

# CS3410: Computer Systems and Organization

LEC11: RISC-V Data Transfer + Control Flow

Professor Giulia Guidi Wednesday, October 1, 2025



### Plan for Today

- Review: RISC-V Data Transfer Vol. I
- RISC-V Data Transfer Vol. II
- RISC-V Control Flow or Decision Making

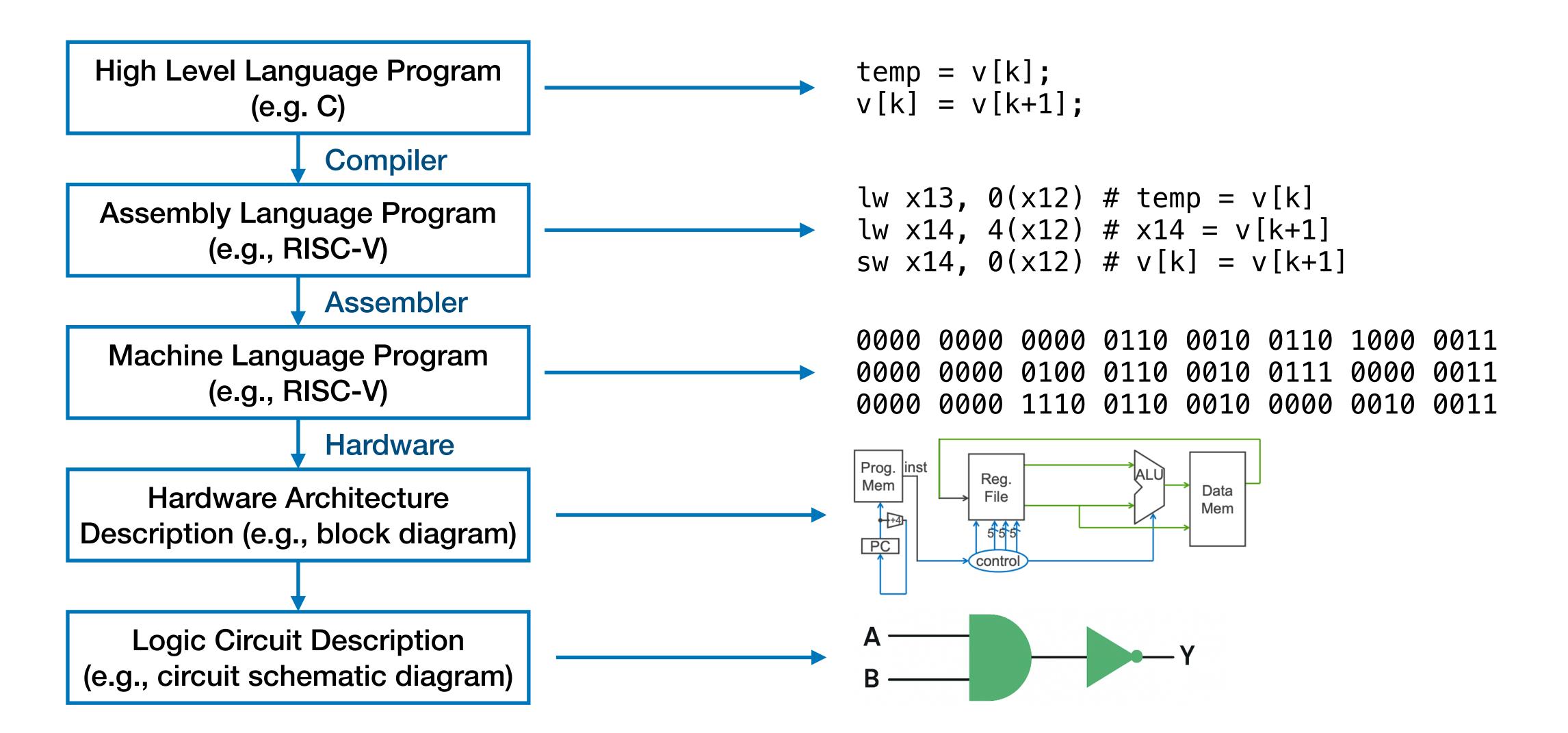
PSA: Prelim & Prelim Survey [due **Friday 10/3**]
Please go to **OH** or post on **Ed** 



