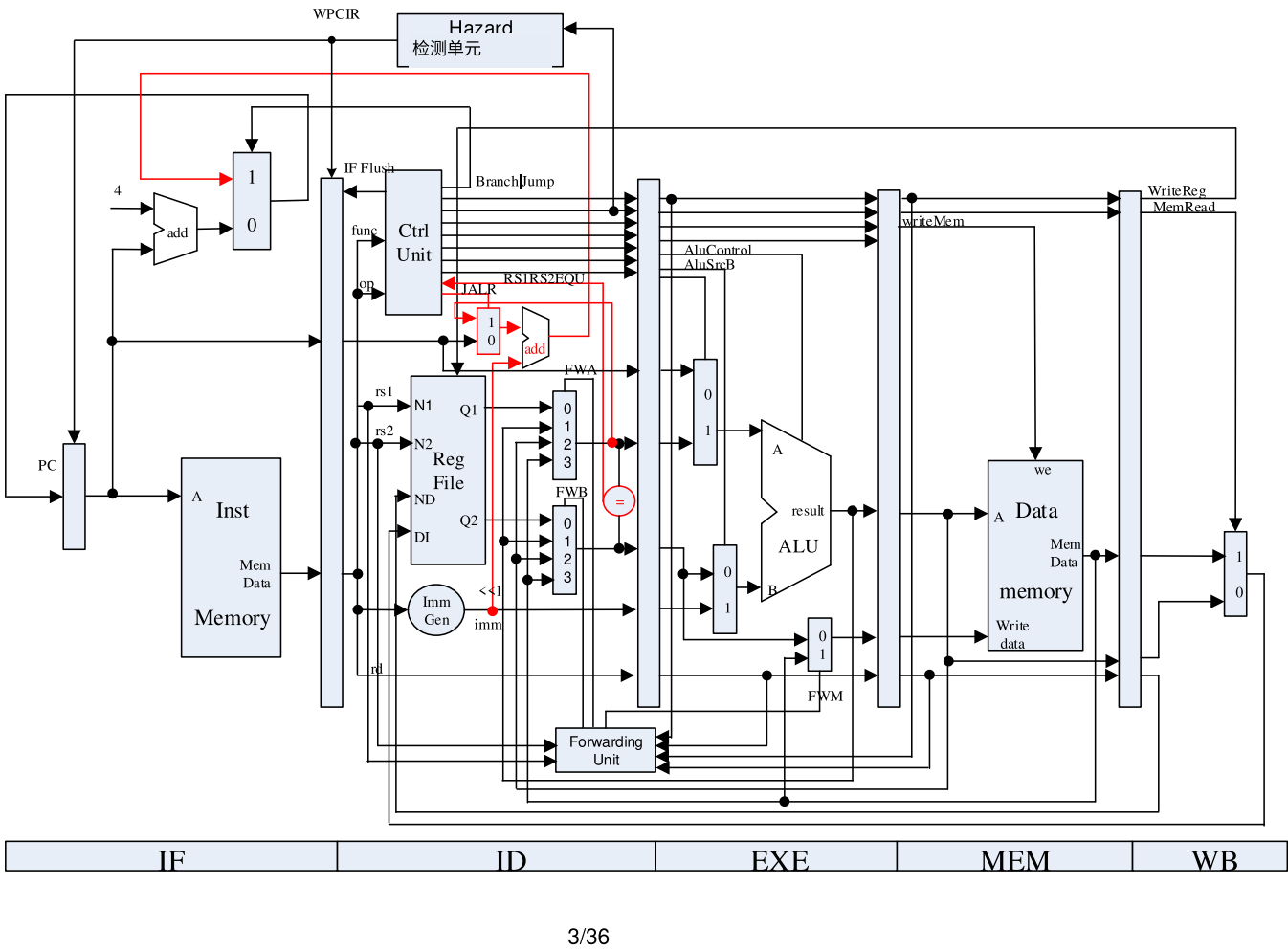


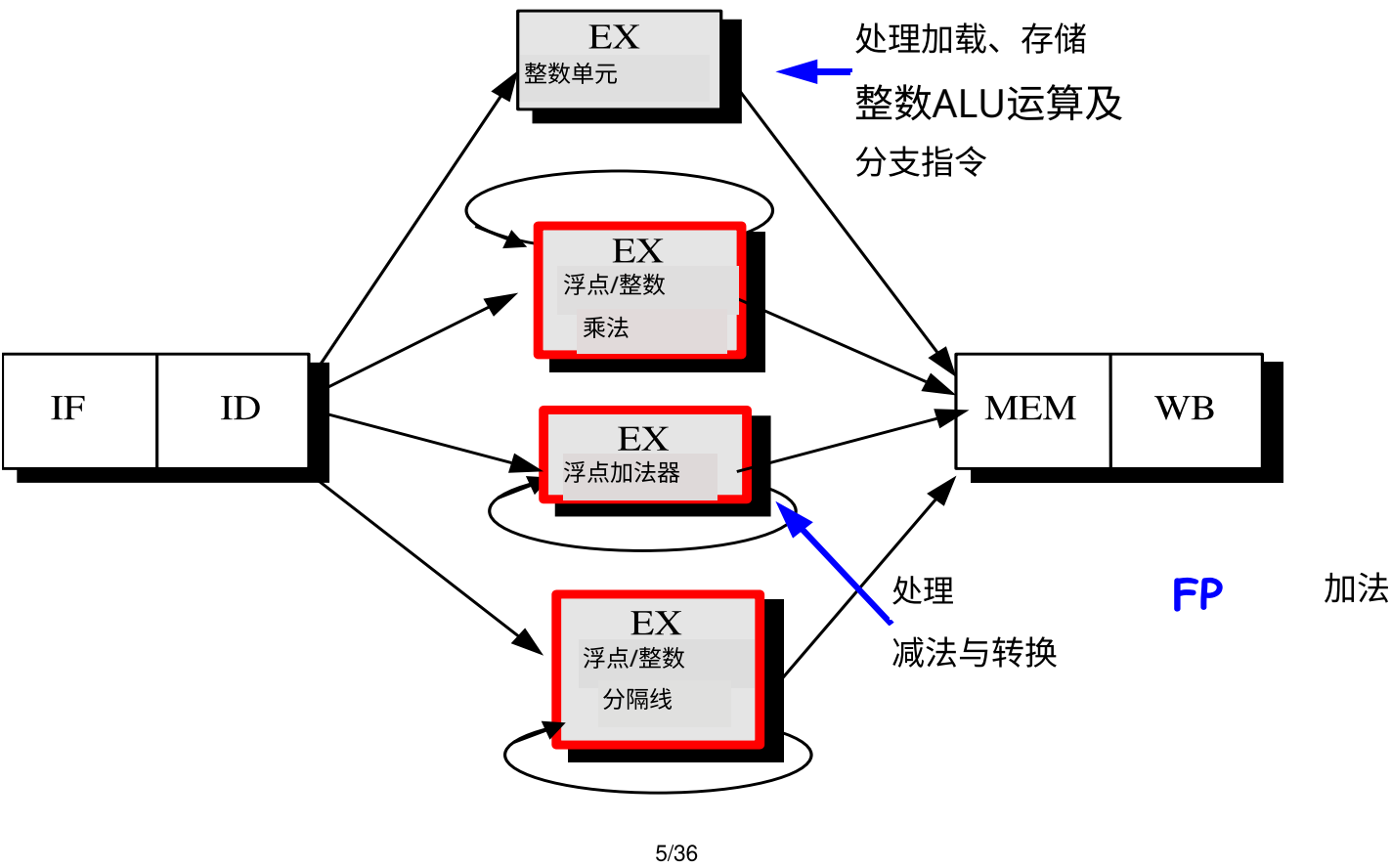
Ch3--
ILP及其
探索

第3章-0 扩展五级流水线以支持多周期操作
附录C5-C6

支持RISC V的流水线CPU



带浮点单元的五级流水线



功能单元的延迟与启动间隔

| 功能单元 | 延迟周期 | 启动间隔 |
|----------------|------|------|
| 整数算术逻辑单元 | 0 | 1 |
| 数据存储器（整数与浮点加载） | 1 | 1 |
| 浮点加法 | 3 | 1 |
| 浮点乘法（含整数乘法） | 6 | 1 |
| 浮点除法（含整数除法） | 24 | 25 |

注：延迟周期=功能单元时间-1个时钟周期

结构冲突

- 这些是硬件资源使用上的冲突
- 增加额外硬件资源；功能单元全流水线化；否则仍需停顿

•允许存在结构冲突的机器，因其发生频率不高»数据冲突

•指令依赖于尚未就绪（未完成计算或存储）的前序运算结果

•停顿；双缓冲；转发路径；编译器调度

控制冒险

- 分支条件与分支程序计数器未能及时就绪，导致下一时钟周期无法获取指令
- 冲刷流水线；预测执行、预测不执行、延迟分支
- 尽可能前移目标地址计算与条件比较的时序

扩展MIPS流水线以处理多周期操作

□ 处理浮点运算（或复杂运算）的替代方案

- 在1或2个时钟周期内完成操作，
 - 这意味着采用低速时钟，
 - 和/或在浮点单元中使用大量逻辑电路。
- > 允许操作具有更长延迟
 - 执行周期可按需重复多次以完成操作
- 可配置多个浮点单元

