



# 计算机组成原理

## 第 18、19讲

左德承

哈尔滨工业大学计算学部  
容错与移动计算研究中心

## 2. 中断系统需解决的问题

- (1) 各中断源 如何 向 CPU 提出请求？
- (2) 各中断源 同时 提出 请求 怎么办？
- (3) CPU 什么 条件、什么 时间、以什么 方式 响应中断？
- (4) 如何 保护现场？
- (5) 如何 寻找入口地址？
- (6) 如何 恢复现场，如何 返回？
- (7) 处理中断的过程中又 出现新的中断 怎么办？

硬件 + 软件

## 二、中断请求标记和中断判优逻辑

### 1. 中断请求标记 **INTR**

一个请求源 一个 **INTR** 中断请求标记触发器

多个**INTR** 组成 中断请求标记寄存器

1	2	3	4	5			<i>n</i>
掉电	过热	主存读写校验错	阶上溢	非法除法		键盘输入	打印机输出

**INTR** 分散 在各个中断源的 接口电路中

**INTR** 集中在 **CPU** 的中断系统 内

## 2. 中断判优逻辑

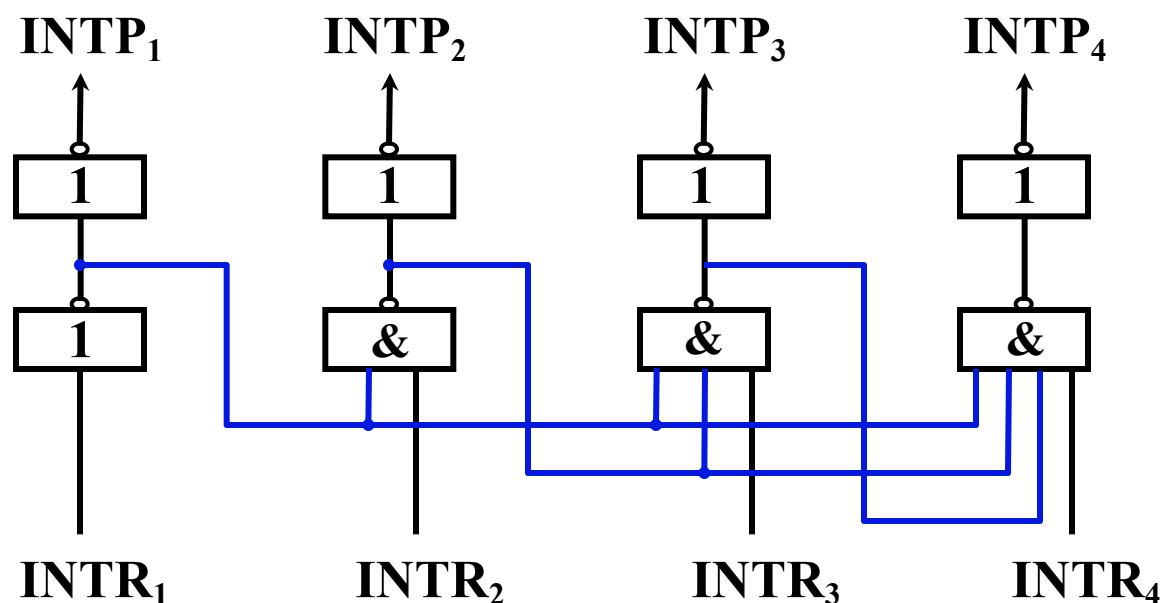
## 8.4

### (1) 硬件实现（排队器）

① 分散 在各个中断源的 接口电路中 链式排队器

参见 第五章

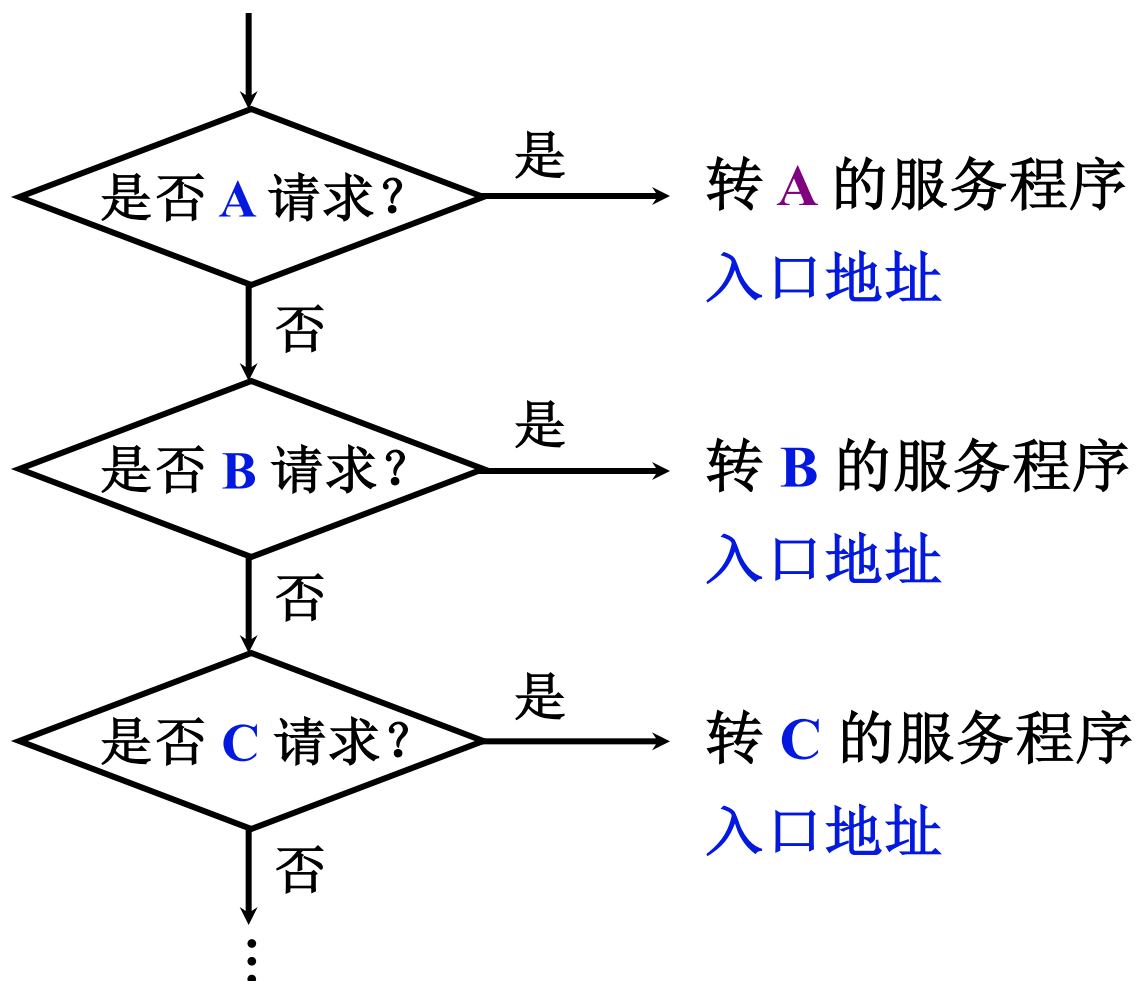
② 集中 在 CPU 内



$INTR_1$ 、 $INTR_2$ 、 $INTR_3$ 、 $INTR_4$  优先级 按 降序 排列

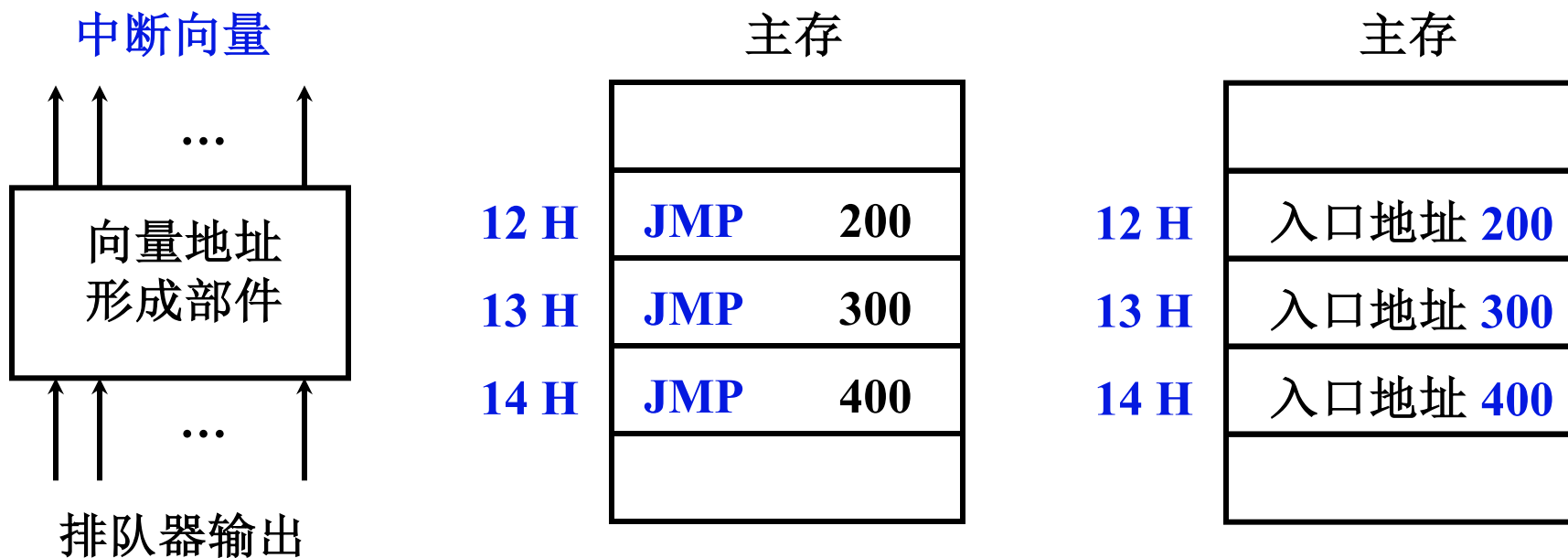
## (2) 软件实现（程序查询）

A、B、C 优先级按 降序 排列



# 三、中断服务程序入口地址的寻找 8.4

## 1. 硬件向量法



向量地址 12H、13H、14H

入口地址 200、300、400

## 2. 软件查询法

八个中断源 1, 2, ... 8 按 降序 排列

中断识别程序（入口地址 **M**）

地 址	指 令	说 明
<b>M</b>	<b>SKP</b> DZ 1 <sup>#</sup>	1 <sup>#</sup> D = 0 跳（D为完成触发器）
	<b>JMP</b> 1 <sup>#</sup> SR	1 <sup>#</sup> D = 1 转1 <sup>#</sup> 服务程序
	<b>SKP</b> DZ 2 <sup>#</sup>	2 <sup>#</sup> D = 0 跳
	<b>JMP</b> 2 <sup>#</sup> SR	2 <sup>#</sup> D = 1 转2 <sup>#</sup> 服务程序
	⋮	
	<b>SKP</b> DZ 8 <sup>#</sup>	8 <sup>#</sup> D = 0 跳
	<b>JMP</b> 8 <sup>#</sup> SR	8 <sup>#</sup> D = 1 转8 <sup>#</sup> 服务程序

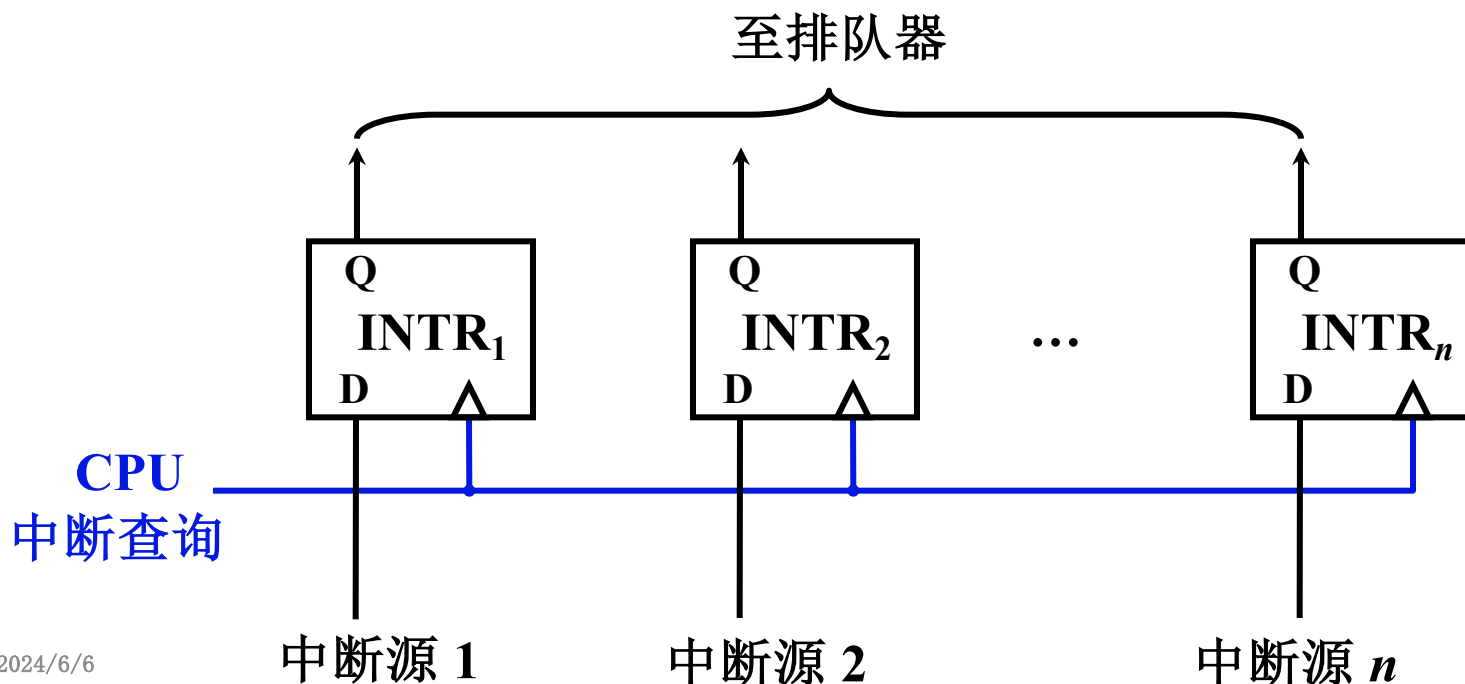
## 四、中断响应

### 1. 响应中断的条件

允许中断触发器  $EINT = 1$

### 2. 响应中断的时间

指令执行周期结束时刻由CPU发查询信号







# 五、保护现场和恢复现场

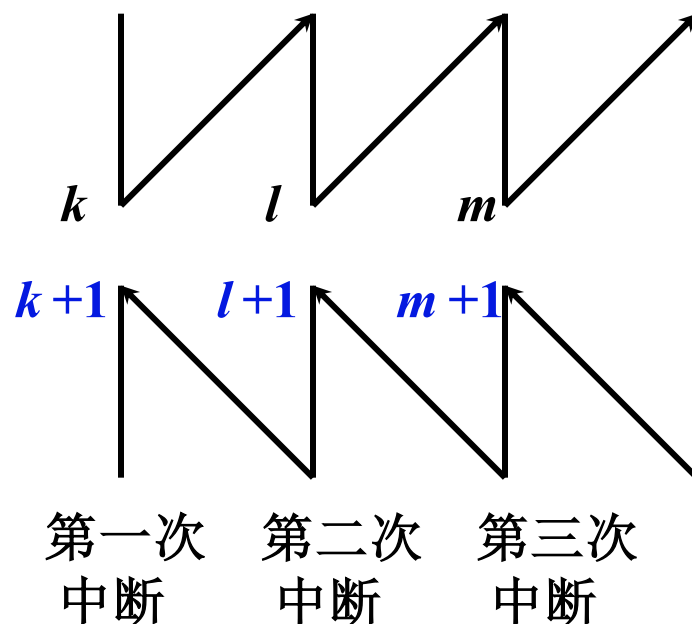
8.4

1. 保护现场 { 断点 中断隐指令 完成  
寄存器 内容 中断服务程序 完成
2. 恢复现场 中断服务程序 完成



# 六、中断屏蔽技术

## 1. 多重中断的概念



程序断点  $k+1$ ,  $l+1$ ,  $m+1$

# 划分优先级的一般规律

## 8.4

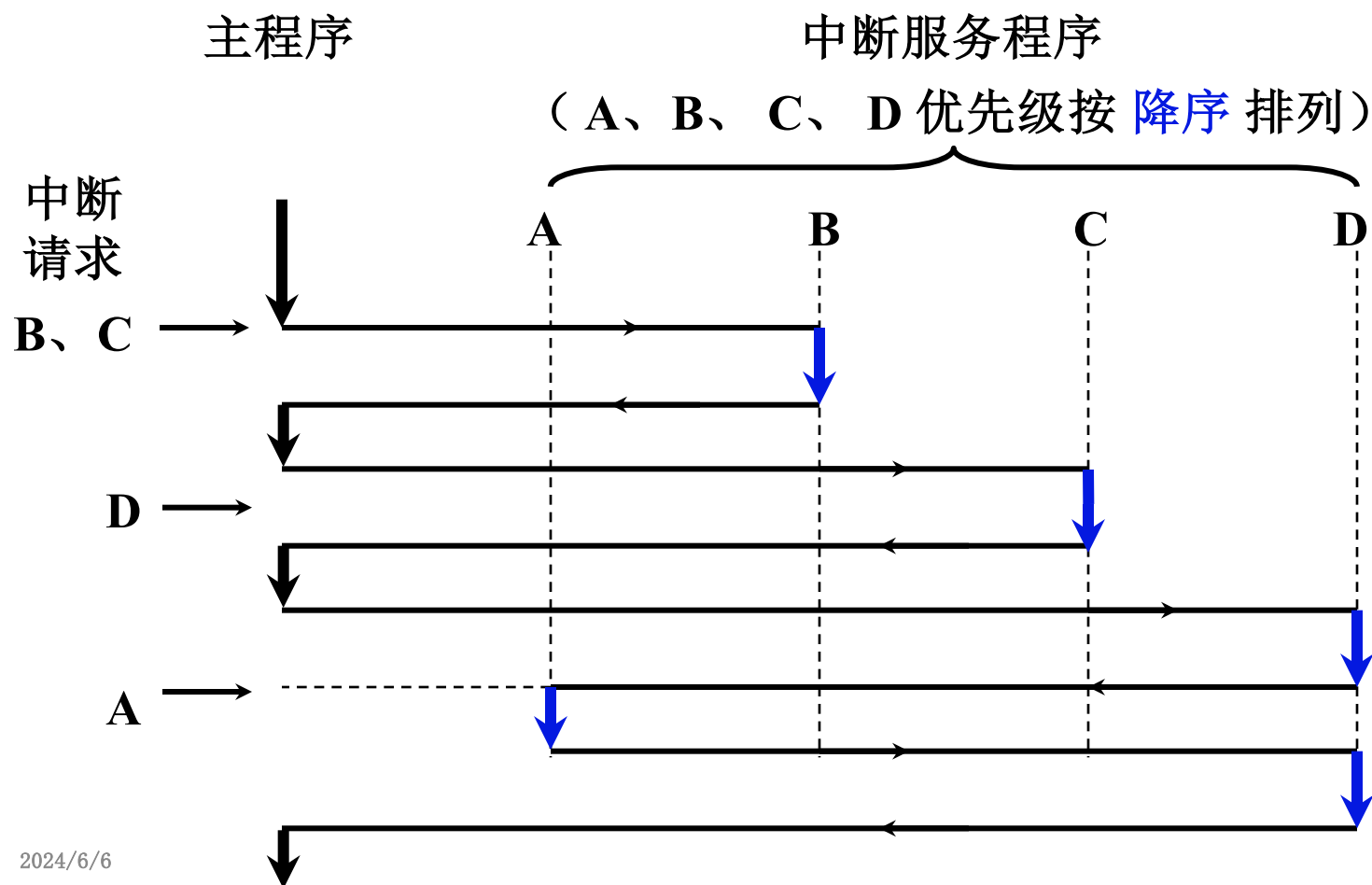
- 硬件故障中断属于最高级，其次是程序错误中断
- 非屏蔽中断优于可屏蔽中断
- DMA请求优先于I/O设备传送的中断请求
- 高速设备优于低速设备
- 输入设备的中断优于输出设备
- 实时设备优先于普通设备

## 2. 实现多重中断的条件

## 8.4

(1) 提前 设置 开中断 指令

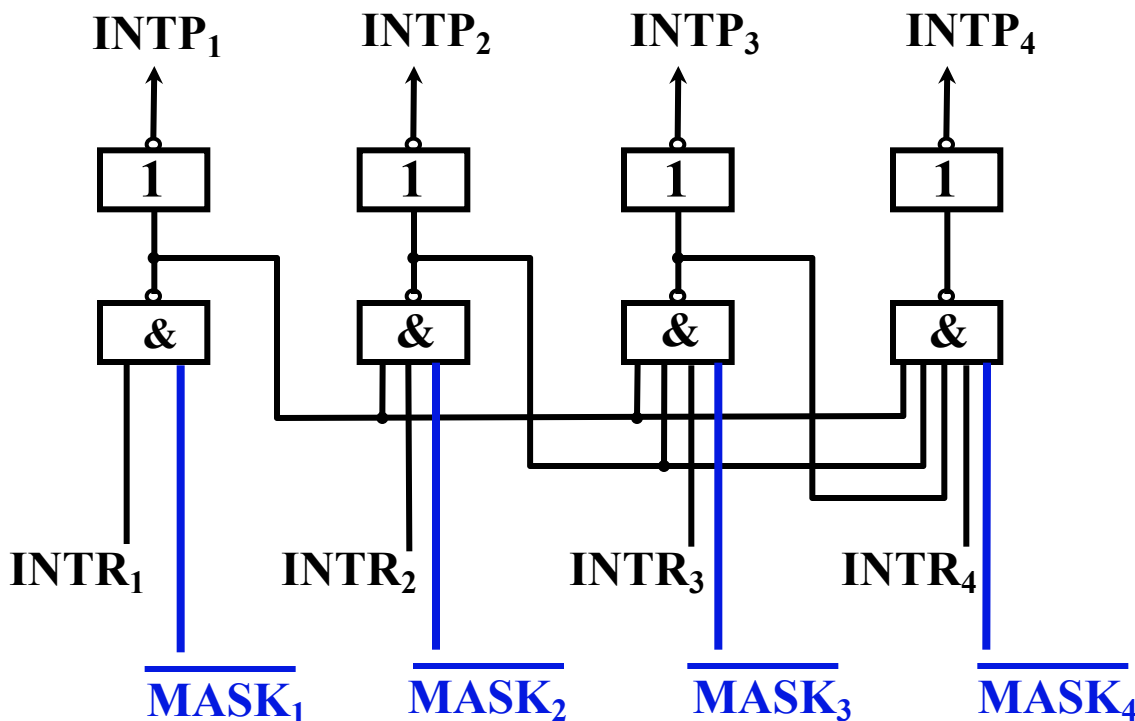
(2) 优先级别高 的中断源 有权中断优先级别低 的中断源



### 3. 屏蔽技术

## 8.4

#### (1) 屏蔽触发器的作用



$MASK = 0$  (未屏蔽)

