

王道考研——组成原理

WWW.CSKAOYAN.COM

第五章 中央处理器

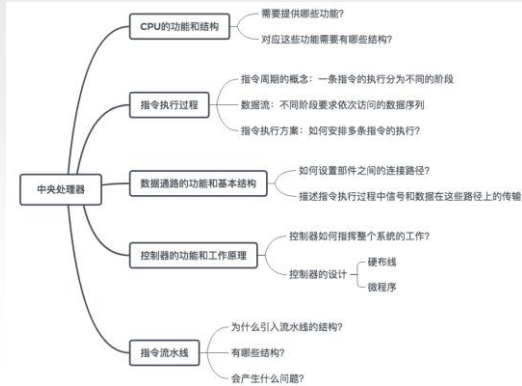
本节内容

中央处理器

指令流水线
基本概念
性能指标

王道考研/CSKAOYAN.COM

本章总览



王道考研/CSKAOYAN.COM

指令流水的定义

一条指令的执行过程可以分成多个阶段（或过程）。
根据计算机的不同，具体的分法也不同。

取指	分析	执行
----	----	----

取指：根据PC内容访问主存储器，取出一条指令送到IR中。

分析：对指令操作码进行译码，按照给定的寻址方式和地址字段中的内容形成操作数的有效地址EA，并从有效地址EA中取出操作数。

执行：根据操作码字段，完成指令规定的功能，即把运算结果写到通用寄存器或主存中。

设取指、分析、执行3个阶段的时间都相等，用t表示，按以下几种执行方式分析n条指令的执行时间：

1. 顺序执行方式 $\text{总耗时} T = n \times 3t = 3nt$

取指k	分析k	执行k	取指k+1	分析k+1	执行k+1
-----	-----	-----	-------	-------	-------

传统冯诺依曼机采用顺序执行方式，又称串行执行方式。

优点：控制简单，硬件代价小。

缺点：执行指令的速度较慢，在任何时刻，处理机中只有一条指令在执行，各功能部件的利用率很低。

王道考研/CSKAOYAN.COM

特点：每个阶段用到的硬件不一样。

指令流水的定义

1. 顺序执行方式 总耗时 $T = n \times 3t = 3nt$



传统冯 诺依曼机采用顺序执行方式，又称串行执行方式。

优点：控制简单，硬件代价小。

缺点：执行指令的速度较慢，在任何时刻，处理机中只有一条指令在执行，各功能部件的利用率很低。

2. 一次重叠执行方式 总耗时 $T = 3t + (n-1) \times 2t = (1+2n)t$



优点：程序的执行时间缩短了1/3，各功能部件的利用率明显提高。

缺点：需要付出硬件上较大开销的代价，控制过程也比顺序执行复杂了。

王道考研/CSKAOYAN.COM

指令流水的定义

1. 顺序执行方式 总耗时 $T = n \times 3t = 3nt$



传统冯 诺依曼机采用顺序执行方式，又称串行执行方式。优点：控制简单，硬件代价小。

缺点：执行指令的速度较慢，在任何时刻，处理机中只有一条指令在执行，各功能部件的利用率很低。

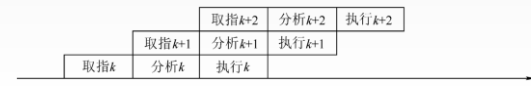
2. 一次重叠执行方式 总耗时 $T = 3t + (n-1) \times 2t = (1+2n)t$



优点：程序的执行时间缩短了1/3，各功能部件的利用率明显提高。

缺点：需要付出硬件上较大开销的代价，控制过程也比顺序执行复杂了。

3. 二次重叠执行方式 总耗时 $T = 3t + (n-1) \times t = (2+n)t$



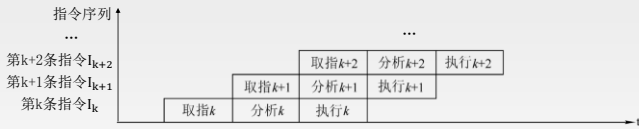
与顺序执行方式相比，指令的执行时间缩短近2/3。这是一种理想的指令执行方式，在正常情况下，处理机中同时有3条指令在执行。