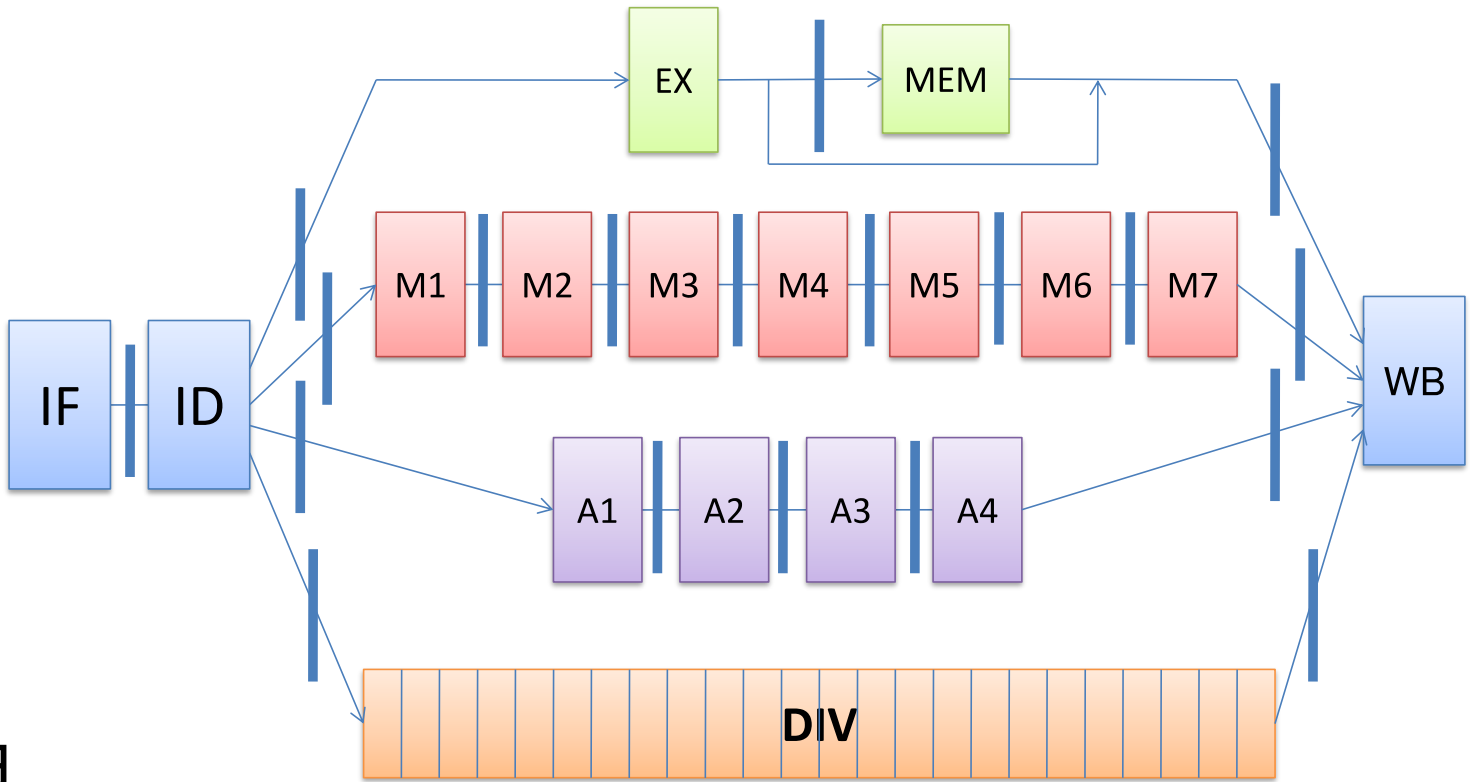
对部分浮点运算单元进行流水线化

□ 两种术语

- > 延迟周期----从产生结果的指令到使用该结果的指令之间的间隔周期数
- > 启动间隔----同一运算单元连续执行指令间必须间隔的周期数

- 全流水线单元的启动间隔为1
- 非流水线单元的启动间隔恒等于延迟周期加1

流水线支持多个未完成的浮点运算操作



规格说明

- ☐ 内存带宽：双字/周期

☐ 需新增流水线锁存器：

> M1/M2、M2/M3、M3/M4、M4/M5、M5/M6、M6/M7

> A1/A2, A2/A3, A3/A4

☐ 需新增连接寄存器：

> ID/EX、ID/M1、ID/A1、ID/DIV

> 执行/访存阶段、M7/访存、A4/访存、除法/访存

☐ 由于除法单元未采用流水线结构，可能发生结构冲突。

☐ 因指令执行时间差异，单个周期内所需的寄存器写入次数可能超过1次

☐ 新增数据冲突：乱序写回可能导致写后写冲突

☐ 操作延迟增长将导致RAW冲突停顿周期更为频繁。

无序完成导致的异常问题

9/36

浮点寄存器写入端口的结构性冲突

| 指令 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------------------|----|----|----|----|-----|-----|----|-----|-----|----|
| 浮点乘法指令 F0,F4,F6 | IF | ID | M1 | M2 | M3 | M4 | M5 | M6 | M7 | WB |
| | | IF | ID | EX | MEM | WB | | | | |
| | | | IF | ID | EX | MEM | WB | | | |
| 浮点加法指令 F2,F4,F6 | | | | IF | ID | A1 | A2 | A3 | A4 | WB |
| | | | | | IF | ID | EX | MEM | WB | |
| 加载双精度浮点数 F8, Q(R2) | | | | | | IF | ID | EX | MEM | WB |

11/36

数据冲突类型

指令流

指令B

指令A

CPU

☐ RAW(写后读)真依赖

> 指令A写入Rx，指令B读取Rx

☐ WAW(写后写)输出依赖

> 指令A写入Rx，指令B写入Rx

☐ WAR(读后写)反依赖

> 指令A读取Rx，指令B写入Rx

☐ 冒险类型根据必须保持的执行顺序命名由流水线

加法指令 x5, x8,x10

乘法指令 x6, x5, x9

加法指令 x5, x8, x10

乘法指令 x5, x6, x9

加法指令 x5, x8, x9

乘法指令 x8, x6, x10

13/36

写后写依赖-命名依赖1

- ☐ B试图在A写入前覆盖操作数。→ 当指令 B 执行后，寄存器本应存储B的结果，实际却保留了A的结果。
- ☐ 这种情况仅发生在多级写入的流水线中，或变长流水线（如浮点流水线）。