$INTR$  能被置“1”

$MASK_i = 1$  (屏蔽)

$INTP_i = 0$  (不能被排队选中)

## (2) 屏蔽字

16个中断源 1, 2, 3, ..., 16 按 降序 排列

优先级	屏蔽字															
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1
5	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1
6	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1
⋮	⋮															
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

### (3) 屏蔽技术可改变处理优先等级

## 8.4

响应优先级      不可改变

处理优先级      可改变（通过重新设置屏蔽字）

中断源	原屏蔽字	新屏蔽字
A	1 1 1 1	1 1 1 1
B	0 1 1 1	0 1 0 0
C	0 0 1 1	0 1 1 0
D	0 0 0 1	0 1 1 1

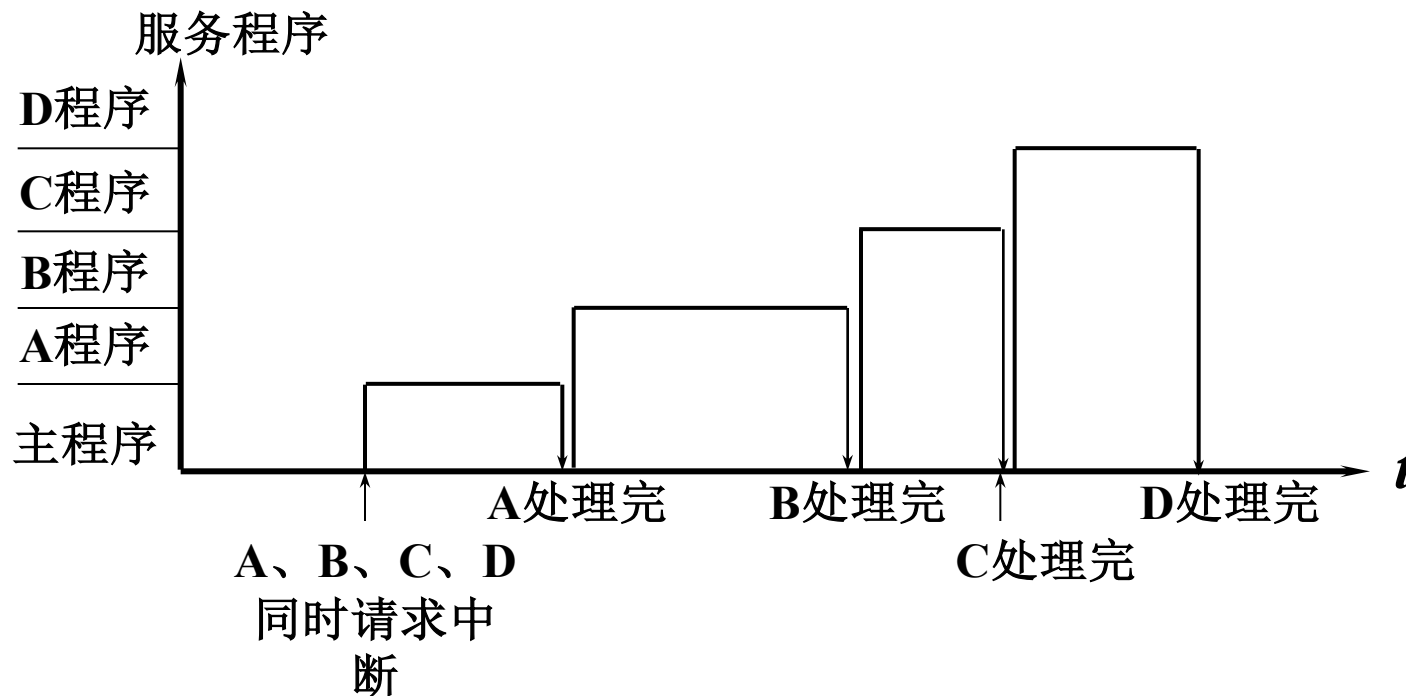
响应优先级 **A→B→C→D** 降序排列

处理优先级 **A→D→C→B** 降序排列



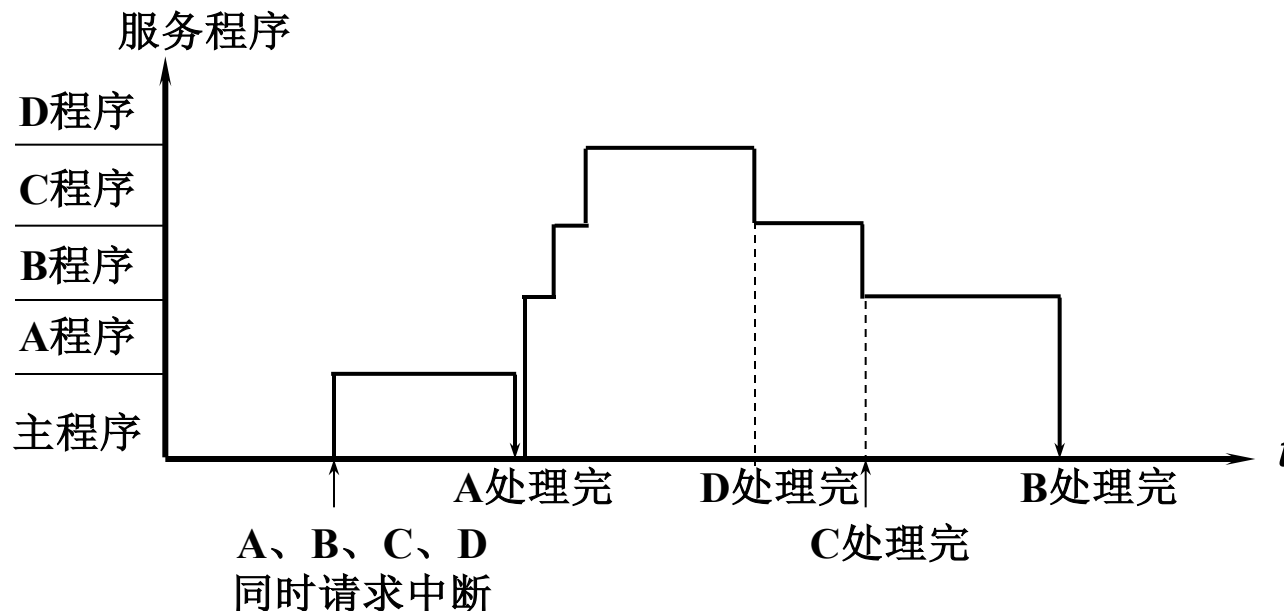
### (3) 屏蔽技术可改变处理优先等级

## 8.4



CPU 执行程序轨迹（原屏蔽字）

### (3) 屏蔽技术可改变处理优先等级

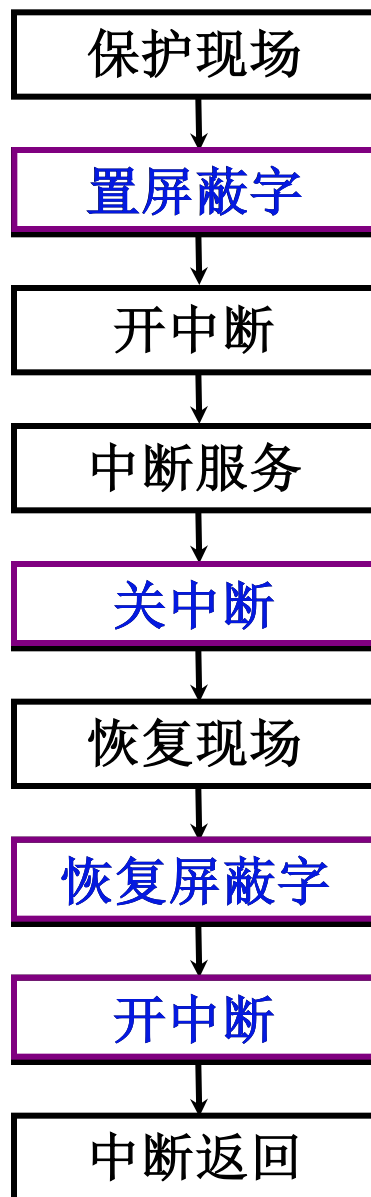


CPU 执行程序轨迹（新屏蔽字）

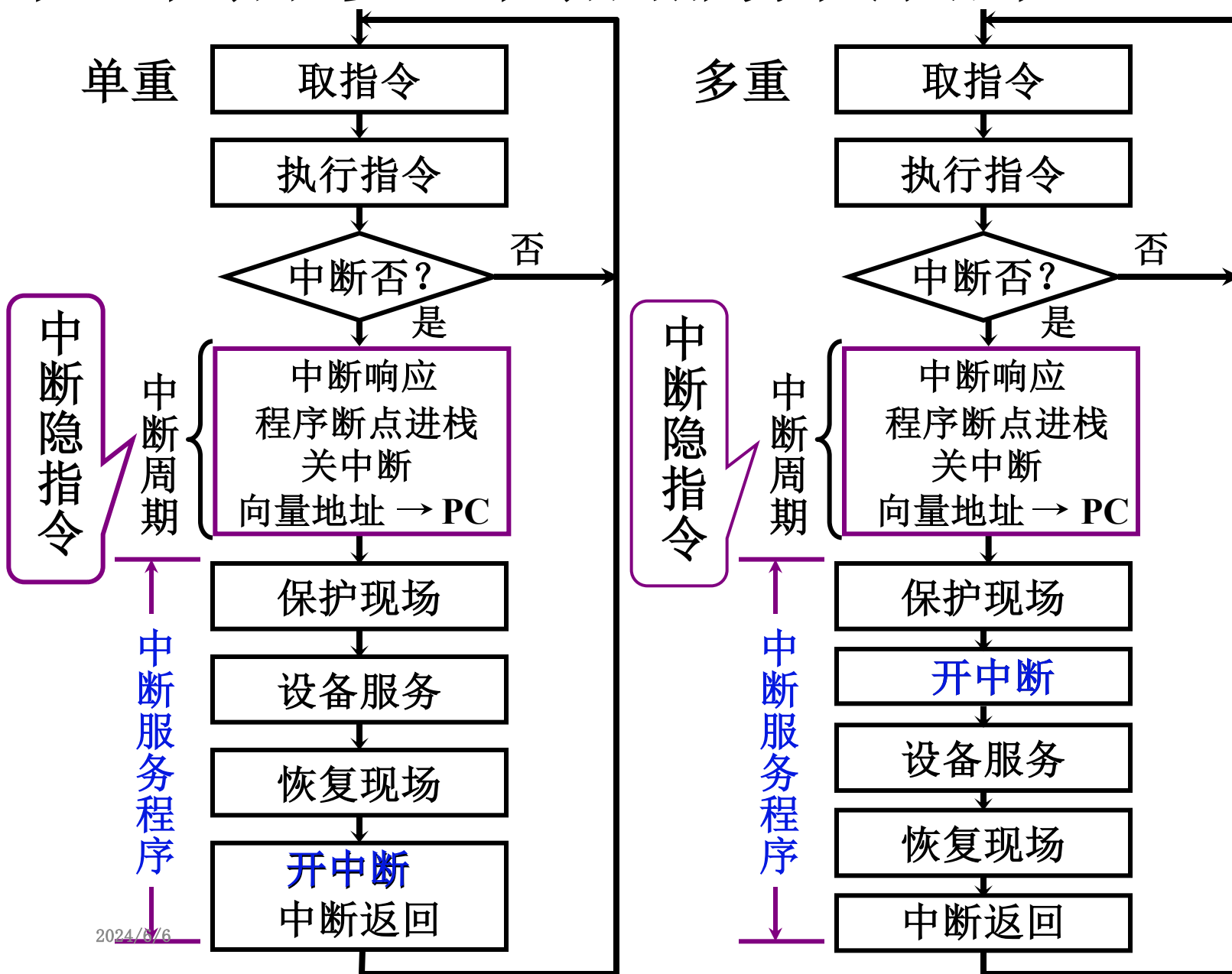
### (4) 屏蔽技术的其他作用

可以 **人为地屏蔽** 某个中断源的请求  
便于程序控制

## (5) 新屏蔽字的设置



# 单重中断和多重中断的服务程序流程



## 4. 多重中断的断点保护

(1) 断点进栈                      中断隐指令 完成

(2) 断点存入 “0” 地址      中断隐指令 完成

中断周期       $0 \rightarrow \text{MAR}$

命令存储器写

$\text{PC} \rightarrow \text{MDR}$       断点  $\rightarrow \text{MDR}$

$(\text{MDR}) \rightarrow \text{存入存储器}$

三次中断，三个断点都存入 “0” 地址

？ 如何保证断点不丢失？

### (3) 程序断点存入 “ 0 ” 地址的断点保护8.4

地 址	内 容	说 明
0	××××	存程序断点
5	JMP SERVE	5 为向量地址
SERVE	STA SAVE	保护现场
	⋮	
置屏蔽字	LDA 0	} 0 地址内容转存
	STA RETURN	
	ENI	开中断
	⋮	} 其他服务内容
	LDA SAVE	
	JMP @ RETURN	恢复现场
SAVE	××××	间址返回
RETURN	××××	存放 ACC 内容
		转存 0 地址内容

