



# 计算机组成原理

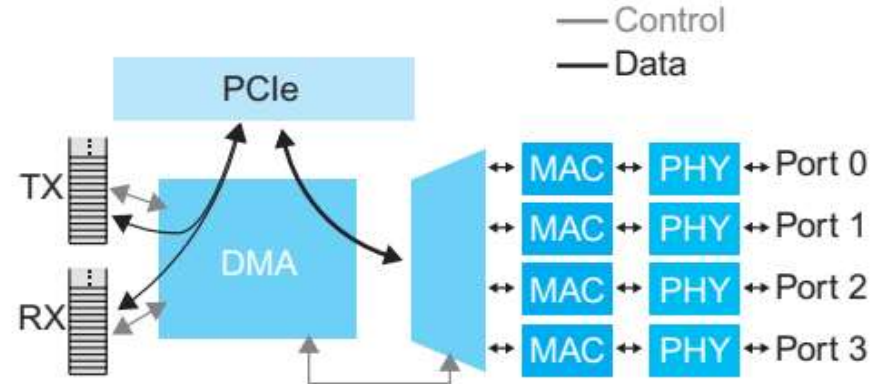
## 系统互连与通信——总线 (RV \$6.9, 唐第3章)

李曦 ([llxx@ustc.edu.cn](mailto:llxx@ustc.edu.cn))



# RV \$6.9: 总线与I/O

- 示例: NetFPGA 10G Ethernet NIC
  - an FPGA-based open platform for network research and classroom experimentation
- 总线
  - PCI: 并行总线
  - PCIe总线: 串行总线
- I/O机制
  - MMIO
  - Polling, Interrupt-driven I/O, DMA
  - 设备驱动程序



# 系统部件互连 (interconnect)

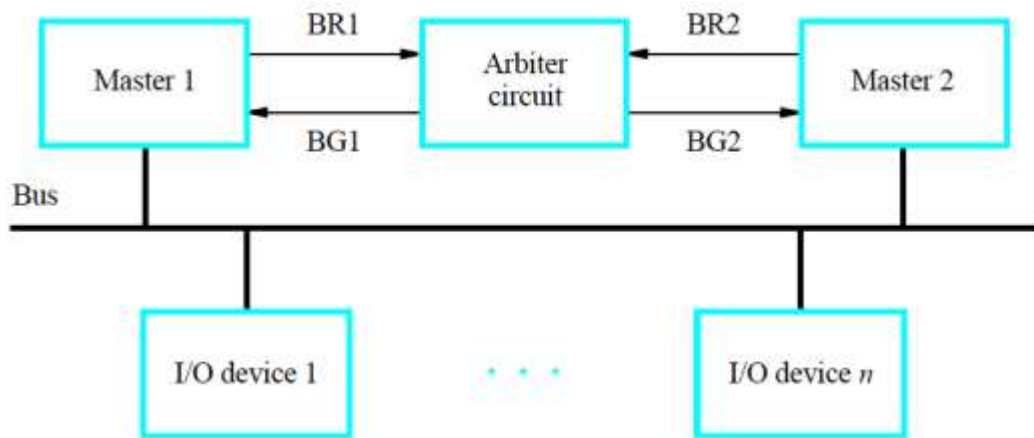
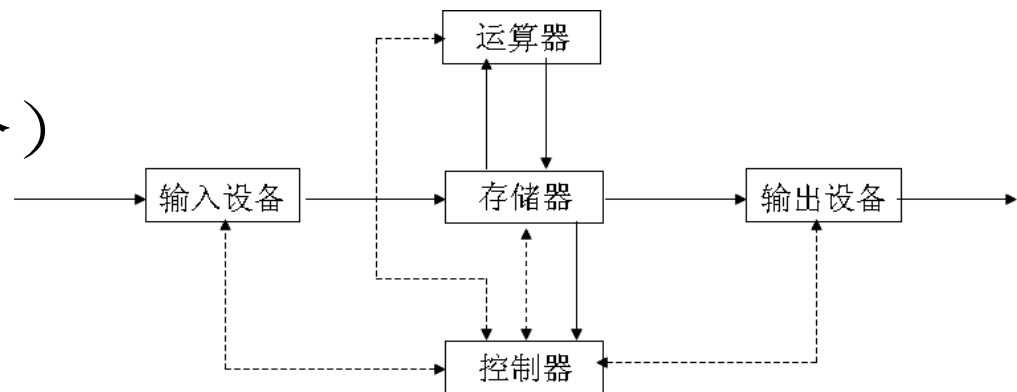
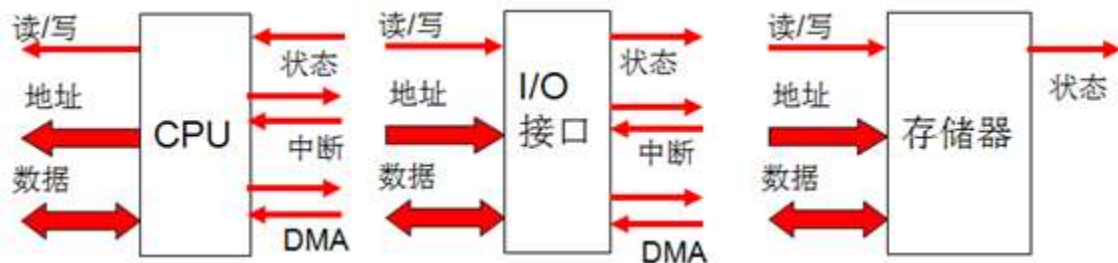


- 信号线wire

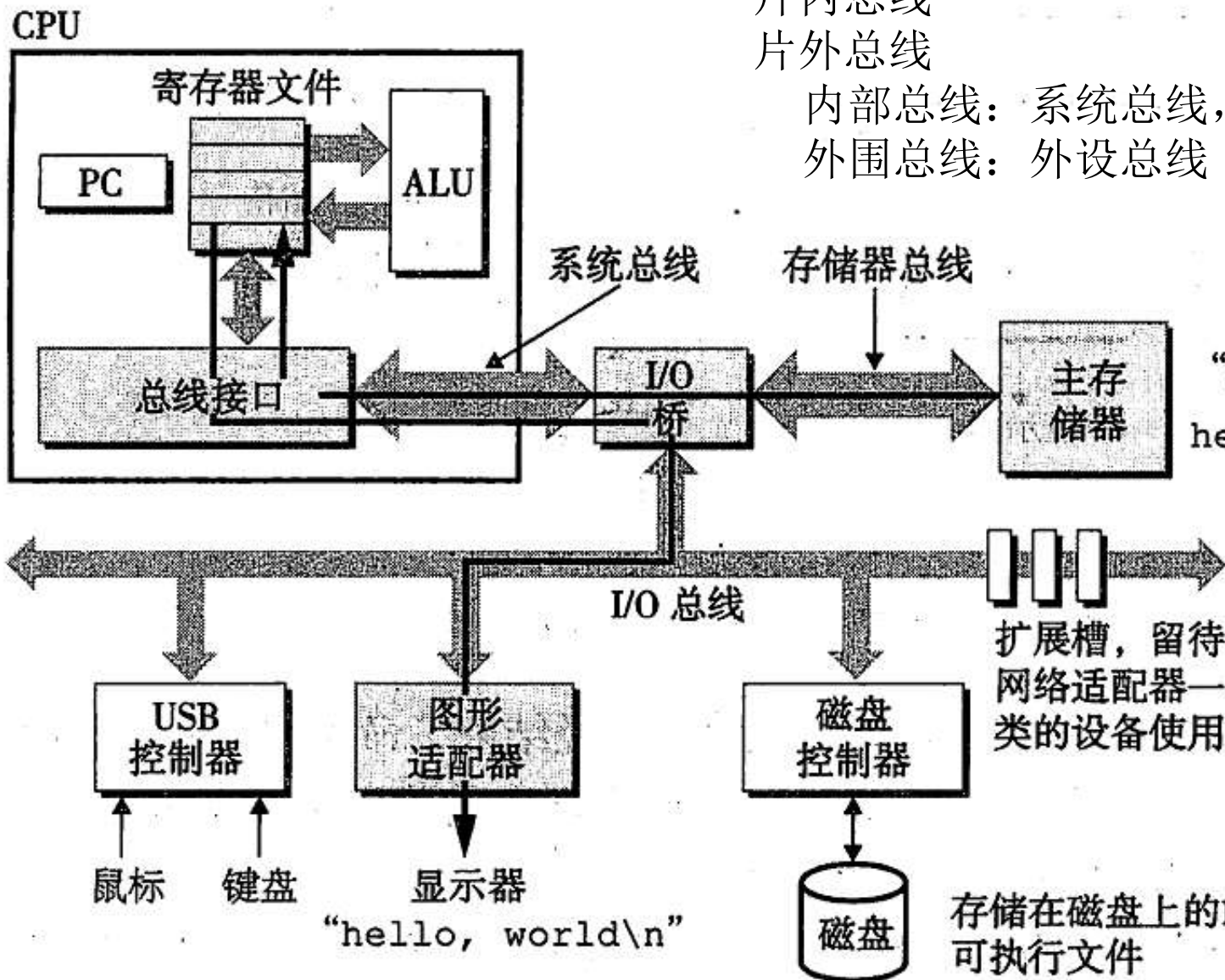
- 地址
- 数据
- 控制 (状态, 命令)

- 互连模式

- 点对点: 独占式
- 总线: 共享式
- 混合



# "hello, world\n"



片内总线  
片外总线  
内部总线: 系统总线, I/O总线  
外围总线: 外设总线

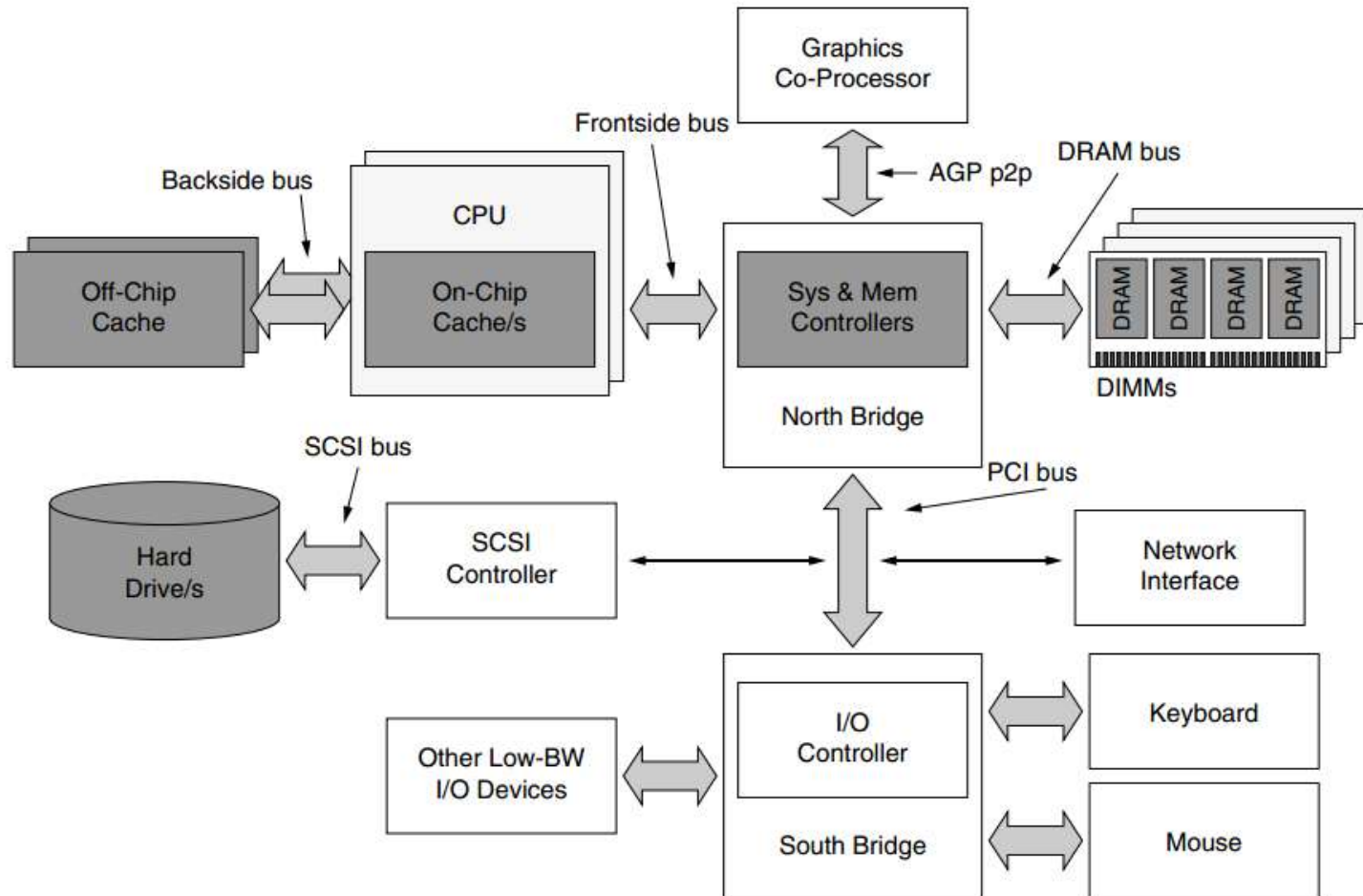
"hello, world\n"  
hello 代码

扩展槽, 留待  
网络适配器一类的设备使用



存储在磁盘上的hello可执行文件

# PC机中的总线：层次化拓扑结构

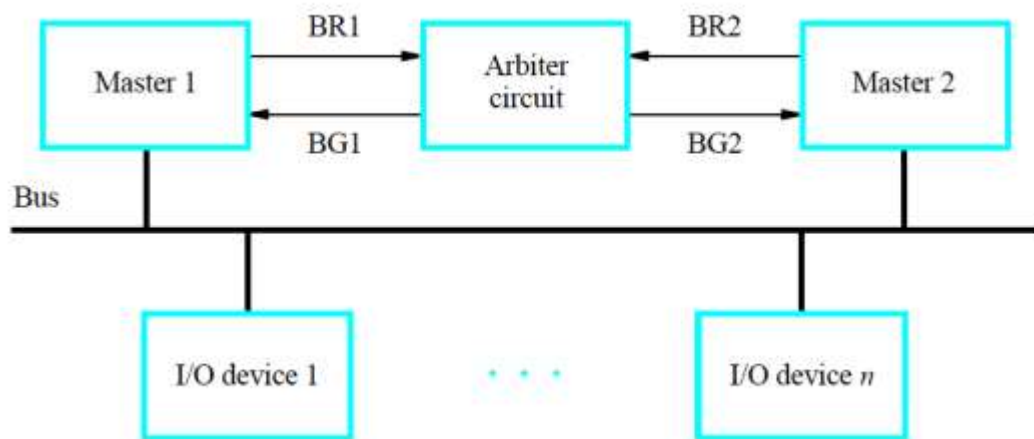
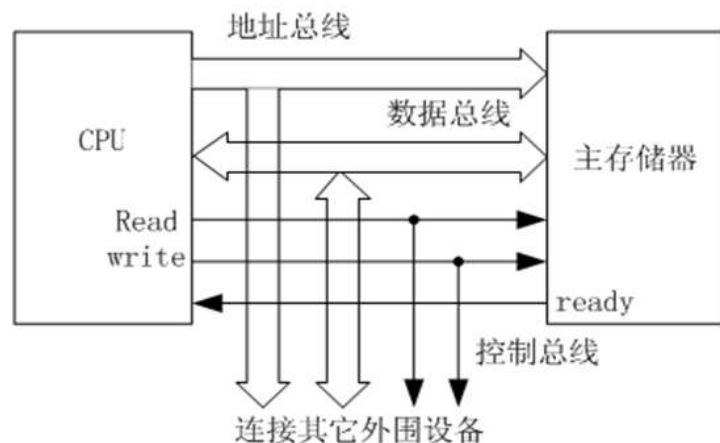


**FIGURE 0v.3:** Typical PC organization. The memory subsystem is one part of a relatively complex whole. This figure illustrates a two-way multiprocessor, with each processor having its own dedicated off-chip cache. The parts most relevant to this text are shaded in grey: the CPU and its cache system, the system and memory controllers, the DIMMs and their component DRAMs, and the hard drive/s.

# 总线



- 定义：连接多个部件的**传输线**
  - 数据、地址、控制（状态，命令）
- 总线的关键特征：多设备**共享**的传输线。
  - 主设备**master**：任意时刻只能有一个设备向总线**发送**信息
    - 系统瓶颈
  - 从设备**slave**：多个部件可以同时从总线**接受**相同的信息
    - 广播式
- 总线拓扑：单总线，多总线（层次化）
- 总线周期：
  - 数据传输过程
  - **同步**方式
- 总线**仲裁**
  - 单/多**master**系统
  - 集中式，分布式（**CAN**）
- 总线标准





# 总线传输过程：总线周期



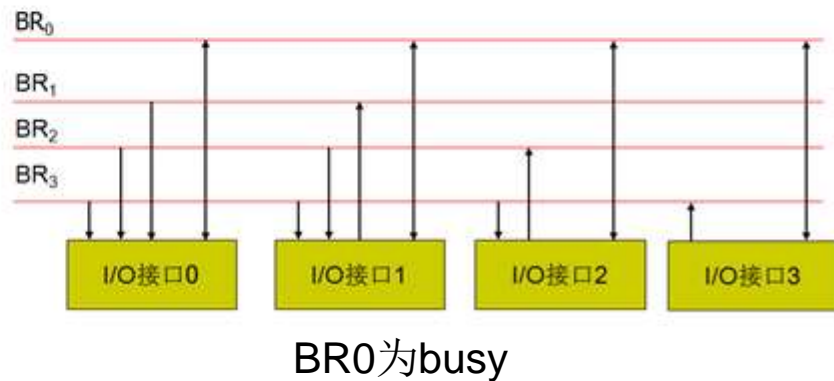
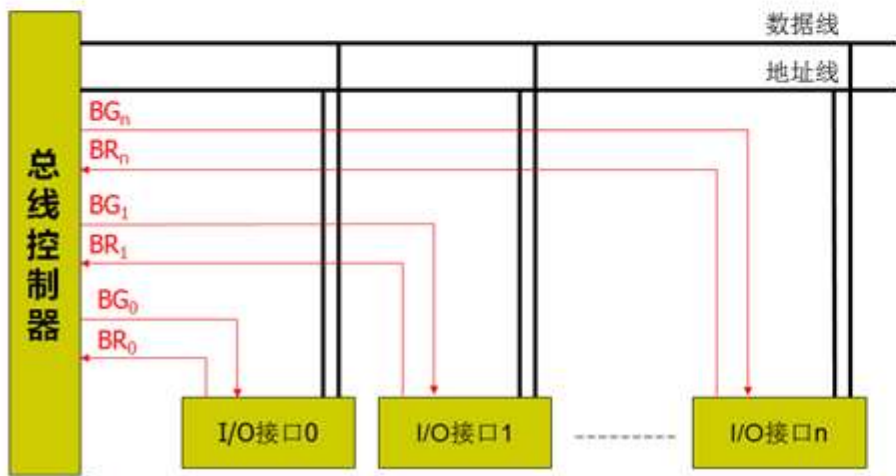
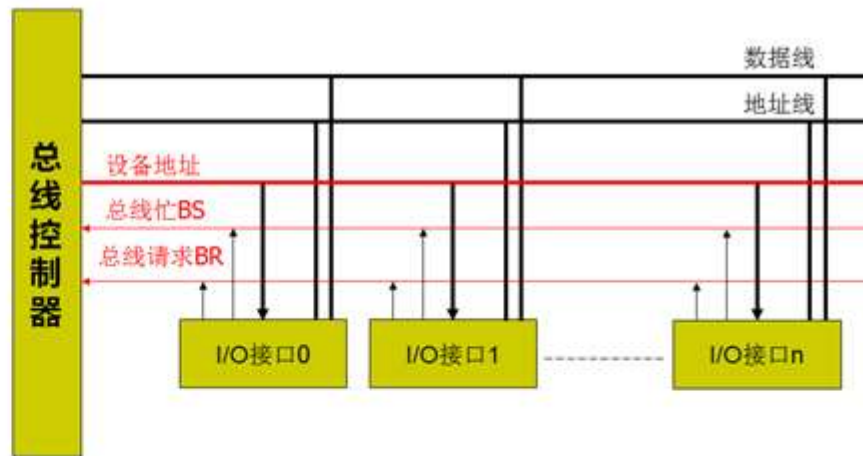
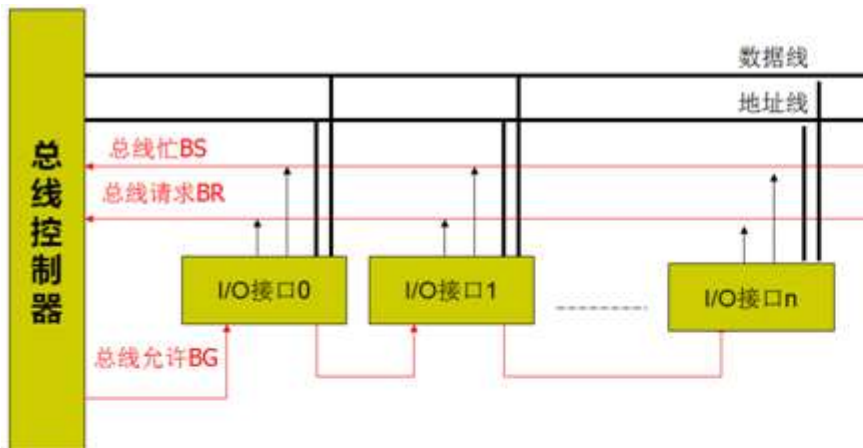
- 完成一次总线传输可分为3个阶段
  - 申请阶段：仲裁算法（集中式、分布式）
    - 主设备提出申请，经总线仲裁机构决定将下一传输周期的总线使用权授予某一申请者。
    - 链式查询、计数器定时查询、独立请求
  - 数据传输：同步控制，传输模式（串/并、单字/突发）
    - 寻址阶段：取得了使用权的主设备，通过总线发出本次打算访问的从设备的存储地址或设备地址及有关命令，启动从设备。
    - 传数阶段：主设备和从设备进行数据交换。
  - 结束阶段：从总线上撤除有关信息，让出使用权。
- 对于独占系统，只有数据传输阶段
  - 独占系统：只有一个主设备的简单系统，如8086



