MANAV RACHNA vidyanatariksha

MANAV RACHNA UNIVERSITY
(FORMERLY MANAV RACHNA COLLEGE OF ENGINEERING
NAAC ACCREDITED 'A' GRADE INSTITUTION)
Declared as State Private University under section 2f of the UGC act, 1956

DEPARTMENT OF COMPUTER ENGINEERING

"T3 Examination, May- 2018"

Semester: 4th
Subject: Big Data
Branch: CE
Course Type: Core
Time: 3 Hours
Max.Marks: 60

Date of Exam: 15/05/2018
Subject Code: CSH628-T
Session: II
Course Nature: Hard
Program: M.Tech
Signature: HOD/Associate HOD:

Note:
Part A: All questions are compulsory. Each Question carries 2 marks.
Part B: Attempt any two questions. Each question carries 10 marks.
Part C: Attempt any two questions. Each question carries 10 marks.

PART-A

Q1.

(a) What is the motivation behind using Sequence files and ORCFILE in Hadoop?
(b) List the various Hadoop Built-In Counters.
(c) What does it mean to be "Pigs Are Domestic Animals"? Explain.
(d) What is the problem in having lots of small files in HDFS?
(e) Explain job scheduling through Job Tracker.
(f) Can a hive query read and write to the same table?
(g) What do you understand by Efficiency of a Program Code?
(h) What is Hadoop and its components.
(i) What is the best way to extract large amounts of data from Hive?
(j) Describe DDL statement in hive.

PART-B

Q2.

a) Highlight the major differences between functional programming and object-oriented programming.
b) List any five real-time industry applications of Hadoop.

Q3.

a) List the various HDFS commands that enable us to work in a distributed environment.
b) "YARN enhances a Hadoop compute cluster". Justify this statement.

****

Q4.

a) How do we perform a Map-Reduce Join? Illustrate it with an example.
b) There are various daemons related to HDFS. Discuss.

PART-C

Q5. Write short notes on the following:

1) Business Intelligence on Hadoop
2) PIG
3) Workflow between Hive and Hadoop framework

Q6. How Facebook uses Hadoop and Hive? Discuss. The discussion should include the key details such as Motivation, Metadata, Architecture, Performance, Pros and Cons, and Application.

Q7. List the commands for the following operations:

1. Insert data into tables directly from SQL.
2. Overwrite any existing data in the table or partition
3. Move data files into locations corresponding to Hive tables
4. Modify the structure, metadata or data of the table
5. Copy the schema of an existing table, not the data

Explain each command in detail.

****