Stored program concept and von Neumann architecture

3.1.1a · Computers · revision notes

The stored program concept means that a program's instructions and its data are held together in main memory, in binary, and the CPU fetches and executes the instructions one at a time. The von Neumann architecture is the computer design built on this idea: a CPU and main memory joined by buses. It comes up on Paper 1 as short recall, diagram-labelling and 2 to 4 mark 'Describe' and 'Explain' questions.

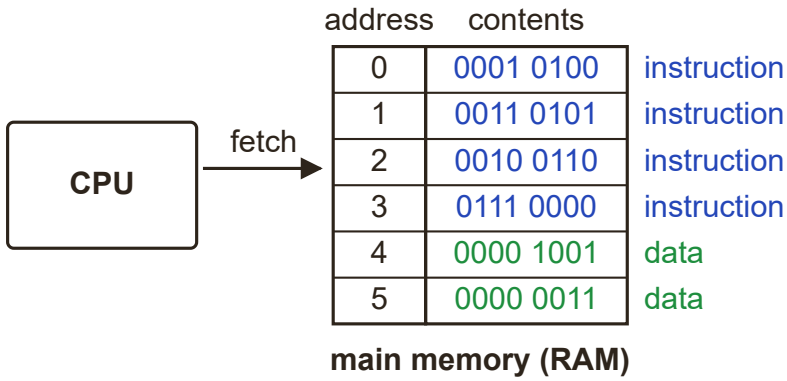
Key facts

- Stored program concept: instructions and data stored together in main memory, in binary
- Von Neumann architecture: CPU (CU, ALU, registers, clock) + main memory + buses
- The CPU fetches directly only from main memory, never from secondary storage
- RAM is volatile: it loses its contents when the power is off
- Instructions are fetched, decoded and executed one at a time, in sequence
- Main memory (RAM): holds the instructions and data currently in use
- Address bus: one-way, CPU to memory. Data bus and control bus: two-way
- Grade 7+ only A jump instruction changes the address in the program counter, so instructions can be repeated or skipped

Notes

The stored program concept

- A program is a sequence of instructions. In the **stored program concept**, the instructions and the **data** they use are stored **together in main memory (RAM)**.

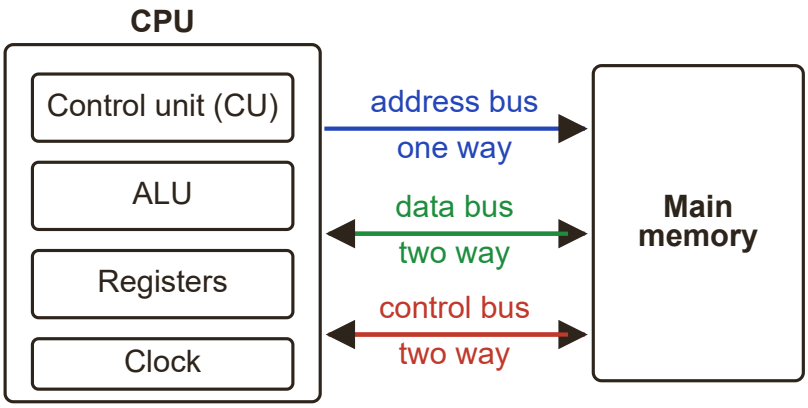


Instructions (blue) and data (green) sit together in main memory, all in binary. Each location has its own address

- Both are stored as **binary** patterns. Instructions in binary form are called **machine code**.
- The CPU **fetches** an instruction from memory, **decodes** it and **executes** it, then fetches the next one.
- Instructions are normally executed **sequentially**: one after another, in the order they are stored.
- To carry out a different task, you just load a different program into memory. The hardware does not change. Some early computers had to be rewired for each new task.

The von Neumann architecture

- The **CPU** contains the **control unit** (CU), the **arithmetic logic unit** (ALU), **registers** and a **clock**.
- **Main memory** is made of numbered locations. Each location has its own unique **address**.
- The CPU and main memory are joined by three **buses** (groups of wires): the **address bus**, the **data bus** and the **control bus**.
- The address bus is one-way (from the CPU to memory). The data bus and control bus carry signals both ways.



The von Neumann architecture: one memory, joined to the CPU by three buses

The role of main memory

- Main memory holds the operating system, programs and data that are **currently in use**.
- The CPU can only fetch instructions and data **directly from main memory**, not from secondary storage. So a program is copied from secondary storage into RAM before it runs.
- RAM is much faster to access than secondary storage, so the CPU spends less time waiting. The CPU can also write results back to main memory.
- RAM is **volatile**: it loses its contents when the power is off, so programs and files are kept permanently on secondary storage.