# 总线判优： 复杂性、公平性、可靠性、可扩展性

集中式系统：链式查询、计数器定时查询、独立请求

分布式：各节点有各自ID和仲裁器。按ID优先级高者占用总线。例：自举式







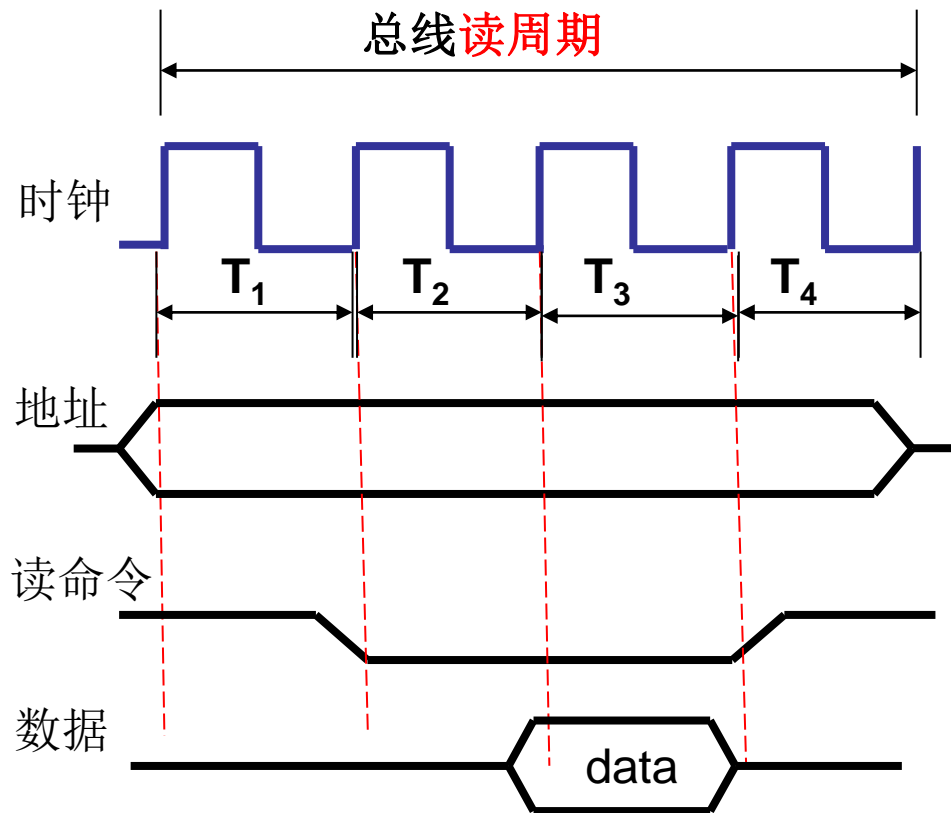
# 总线通信控制：同步通信方式

- 通信双方由**统一时标**控制数据传送称为**同步通信**。
  - 时标通常由CPU的总线控制部件发出，送到总线上的所有部件；
  - 也可以由每个部件各自的时序发生器发出，但是必须有总线控制部件发出的时钟信号对它们进行同步。

- 对于**读**命令，其传输周期为：

- $T_1$ : 主模块发地址
- $T_2$ : 主模块发读命令
- $T_3$ : 从模块提供数据
- $T_4$ : 主模块撤销读命令

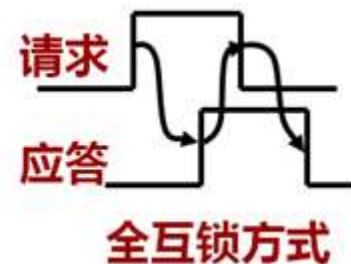
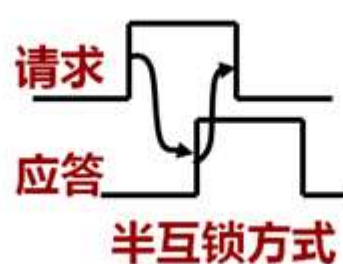
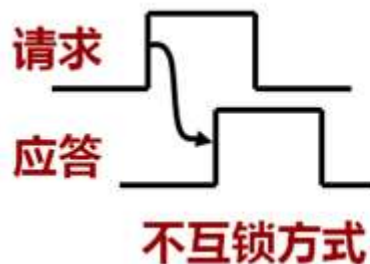
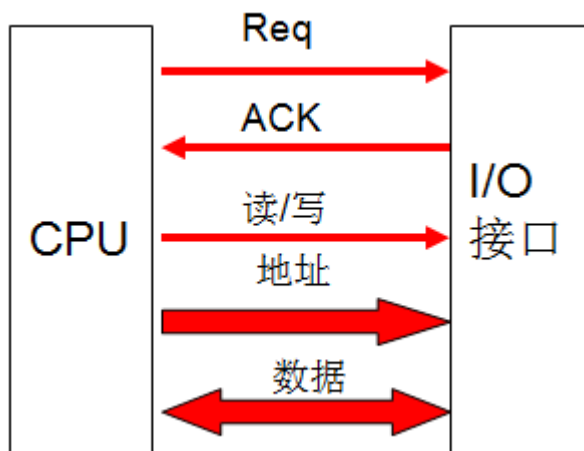
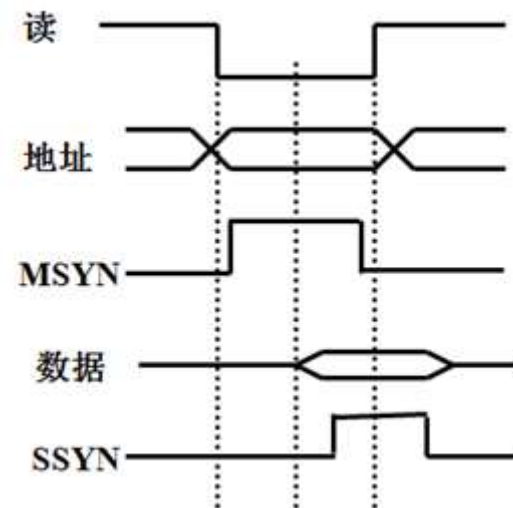
- 仲裁？
- 写操作时序图？



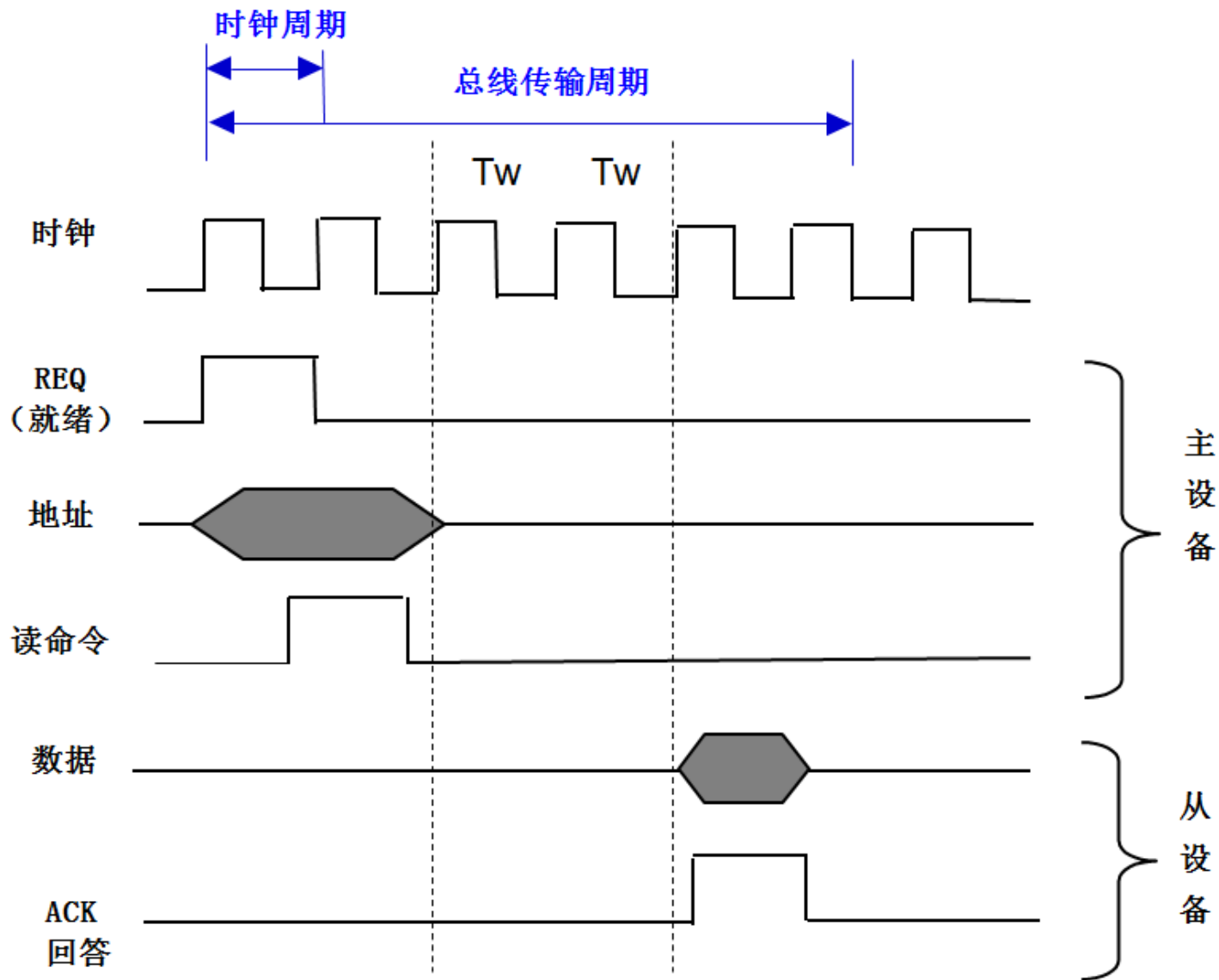


# 总线通信控制：异步通信方式

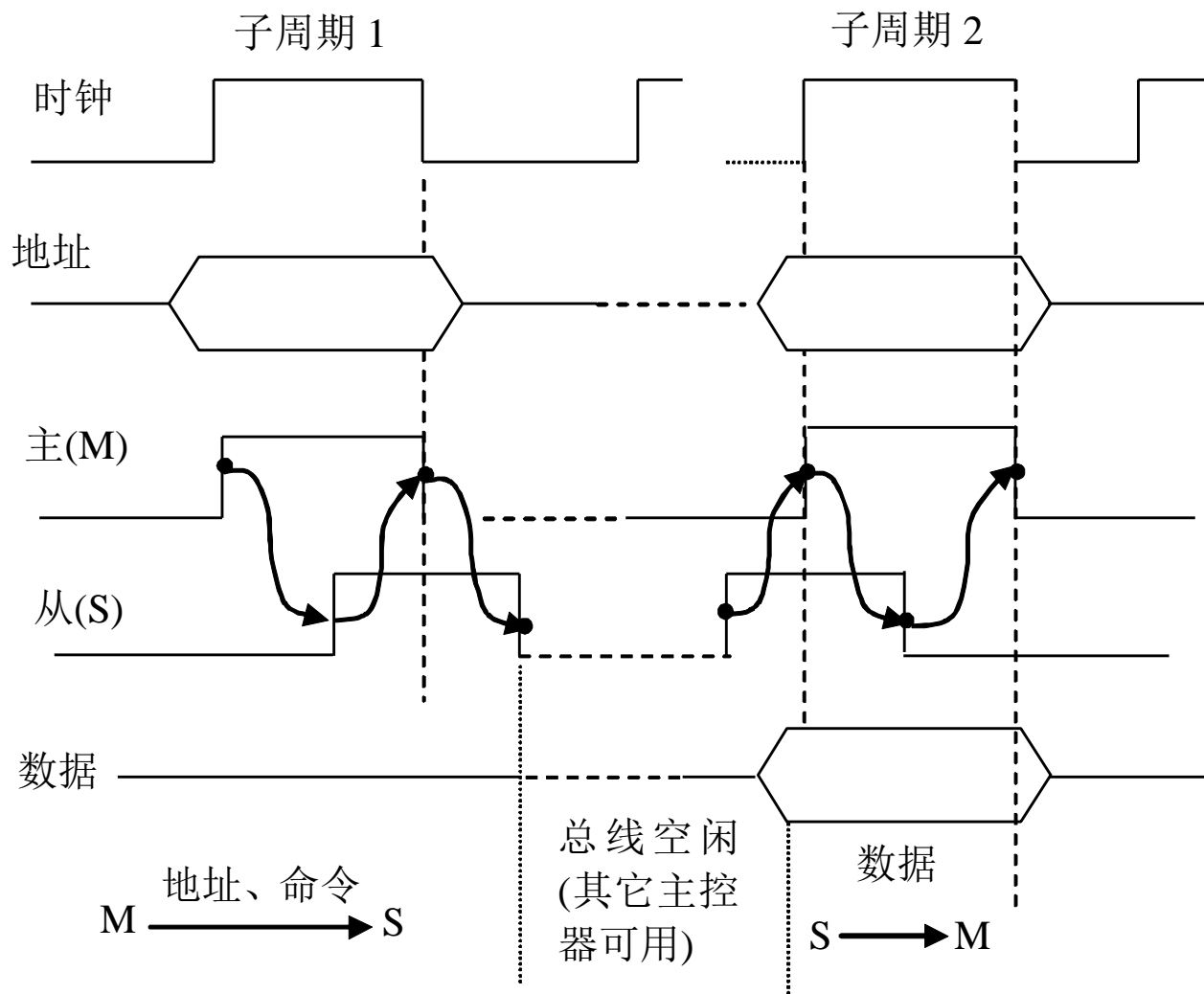
- 没有公共的时钟，采用应答方式（握手）
  - 主设备Request (MSYN)，从设备响应Ack (SSYN)；
- 确认总线周期的开始和结束
  - 三种应答：不互锁，半互锁，全互锁
  - 例：PCI半互锁



# 半同步通信协议

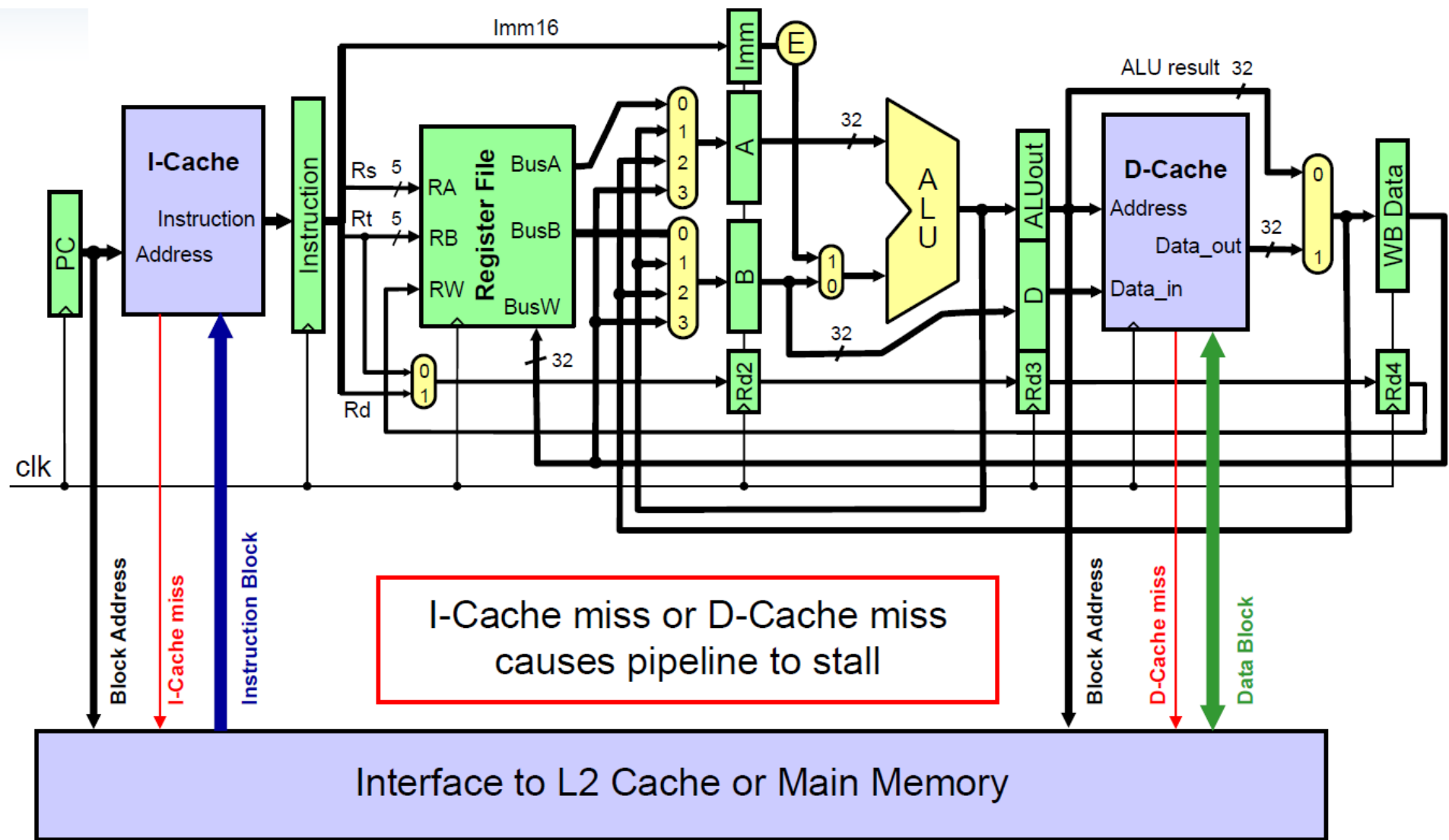


# 分离事务通信协议：总线利用率



总线事务 (**Bus transaction**)：在一个总线周期中发生的一系列活动。  
典型：请求操作、裁决操作、地址传输、数据传输、总线释放

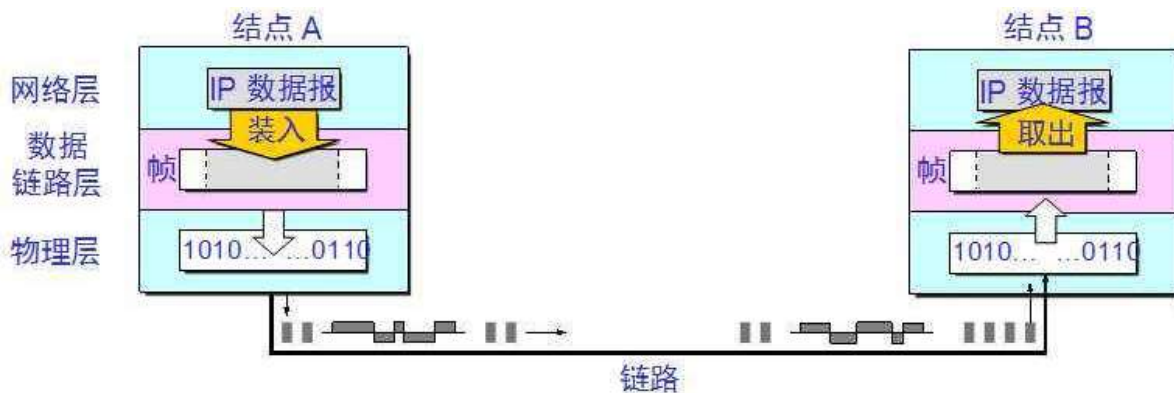
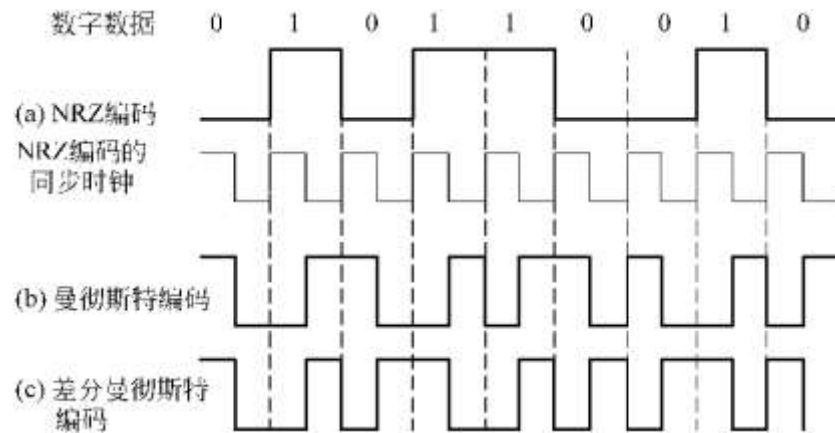
# Load/store访存需要多少个总线周期?





