Intro to Architecture of Computers – II The Von Neumann Model

Professor Hugh C. Lauer CS-2011, Machine Organization and Assembly Language

(Slides include copyright materials from *Computer Systems: A Programmer's Perspective*, by Bryant and O'Hallaron, and from *The C Programming Language*, by Kernighan and Ritchie)

Prior to John Von Neumann

- "Programs" for mechanical and electro-mechanical calculators were
 - Wired into plug-board or
 - Coded onto fixed punched cards or paper tape
 - Executed strictly as finite state machines

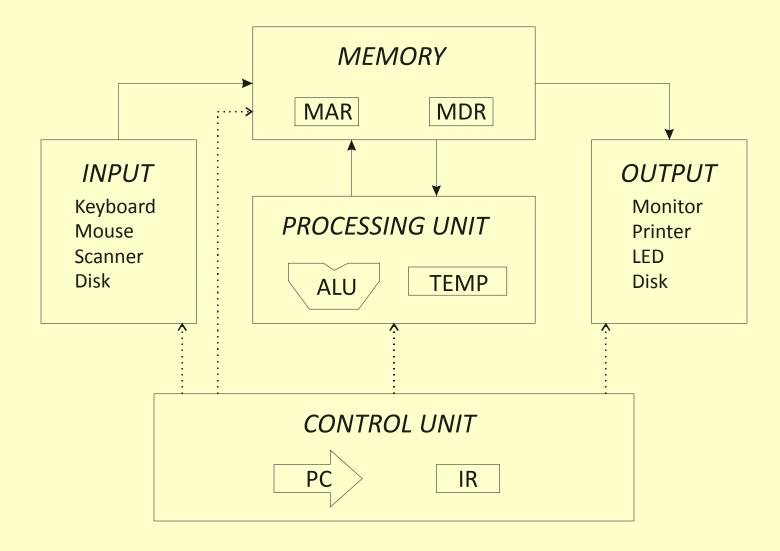
Von Neumann's innovation

- Store programs in memory of the machine
- Enable programs to update their own code!

Turing's contribution

 These two elements constituted a realistic implementation of his mathematical model of computability

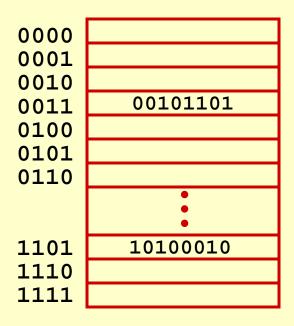
Von Neumann Model



Memory

- **2** $^k \times m$ array of stored bits
- Address
 - unique (k-bit) identifier of location
- Contents
 - m-bit value stored in location
- Basic Operations:
- LOAD
 - read a value from a memory location
- STORE

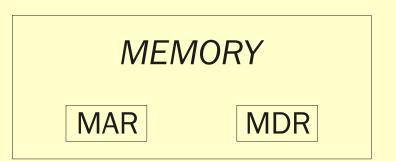
write a value to a memory location



In reality, the Memory Hierarchy

Interface to Memory

- How does processing unit get data to/from memory?
- MAR: Memory Address Register
- MDR: Memory Data Register
- To LOAD a location (A):
 - 1. Write address (A) into MAR
 - 2. Send a "read" signal to memory
 - 3. (After the passage of time) Read data from MDR
- To STORE a value (X) to a location (A):
 - 1. Write data (X) to MDR
 - 2. Write address (A) into MAR
 - 3. Send a "write" signal to the memory



Processing Unit

Functional Units

- ALU = Arithmetic and Logic Unit
- could have many functional units.
 some of them special-purpose
 (multiply, square root, ...)
- x86-64 has about 6-8 units

Registers

- Small, temporary storage
- Operands and results of functional units
- IA32 has eight %eax, %ecx, %edx, %ebx, %esi, %edi, %ebp, %esp

Word Size

- number of bits normally processed by ALU in one instruction
- also width of registers
- IA32 is 32-bits; x86_64 is 64-bits



Input and Output

- Devices for getting data into and out of computer memory
- Each device has its own interface, usually a set of registers like the memory's MAR and MDR

INPUT

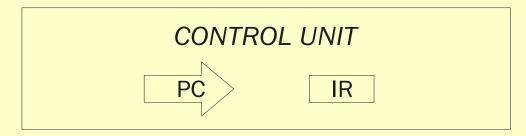
Keyboard Mouse Scanner Disk OUTPUT

Monitor Printer LED Disk

- Laptops and desktop have keyboards and mice
- Smartphones, game consoles have other devices
- Embedded systems have digital channels, analog channels, etc.
- monitor: data register (DDR) and status register (DSR)
- Some devices provide both input and output
 - disk, network
- Program that controls access to a device is usually called a *driver*.

