# Assignment (Video- 10 to 12): Database System Impl. (COP6726)

**Name:** Vikas Chaubey**, UFID:** 3511 5826**, Email:** vikas.chaubey@ufl.edu

1) **NUMA Architecture:** NUMA is Non-uniform memory access. This is a memory architecture model. This memory model is commonly used in computer systems which have more umber of processors units. NUMA memory architecture provides faster access to the memory for processor compared to all other traditional memory models. In this memory model a processor can access memory available to itself which is called local memory in much faster manner compared to the memory available non locally. NUMA architecture assigns a specific local memory to each processor for its local use, this memory is not shared by other processing units or processors. In this way when any data is processed it could be kept in local memory , hence the processor does not need to access non local memory to access data this has two benefits first one is that memory access is fast and second advantage is that the main memory accesses by processor decreases which results into lesser number of main memory access requests to main memory. This reduces multiple access requests reaching the main memory at the same time and hence the processing delays are also reduced. This architecture can significantly improve performance of the system compared to shared memory architecture , in case of shared memory the access times are longer this is because all processor units have to access the same memory again and again when they require data , this increases the chances of several access requests reaching the memory at the same time and hence this results into associated delays because not all the request can be served at the same time , there is always a limit on concurrent memory access. In NUMA architecture this issue is resolved by using local processor memory hence performance is improved. NUMA architecture was developed largely due to the advent of modern microprocessors that are faster than memory speeds. As a result, local memory is placed closer to the processor, shortening signal path lengths and reducing latency times with fewer delays in accessing memory.

2) **Mechanical Disc:** A mechanical disk commonly known as a hard disk drive (HDD) in computer systems is an electro-mechanical data storage device. These hard disk devices are nonvolatile memory that means they can retain stored data even after the power is turned off. The hard disk drive is the mass storage device used in different computer system to store data. Hard disk drives are mechanical in nature that means they store digital , it has a rotating platter which is magnetic , the data is written and read from these rotating magnetic platters using a magnetic head , these magnetic heads are actually arranged in the disk using a moving actuator arm. The memory head writes and reads data in random access manner in forms of memory blocks on the magnetic memory. This memory was introduced in 1956 by IBM and at that time it was main dominant secondary storage for all general-purpose computers at that time. Mechanical memory storage is the cheapest form of storage available in the market compared to other memory types such as SSD memories. SSD memories use NAND flash memory. Hence SSDs are comparatively faster compared to mechanical discs.

The primary advantage of the HDDs is its capacity and its performance. HDDs can have capacity from many GBs to many Terabytes, no other memory can provide such capacity in current scenario

3) **External Sort:** when the large amount of unordered data needs to be sorted then external sorting is used. when the amount of data which needs sorting is huge then in that case it is not possible to import data on the main memory and use traditional sorting techniques to sort it. It is because main memory has limited storage capacity. In that case to do sorting a different technique is required. This technique is called external sorting because data is kept in the hard disk and chunks of data are moved to main memory to perform sorting operation. External sorting uses a hybrid sort merge strategy. In the sorting phase, chunks of data small enough to fit in main memory are read, sorted, and written out to a temporary file. In the merge phase, the sorted sub-files are combined into a single larger file. for example external merge sort algorithm , first divides the whole file data into small chunks which are small enough so that they can fit on the main memory , these chunks are called runs , these runs are sorted individually and they when all runs are sorted these are merged together to form bigger runs , this is done until the whole file is sorted. External sorting algorithms generally fall into two types, distribution sorting, which resembles quicksort, and external merge sort, which resembles merge sort.

    a) **Distribution Sort (Resembles Quick Sort):** the size of working memory area is M. this is the size of the element. additional memory is required to read in the element without overwriting the buffer. the M/2 first and last elements are read into a buffer here the buffer acts like a pivot in case of quick sort, and these elements are sorted. then the next element is read either from beginning or the end to balance the writing. if the next element which is being read is less than the least of the elements present in the buffer then it is written to beginning in available space if the element is bigger than the greatest element in the buffer. If the next element is less than the least of the buffer, write it to available space at the beginning. If the read element is greater than the greatest element in buffer, then it is written to the end. Otherwise write the greatest or least of the buffer and put the next element in the buffer. Keep the maximum lower and minimum upper keys written to avoid resorting middle elements that are in order. When this is done, buffer is written. the smaller partitions are sorted recursively and looped to sort the remaining partitions. When the subarrays are less than the block size, sorting can be done quickly because all reads and writes are done in the cache, and in the external memory model requires O(1) operations.

    b) **External Merge Sort:** External merge sorting uses a hybrid sort merge strategy. In the sorting phase, chunks of data small enough to fit in main memory are read, sorted, and written out to a temporary file. In the merge phase, the sorted sub-files are combined into a single larger file. for example external merge sort algorithm , first divides the whole file data into small chunks which are small

enough so that they can fit on the main memory , these chinks are called runs , these runs are sorted individually and they when all runs are sorted these are merged together to form bigger runs , this is done until the whole file is sorted.