# 第 3 章 系统总线

## 3.1 总线的基本概念

## 3.2 总线的分类

## 3.3 总线特性及性能指标

## 3.4 总线结构

## 3.5 总线控制

# 3.1 总线的基本概念

一、为什么要用总线

二、什么是总线

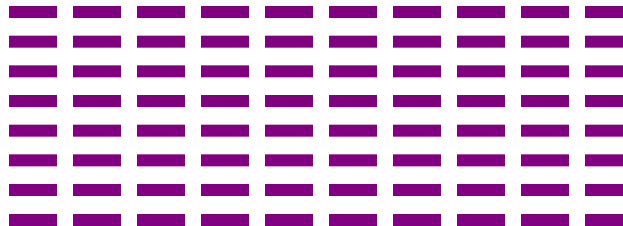
总线是连接各个部件的信息传输线，  
是 各个部件共享的传输介质

三、总线上信息的传送

串行



并行



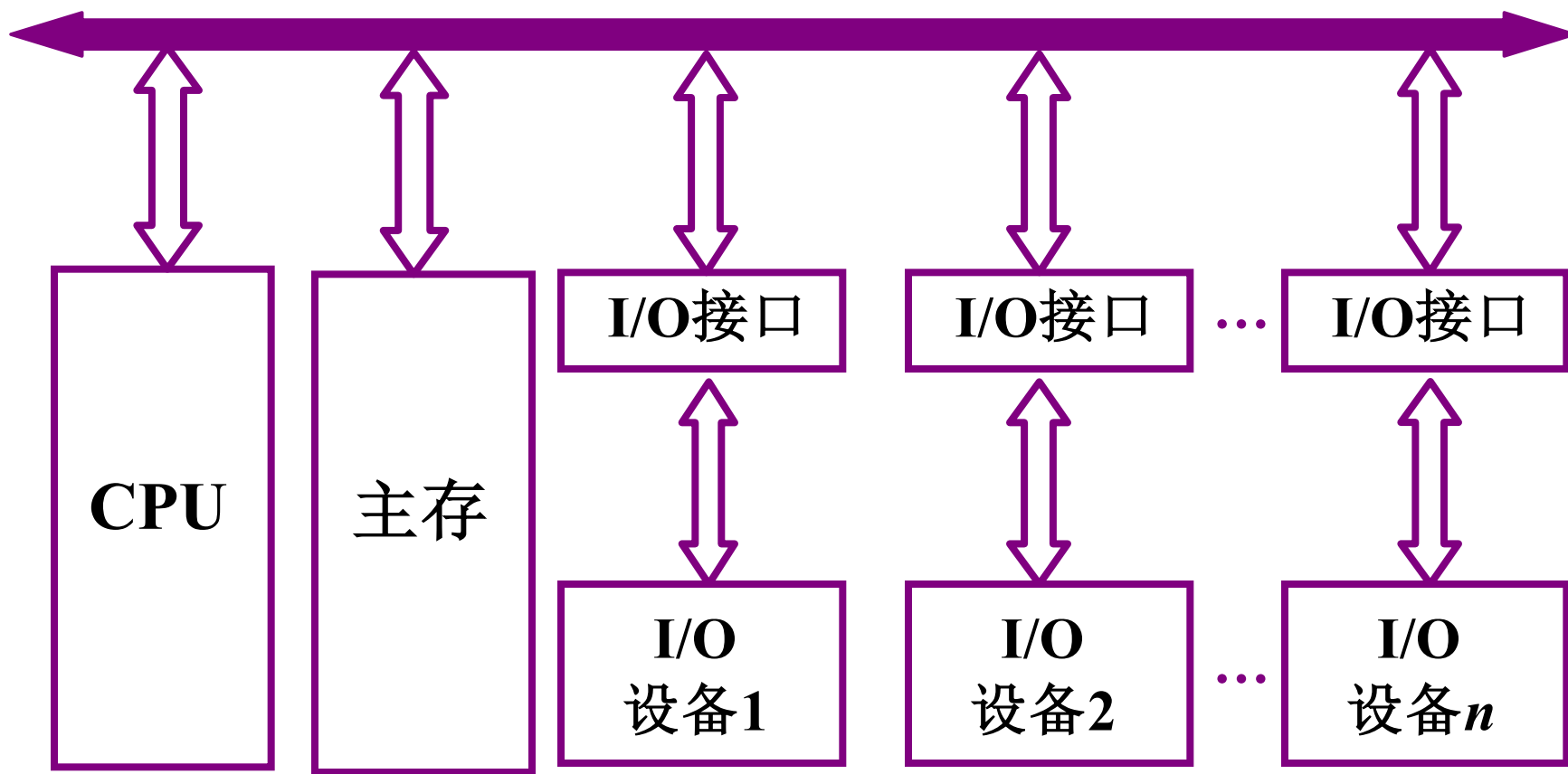


# 四、总线结构的计算机举例

## 3.1

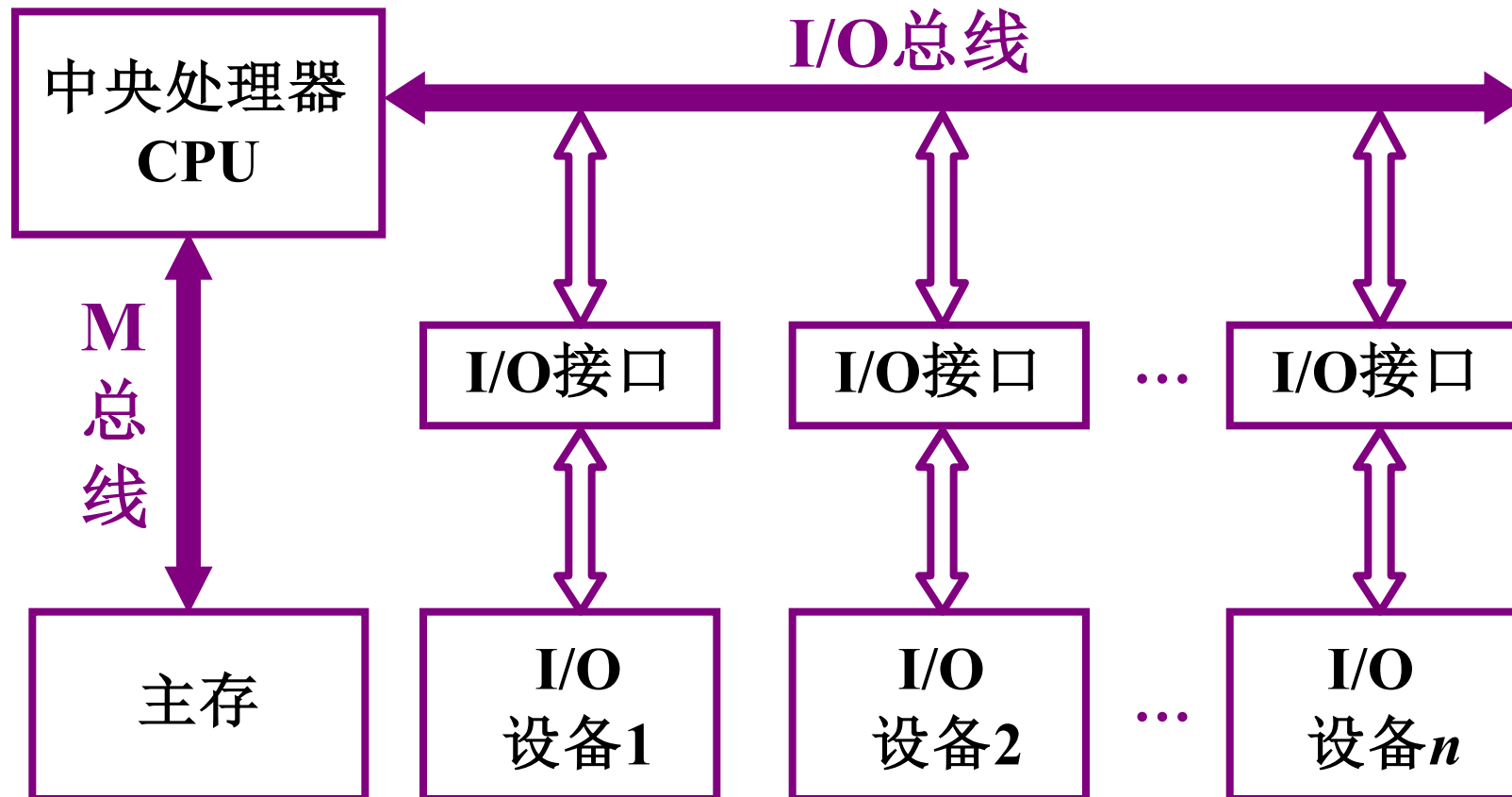
### 1.单总线结构框图

单总线（系统总线）

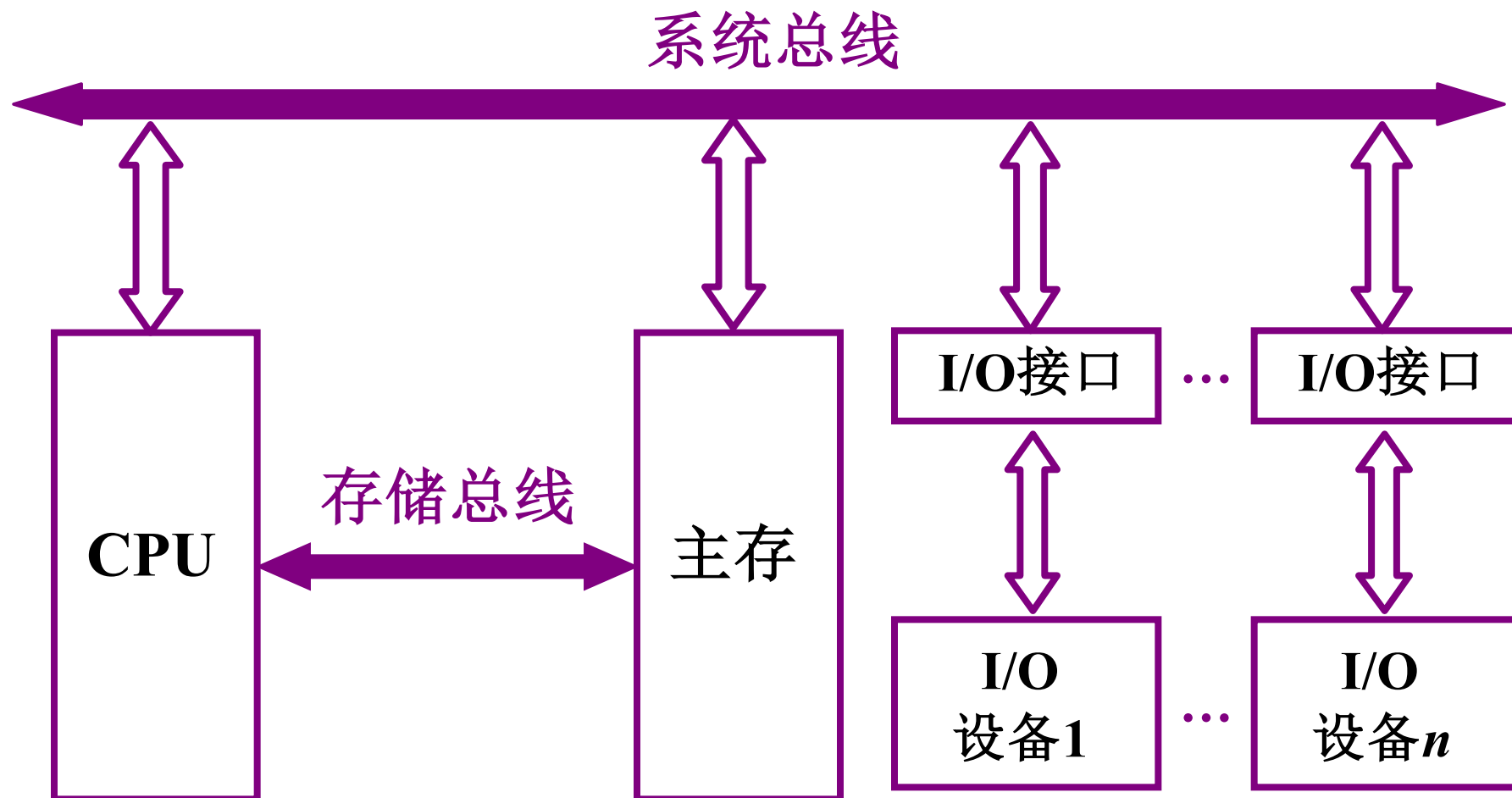


## 2.1 面向 CPU 的双总线结构框图

3.1

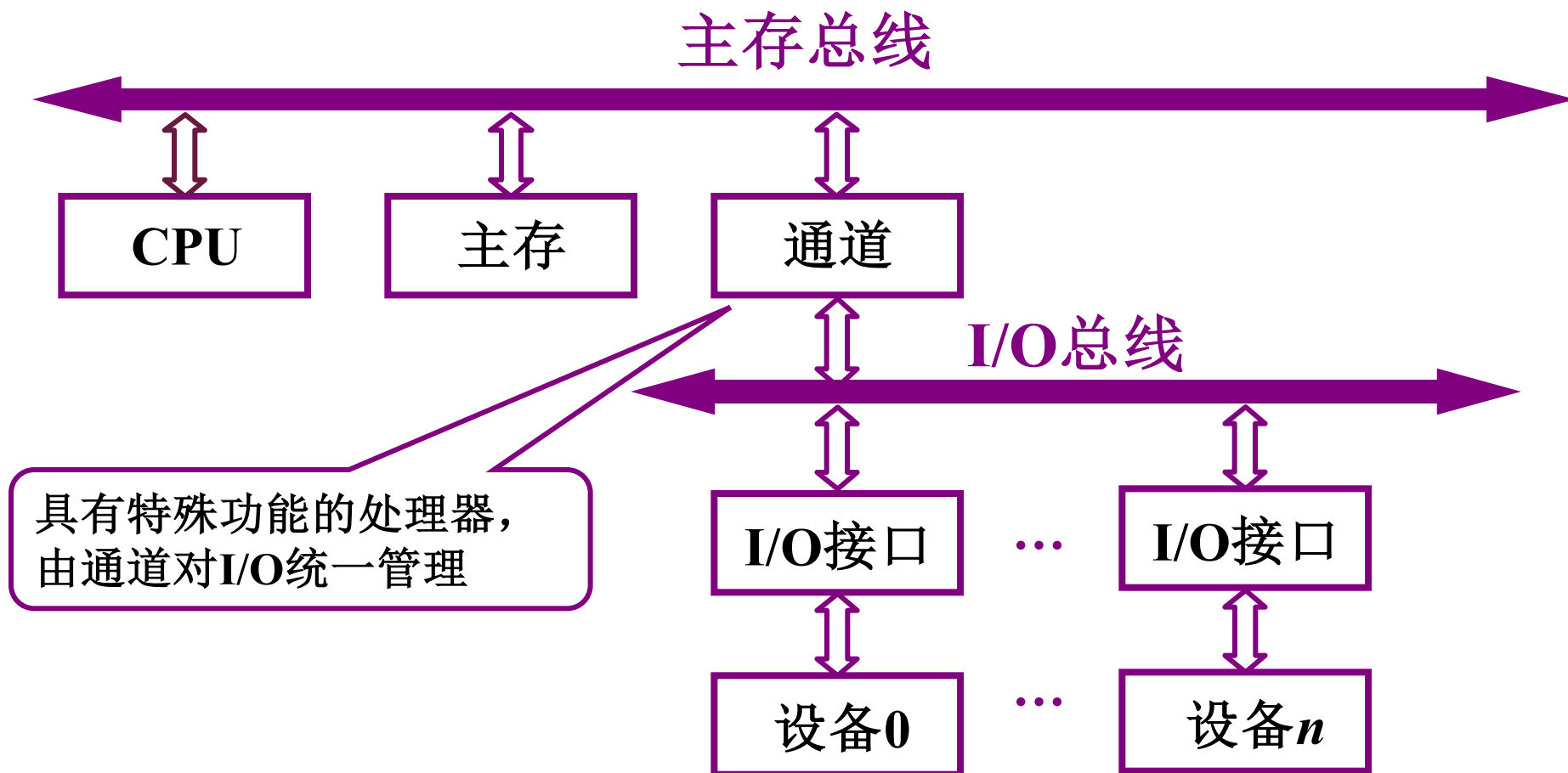


## 2.2 以存储器为中心的双总线结构框图

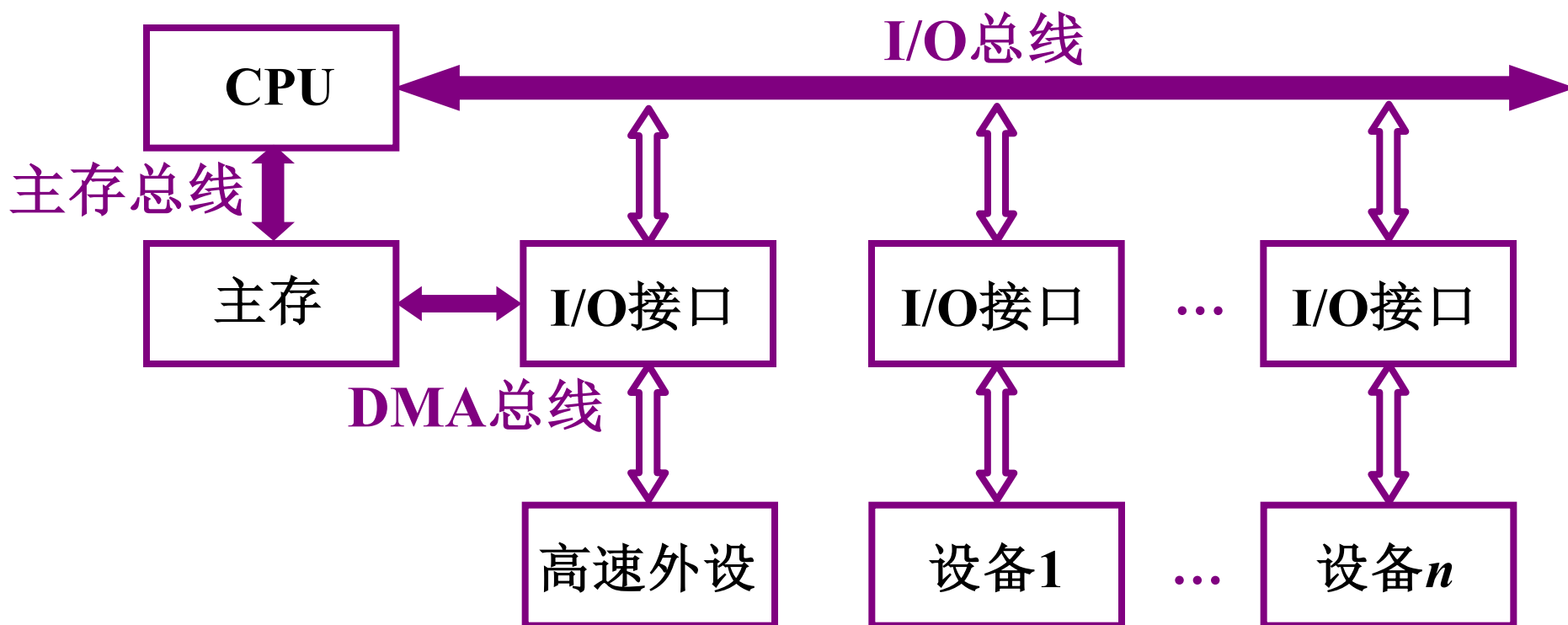


## 2.3 具有通道的双总线结构

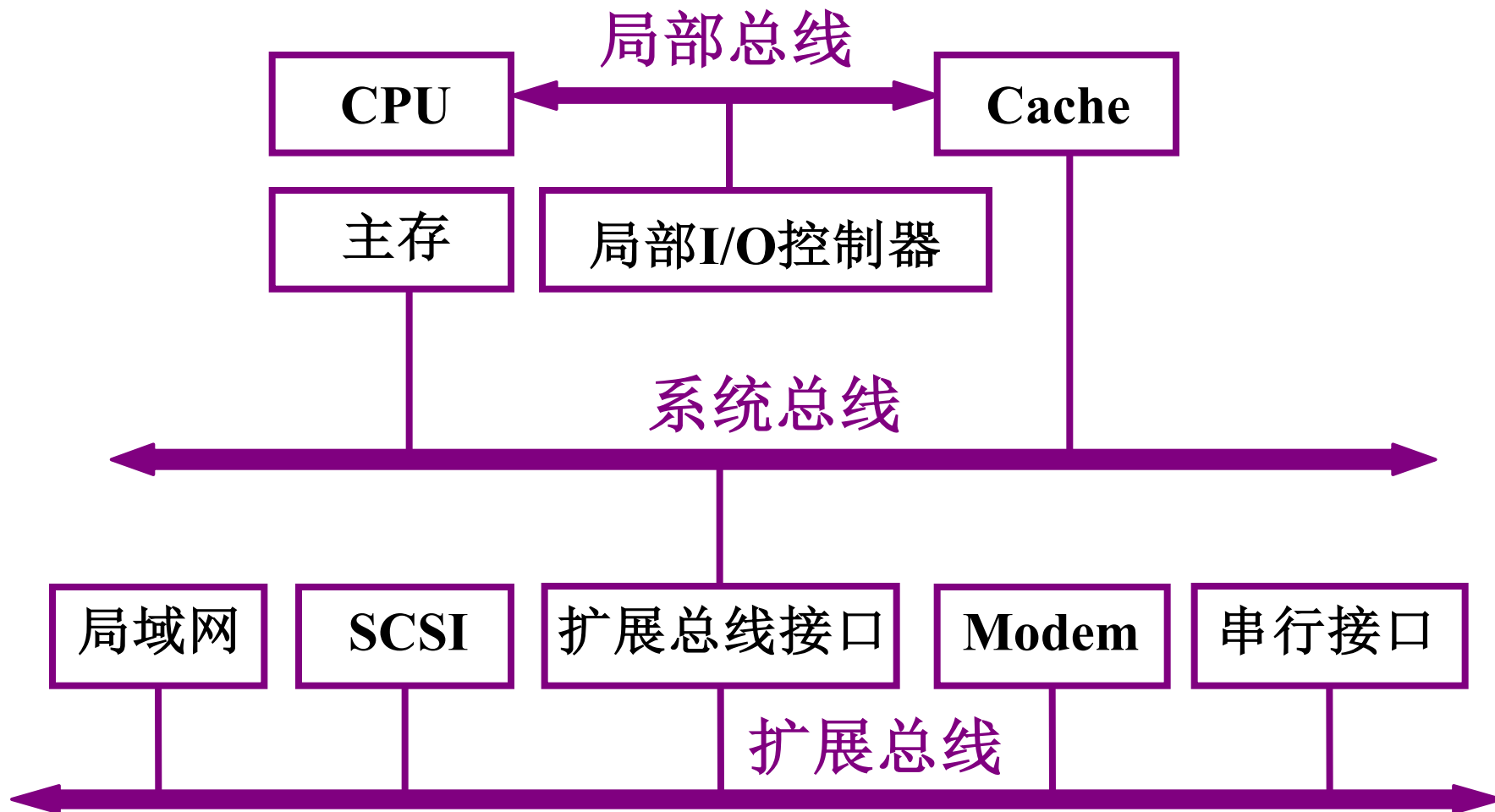
### 3.1



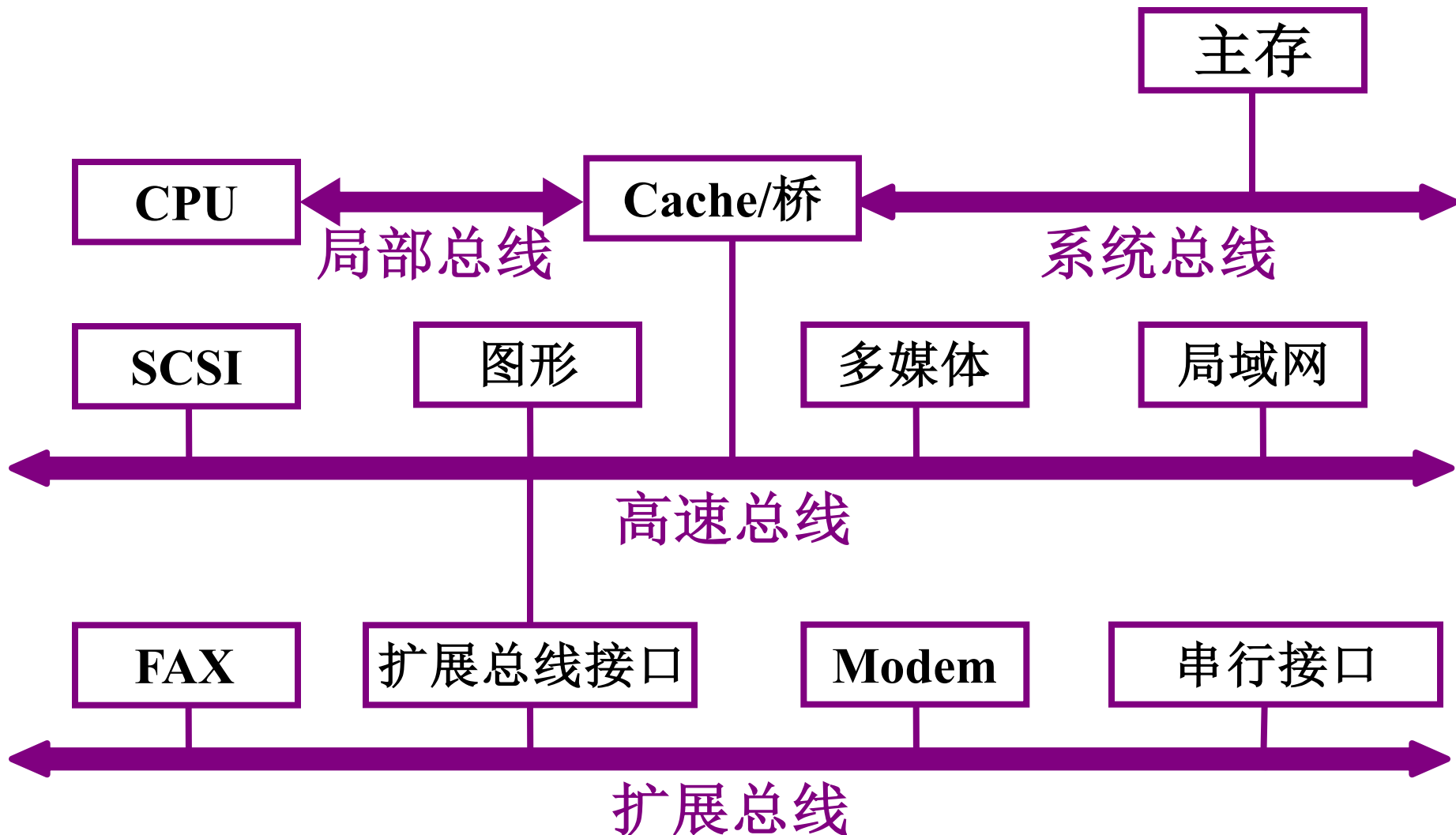
## 3.1 三总线结构



## 3.2 三总线结构的又一形式



## 4. 四总线结构



## 3.2 总线的分类

1. 片内总线      芯片内部 的总线

2. 系统总线      计算机各部件之间 的信息传输线

数据总线      双向 与机器字长、存储字长有关

地址总线      单向 与存储地址、I/O地址有关

控制总线      有出 有入

中断请求、总线请求

存储器读、存储器写  
总线允许、中断确认



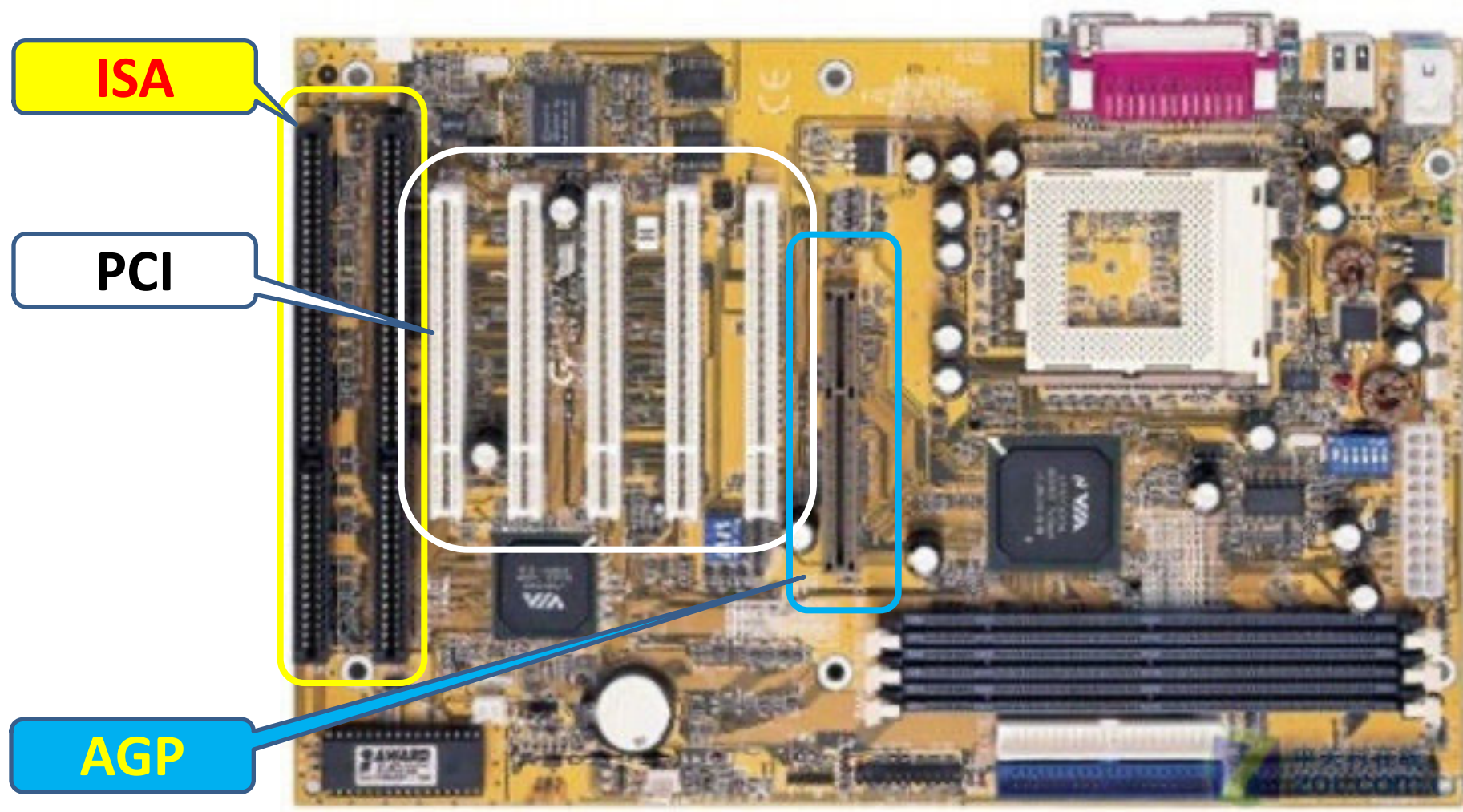
## 3. 通信总线

用于 计算机系统之间 或 计算机系统  
与其他系统（如控制仪表、移动通信等）  
之间的通信

传输方式 { 串行通信总线  
并行通信总线

## 4. 常见总线接口举例

3.2





## 4. 常见总线接口举例

3.2

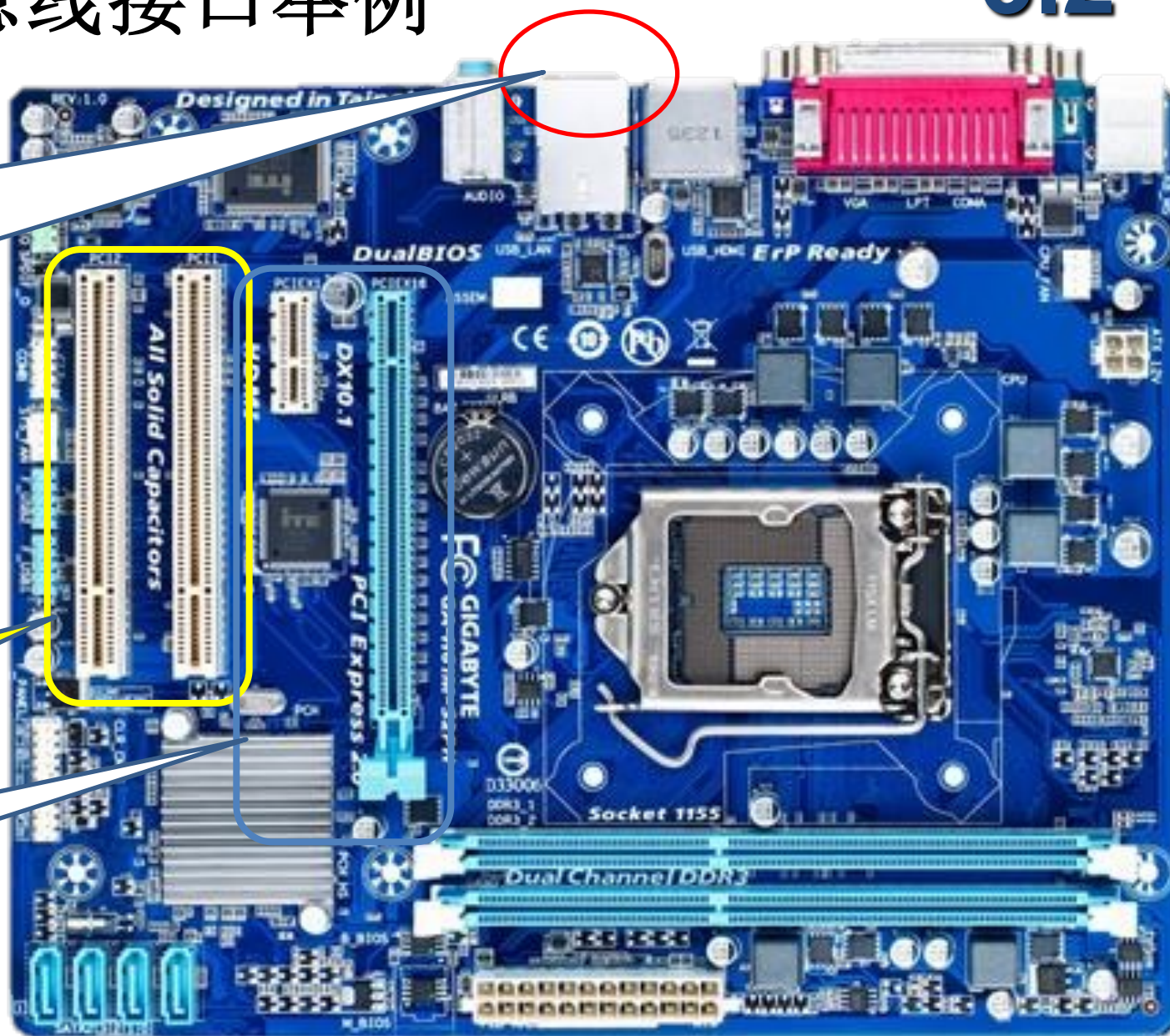


**USB\***

\* USB应归为  
通信总线

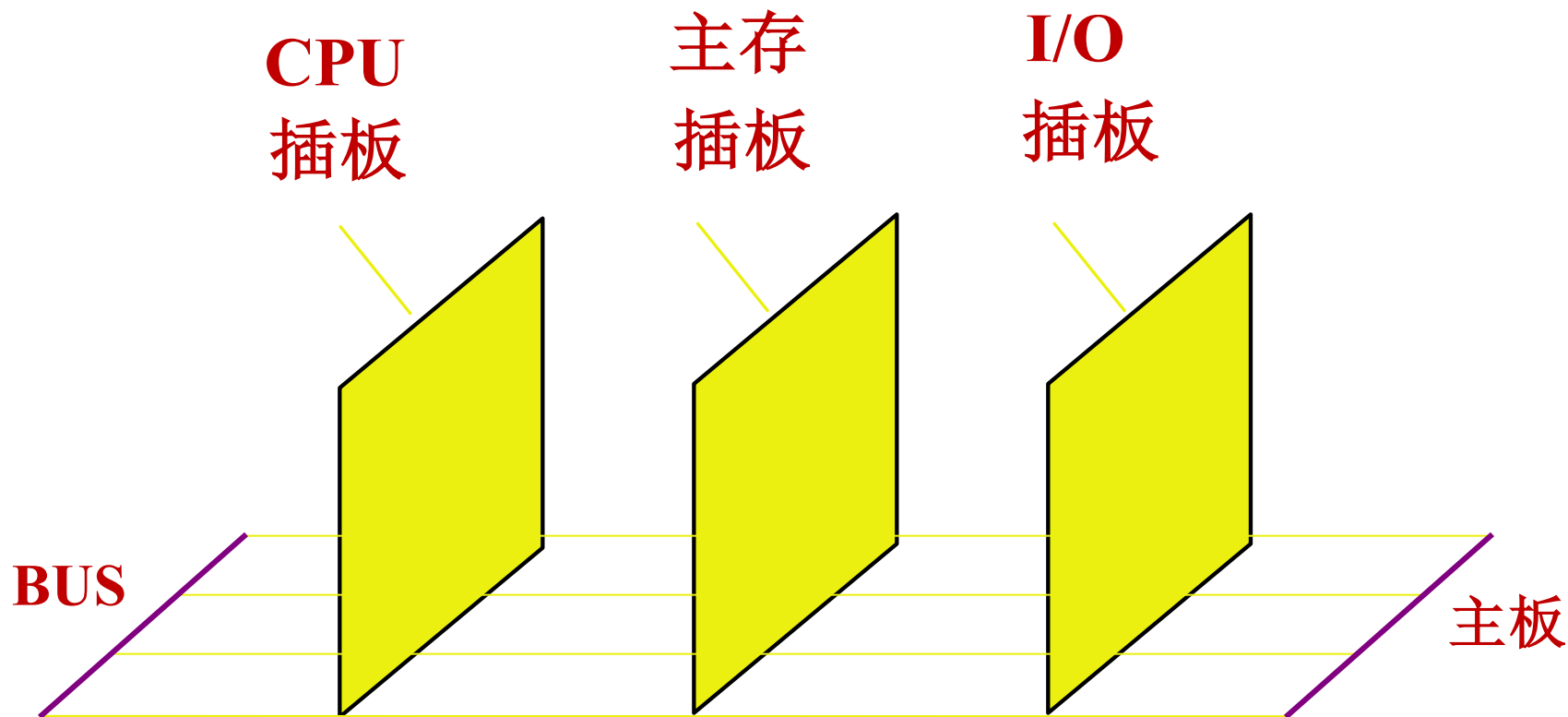
**PCI**

**PCI-Express  
(PCI-E)**



# 3.3 总线特性及性能指标

## 一、总线物理实现



## 二、总线特性

- |         |                            |
|---------|----------------------------|
| 1. 机械特性 | 尺寸、形状、管脚数 及 排列顺序           |
| 2. 电气特性 | 传输方向 和有效的 电平 范围            |
| 3. 功能特性 | 每根传输线的 功能 { 地址<br>数据<br>控制 |
| 4. 时间特性 | 信号的 时序 关系                  |

## 三、总线的性能指标

1. 总线宽度      数据线的根数
2. 标准传输率      每秒传输的最大字节数 (MBps)
3. 时钟同步/异步      同步、不同步
4. 总线复用      地址线 与 数据线 复用
5. 信号线数      地址线、数据线和控制线的 总和
6. 总线控制方式      突发、自动、仲裁、逻辑、计数
7. 其他指标      负载能力

# 总线的性能指标

## 3.3

- 总线工作频率

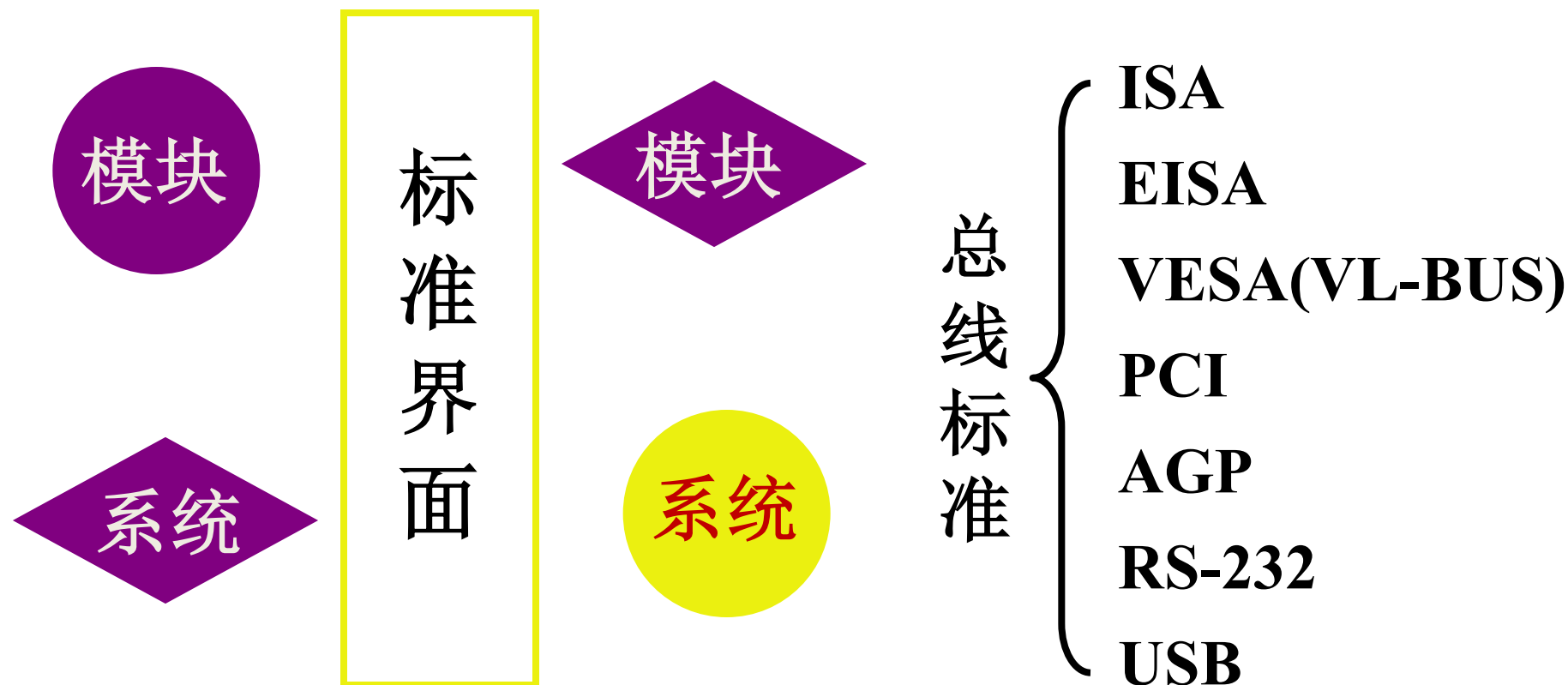
