编号

**南京航空航天大学**

**机组实验**

|  |  |
| --- | --- |
| 题 目 | 11条指令流水线的实现 |

|  |  |
| --- | --- |
| 学生姓名 | 沈王雄 |
| 学 号 | 161510326 |
| 学 院 | 计算机科学与技术学院 |
| 专 业 | 计算计科学与技术专业 |
| 班 级 | 1615104 |
| 指导教师 |  |

二〇一七年六月

**目录**

1. **实验背景介绍：............................................**
2. **实验要求(功能设计要求)....................................**
3. **模块化和层次化设计说明....................................**
4. **流水线的冒险及处理：......................................**
5. **代码测试与结果分析.........................................**
6. **心得体会 ...................................................**
7. **参考文献.....................................................**
8. **实验背景：**
   1. MIPS指令的优异性：MIPS是世界上很流行的一种RISC处理器。MIPS的意思是“无内部互锁流水级的微处理器”(Microprocessor without interlocked piped stages)，其机制是尽量利用软件办法避免流水线中的数据相关问题。它最早是在80年代初期由斯坦福(Stanford)大学Hennessy教授领导的研究小组研制出来的。MIPS公司的R系列就是在此基础上开发的RISC工业产品的微处理器。这些系列产品为很多计算机公司采用构成各种工作站和计算机系统。精简指令集这种设计思路对指令数目和寻址方式都做了精简，使其实现更容易，指令并行执行程度更好，编译器的效率更高。它能够以更快的速度执行操作。1.统一指令编码（例如，所有指令中的op-code永远位于同样的位元位置、等长指令），可快速解译；2.泛用的暂存器，所有暂存器可用于所有内容，以及编译器设计的单纯化（不过暂存器中区分了整数和浮点数）；3.单纯的寻址模式（复杂寻址模式以简单计算指令序列取代）4.硬件中支援少数资料型别（例如，一些CISC电脑中存有处理字节字串的指令。这在RISC电脑中不太可能出现）。而CISC(复杂指令集计算机)和RISC(精简指令集计算机)是当前CPU的两种架构。它们的区别在于不同的CPU设计理念和方法。早期的CPU全部是CISC架构，它的设计目的是要用最少的机器语言指令来完成所需的计算任务。RISC和CISC是设计制造微处理器的两种典型技术，虽然它们都是试图在体系结构、操作运行、软件硬件、编译时间和运行时间等诸多因素中做出某种平衡，以求达到高效的目的，但采用的方法不同，因此，在很多方面差异很大，它们主要有：

（1）指令系统：RISC设计者把主要精力放在那些经常使用的指令上，尽量使它们具有简单高效的特色。对不常用的功能，常通过组合指令来完成。因此，在RISC机器上实现特殊功能时，效率可能较低。但可以利用流水技术和超标量技术加以改进和弥补。而CISC计算机的指令系统比较丰富，有专用指令来完成特定的功能。因此，处理特殊任务效率较高。

（2）存储器操作：RISC对存储器操作有限制，使控制简单化；而CISC机器的存储器操作指令多，操作直接。

（3）程序：RISC汇编语言程序一般需要较大的内存空间，实现特殊功能时程序复杂，不易设计；而CISC汇编语言程序编程相对简单，科学计算及复杂操作的程序社设计相对容易，效率较高。

（4）中断：RISC机器在一条指令执行的适当地方可以响应中断；而CISC机器是在一条指令执行结束后响应中断。

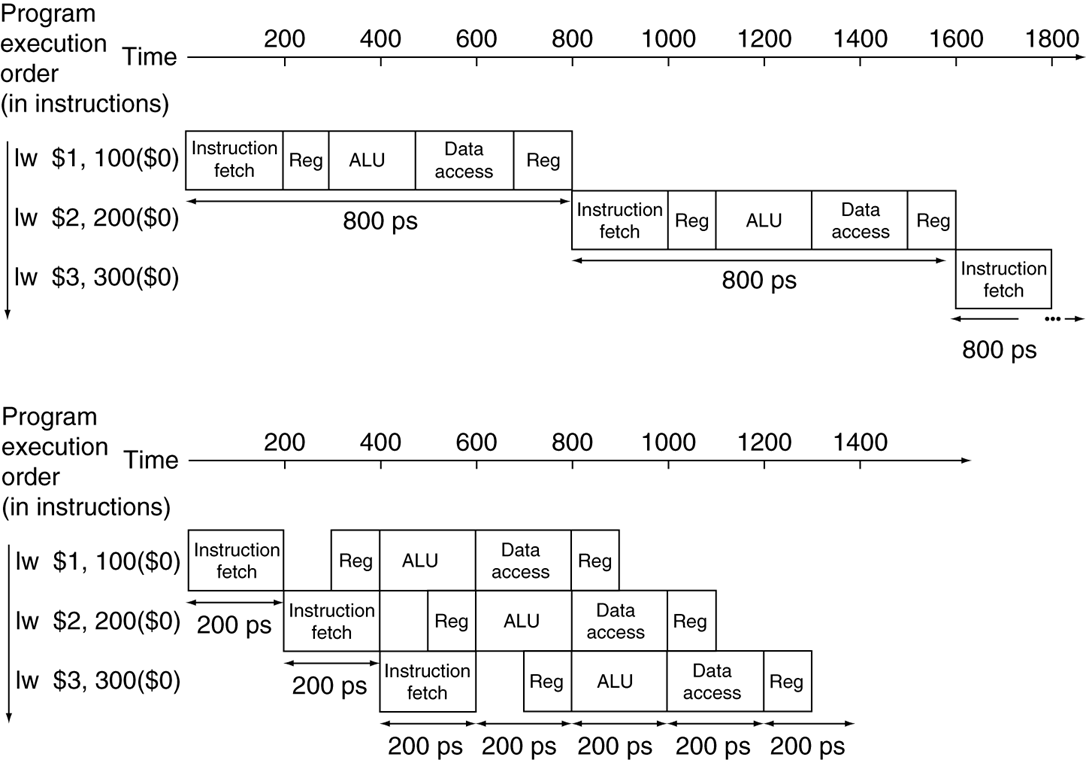
（5）CPU：RISCCPU包含有较少的单元电路，因而面积小、功耗低；而CISCCPU包含有丰富的电路单元，因而功能强、面积大、功耗大。

（6）设计周期：RISC微处理器结构简单，布局紧凑，设计周期短，且易于采用最新技术；CISC微处理器结构复杂，设计周期长。

（7）用户使用：RISC微处理器结构简单，指令规整，性能容易把握，易学易用；CISC微处理器结构复杂，功能强大，实现特殊功能容易。

（8）应用范围：由于RISC指令系统的确定与特定的应用领域有关，故RISC机器更适合于专用机；而CISC机器则更适合于通用机。

2、单周期实验的重要性：基于MIPS的单周期CPU的设计与实现，可以让我们真正的走进计算机CPU的内部世界观察，每一条指令的执行过程，不仅仅是对课堂知识的一个简单的回顾理解，也是从最为根本的门电路去理解计算机。也为之后的多周期，及流水线的实现就是对单周期实现的扩展使得效率的成倍提升。



3、国家安全战略和民生工程等方面，面临挑战：

电站可被远程控制；公务车有处理器后门，可被监听或控制；…….

