

Course Name – Cloud Computing

Course Code - BCA50115

Multiple Choice Type Questions

1.

(i) If switches are used to replace hubs on a network, which of the following statements is true?

- | | |
|--|--|
| a) The number of broadcast domains will increase | b) The number of collision domain will increase |
| c) The number of collision domains will decrease | d) The number of broadcast domains will decrease |

(ii) Choose the correct difference between a switch and a hub in networking.

- | | |
|---|---|
| a) A switch broadcasts data to all ports | b) A hub is smarter than a switch |
| c) A switch sends data only to the destination port | d) A switch connects directly to the internet |

(iii) Which of the following function of the Data Link Layer in the OSI model?

- | | |
|--|---------------------------|
| a) Manages file compression | b) Handles IP addressing |
| c) Provides node-to-node data transfer | d) Controls email routing |

(iv) Choose the correct option: Differentiate between TCP and UDP is _____.

- | | |
|---|--|
| a) TCP is connectionless; UDP is connection-oriented | b) Both guarantee delivery |
| c) TCP provides reliable communication; UDP is faster | d) UDP is only used in cloud computing |

(v) Tell which of the following is a network protocol.

- | | |
|--------|---------|
| a) RAM | b) HTML |
| c) CPU | d) FTP |

(vi) Which network covers a larger area between LAN and MAN?

- | | |
|--------------------------------------|---|
| a) LAN covers a larger area than MAN | b) MAN connects networks across countries |
|--------------------------------------|---|

- c) LAN is limited to a building; MAN covers a city d) Both are used only in cloud computing

(vii) Tell the purpose of IP addressing in networking.

- a) To name software applications b) To ensure unique device identification and communication
c) To provide login credentials d) To convert files into binary

(viii) What is the disadvantage of using a connection-oriented protocol such as TCP?

- a) Packet acknowledgement might add overhead b) Packets are not tagged with sequence number
c) Loss or duplication of data packets is more likely to occur d) The application layer must assume responsibility for the correct sequencing of the data packets

(ix) Which multiple access technique is used by IEEE 802.11 standard for wireless LAN?

- a) CDMA b) CSMA/CA
c) ALOHA d) None of these

(x) Which device allows wireless devices to connect to a wired network using Wi-Fi or related standards?

- a) Router b) Switch
c) Access Point d) None of these

(xi) Location transparency allows: I. Users to treat the data as if it is done at one location. II. Programmers to treat the data as if it is at one location. III. Managers to treat the data as if it is at one location. Which one of the following is correct?

- a) I, II, III b) I and II only
c) II and III only d) II only

(xii) In a distributed computing environment, distributed shared memory is used which is_____.

- a) Logical combination of virtual memories on the nodes. b) Logical combination of physical memories on the nodes.
c) Logical combination of the secondary memories on all the nodes. d) Both 1 and 3.

(xiii) Which of the following is a challenge in designing distributed systems?

- a) Ensuring data consistency b) Managing network latency
c) Handling concurrency issues d) All are correct.

(xiv) Which of the following is a key characteristic of a distributed system?

- a) Centralized control b) Single point of failure
c) Resource sharing d) Tight coupling between components

(xv) What is the purpose of a \"consensus algorithm\" in a distributed system?

- a) To ensure all nodes agree on a single value
- b) To replicate data across multiple nodes
- c) To detect and recover from failures
- d) To manage network traffic

(xvi) Which of the following is a characteristic of a \"loosely coupled\" distributed system?

- a) High degree of inter-component dependency
- b) Components can operate independently with minimal coordination
- c) Components rely on a central authority for coordination
- d) Components share a common memory space

(xvii) Choose the term that describes the ability of a system to adapt to increased service load.

- a) Capacity
- b) Tolerance
- c) Scalability
- d) None of these

(xviii) What is the core technology that enables cloud computing to abstract and virtualize resources, allowing them to be pooled and shared?

- a) Blockchain
- b) Quantum computing
- c) Virtualization
- d) Augmented reality

(xix) Who is the father of cloud computing?

- a) Sharon B. Codd
- b) Edgar Frank Codd
- c) J.C.R. Licklider
- d) Charles Bachman

(xx) Which of the following is an example of a PaaS cloud service?

- a) Heroku
- b) AWS Elastic Beanstalk
- c) Windows Azure
- d) All of the above

(xxi) Which of the following is an example of an IaaS Cloud service?

- a) DigitalOcean
- b) Linode
- c) Rackspace
- d) All of the above

(xxii) Point out the wrong statement.

