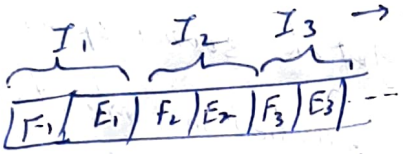
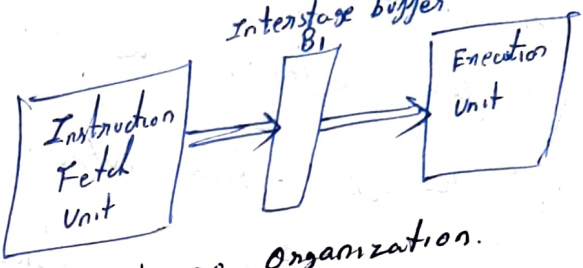
PIPELINING-

Pipelining is a particularly effective way of organizing concurrent activity in a computer system.

In this, the execution of Next instruction is started before completing the execution of current instruction. (ie) the Next instruction will be fetched while executing the current instruction.



Sequential execution.



a) Hardware Organization.

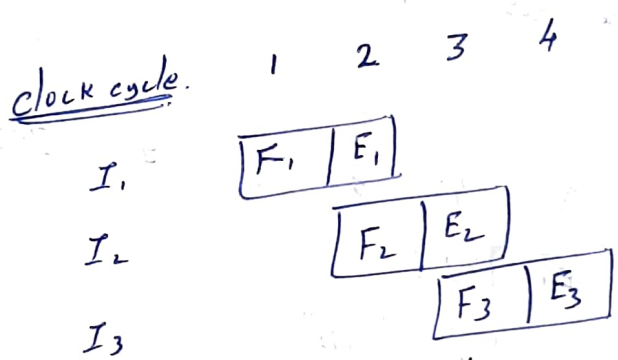


Fig 8.1
2 stage Pipelining.

b) Pipelined Execution.

It has 2 h/w (hardware) unit for Fetching & executing instructions seperately.

Thus the Fetched instruction (I_1) is stored in Interstage buffer & it will be used for execution. So, the Next instruction (I_2) can be fetched during its execution.

In this manner, both the fetch & execute units are kept busy all the time.

The completion rate of instruction execution will be twice that of Sequential operation.

* Interstage storage Buffer (B_i) ⇒ is needed to hold the information being passed from one stage to the next.

* New information is loaded into this buffer at the end of each clock cycle.

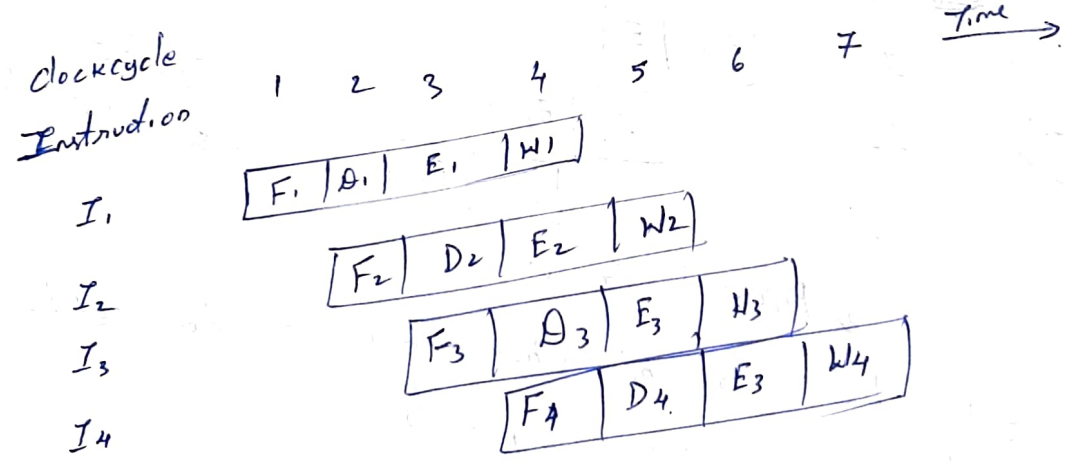
A Pipelined Processor may process each instruction in four steps, as Follows.

(F) Fetch - Read instruction from memory.

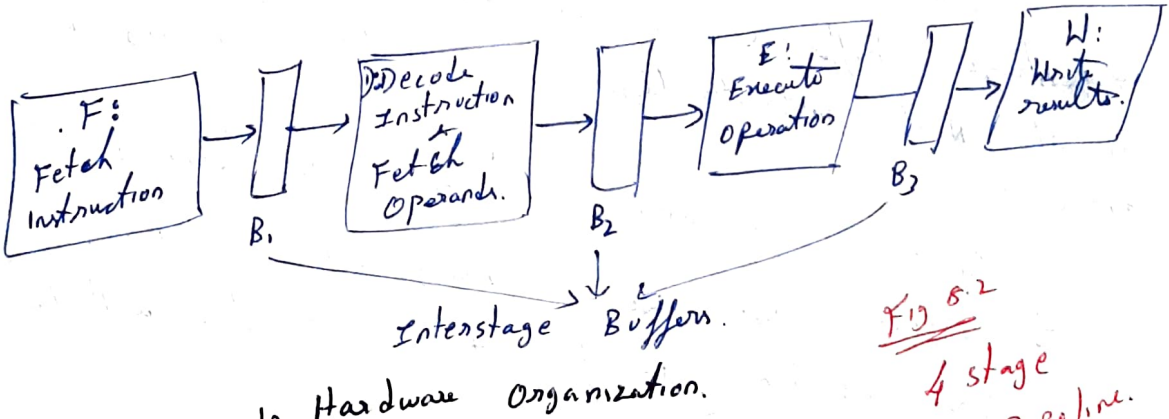
(D) Decode - Decode instruction & fetch source operand(s).

