

Ch3.CPU设计-单周期、多周期 王超

中国科学技术大学计算机学院 高能效智能计算实验室 2024年春

主要内容



- 口处理器的基本结构
- 口单周期CPU设计
 - ✓功能部件
 - ✓数据通路
- 口多周期处理器设计
 - ✓功能部件
 - ✓数据通路
 - ✓控制信号

指令数据通路的Overview



- □CPU性能因素
 - ✓指令数目:取决于ISA和编译器
 - ✓ CPI 和时钟周期长度:取决于CPU硬件
- □我们将要考察三个RISC-V的实现
 - ✓ 一个是简化版本
 - ✓ 一个是多周期版本
 - ✓ 一个是更实际的流水线版本
- □用ISA的简单子集展现CPU设计的多数方面
 - ✓ 存储器访问: ld, sd
 - ✓ 算术/逻辑运算: add, sub, and, or
 - ✓控制转移: beq

指令数据通路Overview



- 口实现不同指令的多数工作都是相同的,与指令类型无关
 - ✓ 取指:将PC送往MEM
 - ✓ 取数:根据指令字中的地址域读寄存器
- ■执行操作各个指令不同,但同类指令非常类似
 - ✓ 算术逻辑指令
 - R类ALU指令 add x1, x2, x3
 - I类ALU指令 addi x3,x3,4
 - ✓ 不同类型指令也有相同之处, 如都要使用ALU
 - 访存指令使用ALU计算地址 lw x1,100(x2)
 - 算逻指令使用ALU完成计算 add x1, x2, x3
 - 分支指令使用ALU进行条件比较 beq x1, x2, name
 - ✓ 其后,各个指令的工作就不同了
 - 访存指令对存储器进行读写 lw x1,100(x2)
 - 算逻指令将ALU结果写回寄存器 add x1, x2, x3
 - 分支指令将基于比较结果修改下一条指令的地址 beq x1, x2, name

RISC-V指令编码格式总结



Name		eld				Comments	
(Field Size)	7 bits	5 bits	5 bits	3 bits	5 bits	7 bits	
R-type	funct7	rs2	rs1	funct3	rd	opcode	Arithmetic instruction format
I-type	immediate[11:0]		rs1	funct3	rd	opcode	Loads & immediate arithmetic
S-type	immed[11:5]	rs2	rs1	funct3	immed[4:0]	opcode	Stores
SB-type	immed[12,10:5]	rs2	rs1	funct3	immed[4:1,11]	opcode	Conditional branch format
UJ-type	immediate[20,10:1,11,19:12]				rd	opcode	Unconditional jump format
U-type	immediate[31:12]				rd	opcode	Upper immediate format

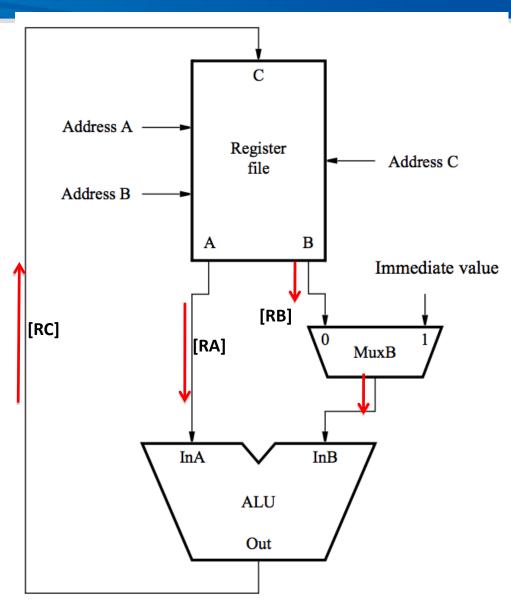
R-type Instructions	funct7	rs2	rs1	funct3	rd	opcode	Example	
add (add)	0000000	00011	00010	000	00001	0110011	add x1, x2, x3	
sub (sub)	0100000	00011	00010	000	00001	0110011	sub x1, x2, x3	
I-type Instructions immediate		diate	rs1	funct3	rd	opcode	Example	
addi (add immediate)	001111	101000	00010	000	00001	0010011	addi x1, x2, 1000	
1d (load doubleword)	001111101000		00010	011	00001	0000011	ld x1, 1000 (x2)	
S-type Instructions	immed -iate	rs2	rs1	funct3	immed -iate	opcode	Example	
sd (store doubleword)	0011111	00001	00010	011	01000	0100011	sd x1, 1000(x2)	

寄存器计算指令的数据流



□源操作数和目的 操作数都在寄存 器中

add x1, x2, x3; x2+x3->x1

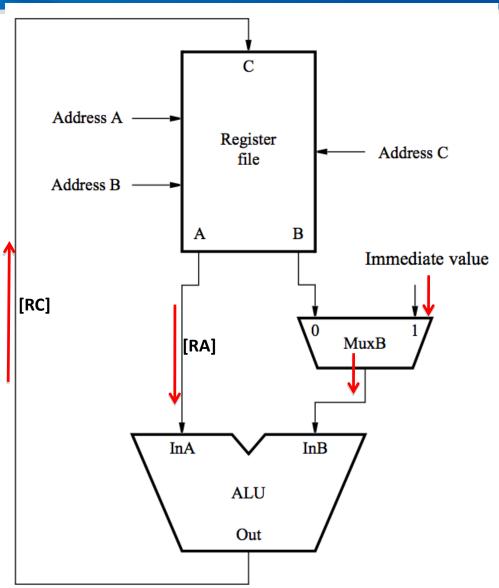


立即数计算指令的数据流



□有一个源寄存器在 立即数寄存器中

addi x3, x3, 4; x3 = x3 + 4



CPU指令数据通路总图

immed[11:5]

immed[12,10:5]

rs2

rs2

rs1

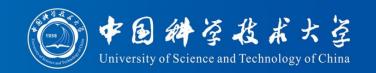
rs1

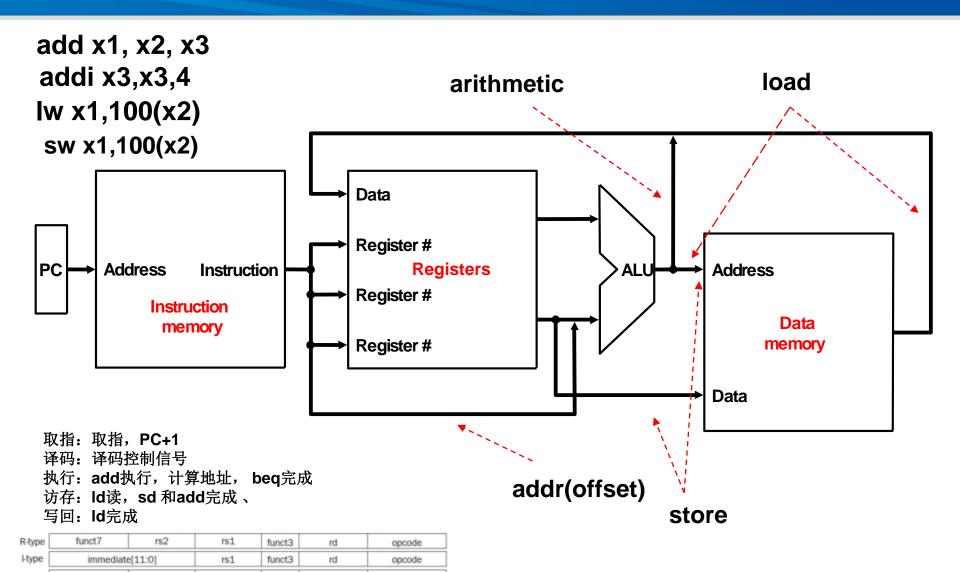
immed[4:0]

immed[4:1.11]