- 早期的总线通常一个时钟周期传送一次数据，此时，工作频率等于总线时钟频率；
- 现在有些总线一个时钟周期可以传送2次或4次数据，因此，工作频率是时钟频率的2倍或4倍。

- 总线带宽

- 总线的最大数据传输率
- 对于同步总线，总线带宽计算公式： $B=W \times F/N$   
W-总线宽度；F-总线时钟频率；  
N-完成一次数据传送所用时钟周期数。

F/N实际上就是总线工作频率

## 四、总线标准





# 系统总线的进化历程

## 3.3

年份	标准	时钟频率	传输位宽	传输位数/周期	带宽
1981	ISA	8MHz	8	0.5	8MB/s
1988	EISA	8.33MHz	32	1	33.33MB/s
1993	PCI	33MHz	32	1	133MB/s
1996	AGP 1x	66MHz	32	1	266MB/s
	AGP 2x/4x/8x	66MHz	32	2/4/8 (DDR/QDR/ODR)	(533/1,066/2,133)MB/s
1998	PCI-X	133MHz	64	1	1,066MB/s
2004	PCI-E Gen1 x1	2.5GHz	1	1	250MB/s
	PCI-E Gen1 x2/4/8/16	2.5GHz	1	1	(0.5/1/2/4)GB/s
2007	PCI-E Gen2 x1	5GHz	1	1	500MB/s
	PCI-E Gen2 x2/4/8/16	5GHz	1	1	(1/2/4/8)GB/s
2010	PCI-E Gen3 x1	8GHz	1	1	1GB/s
	PCI-E Gen3 x2/4/8/16	8GHz	1	1	(2/4/8/16)GB/s
2017	PCI-E Gen4 x1	16GHz	1	1	2GB/s
	PCI-E Gen4 x2/4/8/16	16GHz	1	1	(4/8/16/32)GB/s
2019	PCI-E Gen5 x1	32GHz	1	1	4GB/s
	PCI-E Gen5 x2/4/8/16	32GHz	1	1	(8/16/32/64)GB/s

## 四、常见的总线标准

3.3

### 1. ISA（Industry Standard Architecture）总线

1981

支持**24**位地址线

支持**8**位（PC）/**16**位（PC/AT）数据线

总线时钟频率**8MHz**

用于**80286**计算机

### 2. EISA（Extended ISA）总线

1988

支持**32**位地址线和数据线

总线时钟频率**8MHz**

**DMA**方式下可达**33MB/s**传输速率

用于**80386/80486**计算机

## 四、常见的总线标准

## 3.3

### 3. VL-BUS (VESA Local Bus) 总线

1992

与EISA兼容、外加主存总线以实现高速传输  
总线时钟频率与CPU外频同步 (25-40MHz)  
主存总线驱动能力有限, 最多接3个扩展卡  
用于80486计算机

### 4. PCI (Peripheral Component Interconnect) 总线

1993

总线时钟频率33/66MHz, 并与CPU独立  
总线宽度32位/64位

数据传输率132MB/s起, 所有设备共享  
即插即用 (自动分配地址空间、中断号等)

1996

变种: AGP (Accelerated Graphics Port)

## 四、常见的总线标准

### 5. PCI-X (PCI eXtended) 总线

1998

与PCI总线物理兼容

总线时钟频率支持66/100/133MHz

支持DDR和QDR技术, 最高传输率533MB/s

多用于服务器和高端PC用户

### 6. PCI-E (PCI Express) 总线

2004

高速串行总线

可包含多个数据通道x1/x2/x4/x8/x16

数据传输率250MB/s起

软件与传统PCI兼容

2011

变种: NVMe (用于高速SSD, 提高传输率)