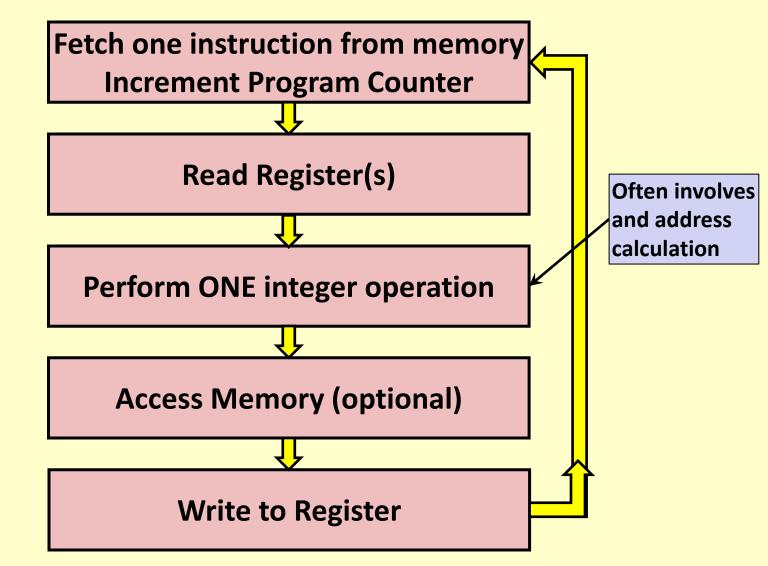
Control Unit

Orchestrates execution of the program



- Instruction Register (IR) contains the <u>current instruction</u>.
- Program Counter (PC) contains the <u>address</u> of the next instruction to be executed.
- Control unit:
 - reads an instruction from memory
 - the instruction's address is in the PC
 - interprets the instruction, generating signals that tell the other components what to do
 - an instruction may take many *machine cycles* to complete

Execution Model for Modern Computers



From an earlier topic

valE, valM

Addr, Data

aluA, aluB

srcA, srcB

dstA, dstB

valM

Data

memory

valE

valA, valB

PC increment

Register B

ALU

Sequential Hardware Structure

State

- Program counter register (PC)
- Condition code register (CC)
- Register File
- Memories
 - Access same memory space
 - Data: for reading/writing program data
 - Instruction: for reading instructions

Instruction Flow

- Read instruction at address specified by Decode
 PC
- Process through stages
- Update program counter

Memory **Execute** icode ifun Instruction | **Fetch**

PC

Write back

Figure 4.22, p. 376

Sequential Stages

Fetch

- Read instruction from instruction memory
- Increment program counter (%eip)

Decode

Read registers named in instruction

Execute

Compute value or address

Memory

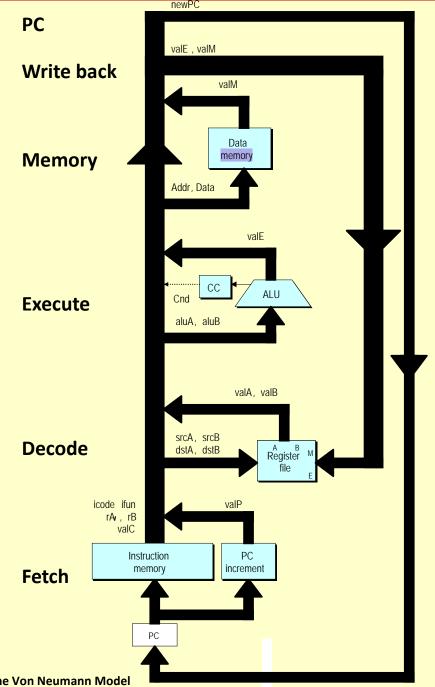
Read or write data

Write Back

Write program registers

PC

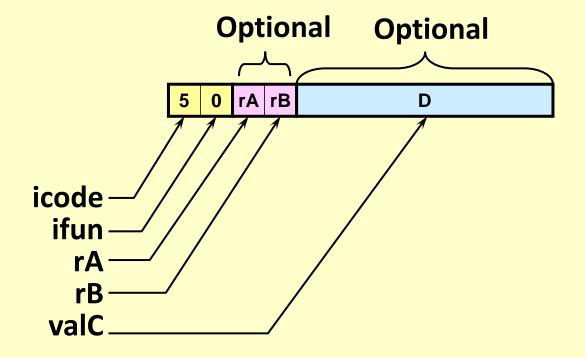
Update program counter (incremented or jump/call/ret target)



Instruction

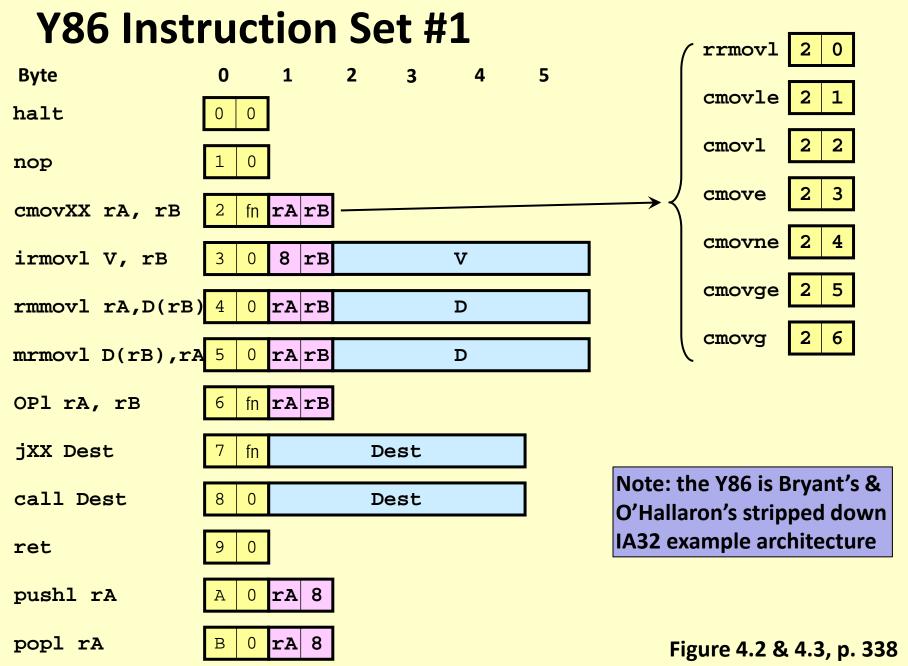
- The instruction is the fundamental unit of work.
- Specifies two things:
 - opcode: operation to be performed
 - operands: data/locations to be used for operation
- An instruction is encoded as a <u>sequence of bits</u>. (Just like data!)
 - May be fixed or variable length
 - Control unit interprets instruction; generates sequence of control signals to carry out operation.
 - Operation is either executed completely, or not at all.
- A computer's instructions and their formats is known as its Instruction Set Architecture (ISA).

