

**GANPAT UNIVERSITY**  
**B. TECH (CE/IT/CE-AI) SEM-VII REGULAR EXAMINATION NOV-DEC 2022**  
**2CEIT702: Big Data Analytics**

**TIME: 3 HRS**

**TOTAL MARKS: 60**

**Instructions:** (1) This Question paper has two sections. Attempt each section in a separate book.  
(2) Figures on right indicate marks.  
(3) Be precise and to the point in answering the descriptive questions.

**SECTION: I**

- Q.1** [A] Compare the 'Small data', 'Large data', and 'Big data' based on the following: (6)  
1. Characteristics  
2. Typical Tools  
3. Analytical Method  
4. Examples
- [B] Explain the concept of Distributed File System. (4)

**OR**

- Q.1** [A] List out the key drivers of big data. (6)  
[B] Explain the design principle of the Google file system with a suitable diagram. (4)
- Q.2** [A] Explain the Job-Tracker and Task-Tracker of Map-Reduce with a suitable diagram. (5)  
[B] Explain the working steps of YARN with a suitable diagram. (5)

**OR**

- Q.2** [A] Discuss the Zookeeper data model with a suitable diagram. (5)  
[B] Explain the features of Apache spark. (5)
- Q.3** [A] Explain the Map-Reduce of Hadoop with a suitable diagram. (5)  
[B] Discuss the advantages of Apache spark. (5)

## SECTION: II

- Q.4** [A] Define Spark's basic building block from a programming perspective and write why spark is known as Lazy Evaluator. (5)  
[B] Describe any five big data challenges. (5)

**OR**

- Q.4** [A] Illustrate gossip protocol in Apache Cassandra with examples. (5)  
[B] Define Kafka's message Anatomy and explain its main components. (5)

- Q.5** [A] Plan the top-down hierarchy of the data model in Cassandra and Explain. (5)  
[B] Discuss the BASE approach in association with NoSQL Concept. (5)

**OR**

- Q.5** [A] Differentiate the SQL and NoSQL. (5)  
[B] Create your own use case and explain it in relation to Apache Kafka Data Model. (5)
- Q.6** [A] How do clients connect to a Kafka Cluster (bootstrap server)? (6)  
[B] Name two techniques to aid the scalability of distributed systems. Provide an example for each. (4)

-----**END OF PAPER**-----