

王道计算机考研强化课

计组题型总结

五段式指令流水线

考研大纲要求&备考方法

考察简单的概念题

(六) 指令流水线

指令流水线的基本概念； 指令流水线的基本实现

五段式指令流水线。指令按序发射，按序完成

结构冒险、数据冒险和控制冒险的处理； 超标量和动态流水线的基本概念

结合MIPS指令序列分析并处理冒险（重点&难点）

考察简单的概念题

备考方法：

- 1.简单的概念题——复习王道书指令流水线课后小题即可
- 2.分析指令流水线是否发生冒险，以及冒险的处理。
 - a. 结构冒险——较少考察。通常考概念、简答。
 - b. 数据冒险——最常考。常结合MIPS指令序列考察
 - c. 控制冒险——较常考。常结合MIPS指令序列考察

结构冒险及其处理



主存/Cache
(存放指令、数据)

时钟	1	2	3	4	5	6	7	8	9
指令1	IF	ID	EX	M	WB				
指令2		IF	ID	EX	M	WB			
指令3			IF	ID	EX	M	WB		
指令4				IF	ID	EX	M	WB	
指令5					IF	ID	EX	M	WB

结构冒险及其处理

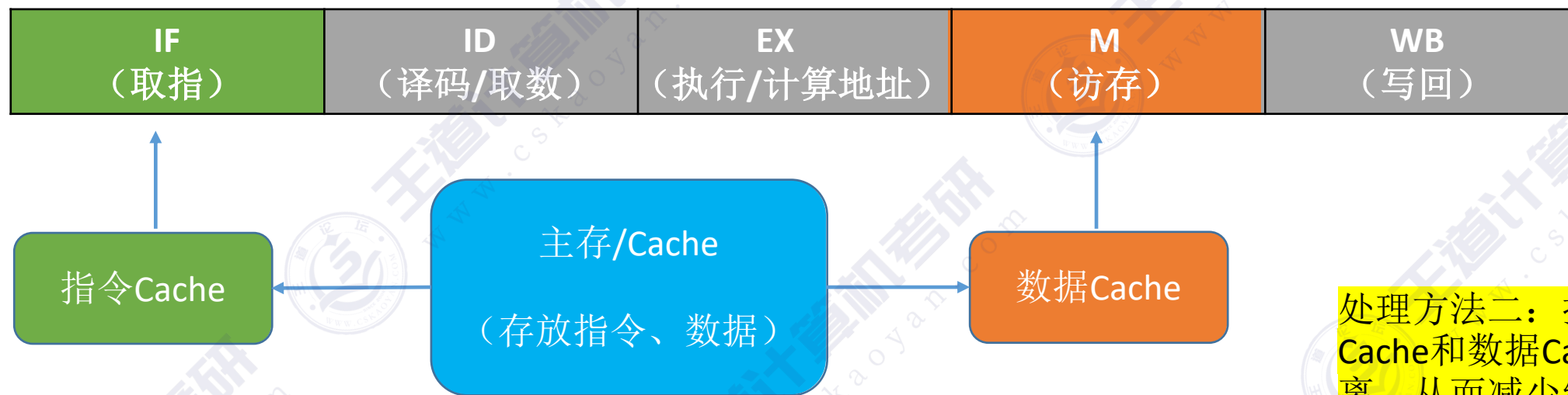


主存/Cache
(存放指令、数据)

处理方法一：可以通过硬件阻塞来解决结构冒险

时钟	1	2	3	4	5	6	7	8	9	10	11	12
指令1	IF	ID	EX	M	WB							
指令2		IF	ID	EX	M	WB						
指令3			IF	ID	EX	M	WB					
指令4				阻塞	阻塞	阻塞	IF	ID	EX	M	WB	
指令5								IF	ID	EX	M	WB

结构冒险及其处理



处理方法二：指令Cache和数据Cache分离，从而减少结构冒险

时钟	1	2	3	4	5	6	7	8	9
指令1	IF	ID	EX	M	WB				
指令2		IF	ID	EX	M	WB			
指令3			IF	ID	EX	M	WB		
指令4				IF	ID	EX	M	WB	
指令5					IF	ID	EX	M	WB