15/36

按序发射

乱序完成

| 指令 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|-------|----|----|----|----|----|----|-----|-----|----|----|----|----|
| MUL.D | IF | ID | M1 | M2 | M3 | M4 | M5 | M6 | M7 | WB | | |
| ADD.D | | IF | ID | A1 | A2 | A3 | A4 | WB | | | | |
| MUL.D | | | IF | ID | M1 | M2 | M3 | M4 | M5 | M6 | M7 | WB |
| LD.D | | | | IF | ID | EX | MEM | WB | | | | |
| SD.D | | | | | IF | ID | EX | MEM | WB | | | |

- 1、单时钟周期多次写入
- 2、乱序写入：不同/相同目标寄存器

10/36

如何解决写端口冲突？

- ☐ 增加写端口数量X

> 毫无吸引力！

> 稳态使用率接近1，无实际价值

☐ 通过串行化写入操作检测并插入停顿

> 在指令译码阶段跟踪写端口使用并实施停顿指令在发出前

•附加硬件：移位寄存器+写冲突逻辑

•移位寄存器追踪已发出指令对寄存器的占用情况文件，并在每个时钟周期右移1位

•停顿可能加剧数据冲突

•所有互锁检测和停顿插入都发生在ID阶段

> 当冲突指令试图进入MEM或WB阶段时将其暂停。

•此时易于检测冲突

•由于暂停可能发生在两处，增加了流水线控制复杂度
- 12/36
- RAW相关性：真实数据依赖
- ☐ B 在 A 写入寄存器前尝试读取，获取到旧值。

☐ 这是常见情况，转发有助于解决该问题。
-
- 14/36
- WAR依赖 - 命名依赖2
- ☐ B 试图在 A 读取前写入寄存器。

☐ 此种情况下，A使用了新的（错误）值。

☐ 此类冒险情况较为罕见，因为多数流水线会尽早读取值并延迟写入结果。

☐ 但对于采用复杂寻址模式（如自动递增）的CPU仍可能发生。
-
- 16/36

由RAW冒险引发的停顿

| 指令 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|-----------------------|----|----|----|-------|----|-------|----|----|----|----|-----|----|----|----|----|-----|
| | | | | | | | | | | | | | | | | |
| 从R2寄存器偏移0处加载双精度浮点数到F4 | IF | ID | EX | MEM | WB | | | | | | | | | | | |
| 双精度浮点乘法 F0 = F4 × F6 | | IF | ID | 流水线停顿 | M1 | M2 | M3 | M4 | M5 | M6 | M7 | WB | | | | |
| 双精度浮点加法 F2 = F0 + F8 | | | IF | 流水线停顿 | ID | 流水线停顿 | 失速 | 摊位 | 畜栏 | 拖延 | 抛锚 | A1 | A2 | A3 | A4 | WB |
| SD.D F2, 0(R2) | | | | | IF | 失速 | 摊位 | 畜栏 | 拖延 | 抛锚 | 小隔间 | ID | EX | 停顿 | 阻塞 | MEM |

解决WAW冒险

- ❑ 将可能"越过"前一条指令的后续指令阻塞，直到前一条指令进入MEM阶段
- ❑ 取消前一条指令的WB阶段
- ❑ 这两项操作都可在ID阶段完成，即当LD指令即将发射时
- ❑ 由于纯WAW冒险并不常见，任一种方法都可行。

