроцессор i80286 мог обмениваться с периферийными устройствами 2-байтными сообщениями. 24 адресные линии позволяли в защищенном режиме обращаться уже к 16 Мбайтам памяти. В микропроцессоре i80286 впервые на уровне микросхем были реализованы мультизадачность и управление виртуальной памятью. При тактовой частоте 8 МГц достигалась производительность 1,2 MIPS.

В октябре 1985 года фирмой Intel был анонсирован первый 32-разрядный микропроцессор i80386. Новый чип содержал примерно 275 тысяч транзисторов. Полностью 32-разрядная архитектура (32-разрядные регистры и 32-разрядная внешняя шина данных) в новом микропроцессоре была дополнена расширенным устройством управления памятью MMU (Memory Management Unit), которое помимо блока сегментации (Segmentation Unit) было дополнено блоком управления страницами (Paging Unit). Это устройство

позволяло легко переставлять сегменты из одного места памяти в другое (свопинг) и освобождать драгоценные килобайты памяти. На тактовой частоте 16 МГц быстродействие нового процессора составило примерно 6 MIPS.

В реальном режиме (после включения питания) микропроцессор i80386 работал как «быстрый 18088» (адресное пространство 1 Мбайт, 16-разрядные регистры). Защищенный режим был полностью совместим с аналогичным режимом в i80286. Тем не менее в этом же режиме i80386 мог выполнять и свои «естественные» 32-разрядные программы. Вспомним, что 32 адресные линии микропроцессора позволяют физически адресовать 4 Гбайта памяти. Кроме того, был введен новый режим — виртуального процессора (V86). В этом режиме могли одновременно выполняться несколько задач, предназначенных для i8086.

Более дешевая альтернатива 32-разрядному процессору i80386, который впоследствии получил окончание DX, появилась только в июне 1988 года. Это был микропроцессор i80386SX. В отличие от своего старшего «собрата» новый микропроцессор использовал 16-разрядную внешнюю шину данных и 24-разрядную адресную (адресуемое пространство — 16 Мбайт). Это было особенно удобно для стандартных PC/AT, системная шина которых использует, как известно, только 16 линий данных. Благодаря дешевизне нового изделия многие производители «железа» стали заменять теперь уже устаревший микропроцессор i80286 на более производительный i80386SX. Одним из решающих факторов для замены была полная совместимость 32-разрядных микропроцессоров: программное обеспечение, написанное для i80386DX, корректно работало и на i80386SX. Дело в том, что внутренние регистры их были полностью идентичны. Надо отметить, что уже к концу 1988 года микропроцессор i80386SX выпускался в количествах, существенно превосходящих рекордные показатели для i80386DX. Кстати, говорят, что индекс SX произошел от слова SiXteen (шестнадцать), поскольку разрядность внешней шины данных нового тогда процессора была именно такой. В дальнейшем, правда, для 486-х процессоров SX стал означать отсутствие математического сопроцессора.

На осенней выставке Comdex в 1989 году фирма Intel впервые анонсировала микропроцессор 486DX, который содержал

более миллиона транзисторов (а точнее, 1,2 миллиона) на одном кристалле и был полностью совместим с процессорами ряда x86. Напомним, что на кристалле первого члена этого семейства — микропроцессора i8088 — насчитывалось около 29 тысяч транзисторов. В борьбе с микропроцессорами-клонами фирма Intel намеренно убрала из названия нового устройства число 80. Новая микросхема впервые объединила на одном чипе такие устройства, как центральный процессор, математический сопроцессор и кэш-память. Использование конвейерной архитектуры, присущей RISC-процессорам, позволило достичь четырехкратной производительности обычных 32-разрядных систем. Это связано с уменьшением количества тактов для реализации каждой команды. 8-Кбайтная встроенная кэш-память ускоряет выполнение программ за счет промежуточного хранения часто используемых команд и данных. На тактовой частоте 25 МГц микропроцессор показал производительность 16,5 MIPS. Созданная в июне 1991 года версия микропроцессора с тактовой частотой 50 МГц позволила увеличить производительность еще на 50%. Встроенный математический сопроцессор существенно облегчил и ускорил математические вычисления. Однако впоследствии стало ясно, что подобный сопроцессор необходим всего лишь 30% пользователей.

Появление нового микропроцессора i486SX фирмы Intel вполне можно было считать одним из важнейших событий 1991 года. Уже предварительные испытания показали, что компьютеры на базе i486SX с тактовой частотой 20 МГц работают быстрее (примерно на 40%) компьютеров, основанных на i80386DX с тактовой частотой 33 МГц. Микропроцессор i486SX, подобно оригинальному i486DX, содержал на кристалле кэш-контроллер и кэш-память, а вот математический сопроцессор у него был заблокирован. Значительная экономия (благодаря исключению затрат на тестирование сопроцессора) позволила фирме Intel существенно снизить цены на новый микропроцессор. Надо сказать, что если микропроцессор i486DX был ориентирован на применение в сетевых серверах и рабочих станциях, то i80486SX послужил отправной точкой для создания мощных настольных компьютеров. Вообще говоря, в семействе микропроцессоров i486 предусматривалось несколько новых возможностей для построения мультипроцессорных систем: соответствующие команды поддерживают

механизм семафоров памяти, аппаратно реализованное выявление недостоверности строки кэш-памяти обеспечивает согласованность между несколькими модулями кэш-памяти и т.д. Для микропроцессоров семейства i486 допускалась адресация физической памяти размером 4 Гбайта и виртуальной памяти размером 64 Тбайта.

К концу 1991 года 32-разрядные микропроцессоры стали стандартными для компьютеров типа лэптоп и ноутбук, однако обычные микросхемы i80386DX/SX не полностью отвечали требованиям разработчиков портативных компьютеров. Для удовлетворения потребностей этого сегмента рынка в 1990 году фирмой Intel был разработан микропроцессор i80386SL, который содержал примерно 855 тысяч транзисторов. Данный микропроцессор представлял собой интегрированный вариант микропроцессора i80386SX, базовая архитектура которого была дополнена еще несколькими вспомогательными контроллерами. По существу, все компоненты, необходимые для построения портативного компьютера, сосредоточены в двух микросхемах: микропроцессоре i80386SL и периферийном контроллере i82360SL. В набор i386SL впервые было введено новое прерывание, называемое System Management Interrupt (SMI), которое использовалось для обработки событий, связанных, например, с управлением потребляемой мощностью. Вместе с математическим сопроцессором i80387SL данный набор микросхем позволял создать 32-разрядный компьютер на площади, не намного превышающей размер игровой карты.

1.3.3 Процессоры с умножением частоты

В марте 1992 года фирма Intel объявила о создании второго поколения микропроцессоров 486. Эти микропроцессоры, названные i486DX2, обеспечили новую технологию, при которой скорость работы внутренних блоков микропроцессора в два раза выше скорости остальной части системы. Тем самым появилась возможность объединения высокой производительности микропроцессора с внутренней тактовой частотой 50(66) МГц и эффективной по стоимости 25/33-мегагерцевой системой. Новые микросхемы по-прежнему включали в себя центральный процессор, математический сопроцессор и кэш-память на 8 Кбайт. Компью-

теры, построенные на базе микропроцессоров i486DX2, работали приблизительно на 70% производительнее основанных на микропроцессорах i486DX первого поколения. Несколько позже появились процессоры i486SX2, в которых, как следует из названия, отсутствует встроенный сопроцессор.

Технология умножения частоты (не только в два, но и, например, в полтора, два с половиной или три раза) находила и находит широкое применение практически во всех современных процессорах. Так, после DX2 фирма Intel выпускала серию микропроцессоров с умножением частоты в три раза — DX4 (кодовое название P24C). Процессоры этого семейства — 486DX4-75 и 486DX4-100 имели кэш-память 16 Кбайт и были предназначены для установки в системные платы, работающие на тактовой частоте 25 и 33 МГц. Напряжение питания этих процессоров составляло 3,3 В, количество транзисторов на кристалле — 1,6 миллиона.

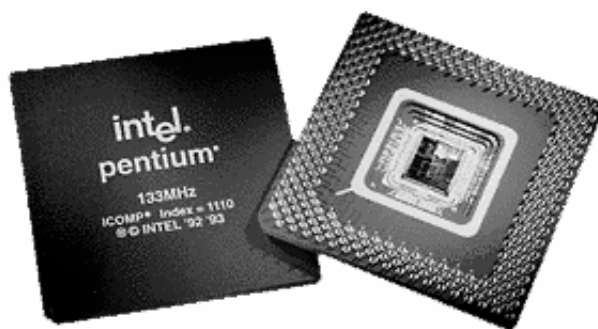


Рис. 2 — Процессор INTEL PENTIUM 133 МГц

В марте 1993 года фирма Intel объявила о начале промышленных поставок 66- и 60-Мегагерцевых версий процессора Pentium, известного ранее как 586 или P5. Название нового микропроцессора является зарегистрированной торговой маркой корпорации Intel. Таким образом, в системах Intel Inside микропроцессор 586 фигурировать перестал. Системы, построенные на базе Pentium, были полностью совместимы со 100 миллионами персональных компьютеров, использовавших микропроцессоры i8088, i80286, i80386, i486. Новая микросхема содержала около 3,1 миллиона транзисторов, имела 32-разрядную адресную и 64-разрядную внешнюю шину данных, что обеспечивало обмен дан-

ными с системной платой со скоростью до 528 Мбайт/с. В отличие от процессоров семейства 486-х, для производства которых использовалась CMOS-технология, для Pentium фирма Intel применяла 0,8-микронную BiCMOS-технология. Pentium с тактовой частотой 66 МГц имел производительность около 112 MIPS (миллионов инструкций в секунду). Суперскалярная архитектура содержала два пятиступенчатых блока исполнения, работавших независимо и обрабатывавших две инструкции за один такт синхронизации. Pentium имел два отдельных 8-Кбайтных кэша: один для команд и один для данных. Одним из наиболее интересных новшеств, использованных в Pentium, являлась небольшая кэш-память, называемая Branch Target Buffer — БТВ (буфер меток переходов), которая позволяет динамически предсказывать переходы в исполняемых программах. По скорости выполнения операций с плавающей точкой Pentium оставлял далеко позади всех своих «собратьев по классу» — 1486DX-33 (почти в 10 раз), 1486DX2-66 (2,5 раза). Это достигалось, в частности, благодаря реализации оптимизированных алгоритмов, а также специализированным аппаратным блокам сложения, умножения и деления с восьмиступенчатой конвейеризацией, что позволяло выполнять операции с плавающей точкой за один такт. Как известно, в процессорах i486 специального конвейера для устройств с плавающей точкой предусмотрено не было. В дальнейшем выпускались версии процессоров второго поколения Pentium (P54C) с внутренним умножением частоты в 1,5 и 2 раза (кодовое название P54C) на 75/50, 90/60, 100/66, 120/60, 133/66, 150/60 и 166/66 МГц. Объявление первых моделей этих процессоров произошло в марте 1994 года. Для снижения рассеиваемой мощности с 13 до 4 Вт напряжение питания для P54C было снижено до 3,3 Вольт. Три режима энергопотребления были рассчитаны на максимальный ток 1 А, 50 мА и 100 мкА. Количество выводов возросло до 296 (рисунок 2). Размеры и тип корпуса также были изменены. Для производства кристалла была использована 0,6-микронная BiCMOS-технология. Количество транзисторов увеличено до 3,3 миллиона.

1.3.4 Процессоры-клоны

Большая часть рынка процессоров принадлежит фирме AMD, которая начинала с производства процессоров по технологии Intel, хотя и вносила в их конструкцию значительные усовершенствования. К примеру, производство 386-х процессоров-клонов AMD практически все время сопровождалось судебными исками со стороны Intel.

Несмотря на то, что основной упор сегодня фирмы AMD и Cyrix делают на процессоры с умножением частоты, в спектре их продукции по-прежнему присутствуют и «обычные» микросхемы. Указанные фирмы всегда старались наладить выпуск процессоров, отсутствующих в производственной программе Intel, и это им хорошо удавалось. Заслуженной популярностью пользовались, например, процессоры DX и SX, рассчитанные на внешнюю тактовую частоту 40 МГц. Эти изделия по производительности превосходили I486DX-33 и I486SX-33 и предлагались по более низким ценам. То же самое можно сказать о процессорах с умножением частоты, например Am486DX2-80. Кроме того, цены на микроэлектронные изделия AMD и Cyrix обычно были, да и сейчас ниже, чем на аналогичную продукцию Intel. Стоит отметить, что с уходом Intel с рынка 486-х процессоров, наряду с изделиями фирм AMD и Cyrix, на нем более активно начали предлагаться электронные компоненты, например, от компаний Texas Instruments, SGS-Thomson, UMC.

Вторым после Intel производителем микропроцессоров по праву считается компания AMD. Как уже отмечалось, хотя между изделиями Texas Instruments и AMD довольно много общего, они унаследовали различную технологическую базу. Что касается фирмы AMD, то, став в свое время безусловным лидером на рынке 386-х процессоров, она после ухода Intel практически добилась того же успеха и на рынке 486-х. Так, микропроцессоры Am486 потребляли меньшую мощность и имели некоторые усовершенствования в архитектуре, касающиеся управления памятью и организации использования ее в мультипроцессорных системах. Am486DX2/DX4 изготавливались с соблюдением технологических норм 0,5 мкм и имели напряжение питания 3,3–3,6 В. Благодаря этому до тактовой частоты 100 МГц на корпусе микросхемы

было достаточно только охлаждающего радиатора, работа с более высокими частотами (включая и 100 МГц) требовала уже микро-вентилятора. Вообще говоря, микросхемы DX2 и DX4 представляли собой один и тот же процессор с переменным коэффициентом внутреннего умножения частоты. Как правило, значение коэффициента выбиралось соответствующей установкой переключки на системной плате компьютера. Хотя данный кристалл в отличие от Intel DX4 имел лишь 8-Кбайтный кэш, как показывает практика, по производительности они примерно равны. Производительность Am486DX4-120 находилась на уровне Pentium 75/90. Выпускались также кристаллы, рассчитанные на тактовые частоты 133 и 160 МГц. В конце 1995 года было объявлено о появлении еще одного нового микропроцессора от AMD — Am5x86. Тактовая частота этой микросхемы составляла 133 МГц, а размер встроенной кэш-памяти с обратной записью достигал 16 Кбайт. Кристаллы производились с соблюдением технологических норм 0,35 мкм.

1.3.5 Теплообмен

Проблема теплообмена стала особенно актуальной с повышением рабочей тактовой частоты процессоров и ужесточением технологических норм при производстве самих кристаллов. Заметим, что микросхема, рассеивающая более 4 Вт мощности, уже требует охлаждения. Имеются данные, показывающие, что снижение рабочей температуры процессора на 10 градусов ведет к удвоению времени его безотказной работы, при этом скорость движения электронов в полупроводниках также повышается почти вдвое.

Стандартным решением для охлаждения микропроцессоров является малогабаритный вентилятор, установленный на радиаторе (Cool-CPU). Двигатель вентилятора запитывается от блока питания компьютера через переходной разъем. Сам радиатор может снизить температуру микросхемы примерно на 20 градусов, а вентилятор позволяет довести эту цифру до 40. Например, за 1994 год было произведено около 40 миллионов подобных устройств.

Кроме обычных малогабаритных вентиляторов для охлаждения процессоров ряд фирм предлагают несколько оригинальных решений, например охлаждающие устройства, использующие термоэлектрический эффект Пельтье (Peltier). Данный эффект заключается в том, что прохождение слабого электрического тока через контакт двух специально подобранных материалов сопровождается поглощением тепла. Подобные термоэлектрические устройства способны снизить температуру на 50–70 градусов. Известны также миниатюрные вентиляторы, которые, используя вихревой, или турбулентный, эффект, могут понижать температуру до 50–55 градусов.

1.3.6 Сопроцессоры

Важнейшей характеристикой любого компьютера является его быстродействие. Причем для ряда решаемых на компьютере задач одним из самых критичных параметров выступает скорость выполнения операций с плавающей точкой. Даже самые мощные универсальные микропроцессоры тратят на такие вычисления достаточно много времени. Поэтому вполне логичным было создание для этой цели специального устройства — микросхемы математического сопроцессора. До недавнего времени математический сопроцессор представлял собой специализированную интегральную схему, работающую во взаимодействии с центральным микропроцессором. Данная микросхема была предназначена только для одного — выполнения математических операций. Во всех микропроцессорах Intel от 486DX и выше сопроцессор интегрирован на кристалл основного процессора.

С другой стороны, хотя одно из значений слова «computer» и определяется как «тот, кто вычисляет», масса современных программных приложений, используемых на персональных компьютерах, вовсе не требует выполнения сложных математических операций. Впрочем, и интуитивно понятно, что для решения одних задач можно вполне обойтись без математического сопроцессора, для других же его отсутствие будет крайне нежелательным. Если не затрагивать специальных физических или математических (наверное, не очень сложных) задач моделирования, решаемых на персональных компьютерах, можно однозначно

сказать о необходимости установки в компьютер математического сопроцессора для работы, например, с 3-мерной графикой, издательскими пакетами, электронными таблицами, пакетами САПР, специальными математическими пакетами и т.п. При работе же с небольшими базами данных или обычными текстовыми редакторами использование сопроцессора не даст никаких ощутимых результатов. Бесплезным окажется сопроцессор в вашем компьютере и при работе с сетевыми операционными системами. По некоторым оценкам только треть всех владельцев персональных компьютеров эффективно используют математические сопроцессоры.

Первым математическим сопроцессором для персональных компьютеров IBM был NDP (Numerical Data Processor или MCP, Math Coprocessor) i8087 фирмы Intel, который появился еще в 1980 году, то есть за год до рождения первой «писишки» IBM. Отношение самой фирмы IBM к «собрату» i8088 было на первых порах достаточно прохладным — панелька (chip socket) для i8087 на системной плате компьютера, как правило, пустовала. Но со временем помимо «чисто интеловских» сопроцессоров в персональных компьютерах появились математические сопроцессоры и ряда других фирм, например ULSI (Ultra Large Scale Integration), AMD или Cyrix.

1.3.7 Что означает термин «сверхмасштабный»

Рассуждая о том, насколько компьютеры каждого нового поколения быстрее своих предшественников, производители часто упоминают о *сверхмасштабной* (superscalar) производительности. Что же означает это загадочное слово?

Грубо говоря, это означает, что если частоты часов старого и нового РС одинаковы, то новый РС все равно будет работать быстрее старого. Если частота часов нового РС в два раза больше, то новый РС будет работать более чем в два раза быстрее старого. То есть *производительность растет не пропорционально частоте, а быстрее*, отсюда и приставка «сверх».

Причин у сверхмасштабности несколько. Обычно она достигается усовершенствованием внутренней структуры процессора, а именно, встраиванием дополнительных исполняющих моду-

лей, увеличением размера конвейера инструкций, улучшением кэша L1 и т.д. Другая причина сверхмасштабной производительности — это умение процессора «мыслить» на несколько ходов вперед, т.е. умение специально выполнять инструкции, очередь которых еще не настала, в надежде на то, что они понадобятся в дальнейшем. Если в дальнейшем дело действительно доходит до этих инструкций, то процессор пролетает их на повышенной скорости, не теряя времени на их выполнение. Есть еще один способ повысить производительность, не трогая частоту, — поставить на материнскую плату дополнительный процессор.

Независимо от причин, термин «сверхмасштабный» (иногда используют не вполне удачную кальку *суперскалярный*) обозначает компьютер, работающий быстрее, чем можно было подумать. Это, конечно, здорово, но нет здесь той запутанности, мистичности (или техничности), которых можно было ожидать от расшифровки столь загадочно звучащего термина.

1.3.8 Тенденции развития микропроцессоров семейств INTEL-AMD

Зачем изучать историю? — Это неинтересно! Зачем знать, в каком году был выпущен тот или иной микропроцессор — скажете Вы. И будете не правы. Ибо только изучая историю, можно понять и предугадать будущее. Обратите внимание — процессор с 16-разрядной шиной данных i8086 появляется на год раньше процессора с 8-разрядной шиной данных i8088. Случайность — скажете Вы. А появление 32-разрядного i80386DX на год раньше 16-разрядного i80386SX? А появление 486 SX позже 486DX? Просто на момент появления i8086 не существовало 16-разрядных приложений. Другими словами, у появившегося процессора просто не было потребителя на тот момент. Точно также большинству задач для 486 процессора не требовался сопроцессор. Таким образом, налицо тенденция развития рынка микропроцессоров — сначала появляется новый процессор, а затем более дешевый упрощенный его вариант, совместимый с существующими программными продуктами и аппаратным обеспечением. Вторая тенденция — это «поддержка программного и аппаратного обеспечения сверху вниз». Это означает, что любой новый процессор,

например PENTIUM IV или DURON, должен поддерживать весь набор инструкций более ранних процессоров, например i8088. Это связано со стремлением разработчиков «привязать» потребителя к конкретной платформе. И пусть этот набор инструкций морально устарел, зато потребителю не приходится переходить на другое программное обеспечение. Обратная сторона этой медали — невозможность кардинальных перемен, ведущих к значительному улучшению архитектуры процессоров.

Вообще производительность процессора можно повысить тремя способами: увеличивая разрядность, повышая тактовую частоту, совершенствуя архитектуру. Разрядность микропроцессоров (РС совместимых) в настоящее время достигла стабильных величин — шины адреса 32 (4 Гб адресного пространства), шины данных 64. Действительно, пока сложно представить себе компьютер с более чем 4 Гб оперативной памяти. Работы в области увеличения тактовых частот процессоров ведутся очень активно, но это экстенсивный путь развития: повышение частоты в 2 раза при прочих равных увеличивает производительность на 70% и имеет физический и технологический пределы. Например, на сегодняшний (или уже вчерашний) день это 1,5–2 ГГц. Следовательно, основной путь развития — оптимизация структуры процессора и за счет этого увеличение производительности. Тенденций здесь несколько:

а) увеличение количества потоков, по которым могут производиться конвейерные вычисления,

б) увеличение внутренней кэш-памяти и деление ее на несколько уровней, что позволяет держать в «быстрой» памяти основные и наиболее часто используемые данные,

с) разбиение «больших» команд на несколько более мелких и удобных для процессора.

Кроме того, есть ряд характерных тенденций в технологии производства микропроцессоров. Это:

- Увеличение производительности требует дальнейшего снижения теплоотдачи процессора, а достигнуть этого можно в основном за счет снижения напряжения питания (с 5 В в 1991 году до 1,5 В в 2004).

- Дальнейшее увеличение производительности возможно

только за счет увеличения количества транзисторов на кристалле, а это требует совершенствования технологии производства и уменьшения расстояния между транзисторами на кристалле (с 0,6 микрон у Pentium I до 0,1 микрон у Pentium IV).

- Совершенствование системы охлаждения микропроцессоров.

Таким образом, можно уже сейчас предположить, каким будет процессор через 2–3 года:

- Технология изготовления 0,1–0,05 микрон.
- Тактовая частота 3–5 ГГц.
- Внутренний кэш 4-го уровня.
- Внешняя шина 200–500 МГц.
- Шина данных 128 (256?) разрядов.
- Набор инструкций RISC процессора, впрочем, эмулирующий все предыдущие версии.

При этом ошибочно было бы предполагать, что нельзя предугадать примерную цену на новый процессор. Дело в том, что цены на все высокотехнологичное оборудование меняются очень быстро. Связано это как с инфляцией, так и с необходимостью продвижения на рынке новых изделий. Таким образом, при себестоимости 5–10 долларов продажная цена процессора может достигать величин 500–2000 долларов.

График зависимости цены микропроцессора в зависимости от его производительности представлен на рис. 3. На этом графике специально нет координатной сетки, поскольку цена меняется практически каждую неделю, и, следовательно, каждый может построить такой график сам для конкретного момента времени, используя любой индекс оценки производительности процессора (например iCOMP). Данную кривую можно аппроксимировать тремя прямыми, первая из которых, почти параллельная оси абсцисс, характеризует процессоры, выпускавшиеся 2–5 лет назад, и уже снятые с производства. Цена на них почти равна себестоимости. Третья, почти параллельная оси ординат, характеризует процессоры, выпускаемые в последнее время, а также только что выпущенные. Цена на них резко завышена. И, наконец, вторая прямая характеризует процессоры, производящиеся более года, но не более двух. Именно эти процессоры и являются оптимальными по соотношению цена/качество.

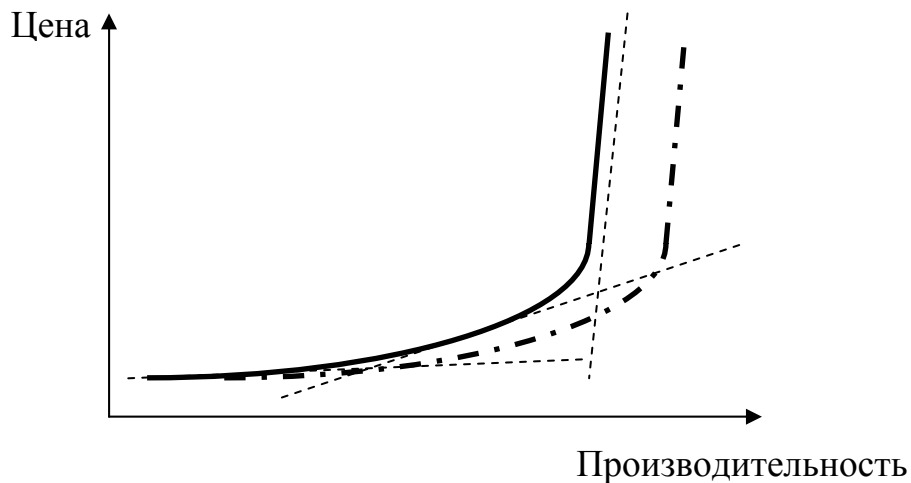


Рис. 3 — Зависимость цены микропроцессора от его производительности.

Следует заметить, что штрих-пунктирной линией на этом же графике можно отметить то, что будет через какой-либо промежуток времени, например полгода. Те процессоры, которые стоили «бешеных» денег, потеряют в цене, зато появятся новые, которые будут стоить столько же, сколько и их предшественники полгода назад. Таким образом, во все времена, современный производительный процессор стоил и будет стоить от 70 до 200 долларов, поскольку цена его связана не с себестоимостью производства (напомню, 5–10 долларов), а с реальным спросом населения.

1.4 Память

Выражаясь языком нашего недавнего классического наследия, процессор, память и устройства ввода-вывода — это «три источника и три составные части» компьютера. Практически все компьютеры используют три вида памяти: оперативную, постоянную и внешнюю.

Оперативная память предназначена для хранения переменной информации, так как допускает изменение своего содержания в ходе выполнения микропроцессором вычислительных операций. Таким образом, этот вид памяти обеспечивает режимы записи, считывания и хранения информации. Поскольку в любой момент времени доступ может осуществляться к произвольно

выбранной ячейке, то этот вид памяти называют также памятью с произвольной выборкой — RAM (Random Access Memory). Для построения запоминающих устройств типа RAM используют микросхемы статической и динамической памяти.

Постоянная память, где хранится такая информация, которая не должна меняться в ходе выполнения микропроцессором программы, имеет собственное название — ROM (Read Only Memory), которое указывает на то, что обеспечиваются только режимы считывания и хранения. Постоянная память обладает тем преимуществом, что может сохранять информацию и при отключенном питании. Это свойство получило название энергонезависимости. Все микросхемы постоянной памяти по способу занесения в них информации (программированию) делятся на масочные (ROM), программируемые изготовителем, однократно программируемые пользователем (Programmable ROM) и многократно программируемые пользователем (Erasable PROM). Последние в свою очередь подразделяются на стираемые электрически и с помощью ультрафиолетового облучения. К элементам EPROM с электрическим стиранием информации относятся и микросхемы флэш-памяти (flash). От обычных EPROM они отличаются высокой скоростью доступа и быстрым стиранием записанной информации.

Внешняя память реализована обычно на магнитных или оптических носителях.

1.4.1 Классификация оперативной памяти

Если от типа процессора зависит количество адресуемой памяти, то быстродействие используемой оперативной памяти в свою очередь во многом определяет скорость работы процессора, а в конечном итоге влияет и на производительность всей системы. Практически любой PC-совместимый компьютер оснащен оперативной памятью, выполненной на микросхемах динамического типа с произвольной выборкой (DRAM, Dynamic Random Access Memory). Каждый бит такой памяти представляется в виде наличия (или отсутствия) заряда на конденсаторе, образованном в структуре полупроводникового кристалла. Другой тип памяти — статический (SRAM, Static RAM) — в качестве элементарной ячейки памяти

использует так называемый статический триггер. Если для реализации одного запоминающего элемента динамической памяти требуется 1–2 транзистора, то для статической их число возрастает до 4–6. Статический тип памяти обладает высоким быстродействием и, как правило, используется в самых «узких» местах системы, например для организации кэш-памяти.

Теперь коротко рассмотрим принцип работы динамической памяти. Единственным способом выяснить, заряжен или разряжен конденсатор, является попытка разрядить его. Если конденсатор был действительно заряжен (то есть хранил единичный бит), то после разряда его, разумеется, надо снова подзарядить. Ячейки памяти динамического типа конфигурируются обычно в матрицу строк и столбцов, причем процесс считывания организуется таким образом, что содержимое целой строки переносится в некий буфер, выполненный на элементах статической памяти. После считывания соответствующего бита содержимое буфера перезаписывается в ту же строку ячеек динамической памяти, то есть производится перезарядка конденсаторов, которые до считывания были в заряженном состоянии.

Не следует также забывать о том, что время хранения заряда конденсатором ограничено (из-за «паразитных» утечек). Таким образом, чтобы не потерять имеющиеся данные, необходимо периодическое восстановление записанной информации, которое выполняется в циклах регенерации (refresh cycle). Кстати, в первых моделях РС для индикации моментов времени, когда возникала необходимость регенерации динамической памяти, использовался специально выделенный таймер. Для считывания содержимого ячеек (которое, разумеется, сопровождается перезаписью информации) применялся один из каналов контроллера прямого доступа DMA. Стоит, правда, отметить, что микросхемы динамической памяти имели встроенные средства регенерации, что уменьшало загрузку процессора. Тем не менее операции разрядки-перезарядки занимают определенное время, которое снижает скорость работы динамической памяти. Это является, пожалуй, одним из основных недостатков динамической памяти, так как по критерию, учитывающему информационную емкость, стоимость и энергопотребление, этот тип памяти во многих случаях предпочтительнее статической.

1.4.2 Корпуса и маркировка

Элементы динамической памяти для персональных компьютеров конструктивно бывают выполнены либо в виде отдельных микросхем в корпусах типа DIP (Dual In line Package), либо в виде модулей памяти типа SIP/SIPP (Single In line Pin Package), типа SIMM (Single In line Memory Module) или DIMM (Dual In line Memory Module). Модули памяти представляют собой небольшие текстолитовые платы с печатным монтажом и установленными на них микросхемами памяти в DIP-корпусах. В большинстве современных модулей памяти используются микросхемы в корпусах для поверхностного монтажа (например, типа SOP — Small Outline Package). Для подключения к системной плате на SIMM- и DIMM-модулях используется печатный («ножевой») разъем, а на SIP-модулях — штыревой. У DIMM-модулей в отличие от SIMM контакты на противоположных сторонах платы электрически не связаны между собой. Это дает возможность практически вдвое увеличить количество выводов модуля. Сами микросхемы памяти также устанавливаются на плате с двух сторон. Такие модули называют двусторонними (double sided) SIMM.

Когда-то наиболее распространенными были 30-контактные SIMM-модули с байтовой организацией и контролем четности (емкостью 256 Кбайт, 1 Мбайт, 4 Мбайта и 16 Мбайт). Здесь стоит, видимо, сказать несколько слов о контроле четности.

Так, при записи байта информации в запоминающее устройство определяется дополнительный контрольный разряд, который вычисляется как сумма по модулю 2 всех информационных битов. Обычно контрольный разряд равен нулю, если число единиц в байте было четным, и наоборот, он равен единице, если число единиц в группе было нечетным. Таким образом, при чтении ранее записанного байта, вновь получив контрольный разряд и сравнив его с уже имеющимся, можно говорить о достоверности получаемой информации. Этот метод нашел широкое распространение для контроля информации в оперативной памяти IBM PC-совместимых компьютеров.

С другой стороны, на каждые 8 информационных разрядов приходится один бит четности, то есть стоимость модуля увеличивается более чем на 10%. Современные технологические успе-

хи производителей микросхем памяти существенно повысили степень их надежности (среднее время безотказной работы составляет теперь несколько десятков лет), что дало возможность исключить бит четности как таковой.

Другая крайность: некоторые изготовители очень мощных персональных компьютеров используют даже коды с коррекцией ошибок (ЕСС), чтобы защитить содержимое RAM так же, как это делается на жестких дисках. Это достигается либо использованием нескольких модулей RAM без четности (но при наличии запасных модулей, где и хранятся данные ЕСС), либо RAM с четностью (при сохранении данных ЕСС в объединенных битах четности).

Надо ли дополнительно платить за память с четностью. Не обязательно. Такую вещь хорошо иметь, но она может и не стоить своих денег. То же можно сказать и про память с ЕСС. Возможно, решающими факторами будут объем памяти PC (чем большей памятью он обладает, тем больше вероятность однажды ошибиться) и серьезность вашего отношения к целостности данных. Чем дороже ваш PC и чем больше у него памяти, тем выгоднее иметь четность или даже ЕСС-защиту. Но даже и не думайте об этом, если вы не защитили данные, установив источник бесперебойного питания (UPS) на вашем PC. Защита с помощью UPS гораздо важнее, и если это для вас слишком дорого, значит, вам не нужна и память с четностью.

Для 32-разрядных процессоров 30-контактные модули должны были устанавливаться на системную плату в количестве, кратном 4 (для получения длины слова в 32 бита). Поскольку на плате обычно присутствует 8 таких разъемов, то максимальный объем памяти в этом случае ограничен 64 Мбайтами.

С появлением Pentium стали применяться 72-контактные 36-битовые модули (32 бита длина слова и по биту контроля четности на каждый байт) (рис. 4). Они обозначались как 256Kx36, 512Kx36, 1Mx36, 2Mx36, 4Mx36, 8Mx36, 16Mx36, что соответствовало емкости 1, 2, 4, 8, 16, 32 и 64 Мбайта. Понятно, что такие модули для 486-х микропроцессоров могут уже устанавливаться и по одному. Макси-

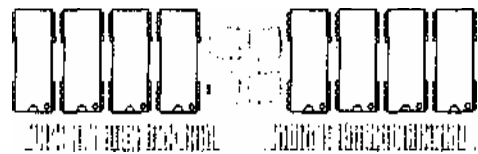


Рис. 4 — Внешний вид модуля памяти

мально достижимый в этом случае объем памяти может быть разным, что зависит не только от количества разъемов под модули, но и от того, с какой максимальной емкостью модули поддерживаются контроллером памяти.

Обычно вся оперативная память персонального компьютера делится на несколько банков, причем вид и тип элементов, используемых в них, зависят от конструкции системной платы и приводятся в ее техническом описании. Банк определяет наименьшее количество памяти, которое может быть адресовано процессором за один раз и соответствует разрядности шины данных этого процессора. Например, микропроцессор 18088 за один раз может адресовать только один байт, i80286 — 16 бит, i80386 — 32, Pentium — 64.

Модули SIMM оказались весьма популярными, однако и им нашлась замена, и они практически совсем уступили место DIMM.

Модуль DIMM (Dual Inline Memory Module — модуль памяти «в две линейки») представляет собой целый блок памяти, все элементы которого соответствуют друг другу автоматически, без особых усилий со стороны пользователя. (Более того, эти модули защищены прорезями-ключами от некорректного использования, так что вы не сможете установить DIMM, рассчитанный на 3,3 вольта, в разъем для 5-вольтового DIMM, и наоборот. У модулей DIMM, кроме того, гораздо больше проводов заземления, чем у SIMM, что становится критичным по мере роста скоростей, на которых работают модули памяти).

Микросхемы памяти — это изделия высочайшего технологического уровня, которые выпускаются небольшим количеством японских, корейских, американских и европейских фирм с мировым именем: Samsung, NEC, Toshiba, Mitsubishi, Oki, Hitachi, Sanyo, Goldstar, Hyundai, Motorola, Texas Instruments, Micron Technology, Siemens. Так, первой фирмой, выпустившей 256-Мбитную микросхему динамической памяти, стала корейская компания Samsung. Некоторые фирмы поставляют не только готовые микросхемы, но и полуфабрикаты — кремниевые пластины со сформированными на них схемами памяти. Они используются другими фирмами, которые монтируют чипы в корпус, тестируют и могут продавать под другой торговой маркой, как пра-

вило, по более низким ценам. Некоторые известные производители компьютеров (brand name) применяют в своих изделиях нестандартные модули памяти, так называемые «собственные» (proprietary memory).

1.4.3 Режимы работы памяти

Напомним, что для адресации к любому биту, например 1-Мбитной микросхемы, требуются 20 адресных линий, в то время как корпус такой микросхемы имеет только 18 выводов. Дело здесь в том, что для этой цели используется так называемое мультиплексирование адресов. Полный адрес ячейки данных делится на два компонента — адрес строки (row address) и адрес столбца (column address). Для сопровождения первого компонента служит сигнал RAS (Row Address Strobe), а второго — сигнал CAS (Column Address Strobe). Вообще говоря, под временем выборки для микросхемы памяти понимается длительность именно сигнала RAS.

В процессе обращения к микросхеме DRAM для записи или считывания информации необходимо сначала подать на ее адресные входы код адреса строки и одновременно с ним (или с некоторой ненормируемой задержкой) сигнал RAS, затем через нормированное время задержки должен быть подан код адреса столбца, сопровождаемый сигналом CAS. Следующее обращение к этой микросхеме возможно только после промежутка времени, в течение которого происходит восстановление (перезарядка) внутренних цепей микросхемы. Это время называют временем перезарядки (prechargeitme), причем оно составляет почти 90% от общего времени выборки.

Одним из способов повышения быстродействия динамической памяти является метод управления памятью с чередованием адресов (interleaving mode). Дело в том, что, используя даже обычные микросхемы DRAM, задержки, связанной со временем перезарядки, можно избежать, если каждые последовательно выбираемые ячейки памяти будут относиться к разным банкам памяти. Таким образом, метод чередования адресов предусматривает деление памяти на блоки (банки), из которых процессор должен считывать данные попеременно. То есть пока считываются

данные из одной группы микросхем, другая группа получает время на перезарядку.

Другим способом повышения быстродействия, требующим, как правило, небольших архитектурных добавок, является метод страничной выборки (paging mode). Он базируется на том, что повторения сигнала RAS также можно избежать, если адреса строк выбираемых ячеек памяти лежат в пределах одной страницы, то есть адрес их строк неизменен. Напомним, что в микросхеме динамической памяти считывание в статический буфер происходит для всей строки целиком, конкретный же бит выбирается уже адресом столбца. Понятно, что если следующий подлежащий считыванию бит находится в той же строке, то нет смысла читать ее в буфер еще раз. Наиболее распространены две разновидности подобного режима: с повторением сигнала стробирования CAS при изменении младшей части адреса и без повторения. В последнем случае быстродействие, естественно, увеличивается.

В стремлении создавать все более быстрые чипы DRAM производители усовершенствовали их во многих аспектах. Каждый носит свое замысловатое название и имеет особые преимущества.

Микросхемы DRAM, реализующие страничный режим, часто называют FPM (Fast Page Mode). Их принцип работы основан на том, что обычно у микросхем DRAM только половина необходимых адресных линий. Одни и те же контакты используются для адресов строк и столбцов, а их назначение определяется временем поступления сигнала. Для ускорения работы, в случае если процессор обращается к нескольким последовательным или близко расположенным ячейкам, чипу памяти адрес строки сообщается только однажды, а затем осуществляется последовательный доступ к ячейкам, расположенным в указанных далее столбцах. Новый адрес строки передается только при переходе к новой строке. Тем не менее использование даже подобных способов повышения быстродействия оперативной памяти, выполненной на стандартных DRAM, не давало требуемых результатов.

Другая стратегия (используемая в EDO DRAM) основана на том, что считанные данные некоторое время могут накапливаться на выходе. Подобный принцип, называемый *ускоренным*

выводом данных (Extended Data Out), позволяет обращаться к значению некоторой ячейки, и в то время как результат предыдущего запроса еще считывается, чип может подготавливать следующий бит данных.

Двухпортовые DRAM позволяют обращаться сразу к двум ячейкам (по двум адресам). Это осуществляется с помощью двух полных наборов схем для чтения данных из битового массива. Когда миллионы битовых ячеек расположены на одном чипе, дополнительные расходы на это незаметны на фоне остальной стоимости. Такие устройства несимметричны. То есть в то время как один порт ввода-вывода позволяет обращаться к произвольной ячейке (и имеет схемы для записи данных в произвольные ячейки), второй используется только для чтения данных и только всей строки за один раз. Такие двухпортовые чипы DRAM, называемые также видео RAM: (Video RAM, VRAM), особенно полезны для кадровых видеобуферов, так как они изначально используются CPU для оперативного чтения и записи, а также для линейного чтения (схемой видеоизображения).

WRAM (Windows RAM) — это специальная версия VRAM, оптимизированная для типа доступа, обычного в ОС Windows и приложений Windows.

1.4.4 Новые типы динамической памяти

Стоит начать с того, что за последние десять лет скорость работы микропроцессоров возросла во много раз больше, чем быстродействие оперативной памяти. Так, high-end-компьютеры образца 1984 года (PC/AT) работали с тактовой частотой 10 МГц. Сегодня внутренняя тактовая частота процессоров Pentium достигает, например, 1 ГГц. За этот же период время доступа микросхем оперативной памяти снизилось со 150 всего до 10–5 нс. Причем заметим, что стандартная схемотехника, применяемая в производстве оперативной памяти, не обещает особых изменений и в будущем. Что касается внешней кэш-памяти (подробнее о ней чуть позже), то, во-первых, она достаточно дорогая, а во-вторых, ее использование оказывается эффективным только до тех пор, пока программное приложение не начинает работать с данными, находящимися в разных частях оперативной памяти.

Ведь, как известно, основной принцип кэш-памяти заключается в том, что 10–20% команд или данных будут необходимы в 80–90% случаев. Кроме того, считается справедливым предположение, что если считываются инструкции или набор данных из одного места памяти, то соседние с ними также скоро будут востребованы. Стоит напомнить, что, например, в микросхемах, совместимых с архитектурой Intel, начиная с 486-х процессоров, для доступа к памяти реализован так называемый пакетный (burst) режим. Он заключается в том, что при необходимости чтения одного слова процессор вместе с ним считывает еще три, расположенных рядом. Обычно время пересылки измеряют в тактах и записывают, например, так: 6-3-3-3. Это означает, что если на первую пересылку данных из памяти потребовалось 6 тактов работы процессора, то на каждую последующую — только по 3.

Итак, если процессору 8088 требовалось четыре такта для передачи данных, то современным процессорам — всего один. Например, Pentium, работающий на внутренней тактовой частоте 100 МГц, с внешней кэш-памятью (время доступа 15 нс) мог обеспечить пакетный режим 3-2-2-2. Для обмена с динамической памятью параметры будут в два-три раза хуже, например 7-3-3-3. Хотя теоретически микропроцессор 486DX2-66 мог бы реализовать пакетный режим 2-1-1-1, но с обычной динамической памятью (время доступа 70 нс) реальные значения будут на уровне 5-2-2-2.

В настоящее время можно выделить два основных схемотехнических решения, используемых для увеличения быстродействия динамической памяти. Одно из них основано на синхронной работе памяти и процессора, что достигается использованием внутренней конвейерной архитектуры и чередованием адресов. Другое решение предполагает включение в структуру динамической памяти определенного количества быстрой статической памяти, которая в данном случае работает примерно как встроенный кэш.

Синхронная DRAM (Synchronous DRAM, SDRAM) и *синхронная графическая RAM* (Synchronous Graphics RAM, SGRAM) — еще две популярных вариации на ту же тему. Оба эти типа однопортовые и поэтому несколько дешевле, чем VRAM или WRAM.

Преимущество SDRAM и SGRAM перед обычной DRAM в том, что эти микросхемы памяти используют тот же сигнал таймера, что и CPU. Это означает, что данные микросхемы памяти готовы к передаче данных тогда, когда CPU их ожидает. Подобные микросхемы используют трехступенчатую конвейерную архитектуру и, кроме того, внутренний доступ типа «пинг-понг» к двум блокам памяти с чередованием адресов. Тактирование микросхем осуществляется внешней частотой для микропроцессоров. Современные SDRAM могут работать на тактовых частотах 66, 75, 83, 100 и 133 МГц. Пионерами в разработке подобных устройств являются фирмы Samsung и NEC. При использовании других типов памяти CPU требуется больше времени, чтобы убедиться, что микросхема передала данные, а это создает непроизводительные задержки. Стоит отметить, что первый доступ в пакетном режиме выполняется медленнее для микросхем SDRAM, нежели, например, EDRAM. Для последующих трех — все наоборот.

Одним из наиболее быстродействующих типов памяти является RDRAM (Rambus DRAM). Это, наверное, наиболее сильно отличающаяся дизайном память, потребовавшая абсолютно новой архитектуры. Кроме того, из всех созданных на сегодняшний день технологий эта обладает наибольшим потенциалом по скорости. Однако этот потенциал стоит недешево. RDRAM должна соединяться с CPU через очень специфическую шину, которая не может быть слишком длинной. Тактируемая частотой 250 МГц 9-разрядная RDRAM достигает пиковой скорости передачи 500 Мбайт/с, а 2 ГГц — 4 Гбайт/с. Вообще говоря, Rambus-архитектура состоит из трех частей: Rambus-интерфейса, Rambus-канала и Rambus-микросхем. Вследствие использования высоких тактовых частот серьезные требования предъявляются к печатным платам и соединениям. Подобная память, впервые разработанная американской компанией Rambus, в настоящее время выпускается такими компаниями, как NEC, Fujitsu и Toshiba. Пройдет время, прежде чем мы увидим RDRAM в качестве основной памяти PC. Несмотря на это, Intel обратилась к RDRAM как к решению проблемы нехватки памяти компьютера в недалеком будущем, а так как Intel является основным производителем материнских плат для PC, можно ожидать претворе-

ния этой затеи в жизнь.

Стоит отметить, что существуют и другие типы памяти, причем более или менее экзотические. Например, в MDRAM (Multibank DRAM) вся память делится на 10 маленьких банков, время доступа к которым существенно меньше времени доступа к одному большому: 15 не против 50 нс. Тем не менее по оценкам экспертов в ближайшее время не следует ожидать ухода от массового использования в персональных компьютерах так называемой EDO (Extended Data Out) DRAM или ее разновидности BEDO (Burst EDO) DRAM. Во-первых, она обеспечивает более высокую скорость передачи (особенно в пакетном режиме), а во-вторых, полностью совместима по выводам с современными SIMM-модулями DRAM. В отличие от обычных микросхем DRAM в EDO DRAM добавлен набор регистров-«защелок», благодаря которым данные на выходе могут удерживаться даже в течение следующего запроса к микросхеме. Такого эффекта можно добиться на обыкновенных DRAM только в режиме чередования адресов. Напомним, что в любом обращении к памяти можно выделить три фазы: начало доступа; период, когда данные становятся действительными, и непосредственно передача. Эти фазы повторяются последовательно для каждой ячейки в считываемой строке. В случае с EDO-памятью временные параметры (а следовательно, и быстродействие) улучшаются за счет исключения циклов ожидания в фазе готовности данных. По некоторым данным EDO DRAM работает быстрее FPM DRAM примерно на 20–25%. Однако при использовании кэш-памяти второго уровня (L2) быстродействие возрастает только на 5%. Стоит отметить, что EDO-память дороже обычных DRAM на 7–10%. Именно поэтому в настоящее время EDO-память рекомендуется использовать в недорогих системах без кэш-памяти, что в данном случае достаточно эффективно. Кроме того, оправдано применение EDO-памяти и в многозадачных системах.

Эта история не имеет конца. Я всего лишь надеюсь дать некоторое представление о том, как выглядит ситуация в настоящий момент. Через несколько лет, без сомнения, перед нами предстанет совершенно иная картина чипов памяти. Можно быть уверенным только в одном, — это будет более быстрая, более

дешевая и более вместительная память. Естественно, программы будущего наверняка будут требовать больше места и скорости, чем мы сможем им предоставить.

1.4.5 Видеопамять

Видеоподсистеме нужна какая-то память, где она могла бы строить изображение, впоследствии выводимое на экран монитора. Эту память называют *буфером видеокadra* (video frame buffer). (Если мониторов два, то у каждого из них, скорее всего, будет свой кадровый буфер.) Кроме того, видеосистемы обычно имеют в своем распоряжении еще и дополнительную (по отношению к кадровому буферу) память.

В современных видеосистемах с высоким разрешением и большой глубиной цветов вполне обычным уже является кадровый видеобуфер размером 8–64 Мб. Еще больше памяти используется в высококачественных (high-end) системах.

И наконец, видеосистема может использовать свою память, не привлекая к этому CPU. Основное назначение кадрового буфера в видеосистеме — это сохранение изображения, находящегося в данный момент на экране, и постоянная подкачка пикселей к монитору, чтобы он мог перерисовывать изображение. Но системе может понадобиться место для временного размещения узора какого-либо шрифта или некоторое рабочее пространство для проведения различных действий графического ускорения. Обычно именно для этого используется незанятая кадровым изображением часть кадрового буфера памяти.

1.4.6 Кэш-память

Функционально кэш-память предназначена для согласования скорости работы сравнительно медленных устройств, таких например, как динамическая память, с относительно быстрым микропроцессором. Дело в том, что работа большинства элементов, на которых построен процессор, во многом похожа на работу ячеек статической памяти — триггеров. Поэтому их быстроедействие существенно выше, нежели элементов динамической памяти. Ис-

пользование кэш-памяти позволяет избежать циклов ожидания в его работе, которые снижают производительность всей системы.

Для пользователей IBM PC-совместимых компьютеров использование кэш-памяти началось еще с 386-х микропроцессоров. Для таких устройств, синхронизируемых, например, частотой 33 МГц, тактовый период составляет приблизительно 30 нс. Обычные микросхемы динамической памяти имеют время выборки от 60 до 100 нс. Отсюда, в частности, следует, что центральный процессор вынужден простаивать 2–3 периода тактовой частоты (то есть имеет 2–3 цикла ожидания), пока информация из соответствующих микросхем памяти установится на системной шине данных компьютера. Понятно, что в это время процессор не может выполнять никакую другую работу. Такая ситуация приводит к снижению общей производительности системы, что, разумеется, крайне нежелательно. Таким образом, узким местом системы становится оперативная динамическая память. Из этого положения существует, казалось бы, простой выход — использовать в качестве основной памяти достаточно быструю статическую. Однако если основную оперативную память выполнить на микросхемах статической памяти, то стоимость компьютера возрастет очень существенно. Таким образом, с помощью технологии обработки, использующей кэш-память, найден определенный компромисс между стоимостью и производительностью системы.