opcode

opcode





CPU指令数据通路总图

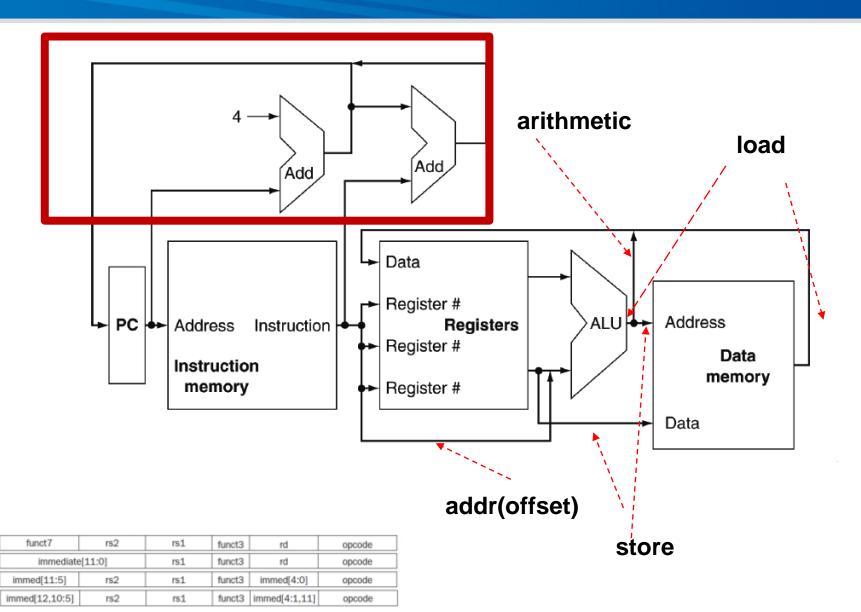
R-type

I-type

S-type

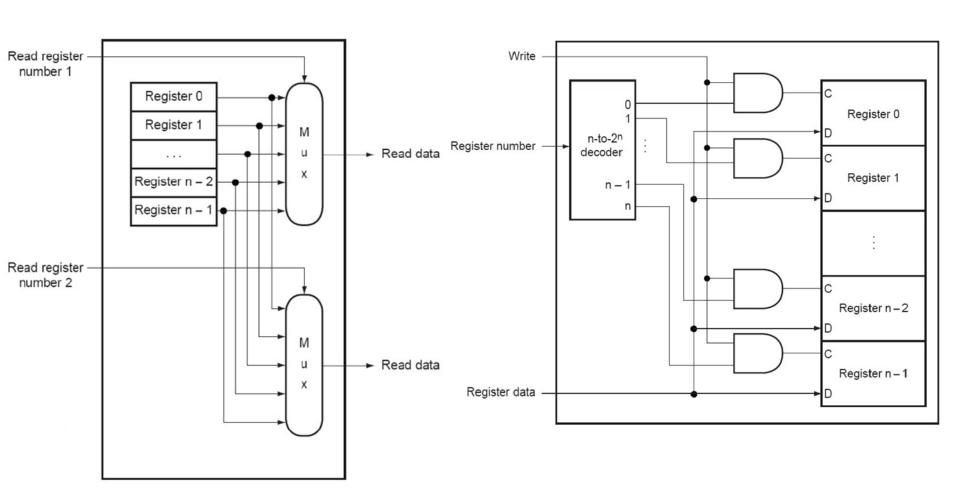
B-type





通路中的存储-RegisterFile

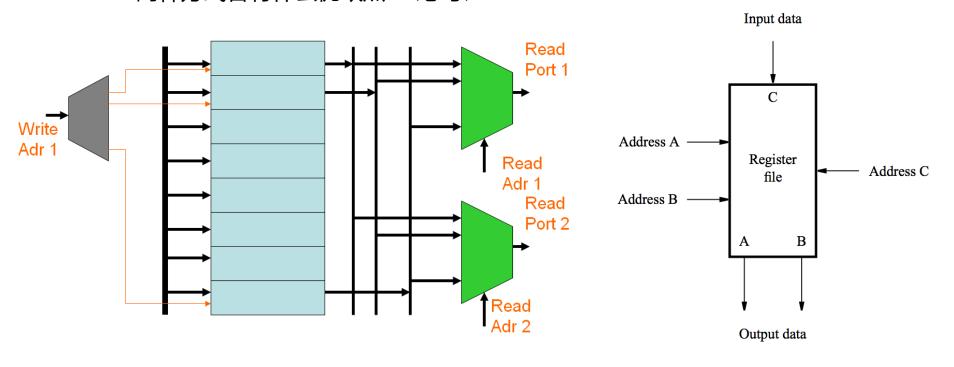




RegFile读写操作

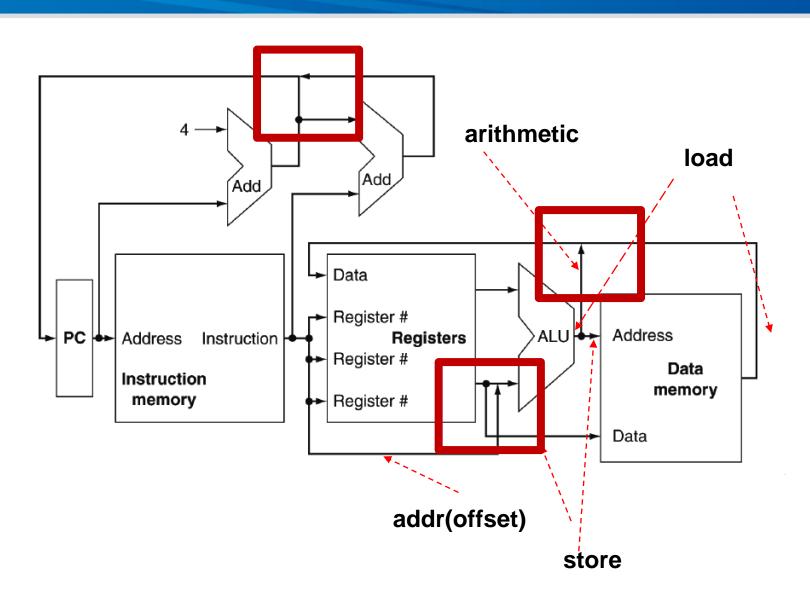


- 寄存器文件在一个周期中有两个读和一个写操作。
 - ✓ 在一个周期内,某个REG可以同时完成读写操作,但读出的是上一个周期写入的值
- □ 寄存器文件不能同时进行读和写操作。两种设计方法:
 - ✓ 后写 (late write): 在前半周期读数据,在后半周期写数据。
 - ✓ 先写 (early write) : 与后写相反。
 - ✓ 两种方式各有什么优缺点? (思考)



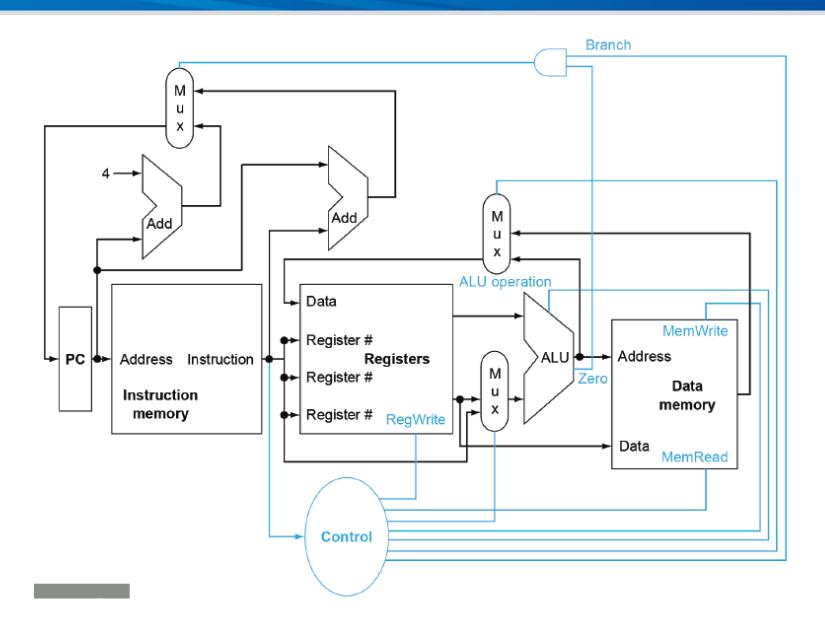
通道中的多路选择器





通道中的多路选择器





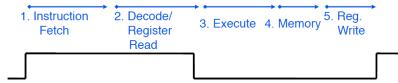


数据通路设计→控制部件设计 控制信号的作用? 控制信号的来源?

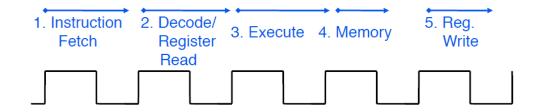
实现 单周期->多周期



- □ 指令周期 = m 机器周期。机器周期 = 时钟周期
- □ 单周期实现:指令周期 = 1 机器周期 = 1 时钟周期
 - ✓ All stages of an instruction are completed within one long clock cycle.
 - ✓ 所需控制信号同时生成



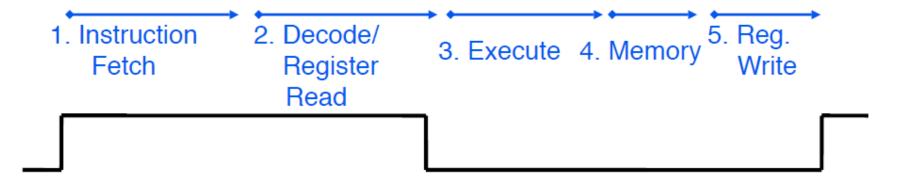
- □ 多周期实现: 指令周期 = 机器周期 = 时钟周期
 - ✓ Only one stage of instruction per clock cycle
 - ✓ 按时钟周期(= 机器周期)生成当前周期所需控制信号





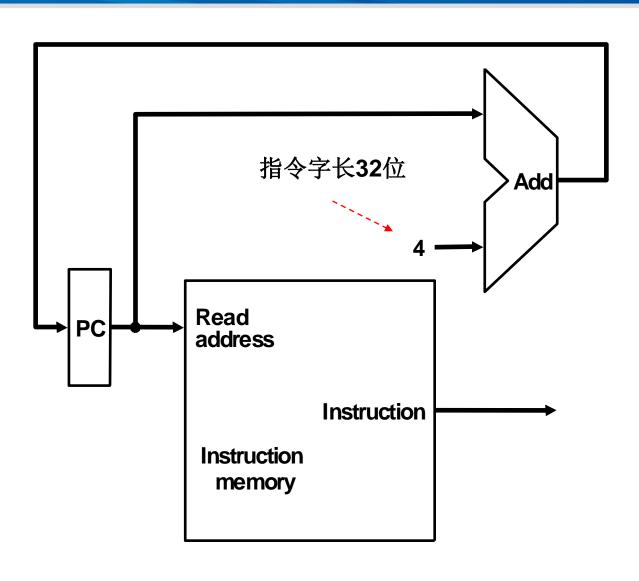
□定长指令周期

- ✓ All stages of ainstructioare completed within one *long* clock cycle.
- ✓采用时钟边沿触发方式
 - 所有指令在时钟的一个边开始执行, 在下一个边结束



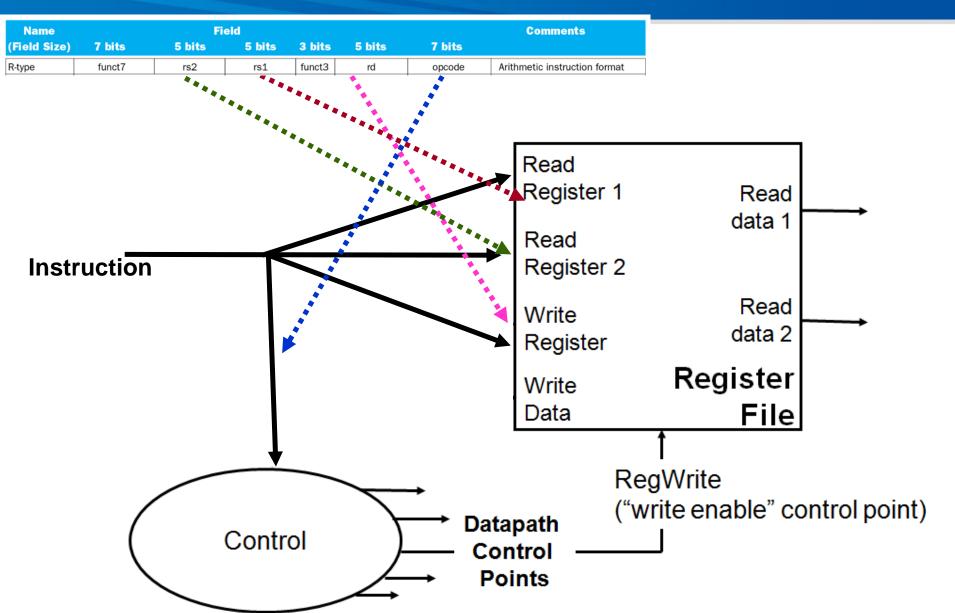
取指 Instruction Fetch





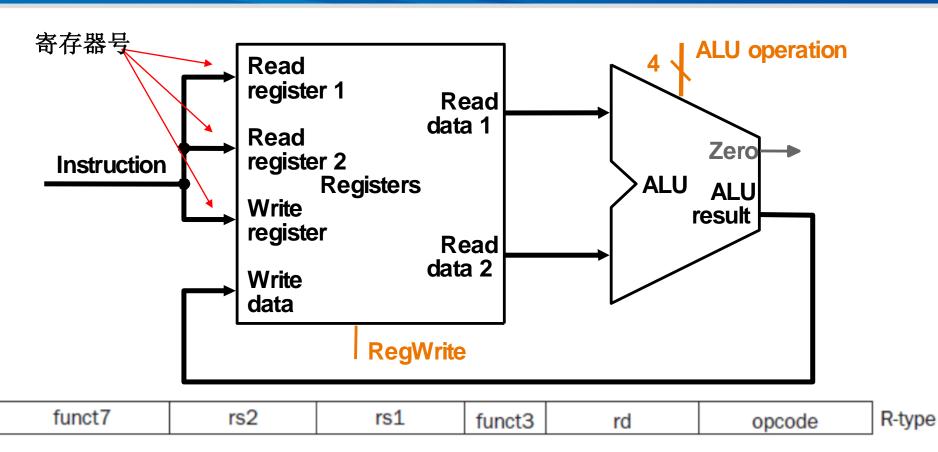
译码 Decode-RISC-V





R-type指令的执行





• 寄存器堆操作

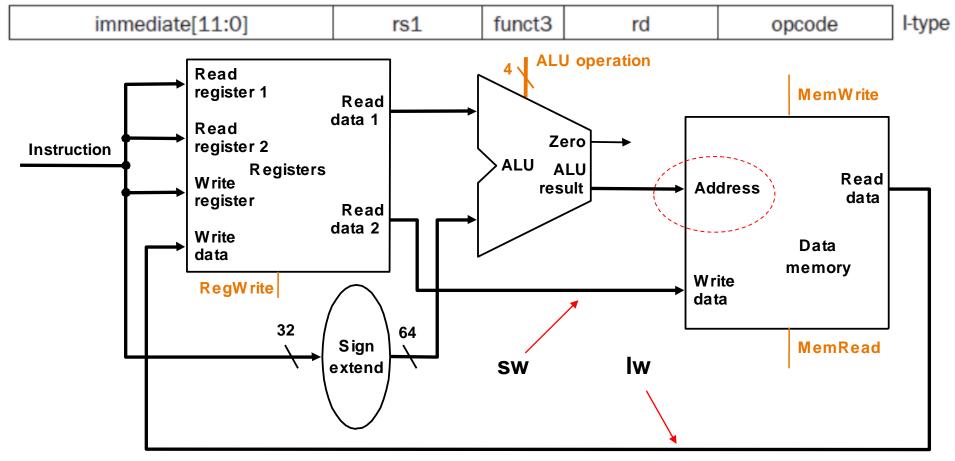
- 读:给出寄存器编号,则寄存器的值自动送到输出端口
- 写:需要寄存器编号和控制信号RegWrite,时钟边沿触发
- 在一个周期内, REG可以同时完成读写操作, 但读出的是上一个周期写入的值(读后写)

