# Ch-7 Introduction to Pentium Architecture

Pentium logos


1993–2003 | 2003–2006 | 2006–2009 | 2009–2013 | 2013-2015 | 2015–2020 | 2020–2023

'entium processor family


Intel's new simplified branding. Image: Intel

# Features of Pentium Processor :

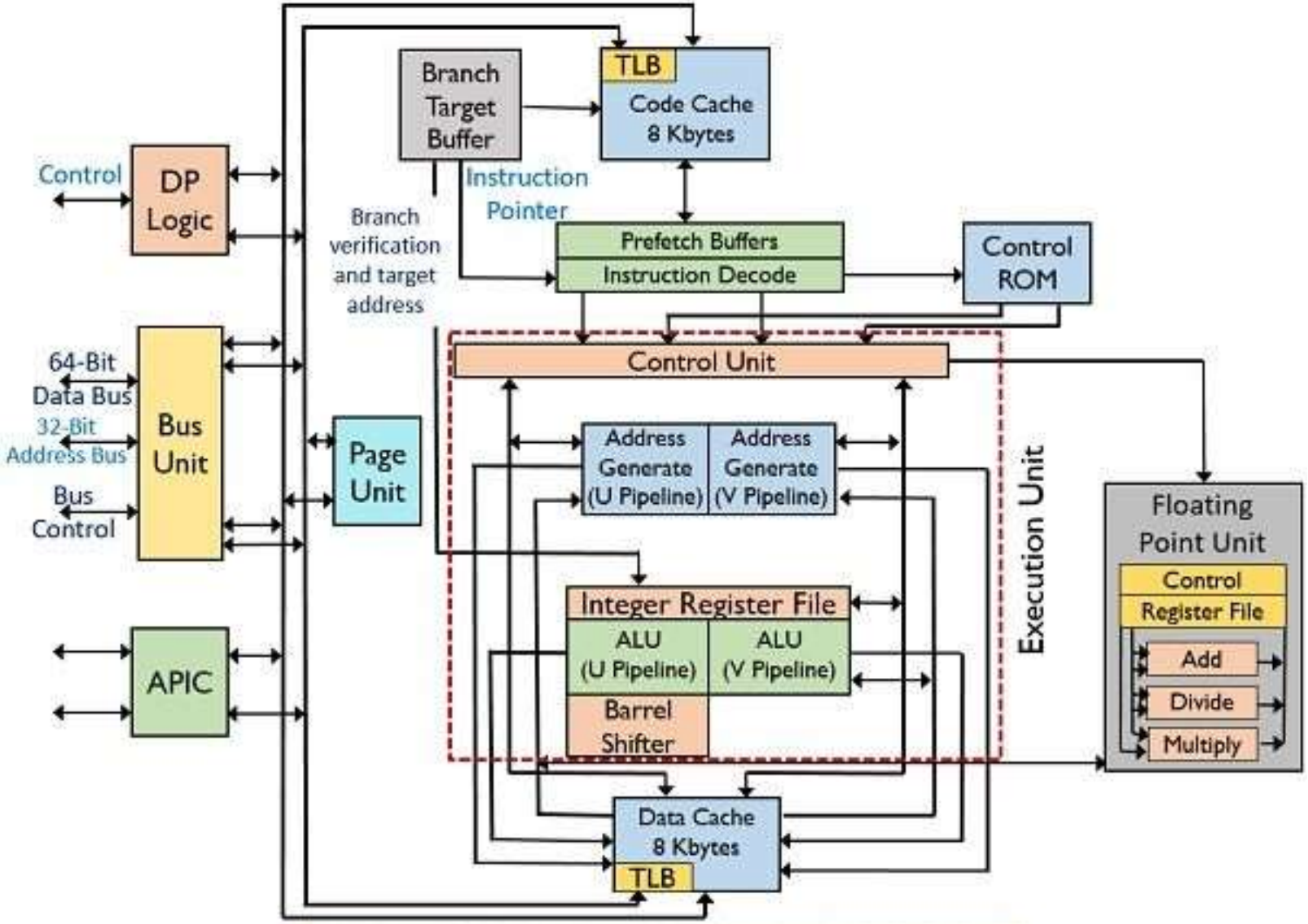
- Pentium Processor is **fifth generation** microprocessor evolved in 1993.
- It has a **32-bit address** bus
- Address bus is 32-bit hence can access **$2^{32}$ = 4GB memory**.
- It has **64-bit data** bus
- Pentium Processor designed to operate from **60MHZ to 233 MHZ**
- **It has 5 stage pipelining .**


