复习pdf+每章章末总结

第一章

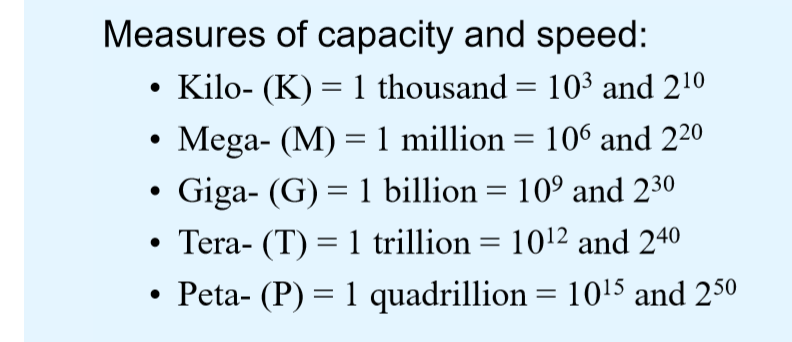
1. 硬件、软件等效性原理 Anything that can be done with software can also be done with hardware, and anything that can be done with hardware can also be done with software.
2. 计算机三个基本组成部分

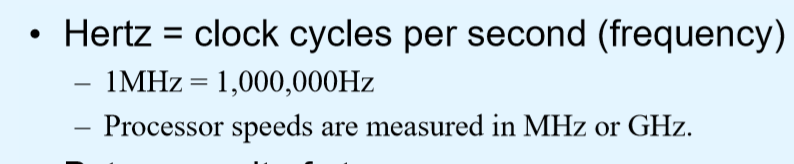
A processor to interpret(解释) and execute(执行) programs

A memory to store both data and programs

A mechanism(机制) for transferring data to and from the outside world

1. 一些度量单位





1Hz代表每秒钟周期震动1次

1. Moore’s Law (1965)

“The density of transistors(晶体管) in an integrated circuit(集成电路) will double every year”

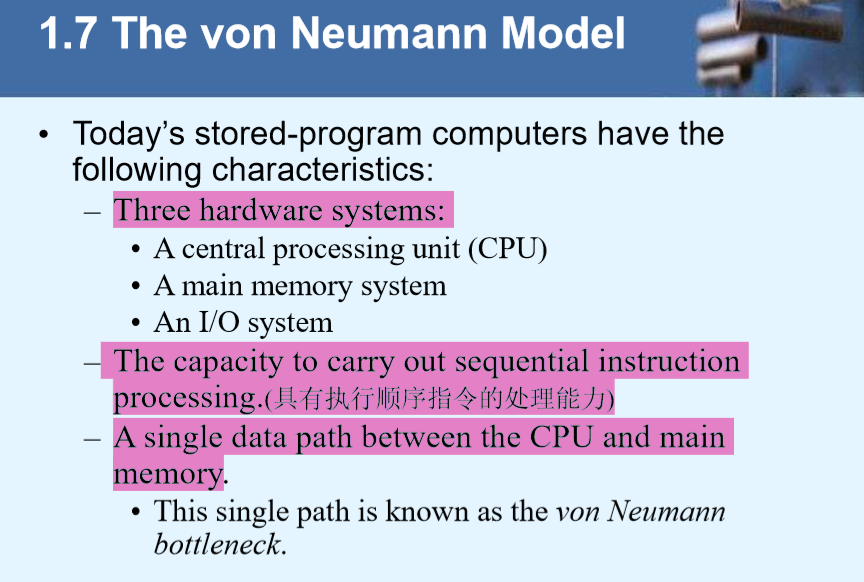
Contemporary(当代) version

The density of silicon chips(硅片) doubles every 18 months

1. The von Neumann Model 冯.诺伊曼模型

main components of a von Neumann::

capacity能力 sequential顺序



1. CPU；ALU???

以下为相对不重要

1. ISA:（instruction set architecture）指令集体系结构

ISA是在机器上运行的所有软件和执行这些软件的硬件之间的协定接口，ISA实现人机对话

1. 两种类型的计算机存储器???：ROM 和 RAM (内存 硬盘 缓存)
2. Rock定律：The cost of capital equipment to build semiconductors will double every four years
3. 取指-译码-执行 周期的工作原理：

第四章

1. CPU概念、功能The computer’s CPU fetches, decodes, and executes program instructions
2. CPU= datapath ( 数据通道)+control unit（控制单元）
3. datapath = arithmetic-logic unit（ALU算术逻辑单元） + storage units (registers寄存器)
4. The arithmetic-logic unit (ALU) carries out logical and arithmetic operations（逻辑和算术运算） as directed by the control unit.