访存指令的执行-ld



rs: 基址

rt: Id目的寄存器



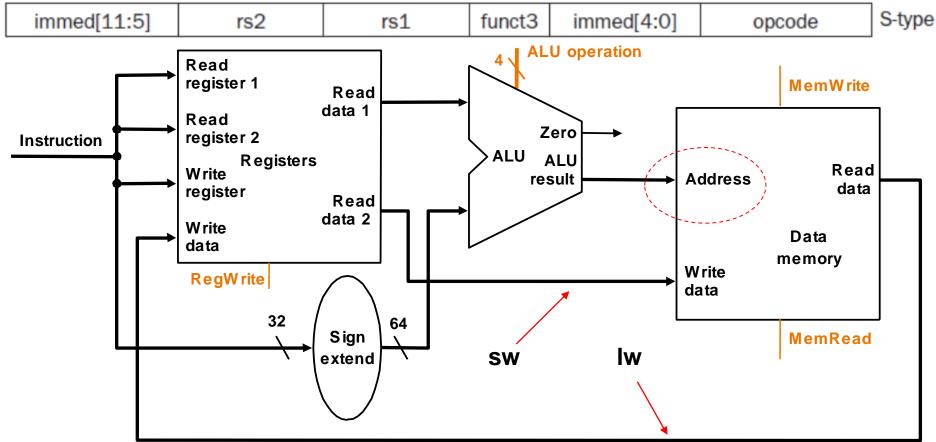
- \square lw x1, offset(x2); M(x2+offset) -> x1
- □需要对指令字中的32位偏移进行64位带符号扩展 (或者16位偏移进行32位扩展)

访存指令的执行-sw



rs:基址 rt:sw源寄存器

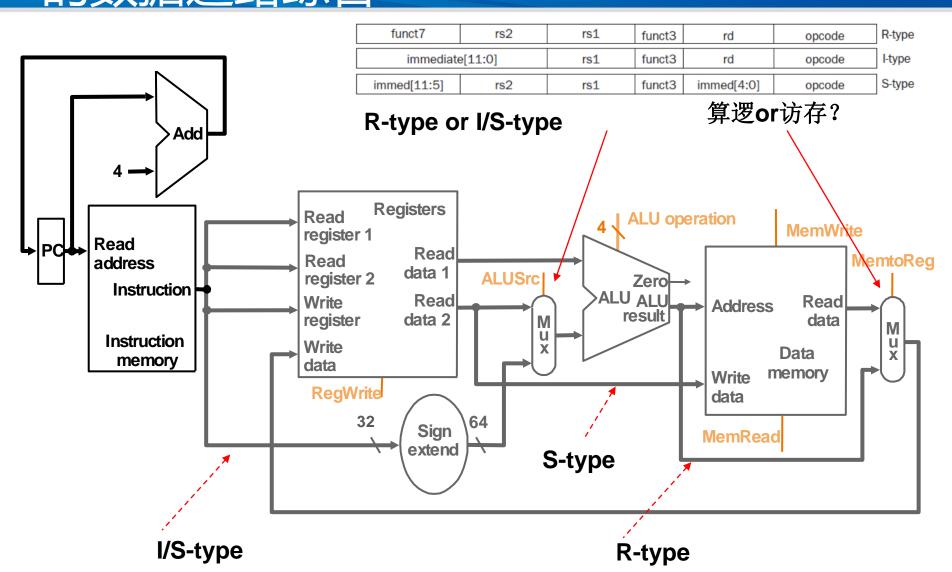
大 字



- \square sw x1, offset(x2); x1 -> M(x2+offset)
- □需要对指令字中的32位偏移进行64位带符号扩展 (或者16位偏移进行32位扩展)

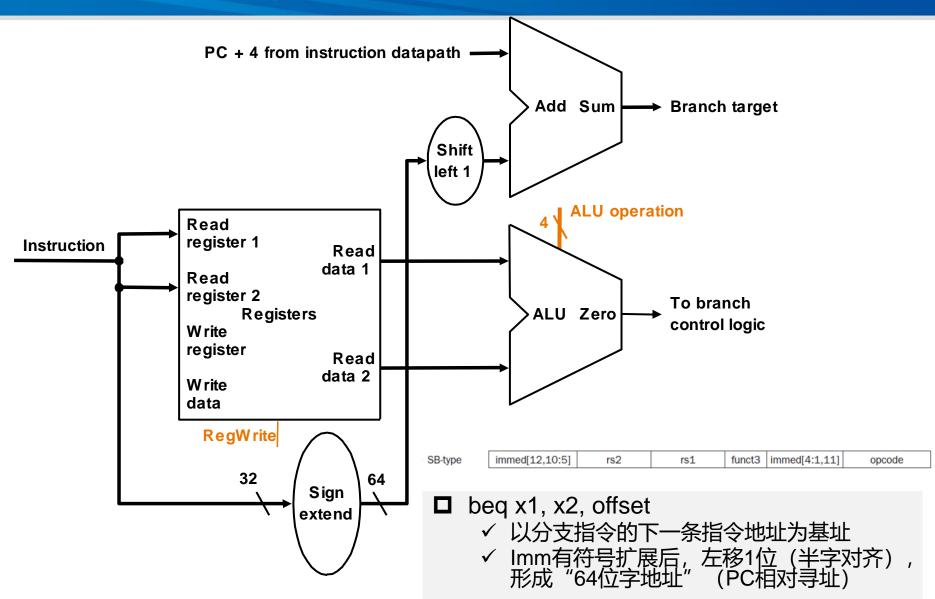
访存指令和算逻指令 的数据通路综合





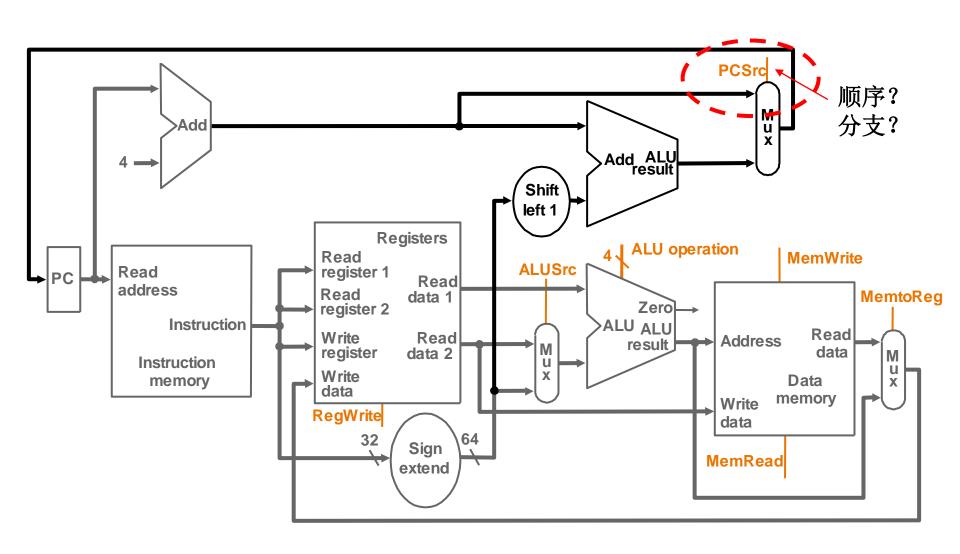
条件转移beq





R-/I-/S/SB-type数据通路总图

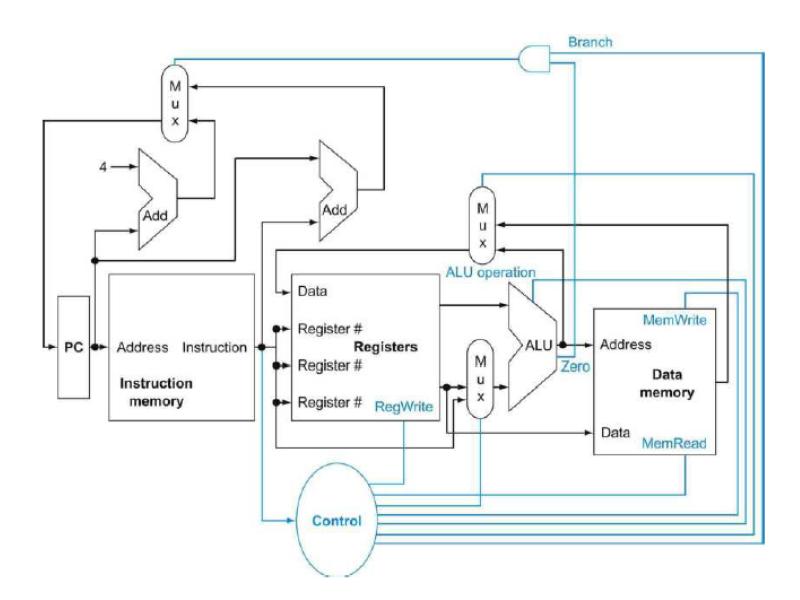




控制信号如何处理?