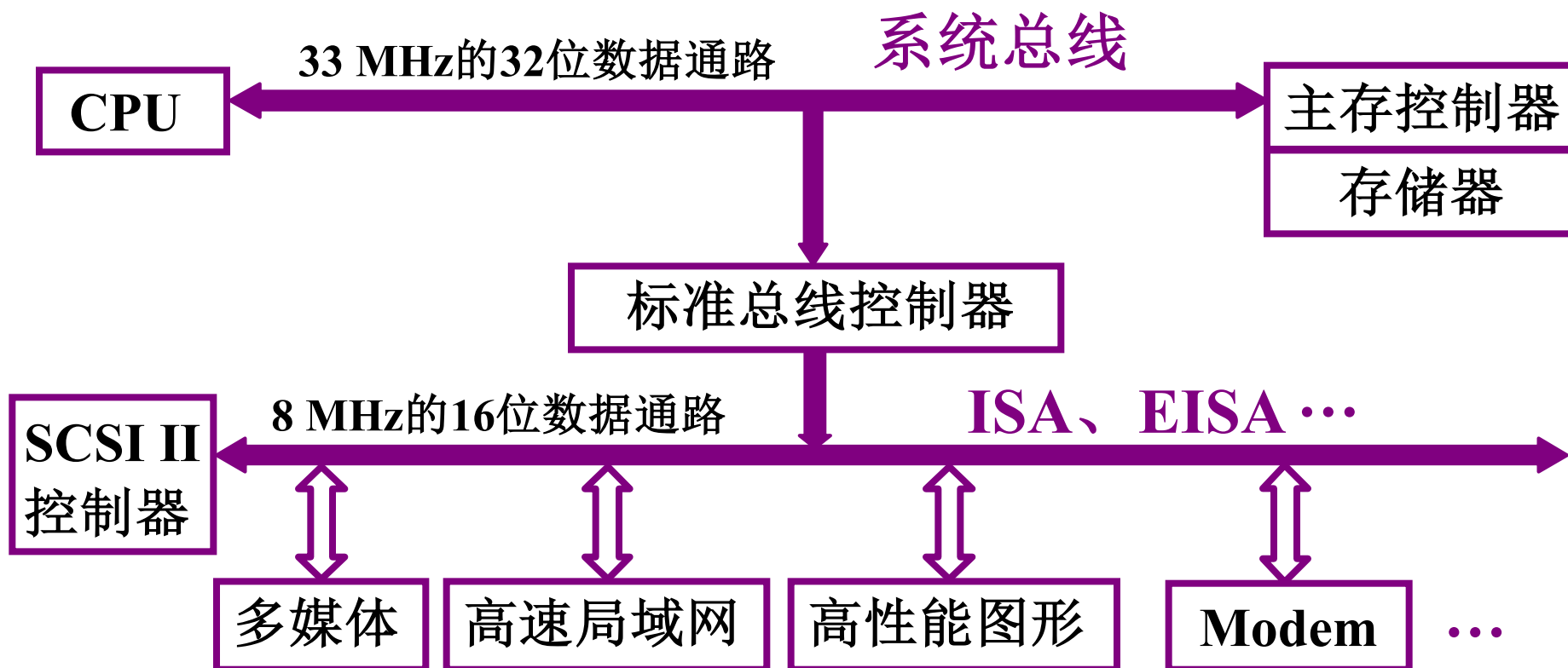
# 3.4 总线结构

## 3.4

### 一、总线结构举例

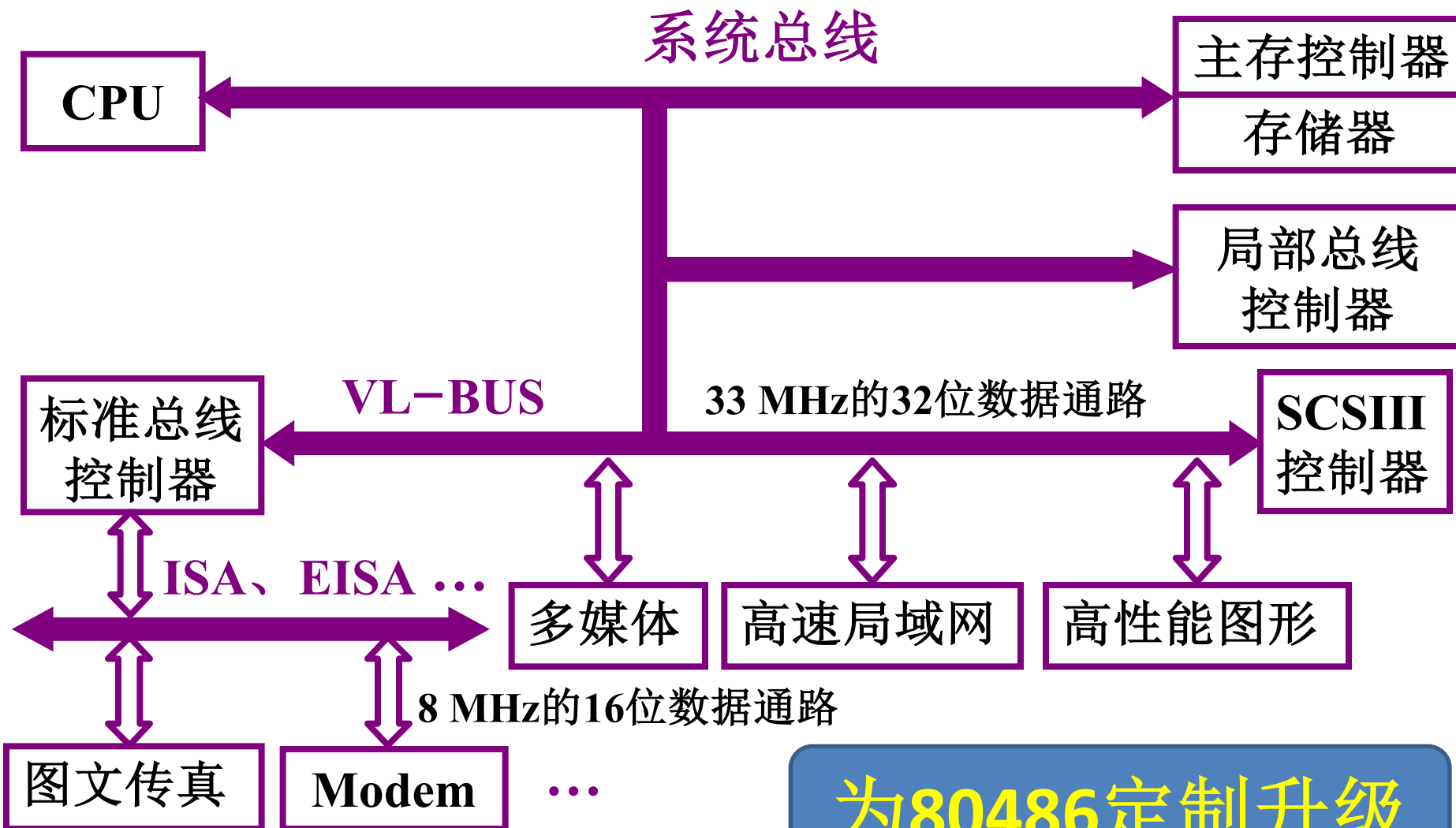
#### 1. 传统微型机总线结构

源于80286



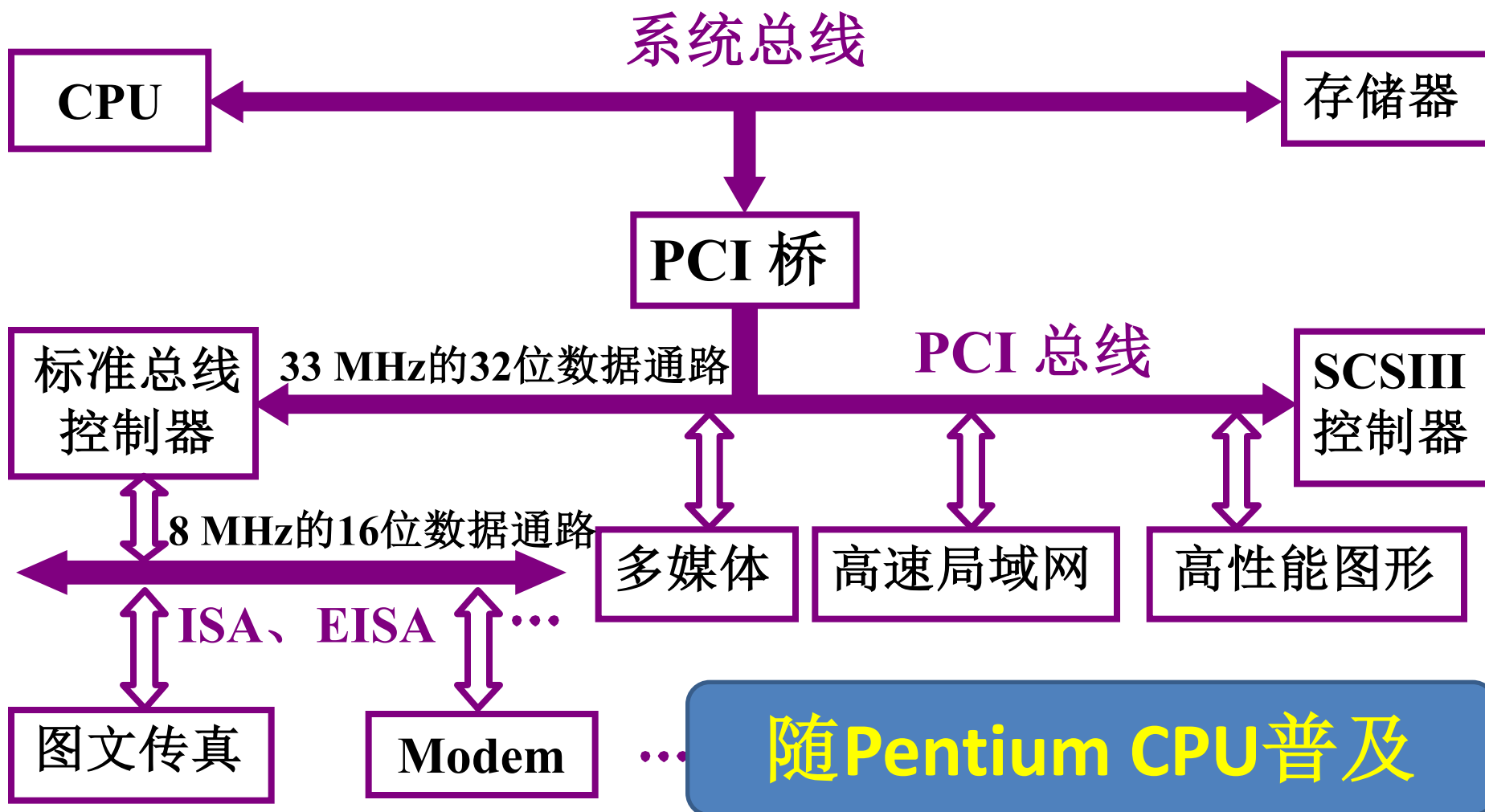
## 2. VL-BUS局部总线结构

## 3.4



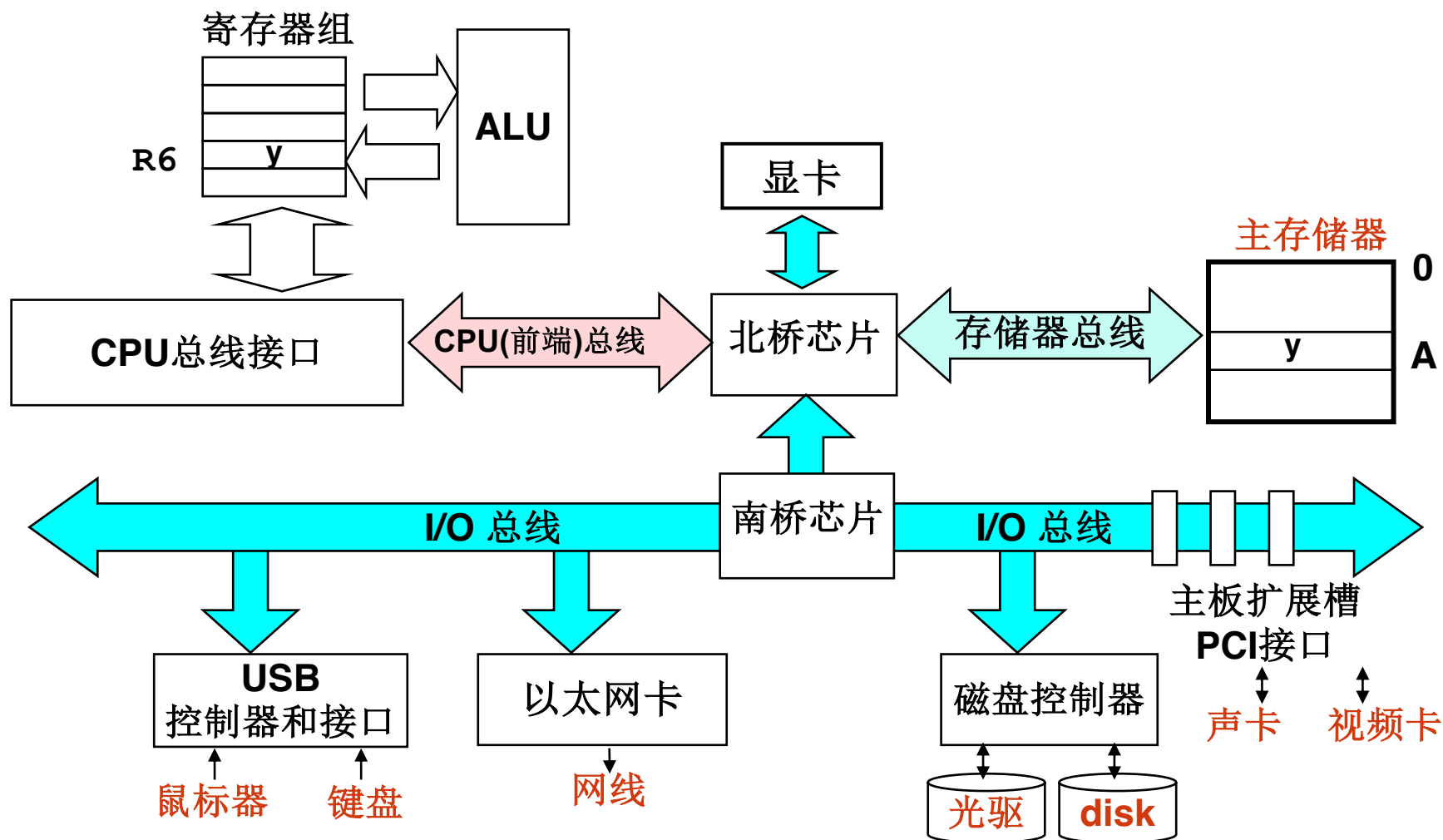
### 3. PCI 总线结构

## 3.4

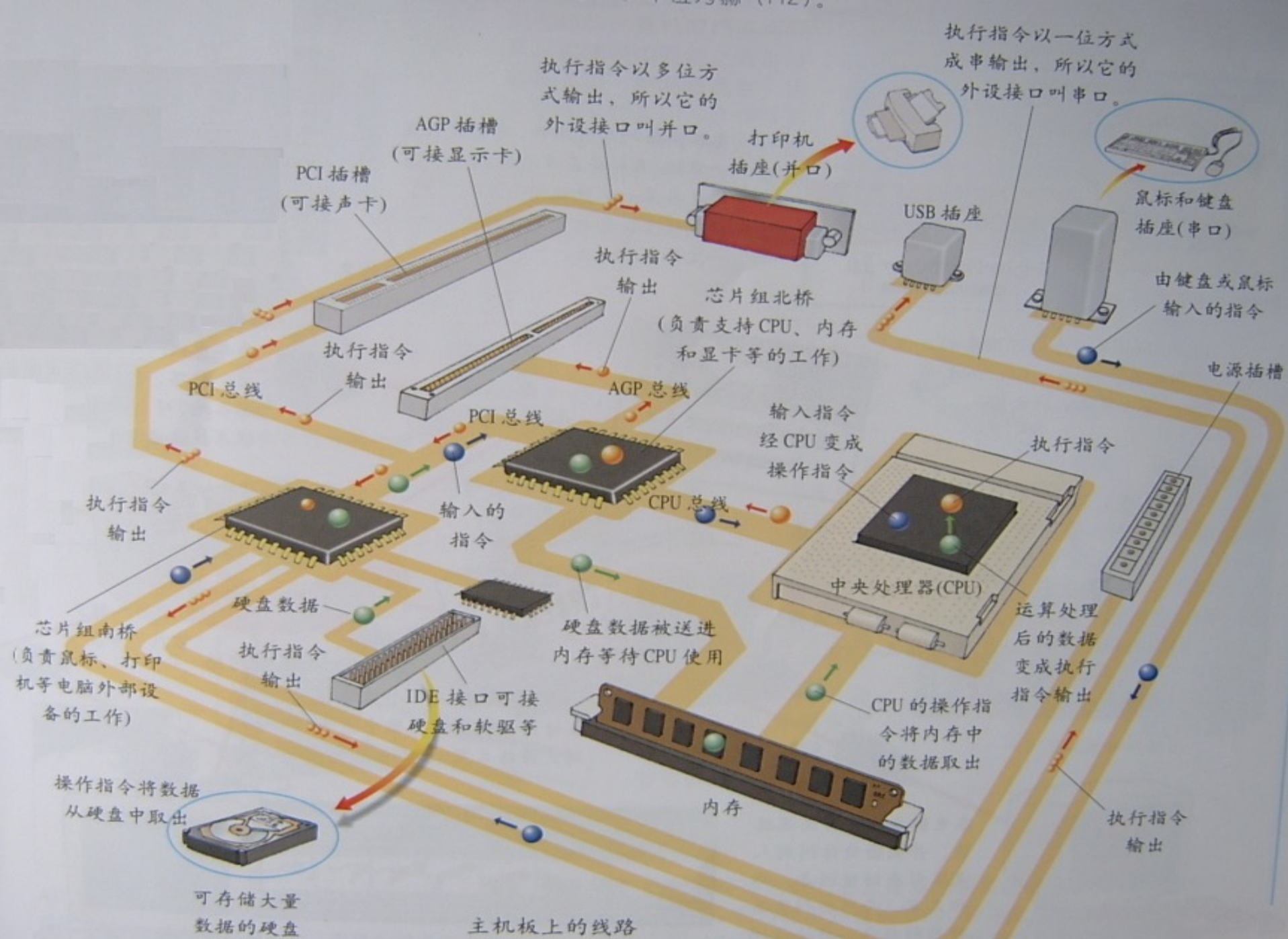


## 二、现代总线结构

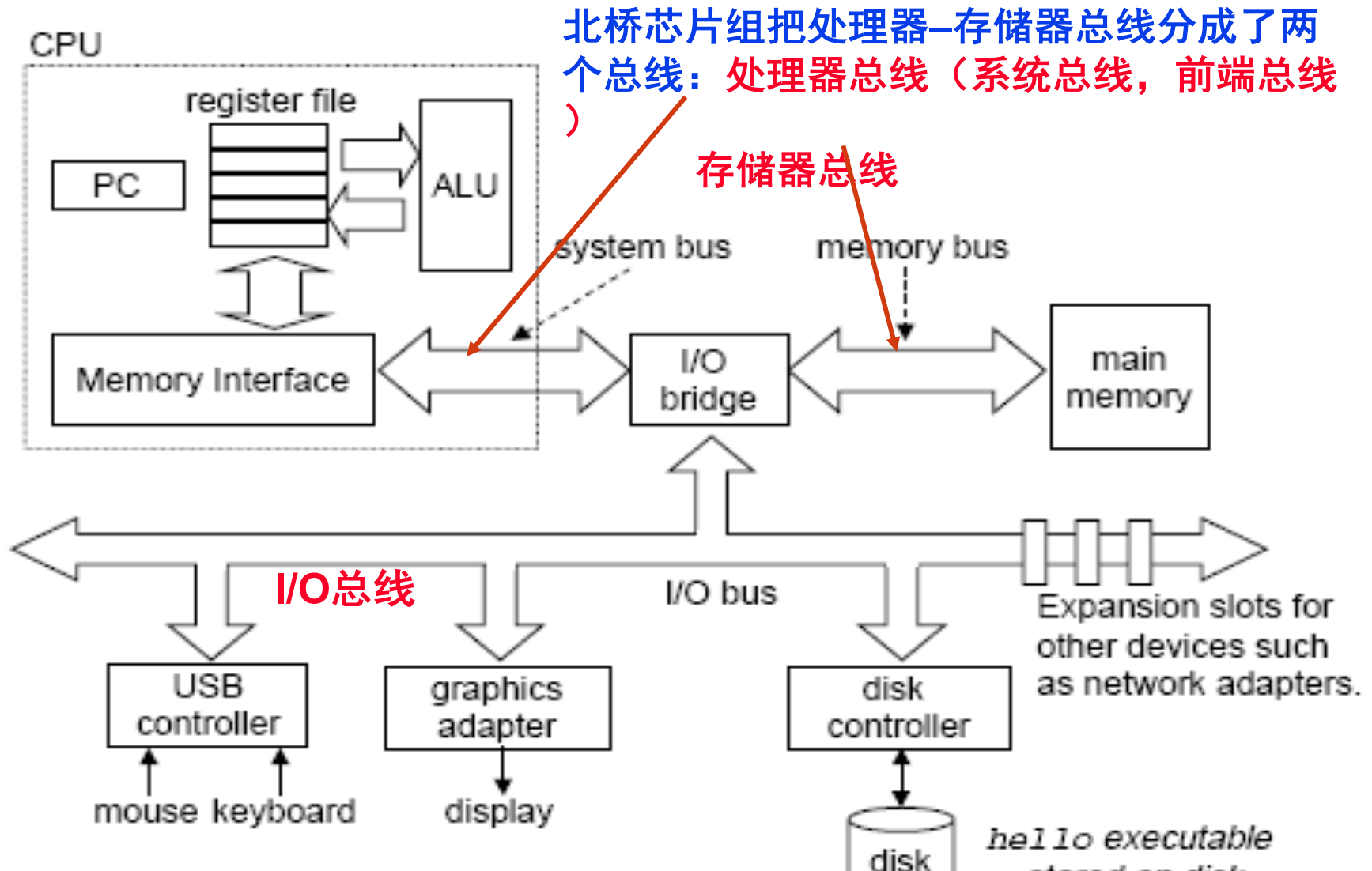
## 3.4







# Intel 体系结构中特指的“系统总线” 3.4



系统总线上传输的信息有哪些？

数据（指令、操作数、中断号）、地址、其他控制/状态/定时等信号！

# 处理器总线

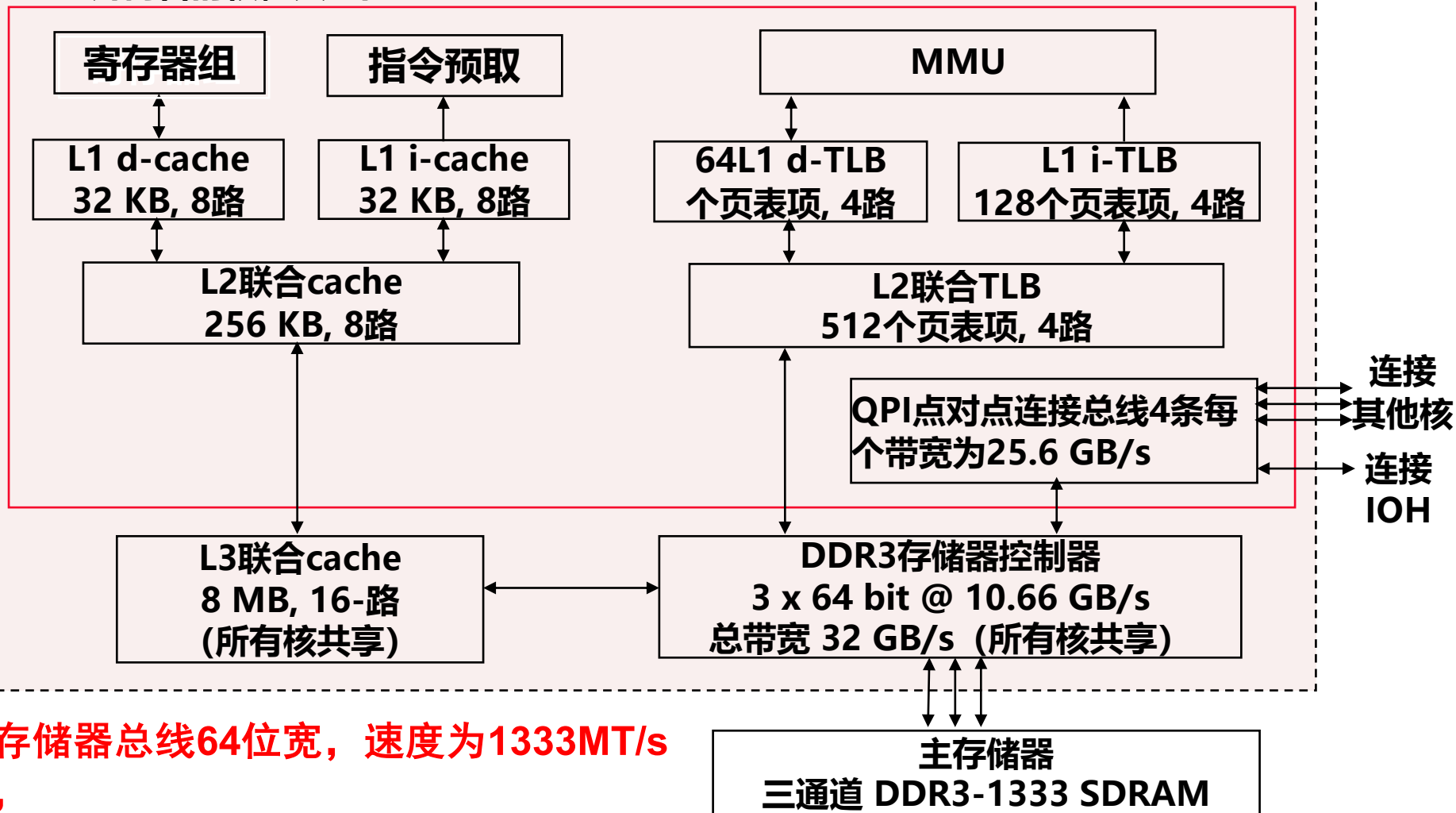
## 3.4

- 前端总线（Front Side Bus, FSB）
  - 并行传输、同步定时方式
  - 早期Intel架构使用，位于CPU芯片与北桥芯片之间互连
  - 从Pentium Pro开始，FSB采用quad pumped技术：每个总线时钟周期传送4次数据。
  - 若工作频率为1333MHz（实际单位应是MT/s，表示每秒传送1333M次数据，实际时钟频率为333MHz），总线宽度为64位，则总线带宽为 $1333\text{MT/s} \times 8\text{B} = 10.5\text{GB/s}$ 。
- QPI（Quick Path Interconnect）总线
  - 目前在Intel架构中CPU芯片内部核之间、CPU芯片之间、CPU芯片与IOH（I/O Hub）芯片之间，都通过QPI总线互连
  - QPI是基于包交换的串行、高速点对点连接：发送方和接收方各有时钟信号，双方同时传输数据（各有20条数据线），每个QPI数据包包含80位，分两个时钟周期传送，每个时钟周期传两次，故每次传20位（16位数据+4位校验位），QPI总线带宽为：每秒传送次数 $\times 2\text{B} \times 2$ 。
  - QPI总线的速度单位（工作频率）为GT/s，表示每秒传送多少G次。若QPI时钟频率为2.4GHz，则速度为4.8GT/s，带宽为 $4.8\text{G} \times 2\text{B} \times 2 = 19.2\text{GB/s}$ 。

# 存储器总线

## 3.4

CPU芯片内含的核，共4个



存储器总线64位宽，速度为1333MT/s

总带宽为： $3 \times 8B \times 1333M = 32GB/s$

从Core i7开始，北桥在CPU芯片内，CPU通过存储器总线（即内存条插槽，图中为三通道插槽）直接和内存条相连。3个存控包含在CPU芯片内。

I/O总线用于为系统中的各种I/O设备提供输入输出通道

I/O总线在物理上可以是主板上的I/O扩展槽，如：

第一代：ISA/EISA总线、VESA总线，早被淘汰

第二代：PCI、AGP、PCI-X，被逐渐淘汰

第三代：PCI-Express（串行总线，主流总线）

PCI-Express总线

两个PCI-Express设备之间以一个链路（link）相连

每个链路包含多条通路（lane），可以是1,2,4,8,16或32条

PCI-Express × n表示一个具有n条通路的PCI-Express链路

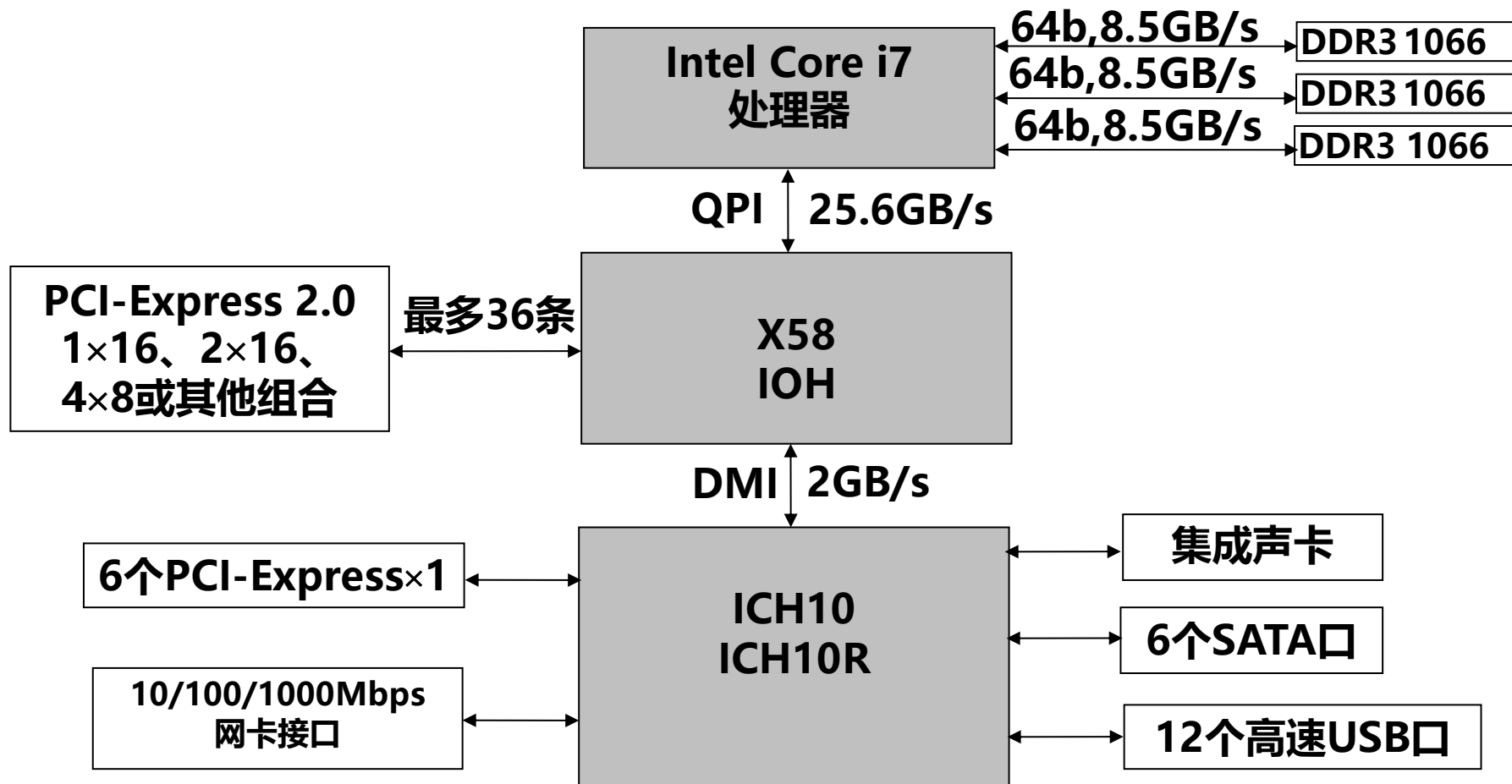
每条通路可同时发送和接受，每个数据字节被转换为10位信息被传输

PCI-Express 1.0下，每条通路的发送和接受速率都是2.5Gb/s，故PCI-Express × n的带宽为： $2.5\text{Gb/s} \times 2 \times n / 10 = 0.5\text{GB/s} \times n$ 。