- a) Azure enables .NET Framework applications to run over the Internet
- b) Cloud Computing has two distinct sets of models
- c) Amazon has built a worldwide network of datacenters to service its search engine
- d) None of the mentioned

(xxiii) Which of the following model attempts to categorize a cloud network based on four dimensional factors?

- a) Cloud Cube
- b) Cloud Square
- c) Cloud Service
- d) All of the mentioned

(xxiv) All cloud computing applications suffer from the inherent _____ that is intrinsic in their WAN connectivity.

- a) noise
- b) propagation
- c) latency
- d) all of the mentioned

(xxv) Which of the following is the correct statement about cloud types?

- a) Cloud Square Model is meant to show is that the traditional notion of a network boundary being the network's firewall no longer applies in cloud computing
- b) A deployment model defines the purpose of the cloud and the nature of how the cloud is located **(Correct Answer)**
- c) Service model defines the purpose of the cloud and the nature of how the cloud is located
- d) All of the mentioned

(xxvi) Identify the correct statement about cloud computing.

- a) Cloud computing relies on a set of protocols needed to manage interprocess communications
- b) Platforms are used to create more complex software
- c) Cloud architecture can couple software running on virtualized hardware in multiple locations to provide an on-demand service
- d) All of the mentioned

(xxvii) Which of the following is the wrong statement about cloud computing?

- a) Private cloud doesn't employ the same level of virtualization
- b) Data center operates under average loads
- c) Private cloud doesn't pooling of resources that a cloud computing provider can achieve
- d) Abstraction enables the key benefit of cloud computing: shared, ubiquitous access

(xxviii) Which of the following is the most essential element in cloud computing by CSA?

- a) Virtualization
- b) Multi-tenancy
- c) Identity and access management
- d) All of the mentioned

(xxix) Which of the following monitors the performance of the major cloud-based services in real time in Cloud Commons?

- a) CloudWatch
- b) CloudSensor
- c) CloudMetrics
- d) All of the mentioned

(xxx) How distributed and centralized systems differ from each other?

- a) Both use one central node
- b) Distributed systems spread tasks across multiple machines
- c) Centralized systems are cloud-based
- d) Distributed systems do not allow communication

(xxxi) How many phases are there in Cloud Computing Planning?

- a) 1
- b) 2

c) 3

d) 4

(xxxii) In distributed systems, links and site failure is detected by

_____.

a) Polling

b) Handshaking

c) Token Passing

d) None of these

(xxxiii) Choose how fault tolerance is achieved in distributed systems.

a) By shutting down all nodes

b) Through data loss

c) By replication and redundancy

d) By using only one central processor

(xxxiv) Which of the following is a workflow control and policy based automation service by CA?

a) CA Cloud Compose

b) CA Cloud Insight

c) CA Cloud Optimize

d) CA Cloud Orchestrate

(xxxv) The Amazon Machine Images are basically virtual appliances that are packaged for running on the _____ nodes grid.

a) Ken

b) Ben

c) Zen

d) Xen

(xxxvi) Out of these combinations, which one is possible in the WebSphere ESB?

a) SOAP/HTTP

b) Governance

c) SOAP/JMS

d) XML/JMS

(xxxvii) Virtualization contributes to Green Computing by:_____

a) Reducing the system's efficiency

b) Reducing the number of physical hardware used

c) Increasing physical hardware uses

d) Decreasing the requirement for individual software purchases

(xxxviii) Which are the three conceptual layers in CometCloud?

a) development, application, physical

b) infrastructure, dynamic, static

c) transport, network, programming

d) programming, service, infrastructure

(xxxix) Private cloud define as_____.

a) A standard cloud service offered via the Internet

b) A cloud architecture maintained within an enterprise data center

c) A cloud service inaccessible to anyone but the cultural elite

d) None of these

(xl) Give an example of a distributed system in real life.

a) A standalone game

b) A local calculator

c) Google Docs

d) MS Paint

(xli) Which of the following benefit is provided by the PaaS service provider?

- a) A larger pool of qualified developers
- b) More reliable operation
- c) A logical design methodology
- d) All of the mentioned

(xlii) Which of the following is the most refined and restrictive cloud service model?

- a) PaaS
- b) IaaS
- c) SaaS
- d) CaaS

(xliii) Which of the following is an example of a SaaS cloud service?

- a) Google Workspace
- b) Microsoft Azure Virtual Machines
- c) Kubernetes Engine
- d) Amazon EC2

(xliv) Which of the following is the Virtual machine conversion cloud?