注：也可以把每条指令的执行过程分成4个或5个阶段，分成5个阶段是比较常见的做法。

王道考研/CSKAOYAN.COM

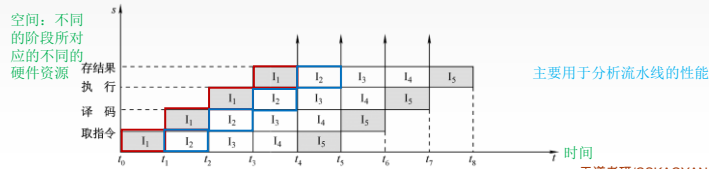
流水线的表示方法

1. 指令执行过程图



主要用于分析指令执行过程以及影响流水线的因素(见下一个视频)

2. 时空图



主要用于分析流水线的性能

王道考研/CSKAOYAN.COM

流水线的性能指标

1. 吞吐率
2. 加速比
3. 效率

王道考研/CSKAOYAN.COM

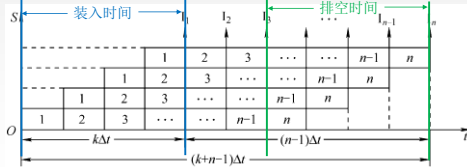
流水线的性能指标

1. 吞吐率 吞吐率是指在单位时间内流水线所完成的任务数量，或是输出结果的数量。

设任务数为 n ；处理完成 n 个任务所用的时间为 T_k

则计算流水线吞吐率（TP）的最基本的公式为 $TP = \frac{n}{T_k}$

理想情况下，流水线的时空图如下：当连续输入的任务 $n \rightarrow \infty$ 时，得最大吞吐率为 $TP_{max} = 1/\Delta t$ 。



$$T_k = (k+n-1) \Delta t$$

流水线的实际吞吐率为

$$TP = \frac{n}{(k+n-1) \Delta t}$$

一条指令的执行分为 k 个阶段，每个阶段耗时 Δt ，一般取 Δt = 一个时钟周期

王道考研/CSKAOYAN.COM

流水线的性能指标

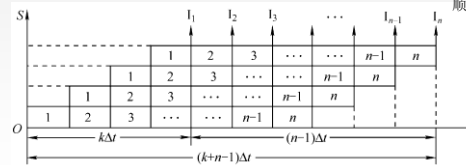
2. 加速比 完成同样一批任务，不使用流水线所用的时间与使用流水线所用的时间之比。

设 T_0 表示不使用流水线时的执行时间，即顺序执行所用的时间； T_k 表示使用流水线时的执行时间

则计算流水线加速比（S）的基本公式为 $S = \frac{T_0}{T_k}$ 当连续输入的任务 $n \rightarrow \infty$ 时，最大加速比为 $S_{max} = k$ 。

理想情况下，流水线的时空图如下：

单独完成一个任务耗时为 $k \Delta t$ ，则顺序完成 n 个任务耗时 $T_0 = nk \Delta t$



$$T_k = (k+n-1) \Delta t$$

实际加速比为

$$S = \frac{kn \Delta t}{(k+n-1) \Delta t} = \frac{kn}{k+n-1}$$

一条指令的执行分为 k 个阶段，每个阶段耗时 Δt ，一般取 Δt = 一个时钟周期

王道考研/CSKAOYAN.COM

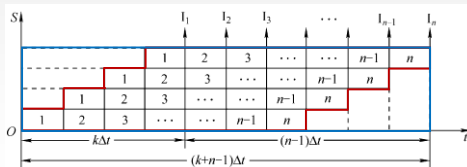
流水线的性能指标

3. 效率 流水线的设备利用率称为流水线的效率。

在时空图上，流水线的效率定义为完成 n 个任务占用的时空区有效面积与 n 个任务所用的时间与 k 个流水段所围成的时空区总面积之比。

则流水线效率（E）的一般公式为 $E = \frac{n \text{个任务占用} k \text{个流水段所围成的时空区有效面积}}{n \text{个任务所用的时间与} k \text{个流水段所围成的时空区总面积}} = \frac{T_0}{k T_k}$

