

Design a Computer System

John von Neumann

(<u>/vpn 'nɔɪmən/</u>) 1903 –1957

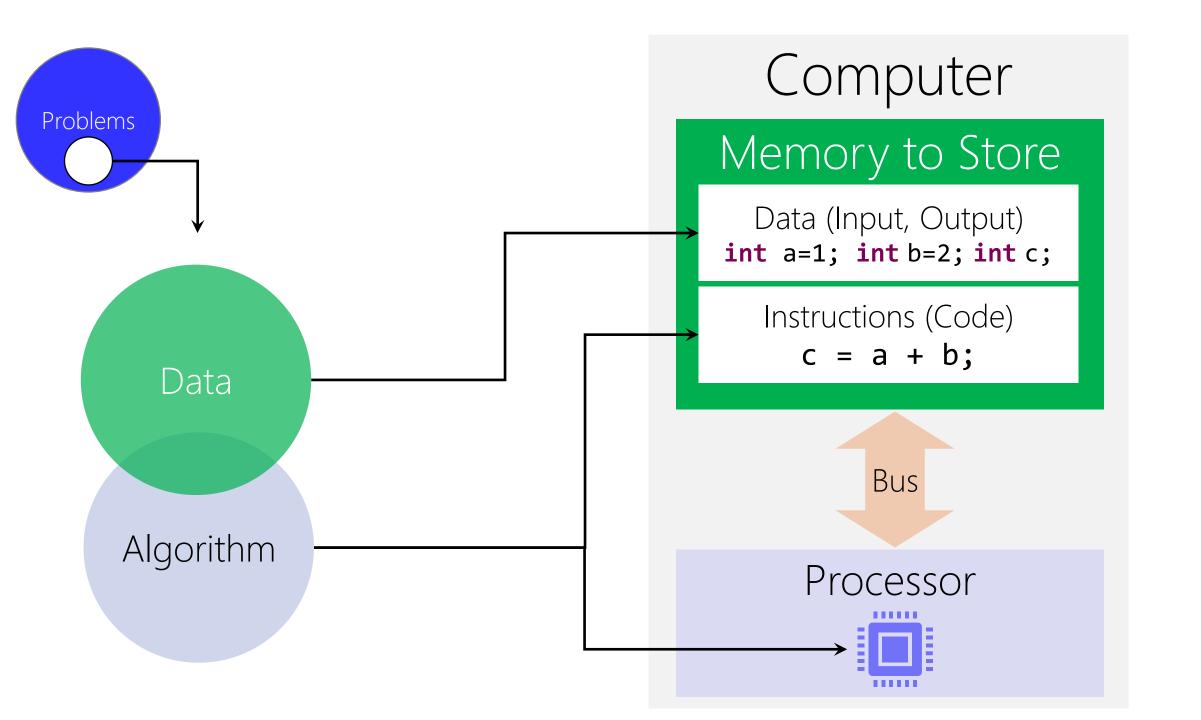
Mathematician, Physicist, Computer Scientist, Engineer

Polymath

He integrated pure and applied sciences. He made major contributions to many fields, including:

- Mathematics
- Physics
- Economics (game theory)
- Computing
- Statistics





von Neumann Architecture

Principles

- Data and instructions are both stored in the main memory
- The content of the memory is addressable by location (regardless of what is stored in that location)
- Instructions are executed sequentially unless the order is explicitly modified

Computer System

Input/Output Devices scanf("%d", &a); scanf("%d", &b); printf("%d", c);



Computer

Memory to Store

```
Data (Input, Output)
int a=1; int b=2; int c;
```

Instructions (Code)

$$c = a + b;$$



Processor

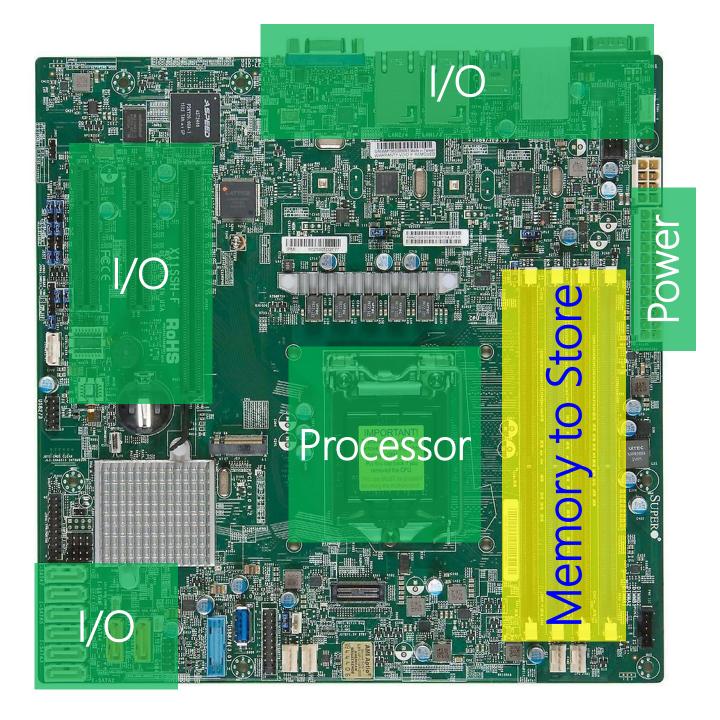


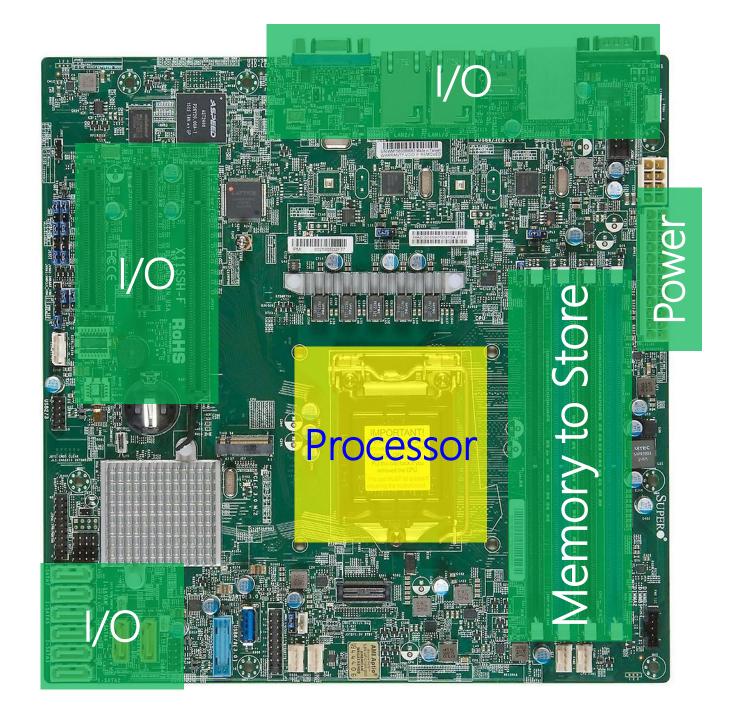
Bus

Permanent
Storage
fprintf()
fscanf()
fread()

fwrite()

fseek()





Do Calculation

Calculation on Data → Calculation on Numbers

Number → Binary

Computer

Memory to Store

```
Data (Input, Output)
int a=1; int b=2; int c;
```

Instructions (Code)

$$c = a + b;$$



Processor



- 1. We instruct the processor what to do!
- 2. Write instructions (programs)
- 3. Processor does our instructions

Instruction Set

- 1. Cannot instruct a processor to do whatever we want!
- 2. Any processors have limitations.
 - 1. Some can only do addition,
 - 2. Some can do both addition and subtraction, but no division



Intel® 64 and IA-32 Architectures Software Developer's Manual

Volume 2 (2A, 2B, 2C & 2D): Instruction Set Reference, A-Z

https://www.intel.com/content/dam/www/public/us/en/documents/manuals/64-ia-32-architectures-software-developer-instruction-set-reference-manual-325383.pdf

NOTE: The Intel 64 and IA-32 Architectures Software Developer's Manual consists of three volumes: Basic Architecture, Order Number 253665; Instruction Set Reference A-Z, Order Number 325383; System Programming Guide, Order Number 325384. Refer to all three volumes when evaluating your design needs.