Заметим, что работа кэш-памяти практически «прозрачна» (то есть невидима) для пользователя. Тем не менее она очень важна. В IBM PC-совместимых компьютерах технология использования кэш-памяти помимо обмена данными между микропроцессором и основной оперативной памятью находит применение также между основной оперативной памятью и внешней (накопителями на сменных и несменных носителях).

Архитектура кэш-памяти

Прежде чем говорить об архитектуре кэш-памяти, введем такое понятие, как длина строки-кэша (cache-line). Если при обмене данными между устройствами речь обычно идет о блоке информации, то для кэш-памяти существует некий набор данных, называемых строкой. Итак, архитектура кэш-памяти определяется тем, каким образом достаточно большая по размеру основная

память отображается на сравнительно небольшой кэш. Существуют, вообще говоря, три разновидности отображения: кэш-память с прямым отображением (direct-mapped cache), частично-, или наборно-ассоциативная (set-associative cache) и полностью ассоциативная (fully associative cache). Все эти архитектуры так или иначе используются для построения кэш-памяти современных микропроцессоров.

Кэш-память с прямым отображением

Самой простой организацией обладает кэш-память с прямым отображением. В этом случае адрес памяти полностью определяет используемую строку кэша. Таким образом, один или несколько блоков оперативной памяти строго соответствуют одной строке кэша, однако поскольку занимать ее в одно и то же время может только один из них, то для каждой строки используется специальный признак — тег (tag).

Преимуществом реализации такого типа архитектуры являются довольно низкие затраты, поскольку, по сути, требуется всего лишь одна операция сравнения (для тегов). Недостатки ее, впрочем, также очевидны. Например, если два блока данных основной памяти, используемые одинаково часто, претендуют на одну и ту же строку в кэше. Внешняя кэш-память с прямым отображением использовалась, например, вместе с 386-ми процессорами (кэш-контроллер 182385), а внутренняя — в микропроцессоре DEC Alpha 21064.

Полностью ассоциативная архитектура

Другим типом архитектуры является полностью ассоциативная кэш-память. В этом случае любой блок памяти может занимать любую строку кэша. Полный адрес памяти делится только на две части: младшие разряды — смещение в строке и старшие разряды — информация о теге. В этой архитектуре решена проблема конфликтов адресов, однако сама кэш-память требует для своей реализации больших аппаратных затрат, поскольку значения тегов должны уже сравниваться для всех линий кэша. Тем не менее микропроцессор 6x86 фирмы Сугіх имел вторичную встроенную 256-байтную кэш-память для команд, которая имела полностью ассоциативную архитектуру.

Наборно-ассоциативная архитектура

Разумным компромиссом между двумя рассмотренными архитектурами является наборно-ассоциативная организация кэш-памяти. В этом случае несколько линий (две, четыре, пять, восемь) объединяются в наборы, и средние биты адреса памяти определяют уже не конкретную линию (как в прямом отображении), а набор. Сравнение тегов (со значением старших разрядов) производится только для линий, входящих в набор. Подобную архитектуру имеет подавляющее число процессоров, например 486DX, 486DX2, Intel DX4, Pentium и т.д.

По количеству линий кэша, входящих в набор, подобная архитектура может называться 2-входовой (2-way set associative), 4-входовой (4-way set associative) и т.д.

Обновление информации

Каждый раз, когда микропроцессору требуется информация, отсутствующая в кэше (cache-miss), он вынужден обращаться через системную шину к основной оперативной памяти. После этого обычно решается, должна ли происходить замена строки в кэш-памяти и какая конкретно строка кэша будет заменена. В подавляющем большинстве случаев об этом заботится встроенный в контроллер LRU-алгоритм (Last Recently Used), который обновляет именно ту строку кэша, которая используется менее интенсивно.

1.5 BIOS и CMOS RAM

Базовая система ввода-вывода BIOS (Basic Input Output System) называется так потому, что включает в себя обширный набор программ ввода-вывода, благодаря которым операционная система и прикладные программы могут взаимодействовать с различными устройствами как самого компьютера, так и подключенными к нему. Большинство современных видеоадаптеров, а также контроллеры накопителей имеют собственную систему BIOS, которая обычно дополняет системную. Во многих случаях программы, входящие в конкретную BIOS, заменяют соответствующие программные модули основной BIOS. Вызов программ BIOS, как правило, осуществляется через программные или аппа-

ратные прерывания.

Система BIOS, помимо программ взаимодействия с аппаратными средствами на физическом уровне, содержит программу тестирования при включении питания компьютера POST (Power-On-Self-Test) и программу начального загрузчика. Последняя программа необходима для загрузки операционной системы с соответствующего накопителя.

Система BIOS в IBM PC-совместимых компьютерах реализована в виде одной или двух микросхем, установленных на системной плате компьютера. Название «ROM BIOS» в настоящее время не совсем справедливо, ибо «ROM» предполагает использование постоянных запоминающих устройств (ROM — Read Only Memory), а для хранения кодов BIOS применяются в основном перепрограммируемые (стираемые электрически или с помощью ультрафиолетового излучения) запоминающие устройства, например ППЗУ ЭС (Electrically Erasable Programmable ROM, EEPROM). Мало того, наиболее перспективной для хранения системы BIOS является сейчас флэш-память.

Поскольку содержимое ROM BIOS фирмы IBM было защищено авторским правом (то есть его нельзя подвергать копированию), то большинство других производителей компьютеров вынуждены были использовать микросхемы BIOS независимых фирм, разумеется, практически полностью совместимые с оригиналом. Наиболее известны из этих фирм три: American Megatrends Inc. (AMI), Award Software и Phoenix Technologies. Заметим, что конкретные версии BIOS неразрывно связаны с набором микросхем (chipset), размещенных на системной плате.

Система BIOS в компьютерах неразрывно связана с аббревиатурой CMOS RAM. Под этим понимается «неизменяемая» память, в которой хранится информация о текущих показаниях часов, значении времени для будильника, конфигурации компьютера: количестве памяти, типах накопителей и т.д. Именно в такой информации нуждаются программные модули системы BIOS. Своим названием CMOS RAM обязана тому, что выполнена на основе КМОП-структур (CMOS — Complementary Metal-Oxide-Semiconductor), которые, как известно, отличаются малым энергопотреблением. Поскольку в данной микросхеме реализованы также часы реального времени RTC (Real Time Clock), то ее

обычно называют RTC CMOS RAM. В современных компьютерах функции этой микросхемы включены в одну из вспомогательных микросхем, например периферийного контроллера. CMOS-память энергонезависима только постольку, поскольку постоянно подпитывается, например от аккумулятора, расположенного на системной плате, или батареи гальванических элементов, как правило, смонтированной на корпусе системного блока. Заметим, что большинство системных плат допускают питание CMOS RAM как от встроенного, так и от внешнего источников.

В системе BIOS имеется программа, называемая Setup, которая может изменять содержимое CMOS-памяти. Вызывается эта программа определенной комбинацией клавиш, которая обычно высвечивается на экране монитора после включения питания компьютера. Некоторые из данных комбинаций для различных фирм-производителей приведены в таблице 1. Заметим, что войти в программу Setup можно либо после успешной загрузки компьютера (нажав соответствующую комбинацию клавиш), либо после возникновения ошибки (клавиши F1 или F2). Во время загрузки компьютера запустить программу Setup можно, например для системы BIOS фирмы AMI, нажав клавишу DEL.

Таблица 1

Комбинации клавиш для входа в BIOS различных фирм

Фирма-изготовитель	Комбинация клавиш	Примечание
Phoenix Technologies	CTL-ALT-S	или F2 после ошибки
American Megatrends	DEL	в течение загрузки
Award Software	CTL-ALT-ESC	—
Chips&Technologies	CTL-ALT-S	или F2 после ошибки
Zenith	CTL-ALT-INS	—
Quadtel	CTL-ALT-S	или F2 после ошибки

Программы Setup многих фирм могут не только выполнять стандартные установки, доступные практически на любой AT/286, но и имеют ряд дополнительных возможностей.

Расширенные установки Advanced CMOS Setup или Advanced ChipSet Setup включают в себя дополнительные воз-

возможности конфигурирования системной платы, которые во многом зависят от используемого набора вспомогательных микросхем. Наиболее общими являются обычно такие возможности, как допустимая скорость ввода символов с клавиатуры (по умолчанию 15 символов в секунду), тестирование памяти выше границы 1 Мбайт, разрешение использования арифметического сопроцессора, приоритет или последовательность загрузки (то есть попытка загрузки компьютера сначала с накопителя со сменным, а затем несменным носителем или наоборот), установка определенной тактовой частоты микропроцессора при включении, разрешение парольной защиты, запрет контроля четности памяти и т.д. Кроме того, возможна установка тактовой частоты системной шины, а также числа тактов ожидания (или временной задержки) для микропроцессора при обращении к устройствам ввода-вывода, оперативной и/или кэш-памяти.

Заметим, что в случае повреждения микросхемы CMOS RAM (или разряда батареи/аккумулятора) программа Setup имеет возможность воспользоваться некой информацией по умолчанию (BIOS Setup Default Values), которая хранится в таблице соответствующей микросхемы ROM BIOS. Кстати, на некоторых материнских платах питание микросхемы CMOS RAM может осуществляться как от внутреннего, так и от внешнего источников. Выбор определяется установкой соответствующей перемычки.

Обычно программа Setup поддерживает также установки, связанные с шиной PCI и автоматической конфигурацией системы Plug and Play.

Задание полной конфигурации компьютера осуществляется не только установками из программы Setup, но и замыканием (или размыканием) соответствующих перемычек на системной плате. Назначение каждой из них указано в соответствующей документации. Они определяют тип процессора, наличие внешнего сопроцессора, тактовую частоту, размер внешней кэш-памяти, разрешения локальной шины. Причем подобные перемычки могут использоваться даже на системных платах с автоматической конфигурацией, например для выбора внешней тактовой частоты процессора, значения коэффициента умножения частоты, напряжения питания процессора и т.д.

1.6 Вспомогательные микросхемы

Хотя микропроцессор и память являются главными компонентами системной платы, но сами по себе они — еще не компьютер. Для создания полновесной микропроцессорной системы необходимы вспомогательные микросхемы, такие, как тактовый генератор, таймер, различные контроллеры, буфера адреса и данных и т.п.

1.6.1 Тактовый генератор

Большинство логических элементов компьютера разработано таким образом, чтобы работать синхронно, то есть по определенным тактовым сигналам. Именно генератор тактовой частоты вырабатывает специальные импульсы, служащие отсчетами времени для всех электронных устройств на системной плате. В качестве главного элемента генератора используется кристалл кварца, который на своей резонансной частоте обладает стабильностью 10^{-6} .

Итак, для получения стабильной опорной частоты на системной плате компьютера могут находиться один или два кварцевых генератора. Генерируемая ими тактовая частота используется, в частности, для синхронизации работы микропроцессора и системной шины. Как правило, имеются фиксированные величины тактовых частот для каждой модели микропроцессора. Например, существуют модели микропроцессоров, работающие на внешних тактовых частотах 33, 40, 50, 60, 66, 100, 133 МГц. Повышение частоты свыше установленного предела может привести к возникновению ошибок и сбоям в работе, причем не только микропроцессора. Кстати, заметим, что для многих микропроцессоров существует и нижний уровень ограничения на тактовую частоту. Дело в том, что отдельные узлы микропроцессора могут быть построены по принципу динамической памяти и требовать постоянной регенерации.

Системная шина компьютера может тактироваться сигналом либо CLK2IN, либо CLK2IN/2, либо ATCLK (BUSCLK). Для каналов прямого доступа в память DMA на системной плате используется еще один тактовый сигнал — SCLK, который зависит от частоты сигналов CLK2IN и ATCLK. Для часов же реального времени на системной плате используется свой «часовой» кварц (обычно 32 768 кГц).

1.6.2 Контроллер прерываний

В первых компьютерах IBM PC использовалась микросхема контроллера прерываний I8259 (Interrupt Controller), которая имеет восемь входов для сигналов прерываний (IRQ0-IRQ7). Как известно, в одно и то же время микропроцессор может обслуживать только одно событие, и в этом ему помогает контроллер прерываний, который устанавливает для каждого из своих входов определенный уровень важности — приоритет. Наивысший приоритет имеет линия запроса прерывания IRQ0, а наименьший — IRQ7, то есть приоритет убывает в порядке возрастания номера линии.

В IBM PC/AT восьми линий прерывания оказалось уже недостаточно, и их количество было увеличено до 15 путем каскадного включения двух микросхем контроллеров прерываний I8259. Такое каскадное включение осуществлялось путем соединения выхода второго контроллера ко входу IRQ2 первого. Важно понять следующее: линии прерывания IRQ8-IRQ15 (то есть входы второго контроллера) имеют приоритет ниже, чем IRQ1, но выше IRQ3.

В таблице 2 приведено распределение номеров прерываний по устройствам для большинства IBM PC/AT-совместимых компьютеров.

Для компьютеров, совместимых с PC/AT,

Таблица 2

Таблица прерываний

IRQ	INT	Доступно	Использование
0	08h	нет	Системные часы (18,2 Гц)
1	09h	да	Клавиатура
2	0Ah	да	Второй блок IRQ8-15
8	70h	нет	Таймер (1024 Гц)
9	71h	да	Переадресовано на IRQ2
10	72h	да	-
11	73h	да	-
12	74h	да	-
13	75h	нет	Сопроцессор
14	76h	да	Контроллер винчестера
15	77h	да	-
3	0Bh	да	COM2 или COM4
4	0Ch	да	COM1 или COM3
5	0Dh	да	LPT2
6	0Eh	да	Контроллер флоппи
7	0Fh	да	LPT1

обычно свободны всего четыре линии запроса прерываний: IRQ 10, 11, 12 и 15. Это, разумеется, касается только полноразмерных плат (16 разрядов данных). Для шин ISA и PCI используется одинаковый набор IRQ. Однако для PCI обычно используются прерывания, активные не по фронту, а по уровню, хотя и первая возможность также не исключается. Это позволяет использовать одну линию для передачи нескольких сигналов прерываний. Обычно для устройств PCI автоматически используются оставшиеся прерывания после установки ISA-плат.

Помимо прерываний есть еще один способ вызывать какие-то события внутри PC без прямого вмешательства CPU. Это канал DMA.

1.6.3 Контроллер прямого доступа в память

Пересылка байта информации из одной области памяти в другую или из памяти в порт обычно проходит в два шага. На первом шаге CPU получает этот байт из памяти и записывает в один из своих регистров. На втором шаге он записывает байт в место назначения.

Немного подумав, сразу же можно назвать два основных недостатка подобного подхода. Первый из них заключается в том, что во время перемещения байта CPU ничем другим заниматься не может. Второй — на перемещение байта требуется два шага. Кажется, мелочь, но если нужно переместить не один байт, а целый блок, мелочью это уже не назовешь.

Внутри PC есть несколько типов устройств, которым нужно пересылать именно большие группы байт. Устройство чтения гибких дисков, например. Звуковые карты также очень интенсивно пересылают байты, настолько интенсивно, что часто они используют сразу несколько каналов DMA. Многие сканеры пользуются каналом DMA, и кто знает, какие еще подобные устройства появятся в дальнейшем.

Осознавая важность проблемы, разработчики первого PC решили завести дополнительный микропроцессор, называемый *контроллером прямого доступа к памяти*. Этот микропроцессор занимается тем, что по команде CPU перемещает заданное количество байт из одной последовательности адресов памяти в ука-

занный порт, начиная с определенного адреса памяти.

Он может и принять из порта последовательность байт определенной длины и поместить ее в память, начиная с указанного адреса. Все подобные пересылки происходят через каналы DMA, иначе говоря, каждой пересылкой занимается определенная часть контроллера DMA. Какая именно это будет часть, т.е. какой канал займется этим, решает не сам контроллер, а устройство (CPU или какое-либо устройство ввода-вывода), запросившее доступ к нему. А именно, перед тем как пользоваться DMA-контроллером для пересылки данных, такое устройство обязано сказать контроллеру, какой канал будет им использоваться. Так что каналы DMA — это ресурс, требующий аккуратного обращения, иначе при обращении к ним могут начаться конфликты устройств, которые ни к чему хорошему не приведут.

Разумеется, DMA подверглась улучшению, позволившему более эффективно использовать пропускную способность существующих шин. Новейшая версия, названная UltraDMA, может передавать данные IDE-устройствам со скоростью до 33 Мб/с (мегабайт в секунду), т.е. в два раза быстрее старой версии.

1.6.4 Другие вспомогательные микросхемы

К вспомогательным микросхемам в первую очередь можно отнести таймеры (реализованные раньше на микросхеме i8254) и часы реального времени (MC146818A). В зависимости от типа процессора на системной плате могут располагаться контроллеры шины и памяти, системный и периферийный контроллеры, кэш-контроллер, а также буфера для данных и адресов.

1.6.5 Асинхронные и синхронные компьютеры

Несмотря на всю потрясающую скорость, электронным схемам все-таки требуется некоторое время на выполнение своих операций. Как много — варьируется от типа к типу. По большому счету, можно сделать компьютер вообще без часов. Скорость работы такого компьютера будет полностью зависеть от скорости работы его частей: чем быстрее части, тем быстрее компьютер. Данные, запрошенные из памяти, будут использованы, как только

они появятся там, где их ждут. Результаты вычислений будут отправлены обратно в память, как только они станут доступны. Очень быстрые части, уже выполнившие свою работу, не будут простаивать без дела, дожидаясь, пока стрелки на главных часах разрешат им двигаться дальше. Спроектировать и сделать такой асинхронный компьютер можно, но сложно. Гораздо легче спроектировать компьютер, все части которого работают синхронно с некоторыми главными часами. Если эти часы тикают достаточно медленно, то можно быть уверенным, что все части компьютера смогут закончить назначенные им задачи до поступления сигнала продвинуться еще на один шаг вперед — следующего такта часов. Синхронные компьютеры настолько легче в проектировании и производстве, что практически каждый компьютер являет собой пример подобного подхода.

1.6.6 Часы внутри CPU

Самые знаменитые часы — это часы, тикающие внутри CPU. Это те самые 266 МГц (или 300 МГц, или 1800 МГц, или те, что стоят внутри вашего PC), о которых вы столь много слышите. От них зависит, насколько быстро работают самые быстрые части вашего PC. На сегодня лишь некоторые схемы CPU работают с этой скоростью, остальные части PC даже близко к ней не приближаются, исключая процессор Pentium Pro, кэш L2 которого, расположенный внутри модуля CPU, также работает с этой скоростью. В Pentium II кэш L2 работает на половине частоты часов CPU. Во всех остальных процессорах x86 внешний кэш работает с такой же скоростью, как шина памяти.

Шина от CPU к основной памяти работает со скоростью, в несколько раз меньшей скорости часов внутри CPU. Или, иными словами, CPU работает со скоростью, в несколько раз превышающей скорость внешней шины. Микросхема часов, контролирующая эту скорость, расположена вне CPU, и CPU просто синхронизирует свои не слишком точные часы по сигналу внешних часов.

Стандартная линейка частот тактовых генераторов: 50, 60, 66, 100, 133 МГц.

Те же часы, что управляют CPU, также управляют модулями

основной памяти и сопутствующими им микросхемами. Часто только внешняя (L2 или L3) память реально способна успевать за этими быстрыми часами. Быстрые микросхемы DRAM можно заставить работать медленнее, вставив одно или несколько *состояний ожидания* (wait states). Это паузы между интервалом, в течение которого CPU или внешний кэш-контроллер спрашивает что-либо у микросхемы памяти, и интервалом, в течение которого они ожидают получения ответа на свой запрос.

Многие подсистемы PC должны работать синхронно с часами, идущими с другой частотой. Например, частота перемещения электронного луча вдоль и вниз по сетке монитора в процессе развертки изображения устанавливается Пользователем (в пределах возможного, конечно). Жестким дискам нужны часы с частотой, в определенное количество раз превышающей частоту вращения их магнитных пластин. Модему тоже нужны часы — чтобы передавать и принимать данные согласно скорости соединения. Вообще, сколько различных устройств со своими возможностями и запросами (и микропроцессорами), столько и часов внутри PC. (Никакое устройство не может работать быстрее, чем позволяют его возможности, и в некоторых случаях скорость работы устройства должна поддерживаться на некотором стандартном для этого устройства уровне, не быстрее и не медленнее.)

1.7 Набор микросхем, или chipset

В современных компьютерах, конечно, давно не используются, например, отдельные чипы контроллеров, однако все их функции реализованы в микросхемах системных и периферийных контроллеров (это уже сверхбольшие интегральные схемы), которые, тем не менее, обеспечивают со своими предшественниками полную программную совместимость. Да и вообще, системная плата современного компьютера отличается от аналогичной платы более ранней модели в первую очередь тем, что большое количество микросхем средней степени интеграции (MSI, Medium-Scale Integration), на которых были выполнены основные функциональные узлы компьютера, теперь заменены на несколько (одну—четыре) сверхбольших интегральных схем (VLSI, Very Large-Scale Integration). Такие VLSI, реализующие функции

прежних микросхем MSI, называются набором микросхем, или chipset (чипсет). Преимущество применения таких наборов очевидно. Во-первых, они занимают меньше места, во-вторых, потребляют меньший ток и, в-третьих, гораздо надежнее (ведь в общем случае надежность устройства обратно пропорциональна количеству входящих в него микросхем).

Наиболее известные наборы микросхем выпускают сегодня такие компании, как Intel, AMD, UMC, SiS, хотя ими, конечно, дело не ограничивается.

В большинство наборов тем или иным образом входит периферийный контроллер, например микросхема 82C206 или ей подобная. Обычно такая микросхема функционально содержит два контроллера прерываний типа 8259, два контроллера прямого доступа к памяти типа 8237, таймер типа 8254, часы реального времени и более 100 байт CMOS RAM для хранения системной конфигурации.

Поскольку основной задачей чипсета является обеспечение бесперебойной и быстрой работы процессора с периферийными устройствами, вся его структура посвящена выполнению именно этой задачи.

1.8 Системные и локальные шины

Основным функционалом системной шины является передача информации и данных между базовым микропроцессором и остальными электронными компонентами компьютера. По этой шине осуществляется также адресация устройств и происходит обмен специальными служебными сигналами. Таким образом, упрощенно системную шину можно представить как совокупность сигнальных линий, объединенных по их назначению (данные, адреса, управление). Передачей информации по шине управляет одно из подключенных к ней устройств или специально выделенный для этого узел, называемый арбитром шины.

Системная шина IBM PC и IBM PC/XT была предназначена для одновременной передачи только 8 бит информации, так как используемый в компьютерах микропроцессор 18088 имел 8 линий данных. Кроме того, системная шина включала 20 адресных линий, которые ограничивали адресное пространство пределом в 1 Мбайт. Для работы с внешними устройствами в этой шине были преду-

смотрены также 4 линии аппаратных прерываний (IRQ) и 4 линии для требования внешними устройствами прямого доступа в память (DMA, Direct Memory Access). Для подключения плат расширения использовались специальные 62-контактные разъемы. Системная шина и микропроцессор синхронизировались от одного тактового генератора с частотой 4,77 МГц. Таким образом, теоретически скорость передачи данных могла достигать более 4,5 Мбайта/с.

1.8.1 Шина ISA

Появление PC/AT, использующего микропроцессор i80286 с 16 разрядной шиной данных, потребовало разработки новой системной шины ISA (Industry Standard Architecture), полностью реализующей возможности упомянутого микропроцессора. Она отличалась наличием дополнительного 36-контактного разъема для соответствующих плат расширения. За счет этого количество адресных линий было увеличено на четыре, а данных — на восемь. Теперь можно было передавать параллельно уже 16 разрядов данных, а благодаря 24 адресным линиям — напрямую обращаться к 16 Мбайтам системной памяти. Количество линий аппаратных прерываний в этой шине было увеличено с 7 до 15, а каналов DMA — с 4 до 7. Надо отметить, что новая системная шина ISA полностью включала в себя возможности старой 8-разрядной шины, то есть все устройства, используемые в PC/XT, могли без проблем применяться и в PC/AT 286. Системные платы с шиной ISA уже допускали возможность синхронизации работы самой шины и микропроцессора разными тактовыми частотами, что позволяло устройствам, выполненным на платах расширения, работать медленнее, чем базовый микропроцессор. Это стало особенно актуальным, когда тактовая частота процессоров превысила 10—12 МГц. Теперь системная шина ISA стала работать асинхронно с процессором на частоте 8 МГц. Таким образом, скорость передачи могла достигать 16 Мбайт/с.

1.8.2 Шина EISA

С появлением новых микропроцессоров, таких, как i80386 и i486, стало очевидно, что одним из вполне преодолимых пре-

пятствий на пути повышения производительности компьютеров с этими микропроцессорами является системная шина ISA. Дело в том, что возможности этой шины для построения высокопроизводительных систем следующего поколения были практически исчерпаны. Новая системная шина должна была обеспечить больший возможный объем адресуемой памяти, 32-разрядную передачу данных, в том числе и в режиме DMA, улучшенную систему прерываний и арбитраж DMA, автоматическую конфигурацию системы и плат расширения. Такой шиной для IBM PC-совместимых компьютеров стала EISA (Extended Industry Standard Architecture). В EISA-разъем на системной плате компьютера помимо, разумеется, специальных EISA-плат могла вставляться либо 8-, либо 16-разрядная плата расширения, предназначенная для обыкновенной PC/AT с шиной ISA. Это обеспечивалось простым, но поистине гениальным конструктивным решением. EISA-разъемы имеют два ряда контактов, один из которых (верхний) использует сигналы шины ISA, а второй (нижний) — соответственно EISA. Контакты в соединителях EISA расположены так, что рядом с каждым сигнальным контактом находится контакт «Земля». Благодаря этому сводится к минимуму вероятность генерации электромагнитных помех, а также уменьшается восприимчивость к таким помехам.

Шина EISA позволяла адресовать 4-Гбайтное адресное пространство, доступное микропроцессорам 180386/486. Стандарт EISA поддерживал многопроцессорную архитектуру для «интеллектуальных» устройств (плат), оснащенных собственными микропроцессорами. Поэтому данные, например, от контроллеров жестких дисков, графических контроллеров и контроллеров сети могли обрабатываться независимо, не загружая при этом основной процессор. Теоретически максимальная скорость передачи по шине EISA в так называемом пакетном режиме (burst mode) могла достигать 33 Мбайт/с. В обычном (стандартном) режиме она не превосходила, разумеется, известных значений для ISA.

На шине EISA предусматривался метод централизованного управления, организованный через специальное устройство — системный арбитр. Таким образом поддерживается использование ведущих устройств на шине, однако возможно также предоставление шины запрашивающим устройствам по циклическому

принципу.

Для компьютеров с шиной EISA было предусмотрено автоматическое конфигурирование системы. Каждый изготовитель плат расширения для компьютеров с шиной EISA поставлял вместе с этими платами и специальные файлы конфигурации. Информация из этих файлов использовалась на этапе подготовки системы к работе, которая заключается в разделении ресурсов компьютера между отдельными платами. Для «старых» плат адаптеров пользователь должен сам подобрать правильное положение DIP-переключателей (рис. 5) и перемычек, однако сервисная программа на EISA-компьютерах позволяла отображать установленные положения соответствующих переключателей на экране монитора и давала некоторые рекомендации по правильной их установке. Помимо этого в архитектуре EISA предусматривалось выделение определенных групп адресов ввода-вывода для конкретных слотов шины — каждому разъему расширения отводится адресный диапазон 4 Кбайта, что также позволяло избежать конфликтов между отдельными платами EISA.

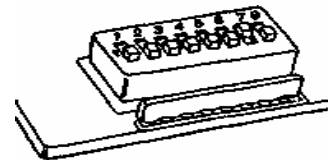


Рис. 5 — Модуль переключателей

Компьютеры, использовавшие системные платы с шиной EISA, были достаточно дороги. К тому же шина по-прежнему тактировалась частотой около 8–10 МГц, а скорость передачи увеличивалась в основном благодаря увеличению разрядности шины данных.

1.8.3 Локальные шины

Разработчики компьютеров, системные платы которых основывались на микропроцессорах i80386/486, стали использовать отдельные шины для памяти и устройств ввода-вывода, что позволило максимально задействовать возможности оперативной памяти, так как именно в данном случае память может работать с наивысшей для нее скоростью. Тем не менее при таком подходе вся система не может обеспечить достаточной производительности, так как устройства, подключенные через разъемы расширения, не могут достичь скорости обмена, сравнимой с процессором. В основном это касается работы с контроллерами накопителей и видеоадаптерами. Для

решения возникшей проблемы стали использовать так называемые локальные (local) шины, которые непосредственно связывают процессор с контроллерами периферийных устройств.

1.8.4 Шины VL-bus и PCI

Практически одновременно, появились две локальные шины, признанные промышленными: VL-bus (или VLB), предложенная ассоциацией VESA (Video Electronics Standards Association), и PCI (Peripheral Component Interconnect), разработанная фирмой Intel. Обе эти шины были предназначены, вообще говоря, для одного и того же — для увеличения быстродействия компьютера, позволяя таким периферийным устройствам, как видеоадаптеры и контроллеры накопителей, работать с тактовой частотой до 33 МГц и выше. Обе шины использовали разъемы типа MCA. На этом, впрочем, их сходство и заканчивалось, поскольку требуемая цель достигалась разными средствами.

Если VL-bus являлась, по сути, расширением шины процессора (вспомним шину IBM PC/XT), то PCI по своей организации более тяготеет к системным шинам, например к EISA, и представляет собой абсолютно новую разработку. Строго говоря, PCI относится к классу так называемых mezzanine-шин, то есть шин «пристроек», поскольку между локальной шиной процессора и самой PCI находится специальная микросхема согласующего «моста» (bridge).

Так как VL-bus продолжает шину процессора без промежуточных буферов, ее схемная реализация оказывается более дешевой и простой. Первая спецификация VESA, в частности, предусматривала, что к шине, которая являлась локальной 32-разрядной шиной системного микропроцессора, могло подключаться до трех периферийных устройств. Некоторые изготовители, впрочем, были убеждены, что добиться устойчивой работы трех устройств на высоких частотах вообще невозможно, и устанавливали на свои платы только 2 слота. Ограничение на число устройств было связано с тем, что электрическая нагрузочная способность на сигнальные линии любого процессора весьма невелика, и, по сути своей, любое новое подключаемое устройство являлось шунтом для предыдущих.

В качестве устройств, подключаемых к VL-bus, выступали контроллеры накопителей, видеоадаптеры и сетевые платы. Конструктивно VL-bus выглядит как короткий соединитель типа MCA (112 контактов), установленный, например, рядом с разъемами расширения ISA или EISA. При этом 32 линии используются для передачи данных и 30 — для передачи адреса. Максимальная скорость передачи по шине VL-bus теоретически может составлять около 130 Мбайт/с. Стоит отметить, что на VL-bus не был предусмотрен арбитр шины. К счастью, большинство подключаемых к ней устройств являлись «пассивными», то есть сами не инициировали передачу данных. Тем не менее во избежание возможных конфликтов между подключенными к шине устройствами в спецификации выделялись «управляющие» (master) и «управляемые» (slave) адаптеры. Для «управляющих» устройств на системных платах обычно были определены свои «мастер-слоты». По замыслу разработчиков подобные «управляющие» устройства могли осуществлять арбитраж на шине.

После появления процессора Pentium ассоциация VESA приступила к работе над новым стандартом VL-bus (версия 2). Он предусматривал, в частности, использование 64-разрядной шины данных и увеличение количества разъемов расширения (предположительно три разъема на 40 МГц и два на 50 МГц). Ожидаемая скорость передачи теоретически должна была возрасти до 400 Мбайт/с.

Спецификация шины PCI обладает несколькими преимуществами перед основной версией VL-bus. Так, использовать PCI можно вне зависимости от типа процессора. Специальный контроллер заботится о разделении управляющих сигналов локальной шины процессора и PCI-шины и, кроме того, осуществляет арбитраж на PCI. Именно поэтому данная шина может использоваться и в иных компьютерных платформах. Следует отметить, что гибкость и быстроедействие этой шины предполагают и большие аппаратные затраты, чем для VL-bus. Тем не менее шина PCI стала практическим стандартом для систем на базе Pentium и не менее успешно используется в других компьютерах, даже и не PC совместимых.

В соответствии со спецификацией PCI к шине могут подключаться до 10 устройств. Это, однако, не означает использова-

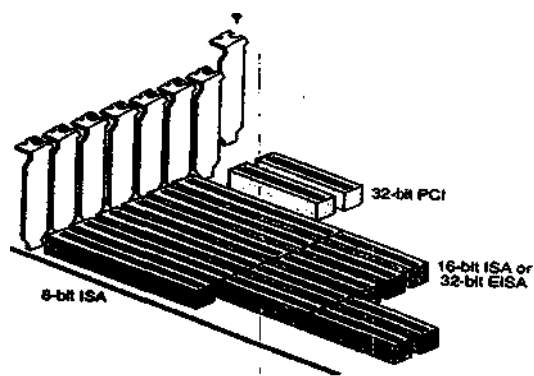


Рис. 6 — Слоты расширения

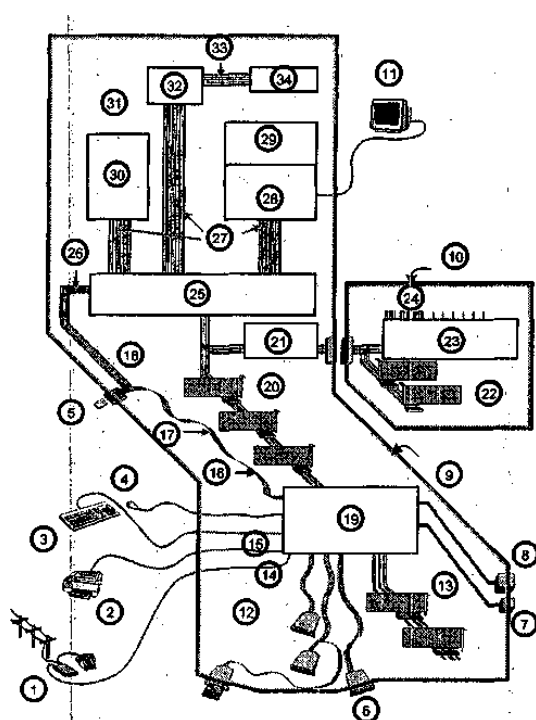
ния такого же числа разъемов расширения — ограничение относится к общему числу компонентов, в том числе расположенных на системной плате. Поскольку каждая плата расширения PCI может разделяться между двумя периферийными устройствами, то уменьшается общее число устанавливаемых разъемов. В отличие от VL-bus шина PCI работает на фиксированной тактовой частоте 33 МГц и предусматривает напряжение питания для контроллеров как 5, так и 3,3 В, а также обеспечивает режим их автоконфигурации (plug and play — «включай и работай»). Заметим, что, например, PCI-карты, рассчитанные на напряжение 5 В, могут вставляться только в соответствующие слоты, которые конструктивно отличаются от слотов для напряжения 3,3 В. Впрочем, имеются и так называемые универсальные PCI-адаптеры, которые работают в любом из слотов. Шина PCI может использовать 124-контактный (32-разрядная) или 188-контактный разъем (64-разрядная передача данных), при этом теоретически возможная скорость обмена составляет соответственно 132 и 264 Мбайта/с.

Примерный вид такой платы приведен на рис 6.

1.8.5 Шина PCI как ось «Север — Юг» в РС

Шина PCI представляет собой мощное средство взаимодействия различных компонентов РС как расположенных внутри системного блока, так и находящихся за его пределами. Intel воспользовалась этим преимуществом, чтобы разделить различные функции, возложенные на чипсет материнской платы, на две группы. Она интегрировала одну группу функций в микросхему, назвав ее Northbridge (Северный мост). Другая группа функций была интегрирована (фирмой Intel и некоторыми из ее конкурентов) в микросхему под названием Southbridge (Южный мост). Эти микросхемы соединяются шиной PCI. Данное соединение (и многое другое) показано на рис. 7.

Тщательный анализ этого рисунка поможет вам глубже понять архитектуру современных РС. На этом рисунке показано, что наш гипотетический современный РС (который может быть как настольным компьютером, так и лэптопом) соединен со «стыковочной» станцией. Смысл такой организации в том, чтобы компьютер сохранял работоспособность при отключении от этой стыковочной станции, но при подключении к ней его возможности существенно возрастали бы. Приведем подробное описание каждого устройства, изображенного на этом рисунке, начав с описания микросхемы Northbridge и новейшего способа подключения видеоподсистемы к центральному процессору.



1. Модем.
2. Принтер.
3. Клавиатура.
4. Мышь
5. Отсек PC Card.
6. Дисководы гибких дисков.
7. USB.
8. IEEE1394.

9. Портативный РС или настольный системный блок.
10. Стыковочная станция.
11. Монитор РС (видео-дисплей).
12. CD-ROM на IDE.
13. Шина и карты ISA 16 Кбит; 8,3 МГц.
14. Посл. порт. 15. Параллельный порт.
16. Шина Card Bus PC-карты 16 Кбит; 8,3 МГц.
17. Шина Zoomed Video 32 Кбита; 33 МГц.
18. Микросхема Southbridge.
19. Интерфейс PCI-ISA.
20. Шина и карты PCI 32 Кбита; 33 МГц.
21. Микросхема моста PCI-PCI.
22. Вторичная шина PCI.
23. Микросхема Southbridge.
24. IDE, FDD и другие порты.
25. Микросхема Northbridge Интерфейс CPU-PCI.
26. 32 Кбита; 33 МГц;
27. 64 Кбита; 66 МГц.
28. Графический акселератор.
29. Videобуфер (RAM).
30. Главная память (RAM и ROM).
32. CPU.
33. 64 Кбита; 300 МГц или 150 МГц.
34. Кэш L2.

Рис. 7 — Архитектура РС

Во всех современных РС соединение центрального процессора с оперативной памятью осуществляется через более быструю и более широкую шину данных, чем это может быть обеспечено при использовании для этого соединения шины PCI. Эта быстрая шина называется *внутренней* (system) или *главной* (host) шиной. Центральный процессор не может непосредственно подключаться к модулям памяти. Для этого, по крайней мере, необходимо использование буферов и декодеров адресов памяти. Точно так же он не может непосредственно подключаться к линиям шины PCI. (Все сигналы, необходимые для обеспечения взаимодействия со всеми остальными компонентами РС, присутствуют на контактах разъемов шины PCI, поэтому Intel решила подключать все эти компоненты к шине PCI, а не прямо к микросхеме Northbridge.

Электронные схемы интерфейса, необходимые для решения этих задач, иногда называемые частью *связующей логики материнской платы* (motherboard glue logic), были интегрированы в один очень большой кристалл интегральной микросхемы (сложность которого сравнима со сложностью центрального процессора). Intel назвала эту микросхему Northbridge. В процессе разработки этой микросхемы у разработчиков возникла идея реализовать на ней поддержку другой, очень быстродействующей шины данных специального назначения. Они назвали ее интерфейс *передовым портом графики* (Advanced Graphics Port, AGP).

Шина AGP представляет собой шину, способную работать с частотой внутренней шины материнской платы, имеющую то же количество линий для передачи данных (в настоящее время это означает параллельную передачу 32 бит данных с тактовой частотой 66 МГц) и осуществляющую соединение микросхемы Northbridge с графическим акселератором. Графический акселератор может быть установлен на материнской плате (как это показано на рис. 7) или реализован в виде съемной карты расширения, вставляемой в специальный разъем AGP.

В этом случае на графическом акселераторе установлен банк локальной памяти, к которому имеет доступ только сам акселератор, он является дополнением к буферу изображения, который все еще логически размещен в диапазоне адресов оперативной памяти.

Кроме того, видеоадаптер может получить высокоскоростной доступ к оперативной памяти РС, используя для/этого шину AGP, микросхему Northbridge и шину оперативной памяти.

1.8.6 Шина L2: обходной путь для повышения производительности центрального процессора

Существует другой способ, позволяющий повысить быстродействие РС. Этим способом является использование специальной шины от центрального процессора до кэш-памяти второго уровня. Например, в процессоре Pentium Pro кэш-память второго уровня работала с тактовой частотой микросхемы центрального процессора. В Pentium II она работает на половине тактовой частоты центрального процессора. В любом случае она работает намного быстрее, чем внутренняя шина системы, шина AGP или шина оперативной памяти.

Шина кэш-памяти второго уровня не только имеет более высокую тактовую частоту, чем шина, связывающая центральный процессор и микросхему Northbridge, она еще и работает независимо от этой шины. Это означает, что различные данные могут одновременно передаваться по каждой из этих шин.

На рис. 7 приведена конфигурация системы с несколькими шинами PCI. Основная шина PCI подключена к микросхеме Northbridge, и в разъемы этой шины установлены три карты расширения в формате PCI. На другом конце этой шины расположена микросхема Southbridge, назначение которой будет вскоре рассмотрено.

Обратите внимание, что у шины PCI есть ответвление, идущее направо. Это ответвление проходит через микросхему «моста от одной шины PCI к другой» к специальному разъему, разработанному для использования в составе порта стыковочной станции. Порт стыковочной станции имеет соответствующий разъем, через который в нее кратчайшим путем передаются сигналы основной шины PCI. Существуют два способа осуществления подобного соединения: один из них предполагает использование дополнительной микросхемы Southbridge; другой из них построен на увеличении числа разъемов шины PCI.

На «южном» конце шины PCI расположена микросхема Southbridge. Эта микросхема содержит всю необходимую интерфейсную логику для передачи сигналов с шины PCI на много более медленной шине ISA и на другие интерфейсы, которые могут использоваться в PC. Шина ISA, показанная на рисунке, имеет всего пару разъемов, поскольку большинство функций, которые раньше были возложены на карты расширения шины ISA, в настоящее время реализуются соответствующими портами или картами расширения в формате PCI.

К микросхеме Southbridge подключаются дисководы гибких дисков. Эта микросхема обеспечивает поддержку двух EIDE-каналов (первичного и вторичного), к каждому из которых может быть подключено до двух EIDE-устройств. Кроме того, на ней реализована отдельная шина клавиатуры, шина мыши, стандартный (возможно ECP или EPP) параллельный порт и один или несколько стандартных последовательных портов. В дополнение ко всем перечисленным интерфейсам на этой микросхеме может быть еще интерфейс шины, позволяющий осуществлять связь с разъемами универсальной последовательной шины (USB) и с разъемами интерфейса IEEE 1394 (Firewire). И, наконец, микросхема Southbridge может иметь интерфейс для отсека подключения PC-карт — в приведенном на рисунке случае интерфейс Card Bus (32-разрядный) с шиной Zoomed Video подключается непосредственно к видеоподсистеме.

Интерфейсы ISA, IDE и PC-карт работают с тактовой частотой 8,33 МГц (это составляет ровно одну четверть тактовой частоты шины PCI). Другие порты работают с еще меньшими тактовыми частотами.

Конечно, главным событием, связанным с развитием PC за прошедшее десятилетие, явилась скорость, с которой росло их быстродействие, повышалась их сложность и снижалась их цена по сравнению с предшествовавшими моделями. Одним из очевидных последствий этого процесса является то, что по мере того, как эти маленькие коробочки становились все более способными и все более доступными, все большее число людей желало их приобрести. Именно поэтому PC в настоящее время превратился в товар массового спроса. И большинство усилий людей, занятых в области разработки и производства PC, направлено на

то, чтобы еще больше людей использовали компьютеры для решения все более широкого круга задач.

Но при этом назревает проблема. Уже сейчас в РС насчитывается дюжина различных интерфейсов ввода-вывода специального назначения. Каждый из них отличается от всех остальных. И большинство из них, если не все, необходимы, чтобы обеспечить обмен РС информацией со всеми аппаратными средствами, которые вы собираетесь к нему подключить.

Создается впечатление, что перед рассветом тьма гуще. Мы сталкиваемся с агрессивной рекламой «объединения» РС с телефонами и другими «бытовыми электроприборами», такими, как видеоманитофоны, телевизоры, радиоприемники, стиральные машины, микроволновые печи, тостеры, — вы сами можете продолжить этот список. Возможно, в недалеком будущем вы сможете купить подобное устройство, которое будет работать под управлением РС или передавать в него информацию. Мы не можем создавать специальный интерфейс для каждого из этих бытовых приборов и при этом не можем использовать для их подключения имеющиеся в настоящее время интерфейсы ввода-вывода информации. Нужно что-то радикально иное. Решение, к которому пришли некоторые очень умные люди, состоит в замене всех существующих интерфейсов ввода-вывода специального назначения на один или два действительно универсальных интерфейса. Да, это один или два *типа* действительно универсальных интерфейсов.

1.8.7 Последовательные шины (USB и IEEE 1394)

Шина SCSI обладает многими полезными свойствами. Шина PCI обладает рядом других полезных свойств. Объединяя лучшие черты каждой из шин и используя накопленный ими опыт, инженеры создали интерфейсы, которые по основным параметрам превосходят все предыдущие до них.

Одним из них является *универсальная последовательная шина* (Universal Serial Bus, USB). Другой стандарт во многом аналогичен первому, но оптимизирован для обеспечения более высоких скоростей передачи данных, в настоящее время этот стандарт называется по имени разработавшего его комитета

по стандартам IEEE 1394. (Фирма Apple Computer, участвовавшая в разработке этого стандарта, дала ему более благозвучное имя «Firewire», но они оформили это название в качестве товарного знака и пока никому не позволяют его использовать.)

У этих двух стандартов, возможно, намного больше общего, чем различий.

Прежде всего, оба они — стандарты последовательного интерфейса. Это означает, что за один такт они передают в канал всего один бит информации. Это кажется медленным, пока вы не узнаете, что в самом медленном режиме скорость передачи составляет 1/2 миллиона бит каждую секунду. В самом быстром режиме скорость передачи превышает тысячу миллионов бит в секунду. Даже если в среднем для передачи одного байта требуется передать примерно 10 бит, такие скорости впечатляют.

Кроме того, к каждой шине можно подключить целый пучок различных периферийных устройств, начиная с клавиатуры, мыши, «интеллектуального» громкоговорителя или «интеллектуального» микрофона и заканчивая видеокамерой, проигрывателем DVD-дисков или жестким диском. Фактически, можно подключать все перечисленные и многие другие устройства одновременно.

И вот тут начинают проявляться различия между этими шинами. Шина USB специально разработана для периферийных устройств, характеризующихся низкой или средней скоростями передачи данных. Например, клавиатура должна посылать в РС несколько бит информации в секунду и получать еще меньшие объемы информации обратно. Она представляет собой типичное медленное устройство. В *интеллектуальный* (smart) громкоговоритель (под которым я подразумеваю устройство, способное принимать цифровые данные и преобразовывать их в звуки) будет поступать гораздо больше бит в секунду, но нескольких сотен тысяч вполне достаточно. Он представляет собой типичное периферийное устройство со средним быстродействием.

Шина USB позволяет обмениваться информацией с этими устройствами в двух режимах. В медленном режиме она осуществляет связь со скоростью 1,5 Мбит/с. В режиме среднего быстродействия она работает в восемь раз быстрее, и скорость передачи данных составляет 12 Мбит/с. Если преобразовать эти скорости передачи данных в байты в секунду, вы получите величины

от 150 до 200 Кб/с (что приблизительно соответствует предельной скорости передачи стандартного последовательного порта) и даже немного выше 1 Мб/с (что близко к предельной скорости передачи параллельного порта в режиме EPP).

С другой стороны, интерфейс IEEE 1394 используется для высокоскоростной передачи больших объемов информации. Минимальная скорость передачи данных по этой шине примерно в восемь раз выше, чем максимальная скорость передачи по шине USB. А предполагаемая максимальная скорость передачи (пока еще достигнутая только в лабораторных условиях, но ожидаемая в настольных системах через несколько лет) может далеко обойти скорость самых быстрых шин SCSI.

Вернемся к тому, что объединяет эти шины: они способны самостоятельно выбирать ширину используемой ими полосы частот. Это означает, что они позволяют любому устройству передавать данные с максимально возможной для него скоростью передачи. Но они оставляют в резерве достаточный запас пропускной способности, чтобы поддерживать обмен данными с другими устройствами на шине. Такое устройство, как видеокамера, которая просто *обязана* передать все имеющиеся данные на шину, поскольку ей их негде хранить, может запрашивать и должно быть уверено в том, что ему будет предоставлен канал передачи данных достаточной пропускной способности, чтобы успеть передать очередной кадр изображения до того, как будет закончено формирование следующего кадра. Другие устройства, которые могут позволить себе ожидание в очереди на обслуживание, могут быть вынуждены ждать в этой очереди, но их потребности в скорости передачи данных будут, в конечном счете, учтены, поскольку общая ширина полосы частот (количество бит, передаваемых в секунду) шины достаточна для передачи всей информации.

1.8.8 Стандарт PCMCIA

Устройства, соответствующие первой версии стандарта PCMCIA, задумывались как альтернатива относительно тяжелым и энергоемким приводам флоппи-дисков в портативных компьютерах. «Загадочная» аббревиатура PCMCIA означает не что иное, как Personal Computer Memory Card International Association. Се-

годня данный стандарт поддерживают уже более 300 производителей. PCMCIA-устройства размером с обычную кредитную карточку являются альтернативой обычным платам расширения, подключаемым к системной шине. Сегодня в этом стандарте выпускаются модули памяти, модемы и факс-модемы, SCSI-адаптеры, сетевые карты, звуковые карты, винчестеры и т.д. Особой популярностью пользуются PCMCIA-карты флэш-памяти, которые не теряют информацию при выключении питания, обладают высоким быстродействием и могут быть использованы в качестве винчестера без движущихся частей.

Стандарт PCMCIA для связи между PC Card и соответствующим устройством (адаптером или портом) компьютера определяет 68-контактный механический соединитель. На нем выделены 16 разрядов под данные и 26 разрядов под адрес, что позволяет непосредственно адресовать 64 Мбайта памяти. Хотя некоторые выводные контакты предназначены для сигналов, необходимых при работе с памятью, эти же контакты могут использоваться и для иных сигналов, рассчитанных на работу с устройствами ввода-вывода. Разумеется, перед этим происходит так называемая переконфигурация выводов.