理想情况下，流水线的时空图如下：

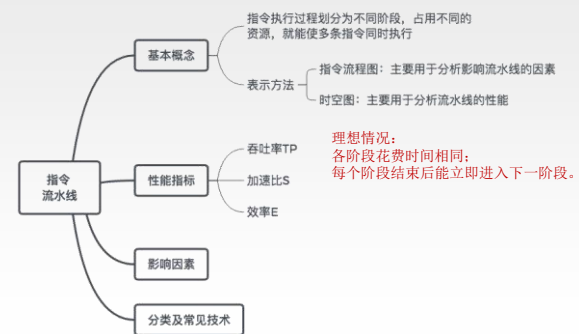


当连续输入的任务 $n \rightarrow \infty$ 时，最高效率为 $E_{max} = 1$ 。

一条指令的执行分为 k 个阶段，每个阶段耗时 Δt ，一般取 Δt = 一个时钟周期

王道考研/CSKAOYAN.COM

本节回顾



王道考研/CSKAOYAN.COM

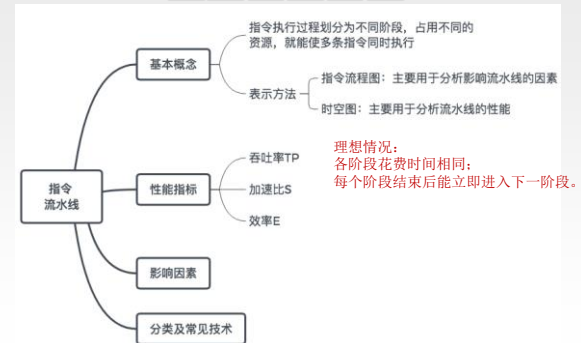
本节内容

中央处理器

指令流水线
影响因素
分类

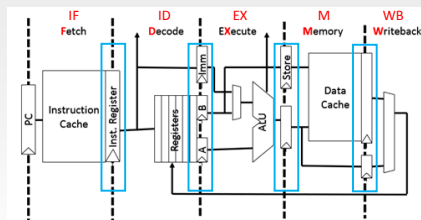
王道考研/CSKAOYAN.COM

本节总览



王道考研/CSKAOYAN.COM

机器周期的设置



各部件实际耗时： 100ns 80ns 70ns 50ns 50ns

流水线每一个功能段部件后面都要有一个缓冲寄存器，或称为锁存器。其作用是保存本流水段的执行结果，提供给下一流水段使用。

为方便流水线的设计，将每个阶段的耗时取成一样，以最长耗时为准。即此处应将机器周期设置为100ns。

王道考研/CSKAOYAN.COM

影响流水线的因素

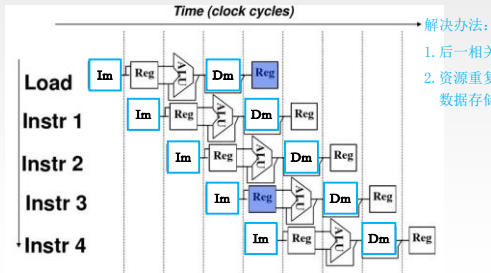
1. 结构相关（资源冲突）
2. 数据相关（数据冲突）
3. 控制相关（控制冲突）

王道考研/CSKAOYAN.COM

影响流水线的因素

1. 结构相关（资源冲突）

由于多条指令在同一时刻争用同一资源而形成的冲突称为结构相关。



解决办法:

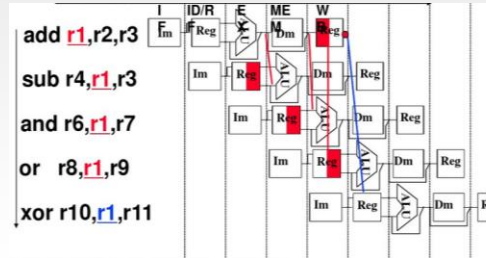
1. 后一相关指令暂停一周
2. 资源重复配置:
数据存储器+指令存储器

王道考研/CSKAOYAN.COM

影响流水线的因素

2. 数据相关（数据冲突）

数据相关指在一个程序中，存在必须等前一条指令执行完才能执行后一条指令的情况，则这两条指令即为数据相关。



解决办法:

1. 把遇到数据相关的指令及其后续指令都暂停一至几个时钟周期，直到数据相关问题消失后再继续执行。可分为硬件阻塞(stall)和软件插入“NOP”两种方法。
2. 数据旁路技术。
3. 编译优化：通过编译器调整指令顺序来解决数据相关。

王道考研/CSKAOYAN.COM

影响流水线的因素

2. 数据相关（数据冲突）

例题. 假设某指令流水线采用“按序发射，按序完成”方式，没有采用转发技术处理数据相关，并且同一寄存器的读和写操作不能在同一个时钟周期内进行。若高级语言程序中某赋值语句为 $x = a + b$ ， x 、 a 和 b 均为int型变量，它们的存储单元地址分别表示为 $[x]$ 、 $[a]$ 和 $[b]$ 。该语句对应的指令序列及其在指令流中的执行过程如下图所示。

```

I1 LOAD R1, [a] ([a]) -> R1
I2 LOAD R2, [b] ([b]) -> R2
I3 ADD R1, R2 (R1) + (R2) -> R2
I4 STORE R2, [x] (R2) -> [x]

```

I3与I1和I2存在数据相关，

则这4条指令执行过程中I3的ID段和I4的IF段被阻塞的原因各是什么？I4和I3存在数据相关。

指令	1	2	3	4	5	6	7	8	9	10	11	12	13	14
I ₁	IF	ID	EX	M	WB									
I ₂		IF	ID	EX	M	WB								
I ₃				IF			ID	EX	M	WB				
I ₄											ID	EX	M	WB

王道考研/CSKAOYAN.COM

影响流水线的因素

2. 数据相关（数据冲突）

数据的基本操作：读（R）、写（W）

冲突的基本类型：RAW、WAR、WAW

RAW

注：“按序发射，按序完成”时，只可能出现RAW相关。

I1: ADD R5, R2, R4; (R2)+(R4) -> R5

I2: ADD R4, R5, R3; (R5)+(R3) -> R4

WAR

I1: STA M, R2;

(R2) -> M, M为主存单元

乱序发射，编写程序的时候希望I1在I2前完成，

I2: ADD R2, R4, R5;

(R4)+(R5) -> R2

但优化手段导致I2在I1前发射。

WRW

I1: MUL R3, R2, R1;

(R2)*(R1) -> R3

存在多个功能部件时，后一条指

I2: SUB R3, R4, R5;

(R4)-(R5) -> R3

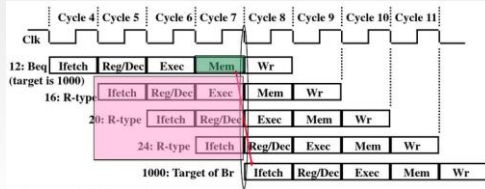
令可能比前一条指令先完成。

王道考研/CSKAOYAN.COM

影响流水线的因素

3. 控制相关（控制冲突）

当流水线遇到转移指令和其他改变PC值的指令而造成断流时，会引起控制相关。

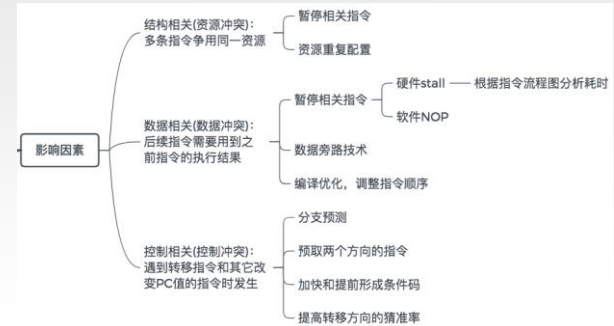


解决办法：

1. 尽早判别转移是否发生，尽早生成转移目标地址
2. 预取转移成功和不成功两个控制流方向上的目标指令
3. 加快和提前形成条件码
4. 提高转移方向的猜准确率

王道考研/CSKAOYAN.COM

影响流水线的因素



王道考研/CSKAOYAN.COM

流水线的分类

1. 部件功能级、处理机级和处理机间级流水线

根据流水线使用的级别的不同，流水线可分为部件功能级流水线、处理机级流水线和处理机间流水线。
部件功能级流水就是将复杂的算术逻辑运算组成流水线工作方式。例如，可将浮点加法操作分成求阶差、对阶、尾数相加以及结果规格化等4个子过程。

