PMCAT303 - Paper-III : Big Data & Hadoop

P. Pages : 2

Time : Three Hours

2381

GUG/S/23/13649

Max. Marks : 80

––

Notes : 1. All questions are compulsory and carry equal marks.
2. Draw neat and labelled diagrams wherever necessary.
3. Avoid vague answers and write answers relevant and specific to questions only.

Either:

1. a) What is distributed File System? Explain the significance of Four V's in Big data. **8**

b) Explain in brief about Big Data Analysis? **8**

OR

c) What is Big Data? Explain Characteristics of Apache Hadoop? **8**

d) Explain Map Reduce concept and all building blocks of Hadoop. **8**

Either:

2. a) Explain Hadoop Shell Command in detail? **8**

b) Write about SSH & Hadoop Configuration. **8**

OR

c) Explain in brief about Hadoop Architecture? **8**

d) Define Map Reduce? Explain the implementation of a map reduce with suitable example. **8**

Either:

3. a) Explain in detail. **8**

i) MRv2

ii) YARN.

b) What are the New Features of Hadoop 2.0? Explain? **8**

OR

c) Explain the concept of HDFS? **8**

d) What are the components of Hadoop ecosystem? Explain. **8**

Either:

4.	a)	Explain Sorting and Aggregating in HiveQL?	**8**
	b)	How the Schema Design is done in HBase? Explain in detail.	**8**

<div align="center">

OR

</div>

	c)	What is Zookeeper and how it helps in monitoring a Cluster?	**8**
	d)	Write a short note on PIG?	**8**
5.		Attempt **all** the questions.	
	a)	Explain Data serialization?	**4**
	b)	Write a short note on Name Node and Data Node?	**4**
	c)	Write about Schedulers in Hadoop.	**4**
	d)	How to Build Applications with Zookeeper?	**4**

<div align="center">

****

</div>