

中国科学技术大学

2020 年春季学期期末考试试卷

课程名称: 计算机体系结构 课程代码: 011135
 开课院系: 计算机科学与技术学院 考试形式: 闭卷
 姓 名: 学 号:

1、(10 分) 某台主频为 500MHz 的计算机执行标准测试程序, 程序 gcc 中的指令类型、执行条数百分比和时钟周期数如表 1 所示。

表 1 gcc 中的指令类型和执行数量百分比

指令类型	执行条数百分比(%)
load	26.5
store	10.3
add	21.1
sub	1.7
mul	1.4
compare	2.8
load imm	4.8
cond branch	9.3
cond move	0.4
jump	0.8
其他	20.9

表 2 不同类型指令执行的时钟周期

指令类型	时钟周期
所有 ALU 指令	1.0
loads-stores (载入-存储)	1.4
Jump (无条件跳转)	1.2
Conditional branches (条件分支)	
Taken (选中)	2.0
Not Taken (未选中)	1.5

假设: 60%的 conditional branch 指令结果是 taken。所有表一里面的“其他”指令都为 ALU 指令。求该计算机执行 gcc 的有效 CPI。请注意不需要把最终结果算出来, 列出式子即可。

2、（10分）简要回答下列有关ISA的问题：

- (1) RISC-V 指令集支持变长指令吗？其基础指令集（RV32I）是定长还是变长？
- (2) 在早期的 RISC 指令集架构中，ISA 对位置无关代码(position-independent code)支持不足。请给出一个位置有关代码的例子。在 RISC-V 指令集中，是如何改善这一点的？请举例说明。

3、（10分）简要回答下列有关流水线的问题：

- (1) 假设我们有一个 5 阶段基本流水线（IF, ID, EXE, MEM, WB）。程序 P1 在该流水线执行没有任何相关。P1 共有 n 条指令，请问需要多少时钟周期完成该程序的执行。CPI 为多少？
- (2) 假设同样的流水线。程序 P2 只包含 n 条整数运算指令，其中 60%的指令只和其前面紧邻的指令有数据相关（data hazard）；20%的指令和前面紧邻的第 2 条指令有数据相关。程序 P2 没有任何其他的相关。P2 共有 n 条指令，请问需要多少时钟周期完成该程序的执行？CPI 为多少？
- (3) 假设该 5 阶段流水线拥有旁路定向能力（forwarding）。程序 P3 的指令中 40%是 load 指令，其中 60%的 load 指令和其后紧邻的指令有相关（RAW 相关）。P3 共有 n 条指令，请问需要多少时钟周期完成该程序的执行？CPI 为多少？

4、（10分）简要回答下列有关指令级并行的问题：

- (1) 简短描述Tomasulo是如何处理RAW、WAR、WAW 相关（Hazards）的。
- (2) 请简短描述何为精确中断(precise interrupts)。并请简短描述Tomasulo是如何实现精确中断的。

5、（10分）回答下列有关指令级并行的问题：

假设我们有如下scoreboard结构

Instruction Status:

instruction	i	j	k	Issue	Read operand	Execution complete	Write result
LD	F6	34+	R2				
LD	F2	45+	R3				
MULT	F0	F2	F4				
SUBD	F8	F6	F2				
DIVD	F10	F0	F6				
ADDD	F6	F8	F2				

Functional unit status:

	latency	Busy	Op	Dest Fi	S1 Fj	S2 Fk	FU Qj	FU Qk	Fj? Rj	Fk? Rk
Integer	1									
Mult1	6									
Mult2	6									
Add	2									
Divide	30									

Register result status:

	F0	F2	F4	F6	F8	F10	F12	...	F30
FU									

请填写好所有指令执行完成后的每条指令instruction status的状态细节。（只需要填写 Instruction Status 表格）

6、（10 分）简要回答以下关于 Cache 的问题：

直接映像 Cache 和两路组相联 Cache，试问他们对 CPU 性能的影响？先求平均访存时间，然后再计算 CPU 性能。分析时请用以下假设：

- （1）理想 Cache(命中率为 100%) 情况下 CPI 为 1.0，时钟周期为 0.35ns，平均每条指令访存 1.4 次
- （2）两种 Cache 容量均为 128KB，块大小都是 64B
- （3）采用组相联时，由于多路选择器的存在，时钟周期增加到原来的 1.35 倍
- （4）两种结构的失效开销都是 65ns (hint：应取整为整数个时钟周期)
- （5）命中时间为 1 个 cycle，128KB 直接映像 Cache 的失效率为 2.1%，相同容量的两路组相联 Cache 的失效率为 1.9%

