**Session05 Assignment 2 Solutions**

1. Explain the importance of below 4 demons in job execution with minimum of 5 points

I ) Name node ::

The namenode manages the filesystem namespace. It maintains the filesystem tree and the metadata for all the files and directories in the tree. This information is stored persistently on the local disk in the form of two files: the namespace image and the edit log. The namenode also knows the datanodes on which all the blocks for a given file are located; however, it does not store block locations persistently, because this information is reconstructed from datanodes when the system starts.

II )Data node ::

Data node are the work horses of the hadoop cluster. The data in HDFS is actually stored in data nodes memory. Data node runs node manager which is responsible for managing local resources available at data node.

III ) Resource Manager : Resource manager is a YARN daemon which is one per cluster. Yarn invokes Application Master process to read and write data to HDFS. Resource manager finds node manager that can launch application master in its container.

IV) Node manager : Node manager is a YARN daemon which is one per node. Node manager manager resources at local site. It sends heartbeat signal to namenode every periodic interval to notify its availability.