



MERU UNIVERSITY OF SCIENCE AND TECHNOLOGY
P.O. Box 972-60200 – Meru-Kenya
Tel: +254(0) 799 529 958, +254(0) 799 529 959, + 254 (0) 712 524 293,
Website: info@must.ac.ke Email: info@must.ac.ke

University Examinations 2023/2024

**SECOND YEAR SECOND SEMESTER EXAMINATION FOR THE DEGREE OF
BACHELOR OF SCIENCE DATA SCIENCE**

CDS 3251: BIG DATA ARCHITECTURE

DATE: APRIL 2024

TIME: 2 HOURS

INSTRUCTIONS: *Answer question **one** and any other **two** questions*

QUESTION ONE (30 MARKS)

- (a) Briefly explain any three big data challenges. [6 marks]
- (b) Discuss any three business opportunities offered by big data . [6 marks]
- (c) Discuss any three components of Hadoop ecosystem. [6 marks]
- (d) Discuss main components of the following architectures
- i. Kappa
 - ii. Lambda
 - iii. Event driven [6 marks]
- (e) Briefly explain main components of bigdata system and state any common application that implements it. [6 marks]



MUST is ISO 9001:2015 and



ISO/IEC 27001:2013 CERTIFIED

QUESTION TWO (20 MARKS)

- (a) Discuss any two challenges of dealing with unstructured data. [4 marks]
- (b) Using a simple diagram discuss Architecture of Hadoop MapReduce and state any of its two shortcomings as compared with Apache SPARK. [7 marks]
- (c) Discuss any three of the following Big Data processing systems and platforms.
- i. PIG
 - ii. STORM
 - iii. SPARK [9 marks]

QUESTION THREE (20 MARKS)

- (a) Discuss any three concerns for any Big Data system. [6 marks]
- (b) Discuss any four main components of MongoDB. [8 marks]
- (c) Briefly explain workflow of RDD as used in SPARK applications [6 marks]

QUESTION FOUR (20 MARKS)

- (a) Briefly explain the concept of features selection while dealing with Bigdata. [2 marks]
- (b) Discuss any three Big Data analytical type and state design consideration for such systems. [8 marks]
- (c) Differentiate between Hadoop Distributed file system from traditional Network file system. [4 marks]
- (d) Outline roles of Map Reduce in Hadoop and outline any two of its core functions. [6 marks]



QUESTION FIVE (20 MARKS)

- a) State roles of Cloud computing in Big Data Analytics [6 marks]
- b) Discuss roles of orchestration components in Big Data Applications. Give example of such software components. [4 marks]
- c) Briefly explain the following terms in regards to stream processing
- i. Cold path
 - ii. Hot path analytics [4 marks]
- d) Using a simple diagram explain SPARK architecture. [6 marks]



MUST is ISO 9001:2015 and



ISO/IEC 27001:2013 CERTIFIED