### Review of RISC-V data transfer, so far

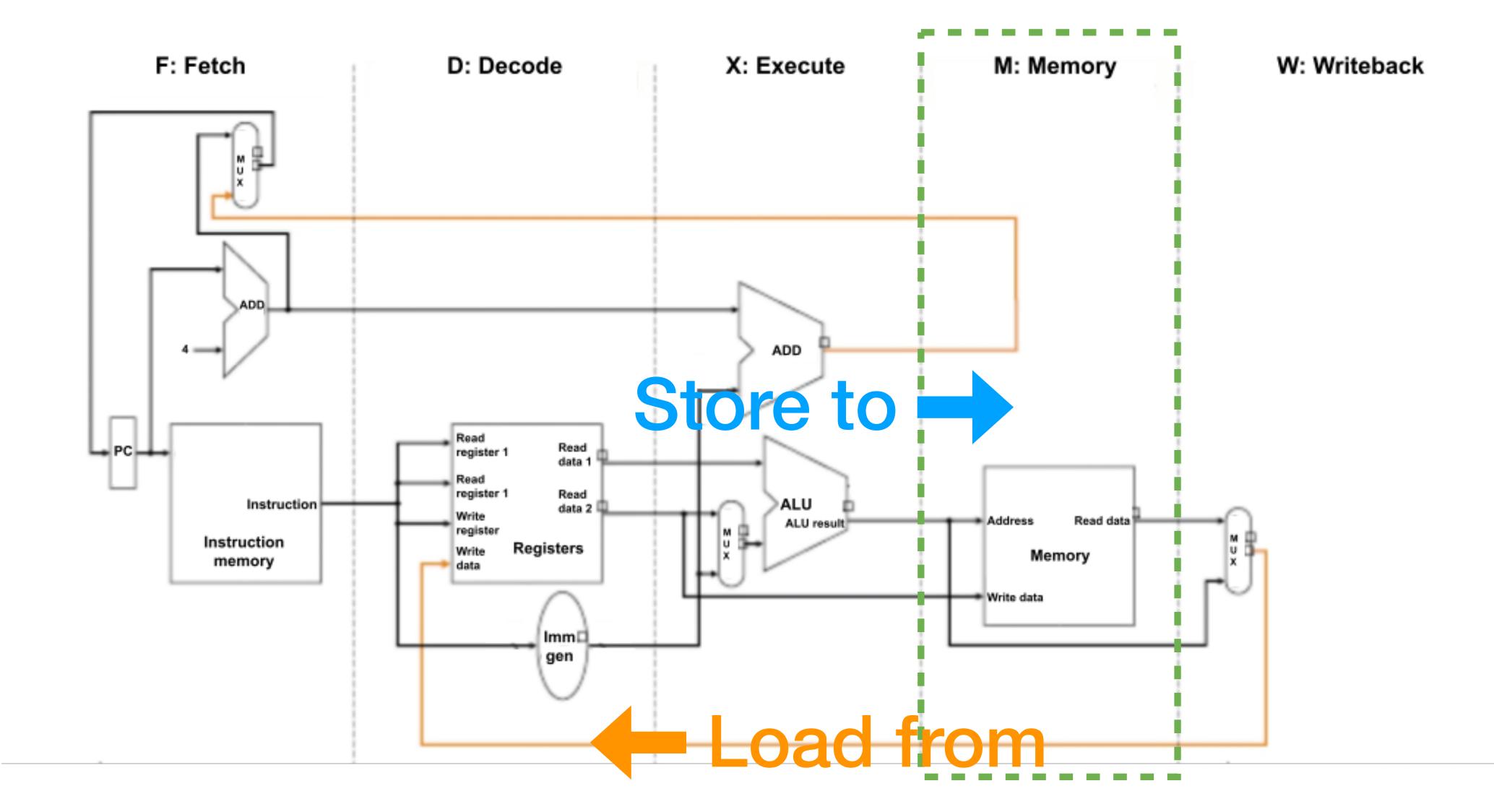


### RISC-V Assembly





### Data Transfer: Load from and Store to Memory



### Load from Memory to Register

#### C code: address value (hex) int A[100]; g = h + A[3];1000 0x03 (LSB) 1 word = 4 bytes1001 0x87 Using "load word" (lw) in RISC-V: 1002 0xC7 (MSB) <- 1 word (bytes 1000-1003) 1003 000lw x10, 12(x15) # Reg x10 gets 1004 (LSB) 0x33 add x11, x12, x10 # g = h + A[3]1005 0x06 1006 0xA6 1007 00x0 (MSB) < -1 word (bytes 1004-1007)

→ lw updates x10 in the register file



### Load from Memory to Register

In this example, we assumed we knew x15 was the base address of A [0] Code (or compiler) must have loaded the base address into x15

```
1 \text{ word} = 4 \text{ bytes}
```

Using "load word" (lw) in RISC-V:

```
la x15, A \# x15 = &A[0] (address of first element)
lw x10, 12(x15) \# Reg x10 gets A[3] x15: address in memory (pointer to A[0])
add x11, x12, x10 \# g = h + A[3] 12: offset in bytes but we load one word at a time
```

→ lw updates x10 in the register file



# Load from Memory to Register

In this example, we assumed we knew x15 was the base address of A [0] Code (or compiler) must have loaded the base address into x15

```
1 word = 4 bytes
Using "load word" (\lambda w) in RISC-V:
```

```
la rd, imm \rightarrow lui rd, imm # put 20-bit imm into the top 20 bit of rd addi rd, rs1, imm # put 12-bit imm into the low 12 bit of rd
```



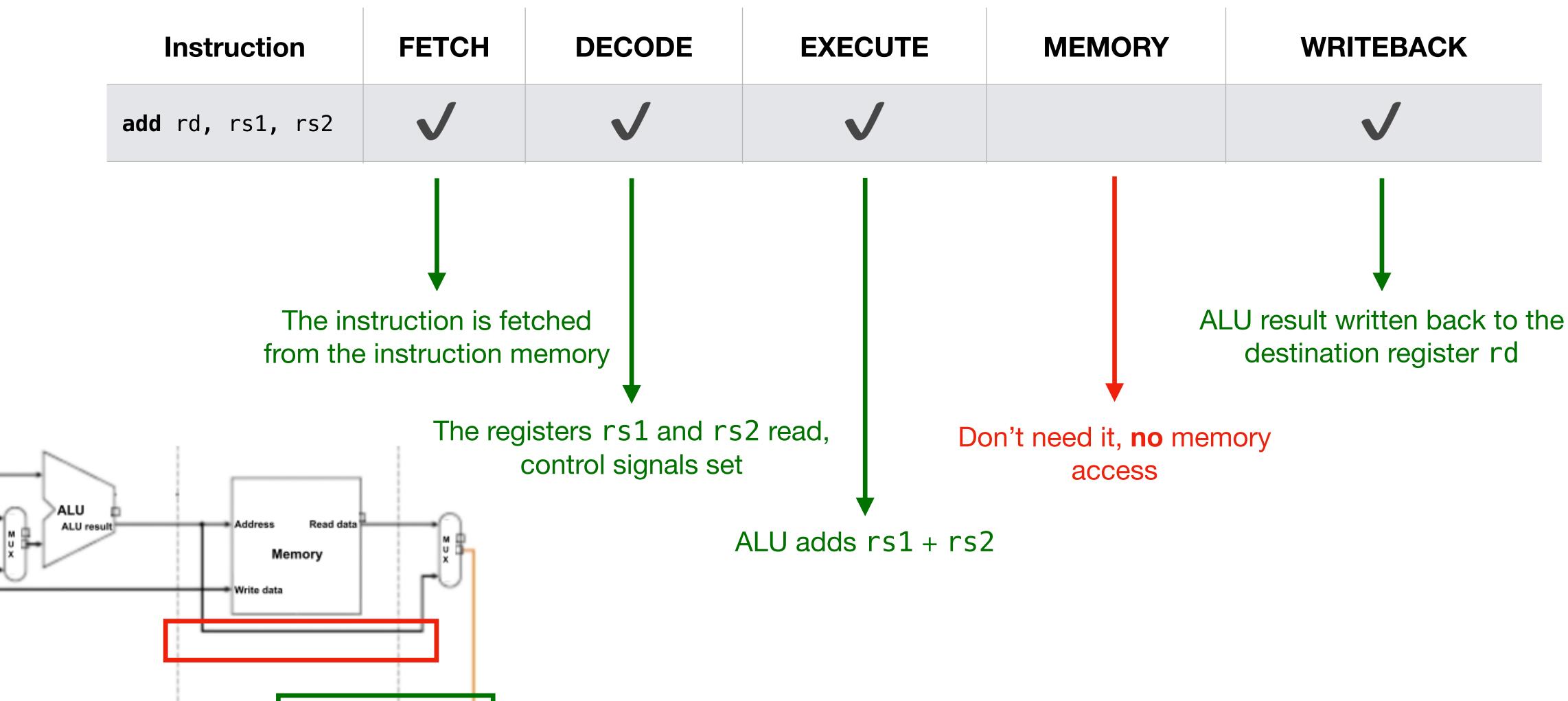
### Store from Register to Memory

```
lw x10, 12(x15) # Temp reg x10 gets A[3] x15 + 12 add x11, x12, x10 # Temp reg x11 gets h + A[3] Offset must be a multiple of 4 sw x11, 40(x15) # A[10] = h + A[3] x15 + 40
```