选择最容易实现的方式。

- ❑ RISC V流水线的最简解决方案是：当指令要写入的寄存器与已发射指令相同时，将其阻滞在ID阶段。

需在ID阶段进行检测

- ❑ 检测结构冒险
 - 除法等器等非全流水功能单元
- ❑ 检测RAW冒险

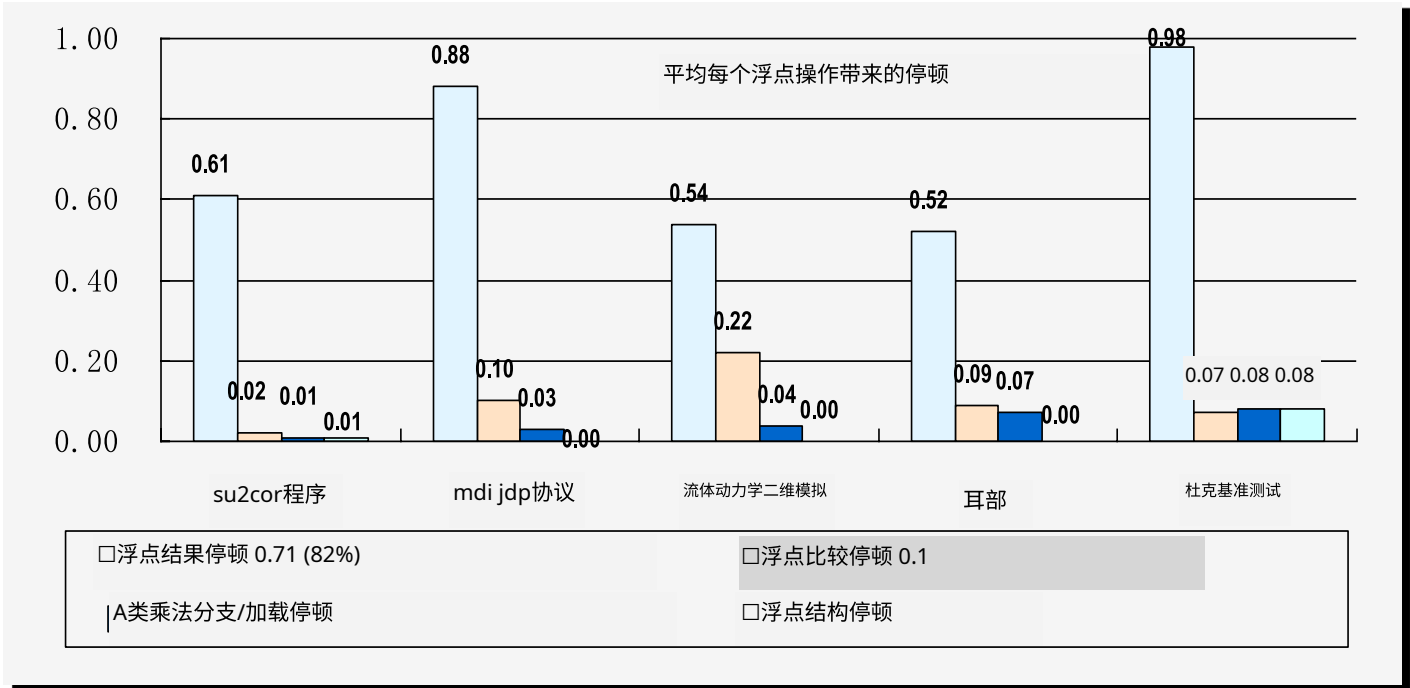
CPU会简单地将指令在ID阶段暂停，直到满足以下任一条件： - 其源寄存器不再被列为任何执行流水线寄存器（位于 M 和 A 阶段之间的寄存器）中的目标寄存器；或- 其源寄存器不再被列为EX/MEM寄存器中的加载目标。

- ❑ 检查WAW冲突
 - 检查A1至A4、Divide或M1至M7阶段的指令是否存在相同目标寄存器（需检查流水线寄存器）
 - 若无WAW冲突，则预留寄存器写入端口。

- ❑ 必要时在ID阶段暂停指令

注：仅预留写入端口
仅当ID阶段指令可被发射时执行

浮点流水线性能



WAW冒险

| 指令 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 16 |
|--------------------|---|----|----|----|-----|-----|----|----|----|----|----|----|----|
| 浮点乘法 F0, F4, F6 IF | | ID | M1 | M2 | M3 | M4 | M5 | M6 | M7 | WB | | | |
| | | IF | ID | EX | MEM | WB | | | | | | | |
| | | | IF | ID | EX | MEM | WB | | | | | | |
| 浮点加法 F2, F4, F6 | | | | IF | ID | 停顿 | A1 | A2 | A3 | A4 | WB | | |
| 双字加载 F2, 0(R2) | | | | | IF | 档位 | ID | 失速 | 畜栏 | EX | DM | WB | |
| | | | | | | | IF | 摊位 | 拖延 | ID | EX | DM | WB |
| 双字加载 F8, 0(R2) | | | | | | | | | | IF | ID | EX | DM |