举例：2012年伊朗核离心装置遭到“震网”病毒攻击，20%的离心机报废，监控录像被篡改，而遭攻击的电脑是物理隔离的

基本现状：国家急需计算机系统方面的人才！！！

4、神威·太湖之光超级计算机

（a）、安装了40960个中国自主研发的“申威26010”众核处理器，采用64位自主申威指令系统，峰值性能为12.5亿亿次/秒，持续性能为9.3亿亿次/秒。2016年6月20日在法兰克福世界超算大会上，国际TOP500组织发布的榜单显示，“神威·太湖之光”超级计算机系统速度比第二名“天河二号”快出近两倍，其效率也提高3倍. “天河二号”比第三名美国“泰坦”快近一倍

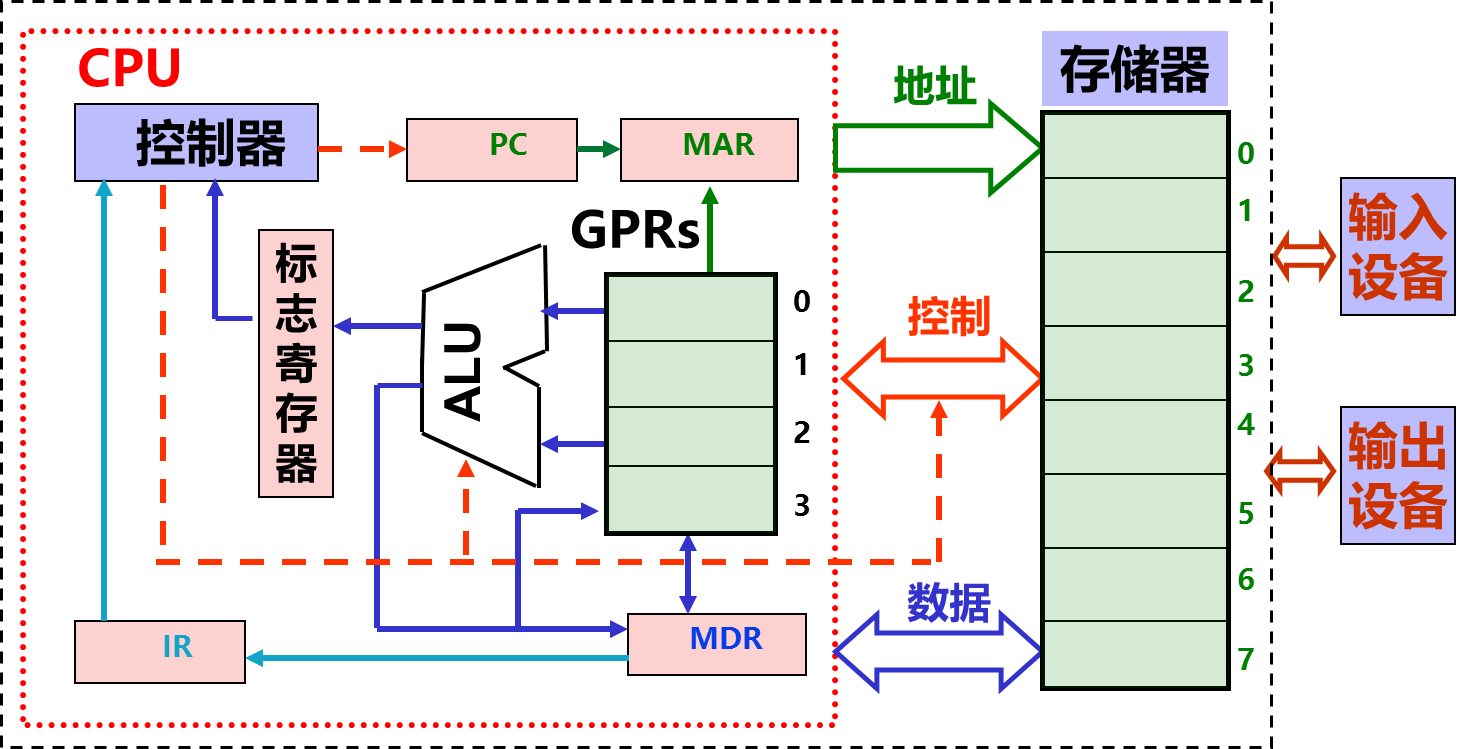
(b)、问题：实用性和效能难令人满意，美国限售Intel高性能处理器芯片

5、龙芯：

(a)、从“北斗”卫星、防火墙等国家安全战略层面，到电子水表、机顶盒等民生项目都有应用，取得长足进步

(b)、软硬件开发人员上万，2013年1.8万片，2014年达35万片

(c)、“长板”与国际先进水平相差2-3年，“短板”相差10年以上。若有10万人基于龙芯开发基础软件，就能很好建立生态系统

6、冯·诺依曼体系结构下图所示：

其存储器结构是将程序指令存储器与数据存储器合并在一起，也就是将指令和数据共同放在一个存储器中，指令存储器和数据存储器共用一条数据总线，这样的结构在早期的微处理器以及单片机中是非常常见。其主要特点有以下几点

（1） 指令与数据共用一个存储器；

（2） 指令与数据的存储地址指向同一存储器的不同物理位置；

（3） 取指令与取数据共用一条总线；

（4） 程序指令和数据的位宽相等。

根据以上总结的特点可以看出冯·诺依曼体系结构其优点在于由于共享数据总线，所以减小了总线的开销，并且能够把 RAM 映射到程序空间。但是其缺点也是显而易见的，处理器在执行命令式，首先必须从存储器中取出指令进行解码，然后再从存储器中取出操作数执行运算，如此操作即使是单指令也要耗费几个甚至几十个时钟周期来执行，所以在高性能和高实时性场合，会在数据传输通道上会出现时序瓶颈，制约微处理器性能。与冯·诺依曼体系结构不同，哈佛结构是一种将程序指令存储和数据存储分开的存储器结构 、

（1） 指令和数据分别存储在指令存储器和数据存储器中；

（2） 指令与数据的地址都是独立编码的；

（3） 指令与数据都是独立访问的；

（4） 微处理器中包含四条独立的总线；

（5） 数据和指令宽度可以不等。

从上面的几点总结可以看出，由于采用地址总线与数据总线相分离的体系结构，所以可以使微处理器能够在一个时钟内同时读取指令和数据，从而缓解了冯·诺依曼体系结构中数据传输通道上会出现的瓶颈效应。从以上对两种体系结构的特点介绍可以发现哈佛结构相对于冯·诺依曼体系结构来说，它具有以下两个特点：首先是它使用两个独立的存储器分别存储指令与数据，其次是它具有两套独立的总线系统，一共四条总线，它们分别担当指令存储器以及数据存储器与微处理之间进行通信的作用，这两套总线系统互相独立。也正是有了这些特点，使得哈佛结构的微处理器拥有较大的存储器带宽，这样可以使得执行指令读取以及数据存储操作时更加方便也更有效率。