- a) Amazon CloudWatch
- b) AbiCloud
- c) BMC Cloud Computing Initiative
- d) None of the mentioned

(xlv) Approximate the market share of AWS in cloud services is _____

- a) ~5%
- b) ~10%
- c) ~30%
- d) ~50%

(xlvi) Choose the estimated cost in a utility computing model if 10 hours of VM usage cost \$0.05/hr.

- a) \$1
- b) \$0.50
- c) \$2
- d) \$5

(xlvii) Choose the correct interpretation of the SLA clause: \"99.9% uptime guarantee.\" What does it imply?

- a) Unlimited usage
- b) 1-hour downtime/year approx.
- c) Only 5% uptime
- d) Free support

(xlviii) Use IaaS to deploy a scalable web app. What will be your first step?

- a) Choose frontend tools
- b) Rent compute resources
- c) Install antivirus
- d) Upload logs

(xlix) Which is the cloud monitoring tools for detecting service downtime?

- a) Google Forms
- b) Google Analytics
- c) AWS CloudWatch
- d) Microsoft Word

(l) Choose the architecture of a hybrid cloud deployment model.

- a) Only internal resources shown
- b) Combination of private and public resources
- c) Only public services shown
- d) Flat network

(li) Choose the number of users a server can handle if each user requires 0.5GB RAM and the server has 16GB RAM.

- a) 10
- b) 20
- c) 32
- d) 40

(lii) Choose how dynamic scalability works in a real-time video streaming app.

- a) Manual reboot
- b) Adding VM instances as needed
- c) Fixed data rate
- d) Single user handling

(liii) Choose a resource pooling architecture in cloud services.

- a) Isolated servers
- b) Shared pool of computing/storage/network resources
- c) One server per app
- d) Flat directory

(liv) Interpret metrics from a cloud monitoring dashboard. Choose a correct observation.

- a) CPU usage 90% = idle
- b) Disk I/O = 0 = active
- c) High latency = network issue
- d) High RAM = low workload

(lv) Choose the concept of IaaS to reduce infrastructure cost.

- a) Buy new servers
- b) Rent virtualized servers
- c) Install on-prem servers
- d) Cancel cloud services

(lvi) Choose the correct storage usage in cloud if each backup is 2GB and 15 backups are stored.

- a) 25GB
- b) 30GB
- c) 50GB
- d) 10GB

(lvii) Choose how SOAP enables service communication.

- a) Encrypts traffic
- b) Transfers email
- c) Defines XML-based messaging protocol
- d) Builds UI

(lviii) Choose the appropriate action to solve the mismatch between expected and actual SLA.

- a) Ignore logs
- b) Review SLA metrics and raise ticket
- c) Switch DNS
- d) Buy more RAM

(lix) Choose the remote administration tools used to restart failed services in a cloud environment.

- a) Remote desktop or SSH
- b) Local OS
- c) USB boot
- d) Mail alerts

(Ix) Which platform scaling in a PaaS model is for mobile apps?

- a) Fix app bugs
- b) Re-deploy to auto-scale platforms
- c) Uninstall platform
- d) Hard code servers

(Ixi) Operate a resource monitoring system to detect overload. What will help?

- a) RAM usage graph
- b) Excel sheet
- c) Router logs
- d) Clock time

(Ixii) What is the difference between SaaS and IaaS via deployment diagram?

- a) Both show VM only
- b) SaaS = full app; IaaS = infra layer only
- c) Both have APIs only
- d) Same layout

(Ixiii) What is the meaning of “cloud elasticity” from a usage graph?

- a) Same load all time
- b) Load drops = less resources auto-used
- c) Constant memory use
- d) Only input bandwidth matters

(Ixiv) What is the function of SLA management system in a real life scenario?

- a) Tracks file access
- b) Ensures service-level agreement is enforced with penalties
- c) Runs auto-emails
- d) Schedules appointments

(Ixv) Which scenario is best suited for hybrid cloud deployment?

- a) A local gaming app
- b) Government confidential data only
- c) A mix of sensitive and non-sensitive workloads
- d) Personal email hosting

(Ixvi) What is the concept of workload distribution in cloud?

- a) Send all load to one server
- b) Evenly distribute processing among multiple resources
- c) Limit processing to edge devices
- d) Pause all workloads temporarily

(Ixvii) _____ & _____ are cloud deployment models as per categories.

- a) Open and closed clouds
- b) Fixed and dynamic clouds
- c) Public, private and hybrid clouds
- d) Central and distributed clouds

(Ixviii) What is the meaning of cloud bursting?

- a) Buying new servers
- b) Bursting data to the public
- c) Expanding workload temporarily to external cloud
- d) Running out of memory

(Ixix) What is the importance of dynamic scalability in the cloud?

