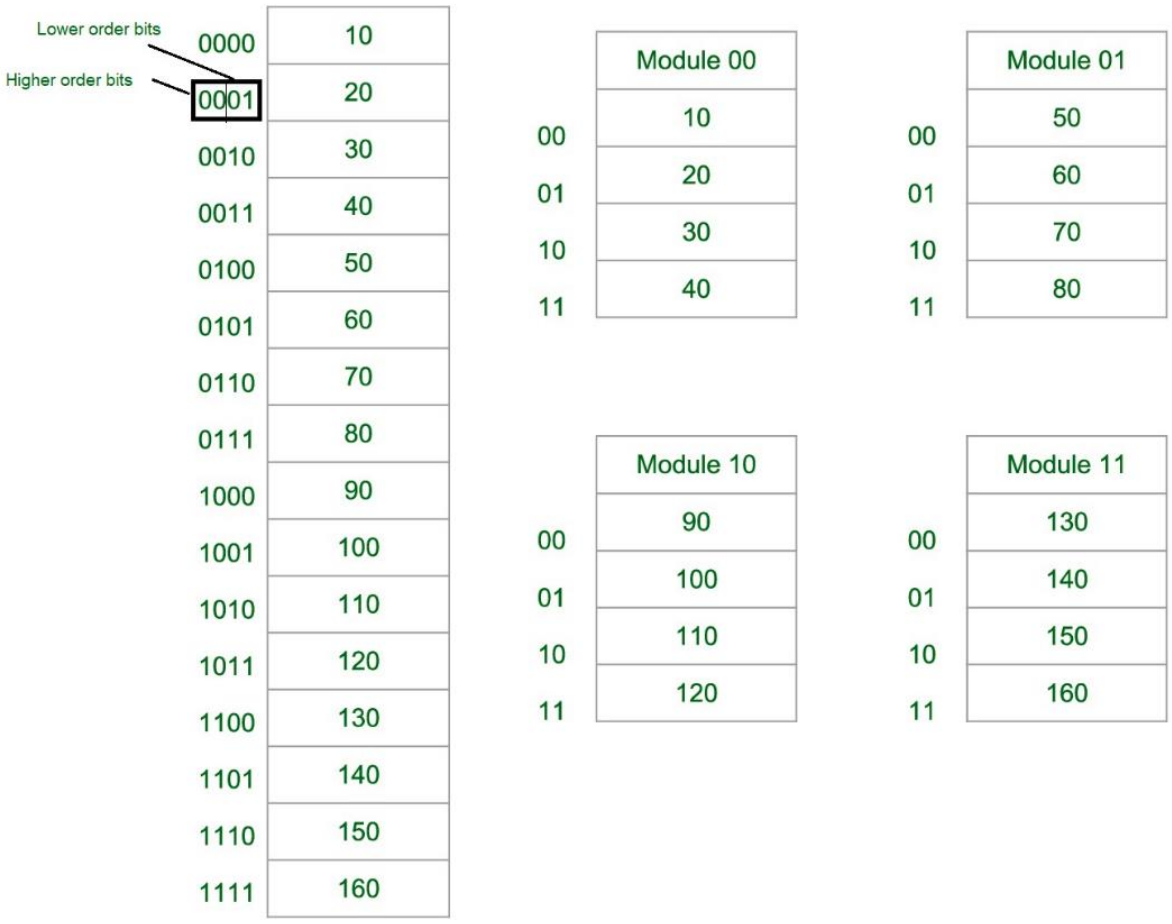
Interleaved Memory

Abstraction is one of the most important aspects of computing. It is a widely implemented Practice in the Computational field.

Memory Interleaving is less or More an Abstraction technique. Though it's a bit different from Abstraction. It is a Technique that divides memory into a number of modules such that Successive words in the address space are placed in the Different modules.

Consecutive Word in a Module:
Figure-1: Consecutive Word in a Module



Let us assume 16 Data's to be Transferred to the Four Module. Where Module 00 be Module 1, Module 01 be Module 2, Module 10 be Module 3 & Module 11 be Module 4. Also, 10, 20, 30….130 are the data to be transferred.