复习控制器-示意图

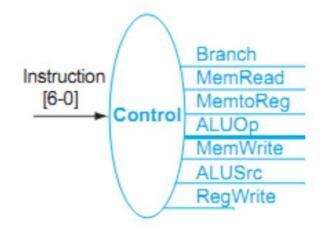




控制器:控制信号生成



- □op域(7位)译码产生的控制信号:7个
 - ✓ RegWrite: 寄存器写操作控制
 - ✓ ALUSrc: ALU的第二个操作数来源
 - R-type指令 (0) 与 I/S/SB-type指令 (1)
 - ✓ ALUOp: R-type指令 (2位)
 - ✓ MemRead: 存储器读控制, load指令
 - ✓ MemWrite: 存储器写控制, store指令
 - ✓ MemtoReg: 目的寄存器数据来源
 - R-type (I类ALU) 指令与load指令二选一
 - ✓ Branch: 是否分支指令 (产生PCSrc)
 - PCSrc: nPC来源控制,顺序与分支成功二选一
 - "是否beq指令 (Branch)" & "ALU的Zero状态有效"



ALU控制信号



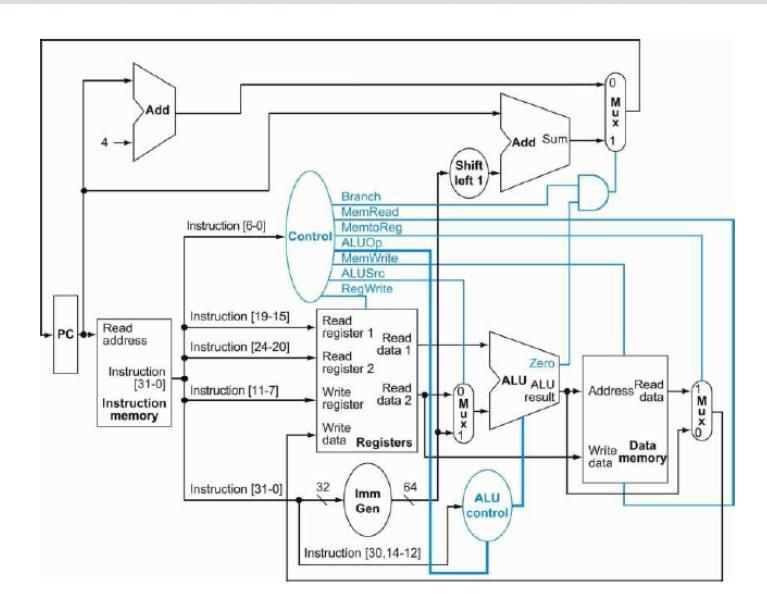
Instruction opcode	ALUOp	Operation	Funct7 field	Funct3 field	Desired ALU action	ALU control input
ld	00	load doubleword	XXXXXXX	XXX	add	0010
sd	00	store doubleword	XXXXXXX	XXX	add	0010
beq	01	branch if equal	XXXXXXX	XXX	subtract	0110
R-type	10	add	0000000	000	add	0010
R-type	10	sub	0100000	000	subtract	0110
R-type	10	and	0000000	111	AND	0000
R-type	10	or	0000000	110	OR	0001

funct7 rs2		rs1	funct3	rd	opcode	R-type
immediate	rs1	funct3	rd	opcode	I-type	
immed[11:5]	rs2	rs1	funct3	immed[4:0]	opcode	S-type
immed[12,10:5]	rs2	rs1	funct3	immed[4:1,11]	opcode	B-type

2位ALUop与Funct7、Funct3组合,产生ALU_ctrl_input (即4位ALU operation)

ALU控制选择





R-type指令的执行过程



■ add x1, x2, x3; x2+x3->x1

口在一个周期内完成如下动作

✓第一步: 取指和PC+4

✓第二步: 读两个源操作数寄存器x2和x3

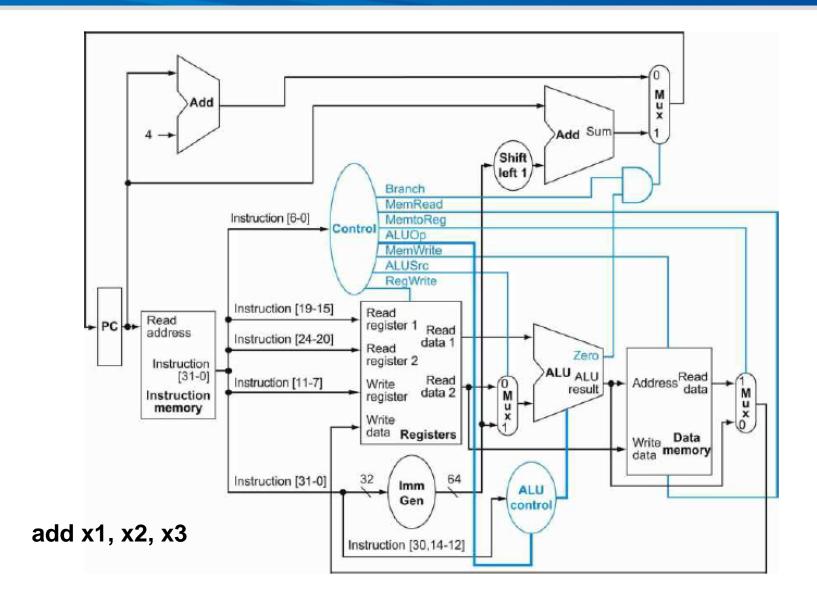
✓第三步:ALU操作

✓第四步: 结果写回目的寄存器x1

	funct7	rs2	rs1	funct3	rd	opcode	R-type
--	--------	-----	-----	--------	----	--------	--------

R-type指令的执行路径





lw指令的执行过程



 \square Iw x1, offset(x2); M(x2+offset) -> x1

□第一步: 取指和PC + 4

□第二步: 读寄存器x2

□第三步: ALU操作完成x2与符号扩展后的offset加

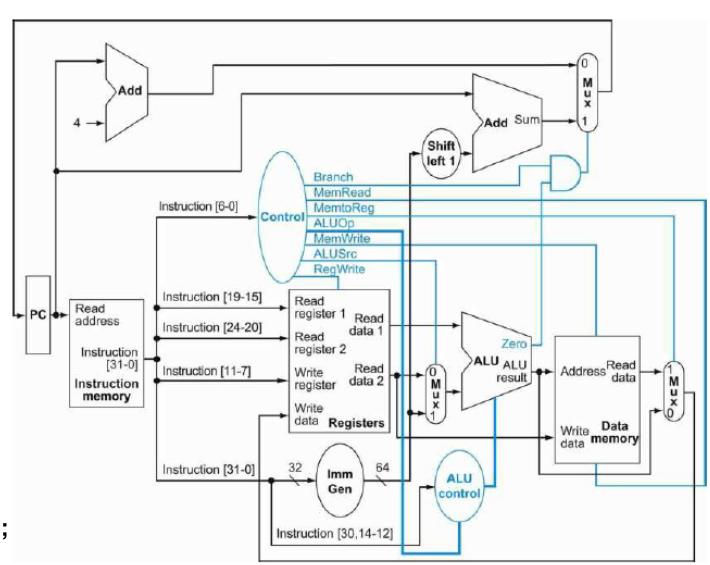
□第四步: ALU的结果作为访存地址, 送往数据MEM

□第五步:内存中的数据送往x1

immediate[11:0]	rs1	funct3	rd	opcode	1-type

lw指令的执行路径





lw x1, offset(x2);

beq指令的执行过程



□beq x1, x2, offset

□第一步: 取指和PC + 4

□第二步:读寄存器x1,x2

□第三步: ALU将x1和x2相减; PC+4与被左

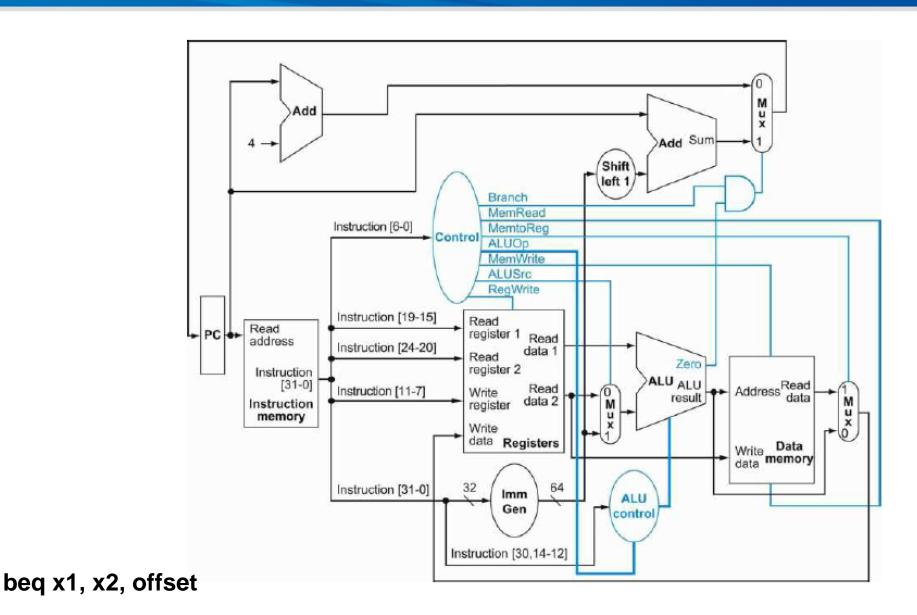
移1位并进行符号扩展后的offset相加,作

为分支目标地址

□第四步: ALU的Zero确定应送往PC的值

beq的执行路径





JAL指令的实现

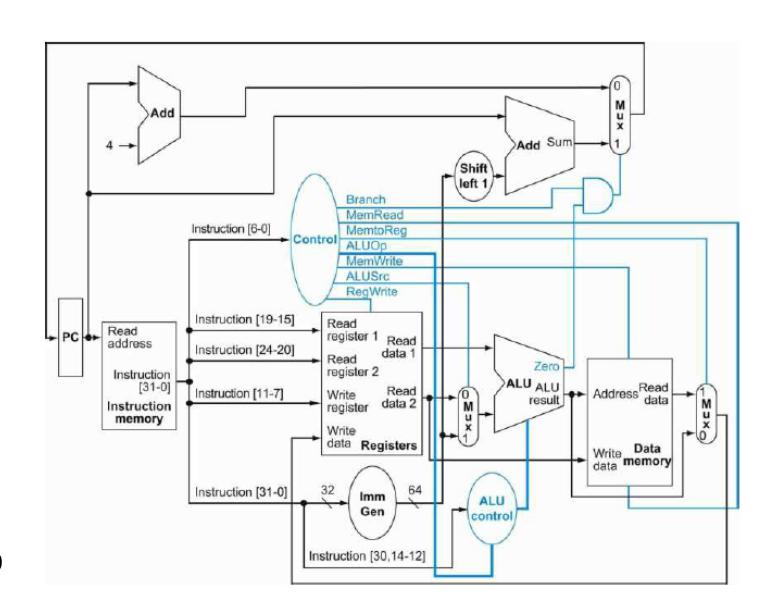




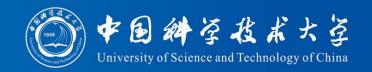
- □UJ无条件转移,关键在于目标地址的生成 ✓PC±2¹⁹
- □增加一个jump指令识别控制