处理机级流水是把一条指令解释过程分成多个子过程，如前面提到的取指、译码、执行、访存及写回5个子过程。

处理机间流水是一种宏流水，其中每一个处理机完成某一专门任务，各个处理机所得到的结果需存放在与下一个处理机所共享的存储器中。

2. 单功能流水线和多功能流水线

按流水线可以完成的功能，流水线可分为单功能流水线和多功能流水线。

单功能流水线指只能实现一种固定的专门功能的流水线；

多功能流水线指通过各段间的不同连接方式可以同时或不同时地实现多种功能的流水线。

王道考研/CSKAOYAN.COM

流水线的分类

3. 动态流水线和静态流水线

按同一时间内各段之间的连接方式，流水线可分为静态流水线和动态流水线。

静态流水线指在同一时间内，流水线的各段只能按同一种功能的连接方式工作。

动态流水线指在同一时间内，当某些段正在实现某种运算时，另一些段却正在进行另一种运算。这样对提高流水线的效率很有好处，但会使流水线控制变得很复杂。

4. 线性流水线和非线性流水线

按流水线的各个功能段之间是否有反馈信号，流水线可分为线性流水线与非线性流水线。

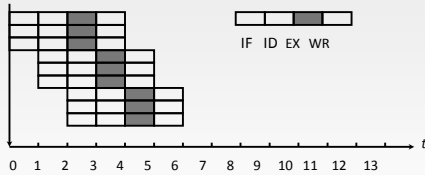
线性流水线中，从输入到输出，每个功能段只允许经过一次，不存在反馈回路。

非线性流水线存在反馈回路，从输入到输出过程中，某些功能段将数次通过流水线，这种流水线适合进行线性递归的运算。

王道考研/CSKAOYAN.COM

流水线的多发技术

1. 超标量技术



每个时钟周期内可并发多条独立指令
要配置多个功能部件

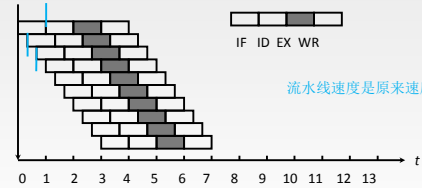
不能调整指令的执行顺序

通过编译优化技术,把可并行执行的指令搭配起来

王道考研/CSKAOYAN.COM

流水线的多发技术

2. 超流水技术



流水线速度是原来速度的3倍

在一个时钟周期内再分段(3段)

在一个时钟周期内一个功能部件使用多次(3次)

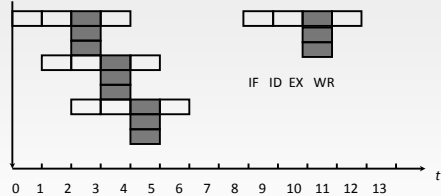
不能调整指令的执行顺序

靠编译程序解决优化问题

王道考研/CSKAOYAN.COM

流水线的多发技术

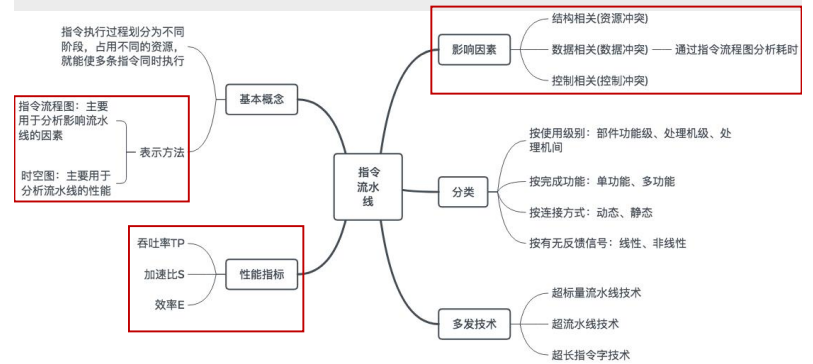
3. 超长指令字



由编译程序挖掘出指令间潜在的并行性,
将多条能并行操作的指令组合成一条
具有多个操作码字段的超长指令字(可达几百位)
采用多个处理部件

王道考研/CSKAOYAN.COM

本节回顾



王道考研/CSKAOYAN.COM