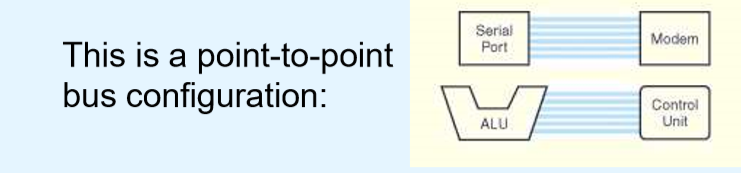
控制单元的任务：1.监视各种指令的执行以及各种信息的传递过程

2.从内存中提取指令，对指令进行译码，确保信息出现在正确的位置

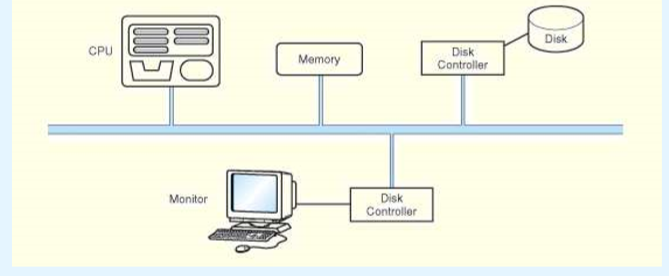
3.通知ALU执行各种操作，执行各种中断服务程序

4.使用程序计数器PC来确定下一条程序指令的位置，使用状态寄存器存储某些特定的状态

1. The control unit determines which actions to carry out according to the values in a program counter register（程序计数器寄存器） and a status register（状态寄存器）
2. A bus is a set of wires that simultaneously（同时） convey a single bit along each line
3. bus master（最初启动的设备） , bus slave（响应主设备请求的设备）
4. 点对点总线



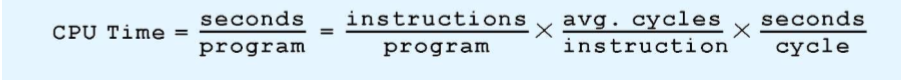
1. 多点总线



1. 传递信息类型不同分为data lines数据总线, control lines控制总线, address lines地址总线 和.电源线
2. synchronous buses同步总线 和 nonsynchronous buses异步总线（也叫asynchronous buses）
3. 总线仲裁：Daisy chain(菊花链仲裁方式): Centralized parallel(集中式平行仲裁方式):

Distributed using self-detection( 采用自选择的分配式仲裁方式):

Distributed using collision-detection(采用冲突检测的分配式仲裁方式):

1. 总线周期-bus cycle 完成总线信息传送所需的时钟脉冲间的时间间隔
2. clock cycle-时钟周期 A fixed number of clock cycles are required to carry out each data movement or computational operation
3. 

CPU时间=指令/程序 x 平均周期/指令 x 秒/周期

1. Memory Organization and Addressing(存储器组成和寻址方式)

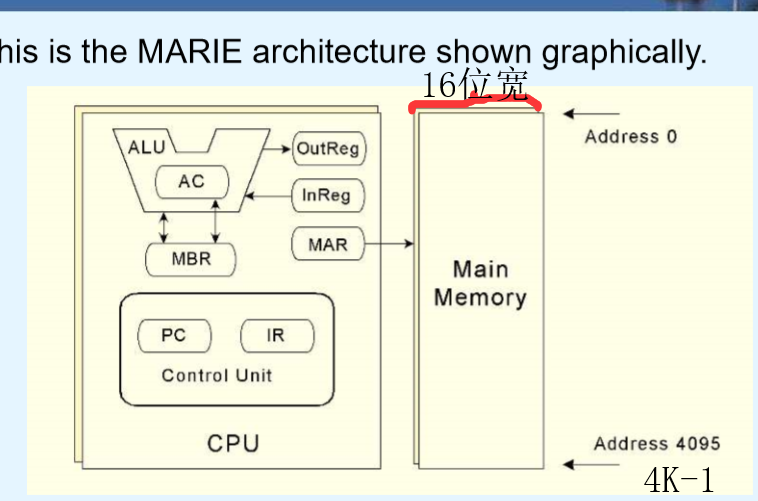
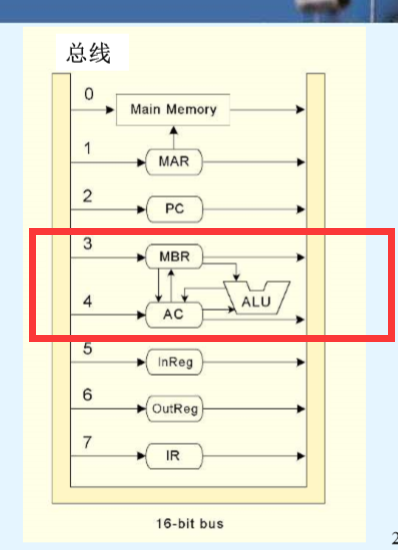
low-order interleaving(低位交叉存储),

high-order interleaving(高位交叉存储),

1. Interrupts： 中断概念、及类型:

概念：改变系统正常执行流程的各种事件

Interrupts can be triggered by I/O requests, arithmetic errors算术错误 (such as division by zero), or when an invalid instruction无效指令 is encountered

1. Memory address register(存储器地址寄存器), MAR 用来保存被引用数据的存储器地址，12位，保存指令或者指令操作数
2. Memory buffer register(存储器缓冲寄存器), MBR 用来保存刚从存储器中存取或者将要写入存储器的数据，16位，(holds the data after its retrieval检索 from, or before its placement in memory.)
3. Program counter, PC,程序将要执行的下一条指令的地址，12位
4. Instruction register, IR，程序将要执行的下一条指令，16位
5. 
6. 输入输出寄存器是8位
7. Accumulator(累加器), AC 用来保持CPU需要处理的数据，是一个通用寄存器
8. MARIE指令格式=4位操作码+12位操作数
9. 
10. Load X
11. Store X
12. Add X
13. Subt X
14. Skipcond X ？？？
15. 取值-译码-执行周期fetch-decode-execute