На стороне модуля PC Card расположен соединитель-розетка (female), а на стороне компьютера — соединитель-вилка (male). Кроме того, стандарт определяет три различные длины контактов соединителя-вилки. Такое решение легко объяснимо. Поскольку подключение и отключение PC Card может происходить при работающем компьютере (так называемая горячая замена), то для того, чтобы на модуль сначала подавалось напряжение питания, а лишь затем напряжение сигнальных линий, соответствующие контакты выполнены более длинными. Понятно, что при отключении PCMCIA-модуля все происходит в обратном порядке.

Помимо габаритных размеров стандарт PCMCIA предписывает размещение переключателя защиты записи, внутреннего источника тока, марки изготовителя, в случае если таковые имеются. Надо отметить, что «теплолюбивые» PC Cards должны нормально функционировать при температуре от 0 до 55 градусов по Цельсию.

1.9 Выбор системной платы

Тип материнской платы определяют прежде всего базовый микропроцессор и системная шина. От типа материнской платы и используемого набора вспомогательных микросхем зависит наличие той или иной локальной шины. Системная плата имеет конечное число не только разъемов расширения, но и установочных мест для строго определенных модулей памяти. Стоит обратить внимание на то, какие разъемы предназначены для модулей памяти (количество выводов, частота шины, другие параметры).

Тип системной платы определяет возможность замены одного микропроцессора на другой, либо обновления только самого микропроцессора, либо всего процессорного модуля (например, с внешней кэш-памятью). Поскольку системные платы типа All-In-One интегрируют на себе как контроллер жесткого диска, так и видеоадаптер, то в минимальной конфигурации все имеющиеся слоты расширения остаются свободными. Более того, использование встроенных адаптеров позволяет избежать проблем, связанных с совместимостью оборудования разных производителей, а также повысить надежность. Ряд системных плат включают в себя, например, встроенный звуковой модуль и/или сетевую карту и даже в ряде случаев встроенный SCSI-адаптер.

Не стоит приобретать системную (да, впрочем, и любую) плату, если вы увидите на ней соединения, выполненные навесными проводниками. Это говорит о ее низком качестве. Некоторые системные платы поддерживали несколько типов процессоров от разных производителей, например Intel и AMD. Такие платы называли обычно универсальными или наращиваемыми (upgradable).

Среди производителей системных плат типа brand name можно отметить, например, такие компании, как Intel, Micronics, Ivill, ASUSTek. Неплохо зарекомендовали себя подобные изделия от фирм Acer, FIC, DataExpert и других.

Что касается системных плат на базе процессоров Pentium или AMD, то здесь замена процессора на более производительный также выполняется относительно просто. Гнездо ZIF позволяет легко удалить старую и вставить новую микросхему. Единственно на что следует обратить особое внимание, это предель-

ная тактовая частота процессора и верная установка соответствующих перемычек.

Таким образом, фирмы-производители материнских плат можно разбить на три категории:

Первая, так называемые «brand name», к ним относятся:

- ASUS Technology (Тайвань).

Их отличительные особенности:

- Специальная микросхема аппаратного мониторинга системы.
- Хорошая сервисная поддержка (обновления на сайте и т.д.).
- Использование более качественных элементов.
- Возможность всегда развести прерывания по слотам расширения.

- ABiT (Тайвань).

- Полностью программная настройка всех параметров, режим «soft menu».

• Наличие резервной микросхемы BIOS (Dual BIOS).

- INTEL.

- Хорошее качество изготовления, хотя и невысокая скорость.

• Продукция указанных фирм отличается высоким качеством изготовления, высокой надежностью и хорошими службами технической поддержки.

Вторая категория — «средний класс»:

- Giga Bite GA , Chain Tech, ACORP и др.

Представители этой категории, как правило, достаточно высокого качества, недорогие.

К третьей категории относятся фирмы, выпускающие продукцию дешевую и не всегда качественную. Это Lucky Star (Formoza), Tomato Board и т.д.

1.10 Накопители

Для хранения программ и данных в IBM PC-совместимых персональных компьютерах используют различного рода накопители, общая емкость которых, как правило, в сотни или тысячи раз превосходит емкость оперативной памяти. По отношению к компьютеру накопители могут быть внешними или встраиваемыми.

мыми (внутренними). В первом случае такие устройства имеют собственный корпус и источник питания, что экономит пространство внутри корпуса компьютера и уменьшает нагрузку на его блок питания. Встраиваемые накопители крепятся в специальных монтажных отсеках (drive bays) и позволяют создавать компактные системы, которые совмещают в системном блоке все необходимые устройства (рис. 8).

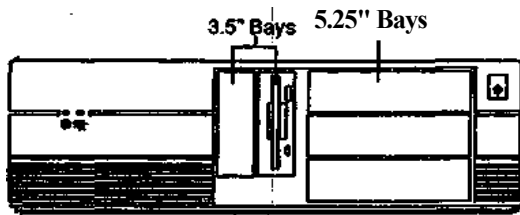


Рис. 8 — Системный блок

Сам накопитель можно рассматривать как совокупность носителя и соответствующего привода. В связи с этим различают накопители со сменным и несменным носителем. В зависимости от типа носителя все накопители можно

подразделить на накопители на магнитной ленте и дисковые накопители. Накопители на магнитной ленте в свою очередь бывают двух видов: накопители на полудюймовых девятидорожечных лентах, работающие в старт-стопном режиме, и стримеры, работающие в потоковом (инерционном) режиме. Накопители на магнитной ленте называют также устройствами последовательного доступа, так как обратиться к удаленным фрагментам данных можно только после считывания менее удаленных. Накопители же на дисках, как правило, являются устройствами произвольного доступа, поскольку интересующие данные могут быть получены без обязательного прочтения им предшествующих.

По способу записи и чтения информации на носитель дисковые накопители можно подразделить на магнитные, оптические и магнитооптические. Среди дисковых накопителей можно выделить:

- накопители на флоппи-дисках;
- накопители на флоптических дисках;
- накопители на несменных жестких дисках (винчестеры);
- накопители на сменных жестких дисках;
- накопители на сменных гибких дисках, использующие эффект Бернулли;
- накопители на магнитооптических дисках;
- накопители на оптических дисках с однократной записью и многократным чтением WORM (Write Once Read Many);

- накопители на оптических компакт-дисках CD-ROM (Compact Disk ROM).

Сразу отметим, что мир накопителей со сменным носителем гораздо шире и многообразнее остальных. Появились сменные винчестеры, которые чаще всего используются в портативных компьютерах. Такое многообразие сменных накопителей связано, видимо, с несколькими причинами. Во-первых, каждый пользователь персонального компьютера знает: какова бы ни была емкость винчестера, наступит время, когда он заполнится до отказа. С другой стороны, чисто теоретически емкость накопителя со сменным носителем, вообще говоря, не имеет предела. Во-вторых, довольно остро стоит проблема архивирования и резервного копирования накапливаемой информации. Исторически она решается с использованием сменных носителей. В-третьих, поскольку IBM PC-совместимый компьютер все-таки персональный, то довольно часто требуется определенный уровень защиты используемых данных. Разумеется, сменные носители — наиболее подходящее средство для обеспечения секретности при хранении частной, служебной и иной закрытой информации. В-четвертых, в ряде случаев с помощью сменных носителей вопрос переноса нескольких единиц, десятков и даже сотен мегабайт данных решается довольно просто. Это, конечно, далеко не все имеющиеся причины, и при желании к ним можно добавить еще несколько.

1.10.1 Винчестеры

За редким исключением практически все IBM PC-совместимые компьютеры имеют в своем составе хотя бы один накопитель на жестких несменных дисках, иначе называемый винчестером. История появления этого устройства такова. В 1973 году на фирме IBM по новой технологии был разработан первый жесткий диск, который мог хранить до 16 Кбайт информации. Поскольку этот диск имел 30 цилиндров (дорожек), каждая из которых была разбита на 30 секторов, то поначалу ему присвоили незамысловатое название — 30/30. По аналогии с автоматическими винтовками, имеющими калибр 30/30, такие жесткие диски получили прозвище «винчестер».

1.10.2 Как устроен винчестер

Итак, накопитель содержит один или несколько дисков (platters), то есть это носитель, который смонтирован на оси-шпинделе, приводимом в движение специальным двигателем (часть привода). Скорость вращения двигателя для обычных моделей составляет около 5400 об/мин. Понятно, чем выше скорость вращения, тем быстрее считывается информация с диска (разумеется, при постоянной плотности записи), однако пластины носителя при больших оборотах могут физически просто разрушиться. Тем не менее в современных моделях винчестеров скорость вращения достигает уже 5400, 7200, 10 000 и даже 15 000 об/мин.

Сами диски представляют собой обработанные с высокой точностью керамические или алюминиевые пластины, на которые нанесен специальный магнитный слой (покрытие). Надо отметить, что за последние годы технология изготовления этих деталей ушла далеко вперед.

Количество дисков может быть различным — от одного до пяти и выше, число рабочих поверхностей при этом соответственно в два раза больше, правда не всегда. Иногда наружные поверхности крайних дисков или одного из них не используются для хранения данных, при этом число рабочих поверхностей уменьшается и может оказаться нечетным (рис. 9).

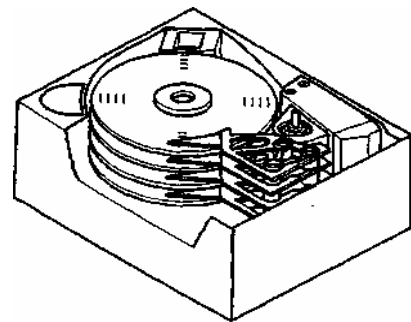


Рис. 9 — Жесткий диск

Наиболее важной частью любого накопителя являются головки чтения-записи (read-write head). Как правило, они находятся на специальном позиционере, который напоминает рычаг звукоснимателя на проигрывателе грампластинок (тонарм) (рис. 10). Это и есть вращающийся позиционер головок (head actuator). Существуют также линейные позиционеры, по принципу движения напоминающие тангенсальные тонармы.

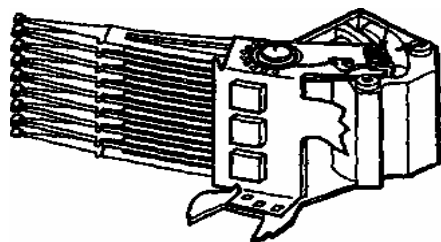


Рис. 10 — Позиционер

Заметим, что в современных винчестерах головки как бы «летят» на расстоянии доли микрона (обычно около 0,13 мкм) от поверхности дисков, разумеется, не касаясь их. Кстати, в жестких дисках выпуска 1980 года это расстояние составляло 1,4 мкм, в перспективных же моделях ожидается его уменьшение до 0,005 мкм. На первых моделях винчестеров позиционер головок перемещался обычно с помощью шагового двигателя. В настоящее время для этой цели используются преимущественно линейные (типа voice coil, или «звуковая катушка») двигатели, иначе называемые соленоидными. К их преимуществам можно отнести относительно высокую скорость перемещения, практическую нечувствительность к изменениям температуры и положениям привода. Кроме того, при использовании соленоидных двигателей реализуется автоматическая парковка головок записи-чтения при отключении питания винчестера. В отличие от накопителей с шаговым двигателем не требуется периодическое переформатирование поверхности носителя.

Кроме вышеперечисленного, внутри любого винчестера находится печатная плата с электронными компонентами, которые необходимы для нормального функционирования устройства привода. Так, например, электроника расшифровывает команды контроллера жесткого диска, стабилизирует скорость вращения двигателя, генерирует сигналы для головок записи и усиливает их от головок чтения и т.п. В настоящее время в ряде винчестеров применяются даже цифровые, сигнальные процессоры DSP (Digital Signal Processor).

Непременными компонентами большинства винчестеров являются специальные внутренние фильтры. По понятным причинам большое значение для работы жесткого диска имеет чистота окружающего воздуха, поскольку грязь или пыль могут вызвать соударение головки с диском, что однозначно приведет к выходу его из строя (рис. 11). Габаритные размеры современных винчестеров характеризуются так

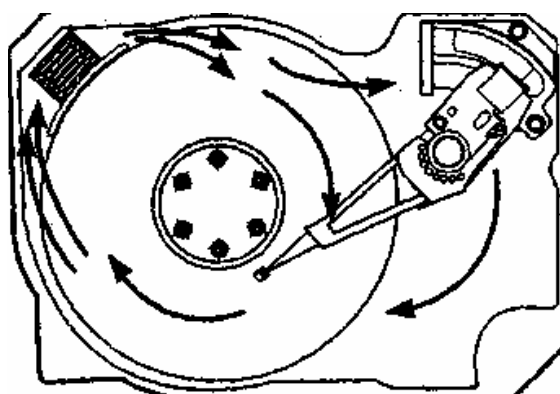


Рис. 11 — Воздушные потоки внутри жесткого диска

как работы жесткого диска имеет чистота окружающего воздуха, поскольку грязь или пыль могут вызвать соударение головки с диском, что однозначно приведет к выходу его из строя (рис. 11). Габаритные размеры современных винчестеров характеризуются так

называемым форм-фактором. Как правило, форм-фактор указывает горизонтальный и вертикальный размеры винчестера. В настоящее время горизонтальный размер жесткого диска может быть определен одним из следующих значений: 1,8; 2,5; 3,5 дюйма (действительный размер корпуса винчестера, разумеется, чуть больше). Вертикальный размер характеризуется обычно такими параметрами, как Full Height (FH), Half-Height (HH), Third-Height (или Low-Profile, LP). Винчестеры «полной» высоты имеют вертикальный размер более 3,25 дюйма (82,5 мм), «половинной» — 1,63 дюйма и «низкопрофильные» — около 1 дюйма. Необходимо помнить, что для установки привода, имеющего меньший форм-фактор, чем монтажный отсек в системном блоке, придется использовать специальные крепления.

1.10.3 Методы хранения информации

Для начала вспомним некоторые основы магнитной записи. Цифровая информация (в виде нулей и единиц) преобразуется в переменный электрический ток, который, как известно, сопровождается переменным магнитным полем, и уже этот переменный ток подается на магнитную головку записи-чтения. Поскольку магнитное покрытие диска представляет собой множество мельчайших областей спонтанной намагниченности (доменов), то под воздействием внешнего магнитного поля, создаваемого головкой, собственные магнитные поля доменов ориентируются в соответствии с его направлением. После снятия внешнего поля на поверхности дисков в результате записи информации остаются зоны остаточной намагниченности в форме концентрических окружностей — это и есть магнитные дорожки. Совокупность таких дорожек, расположенных друг над другом на всех рабочих поверхностях дисков, называют цилиндром. Все концентрические дорожки разбиты на дуги — так называемые сектора, причем сектор является одной из основных единиц записи информации на жесткий диск.

Понятно, что производители винчестеров заинтересованы в размещении как можно большего числа битов на одной дорожке. Здесь им на помощь приходят различные методы кодирования

и *записи* данных. Наиболее распространенным когда-то был способ магнитной записи — модифицированной частотной модуляции (MFM, Modified Frequency Modulation). В соответствии с ним в простейшем случае каждому изменению знака намагниченности на противоположный присваивалось значение бита данных. Другие методы записи, использующие так называемое групповое RLL-кодирование (Run Length Limited), оказались более эффективными по плотности записи и в настоящее время практически вытеснили метод MFM.

Все дорожки магнитного диска на внешних цилиндрах больше, чем на внутренних. Соответственно, при одинаковом количестве секторов на каждой из них плотность записи на внутренних дорожках должна быть больше, чем на внешних. Этот процесс называется прекомпенсацией. В современных винчестерах стал использоваться метод зонно-секционной записи, когда все пространство диска делится на несколько зон, причем во внешних зонах секторов размещается больше, чем во внутренних;. Это, в частности, позволило увеличить емкость жестких дисков примерно на 30%.

В современных винчестерах существует несколько алгоритмов для автоматического замещения дефектных секторов, появляющихся в процессе эксплуатации диска. Один из них заключается в том, что на каждом цилиндре имеется несколько резервных секторов и для каждой зоны — несколько запасных цилиндров. В этом случае сбойный сектор или дорожка могут заменяться на резервные путем переадресации, то есть записи в служебные поля адреса свободного сектора или дорожки. Второй метод использует создание специальной таблицы перекодировки. В этом случае обращения к сбойным секторам вообще не происходит, так как они исключаются из поля доступных секторов диска.

В некоторых винчестерах информация о заголовках секторов (ID или Header) хранится не на поверхности диска, а в специальной полупроводниковой памяти. Благодаря этому повышается полезная емкость диска и, кроме того, по утверждению фирмы IBM (которая и является автором этого нововведения), увеличивается скорость обмена данными и уменьшается время доступа.

1.10.4 Важные параметры

Помимо своих геометрических размеров (форм-фактора) винчестеры да и практически все накопители характеризуются такими параметрами, как емкость, среднее время доступа к данным, скорость передачи данных, среднее время безотказной работы.

Емкость винчестера может указываться как до, так и после форматирования. В последнем случае она, разумеется, несколько меньше. Измеряется емкость в мегабайтах.

Среднее время доступа определяет временной интервал, в течение которого накопитель находит требуемые данные. Это время обычно представляет собой сумму времени, необходимого для позиционирования головок на нужную дорожку и ожидания требуемого сектора. Как правило, эти параметры называют временем поиска и временем латентности. Измеряется данная величина в миллисекундах. Заметим, что среднее время доступа только примерно отражает действительное быстродействие накопителя при работе с тем или иным программным приложением.

Для накопителей могут указываться как внутренняя (от носителя к встроенному интерфейсу привода), так и внешняя скорость передачи данных (от накопителя к системной локальной шине). Последняя величина, разумеется, существенно ниже. В зависимости от типа интерфейса скорость определяется либо в мегабитах, либо в мегабайтах за секунду.

Среднее время безотказной работы MTBF (Mean Time Between Failure) вычисляется обычно как статистическая величина. Берется, допустим, 1000 винчестеров, которые работают по 24 часа в течение месяца. Зная число отказавших за этот месяц винчестеров и общее время работы, определяют MTBF. Таким образом, понятно, как определяется данный параметр для накопителей со временем безотказной работы 200 тысяч часов (то есть более 20 лет).

1.11 Интерфейсы накопителей

Для подключения накопителей к IBM PC-совместимому компьютеру в настоящее время используется либо интерфейс IDE (Integrated Drive Electronics), или ATA (AT Attachment), либо SCSI (Small Computer System Interface), либо оба интерфейса вместе.

1.11.1 Интерфейс ST-506/412

Массовое применение жестких дисков типа винчестер началось только после того, как в 1980 году фирма Shugart Technology (сегодня Seagate Technology) выпустила устройство ST-506 с форм-фактором FH (Full Height) 5,25 дюйма. Форматированная емкость этого винчестера была всего 5 Мбайт, а для связи с компьютером он использовал интерфейс того же названия (ST-506), разработанный фирмой Shugart Technology в конце 70-х годов. Впрочем, «корни» ST-506 можно было обнаружить в двух других интерфейсах: SA450 — для флоппи-дисков и SA1000 — для 8-дюймовых жестких дисков. Как и интерфейс SA450, ST-506 использовал 34-жильный кабель управления (по принципу «дейзи-цепочки», то есть в режиме управляющий-управляемый), а от SA1000 остались 20-жильные «радиальные» кабели для передачи данных индивидуально для каждого накопителя (рис. 12). Впрочем, у оригинального интерфейса ST-506 имелся один существенный недостаток. При поиске нужной дорожки каждая выполняемая приводом команда передвигала головку чтения-записи только на один шаг (подобный принцип работы используется у дисководов флоппи-дисков).

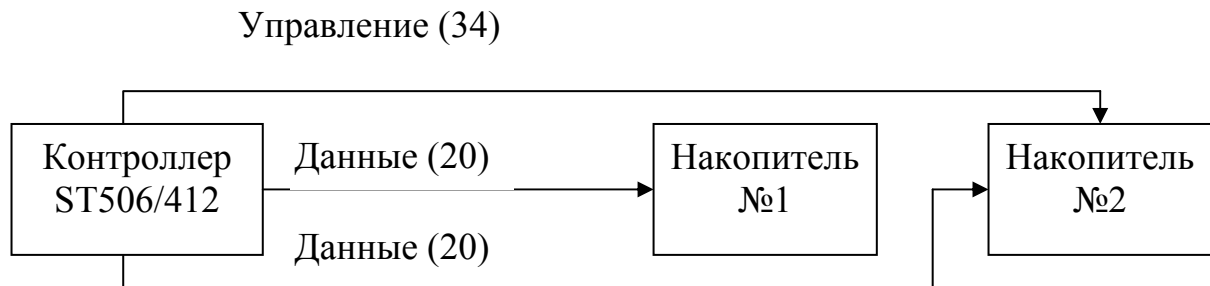


Рис. 12 — Интерфейс ST 506/412

В новом 10-Мбайтном винчестере ST-412, появившемся в 1981 году, эта проблема была решена введением так называемого буферизированного поиска (buffered seek). Теперь за одну команду мог выполняться «длинный» поиск дорожки, например через всю поверхность диска. Кстати, именно этот винчестер выбрала фирма IBM для своего персонального «детища» PC/XT.

Электроника, расположенная на винчестере с интерфейсом

ST506/412, была достаточно проста. Основную нагрузку по обработке данных выполнял сам контроллер. Как уже говорилось, связь между контроллером и винчестером осуществлялась через два плоских кабеля: 20-жильный — для передачи данных и 34-жильный — для управляющих сигналов. Простой набор этих сигналов (Direction In, Step, Head Select и т.д.) затруднял использование накопителей большой емкости. Скорость вращения шпинделя диска винчестера составляла 3600 оборотов в минуту.

Для первых винчестеров ST506/412 применялся способ модифицированной частотной модуляции (MFM), который позволял записывать 17 стандартных 512-байтных секторов на одну дорожку винчестера. Максимально возможная скорость передачи данных, достигаемая в этом случае, подсчитывается очень просто. Если бы контроллер винчестера успевал читать последовательно один за другим сектора дорожки, она составила бы $(17 \times 512 \times 8 \times 3600) : 60 = 4\,177\,920$ Мбит/с, то есть около 5 Мбит/с.

Хотя метод MFM-кодирования относительно прост, надежен и не требует больших затрат при своей реализации, он является далеко не самым лучшим по плотности хранения информации. Дело в том, что помимо информационных необходимы биты синхронизации, которые должны храниться наряду с информационными. Заметим, что сигналы, включающие в себя данные и биты синхронизации, передаются по кабелю в аналоговом виде. Разделение этой информации происходит в специальном устройстве — сепараторе, который для интерфейса ST-506/412 находится в контроллере.

Вообще говоря, метод MFM-кодирования позволяет записывать от одного до трех бит данных на один переход намагниченности. Поиск путей повышения плотности записи на винчестере был связан со стабилизацией вращения диска и улучшением качества его рабочих поверхностей. Благодаря этому новый метод кодирования, впервые предложенный фирмой IBM, — 2,7 RLL (или просто RLL) — позволил увеличить емкость дисков почти в 1,5 раза, а скорость передачи данных возросла (количество секторов на дорожку — 26): $(512 \times 26 \times 8 \times 3600) : 60 = 6\,389\,760$ бит/с. Основа метода RLL состоит в перекодировании исходной группы информации и введении избыточности. Чтобы использовать такие винчестеры, необходимы специальные RLL-контроллеры.

Дальнейшим развитием метода записи 2,7 RLL стал так называемый метод Advanced или Enhanced RLL (ARLL, ERLL или 3,9 RLL). Теперь из названия ясно, что изменение магнитного потока возникает не ранее чем после 3 и не позже чем после 9 бит информации. Этот метод позволяет записывать 31 сектор на дорожку и повысить скорость передачи информации до значения $(512 \times 31 \times 8 \times 3600) : 60 = 7\,618\,560$ бит/с.

Максимальные значения емкости винчестеров с интерфейсом ST506/412 составляют для кодирования MFM — 152, для кодирования RLL — 233 Мбайта.

1.11.2 Интерфейс ESDI

Уже в 1985 году в немалой степени благодаря усилиям около 40 компаний, в том числе фирмы Maxtor, появилась первая версия спецификации ESDI, которая определяла, по сути, улучшенную версию интерфейса ST-506/412. Типы разъемов и кабелей, используемых в новом интерфейсе, на первый взгляд ничем не отличались от используемых ранее, однако, разумеется, только на первый взгляд. Одно из существенных отличий интерфейса ESDI состояло в том, что сепаратор данных располагался теперь не на контроллере, а на самом приводе. Такой, казалось бы, простой «перенос» этого устройства давал два больших преимущества. Во-первых, поскольку на сигналах уже никак не сказывались искажения в кабеле, то сепаратор мог быть достаточно точно настроен на индивидуальные характеристики конкретного привода. И во-вторых, отказ от использования передачи низкоуровневых аналоговых сигналов позволил повысить скорость передачи по сравнению с интерфейсом ST-506/412 в два раза — до 10 Мбит/с (теоретически до 24 Мбит/с). Несомненным преимуществом нового интерфейса являлась также возможность подключения жестких дисков большого объема (сотни Гбайт) и оптических накопителей.

1.11.3 Интерфейс SCSI

Интерфейс SCSI был разработан в конце 70-х годов и предложен организацией Shugart Associates первоначально под названием SASI (Shugart Associates System Interface). После стандарти-

зации этого интерфейса в 1986 году уже под «именем» SCSI (читается «скази») он стал одним из важнейших промышленных стандартов для подключения «разумных» периферийных устройств, таких, как винчестеры, стримеры, сменные жесткие и оптические диски и т.п.

Интерфейс SCSI не разрабатывался специально для работы с дисковыми устройствами, он представляет собой миниатюрную сеть, построенную в пределах одного компьютера. Для подсоединения устройства любого типа с интерфейсом SCSI (а жесткие диски являются только одним из многих видов устройств, подключаемых посредством интерфейса SCSI) к PC необходимо наличие специальной платы расширения, называемой *адаптером SCSI* (SCSI Host Adapter). Эта карта расширения используется в качестве моста между системной шиной PC и шиной SCSI. Накопители с интерфейсом SCSI, с другой стороны, обладают большей емкостью и скоростью, но за это придется не только заплатить дополнительные деньги, но и преодолеть трудности, связанные с их установкой в PC. Необходимо отметить, что SCSI — интерфейс системного, а не приборного уровня, поскольку протокол определяет только логический и физический уровень. В отличие от последовательных приборных интерфейсов ST506/412 и ESDI (где информация между накопителем и контроллером передается бит за битом) SCSI осуществляет параллельную передачу данных. Это, в частности, позволяет существенно повысить скорость обмена. Контроллеры SCSI применяются не только в IBM PC-совместимых компьютерах, но и других платформах.

По существу, шина SCSI представляет собой две отдельные шины, объединенные мостом. На рис. 13 эта структура отображена графически. Как видно из этого рисунка, управляющий адаптер шины SCSI подключается к системной шине ISA через ее разъем. К этому адаптеру могут подключаться как внутренние (например, жесткий диск), так и внешние устройства, имеющие разъемы шины SCSI. В некотором смысле шину SCSI можно рассматривать как небольшую локальную вычислительную сеть. Основное различие между шиной SCSI и обычной локальной вычислительной сетью (ЛВС) состоит в том, что ЛВС (Local Area Network, LAN) обычно используется для соединения нескольких PC (или, возможно, для подключения их к файловому серверу),

в то время как PC, на котором установлен адаптер шины SCSI, является единственным универсальным компьютером, который может быть подключен к этой шине.

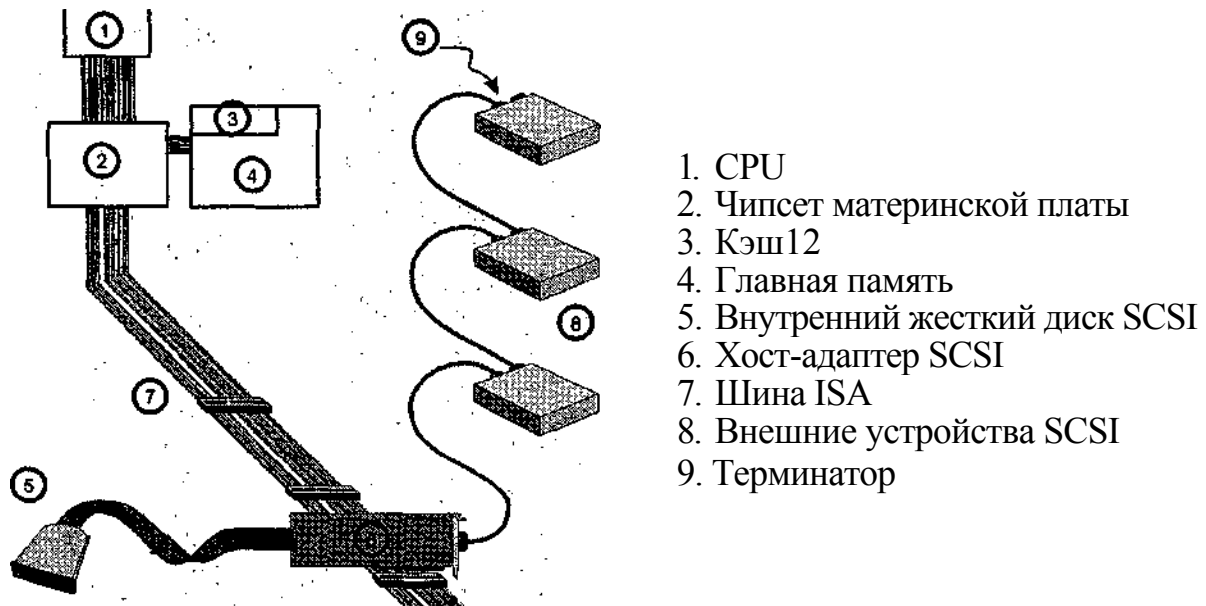


Рис. 13 — Структура шины SCSI

Каждое устройство, имеющее разъем шины SCSI, фактически содержит в своем составе маленький специализированный компьютер, реализующий функции контроллера интерфейса SCSI. Но этот компьютер может решать только одну задачу, а именно, обеспечение связи данного периферийного устройства с адаптером шины, установленным в PC, и, возможно, с другими устройствами, подключенными к шине SCSI.

Следует заметить, что SCSI не накладывает никаких ограничений на связь между контроллером и периферийным устройством. Устройства, подключаемые к шине SCSI, могут выступать в двух ипостасях: Initiator (ведущий, например компьютер) и Target (ведомый, например винчестер), причем одно и то же устройство может быть как ведущим, так и ведомым. В стандарте выделяются четыре схемы подключения устройств: один ведущий и один ведомый, один ведущий и несколько ведомых, несколько ведущих и один ведомый, несколько ведущих и несколько ведомых. К шине одновременно может быть подключено до восьми устройств, в том числе основной (хост) адаптер SCSI. Если необходимо подключить более семи устройств, то следует использо-

вать второй хост-адаптер. Большинство систем позволяет использовать до 4 хост-адаптеров, таким образом, общее количество периферийных устройств достигает 28. Однако на практике не рекомендуется «смешивать» адаптеры различных фирм-производителей.

Хост-адаптер SCSI имеет собственную BIOS (базовую систему ввода-вывода), которая занимает обычно 16 Кбайт в верхней области памяти UMB (Upper Memory Block). Замена системной BIOS позволяет адаптеру работать не с двумя, а с семью приводами (например, Adaptec I1540C). Тем не менее стоит отметить, что некоторые адаптеры ограничиваются поддержкой только двух накопителей. В этом случае, правда, можно воспользоваться специальным программным драйвером.

Для осуществления обмена с процессором адаптер SCSI использует такие системные ресурсы, как порты ввода-вывода, прерывания IRQ и каналы прямого доступа в память DMA.

Физически SCSI-шина представляет собой плоский кабель с 50-контактными разъемами для подключения периферийных устройств. Хост-адаптер, как правило, имеет разъемы для подключения не только встраиваемых, но и внешних SCSI-устройств. Стандарт SCSI определяет два способа передачи сигналов — синфазный и дифференциальный. В первом случае сигналы на линиях имеют ТТЛ-уровни (транзисторно-транзисторной логики) и отрицательную логику, то есть логической единице соответствует низкий уровень напряжения. Длина кабеля в этом случае ограничена 6 м. Версии шины SCSI с дифференциальной передачей сигнала («токовой петлей») дают возможность увеличить длину шины до 25 м. К тому же все «недисковые» SCSI-устройства используют, как правило, интерфейс с синфазными сигналами (общей «землей»). Разумеется, что при использовании устройств как с синфазным, так и с дифференциальным способами передачи сигналов для каждого из них необходим свой хост-адаптер. Хост-адаптер (host adapter) SCSI, который вставляется в разъем системной шины PC, представляет собой аппаратный интерфейс, осуществляющий передачу данных между системной шиной PC и устройствами, подключенными к шине SCSI. В его задачу входит посылка и получение сообщений, посылаемых на языке интерфейса SCSI по шине SCSI, и одновременно посыл-

ка и получение сообщений по системной шине PC, использующей свой протокол обмена информацией.

Поскольку поддержка функциональных возможностей интерфейса SCSI не была изначально предусмотрена в PC, в BIOS материнской платы нет программного обеспечения для реализации этой функции. Поэтому, чтобы адаптер шины SCSI мог работать, на нем должна быть установлена микросхема ROM, в которой записана его собственная BIOS, или, прежде чем вы сможете получить доступ к устройствам, подключенным к шине SCSI, вам придется загрузить соответствующее программное обеспечение с не SCSI-диска.

Чтобы гарантировать качество сигналов на магистрали SCSI, линии шины должны иметь согласование с обеих сторон (наборы согласующих резисторов, или терминаторы). На это следует обращать особое внимание при подключении новых SCSI-устройств. Терминаторы должны быть установлены на хост-адаптере и на последнем устройстве в «гирлянде» магистрали. Обычно используется один из трех методов согласования: пассивное согласование линии при помощи резисторов; улучшенное согласование с исключением перегрузок FPT (Force Perfect Termination), с применением ограничительных диодов; активное согласование, использующее регуляторы напряжения.

SCSI-шина имеет восемь линий данных, сопровождаемых линией четности, и девять управляющих линий. На шине, предназначенной для синфазной передачи, все нечетные контакты должны быть подключены к контакту «корпус» (нулевой потенциал). Многие адаптеры имеют переключки для разрешения или запрещения использования сигнала четности.

Каждое устройство на магистрали SCSI имеет свой адрес (SCSI ID) в диапазоне от 0 до 7. В качестве адреса платы хост-адаптера обычно используется SCSI ID=7.

Обмен между устройствами на магистрали SCSI происходит в соответствии с протоколом высокого уровня. Стандарт SCSI содержит нормированный список команд CCS (Common Command Set). Этот универсальный набор команд обеспечивает доступ к данным с помощью адресации логических, а не физических блоков, как, например, в интерфейсе ESDI. Программное обеспечение для интерфейса SCSI не оперирует физическими ха-

рактическими винчестера (то есть числом цилиндров, головок и т.д.), а имеет дело только с логическими блоками. Понятно, что именно это и дает возможность работать практически с любыми блочными устройствами.

На магистрали SCSI возможны синхронные и асинхронные передачи. Конечно, передача данных в синхронном режиме осуществляется быстрее. Так, скорость передачи данных по шине 8-разрядной SCSI достигает 1,5 Мбайта/с в асинхронном и 3–4 Мбайт/с в синхронном режиме.

Дальнейшим развитием спецификации SCSI стал стандарт SCSI-2. Он предлагал для данного интерфейса большую гибкость и производительность. В спецификацию CSS были включены команды, поддерживающие такие устройства, как приводы CD-ROM, сканеры, коммуникационные устройства, оптические накопители.

Для повышения производительности в спецификацию SCSI-2 был введен так называемый широкий (Wide) вариант шины данных, предусматривающий наличие дополнительных 24 информационных линий. Так появился Wide SCSI-2. В этом случае 8-разрядные устройства продолжают взаимодействовать по магистрали SCSI, использующей 50-проводный кабель (кабель А), а устройства большей разрядности связываются дополнительным 68-проводным кабелем (кабель В) (рис. 14). Надо отметить, что ограничение «широкого» варианта SCSI-2 состоит в том, что он может надежно работать только при реализации дифференциального варианта SCSI.

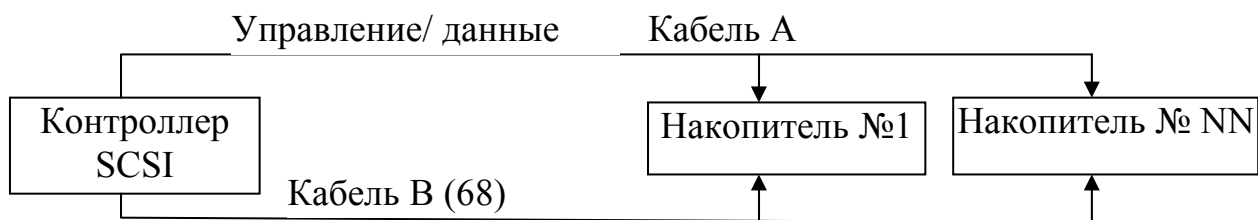


Рис. 14 — Интерфейс SCSI

Для повышения пропускной способности магистрали SCSI было предложено увеличить тактовую частоту обмена примерно в два раза за счет сокращения критических временных парамет-

ров шины, применения новейших БИС, высококачественных кабелей и активного согласования линий. Реализуемый таким образом «скоростной» (Fast) SCSI-2 повысил производительность магистрали до 10 Мбайт/с. Совместное использование Fast и Wide (32-разряда) SCSI-2 теоретически позволяет достичь быстродействия до 40 Мбайт/с. Кстати, на практике вместо кабелей А и В производители используют кабели Р (16- и 32-разрядный Wide SCSI) и Q (32-разрядный Wide SCSI), которые определены только в стандарте SCSI-3. Дело в том, что 68-контактный Р-кабель заменяет кабели А и В для 16-разрядного варианта Wide SCSI-2. Новый вариант интерфейса — Ultra SCSI для 8-разрядной передачи обеспечивает скорость 20 Мбайт/с, а для 16-разрядной — 40 Мбайт/с.

Разумеется, максимальная производительность может быть достигнута, когда плата хост-адаптера реализована для системной EISA-шины и/или локальных VL-bus- и PCI-шин. Устаревшая ISA-шина слишком медленна даже для варианта Fast SCSI-2, поскольку может обеспечить скорость обмена около 2 Мбайт/с. Кстати, одно из преимуществ использования SCSI-винчестеров состоит в том, что они спокойно «уживаются» с другими типами интерфейсов: ST506/412, ESDI, ATA, так как могут использовать отличные от них системные ресурсы компьютера.

Разумеется, не все различия между интерфейсами SCSI и SCSI-2 перечислены здесь. Ряд изменений внесен и в логический протокол, например введена возможность организации очереди команд. Заметим, что окончательная редакция стандарта SCSI-2 была принята только в январе 1994 года, а в 1995 была утверждена уже спецификация SCSI-3. Она отличается возможностью подключать большее количество устройств (более 7) и допускает использование более длинного кабеля. В частности, в спецификации выделен последовательный вариант SCSI. Возможны три варианта его реализации: Serial Storage Architecture, P1394, или FireWare, и Fibre Channel. Последний, кстати, основан на оптических линиях связи. Скорость передачи от 51 Мбита/с до 1 Гбита/с.

Здесь следует также отметить, что существует спецификация ASPI (Advanced SCSI Programming Interface), которую разработала фирма Adaptec — ведущий производитель адаптеров

SCSI. ASPI определяет стандартный Программный интерфейс для хост-адаптера SCSI, то есть позволяет ему общаться с соответствующей операционной системой. Программные модули ASPI (вторичные драйверы) устанавливаются для каждого отдельного устройства. Они обеспечивают взаимодействие периферийных устройств с главным адаптером. Итак, основным программным модулем ASPI является ASPI хост-менеджер. С ним связываются вторичные драйверы ASPI, например для таких устройств, как приводы CD-ROM, оптические и сменные жесткие диски, сканеры и т.д.

В случае если производитель SCSI-устройства поставляет ASPI-совместимый драйвер, гарантируется его совместимость со всеми хост-адаптерами или интерфейсными картами Adaptec и большинства других производителей. К сожалению, в ряде случаев производители (например, приводов CD-ROM) поставляют свою карту контроллера с собственным (несовместимым с ASPI) драйвером, называя интерфейс SCSI. Это следует иметь в виду, если вы хотите подключить к SCSI другие устройства.

Другая спецификация — SCAM (SCSI Configuration Auto Magically) — позволяет упростить настройку SCSI-устройств и скрыть от пользователя некоторые ее детали.

Ведущими производителями адаптеров SCSI помимо Adaptec являются такие фирмы, как Future Domain, Distributed Processing Technology (DPT), NCR и т.д.

1.11.4 Интерфейс ATA

Вообще говоря, первыми устройствами типа IDE были накопители на платах расширения (HardCard), то есть устройства, объединяющие привод и контроллер. Так, в отделении Plus Development компании Quantum было разработано устройство, где на одной плате расширения находились 3,5-дюймовый винчестер с интерфейсом ST506/412 и его контроллер.

Интерфейс IDE (в современном понимании этого термина) был предложен пользователям AT- и XT-совместимых компьютеров летом 1988 года известной сегодня фирмой Conner Peripherals как недорогая альтернатива интерфейсам ESDI и SCSI. В его создании непосредственное участие приняли так-

же фирмы CDC, Compaq Computers и Western Digital. Первый документ CAM ATA (Common Access Method AT Attachment), регламентирующий новый интерфейс, был представлен в 1989 году. Стандартизован ATA-интерфейс был в 1990 году (ANSI X3.221). Этот стандарт определяет временные соотношения сигналов интерфейса, спецификацию кабеля, сигналы на разъеме и т.д. Поскольку ATA-спецификация была долгое время весьма «расплывчатым» документом, а большинство производителей тем не менее выпускали ATA-винчестеры, то с первыми подобными устройствами часто возникали проблемы.

Как уже говорилось, отличительной особенностью нового интерфейса является реализация функций контроллера в самом накопителе. Таким образом, если винчестер имел IDE-интерфейс, это означало, что большая часть компонентов контроллера, совместимого по адресам с интерфейсом ST506/412, расположена вместе с электронной частью винчестера. Кстати, та плата с электронными компонентами, которая обычно включается между системной шиной компьютера и самим накопителем, контроллером, вообще говоря, не является. Как правило, она выполняет функции дешифратора базовых адресов контроллера и формирователя интерфейсных сигналов. Видимо, правильнее называть эту плату адаптером.

Как и любой винчестер, накопитель ATA имеет пакет магнитных дисков, блок магнитных головок, систему позиционирования и канал считывания-записи. Кроме этого, его важнейшими составными частями являются сепаратор данных и однокристалльный микроконтроллер.

Отметим, что вся служебная информация, необходимая для функционирования ATA-винчестеров, записывается на магнитные диски на заводе-изготовителе. Эта информация может быть нескольких типов: сервисная информация, рабочие программы, формат нижнего уровня, паспорт диска и таблица сбойных секторов. В частности, сервисная информация необходима для работы сервосистемы привода магнитных головок. В накопителях ATA используются как выделенные, так и встроенные сервосистемы, тем не менее в современных моделях предпочтение отдается последним. Рабочие программы предназначены для управляющего микропроцессора и представляют собой необходимый набор кодов для работы аппаратуры накопителя. Паспорт диска накопите-

ля АТА содержит справочную информацию о конфигурации и характеристиках накопителя, а также название модели и ее серийный номер. Таким образом, вся информация о «геометрии» АТА-винчестера (количестве цилиндров, головок, секторов) хранится на самом диске. В частности, паспорт диска предназначен для автоматического конфигурирования системы и настройки программного обеспечения при работе с АТА-накопителем.

В накопителях АТА избыточность по емкости скрыта от пользователя. Часть этой избыточной емкости отводится для рабочих программ, паспорта диска и таблицы сбойных секторов. Оставшаяся часть резервируется для замены сбойных секторов. Кстати, заполнение таблицы сбойных секторов производится на заводе-изготовителе при форматировании винчестера. Таким образом, если происходит обращение к сбойному сектору, то накопитель сам переадресует обращение к резервному. По этой причине все новые АТА-накопители не имеют ни одного сбойного сектора.

Подсоединение АТА-винчестера к компьютеру выполняется посредством 40-контактного плоского кабеля. Рекомендуемая длина кабеля не должна превышать 50 см.

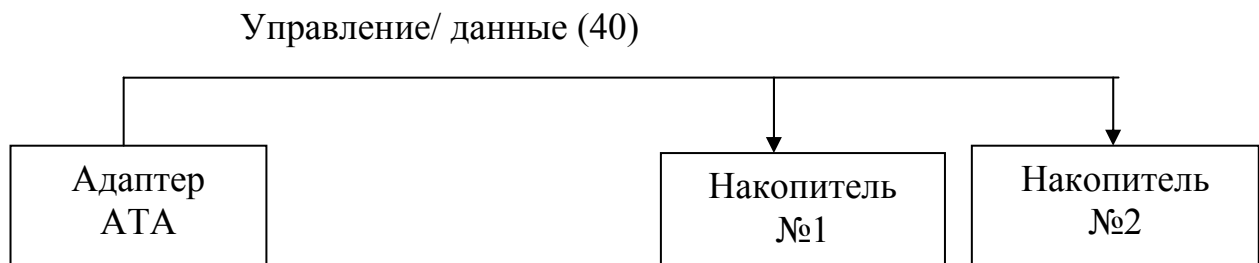


Рис. 15 — Интерфейс АТА

Благодаря тесному взаимодействию пары контроллер-винчестер предусматривается, как правило, ряд особенностей, повышающих производительность винчестера, например использование аппаратной кэш-памяти для получения коэффициента чередования 1:1, режима трансляции физических параметров диска в логические, что позволяет использовать «нестандартные» параметры накопителя, и т.п. Кстати, современные накопители поддерживают так называемый универсальный режим трансляции, когда основным критерием для выбора логических параметров винчестера (количество цилиндров, головок, секторов на дорожке) является

общее количество секторов на носителе. Как известно, под трансляцией понимают замену физических параметров накопителя (число головок, цилиндров, секторов на дорожку) логическими.

Теоретически скорость обмена данными для накопителей АТА составляет от 5 до 10 Мбайт/с. Наивысшей производительностью обладают системы с кэшированными адаптерами и использованием локальных шин. С массовым выпуском малогабаритных компьютеров типа лэптоп и ноутбук большее значение стали иметь размеры как самого винчестера, так и его контроллера, поэтому новая концепция интерфейса стала доминировать и в этой области.

В соответствии со спецификацией к одному разъему АТА можно подключить два винчестера, используя соединение «дэйзи-цепочка» (управляющий — Master, управляемый — Slave) (рис. 15). Правда, все контроллеры допускают подключение четырех устройств. «Трюк» в данном случае состоит в использовании адресов вторичного контроллера.

Так как винчестеры с АТА используют те же адреса, что и винчестеры с интерфейсом ST506/412, то, как правило, их нельзя использовать вместе. Еще одним ограничением АТА-винчестеров является верхний предел их емкости. Так как все операции ввода-вывода данных устройств выполняются через прерывание 13h системной BIOS, общая емкость АТА-винчестеров не может превышать примерно 1 Гбайта (около 504 Мбайт на один накопитель). Чтобы обойти это ограничение, разработчики шли на различные «трюки», которые, например, позволяли представить один физический диск как два.

Использовались и особенности самого интерфейса. Дело в том, что в одном из регистров интерфейса АТА бит номера устройства следует непосредственно за тремя битами номера головки. Следовательно, можно предположить, что подключены не два накопителя, а один, но с 32 головками.

1.11.5 Интерфейс Enhanced IDE и другие

«Мучения», связанные с ограниченной емкостью АТА-винчестеров, закончились, когда компания Western Digital разработала новый интерфейс Enhanced IDE. Его спецификацию под-

держали практически все ведущие компании по производству жестких дисков. Тем не менее, например, фирма Seagate Technology использует для подобного интерфейса собственные названия: Fast ATA и Fast ATA-2. Надо сказать, что основные особенности этих интерфейсов совпадают со спецификацией ATA-2 и новым ANSI-стандартом — ATA-3 (ATAPI). Заметим, что Seagate Technology и Western Digital (как, впрочем, Conner и Quantum) являются членами комитета SFF (Small Form Factor), который, собственно, и дает предложения по стандартизации для ANSI от имени производителей.

Вообще говоря, интерфейс Enhanced IDE имеет четыре основные особенности:

- использование IDE-накопителей емкостью свыше 504 Мбайт,
- более производительные режимы обмена данными,
- подключение к одному адаптеру до четырех устройств,
- поддержка периферийных устройств, отличных от жестких дисков.

Спецификация Enhanced IDE позволяет не только увеличить количество подключаемых устройств, но и использовать другие типы устройств, например приводы CD-ROM или стримеры. В частности, Western Digital для поддержки накопителей CD-ROM с интерфейсом IDE предлагает протокол ATAPI (ATA Packed Interface). ATAPI является расширением протокола ATA и требует незначительных изменений в системной BIOS. В общем случае применяется специальный драйвер.

Если интерфейс Fast ATA поддерживает режимы PIO Mode 3 и Multiword DMA Mode 1, то Fast ATA-2 реализует уже другие новые режимы: PIO Mode 4 с максимальной скоростью обмена 16,6 Мбайт/с и Multiword DMA Mode 2 (16,7 Мбайт/с). В перспективе речь идет уже о PIO Mode 5 (свыше 20 Мбайт/с). Стоит отметить, что единственным требованием спецификаций Fast ATA является только реализация режимов обмена данными, то есть система BIOS, предназначенная для Fast ATA, необязательно должна поддерживать одновременно четыре устройства, в том числе и отличные от жестких дисков, а также емкость свыше 504 Мбайт.

1.11.6 Кэширование диска

Принцип кэширования, используемый для оперативной динамической памяти, во многом похож на принцип кэширования жесткого диска, хотя понятно, что способы доступа к диску и памяти, вообще говоря, сильно различаются. Если время доступа к любой из ячеек оперативной памяти имеет одинаковое, по крайней мере, постоянное для данного компьютера, значение, то время доступа к различным блокам информации на винчестере в общем случае будет, увы, различным. Во-первых, нужно затратить некоторое время, чтобы магнитная головка записи-чтения подошла к искомой дорожке. Во-вторых, поскольку при движении головки вибрирует, то необходимо некоторое время, чтобы она успокоилась. В-третьих, искомый сектор может оказаться под головкой также спустя лишь некоторое время.

