**14CS/IT 703**

**Hall Ticket Number:**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **IV/IV B.Tech (Regular / Supplementary) DEGREE EXAMINATION** | | | | | | | |
| **January, 2021** | | | **Common to CSE and IT** | | | | |
| **Seventh Semester** | | | **Advanced Data Analytics** | | | | |
| **Time:** Three Hours | | | | **Maximum :** 60 Marks | | | |
| *Answer ALL Questions from PART-A.* | | | | | (1X12 = 12 Marks) | | |
| *Answer* ***ANY FOUR*** *questions from PART-B.* | | | | | (4X12=48 Marks) | | |
| **Part - A** | | | | | | | |
| 1 | Answer all questions | | | | | (1X12=12 Marks) | |
|  | a) | What is Big Data? | | | | |  |
|  | b) | What are the applications of big data? | | | | |  |
|  | c) | Define namenodes and datanodes in HDFS. | | | | |  |
|  | d) | Why blocks are large in HDFS? | | | | |  |
|  | e) | What is the default scheduler in Hadoop? | | | | |  |
|  | f) | What is the role of Jobtracker and Tasktracker in Map Reduce? | | | | |  |
|  | g) | List the failures in classical Map Reduce. | | | | |  |
|  | h) | List the built-in counters in Map Reduce. | | | | |  |
|  | i) | How to display the records in a relation R using Pig Latin. | | | | |  |
|  | j) | Compare Managed table with external table in Hive. | | | | |  |
|  | k) | How to read and print a text file using spark? | | | | |  |
|  | l) | Write any one application of Apache Sqoop? | | | | |  |
| **Part - B** | | | | | | | |
| 2 |  | Show how a client read and write data in HDFS with suitable code. | | | | | 12M |
|  | | | | | | | |
| 3 | | Discuss the steps involved in designing Hadoop Distributed File System and Give the design of HDFS. | | | | | 12M |
|  | | | | | | | |
| 4 | a) | Analyze the data with Hadoop using Map and Reduce with an example. | | | | | 6M |
|  | b) | Explain in detail about MapReduce data flow with single and multiple reduce tasks. | | | | | 6M |
|  | | | | | | | |
| 5 | a) | How do you support the statement that, “YARN is better than MapReduce 1”? | | | | | 6 M |
|  | b) | How YARN runs a MapReduce application? | | | | | 6 M |
|  | | | | | | | |
| 6 | a) | Explain about Pig Latin expressions and Pig Latin types. | | | | | 6M |
|  | b) | Explain about data processing operators in Pig. | | | | | 6M |
|  | | | | | | | |
| 7 |  | Write about the Hive architecture and how it works? | | | | | 12M |
|  | | | | | | | |
| 8 | a) | Explain the Anatomy of spark job run. | | | | | 6M |
|  | b) | Explain briefly about task scheduling and task execution in spark. | | | | | 6M |
|  | | | | | | | |
| 9 | a) | Describe Sqoop import command. | | | | | 6M |
|  | b) | How Sqoop export command works? | | | | | 6M |

****