WALCHAND COLLEGE OF ENGINEERING

(Government Aided Autonomous Institute)
Visharambag, Sangli – 416415
Second Year B.Tech. Information Technology
ESE , EVEN SEMESTER, AY 2022-23
Computer Architecture (6IT222)

MODEL ANSWER

Q1. A) Steps to fetch a word from memory

To fetch a word of information from memory, the processor has to specify the address of the memory location where this information is stored and request a Read operation. It has four control signals: MDR_{in} and MDR_{out}, control the connection to the internal bus, and MDR_{inE} and MDR_{outE} control the connection to the external bus. These actions may be carried out as separate steps, but some can be combined into a single step. Each action can be completed in one clock cycle, except action 3 which requires one or more clock cycles, depending on the speed of the addressed device.
The memory read operation requires three steps, which can be described by the signals being activated as follows: 1. R_{lout}, MAR_{in}, Read , 2. MDR_{inE}, WMFC , 3. MDR_{out}, $R2_{in}$
where WMFC is the control signal that causes the processor's control circuitry to wait for the arrival of the MFC signal.

Q1. B) Significance of locality of reference in Cache Memory.

The references to memory at any given time interval tend to be confined within a localized areas. Analysis of programs shows that most of their execution time is spent on routines in which instructions are executed repeatedly. Locality of reference is manifested in two ways :
1. Temporal- means that a recently executed instruction is likely to be executed again very soon. The information which will be used in near future is likely to be in use already (e.g. reuse of information in loops)
2. Spatial- means that instructions in close proximity to a recently executed instruction are also likely to be executed soon
If a word is accessed, adjacent (near) words are likely to be accessed soon (e.g. related data items (arrays) are usually stored together; instructions are executed sequentially)
If active segments of a program can be placed in a fast (cache) memory , then total execution time can be reduced significantly
Temporal Locality of Reference suggests whenever an information (instruction or data) is needed first , this item should be brought in to cache
Spatial aspect of Locality of Reference suggests that instead of bringing just one item from the main memory to the cache ,it is wise to bring several items that reside at adjacent addresses as well (i.e. a block of information)

Q1. C)Addressing Modes

1] Register mode
2] Absolute mode
3] Immediate mode
4] Indirect adressing mode

Q2 A) Actions needed to execute instruction Move (R1),R2.

1. MAR [RI]
2. Start a Read operation on the memory bus

3. Wait for the MFC(Memory Function Completed) response from the memory
4. Load MDR from the memory bus
5. R2 [MDR]

Q2. B) Differentiate between Isolated I/O and Memory Mapped I/O.

Isolated I/O : The I/O in which one common bus is used for memory and I/O but there are separate read and write controls for I/O and memory transfer is called isolated I/O. This configuration isolates all I/O interface address from the address assigned to memory. Therefore, this method is called isolated I/O. The I/O read and I/O write control lines are enabled during an I/O transfer. The memory read and memory write control lines are enabled during a memory transfer.

Memory Mapped I/O: Same address can be used for either memory and I/O transfer, only control line identifies whether the transfer is I/O or memory. The I/O in which one common bus is used for memory and I/O bus with common control lines is called memory mapped I/O. It uses common read/write instruction for memory and I/O operation. It uses same address space for both memory and I/O and treats interface register as a part of memory. Same instructions are used for memory and I/O. No need of separate control lines for I/O and memory operation.

Q2. C) Control Word in Micro program.

It is a word whose individual bits represent various control signals in processor
Each of control step in control sequence defines unique combination of 1s and 0s in CW
A sequence of CWs corresponding to the control sequence of machine instruction constitutes microroutine for that instruction. And individual control words in this micro routine are referred to as microinstructions. Control unit generates control signals by reading sequentially CWs of corresponding microroutine from control store.

Q2. D) Factors which affect performance of pipeline in processor.

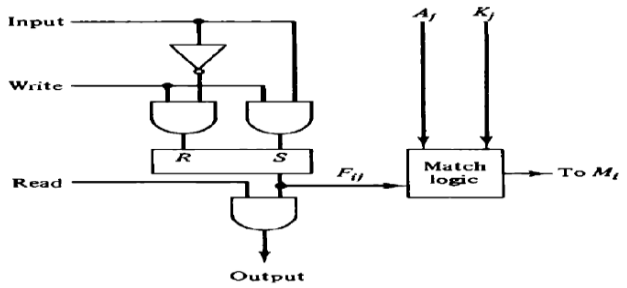
1. Timing Variations
All stages cannot take same amount of time. This problem generally occurs in instruction processing where different instructions have different operand requirements and thus different processing time.
2. Data Hazards
When several instructions are in partial execution, and if they reference same data then the problem arises. We must ensure that next instruction does not attempt to access data before the current instruction, because this will lead to incorrect results.
3. Branching
In order to fetch and execute the next instruction, we must know what that instruction is. If the present instruction is a conditional branch, and its result will lead us to the next instruction, then the next instruction may not be known until the current one is processed.
4. Interrupts
Interrupts set unwanted instruction into the instruction stream. Interrupts effect the execution of instruction.
5. Data Dependency
It arises when an instruction depends upon the result of a previous instruction but this result is not yet available.