### CPU 5 Stages in RISC-V





### Poll

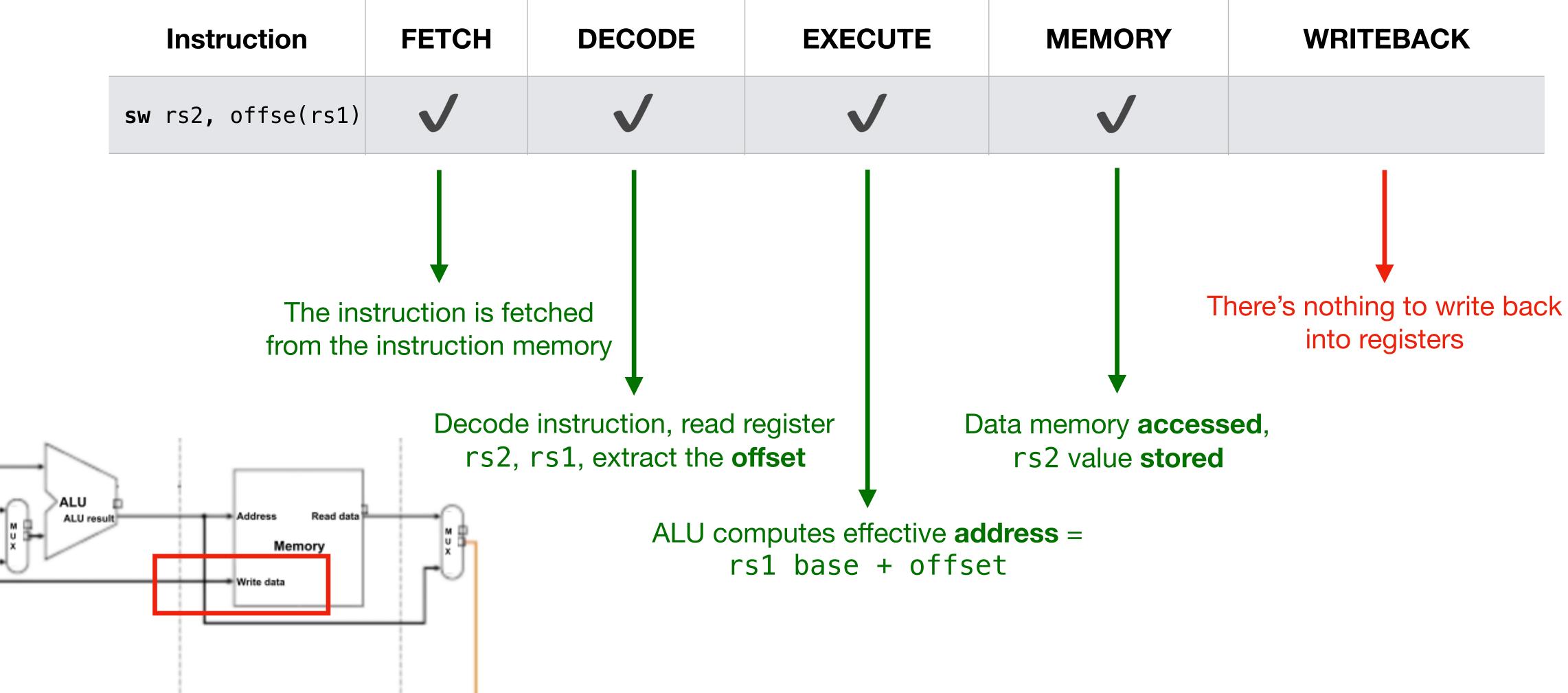
Q: Please select the CPU stage **not** involved in the operation: sw x10, 36(x5)