第一条指令装入pc

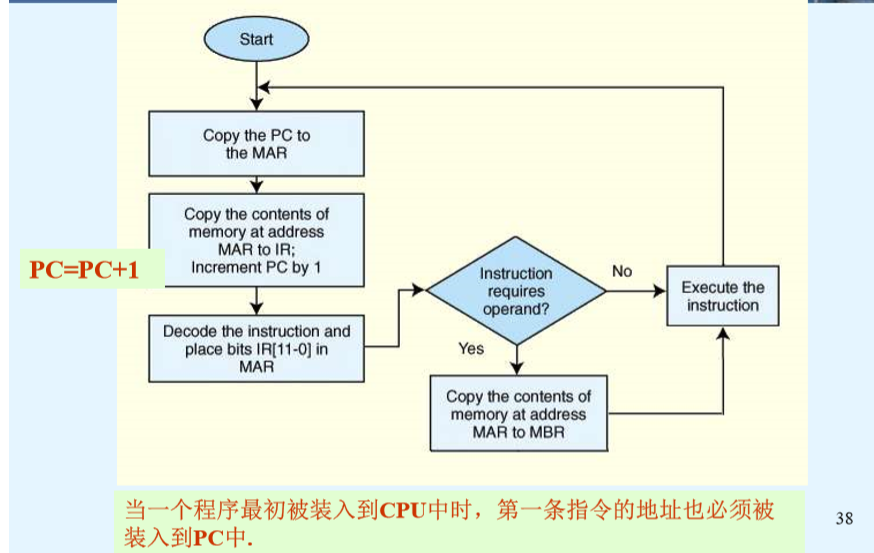
复制PC到MAR

复制MAR地址处的存储器内容到IR，增量PC加1

指令译码和将IR[11-0]放入MAR

如果指令需要操作数，复制MAR地址的存储器内容到MBR

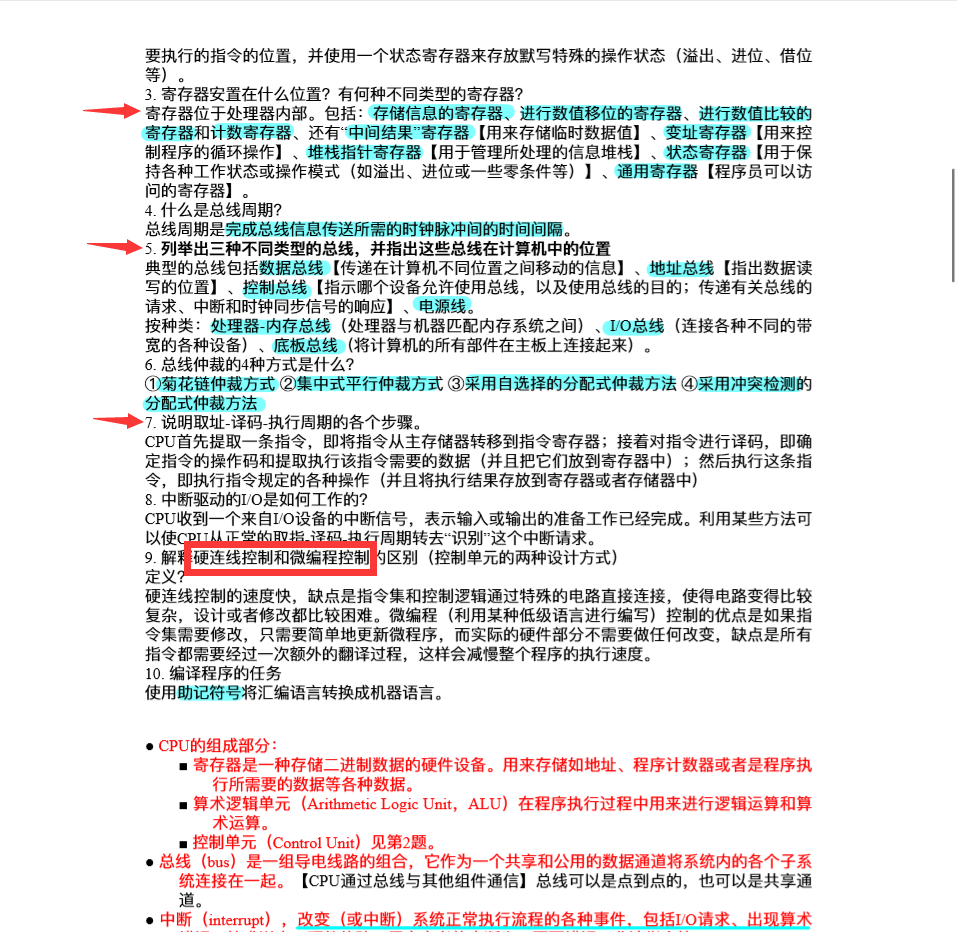
执行指令



1. 编译汇编程序过程：两次扫描

第一次通读，编译程序会建议一组成为符号表的对应关系

第二次通读，编译程序使用符号表来填充空白地址，并且生成相应的机器语言指令

1. 
2. Assemblers(汇编程序 汇编器)

