Name Node:

1. Name Node is Master Node in hadoop architecture
2. Name node contains metadata about the files. Meaning it contains about the files that are stored onto hadoop cluster. It writes this information into a file called “Fsimage” and “editlog”. Editlog file contains incremental information
3. When Name Node starts it will read the contents of both Fsimage plus editlog into it’s memory and then finally writes the up-to-date information back to fsimage
4. If one of the Data Node is down in the cluster then Name Node will not write the updated information back to Fsimage, instead it will write the new block information [Delta information] to “editlogs”. This editlog file contains delta information

Date Node:

1. Data Node is a Slave node.
2. Data Node is the where the actual blocks of the files are present in the hadoop cluster
3. Data Node continuously send its heart beat signal’s to Name node so that Name Node understand Data Node is alive.
4. When Name Node requests the Data Node to store the input file, Data Node basing on the replication factor it will copy the file to other available Data Nodes.
5. For example if Name Node requests the Data Node 1 to store the input file, Data Node1 copies the blocks of the file onto it and requests next available Data Node2 to copy the same block onto it. Data Node2 copes the block and sends acknowledgement back to Data Node1. Similarly Data Nod1 requests Data Node3 to store the block into it; Data Node3 copies the block and sends acknowledgement back to Data Node1.
6. Data Node1 finally sends the block information back to Name Node. Data Nodes continuously sends their heart beat singles back to Name Node to say that they are alive.

Resource Manager:

1. Resource Manager is part of YARN and it is hadoop 2.0 architecture
2. Resource Manager knows about all the nodes in the cluster. Meaning it will have the information which all nodes are free and which all nodes are busy in the cluster
3. Resource Manager is the ultimate authority that arbitrates resources among all the applications in the system.
4. RM has two main components:
   1. Scheduler : Scheduler is responsible for allocating resources to various applications subject to queues. The scheduler can only does one job: Scheduling only allocates resources to various applications in the cluster
   2. Application Manager: It overall manages each application master
5. Resource Manager assigns the slots of the cluster to applications Through Application Master

Node Manager:

1. Node Manger is an agent which runs on every single machine in the hadoop cluster. The Job of Node Manager is manage nodes clots or container