Instruction Decoding

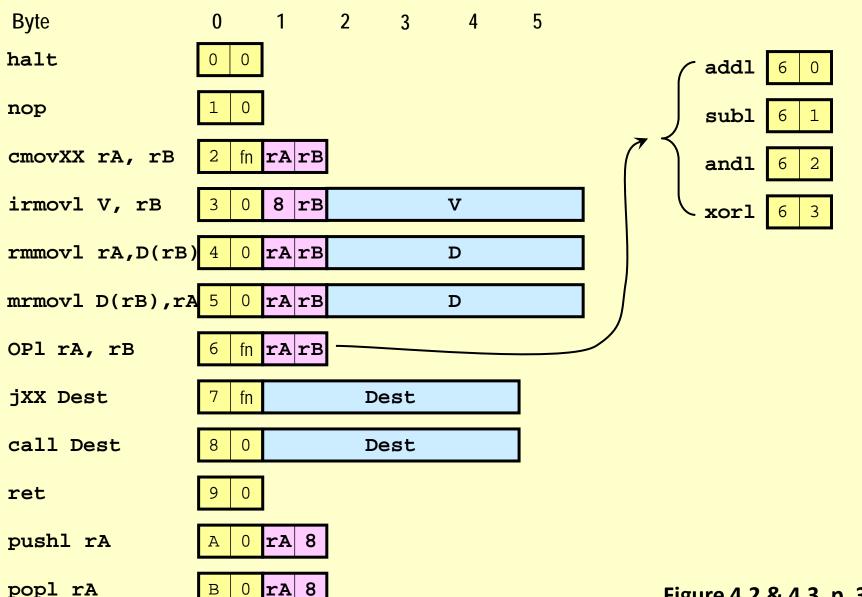


Instruction Format

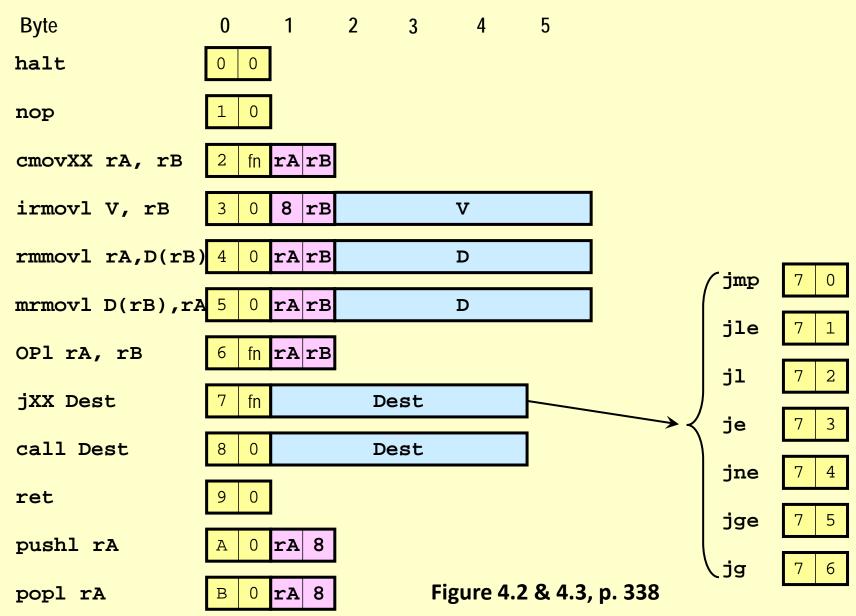
- Instruction byte icode:ifun
- Optional register byte rA:rB
- Optional constant word valC



Y86 Instruction Set #2



Y86 Instruction Set #3



Executing Arithmetic/Logical Operation

OP1 rA, rB 6 fn rA rB

■Fetch

Read 2 bytes

■Decode

Read operand registers

■Execute

- Perform operation
- Set condition codes

■Memory

Do nothing

■Write back

Update register

■PC Update

Increment PC by 2

Stage Computation: Arith/Logical Ops

	OPI rA, rB	
	icode:ifun $\leftarrow M_1[PC]$	
Fetch	$rA:rB \leftarrow M_1[PC+1]$	
	valP ← PC+2	(
Decode	valA ← R[rA]	١
Decode	valB ← R[rB]	
Execute	valE ← valB OP valA	
Execute	Set CC	!
Memory		
Write	R[rB] ← valE	
back		
PC update	PC ← valP	ı

Read instruction byte Read register byte

Compute next PC

Read operand A

Read operand B

Perform ALU operation

Set condition code register

Write back result

Update PC

- Formulate instruction execution as sequence of simple steps
- Use same general form for all instructions

Executing rmmovl

rmmovl rA, D(rB) 4 0 rA rB D

- **■**Fetch
 - Read 6 bytes
- **■**Decode
 - Read operand registers
- **■**Execute
 - Compute effective address

- **■**Memory
 - Write to memory
- **■**Write back
 - Do nothing
- **■PC Update**
 - Increment PC by 6

Stage Computation: rmmovl

	rmmovl rA, D(rB)	
	icode:ifun $\leftarrow M_1[PC]$	Read instruction byte
Fetch	$rA:rB \leftarrow M_1[PC+1]$	Read register byte
	$valC \leftarrow M_4[PC+2]$	Read displacement D
	valP ← PC+6	Compute next PC
Decode	$valA \leftarrow R[rA]$	Read operand A
	valB ← R[rB]	Read operand B
Execute	valE ← valB + valC	Compute effective address
Memory	$M_4[valE] \leftarrow valA$	Write value to memory
Write		
back		
PC update	PC ← valP	Update PC