- **It is 2 way superscalar : U pipe , V pipe**
- **On chip floating point unit**
- **Branch prediction logic : 256 entry BTB**
- **On chip cache memory**
- **Instruction cache** has **8K bytes** memory
- **Data cache** has **8K bytes** memory

- **Cache is** a small amount of memory which **is** a part of the **CPU** - closer to the **CPU** than RAM .
- It **is** used to temporarily hold **instructions** and **data** that the **CPU is** likely to **reuse**.

State any 3 / 4 features of Pentium processor. 3 / 4 Mks

# Pentium superscalar architecture :



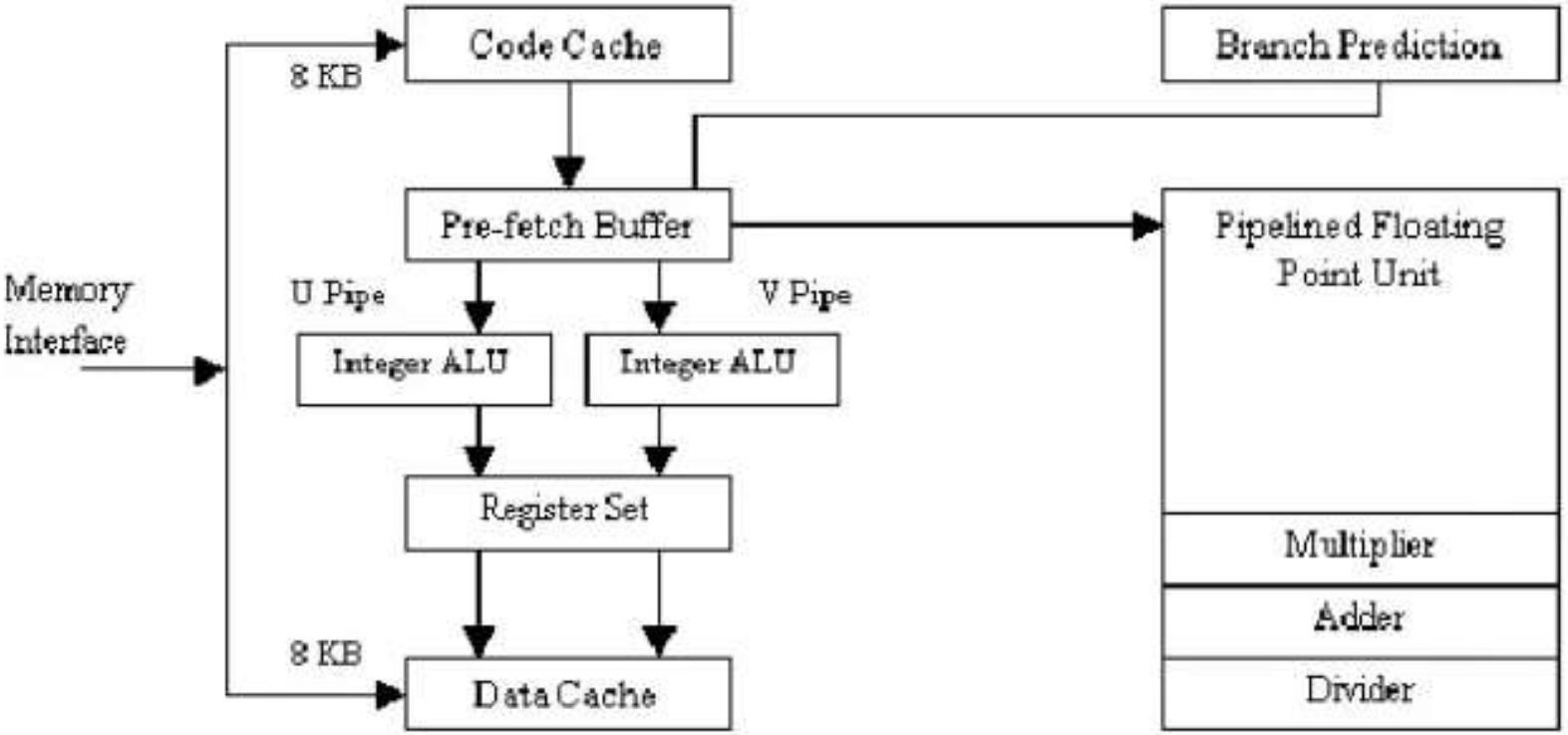

**Architecture of Pentium Processor**

- *The various functional units are as follows*:
- Bus unit
- Paging unit
- Control ROM
- Prefetch buffer
- Execution unit with two integer pipeline (U-pipe and V-pipe)
- Code cache
- Data cache
- Instruction decode
- Branch target buffer
- Dual processing logic
- Advanced programmable interrupt controller
- TLB (translational lookaside buffer)