PollEv.com/gguidi
Or send gguidi to 22333

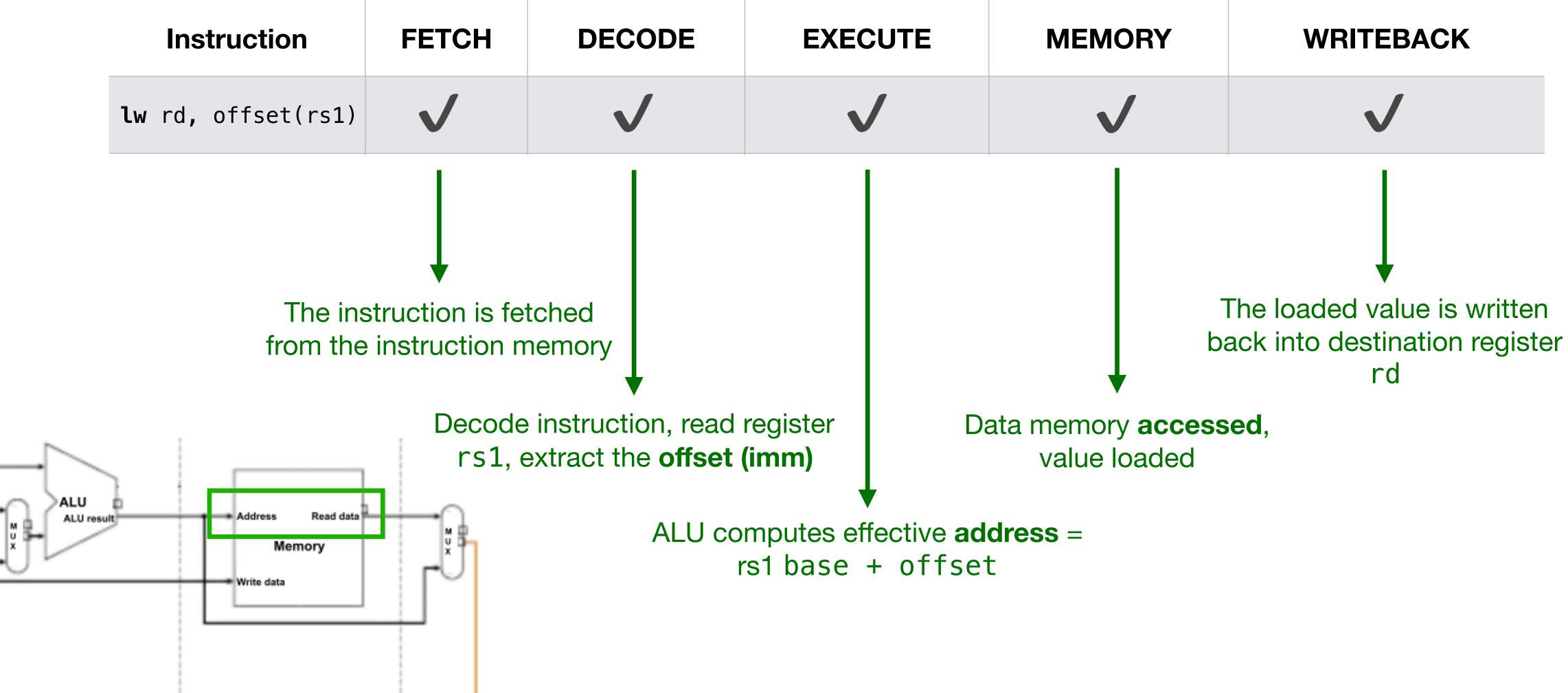


### CPU 5 Stages in RISC-V





### CPU 5 Stages in RISC-V





Ok, back to data transfer — Vol. II



"load word" "store word"

"load byte"

In addition to lw and sw, RISC-V has lb and sb "store byte"



"load word" "store word"

"load byte"

In addition to lw and sw, RISC-V has lb and sb "store byte"

Uses same format as lw and sw:

• E.g., lb x10, 3(x11) pointer to memory

offset in bytes (doesn't have to be a multiple of 4 because I'm only loading 1 byte, i.e., one line in memory)

- 1. Compute the effective address = content of x11 + 3
  - Let us assume 3 (x11) contains the value 4 (decimal) and we use 16-bit register
  - 4 (decimal) = 0000 0000 0000 0100 (16-bit binary)



"load word" "store word"

"load byte"

In addition to lw and sw, RISC-V has lb and sb "store byte"

Uses same format as lw and sw:

• E.g., lb x10, 3(x11) pointer to memory

offset in bytes (doesn't have to be a multiple of 4 because I'm only loading 1 byte, i.e., one line in memory)

- 1. Compute the effective address = content of x11 + 3
  - Let us assume 3 (x11) contains the value 4 (decimal) and we use 16-bit register
  - 4 (decimal) = 0000 0000 0000 0100 (16-bit binary)
- 2. Then, **load 1 byte** from memory at that address
  - The loaded **byte** is 0000 0100 (8-bit binary)



"load word" "store word"

"load byte"

In addition to lw and sw, RISC-V has lb and sb "store byte"

Uses same format as lw and sw:

• E.g., lb x10, 3(x11) pointer to memory

offset in bytes (doesn't have to be a multiple of 4 because I'm only loading 1 byte, i.e., one line in memory)

- 2. Then, **load 1 byte** from memory at that address
  - The loaded **byte** is 0000 0100 (8-bit binary)
- 3. Finally, sign-extend the byte to 16 bits (we assumed x10 is a 16-bit register)
  - The final value in x10 is 0000 0000 0000 0100 (16-bit binary) = 4 (decimal)



"load word" "store word"

"load byte"

In addition to lw and sw, RISC-V has lb and sb "store byte"

Uses same format as lw and sw:

• E.g., lb x10, 3(x11) pointer to memory

offset in bytes (doesn't have to be multiple of 4)

- 1. Compute the effective address = content of x11 + 3
  - Ok but what if 3(x11) contains the value 3410 (decimal) and we use 16-bit register
  - 3410 (decimal) = 0000 1101 0101 0010 (16-bit binary)
- 2. Then, **load 1 byte** from memory at that address
  - The loaded **byte** is 0101 0010 (8-bit binary)



"load word" "store word"

"load byte"

In addition to lw and sw, RISC-V has lb and sb "store byte"

Uses same format as lw and sw:

• E.g., lb x10, 3(x11) pointer to memory

offset in bytes (doesn't have to be multiple of 4)

- 2. Then, **load 1 byte** from memory at that address
  - The loaded byte is 0101 0010 (8-bit binary)
- 3. Finally, sign-extend the byte to 16 bits (we assumed x10 is a 16-bit register)
  - The final value in x10 is 0000 0000 0101 0010 (16-bit binary)
  - Ops! 0000 0000 0101 0010 (16-bit binary) = 82 (decimal) != 3410 (decimal)



### lbu

"load word" "store word"

"load byte"

In addition to lw and sw, RISC-V has lb and sb "store byte"

Uses same format as lw and sw:

• E.g., lb x10, 3(x11) pointer to memory

offset in bytes (doesn't have to be multiple of 4)

lbu = unsigned load byte It doesn't need to preserve the sign



### lbu

"load word" "store word"

"load byte"

In addition to lw and sw, RISC-V has lb and sb "store byte"

Uses same format as lw and sw:

• E.g., lb x10, 3(x11) pointer to memory

offset in bytes (doesn't have to be multiple of 4)