第五章

1. 指令集在特征上的区别：

1.每一条指令的bits

2.操作数在CPU中的存储方式（堆栈结构和寄存器）

3.指令的操作数数目

4.操作数的位置(寄存器-寄存器 寄存器-存储器 存储器-存储器)

5.操作（操作的类，指令是否可以访问存储器）

6.操作数的类型和长度（操作数可以是地址 数字 字符）

2. 指令系统体系结构（ISA）的衡量因素：

1.程序执行指令时所占的内存空间

2.指令系统的复杂程度

3.指令的长度（越短越好，固定长度译码容易）

4.指令系统中指令的数目

3. CPU数据的存储方式

堆栈体系结构A stack architecture（编码密度好，不能对堆栈进行随机访问，很难产生高效率的编码，）

累加器体系结构An accumulator architecture（操作数隐含在累加器中，降低机器的内部复杂性，指令非常短，非常频繁的访问存储器）（1个通用寄存器）

通用寄存器（GPR）体系结构A general purpose register architecture.（最广泛接受的）（很多个通用寄存器）

4. Instructions can be fixed length or variable length.

fixed length：固定长度浪费一些存储空间，但是指令执行速度快，指令流水线上效果更好

variable length：译码复杂，但是节省存储空间

5. To enrich the instruction set for a fixed length instruction set, expanding opcodes (扩展操作码)can be used.扩展操作码译码复杂，

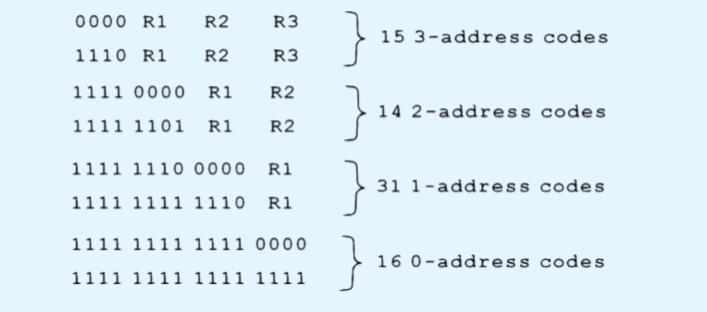
扩展操作码：(有多少条地址编码确定的时候就是看比如1地址和2地址编码里不能有冲突就可以)

15条3地址编码

14条2地址编码

31条1地址编码

16条0地址编码



6. 寻址方式：指定指令中操作数的位置的方法(就是找操作数)

– Immediate 立即寻址

– Direct 直接寻址

– Register 寄存器寻址

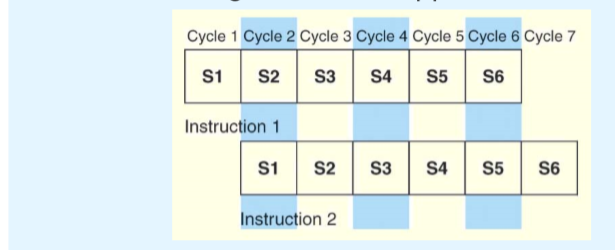
– Register Indirect

– Indirect 间接寻址

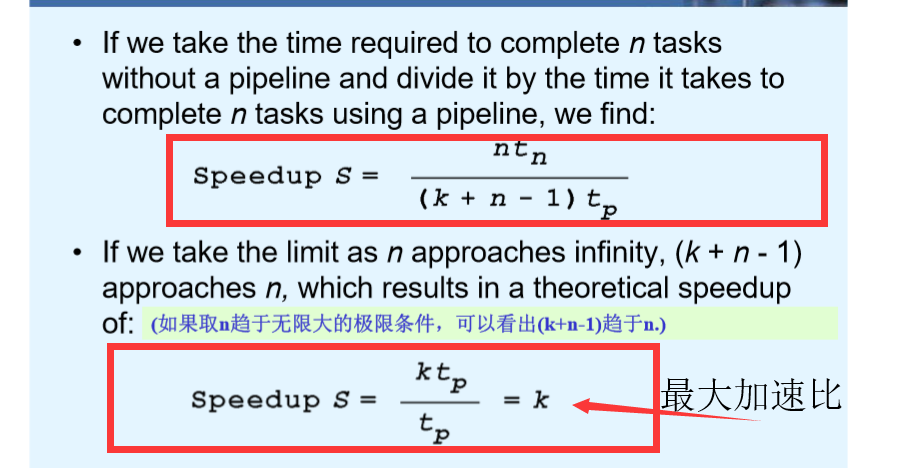
– Indexed 变址寻址(存的是偏移量)

– Based 基址寻址(存的是首地址)(变址和基址都是简单加起来，只是意义不同)

– Stack 堆栈寻址

7. 