# 总线数据传输方式(物理层)

- 总线传输方式
  - 串行总线、并行总线
    - 传输一字节使用的**信道数**
  - 串行、并行传输
    - 串并转换问题
- 数据编码方式
- 数据传输模式
  - 单字节传输
  - 突发传输

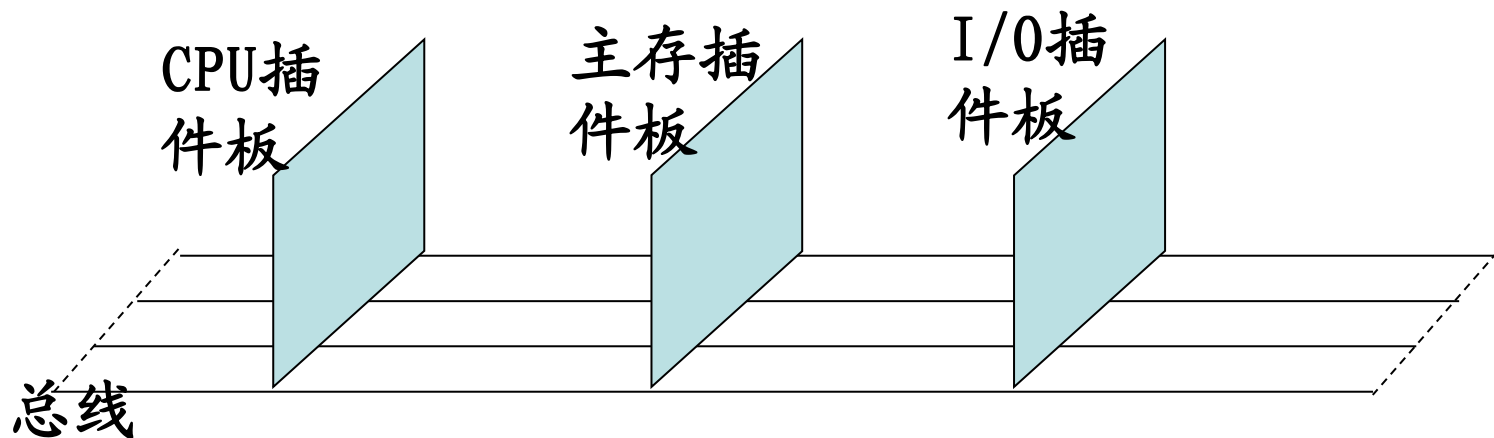






# 总线特性：标准

- 从物理角度来看，总线就是一组电导线。



- 为了保证机械上的可靠连接，必须规定其**机械特性**；
- 为了保证电气上正确连接，必须规定其**电气特性**；
- 为了保证正确地连接不同部件，还需规定其**功能特性**和**时间特性**。
  - 数据、地址、控制



# 总线性能指标

- **信号线数**：即地址总线、数据总线和控制总线三种总线的根数总和。
- **总线宽度**：是指数据总线的根数，用bit（位）表示，如8位、16位、32位、64位（也即8根、16根、32根、64根数据线）。
- **最大传输率（总线带宽）**：总线本身所能达到的最高传输速率，用MB/s（每秒多少兆字节）表示。
  - 例：总线工作频率33.3MHz，总线宽度32位，则最大传输率 =  $33.3 \times 32 / 8 = 132 \text{MB/s}$ 。
- **时钟同步/异步**：总线上的数据与时钟同步工作的总线称为同步总线，与时钟不同步工作的总线称为异步总线。
- **负载能力**：通常用可连接扩增电路板数来反映总线的负载能力。



# 例1

- **题目：**某总线在一个总线周期中并行传送4个字节的数据，假设一个**总线周期**等于一个**总线时钟周期**，总线时钟频率为33MHz，则**总线带宽**是多少？  
~~如果一个总线周期中并行传送64位数据，总线时~~

**提示：**此题主要是考查对总线带宽的理解。

$$\begin{aligned}\text{总线带宽} &= \text{一次传输的字节数} / \text{总线周期} \\ &= (\text{总线宽度} / 8) * \text{总线时钟频率}\end{aligned}$$

- **解：**设总线带宽用 $D_r$ 表示，总线时钟周期用 $T=1/f$ 表示，一个总线周期传送的数据量用 $D$ 表示，

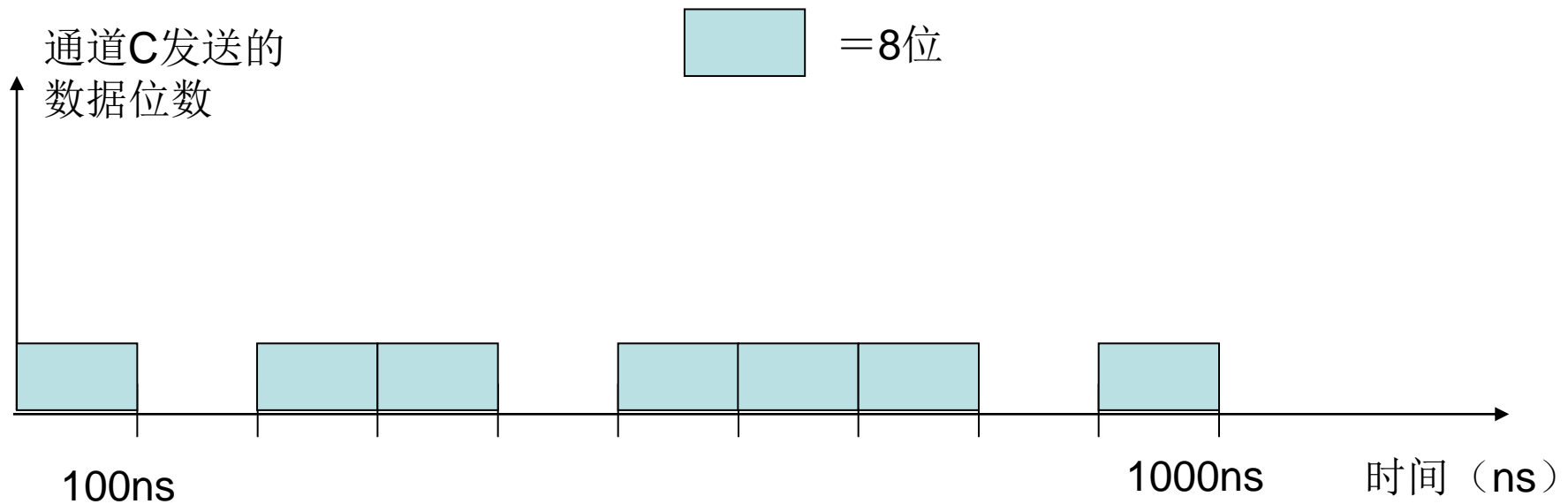
$$\begin{aligned}(1) \text{ 根据定义可得 } D_r &= D/T = D \times 1/T = D \times f \\ &= 4\text{B} \times 33 \times 1000000/\text{s} = 132\text{MB/s}\end{aligned}$$

$$\begin{aligned}(2) \text{ 64位} &= 8\text{B}, \quad D_r = D \times f \\ &= 8\text{B} \times 66 \times 1000000/\text{s} = 528\text{MB/s}\end{aligned}$$



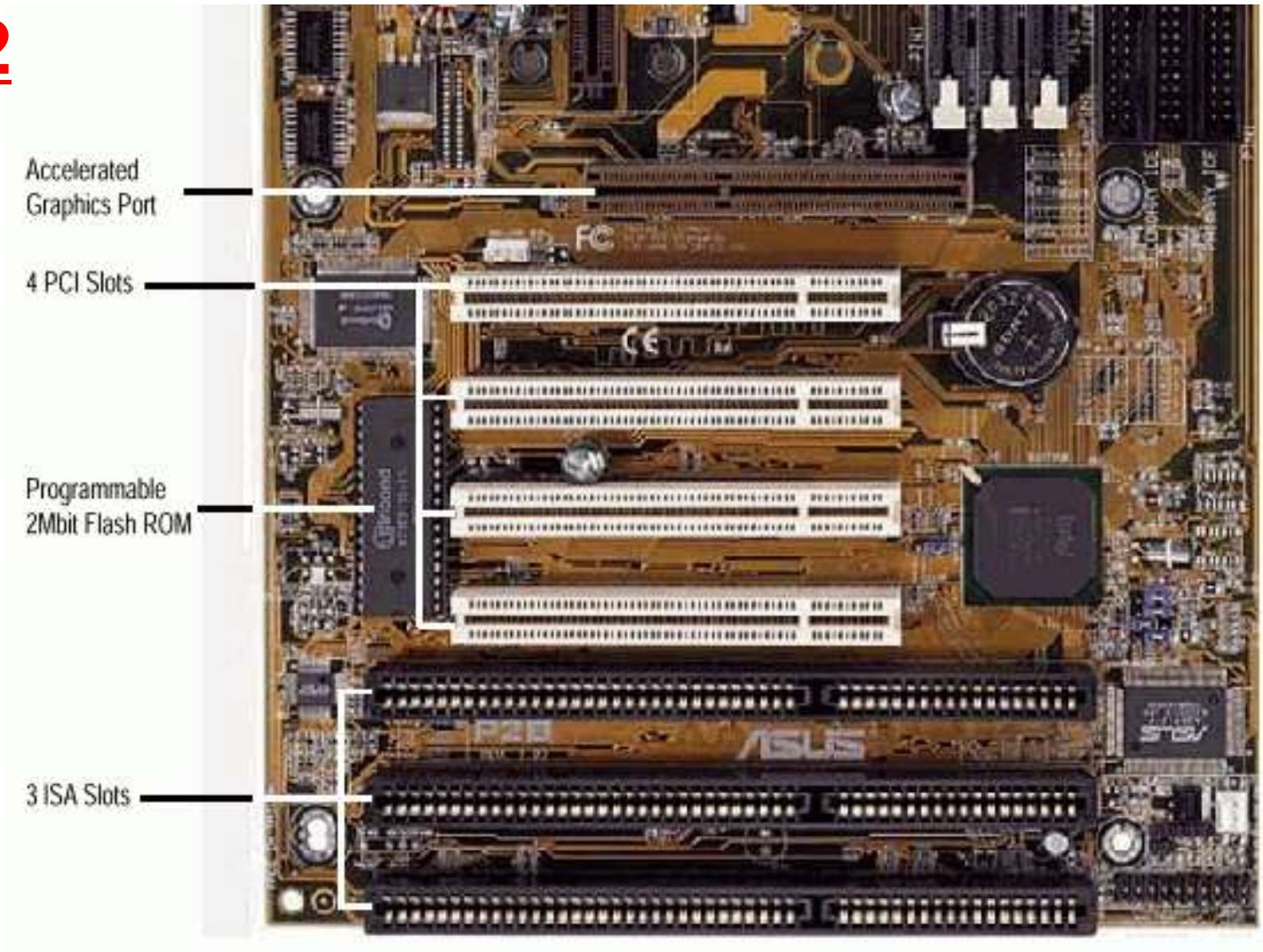
# 通信速度

- 平均速率 =  $56\text{bits}/1000\text{ns} = 56\text{Mb/s}$ 
  - avgrate
- 最大速率 =  $8\text{bits}/100\text{ns} = 80\text{Mb/s}$ 
  - peakrate: 包长/发送时间



# 目前流行的总线标准

- RS232
- ISA
- EISA
- STD
- PCI
- AGP
- SCSI
- USB
- .....







# PCI总线(外围部件互连)

- 特点：**不依附于**特定处理器，主从设备共享
  - 独立请求仲裁，半同步，半互锁，总线事务
  - Burst（cacheline为单位），单字为Burst的特例
  - pin configuration: 复用（Low Pin Count）
  - 对AD线的有效信息进行奇偶校验
  - **中断共享**（支持大量中断设备）
  - 即插即用(Plug and Play)

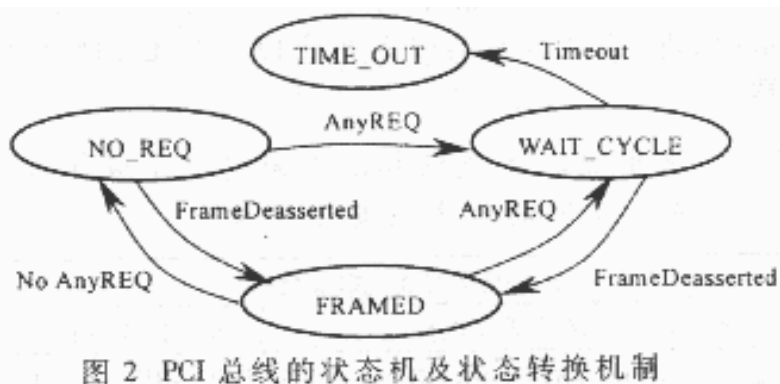
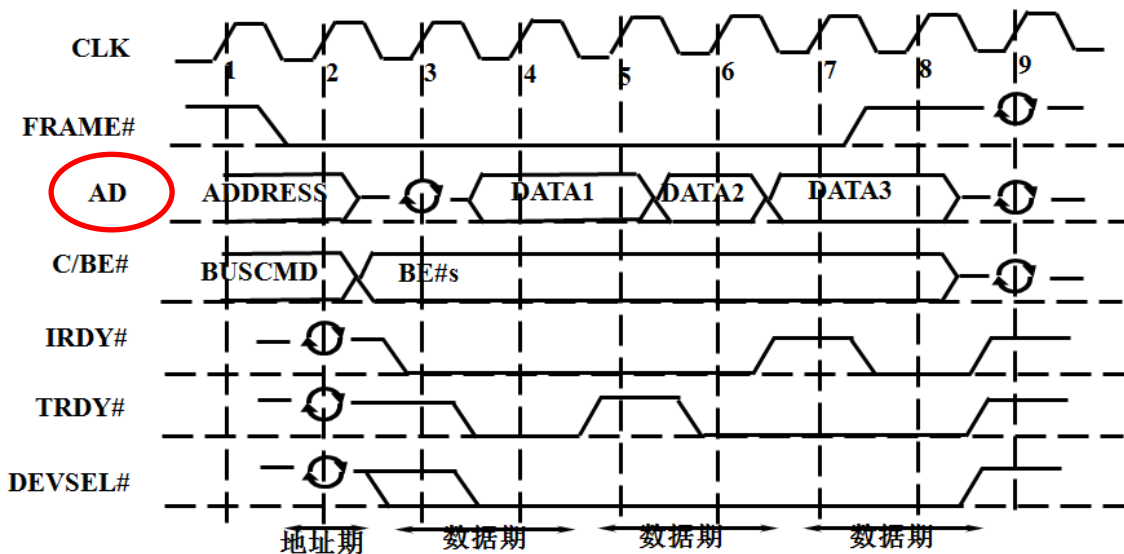


图 2 PCI 总线的状态机及状态转换机制





# PCI仲裁—独立请求仲裁

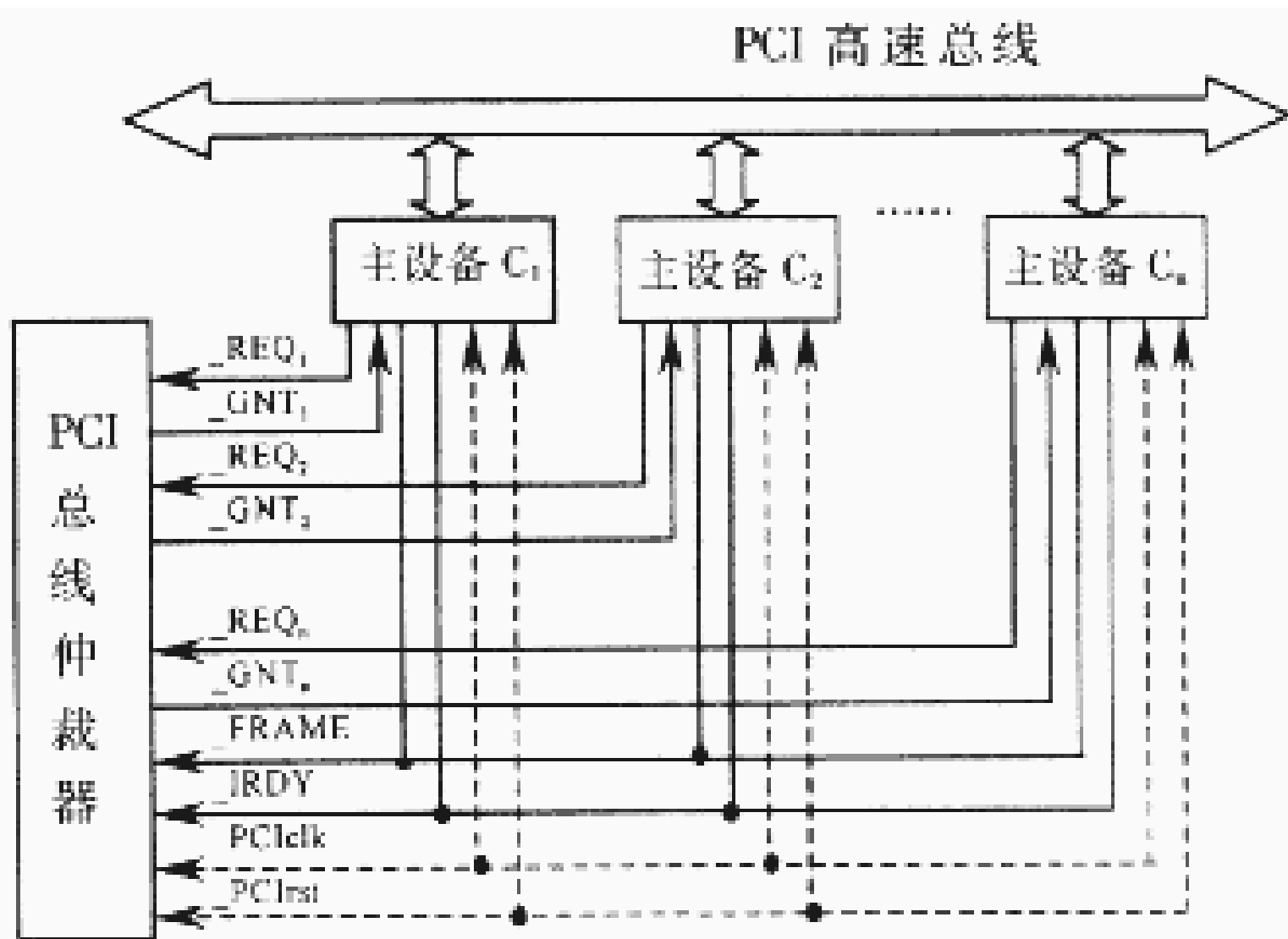
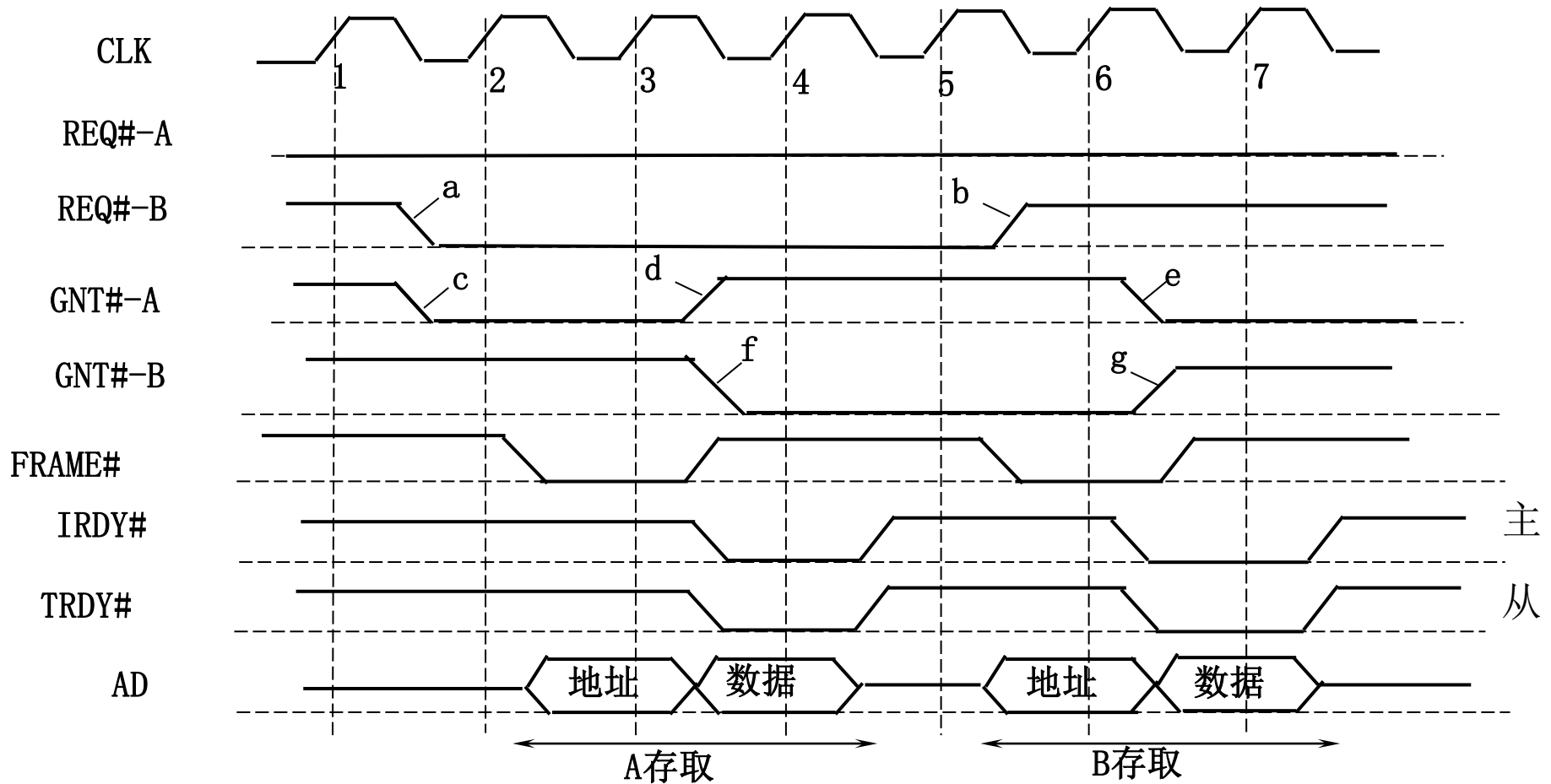