Assembly Language

Writing programs using the Instruction Set of a particular processor

- RISC (/rɪsk/): Reduced Instruction Set Computer
- CISC (/'sɪsk/): Complex Instruction Set Computer

Design Processor RISC

Small but highly optimized set of instructions e.g., integer calculation

https://en.wikipedia.org/wiki/Reduced_instruction_set_computer

Design Processor CISC

Large and NOT highly optimized set of instructions e.g., integer and floating-point calculations

https://en.wikipedia.org/wiki/Complex_instruction_set_computer

Current Processors

x86, x64

By Intel and AMD instruction set size: ~981

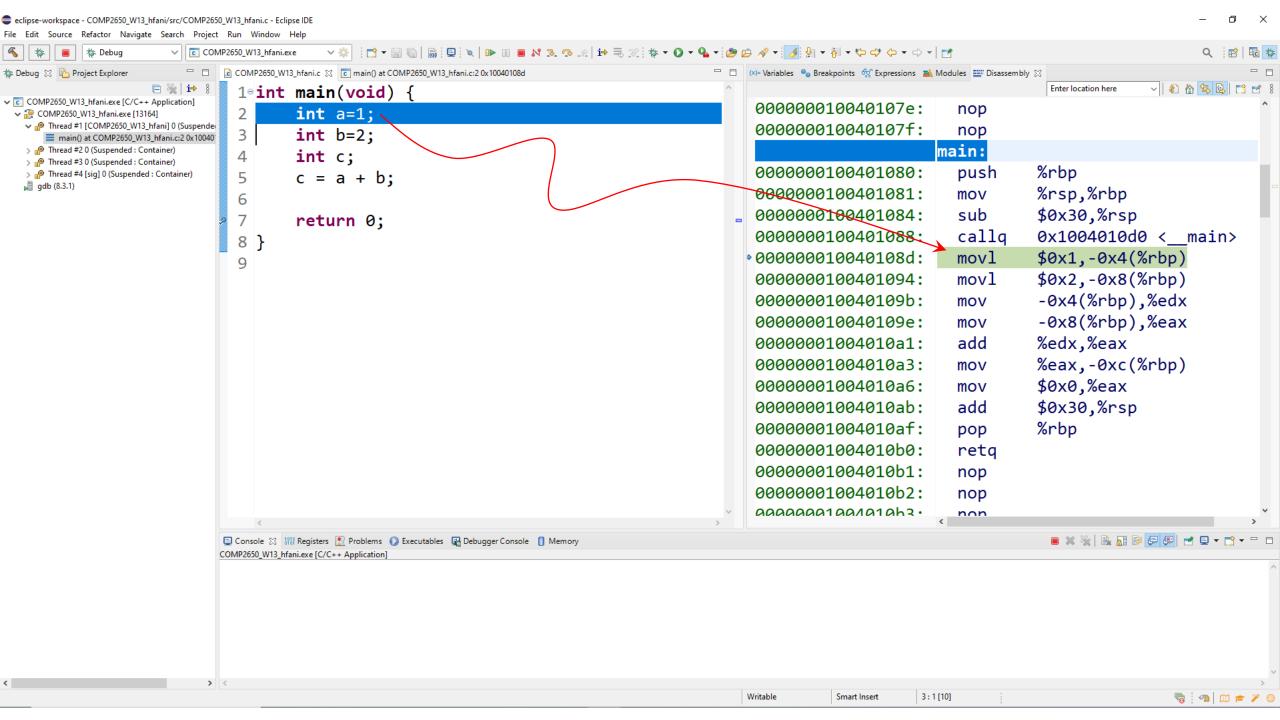
How Many x86-64 Instructions Are There Anyway?

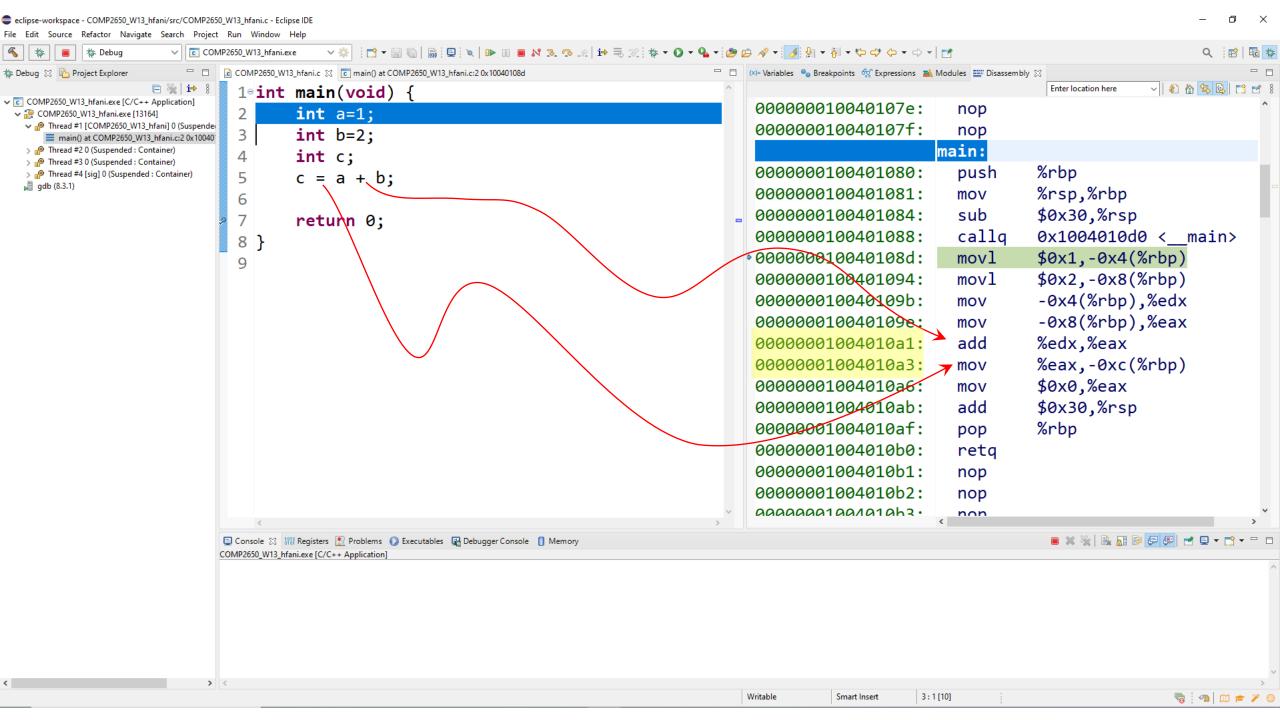
Human Natural Language	Туре	Example
High Level Programing Language	Imperative (What)	SQL SELECT * FROM Students
Middle Level Programming Language	Declarative (What + How)	<pre>C, C++, Java, Python int a = 1; a = a + 1;</pre>
Low Level Programming Language (Assembly Language)		Assembly x86, Assembly Z80 MOV 1, AX INC AX
Binary Digits (<mark>Machine Language</mark>)		00010011 00001011

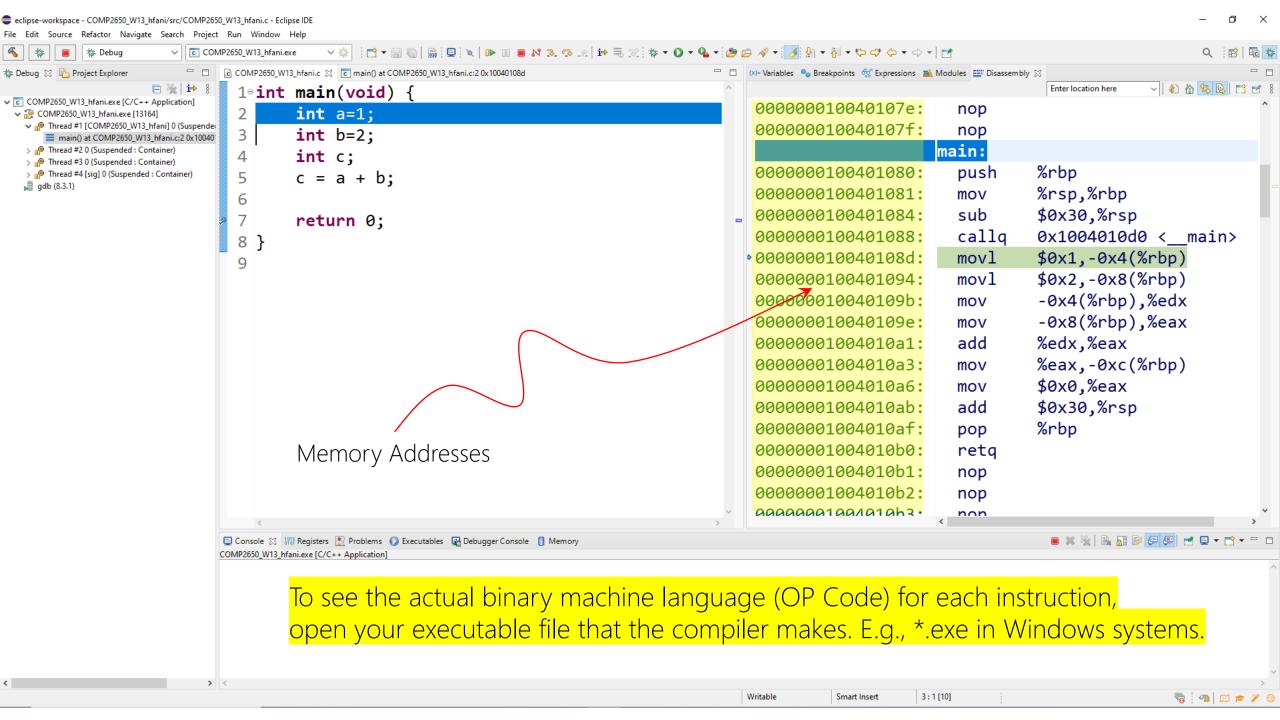
Human Natural Language			Example
High Level Programing Language	Impera		SQL SELECT * FROM Students
Middle Level Programming Language	Declar	Compilers Interpreters	<pre>C, C++, Java, Python int a = 1;</pre>
Low Level Programming Language (Assembly Language)			Assembly x86, Assembly Z80 MOV 1, AX INC AX
Binary Digits (Machine Language)			00010011 00001011