- a) To manually upgrade servers
- b) To reduce load on user PCs

- c) To auto-adjust resources as demand changes
- d) To increase latency

(lxx) Choose the correct concept of elasticity in cloud computing.

- a) Static allocation of resources
- b) Changing IP addresses often
- c) Automatic expansion/shrinking of resources
- d) Slowing down VMs

(lxxi) Choose the purpose of service load balancing in cloud platforms.

- a) Limiting services
- b) Evenly distributing traffic among services
- c) Encrypting all data
- d) Stopping all services

(lxxii) Choose the correct distinction between horizontal and vertical scaling in cloud computing.

- a) Horizontal = add power; vertical = add nodes
- b) Both are the same
- c) Horizontal = add nodes; vertical = add power/resources
- d) Vertical only exists in physical servers

(lxxiii) Infer the key benefit of cloud-based application deployment.

- a) Higher hardware cost
- b) Scalability, remote access, speed
- c) Offline-first application model
- d) Local-only software

(lxxiv) What is the role of SOAP in service-oriented architecture?

- a) Frontend UI
- b) Defines web service protocols for communication
- c) Manages email clients
- d) Executes scripts offline

(lxxv) Which of the following function is service middleware?

- a) Hardware driver manager
- b) Software glue enabling services to communicate
- c) Firewall interface
- d) IP translator

(lxxvi) In reference to Big Data, consider the following databases: (A) Memcached (B) CouchDB (C) InfiniteGraph Choose the most appropriate answer from the options given below:

- a) Marketing plan
- b) Storage policy
- c) Framework for delivering computing services
- d) Installation script

(lxxvii) _____ is the classifications of IaaS, PaaS, and SaaS by service levels.

- a) Network, storage, CPU
- b) Platform, GUI, network
- c) Infrastructure, platform, software services
- d) Code, design, deploy

(lxxviii) Tell what is provided by IaaS to users.

- a) Ready-to-use applications
- b) App hosting platforms only
- c) Virtualized hardware infrastructure
- d) Code repositories

(lxxix) Why SaaS is ideal for non-technical users?

- a) It requires Linux knowledge
- b) It's free software
- c) It offers ready-to-use apps over the internet
- d) It requires cloud programming

(lxxx) What PaaS allows developers to do?

- a) Write kernel modules
- b) Build and deploy apps without managing infrastructure
- c) Manually manage cloud servers
- d) Patch firewalls

(lxxxi) Translate the phrase "administering cloud services" as _____,

- a) Watching video tutorials
- b) Managing, monitoring, and configuring cloud environments
- c) Opening emails
- d) Playing multiplayer games

(lxxxii) Which of the following is different between IaaS and SaaS?

- a) Both offer only data storage
- b) IaaS gives infrastructure; SaaS gives end-user software
- c) SaaS is offline; IaaS is online only
- d) IaaS offers databases directly

(lxxxiii) Outline which steps are involved in monitoring cloud resources.

- a) Logging out users
- b) Checking dashboards, alerts, usage metrics
- c) Turning off VMs randomly
- d) Installing on-premise servers

(lxxxiv) Choose how utility computing is linked to cloud pricing.

- a) Monthly fixed payment
- b) Pay-per-use based on resource consumption
- c) Free-tier use
- d) Unlimited subscriptions

(lxxxv) _____ the goal of elastic computing in cloud environments.

- a) Make systems rigid
- b) Reduce bandwidth
- c) Adjust computing power dynamically
- d) Boost energy usage

(lxxxvi) Which is the purpose of a resource management system?

- a) Stores passwords
- b) Allocates and optimizes use of cloud resources
- c) Encrypts emails
- d) Deletes files

(lxxxvii) Choose the correct classification of cloud computing platforms based on their function.

- a) SaaS for infrastructure, PaaS for storage, IaaS for games
- b) IaaS for infrastructure, PaaS for application development, SaaS for end-user software
- c) PaaS for end-user apps, SaaS for virtualization, IaaS for databases
- d) IaaS for coding tools, SaaS for networking, PaaS for emails

(lxxxviii) What is the SLA in a cloud context?

- a) Server Log App
- b) Security Log Audit
- c) Agreement on service quality and availability
- d) System Language Adapter

(lxxxix) SLA management is vital for enterprises because its_____.

- a) Avoids vendor lock-in
- b) Ensures performance guarantees, uptime, and penalties
- c) Allows free internet
- d) Tracks customer IDs

(xc) Tell the function of remote administration in the cloud.