图 1 PCI 总线仲裁机制示意图



# PCI总线单字（节）数据传输模式



REQ-GNT仲裁：B优先级高于A

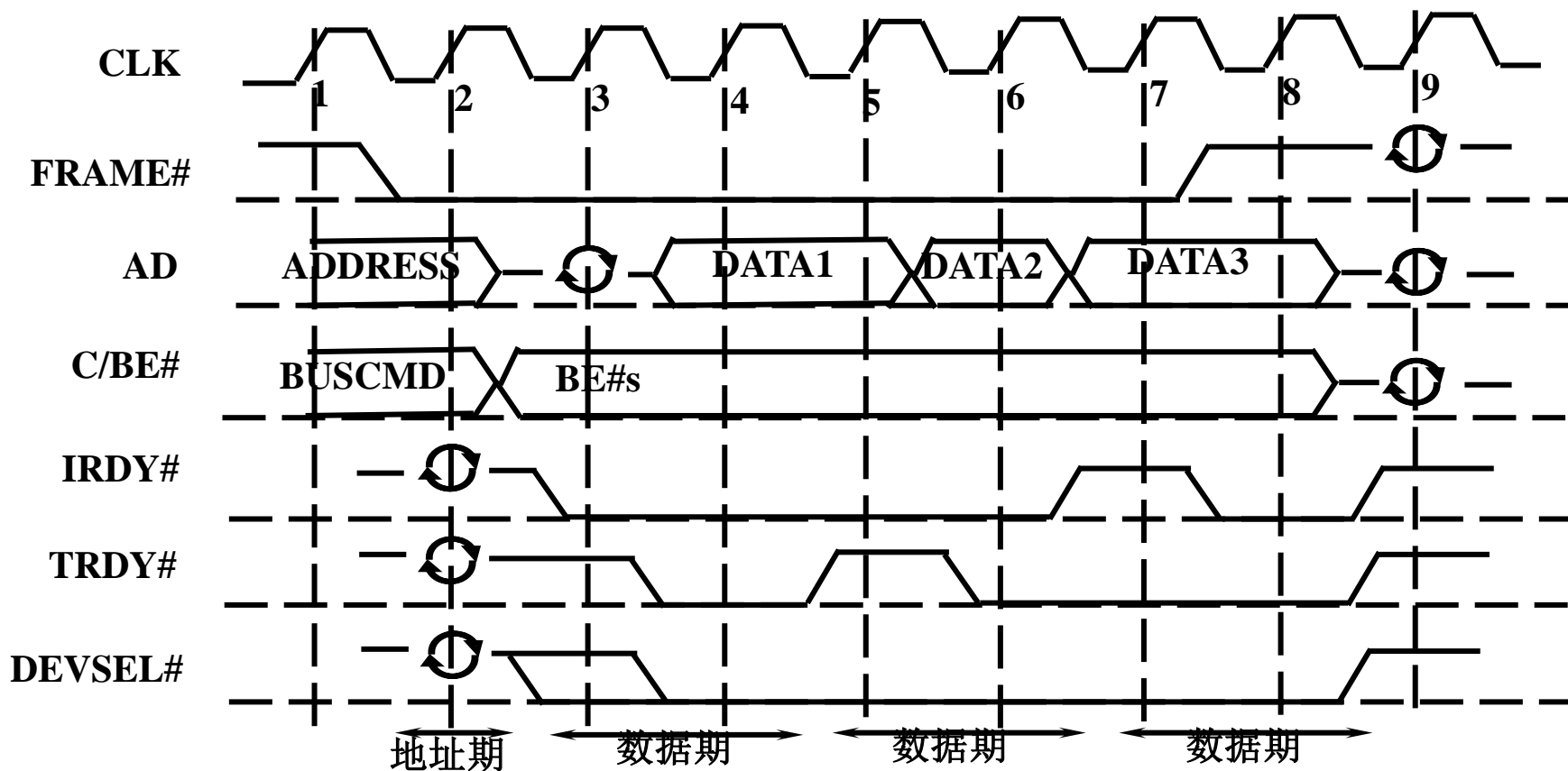
FRAME#：指示总线周期，由主设备控制

AD：地址/数据复用。

半同步：准备好——即主设备IRDY#与从设备TRDY#均有效。

等待——从设备可置TRDY#无效，通知主设备插入等待周期。

# PCI总线突发数据传输模式



**C/BE#:** 总线命令/字节允许

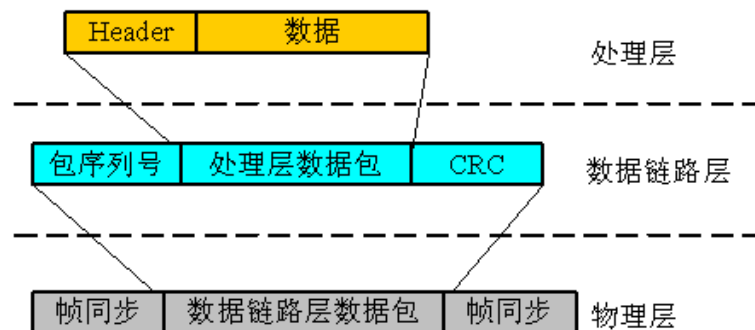
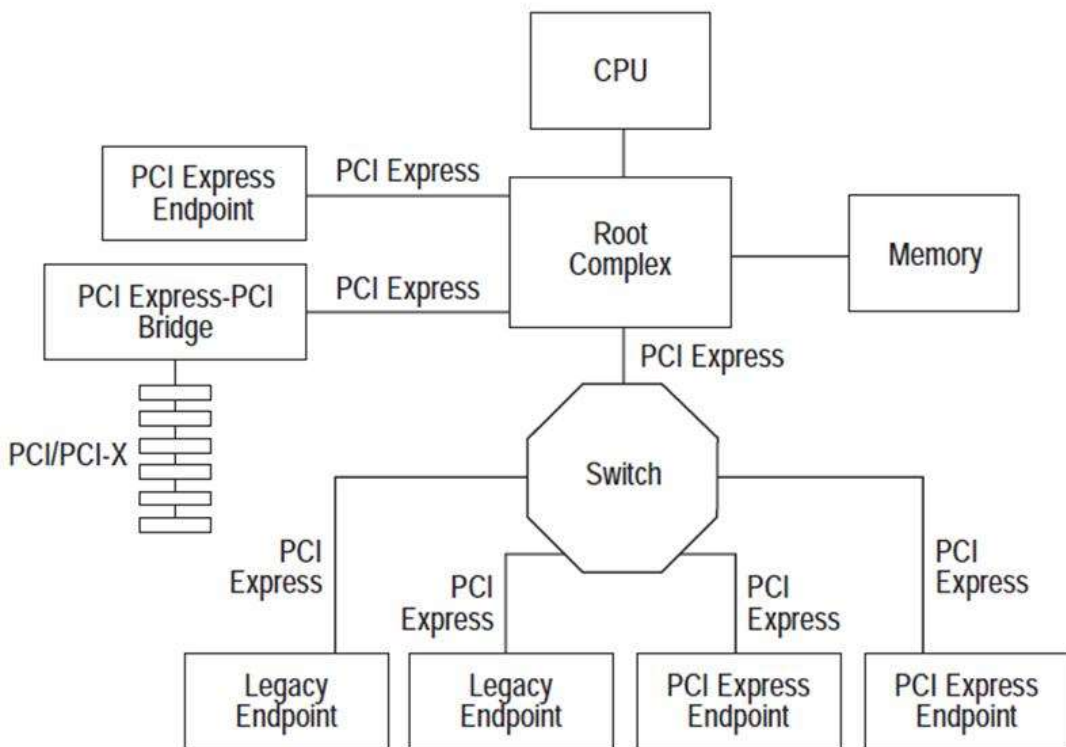
地址期指示总线周期类型，如内存读写、I/O读写、中断确认

数据期指示各字节有效与否

**FRAME#:** 指示总线周期，长度无限制，由主设备控制

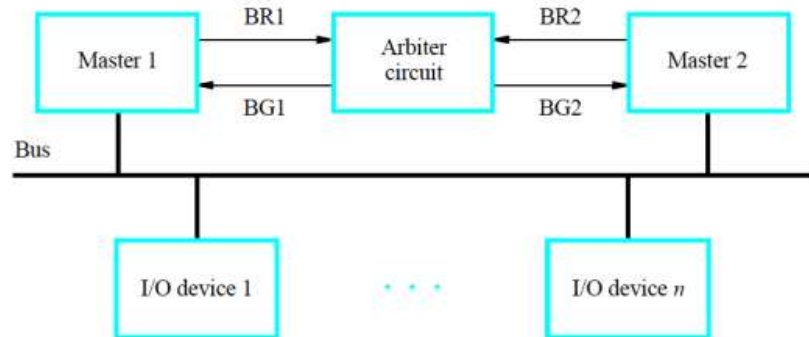
**DEVSEL#:** 经地址和命令字译码，通知主设备从设备被选中，确认总线周期开始

# PCIE (PCI Express, intel, 2004)



- 采用星型拓扑 (点对点)
  - 根组件Root Complex, RC
  - 交换器Switch, SW
  - 端设备, EP, <总线号, 设备号>
    - 可为多功能设备: 功能号
- 串行总线: 4线 (两收两发)
  - 第一级RC所在的层级, 定义为 Bus0
  - 与PCI总线的寻址方式兼容
- 事务层、数据链路层和物理层
  - 数据报

# 总线驱动程序

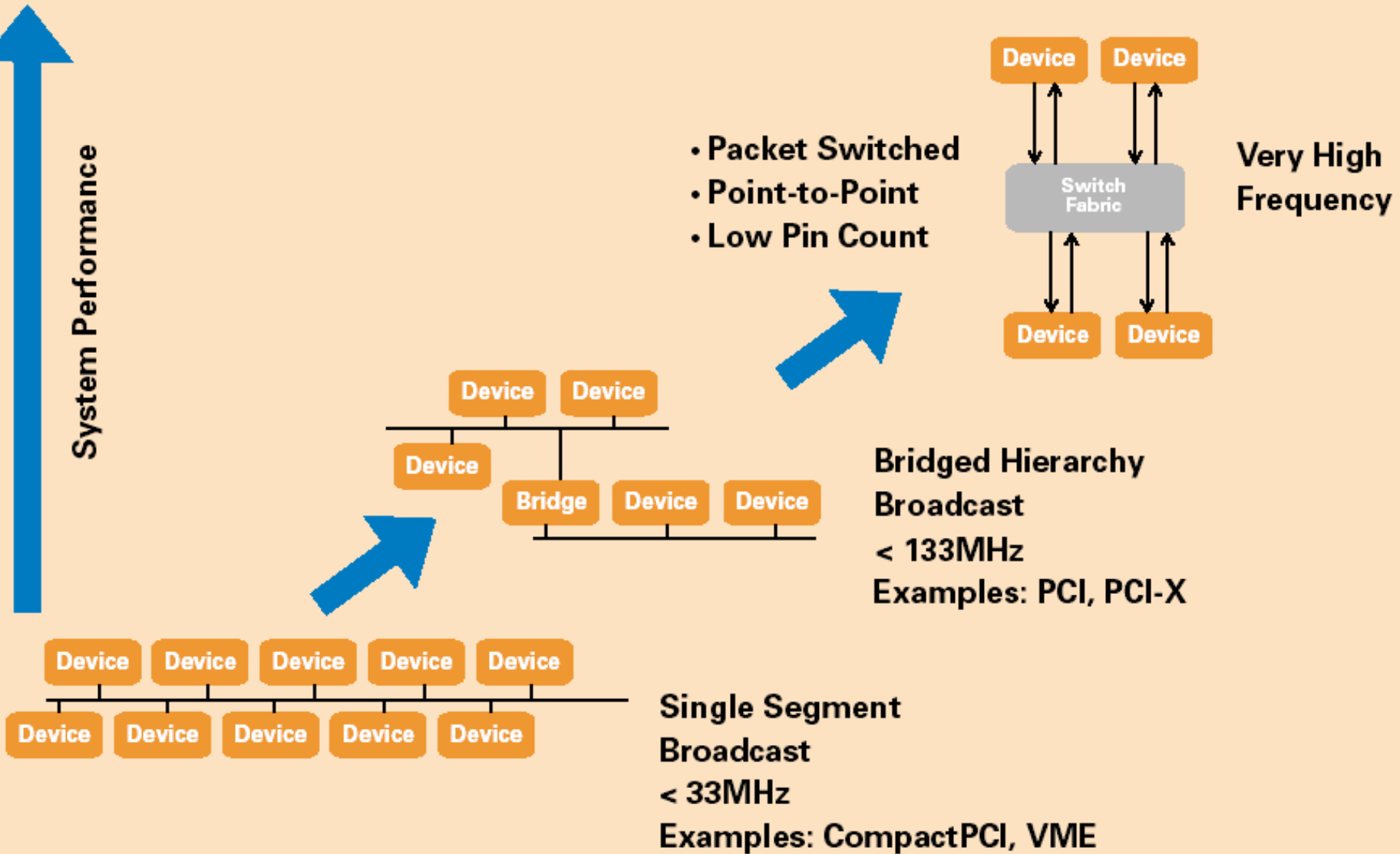


- 设备抽象
  - 代表总线上的设备执行特定的操作
    - 是**总线控制器**、适配器等**的设备驱动程序**
    - 注意：各类外设的读写请求由各自的**设备驱动程序**处理！
- 主要任务（例：**PCI总线**）
  - 枚举其总线上的设备
  - 响应 **PnP** 和电源管理**IRP**（**I/O请求包**）
  - 总线的多路访问（对某些总线）
  - 管理其总线上的设备
- 典型
  - **PCI, PnP ISA, SCSI, USB**

# 总线系统及其发展趋势：分段



System Performance

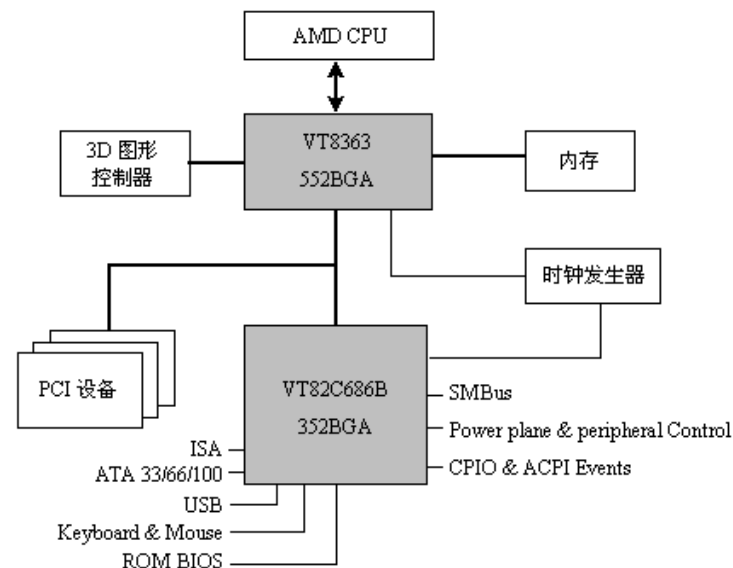






# 小结

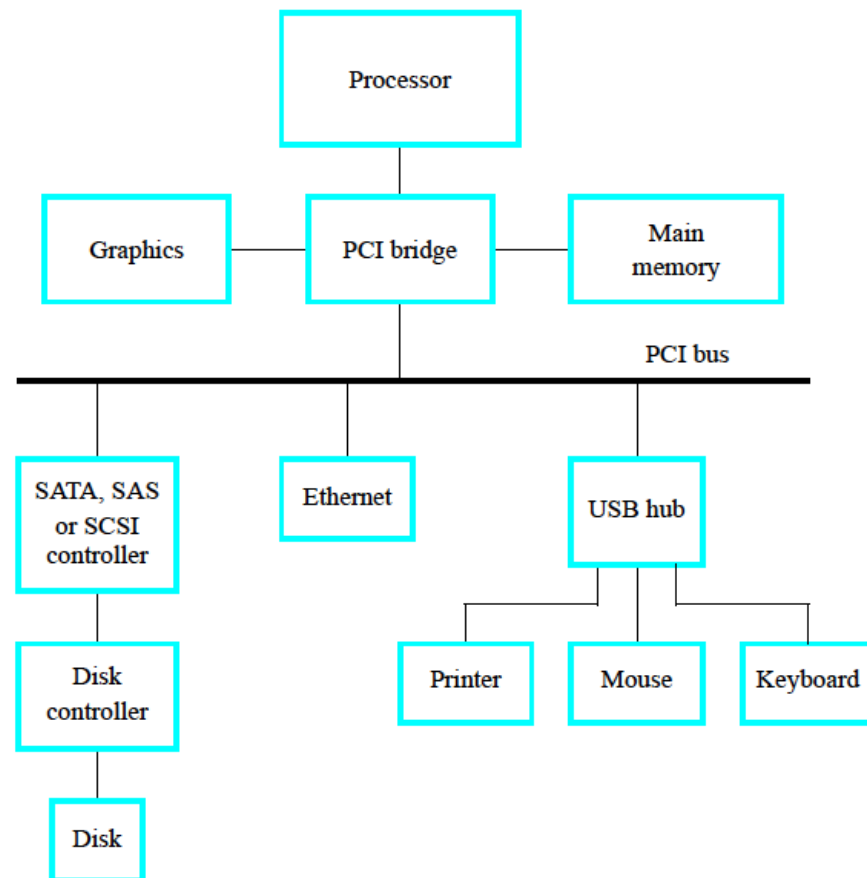
- 总线分类、特性与性能指标，拓扑结构
- 总线通信控制
  - 传输过程
  - 同步控制：时序图
    - 总线时钟宽度？
- 总线数据传输
  - 串并行方式，编码方式
  - 数据传输模式：单字节、突发式
- 总线控制器：功能？如何实现？
  - 总线仲裁
- 软硬件接口：总线驱动程序？
- 作业
  - 唐2： 3.4、3.8、3.12（选）、3.14、3.16



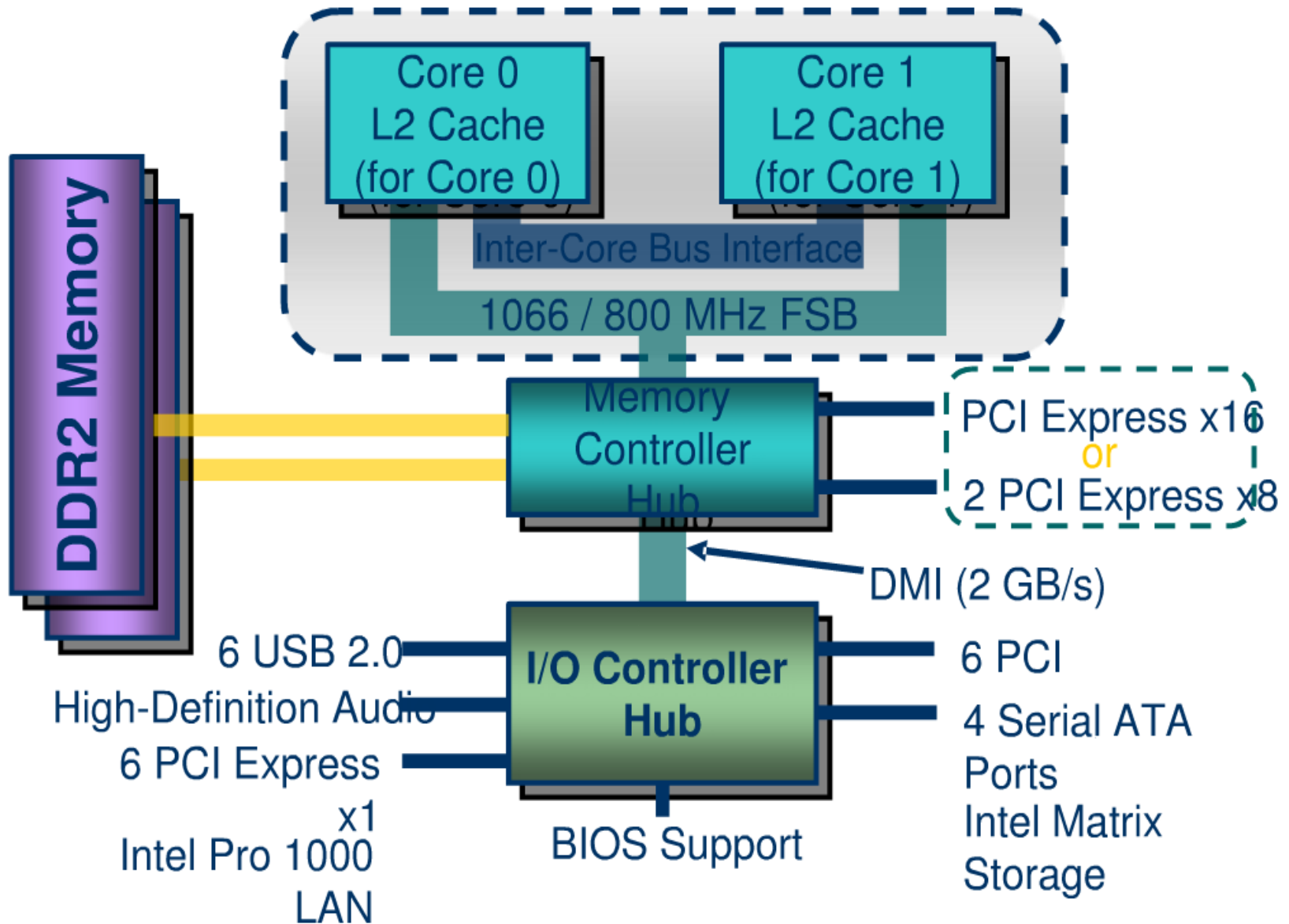


Thank you

# 总线拓扑结构：多总线



# 计算机系统中的总线



# PCI总线

- 一次一对设备使用总线
  - 不支持广播式？
  - 半同步方式： **IRDY#**及**TRDY#**
  - 主设备启动总线周期
    - 发出**FRAME#**信号，表明一次访问（总线周期）开始
  - 单字模式，突发模式
  - 对**AD**线的有效信息进行奇偶校验
- 集中式仲裁
  - 隐式仲裁：仲裁可在上一次访问期间完成，不占总线周期
    - **\_GNT**被撤消而**\_FRAM**有效，指示当前总线访问合法，并将继续
  - 超时回收：占用者空闲**16**个时钟，则可回收
  - 缺省占用：无请求时将使用权“停靠”于某设备
  - 调度策略：基于访问（等待时间小），而不是基于时间片
    - 仲裁器仲裁算法与**PCI**规范无关！
      - 按优先级排序：固定优先级，循环优先级

