## **Question Paper**



## MANIPAL UNIVERSITY

## SCHOOL OF INFORMATION SCIENCES (SOIS) FIRST SEMESTER MASTER OF ENGINEERING - ME (BIG DATA AND DATA ANALYTICS) DEGREE EXAMINATION- APRIL 2017 Friday, 28,2017 Time : 10:00AM- 1:00PM

DevOps for Big Data Systems [BDA 615.1]

Marks: 100

Duration: 180 mins.

## Answer all the questions.

- Compare the following three data collection (10) methods:
  (a) Senser based methods
  - (a) Sensor-based methods
  - (b) log-file based methods
  - (c) Web-crawler based methods
- Bring out four important differences between a <sup>(10)</sup> traditional SQL DBMS and a column-oriented data store such as Google's BigTable.
- <sup>3)</sup> In the context of the consistency model <sup>(10)</sup> supported by GFS, state the
  - (a) guarantees that GFS offers.
  - (b) implications for the GFS applications.
- <sup>4)</sup> Describe how atomic record appending works <sup>(10)</sup> in GFS. What are the key technical challenges in achieving this model.
- <sup>5)</sup> Provide the formal definition and an example <sup>(10)</sup> for each of the following operators found in the algebra of Grammar of Graphics:
  - (a) Nest operator
  - (b) Blend operator
- 6) Show with a good example, how the Nest (10) algebraic operator may be used to represent the **Nested dot plot**.
- <sup>7)</sup> Show the contexts in which you would choose <sup>(10)</sup>

- (a) Stacked graphs
- (b) Small-multiples
- (c) Horizon graphs

Give an example for each to justify your response.

- <sup>8)</sup> With suitable examples, state at least two <sup>(10)</sup> techniques for visualizing network of relationships. What are the technical challenges in each of these techniques?
- <sup>9)</sup> With an example for each, state the rationale <sup>(10)</sup> behind the following rules for producing better visualization:
  - (a) Use colors effectively
  - (b) Avoid Chartjunk
  - (c) Message trumps beauty
- <sup>10)</sup> Bring out four important differences between <sup>(10)</sup> the MapReduce programming model, and the Graph Processing (programming) model.