При обращении к оперативной памяти могут читаться или записываться только несколько отдельных байт, в то время как доступ к диску всегда происходит секторами. Если размер сектора в случае использования любой версии MS-DOS составляет 512 байт, то наименьший размер кэш-памяти также должен быть 512 байт.

При работе с многозадачными системами выгодно иметь винчестер с мультисегментной кэш-памятью, которая для каждой из задач отводит свою часть кэша (сегмент). В адаптивной мультисегментной кэш-памяти для повышения производительности число и размеры сегментов могут изменяться.

1.11.7 Выбор интерфейса винчестера

Немаловажный вопрос — это интерфейс приобретаемого винчестера. Сразу заметим, что про интерфейс ST506/412 нужно забыть.

Итак, какой из интерфейсов предпочтительней использовать в IBM PC-совместимых компьютерах? Ответ на этот вопрос не так прост, как кажется на первый взгляд. Хотя теоретически интерфейс SCSI может обеспечить скорость обмена несколько выше, нежели АТА, на практике все обстоит немного сложнее. Не следует забывать, например, тот факт, что АТА-интерфейс использует в основном программный ввод-вывод, а SCSI-вин-

честеры в большинстве случаев — передачу данных по прямому доступу к памяти. В однопользовательских системах программный ввод-вывод часто оказывается гораздо эффективнее. Это особенно четко проявляется для каптированных адаптеров при использовании улучшенных алгоритмов кэширования. Преимущество SCSI-адаптеров неоспоримо в первую очередь в многозадачных и многопользовательских системах. Дело в том, что команды для SCSI-устройства могут быть построены в очередь, что освобождает процессор для выполнения других операций.

Использование такого преимущества SCSI-интерфейса, как общение устройств, минуя процессор, часто бывает затруднительно. Как правило, SCSI-устройства разных фирм-производителей «не знают», какой набор команд имеет каждое из них. Несмотря на то, что спецификация SCSI-3 определила минимально необходимый список команд, однако это не дает гарантии полной совместимости. Что же касается стоимости, то большинство современных IBM PC-совместимых компьютеров уже имеют АТА-адаптер, чего не скажешь о SCSI.

Установка нового (или второго) АТА-привода достаточно проста. В большинстве случаев справедлив принцип «включай и работай». Для нормальной работы в файлы конфигурации системы компьютера обычно не требуется добавлять никаких дополнительных программных драйверов. Об этом мы поговорим чуть ниже.

Для SCSI-адаптера процесс установки более сложен. Во-первых, стоит помнить о разделяемых системных ресурсах: портах ввода-вывода, прерываниях IRQ, каналах прямого доступа к памяти DMA, областях в верхней памяти UMB. Во-вторых, требуется верно определить SCSI ID для конкретного устройства, в-третьих, помнить о сигнале четности (запретить или разрешить), установке терминаторов и т.д. Кроме того, файлы конфигурации обязательно должны быть дополнены соответствующими программными драйверами адаптера и устройств. Определенным преимуществом SCSI над АТА является возможность одновременного подключения до семи различных устройств, поскольку АТА-2 только отчасти решает эту проблему. Тем не менее, если вы все-таки планируете приобрести SCSI-контроллер, обратите внимание на поддержку ASPI и SCAM.

Реальные скорости передачи данных по интерфейсам весьма далеки от теоретических значений, поэтому не стоит особенно обольщаться количеством мегабайт в секунду.

1.12 Другие виды накопителей

1.12.1 Флоппи-диски

До настоящего времени приводами для флоппи-дисков оснащается большинство РС-совместимых компьютеров. Они используются как для архивирования и хранения небольших объемов информации, так и для ее переноса с одного компьютера на другой.

История гибкого магнитного (флоппи) диска началась с того момента, когда магнитный слой нанесли на тонкую майларовую основу, подобную той, что используется в магнитной ленте. Чтобы не поцарапать и не испачкать поверхность носителя, диск поместили в достаточно жесткий пластиковый конверт, внутри которого он мог свободно вращаться. Первые флоппи-диски имели диаметр 8 дюймов (около 200 мм) и использовались на больших и мини-компьютерах.

Уже на первых IBM PC использовались приводы для дисков диаметром 5,25 дюйма (133 мм), которые впервые появились в 1976 году. Первоначально на одном таком диске можно было записать всего 160 Кбайт информации, причем магнитный слой был нанесен только с одной стороны основы носителя. После того как магнитный слой стали наносить на пластиковую основу с обеих сторон, емкость носителя удвоилась. Соответственно, привод стал использовать уже две головки. По мере развития технологии стала увеличиваться плотность записи, появились дискеты емкостью 360 Кбайт, а затем и 1,2 Мбайта.

Следующим этапом стали дискеты диаметром 3,5 дюйма (89 мм). Емкость этих «малюток» сначала составляла 720 Кбайт, но вскоре достигла величины 1,44 Мбайта. При такой плотности записи защита магнитного слоя становится особенно актуальной, поэтому сам магнитный диск был спрятан в прочный пластмассовый корпус, а зона контакта головок с его поверхностью закрыта от случайных прикосновений специальной шторкой, которая

отодвигается только внутри накопителя. Вообще говоря, первый 3,5-дюймовый привод и соответствующий микрофлоппи-диск (micro floppy disk) были разработаны в 1980 году фирмой Sony. Несколько позже эта система была принята в качестве стандарта такими организациями, как ISO и ANSI. Кстати, фирма IBM приняла решение использовать 3,5-дюймовые приводы в компьютерах серии PS/2 только в 1987 году.

Привод флоппи-диска (или просто дискеты) во многом похож на привод жесткого диска. Здесь также имеются два двигателя: один обеспечивает стабильную скорость вращения вставленной в накопитель дискеты, а второй перемещает головки записи-чтения. Скорость вращения первого двигателя зависит от типа дискеты и составляет от 300 до 360 об/мин. Двигатель для перемещения головок в этих приводах всегда шаговый. С его помощью головки перемещаются по радиусу от края диска к его центру дискретными интервалами. В отличие от привода винчестера головки в данном устройстве не «летят» над поверхностью флоппи-диска, а касаются ее.

Работой всех узлов привода управляет соответствующий контроллер. Он включает и выключает двигатель вращения диска, задерживает его выключение на несколько секунд для ускорения доступа к данным в случае повторного обращения. Контроллер по индексной метке находит нужную дорожку и устанавливает на нее головку записи-чтения, а также проверяет, закрыт или заклеен вырез в пластиковом конверте диска, запрещая в положительном случае операцию записи.

Стандартным интерфейсом для всех приводов в IBM PC-совместимых компьютерах является SA-400 (Shugart Associates), который был разработан еще в начале 70-х годов. Контроллер соединен с накопителем посредством 34-контактного кабеля. К одному контроллеру обычно подключаются два привода, но вообще-то данный интерфейс позволяет подключать до четырех приводов. Большинство приводов с форм-фактором 3,5 дюйма используют так называемое распределенное согласование. Делается это следующим образом. Общее согласующее сопротивление должно находиться в пределах от 100 до 300 Ом. На контроллере и накопителе с форм-фактором 5,25 дюйма установлены согласующие резисторы номиналом 330 Ом, а на 3,5-дюймовом

приводе — 1500 Ом. Поскольку согласующие резисторы приводов и контроллера включены параллельно, то общее сопротивление определяется по формуле

$$1/R = 1/R_1 + 1/R_2 + 1/R_3 \text{ и т.д.}$$

Итак, IBM PC-совместимые компьютеры оперируют данными, записанными на дискетах размерами 5,25 и/или 3,5 дюйма. Вообще говоря, в современных компьютерах накопители для 5,25-дюймовых дискет уже не устанавливаются. Разумеется, для каждого из типоразмеров дискет существуют специальные приводы соответствующего форм-фактора. Информация на дискеты записывается с двойной плотностью записи методом модифицированной частотной модуляции (MFM).

Некоторая информация о носителях, используемых в IBM PC-совместимых компьютерах, приведена в таблице 3.

Таблица 3

Параметры дисководов для флоппи-дисков

Параметры	Дисководы			
	5,25 дюйма		3,5 дюйма	
	DS/DD	DS/HD	DS/DD	DS/HD
Число дорожек	40	80	80	80
Количество секторов на одну дорожку	9	15	9	18
Число сторон	2	2	2	2
Плотность записи: tpi bpi	48/96	96/1 00	135 8700	135 17500
	5500	9800		
Емкость, Кбайт	360	1200	720	1440
Дескриптор носителя	FDh	F9h	F9h	FOh

Кстати, для того чтобы продлить срок службы приводов флоппи-дисков, рекомендуется хотя бы раз в неделю производить их профилактику, используя для этого специальную «чистящую» дискету; для каждого типоразмера существует своя дискета. В большинстве случаев именно загрязнение магнитных головок приводов приводит к ошибкам при записи или считывании информации с дискеты.

1.12.2 Стримеры

Носители на магнитной ленте применяются в компьютерах еще с начала 50-х годов. В те времена подобные устройства стали приходить на смену «бумажным» носителям информации — перфолентам и перфокартам. Как устройство длительного хранения данных магнитная лента до сих пор сохраняет свое значение. В настоящее время чаще всего применяется несколько типов устройств, использующих в качестве носителя информации магнитную ленту. Более всего известны, конечно, накопители, использующие полудюймовые ленты, намотанные на бобины или катушки (half-inch reel-to-reel). Остальные типы устройств (кстати, более современные) используют для намотки ленты не отдельные бобины, а специальные кассеты — картриджи, или компакт-кассеты. Как известно, полудюймовые ленты, используемые в соответствующих накопителях, называют также девятидорожечными. Накопители на полудюймовых лентах в основном используются в составе мини- и универсальных ЭВМ (mainframes).

Под стримером (streamer) понимается просто лентопротяжный механизм, работающий в инерционном режиме, и не более того. Так что то, намотана лента на бобину или заключена в кассету (картридж), к названию «стример» ровно никакого отношения не имеет. Но поскольку практически во всех лентопротяжных механизмах современных средств резервного копирования используется именно этот режим работы, сами накопители и называют стримерами (иногда «потокowymi» лентами).

1.12.3 Накопители, использующие эффект Бернулли

Специалисты американской фирмы Iomega нашли красивое решение для минимизации расстояния между магнитным слоем носителя и головкой записи-считывания. Они использовали для регулирования этого расстояния известное соотношение Бернулли: давление на поверхность, создаваемое потоком движущейся жидкости или газа, зависит от скорости этого потока. Причем чем быстрее движется газ или жидкость, тем меньше давление на поверхность, то есть больше подъемная сила. Внешне новый носитель данных — Bernoulli Cartridge — выглядит как увели-

ченая до 5,25 дюйма обычная 3,5-дюймовая дискета. Движение воздуха в системе привод-носитель создается благодаря быстрому вращению диска в накопителе. Неподвижный гибкий диск с магнитным носителем прогибается под тяжестью собственного веса и, поскольку он расположен ниже головки, отдаляется от нее. При оптимальной скорости вращения магнитный слой носителя и головку разделяет крошечная прослойка воздуха (три миллионных миллиметра). Головка «летит» над рабочей поверхностью носителя, и, как следствие, отсутствует ее износ. При снижении количества оборотов (например, из-за отключения электропитания), ударах и вибрациях расстояние между поверхностью магнитного носителя и универсальной головкой автоматически увеличивается.

Первые модели накопителей Bernulli были выпущены фирмой Iomega в 1986 году. Тогда каждый сменный диск имел емкость всего 20 Мбайт. Сравнительно быстро миновав рубеж в 90 Мбайт, в 1992 году она достигла уже 150 Мбайт. Сегодня же речь идет о сменных удароустойчивых носителях емкостью примерно 1,5 Гбайт. Как и для многих других накопителей, для устройств типа Bernulli имеется несколько вариантов исполнения: встраиваемое, внешнее, с одиночным и сдвоенным приводами. В качестве интерфейса связи между накопителем и компьютером чаще всего используется SCSI. Однако есть варианты, рассчитанные и на интерфейс IDE.

Довольно популярным типом привода, использующим эффект Бернулли, стал накопитель Zip. Вес этого компактного устройства не превышает 400 г. В зависимости от версии к компьютеру оно может подключаться либо через параллельный, либо через SCSI-интерфейс. Причем скорость передачи данных достигает 1,1 Мбайта/с, емкость 3,5-дюймового картриджа составляет от 100 Мбайт, скорость вращения диска — около 3 тысяч оборотов в минуту, а время доступа — около 30 мс.

1.12.4 Накопители на компакт-дисках

Благодаря малым размерам, большой емкости (до 600 Мбайт) надежности и долговечности компакт-дисков (Compact Disk ROM или CD-ROM) подобные накопители с успехом применяются в

качестве устройств внешней памяти. Стоит отметить, что в последнее время CD-ROM используются как стандартное средство для распространения пакетов программ (CD-ROM Edition). Таким образом, накопители на компакт-дисках стали фактически стандартным устройством для персональных компьютеров.

Процесс изготовления самого компакт-диска состоит из нескольких этапов, включающих в себя подготовку информации для мастер-диска (первых образцов), изготовление мастер-диска и матриц, тиражирование компакт-дисков. Закодированная информация наносится на мастер-диск лазерным лучом, который создает на его поверхности микроскопические впадины, разделяемые плоскими участками. Цифровая информация представляется чередованием впадин (неотражающих пятен) и отражающих свет островков. Копии негатива мастер-диска (матрицы) используются для прессования самих компакт-дисков. Тиражируемый компакт-диск состоит из поликарбонатной основы, отражающего и защитных слоев. В качестве отражающей поверхности обычно используется напыленный алюминий. Диаметр такого диска — либо 5,25, либо 3,5 дюйма. В отличие, например, от винчестеров, дорожки которых представляют концентрические окружности, компакт-диск имеет всего одну физическую дорожку в форме непрерывной спирали, идущей от наружного диаметра диска к внутреннему.

Выпускаются устройства, которые позволяют пользователю самостоятельно записывать (или дописывать) специальные компакт-диски. В отличие от обычных данные диски имеют отражающий слой, выполненный из золота. Это так называемые перезаписываемые компакт-диски CD-R (CD-Recordable). Устройства для записи выпускаются фирмами Pinnacle Micro и Ricoh. Читать дописанные CD-ROM можно на обычных приводах компакт-дисков, удовлетворяющих спецификации CD-ROM XA.

Считывание информации с компакт-диска происходит при помощи лазерного луча, который, попадая на отражающий свет островок, отклоняется на фотодетектор, интерпретирующий это как двоичную единицу. Луч лазера, попадающий во впадину, рассеивается и поглощается — фотодетектор фиксирует двоичный ноль.

Основным стандартом, который определяет логический и файловый форматы записи компакт-дисков, является междуна-

родная спецификация ISO 9660. Имеется и ряд ее расширений, например CD-ROM XA. (extended Architecture). Ряд других стандартов, касающихся компакт-дисков, изложен в документах, называемых «книгами». Так, «Белая Книга» (White Book) определяет основные параметры видео CD — компакт-диска, на котором можно было хранить 72 минуты высококачественного видео вместе со стереозвуком.

В то время как все магнитные диски вращаются с постоянным числом оборотов в минуту, то есть с неизменной угловой скоростью (CAV, Constant Angular Velocity), компакт-диск в своем приводе вращается обычно с переменной угловой скоростью, чтобы обеспечить постоянную линейную скорость при чтении (CLV, Constant Linear Velocity). Таким образом, чтение внутренних секторов осуществляется с увеличенным, а наружных — с уменьшенным числом оборотов. Именно этим обуславливается достаточно низкая скорость доступа к данным для компакт-дисков по сравнению, например, с винчестерами. Для различных моделей она колеблется от 50 до 400 мс.

Скорость передачи данных для привода, определяемая скоростью вращения диска и плотностью записанных на нем данных, составляет не менее 150 Кбайт/с. С появлением мощных малогабаритных двигателей предыдущие модели накопителей начали постепенно вытесняться приводами, которые используют технологию увеличения скорости вращения диска в несколько раз, например в 24, 48, 60. В этих случаях скорость передачи достигает 3600, 7200, 9000 Кбайт/с. Коэффициент увеличения скорости не обязательно является целым числом.

Свой резерв создатели компакт-дисков высокой емкости HD-CD (High Density CD) видят в уменьшении длины волны лазера. Этот же параметр во многом связан со скоростью передачи информации.

1.12.5 Магнитооптические накопители

Видимо, одними из самых жизнеспособных устройств, предназначенных для хранения данных, являются накопители, использующие магнитооптические диски. Здесь были объединены достижения магнитной и оптической технологий. Новые уст-

ройства сочетают портативность флоппи-диска, среднюю скорость работы винчестера, надежность оптического компакт-диска и емкость, сравнимую с кассетой хорошего стримера.

Такие накопители записывают данные, представленные в виде колебаний магнитного поля, на соответствующий носитель при помощи дополнительного магнитного поля (поля смещения) и луча лазера. Поверхность носителя покрыта магнитным материалом (пленкой), который при обычной температуре (из-за высокой коэрцитивной силы) не может быть перемагничен приложенным к нему полем смещения. При нагревании, достигнув температуры Кюри (примерно 145 градусов), коэрцитивная сила ослабевает, и соответствующий участок перемагничивается должным образом. Следует заметить, что запись обычно идет в два приема: стирание информации и затем новая запись. Правда, уже появились новые модели накопителей, осуществляющих запись за один проход. Считывание данных с носителя происходит только при помощи луча лазера, но уже меньшей мощности. Здесь, как правило, используется эффект Керра, заключающийся в изменении плоскости поляризации отраженного луча в зависимости от направления магнитного поля.

У магнитооптического диска в отличие от компакт-диска обычно используются обе стороны. Емкость одного двустороннего носителя может достигать от сотен мегабайт до нескольких гигабайт, причем его геометрические размеры соответствуют флоппи-дискам 3,5 или 5,25 дюйма. Формат нижнего уровня магнитооптического носителя определяют жесткое разбиение на сектора, наличие таблиц дефектных секторов и число логических дисков. Наиболее распространенным стандартом такого формата служит спецификация ISO-ANSI. Диски размером 3,5 дюйма стандартизованы в большей степени и имеют один жесткий формат секторов — 512 байт. Количество секторов постоянно и равно 25. Согласно стандартам ISO/ECMA 3,5-дюймовый диск может хранить 128 или 230 Мбайт. Разумеется, для повышения емкости носителей существуют и специальные системные форматы. Например, на 3,5-дюймовый диск уже записывают до 650 Мбайт, а на 5,25-дюймовый — почти 3 Гбайта. В качестве примера можно привести магнитооптический накопитель Арех фирмы

Pinnacle Micro, который использует 5,25-дюймовый картридж емкостью 4,6 Гбайта. За счет использования однократной записи время поиска составляет около 17,5 мс.

Кстати, производительность магнитооптических накопителей существенно повышается за счет использования встроенной кэш-памяти, объем которой в ряде устройств может достигать нескольких мегабайт.

Так как между головкой и поверхностью самого носителя нет непосредственного контакта, вероятность механического повреждения весьма мала. Стоит также отметить, что магнитооптические носители в отличие от обычных магнитных значительно меньше подвержены влиянию внешних электромагнитных полей. По разным оценкам срок гарантированного хранения информации на магнитооптических носителях колеблется от 30 до 50 лет и более.

Для работы с библиотеками оптических дисков выпускают так называемые дисковые автоматы (jukeboxes), которые обеспечивают автоматический поиск и подачу требуемых дисков. Вместимость такого устройства может составлять от нескольких единиц до сотен оптических дисков общей емкостью 1 Гбайт. Электромеханический привод управляется от компьютера и позволяет выбирать, переворачивать и устанавливать диски в пределах нескольких секунд.

Наиболее известными производителями магнитооптических устройств являются фирмы Fujitsu, Maxoptix, Ricoh и другие.

1.13 Видеоподсистемы

Одной из наиболее важных составных частей любого персонального компьютера является его видеоподсистема, под которой обычно подразумевают монитор, плату видеоадаптера и набор соответствующих программ-драйверов, поставляемых в комплекте с адаптером или в составе прикладных пакетов. Оба устройства (монитор-видеоадаптер) очень плотно взаимодействуют друг с другом, поэтому, говоря об одном из них, часто приходится упоминать и другое.

1.13.1 Принцип работы мониторов

За исключением портативных и LCD, подавляющее большинство современных настольных компьютеров используют мониторы на базе электронно-лучевых трубок (ЭЛТ). Кстати, ЭЛТ была изобретена почти сто лет назад (в 1897 году) немецким ученым Карлом Фердинандом Брауном. Принцип действия мониторов на базе ЭЛТ мало чем изменился с тех пор и заключается в том, что испускаемый электродом (электронной пушкой) пучок электронов, попадая на экран, покрытый люминофором, вызывает его свечение. На пути пучка электронов находятся дополнительные электроды: отклоняющая система, позволяющая изменять направление пучка, и модулятор, регулирующий яркость получаемого изображения. Заметим, что любое текстовое или графическое изображение на экране монитора компьютера (как, впрочем, и телевизора) состоит, вообще говоря, из множества дискретных точек люминофора, называемых также пикселями (pixel — picture element). Поэтому такие дисплеи называют еще растровыми. Электронный луч периодически сканирует весь экран, образуя на нем близкорасположенные строки развертки. Именно этот шаблон и называется растром. По мере движения луча по строкам видеосигнал, подаваемый на модулятор, изменяет яркость определенных пикселей, образуя некоторое видимое изображение. Разрешающая способность монитора определяется числом элементов изображения, которые воспроизводятся по горизонтали и вертикали, например 640x480 или 1024x768 пикселей. Сразу же заметим, что существуют несколько обычных типоразмеров экранов мониторов, используемых для IBM PC-совместимых персональных компьютеров: 9, 14, 15, 17, 19, 20 и 21 дюйм (по диагонали).

Для формирования раstra в мониторе используются специальные управляющие сигналы. В цикле сканирования луч движется по зигзагообразной траектории от левого верхнего угла экрана к нижнему правому. Прямой ход луча по горизонтали осуществляется сигналом строчной (горизонтальной), а по вертикали — кадровой (вертикальной) развертки. Перевод луча из крайней правой точки строки в крайнюю левую точку следующей строки (обратный горизонтальный ход луча) и из крайней правой позиции

последней строки экрана в крайнюю левую позицию первой строки (вертикальный обратный ход луча) осуществляется специальными сигналами обратного хода.

1.13.2 Важные параметры

Итак, наиболее важными для монитора являются следующие параметры: частота вертикальной синхронизации (кадровая развертка), частота горизонтальной синхронизации (строчная развертка) и полоса пропускания видеосигнала. Кадровая частота измеряется обычно в герцах и во многом определяет устойчивость изображения (отсутствие мерцаний). Как известно, человеческий глаз воспринимает смену изображений с частотой выше 20–25 Гц практически как непрерывное движение. Чем выше частота кадров, тем устойчивее изображение. Так, организация VESA (Video Electronics Standards Association) рекомендовала использовать для разрешения 640х480 и 800х600 частоту кадровой развертки не ниже 72 Гц, а для разрешения 1024х768 — 70 Гц. В настоящее время речь уже идет о частоте 100 Гц. Однако повышение частоты требует увеличения частоты строчной развертки, так как уменьшается время, отводимое на формирование каждой точки изображения. Частота строк в килогерцах, вообще говоря, определяется произведением частоты вертикальной развертки на количество выводимых строк в одном кадре (разрешающая способность по вертикали). Полоса видеосигнала, измеряемая в мегагерцах, определяет самые высокие частоты в видеосигнале. Приблизительно эта величина может быть рассчитана как произведение количества точек в строке (разрешающая способность по горизонтали) на частоту строчной развертки. Иными словами, этот параметр отражает число точек в строке, которое монитор может воспроизвести за одну секунду.

Не менее значимым фактором, чем частота кадровой развертки, является способ формирования изображения на экране монитора в режимах высокого разрешения — построчный (прогрессивный) или чересстрочный. При построчном способе формирования изображения все строки кадра выводятся в течение одного периода кадровой развертки, при чересстрочном — за один период кадровой развертки выводятся четные строки

изображения, а за следующий — нечетные. Таким образом, говорят, что один кадр делится на два поля. Последний способ, кстати, используется и в телевидении. Совершенно нетрудно заметить, что в случае чересстрочной развертки эквивалентная частота кадров снижается вдвое. Это позволяет достаточно легко увеличивать разрешающую способность монитора, хотя и в ущерб качеству изображения.

Цветной монитор работает примерно так же, как и монохромный (черно-белый). Вообще говоря, монохромные мониторы могут использовать люминофор не только белого, но и, например, янтарного цвета. Итак, в случае цветного монитора имеются уже три электронные пушки с отдельными схемами управления, а на поверхность экрана нанесен люминофор трех основных цветов: R (Red, красный), G (Green, зеленый) и B (Blue, синий). Таким образом, каждая пушка должна «стрелять» только по своей «цели». Для этого в каждом цветном кинескопе имеется либо тeneвая маска, либо (в кинескопах Trinitron) так называемая апертурная решетка. Они служат для того, чтобы лучи электронных пушек попадали только в точки люминофора соответствующего цвета. Если тeneвая маска содержит систему отверстий, то апертурная решетка образует систему щелей, выполняющих ту же функцию. В современных мониторах используется так называемая усовершенствованная тeneвая маска (*advanced shadow mask*). Обычно маска для монитора делается из специального сплава — инвара, который имеет очень маленький коэффициент температурного линейного расширения. Таким образом, даже при нагревании маски четкость изображения сохраняется. Тем не менее, в мониторах используются специальные схемы динамической фокусировки и сведения лучей. Отверстия маски соответствуют точкам люминофора, нанесенным на обратной стороне плоскости экрана. Величина точки люминофора, необходимая для обеспечения требуемого разрешения, зависит от размеров экрана. Чем больше нужно разместить точек и чем меньше экран, тем плотнее приходится располагать точки.

При прочих равных условиях четкость изображения на мониторе тем выше, чем меньше размеры точек люминофора на внутренней поверхности экрана. Обычно говорят не о размерах самих точек, а о расстоянии между ними (*dot pitch*). Этот пара-

метр для различных моделей мониторов может лежать в диапазоне от 0,41 до 0,18 мм, однако для хороших моделей диапазон существенно сужается — от 0,25 до 0,18 мм. Интересно отметить, как примерно определяется расстояние между точками. Для обычного 14-дюймового монитора размер экрана в ширину составляет около 265 мм. Режим 640х480 требует отображения 640 точек на одну линию. Следовательно, расстояние между точками должно быть не более 0,41 мм: 265/640 (табл. 4).

Таблица 4

**Зависимость между разрешением и расстоянием
между точками люминофора**

Размер экрана, дюйм	640X480	800X600	1024X768	1280X1024
14	0,32	0,28	0,22	0, 18
15	0,38	0,30	0,24	0, 19
17	6,43	0,34	0,27	0,22
21	0,50	0,40	0,31	0,25

Мультичастотные (мультисканирующие) мониторы, именуемые иногда Multisync — по названию мониторов, выпускаемых фирмой NEC, обладают способностью настраиваться на произвольные значения частот синхросигналов из некоторого заданного диапазона, например 30–64 кГц для строчной и 50–100 Гц — для кадровой развертки.

Основным управляющим сигналом для мониторов является видеосигнал, поступающий с платы видеоадаптера.

1.13.3 «Зеленые» мониторы

Напомним, что оборудование, удовлетворяющее программе Energy Star (обычно называемое «зеленым»), должно потреблять в среднем (в режиме холостого хода) не более 30 Вт, не использовать токсичных материалов и допускать 100-процентную утилизацию после истечения срока службы. Требования агентства ЕРА учитываются при разработке различных промышленных

стандартов. Так, например, ассоциация VESA приняла стандарт на систему управления энергопотреблением дисплея DPMS (Display Power Managment Signaling).

Данная спецификация, во-первых, описывает метод выведения монитора из режима активной работы, причем эта операция осуществляется в несколько последовательных этапов, на каждом из которых происходит определенное сокращение потребляемой мощности. Во-вторых, этот же стандарт предлагает метод опознавания частотных сигналов строчной и кадровой разверток, используемый для переключения из режима в режим при снижении расходуемой мощности.

Итак, спецификация DPMS устанавливает 4 режима потребления мощности для монитора: On (максимальной нагрузки), Standby (ожидания, или резервный), Suspend (приостановки работы) и Off (отключения). Режимы перечислены в порядке убывания потребляемой мощности. Кстати, напомним, что по оценкам экспертов, например, на 14-дюймовый монитор приходится до 60% расходуемой компьютером энергии.

В режиме On компьютер включен и активно работает, а энергопотребление определяется характеристиками используемого оборудования. В режиме Standby отключаются видеосигнал и горизонтальная синхронизация, а контрастность и яркость удерживаются на минимальном уровне. Нажатие любой клавиши или движение мышью возвращает монитор в режим On. Потребление энергии в режиме Standby снижается примерно на 20%. В режиме Suspend включается горизонтальная, но отключается вертикальная синхронизация. Высокое напряжение в мониторе отключается. В режим On монитор может вернуться только спустя несколько секунд после соответствующей команды. Потребление энергии в Suspend-режиме снижается примерно на 70%. Режим Off обеспечивает максимум сохранения энергии. В мониторе отключены все блоки, за исключением DPMS-контроллера. Уровень потребления падает примерно до 5% от состояния On. Время возвращения к нормальной работе более продолжительное, чем из режима Suspend.

Кстати, реализация спецификации DPMS не является единственным способом уменьшения энергопотребления монитора. Более простой и не требующий специальных аппаратных средств

способ — так называемое гашение экрана. В этом случае монитор реагирует на пропадание одного из сигналов основных цветов (обычно синего) переходом в один из режимов энергосбережения.

1.13.4 Плоские экраны

Наиболее важными для монитора являются такие параметры, как хорошая «читаемость» изображения, высокая разрешающая способность, быстрая реакция на изменение изображения, низкое энергопотребление, малый вес, простое электронное управление, воспроизведение широкой цветовой гаммы или оттенков серого и, при всем при этом, невысокая стоимость.

За последние годы появился достаточно широкий спектр так называемых плоских экранов. Плоские дисплеи могут использовать различную технологию, например газоплазменную, электролюминесцентную и жидких кристаллов.

К основным недостаткам газоплазменных дисплеев можно отнести необходимость использования высокого напряжения (около 200 В) и небольшой срок службы. Этот тип дисплеев обеспечивает, как правило, только янтарный цвет экрана. Электролюминесцентные дисплеи хотя и обладают хорошими характеристиками, но достаточно дороги и в основном используются в военной технике. Наиболее распространенный цвет экрана электролюминесцентных дисплеев — желтый. Последний тип плоских экранов — жидкокристаллические (ЖК). Они составляют сейчас более половины рынка мониторов компьютеров. Хотя технология жидких кристаллов применяется относительно недавно, в ее развитии произошел уже заметный прогресс.

1.13.5 Жидкокристаллические экраны

Как известно, обычные кристаллы обладают симметрией атомной структуры и анизотропией физических свойств, то есть неодинаковостью этих свойств в зависимости от направления. Анизотропия характерна, например, для упругих, тепловых и оптических свойств различных кристаллов. Поскольку жидкости, газы и аморфные тела в отличие от кристаллов не имеют симмет-

ричной структуры молекул, то обладают поэтому изотропией свойств.

Современная наука определяет жидкий кристалл как некоторое мезоморфное состояние, в котором вещество обладает свойствами жидкости (текучестью) и некоторыми свойствами твердых кристаллов (например, анизотропией).

В конце 1966 года американской фирмой RCA (Radio Corporation of America) был продемонстрирован первый LC-дисплей. Английское название Liquid Crystal (сокращенно LC) соответствует русскому «жидкий кристалл» — ЖК. В 1971 году учеными США были предложены так называемые Twisted-Nematic-ячейки (TN-элементы).

Заметим, что для изготовления ЖК-экранов используют так называемые нематические кристаллы, молекулы которых имеют форму палочек или вытянутых пластинок. Допустим, между двумя стеклами с прозрачными электродами находятся молекулы жидких кристаллов (слой вещества составляет приблизительно 5—10 мкм). В отсутствие электрического поля молекулы этого вещества образуют спирали, скрученные на 90 градусов (кстати, отсюда и их название Twisted). В результате такой ориентации молекул плоскость поляризации проходящего через ЖК-элемент света поворачивается примерно на этот же угол. Если на входе и выходе этого элемента поместить поляризаторы, смещенные относительно друг друга также на угол 90 градусов, то свет беспрепятственно может проходить через этот элемент. Если же к прозрачным электродам приложено напряжение, спираль молекул распрямляется (они просто ориентируются вдоль поля). Поворота плоскости поляризации уже не происходит, и, как следствие, выходной поляризатор не пропускает свет. Кстати, если несколько изменить конструкцию элемента (используя зеркало на выходе второго поляризатора), то темный или светлый сегмент можно увидеть и в отраженном свете. Примером в данном случае может служить ЖК-индикатор наручных электронных часов.

Молекулы элемента Super-Twisted-Nematic (STN) закручены на угол от 180 до 270 градусов, за счет чего несколько улучшается контраст изображения. Если при использовании TN-элементов контраст определяется как 3:1, то есть освещенная точка в три раза светлей темной, то для STN-элементов это соотношение мо-

жет составлять уже 10:1 и даже выше. Однако в STN-элементах из-за эффекта эллиптической поляризации проявляется некоторый сдвиг цветов. В этом случае чисто белый цвет может становиться, например, бледно-оранжевым, а черный — зелено-голубым (цвет циан). Зачастую дисплеи, использующие STN-элементы, называют поэтому также Blue-Mode-STN-LCD.

Экран ЖК-дисплея имеет, как правило, либо заднюю подсветку (backlight, или просто backlit), либо боковую (sidelight, или просто sidelit). Для экономии заряда гальванических элементов могут использоваться ЖК-экраны без подсветки, которые работают только в отраженном свете (reflective LCD).

Таким образом, каждая точка изображения на ЖК-дисплее представляет из себя соответствующий TSTN-элемент, а весь экран ЖК-дисплея — по сути, матрица этих элементов. В настоящее время существуют два основных метода, используемых для адресации ЖК-элементов: прямой (или пассивный) и косвенный (или активный).

При прямой адресации элементов матрицы каждая выбираемая точка изображения активируется подачей напряжения на соответствующий адресный (прозрачный) проводник-электрод для строки и соответственно для столбца. При таком способе управления точкой изображения говорят также, что используется пассивная матрица (passive matrix) ЖК-элементов. Очевидно, что такому методу управления присущи определенные недостатки. В частности, практически невозможно достичь высокого контраста изображения, так как электрическое поле возникает не только в точке пересечения адресных проводников, но и на всем пути распространения тока. Эта проблема вполне разрешима при использовании так называемой активной матрицы (active matrix) ЖК-элементов, когда каждой точкой изображения управляет свой электронный переключатель. Информация об изображении (видеосигнал) выдается построчно на все соответствующие столбцы матрицы экрана, а выбор необходимой точки изображения в строке происходит через соответствующий электронный переключатель.

Каждую ячейку такого экрана можно представить в виде простой схемы замещения, представляющей собой RC-цепочку. Благодаря видеосигналу конденсатор заряжается, а через очень

большое параллельно включенное сопротивление достаточно долго разряжается. Поскольку время разряда в несколько раз превышает время, через которое видеосигнал повторяется, то изображение получается устойчивым и контрастным. Возможность изменения амплитуды напряжения видеосигнала позволяет использовать в воспроизводимом изображении оттенки цвета. Следует, конечно, помнить о том, что стоимость реализации активной матрицы экрана существенно выше, нежели пассивной. Обычно активные матрицы реализованы на основе тонкопленочных полевых транзисторов (TFT, Thin Film Transistor), хотя это и не единственно возможное решение.

При применении активных матриц дисплеев большое значение имеют такие параметры, как малое время отклика (Response Time) и большой угол зрения (Viewing Angle). При использовании пассивных матриц ЖК-элементов значение первого параметра лежит в пределах 250–300 мкс, тогда как у активных матриц его значение составляет около 25 мкс. Изображение на экране с TFT-матрицей можно различить даже под углом 75 градусов, а, например, на дисплее с пассивной матрицей этот угол не превышает 45 градусов.

В цветных ЖК-дисплеях с активной матрицей, как обычно, каждый элемент изображения (или пиксел) состоит из трех точек — синей, зеленой и красной. Соответственно для каждой точки, составляющей элемент изображения, используются свой TFT-транзистор и фильтр (Red, Green или Blue). Общее количество транзисторов в случае VGA-дисплея (640 на 480 на 3) составляет 921 600. Кстати, надо заметить, что поскольку цветные фильтры поглощают свет от лампы подсветки довольно существенно, то ее (или их) мощность должна быть раз в пять выше, чем для обычных монохромных дисплеев.

1.14 Видеоадаптеры

Все современные видеоподсистемы могут работать в одном из двух основных видеорежимов: текстовом или графическом. В текстовом режиме, называемом иногда символьным, экран монитора разбивается на отдельные символьные позиции, в каждой из которых может выводиться один символ. Для преобразования

кодов символов, хранимых в видеопамяти адаптера, в точечные изображения на экране служит так называемый знакогенератор, который обычно представляет из себя ПЗУ, где хранятся изображения символов, «разложенные» по строкам. При получении кода символа знакогенератор формирует на своем выходе соответствующий двоичный код, который затем преобразуется в видеосигнал.

В графическом режиме для каждого пиксела отводится от одного (монохромный режим) до нескольких бит (обычно цветной). Графический режим часто называют режимом с адресацией всех точек, поскольку только в этом случае имеется доступ к каждой точке изображения. Исторически сложилось так, что начальные адреса в видеобуфере для текстовых и графических режимов не совпадают. Большинство текстовых режимов имеют стартовый адрес памяти BSOOOh, а графические — AOOOOh.

Максимальное разрешение и количество воспроизводимых цветов конкретной видеоподсистемы в первую очередь зависят от общего объема видеопамяти и количества бит, приходящихся на один элемент изображения. Так, если для отображения одного пиксела отводится один бит, то понятно, что можно обеспечить только монохромный режим (0 или 1, есть точка или нет точки); если более одного, то имеется возможность воспроизводить либо оттенки серого за счет изменения яркости (00 — черный цвет, 01 — слабое свечение, 10 — обычное, 11 — яркое), либо цвета. Например, четырехцветный графический режим может включать поддержку черного, зеленого, красного и коричневого цветов. Разумеется, возможна и смена палитры.

Как известно, для формирования цвета элемента изображения на экране монитора обычно используются три основных цвета — R, G и B. Кроме них, некоторые видеоадаптеры, работающие с уровнями ТТЛ, вырабатывают также сигнал интенсивности (яркости) изображения — Intensity. Это дает возможность реализовать так называемую 16-цветную палитру IRGB. Другие адаптеры (также с цифровыми уровнями сигналов) кодируют каждый цвет не одним, а двумя независимыми сигналами — основным (R, G или B) и дополнительным с меньшей интенсивностью (r, g или b). Все возможные комбинации позволяют отображать уже

64 цвета, хотя одновременно по-прежнему только 16, поскольку для кодирования цвета отводится всего 4 бита. Видеоадаптеры с выходным аналоговым сигналом подают на соответствующий монитор сигналы R, G|B и, разумеется, сигналы синхронизации.

Введем еще одно понятие — «видеостраница» (далее просто страница). Размер памяти, необходимой для заполнения экрана, особенно в текстовых режимах, намного меньше возможной емкости видеобуфера (128 Кбайт). Например, в текстовом режиме 3 необходимо всего около 4 Кбайт: $80 \times 25 \times 2 = 4000$ байт (помним о том, что каждый символ занимает два байта). Именно эта величина и определяет размер страницы для режима 3. Для режима, определяющего 25 строк при 40 столбцах, размер страницы будет составлять уже около 2 Кбайт: $40 \times 25 \times 2 = 2000$ байт. Если реальная емкость видеобуфера превышает размер одной страницы, в нем можно организовать несколько страниц, причем в любой момент времени на экран может выводиться содержимое только одной, активной (или текущей), страницы. Впрочем, остальные страницы в любой момент времени остаются доступными процессору. Например, в режиме 3 средства ROM BIOS позволяют работать с 8 страницами (0–7). Заметим, что концепция разбиения видеопамати на страницы сохраняется и в графических режимах.

Одной из существенных особенностей некоторых видеоадаптеров является поддержка загружаемых символьных наборов, определенных, например, самим пользователем.

Как известно, в первой IBM PC на экране монитора могла отображаться только алфавитно-цифровая информация. Ни возможности вывода графики, ни тем более изменения цветов предусмотрено не было. Первый видеоадаптер имел полное название Monochrome Display and Parallel Printer Adapter (MDPPA), которое чаще использовалось в сокращенном виде — MDA. Разрешающая способность адаптера MDA позволяла отображать на мониторе 720 точек по ширине и 350 точек (пикселей) по высоте экрана. Поскольку, как уже было сказано, графического режима в адаптере предусмотрено не было, то алфавитно-цифровая информация отображалась на экране в 25 строк по 80 символов в каждой. Возможность использования таких атрибутов, как нега-

тивное изображение, повышенная яркость, подчеркивание и мерцание, в некоторой степени компенсировала отсутствие цветов.

Тем не менее спустя всего несколько месяцев после выпуска первой «писишки» с MDA фирма IBM разработала видеоадаптер, который поддерживал не только графическое изображение, но и цвета, что, кстати, особо подчеркивалось даже в его названии. Адаптер CGA (Color Graphics Adapter) обеспечивал отображение четырех цветов при разрешающей способности 320 на 200 пикселей. Заметим, кстати, что именно в то время и появились первые игры для IBM PC.

Чуть позже стало понятно, что графика на CGA, даже цветная, не всегда удовлетворяет решаемым задачам, в частности, из-за низкой разрешающей способности. Первый видеоадаптер для IBM PC, в какой-то мере отвечавший этим нуждам, был создан на фирме Hercules в 1982 году. Этот адаптер — HGC (Hercules Graphics Card) — поддерживал на монохромном мониторе разрешение 720 на 350 точек.

Новой разработкой фирмы IBM стал улучшенный графический адаптер EGA (Enhanced Graphics Adapter), который появился на свет уже в 1984 году. Этот адаптер не только позволял полностью эмулировать все режимы работ предыдущих адаптеров (MDA, CGA), но и, разумеется, обладал дополнительными возможностями. Например, при разрешающей способности 640 на 350 пикселей он мог одновременно воспроизводить 16 цветов из палитры в 64 цвета (именно для этого адаптера использовались сигналы Rr Gg Bb).

Несколько незаметно, по крайней мере в нашей стране, начал и закончил свое существование видеоадаптер PGA (Professional Graphics Adapter), который сделал следующий шаг в развитии возможностей адаптера EGA. Так, при разрешающей способности 640 на 480 пикселей на экране могло одновременно воспроизводиться 256 цветов из 4096 возможных.

Видеоадаптер VGA (Video Graphics Array) был объявлен фирмой IBM еще в 1987 году, и до недавнего времени он по-прежнему оставался одним из самых распространенных. При создании этого устройства была обеспечена его полная совместимость сверху вниз с адаптером (EGA), что позволило сохранить преемственность существующего программного обеспечения.

Немудрено поэтому, что вскоре VGA стал фактическим стандартом, включающим в себя все режимы предыдущих адаптеров и расширяющим их возможности по разрешающей способности и количеству воспроизводимых цветов. Так, при использовании адаптера VGA обеспечивается разрешение 640 на 480 пикселей и на экране монитора может воспроизводиться 16 цветов. Все режимы, исключая графические с разрешением 640 на 480 пикселей, используют вертикальную развертку с частотой 70 Гц, что существенно снижает ощущаемое пользователем мерцание экрана. Частота развертки для режима 640 на 480 точек составляет только 60 Гц. Частота строчной развертки равняется 31,5 кГц, что, кстати, вдвое больше американского телевизионного стандарта NTSC (15,7 кГц). Основными узлами VGA-адаптера являются собственно видеоконтроллер (как правило, заказная БИС — ASIC), видеоBIOS, видеопамять, специальный цифроаналоговый преобразователь с небольшой собственной памятью (RAMDAC, Random Access Memory Digital to Analog Converter), кварцевый осциллятор (один или несколько) и микросхемы интерфейса с системной шиной.

Создание клонов коснулось не только персональных компьютеров фирмы IBM, но и разработанных ею видеоадаптеров. Тем не менее копирование было достаточно точным, то есть можно сказать, что соблюдался некий стандарт «по умолчанию».

1.14.1 Видеоадаптеры SVGA

После того как стало ясно, что стандарт VGA практически полностью себя исчерпал, большинство независимых разработчиков начали его улучшать как за счет увеличения разрешающей способности и количества воспроизводимых цветов, так и введения дополнительных возможностей. Например, некоторыми отличительными особенностями видеоадаптеров SVGA принято считать следующие:

- использование 16- и 256-цветных режимов с разрешением 1024 на 768 точек;
- применение новых цифроаналоговых преобразователей RAMDAC (8 разрядов на цвет вместо 6);
- от 512 Кбайт до 4 Мбайт видеопамати;

- поддержка аппаратного курсора;
- 16-разрядная шина данных (у VGA обычно 8);
- 16-разрядный видеоBIOS (у VGA обычно 8);
- двухпортовые микросхемы памяти (VideoRAM);
- программируемый кварцевый генератор и т.д.

Вообще говоря, прародителем SVGA можно считать фирму NEC, разработавшую VGA-совместимый графический адаптер, который имел дополнительный режим с разрешением 800 на 600 точек и одновременным отображением 256 (из 256 К возможных) цветов (табл. 5). Хотя все производители обеспечивали совместимость своих изделий с VGA, дополнительные видеорежимы и возможности адаптеров зачастую не совпадали, поскольку каждый считал нужным делать это по-своему. Немудрено, что уже само понятие SVGA, не связанное жестко с конкретными режимами работы адаптера, вносило серьезную неразбериху. В октябре 1989 года ассоциация VESA предложила свой стандарт на новые видеоадаптеры — VESA BIOS Extension, который в настоящее время поддерживается большинством фирм-производителей. Сначала VESA рекомендовала использовать режим с разрешением 800 на 600 точек и поддержкой 16 цветов как стандартный. Затем последовали 256-цветные режимы с разрешением 640 на 480, 800 на 600 и 1024 на 768 точек, а также 16-цветный режим с разрешением 1024 на 768 пикселей и т.д.

Таблица 5

**Зависимость количества воспроизводимых цветов
от типа видеоадаптера**

Количество цветов	Разрешения					
	320x200	640x200	640x350	640x480	800x600	1024x768
2	CGA	CGA	EGA	VGA	SVGA	SVGA
4	CGA	EGA	EGA	VGA	SVGA	SVGA
16	EGA	EGA	EGA	VGA	SVGA	SVGA
256	VGA	SVGA	SVGA	SVGA	SVGA	SVGA

1.14.2 Проблемы цветопередачи

Цветовую гамму изображения можно расширить только при увеличении количества разрядов, приходящихся на один цвет. Например, стандартный VGA-режим 12h (16 цветов, 640 на 480 точек) использует 4 бита на пиксел (работает метод bit-mapped, 4 битовые плоскости), а режим 13h — 8 бит на пиксел. Количество одновременно воспроизводимых цветов подсчитывается достаточно просто: цифра 2 возводится в степень, равную количеству бит на один пиксел. Теоретически, конечно, объем памяти видеоадаптера можно увеличить до нужных размеров, чтобы использовать требуемое количество разрядов для кодирования цвета, однако в этом случае организация видеопамати должна соответствовать определенным стандартам. Для персональных компьютеров таких стандартов создано уже несколько. Например, видеоадаптеры поддерживают одновременное воспроизведение 32 768 или 65 536 цветов. Видеорежимы, в которых используются 15 или 16 разрядов для кодирования цвета, называют HighColor. В видеоадаптерах с точной цветопередачей (достаточной даже для обработки профессиональных высококачественных фотографий) применяют 24-разрядное кодирование цвета, по 8 разрядов на каждую составляющую. В этом случае возможно воспроизведение 16,8 миллиона цветов одновременно (а точнее, 16 777 216). Такой видеорежим называется TrueColor.

Хотя общий объем памяти и ограничивает количество цветов, которые может создавать видеоадаптер, сам спектр оттенков, получаемых на экране монитора, зависит от используемого RAMDAC. До сих пор мы в основном уделяли внимание разрешающей способности и количеству воспроизводимых цветов, совершенно «забывая» о скорости работы видеоадаптера.

Понятно, что чем большее количество памяти занимает изображение, тем большее число байт необходимо обрабатывать и пересылать на монитор, причем за время, ограниченное прямым ходом кадровой развертки. Следует также помнить о том, что сама видеопамать — это ресурс, который разделяют между собой микропроцессор системы и видеоконтроллер.

Таблица 6

**Объем видеопамати в зависимости от разрешения
и количества цветов**

Разрешаемая способность	Количество бит	Количество цветов	Минимальный объем видеопамати, Мб
640х480	4	16	0,25
640х480	8	256	0,50
640х480	16	65 536	1,00
640х480	24	16777216	1,00
800х600	4	16	0,25
800х600	8	256	0,50
800х600	16	65536	1,00
800х600	24	16777216	1,50
1024х768	4	16	0,50
1024х768	8	256	1,00
1024х768	16	65 536	1,50
1024х768	24	16 777216	2,50
1280х1024	4	16	1,00
1280х1024	8	256	1,50
1280х1024	16	65536	2,50
1280х1024	24	16 777 216	4,00
1600х1200	4	16	1,00
1600х1200	8	256	2,00
1600х1200	16	65 536	4,00
1600х1200	24	16777216	6,00

Стоит обратить внимание, что обычная микросхема фрейм-контроллера является, вообще говоря, достаточно пассивным устройством. Все операции по записи и модификации данных в видеопамати выполняет сам процессор системы. Следовательно, чем быстрее используемый микропроцессор, тем быстрее начинает работать и видеоподсистема компьютера. Однако и здесь существует определенный предел, который связан с конечным быстродействием системной шины, через которую и происходит обмен между процессором и видеоадаптером. Так как, например, системная шина ISA работает на тактовой частоте 8 МГц,

использование локальной шины VL-bus с тактовой частотой 50 МГц позволило ускорить работу примерно в 6 раз.

Однако более логичным выходом из сложившейся ситуации было бы использование более «интеллектуального» видеоконтроллера, который смог бы разгрузить основной процессор от некоторых рутинных операций. Подавляющее большинство видеоадаптеров базируются на так называемых ускорителях (акселераторах) или, но реже, на графических сопроцессорах.