Use ALU for address computation

Executing popl

popl rA b 0 rA 8

- **■**Fetch
 - Read 2 bytes
- **■**Decode
 - Read stack pointer
- **■**Execute
 - Increment stack pointer by 4

■Memory

Read from old stack pointer

■Write back

- Update stack pointer
- Write result to register

■PC Update

Increment PC by 2

Stage Computation: popl

	popl rA	
Fetch	icode:ifun $\leftarrow M_1[PC]$ rA:rB $\leftarrow M_1[PC+1]$ valP $\leftarrow PC+2$	
Decode	$valA \leftarrow R[%esp]$ $valB \leftarrow R[%esp]$	
Execute	valE ← valB + 4	
Memory	$valM \leftarrow M_4[valA]$	
Write	R[%esp] ← valE	
back	$R[rA] \leftarrow valM$	
PC update	PC ← valP	

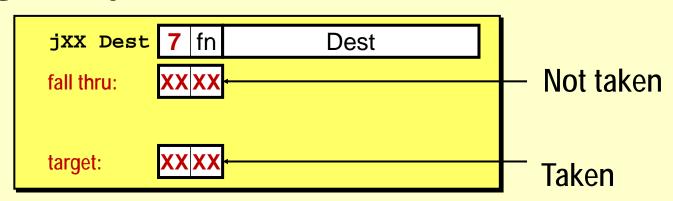
Read instruction byte Read register byte

Compute next PC
Read stack pointer
Read stack pointer
Increment stack pointer

Read from stack
Update stack pointer
Write back result
Update PC

- Use ALU to increment stack pointer
- Must update two registers
 - Popped value
 - New stack pointer

Executing Jumps



■Fetch

- Read 5 bytes
- Increment PC by 5

■Decode

Do nothing

Execute

 Determine whether to take branch based on jump condition and condition codes

■Memory

Do nothing

■Write back

Do nothing

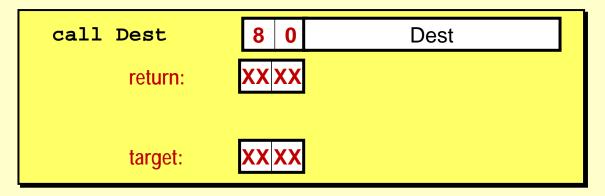
■PC Update

 Set PC to Dest if branch taken or to incremented PC if not branch **Stage Computation: Jumps**

	jXX Dest	•
Fetch	icode:ifun ← M ₁ [PC]	Read instruction byte
	$valC \leftarrow M_4[PC+1]$	Read destination address
	valP ← PC+5	Fall through address
Decode		
Execute	$Cnd \leftarrow Cond(CC,ifun)$	Take branch?
Memory		
Write		
back		
PC update	PC ← Cnd ? valC : valP	Update PC

- Compute both addresses
- Choose based on setting of condition codes and branch condition

Executing call



■Fetch

- Read 5 bytes
- Increment PC by 5

■Decode

Read stack pointer

■Execute

Decrement stack pointer by 4

■Memory

 Write incremented PC to new value of stack pointer

■Write back

Update stack pointer

■PC Update

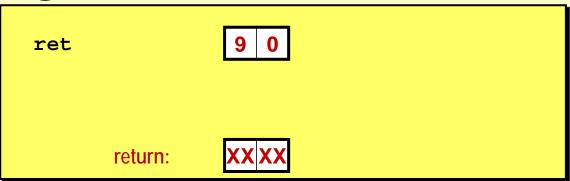
Set PC to Dest

Stage Computation: call

	call Dest	
Fetch	icode:ifun $\leftarrow M_1[PC]$ valC $\leftarrow M_4[PC+1]$	Read instruction byte Read destination address
	valP ← PC+5	Compute return point
Decode	valB ← R[%esp]	Read stack pointer
Execute	valE ← valB + -4	Decrement stack pointer
Memory	$M_4[valE] \leftarrow valP$	Write return value on stack
Write	R[%esp] ← valE	Update stack pointer
back		
PC update	PC ← valC	Set PC to destination

- Use ALU to decrement stack pointer
- Store incremented PC

Executing ret



■Fetch

Read 1 byte

■Decode

Read stack pointer

■Execute

Increment stack pointer by 4

■Memory

 Read return address from old stack pointer

■Write back

Update stack pointer

■PC Update

Set PC to return address

Stage Computation: ret

	ret	
Fetch	icode:ifun ← M ₁ [PC]	
Decode	$valA \leftarrow R[%esp]$ $valB \leftarrow R[%esp]$	
Execute	valE ← valB + 4	
Memory	$valM \leftarrow M_4[valA]$	
Write	R[%esp] ← valE	
back		
PC update	PC ← valM	

Read instruction byte

Read operand stack pointer Read operand stack pointer Increment stack pointer

Read return address Update stack pointer

Set PC to return address

- Use ALU to increment stack pointer
- Read return address from memory

Computation Steps

•		OPI rA, rB
Fetch	icode,ifun	icode:ifun ← $M_1[PC]$
	rA,rB	$rA:rB \leftarrow M_1[PC+1]$
reicii	valC	
	valP	vaIP ← PC+2
Decode	valA, srcA	valA ← R[rA]
	valB, srcB	valB ← R[rB]
Execute	valE	valE ← valB OP valA
	Cond code	Set CC
Memory	valM	
Write	dstE	R[rB] ← valE
back	dstM	
PC update	PC	PC ← valP

Read instruction byte Read register byte [Read constant word] Compute next PC Read operand A Read operand B Perform ALU operation Set condition code register [Memory read/write] Write back ALU result [Write back memory result] **Update PC**