- The bus unit of the architecture sends the control signal and fetches code and data from external memory and IO devices.
- The size of the external data bus is 64-bit through which burst read and burst write-back cycles can be achieved.
- The paging unit in the architecture provides optional extensions of around 2 to 4 Mb page sizes.
- In order to load the instructions into the execution unit, code cache, branch target buffer and prefetch buffers operate together.
- The code cache or the external memory holds the instructions from where these are fetched. While the branch target buffer holds the address of the respective branch and the TLB (translational lookaside buffer) within the code cache converts the linear address into the physical address that is used by the code cache.

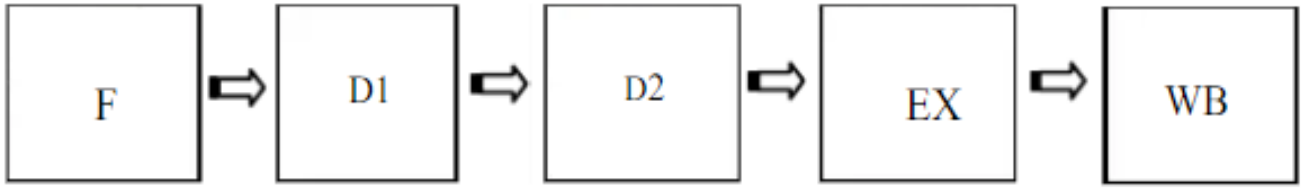
- This processor contains pairs of prefetch buffers having a size of 32-byte that combinedly operate with branch target buffer. Both the buffers operate independently but not at the same time.
- The execution unit within the Pentium microprocessor contains two integer pipelines namely U-pipe and V-pipe and each one has its separate ALU. There are five stages in which these pipelines operate, namely, prefetch, decode-1, decode-2, execute, writeback. The U-pipe is responsible for executing all integer as well as floating-point instructions while V-pipe executes simple integer and some floating-point instructions.

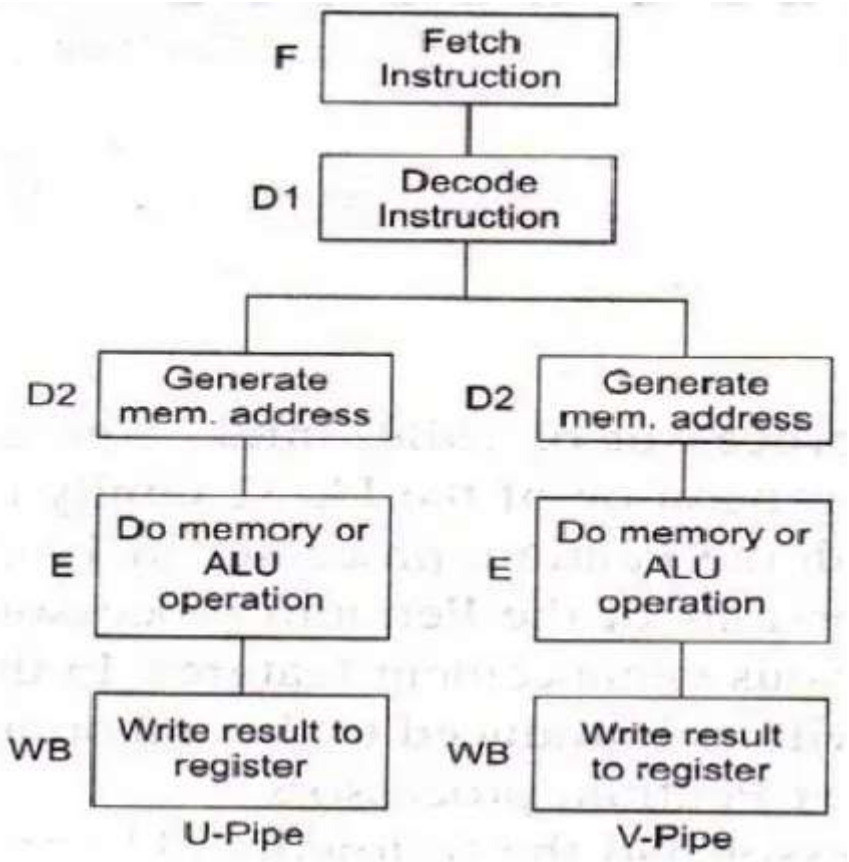
# Pentium superscalar architecture :