因此，为了实现数据的并行处理，采用流水线的方式，提高程序，数据的突出率，使得硬件资源得到充分的利用，进而提高处理器的效率，由于流水线是基于单周期的cpu的设计，因此对于现代流水线的设计依然沿袭单周期的模块设计，但值得注意的是对于流水线的设计，由于要实现对硬件资源的充分，不可避免的就是在数据流动中，出现了数据的逆向流动进而破坏了流水线的数据通路，分别可以分为结构冒险，数据冒险，控制冒险三大类，再设计过程中应该注意避免，以防出现无法预知的错误。

1. **实验要求（功能设计要求）：**

（一）、功能设计说明：

1. 完成以下指令集。

a) 书上的11条指令：add，sub，and，or，slt等多条指令以及lw，sw，beq和J指令。

b) 不支持溢出。

2. 处理器为流水线设计：

a) 数据通路由如下module组成：PC(程序计数器)、NPC(NextPC计算单元)（5.19）、Register File (寄存器文件、寄存器堆)、ALU(算术逻辑单元)（5.13）、EXT(扩展单元)（5.18）、IM(指令存储器)、DM(数据存储器)。

b) IM：容量为4KB(32bit×1024字)。

c) DM：容量为4KB(32bit×1024字)。

d) forwarding单元（处理数据冒险）

e） 数据冒险检测单元

f) IF\_ID\_reg寄存器文件

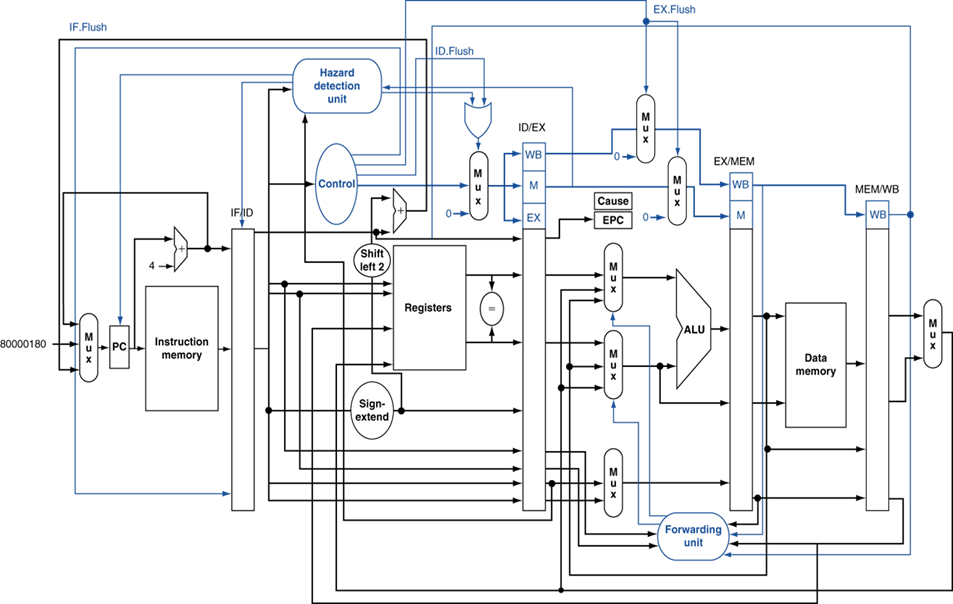
g) ID\_Exe\_reg 寄存器文件

h) Exe\_Mem\_reg 寄存器文件

i) Mem\_wr\_reg 寄存器文件

3、流水线的数据通路架构图（来源于计算机组成与设计:硬件/软件接口）。

* 1. 基于以下的数据通路实现上述11条指令集。数据通路



* 1. 模块分解：

PC

**IF阶段**

Im\_4k

Iunit

IF\_ID\_reg

Im\_test\_code

decode

**ID阶段**

Ctrl

regfile

Hazrd\_det

datapath

Ext

testbench

ID\_EX\_reg

Mux3

ExetUnit

Alu

Exe阶段

Mux1

Frowding

Mux2

EX\_Mem\_reg

Memj阶段

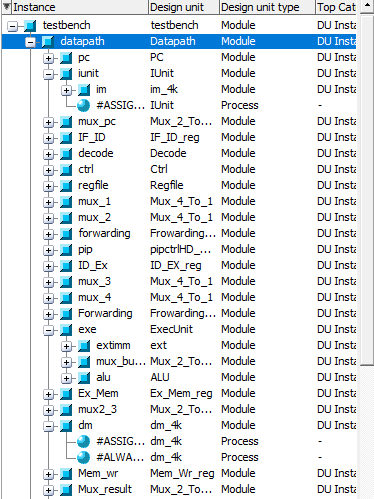
Dm\_4k

RegWr阶段

Mux2

Mem\_wr\_reg

* 1. 实现结果：



4、code.txt中存储的是指令码

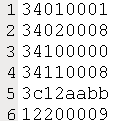
1. 用VerilogHDL建模IM时，必须以读取文件的方式将code.txt中指令加载至IM中。
2. code.txt的格式如 Figure1所示。每条指令占用1行，指令二进制码以文本方式存储。

Figure1code.txt文件格式

1. 为使得代码更加清晰可读，建议多使用宏定义，并将宏定义组织在合理的头文件中。
2. PC复位后初值为0x0000\_3000，目的是与MARS的Memory Configuration相配合。
   1. 也可以用测试程序将通过MARS产生，其配置模式如
   2. Figure2所示。
3. **模块化和层次化设计说明:**

基本模块与单周期的设计基本一致 ,因此在此处只对,流水线中的特有的模块进行分析.因为单周期正在实现的数据通路中的控制信号是不变的,因此,在流水线的设计中可将每一条指令在各个阶段中的控制信号,保存下来以便在之后的数据通路中使用.所以对于五级流水线可设计四个寄存器对中间数据的存储.

(1)、IF\_ID\_reg:

Clk

o\_PC+4

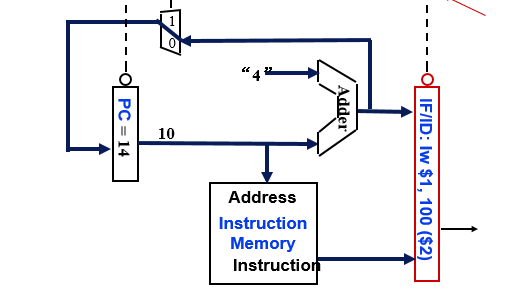
i\_PC+4

o\_instruction

i\_instruction

因此,可设计一个If\_Id\_reg [63:0].的寄存器用于存储取指令阶段产生的pc+4,以及读出的指令。值得注意的是每次都是在时钟下降沿，将数据读出寄存器。

