

School of Computing Science and Engineering

B.Tech CSE
ETE - Jun 2023

Time : 3 Hours

Marks : 100

Sem IV - E1UA402T - Foundation of Big Data

Your answer should be specific to the question asked

Draw neat labeled diagrams wherever necessary

1. Explain all the steps involved in setting up a Hadoop Cluster. K3 CO3 (5)
2. Discuss the NoSQL data stores and their characteristic features. K1 CO1 (5)
3. Explain the advantages of Zookeeper in Apache Hadoop. K2 CO2 (5)
4. Explain Hive Integration and work flow steps involved with a diagram. K1 CO1 (10)
5. What is HDFS? What is its role in Hadoop? Draw a diagram of HDFS components and discuss the working of each component in detail. K3 CO3 (10)
6. Explain the following concepts with the help of examples. Draw neat diagrams to state the difference between them:
(a) Replication (b) Sharding (c) Sharding with replication. K2 CO2 (10)
- 7) What is the difference between Hadoop 2.0 and Hadoop 1.x. What are the new functionalities and components included in Hadoop 2.0? K5 CO4 (10)

OR

- What is Yarn? Explain the function of every component of the Yarn with a diagram. K5 CO4 (10)
- 8) What are the different types of failures that can occur in Hadoop? Explain these failures in detail. What methods does Hadoop follow to handle these failures? K4 CO4 (15)

OR

- Describe the characteristics of Big Data. Illustrate the challenges associated with Big Data. Consider an example of an E-commerce company to explain the functions and benefits of Big Data. K4 CO4 (15)
9. What do you mean by HDFS Federation? Explain its working in detail. Explain its advantages in Hadoop System. K4 CO3 (15)
 10. What is Hadoop ecosystem? What different layers and components exist in it? What is the function of each layer and component? Explain in detail with a diagram. K3 CO3 (15)