**Namenode:** 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.

**Datanode:** It is the Slave node that contains the actual data. It reports information of the blocks it contains to the NameNode in a periodic fashion. A file is split into one or more blocks and these blocks are stored in a set of DataNodes.

**Resource Manager:** The ResourceManager is the ultimate authority that arbitrates resources among all the applications in the system. It works together with the per-node NodeManagers (NMs) and the per-application ApplicationMasters (AMs).

**Node manager:** The NodeManager is the per-machine framework agent who is responsible for containers, monitoring their resource usage (cpu, memory, disk, network) and reporting the same to the ResourceManager/Scheduler.