| NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA (An Autonomous Institute Affiliated to AKTU, Lucknow) M. Tech (Integrated) SEM: VI - THEORY EXAMINATION (2023 - 2024) Subject: Big Data Time: 3 Hours Max. Marks: 100 General Instructions: Max. Marks: 100 IMP: Verify that you have received the question paper with the correct course, code, branch etc. I. This Question paper comprises of three Sections -A, B, & C. It consists of Multiple Choice Questions (MCQ's) & Subjective type questions. 2. Maximum marks for each question are indicated on right -hand side of each question. 3. Illustrate your answers with neat sketches wherever necessary. 4. Assume suitable data if necessary. 5. Preferably, write the answers in sequential order. 6. No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked. 20 1. Attempt all parts:- 1 1-a. Which platform is used to run hadoop? (CO1) 1 (a) Debian 20 (b) Cross platform 1 (c) Bare metal 1 (d) Unix like 1 1-b. Identify the different features of Big Data Analytics. (CO1) 1 (a) Open source . (b) Data recovery . . (| Print | ed Pa | ge:- 03 Subject Code:- AMICSAI0621 |
|---|--------|------------|--|
| (An Autonomous Institute Affiliated to AKTU, Lucknow) M. Tech (Integrated) SEM: VI - THEORY EXAMINATION (2023 - 2024) Subject: Big Data Time: 3 Hours Max. Marks: 100 General Instructions: Max. Marks: 100 Guestion paper comprises of three Sections -A, B, & C. It consists of Multiple Choice Questions (MCQ's) & Subject: by questions. Maximum marks for each question are indicated on right -hand side of each question. 3. Illustrate your answers with neat sketches wherever necessary. 4. Assume suitable data if necessary. 5. Preferably, write the answers in sequential order. 6. No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked. SECTION-A 20 1. Attempt all parts:- 1 1-a. Which platform is used to run hadoop? (CO1) 1 (a) Debian 1 (b) Cross platform 1 (c) Bare metal 1 (d) Unix like 1 1-b. Identify the different features of Big Data Analytics. (CO1) 1 (a) Open source 1 (b) Data recovery 1 (c) Scalbility 1 (d) All of the above 1 1- | | | Roll. No: |
| (An Autonomous Institute Affiliated to AKTU, Lucknow) M. Tech (Integrated) SEM: VI - THEORY EXAMINATION (2023 - 2024) Subject: Big Data Time: 3 Hours Max. Marks: 100 General Instructions: A Bubiestion paper comprises of three Sections -A, B, & C. It consists of Multiple Choice Questions (MCQ's) & & Subject: Buy equestion are indicated on right -hand side of each question. A Subject: Buy equestion are indicated on right -hand side of each question. Sections -A, B, & C. It consists of Multiple Choice Questions (MCQ's) & Not Written material after a blank sheet will not be evaluated/checked. 6 1 | NC | DIDA | INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA |
| SEM: VI - THEORY EXAMINATION (2023 - 2024) Subject: Big Data Time: 3 Hours Max. Marks: 100 General Instructions: IMP: Verify that you have received the question paper with the correct course, code, branch etc. 1. This Question paper comprises of three Sections -A, B, & C. It consists of Multiple Choice Questions (MCQ's) & Subjective type questions. 2. Maximum marks for each question are indicated on right -hand side of each question. 3. Illustrate your answers with neat sketches wherever necessary. 4. Assume suitable data if necessary. 5. Preferably, write the answers in sequential order. 6. No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked. SECTION-A 1. Attempt all parts:- 1-a. Which platform is used to run hadoop? (CO1) 1 (a) Debian 1 (b) Cross platform 1 (c) Bare metal 1 (d) Unix like 1 1-b. Identify the different features of Big Data Analytics. (CO1) 1 (a) Open source . . (b) Data recovery . . . (c) Scala | 1,0 | | |
| Subject: Big Data Max. Marks: 100 General Instructions: Maximum marks for each question paper with the correct course, code, branch etc. 1. This Question paper comprises of three Sections -A, B, & C. It consists of Multiple Choice Questions (MCQ' s) & Subjective type questions. 2. Maximum marks for each question are indicated on right -hand side of each question. 3. 3. Illustrate your answers with neat sketches wherever necessary. 4. 4. Assume suitable data if necessary. 5. 5. Preferably, write the answers in sequential order. 6. 6. No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked. 20 SECTION-A 20 1. Attempt all parts: 1 1-a. Which platform is used to run hadoop? (CO1) 1 (a) Debian 1 (b) Cross platform 1 (c) Bare metal 1 | | | |