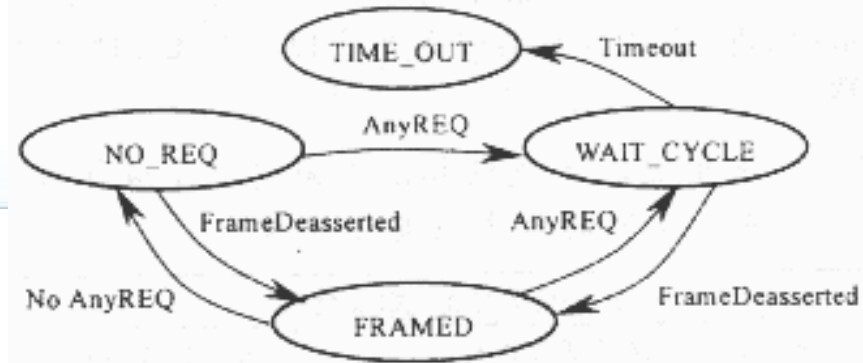


图 2 PCI 总线的状态机及状态转换机制



# PCI配置空间

- PCIe设备有三个独立的物理地址空间
  - 设备存储器空间（memory）、IO空间和配置空间（config）。

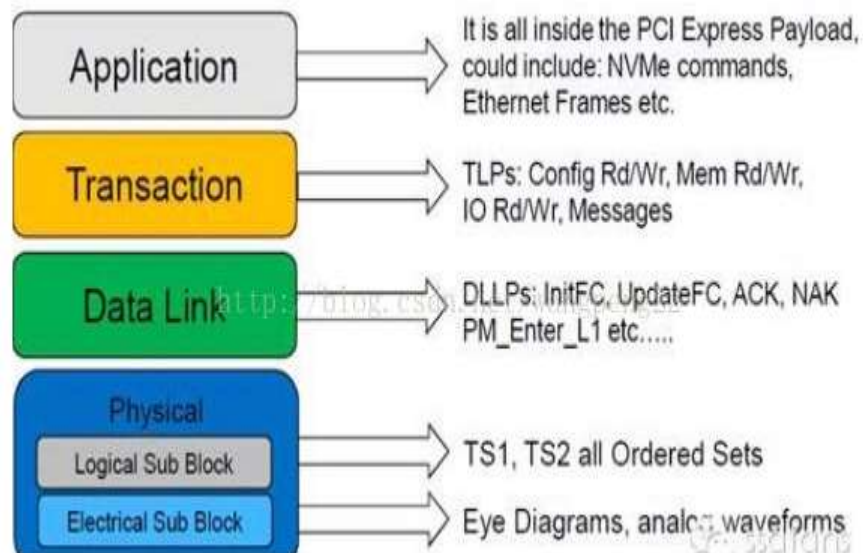
24	23	16	15	8	7	0	
Device ID				Vendor ID			0x00
Status				Command			0x04
Class Code					Revision ID		0x08
BIST	Header Type		Latency Timer	Cache Line Size		0x0C	
Base Address Register 0							0x10
Base Address Register 1							0x14
Base Address Register 2							0x18
Base Address Register 3							0x1C
Base Address Register 4							0x20
Base Address Register 5							0x24
Cardbus CIS Pointer							0x28
Subsystem ID				Subsystem Vendor ID			0x2C
Expansion ROMBase Address							0x30
Reserved					Capabilities Pointer		0x34
Reserved							0x38
MAX_Lat	Min_Gnt		Interrupt Pin	Interrupt Line		0x3C	



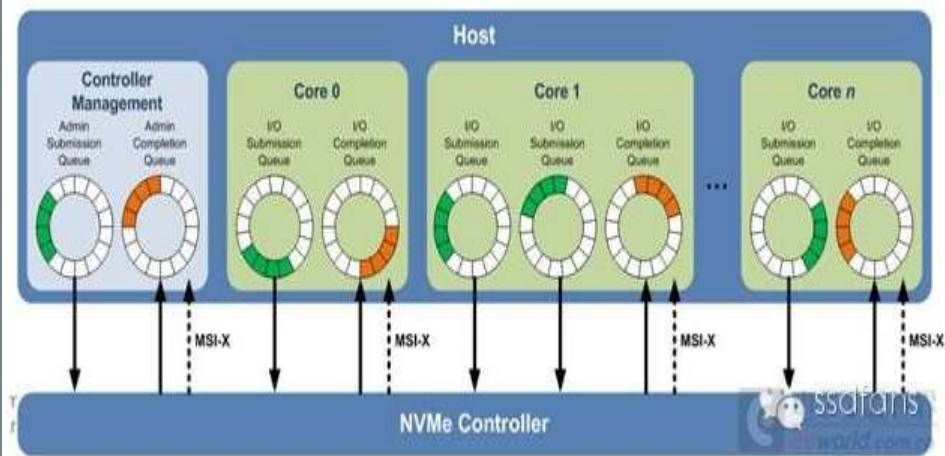
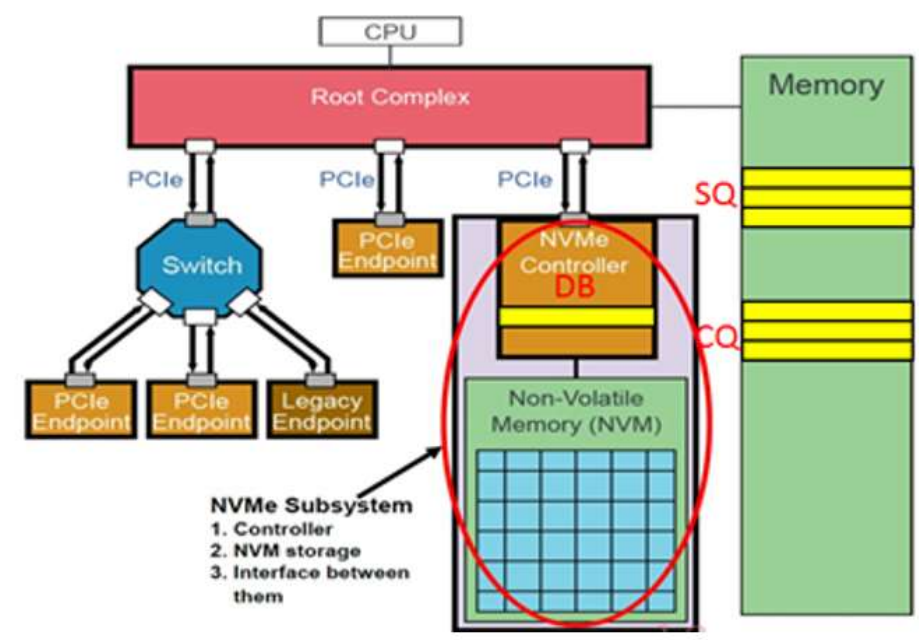


# PCIe设备访问

- TLP(Transaction Layer Packet)
- 在MSI中断方式下，设备通过向OS预先分配的主存空间写入特定数据的方式请求CPU的中断服务，为PCIe系统首选的中断信号机制



# 例：SSD与PCIE



- 基于NVMe(Non-Volatile Memory express)
  - 一种Host与SSD之间通讯的协议
- 使用MSI-X中断机制

# PCIe带宽



## PCI Express Example Connectors

**x1**

**BANDWIDTH**

Single direction: 2.5 Gbps/200 MBps  
Dual Directions: 5 Gbps/400 MBps



**x4**

**BANDWIDTH**

Single direction: 10 Gbps/800 MBps  
Dual Directions: 20 Gbps/1.6 GBps



**x8**

**BANDWIDTH** Single direction: 20 Gbps/1.6 GBps  
Dual Directions: 40 Gbps/3.2 GBps

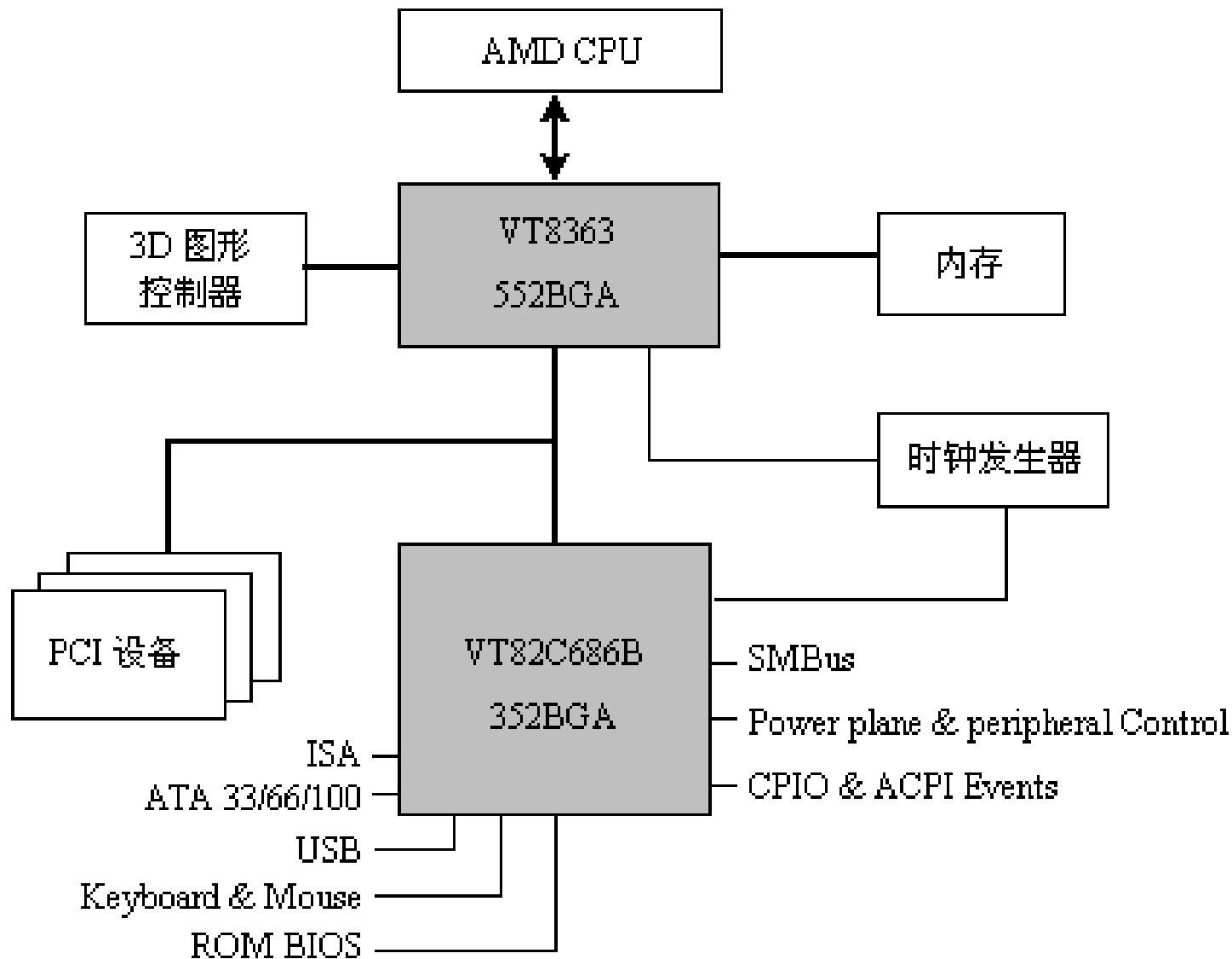


**x16**

**BANDWIDTH** Single direction: 40 Gbps/3.2 GBps  
Dual Directions: 80 Gbps/6.4 GBps



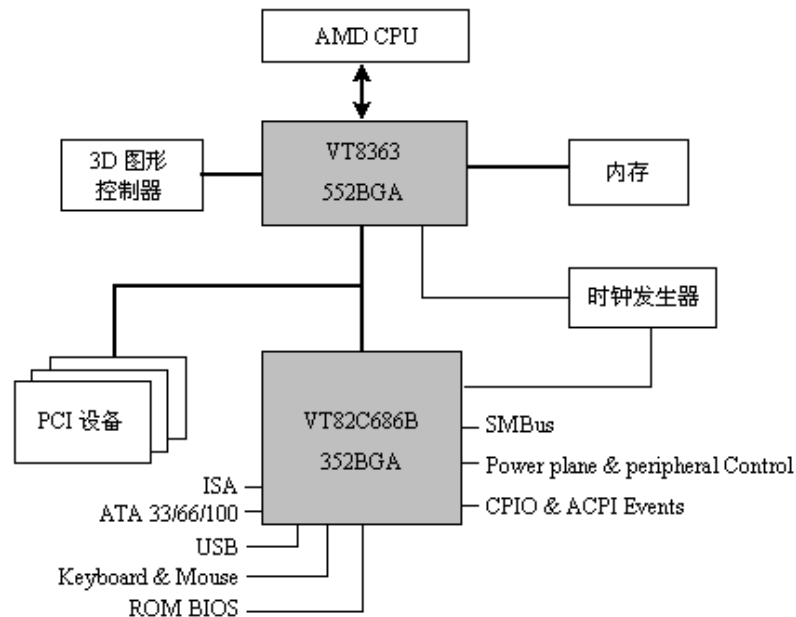
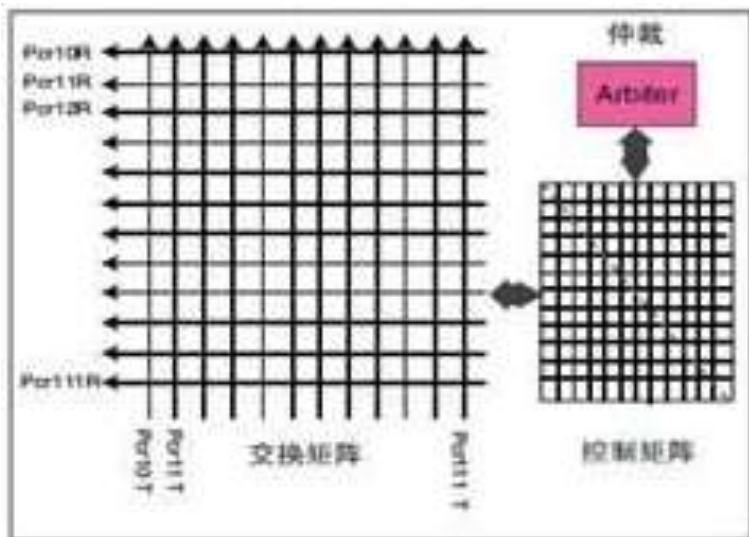
# 南北桥结构的KT133芯片组



# 前端总线（Front Side Bus, FSB）



- 将CPU连接到北桥芯片的总线
  - 基于PACKET的**网络式**通信方式
- 最新的INTEL和AMD多核处理器都不用FSB了
  - 官方名字叫做CSI（Common System Interface）
    - 用来实现芯片之间的直接互联，而不是在通过FSB连接到北桥
  - INTEL技术称QPI（快速通道互联，QuickPath Interconnect）
    - 无论是速度、带宽、每个针脚的带宽、功耗等都超越HT总线。
  - AMD称HT（HyperTransport）



# 总线事务 (Bus transaction)



- 在一个总线周期中发生的一系列活动
  - 典型的总线事务：请求操作、裁决操作、地址传输、数据传输和总线释放。
- Pentium Pro处理器总线事务类型有11种。
  - 延迟回答：当一个从设备需要花很长时间才能完成某个事务时，就通过发出该“延迟回答”事务，以“分离事务”方式来处理该事务。
  - 中断响应：当处理器响应从8259中断控制器送来的中断请求而要去读中断向量时。
  - 特殊事务：当处理器要广播一条与某个内部事件（如 shutdown、halt等）有关的消息时。
  - 分支跟踪消息：在指令执行过程中当一个转移目标的条件被满足时。它将送出转移指令的地址和转移到的目标指令的地址。



# Pentium Pro总线事务（续）



- 存储器读并无效：处理器产生该“存储器读并无效”事务，主要是为了对一个**Cache**行进行独占访问。
- 存储器代码读：当处理器要从存储器中取指令时。
- 存储器数据读：当执行到某指令要从存储器中取数据时。
- 存储器写（不可重试）：当处理器要写回一个更新的**Cache**行到存储器（淘汰）以便为一个新行腾出空间时，由处理器发出该事务。
- 存储器写（可重试）：当处理器执行到某指令需写数据到存储器中去时，由处理器发出该事务。
- I/O读：当处理器执行**IN**或**INS**指令，要从某个I/O设备读取数据或状态时。
- I/O写：当处理器执行**OUT**或**OUTS**指令，要将数据或命令写到某个I/O设备上时。



# 总线事务的操作阶段

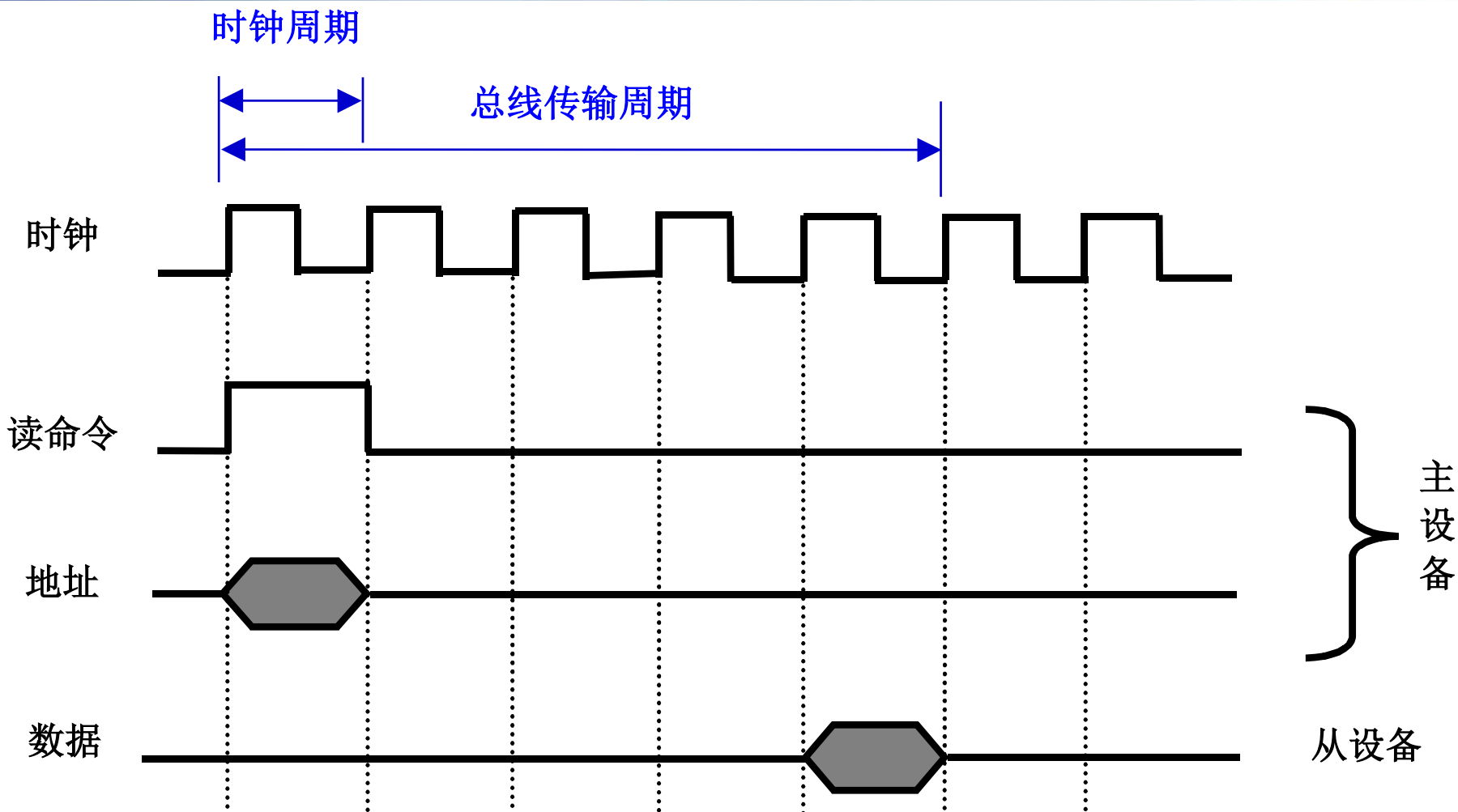
## ◆ Pentium Pro每个总线事务包含五个操作阶段

- 请求阶段（地址阶段）
  - 送出地址信息和事务类型信息，以及有关事务的其他信息。
- 检错阶段
  - 对请求阶段送出的地址和请求信息所对应的奇偶校验信号进行检测。
- 侦听阶段
  - 对请求阶段送出地址中所对应的数据，检查其在各个**Cache**中的命中状态，确定以后阶段如何处理该事务。
- 响应阶段
  - 根据请求的事务类型、检错和侦听结果，确定如何响应当前事务。
    - 返回的响应结果可以是重试事务、延迟事务、硬件错、无数据传送事务、回写（**Writeback**）事务或正常数据传送事务等。
- 数据阶段
  - 送数据到数据总线上，或从数据总线上取数据

