6M 6M

6M 6M

TT-11/TC*-1 -4 NT1			<b>CS/IT 42</b> 2	
На		Γicket Number:		
		IV/IV B.Tech (Supplementary) DEGREE EXAMINATION		
		mber, 2019		
	_	Semester Chree Hours  Advanced Data Analy Maximum: 60 M		
Ans	swer	Question No.1 compulsorily. $(1X12 = 12 \text{ M})$	[arks]	
Ans	swer	ONE question from each unit. (4X12=48 M	[arks]	
1	An	nswer all questions (1X12=12 Ma	rks)	
	a)	What is Big Data?		
	b)	What are the applications of big data?  Define namenodes and datanodes in HDFS.		
	c) d)	Why blocks are large in HDFS?		
	e)	What is the default scheduler in Hadoop?		
	f)	What is the role of Job tracker and Task tracker in Map Reduce?		
	g)	List the failures in classical Map Reduce.		
	h) i)	List the built-in counters in Map Reduce. Purpose of Filter function in Pig.		
	i) j)	List the complex data types in Hive.		
	k)	Write the scala program to find maximum temperature.		
	I)	Write any one application of Apache Swoop.		
		UNIT I		
2	a)	Show how a client read and write data in HDFS with suitable code.	6M	
	b)	Discuss in detail about parallel copying with distep in HDFS.	6M	
2		(OR) Discuss the stars involved in deciming Heleon Distributed File System and Cive the decim of	121/1	
3		Discuss the steps involved in designing Hadoop Distributed File System and Give the design of HDFS.	12M	
		UNIT II		
4	a)	Analyze the data with Hadoop using Map and Reduce with an example.	6M	
	b)		6M	
_	- \	(OR)	1214	
5	a)	Explain briefly about shuffle and sort in Map Reduce.  UNIT III	12M	
6	a)		6M	
	b)	Explain about data processing operators in Pig.  (OR)	6M	
7	a)	· /	6M	
-	b)	Write the difference between the Traditional Databases and Hive.	6M	
		UNIT IV		



(OR)

Explain the Anatomy of spark job run. Explain briefly about task scheduling and task execution in spark.

a) Describe Sqoop import command.b) How Sqoop export command works?

8

9

a)