真题实战：结构冒险

2014

16. 采用指令 Cache 与数据 Cache 分离的主要目的是 ()。
- A. 降低 Cache 的缺失损失
 - B. 提高 Cache 的命中率
 - C. 降低 CPU 平均访存时间
 - D. 减少指令流水线资源冲突

同义术语：结构冒险=结构冲突=资源冲突

只要考察五段式指令流水线，都可以默认数据 Cache 和指令 Cache 分离

真题窥探：结构冒险

44. (12 分) 某 16 位计算机中，带符号整数用补码表示，数据 Cache 和指令 Cache 分离。

题 44 表给出了指令系统中部分指令格式，其中 Rs 和 Rd 表示寄存器，mem 表示存储单元地址，(x) 表示寄存器 x 或存储单元 x 的内容。

该计算机采用 5 段流水方式执行指令，各流水段分别是取指 (IF)、译码/读寄存器 (ID)、执行/计算有效地址 (EX)、访问存储器 (M) 和结果写回寄存器 (WB)，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，并且同一个寄存器的读和写操作不能在同一个时钟周期内进行。请回答下列问题。

题 44 表 指令系统中部分指令格式

名称	指令的汇编格式	指令功能
加法指令	ADD Rs, Rd	$(Rs) + (Rd) \rightarrow Rd$
算术/逻辑左移	SHL Rd	$2 * (Rd) \rightarrow Rd$
算术右移	SHR Rd	$(Rd) / 2 \rightarrow Rd$
取数指令	LOAD Rd, mem	$(mem) \rightarrow Rd$
存数指令	STORE Rs, mem	$(Rs) \rightarrow mem$

- 1) 若 int 型变量 x 的值为 -513，存放在寄存器 R1 中，则执行指令“SHR R1”后，R1 的内容是多少？（用十六进制表示。）
- 2) 若某个时间段中，有连续的 4 条指令进入流水线，在其执行过程中没有发生任何阻塞，则执行这 4 条指令所需的时钟周期数为多少？
- 3) 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

橙色：可能导致数据冒险

绿色：可能导致控制冒险

如何分析 数据冒险、控制冒险？



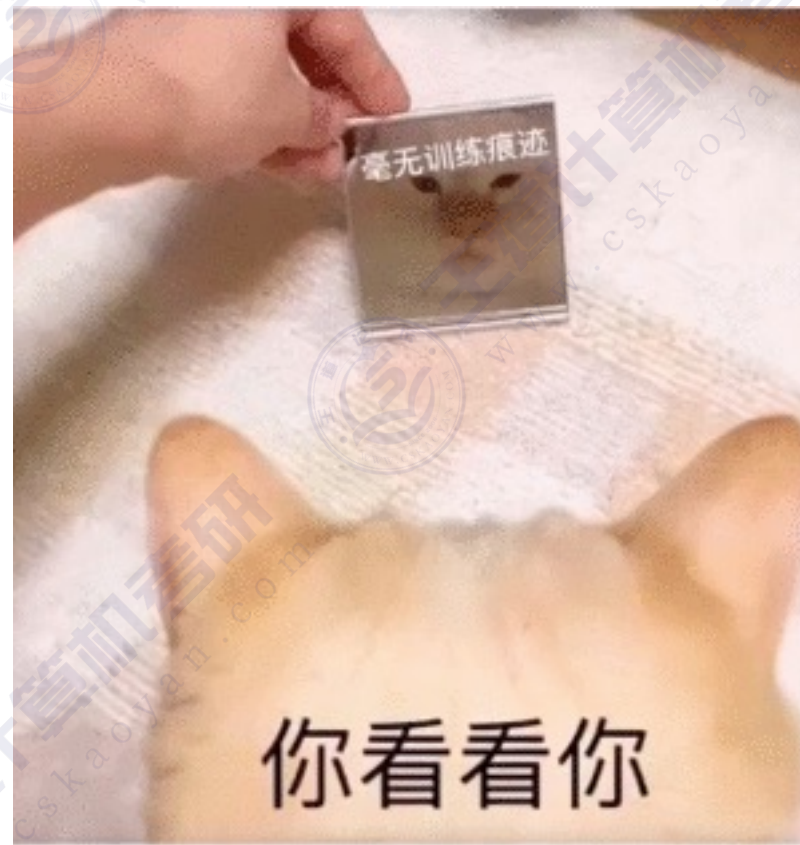
	IF (取指)	ID (译码/取数)	EX (执行/计算地址)	M (访存)	WB (写回)
Load指令	取指令	译码、取数（可能会读寄存器）	计算访存地址	从主存（Data Cache）读数据	将数据写回某个寄存器
Store指令	取指令	译码、取数（可能会读寄存器）	计算访存地址	往主存（Data Cache）写数据	空
转移类指令	取指令	译码、取数（可能会读寄存器）	计算转移地址	更新PC值	空
运算类指令	取指令	译码、取数（可能会读寄存器）	执行（算术运算、逻辑运算等）	空	将数据写回某个寄存器