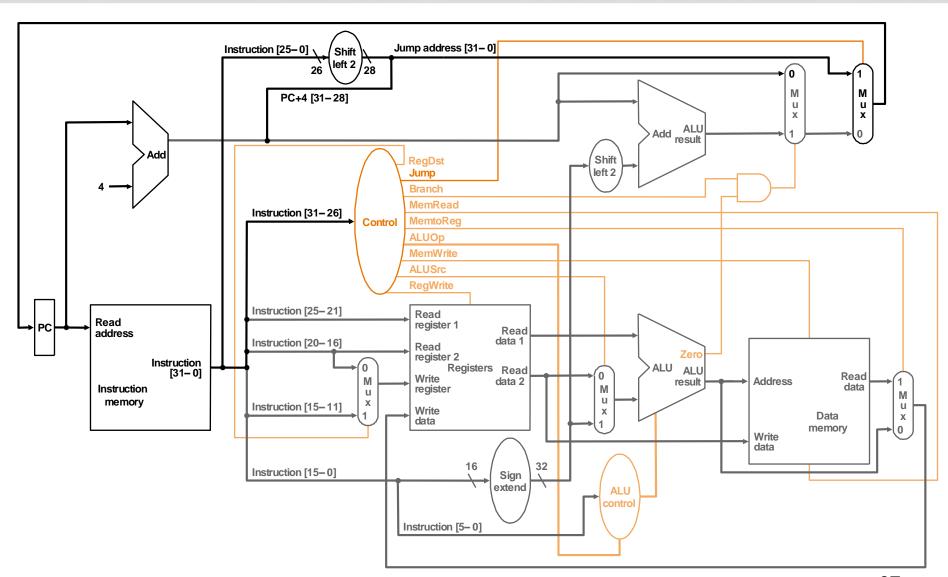
JAL指令应该如何实现





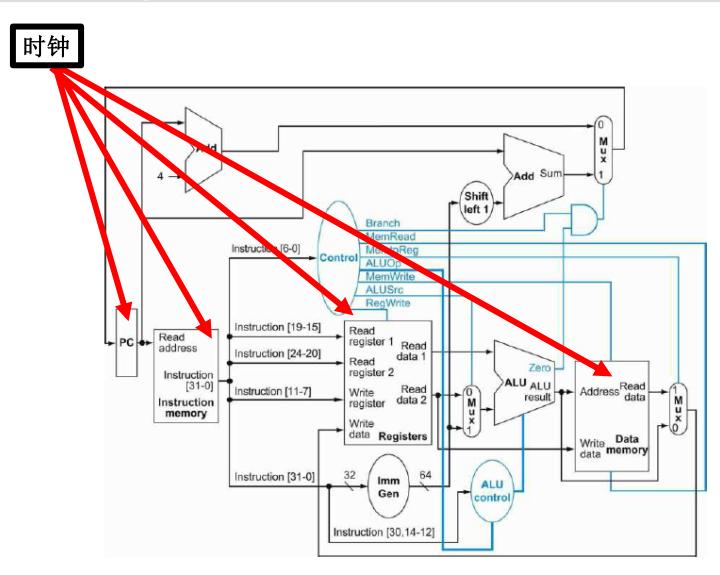
参考: JAL指令的MIPS实现





Single-Cycle Processor: Clocking





时钟周期的确定方法



- □时钟周期由关键路径确定
- □关键路径的选择=哪种指令执行时间最长?
 - ✓ PC时钟初始化=30ps 读内存=250ps
 - ✓寄存器读=150ps 寄存器写=20ps
 - ✓ALU计算=200ps 多个多路选择器共25ps (几个?)
- □画出关键路径,并计算主频

```
T_C = t_{pcq-PC} + t_{mem-I} + t_{RFread} + t_{ALU} + t_{mem-D} + t_{mux} + t_{RFsetup}
= (30 + 250 + 150 + 200 + 250 + 25 + 20) ps
= 925 ps
= 1.08 GHz
```

定长单周期or不定长单周期?



□ 设程序中load有24%, store有12%, R-type有44%, beq有18%, jump有2%。假设load, store,R,beq,jump指令的时间分别是8,7,6,5,2ns(如下表)。如果时钟周期固定,单周期的时钟为8ns;如果时钟周期不定长,单周期的时钟可以是2ns~8ns。试比较时钟定长单周期实现和不定长单周期实现的性能。

指令	inst MEM	Reg Read	ALU	Data MEM	Reg Write	Total
R-Type	2	1	2		1	6ns
lw	2	1	2	2	1	8ns
sw	2	1	2	2		7ns
beq	2	1	2			5ns
JAL	2					2ns

单周期实现



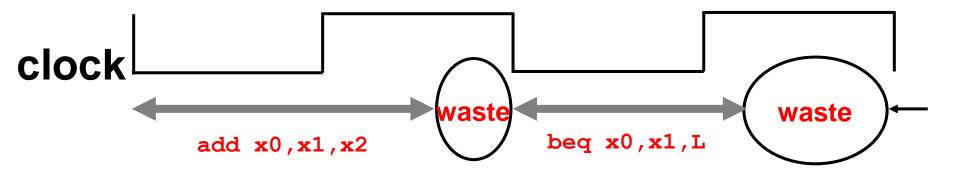
- □程序执行时间 = 指令数×CPI×时钟周期时间
- □指令周期 = 机器周期 = 时钟周期 ✓ CPI = 1
- □平均指令执行时间 = 8×24% + 7×12% + 6×44 % + 5×18% + 2×2% = 6.3ns
- □因此,变长单周期实现较定长单周期实现快 8/6.3=1.27倍

变长单周期实现?

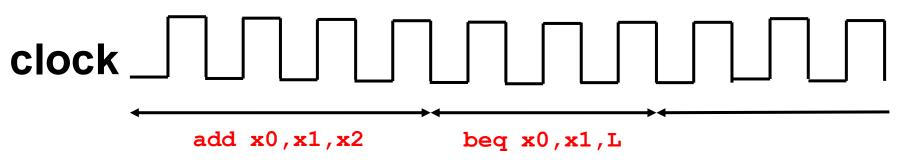
定长单周期 vs. 不定长多周期



Single-cycle Implementation: 定长指令周期



Multicycle Implementation: 不定长指令周期



Multicycle Implementation:
 less waste = higher performance

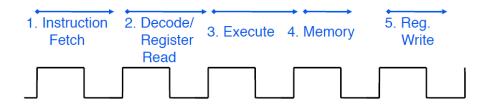
多周期实现



- □根据指令执行所使用的功能部件,将执行过程划分 成多个阶段,每个阶段—个周期(机器周期)
 - ✓早结束的指令可以提升性能
- □时钟周期确定
 - ✓每个周期的工作尽量平衡
 - ✓ 假设一个周期内可以完成



- 一次寄存器访问(2 reads or one write), or
- 一次ALU操作



指令执行的多个周期



3. Execute 4. Memory

- □共5个阶段(周期)
 - ✓取指
 - ✓译码阶段、计算beq目标地址
 - ✓执行:R-type指令执行、访存地址计算,分支完成阶段

1. Instruction

Fetch

2. Decode/

Read

Register

- ✓ 访存: Iw读, store和R-type指令完成阶段
- ✓写回: lw完成阶段
- 口注意
 - ✓ 定长机器周期: 机器周期=时钟周期
 - ✓ 不定长指令周期:分别为3、4、5个机器周期
- □控制器根据机器周期标识发出控制信号

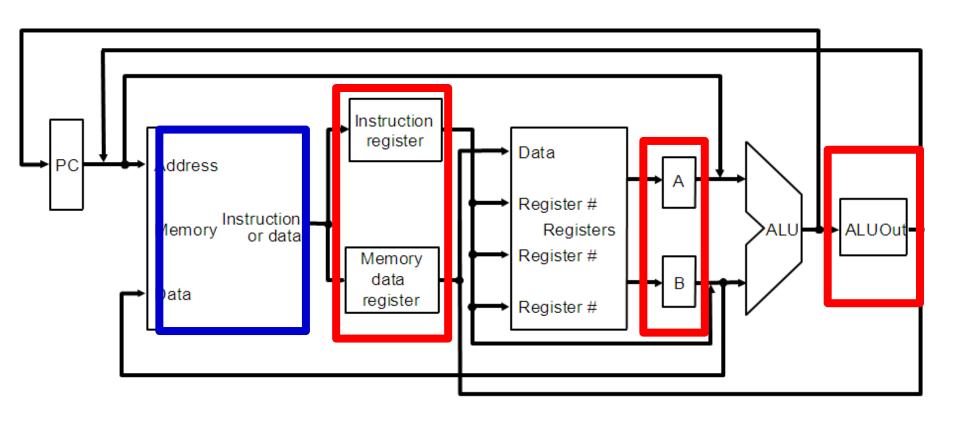
5. Reg.

Write

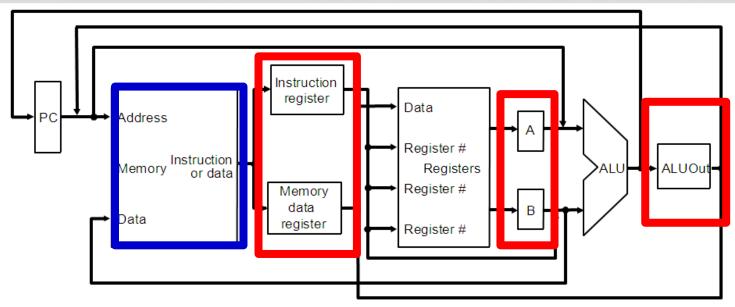
多周期简单数据通路



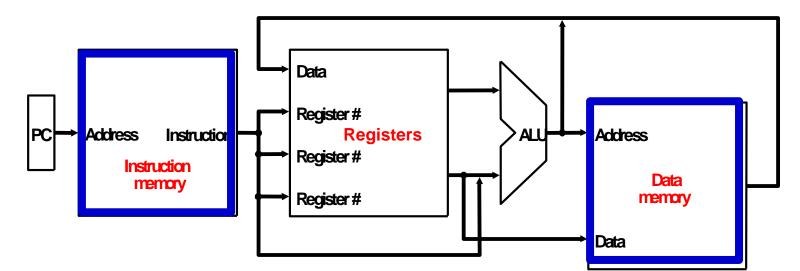
用哪些寄存器来实现多周期



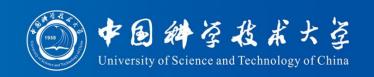
单周期和多周期简单数据通路区别 ・ 回神る投ボスタ University of Science and Technology of China