- All instructions follow same general pattern
- Differ in what gets computed on each step

Computation Steps

•		call Dest
Fetch	icode,ifun	icode:ifun ← $M_1[PC]$
	rA,rB	
i etcii	valC	$valC \leftarrow M_4[PC+1]$
	valP	valP ← PC+5
Decode	valA, srcA	
Decode	valB, srcB	$valB \leftarrow R[\$esp]$
Execute	valE	valE ← valB + -4
	Cond code	
Memory	valM	$M_4[valE] \leftarrow valP$
Write	dstE	R[%esp] ← valE
back	dstM	
PC update	PC	PC ← valC

Read instruction byte [Read register byte] Read constant word Compute next PC [Read operand A] Read operand B Perform ALU operation [Set condition code reg.] [Memory read/write] [Write back ALU result] Write back memory result **Update PC**

- All instructions follow same general pattern
- Differ in what gets computed on each step

Computed Values

■Fetch

icode Instruction code

ifun Instruction function

rA Instr. Register A

rB Instr. Register B

valC Instruction constant

valP Incremented PC

■Decode

srcA Register ID A

srcB Register ID B

dstE Destination Register E

dstM Destination Register M

valA Register value A

valB Register value B

■Execute

valE ALU result

Cnd Branch/move flag

■Memory

valM Value from memory

Sequential Hardware Implementation

Key

- Blue boxes: predesigned hardware blocks
 - E.g., memories, ALU
- Gray boxes: control logic
 - Describe in HCL
- White ovals: labels for signals
- Thick lines: 32-bit word values
- Thin lines: 4-8 bit values
- Dotted lines: 1-bit values

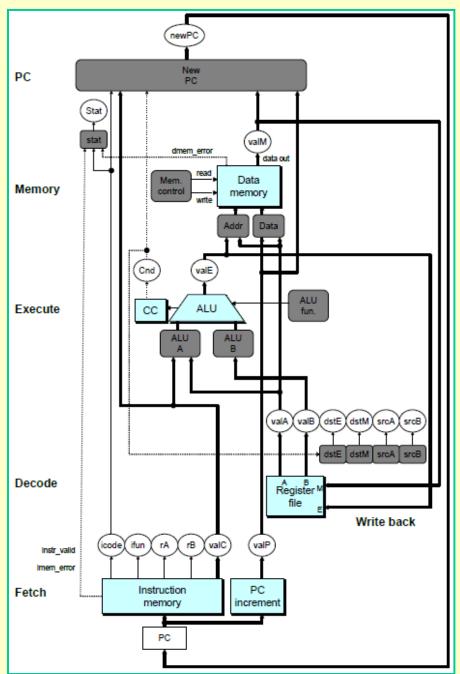
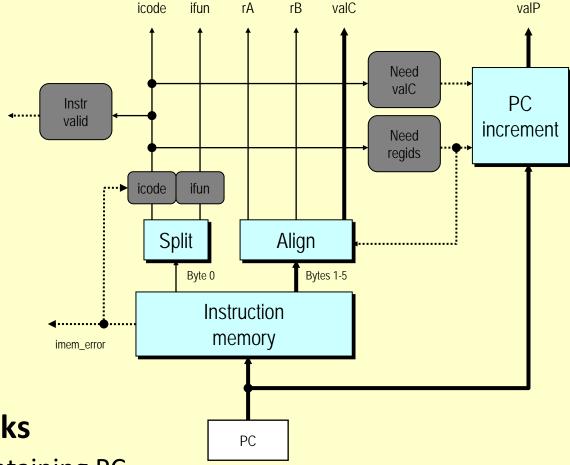


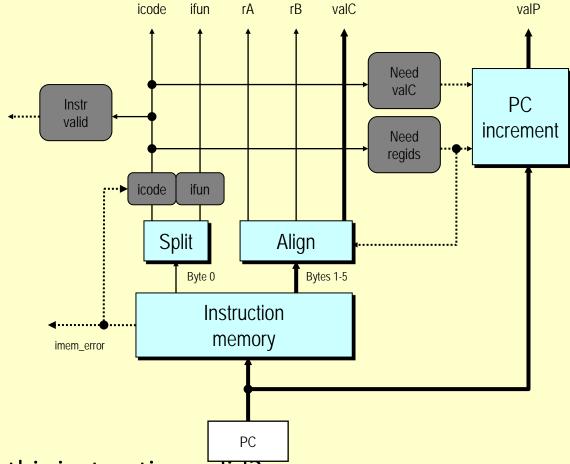
Figure 4.23, p. 378

Fetch Logic



- Predefined Blocks
 - PC: Register containing PC
 - Instruction memory: Read 6 bytes (PC to PC+5)
 - Signal invalid address
 - Split: Divide instruction byte into icode and ifun
 - Align: Get fields for rA, rB, and valC