上面这个图是6级流水线



tn是完成一项(n=1)任务的时间

tp是一个时钟周期(time per stage)

8. 流水线冲突(Pipeline conflicts)：

资源冲突(resource conflicts)数据存储又要存储指令

数据关联(data dependency)指令尚未执行完，后一条指令需要该指令的结果作为操作数

条件分支语句(分支转移delayed branch)

9. pipelining(流水线概念)：计算机使用时钟脉冲来精确控制各个操作步骤的顺利执行，但是有实还可以使用一些额外的脉冲来控制某个操作步骤中的一些小细节，有些CPU会将取指-译码-执行周期分成一些较小的步骤，其中某些较小的步骤可以并行执行，这种时间的交叠可以加快CPU的执行速度，这种方法成为流水线

10. 一些列指令的第一条指令不会被加速，从第二条指令开始会被加速，并行计算减少运行时间

11. evaluating postfix notation后缀表达式

第六章

1. TYPES OF MEMORY：

•**RAM** (random access memory)随机存储器(易失性) and **ROM** (read-only memory)(非失易)只读存储器.

•**SRAM(非常快)** (制作cache memory )is faster and much more expensive than **DRAM(cpu必须等待刷新完成才能访问DRAM，慢，又刷新时间)**; however, designers use DRAM because it is much denser (can store many bits per chip存储密度高), uses less power功耗低, and generates less heat 产热小than SRAM.

•**ROM分为ROM, PROM, EPROM, EEPROM， 闪存**

2. THE MEMORY HIERARCHY ESSENTIAL 存储器体系结构

•Hit命中：要请求的数据就驻留在访问的存储器层

Miss,缺失：要请求的数据就没有驻留在访问的存储器层

Hit rate,命中率：找到所需数据的百分比

Miss rate, 缺失率：1-命中率

Hit time命中时间：CPU取得所请求的数据的需要的时间

Miss penalty缺失损失：cpu处理缺失事件需要的时间：1新的数据代替数据块2.数据块传送到处理器

•Locality of Reference:

Temporal locality,时间局部性：最近访问的内容很可能在不久将来会再次访问

Spatial locality, 空间局部性：可能会集中访问一个范围内的数据

Sequential locality顺序局部性：访问存储器的指令倾向于按顺序执行

3. CACHE MEMORY高速缓存：

Direct Mapped Cache直接映射高速缓存

Fully Associative Cache全关联高速缓存

Set Associative Cache组关联高速缓存

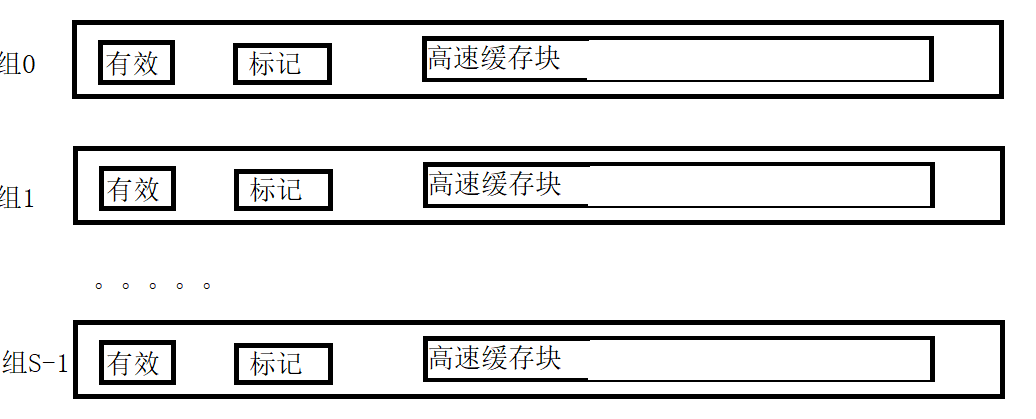
找到请求的字就是三个步骤：1.组选择2.行匹配 3.字抽取

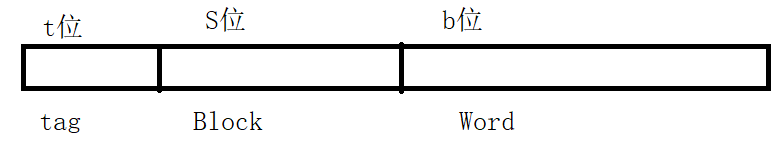
例子如下，地址域划分如下

Direct Mapped Cache直接映射高速缓存

每个组只有一行的高速缓存是Direct Mapped Cache直接映射高速缓存

缓存块构成：

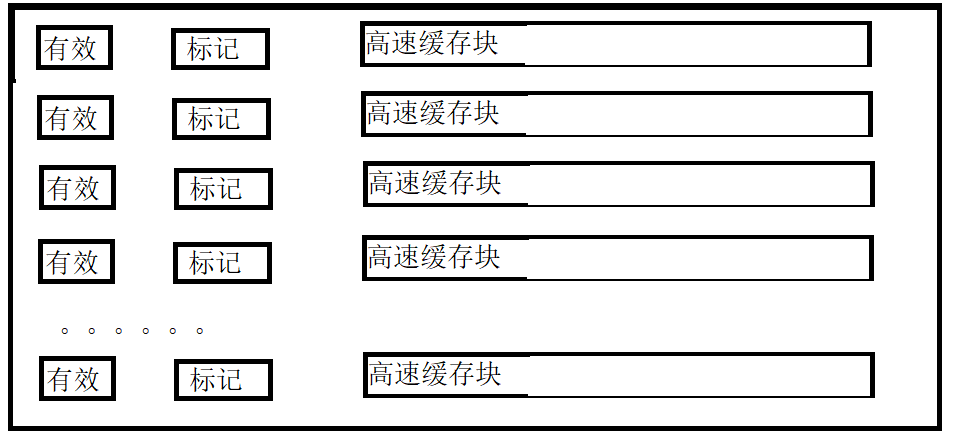


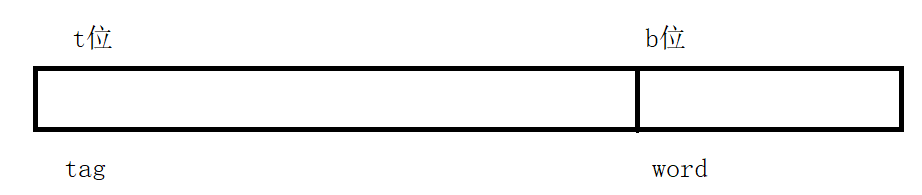


高速缓存块中有很多个字(很多个块)

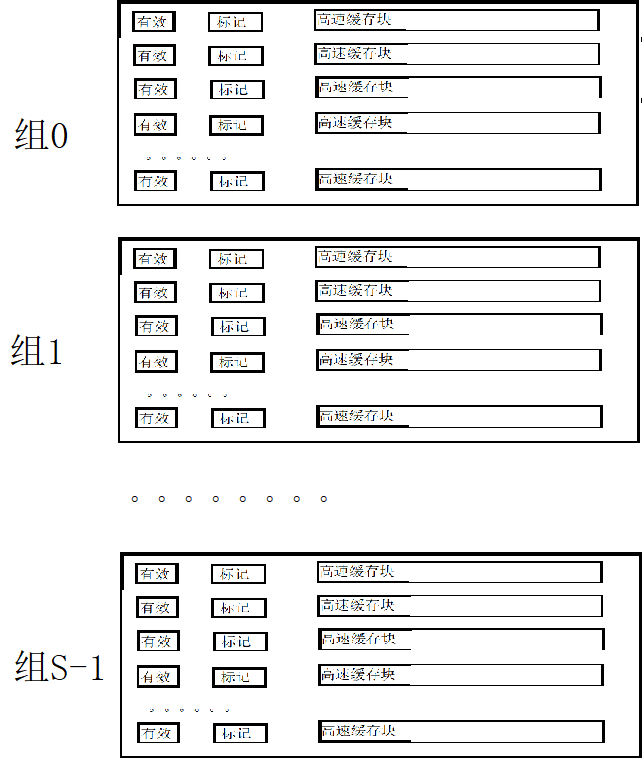
Fully Associative Cache全关联高速缓存

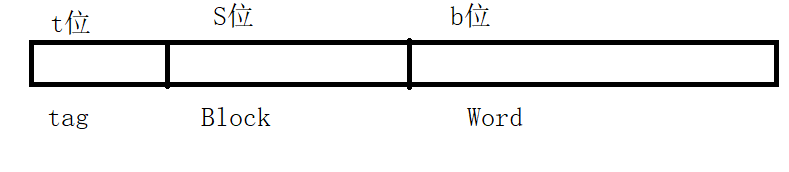
缓存块构成(一个组)



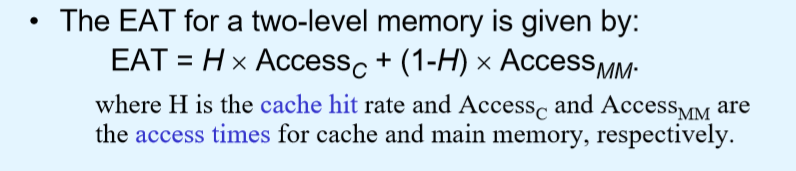


Set Associative Cache组关联高速缓存





4. 有效存取(访问)时间（effective access time 即EAT）



5. 程序局部性不好的时候高速缓存就会失效

6. 虚拟存储器VIRTUAL MEMORY

Virtual address, 进程的逻辑地址或程序地址

Physical address 物理存储器的实际地址

Mapping, 映射

Page frames, 页帧(主存储上的、物理存储器上的)