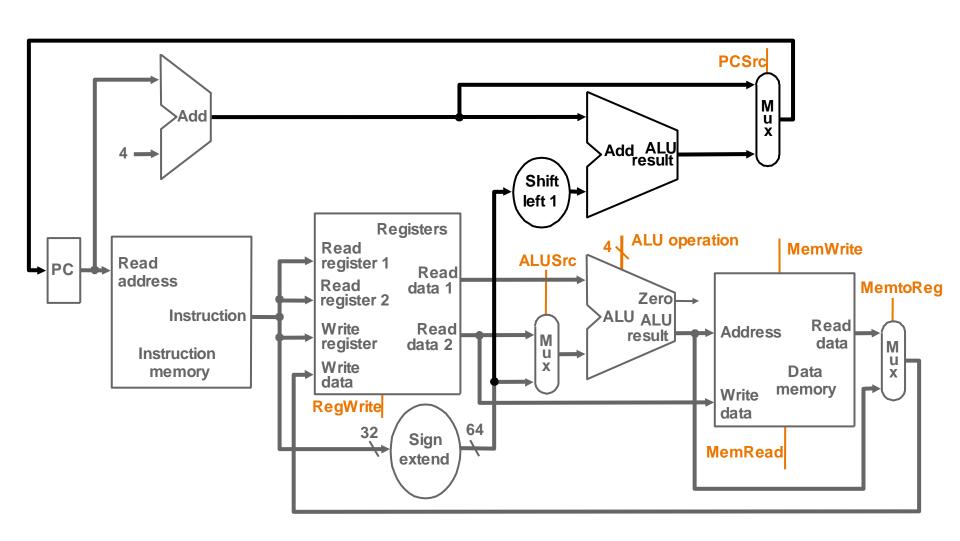


ALU、寄存器、存储器、控制信号



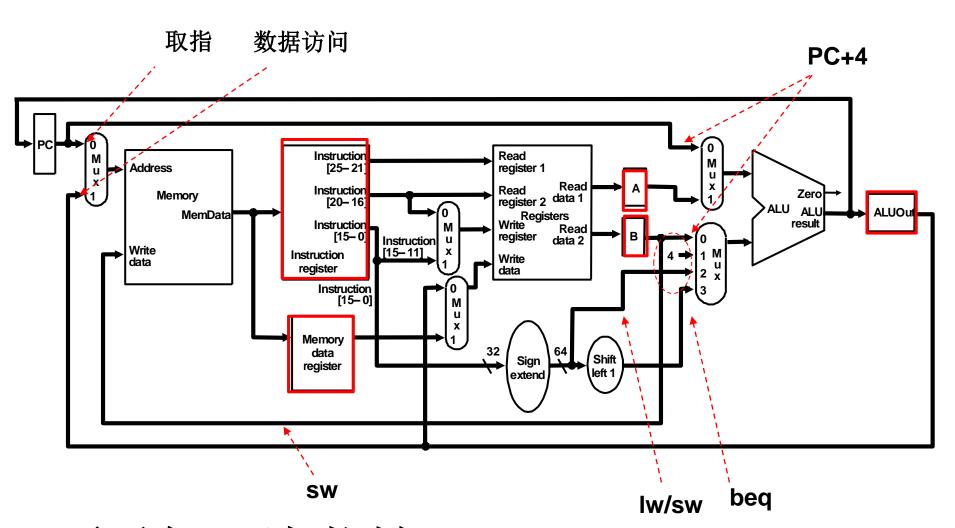
单周期数据通路





多周期数据通路 (MIPS/RISC-V)

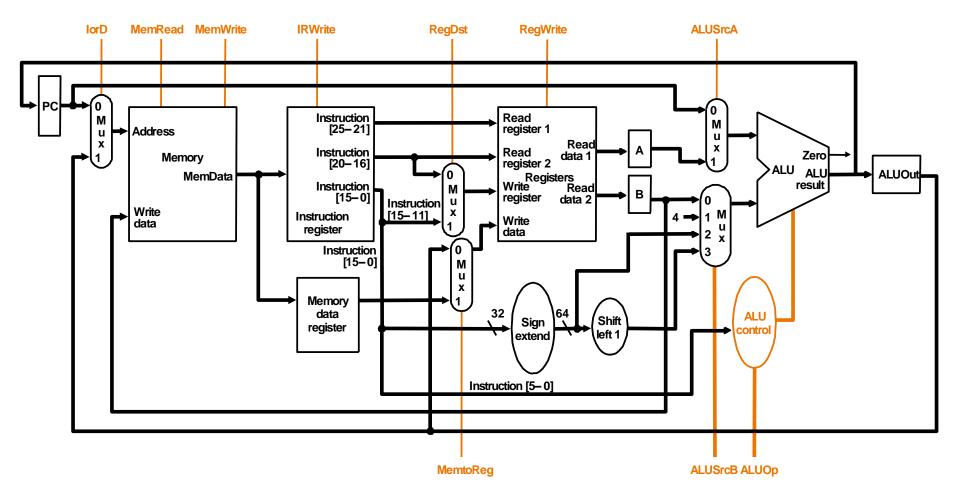
Name		Fi	eld				Comments
(Field Size)	7 bits	5 bits	5 bits	3 bits	5 bits	7 bits	
R-type	funct7	rs2	rs1	funct3	rd	opcode	Arithmetic instruction format
I-type	immediate[11:0] rs1			funct3	rd	opcode	Loads & immediate arithmetic
S-type	immed[11:5]	rs2	rs1	funct3	immed[4:0]	opcode	Stores
SB-type	immed[12,10:5]	rs2	rs1	funct3	immed[4:1,11]	opcode	Conditional branch format
UJ-type	immediate[20,10:1,11,19:12]				rd	opcode	Unconditional jump format
U-type	immediate[31:12]				rd	opcode	Upper immediate format



需要在哪里加控制?



多周期控制信号

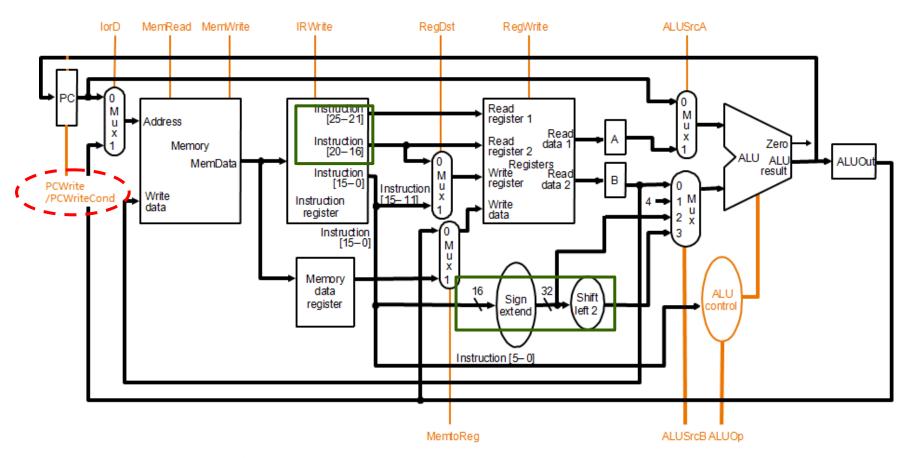


思考:为何单周期实现中beq=NPC+offset,而多周期是PC+offset?

还需要对PC的写控制信号



□ "在一个指令周期内,PC不能变"

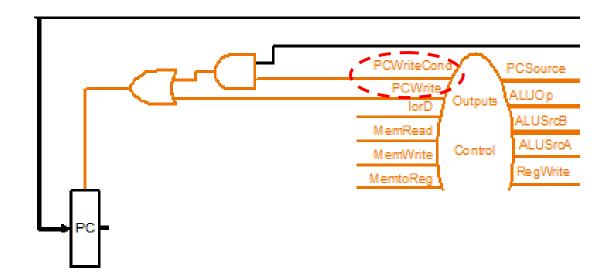


□思考:为何单周期不需要PC写控制?

PC的写控制

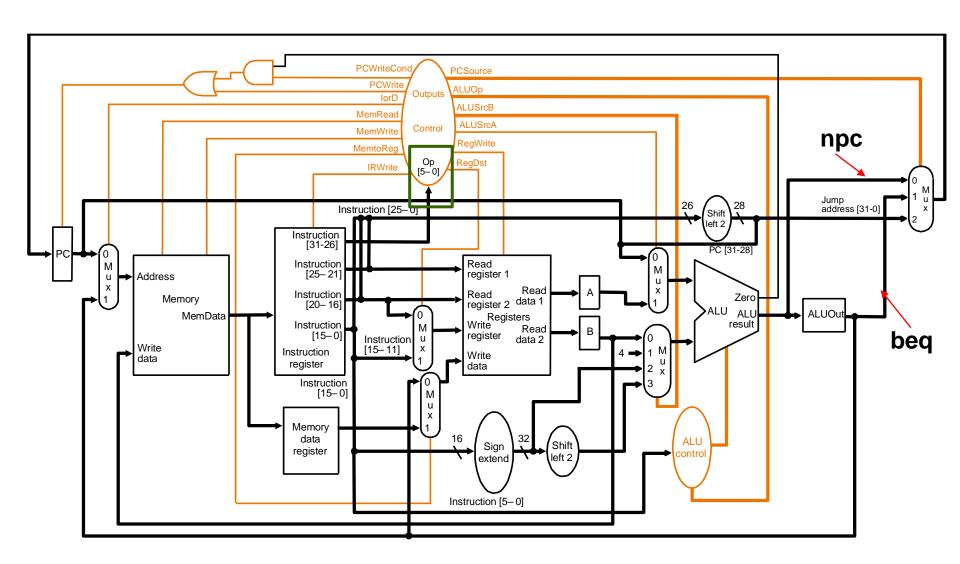


- □3种情况
 - ✓ ALU: PC + 4的输出直接存入PC,无条件写(取指阶段)
 - ✓ ALU: beq指令的目标地址(译码阶段)
 - ✓ ALU: JAL指令 UJ-type (可在译码阶段)
- □需要两个写控制
 - ✓ 无条件写PCWrite: PC+4, JAL
 - ✓ 有条件写PCWriteCond: beq



主控制部件



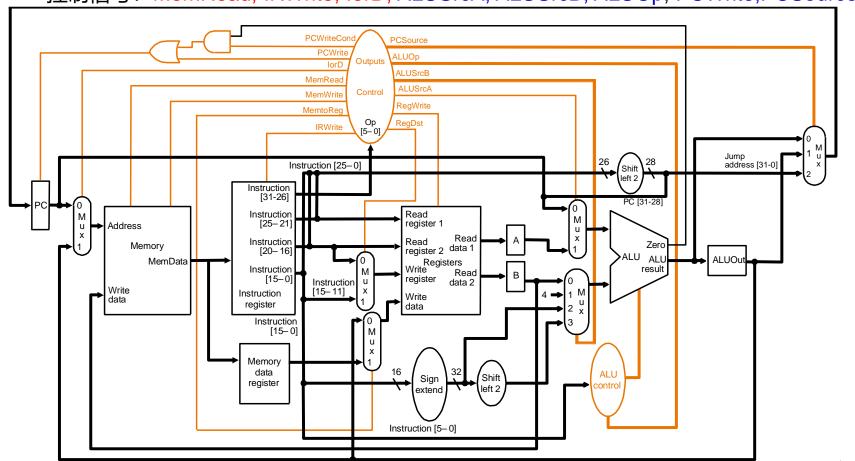


取指阶段



□ 取指

- ✓ 根据PC从MEM中取指, IR=MEM[PC]
- ✓ 计算NPC, PC=PC+4
- ✓ 控制信号: MemRead, IRWrite, IorD, ALUSrcA, ALUSrcB, ALUOp, PCWrite, PCSource



译码阶段



□ 指令译码和读寄存器

✓ 将rs和rt送往A和B: A=Reg[IR[25-21]], B=Reg[IR[20-16]]