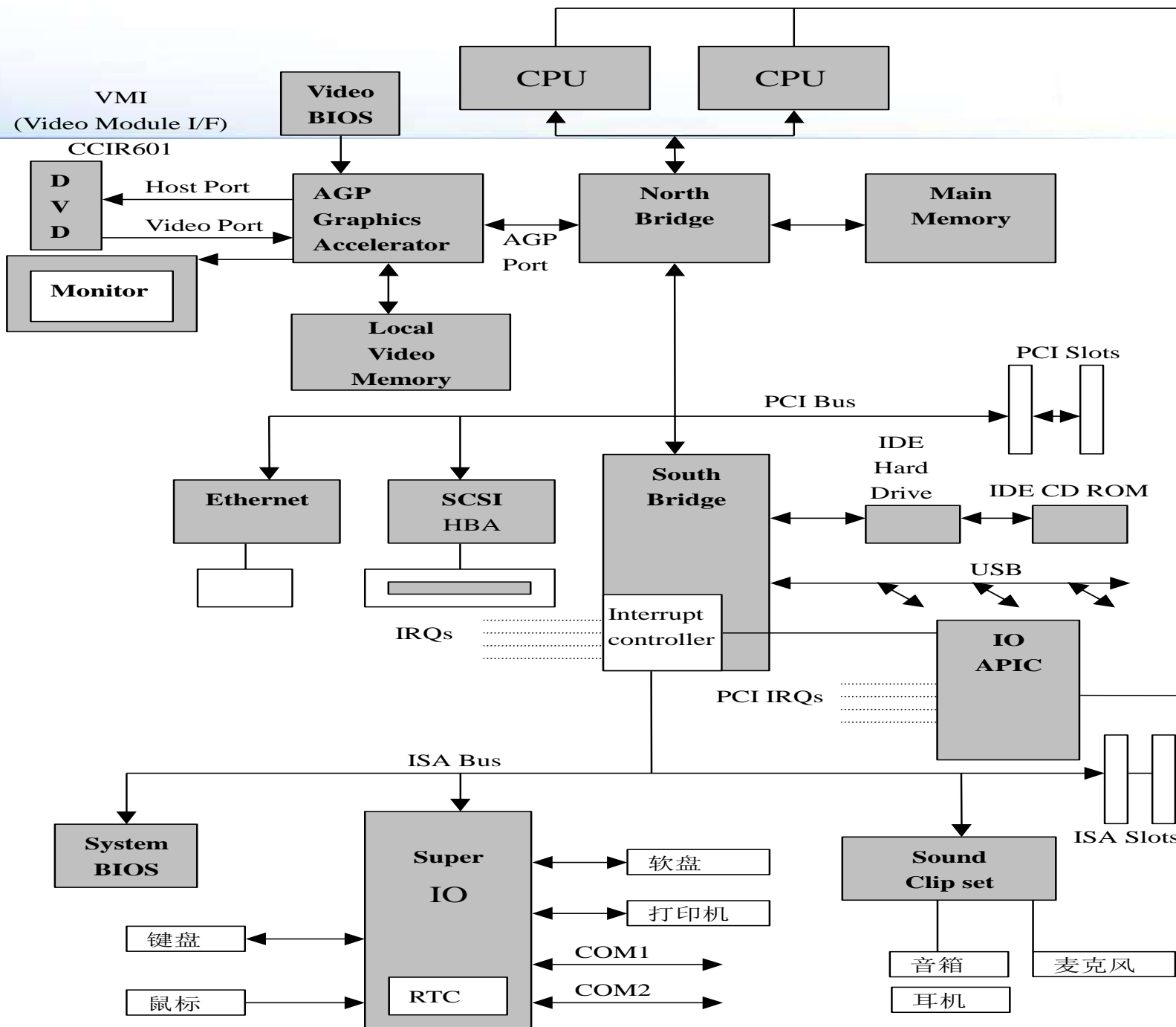




# 总线周期、时钟周期



同步通信协议（读操作）



多处理器系统的PCI配置

# 下一代总线的发展状况

- 基于IA体系的总线也在向着多处理器、跨系统、平滑网络与计算机系统间的连接的方向发展

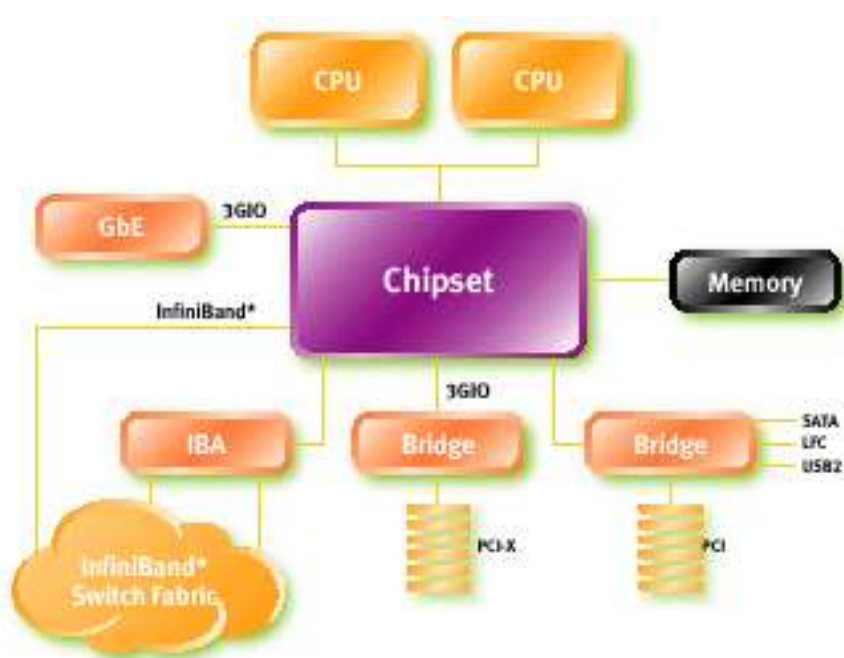


Figure 5. 3GIO-based Server/Workstation System.

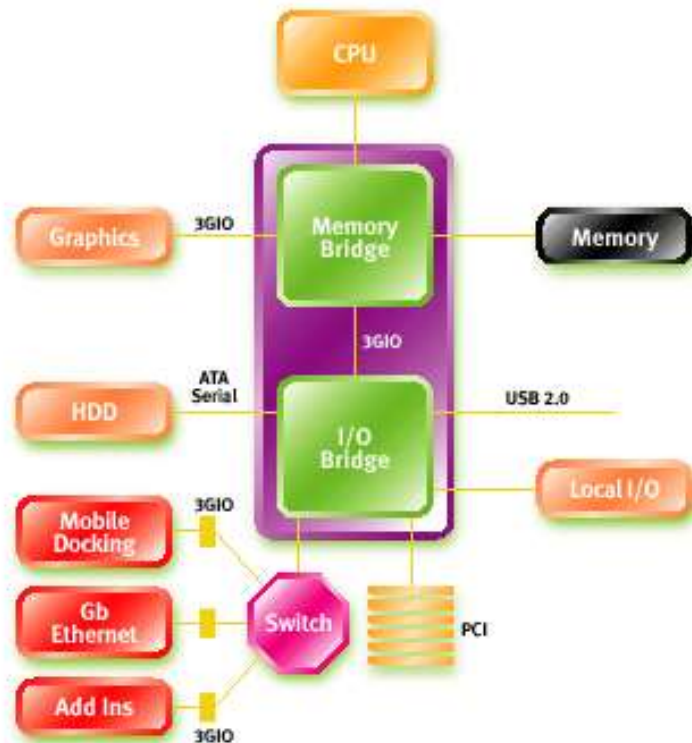
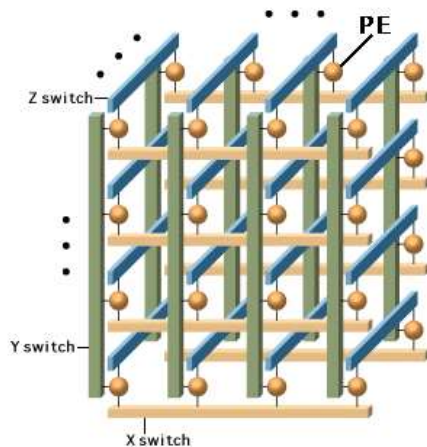


Figure 4. General purpose desktop/mobile I/O interconnect for 2003 and beyond.

# 3GIO总线



- 3GIO总线（Arapahoe总线）是一个双向对称总线，它具有高达2.5 GB/s的数据传输率。
  - 采用Host主桥和交换器。具有可变带宽。支持32-或者 64-bit寻址方式。
- NGIO总线改变了CPU传输数据的方式
  - 在CPU和外部设备之间不进行同步数据传输，而是将信息打成数据包在目标通道适配器和主通道适配器之间发送。
  - 这种异步通讯可将CPU从对速度较慢的外围设备数据的等待中解放出来
    - 这在多处理器系统中尤为重要，因为多处理器系统中各CPU间要为使用较慢的外围总线而展开竞争。





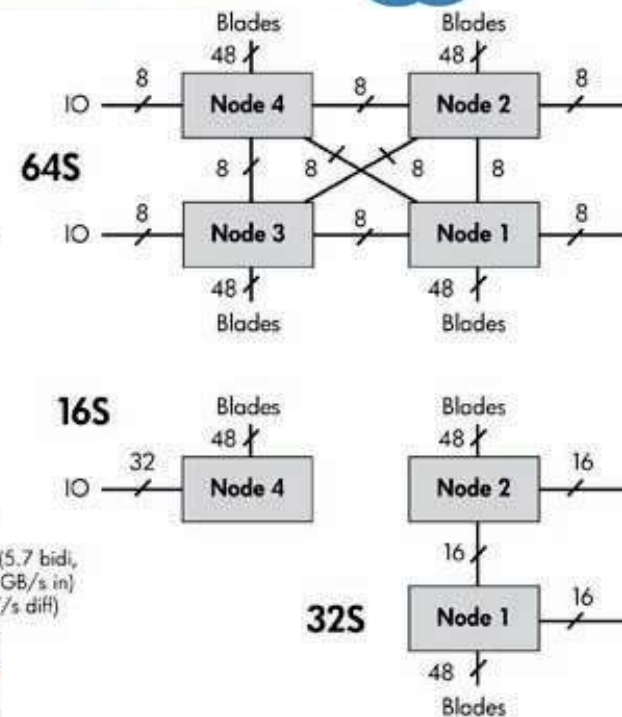
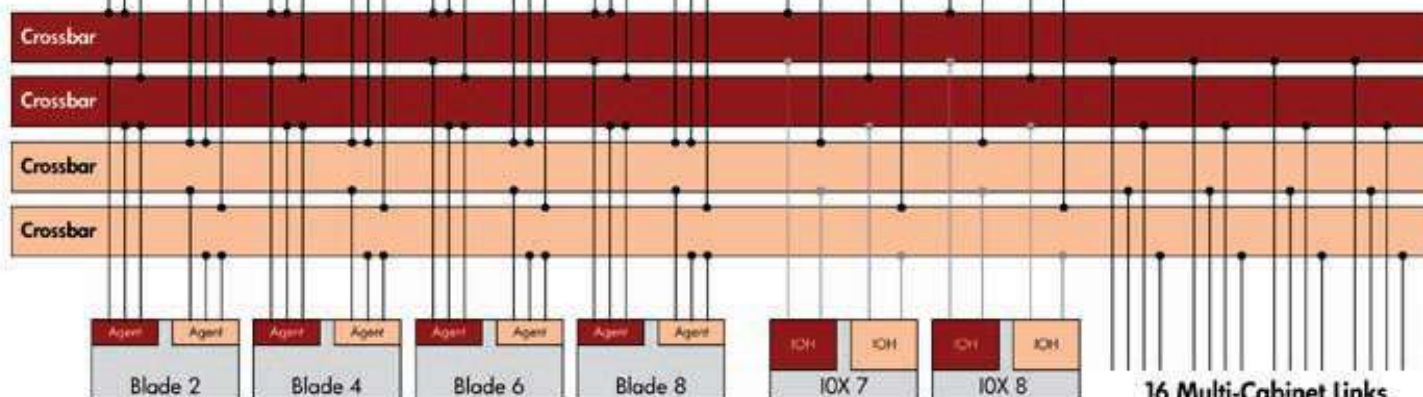
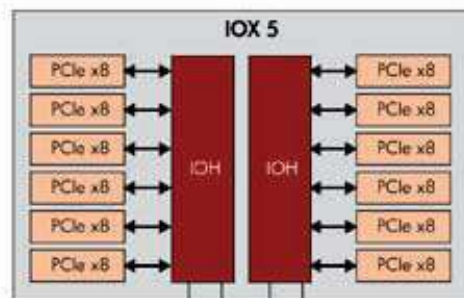
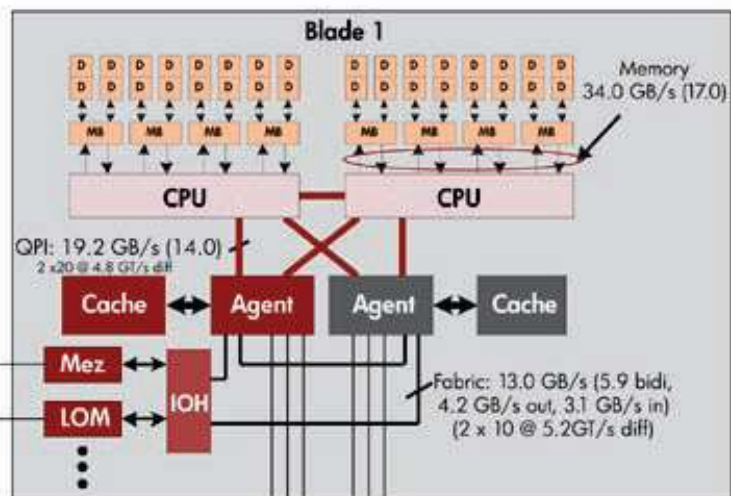
# Backplane Bus

- 刀片服务器是一种低成本服务器平台（机箱）
  - 每一块“刀片”是一块系统主板，配置了CPU、内存、磁盘和网卡等设备
  - 每一个主板运行自己的系统，服务于指定的不同用户群
  - 可以通过刀片服务器中集成的交换“背板”形成星形连接网络，实现与其他“刀片”交换。
    - **Compact PCI**: 同时支持多达256个的标准PCI总线设备
  - 专门为特殊应用行业和高密度计算环境设计





# Crossbar



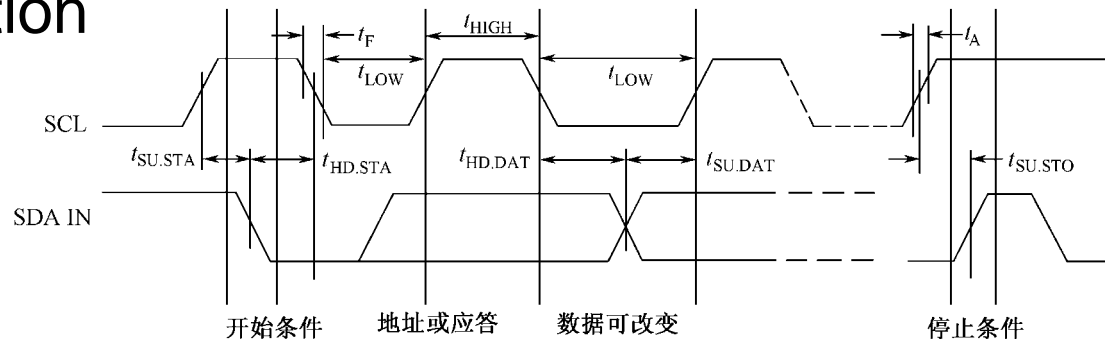
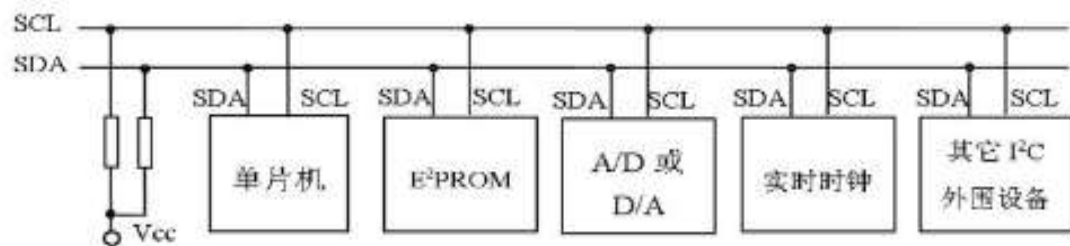
Grey links are used for multi-cabinet on 64S

These links can be used for I/O on 16S

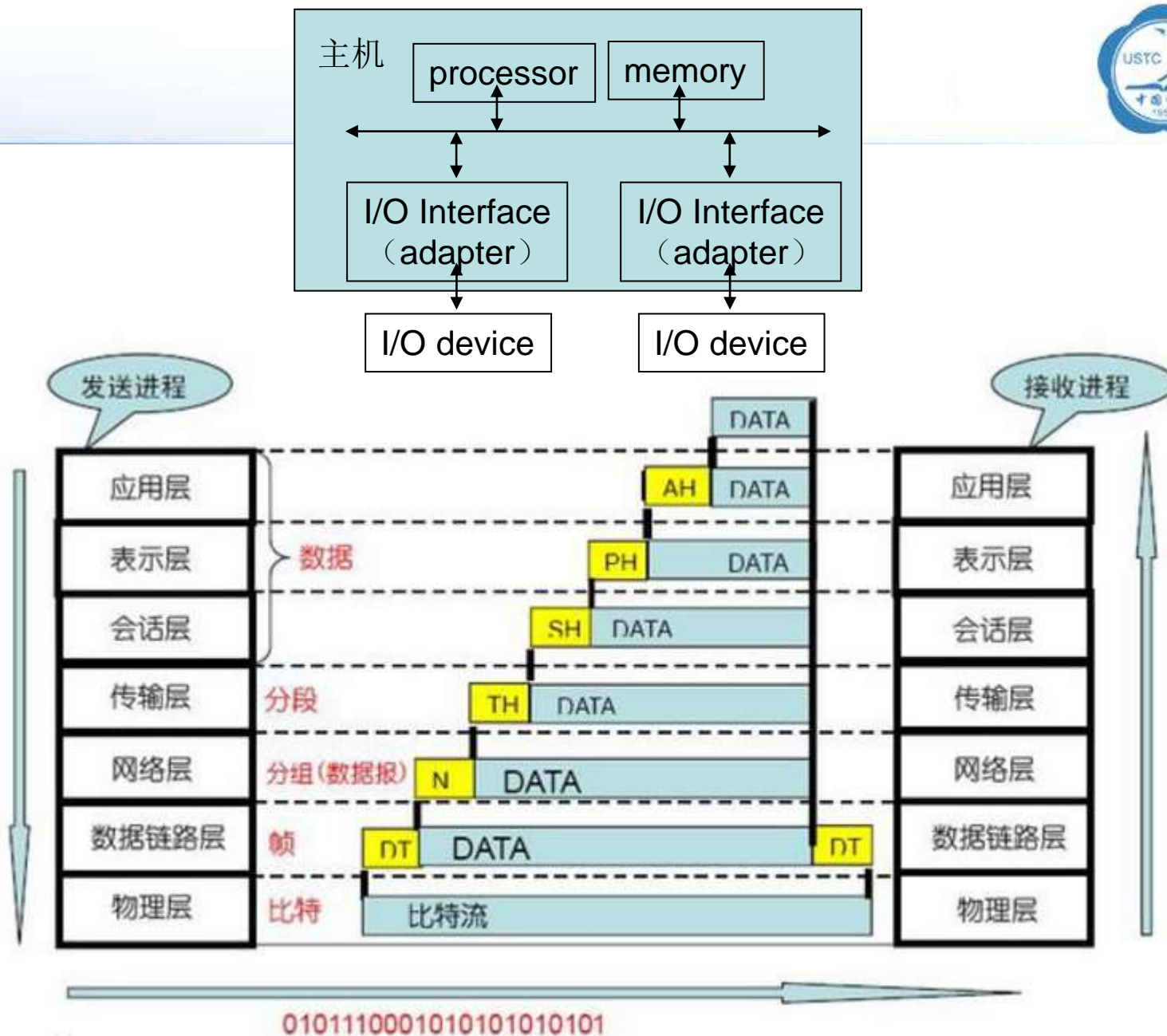


# I2C(Inter-IC)总线

- I2C(Inter-Integrated Circuit): 两线式**串行总线**
  - 双向数据线Serial Data(SDA)、时钟线Serial Clock(SCL)
  - 允许不同IC以不同波特率进行通信
  - 并联在一条总线上的每个IC都有唯一的地址
- 总线仲裁: 基于“线与”逻辑, “低电平优先”
  - SCL线时钟同步
    - 由主设备发出
  - SDA线仲裁
- 数据传输
  - 开始信号start condition
  - 结束信号stop condition
  - 应答信号
    - 一个Byte一个ACK