From the figure above in Module 1, 10 [Data] is transferred then 20, 30 & finally, 40 which are the Data. That means the data are added consecutively in the Module till its max capacity.

Most significant bit (MSB) provides the Address of the Module & the least significant bit (LSB) provides the address of the data in the module.

For **Example**, to get 90 (Data) 1000 will be provided by the processor. This 10 will indicate that the data is in module 10 (module 3) & 00 is the address of 90 in Module 10 (module 3). So,

Module 1 Contains Data : 10, 20, 30, 40

Module 2 Contains Data : 50, 60, 70, 80

Module 3 Contains Data : 90, 100, 110, 120

Module 4 Contains Data : 130, 140, 150, 160

Consecutive Word in Consecutive Module:

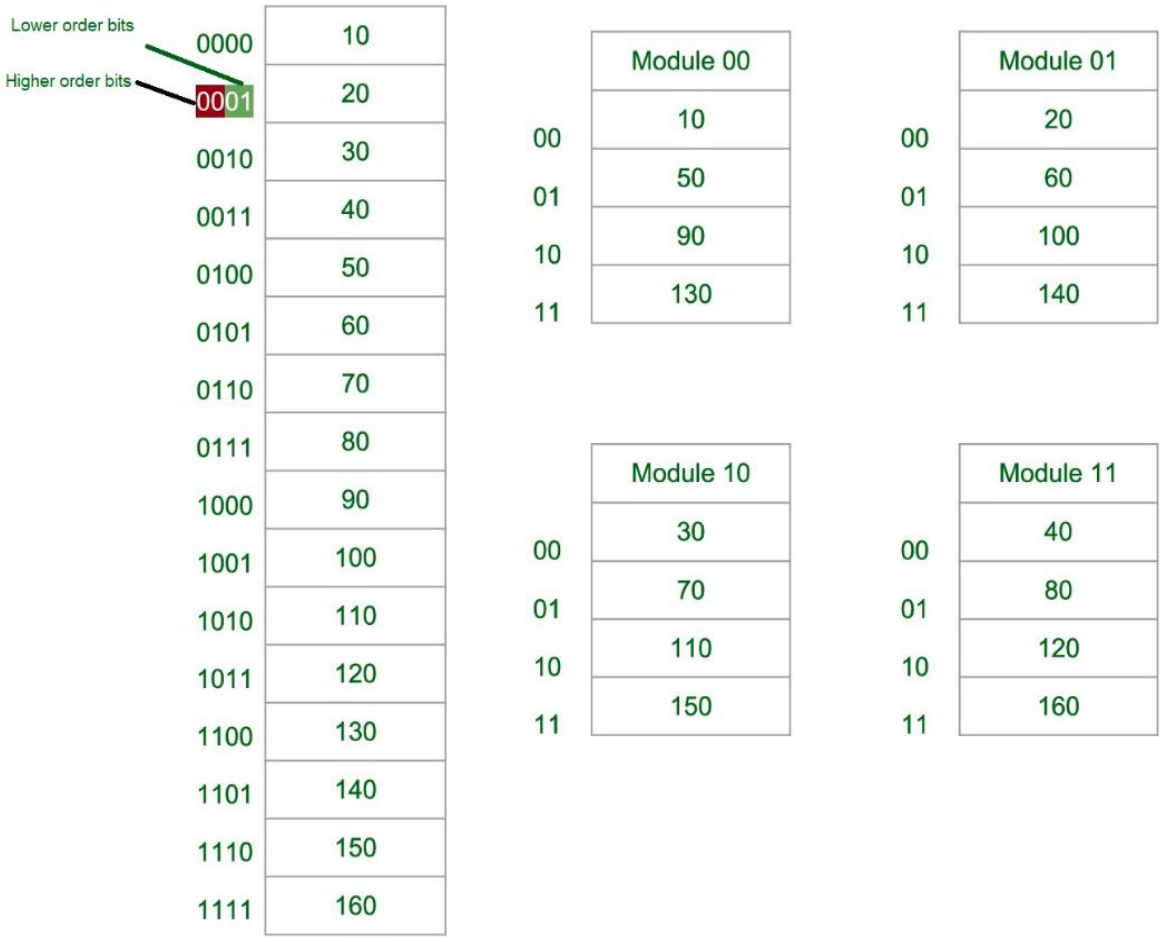


Figure-2: Consecutive Word in Consecutive Module

Now again we assume 16 Data's to be transferred to the Four Module. But Now the consecutive Data are added in Consecutive Module. That is, 10 [Data] is added in Module 1, 20 [Data] in Module 2 and So on.