C/C++

Hossein's Computer







- 4 u**acknulnul**è**nulnulnul**e1à1ò1éè¤**nulnulnul**e1à1ò1éè"**nulnul**nule1à1ò1éè¬nulnulnule1à1ò1éhfä(é¬nulnulnulleno©sinulnul</mark>1òh
- DLENULNULNULÉESCNULNULNULÍÍ. SIUS "NULNULNULNULNULÄŸ%. pNULNULŸ%&pNULNULHÍì (1ÒèUNULNULŸNAKDC3pNULNULÄÄÄÄVSHÍÌ (H‰ÎH‰Ó¹BSNULNULNULÈĒEOTNULNULH...ÛSI "DC1STXNULNULH°NULNULNULNULNULTSOHNULNULH‰s@H5ŒŸŸŸH
- 6 eENQNULNULH%"ðNULNULNULHNAKoENQNULNULH%SPH<NAKxoNULNULH%°€NULNULNULH5nÿÿÿH‰KHH

H5úETXNULNULH‰S (HNAK?EOTNULNULHÇETXNULNULNULNULNULÇCDLEBELNULNULNULH‰K0H‰sxH‰"CANSOHNULNULH...ÀSI,,½NULNULNULH<€àNULNULNULH...ÀSI,,-NULNULNUL< NAKETX^NULNUL...ÒSI...ŸNULNULNULH∫=ETXSONULNULNULSI,,

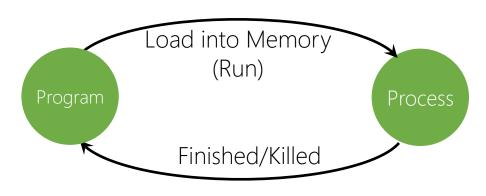
- 9 STX NUL NUL Hf=ý
- 10 NULNULNULSI"OSTXNULNULHf=÷
- 11 NULNULNULSI,,QSTXNULNULHf=ñ
- 12 **NULNULNILSI**,,3**STXNULNUL**Hf=ë
- 13 **NULNULNULSI,,NAKSTXNULNUL**Hf=å
- 14 NULNULNULSI,, ÷SOHNULNULHf=B
- 15 NULNULNULSI,, Ù SOHNULNULH f = Ù 16 NULNULNULSI,,, » SOHNULNUL f SI o ENO"
- 17 NULNULSIDC1NULfSIo
- 19 NULNULSIDC1HDLEfSIoNAKœ
- 20 NULNULSIDC1P fSIoGS

Program vs. Process

- Program: dead body of the instructions stored in permanent place (flash drive, ssd, hard disk, paper!)
- Process: live body of the instructions stored in memory, ready to be fetched by the Processor and be executed!

Program vs. Process

```
C, C++, Java, Python
int a = 1;
a = a + 1;
Assembly x86, Assembly Z80
MOV 1, AX
INC AX
00010011
00001011
```



00010011 00001011	
00010011 00001011	
00010011 00001011	

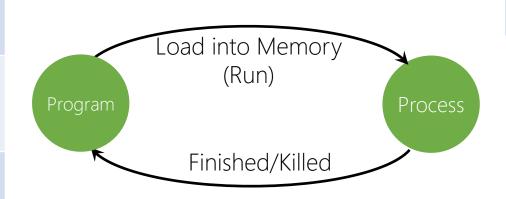
Compilers vs. Interpreters

Instruction Encoders

C, C++, Java, Python int a = 1; a = a + 1; Assembly x86, Assembly Z80 MOV 1, AX

00010011 00001011

INC AX



Interpreters

```
int a = 1; \rightarrow 00010011

a = a + 1; \rightarrow 00001011
```

Instruction	Operation Code (OP Code)
c = a + b	000 XXXX YYYY ZZZZ
c = a - b	001 XXXX YYYY ZZZZ
c = a / b	010 XXXX YYYY ZZZZ
c = a * b	011 XXXX YYYY ZZZZ
c = {number}	100 XXXX XXX XXXX

Instruction Encoders very simplified version!

Processor can only see binary-coded instructions

Processor only understand machine language

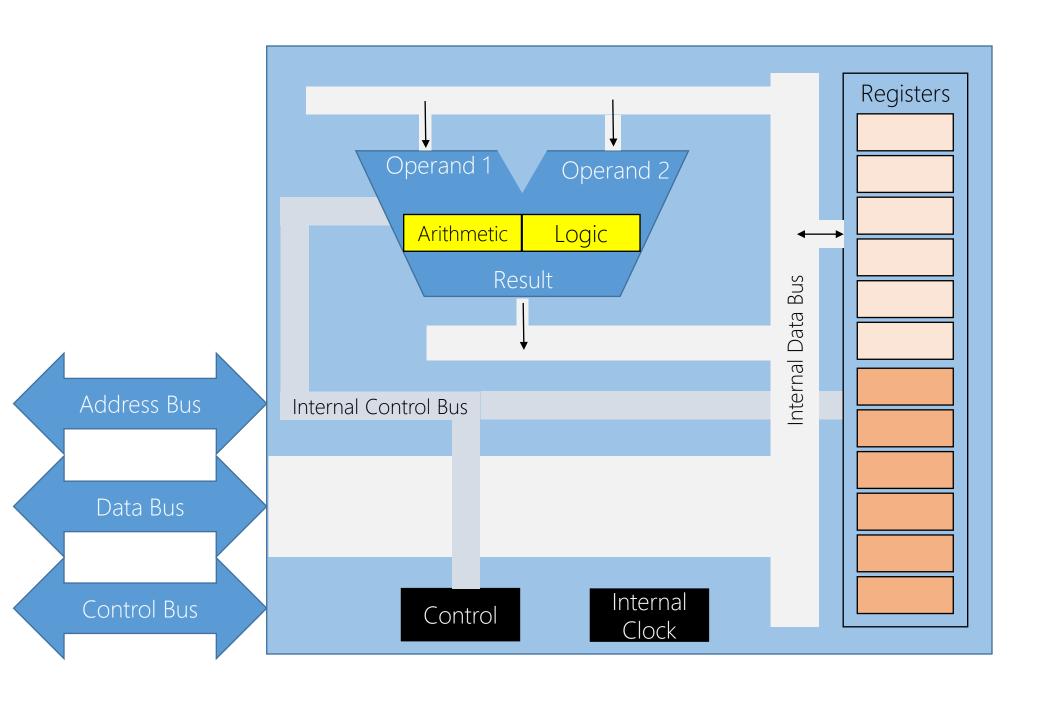
Instruction Set →Instruction Decoder

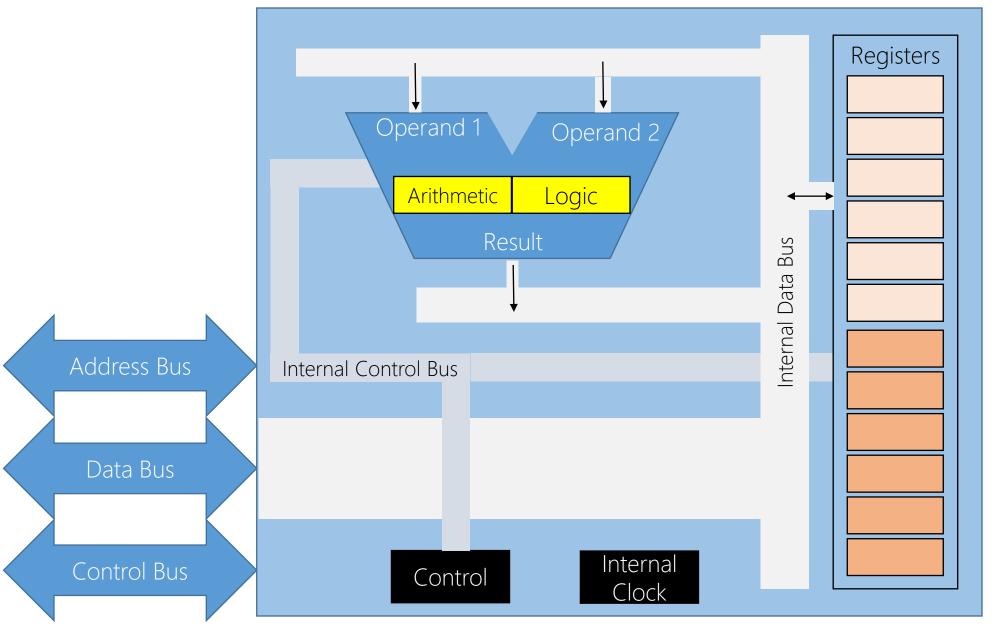
Operation Code (OP Code)	What to do?
000 XXXX YYYY ZZZZ	 Fetch the first operand from memory at XXXX address Store the first operand inside somewhere (AX) Fetch the second operand from memory at YYYY address Store the first operand inside somewhere else (BX) Use the n-bit Adder to add AX and BX Store the result inside somewhere else (CX) Push CX to memory at ZZZZ address
001 XXXX YYYY ZZZZ	
010 XXXX YYYY ZZZZ	
011 XXXX YYYY ZZZZ	
100 XXXX XXXX	

Instruction Decoder very simplified version!