（2）、Iunit模块：取指令并送到IF\_ID\_reg中



取指令部件

(3)、ID\_EX\_reg

W’B

**Ctrl**

M

Ex

o\_imm16

i\_imm16

o\_busA

i\_busA

o\_Rt

o\_busB

i\_busB

i\_Rt

i\_Rs

o\_Rs

Clk

|  |
| --- |
| Memtoreg |
| RegWr |

|  |
| --- |
| Extop |
| ALUsrc |
| ALUCtr |
| RegDst |

其中：

|  |
| --- |
| Memwr |

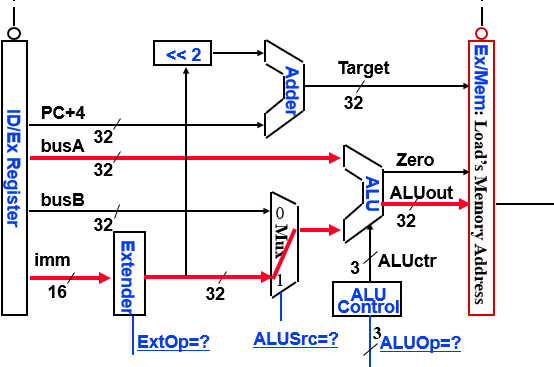
其中，第一个为Mem,

第二个为 Wr

的第三个为Exe

的控制信号。

(4)、ExeUnit模块：



对于基本流水线，Ex阶段分别执行各自的操作：

（1）计算内存地址（lw,sw）

（2）计算转移目标地址(beq,jump)

（3）一般ALU运算(add,sub,or……)

