**Q1-Explain the core changes made in Hadoop 2.x.**

ANSWER:

1. Single point of failure has been eliminated by introducing High Availability which provides the option of 2 name nodes in same cluster Active/Passive configuration with a hot standby.
2. Number of nodes increased from 4000 to 10000.
3. Resources are dynamic and fine grained. This leads to better cluster utilization.
4. Supports Map Reduce as well as other processing models like MPI (Message Passing Interface) and many others.
5. Allows other applications to integrate with HDFS.

**Q2- Explain the difference between MapReduce 1 and MapReduce 2 / Yarn**

ANSWER:

|  |  |
| --- | --- |
| MapReduce 1 | MapReduce 2 / Yarn |
| JobTracker keeps track of resource utilization and job monitoring | Resource Utilization is taken care by ResourceManager and NodeManager, whereas job monitoring is taken care by ApplicationMaster |
| Suited for maximum of 4000 nodes and 40000 tasks. | It can scale up to 10000 nodes and 100000 tasks |
| TaskTracker is configured with static slots. Moreover, a map tasks cannot run on reduce slot. So cluster utilization is low | Resources are dynamic and finegrained. This leads to better cluster utilization. |
|  |  |