- a) Control systems from any location via network tools
- b) Manually open data centers
- c) Run machines offline
- d) Turn off firewalls

(xci) Tell the role of a data management system in cloud computing.

- a) Manages fonts
- b) Tracks disk errors
- c) Handles storage, backup, replication, and access
- d) Deletes old apps

(xcii) What is the different between monitoring and administration?

- a) Both are unrelated
- b) Monitoring = observe; Admin = configure & control
- c) Admin only backs up data
- d) Monitoring is always online

(xciii) Outline key limitations of cloud computing.

- a) Unlimited bandwidth
- b) Manual scaling only
- c) Data security, internet dependency, vendor lock-in
- d) Free lifetime use

(xciv) Which of the following is NOT a common \"V\" of Big Data?

- a) Volume
- b) Velocity
- c) Veracity
- d) Validity

(xcv) What is the primary purpose of Hadoop Distributed File System (HDFS)?

- a) To manage relational databases
- b) To provide a distributed, fault-tolerant storage system for large files **(Correct Answer)**
- c) To process real-time streaming data
- d) To run SQL queries on structured data

(xcvi) Which of the following is a key characteristic of NoSQL databases compared to traditional relational databases?

- a) Strict schema enforcement
- b) Horizontal scalability and flexible schemas
- c) ACID compliance for all operations
- d) Primarily used for small, structured datasets

(xcvii) Google File System (GFS) is a precursor to which open-source distributed file system?

- a) Apache Cassandra
- b) Apache Spark
- c) Hadoop Distributed File System (HDFS)
- d) MongoDB

(xcviii) Which of the following is a common issue encountered when handling large datasets (Big Data)?

- a) Limited storage capacity on a single machine
- b) Difficulty in processing and analyzing data in a timely manner
- c) Challenges in data governance and security
- d) All of the above

(xcix) Which type of NoSQL database stores data in key-value pairs?

- a) Document-oriented
- b) Graph
- c) Key-value
- d) Column-family

(c) What is a \"big data spreadsheet\" often used for?

- a) Storing and analyzing small, structured datasets on a single machine
- b) Providing a visual interface for exploring and manipulating large datasets, often backed by distributed systems **(Correct Answer)**
- c) Performing complex statistical modeling on real-time data streams
- d) Managing traditional relational database tables

(ci) Which of the following scenario may not be a good fit for HDFS in the Big Data paradigm?

- a) HDFS is not suitable for scenarios requiring multiple/simultaneous writes to the same file
- b) HDFS is suitable for storing data related to applications requiring low latency data access
- c) HDFS is suitable for storing data related to applications requiring high latency data access
- d) None of the above

(cii) Hadoop (a big data tool) works with a number of related tools. Choose from the following the common tools included in Hadoop:

- a) MySQL, Google API and MapReduce
- b) MapReduce, Scala and Hummer
- c) MapReduce, HBase and Hive
- d) MapReduce, Hummer and Heron

(ciii) The data node and name node in HADOOP are:

- a) Worker Node and Master Node respectively
- b) Both Worker Nodes

c) oth Master Nodes

d) None of the above

(civ) Which of the following is a component of Hadoop?

a) YARN

b) HDFS

c) MapReduce

d) All of the options

(cv) Which of the following type of virtualization is found in hypervisors such as Microsoft's Hyper-V?

a) Paravirtualization

b) Full virtualization

c) Emulation

d) None of the mentioned

(cvi) What is HDFS Block in Hadoop?

a) It is the logical representation of data

b) It is the physical representation of data

c) Both of the above

d) None of the above

(cvii) What is HDFS?

a) Storage layer

b) Batch processing engine

c) Resource management layer

d) None of the above

(cviii) Which of the following statement/s is/are true? (i) Facebook has the world's largest Hadoop cluster. (ii) Hadoop 2.0 allows live stream processing of real-time data

a) Neither (i) nor (ii)

b) Both (i) and (ii)

c) (i) only

d) (ii) only

(cix) What is the term used for a collection of large, complex data sets that cannot be processed using traditional data processing tools?

a) Big Data

b) Small Data

c) Mini Data

d) Medium Data

(cx) What is the process of transforming structured and unstructured data into a format that can be easily analyzed?

a) Data Mining

b) Data Processing

c) Data Intregation

d) Data Warehousing

(cxi) Which of the following is a tool used for processing and analyzing Big Data?

a) PostgreSQL

b) Hadoop

c) MySQL

d) Oracle

(cxii) Which of the following is not a common challenge associated with Big Data?

a) Data Quality

b) Data Integration

c) Data Privacy

d) Data Duplication

(cxiii) Apply the map-reduce model to count words in a document. What phase comes first?

