



MBT-003-1082002 Seat No. _____

M. Sc. (IT) (Sem. II) (CBCS) Examination

April / May - 2018

CS - 08 : Introduction to Big Data & Hadoop

Faculty Code : 003

Subject Code : 1082002

Time : $2\frac{1}{2}$ Hours]

[Total Marks : 70

- 1 (A) Objective type questions : 4
- (1) The Hadoop framework itself is mostly written in the Java programming language. true or false?
 - (2) List different types of nosql databases.
 - (3) _____ command to list the blocks that make up each file in the file system.
 - (4) An HDFS cluster has two types of nodes _____ and _____ operating in a master-slave pattern.
- (B) Answer in brief : (Any **One**) 2
- (1) Explain Limitations of traditional large Scale systems.
 - (2) Define name node, data node and checkpoint node.
- (C) Answer in detail : (Any **One**) 3
- (1) Explain any three Linux commands.
 - (2) What is Big Data? Explain characteristics of Big Data.
- (D) Write a note on : (Any **One**) 5
- (1) Explain HDFS architecture.
 - (2) Explain Schema on Read and Schema on Write.

- 2 (A) Objective type questions : 4
- (1) MapRduce was devised by _____
 - (2) _____ part of the MapReduce is responsible for processing one or more chunks of data and producing the output results.
 - (3) Which of the following operator is used to view the map reduce execution plans ?
 - (4) _____ operator is used to review the schema of a relation.
- (B) Answer in brief : (Any **One**) 2
- (1) Explain execution mode of apache pig.
 - (2) Write a short note on Partitioning.
- (C) Answer in detail : (Any **One**) 3
- (1) Write a short note on shuffling and sorting.
 - (2) Explain Describe, explain and illustrate.
- (D) Write a note on : (Any **One**) 5
- (1) Write a short note on MRv2 & YARN.
 - (2) Explain pig architecture.
- 3 (A) Objective type questions : 4
- (1) To select all columns starting with the word 'Sell' from the table GROSS-SELL the query is
 - (2) Give name of complex datatypes in hive.
 - (3) Write a syntax of LOAD DATA statement.
 - (4) Hive is designed for OLAP true or false?
- (B) Answer in brief : (Any **One**) 2
- (1) Explain different modes of executing Hive queries.
 - (2) Write a short on partitioning in Hive.

- (C) Answer in detail : (Any One) 3
- (1) What is view? Explain views in Hive.
 - (2) Write a short note on Hadoop Streaming.
- (D) Write a note on : (Any One) 5
- (1) Explain architecture of Hive.
 - (2) Explain Map Reduce non-graph algorithm.
- 4 (A) Objective type questions : 4
- (1) Which command is used to disable all the tables matching the given regex ?
 - (2) What is the use of get() method?
 - (3) HBase is a distributed _____ database built on top of the Hadoop file system
 - (4) Hbase column family cannot be renamed. True or false?
- (B) Answer in brief : (Any One) 2
- (1) Write a short note on master and region server.
 - (2) Explain features of Hbase.
- (C) Answer in detail : (Any One) 3
- (1) How to create a table in Hbase? Explain with example.
 - (2) How to enable and disable tables in Hbase? Explain.
- (D) Write a note on : (Any One) 5
- (1) Explain Hbase architecture in detail.
 - (2) Explain describe, alter and drop command in Hbase.

- 5 (A) Objective type questions : 4
- (1) Give a full form of EC2.
 - (2) _____ is used for graph processing.
 - (3) _____ was created to allow you to flow data from a source into your Hadoop environment.
 - (4) _____ is a distributed real-time computation system for processing large volumes of high-velocity data.
- (B) Answer in brief : (Any **One**) 2
- (1) What is use of impala in Hadoop? explain.
 - (2) Explain RDD in spark.
- (C) Answer in detail : (Any **One**) 3
- (1) How to Importing/exporting data across RDBMS and HDFS using Sqoop?
 - (2) Explain graph processing using neo4j.
- (D) Write a note on : (Any **One**) 5
- (1) What is flume? Explain advantage of flume.
 - (2) Explain spark architecture.
-