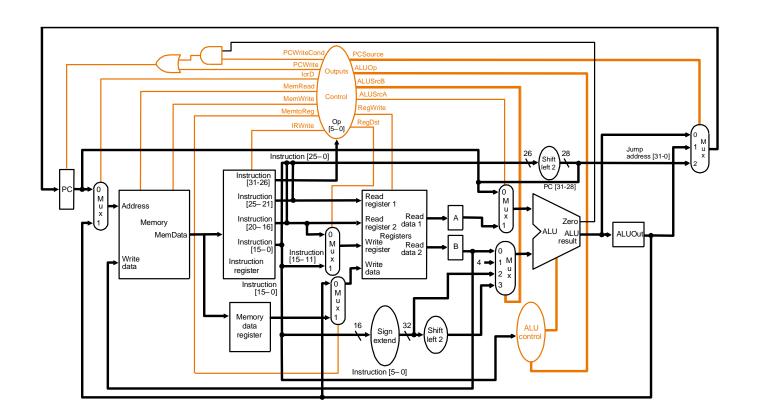
A=rs1 B=rs2

□ 计算beq目标地址

✓ ALUOut=PC+(sign-extend(IR[15-0] <<2)

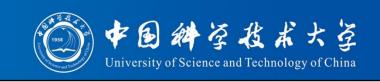
PC+offset

- · 由于此时尚不知是何指令,所以读寄存器和计算分支地址可能无效
- 控制信号: ALUSrcA, ALUSrcB, ALUOp

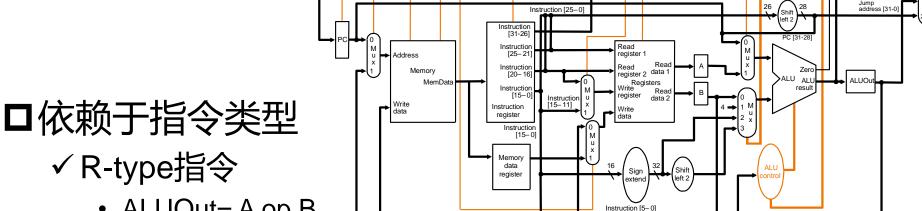


R-type执行、访存地址计算、

分支完成阶段

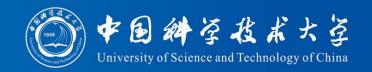


ALUSrcB



- - ALUOut= A op B
 - ✓ 访存指令: 计算访存地址
 - ALUOut=A+(sign-extend(IR[15-0]) imm
 - ✓ beg指令
 - 控制信号: ALUSrcA, ALUSrcB, ALUOp, • if (A == B) PC=ALUOut; PCWriteCond,PCSource, PCWrite
 - ✓ jump指令 PC+offset
 - PC = PC[31-28] || (IR[25-0] << 2) MIPS
 - PC = PC+IR[0-20] RISC-V

R-type和sw完成、Iw读阶段



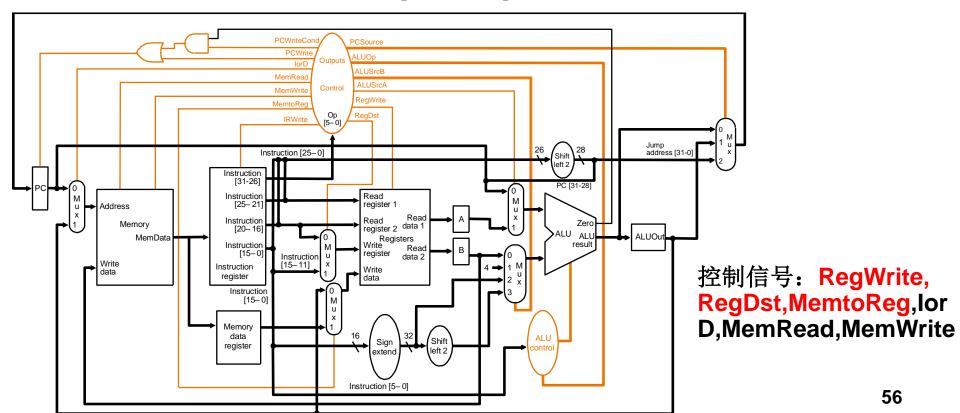
□第四阶段

rd

✓ R-type完成: 结果写回 Reg[IR[15-11]] = ALUOut

✓ sw完成: 写入MEM MEM[ALUOut] = B

✓ lw读: MDR = MEM[ALUOut]



Iw写阶段

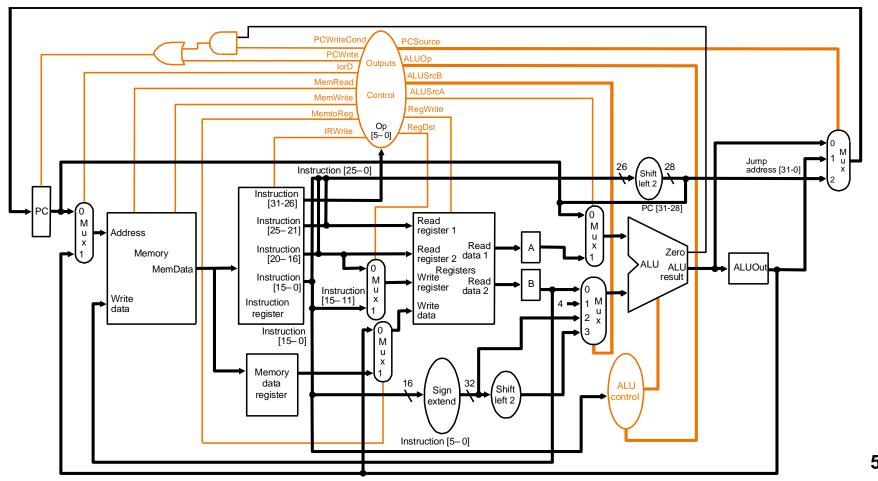


□第五阶段

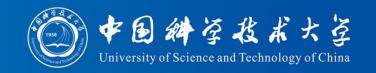
rd

✓ Iw写回: Reg[IR[15-11]] = MDR

控制信号: RegWrite, RegDst,MemtoReg



控制信号与操作的关系

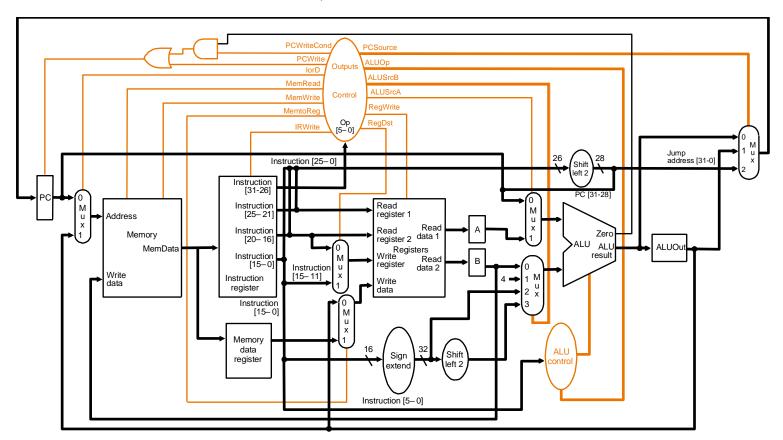


1 ALUSrcA, ALUSrcB, ALUOp (ALU操作= ALU指令or算地址) 2 RegWrite, RegDst,MemtoReg (写 Reg=ALU指令or LW指令)

3. MemRead, IRWrite, IorD (读存储器、写IR=取指)

4 PCWriteCond,PCSource, PCWrite(写PC=顺序、分支) 6. lorD,MemWrite (SW)

5. IorD,MemRead(读存储器=数据)



Multicycle RTL (参考)



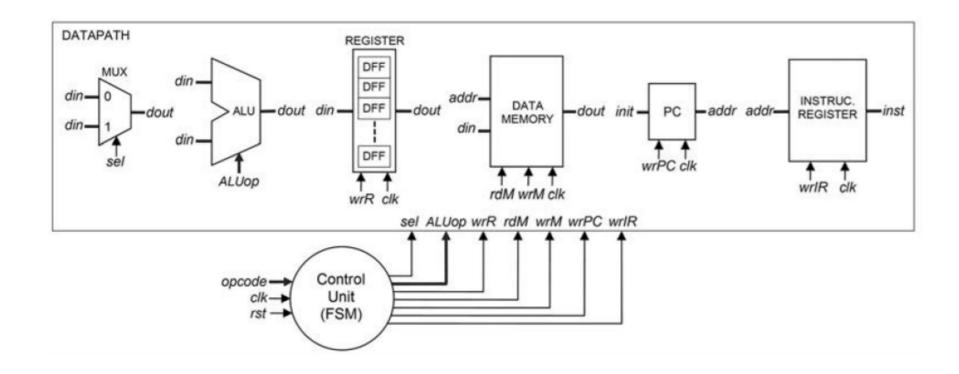
Step	R-Type		lw/sw	beq/bne		j		
IF	IR = Mem[PC] PC = PC + 4							
ID	A = Reg[IR[25-21]] B = Reg[IR[20-16]] A=rs1 B=rs2 B指令PC+offset ALUOut = PC + (SE(IR[15-0]) << 2)							
EX	ALUOut = A op B		ALUOut = A+SE(IR[15-0]) 指令PC+offset	If (A= the PC = AL	en ´	PC = PC[31-28] (IR[25-0]<<2		
MEM	Reg[IR[15-11]] = ALUOut		R=Mem[ALUOut] em[ALUOut] = B		E.	跳转指令PC+offset		
WB		Reg[IR[20-16]] = MDR	Ł				

rd

控制信号产生



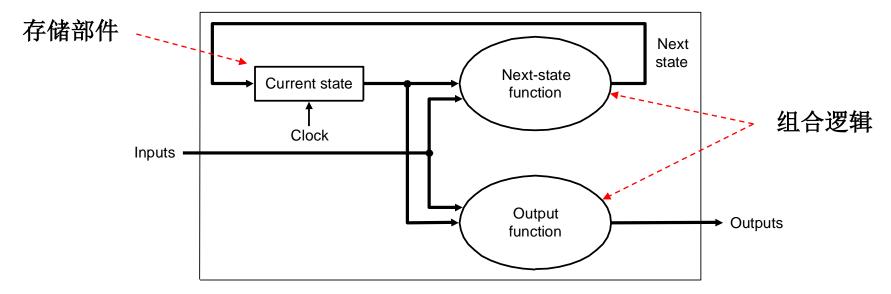
☐ FSM: Finite State Machine



FSM控制部件实现

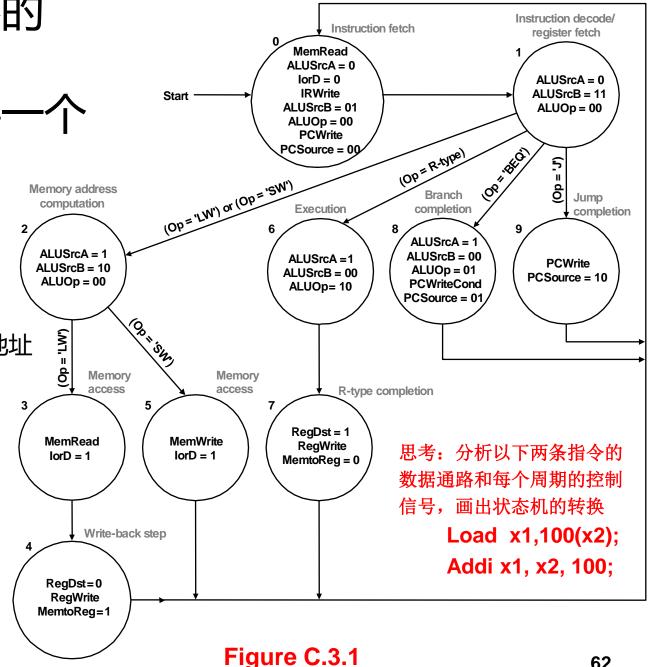


- ☐ FSM: Finite State Machine
- Moore型 (Edward Moore), Mealy型 (George Mealy)
 - ✓ Moore型速度快(输出与输入无关,可以在周期一开始就提供控制信号),输出与时钟完全同步。
 - ✓ Mealy型电路较小。输出与时钟不完全同步(一个tick内随输入而变)。
 - ✓ 两种状态机可以相互转换。
- □ EDA工具可以根据FSM自动综合生成控制器