| Time: 3 Hours Max. Marks: 100 General Instructions: IMP: Verify that you have received the question paper with the correct course, code, branch etc. I. I. This Question paper comprises of three Sections -A, B, & C. It consists of Multiple Choice Questions (MCQ's) & Subjective type questions. Maximum marks for each question are indicated on right -hand side of each question. 3. Illustrate your answers with neat sketches wherever necessary. Assume suitable data if necessary. 5. Preferably, write the answers in sequential order. 6. No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked. SECTION-A 20 1. Attempt all parts:- 1 1-a. Which platform is used to run hadoop? (CO1) 1 (a) Debian 1 (b) Cross platform 1 (c) Bare metal 1 (d) Unix like 1 1-b. Identify the different features of Big Data Analytics. (CO1) 1 (a) Open source .(CO2) 1 (b) Data recovery .(CO2) 1 (c) Scalability .(CO2) 1 (d) All of the above | | | |
| General Instructions: IMP: Verify that you have received the question paper with the correct course, code, branch etc. 1. This Question paper comprises of three Sections -A, B, & C. It consists of Multiple Choice Questions (MCQ's) & Subjective type questions. 2. Maximum marks for each question are indicated on right -hand side of each question. 3. Illustrate your answers with neat sketches wherever necessary. 4. Assume suitable data if necessary. 5. Preferably, write the answers in sequential order. 6. No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked. SECTION-A 20 1. Attempt all parts:- 20 1-a. Which platform is used to run hadoop? (CO1) 1 (a) Debian 1 (b) Cross platform 20 (c) Bare metal 20 (d) Unix like 1 1-b. Identify the different features of Big Data Analytics. (CO1) 1 (a) Open source (b) 1 (b) Data recovery (c) Scalability (c) Scalability (d) All of the above 1-c. Every TaskTracker is configured with a(CO2) 1< | Tim | ne: 3 F | |
| 1. This Question paper comprises of three Sections -A, B, & C. It consists of Multiple Choice Questions (MCQ's) & Subjective type questions. 2. Maximum marks for each question are indicated on right -hand side of each question. 3. Illustrate your answers with neat sketches wherever necessary. 4. Assume suitable data if necessary. 5. Preferably, write the answers in sequential order. 6. No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked. SECTION-A 20 1. Attempt all parts:- 1 1-a. Which platform is used to run hadoop? (CO1) 1 (a) Debian (b) Cross platform 1 (c) Bare metal (c) Unix like 1 1-b. Identify the different features of Big Data Analytics. (CO1) 1 (a) Open source (c) Scalability 1 (d) All of the above | | | |
| Questions (MCQ's) & Subjective type questions. 2. Maximum marks for each question are indicated on right -hand side of each question. 3. Illustrate your answers with neat sketches wherever necessary. 4. Assume suitable data if necessary. 5. Preferably, write the answers in sequential order. 6. No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked. SECTION-A 20 1. Attempt all parts:- 1 1-a. Which platform is used to run hadoop? (CO1) 1 (a) Debian 1 (b) Cross platform 1 (c) Bare metal 1 (d) Unix like 1 1-b. Identify the different features of Big Data Analytics. (CO1) 1 (a) Open source 1 (b) Data recovery (c) Scalability (c) Scalability 1 1 (d) All of the above 1 1 1-c. Every TaskTracker is configured with a | IMP: | Verif | y that you have received the question paper with the correct course, code, branch etc. |
| 2. Maximum marks for each question are indicated on right -hand side of each question. 3. Illustrate your answers with neat sketches wherever necessary. 4. Assume suitable data if necessary. 5. Preferably, write the answers in sequential order. 6. No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked. SECTION-A 20 1. Attempt all parts:- 1-a. Which platform is used to run hadoop? (CO1) (a) Debian (b) Cross platform (c) Bare metal (d) Unix like 1-b. Identify the different features of Big Data Analytics. (CO1) (a) Open source (b) Data recovery (c) Scalability (d) All of the above 1-c. Every TaskTracker is configured with a(CO2) 1 (a) Maximum memory available in the node (b) Not limited (c) set of slots (d) As decided by the job tracker | | - | |