7、（10 分）简要回答以下关于 GPU 的问题：

假设我们调度 8K CUDA threads(64 CUDA thread blocks, with 128 threads per block)在 NVIDIA GTX 680, 该 GPU 有 8 个 SM 核，每个核有 16 个 SP。

- （1）如果一个 warp 有 32 个 threads，每个 SM 核执行多少个 warp？
- （2）在每个 warp 里，我们挑选两条可执行指令来执行，这是哪种并行？
- （3）一个 warp 里的所有 threads 在每个时钟周期执行同一条指令。这是哪种并行？

8、（10 分）考虑如下运行在主频为 1GHz、向量寄存器长度为 64 的 VMIPS 机器上的向量指令序列：

LV V1, Ra
LV V3, Rc
MULV.D V2, V1, V3
ADDV.D V4, V2, V3
SV Rb, V4

各功能部件的启动时间如表 2 所示。

表 2 功能部件启动时间

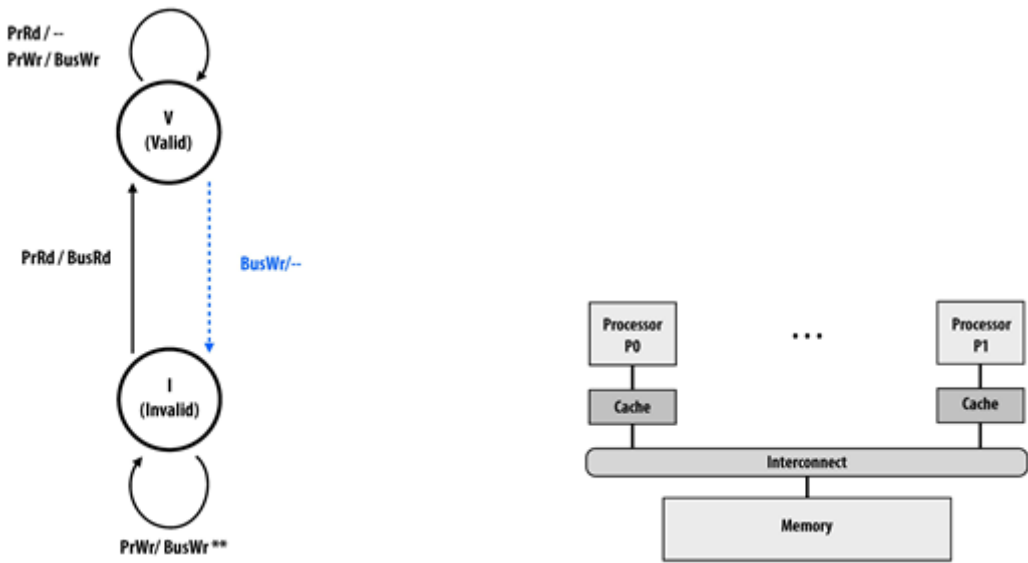
Unit	Start-up overhead (cycles)
Load and Store unit	8
Multiply unit	7
Add unit	4

假设采用分段开采 (strip-mining) 技术来进行向量长度为 200 的向量运算，分段开采的额外开销（循环所涉及的开销）为 10 个时钟周期。

- （1）假设不使用链接技术，并且只有单条存储器访问流水线，则该向量指令序列执行需要多少个 cycles？
- （2）如果使用链接技术，并且只有单条存储器访问流水线，则在考虑功能部件启动时间时，计算平均产生一个结果所花费的时钟周期数。
- （3）假设 VMIPS 有三条访存流水线，并且采用链接技术。如果在执行过程中没有存储器访问冲突，试计算平均产生一个结果所需要的时钟周期数。

9、（10 分）简要回答以下关于 Cache Coherence 的问题：

（1）在一个简单的实现里，我们假设一个 write-through cache。我们的 coherence 协议和系统架构如下图所示：



请填写以下表格：

Action	Bus activity	P0\$	P1\$	Mem[x]
		--	--	0
P0 load x	BusRd	0	--	0
P1 load x				
P0 write 200 to x				
P1 load x				

(2) 请简要解释 invalidation（写作废）和 update（写更新）两种协议各自的优缺点。

10、（10 分）考虑以下一致性问题。

（1） 初始值： A=0； B=0；

Thread 1	Thread 2	Thread 3
<pre>A = 1;</pre>	<pre>while (!A); B = 0;</pre>	<pre>while (B) ; print A; print B;</pre>

在一个满足顺序一致性(sequential consistency)的系统中，以上代码的输出为何？

(2) TSO 代表 Total Store Ordering，其放松了 W→R 关系。请问在一个 TSO 系统中，以上代码的输出为何？