注：熟悉四类指令在各个段分别做了什么

自主训练

先集中训练以下题目，找找感觉：

- 真题实战——2016年 19题
- 真题实战——2019年 18题
- 真题实战——2023年 19题
- 真题实战——2012年 44题 (2)(3)(4) 小问
- 真题实战——2014年 44题 (4) 小问



数据冒险的分析和处理

考数据冒险？

1. 分析是否存在数据冒险？

核心特点：两条指令，前“写”后“读”同一个寄存器。

分析方法：从第一条指令开始，结合注释，观察该指令“写”了某个寄存器，并观察后续相邻的3条指令是否“读”了同一个寄存器。若发现此类情况，则存在数据冒险。

2. 常见的解决数据冒险的方法

- a. 硬件阻塞：将“读寄存器指令”的ID段硬件阻塞若干个时钟，确保其ID段在“写寄存器指令”的WB段之后
- b. 转发（旁路）技术
 - i. 转发（旁路）技术可以解决大部分数据冒险？
 - ii. 转发（旁路）技术不能解决由 Load 指令写寄存器引起的 Load-use 数据冒险？

控制冒险的分析和处理

考控制冒险？

1.分析是否存在控制冒险？

只有**转移类指令**的执行才会引发控制冒险。因此只需观察指令序列中是否包含“转移类指令”即可。

2.常见的解决控制冒险的方法

硬件阻塞：将转移类指令**后一条指令**的IF段硬件阻塞3个时钟

19. 在无转发机制的五段基本流水线（取指、译码/读寄存器、运算、访写回寄存器）中，下列指令序列存在数据冒险的指令对是（ ）。

I1: add R1, R2, R3 ;(R2) + (R3) → R1

I2: add R5, R2, R4 ;(R2) + (R4) → R5

I3: add R4, R5, R3 ;(R5) + (R3) → R4

I4: add R5, R2, R6 ;(R2) + (R6) → R5

A. I1 和 I2

B. I2 和 I3

C. I2 和 I4

D. I3 和 I4

无转发机制，只能用硬件阻塞来处理数据冒险

真题实战：数据冒险

2016

19. 在无转发机制的五段基本流水线（取指、译码/读寄存器、运算、访写回寄存器）中，下列指令序列存在数据冒险的指令对是（ ）。

I1: add R1, R2, R3 ;(R2) + (R3)→R1

I2: add R5, R2, R4 ;(R2) + (R4)→R5

I3: add R4, R5, R3 ;(R5) + (R3)→R4

I4: add R5, R2, R6 ;(R2) + (R6)→R5

A. I1 和 I2

B. I2 和 I3

C. I2 和 I4

D. I3 和 I4

时钟 1 2 3 4 5 6 7 8 9 10 11 12

指令1

IF	ID	EX	M	WB							
	IF	ID	EX	M	WB 写R5						
		IF	阻塞	阻塞	阻塞	ID 读R5	EX	M	WB		
						IF	ID	EX	M	WB	

指令2

指令3

指令4

18. 在采用“取指、译码/取数、执行、访存、写回”5段流水线的处理器中，执行如下指令序列，其中 s0、s1、s2、s3 和 t2 表示寄存器编号。

```
I1: add s2,s1,s0      //R[s2]←R[s1]+R[s0]
I2: load s3,0(t2)     //R[s3]←M[R[t2]+0]
I3: add s2,s2,s3      //R[s2]←R[s2]+R[s3]
I4: store s2,0(t2)    //M[R[t2]+0]←R[s2]
```

