33 – Processor

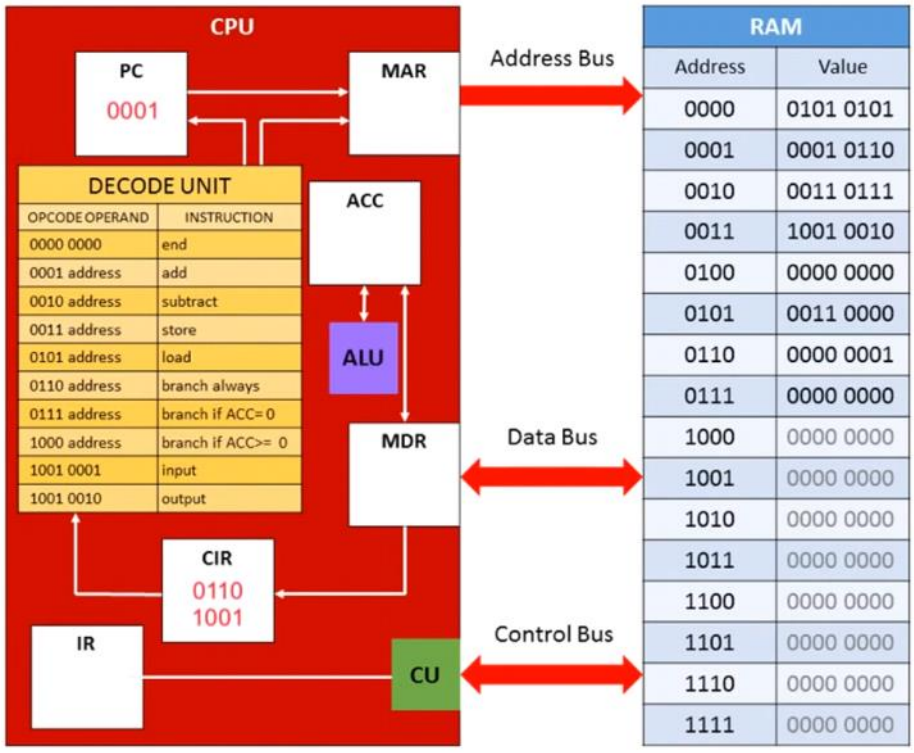
Thursday, January 10, 2019 9:26 AM

Fetch-decode-execute cycle

The running of every program follows a **fetch-decode-execute** cycle.

- Fetch – the processor takes the address of the next instruction from the PC and places it in the MAR. A control signal then tells RAM to read this address from the address bus and place the data at that memory location on the data bus. This data is then placed in the MDR. This is then copied into the CIR we have fetched an instruction to be processed. It includes both the opcode and the operand (e.g. `ADD 5` may be encoded).
- Decode – the op-code which is copied into the CIR is processed so the computer knows what to do (e.g. a code may represent BRANCH ALWAYS).
- Execute – the processor carries out the instruction – this may involve the ALU placing data into the ACC (accumulator) register for future instructions to use.

There is also an additional **interrupt** part of the cycle after the execute. Here, the processor checks for any interrupts and depending on their priority, may decide to halt what it is doing and handle the interrupt by looking up what it needs to do in an interrupt lookup table where the **interrupt service routine** will be followed.



Some terminology

The control unit is the part of the processor that coordinates the execution part of the FDE cycle. It tells the computer's memory, the ALU and I/O devices how to respond to given instructions.

The arithmetic logic unit (ALU), carries out two types of operations as the name suggests. It can carry out arithmetic (mathematical) operations such as adding, multiplying etc, as well as logic operations – determining if numbers are equal, less than, or greater than each other. It is sent and operation code (op-code) and the operands.

The internal clock of a computer generates the signal that is used to synchronise the operation of the processor and the movement of the data around the other components of the computer.

Registers

There are a varying number of general purpose registers that can be addressed and used to store data between calculations. The number of these quick-access processor registers varies between processes and can affect their performance

In addition to these general purpose registers, there are dedicated ones. For instance:

- The status register – holds information about parts of the processor. E.g. one bit may signify if the result of the last ALU calculation 0?
- The interrupt register (IR) – hold information about any data that has been received from peripherals – linked to the control unit (CU).

As well as four registers involved in the FDE cycle:

- The Current Instruction Register (CIR)
- The Program Counter (PC)
- The Memory Data/Buffer Register (MDR / MBR)
- The Memory Address Register (MAR)

Some Interrupts

vector number	description
0	divide error
1	debug exception
2	null interrupt
3	breakpoint
4	INTO-detected overflow
5	bound range exception
6	invalid opcode
7	device not available
8	double fault
9	coprocessor segment overrun (reserved)
10	invalid task state segment
11	segment not present
12	stack fault
13	general protection
14	page fault
15	(Intel reserved, do not use)
16	floating-point error
17	alignment check
18	machine check
19–31	(Intel reserved, do not use)
32–255	maskable interrupts

Cache

Cache is a software or hardware component of a computer that sits between the processor and RAM. It stores results of previous computations or recently retrieved data from RAM. This means that when data requested from RAM is already held in cache, it can be returned more quickly. Thus the more requests that can be served from the cache, the faster the system performs.

Speeds of Memory

- Cache ~ 1ns
- RAM ~ 100ns
- HDD ~ 10ms
- SSD ~ 10µs

Factors affecting processor performance:

• Multiple Cores – computational work load is shared between different separate processors so multiple operations can be completed faster

• Cache Memory – more of this faster memory (1ns vs 100ns for RAM) means temporary data can be stored between calculations.

• Clock Speed – operations only happen on the edge of the clock signal. So long as the physical hardware can handle it (EM interference needs time to settle), a faster clock speed means more computations can be handled in a given amount of time.

• Word Length – a larger word size means more data can be processed at once so less computations are required for large operations.

• Address Bus Width – allows for more memory to be addressed to so more RAM can be worked with.

• Data Bus Width – this must be at least as big as word length. If it is bigger however, then more data packets can be sent in a given time to the processor as they could, in theory be sent in parallel. It is especially useful for multi-core processors.