lbu = unsigned load byte

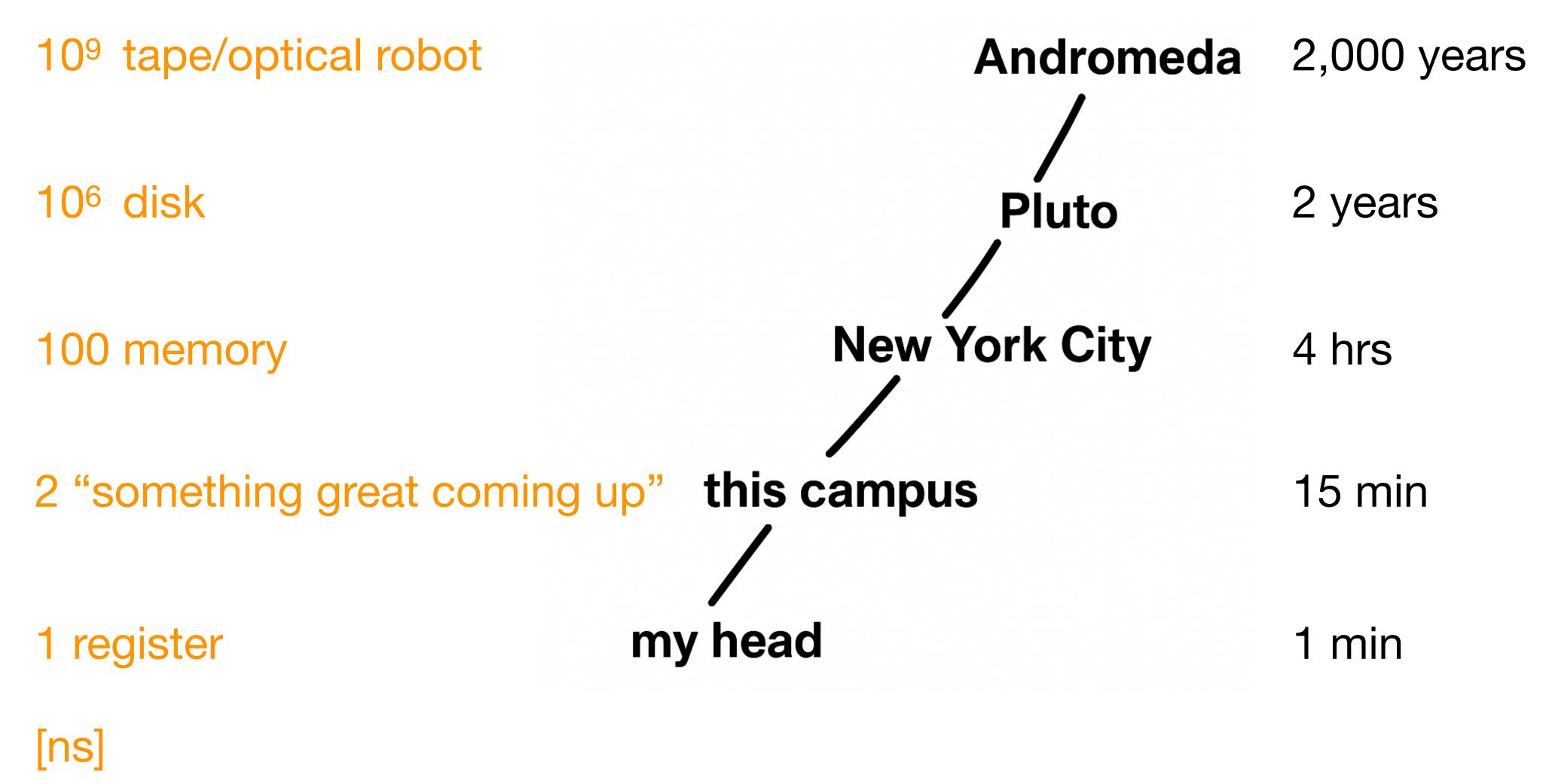
It doesn't need to preserve the sign: zero extension

But no sbu, why?

It doesn't matter! You're just writing the low 8 bits of a register directly to memory, so **no** extension



# Register versus Memory





#### addi

#### The following two instructions:

```
lw x10, 12(x15) # temp reg x10 gets A[3] add x12, x12, x10 # reg x12 = reg x12 + A[3]
```

#### Replace addi:

```
addi x12, value # put value in A[3]
```

This involves going to New York City (load from memory)

The add immediate is so common that is deserves its own instruction



### Poll

```
addi x11, x0, 0x49C
sw x11, 0(x5)
lb x12, 0(x5)
```

**Reminder:** 1 word = 4 bytes

**Reminder:** 1 hex digit = 4-bit pattern

Let us assume 32-bit register



PollEv.com/gguidi
Or send gguidi to 22333



```
addi x11, x0, 0x49C \longrightarrow addi = add immidiate x0 is always 0! So this just loads 0x49C (hex) into x11
```

**Reminder:** 1 word = 4 bytes

**Reminder:** 1 hex digit = 4-bit pattern

Let us assume 32-bit register



**Reminder:** 1 word = 4 bytes

**Reminder:** 1 hex digit = 4-bit pattern

Let us assume 32-bit register

So this just loads 0x49C (hex) into x11

sw = store word (32-bit)

It takes the value in x11 and writes 1 word to memory starting at address in x5

In memory at x5, the 32-bit value 0x0000049C is stored

→ SW does **not** update any register in the register file; only memory is updated



addi x11, x0, 0x49C sw x11, 0(x5) tb x12, 0(x5)

**Reminder:** 1 word = 4 bytes

**Reminder:** 1 hex digit = 4-bit pattern

Let us assume 32-bit register

So this just loads 0x49C (hex) into x11

sw = store word (32-bit)

It takes the value in x11 and writes 1 word to memory starting at address in x5

In memory at x5, the 32-bit value 0x000049C is stored

memory address	(little endian)	value
----------------	-----------------	-------



**Reminder:** 1 word = 4 bytes

**Reminder:** 1 hex digit = 4-bit pattern

Let us assume 32-bit register

→ lb updates x12 in the register file

So this just loads 0x0000049C (hex) into x11 x5 stores 0x0000049C

lb = load byte (8-bit), sign-extend to 32-bit

This loads the **first byte** from memory at x5

$$x5 + 0 = 0 \times 9C$$

nemory address (little endian)	value
x5 + 0	0x9C
x5 + 1	0x04
x5 + 2	0x00
x5 + 3	0x00



**Reminder:** 1 word = 4 bytes

**Reminder:** 1 hex digit = 4-bit pattern

Let us assume 32-bit register

So this just loads 0x0000049C (hex) into x11 x5 stores 0x0000049C

lb = load byte (8-bit), sign-extend to 32-bit

This loads the **first byte** from memory at x5

$$x5 + 0 = 0x9C$$
 $0x9C_{16} = 10011100_2 = -100_{10}$ 

MSB is 1 → negative!