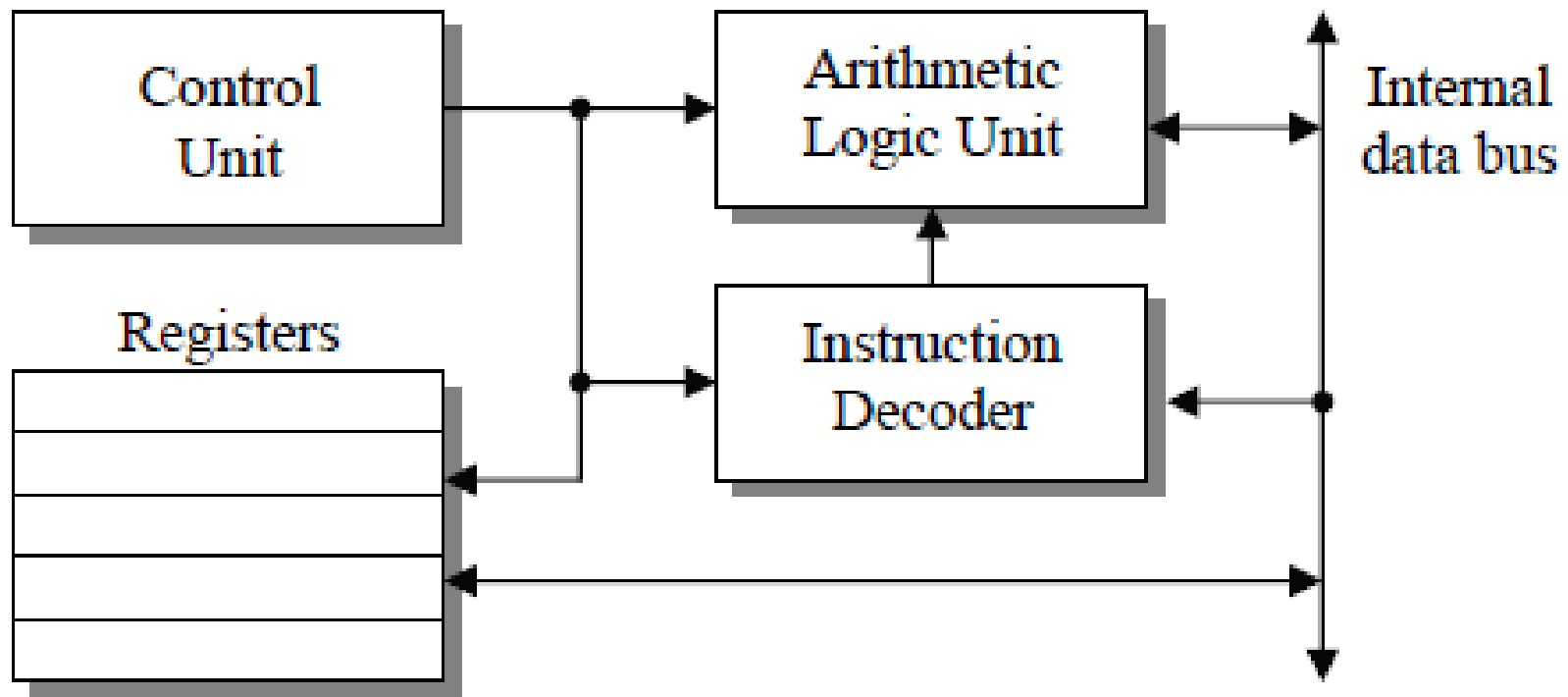




# Generic Block Diagram of a CPU



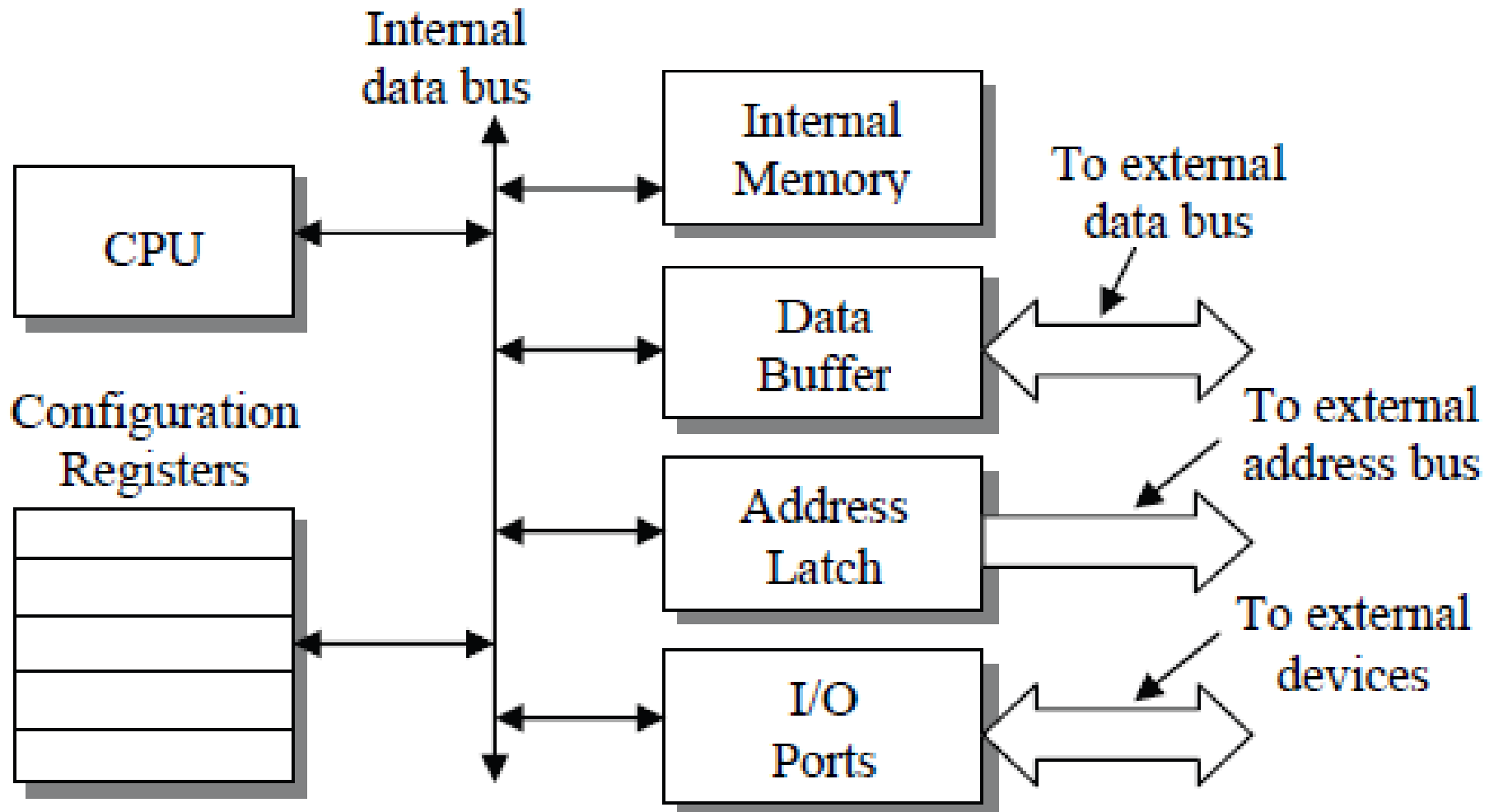
- 片内总线



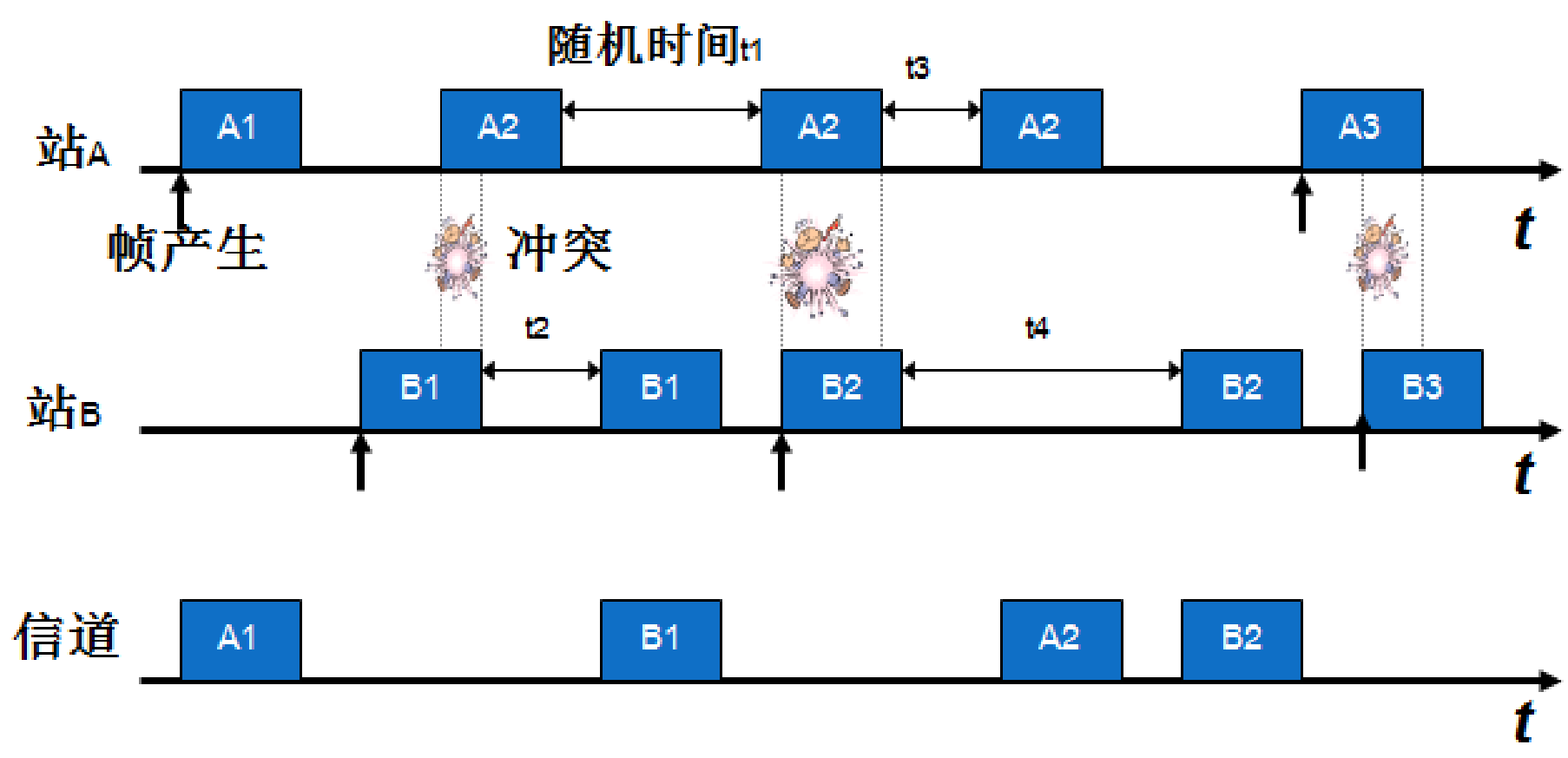


# Generic Block Diagram of a Processor

- 片外总线：内部总线，外部（外围）总线



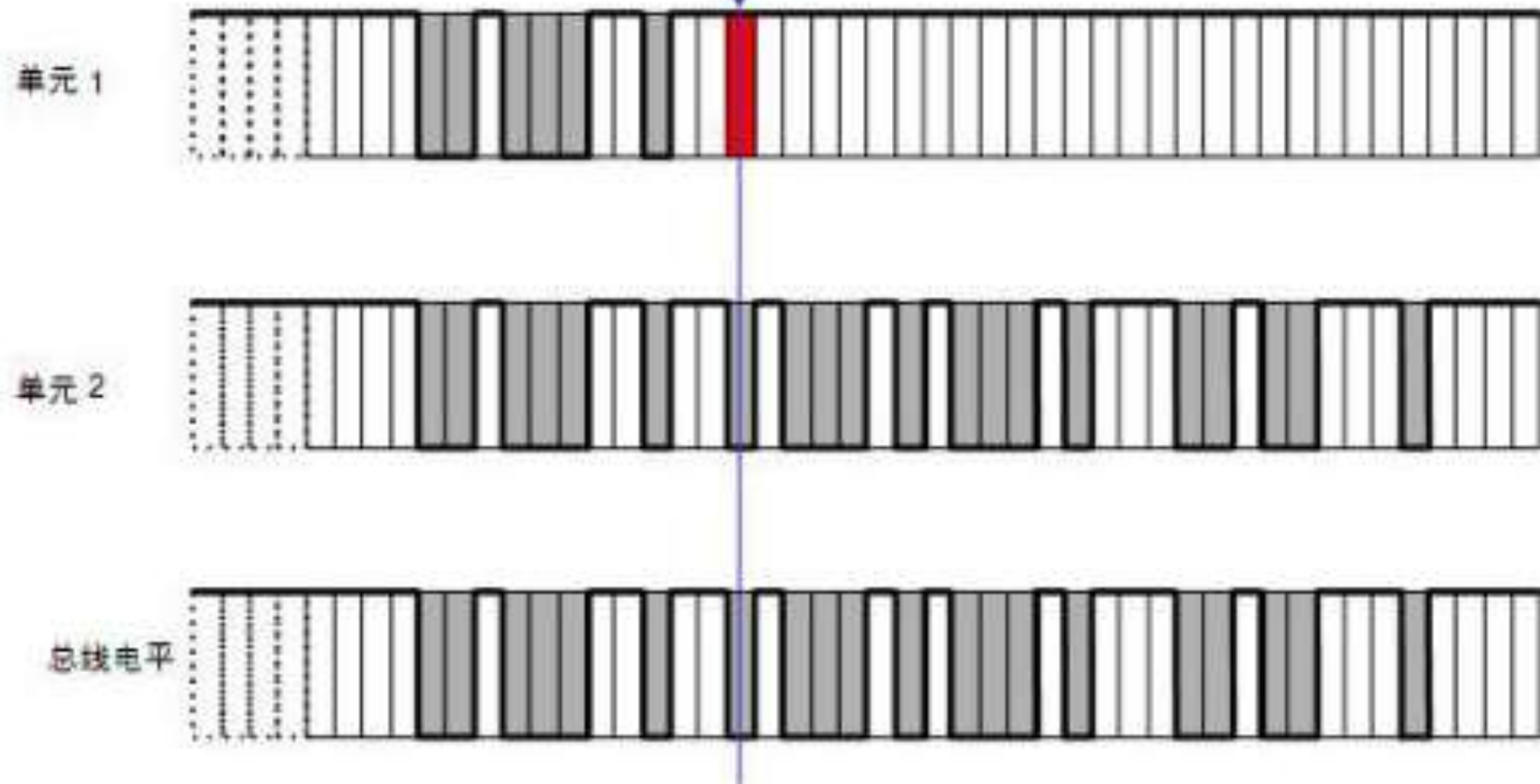
# 分布式系统：无全局时钟，事件触发



# CAN仲裁: CSMA/CA, 2胜出!

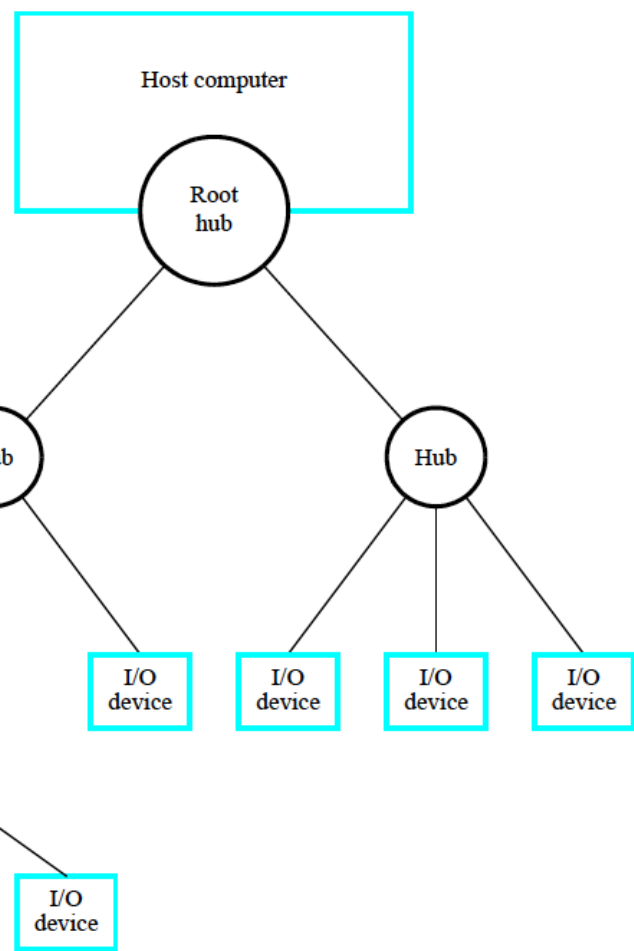
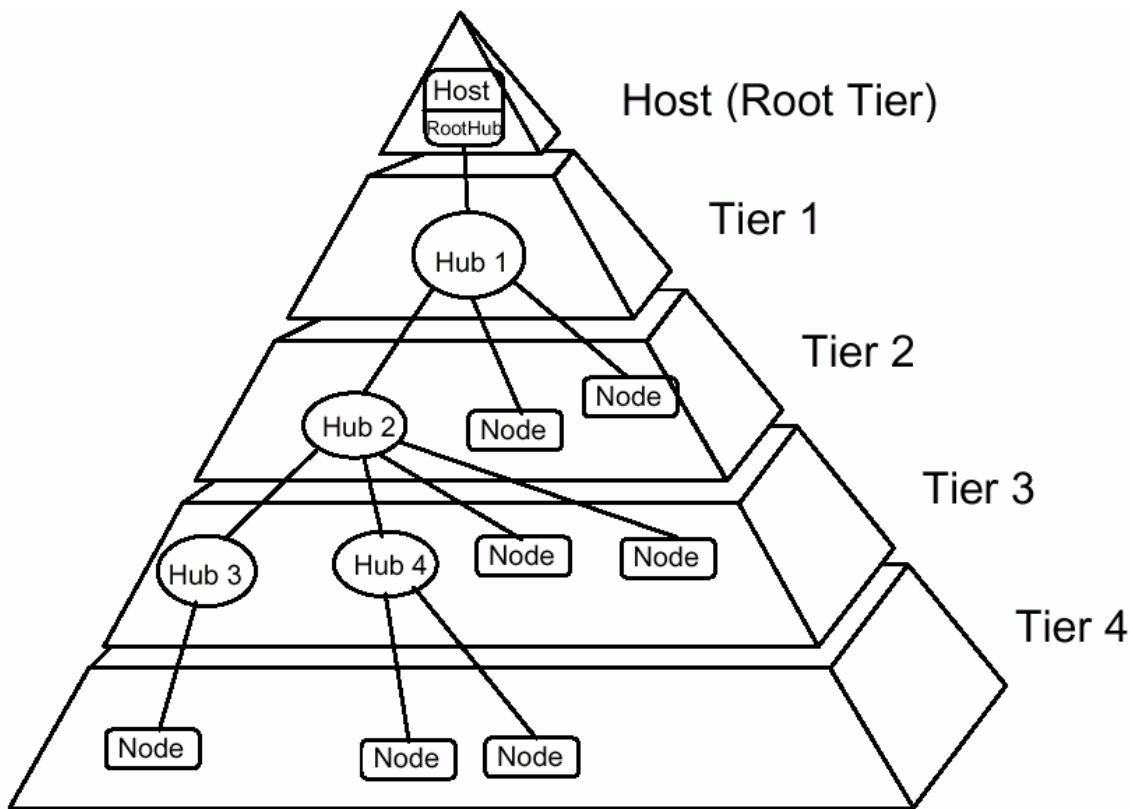
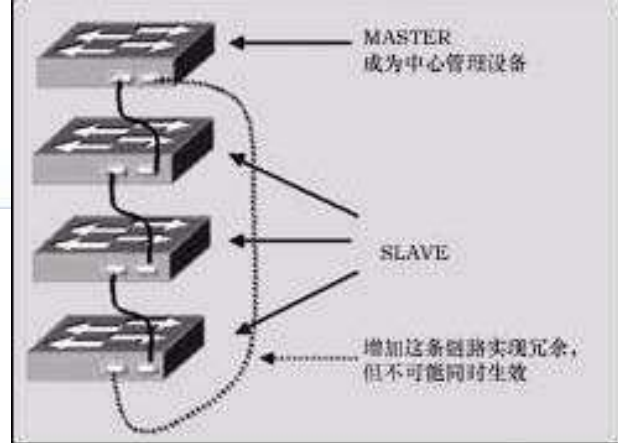


仲裁失利, 从下一个位开始转为接收状态工作



- 分布式仲裁: 冲突检测/冲突避免、并行竞争

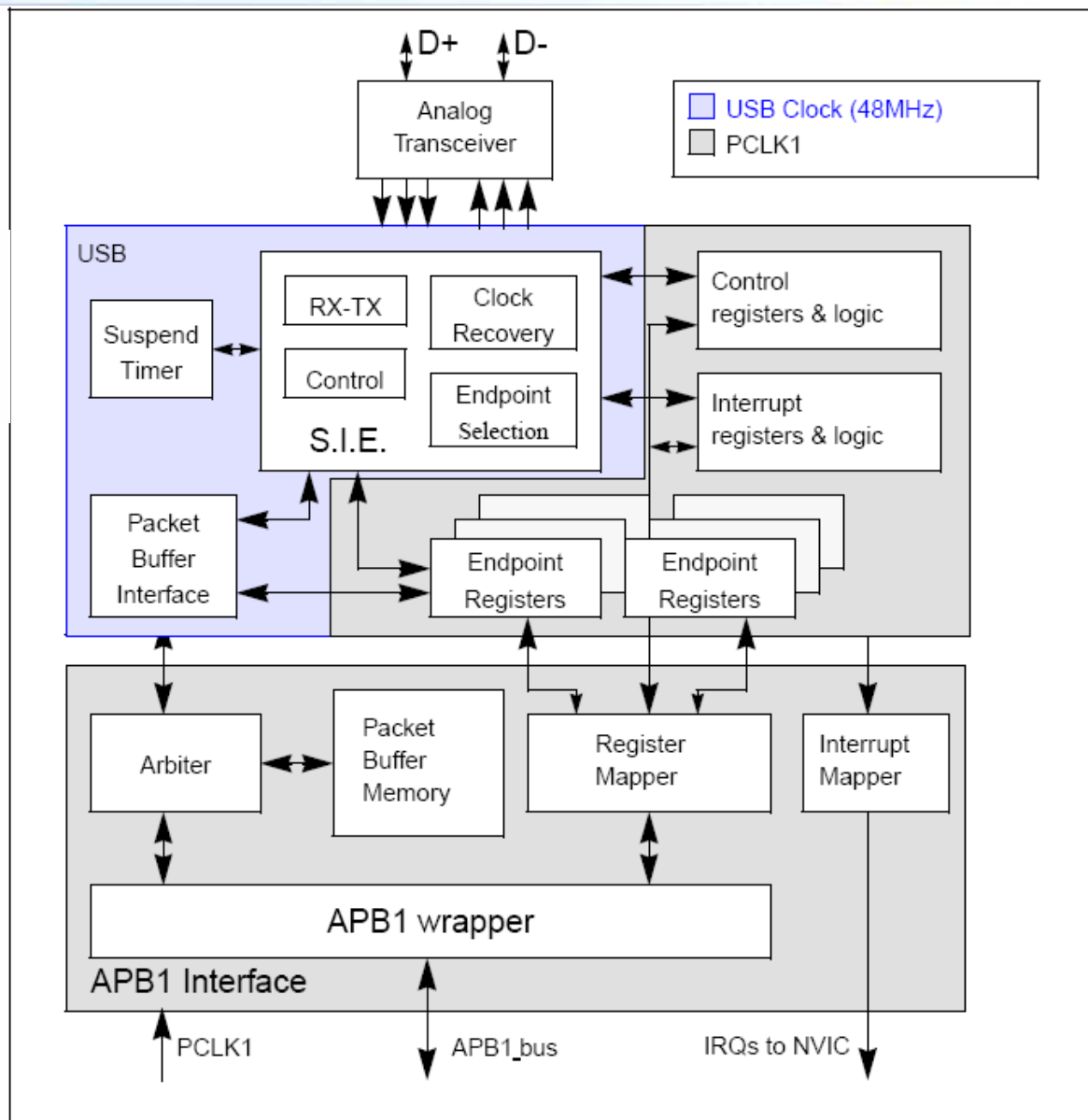
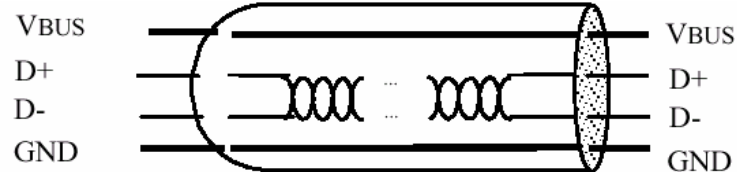
# USB总线拓扑结构