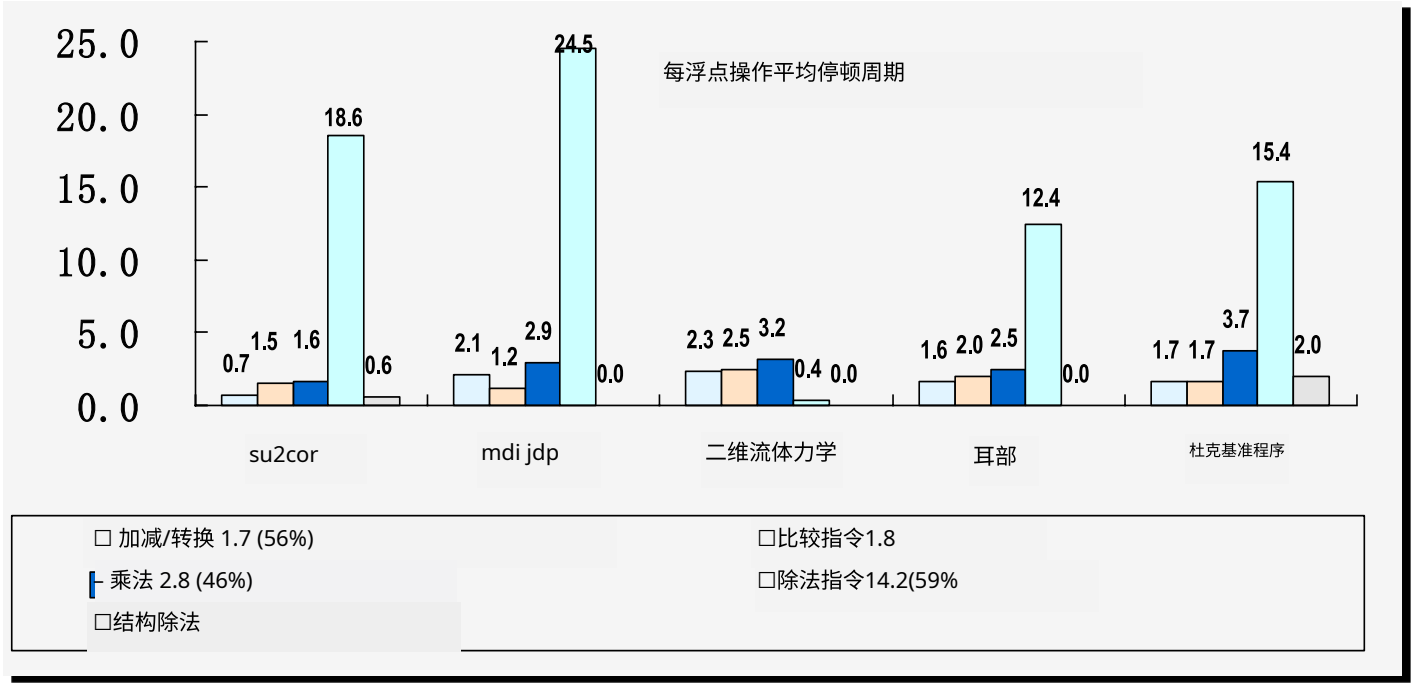
* 假设ADD指令后两条指令存在 n 写寄存器冲突。
1836

还可能存在哪些其他冲突？

- ❑ 浮点指令间的冲突
- ❑ 浮点指令与整型指令间的冲突

➢ 由于存在两个寄存器文件，仅浮点加载/存储及浮点寄存器向整型寄存器的移动会产生冲突

浮点流水线性能



MIPS R4000流水线

- ❑ IF-指令获取前半周期。进行程序计数器选择，启动缓存访问。

- ❑ IS-指令获取后半周期。
 - 这使得缓存访问可占用两个时钟周期。
- ❑ RF-译码与寄存器读取、冒险检测、一级缓存命中判断。

- ❑ EX-执行阶段：地址计算、算术逻辑单元操作、分支目标计算及条件评估。

❑ DF/DS/TC

- 前两个周期从缓存中获取数据。
- 第三周期需进行标记检查以确定缓存访问是否命中。

- ❑ WB-为加载和R-R操作写回结果。

可能的停顿与延迟

加载延迟：两个周期

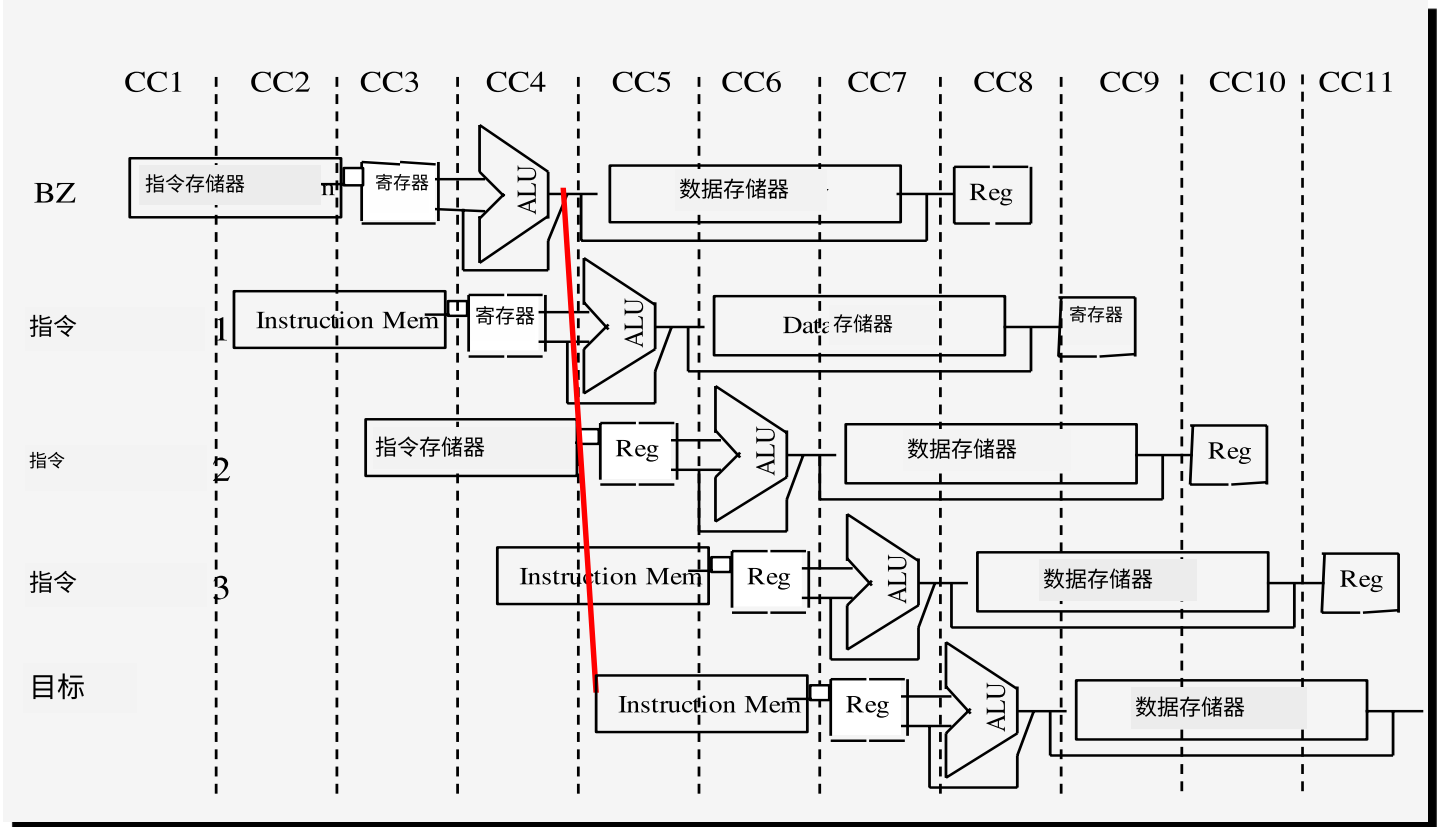
表面上看延迟可能是三个周期，因为标记检查需等到TC周期结束时才进行。

>但若TC显示未命中，则需从主存获取数据，并回滚流水线以取得真实值。

示例：加载停顿