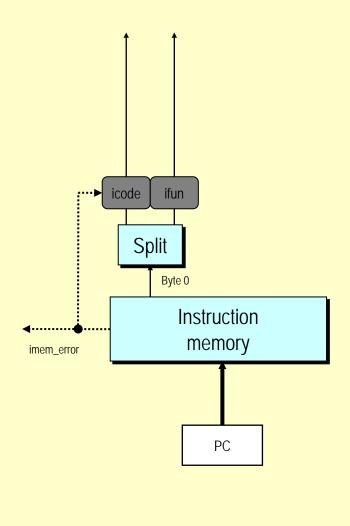
Fetch Logic

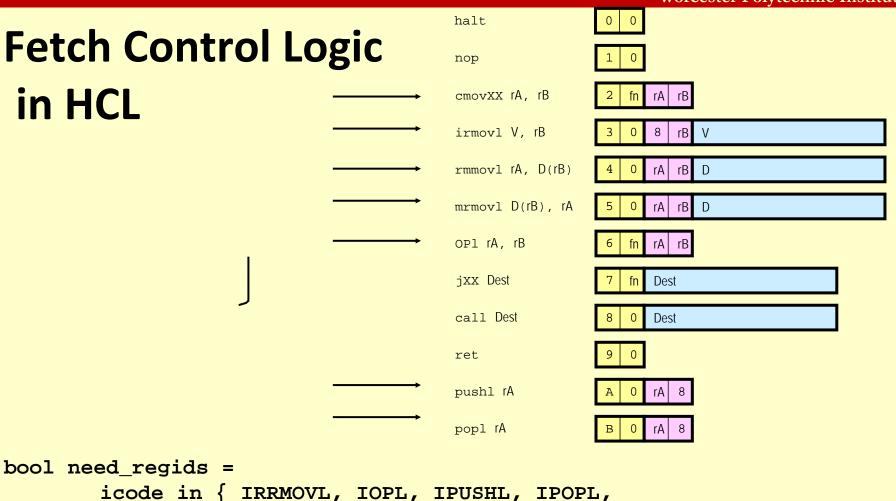


Control Logic

- Instr. Valid: Is this instruction valid?
- icode, ifun: Generate no-op if invalid address
- Need regids: Does this instruction have a register byte?
- Need valC: Does this instruction have a constant word?

Fetch Control Logic in HCL





Decode Logic

Register File

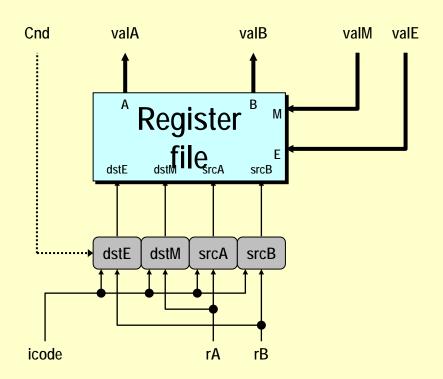
- Read ports A, B
- Write ports E, M
- Addresses are register IDs or 15 (0xF) (no access)

Control Logic

- srcA, srcB: read port addresses
- dstE, dstM: write port addresses

Signals

- Cnd: Indicate whether or not to perform conditional move
 - Computed in Execute stage



A Source

```
OPI rA, rB
            Decode
                          valA \leftarrow R[rA]
                                                             Read operand A
                           cmovXX rA, rB
             Decode
                           valA \leftarrow R[rA]
                                                             Read operand A
                           rmmov1 rA, D(rB)
             Decode
                                                             Read operand A
                           valA \leftarrow R[rA]
                           popl rA
             Decode
                           valA \leftarrow R[\$esp]
                                                             Read stack pointer
                          jXX Dest
             Decode
                                                             No operand
                           call Dest
             Decode
                                                             No operand
                          ret
                           valA \leftarrow R[\$esp]
             Decode
                                                             Read stack pointer
icode in { IRRMOVL, IRMMOVL, IOPL, IPUSHL
                                                            } : rA;
icode in { IPOPL, IRET } : RESP;
1 : RNONE; # Don't need register
```

];

int srcA = [

Worcester Polytechnic Institute

E Destination

```
OPI rA, rB
Write-back
                R[rB] \leftarrow valE
                                                        Write back result
                cmovXX rA, rB
                                                        Conditionally write back
Write-back
                R[rB] \leftarrow valE
                                                        result
                rmmov1 rA, D(rB)
Write-back
                                                        None
                popl rA
Write-back
                R[\$esp] \leftarrow valE
                                                        Update stack pointer
                jXX Dest
Write-back
                                                        None
                call Dest
Write-back
                R[\$esp] \leftarrow valE
                                                        Update stack pointer
                ret
Write-back
                R[\$esp] \leftarrow valE
                                                        Update stack pointer
```

```
int dstE = [
```

```
icode in { IRRMOVL } && Cnd : rB;
icode in { IIRMOVL, IOPL} : rB;
icode in { IPUSHL, IPOPL, ICALL, IRET } : RESP;
1 : RNONE; # Don't write any register
```

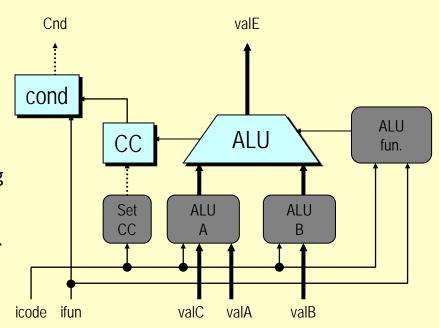
Execute Logic

Units

- ALU
 - Implements 4 required functions
 - Generates condition code values
- CC
 - Register with 3 condition code bits
- cond
 - Computes conditional jump/move flag

Control Logic

- Set CC: Should condition code register be loaded?
- ALU A: Input A to ALU
- ALU B: Input B to ALU
- ALU fun: What function should ALU compute?