Pages, 页 （虚拟存储器上的）

Paging, 分页 （将虚拟页从硬盘复制到主存储器某个页帧的过程）

Fragmentation, 存储碎片

Page fault 缺页（一个请求页在主存储上没有找到的事件，必须将请求页从硬盘复制到存储器）

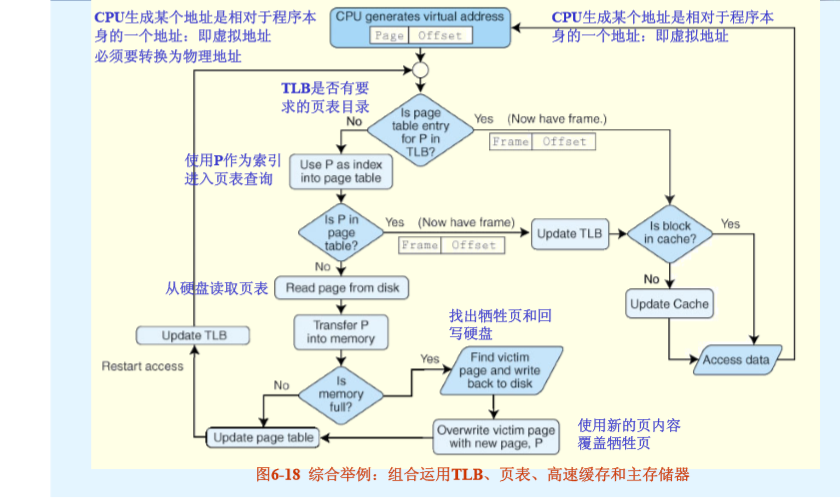
自己看书😳175页 176页

7. 包含虚拟存储器的EAT计算：

处理器每次访问存储器，必须两次对物理存储器进行访问操作：一次是引用页表，一次是访问要求的实际数据

课本上的177页

8. TLB 178页 179页 TLB的构成和页表相同



一些复习pdf上的话

9. Computer memory is organized in a hierarchy, with the smallest, fastest memory at the top and the largest, slowest memory at the bottom

10. Cache memory gives faster access to main memory, while virtual memory uses disk storage to give the illusion of having a large main memory

11. Cache maps blocks of main memory to blocks of cache memory. Virtual memory maps page frames to virtual pages.

12. There are three general types of cache: Direct mapped, fully associative and set associative

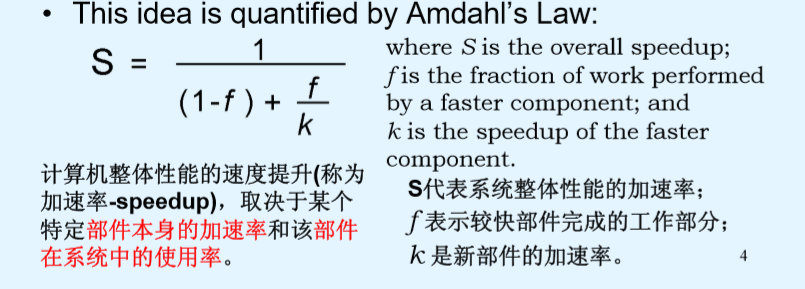
13. With fully associative and set associative cache, as well as with virtual memory, replacement policies must be established

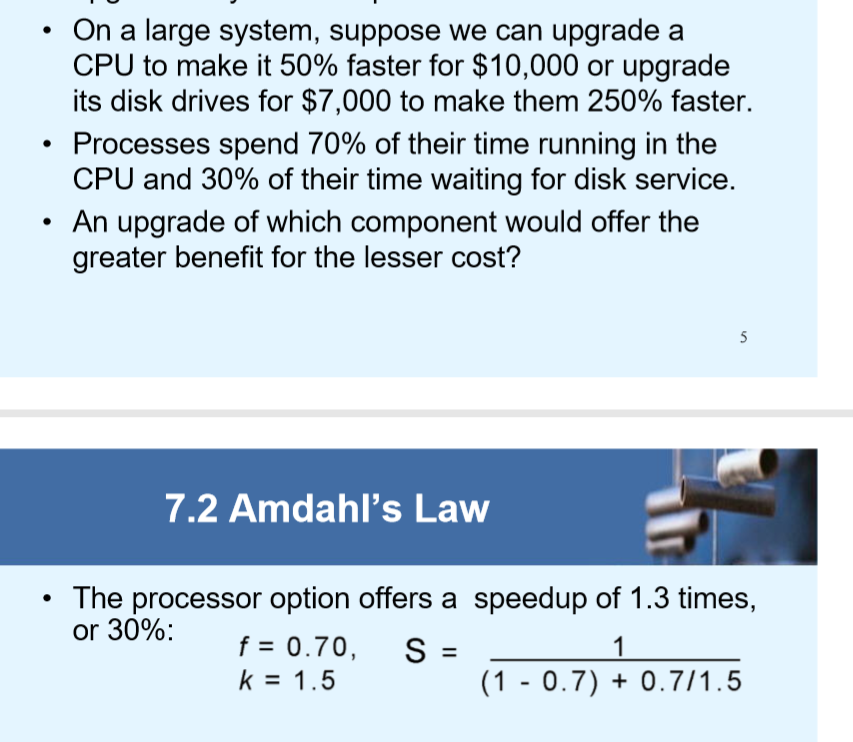
14. Replacement policies include LRU(最近最少被使用), FIFO(先进先出). These policies must also take into account what to do with dirty blocks.

