

Code No: RT4105B

## **R13**

Set No. 1

## IV B.Tech I Semester Supplementary Examinations, February/March - 2018 HADOOP AND BIG DATA

(Common to Computer Science and Engineering and Information Technology)
Time: 3 hours

Max. Marks: 70

Question paper consists of Part-A and Part-B Answer ALL sub questions from Part-A Answer any THREE questions from Part-B \*\*\*\*\*

## PART-A (22 Marks)

1.	a)	Write a short note on Serialization.	[3]
	b)	Explain the working of name node and data node.	[4]
	c)	What is the necessity of driver code?	[4]
	d)	Define Bytes Writable?	[3]
	e)	What is the necessity of pig Latin?	[4]
	f)	Write at least two differences between pig and hive.	[4]
		<b>DADT P</b> (2016 - 48 Manks)	
2	۵)		F01
2.	a)	Discuss in detailed about linked list with an example.	[8]
	b)	What is set and map and explain their operation.	[8]
3.	a)	Explain Google File System architecture with neat diagram.	[8]
	b)	Discuss in detail about Hadoop Distributed File System (HDFS).	[8]
4.	a)	Write short notes on (i) Driver code (ii) Mapper code (iii) Reducer code (iv) Combiner	[8]
		(i) Driver code	
		(ii) Mapper code	
		(iii) Reducer code	
	b)	Write the difference between Old and New Hadoop API for MapReduce	
		Framework.	[8]
5.	a)	How to Implementing a Raw Comparator for speed?	[8]
	b)	Explain about Generic Writable, Writable collections.	[8]
6.	a)	Explain the pig architecture with a neat diagram.	[8]
	b)	Discuss Pig Latin Application Flow.	[8]
7.	a)	Explain briefly about different Hive Data Types.	[8]
٠.	b)	How to Creating and Managing Databases and Tables in hive?	[8]
	$\sigma_j$	110 % to Crowing and Managing Databases and Tubies in inve-	[o]