Акселераторы и графические сопроцессоры повышают быстродействие видеоподсистемы отчасти благодаря сокращению количества информации, передаваемой по системной шине компьютера. Часть изображения может создаваться этими устройствами уже без загрузки основного процессора. Для этого им посылаются специальные команды или даже небольшие подпрограммы (для сопроцессоров). Собственно акселератор представляет из себя специализированный графический сопроцессор, ориентированный на выполнение строго определенного перечня графических операций с ориентацией на конкретные программы и приложения. Соответственно графический сопроцессор — устройство более универсальное, которое можно запрограммировать на выполнение практически любых графических функций. Таким образом, основная разница между сопроцессором и акселератором состоит в степени их программируемости. Поскольку эти устройства оптимизированы именно для выполнения графических операций, то и все такие операции они выполняют быстрее, чем универсальный микропроцессор, кроме того, работают они с ним параллельно.

Широкое применение Windows подтолкнуло развитие видеоадаптеров с акселераторами, в первую очередь ориентированными именно на эту программную среду. Большинство микросхем акселераторов берет на себя выполнение операций перемещения фрагментов растрового изображения (битовых блоков) BitBlt, рисования линий и многоугольников, закрашивания определенным цветом указанных многоугольников, а также поддержку аппаратного курсора.

1.14.3 Графические и мультимедиа-акселераторы

Вообще говоря, для повышения производительности по графическим операциям существуют довольно веские причины. Допустим, при работе в Windows 256 цветов (то есть 8 бит на пиксел) отображаются при разрешающей способности 1280 на 1024 точки. Это значит, что 1,3 Мбайта данных передаются из видеопамяти на экран 72 раза в секунду. Таким образом, необходимо обеспечить скорость передачи не менее 92 Мбайт/с. И это только для пересылки статического изображения. Любые другие операции потребуют дополнительного увеличения скорости доступа конкретных микросхем видеопамяти, но изменив способ ее организации и/или увеличив разрядность шины память-контроллер. Впервые фирма Tseng Labs использовала «чередующуюся» (intereaving) память на обычных DRAM. Вся видеопамять была разбита на два блока по четным и нечетным адресам. Это позволяло посылать из одного блока данные на экран, а во второй блок записывать данные с системной (локальной) шины. Скорость передачи удалось увеличить до 160 Мбайт/с.

Вообще говоря, внутренняя структура микросхем видеоконтроллеров отличается весьма значительно. Впрочем, можно выделить ряд общих задач, которые выполняет каждая из них: связь с основным процессором, коммуникация с видеопамятью и пересылка изображения на экран. Для реализации дополнительных (ускоряющих) функций в каждой микросхеме имеются, разумеется, специальные блоки. Например, тот же графический контроллер MGA фирмы Matrox имеет встроенные функции для трехмерной графики. Большинство современных микросхем акселераторов поддерживают функции энергосбережения.

Жесткие требования, предъявляемые сегодня к видеоподсистеме персонального компьютера, особенно заметны именно для систем, в той или иной степени связанных с мультимедиа. Под мультимедиа-акселераторами понимают устройства, которые помимо ускорения обычных графических операций (перенос блока данных BitBlt, закраска прямоугольников, поддержка аппаратного курсора) могут также выполнять ряд операций по обработке видеоданных. Иными словами, под мультимедиа-акселераторами понимают совокупность программно-аппаратных средств, кото-

рые объединяют базовые возможности графических акселераторов с одной или несколькими мультимедиа-функциями, требующими установки в компьютер дополнительных устройств. Так, к мультимедиа-функциям относятся, например, следующие:

- цифровые фильтрация и масштабирование видео,
- аппаратные цифровые компрессия и декомпрессия видео,
- ускорение графических операций, связанных с трехмерной (3D) графикой,
- поддержка «живого» видео на мониторе,
- наличие композитного видеовыхода,
- вывод TV-сигнала на монитор.

Появление спецификации DCI (Display Control Interface) фирмы Microsoft позволило осуществить воспроизведение полно-скоростного цифрового видео с использованием ресурсов мультимедиа-акселераторов. Вообще говоря, DCI — это интерфейс нижнего уровня, который дает возможность программным средствам для воспроизведения видео воспользоваться имеющимися возможностями аппаратных средств. Так, если DCI-драйвер обнаружил наличие аппаратной поддержки некоторых мультимедиа-функций, он изменяет последовательность выполнения операций и разгружает центральный процессор от выполнения некоторых из них.

Надо сказать, что сейчас большинство фирм, выпускающих микросхемы графических акселераторов, включили в свою производственную программу изделия, выполняющие ряд мультимедиа-функций. В частности, сигнал изображения может преобразовываться из пространства RGB в пространство YUV, над ним могут выполняться такие операции, как сжатие, билинейное масштабирование, линейная интерполяция, фильтрация и растривание (dithering). Кроме того, часто имеются встроенные схемы genlock, позволяющие синхронизировать преобразованный компьютерный сигнал RAMDAC по внешнему (NTSC/PAL) видеосигналу. На кристаллах также интегрируют схемы управления разделяемым фрейм J-буфером. Для наложения графики и видео используют схемы управления типа «хромакей» (chromakey). Многие микросхемы ускоряют алгоритмы декомпрессии стандартных видеокодеков, включая, например, Indeo, Cinepak и MPEG-4.

Из наиболее общих технологических тенденций, доминирующих на рынке мультимедиа-акселераторов, кроме уже перечисленных, можно отметить следующие: использование 32- и 64-разрядных микросхем-контроллеров с чередованием блоков памяти, увеличение объема видеопамати и повышение частоты обновления изображения, внедрение новых стандартов мониторов DPMS и DDC, поддержка воспроизведения цифрового видео, отсутствие единого стандарта на видеопину, ускорение трехмерных (3D) графических операций недорогими средствами.

Несомненно революционным стало появление так называемых 3D (трехмерных) мультимедиа-акселераторов. Подобные платы адаптеров в конце 1995 года стала выпускать фирма Diamond Multimedia Systems, лидирующая на рынке плат графических акселераторов. Данные устройства базируются на технологии, предложенной компанией Nvidia. В частности это позволяет пользователям при работе с Windows 95 получать фотореалистичное трехмерное изображение, быструю двухмерную графику, «живое» видео и табличный волновой синтез звука. 3D-акселераторы, в частности, позволяют достигать разрешения 1024 на 768 точек и более при одновременном воспроизведении 65 тысяч цветов, но, самое главное, получаемое при этом изображение полностью соответствует по скорости реальному действию. В недавнем прошлом такая технология обходилась не менее чем в 2 тысячи долларов, а вот стоимость карт сейчас варьируется в пределах двух сотен долларов.

1.14.4 Видеокабели

Кабели монитор-видеоадаптер различаются в зависимости от уровня применяемых сигналов (ТТЛ или аналоговый). Имеется два стандарта на соединители в таких кабелях: 15-контактный DB-shell для мониторов, использующих аналоговые сигналы (VGA, Super VGA), и 9-контактный DB-shell для мониторов, использующих ТТЛ-сигналы (MDA, CGA и EGA). Как правило, на всех таких кабелях имеются специальные фильтры (выглядят обычно как утолщения), уменьшающие высокочастотные помехи в сигналах изображения. Длина кабелей не превышает обычно одного метра.

Теперь несколько практических советов.

Помните, что размер изображения на экране будет немного меньше, чем заявляемый изготовителем диагональный размер. Правда, ряд современных моделей избавлен от этого недостатка, например SyncMaster фирмы Samsung. При покупке монитора после его включения по возможности расфокусируйте изображение и введите достаточный уровень яркости. После этого с небольшого расстояния посмотрите, видны ли на экране мерцания и пульсации. Чем меньше они заметны, тем меньше вы будете уставать, работая за таким монитором.

Фокусировку монитора лучше всего проверять, загрузив Windows. Поскольку практически все модели мониторов имеют в центре экрана более четкое и резкое изображение, чем по краям, то многочисленные «иконки» в Windows отобразят это достаточно наглядно. Если разница в качестве изображения в центре и на периферии достаточно велика, то покупать такой монитор не стоит. Заметим, что некоторые погрешности изображения (например, «ломанные» линии) могут быть связаны и с подключенным видеоадаптером, поэтому для более объективной оценки следует по возможности его поменять.

Если видеоадаптер приобретается вместе с монитором, то следует проверить видеоподсистему во всех поддерживаемых режимах и убедиться, что изображение является четким и неподвижным.

Как известно, ни один монитор, основанный на электронно-лучевой трубке, не имеет абсолютно плоской поверхности экрана, но о некоторых моделях говорят уже как о мониторах с плоским экраном. Это следует также иметь в виду. Более того, монитор, как известно, не зеркало, поэтому антибликовое покрытие его экрана позволяет обходиться без дополнительных защитных средств и препятствует утомлению глаз, что имеет существенное значение для комфортной работы.

В ряде случаев (отсутствие антибликового покрытия, несоответствие стандарту MPR II и т.д.) вместе с монитором необходимо использовать специальные защитные фильтры.

1.14.5 Видеокарты

В настоящее время имеется большое количество устройств, предназначенных для работы с видеосигналами на IBM PC-совместимых компьютерах. Условно все эти устройства можно разбить на следующие группы: устройства для ввода и захвата видеопоследовательностей (capture-play), фрейм-грабберы (frame grabber), TV-тюнеры, преобразователи сигналов VGA-TV и MPEG-плейеры.

1.14.6 TV-тюнеры

Эти устройства, выполненные обычно в виде карт или бокса (небольшой коробочки), преобразуют аналоговый видеосигнал, поступающий по сети кабельного телевидения от антенны, видеомagniтофона или камкодера (camcorder). Если одни устройства сами выполняют дискретизацию поступающего видеосигнала, то другие требуют для этого отдельной карты. Функции TV-тюнера могут реализовываться и другими устройствами, такими, как MPEG-плейеры или фрейм-грабберы.

Обычно в комплект с TV-тюнерами входит программное обеспечение для Windows, которое включает в себя систему дистанционного управления, позволяющую переключать каналы, устанавливать время таймера, настраивать яркость, контрастность, звук и т.д.

1.14.7 Фрейм-грабберы

Эти устройства объединяют графические, аналого-цифровые и микросхемы для обработки видеосигнала, которые позволяют дискретизировать видеосигнал, сохранять отдельные кадры изображения в собственной памяти (буфере) с последующей записью на диск либо выводить их непосредственно в окно на мониторе компьютера.

Содержимое буфера видеокарты обновляется каждые 40 мс, то есть с частотой смены кадров. Причем вывод видеоинформации происходит в режиме наложения (overlay). Для реализации окна на экране монитора с «живым» видео карта фрейм-граббера соединяется с графическим адаптером через так называемый 26-

контактный feature-коннектор, который находится, как правило, в верхней части платы адаптера (см. рис 16).

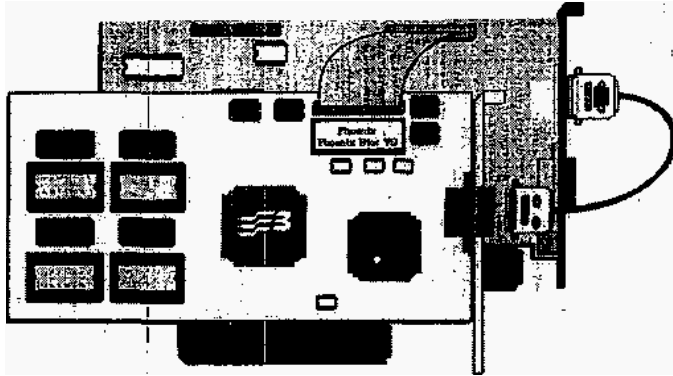


Рис. 16 — Соединение фрейм-граббера с видеоадаптером

Заметим, что специальное программное обеспечение, входящее обычно в комплект с видеокартой, дает возможность выполнять над «захваченным» изображением ряд операций, связанных, например, с его редактированием. Как правило, сохранять отдельные кадры

изображения можно в ряде популярных графических файловых форматов (TIFF, PCX, BMP, GIF и т.д.). Помимо вывода видеoinформации на экран монитора, видеокарты часто выполняют и обратную задачу, то есть преобразовывают компьютерное изображение в композитный телевизионный сигнал. Для современных видеокарт потребителем стандартом стало воспроизведение изображения с качеством видеомаягнитофона (VHS, Video Home System) с разрешением 640 на 480 точек и частотой кадров 25 Гц и выше.

В том случае, когда на жесткий диск необходимо записать не один полный кадр, а их последовательность в течение нескольких десятков секунд, обычные фрейм-грабберы, по понятным причинам, уже не годятся. В данном случае требуются специальные карты для захвата (или вывода) видеопоследовательностей.

1.14.8 Преобразователи VGA-TV

Данные устройства транслируют сигнал о цифровом образе VGA-изображения в аналоговый сигнал, пригодный для вывода на телевизионный приемник. Как правило, преобразователи поддерживают стандарты PAL и NTSC и допускают одновременный вывод изображения на телевизоре и мониторе компьютера. Производители обычно предлагают подобные устройства, выполнен-

ные либо как внутренняя ISA-карта, либо как внешний блок. В последнем случае нет необходимости в свободном слоте на системной плате компьютера, и, кроме того, такое исполнение позволяет подключать преобразователь к портативному компьютеру. Некоторые из преобразователей предлагают такие возможности, как, например, регулировка изображения программным способом или с помощью регулировок (для внешнего исполнения). Ряд преобразователей позволяет накладывать компьютерную графику на внешний видеосигнал, например для создания титров.

1.14.9 MPEG-плейеры

MPEG (Motion Picture Experts Group) — это стандарт, предложенный одноименной организацией для сжатия цифрового видео и звука. Так называемые MPEG-плейеры позволяют воспроизводить последовательности видеоизображений (фильмы), записанные на компакт-дисках, с качеством VHS. Иногда графический адаптер и MPEG-декодер интегрируются на одной плате. Альтернативой MPEG-картам часто выступает программный вариант реализации алгоритма декодирования.

1.15 Порты ввода-вывода

Подсоединение периферийных устройств, таких, как манипулятор типа мышь, внешний модем или принтер, к персональному компьютеру производится через так называемые устройства сопряжения, или адаптеры, на которых реализованы стандартные или специальные интерфейсы. Ранее подобные адаптеры были выполнены в виде отдельных плат ввода-вывода — Input-Output (I/O) Card, вставляемых в разъемы расширения на системной плате. Современные системные платы, как правило, интегрируют все необходимые адаптеры.

Итак, взаимодействие периферийного устройства с адаптером происходит через один (возможно, один из двух) интерфейс, определяющий, в частности, тип и «род» (розетка или вилка, female или male) соединителя, уровни и длительность электрических сигналов, протоколы.

На жаргоне, сложившемся у пользователей IBM PC-совместимых компьютеров, стандартные последовательный и параллельный интерфейсы часто называют портами ввода-вывода. Собственно до последнего времени в качестве последовательного стандартного интерфейса используется разновидность RS-232C (Recommended Standard) — EIA-232D (Electrical Industry Association), а в качестве параллельного — Centronics. Порт называют последовательным, когда информационные биты передаются последовательно один за другим, и параллельным, когда несколько бит данных передаются одновременно. Для подключения джойстика служит специальный аналоговый игровой адаптер — Game Adapter. Если несколько адаптеров (последовательного и параллельного портов, приводов флоппи и жестких дисков) конструктивно выполнены на отдельной плате, она называется, как правило, многофункциональной платой ввода-вывода (Multi I/O Card).

Параллельные порты используются обычно для подключения принтера (в ряде случаев плоттера или сканера). Стандартный интерфейс параллельного порта (Centronics) получил свое название по имени американской фирмы-производителя принтеров, предложившей в свое время собственный интерфейс. Сегодня ситуация несколько изменилась, и, например, ряд принтеров использует более производительные интерфейсы, которые, правда, полностью поддерживают и спецификацию Centronics. Но об этом чуть ниже.

1.15.1 Параллельный порт

Поскольку параллельный порт в IBM PC-совместимом компьютере чаще всего используется для подключения принтера, то его называют также принтер-портом. В MS-DOS компьютер работает максимум с тремя параллельными портами, которые имеют логические имена LPT1, LPT2 и LPT3. В адресном пространстве компьютера резервируются базовые адреса этих портов. Для принтерного порта LPT1 предусмотрено аппаратное прерывание IRQ7, а для LPT2 — IRQ5, хотя на практике они обычно не применяются. Установка базовых адресов портов и возможность использования прерываний настраиваются перестановкой перемычек (jumpers) на плате, описание которых приводится в технической документации для конкретного адаптера, либо программно.

Подсоединение кабеля к адаптеру параллельного интерфейса производится через 25-контактный разъем типа D-shell (DB-25), а со стороны принтера используется специальный 36-контактный разъем типа Centronics. Поскольку частота передаваемых сигналов может достигать десятков килогерц, длина таких кабелей обычно не превышает трех метров. «Фирменные» кабели Centronics бывают длиной 6, 8 и 10 футов (примерно 1,8; 2,4 и 3 м соответственно). Часто ошибки и потеря информации при использовании самодельного кабеля связаны именно с его длиной. Следует помнить, что при передаче данных с TTL-уровнями максимальная длина кабеля не должна превышать 2–3 метров. Обычно для сигналов данных и управления в «фирменном» кабеле используются витые пары, а все проводники заключены в общий экран. Увеличение длины кабеля до десяти и более метров возможно только при использовании специальных усилителей сигналов. Основной функцией таких устройств при этом является сохранение временных соотношений сигналов, их уровней и фронтов. Однако для работы с удаленными принтерами целесообразно использовать последовательный интерфейс.

1.15.2 Последовательный порт

Последовательный интерфейс используется для большинства периферийных устройств, таких, как плоттер, удаленный принтер, мышь, внешний модем, программатор ПЗУ и т.д. До настоящего времени для последовательной связи IBM PC-совместимых компьютеров используются адаптеры с интерфейсом RS-232C (новая версия — EIA-232D). Описание этого интерфейса было опубликовано Американской промышленной ассоциацией еще в 1969 году. Европейским аналогом RS-232 являются два стандарта, разработанных Международным союзом электросвязи (International Telecommunications Union, ITU): V.24 (механические характеристики) и V.28 (электрические характеристики). Хотя первоначально RS-232 был предназначен для связи центральной машины с терминалами, его простота и богатые возможности обеспечили ему более широкое применение. В современном IBM PC-совместимом компьютере, работающем под MS-DOS, может использоваться до четырех последовательных портов, имеющих

логические имена соответственно COM1, COM2, COM3 и COM4. В адресном пространстве IBM PC-совместимых компьютеров последовательный адаптер занимает восемь последовательных адресов, включая базовый. Однако с помощью определенного «трюка» через эти восемь адресов происходит обращение к 11 регистрам, которые программируются соответствующим образом.

По существу, сердцем последовательного адаптера является микросхема UART (Universal Asynchronous Receiver/Transmitter) — универсальный асинхронный приемопередатчик.

Основным преимуществом последовательной передачи является возможность пересылки данных на большие расстояния, как правило, не менее 30 метров. Заметим, что в простейшем случае для приема и передачи через последовательный порт необходимы только три сигнала: TxD (Transmit Data — Передача данных), RxD (Receive Data — Прием данных) и, разумеется, GND (Ground — Земля).

В IBM PC-совместимых персональных компьютерах из 25 сигналов, предусмотренных стандартом RS-232, используются в соответствии с EIA только 9: три, описанных выше, и шесть, объединенных общим названием «handshake (квитирующие) сигналы». При использовании последовательного интерфейса одно из устройств выступает как DTE (Data Terminal Equipment — Оконечное устройство), а другое — как DCE (Data Communication Equipment — Устройство передачи данных). Хотя, вообще говоря, различие между ними состоит только в направлении используемых сигналов. Так, если сигнал для DTE является входным, то для DCE этот же сигнал будет выходным, и наоборот. При обмене данными могут применяться различные протоколы — правила обмена — от простейшего, упомянутого выше и работающего только с тремя сигналами интерфейса, до более сложных, использующих, например, пару квитирующих сигналов RTS-CTS. Различные коммуникационные программы могут применять различные протоколы обмена, поэтому во избежание недоразумений лучше всегда предварительно изучить соответствующие технические описания. Микросхемы UART 8250 и их аналоги рассчитаны на максимальную скорость 38400 бит/с, а 16450 — на 115200 бит/с. Обычно передача данных осуществляется на нескольких дискретных скоростях: 50, 75, 100, 150, 300,

600, 1200, 2400, 4800, 9600 и 19 200 бит/с. Для IBM PC-совместимых компьютеров скорость обмена не превышает 9600 бит/с, поскольку встроенные средства BIOS (коммуникационное прерывание 14h) поддерживают скорости только до этого значения. При специальном программировании регистров можно достигнуть скорости обмена до 115 200 бит/с.

Исходная тактовая частота, используемая для UART, стабилизирована кварцевым генератором (например, 18 432 кГц). Из этой частоты формируются все упомянутые частоты. Последовательный порт работает асинхронно — данные передаются без тактового сигнала. Как известно, в этом случае незначительное различие скоростей приема и передачи не влияет на качество обмена. На обеих сторонах должны быть установлены программно (или с помощью переключателей) следующие параметры: скорость передачи данных, количество битов обмена (от 5 до 8), количество стоповых битов (1 или 2), бит контроля (по четности или нечетности, при 8 битах отсутствует). Передача данных начинается с изменения напряжения на линии с низкого уровня до высокого (с -12 В до $+12\text{ В}$) — так называемый стартовый бит. Стоповые биты передаются низким уровнем напряжения, обычно -12 В . При использовании контроля по четности соответствующий бит выбирается таким образом, что сумма бит данных и бит контроля представляет собой четное число. Аналогичным образом выполняется контроль по нечетности.

1.15.3 USB (Universal Serial Bus)

Физически USB представляет из себя две скрученные пары проводов для передачи данных в каждом направлении (дифференциальное включение) и линию питания ($+5\text{ В}$). Структура стандарта USB была описана выше. Один порт может адресовать до 63 устройств через цепочку концентраторов (hubs). Передача данных может осуществляться как в асинхронном, так и в синхронном режимах. USB может обеспечить скорость до 12 Мбит/с. Одним из преимуществ USB является, в частности, возможность соединения с цифровой телефонной линией без дополнительных плат. Все подключенные к USB устройства должны конфигурироваться автоматически.

1.15.4 Инфракрасный порт

Для связи портативных компьютеров с настольными, а также для подключения к ним лазерных принтеров стал использоваться беспроводной интерфейс, работающий в инфракрасном диапазоне волн. Принцип работы инфракрасного порта довольно прост: светодиод (LED), работающий в инфракрасном диапазоне волн, излучает последовательность импульсов, которую принимает соответствующий фотодиод и затем преобразует обратно в электрические сигналы. Подобная связь имеет ряд преимуществ, например низкая цена и невысокое энергопотребление. К тому же она вполне безопасна для здоровья и не вносит дополнительную лепту в электромагнитную несовместимость устройств.

Соответствующая ассоциация IrDA (Infrared Data Association) была основана еще в 1993 году. Тогда ее членами стали свыше 100 компаний. Целью этой организации было создание стандарта на недорогое соединение с использованием инфракрасного света между различными устройствами. Первый такой стандарт появился в июне 1994 года. Он основывался на технологии, предложенной компанией Hewlett Packard — SIR (Serial Infra Red Technology). В частности данная спецификация определяла следующие ключевые параметры: скорость передачи, расстояние между устройствами и угол передачи. Для обеспечения низкой стоимости SIR базируется на стандартном последовательном порте, который имеется у большинства компьютеров и периферийных устройств. Диапазон скоростей передачи составляет от 9600 до 115 200 бит/с. Невысокое энергопотребление накладывает ограничение на предельное расстояние между устройствами: оно не может превышать одного метра. Довольно небольшой угол (30 градусов) позволяет избегать интерференции волн при работе других SIR-устройств.

В апреле 1995 года ассоциация IrDA предложила два расширения существующего стандарта, которые получили название FIR (Fast IR). Скорости передачи в данном случае могут составлять 1,152 или 4,0 Мбита/с. Однако любая система на базе нового стандарта должна обладать обратной совместимостью для работы на скорости 115,2 Кбит/с.

1.15.5 Игровой порт

Игровой (game) порт, или адаптер, как правило, расположен либо на звуковой, либо на многофункциональной плате ввода-вывода (Multi I/O Card). Основным элементом игрового адаптера является аналого-цифровой преобразователь, который позволяет подключать к компьютеру определенные аналоговые устройства, например джойстик. Этот адаптер принимает до четырех цифровых входов типа «включено-выключено» (например, для нажатия кнопок) и до четырех аналоговых входов (изменение параметров).

Одной из особенностей разъема на game-адаптере является то, что к нему в отличие, например, от разъемов последовательного или параллельного порта могут подключаться два джойстика одновременно. Такое соединение выполняется обычно посредством так называемого Y-кабеля. Для тех, кто увлекается компьютерными играми, нет нужды объяснять необходимость наличия именно двух джойстиков. В адресном пространстве портов для игрового адаптера зарезервировано поле адресов от 200h до 207h. На самом деле для этого адаптера достаточно только одного адреса. Для PC/XT это, как правило, 200h, а для PC/AT — 20In.

1.16 Устройства ввода

1.16.1 Клавиатуры

Как известно, клавиатура является пока основным устройством ввода информации в компьютер. В техническом аспекте это устройство представляет собой совокупность механических датчиков, воспринимающих давление на клавиши и замыкающих тем или иным образом определенную электрическую цепь.

Основным назначением клавиатуры любого компьютера является уведомление его о том, что одна или несколько из ее клавиш нажаты или отпущены. Сведения о том, что клавиша отпущена, не менее важны, чем сведения о том, что она нажата. Это особенно важно для клавиш <Shift> (которые изменяют регистр символов, соответствующих нажимаемым клавишам, если одна из этих кнопок находится в нажатом положении в момент нажатия на клавишу символа), а также для клавиш, одновременное

нажатие которых образует определенную управляющую комбинацию (называемую *аккордом* или *сочетанием клавиш*).

Существуют различные способы реализации клавиатур, отличающиеся способом уведомления пользователя о том, что нужная клавиша была нажата или отпущена. В зависимости от назначения существует множество разновидностей клавиатур. В настоящее время наиболее распространены два типа клавиатур: с механическими и с мембранными переключателями. В первом случае датчик представляет из себя традиционный механизм с контактами из специального сплава. Несмотря на то, что эта технология используется уже несколько десятилетий, фирмы-производители постоянно работают над ее модификацией и улучшением. Стоит отметить, что в клавиатурах известных фирм контакты переключателей имеют позолоченное покрытие, что существенно улучшает электрическую проводимость.

Технология, основанная на мембранных переключателях, считается более прогрессивной, хотя особых преимуществ, вообще говоря, не дает. Переключатель в этом случае может представлять собой набор мембран: активная — верхняя, пассивная — нижняя, которые разделены третьей мембраной — прокладкой. Известны также модули, в которых тонкие посеребренные листки пластика разделены небольшим воздушным зазором и проводящей жидкостью.

Но наиболее распространенными являются клавиатуры, изготовленные с применением формованного эластичного полимера. В этих клавиатурах между клавишами и печатной платой располагается прослойка из эластичного полимера (искусственной резины). На прослойке под клавишами клавиатуры расположены бугорки, образованные за счет изгиба прослойки. При нажатии на клавишу бугорок продавливается и его внутренняя сторона прижимается к печатной плате. При этом токопроводящая прокладка, расположенная с обратной стороны каждого бугорка, замыкает контакты, расположенные на печатной плате, и посылает компьютеру сигнал о нажатии клавиши. При правильном подборе параметров характеристики клавиатур данного типа приближаются к параметрам лучших клавиатур «со щелчком», но при этом имеют гораздо меньшую стоимость. Этим объясняется их широкое распространение.

Как правило, внутри корпуса любой клавиатуры помимо датчиков клавиш расположены электронные схемы дешифрации и микроконтроллер клавиатуры. Обмен информацией между клавиатурой и системной платой осуществляется 11-битовыми блоками (8 разрядов данных плюс служебная информация). Несмотря на то, что физически связь осуществляется только по двум проводам (сигнал и «Земля») последовательно, сам интерфейс отличается от стандартного порта RS-232C. В качестве микроконтроллера клавиатуры обычно используются микросхемы 8048, 8049 (со встроенной ROM-памятью) или совместимые с ними.

Основной принцип работы клавиатуры вместе с микросхемой контроллера заключается в сканировании переключателей клавиш. Замыканию и размыканию любого из этих переключателей (то есть нажатию или отпусканию клавиши) соответствует уникальный цифровой код — скэн-код (размером один байт). Скэн-код — это однобайтное число, младшие семь битов которого представляют собой идентификационный номер, присвоенный каждой клавише. Для клавиатур компьютеров PC и PC/XT старший бит кода говорит о том, была клавиша нажата (1) или отпущена (0). Клавиатура для PC/AT работает немного по-другому. В обоих случаях посылается один и тот же скэн-код, который предваряется кодом F0h, когда клавиша отпускается.

Когда скэн-код попадает в контроллер клавиатуры (8042), тут же инициируется аппаратное прерывание IRQ1. Процессор моментально прекращает свою работу и выполняет процедуру, анализирующую скэн-код. Данное прерывание обслуживается специальной подпрограммой, входящей в состав ROM BIOS. Кстати, эта подпрограмма может быть вызвана также через программное прерывание INT 9. Когда поступает код от клавиши сдвига (Shift) или переключателя (Caps Lock), изменение статуса записывается в специальную область памяти. Во всех остальных случаях скэн-код трансформируется в код символа при условии, что он подается при нажатии клавиши, в противном случае скэн-код отбрасывается. После этого введенный код помещается в буфер клавиатуры, который является областью памяти, способной запомнить до 15 вводимых символов, пока программа не может их обработать.

Трудно описать все многообразие клавиатур, выпускаемых в настоящее время. Даже Microsoft выпустила клавиатуру собственной разработки Microsoft Natural Keyboard, которая имеет 104 клавиши. Коротко перечислим другие интересные клавиатуры.

Существуют клавиатуры с дополнительными программируемыми клавишами. Выпускаются клавиатуры, которые умеют читать штриховые коды (bar code) и магнитные карточки, отвечающие стандарту ISO 3554. Для мультимедиа-приложений разработаны даже клавиатуры со встроенными микрофоном и акустическими системами.

В последнее время большой интерес проявляется к так называемым эргономичным клавиатурам, в конструкции которых учтено большинство рекомендаций медиков и специалистов. Такие клавиатуры обычно разбиты на две секции, причем каждая из них может подстраиваться по боковым и фронтальным разворотам индивидуально. Обе секции имеют площадки для отдыха кистей рук.

Отечественных пользователей IBM PC-совместимых компьютеров, видимо, могли бы заинтересовать клавиатуры, называемые обычно Coffee Proof. Дело в том, что конструкция их корпуса такова, что не позволяет жидкости (например, пролитому кофе) проникнуть в электронную «начинку» клавиатуры. При помощи специального устройства клавиши легко снимаются, а влага удаляется обыкновенной тряпкой.

1.16.2 Мыши и трекболы

Трекболы и мышки являются манипуляторными устройствами ввода информации в компьютер. Разумеется, полностью заменить клавиатуру они не могут. Первую компьютерную мышку создал молодой талантливый ученый Дуглас Энджелбарт. Произошло это событие в 1963 году в Стэнфордском исследовательском центре. Первый трекбол (trackball) появился значительно позже на фирме Logitech.

Не секрет, что своей популярностью мышка обязана главным образом громадному спросу на прикладные графические программные системы, а также широкому распространению

графического интерфейса пользователя, чему в немалой степени способствует экспансия Windows.

Теперь немного о мышью «анатомии». Как известно, первая мышка каталась на двух колесиках, которые были связаны с осями переменных резисторов. Перемещение такой мышки было прямо пропорционально изменению сопротивления переменных резисторов. В дальнейшем конструкция мышки претерпела значительные изменения. Колесики (ролики) были перенесены внутрь корпуса, а с поверхностью стал соприкасаться тяжелый обрезиненный или просто сделанный из твердой резины шарик сравнительно большого диаметра. Ось вращения одного из роликов вертикальна, а другого — горизонтальна. Ролики, прижатые к поверхности шарика, установлены на оси с датчиками, с помощью которых и определяются направление и скорость перемещения мыши (рис. 17).

Некоторое время в качестве датчиков использовались непроводящие диски с нанесенными печатным способом контактами, которые поочередно могли соприкасаться с одним неподвижным контактом. При таком внутреннем устройстве мышка была практически полностью механической. Но, как известно, механика — вещь не очень надежная, поэтому впоследствии (да и до сих пор) подавляющее число компьютерных мышек стали использовать оптико-механический принцип кодирования перемещения. На смену механическим шифраторам пришли оптопары: светодиод-фотодиод, или фоторезистор, а в некоторых случаях даже фототранзистор. Такая пара располагается по разные стороны диска с прорезями. Порядок, в котором освещаются фоточувствительные элементы, определяет направление перемещения мыши, а частота приходящих от них импульсов — скорость.

Сегодня не все мышки используют одинаковый способ перемещения. Например, мышка Honeywell имеет уникальный запатентованный дизайн: вместо обычного шара ей «приделаны»

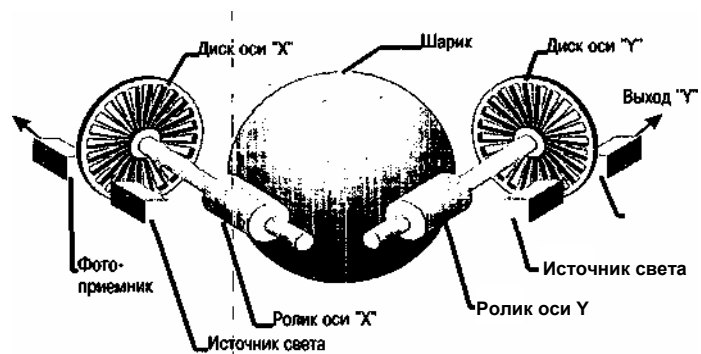


Рис. 17 — Вид роликов мыши

две «ножки». Эти «ножки» являются частью X-Y-механизма оптического-механического шифратора.

Но существуют и полностью оптические мышки. Полностью оптическая мышка является более сложным и дорогим устройством, однако она обладает и существенными достоинствами. Во-первых, в ней отсутствуют движущиеся части, что делает эту мышку практически безотказной, и, во-вторых, как правило, такая мышка обеспечивает более «тонкое» управление курсором на экране. Наиболее распространены мышки либо с двумя, либо с тремя кнопками.

В настоящее время можно выделить три основных способа подключения мышки: через последовательный порт (интерфейс RS-232C, EIA-232D), USB, PS/2. Связи между мышкой и компьютером обеспечивает уже упоминаемая микросхема контроллера 8042.

Беспроводные (cordless) мышки используют передачу данных в радио- или инфракрасном диапазоне волн на расстоянии 1,5–2 м. Во избежание интерференции каждая такая мышь может использовать не один (до четырех) частотный канал. Широкого распространения эти устройства пока не получили. Таким образом, наибольший интерес для массового пользователя представляют только те мышки, которые подключаются через последовательный порт компьютера.

Драйвер определяет направление движения мышки: вверх или вниз, вправо или влево. Это вполне возможно сделать, поскольку 8-разрядные приращения перемещений кодируются в дополнительном коде, и, соответственно, максимальный диапазон перемещения составляет от –128 до +127 единиц. С учетом скорости передачи за каждые 20 мс мышка может передвигаться на 0,62 дюйма. Эта величина вполне достаточна даже при быстром перемещении мышки по коврику. Следует отметить, что, например, некоторые мышки Logitech, использующие последовательный интерфейс, работают на скорости 2400 бит/с.

Современные мышки от Microsoft и Logitech имеют оптимальное аппаратное разрешение 400 dpi. Когда иные фирмы декларируют разрешение на уровне 1800 dpi, то речь, видимо, идет о программном разрешении, то есть о значении, которое может обеспечить соответствующий драйвер.

1.16.3 Сканеры

Сканером называется устройство, позволяющее вводить в компьютер образы изображений, представленных в виде текста, рисунков, слайдов, фотографий или другой графической информации. Кстати, несмотря на обилие различных моделей сканеров, в первом приближении их классификацию можно провести всего по нескольким признакам (или критериям). Во-первых, по степени прозрачности вводимого оригинала изображения, во-вторых, по кинематическому механизму сканера (конструкции механизма движения), в-третьих, по типу вводимого изображения и, в-четвертых, по особенностям программного и аппаратного обеспечения сканера.

1.16.4 Джойстики

Как известно, английское слово joystick состоит из двух слов: joy (радость) и stick (палка). Вообще говоря, джойстик является аналоговым координатным устройством ввода информации. Впрочем, первые модели джойстиков были, можно сказать, «цифровые». Дело в том, что они были основаны на нескольких микропереключателях. При перемещении рукоятки джойстика в зависимости от направления замыкался тот или иной переключатель. Практически любую современную модель джойстика технически можно представить как два реостатных датчика. Рукоятка джойстика связана с двумя переменными резисторами. Один резистор определяет перемещение по координате X, а другой — по Y. В задачу адаптера джойстика входит преобразование изменения параметра сопротивления в соответствующий цифровой код. Разумеется, что разнообразие дизайна джойстиков практически не влияет на их внутреннее устройство.

1.17 Устройства вывода

1.17.1 Принтеры

Пожалуй, самым популярным устройством вывода информации для IBM PC-совместимых компьютеров является принтер. Делаете вы высококачественные иллюстрации или просто выво-

дите на печать текстовые данные, вам в любом случае требуется устройство, именуемое принтером. Разумеется, в каждом конкретном случае может потребоваться модель печатающего устройства, которая использует соответствующую технологию печати, имеет определенную производительность, разрешающую способность и т.п.

1.17.2 Классификация принтеров

Большинство пользователей подразделяет все принтеры только на матричные, струйные и лазерные. Это не совсем так. Стоит начать с того, что практически все современные принтеры (используемые с IBM PC-совместимыми компьютерами) могут быть отнесены к матричным устройствам. Как известно, идея матричных печатающих устройств заключается в том, что все мыслимые (и немыслимые) знаки воспроизводятся ими из набора отдельных точек, наносимых на бумагу тем или иным способом.

Поэтому точнее сказать, что все печатающие устройства можно подразделить на последовательные, строчные и страничные. Принадлежность принтера к той или иной из перечисленных групп зависит от того, формирует он на бумаге символ за символом или сразу всю строку, а то и целую страницу. В свою очередь в каждой группе можно выделить устройства ударного (impact) и безударного (non-impact) действия. Далее принтеры можно подразделить на матричные и символьные (сейчас, кстати, крайне редкие), и только после этого речь может идти об используемой технологии печати (рис. 18). К слову сказать, все принтеры безударного действия являются матричными. Поэтому, вообще говоря, их всех можно назвать матричными печатающими устройствами, потому как даже страничный лазерный принтер формирует изображение из отдельных точек (по строкам) и, разумеется, является матричным. Но называть все принтеры матричными, видимо, все же не стоит.

На практике сложилось так, что, когда говорят о матричных принтерах, обычно имеют в виду устройства ударного действия (impact dot matrix), например всем известные модели Epson'ов, Star'ов и Microlin'ов. Будем придерживаться этой терминологии и мы.

Поскольку трактовка терминов в русском языке еще до конца не устоялась, то по тексту в скобках мы будем использовать и оригинальные названия. Это поможет ориентироваться в англоязычной документации на печатающие устройства.

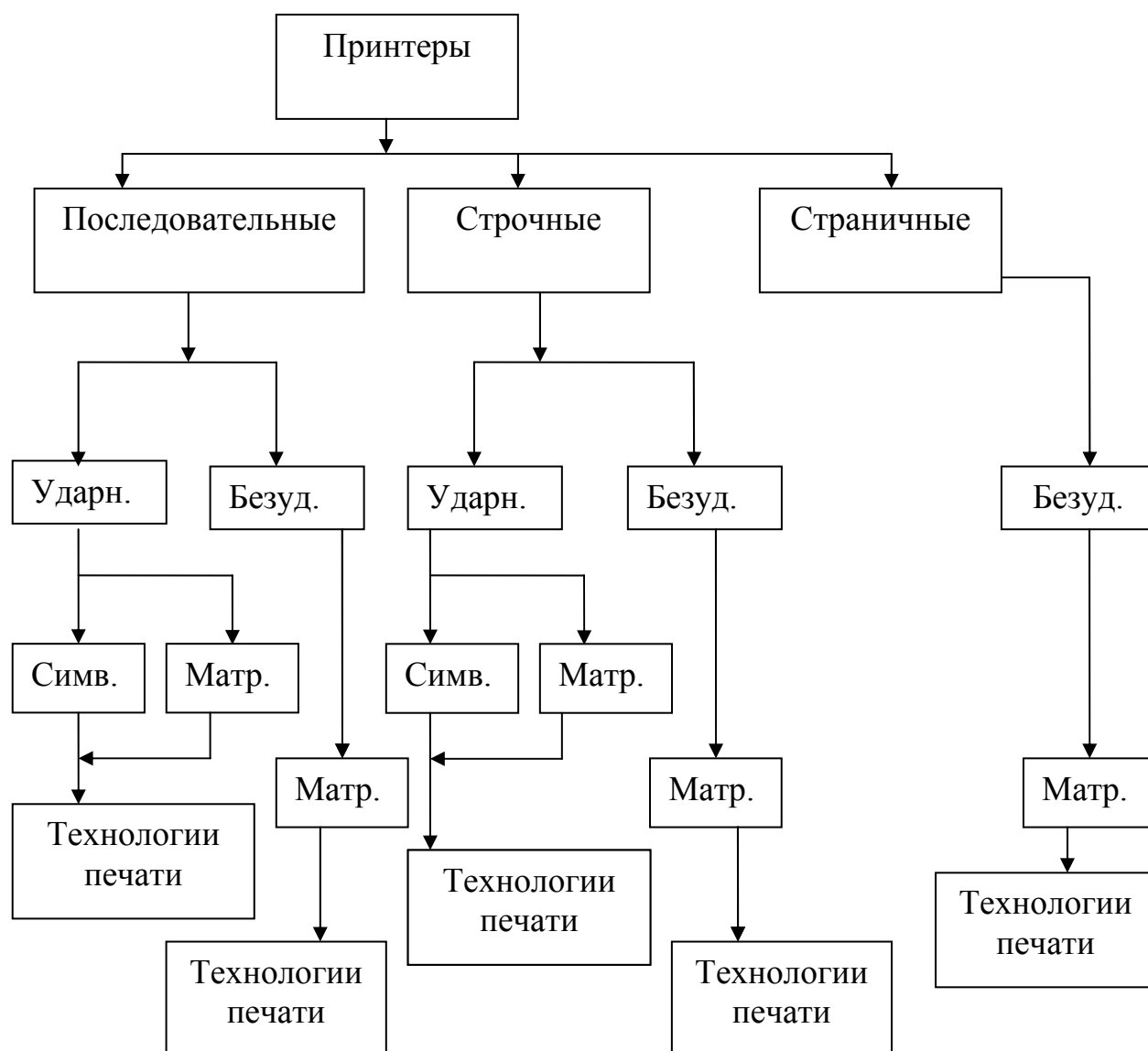


Рис. 18 — Классификация принтеров

1.17.3 Матричные печатающие устройства

Большинство принтеров, работающих (и продаваемых) сейчас с IBM PC-совместимыми компьютерами в нашей стране, могут быть причислены к группе последовательных ударных матричных печатающих устройств (impact dot matrix). Вертикальный

ряд (два ряда) игл, или молоточков, «вколачивает» краситель с ленты прямо в бумагу, формируя последовательно символ за символом. Такое засилье (рис. 19) «игольчатых» вполне объясняется приемлемым качеством их печати, невысокой ценой расходных материалов (красящей ленты) и используемой бумаги, да и са-

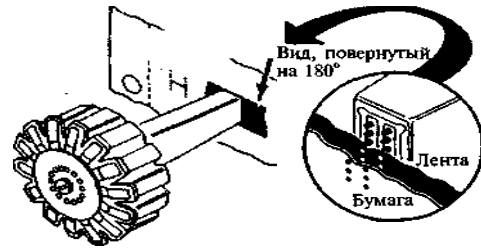


Рис. 19 — Головка матричного принтера

мих устройств. Кстати, для этих принтеров обычно возможно использование как форматной, так и рулонной бумаги. Головка принтера может быть оснащена 9, 18 или 24 иглами. Существуют модели принтеров как с широкой (формат А3), так и с узкой (формат А4) кареткой (рис. 20). Высокое качество печати достигается в режимах NLQ (Near Letter Quality) для 9-игольчатых (почти машинописное) и LQ (Letter Quality) — для 24-игольчатых принтеров. Скорость печати для высокопроизводительных моделей может составлять до 380 знаков в секунду. На рынке последовательных ударных матричных принтеров лидируют фирмы Epson, Star Micronics, Okidata, Mannesman Tally, Samsung, Panasonic. Подача бумаги является важным фактором, влияющим на выбор матричного принтера. Хотя основным преимуществом моделей матричных устройств является возможность печати на рулонной бумаге, следует обращать внимание на наличие в комплекте автоматического загрузчика для форматных листов. Стоит отметить, что в современных моделях могут одновременно использоваться несколько способов подачи бумаги.

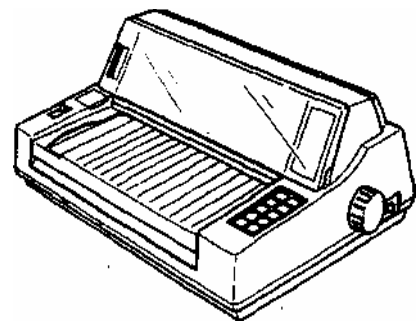


Рис. 20 — Матричный принтер

К недостаткам ударных принтеров прежде всего относится высокий уровень шума. Фирмы—производители устройств, использующих эту технологию, применяют различные технические решения, чтобы по возможности уменьшить шум. Так, фирма Panasonic разработала малошумящий принтер со специальным расположением иглоков в печатающей головке.

1.17.4 Струйные принтеры

Итак, струйные принтеры относятся к безударным печатающим устройствам. Обычно безударными принтерами называются такие устройства, у которых носитель печатаемой информации не касается бумаги. Не требуется, конечно, пояснять, что данные устройства работают практически бесшумно, что является одним из их несомненных преимуществ по сравнению с ударными. Струйные чернильные принтеры относятся, как правило, к классу последовательных матричных безударных печатающих устройств. Если продолжить уточнение признаков принадлежности печатающих устройств к отдельным группам, можно сказать, что последовательные безударные матричные струйные чернильные (liquid ink jet) принтеры в свою очередь подразделяются на устройства непрерывного (continuous drop, continuous jet) и дискретного (drop-on-demand) действия. Последние в своей работе опять же могут использовать либо «пузырьковую» технологию (bubble-jet или thermal ink-jet), либо пьезоэффект (piezo ink-jet). Кстати, «первопроходцами» этих технологий стали фирмы Canon и Epson. Почти все современные устройства этого класса используют именно две последние технологии печати. У чернильных устройств, как, впрочем, и у ударных матричных принтеров, печатающая головка движется только в горизонтальной плоскости, а бумага подается вертикально. Сопла (канальные отверстия) на печатающей головке, через которые разбрызгиваются чернила, соответствуют «ударным» иглам. Количество сопел у разных моделей принтеров, как правило, может варьироваться от 12 до 64. Поскольку размер каждого сопла существенно меньше диаметра иглы (тоньше человеческого волоса), а количество сопел может быть больше, то и получаемое изображение (теоретически) должно быть четче. Как правило, максимальная разрешающая способность достигает значения около 360 точек на дюйм. К сожалению, реальная картина часто отличается от теоретической, поскольку очень многое зависит от качества используемой бумаги.

1.17.5 Лазерные и LEty-принтеры

В лазерных (laser) принтерах используется электрографический принцип создания изображения — примерно такой же, как и в копировальных машинах. Наиболее важными частями лазерного принтера можно считать фотопроводящий цилиндр (печатающий барабан), полупроводниковый лазер и прецизионную оптико-механическую систему, перемещающую луч. Мощный полупроводниковый лазер генерирует тонкий световой луч, который, отражаясь от вращающегося зеркала, формирует электронное изображение на светочувствительном фотоприемном барабане. Барабану предварительно сообщается некий статический заряд. Для получения изображения лазер должен включаться и выключаться, что обеспечивается специальной управляющей электроникой принтера. Вращающееся зеркало служит для разворота луча лазера на новую строку, формируемую на поверхности печатающего барабана. Когда луч лазера попадает на предварительно заряженный барабан, заряд «стекает» с освещенной поверхности. Таким образом, освещаемые и не освещаемые лазером участки барабана имеют разный заряд. В зависимости от того, как (положительно или отрицательно) заряжены частицы порошкообразного тонера, они будут притягиваться и прилипать к барабану только в областях с разноименным зарядом. После формирования каждой строки специальный шаговый двигатель поворачивает барабан так, чтобы можно было формировать следующую строку. Это смещение равняется разрешающей способности принтера и может составлять, например, 1/300, 1/600 или 1/1200 дюйма. Данный этап работы во многом напоминает построение изображения на экране монитора (растрирование).

Когда изображение на барабане построено и он покрыт тонером, подаваемый лист заряжается таким образом, чтобы тонер с барабана притягивался к бумаге. После этого изображение закрепляется на ней за счет нагрева частиц тонера до температуры плавления. Окончательную фиксацию изображения осуществляют специальные резиновые валики, прижимающие расплавленный тонер к бумаге.

На рынке лазерных принтеров можно выделить печатающие устройства малого быстродействия (скорость вывода — 4–6

страниц в минуту), принтеры среднего быстродействия (7–11 страниц в минуту) и принтеры коллективного использования, так называемые сетевые принтеры (более 12 страниц в минуту).

К наиболее важным функциональным возможностям принтеров относятся такие, как поддержка технологии повышения разрешающей способности, наличие масштабируемых шрифтов (PostScript, TrueType), объем оперативной памяти и т.п. Безусловным лидером на рынке лазерных принтеров малого быстродействия является фирма Hewlett Packard. Хорошо зарекомендовали себя также модели фирм Lexmark, Canon, Epson, QMS, Minolta.