- Processors capable of parallel instruction execution of multiple instructions are known as Superscalar.
- The Pentium processor is a superscalar machine, capable of executing two instructions in parallel.
- The process of issuing two instructions in parallel is called as dual pairing.
- Given figure presents a block diagram overview of the Pentium Processor including two instruction pipelines because of dual pipelining feature.
  1. "U" pipe : The U-pipe can execute all integer and floating point instructions.
  2. "V" pipe : The V-pipe can execute simple integer instructions and floating point instructions.
- There are two separate cashes are there in the block diagram :
  1. Instruction cache (code cache)
  2. Data cache
- There are 2 separate 32 bit ALU to perform 64 bit data operation.
- 5-stage or dual pipelining and branch prediction are two advance and important features of Pentium processor.

# Pipelining :




**The Pipeline Stages**

F - Instruction Fetch Stage
D1 - Decode Stage 1
D2 - Decode Stage 2
E - Execute Memory or ALU
WB - Write back to destination

# Pipelining :

Each of the two pipelines have five stages:

1. **Prefetch (F) :** Instructions are prefetched from instruction cache or memory.
2. **D1 instruction decode :** In the D1 stage, the processor decodes the instruction to generate a control word. A single control word executes instruction directly.
3. **D2 Address generate :** In the D2 stage, the processor decodes the control word from D1 for use in the EX stage.
4. **Execute (EX) :** In this stage, the processor either accesses the data cache or calculates results in the ALU.
5. **Write Back (WB) :** In the WB stage, the processor updates the registers and flags with the instructions result.

# Instruction Branch Prediction:

**Why do we need branch prediction?**

1. The gain produced by Pipelining can be reduced by the presence of program transfer instructions eg JMP, CALL, RET etc .

2. They change the sequence causing all the instructions that entered the pipeline after program transfer instructions invalid.

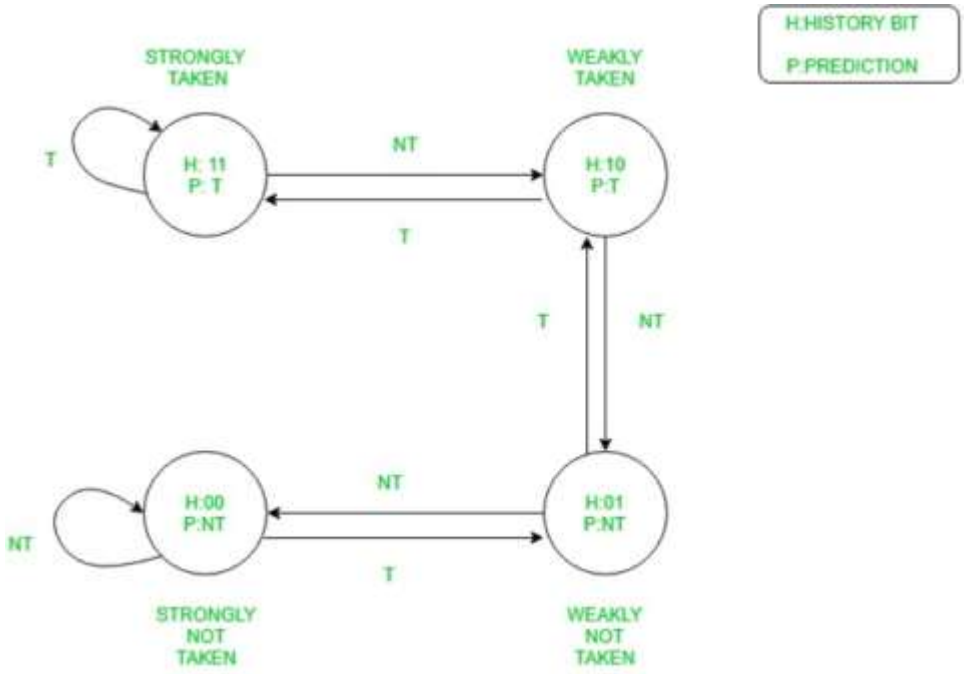
3. Thus no work is done as the pipeline stages are reloaded.