Operation Code (OP Code)	What to do?
000 XXXX YYYY ZZZZ	1) Fetch the first operand from memory at XXXX address 2) Store the first operand inside somewhere (AX) 3) Fetch the second operand from memory at YYYY address 4) Store the first operand inside somewhere else (BX) 5) Use the n-bit Adder to add AX and BX 6) Store the result inside somewhere else (CX) 7) Push CX to memory at ZZZZ address
001 XXXX YYYY ZZZZ	
010 XXXX YYYY ZZZZ	
011 XXXX YYYY ZZZZ	
100 XXXX XXXX	

Instruction Decoder very simplified version!





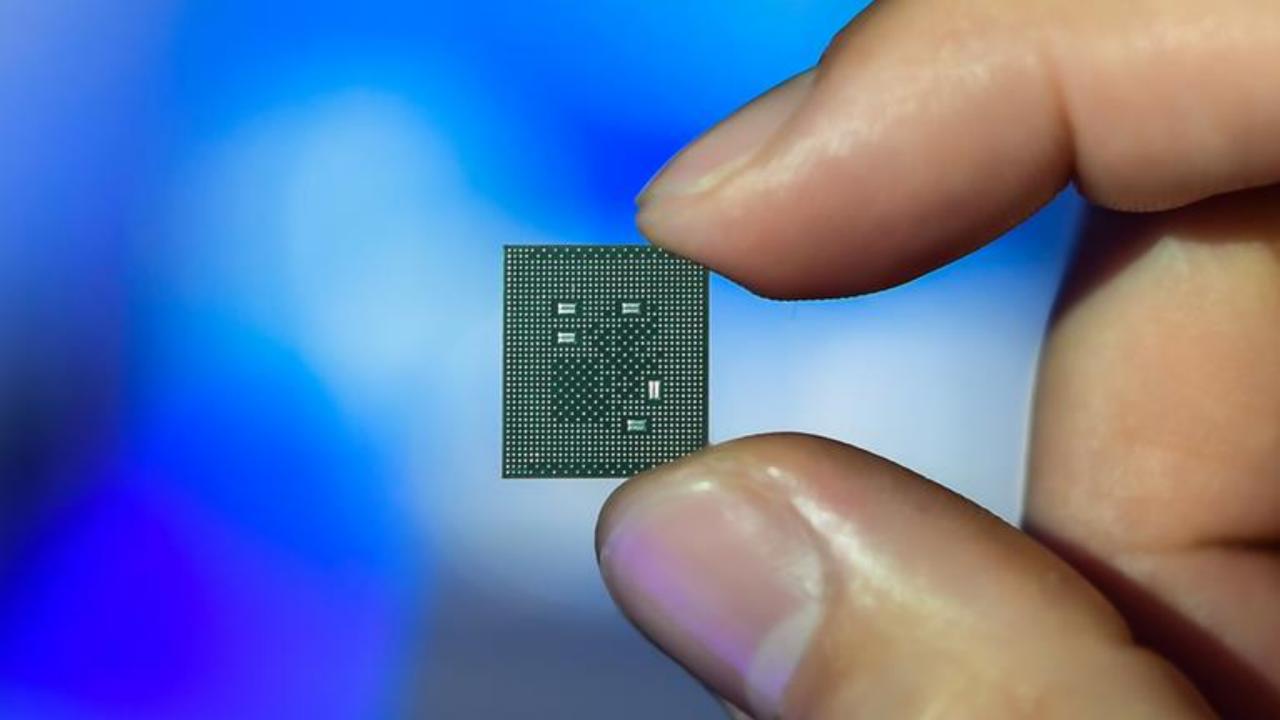
City

- Streets

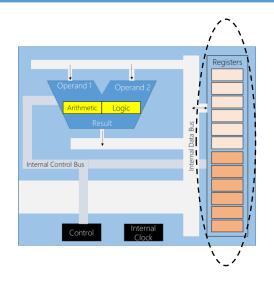
1-way

2-way

- Gates
- Bridges
- Warehouses
- Policing



Registers Some vectors of FFs

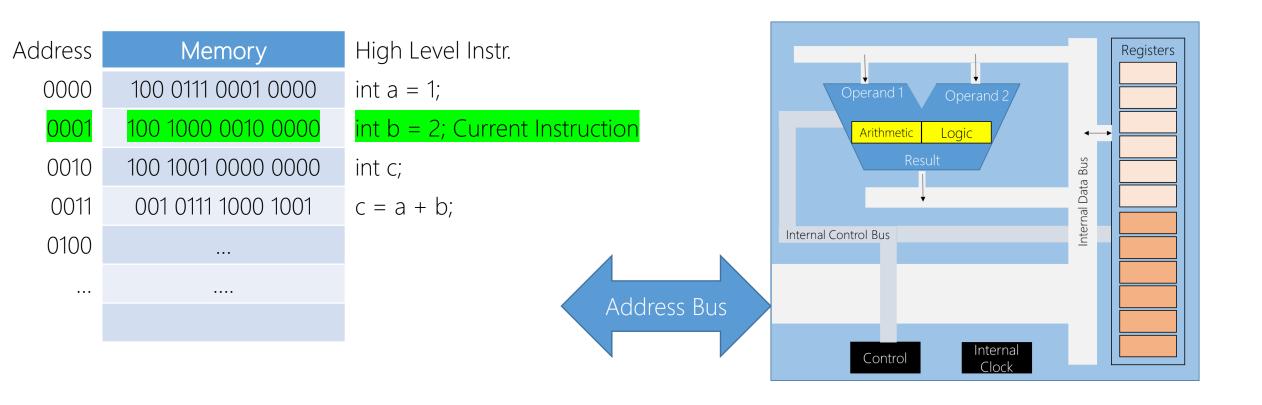


aka. Program Counter (PC), Instruction Address Register (IAR), Instruction Counter

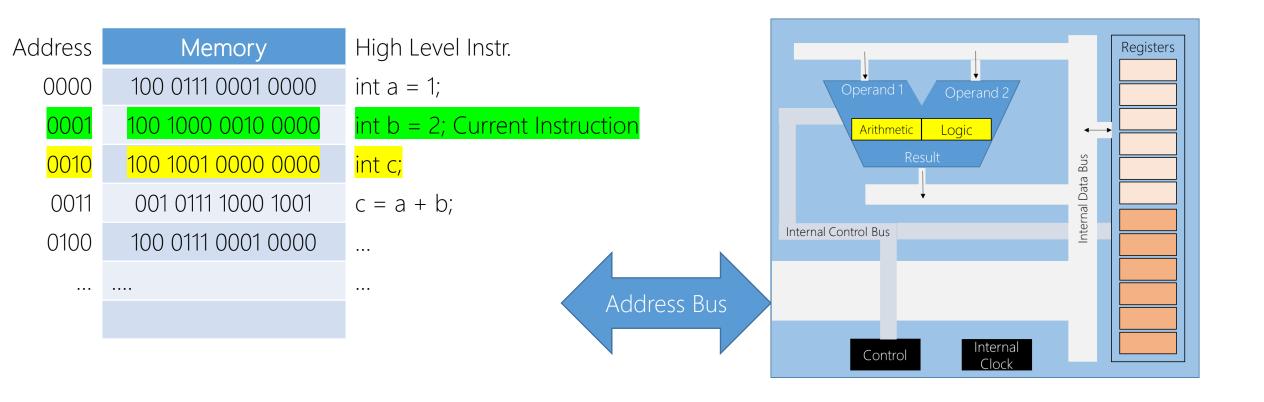
Keep track of the address of next instruction in memory

aka. Program Counter (PC), Instruction Address Register (IAR), Instruction Counter

Keep track of the address of next instruction in memory Why is it called 'counter'?