Кстати, первый принтер с действительным разрешением 1200 на 1200 dpi выпустила фирма Lexmark. Как правило, производители принтеров указывают, что их изделия обеспечивают так называемое алгоритмическое разрешение 1200 точек на дюйм. А это обычно означает, что разрешение составляет все-таки 1200 на 600 точек на дюйм, то есть шаг вращения барабана остается равным 1/600 дюйма. Алгоритмическое разрешение достигается за счет того, что механизм подобных принтеров позволяет слегка изменять положение луча по вертикали (напомним, что развертка луча бывает, как правило, только горизонтальной). В результате темная точка на бумаге появляется либо в верхней, либо в нижней части прямоугольника высотой 1/600 дюйма. Хотя подобная технология действительно позволяет сделать края изображения более гладкими, однако понятно, что в столбике высотой один дюйм чисто физически не может быть больше 600 точек. А ведь максимальное количество точек важно не только для хорошего качества черного цвета, но и для передачи полутоновых изображений.

Кроме лазерных принтеров (рис. 21), существуют так называемые LED-принтеры (Light Emitting Diode), которые получили свое название из-за того, что полупроводниковый лазер в них был заменен «гребенкой» мельчайших светодиодов. Разумеется, в данном случае не требуется сложная оптическая

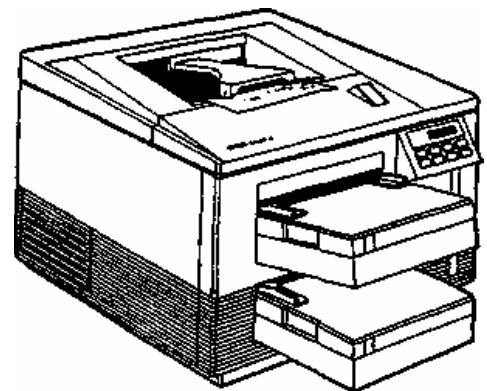


Рис. 21 — Лазерный принтер

система вращающихся зеркал и линз. Изображение одной строки на светочувствительном барабане формируется одновременно. Одним из лидеров на рынке LED-принтеров можно назвать фирму Okidata, хотя подобные устройства сегодня выпускают еще ряд фирм.

1.17.6 Цветная печать

На сегодняшний день для цветной печати используются все описанные выше технологии плюс еще две: принтеры с термосублимацией красителя (dye sublimation) и с изменением фазы красителя (phase-change ink-jet), но о них мы поговорим чуть позже. Теперь же коротко остановимся на технологии цветопередачи для печатающих устройств.

Напомним, что при использовании RGB-модели цветообразования — Red (красный), Green (зеленый) и Blue (синий) — указанные цвета называются первичными, поскольку путем сложения соответствующего их количества можно получить любой другой цвет вплоть до белого. RGB-модель цветообразования применяется во всех мониторах персональных компьютеров и называется также аддитивной (addition — сложение). Кстати отметим, что черный цвет для монитора, вообще говоря, не чернее поверхности его экрана в выключенном состоянии.

Печатающие устройства работают с другими первичными цветами и используют соответственно иную модель цветообразования — так называемую субтрактивную (subtraction — вычитание). Первичными цветами для цветных принтеров являются зелено-голубой (Cyan), светло-пурпурный (Magenta) и желтый (Yellow). Наложение двух из этих первичных цветов в данном случае дает красный, зеленый или голубой цвет, а смешение всех трех первичных цветов субтрактивной модели — черный цвет. Поясним, почему собственно различаются модели цветообразования для мониторов и принтеров. Напомним, что наши глаза являются сложной оптической системой, которая воспринимает излучаемый или уже отраженный от освещаемых предметов свет, разумеется, если они сами его не излучают. Цвет, как известно, определяется длиной волны электромагнитного излучения, определенный частотный спектр которого и представляет для нас

видимый свет. Нанесенные на экран точки люминофора воспринимаются именно того цвета, какой они и излучают. Краситель же, нанесенный на бумагу, напротив, действует как фильтр, поглощая (вычитая) одни и отражая другие длины электромагнитных волн. Насыщенность цвета (розовый, красный, пурпурный) зависит от количества белого цвета. Таким образом, промежуточные цвета при выводе изображения, например розовый, получаются, как правило, путем пропуска (не печати) нескольких точек. Собственно, это обычный подход, связанный с растриванием изображения (dithering), то есть оттенки соответствующего цвета получаются путем группировки нескольких точек изображения в псевдопиксели размером 2х2, 3х3 и более точек.

1.17.7 Модемы

Для связи удаленных компьютеров друг с другом могут использоваться обычные телефонные сети, которые в той или иной степени покрывают территории большинства государств. Единственной проблемой в этом случае является преобразование цифровых (дискретных) сигналов, с которыми оперирует компьютер, в аналоговые (непрерывные). Как известно, до недавнего времени широкое распространение имели именно аналоговые телефонные сети, например PSTN (Public Switchable Telephone Network) — коммутируемая телефонная сеть общего назначения. Здесь на помощь компьютерам приходят модемы. Итак, модемом называется устройство, способное осуществлять модуляцию и демодуляцию информационных сигналов (МОдуляция-ДЕМОдуляция) (рис. 22). Собственно работа модулятора модема заключается в том, что поток битов из компьютера преобразуется в аналоговые сигналы, пригодные для передачи по телефонному каналу связи. Понятно, что демодулятор модема выполняет обратную задачу.

Таким образом, данные, подлежащие передаче, преобразуются в аналоговый сигнал модулятором модема «передающего» компьютера. Принимающий модем, находящийся на противоположном конце линии, «слушает» передаваемый сигнал и преобразует его обратно в цифровой при помощи демодулятора. После

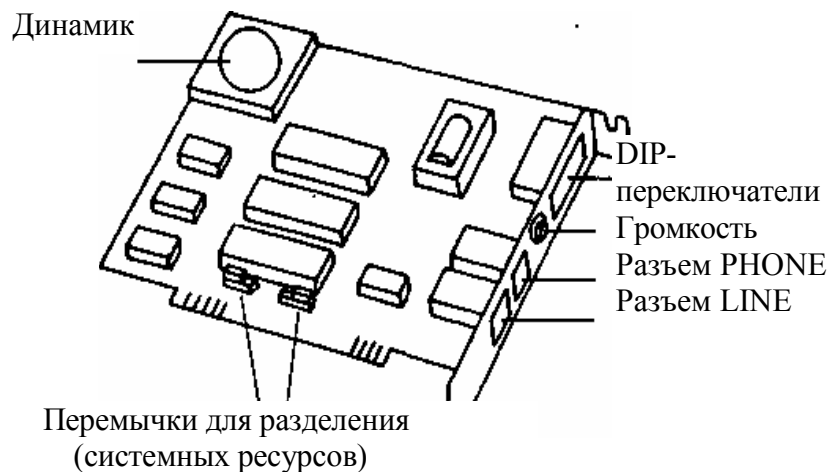


Рис. 22 — Модем

того, как эта работа выполнена, информация может передаваться в принимающий компьютер. Режим работы, когда передача данных осуществляется только в одном направлении, называется полудуплексным (half duplex). Вообще говоря, оба компьютера, как правило, могут одновременно обмениваться информацией в обе стороны. Этот режим работы называется полным дуплексом или просто дуплексом (full duplex).

Можно выделить некоторые основные этапы работы модема. Первым делом модем принимает данные, поступающие из компьютера, после чего разделяет их на исполняемые команды и информацию, которую необходимо передать в линию. Сразу же заметим, что большинство современных модемов используют так называемый набор команд AT (сокращение от слова ATtention). Поскольку этот набор команд был в свое время разработан фирмой Hayes Microcomputer Product, то использующие его модемы называют Hayes-совместимыми. Сегодня они составляют подавляющее большинство среди подобных устройств.

Кроме собственно модуляции и демодуляции сигналов, модемы могут выполнять сжатие и декомпрессию пересылаемой информации, а также заниматься поиском и исправлением ошибок, возникнувших в процессе передачи данных по линиям связи.

Модемы могут отличаться друг от друга, например, по методам модуляции. Ведь, как известно, у одного и того же сигнала, определенного во времени, можно модулировать амплитуду, частоту и фазу. Наиболее известны три метода модуляции: FSK

(Frequency Shift Keying), PSK (Phase Shift Keying) и QAM (Quadrature Amplitude Modulation). FSK является разновидностью частотной модуляции (ЧМ), а PSK — фазовой (ФМ). В методе квадратурной амплитудной модуляции QAM одновременно изменяются фаза и амплитуда сигнала, что позволяет передавать большее количество информации. В современных модемах используется так называемая модуляция с решетчатым кодированием TCQAM (Trellis Coded QAM) или просто TCM. Теперь о скорости передачи.

Одной из основных характеристик модема является скорость модуляции (modulation speed), которая определяет физическую скорость передачи данных без учета исправления ошибок и сжатия данных. Единицей измерения этого параметра является количество бит в секунду (бит/с). Скорость модуляции не следует путать с пропускной способностью канала (throughput), которая может быть меньше или больше скорости модуляции в зависимости от качества линии, применения коррекции ошибок и сжатия передаваемых данных. Поскольку скорость передачи данных может измеряться как в битах в секунду, так и в бодах, то следует отметить, что это, вообще говоря, — единицы разные. Дело в том, что бод определяет число изменений (модуляций) сигнала в секунду. Однако в зависимости от способа модуляции каждое изменение сигнала может соответствовать не только одному, но и большему количеству бит.

1.17.8 Факс-модемы

Заметим, что сама идея передачи изображений по линиям проводной связи не нова и впервые была предложена еще в 1842 году, а первая работающая факсимильная машина появилась на свет спустя 12 лет. Правда, затем эта идея была забыта почти на столетие.

Система должна обеспечивать сканирование документа на передающей стороне, преобразование информации в форму, пригодную для передачи по имеющемуся каналу связи, и формирование на бумажном носителе на приемной стороне дубликата — факсимиле — исходного документа. Факсимильные машины стали популярны во второй половине 60-х годов, когда МККТТ при-

нял набор стандартов для телефаксов, названных Group 1 и Group 2 (или G1, G2). Однако телефаксы, отвечающие требованиям G1 или G2, были исключительно аналоговыми и страдали неточностью передачи информации. Существенный скачок в развитии факсимильной связи произошел в 1980 году, когда ССИТТ ввел стандарт Group 3 (G3), который определял методы цифрового сканирования и сжатия информации. Иными словами, избыточность информации в сигнале снижалась еще до ее передачи (модуляции). В отличие от первых аналоговых телефаксов, работающих со скоростью 300 бит/с, аппарат, отвечающий стандарту G3, передает информацию со скоростью 9600 бит/с, что позволяет отправлять страницу документа менее чем за минуту (обычно за 15–20 секунд). Сам процесс модуляции-демодуляции сигнала определяется стандартом V.29.

Эффективные методы сжатия позволяют в некоторых случаях сократить количество передаваемой информации на 70–80%.

1.17.9 Звуковые карты

Как явствует из названия, звуковые карты используются для записи и воспроизведения различных звуковых сигналов: речи, музыки, шумовых эффектов. Любая современная звуковая карта может использовать несколько способов для воспроизведения звука. Одним из простейших способов является преобразование ранее оцифрованного непрерывного (аналогового) сигнала снова в аналоговый. Для этого используются микросхемы ЦАП (цифро-аналоговых преобразователей). Цифровые выборки реального звукового сигнала хранятся в памяти компьютера (например, в виде WAV-файлов) и преобразуются в аналоговый сигнал через ЦАП по мере необходимости. Кстати, для записи сигнала применяются микросхемы аналогоцифровых преобразователей (АЦП), то есть устройства, способные формировать из аналогового эквивалентный цифровой сигнал.

Качество записываемого и воспроизводимого сигналов зависит от разрядности (бит) и частоты преобразования (кГц) применяемых АЦП и ЦАП. В зависимости от разрядности АЦП-ЦАП карты условно подразделяются на 8- и 16-разрядные (вообще говоря, АЦП-ЦАП бывают также 10- и 12-разрядными). 8-разрядное

преобразование может обеспечить качество звучания кассетного магнитофона, а 16-разрядное ассоциируется обычно с качеством аудио-компакт-диска. Аппаратные средства, необходимые для прямой записи и воспроизведения сигнала, часто называют цифровым аудиоканалом (digital audio channel).

Другой способ воспроизведения звука заключается в его синтезе. При поступлении на синтезатор некоторой управляющей информации по ней формируется соответствующий выходной сигнал. В настоящее время применяются две основные формы для синтеза звукового сигнала. Это синтез с использованием частотной модуляции, или FM-синтез, и синтез с применением таблицы волн (WaveTable) — так называемый табличный, или WT-синтез. Разумеется, поскольку синтез в звуковых картах может быть только цифровой, то для преобразования выходной информации в непрерывный сигнал также используются микросхемы ЦАП. Совокупность микросхемы синтезатора и ЦАП называют набором. Так, если набор OPL 3 обеспечивает только FM-синтез, то OPL 4 поддерживает как FM-, так и WT-синтез. В последнем случае звучание наиболее приближено к естественному.

Кстати, управляющие команды для синтеза звука могут поступать на звуковую карту не только от компьютера, но и от другого, например MIDI, устройства. Собственно MIDI (musical instruments digital interface) определяет протокол передачи команд по стандартному интерфейсу. Вообще говоря, MIDI-сообщение содержит ссылки на ноты, а не запись музыки как таковой. В частности, когда звуковая карта получает подобное сообщение, оно расшифровывается (какие ноты каких инструментов должны звучать) и отрабатывается на синтезаторе. В свою очередь компьютер может через MIDI управлять различными «интеллектуальными» музыкальными инструментами с соответствующим интерфейсом.

Отдельные узлы звуковой карты (фильтры, выходной усилитель) могут использоваться при воспроизведении звука с аудио-компакт-диска, при этом интерфейс соответствующего привода CD-ROM может располагаться также на звуковой карте.

Некоторые звуковые карты оснащают сигнальными процессорами DSP (Digital Signal Processor). Это обеспечивает существенное увеличение скорости работы при компрессии и деком-

прессии звуковых файлов для звукового аудиоканала, так как они обычно занимают много места. Небесполезным оказывается использование DSP и при WT-синтезе. Как правило, настоящий DSP — достаточно дорогое устройство, поэтому сразу устанавливается только на профессиональных музыкальных картах. Большинство пользователей пока могут спокойно обходиться без DSP таким же образом, как масса людей, не использующих в своей работе математического сопроцессора в компьютере

Стереозвучание — далеко не предел в стремлении к естественному звучанию. Одним из путей реализации этого стремления стала технология так называемого объемного, или трехмерного, звучания (иначе, 3D-звучания), названного так по аналогии с трехмерным изображением. Например, карты серии Sound Blaster 16 ASP (со встроенным довольно специализированным DSP) могут загружать специальное программное обеспечение (по лицензии фирмы Q-Sound), которое и позволяет получить объемное 3D-звучание. Стоит отметить, что программные приложения должны быть специально ориентированы на Q-Sound. В последнее время стали появляться звуковые карты, оснащаемые дочерними платами, которые обеспечивают для большинства приложений так называемый псевдо 3D-эффект.

1.17.10 Акустические системы

Акустические системы (динамики или колонки) являются, вообще говоря, неотъемлемой частью звуковой карты, за исключением того случая, когда вы предпочитаете слушать музыку через головные телефоны (наушники). В настоящее время существуют две основные разновидности акустических систем: со встроенным выходным усилителем (так называемые активные системы) и без него (пассивные системы). Если пассивные системы подключаются только к соответствующему выходу звуковой карты, то для активных необходим дополнительный источник энергии. В качестве такого источника может выступать либо батарея гальванических элементов, либо блок питания, который в свою очередь может быть как встраиваемым, так и внешним. Кроме регулировки громкости активные системы имеют обычно 3-полосный эквалайзер. Вместе с компьютером необходимо использовать только

экранированные (shielded) колонки — они могут быть расположены непосредственно рядом с монитором и не приведут к постепенному размагничиванию ЭЛТ. Как мы уже отмечали, существуют не только мониторы со встроенными акустическими системами, но даже и клавиатуры. Разумеется, для получения высококачественного звучания стереосистемы должны быть внешними.

2 ФИЗИЧЕСКОЕ ОКРУЖЕНИЕ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ

К сожалению, при эксплуатации вычислительной техники ее физическому окружению уделяется, как правило, мало внимания. Тем не менее исследования доказали, что подавляющее большинство поломок или сбоев в работе вычислительных систем происходит именно из-за проблем, возникающих в ее физическом окружении.

Классифицировать их можно следующим образом:

- климатические (окружающая среда),
- проблемы электропитания,
- электромагнитные,
- электростатические.

Рассмотрим каждый из пунктов более подробно.

2.1 Климатические условия

Большинство компонентов, составляющих современную вычислительную технику, рассчитаны для работы в нормальных климатических условиях (температура окружающего воздуха 15–40 °С, влажность и т.д.), за исключением устройств, специально разработанных для работы в других условиях. Несоблюдение этих требований может привести как к выходу из строя всей системы в целом, так и отдельных ее компонентов.

Кроме того, ряд устройств, имеющих повышенную теплоотдачу, в настоящее время снабжается специальными системами охлаждения. Выход из строя этих систем может привести к перегреву устройства и, возможно, его поломке. Даже если этого не произойдет, велика опасность повышения температуры внутри корпуса всего устройства, что может создать эффект «chip steer», когда микросхемы или другие компоненты самодемонтируются либо теряют контакт с разъемом.

В качестве рекомендаций по предотвращению подобных проблем необходимо указать следующее:

- регулярно (раз в 3–4 месяца) проводить ревизию систем охлаждения и теплоотвода (вентиляторов и т.д.), вовремя очи-

щать накопившуюся на них пыль и грязь;

- следить за качеством воздуха в помещении, температурой и влажностью;
- курение в помещении, в котором установлена вычислительная техника, сокращает срок ее эксплуатации на 40–50% из-за осаждения табачного дыма на контактах и, как следствие, их окисления.

2.2 Проблемы электропитания

К сожалению, вопросам, связанным с электропитанием вычислительных систем, не всегда уделяется должное внимание. А имеющаяся статистика свидетельствует, что по причинам, связанным со сбоями в электросети, в 75 % случаев происходит потеря информации и в 65% — выходит из строя само электронное оборудование. В некоторых случаях убытки от потери информации гораздо больше, нежели затраты на ремонт и замену оборудования, вышедшего из строя.

Важнейшим моментом при оборудовании помещения, в котором будет эксплуатироваться вычислительная техника, является правильная разводка линий электропитания (220 В). Желательно, например, чтобы розетки для подключения персональных компьютеров и периферийных устройств отличались от розеток для подключения мощных бытовых приборов: пылесосов, полотеров и т.п. Все узлы одного компьютера и подключенное к нему периферийное оборудование должны запитываться от одной фазы электросети. Подавляющее большинство компьютерного оборудования имеет разъемы электропитания с тремя (!!!) контактами, третий из которых является заземлением. Ни в коем случае не следует пытаться адаптировать эту «трехточку» в обычную бытовую розетку электропитания, используемую в России. По всем правилам должны быть выполнены и шины «земля»: радикально с одной общей точкой. Для отключения компьютерного оборудования должен использоваться отдельный щит с автоматами защиты и общим рубильником.

Однако, несмотря на соблюдение всех этих простых правил, значительная часть проблем, связанных с качественным электро-

питанием вычислительной техники по-прежнему остается. Их можно классифицировать следующим образом:

- полное отключение сетевого напряжения;
- кратковременные его провалы;
- перенапряжение;
- гармонические искажения;
- электромагнитные и радиочастотные шумы.

Рассмотрим каждую из этих причин более подробно.

Очевидно, что самая простая для понимания неприятность с электропитанием – его полное отключение (blackout). Причины этого не столь уж редкого события достаточно разнообразны: от форс-мажорных до комических. К нему могут привести, например, как авария на электрической подстанции, замыкание проводов в ураган, так и простое отключение рубильника на распределительном щите, или, ставшие в последнее время «модными», неожиданные веерные отключения электроэнергии. К сожалению, вне зависимости от причины, вызвавшей отключение питания, последствия этого бывают, как правило, самыми грустными. Это может быть и полная потеря информации (данных) на «электронных» дисках и в кэш-памяти, а при работе в сетевой операционной системе наиболее вероятен крах таблиц размещения файлов на диске и в самом худшем случае — повреждение электронных элементов самого компьютера.

Не так заметны для пользователя кратковременные «провалы» питающего напряжения в течение долей секунды, так называемые «sags» или «brownout». Причины кратковременного понижения напряжения питающей сети кроются обычно во включении мощных электрических машин, например лифтов, различных двигателей, электрических чайников и т.п. Этот вид помех является основным. Они существуют всегда, поскольку всегда существует необходимость включения либо выключения какой-либо аппаратуры, но нормальное их значение не должно превышать 3–5 % от напряжения питания. «Плавающее» (подвижное во времени, но не периодическое) понижение питающего напряжения (rolling brownout) вызывается обычно включением сварочных аппаратов, компрессоров, приборов кондиционирования воздуха и, как правило, приводит к потере данных, находящихся в энергонезависимой памяти. Кратковременные повышения пи-

тающего напряжения сети также ничего хорошего для вычислительной техники не несут. Различают два вида подобного нарушения электропитания: кратковременное повышение на доли секунды (surge) и импульсное повышение с амплитудой не менее 100% от номинального и длительностью 0.5–100 мкс (spike). Такие импульсные помехи вызываются обычно плохим состоянием осветительной аппаратуры, аварией электрических машин или замыканием проводов. Этот вид помех труднее всего обнаружить, и, следовательно, устранить. Вообще, любое повышение питающего напряжения выше допустимого уровня в пике или по среднеквадратическому значению ведет к повреждению электронных компонентов аппаратуры и преждевременному выходу их из строя. Дело в том, что обычные блоки питания компьютеров и периферийных устройств рассчитаны на работу с питающим напряжением 220–240 В и не обеспечивают защиты от вышеперечисленных факторов.

Способы минимизация риска возникновения подобных помех:

- консультации с фирмами, специализирующимися на данной проблематике (что в России делать пока не принято);
- использование различных устройств защиты.

2.2.1 Классификация устройств защиты

Самую простейшую защиту по питанию обеспечивают так называемые ограничители перенапряжения. Они способны предохранить нагрузку от различного рода выбросов и всплесков питающего напряжения. Подобные устройства выпускают, например, фирмы American Power Conversion (APC), TrippLite, Best Power Technology.

Более высокий уровень защиты обеспечивают устройства нормализации, которые «очищают» питающее напряжение от всевозможных шумов и позволяют регулировать его в некотором диапазоне. Некоторые из них способны предотвратить даже кратковременные провалы в питающем напряжении. В том случае, если в данном устройстве используется технология феррорезонансного преобразования, оно способно обеспечить полную развязку по частоте, не допуская проникновения ВЧ-шумов в це-

в цепи нагрузки.

Большая часть «необъяснимых» повреждений системных, модемных, сетевых и т.п. плат зачастую является следствием импульсов высокого напряжения, попадающих в интерфейсный порт не по цепи питания, а по кабелям данных. Чтобы избежать подобных эффектов, необходимо использовать дополнительные устройства.

Обеспечить работу нагрузки при полном отключении электропитания (blackout) могут только устройства, называемые UPS (Uninterruptible Power Supply) или ИБП (источник бесперебойного питания). Функционально такое устройство почти всегда состоит из устройства подавления помех, зарядного устройства, батареи аккумуляторов (обычно свинцово-кислотных) и преобразователя напряжения. Две наиболее многочисленные группы ИБП составляют устройства, имеющие топологию on-line (постоянно включенные) и off-line или standby (резервные). Схема «классического» ИБП приведена на рисунке 23.

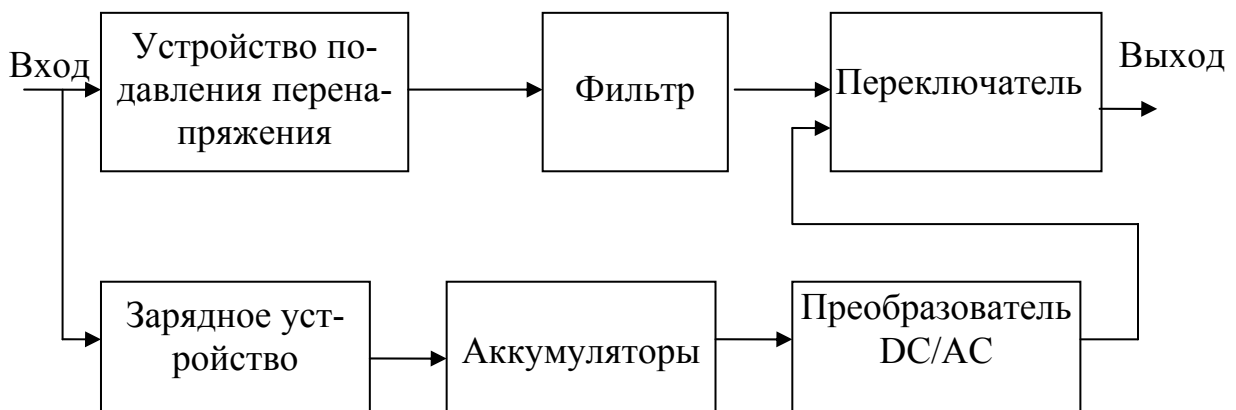


Рис. 23 — Источник бесперебойного питания

ИБП, относящиеся к последней группе, можно в свою очередь разделить на две подгруппы: standby hybrid UPS и standby-ferro UPS (гибридные и феррорезонансные). Существует еще подгруппа устройств, выполненных по топологии line — interactive (интерактивные ИБП), хотя чаще всего их относят к типу standby (или hybrid) UPS.

Постоянно включенные ИБП (работающие в режиме on-line) обеспечивают электроснабжение подключенных устройств от бата-

реи аккумуляторов через преобразователь напряжения независимо от состояния электросети, в то время как резервные UPS переходят на такой режим работы только при полном отключении питающего напряжения.

Для постоянно включенных ИБП в рабочем режиме используется «ветка», включающая в себя зарядное устройство, аккумуляторы и преобразователь. Таким образом, электропитание потребителей происходит от заведомо «чистого» источника и не зависит от «капризов» электросети. Зарядное устройство в этом случае должно быть достаточно мощным, что приводит, естественно, к большим габаритам самого ИБП. В случае выхода из строя какого-либо компонента рабочей ветки подобного ИБП питание потребителей осуществляется в резервном режиме непосредственно от сети по второй «ветке» через стабилизатор. Из-за высоких требований к компонентам (особенно аккумуляторам), классическая схема включения on-line на практике не используется. Вместо нее применяются различные схемы двойного и тройного преобразования напряжения: переменное — постоянное, постоянное — постоянное (с использованием широтно-импульсной модуляции), постоянное — переменное. В случае двойного преобразования выпрямитель преобразует переменный ток в постоянный. При этом зарядное устройство одновременно подзаряжает батарею аккумуляторов. Затем инвертор преобразует постоянный ток в переменный. В случае сбоя питания инвертор питается постоянным током от батареи. В нормальном режиме работы аккумулятор в этом случае не разряжается. Это позволяет гарантировать коэффициент гармонических искажений не более 5% при падении напряжения на входе устройства даже на 50%.

Для резервных ИБП все выглядит с точностью до наоборот: «ветка» со стабилизатором является основной, а с аккумулятором — резервной. В качестве примера резервного ИБП можно привести модели серии Back-UPS фирмы APC или Patriot фирмы Best Power Technology (BPT).

Функциональная блок-схема, поясняющая принцип действия интерактивного ИБП, приведена на рисунке 24.



Рис. 24 — Интерактивный ИБП

Одним из основных отличий от классической топологии standby UPS является наличие узла Smart-Boost. Это позволяет при кратковременных провалах напряжения до 12% от номинального не переходить на питание от аккумуляторов, а «вытягивать» уровень выходного напряжения за счет входного. Преимущества такого решения особенно проявляются в «грязной» электросети, то есть там, где падение напряжения происходит очень часто. В этом случае обычный резервный ИБП работает практически только на аккумуляторах, которые сравнительно быстро разряжаются. Кстати говоря, в современных моделях ИБП Smart-Boost работает не только как стабилизатор, то есть не только увеличивает, но и уменьшает входное напряжение. Другое важное отличие от классической схемы это то, что преобразователь напряжения постоянно подключен к выходу ИБП. Таким образом, при размыкании переключателя на выходе уже присутствует соответствующее напряжение питания. Примеры: Fortress (BPT), Smart UPS (APC). Причем последние наиболее популярны.

Основным узлом феррорезонансных ИБП является феррорезонансный трансформатор, который имеет две первичные обмотки (рис. 25). В нормальном режиме работы напряжение от сети поступает через переключатель на одну из первичных обмоток трансформатора, а в случае сбоя питания от аккумулятора — через преобразователь на другую. В нормальном режиме работы трансформатор выполняет функции стабилизатора и сетевого фильтра. В случае сбоя в электросети энергии, накопленной в магнитном поле трансформатора, хватает на питание нагрузки в течение 10–15 мс. За это время в работу включается инвертор, переключающий питание на резервный режим. Наличие феррорезонанс-

ного преобразования позволяет гарантировать высокий уровень гальванической развязки, практически синусоидальную форму выходного напряжения, а также исключить большинство «неприятностей» в электропитании (особенно импульсные помехи).

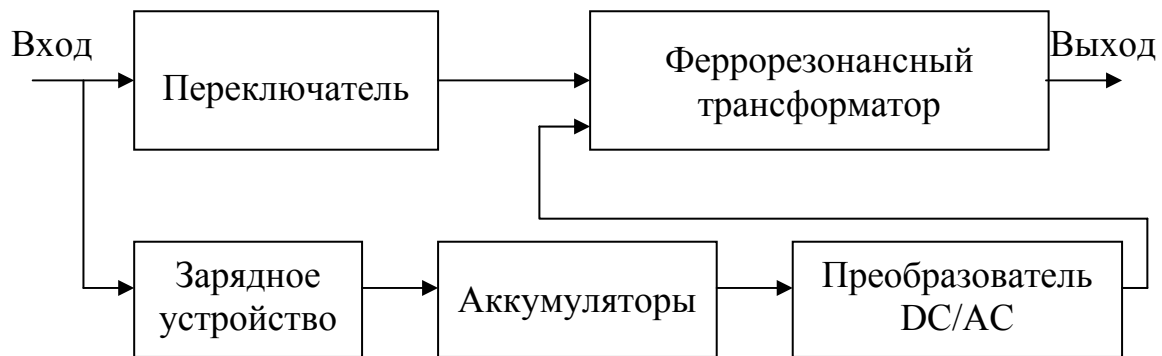


Рис. 25 — Феррорезонансный ИБП

Отдельная группа устройств — непрерываемые аккумуляторные системы UBS (Uninterruptible Battery System). По сути, любая из этих систем это маленькая электростанция: управляемый микропроцессором источник постоянного тока с приводом от двигателя (система «мотор-генератор»). При использовании UBS можно гарантировать надежное электропитание в течение часов, дней и даже недель, при этом не требуется замены дорогостоящих элементов (аккумуляторов).

Для локальных вычислительных систем большое значение имеет автоматический контроль состояния ИБП, подключенного к серверу. С этой целью в сетевые операционные системы включают специальные программы, а ИБП доукомплектовывается соответствующими платами контроля (UPS Monitoring Board). Для соединения с ИБП используется либо специальный интерфейс, либо стандартный последовательный интерфейс RS 232.

В заключение следует отметить, что европейский стандарт EN50091 Part 3 определяет такие категории ИБП, как пассивный резервный (Passive Standby), активный резервный (Active Standby) и постоянного действия (Continuous Operation). Активный ИБП отличается от пассивного наличием узла типа Smart-Boost. Кроме того, в активных ИБП нагрузка питается от UPS

в случае сбоя по питанию только через инвертор. Ключевым элементом ИБП постоянного действия является питание нагрузки через инвертор, который, как правило, от нее гальванически развязан.

2.2.2 Выбор подходящего ИБП

В настоящее время на отечественном рынке предлагаются средства защиты по электропитанию, выпускаемые, например, фирмами American Power Conversion, Fiskars Power Systems, Victron, Upsonic, Leadman Electronics, Para Systems, Tripp Lite, Exide Electronics и рядом других. Для того чтобы выбрать наиболее подходящую для Вас модель, необходимо знать ответы на некоторые вопросы.

Во-первых, какие проблемы с электропитанием необходимо решить. При минимальных требованиях вполне можно обойтись ограничителем перенапряжения, а при самых жестких условиях вам понадобятся непрерываемые аккумуляторные батареи. Таким образом, первое, что нужно сделать, это определить уровень защиты.

Если Вы решили, что без ИБП не обойтись, следующим шагом должен быть выбор топологии UPS. Очень редко, только когда требуется бесперебойное электропитание измерительных устройств, Вам будет необходим постоянно включенный ИБП. Это обусловлено возможной потерей невосполнимой информации при переключении режимов работы. В большинстве случаев вполне достаточным оказывается выбор интерактивного ИБП.

Немаловажным вопросом является мощность подключаемой к ИБП нагрузки. Требуемую величину можно определить, просуммировав мощности всех подключаемых к ИБП устройств. Как известно, потребляемая мощность подобных устройств обычно определяется в ваттах (Вт), а мощность UPS — в вольт-амперах (ВА). Для вычислительной техники пересчет этих значений производится следующим образом: $ВА = Вт \times (1,5 \text{ или } 1,6)$.

Замечание: принтеры, особенно лазерные, подключать к ИБП не рекомендуется. Мощность, которую они потребляют, — велика, а бесперебойная их работа не так уж и важна.

Следующий вопрос — время автономной работы (то есть

при отключении напряжения в электросети) на нагрузке определенной мощности, которое может обеспечить данная модель UPS. Обычно мощность ИБП выбирают таким образом, чтобы она превышала потребляемую на 25–30%. В таком случае время работы от аккумуляторов должно, как минимум, составлять 5–6 минут. В некоторой литературе общую мощность рекомендуют увеличивать вдвое.

Необходимо обратить внимание на форму выходного питающего напряжения, обеспечиваемого конкретным ИБП (лучше всего — «чистый синус»). В случае почти прямоугольной формы генерируемого напряжения (как правило, это дешевые модели) от покупки такого устройства лучше всего воздержаться.

Как известно, все аккумуляторы имеют свой срок службы, по истечении которого их следует заменить. В современных ИБП чаще всего используют различные модели кислотных и никель-кадмиевых аккумуляторов, срок службы которых составляет 3–5 лет и выше. Фирма Fiskas Power Systems предложила технологию АВМ (Advanced Battery Management), которая позволяет осуществлять контроль за аккумуляторной батареей, и подзаряжать ее только в случае необходимости. Применение АВМ позволяет увеличить срок службы батарей более чем на 50%.

2.3 Электромагнитные помехи

Электромагнитными помехами являются шумы низких и высоких частот, токов, наводки от напряжения питания. Их источниками являются низкочастотные электромагнитные колебания от работы больших моторов, дрелей и т.д. (Electromagnetic Interference) или высокочастотные электромагнитные колебания от близлежащих передатчиков, и т.д. (Radio frequency interference). Кстати говоря, Ваш персональный компьютер — тоже источник электромагнитных помех.

Приводят к повреждению или потере данных; обнаруживаются осциллографом. Появляются самостоятельно или вкупе с перенапряжением, бросками напряжения и т.д.

Минимизация риска:

- сетевой кабель не прокладывается возле люминесцентных ламп;

- обеспечение надежного заземления аппаратуры;
- не допущение установки вычислительной техники вблизи мощных радиостанций или других источников EMI/RFI.
- постоянный контроль над низким уровнем излучения своей вычислительной техники;
- установка, если возможно, сеток, защитных экранов и т.п.;
- силовые и сетевые кабели не должны проходить рядом;
- на открытых пространствах для передачи информации необходимо использовать оптоволоконный кабель или радиоканал типа точка-точка. Остальное ведет себя непредсказуемо, а главное не обеспечивает гальваническую развязку.

Кроме того, к электромагнитным помехам относится Crosstalk — интерференция между кабелями, проложенными рядом.

Минимизация:

- физическая защита кабеля;
- использование витой пары;
- использование оптоволоконного кабеля.

2.4 Электростатические разряды (Electro Static Discharge)

Один из самых зловещих бичей вычислительной техники — статическое электричество, повреждающее компьютер и данные. Один из моих знакомых, работающий в отделе технической автоматизации одного из Томских банков, рассказал следующую историю:

«Однажды утром в отделе раздался телефонный звонок. Сотрудница сообщила, что у нее не работает клавиатура. Клавиатуру заменили. На следующий день ситуация повторилась. Так продолжалось три или четыре дня, пока инженеры не обратили внимания на то, что у сотрудницы шерстяные детали одежды соседствовали с синтетическими. Это и являлось причиной появления статического электричества и, как следствие, выхода из строя микросхемы клавиатуры».

Следует заметить, что электростатический потенциал может достигать величины 3-х киловольт, например движущийся человек способен создать разность потенциалов 1кВ (шаговое напряжение).

Это при том, что практически любую вычислительную технику можно вывести из строя потенциалом больше 20 В.

Электростатические заряды обнаруживаются специальной аппаратурой.

Меры защиты:

- кроме выключения кнопки off, при замене плат вынимать из розетки шнур питания;
- заземление всех компонентов, в особенности сетевых плат для коаксиального кабеля;
- при переноске различных компонентов использовать фольгированные пакеты.

Кроме того, существуют так называемые «золотые правила Static Prevention». Они гласят:

- при работе с вычислительной техникой: необходимо следить за заземлением всех компонентов, себя, окружающих;
- не брать руками за токоведущие части;
- перевозить все компоненты в электростатических чехлах;
- осторожно использовать диэлектрики: пластик и т.д. — они могут являться источниками статического электричества;
- не касаться других людей во время их работы с вычислительной техникой и особенно клавиатуры их компьютера;
- не класть компоненты вычислительной техники на токоведущие части (металл);
- не вскрывать вычислительную технику в местах с влажностью более 90%.

Преимущества этой программы:

- менее интенсивно меняется «железо» (аппаратные средства);
- меньше времени простаивает техника (downtime);
- меньше труднообъяснимых проблем difficult-to-trace;
- менее интенсивны визиты служб сервиса.

Как показали исследования, результаты которых были приведены еще в феврале 1987 года журналом «Microservice Management», возврат инвестиций в электростатическую защиту 300–4000 %.

3 КОМПЬЮТЕРНЫЕ СЕТИ

3.1 Понятие компьютерной сети

Самая простая сеть (network) состоит как минимум из двух компьютеров, соединенных друг с другом кабелем, что позволяет им обмениваться данными. Все сети (независимо от сложности) основываются именно на этом простом принципе. Хотя идея соединения компьютеров с помощью кабеля не кажется нам особо выдающейся, в свое время она явилась значительным достижением в области коммуникаций.

Рождение компьютерных сетей было вызвано практической потребностью в совместном использовании данных. Персональный компьютер — прекрасный инструмент для создания документов, подготовки таблиц, графических данных и других видов информации, но сам по себе он не позволяет Вам быстро поделиться результатами своей работы с коллегами. Когда не было сетей, приходилось распечатывать каждый документ, чтобы другие пользователи могли работать с ним или, в лучшем случае — копировать информацию на дискеты. При редактировании копий документа несколькими пользователями было очень трудно со-

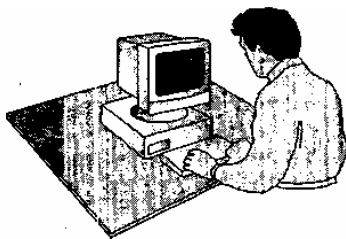


Рис. 26 — Автономная среда

брать все изменения в одном документе. Подобная схема работы называется работой в автономной среде.

Если бы пользователь, изображенный на рис. 26, подключил свой компьютер к другим, он смог бы работать с их данными и их принтерами.

Группа соединенных компьютеров и других устройств называется сетью.

А концепция соединенных и совместно использующих ресурсы компьютеров носит название сетевого взаимодействия.

Компьютеры, входящие в сеть, могут совместно использовать:

- данные;
- сообщения;
- принтеры;
- факсимильные аппараты;

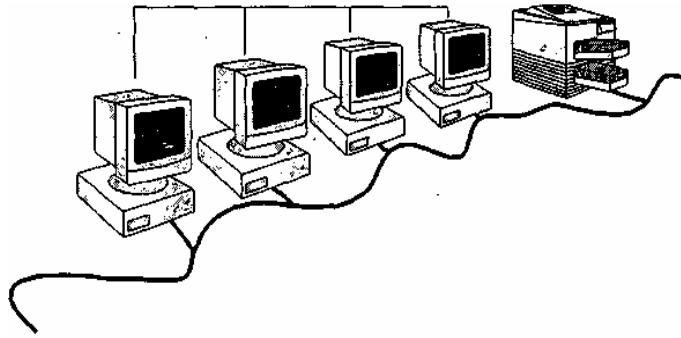


Рис. 27 — Простая сеть

- модемы;
- другие устройства.

Этот список постоянно пополняется, так как возникают новые способы совместного использования ресурсов.

3.1.1 Локальные вычислительные сети

Первоначально компьютерные сети были небольшими и объединяли до десяти компьютеров и один принтер. Технология ограничивала размеры сети, в том числе количество компьютеров в сети и ее физическую длину. Так, в начале 80-х годов самый популярный тип сетей состоял не более чем из 30 компьютеров, а длина кабеля такой сети не превышала 185 м (600 футов). Подобные сети располагались в пределах одного этажа здания или небольшой организации. Для маленьких фирм подобная конфигурация подходит и сегодня. Эти сети называются локальными вычислительными сетями [ЛВС (LAN)].

3.1.2 Расширение компьютерных сетей

Первые типы локальных сетей не могли удовлетворить потребностей крупных предприятий, офисы которых обычно расположены в различных местах. Но как только преимущества компьютерных сетей стали неоспоримы и все больше сетевых программ появлялось на рынке, перед корпорациями — для сохранения конкурентоспособности — встала задача расширения сетей. Так, на основе локальных сетей возникли более крупные системы.

Сегодня, чтобы соединить пользователей из разных городов и государств, географические рамки сетей раздвигаются и ЛВС

перерастает в глобальную вычислительную сеть [ГВС (WAN)], где количество компьютеров уже может варьироваться от десятка до нескольких тысяч.

В настоящее время большинство организаций хранит и совместно использует в сетевой среде огромные объемы жизненно важных данных. Вот почему сети сейчас так же необходимы, как еще совсем недавно были необходимы пишущие машинки и картотеки. Теперь сети позволяют целому ряду пользователей одновременно «владеть» данными и периферийными устройствами. Если нескольким пользователям надо распечатать документ, все они обращаются к сетевому принтеру.

До появления компьютерных сетей сотрудники организаций обменивались информацией, например, так:

- передавали информацию устно (устная речь);
- писали записки или письма (письменная речь);
- записывали информацию на дискету, несли дискету к другому компьютеру и копировали в него данные.

Компьютерные сети экономят бумагу и упрощают обмен информацией, предоставляя пользователям быстрый доступ к любым типам данных.

3.1.3 Приложения

Сети создают отличные условия для унификации приложений (например, текстового процессора). Это значит, что на всех компьютерах в сети выполняются приложения одного типа и одной версии. Использование унифицированного приложения упрощает обслуживание сети. Действительно, проще изучить одно приложение, чем пытаться освоить сразу четыре или пять. Удобнее также иметь дело с одной версией приложения и настраивать компьютеры одинаковым образом.

Другое достоинство сетей — наличие программ электронной почты и планирования рабочего дня. Благодаря им менеджеры могут успешно взаимодействовать с многочисленным штатом своих сотрудников или партнеров по бизнесу, а планирование и управление деятельностью всей компании осуществляется гораздо быстрее и эффективнее.

Все сети имеют некоторые общие компоненты, функции и характеристики. В их числе:

- серверы (servers) — компьютеры, чьи ресурсы доступны сетевым пользователям;
- клиенты (clients) — компьютеры, осуществляющие доступ к сетевым ресурсам, предоставляемым серверами;
- среда передачи (media) — способ соединения компьютеров;
- совместно используемые данные — файлы, предоставляемые серверами по сети;
- совместно используемые периферийные устройства, например принтеры, библиотеки CD-ROM и т.д., — ресурсы, предоставляемые серверами;
- ресурсы — файлы, периферийные устройства и другие элементы, используемые в сети.

Несмотря на отмеченное сходство, сети разделяются на два типа:

- одноранговые (peer-to-peer);
- на основе сервера (server based).

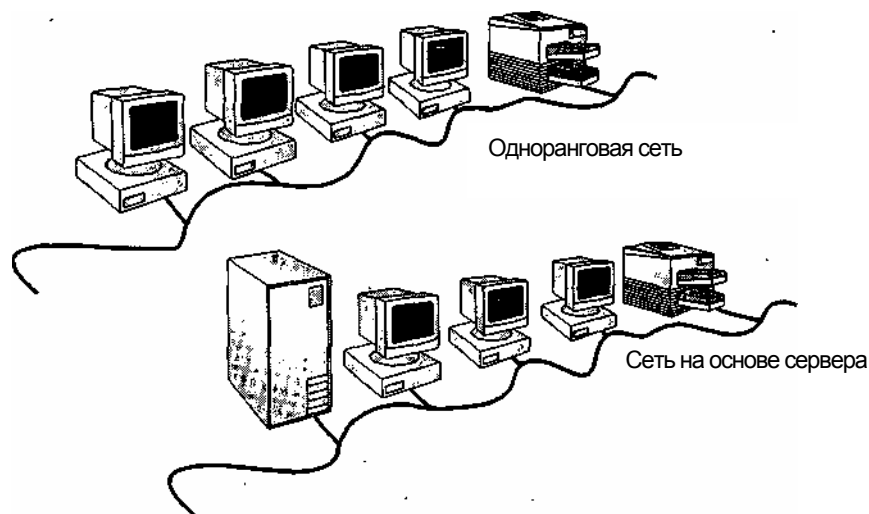


Рис. 28 — Типы сетей

Различия между одноранговыми сетями и сетями на основе сервера принципиальны, поскольку определяют разные возможности этих сетей. Выбор типа сети зависит от многих факторов:

- размера предприятия;
- необходимой степени безопасности;
- вида бизнеса;

- доступности административной поддержки;
- объема сетевого трафика;
- потребностей сетевых пользователей;
- уровня финансирования.

3.1.4 Одноранговые сети

В одноранговой сети все компьютеры равноправны: нет иерархии среди компьютеров и нет выделенного (dedicated) сервера. Обычно каждый компьютер функционирует и как клиент, и как сервер — иначе говоря, нет отдельного компьютера, ответственного за всю сеть. Пользователи сами решают, какие данные на своем компьютере сделать доступными по сети.

3.1.5 Сети на основе сервера

Если к одноранговой сети, где компьютеры выступают в роли и клиентов, и серверов, подключить более 10 пользователей, она может не справиться с объемом поставленных перед ней задач. Поэтому большинство сетей имеет другую конфигурацию — они работают на основе выделенного сервера. Выделенным называется такой сервер, который функционирует только как сервер и не используется в качестве клиента или рабочей станции. Он оптимизирован для быстрой обработки запросов от сетевых клиентов и для повышения защищенности файлов и каталогов.

При увеличении размеров сети и объема сетевого трафика необходимо увеличивать количество серверов. Распределение задач среди нескольких серверов гарантирует, что каждая задача будет выполняться наиболее эффективно.

3.1.6 Специализированные серверы

Круг задач, которые должны выполнять серверы, многообразен и сложен. Для того, чтобы серверы отвечали современным требованиям пользователей, в больших сетях их делают специализированными (specialized). Например, в сети могут работать различные типы серверов:

- Серверы файлов и печати.

Серверы файлов и печати управляют доступом пользователей соответственно к файлам и принтерам. Так, чтобы работать с текстовым процессором, прежде всего Вы должны запустить его на своем компьютере. Документ текстового процессора, хранящийся на сервере файлов, загружается в память вашего компьютера, и теперь Вы можете работать с этим документом на своем компьютере. Другими словами, сервер файлов предназначен для хранения данных.

- Серверы приложений.

На серверах приложений выполняются прикладные части клиент-серверных приложений, а также находятся данные, доступные клиентам. Например, чтобы ускорить поиск данных, серверы хранят большие объемы информации в структурированном виде. Эти серверы отличаются от серверов файлов и печати. В последних файл или данные целиком копируются на запрашивающий компьютер. А в сервере приложений на клиентский компьютер пересылаются только результаты запроса.

Приложение-клиент на удаленном компьютере получает доступ к данным, сохраняемым на сервере приложений. Однако вместо всей базы данных на Ваш компьютер с сервера загружаются только результаты запроса. Например, вместо полного перечня сотрудников Вы можете получить список только тех, кто родился в ноябре.

- Почтовые серверы.

Почтовые серверы управляют передачей сообщений электронной почты между пользователями сети.

- Серверы факсов.

Серверы факсов управляют потоком входящих и исходящих факсимильных сообщений через один или несколько факс-модемов.

- Коммуникационные серверы.

Коммуникационные серверы (серверы связи) управляют проходящим через модем и телефонную линию потоком данных и почтовых сообщений между своей сетью и другими сетями, мейнфреймами или удаленными пользователями.

- Серверы служб каталога.

Каталог (directory) содержит данные о структуре сети, позволяя пользователям находить, сохранять и защищать информа-

цию в сети. Windows NT Server объединяет компьютеры в логические группы — домены (domains), система защиты которых обеспечивает различным пользователям неодинаковые права доступа к сетевым ресурсам.

В большой сети использование серверов разных типов требует особого внимания. Необходимо учитывать все возможные нюансы, которые могут проявиться при разрастании сети, с тем, чтобы изменение специализации какого-либо сервера в дальнейшем не отразилось на работе всей сети.

3.1.7 Комбинированные сети

Существуют и комбинированные типы сетей, сочетающие лучшие качества одноранговых сетей и сетей на основе сервера. Многие администраторы считают, что комбинированные сети наиболее полно соответствуют запросам современных пользователей. Комбинированные сети — наиболее распространенный тип сетей, но для их правильной реализации и надежной защиты необходимы определенные знания и навыки планирования.

3.2 Топология сети

Термин «топология» (topology), или «топология сети», обозначает физическое расположение компьютеров, кабелей и других сетевых компонентов. Топология — это стандартный термин, который используется профессионалами при описании базовой схемы сети. Кроме термина «топология», для описания физической компоновки сети употребляют также следующие:

- физическое расположение;
- компоновка;
- карта;
- схема.

Характеристики сети зависят от типа устанавливаемой топологии. В частности, выбор той или иной топологии влияет на:

- состав необходимого сетевого оборудования;
- возможности сетевого оборудования;
- возможности расширения сети;

- способ управления сетью.

Если Вы поймете, как используются различные топологии, то будете знать, какими возможностями обладают различные типы сетей.

