Code No: **RT4105B** **R13** **Set No. 1**

IV B.Tech I Semester Regular Examinations, November - 2016

HADOOP AND BIG DATA

(Common to Computer Science & 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)	Give the difference between autoboxing and unboxing.	[4]
	b)	How a secondary name node differs from the name node in HDFS.	[4]
	c)	Define the role of combiner and partitioner in a map reduce application.	[4]
	d)	What do you mean by serialization and how should be the RPC serialization format?	[3]
	e)	Define the three key design principles of pig latin.	[3]
	f)	How to create a table by using HIVEQL.	[4]

PART–B *(3x16 = 48 Marks)*

2.	a)	Why linked lists, stacks and queues are called as linear data structures and explain the operations performed on stacks and queues with examples.	[8]
	b)	What is the use of generic methods and generic classes in java and explain the various generic methods and classes supported by java.	[8]
3.	a)	Explain the basic building blocks of Hadoop with a neat sketch.	[8]
	b)	Explain the various operational modes of Hadoop cluster configuration.	[8]
4.	a)	Distinguish between the old and new versions of Hadoop API for Map Reduce frame work.	[8]
	b)	Explain about the implementation of map reduce concept with a small example.	[8]
5.	a)	Explain the significance of Writable interface along with Writable Comparable and comparators w.r.to implementing the serialization.	[8]
	b)	Explain the Writable class hierarchy with a neat sketch.	[8]
6.	a)	Explain the architecture of a pig with a neat sketch.	[8]
	b)	Explain the syntax of a pig program with a suitable example.	[8]
7.	a)	Explain with neat sketch about the configuration of CLI client and WI client while interacting with HIVE.	[8]
	b)	Explain about the various data types supported by HIVEQL with an example.	[8]

Code No: **RT4105B** **R13** **Set No. 2**

IV B.Tech I Semester Regular Examinations, November - 2016

HADOOP AND BIG DATA

(Common to Computer Science & 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)	What is a map and specify the various map implementations in java.	[4]
	b)	Specify the role of name node and data node in HDFS.	[4]
	c)	List the components of a map reduce application that we can develop.	[3]
	d)	List the writable wrapper classes for java primitives.	[3]
	e)	Define the basic syntax of a pig.	[4]
	f)	Define the various file formats supported by HIVE.	[4]

PART–B *(3x16 = 48 Marks)*

2.	a)	What do you mean by linear and non-linear data structures? Specify the sets are comes under linear or non-linear and explain the various types of sets supported by java.	[8]
	b)	What is the advantage of object serialization in java and explain about serializing & de-serializing an object with suitable examples.	[8]
3.	a)	Explain with a neat sketch about the processing of a job in hadoop.	[8]
	b)	List the various operational modes of hadoop cluster configuration and explain in detail about configuring/installing the hadoop in local/standalone mode.	[8]
4.	a)	Explain the role of driver code, mapper code and reducer code within a map reduce program model by a suitable example.	[8]
	b)	Explain the anatomy of map reduce job run.	[8]
5.	a)	What do you mean by a custom writable and explain the implementation of a custom writable with an example.	[8]
	b)	Explain about the implementation of raw comparator and custom raw comparator with an example.	[8]
6.	a)	How the pig programs can be packaged and explain the modes of running a pig script with a neat sketch.	[8]
	b)	Explain about the various data types supported by pig in its data model with an example.	[8]
7.	a)	Explain the steps followed to get SQuirreL running on the Apache HIVE with a neat sketch.	[8]
	b)	What is the use of SerDes (Serializer&Deserializer) in HIVE and explain various types of SerDes'ssupported by HIVE.	[8]

Code No: **RT4105B** **R13** **Set No. 3**

IV B.Tech I Semester Regular Examinations, November - 2016

HADOOP AND BIG DATA

(Common to Computer Science & 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) What is a wrapper class and list out the wrapper classes in java? [3]
   b) Specify the role of job tracker and task tracker in HDFS. [4]
   c) Define the role of mapper code and reducer code in a map reduce application. [4]
   d) How a custom raw comparator differs from the raw comparator. [3]
   e) How the pig programs can be packaged. [4]
   f) Show that how the serializer and deserializer helpful to keep the file formats separate from the record formats. [4]

PART–B*(3x16 = 48 Marks)*

2. a) What is the difference between the 'set' & a 'map' data structure and explain about various map implementations in java with suitable examples. [8]
   b) Specify the difference between a primitive type and a wrapper class. Explain about the conversion from primitive type to wrapper class and vice-versa with an example. [8]

3. a) What are the various operational modes of hadoop cluster configuration and explain in detail about configuring/installing hadoop in fully distributed mode. [8]
   b) Are there any similarities between the GFS & HDFS and explain the GFS architecture with a neat sketch. [8]

4. a) Explain the role of combiner, record reader and partitioner within a map reduce program model of hadoop. [8]
   b) Distinguish between the old and new versions of Hadoop API for Map Reduce frame work. [8]

5. a) Explain the Writable class hierarchy with a neat sketch. [8]
   b) Explain the significance of Writable interface along with WritableComparable and comparators w.r.to implementing the serialization. [8]

6. a) Explain the operators supported by pig w.r.to. data access, transformations and debugging operations. [8]
   b) Explain the syntax of a pig program with suitable example. [8]

7. a) Explain about the various data types supported by HIVEQL with an example. [8]
   b) Explain with neat sketch about the configuration of CLI client and WI client while interacting with HIVE. [8]

Code No:**RT4105B** **R13** **Set No. 4**

IV B.Tech I Semester Regular Examinations, November - 2016

HADOOP AND BIG DATA

(Common to Computer Science & 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)	What is a set and specify the various set implementations in java	[4]
	b)	What do you mean by a block in file system and specify its size?	[3]
	c)	What happens in map phase and reduce phase of a hadoop map reduce frame.	[4]
	d)	Define the following wrappers: Byte writable, Object writable and Generic writable.	[4]
	e)	What is a pig and specify its role in Hadoop?	[3]
	f)	Define the various file formats supported by HIVE.	[4]

PART–B*(3x16 = 48 Marks)*

2.	a)	What is a wrapper class and explain the concept of autoboxing& unboxing with suitable examples.	[8]
	b)	Why stacks, queues and linked lists are called as linear data structures and explain the operations performed on stacks &linked lists with suitable examples.	[8]
3.	a)	Explain the hadoop distributed file system architecture with a neat sketch.	[8]
	b)	How google file system differs from the hadoop file system and explains the google file system architecture with a neat sketch.	[8]
4.	a)	Explain the role of driver code, mapper code and reducer code within a map reduce program model by a suitable example.	[8]
	b)	Explain about the implementation of map reduce concept with a small example.	[8]
5.	a)	Explain about the implementation of raw comparator and custom raw comparator with an example.	[8]
	b)	What do you mean by a custom writable and explain the implementation of a custom writable with an example.	[8]
6.	a)	Explain about the various data types supported by pig in its data model with an example.	[8]
	b)	How the pig programs can be packaged and explain the modes of running a pig script with a neat sketch.	[8]
7.	a)	Explain the architecture of HIVE with a neat sketch.	[8]
	b)	How can we install the Apache Hive on the system – Explain.	[8]