Where is the next instruction to be fetched and execute?



In von Neumann architecture, instructions are executed sequentially!

So, the next would be the address of current + 1

aka. Program Counter (PC), Instruction Address Register (IAR), Instruction Counter

We already designed a 3-bit counter 00, 01, 10, 11

aka. Program Counter (PC), Instruction Address Register (IAR), Instruction Counter

IP is a <mark>n</mark>-bit counter. What is <mark>n</mark>?

aka. Program Counter (PC), Instruction Address Register (IAR), Instruction Counter

IP is a <mark>n</mark>-bit counter. What is <mark>n</mark>?

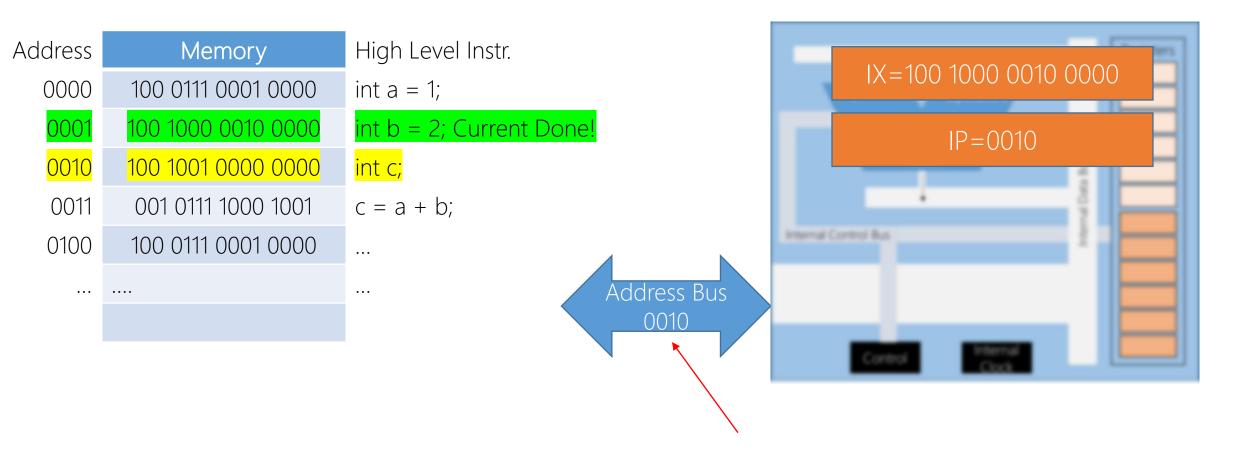
- IP contains a memory address.
- Address bus contains a memory address at a time.
- So, IP is the same size as address bus. Here, 4-bit address bus, we have 4-bit IP.

Where is the **current** instruction?

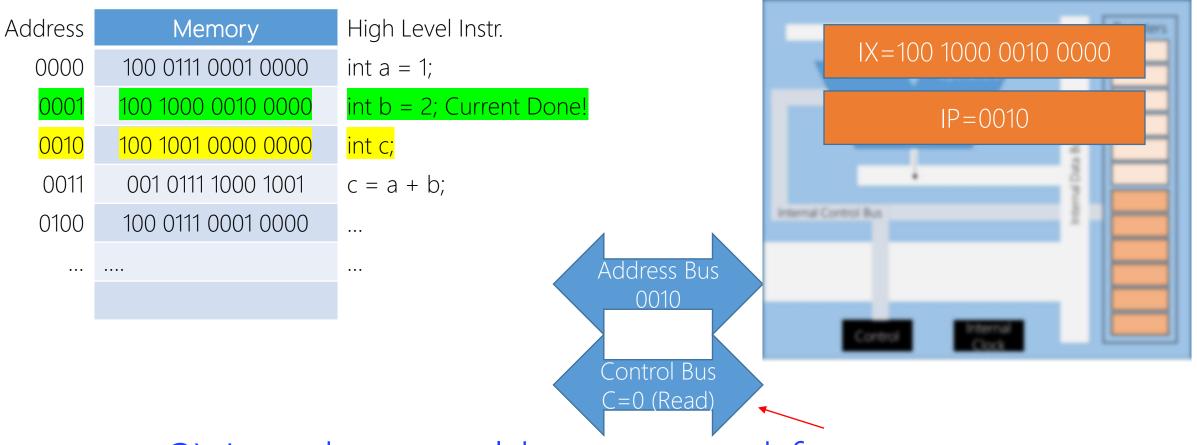
Instruction Register (IR or IX)

- Holds the current instruction that is fetched from memory and is just about to be executed.
- As soon as the instruction placed in IR, the processor starts
 executing the instruction and IP++ to point to the next instruction
 to be executed next.

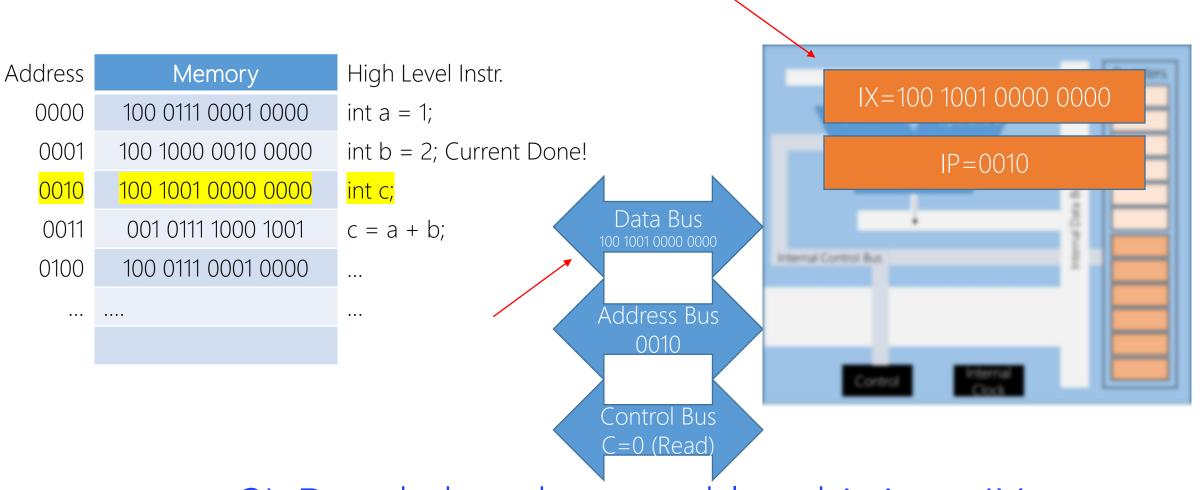
	Manaami				
Address	Memory	High Level Instr.		IX=100 1000 0010 0000	
0000	100 0111 0001 0000	int a = 1;		17-100 1000 0010 0000	
0001	100 1000 0010 0000	int b = 2; Current Instruction		IP=0010	
<mark>0010</mark>	<mark>100 1001 0000 0000</mark>	int c;			
0011	001 0111 1000 1001	c = a + b;			
0100	100 0111 0001 0000		Internal	Control Bus	
		Address Bus			
				Total State of	



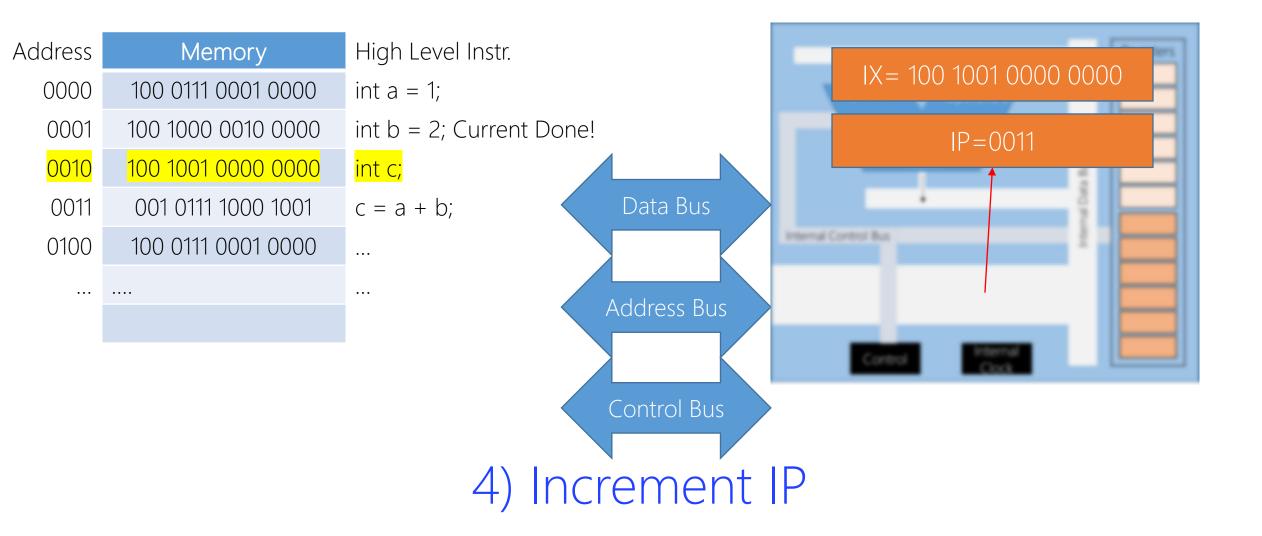
1) Load address bus with IP



2) Load control bus to read from memory



3) Read data bus and load it into IX



Instruction Register (IR or IX)

n-bit Register. What is n?