**Reminder:** 1 word = 4 bytes

**Reminder:** 1 hex digit = 4-bit pattern

Let us assume 32-bit register

#### Recall

Ok, what about hex? E.g., -1 = FFFF

If  $MSB = 0-7 \Rightarrow then append 0s$ 

Otherwise append Fs

So this just loads 0x0000049C (hex) into x11

x5 stores 0x0000049C

b = load byte (8-bit), sign-extend to 32-bit

This loads the **first byte** from memory at x5

$$x5 + 0 = 0x9C$$

$$0 \times 9C_{16} = 10011100_2 = -100_{10}$$

 $x12 \text{ stores } -100_{10} = 0xFFFFFF9C_{16}$ 



#### Data Transfer Conclusion

- The memory is byte-addressable, but lw and sw access one word at a time
- A pointer (used by lw and sw) is just a memory address, we can add to it or subtract from it (using offset)
- Partial memory hierarchy (register versus DRAM)
- Bit sign-extension and zero-extension
- Review of endianess



# Ok, let's move on to RISC-V control flow



# Computer Decisions Making

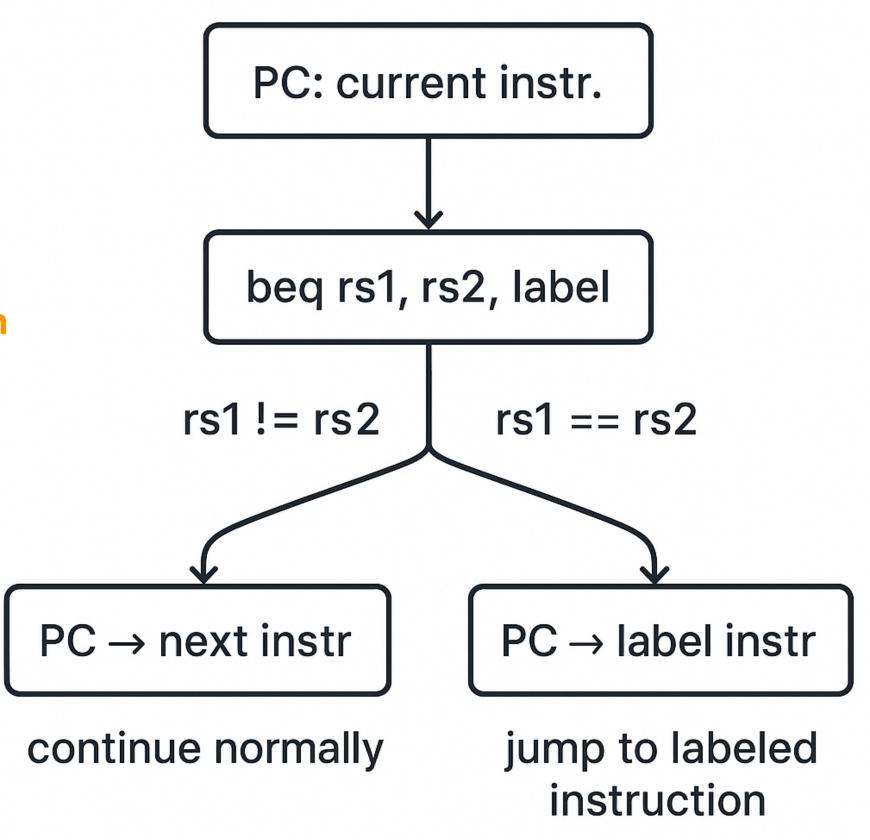
I.e., based on computation, do something different

• In programming languages, if-statement

RISC-V if-statement instruction:

"branch if equal"

 $\rightarrow$  If rs1 == rs2, then go to the instruction at label





### Computer Decisions Making

I.e., based on computation, do something different

• In programming languages, if-statement

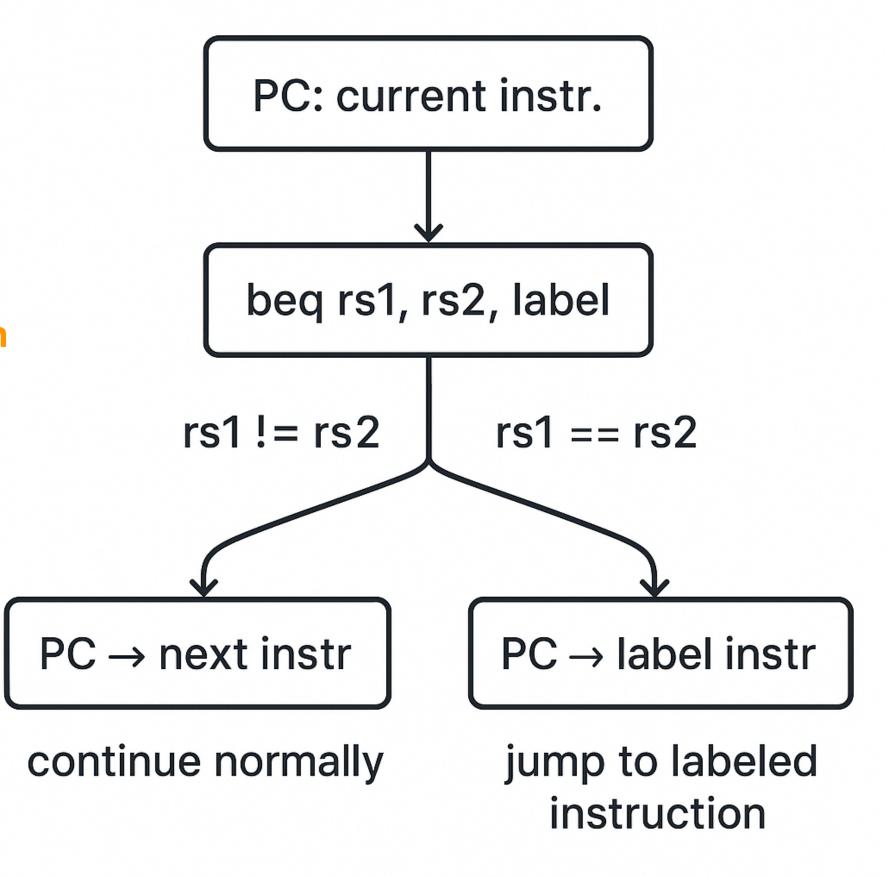
RISC-V if-statement instruction:

"branch if equal"

 $\rightarrow$  If rs1 == rs2, then go to the instruction at label

"branch if **not** equal"

→ If rs1!= rs2, then go to the instruction at label





### Types of Branches

- Branch: change of control flow
- Conditional branch: change control flow depending on outcome of comparison

```
beq, bne
blt, bge "branch if less than", "branch if greater or equal than"

bltu, bgeu "branch if less than" unsigned, "branch if greater than or equal to" unsigned
```