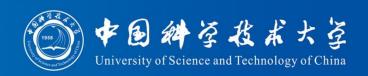


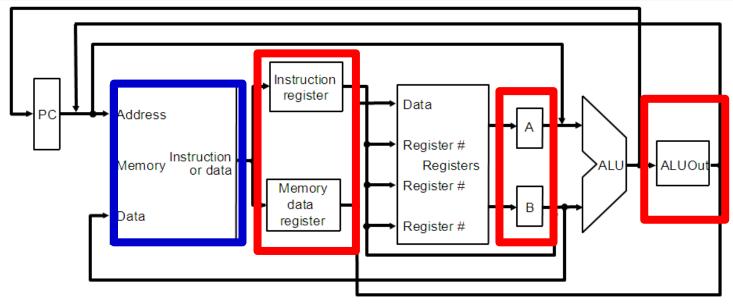
多周期控制器的 MooreFSM, 每个状态需要一 时钟周期。

- □ 取指: PC+4
- □译码: BEQ/Jump算地址
- □执行
 - ✓ R执行
 - ✓ LW/SW算地址
 - ✓ 分支完成
- □访存
 - ✓ R/SW完成
 - ✓ Load访存
- - ✓ Load完成

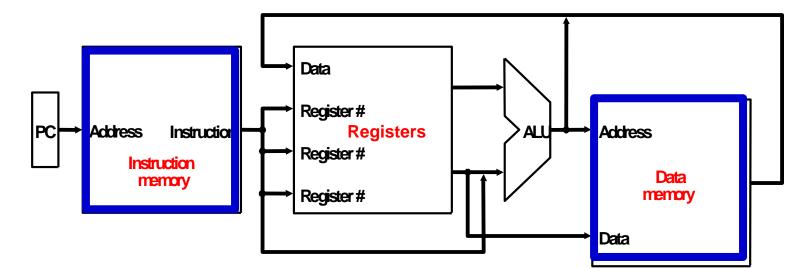


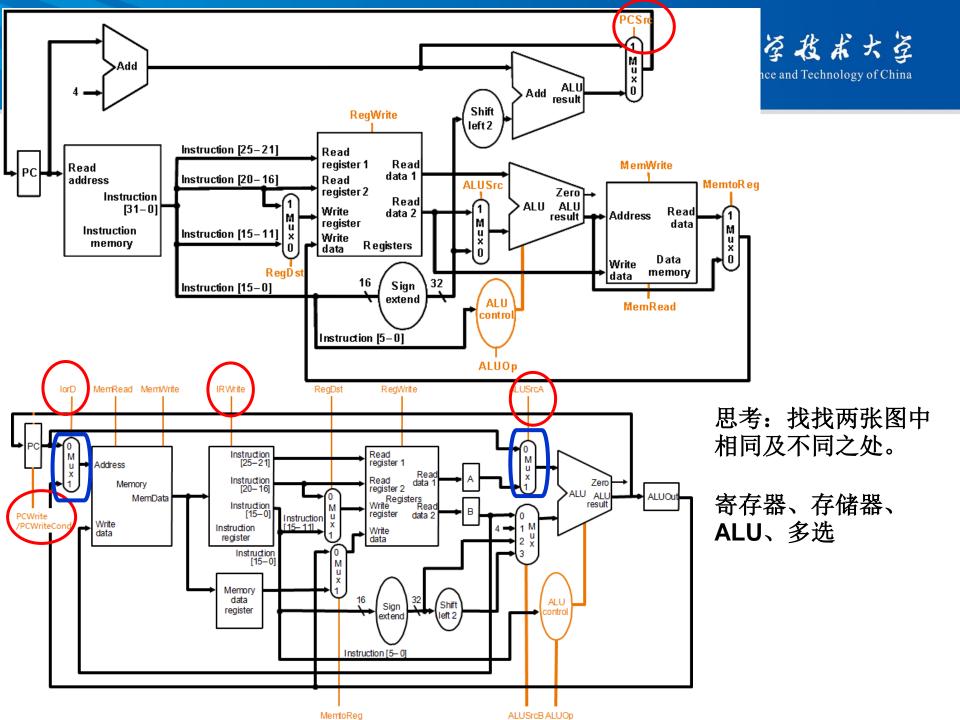
单周期和多周期简单数据通路区别单周期:COD4 多周期: COD3 Ch5

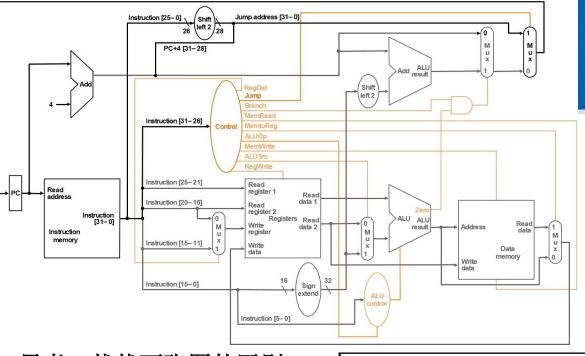




ALU、寄存器、存储器、控制信号



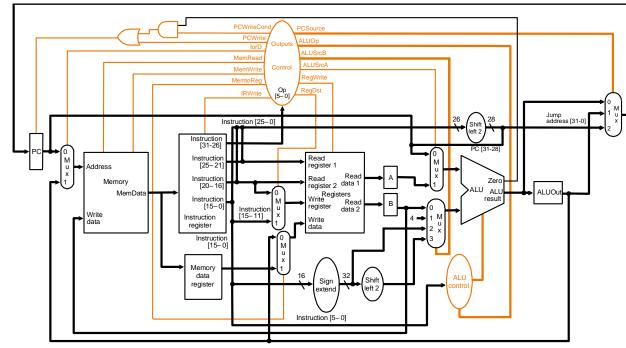






思考: 找找两张图的区别

- 一、数据通路的区别
 - 1.存储器
 - 2.寄存器
 - 3.ALU
 - 4.MUX
- 二、控制信号的区别
 - 1. 寄存器文件写使能
 - 2. 存储器读写使能
 - 3. ALU控制
 - 4. PC/IR 寄存器写
 - 5. 多路选择器 2 时序单周期V.S.状态机



小结



- □ 如何设计一款CPU (单周期or多周期)
 - ✓ 确定指令→实现操作→数据通路→运行部件→RTL 实现→控制信号→综合仿真→开发板调试→性能评 测→迭代
- □ 单周期与多周期
 - ✓ 单周期: 在一个周期内完成指令的所有操作
 - 周期宽度如何确定,能否在一个clk内完成?
 - ✓ 多周期: 一个周期完成指令的一步
 - 与单周期的区别:功能部件复用,中间结果保存
 - 可以实现不定长指令周期,提高性能
 - 按当前周期产生响应的控制信号(状态机)
 - 何时刷新PC,何时访存,何时写寄存器

思考



- □ 单周期: 一个周期内完成指令的所有微操作
 - ✓ 寻址方式是如何实现的?
 - ✓ 周期宽度如何确定?
 - ✓ 能否在一个clk内完成?
 - ✓ 控制逻辑有哪些实现方式?
 - ✓ 单周期的指令和数据内存分开,ALU和加法器分开,原因是什么?
- □ 多周期:一个周期完成指令的一步(微操作)
 - ✓ 与单周期的区别:功能部件复用,中间结果保存
 - ✓ 可以实现不定长指令周期,提高性能
 - ✓ 能否采用单周期的数据通路?
 - ✓ 按当前周期(标识?)产生响应的控制信号
 - ✓ 何时刷新PC:指令周期中PC保持不变如何实现?
- □ 每一类指令的指令周期各含多少个时钟周期?
- □ 比较FSM控制器实现方式的特点。



问:有一段程序共有100条指令,其中Load指令为 20%, Store指令为20%, R类的ALU指令为40%, Beg指令为10%, JUMP指令为10%。这段程序分别 在单周期和多周期的CPU上实现(多周期参考 RISC-V的CPI), 其中单周期CPU时钟为200MHz, 多周期CPU时钟为800MHz(考虑到存储器、ALU实 现复杂),问在何种CPU上运行时快,快多少?

答:略。

Multicycle RTL (参考)



Step	R-Type		lw/sw	beq/bne		j		
IF	IR = Mem[PC] PC = PC + 4							
ID	A = Reg[IR[25-21]] B = Reg[IR[20-16]] A=rs1 B=rs2 B指令PC+offset ALUOut = PC + (SE(IR[15-0]) << 2)							
EX	ALUOut = A op B		ALUOut = A+SE(IR[15-0]) 指令PC+offset	If (A= the PC = AL	en ´	PC = PC[31-28] (IR[25-0]<<2		
MEM	Reg[IR[15-11]] = ALUOut		R=Mem[ALUOut] em[ALUOut] = B		E.	跳转指令PC+offset		
WB		Reg[IR[20-16]] = MDR	Ł				

rd

作业



单周期作业内容:

1、习题:(以英文版为主,中文版作为参考,交)

COD5: 4.1 4.1 EN 4.1 CN figure4.10

COD5: 4.7 4.7 EN 4.7 CN

2、思考题: (交)

单周期处理器在一个周期内完成指令所有的微操作,思考:

- --寻址方式如何实现
- --周期宽度如何确定
- --能否 "在一个clk内完成"
- --能否将两个adder合二为一
- --能否将两个memory合二为一
- 3、根据图<u>4-22</u> ,画出主控制器的PLA实现图(思考,不交)
 - --先写布尔表达式
 - --请参考英文版教材附录C第2节
 - --由于PLA规模不小,可以手绘,也可以用软件画完后打印出来

多周期内容:

- 1. 每一类指令的指令周期各含多少个时钟周期? (交)
- 2、分别分析R/I/S/B-type指令的多周期设计方案中每个周期所用到的功能部件(交)
- 3、调研不同FSM控制器实现方式的特点
- 4、调研水平微指令和垂直微指令的特点(参考唐书第10章内容)

联系方式



Acknowledgements:

This slides contains materials from following lectures:

- Computer Architecture (NUDT)
- CS 152 CS 252 (UC Berkeley)

Research Area:

- ➤ 基于分布式系统, GPU,FPGA的神经网络、图计算加速
- 人工智能和深度学习(寒武纪)芯片及智能计算机

Contact:

- 中国科学技术大学计算机学院
- ▶ 嵌入式系统实验室(西活北一楼)
- 高效能智能计算实验室 (中科大苏州研究院)

cswang@ustc.edu.cn

http://staff.ustc.edu.cn/~cswang