Q3 A) Technique used in Memory Organization to achieve parallelism.

Another approach is to interleave the memory, or split it into "banks" that can be accessed individually. The main benefit is overlapping the latencies of accessing each word. For example, if our main memory has four banks, each one word wide, then we could load four words into a cache block in just 20 cycles. 1 + 15 + (4 x 1) = 20 cycles Our buses are still one word wide here, so four cycles are needed to transfer data to the caches. This is cheaper than implementing a four-word bus, but not too much slower. The magenta cycles represent sending an address to a memory bank.

Each memory bank has a 15-cycle latency, and it takes another cycle (shown in blue) to return data from the memory. This is the same basic idea as pipelining!
As soon as we request data from one memory bank, we can go ahead and request data from another bank as well. Each individual load takes 17 clock cycles, but four overlapped loads require just 20 cycles.

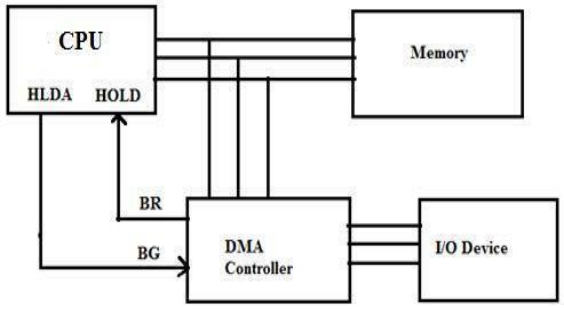
Q3 B) Single cell of Associative memory.



Q3 C) Concept of forwarding used to avoid data hazard

Hardware detects RAW and stalls Assumes register written then read each cycle, + low cost to implement, simple, -- reduces IPC Try to minimize stalls, forward data before it is in the register, + reduces/avoids stalls -- complex, Crucial for common RAW hazards Forwarding is the concept of making data available to the input of the ALU for subsequent instructions, even though the generating instruction hasn't gotten to WB in order to write the memory or registers.

Q3 D) Block diagram of DMA controller and its operation using control signals.



In this transfer mode, DMA controller sends Bus Request (BR) to CPU enabling BR line.
After receiving BR request, CPU acknowledges the DMA controller by sending Read/Write control signal, Device address, Starting address of memory block for data and amount of data to be transferred and disables its control over system bus by issuing Bus Grant (BG) signal to DMA controller. And CPU does other works.
DMA controller then start transfer of data between memory and I/O. After completing all transfer of data, the controller sends interrupt to CPU informing that it has completed the job assigned to it. The CPU finalizes the DMA operation and resumes it operation.

DMA is a sophisticated I/O technique in which a DMA controller replaces the CPU and takes care of the access of both, the I/O device and memory, for fast data transfers.
Using DMA you get the fastest data transfer rates possible.
Momentum behind the DMA: interrupt-driven and Programmed I/O require active CPU intervention (All data must pass through CPU).
transfer rate is limited by processor's ability to service the device and hence CPU is tied up managing I/O transfer.
Removing CPU from the path and letting the peripheral device manage the memory buses directly would improve the speed of transfer.
Bus Request (BR): used by DMA controller to request the CPU for buses. When this input is active, CPU terminates the execution of the current instruction and places the address bus; data bus and read & write lines into high impedance state.
Bus Grant (BG): CPU activates BG output to inform DMA that buses are available (in high impedance state). DMA now take control over buses to conduct memory transfers without processor intervention. When DMA terminates the transfer, it disables the BR line and CPU disables BG and returns to normal operation.