Fetch Instruction Cycle

- 1) Load address bus with IP
- 2) Load control bus to read from memory
- 3) Read data bus and load it into IX
- 4) Increment IP

Memory <mark>Address</mark> Register (MAR)

Holds the address of memory location to Read/Write data from.

Memory Data Register (MDR)

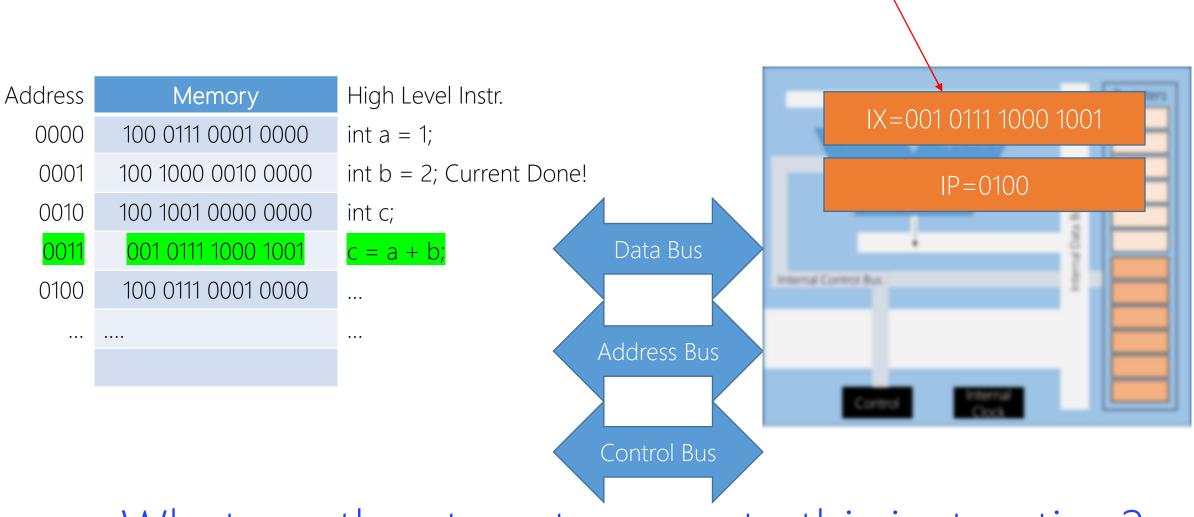
Holds the actual value that is either

- Read from memory, or
- To be written to memory

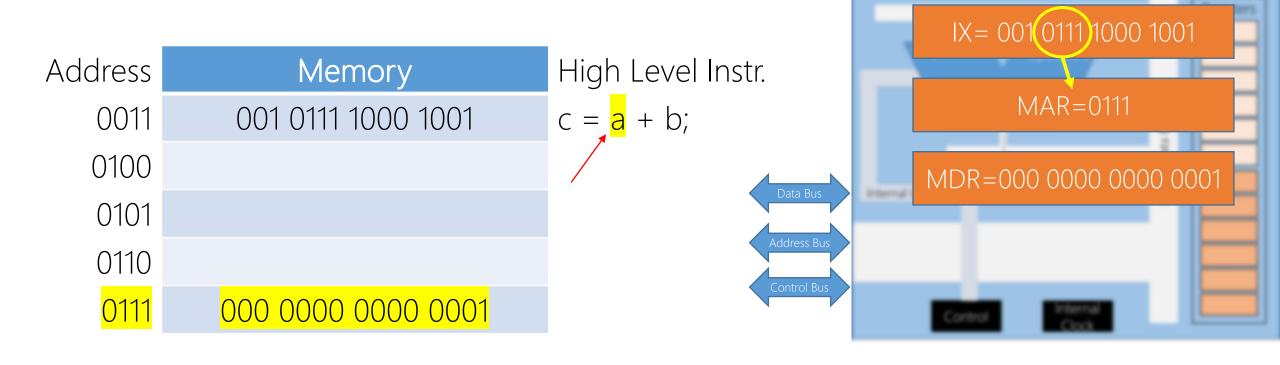
Memory Address Register (MAR) Memory Data Register (MDR)

Work hand-in-hand to facilitate the communication of the processor and the main memory

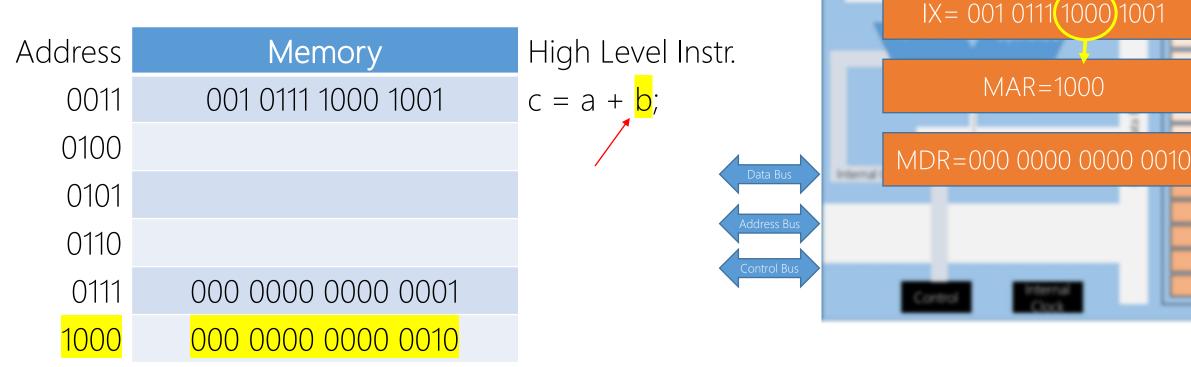
What are their sizes?



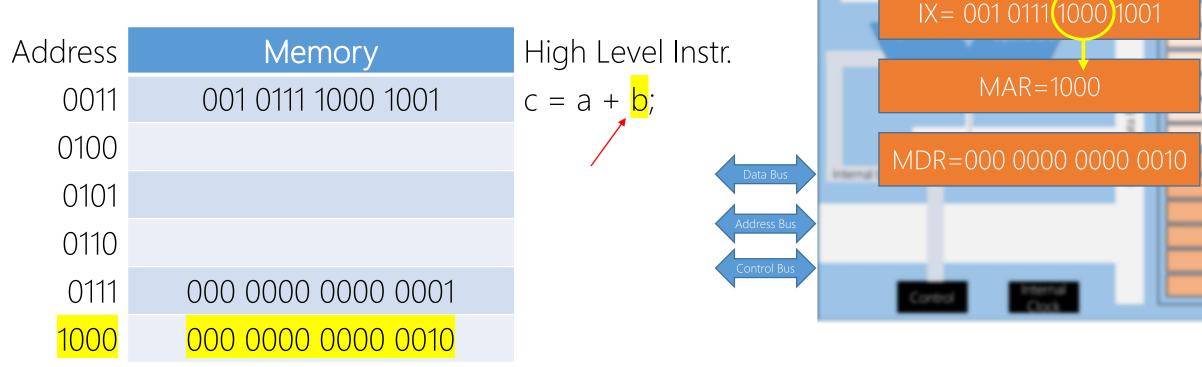
What are the steps to execute this instruction?



