

# **SARDAR VALLABHBHAI PATEL INSTITUTE OF TECHNOLOGY, VASAD.**

Information Technology Department

Subject Name: **BIG DATA ANALYTICS (2171607)**

- 1 How does HDFS ensure data Integrity in a Hadoop Cluster?
- 2 Draw and discuss the architecture of Hive in detail.
- 3 Explain working of Hive with proper steps and diagram.
- 4 Explain Metastore in Hive.
- 5 Elaborate on HivQL data manipulation queries in detail
- 6 What do you mean by HiveQL Data Definition Language? Explain any three HiveQL DDL command with its syntax and example.
- 7 What is Hbase? Discuss in detail the data model and Implementation Aspect.
- 8 Explain Storage mechanism in HBase. Compare Row oriented and Column Oriented database structures.
- 9 Discuss the concept of regions in HBase and Storing Big data with HBase.
- 10 Explain how HBase uses Zookeeper to Build Applications with Zookeeper.
- 11 What is Zookeeper? List the benefits of it.
- 12 Explain Pig data Model in detail and Discuss how it will help for effective data flow.
- 13 Differentiate: Apache pig Vs Map Reduce.
- 14 (i) What is Zookeeper? What are the benefits of Zookeeper?  
(ii) Write a short note on Apache Pig. Enlist applications of Apache Pig.
- 15 What is NoSQL database? List the differences between NoSQL and relational databases. Explain in brief various types of NoSQL databases in practice.
- 16 Explain CRUD operations in MongoDB
- 17 Explain Replication and scaling feature of MongoDB.
- 18 Explain following for MongoDB  
(i) Indexing  
(ii) Aggregation
- 19 Explain following in brief with respect to Mongo DB :
  - ☐ Collections and documents
  - ☐ Indexing and retrieval
  - ☐ Data aggregation
- 20 Requirement specification of blog application in social networking is as follows.
  - ☐ Every post has a unique title, description and url.
  - ☐ Every post can have one or more tags.
  - ☐ Every post has comments given by users along with their name, message, data-time and likes.
  - ☐ On each post, there can be zero or more commentsFor this set of requirements design a Mongo DB schema.
- 21 Requirement specification for a meeting dashboard application in an organization is as follows:
  - ☐ Any member in an organization can host a meeting and send invitations to other members within An organization.
  - ☐ Invitees can accept or reject the meeting with proper reason.
  - ☐ Every meeting has the title, timestamp and place/location associated.
  - ☐ Every meeting has predefined agendas and documents associated.
  - ☐ Meeting discussion concludes with identifying tasks to accomplish.
  - ☐ Every task has title, priority, deadline and note associated with it. Task can be assigned to any Attendee of meeting.For this set of requirements design a Mongo DB schema.