

Reg. No. \_\_\_\_\_

Name: \_\_\_\_\_

**APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY**

SECOND SEMESTER MCA (LATERAL ENTRY) EXAMINATION, AUG 2017

**RLMCA 204: BIG DATA TECHNOLOGIES**

Max. Marks: 60

Duration: 3 Hours

**PART A***Answer All Questions. Each Question carries 3 Marks.*

1. What are the problems that are addressed by using distributed computing in Big Data processing?
2. What is BASE design principle in relation to a database?
3. Explain real-time processing mode.
4. What is the role of Hive in BD ecosystem?
5. How can we optimize MapReduce job using files system?
6. What are the advantages of using distributed file system over conventional databases?
7. What are the characteristics of RDBMS database as an on disk storage device?
8. Explain qualitative analysis with an example.

**PART B***Answer any one question from each module. Each question carries 6 Marks*

## MODULE I

9. a. Which are the three cloud delivery models? (3)
- b. Explain volume, velocity and variety. (3)

OR

10. Explain the cloud deployment models.

## MODULE II

11. Explain peer to peer replication model.

OR

12. With a diagram explain the benefits of comparing Sharding and Replication.

**MODULE III**

13. Explain the HDFS architecture.

OR

14. Draw a neat diagram of the Hadoop ecosystem and list the seven important tools in it.

**MODULE IV**

15. Explain the working of MapReduce.

OR

16. Explain the key features of MapReduce.

**MODULE V**

17. List any six features of NoSQL that enables it to be used for Big Data Storage?

OR

18. How real-time data processing is achieved using In-Memory Data Grid?

**MODULE VI**

19. Explain the use of correlation for data analysis with examples.

OR

20. Explain how time series plots helps in visual analysis of Big Data.

\*\*\*\*