5)、EX\_Mem\_reg

W’B

**Ctrl**

M

O\_overflow

i\_overflow

i\_aluresult

o\_result

o\_RW

i\_Rw

(6)、Mem\_Wr\_reg

W’B

**Ctrl**

O\_overflow

i\_overflow

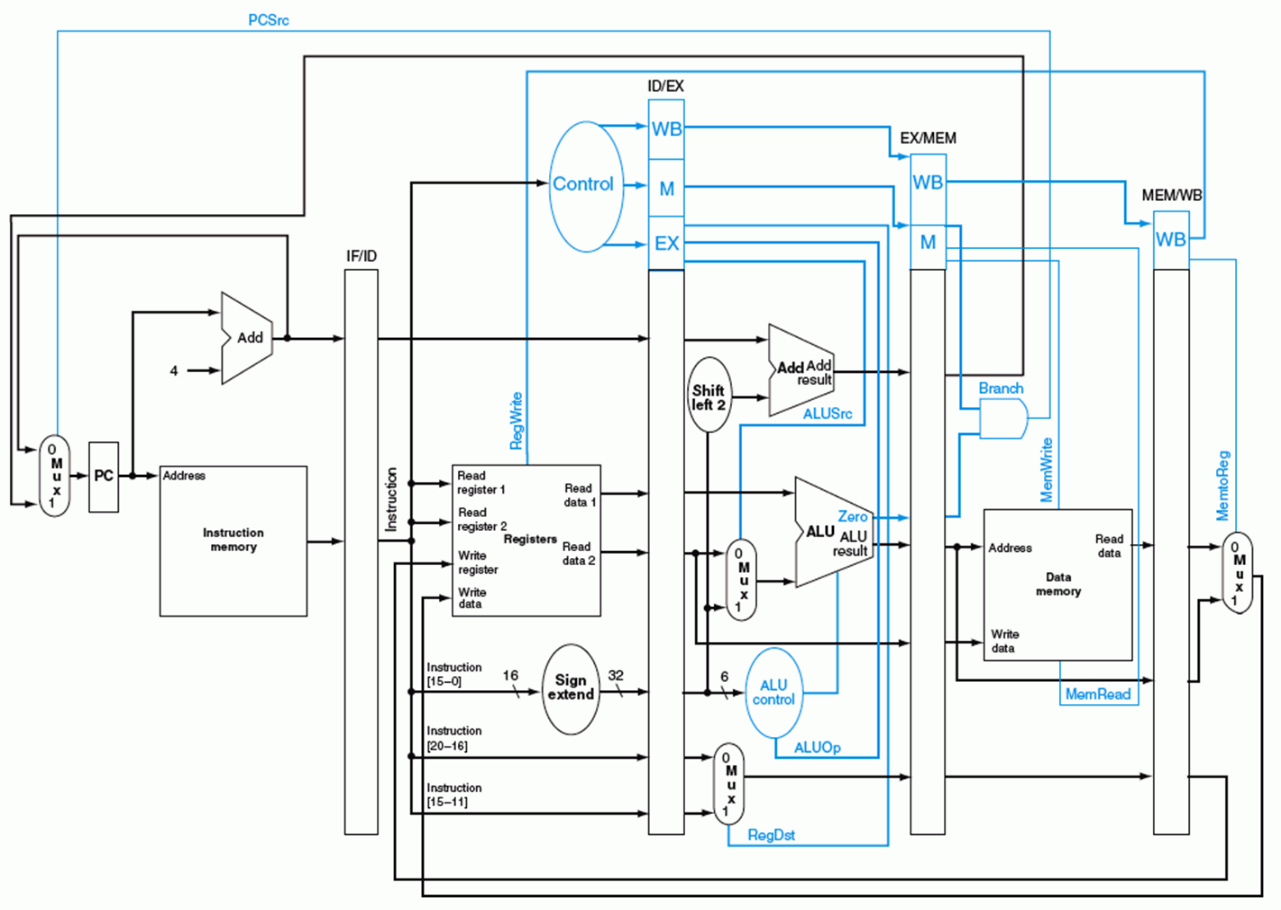
i\_data

o\_data

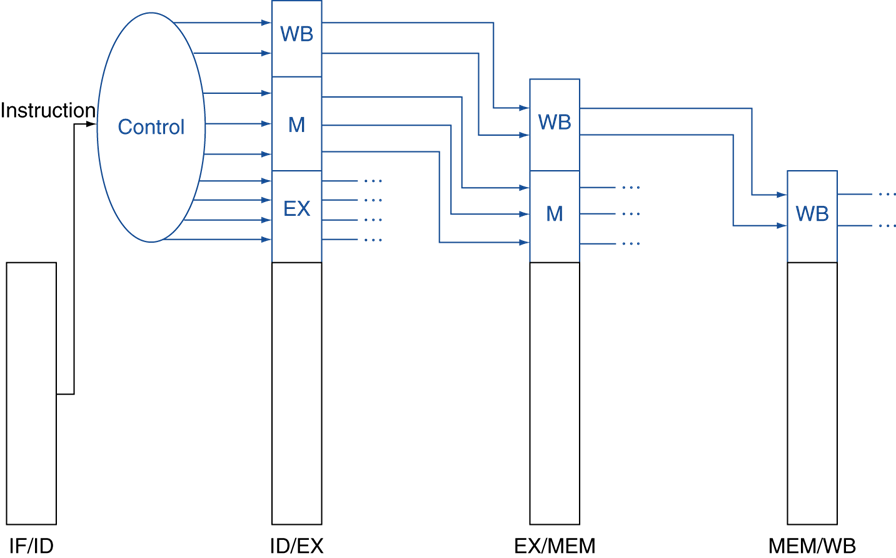
o\_RW

i\_Rw

按照课本上的基本数据通路即可组装为五级基本流水线。



其中对于各个段寄存器的控制指令的变化过程如下：

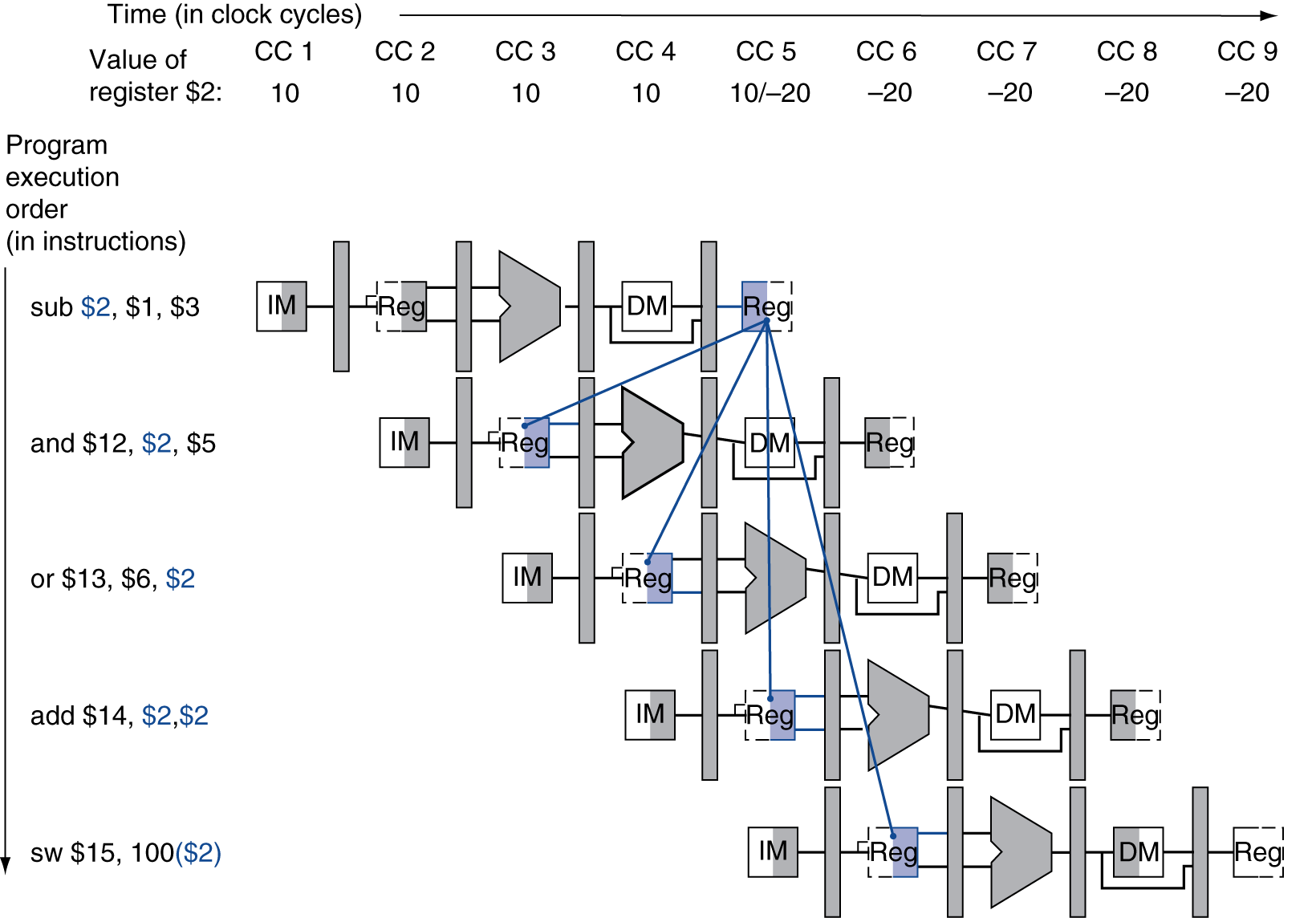


1. **流水线的冒险及处理：**

（1）、结构冒险：同一部件同时被不同指令所使用，也就是由硬件资源竞争造成的。

解决方法：一个部件每条指令只能使用1次，且只能在特定周期使用

设置多个部件，以避免冲突。如指令存储器IM 和数据存储器DM分开。

（2）、数据冒险：后面指令用到前面指令结果数据时，前面指令的结果还没产生

解决方法： 方法1：硬件阻塞（stall）

方法2：软件插入“NOP”指令

方法3：合理实现寄存器堆的读/写操作（不能解决所有数据冒险）:前半时钟周期写，后半时钟周期读，若同一个时钟内前面指令写入的数据正好是后面指令所读数据，则不会发生数据冒险