ALU A Input

```
OPI rA, rB
Execute
                                                    Perform ALU operation
               valE \leftarrow valB OP valA
               cmovXX rA, rB
Execute
               valE \leftarrow 0 + valA
                                                    Pass valA through ALU
               rmmovl rA, D(rB)
Execute
               valE ← valB + valC
                                                    Compute effective address
               popl rA
Execute
               valE ← valB + 4
                                                    Increment stack pointer
               iXX Dest
Execute
                                                    No operation
               call Dest
Execute
               valE \leftarrow valB + -4
                                                    Decrement stack pointer
               ret
Execute
               valE ← valB + 4
                                                    Increment stack pointer
```

```
int aluA = [
    icode in { IRRMOVL, IOPL } : valA;
    icode in { IIRMOVL, IRMMOVL, IMRMOVL } : valC;
    icode in { ICALL, IPUSHL } : -4;
    icode in { IRET, IPOPL } : 4;
    # Other instructions don't need ALU
```

ALU Operation

```
OPI rA, rB
 Execute
                                                     Perform ALU operation
                valE \leftarrow valB OP valA
                 cmovXX rA, rB
 Execute
                 valE \leftarrow 0 + valA
                                                      Pass valA through ALU
                 rmmov1 rA, D(rB)
 Execute
                 valE ← valB + valC
                                                      Compute effective address
                popl rA
 Execute
                 valE \leftarrow valB + 4
                                                      Increment stack pointer
                jXX Dest
                                                      No operation
 Execute
                 call Dest
 Execute
                                                      Decrement stack pointer
                 valE \leftarrow valB + -4
                ret
 Execute
                 valE \leftarrow valB + 4
                                                      Increment stack pointer
int alufun = [
           icode == IOPL : ifun;
           1 : ALUADD;
```

CS-2011, D-Term 2014

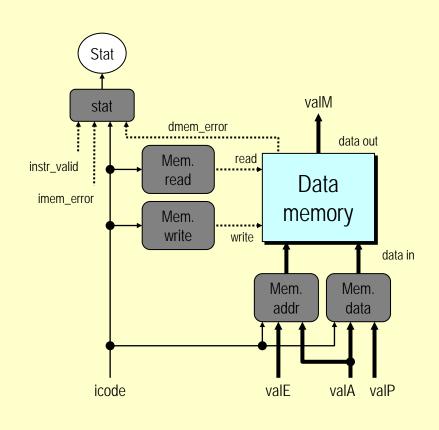
Memory Logic

Memory

Reads or writes memory word

Control Logic

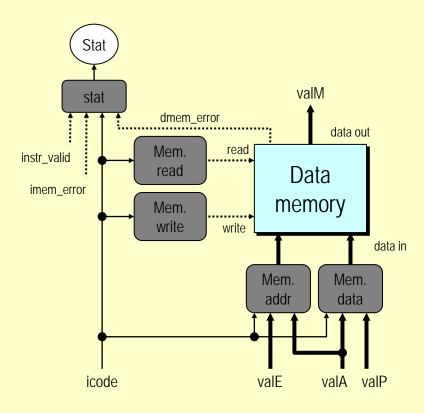
- stat: What is instruction status?
- Mem. read: should word be read?
- Mem. write: should word be written?
- Mem. addr.: Select address
- Mem. data.: Select data



Instruction Status

Control Logic

stat: What is instruction status?



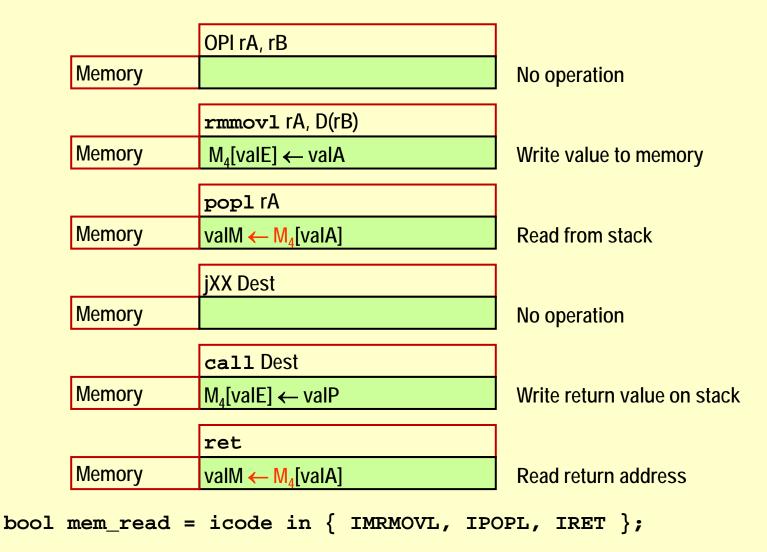
```
## Determine instruction status
int Stat = [
        imem_error || dmem_error : SADR;
        !instr_valid: SINS;
        icode == IHALT : SHLT;
        1 : SAOK;
];
```

Memory Address

	OPI rA, rB	
Memory		No operation
	rmmov1 rA, D(rB)	
Memory	$M_4[valE] \leftarrow valA$	Write value to memory
	popl rA	
Memory	$valM \leftarrow M_4[valA]$	Read from stack
	jXX Dest	
Memory		No operation
	call Dest	
Memory	$M_4[valE] \leftarrow valP$	Write return value on stack
	ret	
Memory	$valM \leftarrow M_4[valA]$	Read return address
	·	

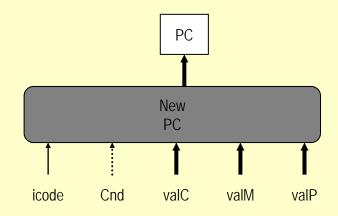
```
int mem_addr = [
    icode in { IRMMOVL, IPUSHL, ICALL, IMRMOVL } : valE;
    icode in { IPOPL, IRET } : valA;
    # Other instructions don't need address
```

