

Nirma University
Institute of Technology
Semester End Examination (IR), December - 2021
B. Tech. in Computer Engineering, Semester-VII
IT7C4 Big Data Analytics

Roll /
Exam No.

Supervisor's Initial
With Date

Time: 2 Hours

Total Marks: 50

Instructions:

1. Attempt all questions.
2. Figures to right indicate full marks.
3. Draw neat sketches wherever necessary.

Q:1 Answer the following questions: (5 X 5) [Any Five]

[25]

- 1 Compare MapReduce with RDBMS.
- 2 Describe the difference between horizontal scaling and vertical scaling. Also list out at least three platforms under each category.
- 3 Consider the application of healthcare analytics. Describe "variety" of data generated in big data.
- 4 Name and elaborate the HDFS commands required to (i) create folder (ii) move file from local file system to HDFS (iii) copy the contents of one file to another file (iv) see the details of all the files and folders
- 5 Explain CAP theorem in detail. Compare NOSQL database with RDBMS.
- 6 What is input split? What is the role of input split in execution of map reduce program?
- 7 What is Cassandra? Write down the queries for CRUD operation in Cassandra.

Q:2 Answer the following questions: (4 X 4) [Any Four]

[16]

Consider the following structure of 'restaurants' collection:

```
{
  "address": {
    "building": "1007",
    "coord": [ -73.856077, 40.848447 ],
    "street": "Morris Park Ave",
    "zipcode": "10462"
  },
  "borough": "Bronx",
  "cuisine": "Bakery",
  "grades": [
    { "date": { "$date": 1393804800000 }, "grade": "A", "score": 2 },
    { "date": { "$date": 1378857600000 }, "grade": "A", "score": 6 },
    { "date": { "$date": 1358985600000 }, "grade": "A", "score": 10 }
  ],
  { "date": { "$date": 1322006400000 }, "grade": "A", "score": 9 },
  { "date": { "$date": 1299715200000 }, "grade": "B", "score": 14 }
],
  "name": "Morris Park Bake Shop",
}
```

```
"restaurant_id": "30075445"  
}
```

- 1
 - a. Write a MongoDB query to display all the documents in the collection restaurants.
 - b. Write a MongoDB query to display the fields restaurant_id, name, borough and cuisine for all the documents in the collection restaurant.
- 2
 - a. Write a MongoDB query to display grades for the past one year.
 - b. Write a MongoDB query to display the fields restaurant_id, name, borough and cuisine, but exclude the field _id for all the documents in the collection restaurant
- 3
 - a. Write a query to find all the document having the cuisine start from letter 'a'.
 - b. Write a MongoDB query to display the fields restaurant_id, name, borough and zip code, but exclude the field _id for all the documents in the collection restaurant
- 4
 - a. Write a MongoDB query to display all the restaurant which is in the borough Bronx.
 - b. Write a query to sort all the documents according to cuisine.
- 5
 - a. Write a query to sort all the documents according to cuisine "Bakery" and restaurant name "Morris Park"
 - b. Write a query to sort all the documents according to restaurant_id

Q:3 Deliberate the detailed flowchart to show how Hadoop runs MapReduce [09]
job using YARN.