**Branch prediction logic:**

- To avoid this problem, Pentium uses a scheme called Dynamic Branch Prediction.
- In this scheme, a prediction is made for the branch instruction currently in the pipeline. The prediction will either be taken or not taken.
- If the prediction is true then the pipeline will not be flushed and no clock cycles will be lost.
- If the prediction is false then the pipeline is flushed and starts over with the current instruction.
- It is implemented using 4 way set associated cache with 256 entries.
- This is called **Branch Target Buffer (BTB)**.
- The directory entry for each line consists of:
  **Valid bit:** Indicates whether the entry is valid or not.
  **History bit:** Track how often bit has been taken.
- Source memory address is from where the branch instruction was fetched. If the directory entry is valid then the target address of the branch is stored in corresponding data entry in BTB.

**Working of Branch Prediction:**

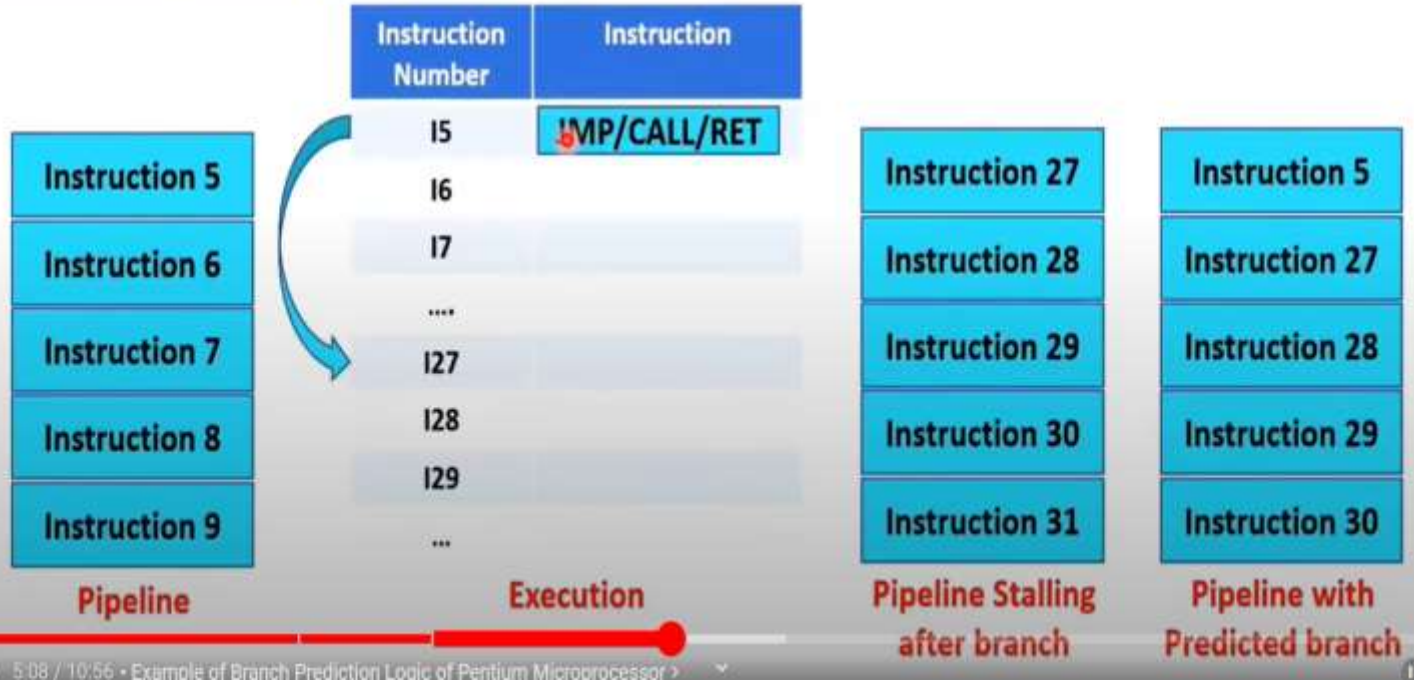
- BTB is a lookaside cache that sits to the side of Decode Instruction(DI) stage of 2 pipelines and monitors for branch instructions.
- The first time that a branch instruction enters the pipeline, the BTB uses its source memory to perform a lookup in the cache.
- Since the instruction was never seen before, it is BTB miss. It predicts that the branch will not be taken even though it is unconditional jump instruction.
- When the instruction reaches the EU(execution unit), the branch will either be taken or not taken.
- If taken, the next instruction to be executed will be fetched from the branch target address.
- If not taken, there will be a sequential fetch of instructions.
- When a branch is taken for the first time, the execution unit provides feedback to the branch prediction.
- The branch target address is sent back which is recorded in BTB.
- A directory entry is made containing the source memory address and history bit is set as strongly taken.