15. All virtual memory must deal with fragmentation, internal for paged memory, external for segmented memory.

第七章

1. Amdahl定律 加速率(就是计算机整体性能的速度提升)S的求法，





2. 输入输出子系统包括I/O systems consist of memory blocks, cabling, control circuitry, interfaces, and media 看课本190页

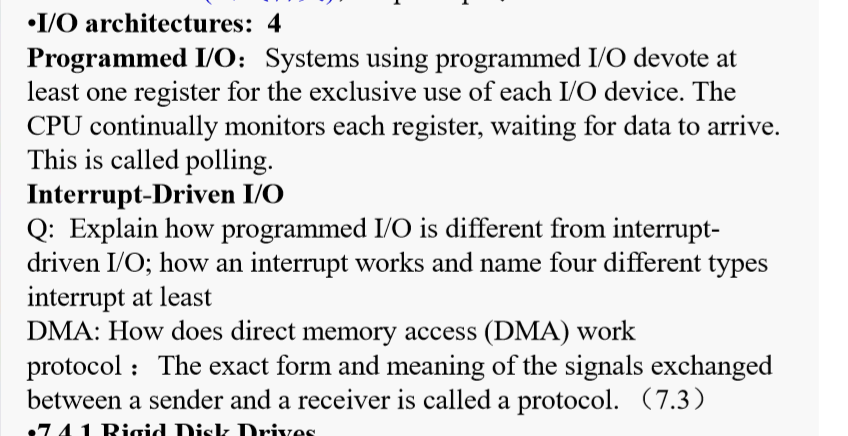
3. I/O control methods include

programmed I/O程序控制IO, 某些专用系统例如自动取款机和一些用来控制或者监视外部事件的系统

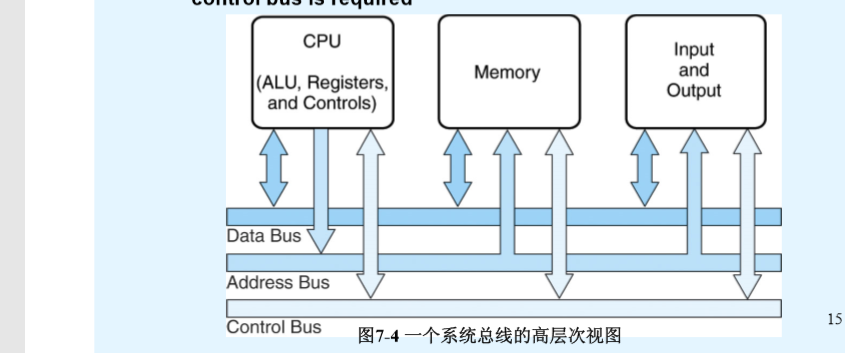
interrupt-based I/O中断控制IO, 个人系统

DMA直接存储器存取, 中小型系统

channel I/O.通道控制IO 大型系统



4. Buses require control lines, a clock, and data lines. Timing diagrams specify operational details.



5. Magnetic disk磁盘驱动器 is the principal(主要的) form of durable持续 storage

6. Disk performance metrics(性能度量) include seek time寻道时间, rotational delay旋转延迟, and reliability estimates可靠性

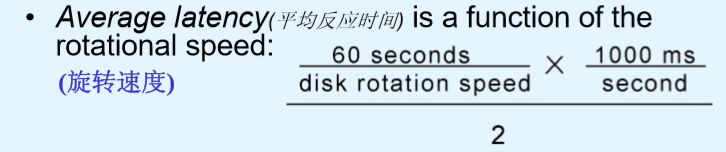
寻道时间seek time： 磁盘驱动器定位到指定磁道上的时间

旋转延迟rotational delay：读写头定位到指定扇区上的时间

存取时间=寻道时间+旋转延迟

传输时间transfer time=存取时间+实际读取数据所需要的时间

反应时间计算公式



7. Magnetic tape磁带 is also an archival存储 medium媒介. Recording methods are track-based基于跟踪的, serpentine(蛇 形扫描), and helical scan(螺旋扫描).

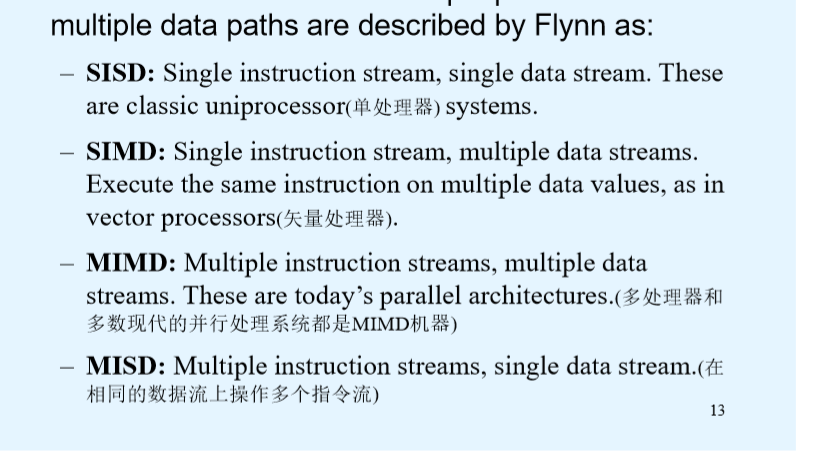
8. RAID 3 and RAID 5 are the most common.

Two approaches to data compression压缩 are statistical data compression统计压缩 and dictionary systems字典系统

9. .压缩系数=1-（压缩后/压缩前）\*100%

第九章

FLYNN分类方法（分类依据：指令的数目和流入处理器的数据流的数目Flynn’s taxonomy considers two factors: the number of instructions and the number of data streams that flow into the processor.）：



第十章 性能度量和分析

优化方案：CPU优化、存储器优化、I/O优化