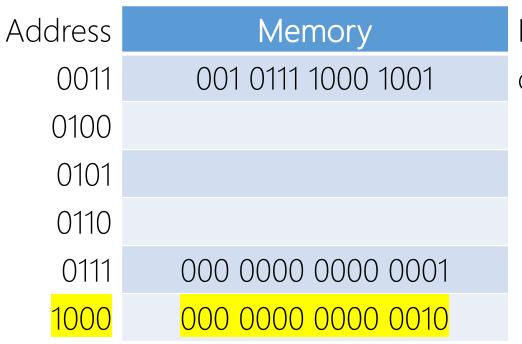
- 1) Fetch the first operand from memory
 - I. Load MAR with the address of 'a' from the instruction op code
 - II. Load address bus with MAR
 - III. Load control bus with C=0 (Read)
 - IV. Load MDR with the content of data bus



- 2) Fetch the second operand from memory
 - I. Load MAR with the address of 'b' from the instruction op code
 - II. Load address bus with MAR
 - III. Load control bus with C=0 (Read)
 - IV. Load MDR with the content of data bus

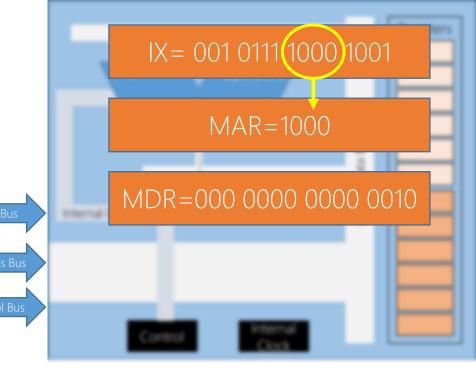


- 2) Fetch the second operand from memory
 - I. Load MAR with the address of 'b' from the instruction op code
 - II. Load address bus with MAR
 - III. Load control bus with C=0 (Read)
 - IV. Load MDR with the content of data bus. Wait! I lost the value of 'a' inside the processor!

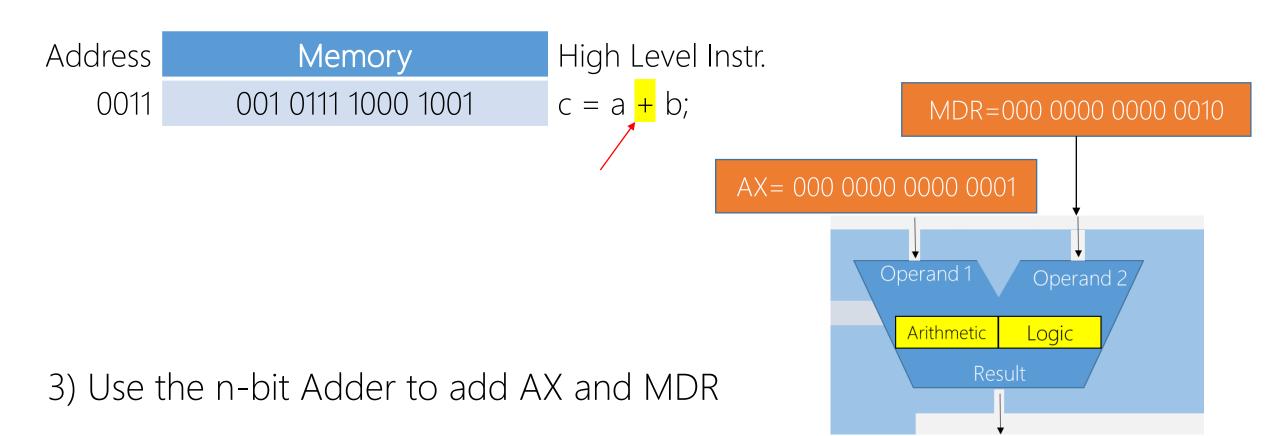


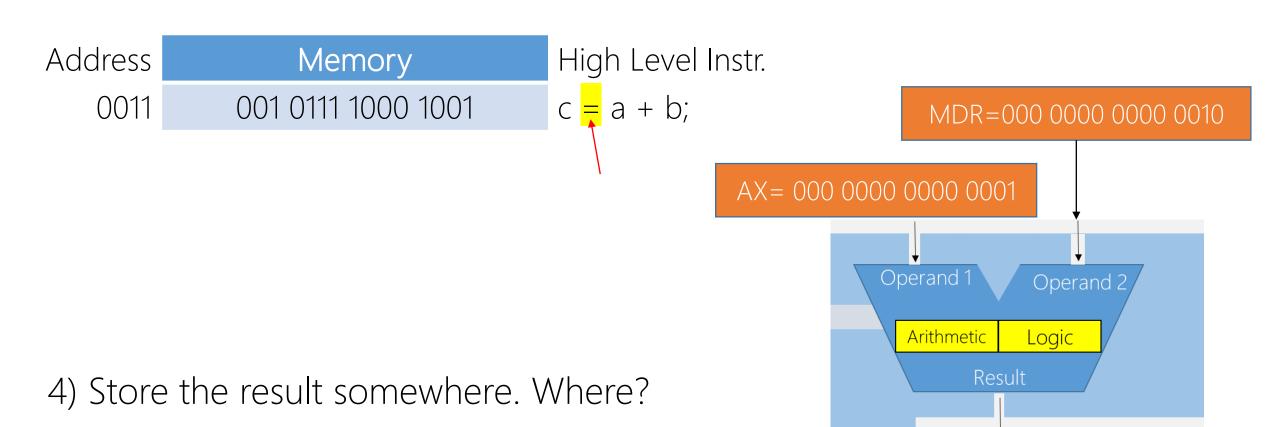
High Level Instr.

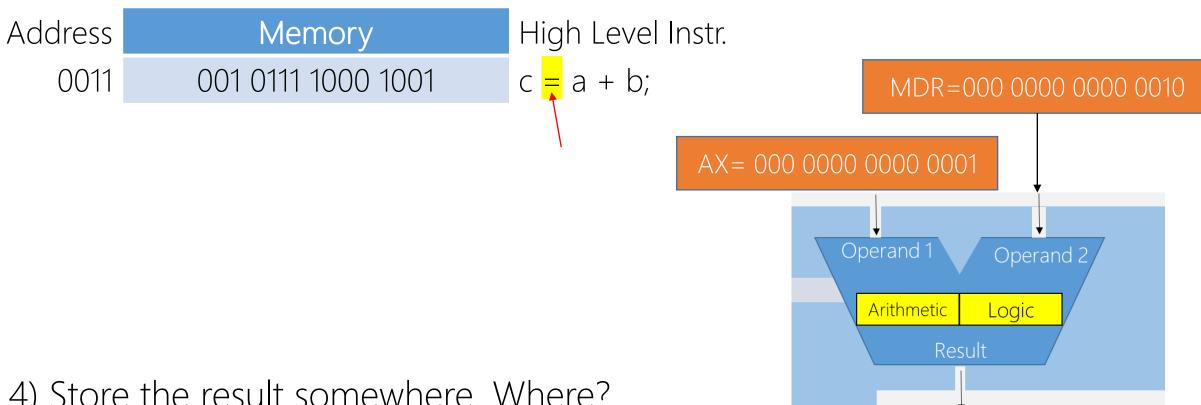
$$c = a + b;$$



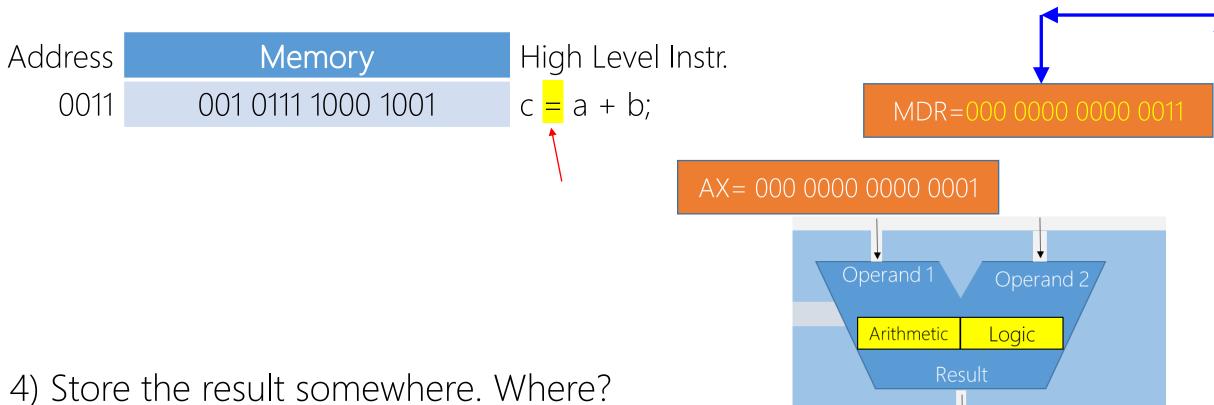
- 2) Fetch the second operand from memory
 - I. Move MDR (value of 'a') to somewhere (AX) first!
 - II. Load MAR with the address of 'b' from the instruction op code
 - III. Load address bus with MAR
 - IV. Load control bus with C=0 (Read)
 - V. Load MDR with the content of data bus.