**PCI-Express 1.0下，PCI-Express × 2的带宽为1GB/s，PCI-Express × 4的带宽为2GB/s，PCI-Express × 16的带宽为8GB/s。**



# 基于Core i7系列处理器的互连结构举例



**QPI总线的带宽为： $6.4GT/s \times 2B \times 2 = 25.6GB/s$**

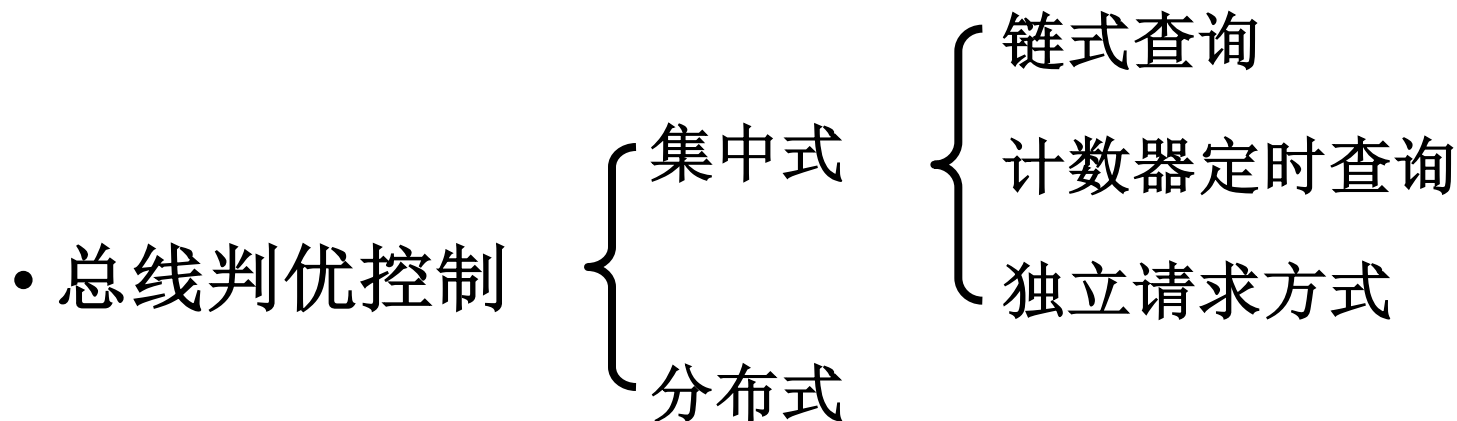
**每个存储器总线的带宽为： $64b/8 \times 1066 MT/s = 8.5 GB/s$  .**

## 3.5 总线控制

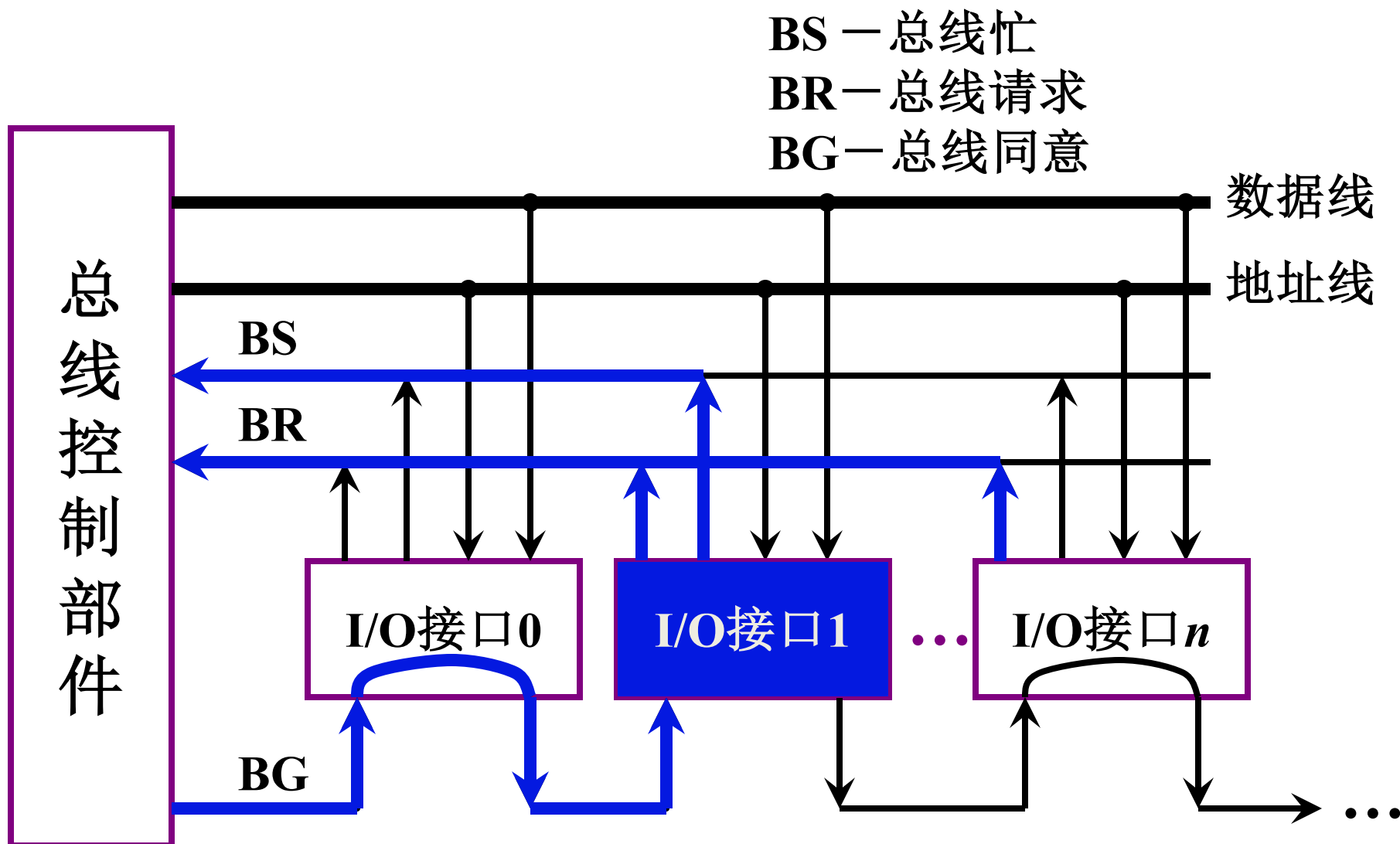
### 一、总线判优控制

#### 1. 基本概念

- 主设备(模块) 对总线有 **控制权**
- 从设备(模块) **响应** 从主设备发来的总线命令



## 2. 链式查询方式



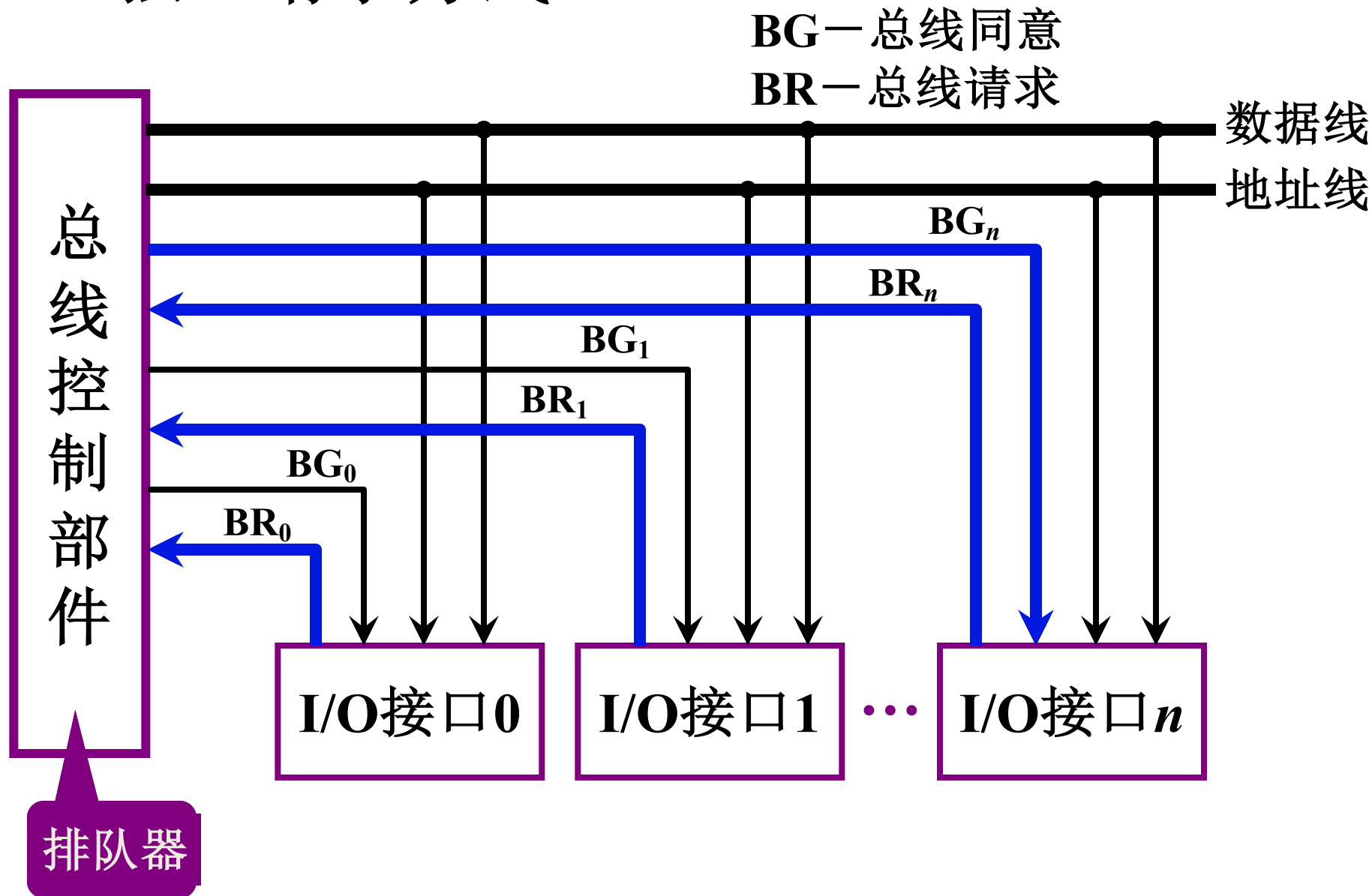


# 3.5



## 4. 独立请求方式


3.5



## 二、总线通信控制


1. 目的 解决通信双方 协调配合 问题

### 2. 总线传输周期

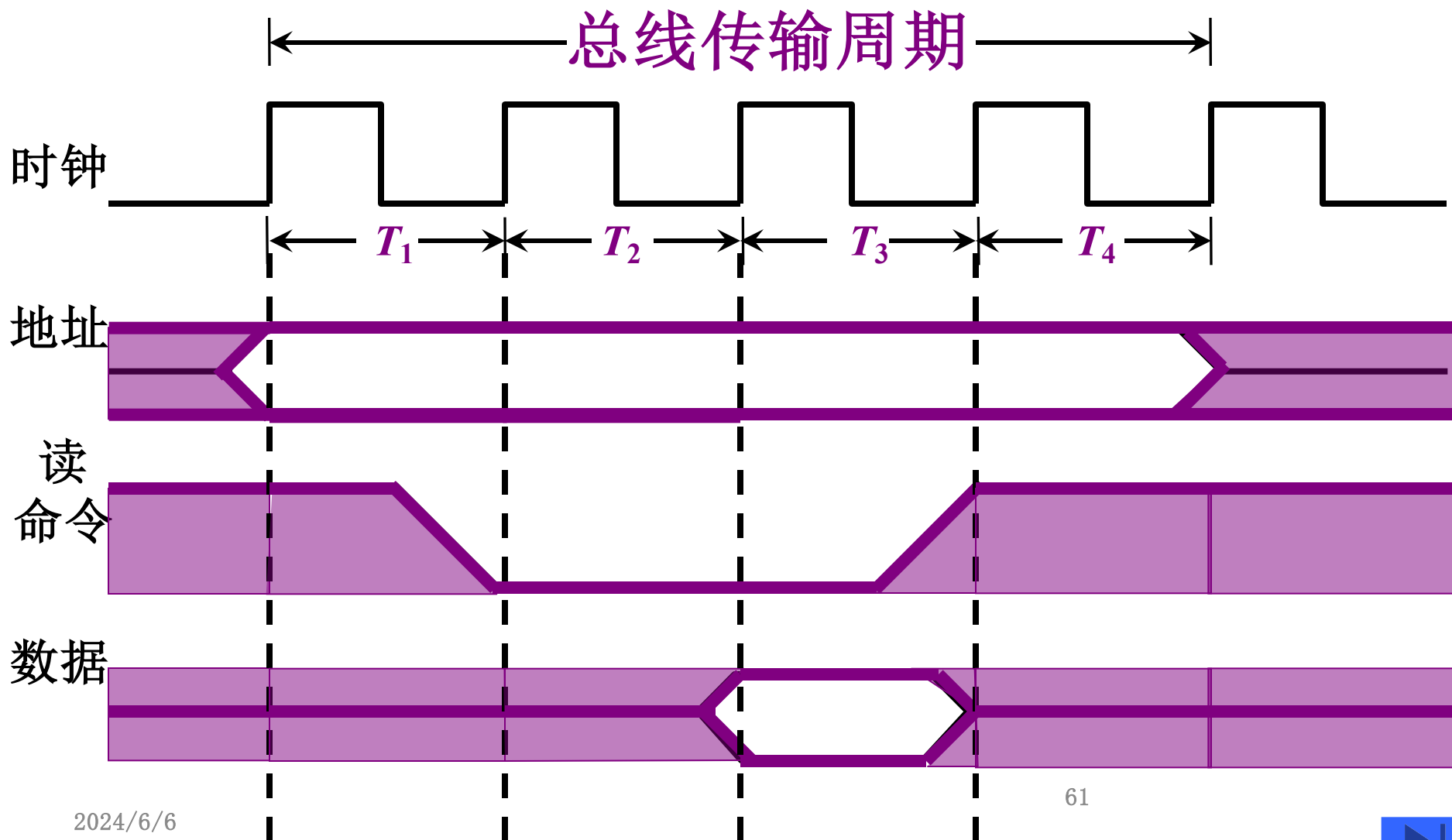


申请分配阶段	主模块申请，总线仲裁决定
寻址阶段	主模块向从模块 给出地址 和 命令
传数阶段	主模块和从模块 交换数据
结束阶段	主模块 撤消有关信息

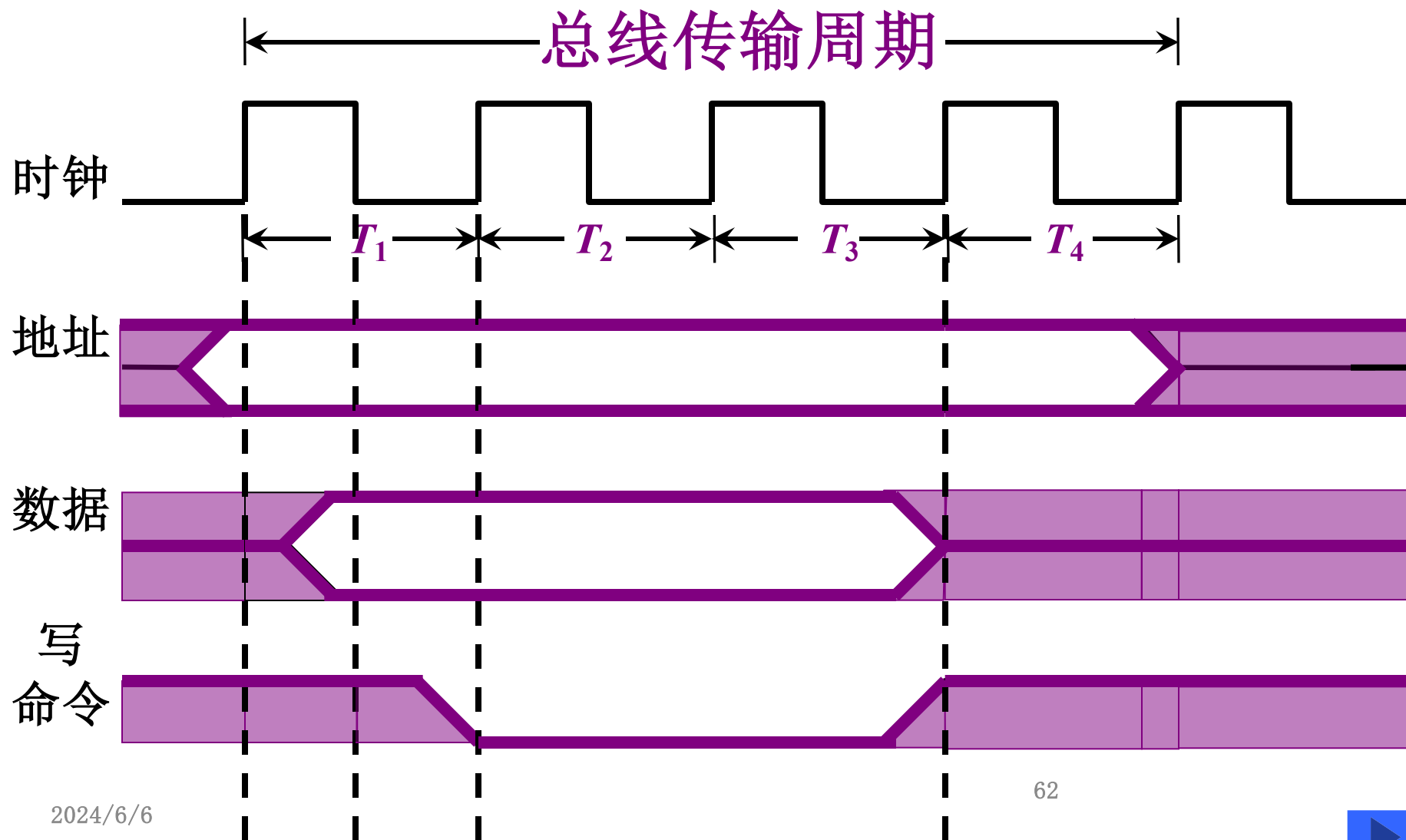
### 3. 总线通信的四种方式

- 
- 同步通信 由 统一时标 控制数据传送
  - 异步通信 采用 应答方式，没有公共时钟标准
  - 半同步通信 同步、异步结合
  - 分离式通信 充分 挖掘 系统 总线每个瞬间 的 潜力