One memory for instructions and data Grade 7+ only

- A binary pattern does not show whether it is an instruction or data. A pattern fetched from the address held in the **program counter** is treated as an instruction; a pattern read while an instruction is being executed is treated as data. Grade 7+ only
- A **jump** (branch) instruction puts a new address in the program counter. This is how a program repeats instructions (a loop) or skips over them. Grade 7+ only
- Drawback: instructions and data travel along the **same buses**, so only one can be fetched at a time and the CPU may have to wait. A program error could also overwrite instructions with data. Grade 8+ only

How to answer each type of question

Describe the stored program concept 2 to 3 marks **Grade 4**

1. Say where: the instructions are stored in main memory.
2. Say what: the instructions are stored together with the data they use.
3. Say how: both are in binary, and the CPU fetches and executes the instructions one at a time, in sequence.

Example. Describe the stored program concept. (3)

Model answer. Any three from: program instructions are stored in main memory (1); in the same memory as the data the program uses (1); both are stored as binary patterns (1); the CPU fetches the instructions one at a time and executes them in sequence (1).

Explain why a program must be in main memory to run 2 marks Grade 5

1. State that the CPU can only fetch instructions and data directly from main memory.
2. Add a linked reason: RAM is much faster than secondary storage, so the CPU is not kept waiting.

Example. A games console keeps its games on an internal solid-state drive (SSD). When a player starts a game, the game is copied into main memory (RAM).
Explain why the game is copied into main memory before it runs. (2)

Model answer. The CPU can only fetch instructions and data directly from main memory, not from the SSD (1). RAM is much faster to access than the SSD, so the CPU does not have to wait as long for each instruction (1).

Explain an advantage of the stored program concept 2 marks Grade 6

1. Say that a new program is loaded into memory in the same way as data.
2. Link it: the CPU runs whatever instructions are in memory, so the task changes without changing the hardware.

Example. Some early computers had to be physically rewired before they could carry out a different task.
Explain how the stored program concept removed the need for this. (2)

Model answer. A new program can simply be loaded into main memory, in the same way as data (1). The CPU fetches and executes whatever instructions are stored in memory, so changing the stored program changes the task without any change to the hardware (1).

Explain a disadvantage of one shared memory 2 marks Grade 8

1. Identify the problem: instructions and data share the same memory and the same buses.
2. Explain the effect: the CPU has to wait, or data could overwrite instructions.

Example. In a von Neumann computer, instructions and data are stored in the same main memory.
Explain **one** disadvantage of this. (2)

Model answer.

Instructions and data travel along the same buses, so an instruction and an item of data cannot be fetched at the same time (1). This means the CPU often has to wait for memory, which limits how quickly programs can run (1).
Or: a program error, or an attacker's input, could overwrite instructions with data (1), so the program could crash or run the attacker's code (1).

Don’t lose marks

- Saying that instructions and data are kept in separate memories. In a von Neumann computer they share the same main memory.
- Saying the CPU runs a program directly from the hard disk or SSD. It must be copied into main memory first.
- Using 'memory' for the hard disk. Main memory is RAM; the hard disk or SSD is secondary storage.
- Describing the concept as 'the program is stored in the computer'. That is too vague: say main memory, together with the data, in binary.
- Saying the address bus carries data, or that it is two-way.

More tips

Memory tricks

- Stored program = instructions + data, same memory, both in binary, run one at a time.
- Von Neumann = CPU + main memory + buses. If you can sketch these boxes and label the three buses, you can answer most questions on it.
- The address bus is a one-way street: addresses only travel out from the CPU. Data and control signals go both ways.
- Sense check: nothing runs straight from the hard disk or SSD. Everything the CPU uses goes through main memory.

Exam technique

- For 'Describe the stored program concept', give separate facts: where (main memory), what (instructions and data together) and how (binary, fetched and executed in sequence).
- Use the exact terms: main memory, binary, fetch, decode, execute, sequentially.
- In a diagram question, the part of the CPU that decodes instructions is the control unit, and the part that does calculations and comparisons is the ALU.
- If the question has a context, such as a games console or a robot, use it in your answer.

What each grade needs

What you need to be able to do, from the first marks up to the top grade.

Grade 3 — **State what main memory holds**
Main memory (RAM) holds the program instructions and the data that are currently in use.

Grade 4 — **Name the parts of the von Neumann architecture**
A CPU (control unit, ALU, registers and clock), main memory (RAM), and the address, data and control buses.

Grade 4 — **Describe the stored program concept**
Instructions and data are stored together in main memory as binary, and are fetched and executed one at a time.

Grade 5 — **Explain why programs are loaded into RAM**
The CPU can only fetch instructions directly from main memory, which is also much faster to access than secondary storage.

Grade 6 — **Explain why new tasks need no rewiring**
A different program is simply loaded into memory, and the CPU executes whatever instructions are stored there.

Grade 7

Explain how a jump changes sequential execution

Instructions normally run in the order they are stored, but a jump puts a new address in the program counter, so instructions can be repeated or skipped.

Grade 8

Explain a drawback of one shared memory

Instructions and data travel on the same buses, so they cannot be fetched at the same time and the CPU may have to wait.

Free notes, Full packs, practice papers and mock exams at chhetriacademy.com. Written for this site; not taken from past papers. Stuck? 1-to-1 tutoring with a free 30-minute trial lesson.