| 3. Illustrate your answers with neat sketches wherever necessary. 4. Assume suitable data if necessary. 5. Preferably, write the answers in sequential order. 6. No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked. 20 1. Attempt all parts:- 1-a. Which platform is used to run hadoop? (CO1) (a) Debian (b) Cross platform (c) Bare metal (d) Unix like 1-b. Identify the different features of Big Data Analytics. (CO1) (a) Open source (b) Data recovery (c) Scalability (d) All of the above 1-c. Every TaskTracker is configured with a(CO2) 1 (a) Maximum memory available in the node (b) Not limited (c) set of slots (d) As decided by the job tracker | | | |
| 4. Assume suitable data if necessary. 5. Preferably, write the answers in sequential order. 6. No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked. SECTION-A 20 1. Attempt all parts:- 1-a. Which platform is used to run hadoop? (CO1) (a) Debian (b) Cross platform (c) Bare metal (d) Unix like 1-b. Identify the different features of Big Data Analytics. (CO1) (a) Open source (b) Data recovery (c) Scalability (d) All of the above 1-c. Every TaskTracker is configured with a(CO2) 1 (a) Maximum memory available in the node (b) Not limited (c) set of slots (d) As decided by the job tracker 1-d. Hadoop is shared-nothing architecture (CO3) 1 | | | |
| 5. Preferably, write the answers in sequential order. 6. No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked. SECTION-A 20 1. Attempt all parts:- 1-a. Which platform is used to run hadoop? (CO1) (a) Debian (b) Cross platform (c) Bare metal (d) Unix like 1-b. Identify the different features of Big Data Analytics. (CO1) (a) Open source (b) Data recovery (c) Scalability (d) All of the above 1-c. Every TaskTracker is configured with a(CO2) (a) Maximum memory available in the node (b) Not limited (c) set of slots (d) As decided by the job tracker 1-d. Hadoop is shared-nothing architecture (CO3) | | | |
| evaluated/checked. 20 SECTION-A 20 1. Attempt all parts:- 1 1-a. Which platform is used to run hadoop? (CO1) 1 (a) Debian 1 (b) Cross platform 1 (c) Bare metal 1 (d) Unix like 1 1-b. Identify the different features of Big Data Analytics. (CO1) 1 (a) Open source 1 (b) Data recovery 1 (c) Scalability 1 (d) All of the above 1 1-c. Every TaskTracker is configured with a(CO2) 1 (a) Maximum memory available in the node 1 (b) Not limited 1 (c) set of slots 1 (d) As decided by the job tracker 1 1-d. Hadoop is shared-nothing architecture (CO3) 1 | | | |
| SECTION-A 20 1. Attempt all parts:- 1 1-a. Which platform is used to run hadoop? (CO1) 1 (a) Debian (b) Cross platform (c) Bare metal (d) Unix like 1-b. Identify the different features of Big Data Analytics. (CO1) 1 (a) Open source (b) Data recovery (c) Scalability (d) All of the above 1-c. Every TaskTracker is configured with a(CO2) 1 (a) Maximum memory available in the node (b) Not limited . (c) set of slots . (d) As decided by the job tracker 1 | | | · · · |
| 1. Attempt all parts:- 1-a. Which platform is used to run hadoop? (CO1) (a) Debian (b) Cross platform (c) Bare metal (d) Unix like 1-b. Identify the different features of Big Data Analytics. (CO1) 1 (a) Open source (b) Data recovery 1 (c) Scalability 1 (d) All of the above 1 1-c. Every TaskTracker is configured with a(CO2) 1 (a) Maximum memory available in the node 1 (b) Not limited 1 (c) set of slots 1 (d) As decided by the job tracker 1 1-d. Hadoop is shared-nothing architecture (CO3) 1 | evalue | ated/c | hecked. |
| 1. Attempt all parts:- 1-a. Which platform is used to run hadoop? (CO1) (a) Debian (b) Cross platform (c) Bare metal (d) Unix like 1-b. Identify the different features of Big Data Analytics. (CO1) 1 (a) Open source (b) Data recovery 1 (c) Scalability 1 (d) All of the above 1 1-c. Every TaskTracker is configured with a(CO2) 1 (a) Maximum memory available in the node 1 (b) Not limited 1 (c) set of slots 1 (d) As decided by the job tracker 1 1-d. Hadoop is shared-nothing architecture (CO3) 1 | SECI | TION- | - A 20 |