Чтобы совместно использовать ресурсы или выполнять другие сетевые задачи, компьютеры должны быть подключены друг к другу. Для этой цели в большинстве сетей применяется кабель.

Однако просто подключить компьютер к кабелю, соединяющему другие компьютеры, недостаточно. Различные типы кабелей в сочетании с различными сетевыми платами, сетевыми операционными системами и другими компонентами требуют и различных методов реализации. Кроме того, каждая топология сети при установке выдвигает ряд условий. Например, применения не только конкретного типа кабеля, но и способа его прокладки. Топология может также определять способ взаимодействия компьютеров в сети. Различным видам топологий соответствуют различные методы взаимодействия, и эти методы оказывают большое влияние на работу сети.

3.2.1 Базовые топологии

Все сети строятся на основе трех базовых топологий, известных как:

- шина (bus);
- звезда (star);
- кольцо (ring).

Если компьютеры подключены вдоль одного кабеля [сегмента (segment)], топология называется «шиной». В том случае, когда компьютеры подключены к сегментам кабеля, исходящим из одной точки [концентратора (hub)], топология называется «звездой». Если кабель, к которому подключены компьютеры, замкнут в кольцо, такая топология носит название «кольца».

Сами по себе базовые топологии несложны, однако на практике часто встречаются довольно сложные комбинации, сочетающие свойства и характеристики нескольких топологий.

3.2.2 Шина

Топологию «шина» часто называют «линейной шиной» (linear bus). В ней используется один кабель, именуемый магистралью или сегментом, к которому подключены все компьютеры сети. Данная топология является наиболее простой и распространенной реализацией сети.

Взаимодействие компьютеров

В сети с топологией «шина» компьютеры адресуют данные конкретному компьютеру, передавая их по кабелю в виде электрических сигналов. Чтобы понять процесс взаимодействия компьютеров по шине, Вы должны уяснить следующие понятия:

- передача сигнала;
- отражение сигнала;
- терминатор.

Передача сигнала

Данные в виде электрических сигналов передаются всем компьютерам сети; однако информацию принимает только тот компьютер, чей адрес соответствует адресу получателя, зашифрованному в этих сигналах. Причем в каждый момент времени вести передачу может только один компьютер.

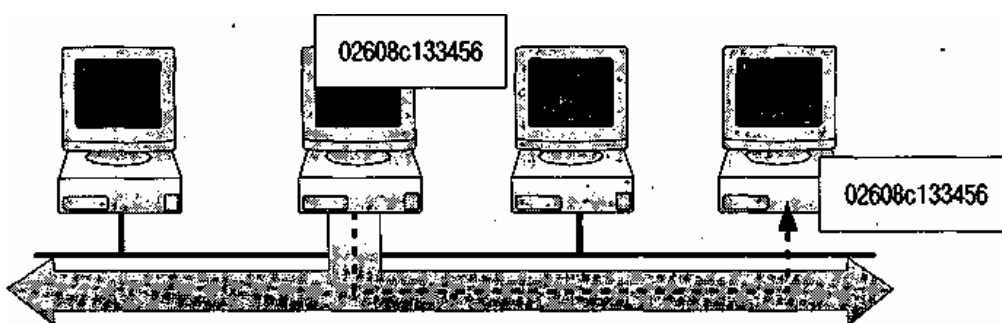


Рис. 29 — Данные посылаются всем компьютерам, но принимает их только адресат

Так как данные в сеть передаются лишь одним компьютером одновременно, ее производительность зависит от количества компьютеров, подключенных к шине. Чем больше компьютеров, тем большее их число ожидает передачи и тем медленнее сеть.

Однако вывести прямую зависимость между пропускной способностью сети и количеством компьютеров в ней нельзя, поскольку, кроме числа компьютеров, на быстродействие сети влияет множество других факторов, например:

- тип аппаратного обеспечения сетевых компьютеров;
- частота, с которой компьютеры передают данные;
- тип работающих сетевых приложений;
- тип сетевого кабеля;
- расстояние между компьютерами в сети.

Шина — пассивная топология. Это значит, что компьютеры только «слушают» передаваемые по сети данные, но не перемещают их от отправителя к получателю. Поэтому если какой-либо компьютер выйдет из строя, то это не скажется на работе сети. В активных топологиях компьютеры регенерируют сигналы и передают их дальше по сети.

Отражение сигнала

Электрические сигналы распространяются от одного конца кабеля к другому. Если не предпринимать никаких специальных мер, сигнал, достигая конца кабеля, будет отражаться и создавать помехи, не позволяя компьютерам осуществлять передачу. Поэтому на концах кабеля электрические сигналы необходимо гасить.

Терминатор

Для того чтобы предотвратить отражение электрических сигналов, на каждом конце кабеля устанавливают терминаторы (terminators), поглощающие эти сигналы.

К любому свободному, т.е. ни к чему не подключенному, концу кабеля нужно подсоединять терминатор.

Нарушение целостности кабеля

Нарушение целостности сетевого кабеля происходит при его разрыве или отсоединении одного из его концов. Возможна также ситуация, когда на одном или нескольких концах кабеля отсутствуют терминаторы, что приводит к отражению электрических сигналов и, как следствие, к выходу из строя сети. Компьютеры остаются полностью работоспособными, но до тех пор, пока сегмент разорван, они не могут взаимодействовать друг с другом.

3.2.3 Звезда

При топологии «звезда» (рис. 30) все компьютеры с помощью сегментов кабеля подключаются к центральному компоненту — концентратору (hub). Сигналы от передающего компьютера поступают через концентратор ко всем остальным. Эта топология возникла на заре вычислительной техники, когда компьютеры подключались к центральному, главному, компьютеру.

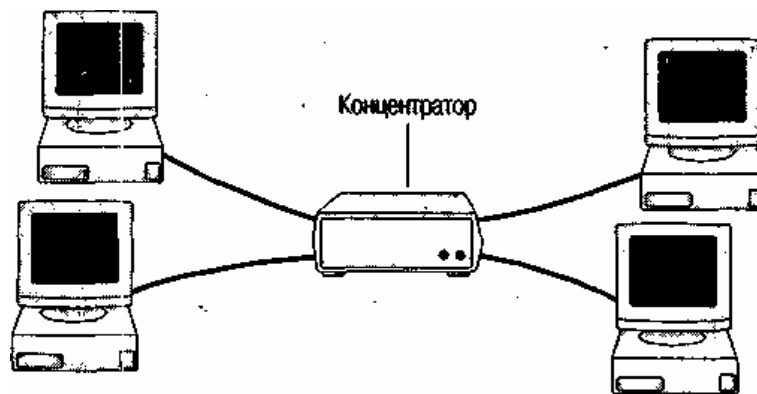


Рис. 30 — Сеть с топологией «звезда»

В сетях с топологией «звезда» подключение компьютеров к сети выполняется централизованно. Но есть и недостатки: так как все компьютеры подключены к центральной точке, для больших сетей значительно увеличивается расход кабеля. К тому же если центральный компонент выйдет из строя, то остановится вся сеть.

А если выйдет из строя только один компьютер (или кабель, соединяющий его с концентратором), то лишь этот компьютер не сможет передавать или принимать данные по сети. На остальные компьютеры в сети этот сбой не повлияет.

3.2.4 Кольцо

При топологии «кольцо» (рис. 31) компьютеры подключаются к кабелю, замкнутому в кольцо. Поэтому у кабеля просто не может быть свободного конца, на который надо поставить терминатор. Сигналы передаются по кольцу в одном направлении и проходят через каждый компьютер. В отличие от пассивной

топологии «шина», здесь каждый компьютер выступает в роли повторителя, усиливая сигналы и передавая их следующему компьютеру. Поэтому, если выйдет из строя один компьютер, прекращает функционировать вся сеть.

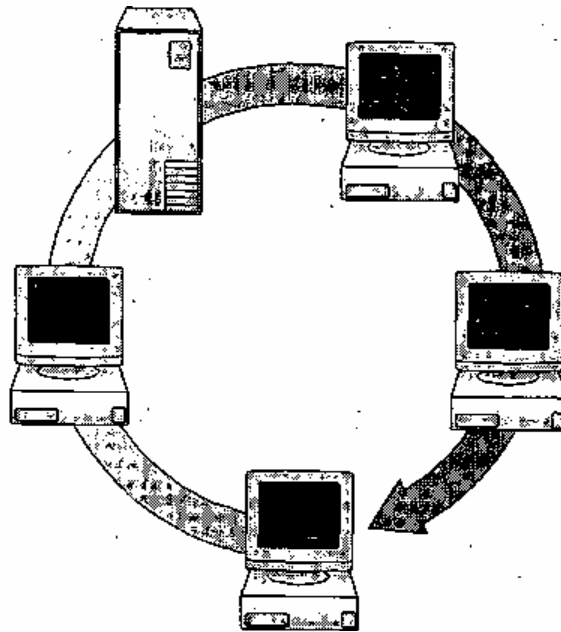


Рис. 31 — Топология сети типа «кольцо»

3.2.5 Комбинированные топологии

Сегодня при компоновке сети все чаще используются комбинированные топологии, которые сочетают отдельные свойства шины, звезды и кольца.

Звезда-шина

«Звезда-шина» (star-bus) — это комбинация топологий «шина» и «звезда» (рис. 32). Обычно схема выглядит так: несколько сетей с топологией «звезда» объединяются при помощи магистральной линейной шины.

В этом случае выход из строя одного компьютера не скажется на работе всей сети — остальные компьютеры по-прежнему взаимодействуют друг с другом. А выход из строя концентратора повлечет за собой отсоединение от сети только подключенных к нему компьютеров и концентраторов.

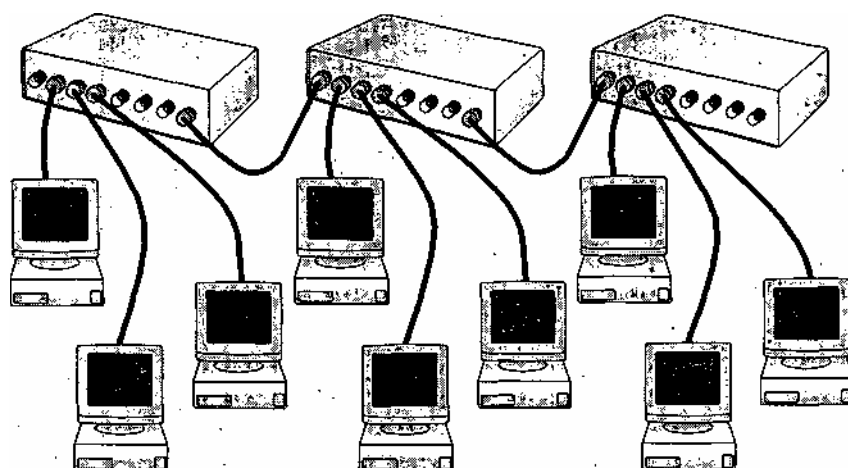


Рис. 32 — Сеть с топологией «звезда-шина»

Звезда-кольцо

«Звезда-кольцо» (star-ring) несколько похожа на «звезду-шину» (рис. 33). И в той и в другой топологиях компьютеры подключаются к концентраторам. Отличие состоит в том, что концентраторы в «звезде-шине» соединены магистральной линейной шиной, а в «звезде-кольце» все концентраторы подключены к главному концентратору, образуя звезду. Кольцо же реализуется внутри главного концентратора.

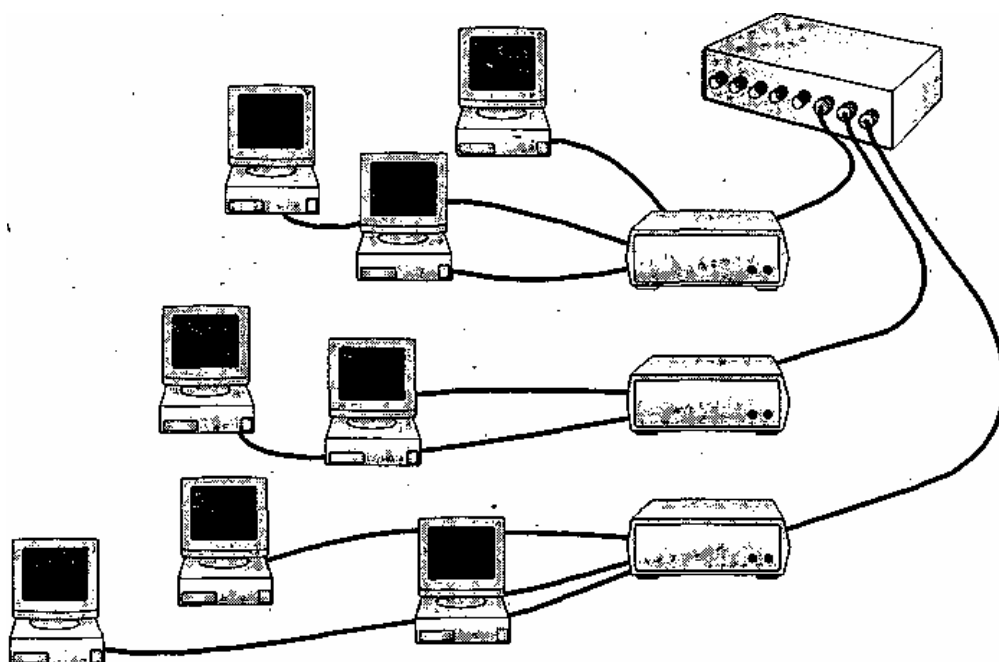


Рис. 33 — Сеть с топологией «звезда-кольцо»

3.2.6 Выбор топологии

Существует множество факторов, которые необходимо учитывать при выборе топологии для каждой конкретной сети. Эта таблица поможет Вам осуществить правильный выбор.

Таблица 7

Различия топологий построения сети

Топология	Преимущества	Недостатки
Шина	Экономный расход кабеля. Сравнительно недорогая и несложная в использовании среда передачи. Простота, надежность. Легко расширяется	При значительных объемах трафика уменьшается пропускная способность сети. Трудно локализовать проблемы. Выход из строя кабеля останавливает работу многих пользователей
Кольцо	Все компьютеры имеют равный доступ. Количество пользователей не оказывает сколько-нибудь значительного влияния на производительность	Выход из строя одного компьютера может вывести из строя всю сеть. Трудно локализовать проблемы. Изменение конфигурации сети требует остановки всей сети
Звезда	Легко модифицировать сеть, добавляя новые компьютеры. Централизованный контроль и управление. Выход из строя одного компьютера не влияет на работоспособность сети.	Выход из строя центрального узла парализует всю сеть

3.3 Кабельные каналы связи

3.3.1 Основные группы кабелей

На сегодняшний день подавляющая часть компьютерных сетей использует для соединения провода или кабели. Они выступают в качестве среды передачи сигналов между компьютерами. Существуют различные типы кабелей, которые обеспечи-

вают нормальную работу всевозможных сетей, от малых до больших.

Однако на практике в большинстве сетей применяются только три основные группы кабелей:

- коаксиальный кабель (coaxial cable);
- витая пара (twisted pair);
- неэкранированная (unshielded);
- экранированная (shielded);
- оптоволоконный кабель (fiber optic).

3.3.2 Коаксиальный кабель

Не так давно наиболее распространенным типом считался коаксиальный кабель. Это объяснялось двумя причинами. Во-первых, он был относительно недорогим, легким, гибким и удобным в применении, а во-вторых, надежным и простым в установке. Самый простой коаксиальный кабель состоит из медной жилы (core), окружающей ее изоляции, экрана в виде металлической оплетки и внешней оболочки. Если кабель, кроме металлической оплетки, имеет и слой фольги, он называется кабелем с двойной экранизацией. При наличии сильных помех можно воспользоваться кабелем с учетверенной экранизацией. Он состоит из двойного слоя фольги и двойного слоя металлической оплетки.

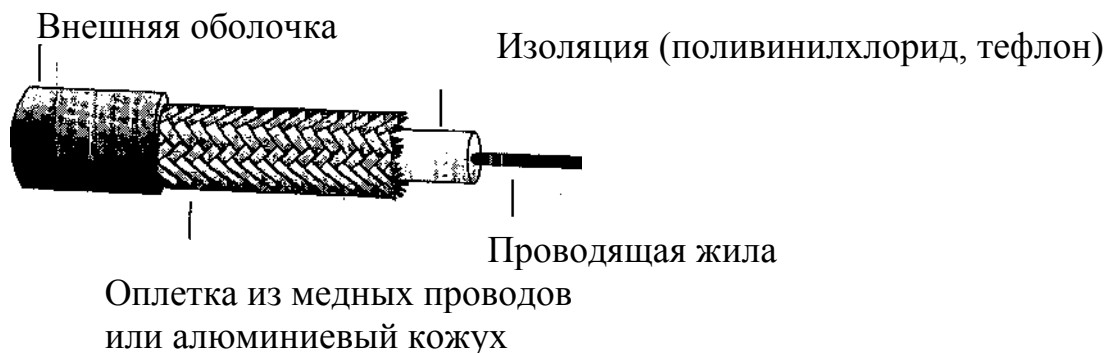


Рис. 34 — Строение коаксиального кабеля

Некоторые типы кабелей покрывает металлическая сетка — экран (shield). Он защищает передаваемые по кабелю данные, поглощая внешние электромагнитные сигналы, которые называются

помехами или шумом. Таким образом, экран не позволяет помехам исказить данные.

Электрические сигналы, кодирующие данные, передаются по жиле. Жила — это один провод (сплошная жила) или пучок проводов. Сплошная жила изготавливается, как правило, из меди.

Жила окружена *диэлектрическим* (dielectric) изоляционным слоем, который отделяет ее от металлической оплетки. Оплетка играет роль «земли» и защищает жилу от электрических шумов (noise) и перекрестных помех (crosstalk). Перекрестные помехи — это электрические наводки, вызванные сигналами в соседних проводах.

Проводящая жила и металлическая оплетка не должны соприкасаться, иначе произойдет короткое замыкание и данные разрушатся. Снаружи кабель покрыт непроводящим слоем — из резины, тефлона или пластика.

Коаксиальный кабель более помехоустойчив, затухание сигнала в нем меньше, чем в витой паре. Затухание (attenuation) — это ослабление сигнала при его прохождении по кабелю.

Как уже говорилось, плетеная защитная оболочка поглощает внешние электромагнитные сигналы, не позволяя им влиять на передаваемые по жиле данные, поэтому коаксиальный кабель можно использовать при передаче на большие расстояния и в тех случаях, когда высокоскоростная передача данных осуществляется на несложном оборудовании.

Типы коаксиальных кабелей

Существует два типа коаксиальных кабелей для построения компьютерных сетей:

- тонкий (thinnet) коаксиальный кабель;
- толстый (thicknet) коаксиальный кабель.

Выбор того или иного типа кабеля зависит от потребностей конкретной сети.

Тонкий коаксиальный кабель

Тонкий коаксиальный кабель — гибкий кабель диаметром около 0,5 см (0,25 дюйма). Он прост в применении и подходит практически для любого типа сети. Подключается непосредственно к плате сетевого адаптера компьютера.

Тонкий коаксиальный кабель способен передавать сигнал на расстояние до 185 м (около 600 футов) без его заметного искажения, вызванного затуханием.

Производители кабелей выработали специальную маркировку для различных типов кабелей. Тонкий коаксиальный кабель относится к группе, которая называется семейством RG-58; его волновое сопротивление равно 50 Ом. Волновое сопротивление (impedance) — это сопротивление переменному току, выраженное в омах. Основная особенность семейства RG-58 — медная жила. Она может быть сплошной или состоять из нескольких переплетенных проводов.

Толстый коаксиальный кабель

Толстый коаксиальный кабель — относительно жесткий кабель с диаметром около 1 см (0,5 дюйма). Иногда его называют «стандартный Ethernet», поскольку он был первым типом кабеля, применяемым в Ethernet — популярной сетевой архитектуре. Медная жила этого кабеля толще, чем у тонкого коаксиального кабеля. Чем толще жила у кабеля, тем большее расстояние способен преодолеть сигнал. Следовательно, толстый коаксиальный кабель передает сигналы дальше, чем тонкий, — до 500 м (около 1640 футов). Поэтому толстый коаксиальный кабель иногда используют в качестве опорного кабеля, магистрали (backbone), который соединяет несколько небольших сетей, построенных на тонком коаксиальном кабеле.

Для подключения к толстому коаксиальному кабелю применяют специальное устройство — трансивер (transceiver).

Трансивер снабжен специальным соединителем, который назван довольно оригинально — «вампир» (vampire tap) или «пронзающий ответвитель» (piercing tap). «Вампир» проникает через изоляционный слой, протыкая его, и вступает в непосредственный физический контакт с проводящей жилой. Чтобы подключить трансивер к сетевому адаптеру, надо кабель трансивера подключить к коннектору AUI-порта сетевой платы. Этот коннектор известен также как DIX-коннектор (Digital Intel Xerox®), в соответствии с названиями фирм-разработчиков, или коннектор DB-15.

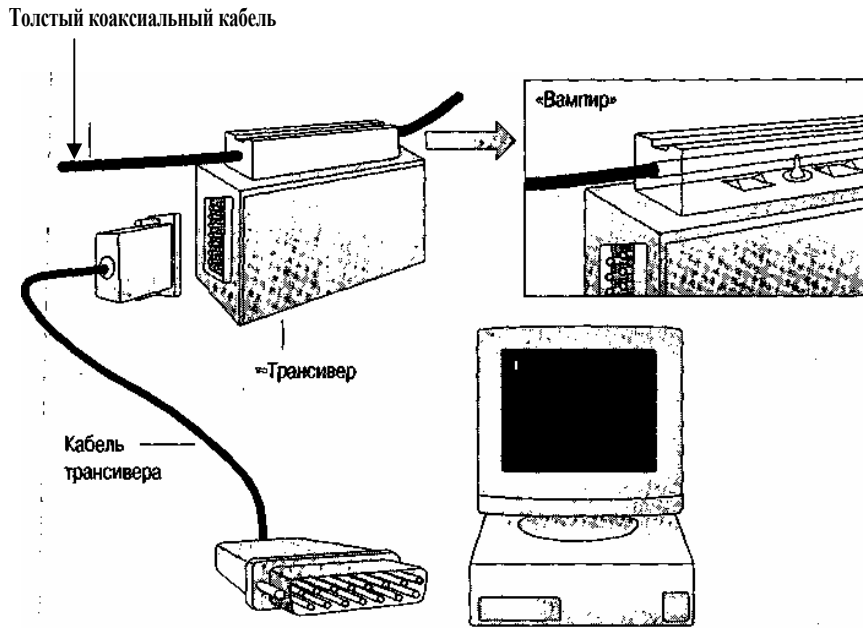


Рис. 35 — Подключение трансивера к толстому коаксиальному кабелю

Сравнение двух типов коаксиальных кабелей

Как правило, чем толще кабель, тем сложнее его прокладывать. Тонкий коаксиальный кабель гибок, прост в установке и относительно недорог. Толстый кабель трудно гнуть, следовательно, его сложнее монтировать. Это очень существенный недостаток, особенно в тех случаях, когда необходимо проложить кабель по трубам или желобам. Толстый коаксиальный кабель дороже тонкого, но при этом он передает сигналы на большие расстояния.

Оборудование для подключения коаксиального кабеля

Для подключения тонкого коаксиального кабеля к компьютерам используются так называемые BNC-коннекторы (British Naval Connector, BNC) (рис. 36). В семействе BNC выделяют несколько основных компонентов:

- BNC-коннектор.

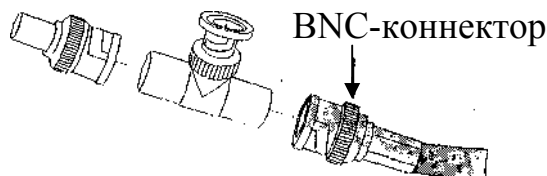


Рис. 36 — BNC-коннектор

BNC-коннектор либо припаивается, либо обжимается на конце кабеля.

- BNC T-коннектор (рис. 37).

T-коннектор соединяет сетевой кабель с сетевой платой компьютера.

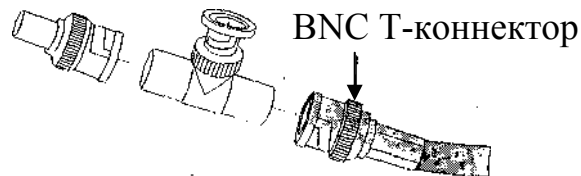


Рис. 37 — BNC T-коннектор

- BNC-терминатор (рис. 38).

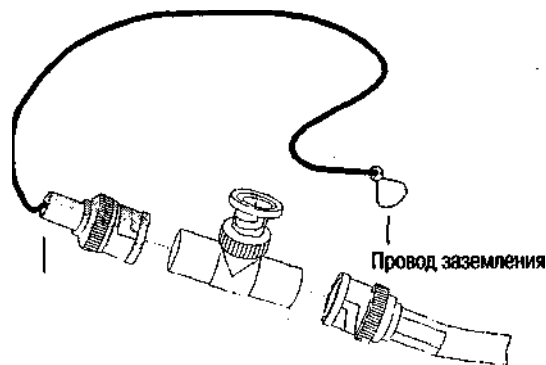


Рис. 38 — BNC-терминатор

В сети с топологией «шина» для поглощения блуждающих сигналов на каждом конце кабеля устанавливаются терминаторы, иначе сеть не будет работать.

3.3.3 Витая пара

Самая простая витая пара (twisted pair) — это два перевитых вокруг друг друга изолированных медных провода. Существует два типа витой пары:

- неэкранированная (unshielded) витая пара (UTP);
- экранированная (shielded) витая пара (STP).



Рис. 39 — Экранированная витая пара

Несколько витых пар проводов часто помещают в одну защитную оболочку. Их количество в таком кабеле может быть разным. Завивка проводов позволяет избавиться от электрических помех, наводимых соседними парами и другими внешними источниками, например двигателями, реле и трансформаторами.

Неэкранированная витая пара

Неэкранированная витая пара (спецификация 10 BaseT) широко используется в ЛВС; максимальная длина сегмента составляет 100 м (328 футов).

Неэкранированная витая пара состоит из двух изолированных медных проводов. Существует несколько спецификаций, которые регулируют количество витков на единицу длины — в зависимости от назначения кабеля. В Северной Америке UTP повсеместно используется в телефонных сетях.

Неэкранированная витая пара определена особым стандартом — Electronic Industries Association and the Telecommunications Industries Association (EIA/TIA) 568 Commercial Building Wiring Standard. EIA/TIA 568, предлагая нормативные характеристики кабелей для различных случаев, гарантирует единообразие продукции. Эти стандарты включают пять категорий UTP.

- Категория 1.

Традиционный телефонный кабель, по которому можно передавать только речь, но не данные. Большинство телефонных кабелей, произведенных до 1983 года, относится к категории 1.

- Категория 2.

Кабель, способный передавать данные со скоростью до 4 Мбит/с. Состоит из четырех витых пар.

- Категория 3.

Кабель, способный передавать данные со скоростью до 10 Мбит/с. Состоит из четырех витых пар с девятью витками на метр.

- Категория 4.

Кабель, способный передавать данные со скоростью до 16 Мбит/с. Состоит из четырех витых пар.

- Категория 5.

Кабель, способный передавать данные со скоростью до 100 Мбит/с. Состоит из четырех витых пар медного провода.

Большинство телефонных систем использует неэкранированную витую пару. Обычно при строительстве новых зданий UTP прокладывают в расчете на будущие потребности. Если установленные во время строительства провода рассчитаны на передачу данных, то их можно использовать и в компьютерной сети. Однако надо быть осторожным, так как обычный телефонный провод не имеет витков и его электрические характеристики могут не соответствовать тем, которые требуются для надежной и защищенной передачи данных между компьютерами.

Одной из потенциальных проблем для любых типов электрических кабелей являются перекрестные помехи. Перекрестные помехи — это электрические наводки, вызванные сигналами в смежных проводах. Неэкранированная витая пара особенно страдает от перекрестных помех. Для уменьшения их влияния используют экран.

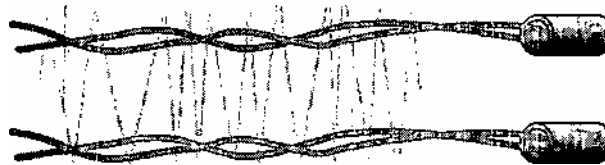


Рис. 40 — Перекрестные помехи — это электрические наводки со стороны соседних линий

Экранированная витая пара

Кабель экранированной витой пары (STP) имеет медную оплетку, которая обеспечивает более надежную защиту от помех. Кроме того, пары проводов STP обмотаны фольгой (рис. 41). В результате экранированная витая пара прекрасно защищает передаваемые данные от внешних помех.

Все это означает, что STP, по сравнению с UTP, меньше подвержена воздействию электрических помех и может передавать данные с более высокой скоростью и на большие расстояния.



Рис. 41 — Экранированная витая пара

Компоненты кабельной системы

Оборудование для подключения.

Для подключения витой пары к компьютеру используются телефонные коннекторы RJ-45. На первый взгляд, они похожи на RJ-11, но в действительности между ними есть существенные отличия.

Во-первых, вилка RJ-45 чуть больше по размерам и не подходит для гнезда RJ-11. Во-вторых, коннектор RJ-45 имеет 8 контактов, а RJ-11 — только 4.

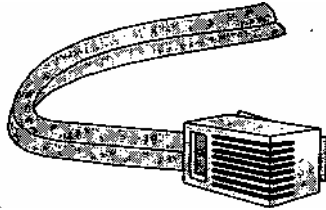


Рис. 42 — Вилка RJ-45

Построить сложную кабельную систему и в то же время упростить работу с ней Вам поможет ряд очень полезных компонентов.

- Распределительные стойки и полки (distribution racks, shelves). Распределительные стойки и полки предназначены для монтажа кабеля. Они позволяют централизованно организовать множество соединений и при этом занимают сравнительно мало места.

- Коммутационные панели (patch panels).

Существуют разные типы коммутационных панелей. Они поддерживают до 96 портов и скорость передачи до 100 Мбит/с.

- Соединители.

Соединители — это одинарные или двойные гнезда RJ-45 на панелях расширения или настенных розетках. Они обеспечивают скорость передачи до 100 Мбит/с.

- Настенные розетки.

Настенные розетки имеют одно или несколько гнезд RJ-45.

3.3.4 Оптоволоконный кабель

В оптоволоконном кабеле цифровые данные распространяются по оптическим волокнам в виде модулированных световых импульсов. Это относительно защищенный способ передачи, поскольку при нем не используются электрические сигналы. Следовательно, к оптоволоконному кабелю невозможно подключиться, не разрушая его, и перехватывать данные, от чего не застрахован любой кабель, проводящий электрические сигналы.

Оптоволоконные линии предназначены для передачи больших объемов данных на очень высоких скоростях, поскольку сигнал в них практически не затухает и не искажается.

Строение

Оптическое волокно — чрезвычайно тонкий стеклянный цилиндр, называемый жилой (core). Он покрыт слоем стекла (оболочкой) с иным, чем у жилы, коэффициентом преломления. Иногда оптоволоконно производят из пластика. Пластик проще в монтаже, но он передает световые импульсы на меньшие расстояния по сравнению со стеклянным оптоволоконном.

Каждое оптоволоконно передает сигналы только в одном направлении, поэтому кабель состоит из двух волокон с самостоятельными коннекторами. Одно из них служит для передачи, а другое — для приема. Жесткость кабеля увеличена покрытием из пластика, а прочность — волокнами из кевлара. Волокна кевлара располагаются между двумя кабелями, заключенными в пластик.

Передача по оптоволоконному кабелю не подвержена электрическим помехам и ведется на чрезвычайно высокой скорости (в настоящее время широко используется скорость в 2000 Мбит/с, получает все большее распространение скорость в 10 Гбит/с и выше). По нему можно передавать световой импульс на многие километры.

Некоторые соображения

Используйте оптоволоконный кабель, если требуется:

- передавать данные с очень высокой скоростью на большие расстояния по защищенной среде.

Не используйте оптоволоконный кабель, если требуется:

- построить сеть при ограниченных денежных средствах (примечание: в настоящее время стоимость оптоволоконного кабеля уже сравнима со стоимостью высококачественного медного кабеля);
- дополнительная подготовка для правильной установки и корректного подключения оптоволоконных сетевых устройств (примечание: хотя на самом деле научиться работать с оптоволоконном не так уж сложно).

3.4 Беспроводная среда передачи данных

Словосочетание «беспроводная среда» может ввести в заблуждение, поскольку означает полное отсутствие проводов в сети. В большинстве случаев это не совсем так. Обычно беспроводные компоненты взаимодействуют с сетью, в которой в качестве среды передачи используется кабель. Такая сеть со смешанными компонентами называется гибридной.

Возможности

Идея беспроводной среды весьма привлекательна, так как ее компоненты:

- обеспечивают временное подключение к кабельной сети;
- помогают организовать резервное копирование в кабельную сеть;
- гарантируют определенный уровень мобильности;
- позволяют снять ограничения на максимальную протяженность сети, накладываемые медными или даже оптоволоконными кабелями.

Применение

Трудность монтажа кабеля — фактор, который дает беспроводной среде неоспоримое преимущество. Беспроводная среда может оказаться особенно полезной в следующих ситуациях:

- в помещениях с большим скоплением народа (например, в приемной);
- для людей, у которых нет постоянного рабочего места (например, для врачей или медсестер);
- в изолированных помещениях и зданиях;

- в помещениях, где планировка часто меняется;
- в строениях (например, памятниках истории или архитектуры), где прокладывать кабель запрещено.

3.4.1 Типы беспроводных сетей

В зависимости от используемой технологии беспроводные сети можно разделить на три типа:

- локальные вычислительные сети;
- расширенные локальные вычислительные сети;
- мобильные сети (переносные компьютеры).

Основные различия между этими типами сетей — параметры передачи. Локальные и расширенные локальные вычислительные сети используют передатчики и приемники, принадлежащие той организации, в которой функционирует сеть. Для переносных компьютеров средой передачи служат общедоступные сети, например телефонная сеть или Интернет.

Локальные вычислительные беспроводные сети

Типичная беспроводная сеть выглядит и функционирует практически так же, как кабельная, за исключением среды передачи. Беспроводной сетевой адаптер с трансивером установлен в каждом компьютере, и пользователи работают так, будто их компьютеры соединены кабелем.

Точки доступа

Трансивер, называемый иногда точкой доступа (access point), обеспечивает обмен сигналами между компьютерами с беспроводным подключением и кабельной сетью. В беспроводных ЛВС используются небольшие настенные трансиверы. Они устанавливаются радиоконтакт с переносными устройствами. Наличие этих трансиверов не позволяет назвать такую сеть строго беспроводной.

Беспроводные локальные сети используют четыре способа передачи данных:

- инфракрасное излучение;
- лазер;
- радиопередачу в узком диапазоне (одночастотная передача);
- радиопередачу в рассеянном спектре.

3.4.2 Инфракрасные беспроводные сети

Все инфракрасные беспроводные сети используют для передачи данных инфракрасные лучи. В подобных системах необходимо генерировать очень сильный сигнал, так как в противном случае значительное влияние будут оказывать другие источники, например свет из окна.

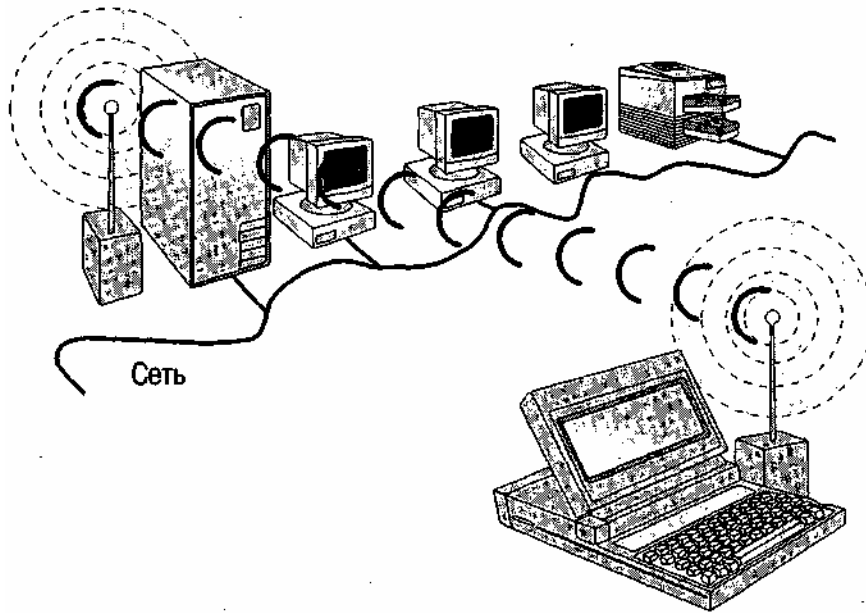


Рис. 43 — Переносной компьютер, подключенный к точке доступа

Этот способ позволяет передавать сигналы с большой скоростью, поскольку инфракрасный свет имеет широкий диапазон частот. Инфракрасные сети способны нормально функционировать на скорости 10 Мбит/с.

Существует четыре типа инфракрасных сетей:

- Сети прямой видимости.

В таких сетях «передача возможна лишь в случае прямой видимости между передатчиком и приемником.

- Сети на рассеянном инфракрасном излучении.

При этой технологии сигналы, отражаясь от стен и потолка, в конце концов достигают приемника. Эффективная область действия ограничена примерно 30 м (100 футами), и скорость передачи невелика (из-за неравномерности сигнала).

- Сети на отраженном инфракрасном излучении.

В этих сетях оптические трансиверы, расположенные рядом с компьютером, передают сигналы в определенное место, откуда они пересылаются соответствующему компьютеру.

- Модулированные оптические сети.

Эти инфракрасные беспроводные сети соответствуют жестким требованиям мультимедийной среды и практически не уступают в скорости кабельным сетям. Хотя скорость инфракрасных сетей и удобство их использования очень привлекательны, возникают трудности при передаче сигналов на расстояние более 30 м (100 футов). К тому же такие сети подвержены помехам со стороны сильных источников света, которые есть в большинстве организаций.

3.4.3 Лазерные технологии организации сетей

Лазерная технология похожа на инфракрасную тем, что требует прямой видимости между передатчиком и приемником. Если по каким-либо причинам луч будет прерван, то это прервет и передачу.

3.4.4 Радиопередача в узком диапазоне (одночастотная передача)

Этот способ напоминает вещание обыкновенной радиостанции. Пользователи настраивают передатчики и приемники на определенную частоту. При этом прямая видимость необязательна, площадь вещания составляет около 46 500 м² (500 000 квадратных футов). Однако поскольку используется сигнал высокой частоты, он не проникает через металлические или железобетонные преграды.

Доступ к такому способу связи осуществляется через поставщика услуг. Связь относительно медленная (около 4,8 Мбит/с).

3.4.5 Радиопередача в рассеянном спектре

При этом способе сигналы передаются на нескольких частотах, что позволяет избежать проблем, присущих одночастотной передаче.

Доступные частоты разделены на каналы. Адаптеры в течение заданного промежутка времени настроены на определенный канал, после чего переключаются на другой. Переключение всех компьютеров в сети происходит синхронно. Данный способ передачи обладает некоторой «встроенной» защитой: чтобы подслушать передачу, необходимо знать алгоритм переключения каналов.

Если необходимо усилить защиту данных от несанкционированного доступа, применяют кодирование.

Скорость передачи в 250 Кбит/с (килобит в секунду) относит данный способ к разряду самых медленных. Но есть сети, которые передают данные со скоростью до 2 Мбит/с на расстояние до 3,2 км (2 миль) — на открытом пространстве и до 120 м (393 футов) — внутри здания.

Это тот случай, когда технология позволяет получить настоящему беспроводную сеть.

3.4.6 Передача «точка-точка»

Данный способ передачи несколько выходит за рамки существующего определения сети. Технология передачи «точка-точка» предусматривает обмен данными только между двумя компьютерами, а не между несколькими компьютерами и периферийными устройствами. Для того чтобы организовать сеть с беспроводной передачей, необходимо использовать дополнительные компоненты, такие, как одиночные трансиверы и хост-трансиверы.

3.4.7 Многоточечное беспроводное соединение

Компонент, называемый беспроводным мостом (wireless bridge), помогает установить связь между зданиями без помощи кабеля. Если обычный мост служит людям для перехода с одного берега реки на другой, то беспроводной мост прокладывает для данных путь между двумя зданиями. Мост AIRLAN/Bridge Plus, например, использует технологию радиопередачи в рассеянном спектре для создания магистрали, соединяющей две ЛВС. Расстояние между ними, в зависимости от условий, может достигать 5 км (3-х миль). Стоимость такого устройства не покажется чрезмерной, поскольку арендовать линии связи больше не надо.

3.5 Назначение платы сетевого адаптера

Платы сетевого адаптера выступают в качестве физического интерфейса между компьютером и средой передачи. Платы вставляются в слоты расширения всех сетевых компьютеров и серверов или интегрируются на материнскую плату.

Для того чтобы обеспечить физическое соединение между компьютером и сетью, к соответствующему разъему или порту платы подключается сетевой кабель.

Назначение платы сетевого адаптера:

- подготовка данных, поступающих от компьютера, к передаче по сетевому кабелю;
- передача данных другому компьютеру;
- управление потоком данных между компьютером и кабелем.

Плата сетевого адаптера, кроме того, принимает данные из кабеля и переводит их в форму, понятную центральному процессору компьютера.

Плата сетевого адаптера состоит из аппаратной части и встроенных программ, записанных в ПЗУ.

3.5.1 Подготовка данных

Перед тем как послать данные в сеть, плата сетевого адаптера должна перевести их из формы, понятной компьютеру, в форму, в которой они могут передаваться по сетевому кабелю.

Внутри компьютера данные передаются по шинам. Шина — это несколько проводников, расположенных параллельно друг другу. Так как линий несколько, то и биты данных передаются по ним блоками, а не последовательно.

В первых персональных компьютерах IBM использовались 8-разрядные шины: они могли передавать блоки по 8 битов данных. IBM PC/AT® имеет 16-разрядную шину — это означает, что она способна передавать сразу 16 бит. Большинство современных компьютеров оснащены уже 64-разрядной шиной. Часто говорят, что данные по шине компьютера передаются параллельно (parallel), так как 16 битов или 64 бита движутся параллельно друг другу. Представьте, что 16-разрядная шина — это 16-

полосная автострада, по которой рядом (параллельно) едут 16 машин, каждая из которых перевозит один бит.

В сетевом кабеле данные должны перемещаться в виде потока битов. При этом говорят, что происходит последовательная (serial) передача, потому что биты следуют друг за другом. Иными словами, кабель — это дорога с одной полосой. По таким «дорогам» данные в каждый момент времени движутся в одном направлении.

Плата сетевого адаптера принимает параллельные данные и организует их для последовательной, побитовой, передачи. Этот процесс завершается переводом цифровых данных компьютера в электрические и оптические сигналы, передающиеся по сетевым кабелям. Отвечает за это преобразование трансивер (приемопередатчик). Плата сетевого адаптера, помимо преобразования данных, должна указать свое местонахождение, или адрес, чтобы ее могли отличить от остальных плат.

Сетевые адреса (network address) находятся в ведении комитета IEEE (Institute of Electrical and Electronics Engineers, Inc). Этот комитет закрепляет за каждым производителем плат сетевого адаптера некоторый интервал адресов. Затем каждый производитель записывает в ПЗУ платы ее уникальный сетевой адрес.

При приеме данных от компьютера и подготовке их к передаче по сетевому кабелю плата сетевого адаптера выполняет и другие действия:

1. Компьютер и плата сетевого адаптера должны быть связаны друг с другом, чтобы осуществлять передачу данных от компьютера к плате. Если плата может использовать прямой доступ к памяти, компьютер выделяет ей некоторую область своей памяти.

2. Плата сетевого адаптера запрашивает у компьютера данные.

3. Шина компьютера передает данные из его памяти плате сетевого адаптера.

Часто данные поступают быстрее, чем их способна передать плата сетевого адаптера, поэтому временно они помещаются в буфер.

3.5.2 Передача и управление данными

Перед тем как послать данные в сеть, плата сетевого адаптера проводит электронный диалог с принимающей платой, во время которого они «обговаривают»:

- максимальный размер (блока передаваемых данных);
- объем данных, передаваемых без подтверждения о получении;
- интервалы между передачами блоков данных;
- интервал, в течение которого необходимо послать подтверждение;
- объем данных, который может принять каждая плата без переполнения буфера;
- скорость передачи.

Если новой (более сложной и быстрой) плате приходится взаимодействовать со старой (медленной) платой, то они должны найти общую для них обеих скорость передачи. Схемы современных плат сетевого адаптера позволяют им приспособиться к медленной скорости старых плат.

Каждая плата оповещает другую о своих параметрах, принимая «чужие» параметры и подстраиваясь к ним. После того как все детали определены, платы начинают обмен данными.

3.6 Работа сети

Работа сети заключается в передаче данных от одного компьютера к другому. В этом процессе можно выделить следующие задачи:

- распознать данные;
- разбить данные на управляемые блоки;
- добавить информацию к каждому блоку, чтобы:
 - указать местонахождение данных;
 - указать получателя;
- добавить информацию синхронизации и информацию для проверки ошибок;
- поместить данные в сеть и отправить их по заданному адресу.

Сетевая операционная система при выполнении всех задач следует строгому набору процедур. Эти процедуры называются протоколами или правилами поведения. Протоколы регламентируют каждую сетевую операцию.

Стандартные протоколы позволяют программному и аппаратному обеспечению различных производителей нормально взаимодействовать. Существует два главных набора стандартов: модель OSI и ее модификация, называемая Project 802. Для изучения технической стороны функционирования сетей необходимо иметь четкое представление об этих моделях.

3.6.1 Эталонная модель взаимодействия открытых систем

В вычислительной технике начало развития идеологии построения открытых систем связывается с созданием компьютеров серии IBM 360, позволявших использовать одно и то же системное и прикладное программное обеспечение на любых компьютерах с IBM-подобной архитектурой. Данный подход оказался достаточно эффективным и получил дальнейшее развитие при построении персональных компьютеров той же фирмы. В частности, это определило одну из причин широкого распространения IBM-совместимых компьютеров на мировом компьютерном рынке.

Как уже отмечалось, современным системам телеобработки свойственен принцип «открытых» систем, поэтому естественным является использование данного подхода и в компьютерных сетях. В рамках сетевых технологий «открытость» систем используется с целью обеспечения возможности подключения к компьютерной сети оборудования различных фирм без дополнительной доработки сетевого программного и аппаратного обеспечения. При этом основным и, пожалуй, единственным условием является то, чтобы подключаемые системы также отвечали требованиям модели взаимодействия открытых систем.

Стремление к максимальному упорядочению и упрощению процессов разработки, модернизации и расширения сетей определило необходимость введения стандартов, регламентирующих принципы и процедуры организации взаимодействия абонентов компьютерных сетей. Интенсивные работы в данном направле-

нии ведутся рядом международных организаций, таких, как Международная организация стандартов (ISO), Международный консультативный комитет по телефонии и телеграфии (ССГТТ), Европейская ассоциация производителей компьютеров (European Computer Manufacture Association — ECMA) и др.

Международной организацией стандартов был специально создан Технический комитет ТС 97, один из подкомитетов которого (подкомитет SC 16) занимается разработкой стандартов для компьютерных сетей. Первой задачей, решенной в рамках стандартизации компьютерных сетей, было определение структуры построения стандартов и принципов организации работ по их созданию. Основопологающим результатом работы в данном направлении явилось создание Стандарта 7498, определяющего так называемую Базовую эталонную модель взаимодействия открытых систем. Впоследствии этот стандарт был принят за основу всеми организациями, занимающимися разработкой стандартов в области компьютерных сетей. Разработку и внедрение эталонной модели взаимодействия открытых систем можно считать одним из важнейших результатов в области стандартизации компьютерных сетей, способствующих широкому внедрению их в различные сферы человеческой деятельности.

Данный стандарт определяет:

- понятия и основные термины, используемые при построении открытых систем;
- описание возможностей и набора конкретных услуг, которые должна предоставлять открытая система;
- логическую структуру открытых систем; протоколы, обеспечивающие услуги открытых систем.

3.6.2 Понятие открытой системы

При рассмотрении взаимодействия структурных элементов компьютерных сетей вводится понятие «система», под которой подразумевается сервер, абонентская или любая другая система, предоставляющая или потребляющая сетевые ресурсы. В соответствии со стандартом 7498 *открытой системой* считается система, отвечающая требованиям эталонной модели взаимодействия открытых систем, реализующая стандартный набор услуг

и поддерживаемая стандартными протоколами. Соблюдение этих требований обеспечивает возможность взаимодействия открытых систем между собой, несмотря на их технические и логические различия в реализации, что является достаточно существенным фактором построения компьютерных сетей. Открытые системы объединяются с помощью сети передачи данных в *открытую компьютерную сеть*. Следует подчеркнуть, что модель взаимодействия открытых систем не рассматривает структуру и характеристики физических средств соединения, а только определяет основные требования к ним. *Основной же задачей модели взаимодействия открытых систем* является описание множества функций, определяющих правила взаимодействия открытых систем. При этом широко используется понятие «*процесс*», определяемый как динамический объект, реализующий целенаправленный акт обработки информации. Подобная формализация позволяет выделить характерные черты процесса взаимодействия систем, независимо от средств его реализации. Дело в том, что при многопользовательском режиме работы, который характерен для современных компьютеров, выполнение одной и той же программы в различные моменты времени может осуществляться по-разному. Это зависит от ряда факторов и в первую очередь — от числа задач в системе, порядка их выполнения и предоставляемых им ресурсов системы. Таким образом, программа не может однозначно определять функционирование систем и порядок их взаимодействия, с этой целью и вводится понятие процесса.

3.6.3 Процессы

Принято подразделять процессы на прикладные и системные. Прикладной процесс отождествляется с реализацией определенных процедур, связанных с обработкой информации при решении пользовательских задач. Системные же процессы определяют выполнение вспомогательных функций, связанных с обеспечением прикладных процессов. К системным процессам относятся: организация связи между прикладными процессами, управление каналами передачи данных, активизация терминалов и др. Процесс, как любой динамический объект, протекает во времени и состоит из этапов инициализации, выполнения и за-

вершения. При этом процесс может порождаться пользователем, системой или другим процессом. Ввод данных, необходимых процессу, и вывод данных производятся в форме сообщений через логические (программно-организованные) точки, называемые *портами*. Различают входные и выходные порты. Через входные порты осуществляется ввод данных для данного процесса, соответственно, через выходные порты текущий процесс выдает результаты обработки данных. Взаимодействие процессов осуществляется путем обмена сообщениями, которые представляют собой блоки данных определенной структуры. Промежуток времени, в течение которого взаимодействуют процессы, принято называть *сеансом* обмена или *сессией*. Во время сеанса обмена процесс формирует сообщение и необходимую для его передачи сопутствующую информацию. В зависимости от решаемой задачи взаимодействующие процессы могут генерироваться в одной или смежных системах.