a) Reduce

b) Filter

c) Map

d) Shuffle

(cxiv) Calculate the total HDFS storage needed for a 1GB file with 3 replicas.

a) 1GB

b) 2GB

c) 3GB

d) 4GB

(cxv) Infer how parallelism helps speed up matrix multiplication.

a) Decreases speed

b) Processes parts simultaneously

d) Adds complexity only

c) Needs more memory

(cxvi) Choose the most fault-tolerant model to solve a big data storage problem.

a) NoSQL

b) RDBMS

c) HDFS

d) CSV

(cxvii) Use HDFS when_____.

a) Files are small

b) Real-time response is critical

c) Batch processing is needed

d) RAM is high

(cxviii) Choose the best approach to implement a distributed solution for sorting 1TB of data.

a) Single-threaded sorting

b) HDFS + MapReduce

c) Python loop

d) Manual indexing

(cxix) Choose the principle used to operate a distributed system.

a) Centralization

b) Redundancy

c) Single thread

d) Static IP

(cxx) Choose the two key components of the HDFS (Hadoop Distributed File System) architecture.

a) NameNode and DataNode

b) Server and Router

c) Block and Bit

d) Heap and Stack

(cxxi) Interpret the NoSQL data model from a JSON structure. What does it represent?

a) Array

b) Key-value pair

c) Columnar

d) Tuple

(cxxii) Illustrate the difference between GFS and HDFS.

a) HDFS is for Google

b) GFS is Hadoop's

c) GFS is proprietary, HDFS is open-source

d) Both are same

(cxxiii) Choose what is improved by applying the concept of sharding in NoSQL databases.

- a) Improves data encryption
- c) Improves query performance and scalability

- b) Improves user interface design
- d) Schema rigidity

(cxxiv) Calculate replication factor storage: File = 500MB, RF = 2. Total?

- a) 250MB
- c) 1GB

- b) 500MB
- d) 2GB

(cxxv) Which of the following is a big data spreadsheet that handles 10 million rows?

- a) Crashes
- c) Uses chunked memory model

- b) Loads slowly
- d) Requires reboot

(cxxvi) Which is solution of the bottleneck problem in distributed programming?

- a) Increase single-thread load
- c) Use one loop

- b) Use parallel tasks
- d) Disable threads

(cxxvii) Use Hadoop for processing which type of data?

- a) Small binary logs
- c) HTML content only

- b) Massive structured data
- d) Temporary RAM cache

(cxxviii) Implement parallelism in data filtering using...

- a) Manual typing
- c) MapReduce

- b) SQL only
- d) Email alerts

(cxxix) Choose the feature used to operate a fault-tolerant distributed storage system.

- a) Encryption only
- c) Manual backup

- b) Data replication
- d) Data compression

(cxxx) Choose the last phase in a high-level workflow of the MapReduce programming model.

- a) Map
- c) Reduce

- b) Shuffle
- d) Cleanup

(cxxxii) Interpret the error message in HDFS when DataNode fails.

- a) Backup enabled
- c) Data replicated

- b) Data loss
- d) File deleted

(cxxxiii) Illustrate how big data affects traditional storage methods.

- a) Makes it easier
- c) Requires distributed models

- b) Slows down access
- d) Reduces RAM

(cxxxiii) Apply a hash function in a NoSQL DB for_____.

- a) Indexing
- c) Replication

- b) Validation
- d) Sorting

(cxxxiv) Calculate map task distribution for 1GB file across 4 nodes.

- a) 4 tasks
- c) 1 task

- b) 2 tasks
- d) 8 tasks

(cxxxv) Scalability using distributed file systems as_____.

- a) Less data allowed
- c) Fewer users supported

- b) Adds more nodes easily
- d) Manual sync

(cxxxvi) Tell slow write problems in big data NoSQL.

- a) Use SSD
- c) Use write buffers

- b) Use indexes
- d) Disable write-ahead log

(cxxxvii) Use BigQuery or similar tools to process_____.

- a) 10 rows
- c) Petabyte-scale data

- b) 100 rows
- d) KB logs only

(cxxxviii) Implement a reducer in Hadoop that counts frequencies. What does it emit?

- a) Input key
- c) File name

- b) Sum of values
- d) Memory size

(cxxxix) Choose the correct use case to operate columnar storage for _____.

- a) Fast random read
- c) Fast analytics

- b) Reduced redundancy
- d) Faster email

(cxl) Tell which option represents parallel write architecture in distributed file systems.

- a) Sequential blocks
- c) Only NameNode writes

- b) Parallel disk writes
- d) FIFO writes

(cxli) Interpret Hadoop job failure. What might be the cause?