| 1-a. Which platform is used to run hadoop? (CO1) (a) Debian (b) Cross platform (c) Bare metal (d) Unix like 1-b. Identify the different features of Big Data Analytics. (CO1) (a) Open source (b) Data recovery (c) Scalability (d) All of the above 1-c. Every TaskTracker is configured with a(CO2) 1 (a) Maximum memory available in the node (b) Not limited (c) set of slots (d) As decided by the job tracker 1-d. Hadoop is shared-nothing architecture (CO3) | | | |
| (a) Debian (b) Cross platform (c) Bare metal (d) Unix like 1-b. Identify the different features of Big Data Analytics. (CO1) (a) Open source (b) Data recovery (c) Scalability (d) All of the above 1-c. Every TaskTracker is configured with a(CO2) 1 (a) Maximum memory available in the node (b) Not limited (c) set of slots (d) As decided by the job tracker 1-d. Hadoop is shared-nothing architecture (CO3) | | - | |
| (b) Cross platform (c) Bare metal (d) Unix like 1-b. Identify the different features of Big Data Analytics. (CO1) (a) Open source (b) Data recovery (c) Scalability (d) All of the above 1-c. Every TaskTracker is configured with a(CO2) 1 (a) Maximum memory available in the node (b) Not limited (c) set of slots (d) As decided by the job tracker 1-d. Hadoop is shared-nothing architecture (CO3) | 1 a. | | |
| (c) Bare metal (d) Unix like 1-b. Identify the different features of Big Data Analytics. (CO1) (a) Open source (b) Data recovery (c) Scalability (d) All of the above 1-c. Every TaskTracker is configured with a(CO2) 1 (a) Maximum memory available in the node (b) Not limited (c) set of slots (d) As decided by the job tracker 1-d. Hadoop is shared-nothing architecture (CO3) | | . , | |
| (d) Unix like1-b.Identify the different features of Big Data Analytics. (CO1)1(a) Open source(b) Data recovery(c) Scalability(d) All of the above(d) All of the above1-c.Every TaskTracker is configured with a(CO2)1(a) Maximum memory available in the node(b) Not limited(c) set of slots(d) As decided by the job tracker1 | | | - |
| 1-b. Identify the different features of Big Data Analytics. (CO1) (a) Open source (b) Data recovery (c) Scalability (d) All of the above 1-c. Every TaskTracker is configured with a(CO2) 1 (a) Maximum memory available in the node (b) Not limited (c) set of slots (d) As decided by the job tracker 1-d. Hadoop is shared-nothing architecture (CO3) | | | |
| (a) Open source (b) Data recovery (c) Scalability (d) All of the above 1-c. Every TaskTracker is configured with a(CO2) 1 (a) Maximum memory available in the node (b) Not limited (c) set of slots (d) As decided by the job tracker 1-d. Hadoop is shared-nothing architecture (CO3) 1 | 1-h | | |
| (b) Data recovery (c) Scalability (d) All of the above 1-c. Every TaskTracker is configured with a(CO2) 1 (a) Maximum memory available in the node (b) Not limited (c) set of slots (d) As decided by the job tracker 1-d. Hadoop is shared-nothing architecture (CO3) 1 | 1 0. | | |
| (c) Scalability (d) All of the above 1-c. Every TaskTracker is configured with a(CO2) 1 (a) Maximum memory available in the node (b) Not limited (c) set of slots (d) As decided by the job tracker 1-d. Hadoop is shared-nothing architecture (CO3) 1 | | . , | |
| (d) All of the above1-c.Every TaskTracker is configured with a(CO2)(a) Maximum memory available in the node(b) Not limited(c) set of slots(d) As decided by the job tracker1-d.Hadoop is shared-nothing architecture (CO3)1 | | . , | |
| 1-c. Every TaskTracker is configured with a(CO2) 1 (a) Maximum memory available in the node .(CO2) 1 (b) Not limited | | | |
| (a) Maximum memory available in the node (b) Not limited (c) set of slots (d) As decided by the job tracker 1-d. Hadoop is shared-nothing architecture (CO3) | 1-c | . , | |
| (b) Not limited (c) set of slots (d) As decided by the job tracker 1-d. Hadoop is shared-nothing architecture (CO3) | 1 0. | | |
| (c) set of slots (d) As decided by the job tracker 1-d. Hadoop is shared-nothing architecture (CO3) | | . , | - |
| (d) As decided by the job tracker1-d. Hadoop is shared-nothing architecture (CO3)1 | | . , | |
| 1-d.Hadoop is shared-nothing architecture (CO3)1 | | | |
| | 1-d | . , | |
| | 1 U. | | |
| (a) TROE (b) FALSE | | (a) (b) | |
| (c) Can't Say | | . , | |