方法4：转发（Forwarding或Bypassing 旁路）技术

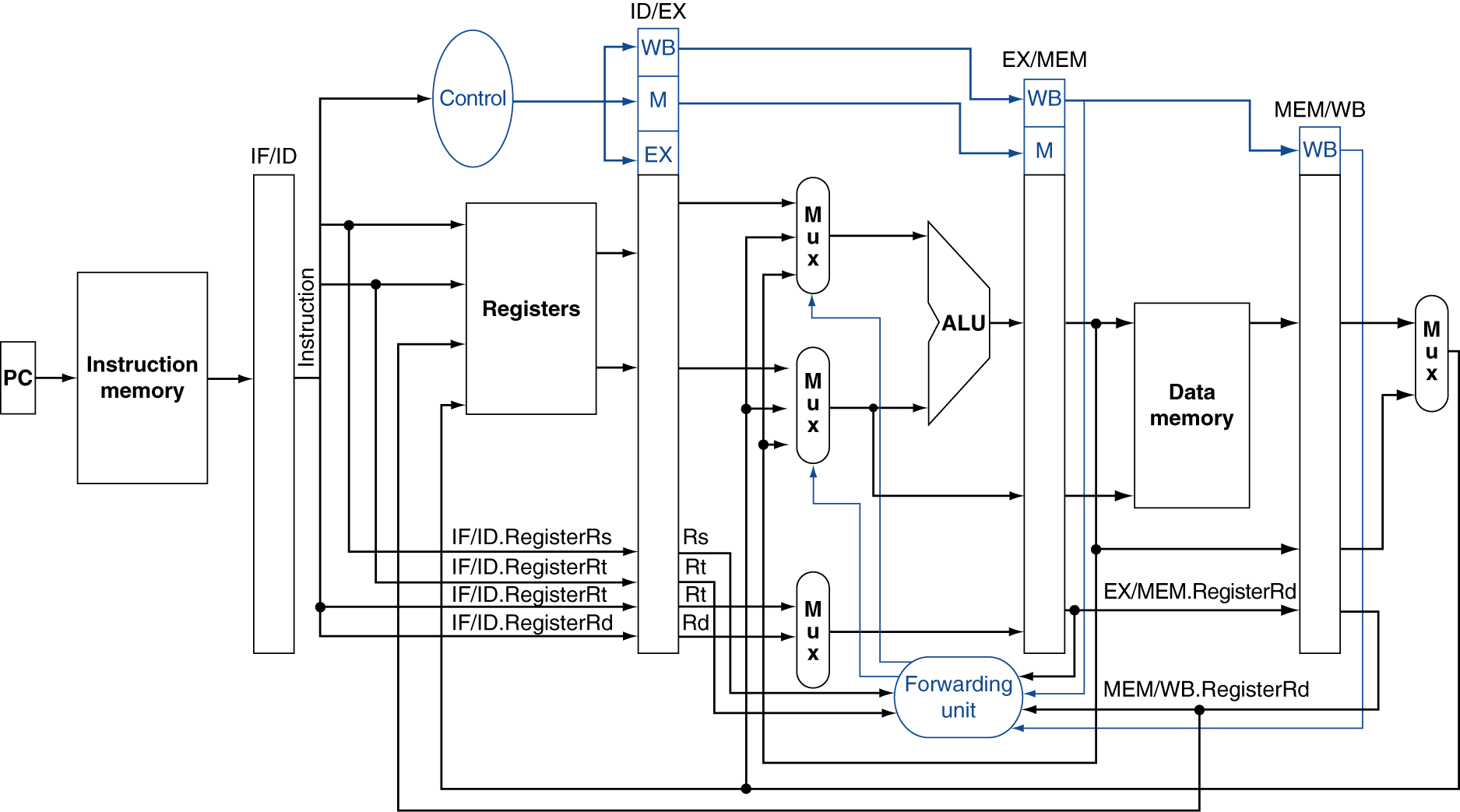
若相关数据是ALU结果， 可通过转发解决

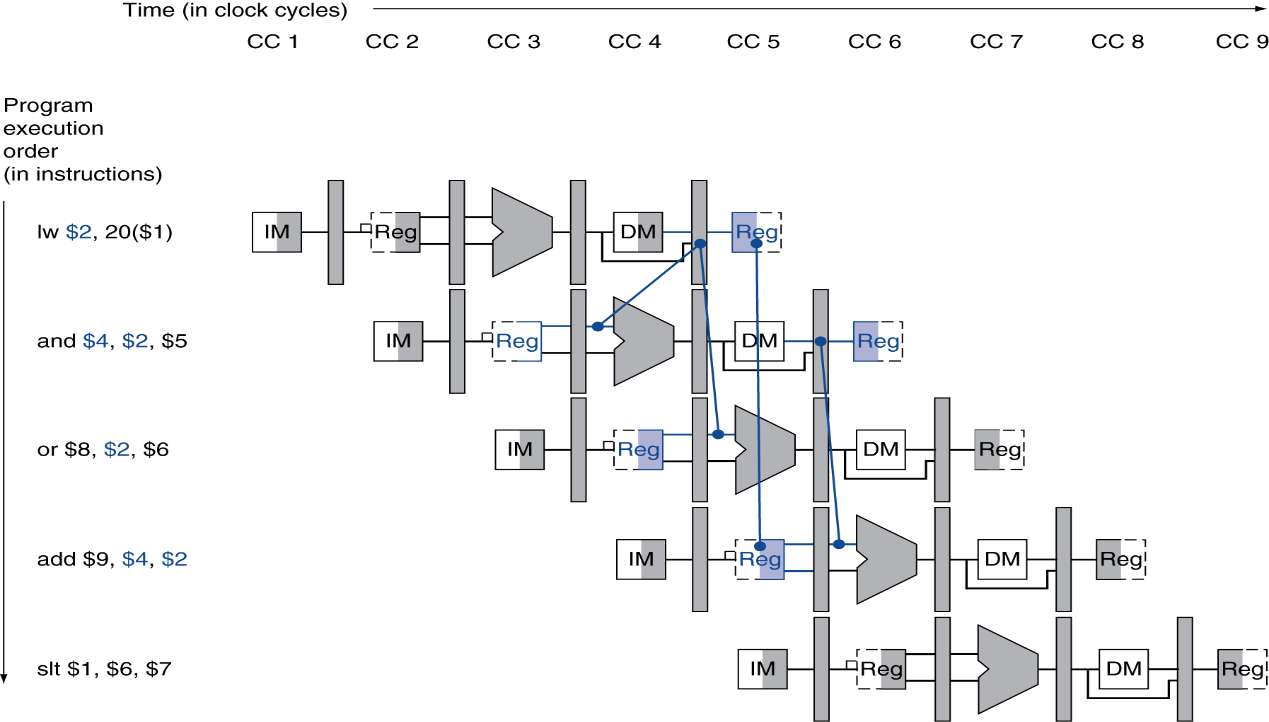
若相关数据是上条指令DM读出内容,不能通过转发解决，随后指令需被阻塞一个时钟 或 加NOP指令

方法5：编译优化：调整指令顺序（不能解决所有数据冒险）

在本次实验中，我们所采用的解决办法为转发加阻塞：

* MEM hazard
* if (MEM/WB.RegWrite and (MEM/WB.RegisterRd ≠ 0)  
   and not (EX/MEM.RegWrite and (EX/MEM.RegisterRd ≠ 0)  
   and (EX/MEM.RegisterRd = ID/EX.RegisterRs))  
   and (MEM/WB.RegisterRd = ID/EX.RegisterRs))  
   ForwardA = 01
* if (MEM/WB.RegWrite and (MEM/WB.RegisterRd ≠ 0)  
   and not (EX/MEM.RegWrite and (EX/MEM.RegisterRd ≠ 0)  
   and (EX/MEM.RegisterRd = ID/EX.RegisterRt))  
   and (MEM/WB.RegisterRd = ID/EX.RegisterRt))  
   ForwardB = 01

