Cache Enhancement

In computing, a cache is a hardware or software component that stores data so that future requests for that data can be served faster; the data stored in a cache might be the result of an earlier computation or a copy of data stored elsewhere.

Cache helps in reducing latency. In computing, memory **latency** is the time (the **latency**) between initiating a request for a byte or word in memory until it is retrieved by a processor. If the data are not in the processor's **cache**, it takes longer to obtain them, as the processor will have to communicate with the external memory cells.

Therefore the purpose of undertaking specifically this project is to improve cache performance and therefore analyse the means to do so which includes pre-fetching techniques and modifying the memory address index.

- The various prefetching techniques that we analyse are:-
    a) **Hardware based prefetching** is typically accomplished by having a dedicated hardware mechanism in the processor that watches the stream of instructions or data being requested by the executing program, recognizes the next few elements that the program might need based on this stream and prefetches into the processor's cache.
        1) **Strided Prefetching** pattern of prefetching instructions is to prefetch addresses that are s addresses ahead in the sequence. It is mainly used when the consecutive blocks that are to be prefetched are s addresses apart.
        2) **Sequential prefetch** is a mechanism that reads consecutive pages into the buffer pool before the pages are required by the application.
- Also several memory index modification techniques will be implemented to reduce line conflicts.

Cache enhancement is an industrial research field that highly borrows from the CSN-221 course that we are pursuing.