•

•

| | (d) | May be | |
|---------------|--|---|---|
| 1-e. | W | Thich is known as horizontal scaling? (CO2) | 1 |
| | (a) | Scaling up | |
| | (b) | Scaling out | |
| | (c) | Scaling | |
| | (d) | Scaling down | |
| 1-f. | W | which one is name of the C++ interface to hadoop mapreduce? (CO3) | 1 |
| | (a) | Hadoop streaming | |
| | (b) | Hadoop Pipes | |
| | (c) | Mahaout | |
| | (d) | Oozie | |
| 1-g. | Which of these an example of SaaS? (CO4) | | |
| | (a) | Google Workspace | |
| | (b) | Dropbox | |
| | (c) | Salesforce | |
| | (d) | All of the above | |
| 1-h. | W | Thich is not a cloud stakeholder? (CO4) | 1 |
| | (a) | Cloud providers | |
| | (b) | Clients | |
| | (c) | End users | |
| | (d) | Cloud users | |
| 1-i. | | he combination of and is often termed the local address the local portion of the IP address. (CO5) | 1 |
| | (a) | Network number and host number | |
| | (b) | Network number and subnet number | |
| | (c) | Subnet number and host numbe | |
| | (d) | Host number | |
| 1 - j. | | Thich is the broadcast address for a Class B network ID using the default abnetmask? (CO5) | 1 |
| | (a) | 172.16.10. 255 | |
| | (b) | 255.255.255. 255 | |
| | (c) | 172.16.255. 255 | |
| | (d) | 172.255.255. 255 | |
| 2. Atte | | Ill parts:- | |
| 2.a. | - | Trite any two real time examples of big data. (CO1) | 2 |
| 2.b. | | xplain the limitation of traditional approach in big data process? (CO2) | 2 |
| 2.c. | | an you explain the components of HDFS architecture.(CO3) | 2 |
| 2.d. | | Trite characteristics of Cloud Computing. (CO4) | 2 |
| <i>∠</i> .u. | ** | The endlowed of cloud computing. (COT) | 4 |

•

•

Page 2 of 3

| 2.e. | Define subnet. (CO5) | 2 | | | |
|---|--|----|--|--|--|
| SECTION-B | | | | | |
| | ver any <u>five</u> of the following:- | | | | |
| 3-a. | Can you explain the digital data types and describe the data classification features? (CO1) | б | | | |
| 3-b. | Explain in detail about Apache hadoop daemons, snowflake and AWS. (CO1) | 6 | | | |
| 3-c. | Define HDFS. Also explain components of hadoop. (CO2) | 6 | | | |
| 3-d. | Explain all the data formats available in hadoop? (CO2) | 6 | | | |
| 3.e. | Define HDFS with some of it's benefits and challenges.(CO3) | 6 | | | |
| 3.f. | Define open stack. Also define it's components. (CO4) | 6 | | | |
| 3.g. | Explain in detail the concept of Subneting. Explain. (CO5) | 6 | | | |
| SECTION-C 50 | | | | | |
| 4. Ansv | ver any <u>one</u> of the following:- | | | | |
| 4-a. | Explain in detail about the cloud computing evolution and draw its proper diagram? (CO1) | 10 | | | |
| 4-b. | Explain about the decrease replication factor of HDFS? Explain what is happen when a file delete and undelete in HDFS? (CO1) | 10 | | | |
| 5. Answer any <u>one</u> of the following:- | | | | | |
| 5-a. | Can you explain in detail about the HDFS architecture. (CO2) | 10 | | | |
| 5-b. | Explain block and block size and why is it needed in hadoop? (CO2) | 10 | | | |
| 6. Answer any <u>one</u> of the following:- | | | | | |
| 6-a. | How does Namenode handles Datanode failure in HDFS? Briefly explain the functions of block? (CO3) | 10 | | | |
| 6-b. | Explain the roles of Namenode, Datanode, Secodary Namenode and ResourceManager? (CO3) | 10 | | | |
| 7. Answer any <u>one</u> of the following:- | | | | | |
| 7-a. | List out features and challenges of GCP and explain in detail the features of any four services provided by GCP? (CO4) | 10 | | | |
| 7-b. | Can you explain how the MapReduce job run in hadoop with an example? (CO4) | 10 | | | |
| 8. Answer any <u>one</u> of the following:- | | | | | |
| 8-a. | Discuss the BigTable concept in cloud with an example? (CO5) | 10 | | | |
| 8-b. | Explain BigQuery with all the basic concepts of BigQuery.CO5 | 10 | | | |
| | | | | | |

•

•