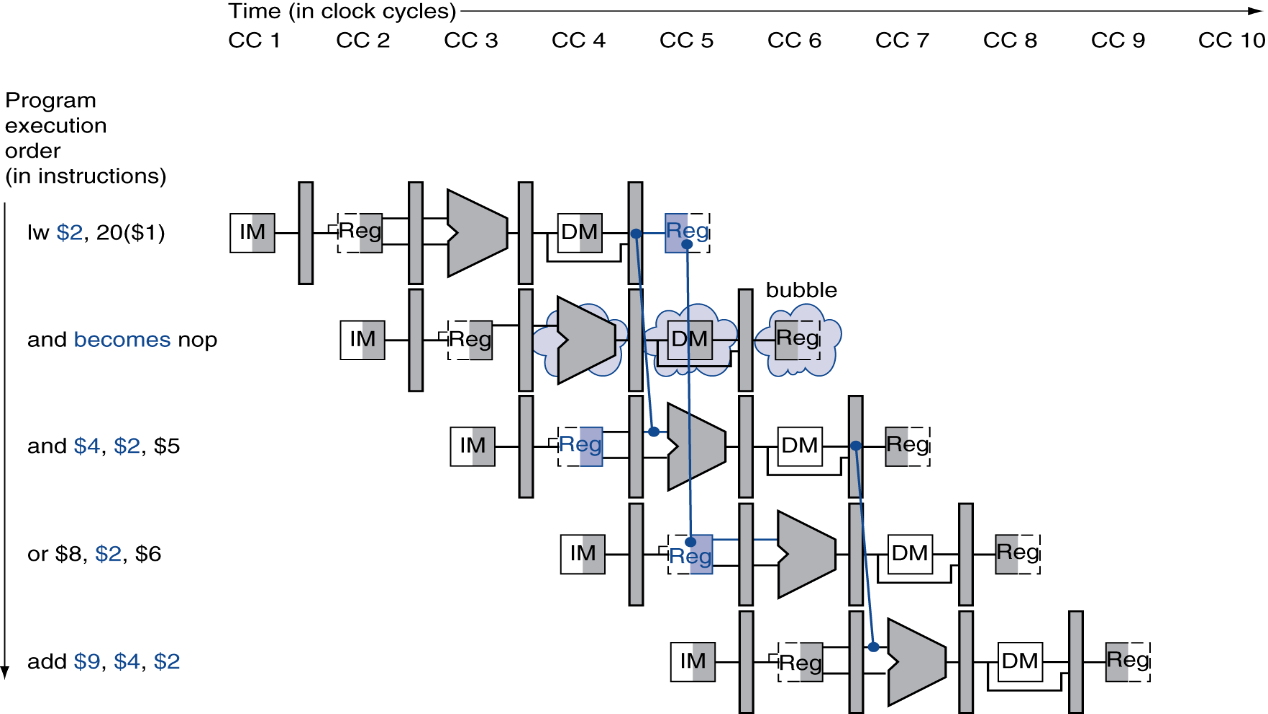


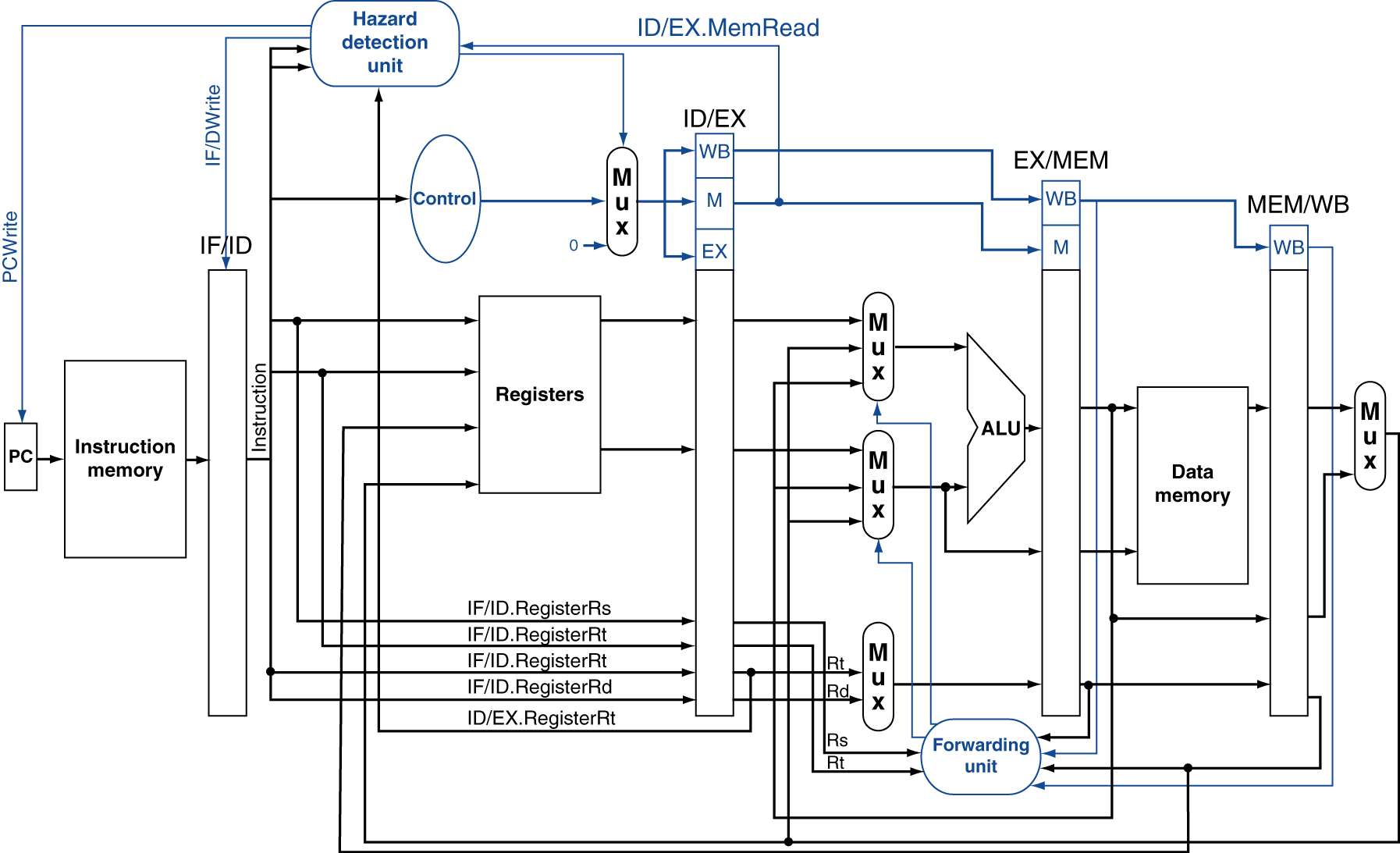


Need to stall for one cycle

但是对于load\_use类型的则需要对其阻塞一个周期：

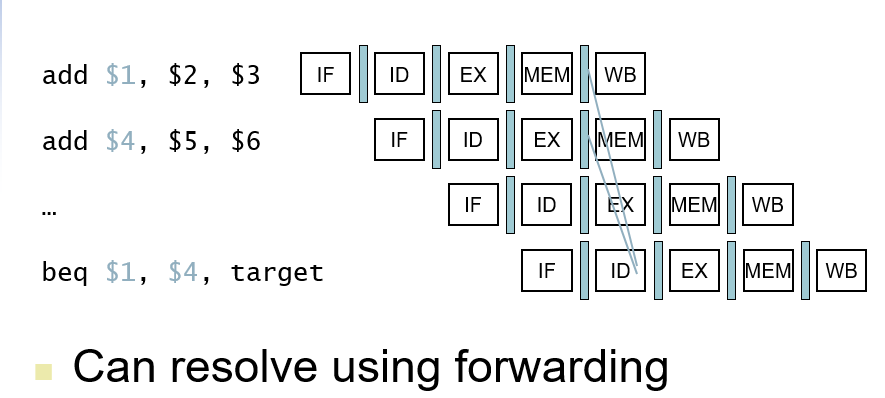
* Check when using instruction is decoded in ID stage
* ALU operand register numbers in ID stage are given by
  + IF/ID.RegisterRs, IF/ID.RegisterRt
* Load-use hazard when
  + ID/EX.MemRead and  
     ((ID/EX.RegisterRt = IF/ID.RegisterRs) or  
     (ID/EX.RegisterRt = IF/ID.RegisterRt))
* If detected, stall and insert bubble

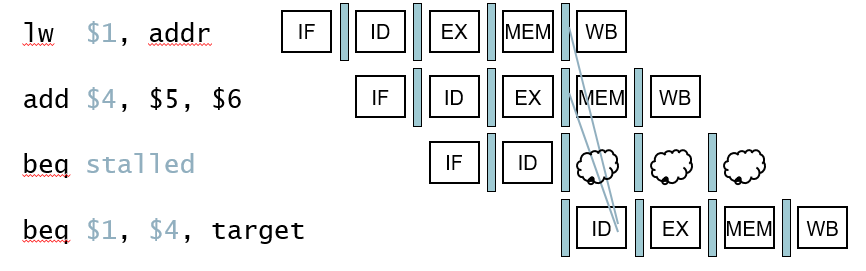




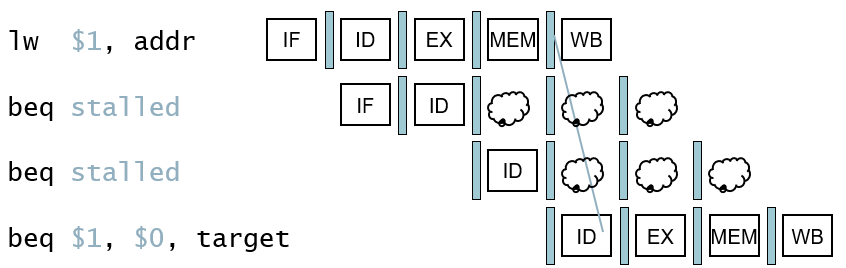
（3）、控制冒险：转移或异常改变执行流程，后继指令在目标地址产生前已被取出；解决方法：采用静态或动态分支预测：编译程序优化指令顺序(分支延迟)