The result of x1 < x2 changes if we use blt (true) vs bltu (false)

```
x1 = 1111_2 \rightarrow 15_{10} if unsigned and -1_{10} if signed x2 = 0000_2 \rightarrow 0_{10}
```



#### Types of Branches

- Branch: change of control flow
- Conditional branch: change control flow depending on outcome of comparison

```
beq, bne
blt, bge "branch if less than", "branch if greater or equal than"

bltu, bgeu "branch if less than" unsigned, "branch if greater than or equal to" unsigned
```

- → beq, bne, blt, etc. do not update any register in the register file
- → Branches only affect program counter (PC)



#### Types of Branches

- Branch: change of control flow
- Conditional branch: change control flow depending on outcome of comparison

```
beq, bne
blt, bge "branch if less than", "branch if greater or equal than"

bltu, bgeu "branch if less than" unsigned, "branch if greater than or equal to" unsigned
```

• Unconditional branch: always branch



## beq

```
beq rs1, rs2, label —— It marks the address of an instruction
```

```
address instruction

-----

1000 beq x10, x12, mahler → Points to 1012 <

1004 sub x5, x5, x1

1008 srli x5, x5, 0x01

1012 add x1, x2, x3 (label) mahler

↓

1 line is 1 instruction, 1 line is 1 word, 1 line "includes" 4 bytes
```

The assembler computes the branch offset:

```
offset = label_address - PC = 0 \times 1012 - 0 \times 1000 = 12
```

...and encodes it in the instruction itself as an imm

In machine code, the instruction no longer says mahler — it has a fixed signed offset (12)

If the **branch is taken**, we **skip** the instruction at  $0 \times 1004$  and the instruction at  $0 \times 1008$  and jump directly to mahler, i.e.,  $0 \times 1012$ 



Instruction	FETCH	DECODE	EXECUTE	MEMORY	WRITEBACK
beq x10, x12, mahler					

address	instruction
1000	beq x10, x12, mahler
1004	sub x5, x5, x1
1008	srli x5, x5, 0x01
1012	add x1, x2, x3 (label) mahler



Instruction	FETCH	DECODE	EXECUTE	MEMORY	WRITEBACK
beq x10, x12, mahler					

Fetch beg at  $0 \times 1000$ ; compute PC + 4 =  $0 \times 1004$  — "next instruction" if branch not taken

```
address instruction
-----

1000 beq x10, x12, mahler

1004 sub x5, x5, x1

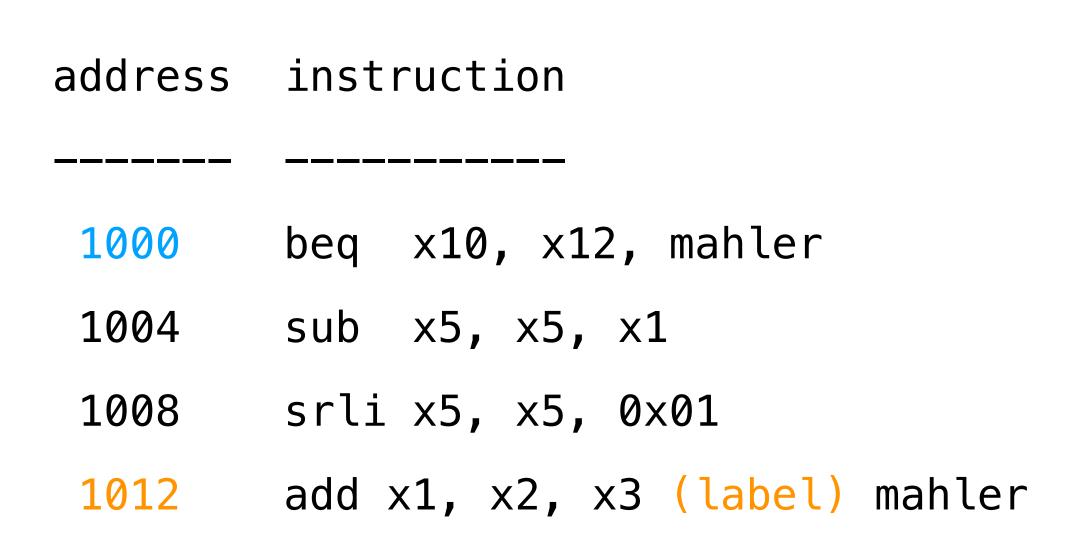
1008 srli x5, x5, 0x01

1012 add x1, x2, x3 (label) mahler
```



Instruction	FETCH	DECODE	EXECUTE	MEMORY	WRITEBACK
beq x10, x12, mahler					

Decode beq, read x10 and x12; sign-extend branch offset = 12



In machine code, the instruction no longer says mahler It has a fixed signed offset (12)



Instruction	FETCH	DECODE	EXECUTE	MEMORY	WRITEBACK
beq x10, x12, mahler					

Compare x10 == x12 and compute branch\_target = PC + offset = 0x1000 + 12 = 0x1012

```
address instruction
-----

1000 beq x10, x12, mahler

1004 sub x5, x5, x1

1008 srli x5, x5, 0x01

1012 add x1, x2, x3 (label) mahler
```



				•	
Instruction	FETCH	DECODE	EXECUTE	MEMORY	WRITEBACK
beq x10, x12, mahler					

If x10 == x12, we update the PC and jump directly to mahler If  $x10 \neq x12$ , we go to the next instruction, i.e., 0x1004

```
address instruction
-----
1000 beq x10, x12, mahler
1004 sub x5, x5, x1
1008 srli x5, x5, 0x01
1012 add x1, x2, x3 (label) mahler
```



Instruction	FETCH	DECODE	EXECUTE	MEMORY	WRITEBACK
beq x10, x12, mahler					

beq doesn't read nor write to data memory

```
address instruction
-----

1000 beq x10, x12, mahler

1004 sub x5, x5, x1

1008 srli x5, x5, 0x01

1012 add x1, x2, x3 (label) mahler
```



Instruction	FETCH	DECODE	EXECUTE	MEMORY	WRITEBACK
beq x10, x12, mahler					

beq doesn't write any register; only PC update occurs if branch taken

```
address instruction
-----

1000 beq x10, x12, mahler

1004 sub x5, x5, x1

1008 srli x5, x5, 0x01

1012 add x1, x2, x3 (label) mahler
```