(E) Execute - Perform the operation specified in instruction.

(W) Write - Store result in the destination location.



a) Instruction Execution.



b) Hardware Organization.

Fig 8.2
4 stage pipeline.

Role of cache Memory -

Each stage in a pipeline is expected to complete its operation in one clock cycle. Hence, the clock period should be sufficiently long to complete the task being performed in any stage. If different units require different amounts of time, the clock period must allow the longest task to be completed. A unit that completes its task early is idle for the remainder of the clock period.

Hence, Pipelining is most effective in improving performance of the tasks being performed in different stages require about the same amount of time.

During fetch cycle, the access time for main memory will be greater compared to execution of other steps inside the processor. So, a cache is included to balance the fetch & processing steps to be more or less equal in duration.

Pipeline Performance -

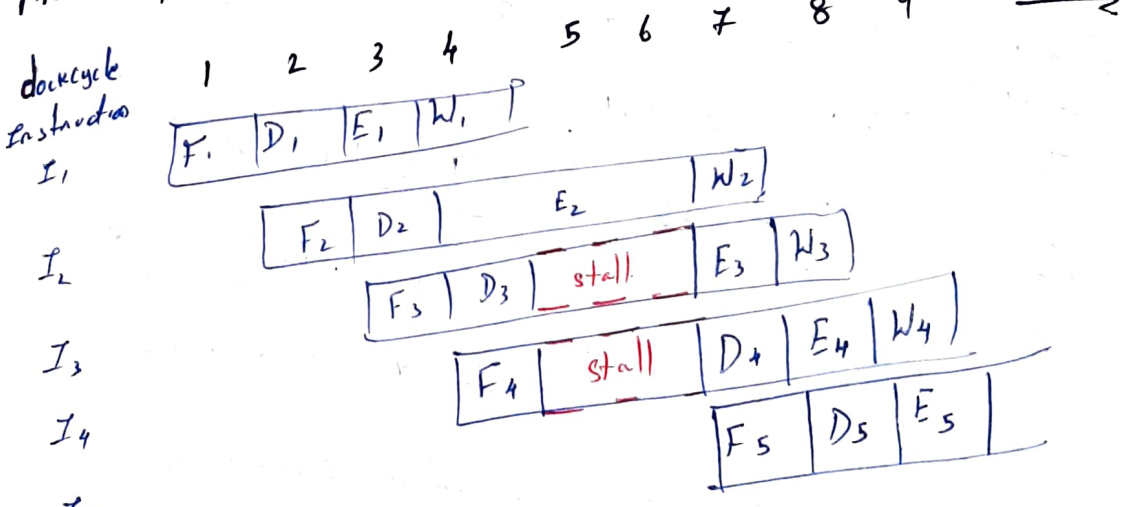


Fig 8.3 execution operation taking more than one clock cycle.

In the Fig 8.3 the pipelined operation is said to have been stalled for two clock cycles.

Any condition that causes the pipeline to stall is called as hazard.

1) Data Hazard -

It is the condition in which either the source or the destination operands of an instruction are not available at the time expected in the pipeline. As a result some operation has to be delayed and the pipeline stalls.

2) Instruction Hazards / Control Hazards -

The pipeline may stall because of a delay in availability of an instruction. This may be a result of a miss in the cache, requiring the instruction to be fetched from the main memory. Such hazards are often called control hazards or instruction hazards.

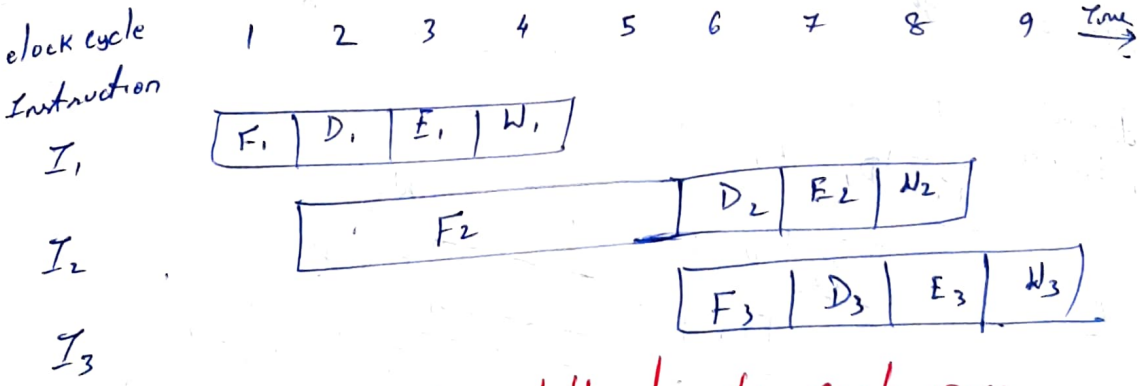


Fig 8.4 Pipeline stall due to cache miss.

a) Instruction Execution clock cycles.

clock cycle.	1	2	3	4	5	6	7	8	9 → Time
Stage.									
F: Fetch.	F_1	F_2	F_2	F_2	F_2	F_3			
D: Decode		D_1	idle	idle	idle	D_2	D_3		
E: Execute			E_1	idle	idle	idle	E_2	E_3	
W: Write				W_1	idle	idle	idle	W_2	W_3

b) Function performed by each processor in successive clock cycles.

The idle periods of the processor units are called stalls. These are often referred to as bubbles in the pipeline.

Once created bubble as a result of delay in one of the pipeline stages, moves downstream until it reaches the last unit.

3) Structural Hazard -

This is the situation when 2 instructions require to use the given hardware resource at the same time. The most common case for this hazard may arise in access to memory.

Many processor use separate instruction and data caches to avoid this delay.

Load x(R1), R2.