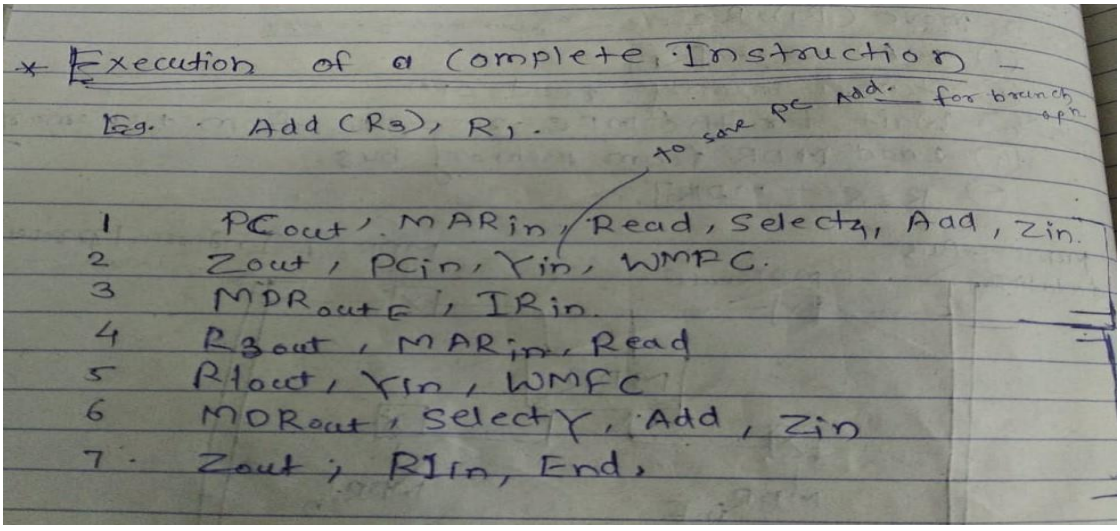
Q4. A) Guard bits: re used to store some intermediate results. Ex Addition and substraction.
Sometimes results may goes more than digits, so to store that result we require Guard bits.
Ex. 2.95 * 10^2
- 2.39 * 10^0

Step one check exponent and make it equal.
i.s. 0.0239 * 10^2
so now significant bits are 002 so what about 39. So here to store 39 guard bits are stored.
. 2.95 * 10^2
- 0.0239 * 10^2

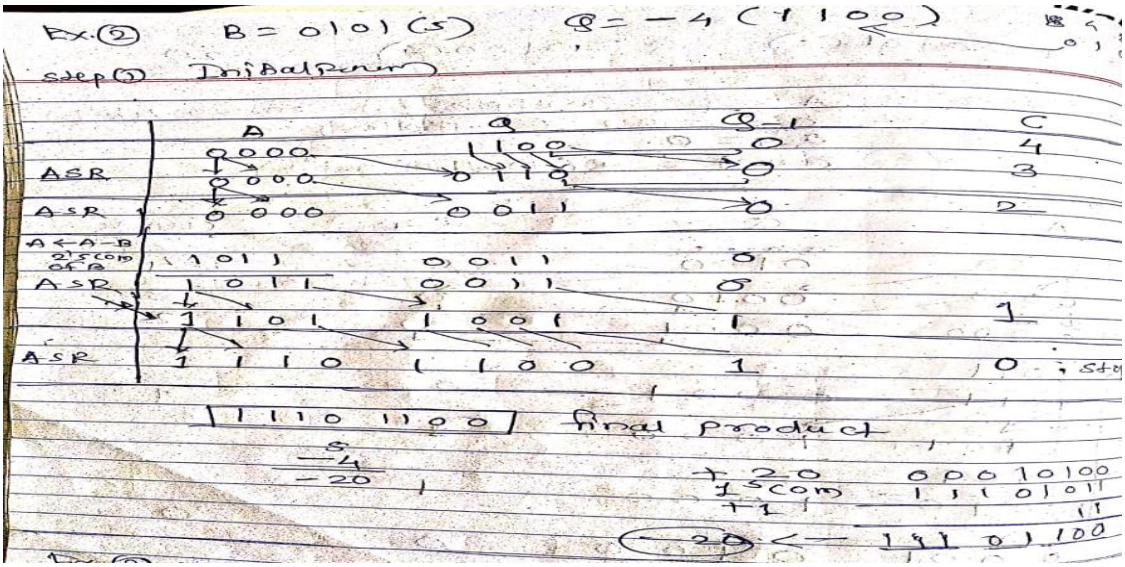
2.9261 * 10^2

Truncating: 1] Chopping : simply cut guard bits.
2] Rounding : Add (R3), R1

Q4. B)



Q4.C)


Ex. (2) B = 0101 (5) Q = −4 (1100)
Step (1) Initialization
A
Q
Q−1
C
ASR
ASR
A ← A − B
2's com of B
ASR
ASR
1 1 1 0 1 1 0 0 final product
−5
−20
+20
1's com
+1
−20 ← 1 1 1 0 1 1 0 0