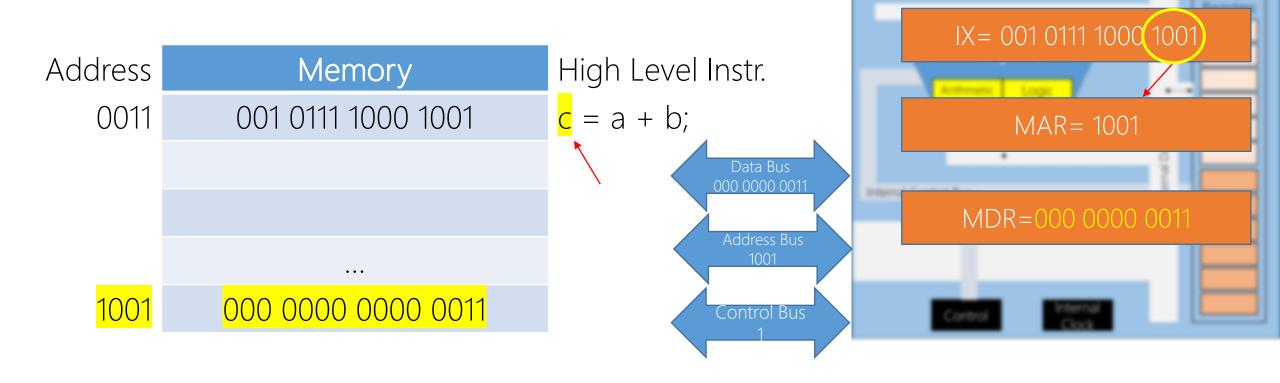


4) Store the result somewhere. Where? We want to write it back to memory at location for 'c' So, better to store it in MDR



What do we want to do next?

We want to write it back to memory at location for 'c' So, better to store it in MDR



- 5) Write the result back to memory at location for 'c'.
 - I. Find the address of 'c' in op code
 - II. Load it into MAR
 - III. Load address bus with MAR
 - IV. Load data bus with MDR
 - V. Load control bus with C=1 (Write)

Accumulator (AX)

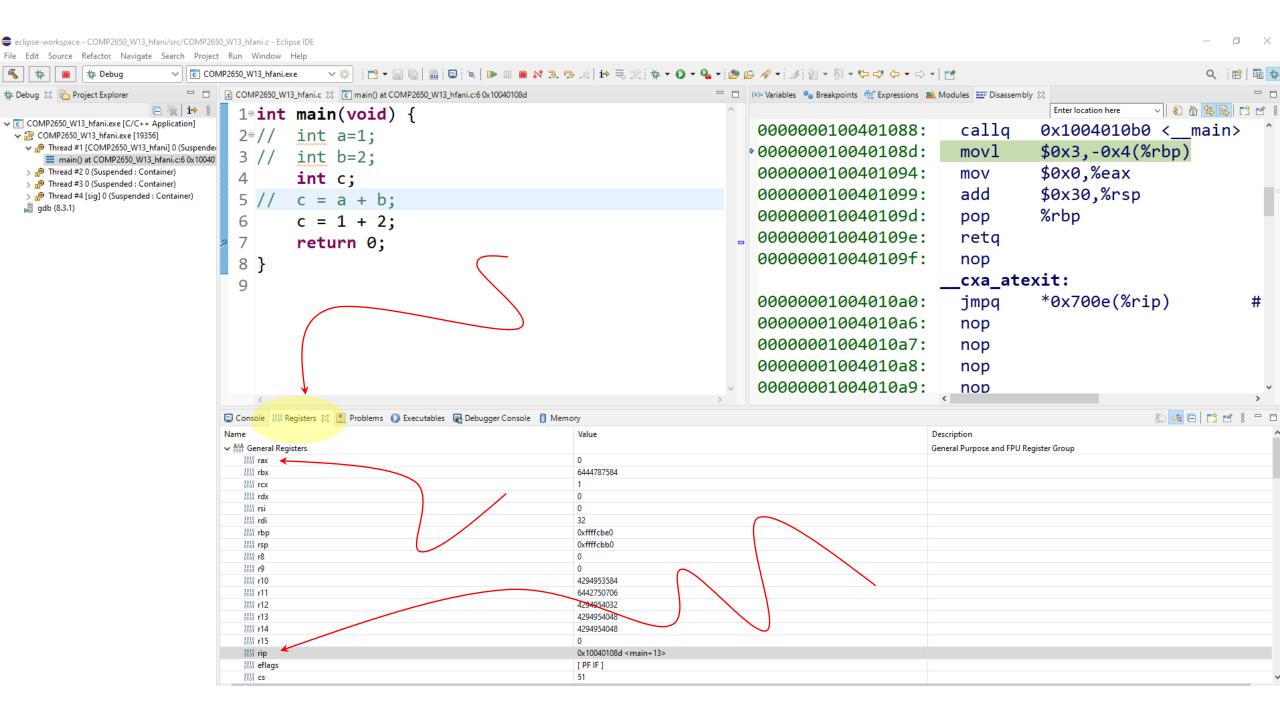
This is the most frequently used register.

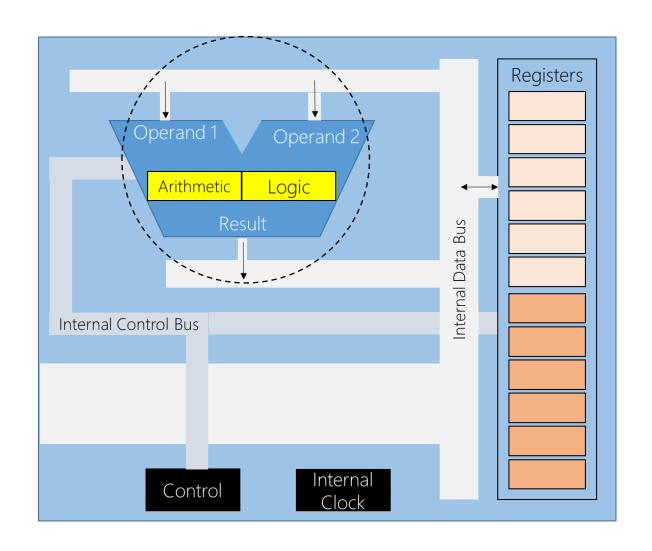
Any intermediate results

somewhere
AX

General Purpose Registers

 $R_1, R_2, ..., R_i$





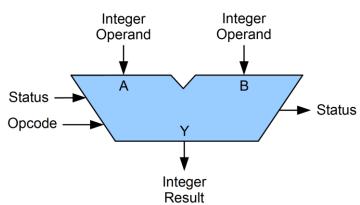
Arithmetic & Logic Unit (ALU)

Arithmetic Unit

Integer Arithmetic (Addition, Subtraction)

- Multiplication: multiple addition
- Division: multiple subtraction

Floating Point Arithmetic (FPU)

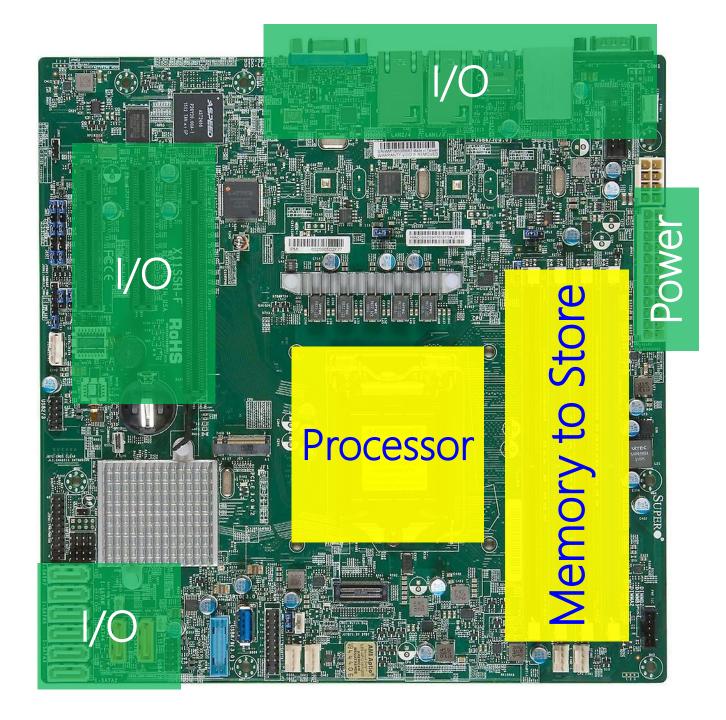


Logic Unit

Bitwise AND, OR, XOR, 1's Comp.

Recap

- 1) Fetch Instruction Cycle
- 2) Execution Cycle
 - Depends on OP Code
 Arithmetic
 Logic



Microprocessor

Central Processing Unit (CPU) Intel (Xeon), AMD (Ryzen)

Graphic Processing Unit (GPU)

nVidia (GeForce) for Video Card (Gaming)

Al Acceleration (Neural Nets)

Tensor Processing Unit (TPU)
Google (TensorFlow Al Lib for Neural Nets)

COMP-2660 Computer Architecture II Microprocessor Programming

Postscript

Office Hours Before Week

- Practice on questions

Office Hours Final Exam

- Review exam questions