- a) Browser crash
- c) RAM upgrade

- b) Mapper issue
- d) HDD replaced

(cxlii) Choose the way for NoSQL handles schema changes.

- a) Fixed schema
- c) Flexible schema

- b) No schema at all
- d) Uses CSV only

(cxliii) Apply replication in GFS for_____

- a) Performance
- c) Fault tolerance

- b) Security
- d) Compression

(cxliv) Calculate which value is storage of overhead in HDFS with 3x replication for 2GB file.

- a) 4GB
- b) 6GB
- c) 2GB
- d) 3GB

(cxlv) Choose the role of NameNode in HDFS.

- a) Stores metadata
- b) Stores actual data
- c) Compresses file
- d) Schedules jobs

(cxlvi) Which of the following is consistency problem in parallel tasks?

- a) Use locks
- b) Remove cache
- c) Disable threads
- d) Use one core

(cxlvii) Use Spark instead of MapReduce when_____.

- a) You want slower performance
- b) You want real-time processing
- c) Batch-only is enough
- d) You use spreadsheets

(cxlviii) Implement big data analysis on...

- a) CPU-only system
- b) Local desktop
- c) Clustered architecture
- d) XML parser

(cxlix) Choose the purpose to operate distributed cache for _____purpose.

- a) Executing batch jobs
- b) Reducing database load and improving performance
- c) File sorting
- d) Compression

(cl) Multiple mappers work in Hadoop as_____.

- a) Single-thread
- b) Parallel threads
- c) Serial blocks
- d) RAM-only

(cli) Choose the correct interpretation: log reports from a Hadoop job where high Map time means _____.

- a) Efficiency
- b) Inefficient mapping or data skew
- c) Success
- d) Disk error

(clii) Choose the way spreadsheet-like tools such as Apache Zeppelin help in big data.

- a) By physically storing distributed data blocks across Hadoop clusters
- b) By visualizing, exploring, and analyzing big data interactively
- c) By replacing all backend processing engines like Spark and Hive
- d) By providing firewalls and security policies for protecting big data systems

(cliii) In the context of big data stores, which of the following best demonstrates how partitioning improves performance and scalability in distributed systems?

- a) By compressing large datasets to reduce storage requirements
- b) By duplicating data across multiple servers to ensure redundancy

c)By dividing a large dataset into smaller, manageable segments distributed across different nodes

d)By encrypting data to enhance security during transmission

Short Answer Type Questions

1. Analyze the limitations of traditional client-server models in handling large-scale distributed systems.
2. Investigate how packet loss affects TCP vs UDP-based cloud applications.
3. Distinguish between BUS Topology and Mesh Topology.
4. Break down how distributed consensus is achieved in cloud storage systems using Paxos or Raft protocols.
5. Interpret real-time cloud billing logs to identify inefficiencies in resource usage across 3 virtual machines.
6. Classify different network topologies and evaluate their suitability for distributed cloud environments.
7. Differentiate performance implications of vertical vs horizontal scaling in cloud systems.
8. Examine a hybrid deployment scenario and deduce its cost/performance trade-offs.
9. What is IP addressing.
10. Analyze a distributed file storage scenario where nodes experience random failures. Propose a data replication strategy.
11. Infer how big data spreadsheets help manage large datasets.
12. Explain how a distributed system enhances resource sharing across multiple machines.
13. Compare circuit switching and packet switching in network communication.
14. Clarify the necessity of cloud computing in the context of modern enterprise IT infrastructure.
15. Interpret the role of DNS in simplifying access to network resources.
16. Illustrate the difference between centralized and distributed computing using real-world examples.
17. Identify the key cloud service providers in India and their core service offerings.
18. Summarize the evolution of cloud computing from mainframes to distributed data centers.
19. Outline the differences in service offerings between AWS and Google Cloud.
20. Distinguish between LAN, MAN, and WAN in terms of geographical coverage and use-cases.