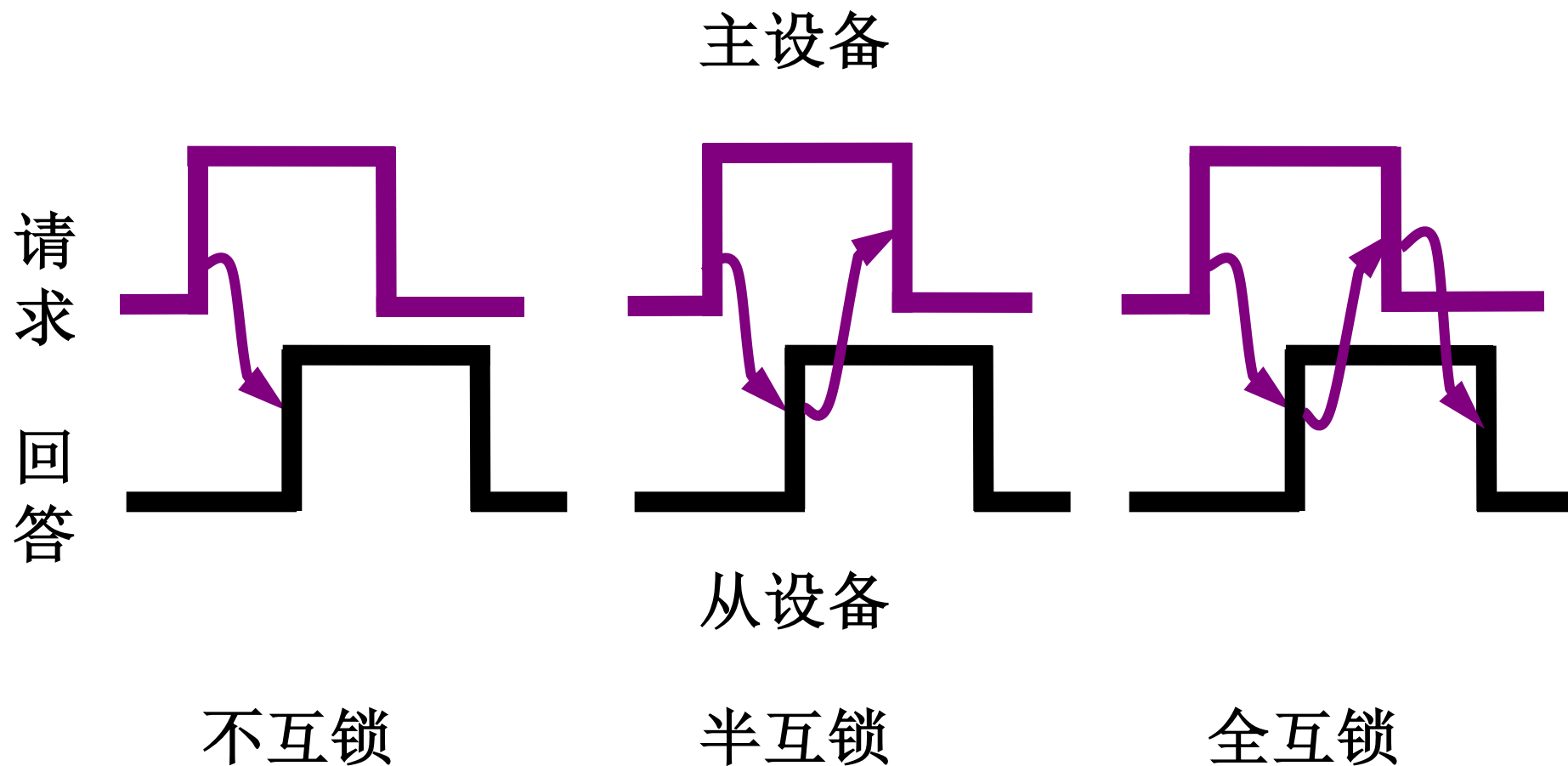
# (1) 同步式数据输入



## (2) 同步式数据输出



### (3) 异步通信



## (4) 半同步通信 (同步、异步 结合) 3.5

同步 发送方 用系统 时钟前沿 发信号

接收方 用系统 时钟后沿 判断、识别

异步 允许不同速度的模块和谐工作

增加一条 “等待” 响应信号  $\overline{\text{WAIT}}$



# 以输入数据为例的半同步通信时序

## 3.5

$T_1$  主模块发地址

$T_2$  主模块发命令

$T_w$  当  $\overline{\text{WAIT}}$  为低电平时，等待一个  $T$

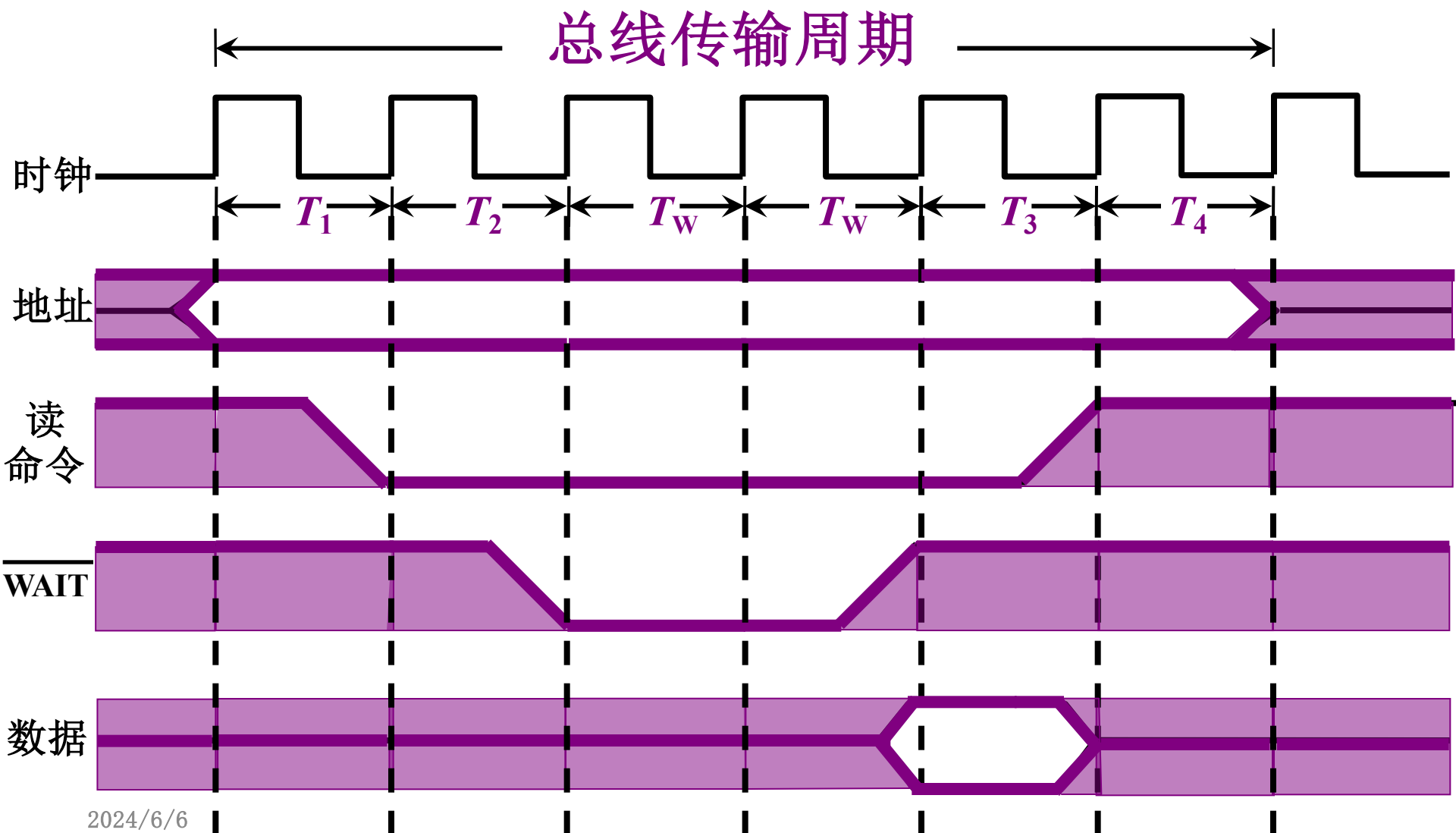
$T_w$  当  $\overline{\text{WAIT}}$  为低电平时，等待一个  $T$

⋮

$T_3$  从模块提供数据

$T_4$  从模块撤销数据，主模块撤销命令

# (4) 半同步通信 (同步、异步 结合) 3.5



## 上述三种通信的共同点

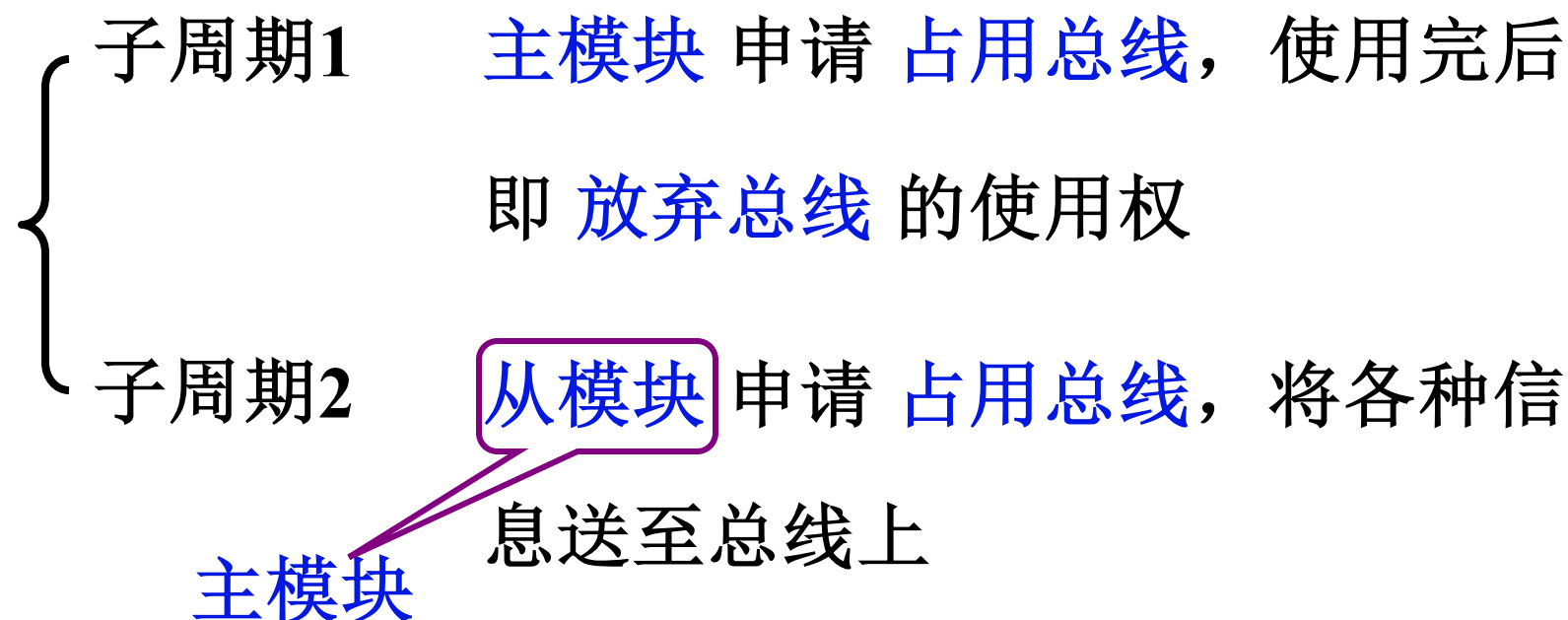
### 一个总线传输周期（以输入数据为例）

- 主模块发地址、命令      占用总线
- 从模块准备数据      不占用总线    总线空闲
- 从模块向主模块发数据    占用总线

## (5) 分离式通信

充分挖掘系统总线每个瞬间的潜力

一个总线传输周期



## 分离式通信特点

1. 各模块有权申请占用总线
2. 采用同步方式通信，不等对方回答
3. 各模块准备数据时，不占用总线
4. 总线被占用时，无空闲

充分提高了总线的有效占用

# 第 5 章 输入输出系统

## 5.1 概述

## 5.2 外部设备

## 5.3 I/O接口

## 5.4 程序查询方式

## 5.5 程序中断方式

## 5.6 DMA方式

# 5.1 概述

## 一、输入输出系统的发展概况

### 1. 早期

分散连接

程序查询方式

CPU 和 I/O设备 串行 工作

### 2. 接口模块和 DMA 阶段

总线连接

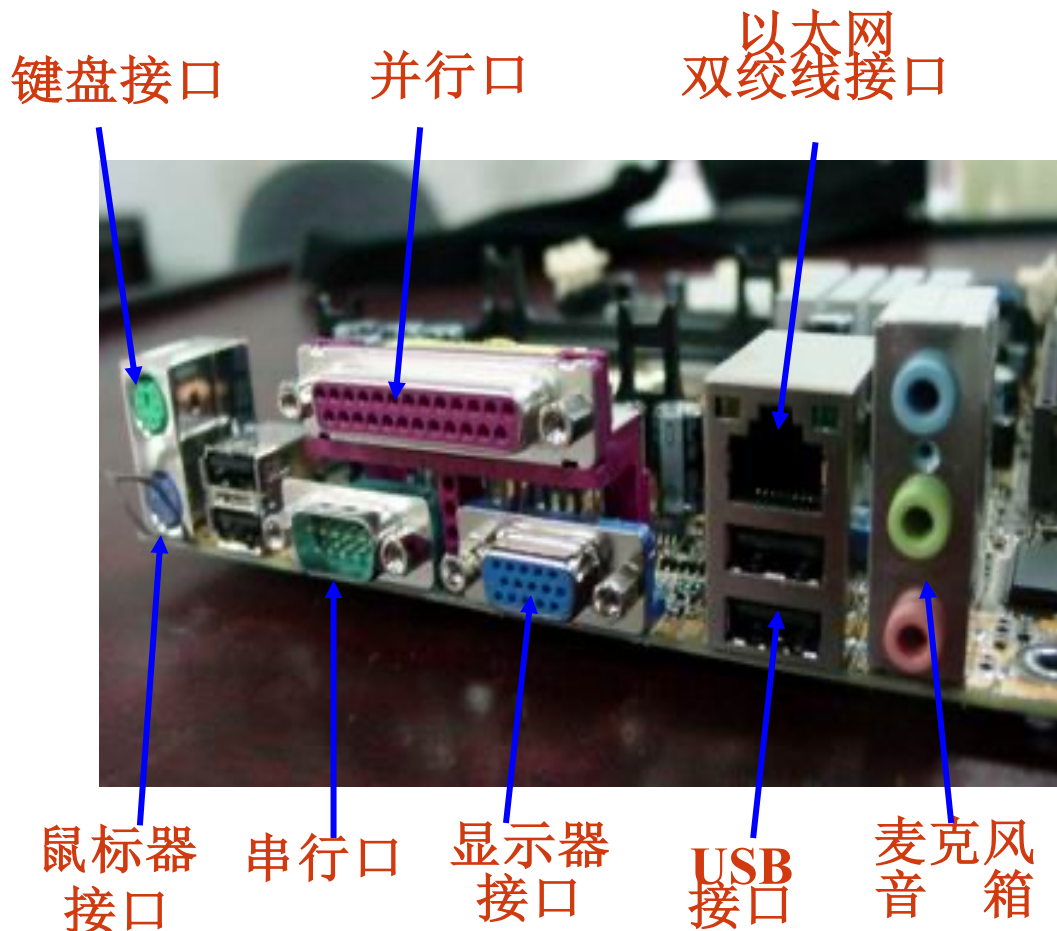
CPU 和 I/O设备 并行 工作 { 中断方式  
DMA 方式

### 3. 具有通道结构的阶段

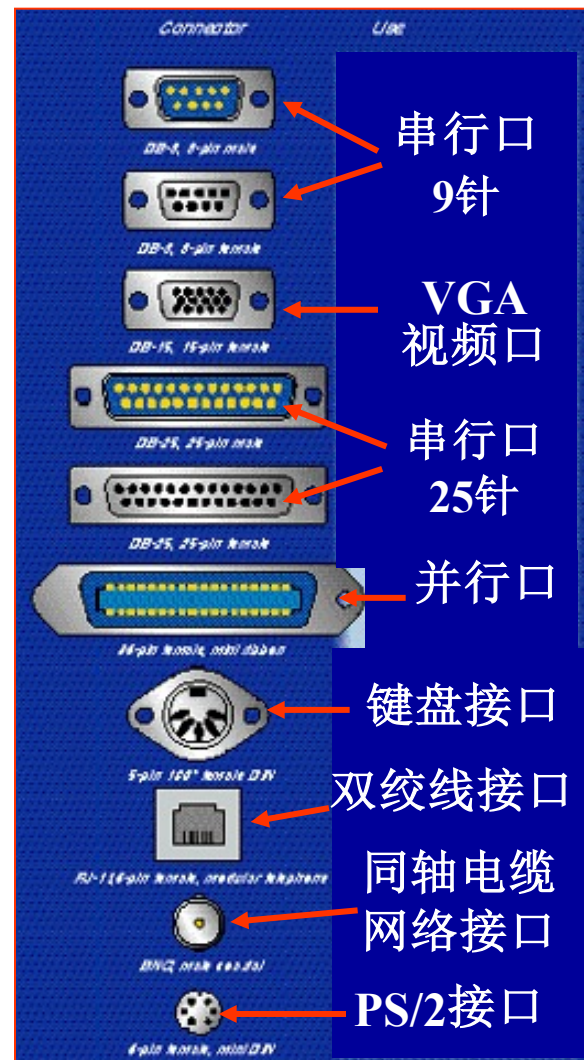
### 4. 具有 I/O 处理机的阶段

# I/O设备接口插座（连接器）

## 5.1



(安装在主板上的I/O设备接口插座)





## 二、输入输出系统的组成

## 5.1

### 1. I/O 软件

(1) I/O 指令      CPU 指令的一部分

操作码	命令码	设备码
-----	-----	-----

(2) 通道指令      通道自身的指令

指出数组的首地址、传送字数、操作命令

如 IBM/370 通道指令为 64 位

### 2. I/O 硬件

设备      I/O 接口

设备      设备控制器      通道

# 三、I/O 设备与主机的联系方式

## 5.1

### 1. I/O 设备编址方式

(1) 统一编址      用取数、存数指令

(2) 不统一编址    有专门的 I/O 指令

### 2. 设备选址

用设备选择电路识别是否被选中

### 3. 传送方式

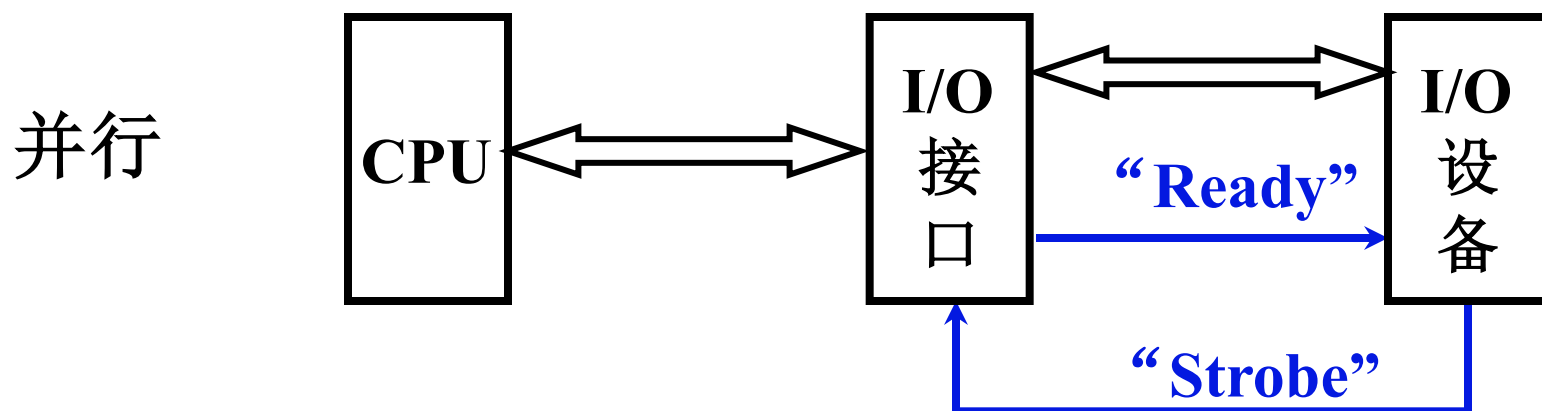
(1) 串行

(2) 并行

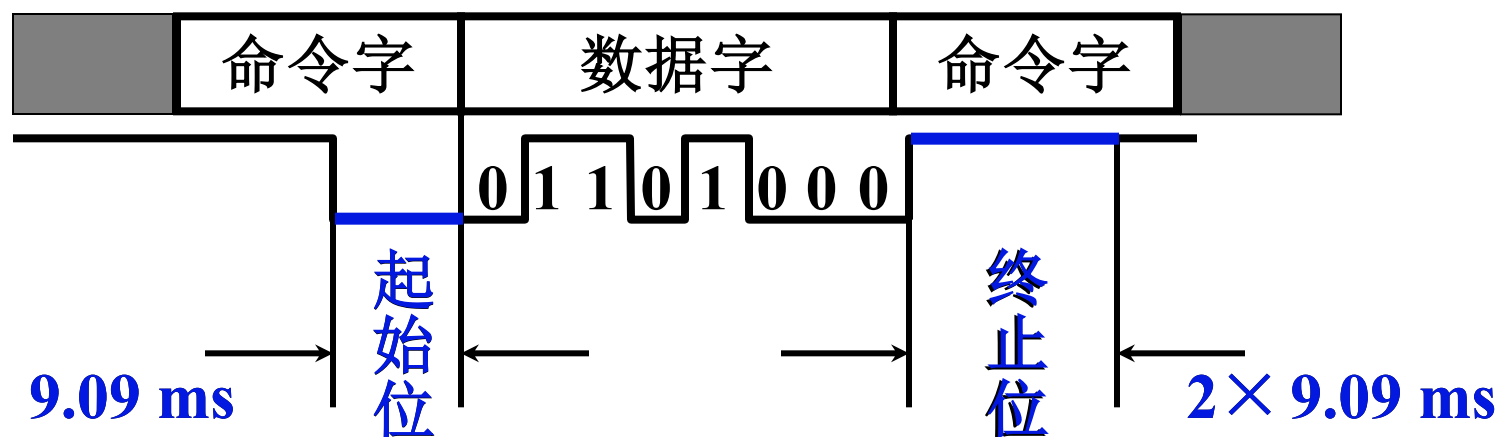
## 4. 联络方式

(1) 立即响应

(2) 异步工作采用应答信号



串行

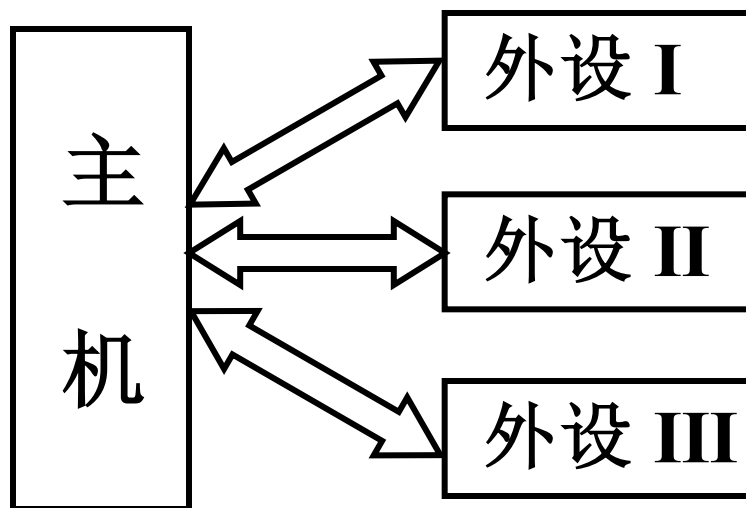


(3) 同步工作采用同步时标

# 5. I/O 设备与主机的连接方式

## 5.1

### (1) 辐射式连接



每台设备都配有一套  
控制线路和一组信号线

不便于增删设备

### (2) 总线连接

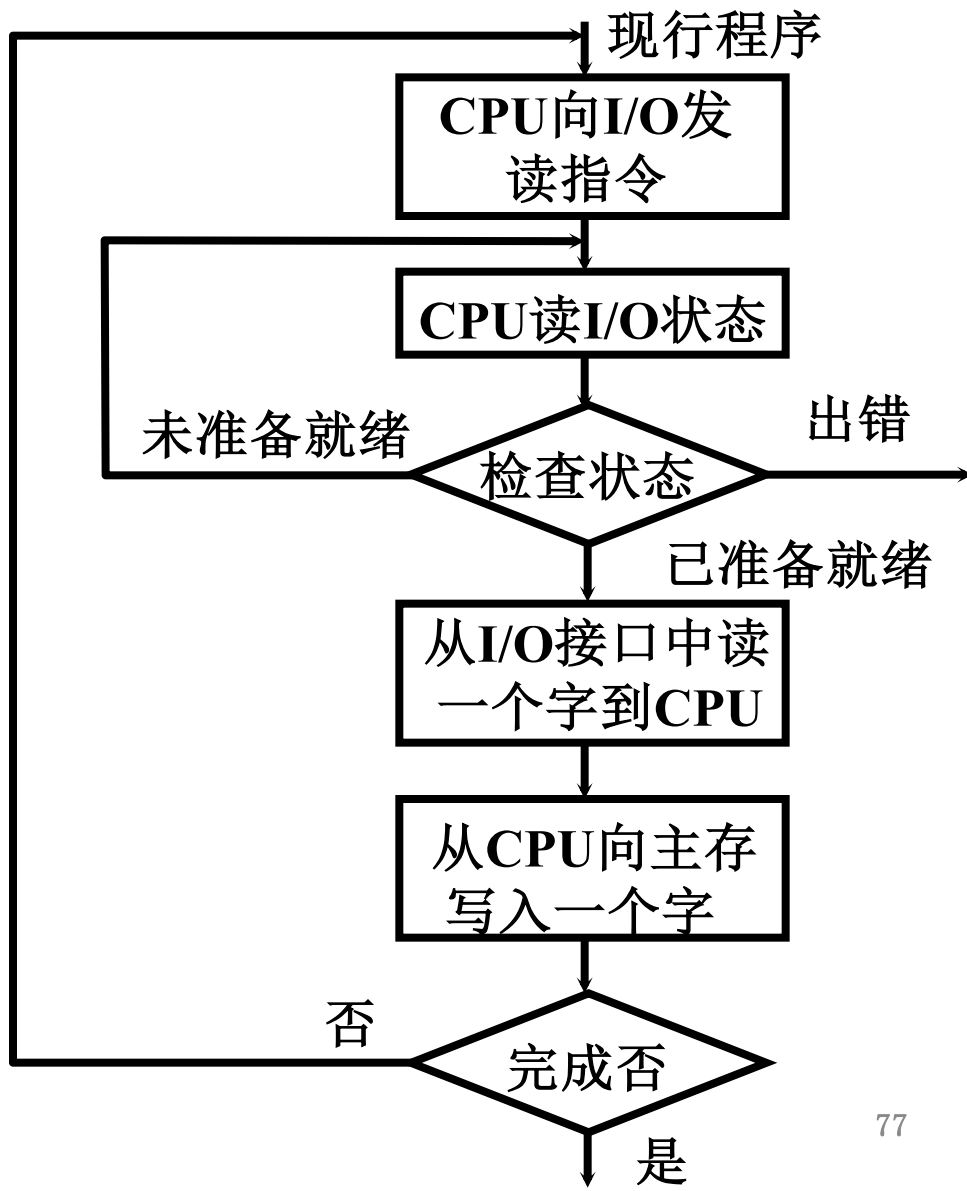
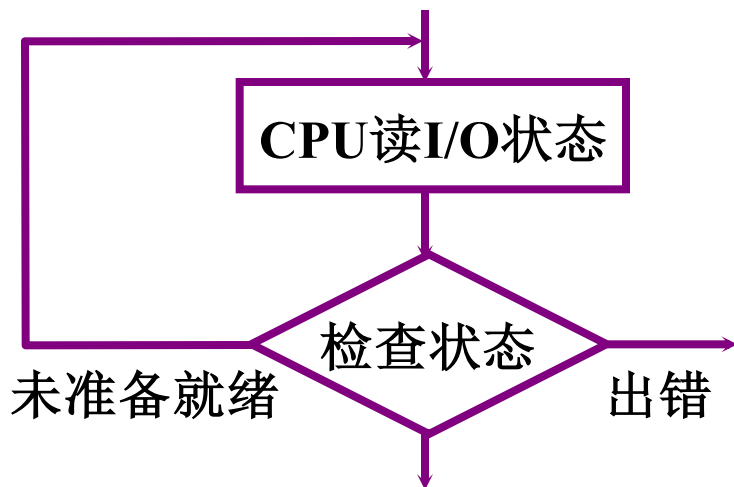
便于增删设备

