1. What is Hadoop framework?
2. On What concept the Hadoop framework works?
3. What is MapReduce ?
4. What is compute and Storage nodes?
5. How does master slave architecture in the Hadoop?
6. How does an Hadoop application look like or their basic components?
7. Explain how input and output data format of the Hadoop framework?
8. What are the restriction to the key and value class ?
9. Explain the WordCount implementation via Hadoop framework ?
10. Which interface needs to be implemented to create Mapper and Reducer for the Hadoop?
11. What Mapper does?
12. What is the InputSplit in map reduce software?
13. What is the InputFormat ?
14. Where do you specify the Mapper Implementation?
15. How Mapper is instantiated in a running job?
16. Which are the methods in the Mapper interface?
17. What happens if you don’t override the Mapper methods and keep them as it is?
18. What is the use of Context object?
19. How does Mapper’s run() method works?
20. What is next step after Mapper or MapTask?

21.How can we control particular key should go in a specific reducer?

22. What is the use of Combiner?

23. How many maps are there in a particular Job?

24.What is the Reducer used for?

25.Explain the core methods of the Reducer?

26.What are the primary phases of the Reducer?

27. Explain the shuffle?

28.Explain the Reducer’s Sort phase?

29.Explain the Reducer’s reduce phase?

30.How many Reducers should be configured?

31.It can be possible that a Job has 0 reducers?

32.What happens if number of reducers are 0?

33.How many instances of JobTracker can run on a Hadoop Cluser?

34. What is the JobTracker and what it performs in a Hadoop Cluster?

35. How a task is scheduled by a JobTracker?

36. How many instances of Tasktracker run on a Hadoop cluster?

37.What are the two main parts of the Hadoop framework?

38.Explain the use of TaskTracker in the Hadoop cluster?

39. What do you mean by TaskInstance?

40.How many maximum JVM can run on a slave node?

41.How does a NameNode handle the failure of the data nodes?

42. Can Reducer talk with each other?

43. Where the Mapper’s Intermediate data will be stored?

44.What is the use of Combiners in the Hadoop framework?

45.What is the Hadoop MapReduce API contract for a key and value Class?

46.What is the meaning of speculative execution in Hadoop? Why is it important?

47.When the reducers are are started in a MapReduce job?

48.What is HDFS ? How it is different from traditional file systems?

49.What is HDFS Block size? How is it different from traditional file system block size?

50.What is a NameNode? How many instances of NameNode run on a Hadoop Cluster?

51.What is a DataNode? How many instances of DataNode run on a Hadoop Cluster?

52.How the HDFS Blocks are replicated?