| 指令 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|-----------|----|----|----|----|----|----|----|----|----|
| LW R1 | IF | IS | RF | EX | DF | DS | TC | WB | |
| 将R1加到R2 | | IF | IS | RF | 暂停 | 暂停 | EX | DF | DS |
| SUB R3,R1 | | | IF | IS | 暂停 | 暂停 | RF | EX | DF |
| OR R4 ,R1 | | | | IF | 暂停 | 停顿 | IS | RF | EX |

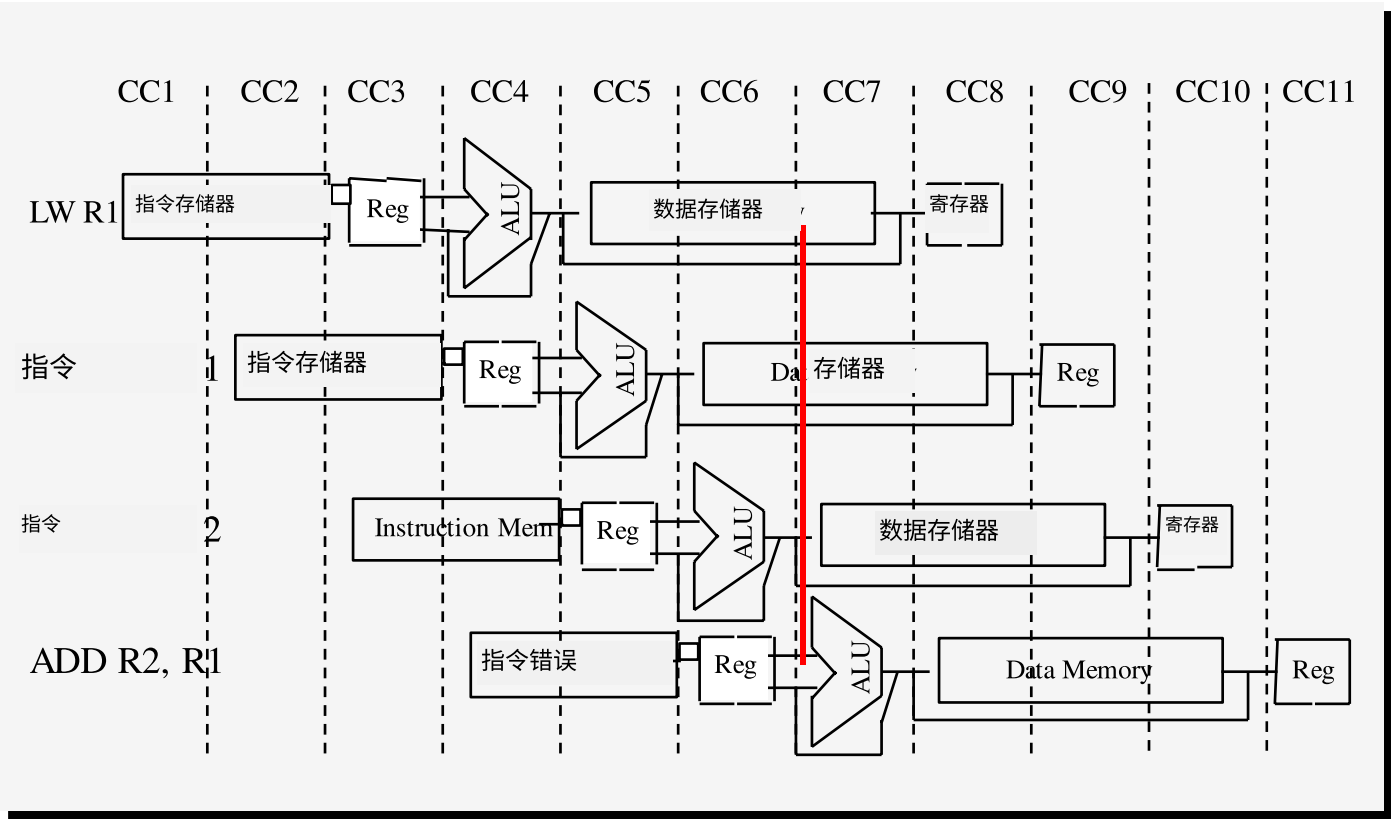
分支延迟：3个停顿周期



FP 8级操作流水线

| 阶段 | 功能单元 | 描述 |
|----|-------|---------|
| A | 浮点加法器 | 尾数加法阶段 |
| D | 浮点除法器 | 除法流水线阶段 |
| E | 浮点乘法器 | 异常测试阶段 |
| M | 浮点乘法器 | 乘法器第一阶段 |
| N | 浮点乘法器 | 乘法器第二阶段 |
| R | 浮点加法器 | 舍入阶段 |
| S | 浮点加法器 | 操作数移位阶段 |
| U | | 解包浮点数 |