下列指令对中，不存在数据冒险的是（ ）。

A. I1 和 I3

B. I2 和 I3

C. I2 和 I4

D. I3 和 I4

本题未说明如何处理数据冒险

真题实战：数据冒险

2019

18. 在采用“取指、译码/取数、执行、访存、写回”5段流水线的处理器中，执行如下指令序列，其中 s0、s1、s2、s3 和 t2 表示寄存器编号。

```
I1: add s2,s1,s0      //R[s2]←R[s1]+R[s0]
I2: load s3,0(t2)     //R[s3]←M[R[t2]+0]
I3: add s2,s2,s3      //R[s2]←R[s2]+R[s3]
I4: store s2,0(t2)    //M[R[t2]+0]←R[s2]
```

假设没有转发技术，采用硬件阻塞处理数据冒险

下列指令对中，不存在数据冒险的是（ ）。

A. I1 和 I3

B. I2 和 I3

C. I2 和 I4

D. I3 和 I4

时钟	1	2	3	4	5	6	7	8	9	10	11	12	13	14
指令1	IF	ID	EX	M	WB 写R[s2]									
指令2		IF	ID	EX	M	WB 写R[s3]								
指令3			IF	阻塞	阻塞	阻塞	ID 读R[s2] 读R[s3]	EX	M	WB 写R[s2]				
指令4							IF	阻塞	阻塞	阻塞	ID 读R[s2]	EX	M	WB

44. (12 分) 某 16 位计算机中，带符号整数用补码表示，数据 Cache 和指令 Cache 分离。题 44 表给出了指令系统中部分指令格式，其中 Rs 和 Rd 表示寄存器，mem 表示存储单元地址，(x)表示寄存器 x 或存储单元 x 的内容。

题 44 表 指令系统中部分指令格式

名称	指令的汇编格式	指令功能
加法指令	ADD Rs, Rd	$(Rs) + (Rd) \rightarrow Rd$
算术/逻辑左移	SHL Rd	$2 * (Rd) \rightarrow Rd$
算术右移	SHR Rd	$(Rd)/2 \rightarrow Rd$
取数指令	LOAD Rd, mem	$(mem) \rightarrow Rd$
存数指令	STORE Rs, mem	$(Rs) \rightarrow mem$

该计算机采用 5 段流水方式执行指令，各流水段分别是取指 (IF)、译码/读寄存器 (ID)、执行/计算有效地址 (EX)、访问存储器 (M) 和结果写回寄存器 (WB)，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，并且同一个寄存器的读和写操作不能在同一个时钟周期内进行。请回答下列问题。

- 1) 若 int 型变量 x 的值为 -513，存放在寄存器 R1 中，则执行指令“SHR R1”后，R1 的内容是多少？（用十六进制表示。）
- 2) 若某个时间段中，有连续的 4 条指令进入流水线，在其执行过程中没有发生任何阻塞，则执行这 4 条指令所需的时钟周期数为多少？
- 3) 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水线中的执行过程如题 44 图所示。

I_1 LOAD R1, [a]
 I_2 LOAD R2, [b]
 I_3 ADD R1, R2
 I_4 STORE R2, [x]

	时间单元													
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							IF				ID	EX	M	WB

题 44 图 指令序列及其执行过程示意图

则这 4 条指令执行过程中， I_3 的 ID 段和 I_4 的 IF 段被阻塞的原因各是什么？

- 4) 若高级语言程序中某赋值语句为 $x = 2 * x + a$ ， x 和 a 均为 unsigned int 类型变量，它们的存储单元地址分别表示为[x]、[a]，则执行这条语句至少需要多少个时钟周期？要求模仿题 44 图画出这条语句对应的指令序列及其在流水线中的执行过程示意图。

44. (12 分) 某程序中有如下循环代码段 P: “for(int i = 0; i < N; i++) sum += A[i];”。假设编译时变量 sum 和 i 分别分配在寄存器 R1 和 R2 中。常量 N 在寄存器 R6 中，数组 A 的首地址在寄存器 R3 中。程序段 P 起始地址为 0804 8100H，对应的汇编代码和机器代码如下表所示。