传输距离不大于5米。

可通过菊花链的形式同时挂接多个（可达127个）USB设备。

# USB外设的结构框图





# USB接口技术



- USB 总线属于一种**轮循**方式的总线，主机控制端口初始化所有的数据传输
- 支持**3**种数据传输速率操作
  - 低速1.5Mb/s，全速12Mb/s和高速480Mb/s。
- 支持**4**种类型的传输模式
  - 控制数据传送：在设备连接时对设备进行设置，还可对指定设备进行控制，如通道控制；
    - 两种类型的通道：流和消息。
  - 批量数据传送：大批量产生并使用的数据，在传输约束下，具有很广的动态范围；
  - 中断数据的传送：由事件产生，数据是少量的，且其数据延迟时间也是有限范围的。
  - 同步数据的传送：由预先确定的传送延迟来填满预定的**USB** 带宽。

# USB数据传输

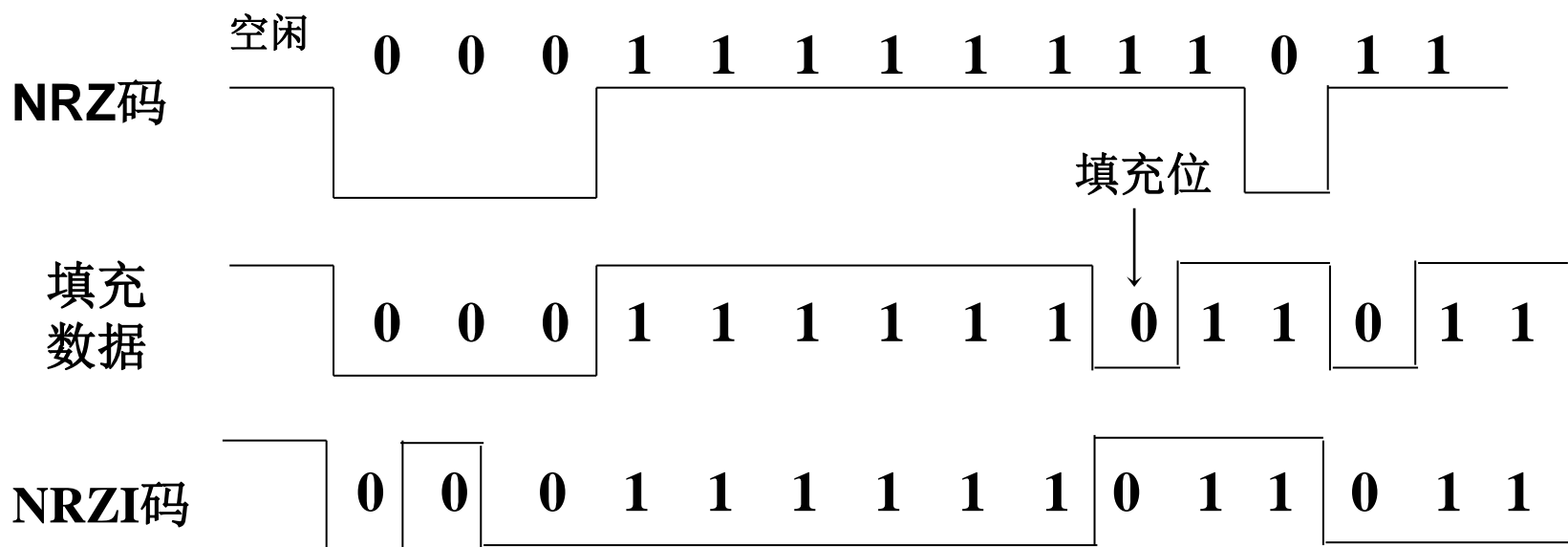
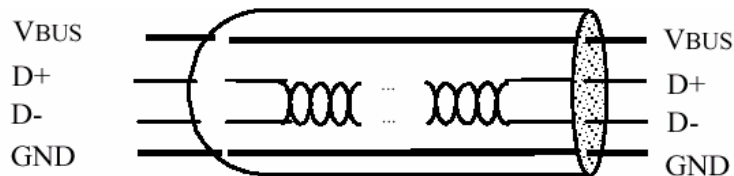


- 每一总线执行动作最多传送三个数据包
  - 标志包(token packet)
    - 在每次传送开始时主机控制器发送，描述传输运作的种类、方向、USB 设备地址和终端号
  - 数据包
  - 握手包
    - 接收端表明是否传送成功
    - 可利用其进行流量控制

# USB总线的编码方式

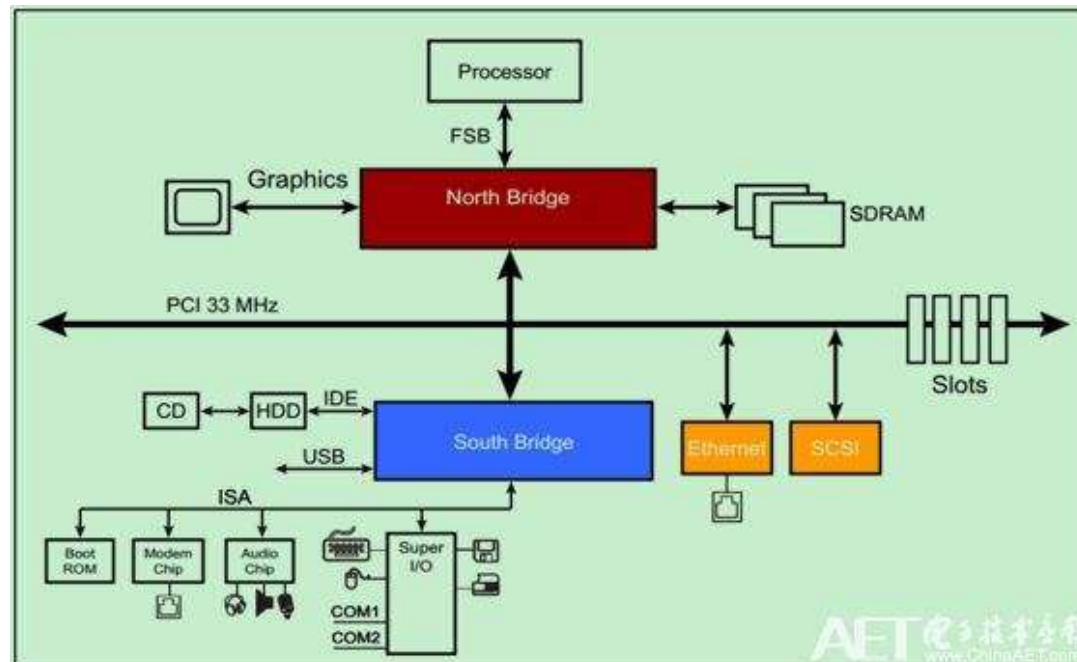


- “见0就翻”



# 总线带宽需求分析

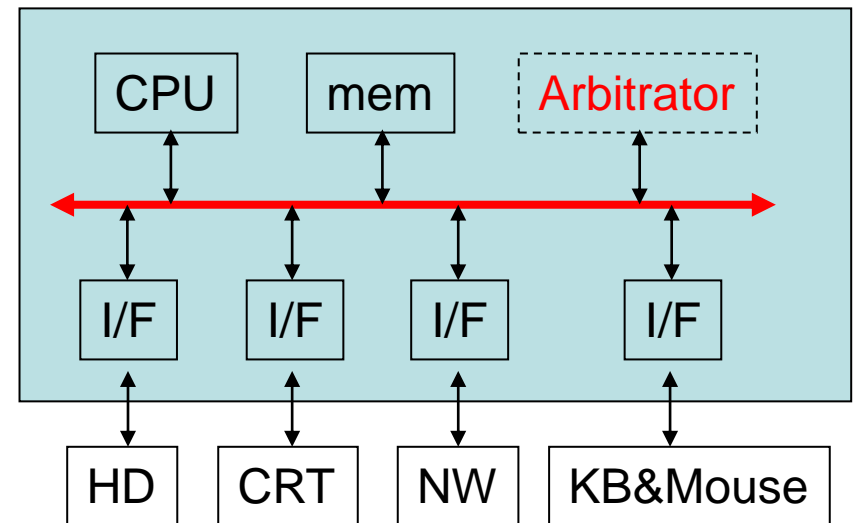
- 例如：系统需求如下
  - 每秒30帧，每帧640\*480，彩色数24位的图象显示卡的数据吞吐量为28MB/s；
  - 100Mbps传输率的光纤网，需总线吞吐量为12.5MB/s；





# 思考：ABC范式@Ixx

- 系统中有哪些功能部件，命名/寻址方式，主从？
- 系统互连方式，数据传输类型
- 总线是什么，为什么需要总线，有哪些类型？
- 总线系统互连拓扑**结构**？
- 总线**传输**过程？
  - 总线**仲裁**的基本策略？
  - 总线传输如何**同步**？
  - 传输**过程控制**？
- 总线标准
  - RS232、ISA、PCIE
- 总线与OS的关系？



# 思考



- 基于访问 or 基于时间片?

- 适于总线的分配策略有哪些?

- “当事件率达到60~70%时，优先级低的结点的服务将得不到保证，整个系统是**不可靠**的”。

- 总线数据传输过程

- 单master系统

- 多master系统

- 集中式：申请，传输，撤销

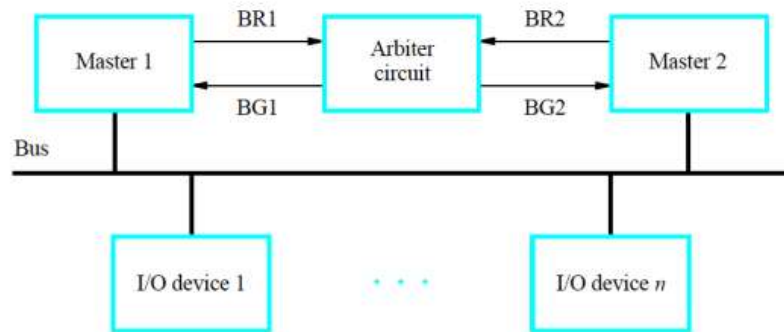
- 分布式：

- 空闲？非空闲则本地buffer，冲突时会出现“碎片”

- » Carrier Sense

- » CAN: ID仲裁

- 发送



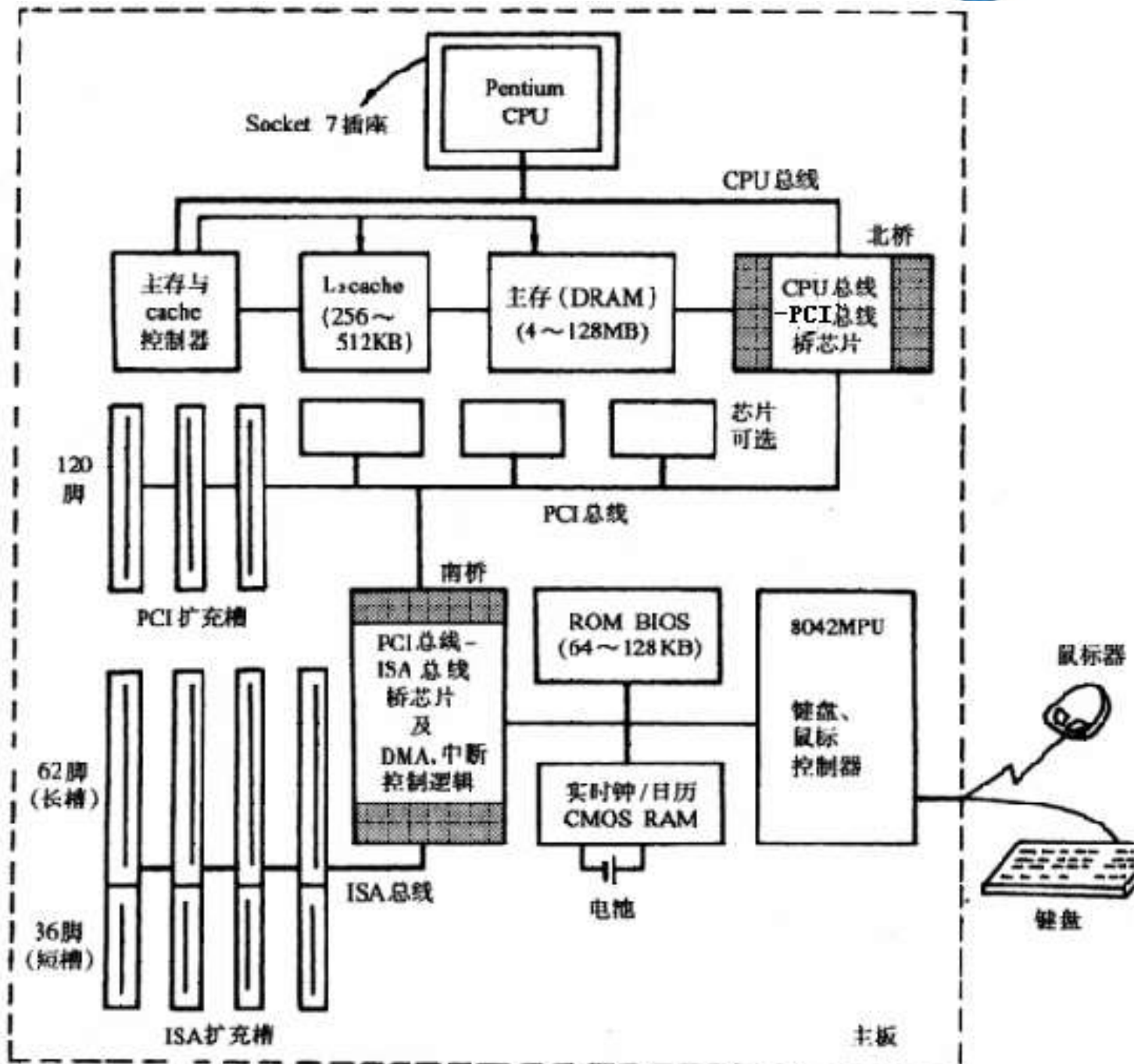
# 思考



- **总线时序**
  - 总线时钟宽度由什么因素决定?
  - 总线周期可否单clk?
  - 允许总线切换/抢占的时刻?
- **PC各设备间采用哪种同步控制方式, 哪种数据传输模式?**
  - 主设备: CPU, DMAC
  - 从设备:
    - MEM
    - Disk
    - Display
    - Keyboard
    - Mouse
    - Network
    - .....
- 对用户而言, 总线是否透明?
- **作业**
  - 3.4、3.8、3.12 (选)、3.14、3.15、3.16



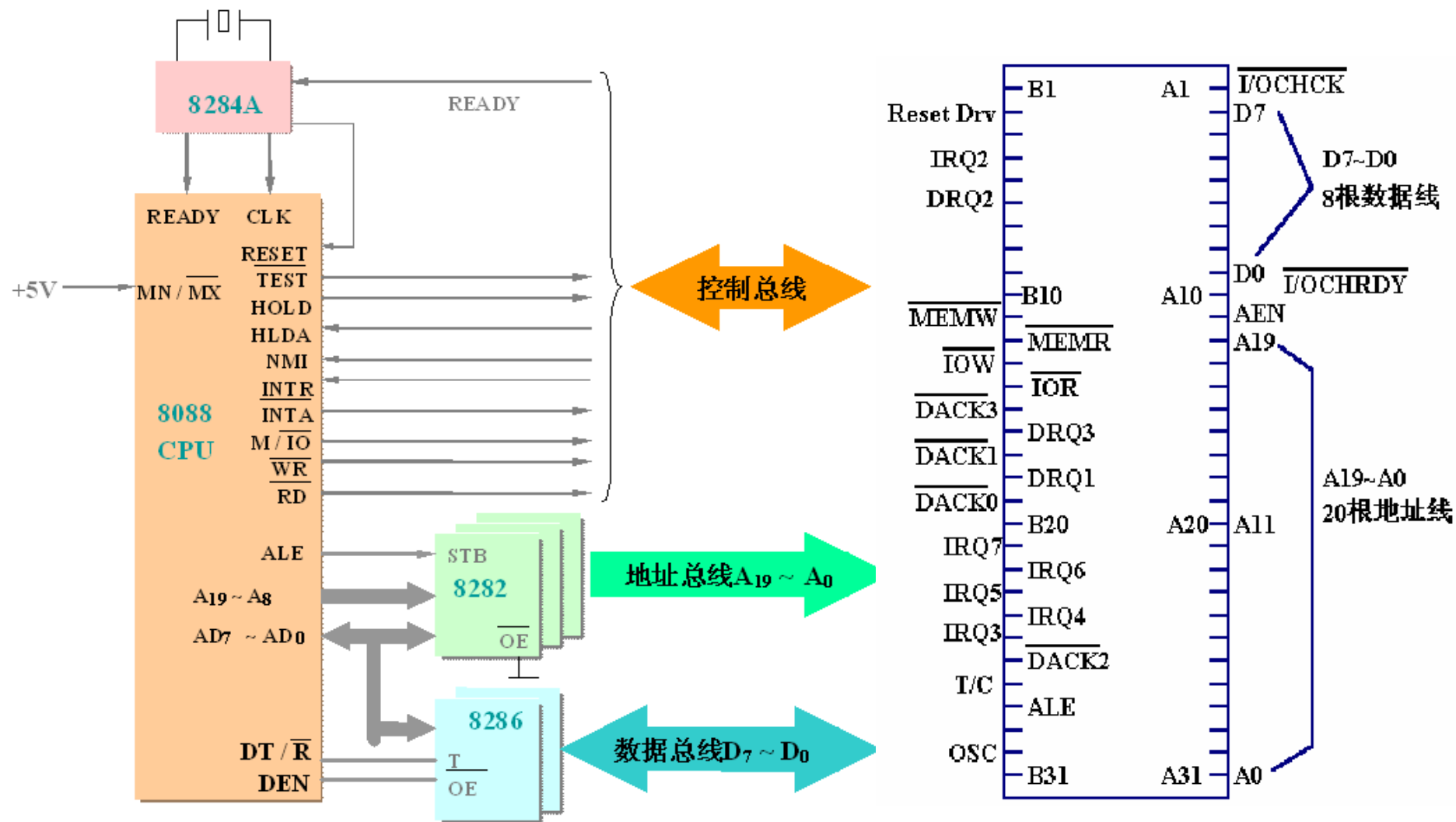
# 总线拓扑结构： 多总线



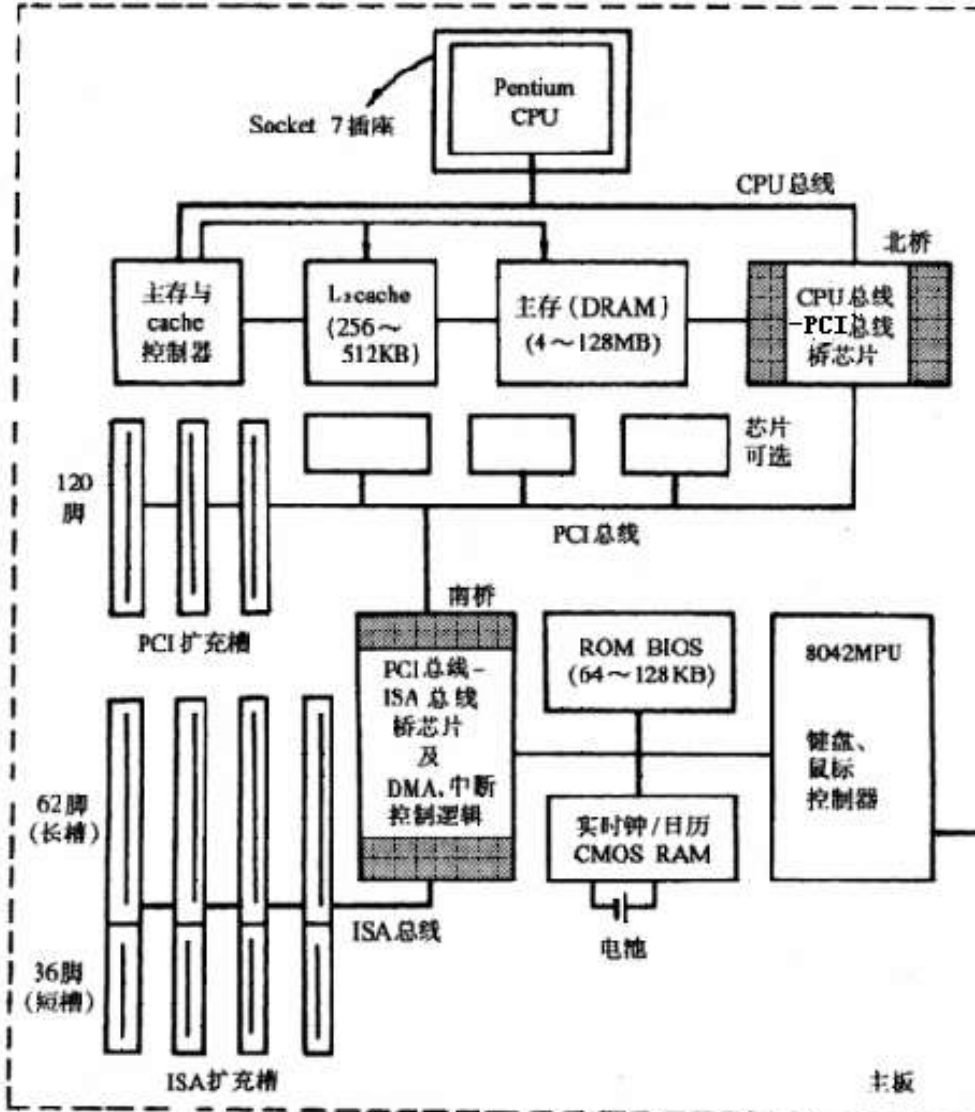
唐图10.19



# 最小组态系统：单master系统



# Pentium计算机主板总线结构图



三总线结构，即：

CPU总线、PCI总线和ISA总线。

主存与cache控制器

北桥（PCI的HOST桥）：

前端总线（CPU总线）— PCI桥  
地址空间转换，仲裁，缓存

南桥：

PCI—ISA桥  
中断、DMA

鼠标器

键盘

唐（p427）



多处理器系统的PCI配置