对于branch指令，之前都是在Mem阶段返回值得，导致索取出的指令浪费的情况，因此可将beq指令提前至ID阶段进行执行以减少计算机资源的浪费。但是将beq指令前置，将会导致beq中再次发生数据冒险。

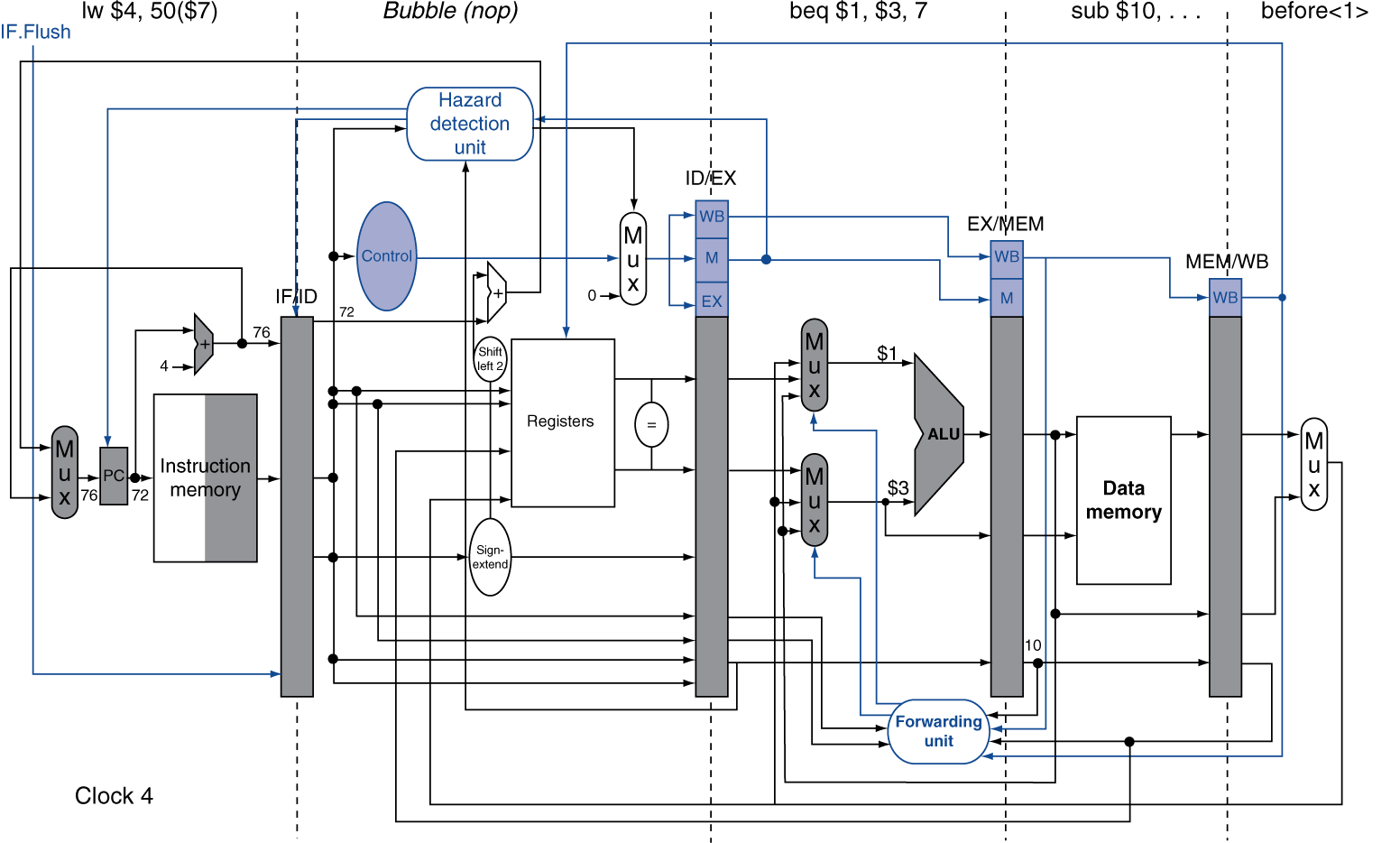




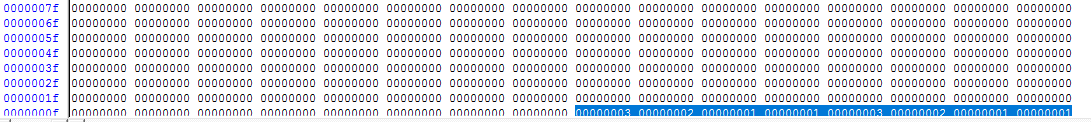
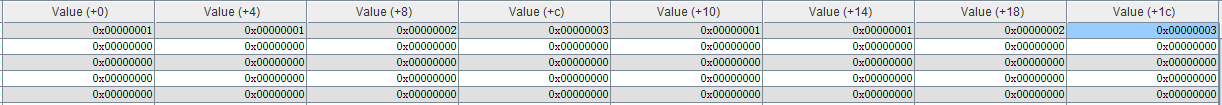
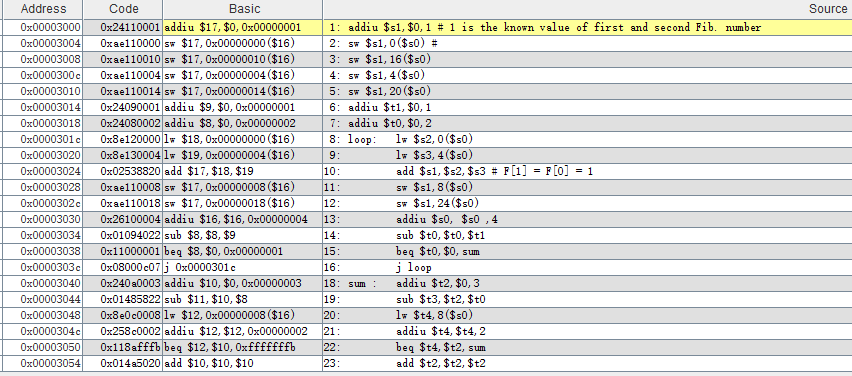
对于以上情况需要阻塞一个周期



需阻塞两个周期



1. **代码测试与结果分析：**



1. **心得体会:**
2. **参考文献：**

袁春风 《计算机系统组成与系统设计》（第二版）清华大学出版社

《深入理解计算机系统》（第三版） 机械工业出版社

《计算机组成——硬件接口与设计》 （第二版） 机械工业出版社