编号	地址	机器代码	汇编代码	注释
1	08048100H	00022080H	loop: sll R4, R2, 2	(R2) << 2 → R4
2	08048104H	00832020H	add R4, R4, R3	(R4) + (R3) → R4
3	08048108H	8C850000H	load R5, 0(R4)	((R4) + 0) → R5
4	0804810CH	00250820H	add R1, R1, R5	(R1) + (R5) → R1
5	08048110H	20420001H	addi R2, R2, 1	(R2) + 1 → R2
6	08048114H	1446FFFAH	bne R2, R6, loop	if (R2) != (R6) goto loop

执行上述代码的计算机 M 采用 32 位定长指令字，其中分支指令 bne 采用如下格式：

31	26	25	21	20	16	15	0
OP	Rs	Rd	OFFSET				

OP 为操作码；Rs 和 Rd 为寄存器编号；OFFSET 为偏移量，用补码表示。请回答下列问题，并说明理由。

- 1) M 的存储器编址单位是什么？
- 2) 已知 sll 指令实现左移功能，数组 A 中每个元素占多少位？
- 3) 表中 bne 指令的 OFFSET 字段的值是多少？已知 bne 指令采用相对寻址方式，当前 PC 内容为 bne 指令地址，通过分析表中指令地址和 bne 指令内容，推断出 bne 指令的转移目标地址计算公式。
- 4) 若 M 采用如下“按序发射、按序完成”的 5 级指令流水线：IF（取值）、ID（译码及取数）、EXE（执行）、MEM（访存）、WB（写回寄存器），且硬件不采取任何转发措施，分支指令的执行均引起 3 个时钟周期的阻塞，则 P 中哪些指令的执行会由于数据相关而发生流水线阻塞？哪条指令的执行会发生控制冒险？为什么指令 1 的执行不会因为与指令 5 的数据相关而发生阻塞？

19. 在采用“取指、译码/取数、执行、访存、写回”5段流水线的RISC处理器中，执行如下指令序列（第一列为指令序号），其中s0、s1、s2、s3和t2表示寄存器编号。

I1	add	s2, s1, s0	// $R[s2] \leftarrow R[s1] + R[s0]$
I2	load	s3, 0(s2)	// $R[s3] \leftarrow M[R[s2] + 0]$
I3	beq	t2, s3, L1	// if $R[t2] = R[s3]$ jump to L1
I4	addi	t2, t2, 20	// $R[t2] \leftarrow R[t2] + 20$
I5	L1:	

若采用转发（旁路）技术处理数据冒险，采用硬件阻塞方式处理控制冒险，则在指令 I1~I4 执行过程中，发生流水线阻塞的指令有（ ）。

- A. 仅 I3 B. 仅 I2、I4 C. 仅 I3、I4 D. 仅 I2、I3、I4

真题实战：数据冒险——采用转发技术

2023

注意：

- i. 转发（旁路）技术可以解决大部分数据冒险
- ii. 转发（旁路）技术不能解决由 Load 指令写寄存器引起的 Load-use 数据冒险

真题实战：数据冒险——采用转发技术

2023

I1 add s2, s1, s0 // $R[s2] \leftarrow R[s1] + R[s0]$
 I2 load s3, 0(s2) // $R[s3] \leftarrow M[R[s2] + 0]$
 I3 beq t2, s3, L1 // if $R[t2] = R[s3]$ jump to L1
 I4 addi t2, t2, 20 // $R[t2] \leftarrow R[t2] + 20$
 I5 L1:

时钟	1	2	3	4	5	6	7	8	9	10	11	12	13	14
指令1	IF	ID 读R[s0] 读R[s1]	EX ALU已 算出 R[s2]	M	WB 写R[s2]									
指令2		IF	ID 读R[s2]	EX 转发技 术运算 R[s2]+0	M 从主存 取数	WB 写R[s3]								
指令3			IF	阻塞	阻塞	阻塞	ID 读R[s2] 读R[s3]	EX	M	WB 写R[s2]				
指令4 或指令							阻塞	阻塞	阻塞	IF	ID 读R[s2]	EX	M	WB