

Important Long Answer Questions

1. List the main characteristics of Bigdata? explain
2. How a single person is contributing towards BigData? Explain.
3. List the main features of a big data in detail?
4. Why big data is important? Describe the use cases for major IT Industries?
5. Explain the advantages and disadvantages of Hadoop?
6. Compare the RDBMS and Hadoop
7. Relate the concept of rack awareness and commodity Hardware with neat diagram?
8. Outline the Hadoop 1.X Architecture with a neat diagram
9. Explain clearly how files are stored in HDFS?
10. Develop a program for word count using MapReduce ?
11. Infer the working principle of MapReduce and show dataflow diagram for it?
12. Organize how to do the storage in File System images, Edit logs, and its uses?
13. Summarize the significance of YARN and its Architecture?
14. Demonstrate how many joins are there in MapReduce and explain it?
15. Explain steps for submission of data to YARN and show in sequence diagram.
16. Identify the minimum requirements for installing Hadoop and steps for installation.
17. Tell about unstructured data? Illustrate with examples?
18. Tell about RDBMS and why it is failing to store Big data explain?
19. Why is Big data? Describe the features of a main big data in detail.
20. List out the various Data types? Explain it ?
21. Explain brief about Apache Hadoop?
22. Explain the advantages and disadvantages of Hadoop?
23. Demonstrate the Apache Hadoop Framework?
24. Illustrate the different V's and explain it briefly?
25. Make use of the mapper and combiner in MapReduce
26. Construct the configure for SSH while installing Hadoop? Explain it with uses?
27. Explain in detail about MapReduce concept with an example?
28. Solve an issue arising from edit logs utilizing check- pointing with a standby Name node?

29. Show Hadoop 1.x and 2.x Architecture and compare the difference between them?
30. Identify the minimum requirements for installing Hadoop and steps for installation.
31. Demonstrate YARN Architecture and explain its components?
32. Identify the modes in which Hadoop can be run? Develop the Pseudo Distributed Mode for Hadoop
33. Show the ease of working with structured data
34. Select the different V's and explain it briefly?
35. Show the sources of structured data and how it generates?
36. Tell about semi-structured data? Illustrate with examples? Characteristics of it
37. Why Hadoop came in existence and explain the Hadoop components?
38. Explain about Hadoop echo system?
39. Illustrate Rack awareness with an example?
40. Explain about Master-slave Architecture with neat diagram
41. Demonstrate various phases in MapReduce with pictorial representation?
42. Explain the various classes and methods in MapReduce implementation?
43. Apply the MapReduce Model for the given input (Dear, Bear, River, Car, Car, River, Deer, Car and Bear)
44. Identify the purpose of using join in MapReduce and explain the types of Joins?
45. Illustrate fundamental idea of Apache Hadoop YARN and its working processer?
46. Relate Hadoop 1.x and 2.x Architecture and explain?
47. Make use of YARN components? Advantages and Disadvantages of YARN.
48. Identify the minimum requirements for installing Hadoop and steps for installation.
49. Find out various sources of Big Data?
50. Tell about working with structured data in detail?
51. Why the Big data? Describe the main features of a big data in detail.
52. What is meant by semi-structured data? Illustrate with examples? Characteristics of it?
53. Who introduced Apache Hadoop? Explain limitations of Hadoop?
54. Explain how Hadoop came in existence and explain the Hadoop components?

55. Illustrate the different V's and explain it briefly?
56. Relate the concept of rack awareness and commodity Hardware with a neat diagram?
57. Explain clearly how files are stored in HDFS?
58. Illustrate the mapper and combiner in MapReduce?
59. Construct the step by step MapReduce concept with an example?
60. Solve an issue arising from edit logs utilizing check- pointing with a standby Name node?
61. Show Hadoop 1.x and 2.x Architecture and compare the difference between them?
62. Demonstrate how many joins are there in MapReduce and explain it?
63. Make use of YARN components? Advantages and Disadvantages of YARN.
64. Identify the modes in which Hadoop can be run? Develop the Pseudo Distributed Mode for Hadoop