Перейдем к рассмотрению модели взаимодействия открытых систем. Основу данной модели составляет концепция многоуровневой организации протоколов, которую можно рассматривать в качестве дальнейшего развития многоуровневой организации протоколов систем телеобработки. Существенной особенностью модели взаимодействия открытых систем является разработка и использование единого подхода к организации протоколов и интерфейсов различных уровней. В соответствии с данной концепцией каждому уровню ставится в соответствие набор определенных функций, связанных с решением конкретной задачи по организации взаимодействия открытых систем. Нумерация уровней осуществляется относительно физических средств соединения, то есть первый номер присваивается физическому уровню, а наибольший номер — прикладному (пользовательскому) уровню. Каждый уровень с меньшим номером считается вспомогательным для смежного с ним более высокого уровня и предоставляет ему определенный набор услуг, называемых *сервисом*. Следует подчеркнуть, что эталонная модель не определяет средства реализации протоколов, а только специфицирует их. Таким образом, функции каждого уровня могут быть реализованы различными аппаратными и программными средствами. Основным условием при этом является то, что взаимодействие между

любыми смежными уровнями должно быть четко определенным, то есть осуществляться через точки доступа посредством стандартного межуровневого интерфейса. *Точка доступа* является портом, в котором объект N-го уровня предоставляет услуги (N+1)- уровню. Это достаточно важное условие определяет возможность изменения протоколов отдельных уровней без изменения системы в целом, что является одним из основных условий построения открытых систем. Заметим, что в случае программной реализации межуровневого интерфейса в качестве портов выступают адреса, по которым заносятся межуровневые сообщения.

В процессе построения любой многоуровневой структуры возникает задача определения оптимального числа ее уровней. Так, при разработке эталонной модели число ее уровней определялось из следующих соображений:

- разбивка на уровни должна максимально отражать логическую структуру компьютерной сети;
- межуровневые границы должны быть определены таким образом, чтобы обеспечивались минимальное число и простота межуровневых связей;
- считается, что большое количество уровней, с одной стороны, упрощает внесение изменений в систему, а с другой стороны, увеличивает количество межуровневых протоколов и затрудняет описание модели в целом.

3.6.4 Модель OSI

В 1978 году International Standards Organization (ISO) выпустила набор спецификаций, описывающих архитектуру сети с неоднородными устройствами. Исходный документ относился к открытым системам, чтобы все они могли использовать одинаковые протоколы и стандарты для обмена информацией.

В 1984 году ISO выпустила новую версию своей модели, названную эталонной моделью взаимодействия открытых систем (Open System Interconnection reference model, OSI). Версия 1984 года стала международным стандартом: ее спецификации используют производители при разработке сетевых продуктов, она лежит в основе построения различных сетей.

Эта модель — широко распространенный метод описания сетевых сред. Являясь многоуровневой системой, она отражает взаимодействие программного и аппаратного обеспечения при осуществлении сеанса связи, а также помогает решить разнообразные проблемы.

В модели OSI сетевые функции распределены между семью уровнями. Каждому уровню соответствуют различные сетевые операции, оборудование и протоколы.

7. Прикладной уровень.
6. Представительский уровень.
5. Сеансовый уровень.
4. Транспортный уровень.
3. Сетевой уровень.
2. Канальный уровень.
1. Физический уровень.

Выше представлена многоуровневая архитектура модели OSI. На каждом уровне выполняются определенные сетевые функции, которые взаимодействуют с функциями соседних уровней, лежащих выше и ниже. Например, Сеансовый уровень должен взаимодействовать только с Представительским и Транспортным уровнями и т.д. Все эти функции подробно описаны.

Нижние уровни — 1-й и 2-й — определяют физическую среду передачи данных и сопутствующие задачи (такие, как передача битов данных через плату сетевого адаптера и кабель). Самые верхние уровни определяют, каким способом осуществляется доступ приложений к услугам связи. Чем выше уровень, тем более сложную задачу он решает.

Каждый уровень предоставляет несколько услуг (т.е. выполняет несколько операций), которые готовят данные для доставки по сети на другой компьютер. Уровни отделяются друг от друга границами — интерфейсами. Все запросы от одного уровня к другому передаются через интерфейс. Каждый уровень при выполнении своих функций использует услуги нижележащего уровня.

3.6.5 Взаимодействие уровней модели OSI

Задача каждого уровня — предоставить услуги вышележащему уровню, «маскируя» детали реализации этих услуг. При

этом каждый уровень на компьютере-отправителе работает так, будто он напрямую связан с таким же уровнем на компьютере-получателе. Однако в действительности связь осуществляется между смежными уровнями одного компьютера — программное обеспечение каждого уровня реализует определенные сетевые функции в соответствии с набором протоколов.

Перед подачей в сеть данные разбиваются на пакеты. Пакет (packet) — это единица информации, передаваемая между устройствами сети как единое целое. Пакет проходит последовательно через все уровни программного обеспечения. На каждом уровне к пакету добавляется некоторая информация, формирующая или адресная, которая необходима для успешной передачи данных по сети.

На принимающей стороне пакет проходит через все уровни в обратном порядке. Программное обеспечение на каждом уровне читает информацию пакета, затем удаляет информацию, добавленную к пакету на этом же уровне отправляющей стороной, и передает пакет следующему уровню. Когда пакет дойдет до Прикладного уровня, вся адресная информация будет удалена, и данные примут свой первоначальный вид.

Таким образом, за исключением самого нижнего уровня сетевой модели, никакой иной уровень не может непосредственно послать информацию соответствующему уровню другого компьютера. Информация на компьютере-отправителе должна пройти через все уровни. Затем она передается по сетевому кабелю на компьютер-получатель и опять проходит сквозь все слои, пока не достигнет того же уровня, с которого она была послана на компьютере-отправителе. Например, если Сетевой уровень передает информацию с компьютера А, она спускается через Канальный и Физический уровни в сетевой кабель, далее по нему попадает в компьютер В, где поднимается через Физический и Канальный уровни и достигает Сетевого уровня.

Примером информации, переданной Сетевым уровнем компьютера А Сетевому уровню компьютера В, могли бы служить адрес и, очевидно, информация контроля ошибок, добавленные к пакету.

Взаимодействие смежных уровней осуществляется через интерфейс. Интерфейс определяет услуги, которые нижний уро-

вень предоставляет верхнему, и способ доступа к ним. Каждому уровню одного компьютера «кажется», что он непосредственно взаимодействует с таким же уровнем другого компьютера.

Далее описывается каждый из семи уровней модели OSI и определяются услуги, которые они предоставляют смежным уровням.

Прикладной уровень

Уровень 7. Прикладной (Application) — самый верхний уровень модели OSI. Он представляет собой окно для доступа прикладных процессов к сетевым услугам. Этот уровень обеспечивает услуги, напрямую поддерживающие приложения пользователя, такие, как программное обеспечение для передачи файлов, доступа к базам данных и электронную почту. Нижележащие уровни поддерживают задачи, выполняемые на Прикладном уровне. Прикладной уровень управляет общим доступом к сети, потоком данных и восстановлением после сбоев связи.

Представительский уровень

Уровень 6. Представительский (Presentation) определяет формат, используемый для обмена данными между сетевыми компьютерами. Этот уровень можно назвать переводчиком. На Представительском уровне компьютера-отправителя данные, поступившие от Прикладного уровня, переводятся в общепонятный промежуточный формат. На этом же уровне компьютера-получателя происходит обратный перевод: из промежуточного формата в тот, который используется Прикладным уровнем данного компьютера. Представительский уровень отвечает за преобразование протоколов, трансляцию данных, их шифрование, смену или преобразование применяемого набора символов (кодовой таблицы) и расширение графических команд. Представительский уровень, кроме того, управляет сжатием данных для уменьшения общего числа передаваемых битов.

На этом уровне работает утилита, называемая редиректором (redirector). Ее назначение — перенаправлять локальные операции ввода/вывода на сетевой сервер.

Сеансовый уровень

Уровень 5. Сеансовый (Session) позволяет двум приложениям на разных компьютерах устанавливать, использовать и завершать соединение, называемое сеансом. На этом уровне выполняются такие функции, как распознавание имен и защита, необходимые для связи двух приложений в сети.

Сеансовый уровень обеспечивает синхронизацию между пользовательскими задачами посредством расстановки в потоке данных контрольных точек (checkpoints). Таким образом, в случае ошибки потребуется заново передать только данные, следующие за последней контрольной точкой. Этот уровень управляет диалогом между взаимодействующими процессами, т.е. регулирует, какая из сторон когда, как долго и т.д. должна осуществлять передачу.

Транспортный уровень

Уровень 4. Транспортный (Transport) располагается ниже Сеансового уровня. Транспортный уровень гарантирует доставку пакетов без ошибок, в той же последовательности, без потерь и дублирования. На этом уровне компьютера-отправителя сообщения переупаковываются: длинные разбиваются на несколько пакетов, а короткие объединяются в один. Это увеличивает эффективность передачи пакетов по сети. На Транспортном уровне компьютера-получателя сообщения распаковываются, восстанавливаются в первоначальном виде и обычно посылается сигнал подтверждения приема.

Транспортный уровень управляет потоком сообщений, проверяет ошибки и участвует в решении проблем, связанных с отправкой и получением пакетов.

Сетевой уровень

Уровень 3. Сетевой (Network) отвечает за адресацию сообщений и перевод логических адресов и имен в физические адреса. Одним словом, исходя из конкретных сетевых условий, приоритета услуги и других факторов здесь определяется маршрут от компьютера-отправителя к компьютеру-получателю. На этом уровне решаются также такие задачи и проблемы, связанные

с сетевым трафиком, как коммутация пакетов, маршрутизация и перегрузки.

Если сетевой адаптер маршрутизатора не может передавать большие блоки данных, посланные компьютером-отправителем, на Сетевом уровне эти блоки разбиваются на меньшие. А Сетевой уровень компьютера-получателя собирает эти данные в исходное состояние.

Канальный уровень

Уровень 2. Канальный (Data Link) осуществляет передачу кадров (frames) данных от Сетевого уровня к Физическому. Кадры — это логически организованная структура, в которую можно помещать данные. Канальный уровень компьютера-получателя упаковывает «сырой» поток битов, поступающих от Физического уровня, в кадры данных. Управляющая информация используется для маршрутизации, а также указывает на тип пакета и сегментацию. Данные — собственно передаваемая информация. CRC (остаток избыточной циклической суммы) — это сведения, которые помогут выявить ошибки, что, в свою очередь, гарантирует правильный прием информации.

Канальный уровень обеспечивает точность передачи кадров между компьютерами через Физический уровень. Это позволяет Сетевому уровню считать передачу данных по сетевому соединению фактически безошибочной.

Обычно, когда Канальный уровень посылает кадр, он ожидает со стороны получателя подтверждения приема. Канальный уровень получателя проверяет наличие возможных ошибок передачи. Кадры, поврежденные при передаче или не получившие подтверждения о приеме, посылаются заново.

Физический уровень

Уровень 1. Физический (Physical) — самый нижний в модели OSI. Этот уровень осуществляет передачу неструктурированного, «сырого» потока битов по физической среде (например, по сетевому кабелю). Здесь реализуются электрический, оптический, механический и функциональный интерфейсы с кабелем. Физический уровень также формирует сигналы, которые переносят данные, поступившие от всех вышележащих уровней.

На этом уровне определяется способ соединения сетевого кабеля с платой сетевого адаптера, в частности количество контактов в разъемах и их функции. Кроме того, здесь определяется способ передачи сигналов по сетевому кабелю.

Физический уровень предназначен для передачи битов (нулей и единиц) от одного компьютера к другому. Содержание самих битов на данном уровне значения не имеет. Этот уровень отвечает за кодирование данных и синхронизацию битов, гарантируя, что переданная единица будет воспринята именно как единица, а не как ноль. Наконец, Физический уровень устанавливает длительность каждого бита и способ перевода бита в соответствующие электрические или оптические импульсы, передаваемые по сетевому кабелю. Простое перечисление уровней недостаточно для определения правил взаимодействия систем, поэтому в рамках эталонной модели также определяются *услуги*, которые должны обеспечивать ее уровни. Услуги, по сути дела, представляют собой функции, выполняемые на заданном уровне.

В частности, физический уровень должен обеспечивать такие виды услуг, как установление и идентификацию физических соединений, организацию последовательностей передачи бит информации; оповещение об окончании связи.

Канальный уровень предоставляет следующие виды услуг: организацию требуемой последовательности блоков данных и их передачу; управление потоками между смежными узлами; идентификацию конечных пунктов канальных соединений; обнаружение и исправления ошибок; оповещение об ошибках, которые не исправлены на канальном уровне.

Сетевой уровень в числе основных услуг осуществляет: идентификацию конечных точек сетевых соединений; организацию сетевых соединений; управление потоками блоков данных; обеспечение последовательностей доставки блоков данных; обнаружение ошибок и формирование сообщений о них; разъединение сетевых соединений.

Транспортный уровень обеспечивает следующие виды услуг: установление и разъединение транспортных соединений; формирование блоков данных; обеспечение взаимодействия сеансовых соединений с транспортными соединениями; управление последовательностью передачи блоков данных; обеспечение це-

лостности блоков данных во время передачи; обнаружение и устранение ошибок, сообщение о неисправленных ошибках; предоставление приоритетов в передаче блоков; передачу подтверждений о принятых блоках; ликвидацию тупиковых ситуаций.

3.6.6 Услуги различных уровней модели OSI

На сеансовом уровне предоставляются услуги, связанные с обслуживанием сеансов и обеспечением передачи данных в диалоговом режиме; установление сеансового соединения; обмен данными; управление обменом; синхронизация сеансового соединения; сообщения об исключительных ситуациях; отображение сеансового соединения на транспортный уровень; завершение сеансового соединения.

Представительный уровень обеспечивает следующие виды услуг: выбор вида представления данных; интерпретация и преобразование передаваемых данных к виду, удобному для прикладных процессов; преобразование синтаксиса данных; формирование данных.

Прикладной уровень обеспечивает широкий набор услуг: управление терминалами; управление файлами; управление диалогом; управление задачами; управление сетью в целом; целостность информации; ряд дополнительных услуг. К дополнительным услугам уровня относятся услуги по организации электронной почты, передачи массивов сообщений и т.п.

Услуги различных уровней определяются с помощью протоколов эталонной модели взаимодействия открытых систем, которые представляют собой правила взаимодействия объектов одноименных уровней открытых систем. При этом по функциональному назначению все протоколы целесообразно разделить на три группы.

Первую группу составляют *протоколы абонентской службы*, соответствующие прикладному, представительскому и сеансовому уровням модели взаимодействия открытых систем. Протоколы этой группы являются сетенезависимыми, то есть их характеристики и структура не зависят от используемой сети передачи данных. Они определяются лишь структурой абонентских систем и поставленными перед ними задачами обработки информации.

Две другие группы протоколов описывают *транспортную службу* компьютерной сети и различаются между собой процедурой доступа к передающей среде. Одна из этих групп определяет систему передачи данных с маршрутизацией информации, а другая — с селекцией информации.

Маршрутизация, представляет собой процедуру определения пути передачи информации в сетях передачи данных и характерна для глобальных и региональных компьютерных сетей, в рамках которых и рассматривается соответствующая группа протоколов.

Под *селекцией* в компьютерных сетях подразумевается процесс выбора очередной абонентской системы, для подключения ее к сети передачи данных с целью обмена информацией. Селекция информации в основном используется в системах передачи данных локальных компьютерных сетей, где и рассматривается третья группа протоколов.

3.7 Доступ абонентских систем к моноканалу

Эффективность взаимодействия абонентских систем в рамках локальной компьютерной сети во многом определяется используемым правилом доступа к общей передающей среде в сетях с шинной и кольцевой топологией или концентратору в древовидных и звездообразных сетях. Правило, с помощью которого организуется доступ абонентских систем к передающей среде, получило название *метода доступа*. В качестве критерия эффективности метода доступа, чаще всего рассматривается *время доступа к передающей среде*, представляющее собой промежуток времени между появлением запроса на передачу данных и собственно началом передачи информации. Значение этого параметра зависит от ряда факторов, в том числе от топологии сети, используемого метода доступа, способа управления сетью и др. В силу большого разнообразия локальных сетей и требований к ним нельзя назвать какой-либо универсальный метод доступа, эффективный во всех случаях. Каждый из известных методов доступа имеет свои определенные преимущества и недостатки. Кратко рассмотрим наиболее распространенные методы доступа.

Как и для всей сети в целом, управление доступом может быть как централизованным, так и децентрализованным. Централизованное управление доступом, как правило, осуществляется специальной управляющей (мониторной или супервизорной) станцией, подключаемой к передающей среде так же, как и любая другая абонентская система. При децентрализованном управлении каждая станция сама принимает решение о возможности доступа к передающей среде.

В зависимости от используемого метода доступа локальные сети делятся на две группы. К первой группе относятся сети с методами детерминированного доступа, ко второй — с методами случайного доступа. Метод детерминированного доступа предполагает наличие определенного алгоритма, на основании которого абонентским системам предоставляется доступ к передающей среде. Например, при централизованном управлении мониторинговая подсистема может последовательно опрашивать каждую из абонентских систем и предоставлять право передачи информации первой желающей абонентской системе. После этого будет опрашиваться следующая абонентская система и так далее. Алгоритм предоставления права передачи информации может быть достаточно гибким и учитывать приоритеты запросов на передачу и их интенсивность. Для нормального функционирования сети необходимо, чтобы она не находилась в режиме насыщения, т.е. нагрузка на сеть не должна превышать ее пропускную способность. В этом случае можно определить минимальное и максимальное значения времени ожидания момента начала передачи информации. Минимальное время ожидания связано с необходимостью опроса абонентских систем на предмет передачи информации. Влияние этой процедуры наиболее ощутимо при низкой интенсивности передачи информации, так как время на опрос затрачивается даже при отсутствии заявок на передачу информации со стороны других абонентских систем. При увеличении интенсивности информационного потока снижается отношение времени опроса ко всему времени передачи информации. Максимальное значение времени ожидания является фиксированной величиной, определяемой при полной загрузке сети, т.е. когда каждая абонентская система готова передавать информацию. При этом предполагается, что следующий запрос на передачу информации

в данной абонентской системе появляется после обслуживания очередного запроса. Возможность установления гарантированного времени доступа является достаточно существенным фактором при работе в режиме реального времени. В общем случае методы детерминированного доступа позволяют учитывать особенности топологии сети и характер передаваемой информации, обеспечивая наиболее эффективное использование передающей среды.

Ко второй группе относятся *методы случайного доступа*, при которых каждая абонентская система произвольным образом, независимо от других систем, может обращаться к моноканалу. При методе случайного доступа возможно одновременное обращение нескольких абонентских систем к общей передающей среде, поэтому данный метод доступа часто называют *методом множественного доступа*. Математически локальная сеть с множественным доступом может быть представлена (рис. 44) в виде с n входными потоками и одним обслуживающим устройством. Количество входных потоков соответствует числу абонентских систем. Интенсивность (X_j) i -го входного потока определяется интенсивностью потока заявок на передачу информации со стороны i -й абонентской системы. Обслуживающее устройство в данном случае представляет собой передающую среду. Время обслуживания заявки соответствует времени передачи блока данных и равно отношению длины блока данных, измеренной в битах, к скорости передачи информации по передающей среде.

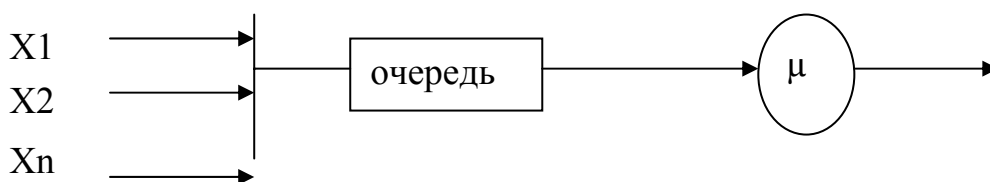


Рис. 44 — Математическая модель локальной сети с множественным доступом, где X — интенсивность входного потока, μ — интенсивность обслуживания

Соответственно, интенсивность обслуживания (μ) является величиной, обратной времени обслуживания. Длина входной очереди и время ожидания обслуживания заявок изменяются по экспоненциальному закону, резко возрастаая при увеличении интенсивности входного потока.

Сравнивая эти две группы методов доступа, можно отметить следующее. Методы случайного доступа проще в реализации, так как не требуют передачи специальной управляющей информации. Они более эффективны при обмене короткими сообщениями и низкой загрузке моноканала. В этом случае доступ к передающей среде осуществляется практически без дополнительных задержек. Методы детерминированного доступа более предпочтительны при обмене длинными сообщениями и повышении уровня загрузки моноканала. Они позволяют также при необходимости организовать приоритетную передачу сообщений.

В процессе работы сети информация от передающей абонентской системы поступает на адаптеры всех абонентских систем, однако воспринимается только адаптером той абонентской системы, которой она адресована. Использование абонентскими системами общей передающей среды предполагает решения задачи организации поочередного доступа к ней. Правило, с помощью которого организуется бесконфликтный доступ абонентских систем к передающей среде, получило название *метода доступа*.

3.7.1 Методы доступа в сетях с шинной топологией

В магистральных локальных сетях используются методы как случайного, так и детерминированного доступа. Появление методов случайного доступа связывают с радиосетью ALOHA, где впервые был использован метод случайного доступа. Абонентские системы передавали информацию в эфир независимо друг от друга. В случае одновременной передачи сообщений несколькими станциями происходило «столкновение» сообщений, подобный конфликт приводил к искажению информации. Во избежание приема ошибочной информации кадр данных дополняется контрольной суммой. Принимающая абонентская система выдает подтверждение только при приеме кадров с правильной контрольной суммой, остальные кадры игнорируются. Это позволяет передающей станции контролировать передачу кадров. Вероятность «столкновения» сообщений зависит от интенсивности обращения абонентских систем к передающей среде и существенно возрастает при ее увеличении. Снижение коэффициента полезного использования моноканала при возрастании количества

«столкновений» как следствие повышения интенсивности запросов на доступ определило поиск возможностей совершенствования метода случайного доступа. Одним из способов снижения конфликтов является предварительное прослушивание передающей среды и начало передачи только при наличии свободного канала. Такой режим передачи получил название *множественного доступа с контролем несущей частоты* (МДКН). Однако и в этом случае из-за конечного времени распространения сигналов невозможно полностью избежать конфликтов. Остановимся более подробно на этом вопросе. На рис. 45 представлен процесс столкновения пакетов.

В начальный момент времени T1-абонентская система B начала передавать информацию. В этот же момент времени абонентская система A прослушивает передающую среду, однако из-за конечного времени распространения сигнала ей не удастся обнаружить сообщение, посылаемое абонентской системой B. В следующий момент времени (T2) абонентская система A начинает передавать информацию, в результате чего в момент времени T3 сообщения «благополучно» сталкиваются. Дальнейшая передача сообщений теряет смысл.

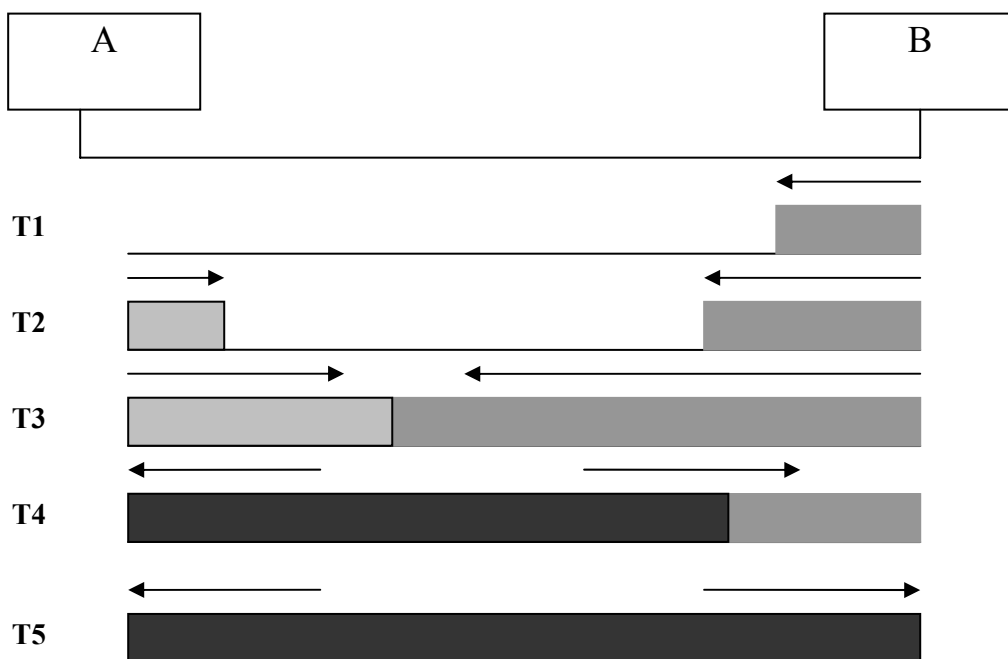


Рис. 45 — Столкновение сообщений

С целью своевременного обнаружения конфликтов абонентская система в процессе передачи информации постоянно контролирует передающую среду и при появлении «столкновения» прекращает передачу. Так, абонентская система *A* прекращает передачу в момент времени T_4 , а абонентская система *B* — в момент времени T_5 . Наличие конфликтов определяется путем сравнения передаваемой информации с информацией в канале передачи. При отсутствии посторонней передачи информация в канале должна соответствовать передаваемой информации, то же касается уровней сигналов. Через некоторый промежуток времени после прекращения передачи конфликтующие абонентские системы осуществляют повторную попытку передачи информации. Время задержки определяется с помощью специальных алгоритмов, направленных на снижение вероятности повторного конфликта. Например, задержка может формироваться так, чтобы ее среднее значение увеличивалось примерно вдвое с каждой новой попыткой занять моноканал. Подобный режим передачи получил название *множественного доступа с контролем несущей частоты и обнаружением столкновений* (МДКН/ОС или CSMA/CD).

Методы детерминированного доступа можно разделить на методы разделения времени и методы передачи полномочий.

Сущность методов разделения времени заключается в разделении времени работы канала связи на отдельные временные интервалы, каждый из которых, согласно определенному правилу, предоставляется какой-либо абонентской системе. Большинство методов разделения времени предусматривает наличие в сети диспетчера, основной функцией которого является контроль и планирование времени доступа. При этом появляется возможность учитывать приоритеты и необходимое время взаимодействия абонентских систем.

Наиболее простым среди методов разделения времени является метод синхронного (циклического) разделения времени. В этом случае цикл (T) обмена с абонентскими системами разбивается на несколько временных интервалов (t), количество которых соответствует числу (n) абонентских систем. Во время цикла обмена каждой абонентской системе предоставляется фиксированный интервал времени, в течение которого она может передавать сообщение. Если у абонентской системы в данный момент времени отсутствует

информация для передачи, то выделенный ей временной интервал не используется. При неравномерном распределении интенсивности обращения абонентских систем к передающей среде эффективность использования канала связи относительно низкая. Она может быть повышена за счет разделения цикла обмена на небольшие интервалы с представлением абонентской системе одного или нескольких интервалов в зависимости от интенсивности обращения абонентской системы к каналу связи.

Эффективность использования моноканала может быть также повышена за счет реализации методов асинхронного разделения времени, основанных на прогнозировании интенсивности запросов доступа к моноканалу со стороны абонентских систем. С помощью специальной процедуры накапливается статистика обращений, на основе которой прогнозируется интенсивность потоков заявок и распределяются временные интервалы между абонентскими системами. Как показывает практика, данный метод временного разделения эффективен лишь при небольшом числе абонентских систем. В локальных сетях с большим числом абонентов достаточно широко используется метод детерминированного доступа, получивший название множественного доступа с передачей полномочий (метод маркерного доступа).

В общем виде алгоритм маркерного доступа достаточно прост: в локальной сети последовательно от одной абонентской системы к другой передается специальная управляющая информация — маркер, при поступлении которого абонентская система получает разрешение на передачу информации. После окончания передачи абонентская система обязана передать маркер следующей абонентской системе. При отсутствии необходимости в передаче сообщения абонентская система, получившая маркер, немедленно передает его следующей абонентской системе. Последняя абонентская система передает маркер первой абонентской системе, образуя, таким образом, логическое кольцо (рис. 46) передачи маркера. При этом передача кадров данных осуществляется в обоих направлениях только получателем на основании сравнения адреса, указанного в передаваемом кадре, с адресом абонентской системы.

Данный способ доступа имеет ряд преимуществ:

- обеспечивает достаточно эффективное использование ре-

сурсов канала передачи данных; предоставляет возможность реализации режима работы в режиме реального времени;

- исключает столкновения сообщений;
- позволяет достаточно просто реализовать приоритетный доступ.

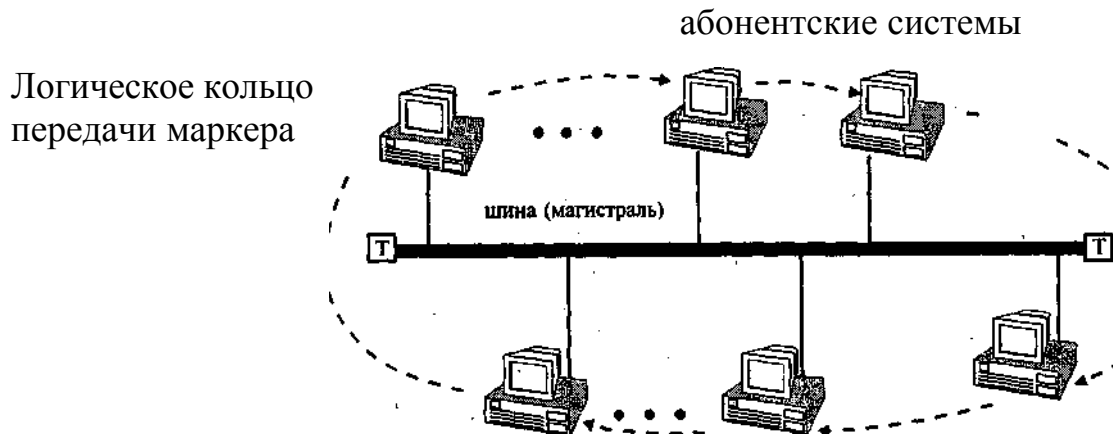


Рис. 46 — Организация логического кольца передачи полномочий (маркера)

К недостаткам метода следует отнести зависимость работы сети от физических характеристик передающей среды, в частности потеря маркера или его раздвоение приводит к неправильной работе сети. Поэтому необходимо с помощью специальных процедур постоянно отслеживать потерю маркера или появление нескольких маркеров.

3.7.2 Методы доступа в кольцевых сетях

В кольцевых локальных сетях используются, как правило, методы детерминированного доступа. Применение методов случайного доступа не имеет смысла при последовательной передаче информации, которой характеризуются кольцевые локальные сети, так как при этом отсутствует возможность прослушивания всего кольца для выявления возможных столкновений сообщений.

Основными методами доступа в локальных сетях с кольцевой структурой являются: метод множественного доступа с вве-

дением задержки, метод циклического доступа (тактируемый) и метод маркерного доступа.

Рассмотрим метод доступа с введением задержки. В данном случае информация между абонентскими системами передается в виде относительно коротких кадров данных фиксированной длины. Название метода связано с тем, что очередной кадр данных из абонентской системы «вклинивается» в поток кадров, поступающих по каналу передачи данных, в результате чего последующие кадры данных задерживаются на время передачи одного кадра. Взаимодействие абонентской системы с передающей средой осуществляется с помощью блока доступа, в состав которого входят: приемник, линия задержки, переключатель, передатчик и буферный регистр. Приемник, наряду с восстановлением физических параметров входных сигналов, осуществляет предварительный анализ поступающих кадров данных. Если кадр адресован данной абонентской системе, то его копия заносится в буферный регистр. Сам же кадр данных через переключатель (Р) передается в передатчик, в котором к нему добавляется служебная информация о состоянии абонентской системы и результате приема данных. Линия задержки служит для хранения информации, поступающей из сети, на время передачи кадра данных из абонентской системы. В регистре признака формируется информация о состоянии абонентской системы и принятом кадре данных.

Основным преимуществом доступа с введением задержки является минимальное время доступа к передающей среде, предельное значение которого равно времени передачи одного кадра. Так как каждая абонентская система может задержать передачу на время одного кадра, то максимальное время между передачами кадров одной абонентской системой определяется произведением числа абонентских систем на длительность передачи кадра. Таким образом, данный способ объединяет преимущества случайного и детерминированного методов доступа, т. к. при низкой нагрузке обеспечивает минимальное время доступа и передачи кадров, а при высокой — гарантированное время доступа. Однако при большом числе абонентов и высокой интенсивности обращения их к передающей среде существенно увеличивается время передачи кадров.

К недостаткам рассмотренного метода относится также блокировка абонентской системы, которая может иметь место в случае искажения или потери кадра данных, переданного этой системой.

Метод тактируемого доступа предполагает (рис. 47) разбиение временного цикла кольца, то есть времени распространения сигнала по кольцу канала связи, на множество равных временных интервалов — тактов (временных сегментов), в каждом из которых помещается по одному кадру. Таким образом, одновременно может передаваться несколько кадров. Количество и длина кадров определяются с учетом основных характеристик сети. Абонентская система может передавать информацию в кольцо только при прохождении через ее блок доступа свободного кадра. Свободные кадры отличаются от занятых значением специального контрольного бита своего заголовка. Единица указывает на то, что данный кадр занят, а ноль — свободен. Например, в сети, представленной на рис. 47 общее число тактов равно десяти. В текущий момент времени через первую абонентскую систему проходит занятый кадр, и она не может передавать информацию. Через вторую абонентскую систему проходит свободный кадр — абонентская система может передавать информацию, однако из-за отсутствия информации эта абонентская система не передает информацию. И, наконец, третья абонентская система, распознав свободный кадр, меняет значение контрольного бита на единицу и заполняет текущий кадр данных.

Адресат, получив кадр данных, копирует его. Освобождение (обнуление) кадров может осуществляться как получателем, так и отправителем информации.

В настоящее время известно много разновидностей данного метода доступа, но все они предполагают разбиение сообщений на пакеты с последующим формированием кадра и эффективны при обмене короткими сообщениями и высокой интенсивности обмена сообщениями.

При обмене большими сообщениями переменной длины предпочтительным является маркерный доступ. Основное отличие маркерного доступа в кольцевой сети от маркерного доступа в сети с шинной топологией заключается в том, что кадры маркера и данных передаются в одном направлении и по физическому

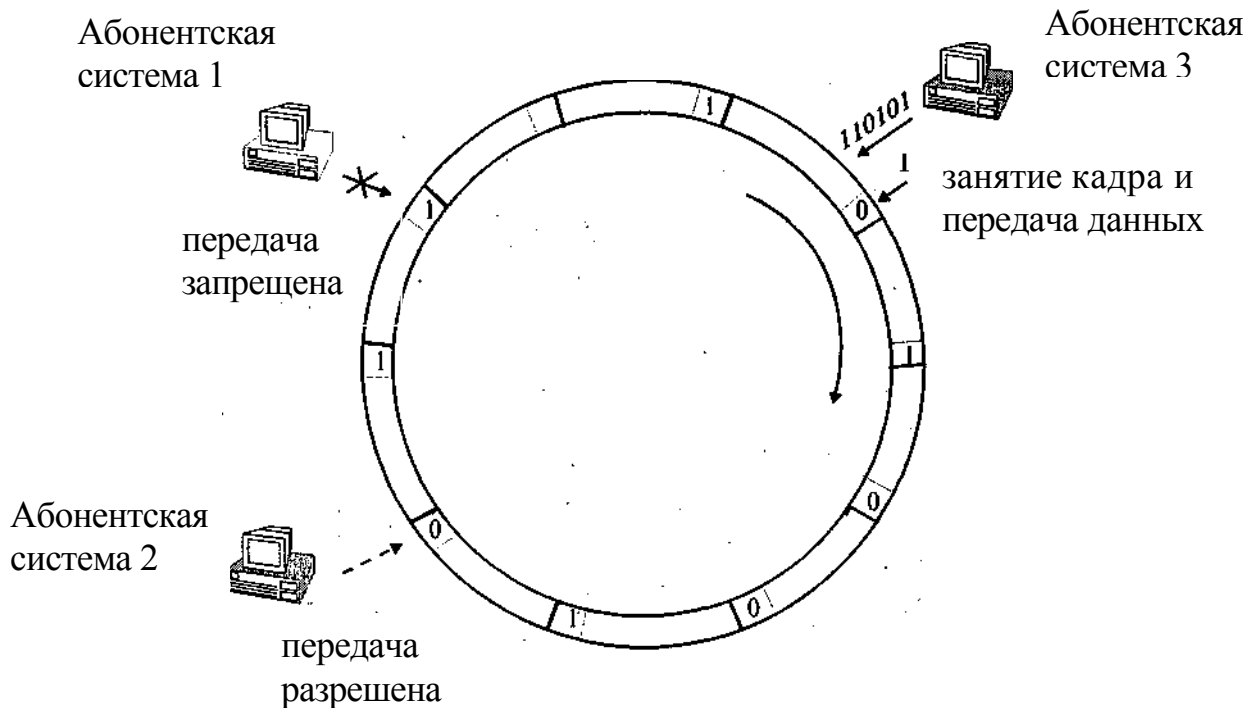


Рис. 47 — Организация тактируемого доступа

кольцу. Передача информации в произвольном направлении, как это происходит в сетях с шинной топологией, исключается. Абонентская система может начать передачу только после получения маркера от предыдущей абонентской системы. Получив маркер, станция посылает в кольцо кадр данных. Передача маркера следующей абонентской системе может осуществляться после возвращения переданного кадра данных либо сразу же после его передачи. Во втором случае говорят о режиме раннего освобождения маркера. При этом каждый последующий кадр данных оказывается помещенным между предыдущим кадром и маркером. Удаление принятых кадров, как правило, осуществляется передающей абонентской системой. В сетях с маркерным доступом необходимо контролировать потерю маркера и удаление полученных пакетов. Более подробно этот вопрос будет рассмотрен ниже.

3.7.3 Модель IEEE Project 802

В конце 70-х годов ЛВС признали наконец в качестве удобного инструмента для ведения бизнеса.

Существенный вклад в развитие стандартов по локальным компьютерным сетям внес Институт инженеров по электротехнике и радиоэлектронике (IEEE) США. В 1980 году в рамках этого института был образован комитет 802, задачей которого является разработка стандартов для локальных компьютерных сетей. Для подготовки проектов отдельных стандартов в рамках комитета 802 были созданы отдельные подкомитеты 802.1 — 802.9, номера которых и были присвоены соответствующим стандартам. Стандарты серии IEEE-802 определяют терминологию, архитектуру и протоколы локальных компьютерных сетей двух нижних уровней Эталонной модели взаимодействия открытых систем. В результате был выпущен Project 802, названный в соответствии с годом и месяцем своего издания (1980 год, февраль).

Хотя публикация стандартов IEEE опередила публикацию стандартов ISO, оба проекта велись приблизительно в одно время и при полном обмене информацией, что и привело к рождению двух совместимых моделей.

Project 802 установил стандарты для физических компонентов сети — интерфейсных плат и кабельной системы, с которыми имеют дело Физический и Канальный уровни модели OSI.

Итак, эти стандарты, называемые 802-спецификациями, распространяются на:

- платы сетевых адаптеров;
- компоненты глобальных вычислительных сетей;
- компоненты сетей, при построении которых используют коаксиальный кабель и витую пару.

На рис. 48 приведено соответствие уровней Эталонных моделей глобальной сети и локальной сети стандарта IEEE-802. Основное отличие заключается в том, что физический и канальный уровни разбиты на подуровни. В то же время верхние уровни не специфицируются. Это объясняется тем, что физический и канальный уровни собственно и определяют локальную сеть. Физический уровень включает подуровни: ПФС — передачи физических сигналов; МСС — модуля сопряжения со средой; ИМС — интерфейса с модулем сопряжения. Подобное разделение физического уровня на подуровни способствует унификации передающей среды. Далее, канальный уровень разбит на два подуровня: УЛК — управления логическим каналом и УДС — управле-

ния доступом к физической среде. В то же время функции управления логическим каналом одинаковы для различных; локальных сетей, поэтому их целесообразно рассматривать отдельно от функций управления доступом к передающей среде, что и реализовано в данном стандарте.

Уровни эталонной модели OSI

Уровни модели локальной сети IEEE

Прикладной	Верхние уровни
Представительный	
Сеансовый	
Транспортный	
Сетевой	
Канальный	УЛК
	УДС
Физический	ПФС
	ИМС
	МСС

Рис. 48 — Соответствие модели глобальной и локальной сетей, где УЛК — управление логическим каналом; УДС — управление доступом к среде; ПФС — передача физических сигналов; ИМС — интерфейс с модулем сопряжения; МСС — модуль сопряжения со средой

802-спецификации определяют способы, в соответствии с которыми платы сетевых адаптеров осуществляют доступ к физической среде и передают по ней данные. Сюда относятся соединение, поддержка и разъединение сетевых устройств.

3.7.4 Категории стандартов ЛВС

Стандарты ЛВС, определенные Project 802, делятся на 12 категорий, каждая из которых имеет свой номер.

802.1 — Объединенные сети.

802.2 — Управление логической связью.

802.3 — ЛВС с множественным доступом, контролем несущей и обнаружением коллизий (Ethernet).

802.4 — ЛВС топологии «шина» с передачей маркера.

802.5 — ЛВС топологии «кольцо» с передачей маркера.

802.6 — сеть масштаба города (Metropolitan Area Network, MAN).

802.7 — Консультативный совет по широковещательной технологии (Broadcast Technical Advisory Group).

802.8 — Консультативный совет по оптоволоконной технологии (Fiber-Optic Technical Advisory Group).

802.9 — Интегрированные сети с передачей речи и данных (Integrated Voice/Data Networks).

802.10 — Безопасность сетей.

802.11 — Беспроводные сети.

802.12 — ЛВС с доступом по приоритету запроса (Demand Priority Access LAN, IGObaseVG-AnyLan).

Структура стандартов IEEE-802 представлена на рис. 49. Стандарт IEEE-802.1 является общим документом, который определяет архитектуру и прикладные процессы системного управления сетью, методы объединения сетей на подуровне управления доступом к передающей среде. Стандарт IEEE-802.2 определяет протоколы управления логическим каналом. Каждый из остальных стандартов, начиная с IEEE-802.3, определяет метод доступа и специфику физического уровня для конкретного типа локальной компьютерной сети. Так, стандарт IEEE-802.3 описывает характеристики и процедуры множественного доступа с контролем передачи и обнаружения столкновений. Стандарт IEEE-802.4 определяет протокол маркерного доступа к моноканалу. Процедуры и характеристики маркерного доступа к кольцевой локальной сети определяется стандартом IEEE-802.5. Для построения локальных сетей, охватывающих площадь радиусом до 25 км и использующих технические средства кабельного телевидения, разработан стандарт IEEE-802.6. В подкомитете IEEE-802.11 разработан стандарт на радиосети для мобильных компьютеров, а в комитете IEEE-802.12 — стандарт на высокоскоростные компьютерные сети IOOVG-AnyLAN.

В 1985 году стандарт IEEE-802 был принят Международной организацией стандартов за основу международных стандартов физического и канального уровней ISO/J[^]DIS 8802/2.2 — ISO/DIS 8802/5. Кроме того, эти стандарты были дополнены стандартом

ISO/DIS 8802/7 на сети с тактируемым методом доступа к кольцу, разработанным на основе протоколов доступа локальной сети Cambridge Ring.

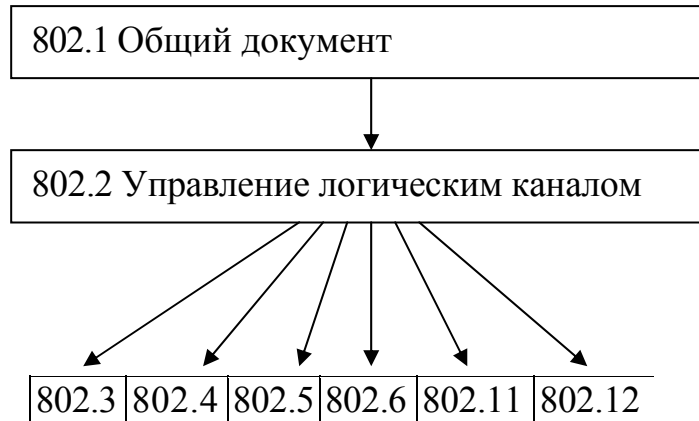


Рис. 49 — Структура стандартов IEEE 802.X

3.7.5 Расширения модели OSI

Два нижних уровня модели OSI, Физический и Канальный, устанавливают, каким образом несколько компьютеров могут одновременно использовать сеть, чтобы при этом не мешать друг другу.

IEEE Project 802 относился именно к этим двум уровням и привел к созданию спецификаций, определивших доминирующие среды ЛВС.

IEEE, подробно описывая Канальный уровень, разделил его на два подуровня:

- Управление логическим каналом (Logical Link Control, LLC) — установление и разрыв соединения, управление потоком данных, упорядочивание и подтверждение приема кадров.
- Управление доступом к среде (Media Access Control, MAC) — управление доступом к среде передачи, определение границ кадров, контроль ошибок, распознавание адресов кадров.

Подуровень Управления логическим каналом устанавливает канал связи и определяет использование логических точек интерфейса, называемых точками доступа к услугам (Service-Access Points, SAP). Другие компьютеры, ссылаясь на точки доступа к услугам, могут передавать информацию с подуровня Управле-

ния логическим каналом на верхние уровни OSI. Эти стандарты определены в категории 802.2.

Подуровень Управления доступом к среде — нижний из двух подуровней. Он обеспечивает совместный доступ плат сетевого адаптера к Физическому уровню. Подуровень Управления доступом к среде напрямую связан с платой сетевого адаптера и отвечает за безошибочную передачу данных между двумя компьютерами сети.

Категории 802.3, 802.4, 802.5 и т.д. определяют стандарты как для этого подуровня, так и для первого уровня модели OSI — Физического.

ЗАКЛЮЧЕНИЕ

Когда РС были просто автономными машинами, изучение их устройства можно было ограничить рассмотрением механизмов действия и назначения их отдельных компонентов. Теперь же, когда РС соединены со всем миром, главным образом через Internet, а локальные сети становятся обычным явлением не только для офиса, но и для дома, необходимо рассмотреть три стороны вопроса: о внутреннем устройстве РС, как такового, о компонентах, обеспечивающих его связь с внешним миром, и о влиянии этого мира на него. Эти стороны, подробно рассмотрены. Мобильные РС захватили новые высоты, заменяя во многих случаях настольные, а не просто дополняя их. У этих портативных машин часто своя собственная история. Кроме того, в этом курсе не уделяется большого внимания операционным системам.

Если Вы владелец персонального компьютера, значит, вы и в ответе за него. Возможно, заплатил за него кто-то другой (вам повезло). Возможно, этот кто-то за ним и ухаживает, и все же, вы — владелец персонального компьютера. Вам принадлежит право (и ответственность) решать, когда и как использовать РС. Не перекладывайте эту ответственность ни на кого.

Делайте свой выбор и настаивайте на нем, если хотите и можете заставить ваш РС работать по-своему.

Пожалуйста, продолжайте изучать РС и все, что с ними связано. Это один из способов (далеко не самый худший) оставаться умственно живым и быть интересным и себе, и людям. Если вы столкнетесь с чем-либо действительно интересным, возможно, чем-то, что, по вашему мнению, поможет другим узнать новое о РС, или с чем-то, восхитившим вас, я буду рад этому.

СПИСОК ЛИТЕРАТУРЫ

Основная литература

1. Борзенко А.Е. IBM PC: устройство, ремонт, модернизация. — 2-е изд., перер. и доп. — М.: ТОО фирма «Компьютер Пресс», 1996. — 344 с.
2. Пилгрим А. Персональный компьютер: [Кн. 2]: Модернизация и ремонт: Пер. с англ. / Под общ. ред. С. Ракова. — СПб.: BHV-Санкт-Петербург, 2000. — 528 с.: ил.
3. Нортон Питер, Гудман Джон. Персональный компьютер: [Кн. 1]: Аппаратно-программная организация: Пер. с англ. — СПб.: BHV-Санкт-Петербург, 2000. — 848 с.: ил.
4. Гук Михаил. Аппаратные средства IBM PC: Энциклопедия. — СПб.: Питер, 2001. — 816 с.: ил.
5. Хоган Том. Аппаратные и программные средства персональных компьютеров. Кн.1. — М.: Радио и связь, 1995. — 384 с.: ил.
6. Хоган Том. Аппаратные и программные средства персональных компьютеров. Кн.2. — М.: Радио и связь, 1995. — 376 с.: ил.
7. Рудометов Евгений, Рудометов Виктор. Аппаратные средства и мультимедиа. — СПб.: Питер, 1999. — 352 с.: ил.
8. Рыбаков М.А. Анатомия персонального компьютера. — М.: Интермеханика, 1990. — 224 с.: ил.
9. Бэрри Нанс. Компьютерные сети: Пер. с англ. — М.: БИНОМ, 1996. — 400 с.: ил.
10. Ветров С.И. Компьютерное «железо». — М.: Солон-Р, 2001. — 560 с.

Дополнительная литература

1. Фролов А.В., Фролов Г.В. Аппаратное обеспечение IBM PC. Ч.1. — М.: Диалог-МИФИ, 1992. — 208 с.: ил.
2. Фролов А.В., Фролов Г.В. Аппаратное обеспечение IBM PC. Ч.2. — М.: Диалог-МИФИ, 1992. — 200 с.
3. Гук Михаил. Процессоры Pentium II, Pentium Pro и просто Pentium. — СПб.: Питер-Ком, 1999. — 288 с.: ил.
4. Линдберг Дж.П. Руководство Novell: Настольная книга администратора Netware 4.1: Пер. с англ. — М.: Лори, 1997. — 582 с.: ил.
5. Novell Education Course 801, Service and Support.
6. Ахметов К.С., Борзенко А.Е. Современный персональный компьютер. — М.: Компьютер-Пресс, 1995. — 317 с.: ил.