加载停顿 - 2个时钟周期



分支延迟：3个周期

分支延迟：三个周期（包含一个分支延迟槽）

该分支在EX阶段被解析，导致3个周期的延迟。首周期可能是常规分支延迟槽（指令必定执行）或分支预测槽（若分支未采纳则取消指令）。

MIPS采用预测不采纳方法，推测因其对硬件需求最低。

分支延迟的流水线状态

| 指令 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|------|----|----|----|----|----|----|----|----|----|
| 分支指令 | IF | IS | RF | EX | DF | DS | TC | WB | |
| 延迟时隙 | | IF | IS | RF | EX | DF | DS | TC | WB |
| 阻塞 | | | 阻塞 | 阻塞 | 阻塞 | 阻塞 | 档位 | 失速 | 摊位 |
| 畜栏 | | | | 拖延 | 抛锚 | 失速 | 货摊 | 拖延 | 畜栏 |
| 分支目标 | | | | IF | IS | RF | EX | DF | |

| 指令 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|--------|----|----|----|----|----|----|----|----|----|
| 分支指令 | IF | IS | RF | EX | DF | DS | TC | WB | |
| 延迟槽 | | IF | IS | RF | EX | DF | DS | TC | WB |
| 分支指令+2 | | | IF | IS | RF | EX | DF | DS | TC |
| 分支指令+3 | | | | IF | IS | RF | EX | DF | DS |

预测不采纳+延迟分支

延迟周期与启动间隔

| 浮点指令 | 延迟周期 | 启动间隔 | 流水线条数 |
|-------|------|------|--|
| 加法/减法 | 4 | 3 | U, S+A, A+R, R+S组合模式 |
| 乘法 | 8 | 4 | U,E+M,M,M,M,N,N+A,R |
| 除法 | 36 | 35 | U,A,R,D ²⁷ , D + A, D + R, D + A, D + R, A, R |
| 平方根 | 112 | 111 | U, E, (A+R) ¹⁰⁸ , A, R |
| 取反 | 2 | 1 | U, S |
| 绝对值 | 2 | 1 | U, S |
| 浮点数比较 | 3 | 2 | U, A, R |

结构冲突-1

| 操作 | 发射/暂停 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|------|-------|---|-----|-----|-----|-----|-----|-----|-----|-----|---|
| 乘法运算 | 发射 | U | M | M | M | M | N | N+A | R | | |
| 加法运算 | 问题 | U | S+A | A+R | R+S | | | | | | |
| | 问题 | | U | S+A | A+R | R+S | | | | | |
| | 问题 | | | U | S+A | A+R | R+S | | | | |
| | 停滞 | | | | U | S+A | A+R | R+S | | | |
| | 停滞 | | | | | U | S+A | A+R | R+S | | |
| | 问题 | | | | | | U | S+A | A+R | R+S | |
| | 问题 | | | | | | | U | S+A | A+R | |

结构冲突-2

| 操作 | 发射/阻塞 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|----|-------|---|-----|-----|-----|---|---|---|-----|-----|---|
| 加法 | 发射 | U | S+A | A+R | R+S | | | | | | |
| 乘法 | 问题 | | U | M | M | M | M | N | N+A | R | |
| | 问题 | | | U | M | M | M | M | N | N+A | R |

结构冲突-3

| 操作 | 问题/停滞 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 |
|----|--------|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|
| 除法 | 第0周期问题 | D | D | D | D | D | D+A | D+R | D+A | D+R | A | R |
| 添加 | 问题 | U | S+A | A+R | R+S | | | | | | | |
| | 问题 | | U | S+A | A+R | R+S | | | | | | |
| | 停滞 | | | U | S+A | A+R | R+S | | | | | |
| | 停滞 | | | | U | S+A | A+R | R+S | | | | |
| | 停滞 | | | | | U | S+A | A+R | R+S | | | |
| | 失速 | | | | | | U | S+A | A+R | R+S | | |
| | 失速 | | | | | | | U | S+A | A+R | R+S | |
| | 失速 | | | | | | | | U | S+A | A+R | |
| | 问题 | | | | | | | | | U | S+A | |
| | 问题 | | | | | | | | | | U | |
| | | | | | | | | | | | | |
| | | | | | | | | | | | | |

结构冲突-4

| 操作 | 问题 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|----|----|---|-----|-----|-----|---|---|---|---|---|---|
| 添加 | 停滞 | U | S+A | A+R | R+S | | | | | | |
| 划分 | 问题 | | U | A | R | D | D | D | D | D | D |
| | 问题 | | | U | A | R | D | D | D | D | D |

性能损失测量