Memory Read



PC Update Logic

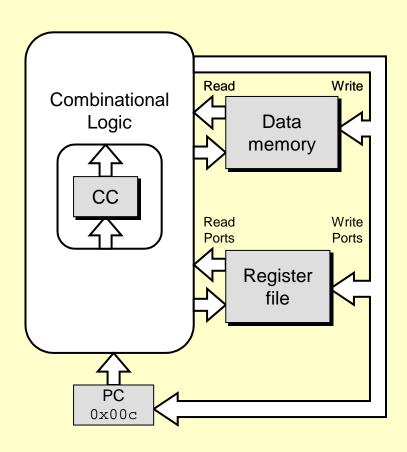
- New PC
 - Select next value of PC



PC Update

```
OPI rA, rB
PC update
                                                  Update PC
              PC \leftarrow valP
              rmmov1 rA, D(rB)
PC update
              PC \leftarrow valP
                                                  Update PC
              popl rA
PC update
              PC \leftarrow valP
                                                  Update PC
              jXX Dest
PC update
              PC \leftarrow Cnd ? valC : valP
                                                  Update PC
              call Dest
PC update
              PC \leftarrow valC
                                                  Set PC to destination
              ret
PC update
              PC \leftarrow valM
                                                  Set PC to return address
int new pc = [
           icode == ICALL : valC;
           icode == IJXX && Cnd : valC;
           icode == IRET : valM;
           1 : valP;
];
```

Sequential Operation



State

- PC register
- Cond. Code register
- Data memory
- Register file

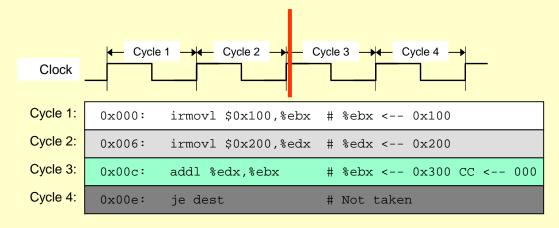
All updated as clock rises

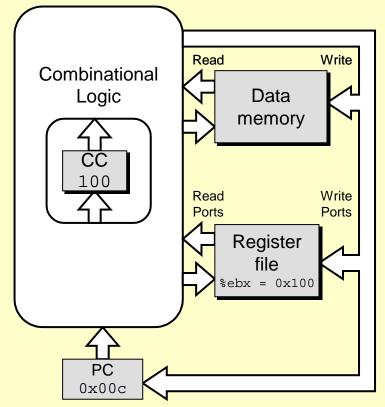
Combinational Logic

- ALU
- Control logic
- Memory reads
 - Instruction memory
 - Register file

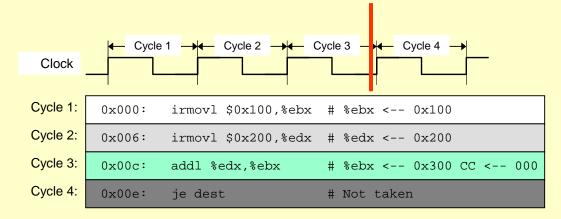


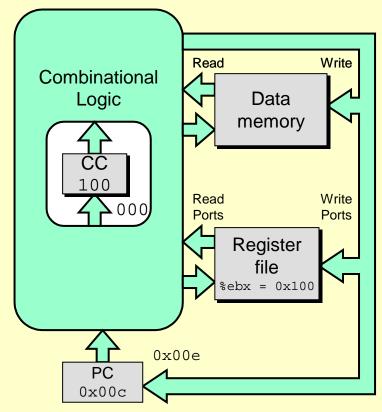




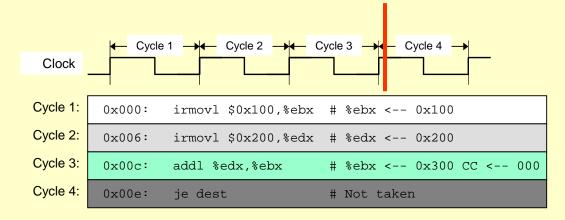


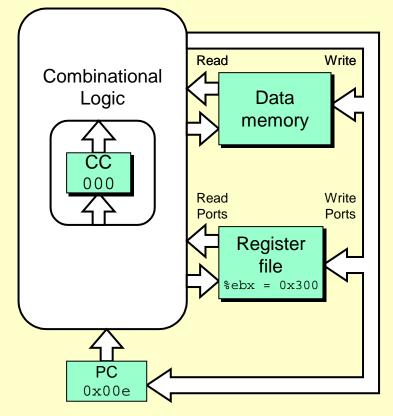
- state set according to second irmovl instruction
- combinational logic starting to react to state changes



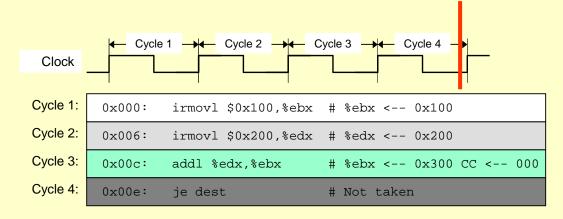


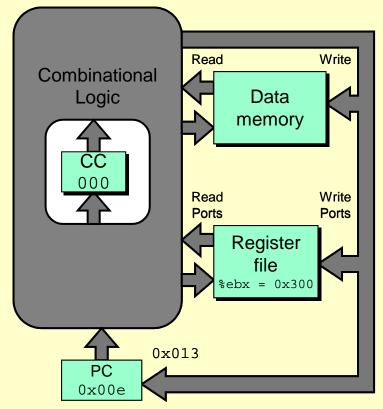
- state set according to second irmovl instruction
- combinational logic generates results for addl instruction





- state set according to addl instruction
- combinational logic starting to react to state changes





- state set according to addl instruction
- combinational logic generates results for je instruction

SEQ Summary

Implementation

- Express every instruction as series of simple steps
- Follow same general flow for each instruction type
- Assemble registers, memories, predesigned combinational blocks
- Connect with control logic

Limitations

- Too slow to be practical
- In one cycle, must propagate through instruction memory, register file, ALU, and data memory
- Would need to run clock very slowly
- Hardware units only active for fraction of clock cycle

Questions?