| History Bits | Resulting Description | Prediction made | If branch taken | If branch not taken |
|---|---|---|---|---|
| 11 | Strongly Taken | Branch Taken | Remains in same state | Downgraded to weakly taken |
| 10 | Weakly Taken | Branch Taken | Upgraded to strongly taken | Downgraded to weakly not taken |
| 01 | Weakly Not Taken | Branch Not Taken | Upgraded to weakly taken | Downgraded to strongly not taken |
| 00 | Strongly Not Taken | Branch Not Taken | Upgraded to weakly not taken | Remains in same state |

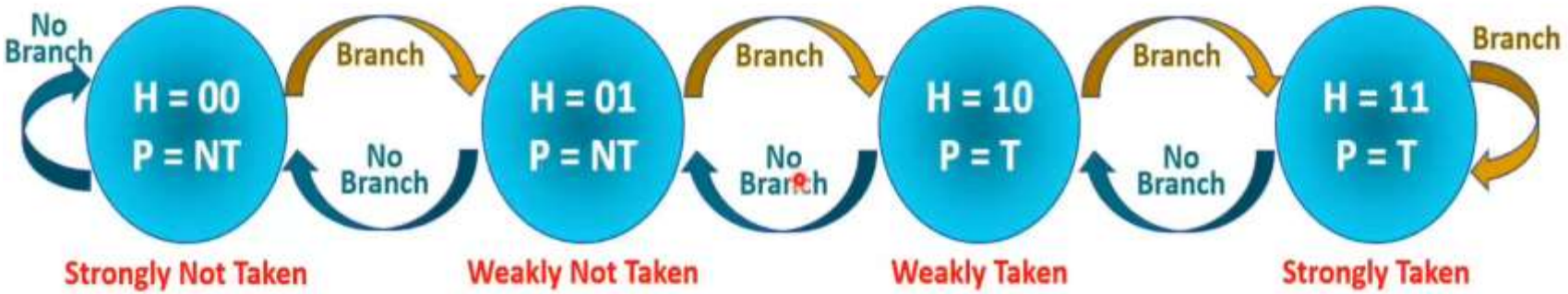
# Branch Prediction Logic of Pentium Microprocessor

- Branch Prediction Unit:
- Branch Prediction Unit plays essential role in the prediction of branch.
- Branch instruction causes pipeline stalling. {Pipeline flushing due to branch instruction, so some clock cycles will be wasted.}
- Frequent pipeline stalling degrades performance of execution.
- So by branch prediction we can predict branch instruction in advance and It gives commands to prefetch buffer regarding branch instructions {CALL, JMP & RET}.



# Branch Prediction Logic

- Pentium uses a scheme called Dynamic Branch Prediction. In this scheme, a prediction is made for the branch instruction currently in the pipeline.
- If the prediction is true then the pipeline will not be flushed and no clock cycles will be lost. If the prediction is false then the pipeline is flushed and starts over with the current instruction.
- It is implemented using 4 way set associated memory with 256 entries. This is called Branch Target Buffer (BTB).
- Here, 4 way set is defined by history bits. Based on history bits prediction is fixed.




| History Bits | Branch Prediction |
|---|---|
| 00 | Not Taken |
| 01 | Not Taken |
| 10 | Taken |
| 11 | Taken |

- For H bits = 00 or 01 branch prediction is not taken and For H bits = 10 and 11 branch prediction is taken.
- For false prediction, Pipeline will gets flushed.
- If branch is executed then H bits data will be incremented by one, it means branch prediction priority will increase.
- If branch is not executed then H bits data will be decremented by one, it means branch prediction priority will decrease.