#### Example if Statement

Let us assume the translations below, compile if block

```
f \rightarrow x10, g \rightarrow x11, h \rightarrow x12, i \rightarrow x13, j \rightarrow x14

if (i == j)

f = g + h;

bne x13, x14, Exit

add x10, x11, x12

Exit: % terminate
```



#### Example if-else Statement

Let us assume the translations below, compile if-else block



#### 3 types of loops:

- while
- do ... while
- for

Each can be rewritten as either of the other two!

```
int main() {
    int i = 0;
    while (i < 10) {
        printf("%d\n", i);
        i++;
        return 0;
}
    return 0;
}</pre>
```

```
int main() {
    int i = 0;
    do {
        printf("%d\n", i);
        i++;
    } while (i < 10);
    return 0;
}</pre>
```



```
int A[20];
// fill A with data
int sum = 0;
for (int i = 0; i < 20; i++)
    sum += A[i];</pre>
```



```
int A[20]; add \times 9, \times 8, \times 0 # \times 9 = &A[0] // fill A with data int sum = 0; for (int i = 0; i < 20; i++) sum += A[i];
```



```
int A[20];
// fill A with data
int sum = 0;
for (int i = 0; i < 20; i++)
    sum += A[i];</pre>
```

```
add x9, x8, x0 # x9 = &A[0] add x10, x0, x0 # sum = 0
```



```
int A[20];
// fill A with data
int sum = 0;
for (int i = 0; i < 20; i++)
    sum += A[i];</pre>
```

```
add x9, x8, x0 # x9 = &A[0] add x10, x0, x0 # sum = 0 add x11, x0, x0 # i = 0
```



```
int A[20];
// fill A with data
int sum = 0;
for (int i = 0; i < 20; i++)
    sum += A[i];</pre>
```

```
add x9, x8, x0 # x9 = &A[0]
add x10, x0, x0 # sum = 0
add x11, x0, x0 # i = 0
addi x13, x0, 20 # x13 = 20
```



```
int A[20];
// fill A with data
int sum = 0;
for (int i = 0; i < 20; i++)
    sum += A[i];</pre>
```

```
add x9, x8, x0 # x9 = &A[0]
add x10, x0, x0 # sum = 0
add x11, x0, x0 # i = 0
addi x13, x0, 20 # x13 = 20
Loop:
```



```
int A[20];
// fill A with data
int sum = 0;
for (int i = 0; i < 20; i++)
    sum += A[i];</pre>
```

```
add x9, x8, x0 # x9 = &A[0]

add x10, x0, x0 # sum = 0

add x11, x0, x0 # i = 0

addi x13, x0, 20 # x13 = 20

Loop:

bge x11, x13, Done
```

You can only compare branches using registers—register to register, **not** e.g., branch less than immediate



```
int A[20];
// fill A with data
int sum = 0;
for (int i = 0; i < 20; i++)
    sum += A[i];</pre>
```

```
add x9, x8, x0 # x9 = &A[0]

add x10, x0, x0 # sum = 0

add x11, x0, x0 # i = 0

addi x13, x0, 20 # x13 = 20

Loop:

bge x11, x13, Done
```

Ok, why bge and not blt?

If that test fails, you never enter the for loop—branches usually point to the **exit**, not the body, and can **save** 1 unconditional jump



```
int A[20];
// fill A with data
int sum = 0;
for (int i = 0; i < 20; i++)
    sum += A[i];</pre>
```

```
add x9, x8, x0 # x9 = &A[0]

add x10, x0, x0 # sum = 0

add x11, x0, x0 # i = 0

addi x13, x0, 20 # x13 = 20

Loop:

bge x11, x13, Done

lw x12, 0(x9) # x12 = A[i]
```



```
int A[20];
// fill A with data
int sum = 0;
for (int i = 0; i < 20; i++)
    sum += A[i];</pre>
```

```
add x9, x8, x0 # x9 = &A[0]

add x10, x0, x0 # sum = 0

add x11, x0, x0 # i = 0

addi x13, x0, 20 # x13 = 20

Loop:

bge x11, x13, Done

lw x12, 0(x9) # x12 = A[i]

add x10, x10, x12 # sum += x12
```



```
int A[20];
// fill A with data
int sum = 0;
for (int i = 0; i < 20; i++)
    sum += A[i];</pre>
```

```
add x9, x8, x0 # x9 = &A[0]
add x10, x0, x0 # sum = 0
add x11, x0, x0 # i = 0
addi x13, x0, 20 # x13 = 20

Loop:
bge x11, x13, Done
lw x12, 0(x9) # x12 = A[i]
add x10, x10, x12 # sum += x12
addi x9, x9, 4 # &A[i+1]
```



```
int A[20];
// fill A with data
int sum = 0;
for (int i = 0; i < 20; i++)
    sum += A[i];</pre>
```

```
add x9, x8, x0 # x9 = &A[0]
 add x10, x0, x0 # sum = 0
 add x11, x0, x0 # i = 0
 addi x13, x0, 20 # x13 = 20
Loop:
 bge x11, x13, Done
 1w \times 12, 0(x9) \# x12 = A[i]
 add x10, x10, x12 # sum += x12
 addi x9, x9, 4 # &A[i+1]
 addi x11, x11, 1 # i++
```



```
int A[20];
// fill A with data
int sum = 0;
for (int i = 0; i < 20; i++)
    sum += A[i];</pre>
```

```
add x9, x8, x0 # x9 = &A[0]
 add x10, x0, x0 # sum = 0
 add x11, x0, x0 # i = 0
 addi x13, x0, 20 # x13 = 20
Loop:
 bge x11, x13, Done
 lw \times 12, 0(x9) \# x12 = A[i]
 add x10, x10, x12 # sum += x12
 addi x9, x9, 4 # &A[i+1]
 addi x11, x11, 1 # i++
  j Loop
```



```
int A[20];
                                           x9, x8, x0 # x9 = &A[0]
                                       add
                                       add x10, x0, x0 # sum = 0
// fill A with data
                                       add x11, x0, x0 # i = 0
int sum = 0;
                                       addi x13, x0, 20 # x13 = 20
for (int i = 0; i < 20; i++)
                                     Loop:
    sum += A[i];
                                       bge x11, x13, Done
                                           x12, 0(x9) # x12 = A[i]
                                       add x10, x10, x12 # sum += x12
                                       addi \times 9, \times 9, 4 # &A[i+1]
                                       addi x11, x11, 1 # i++
       1 conditional jump and 1 unconditional jump
                                     Done:
```