Least Significant Bit (LSB) provides the Address of the Module & Most significant bit (MSB) provides the address of the data in the module.

For **Example**, to get 90 (Data) 1000 will be provided by the processor. This 00 will indicate that the data is in module 00 (module 1) & 10 is the address of 90 in Module 00 (module 1). That is,

Module 1 Contains Data : 10, 50, 90, 130

Module 2 Contains Data : 20, 60, 100, 140

Module 3 Contains Data : 30, 70, 110, 150

Module 4 Contains Data : 40, 80, 120, 160

Why do we use Memory Interleaving? [Advantages]:
Whenever Processor requests Data from the main memory. A block (chunk) of Data is Transferred to the cache and then to Processor. So whenever a cache miss occurs the Data is to be fetched from the main memory. But main memory is relatively slower than the cache. So to improve the access time of the main memory interleaving is used.
We can access all four Modules at the same time thus achieving Parallelism. From Figure 2 the data can be acquired from the Module using the Higher bits. This method Uses memory effectively.

Types of Interleaved Memory

In an operating system, there are two types of interleaved memory, such as:

1. High order interleaving: In high order memory interleaving, the most significant bits of the memory address decides memory banks where a particular location resides. But, in low order interleaving the least significant bits of the memory address decides the memory banks.

The least significant bits are sent as addresses to each chip. One problem is that consecutive addresses tend to be in the same chip. The maximum rate of data transfer is limited by the memory cycle time. It is also known as ***Memory Banking***.

Address bits	25-22	21-0
Use	Bank Select	Address to the chip

Module 0	Module 1	Module 2	Module 3	Module 4	Module 5	Module 6	Module 7
0	4	8	12	16	20	24	28
1	5	9	13	17	21	25	29
2	6	10	14	18	22	26	30
3	7	11	15	19	23	27	31

2. Low order interleaving: The least significant bits select the memory bank (module) in low-order interleaving. In this, consecutive memory addresses are in different memory modules, allowing memory access faster than the cycle time.

Address bits	25-4	3-0
Use	Address to the chip	Bank Select

Module 0	Module 1	Module 2	Module 3	Module 4	Module 5	Module 6	Module 7
0	1	2	3	4	5	6	7
8	9	10	11	12	13	14	15
16	17	18	19	20	21	22	23
24	25	26	27	28	29	30	31

Benefits of Interleaved Memory

An instruction pipeline may require instruction and operands both at the same time from main memory, which is not possible in the traditional method of memory access. Similarly, an arithmetic pipeline requires two operands to be fetched simultaneously from the main memory. So, to overcome this problem, memory interleaving comes to resolve this.

- It allows simultaneous access to different modules of memory. The modular memory technique allows the CPU to initiate memory access with one module while others are

busy with the CPU in reading or write operations. So, we can say interleave memory honors every memory request independent of the state of the other modules.

- So, for this obvious reason, interleave memory makes a system more responsive and fast than non-interleaving. Additionally, with simultaneous memory access, the CPU processing time also decreases and increasing throughput. Interleave memory is useful in the system with pipelining and vector processing.
- In an interleaved memory, consecutive memory addresses are spread across different memory modules. Say, in a byte-addressable 4 way interleave memory, if byte 0 is in the first module, then byte 1 will be in the 2nd module, byte 2 will be in the 3rd module, byte 3 will be in the 4th module, and again byte 4 will fall in the first module, and this goes on.
- An n-way interleaved memory where main memory is divided into n-banks and system can access n operands/instruction simultaneously from n different memory banks. This kind of memory access can reduce the memory access time by a factor close to the number of memory banks. In this memory interleaving memory location, i can be found in bank **i mod n**.