VI SEMESTER B. TECH. (COMPUTER SCIENCE AND ENGINEERING)
MAKEUP EXAMINATIONS, MAY 2018
SUBJECT: ELECTIVE III - BIG DATA INTEGRATION AND PROCESSING [CRA 4002]
REVISED CREDIT SYSTEM

Time: 3 Hours   (22/06/2018)   Max. Marks: 50

Instructions to Candidates:

- ❖ Answer **ALL** the questions.
- ❖ Missing data may be suitably assumed.
- ❖ Draw diagrams wherever applicable.

1A.	Consider the library database : member (**memb_no**, name, age) book (**isbn**, title, authors, publisher) borrowed (**memb_no, isbn**, data) Write the following queries in SQL: a. List the titles of all book published by "McGraw-Hill" b. List the names of members who have borrowed at least 1 book. c. List the name of members who have never borrowed a book. d. List all the titles that are published by publisher whose name starts with "Price". e. List all publishers along with the number of books borrowed for that publisher.	**5**
1B.	How is Big Data Processing different from traditional data processing architecture?	**3**
1C.	List the different aggregate functions in SQL. Illustrate each with an example.	**2**
2A.	With a neat diagram, explain how Aerospike queries streaming data.	**5**
2B.	Using wordcount example, explain three steps that can be applied to achieve data parallel scalability.	**3**
2C.	Explain integrated views.	**2**
3A.	Demonstrate different steps of Data Fusion with suitable examples.	**5**
3B.	Explain Mediated Schema and Attribute Matching with an example.	**3**
3C.	How Schema Mapping is performed during integration of multiple Big Data sources?	**2**
4A.	Explain aggregation framework in MongoDb with an example.	**5**
4B.	Explain how actor-oriented approach of Kepler provides flexibility and improves application programmability?	**3**
4C.	Mention any four shortcomings of Hadoop MapReduce.	**2**
5A.	Explain the functionality of the components in the Spark stack.	**5**
5B.	Describe the immutable property of Spark RDD with an example.	**3**
5C.	Differentiate between Spark Action and Transformation with suitable example.	**2**