# 四、I/O设备与主机信息传送的控制方式 5.1

## 1. 程序查询方式

CPU 和 I/O 串行工作

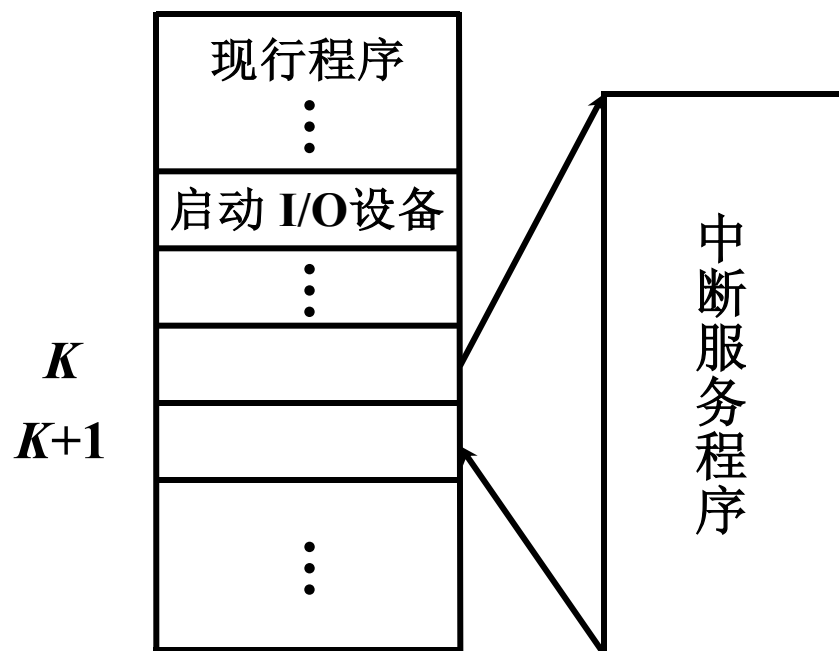
踏步等待



## 2. 程序中中断方式

I/O 工作 { 自身准备      CPU 不查询  
与主机交换信息    CPU 暂停现行程序

**CPU 和 I/O 并行工作**

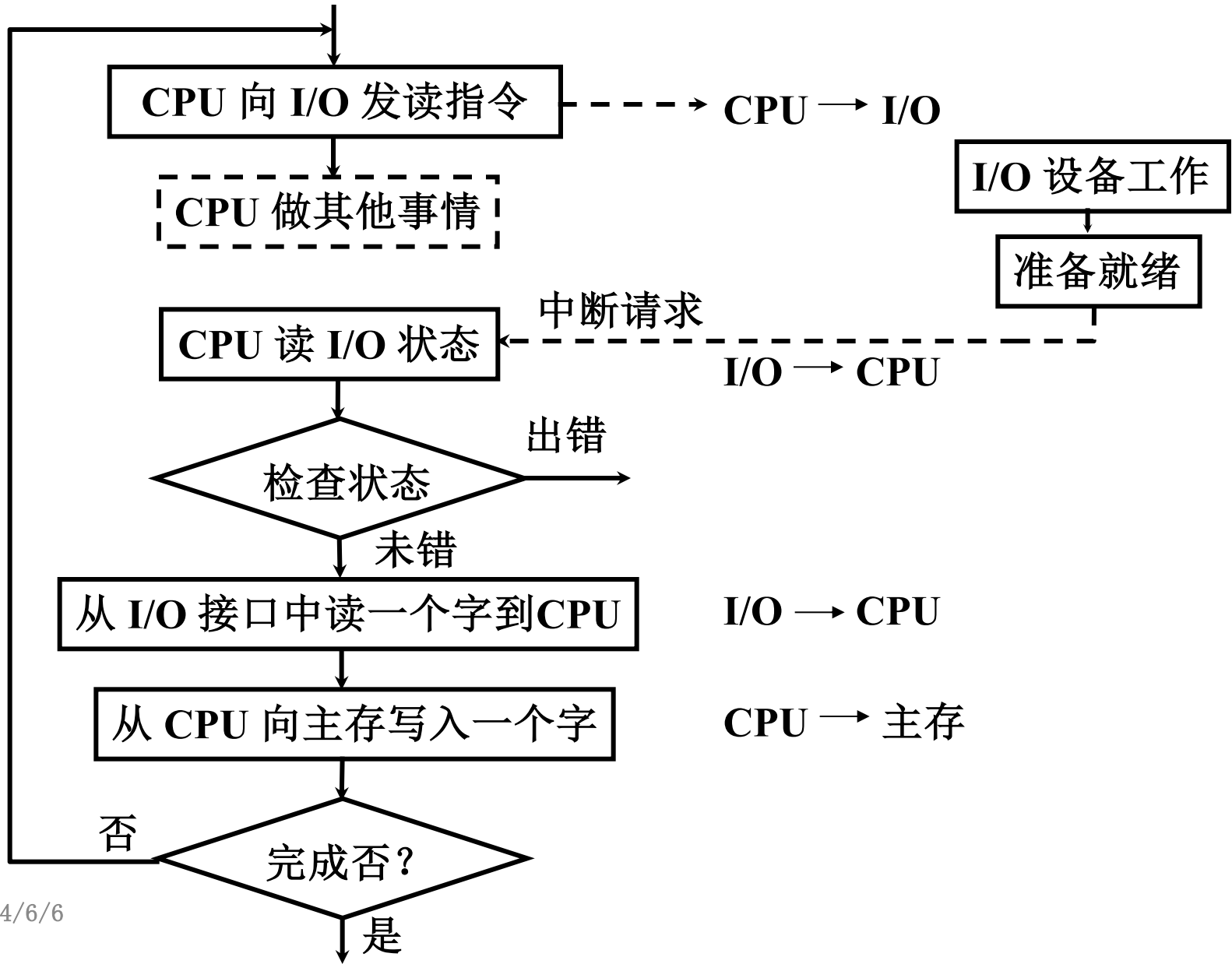


**没有踏步等待现象**

**中断现行程序**

# 程序中中断方式流程

## 5.1



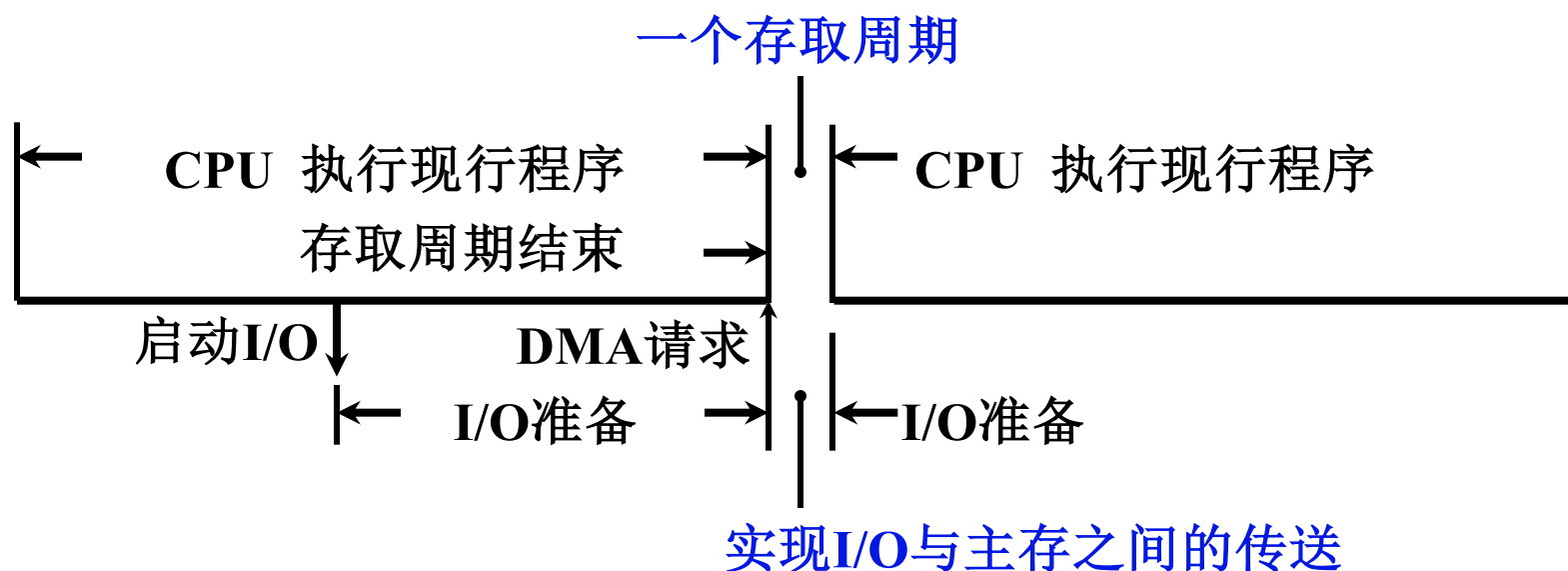
### 3. DMA 方式

主存和 I/O 之间有一条直接数据通道

不中断现行程序

周期挪用（周期窃取）

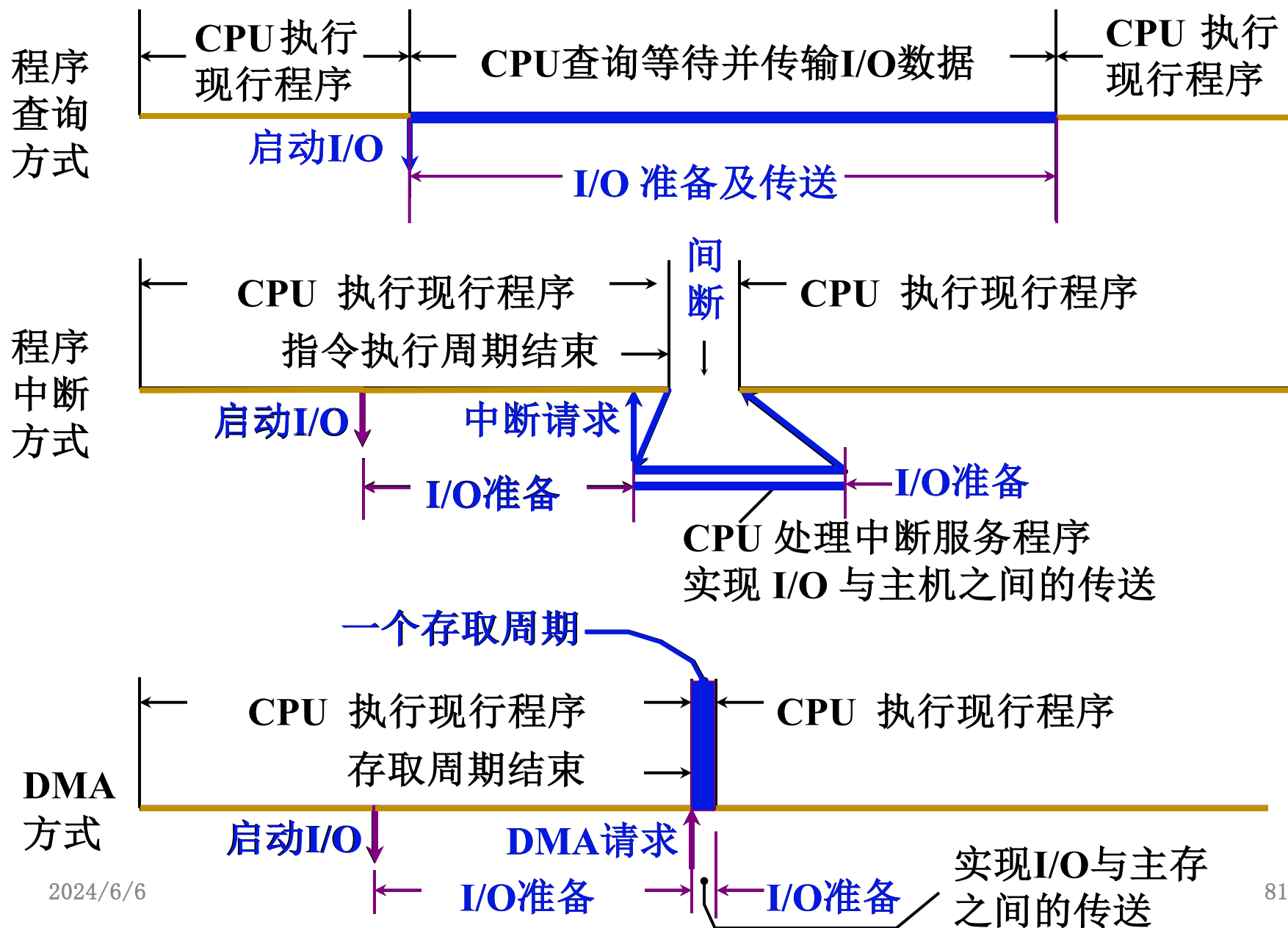
**CPU 和 I/O 并行工作**





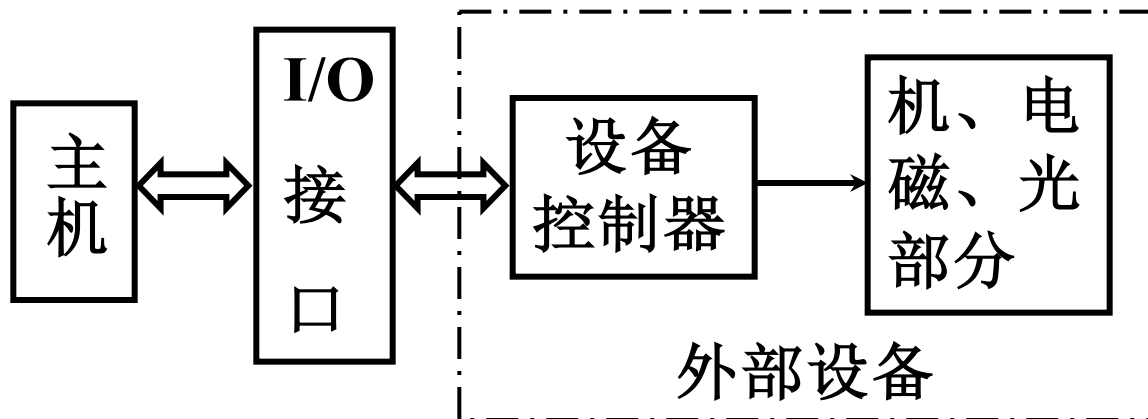
# 三种方式的 CPU 工作效率比较

5.1



# 5.2 I/O设备

## 一、概述



外部设备大致分三类

- |              |               |
|--------------|---------------|
| 1. 人机交互设备    | 键盘、鼠标、打印机、显示器 |
| 2. 计算机信息存储设备 | 磁盘、光盘、磁带      |
| 3. 机—机通信设备   | 调制解调器等        |

## 二、输入设备

### 1. 键盘

按键

判断哪个键按下

将此键翻译成 ASCII 码（编码键盘法）

### 2. 鼠标

机械式      金属球   电位器

光电式      光电转换器

### 3. 触摸屏

## 三、输出设备

### 1. 显示器

(1) 字符显示      字符发生器

(2) 图形显示      主观图像

(3) 图像显示      客观图像

### 2. 打印机

(1) 击打式      点阵式（逐字、逐行）

(2) 非击打式      激光（逐页）喷墨（逐字）

## 四、其他

1. A/D、D/A      模拟/数字（数字/模拟）转换器

2. 终端              由键盘和显示器组成

完成显示控制与存储、键盘管理及通信控制

3. 汉字处理      汉字输入、汉字存储、汉字输出

## 五、多媒体技术

1. 什么是多媒体

2. 多媒体计算机的关键技术

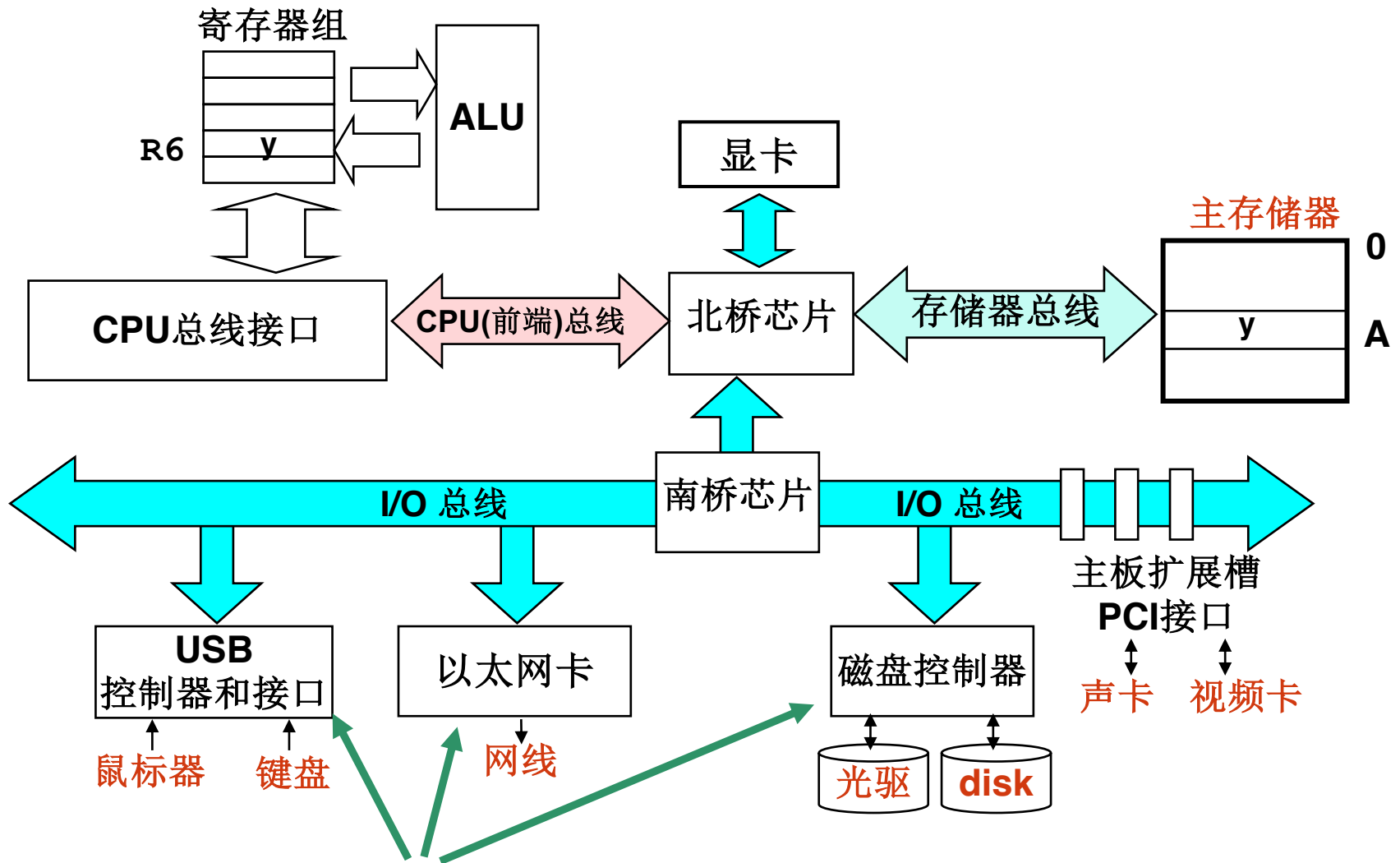
## 5.3 I/O 接口

### 一、概述

为什么要设置接口？

1. 实现设备的选择
2. 实现数据缓冲达到速度匹配
3. 实现数据串一并格式转换
4. 实现电平转换
5. 传送控制命令
6. 反映设备的状态（“忙”、“就绪”、“中断请求”）

# I/O总线、I/O接口与I/O设备的关系 5.3



把I/O控制器和插座合起来称为I/O接口。