21. Compare server virtualization with storage virtualization.
22. Explain different types of virtualization technologies used in cloud environments.
23. Evaluate the importance of virtualization in cloud storage management.
24. What are different storage models used for big data.
25. Outline the key features of distributed programming.
26. Infer the role of parallel programming in cloud computing.
27. Show how NoSQL databases differ from relational databases in cloud applications.
28. Relate big data issues to cloud programming models.
29. Outline the functioning of GFS in distributed storage.
30. Show the connection between cloud scalability and parallel programming.
31. Classify programming models used in cloud environments.
32. Outline the role of HDFS in storing large data sets.
33. Relate the use of big data tools to efficient cloud application development.
34. Assess the resource efficiency of deploying 20 VMs on two hosts, each with 16 cores and 64 GB RAM, with each VM needing 2 cores and 4 GB RAM.
35. Critique the decision to deploy high-I/O database VMs on shared virtualized storage rather than dedicated disks.
36. Assess the performance implications if 10 application VMs share a single 1 Gbps virtual network interface card (vNIC) in a hypervisor setup.
37. Explain how cloud monitoring differ from traditional infrastructure monitoring?
38. Interpret system metrics from a cloud monitoring dashboard showing 85% CPU utilization and 3-second average response time.
39. Analyze a scenario where 10 microservices use a shared PaaS platform and identify potential load balancing challenges.

40. Outline the flow of deploying a web application over cloud using a service middleware component.
41. An organization spends ₹8,00,000 per year on on-premises servers. Moving to a cloud IaaS model costs ₹55,000 per month plus ₹1,20,000 annually for security. Calculate the net annual cost savings.
42. Differentiate between Elastic Computing and Utility Computing in the context of cloud service management with suitable examples.
43. Use elasticity to optimize resource costs for an e-commerce website with fluctuating traffic.
44. Illustrate the working mechanism of cloud monitoring .
45. Demonstrate the use of middleware in connecting distributed cloud applications using SOAP.
46. Apply the concept of service load balancing in deploying a cloud-native banking application across 3 regions.
47. Sketch a cloud deployment diagram showing a hybrid cloud model with workload distribution between on-premise and Azure cloud.
48. Calculate the minimum resource pool required to run 5 applications, where each application requires 1.5 CPU cores, 2 GB RAM, and 7 GB storage. Additionally, include a 10% overhead for resource allocation.
49. Implement cloud service monitoring using AWS CloudWatch for tracking CPU utilization of EC2 instances.
50. Describe how elasticity benefits applications deployed in a public cloud environment.
51. Explain the benefits of cloud monitoring.
52. Compare IaaS, PaaS, and SaaS in terms of developer control and abstraction level.
53. Identify the role of service load balancing in cloud workload distribution.
54. Clarify the function of the resource management system in a cloud platform.
55. Illustrate the difference between cloud bursting and auto-scaling.
56. Describe the advantages of using SOAP as a service technology in enterprise cloud services.
57. Explain how dynamic scalability differs from elasticity in cloud services.
58. Differentiate between network-level, host-level, and application-level security in cloud computing. Provide examples of security measures used at each level.
59. Analyze the role of access control mechanisms in maintaining data security in a cloud environment. How do they differ from authentication methods?

60. Evaluate the effectiveness of existing authentication mechanisms (e.g., multi-factor authentication, biometric access) in securing cloud services against unauthorized access.
61. Compare data privacy issues in mobile and media clouds.

Long Answer Type Questions

1. Describe the key characteristics of distributed systems and how they differ from centralized systems.
2. Illustrate with examples how cloud computing is integrated into real-world applications and services.
3. Investigate how different cloud deployment models (public, private, hybrid) influence organizational decision-making.
4. Analyze how virtualization supports the functionality of distributed and cloud systems.
5. Categorize the main types of network topologies and analyze their impact on performance and fault tolerance.
6. Infer how the evolution from traditional computing to cloud computing has impacted enterprise IT strategies.
7. Analyze how data is transmitted differently in circuit-switched and packet-switched networks.
8. Explain the key characteristics of public, private, and hybrid cloud deployment models with suitable examples.
9. Describe how organizations can use hybrid cloud to balance workload and improve efficiency.
10. Apply the concept of hybrid cloud deployment to design a cloud solution for a university managing public access and secure internal systems.
11. Demonstrate how a hybrid cloud deployment ensures secure data integration and interoperability between systems.
12. You have a VM cluster with a capacity of 20 vCPUs, 30 GB RAM, and 200 GB storage. Can it host 10 apps, each needing 1.8 vCPUs, 2 GB RAM, 12 GB storage, if you add 10% overhead to the pool? Calculate the math.
13. Recommend an appropriate cloud platform solution for a business that requires real-time monitoring and automatic scaling of virtual machines. Support your recommendation with valid reasoning.
14. Calculate the minimum resource pool required to run 7 SaaS applications, where each application needs 1.3 CPU cores, 3 GB RAM, and 4 GB storage. Additionally, include a 20% headroom for overhead.
15. Prioritize components of a cloud monitoring system for ensuring SLA compliance in a logistics company.

16. Assess how an SLA (Service Level Agreement) management system contributes to ensuring performance guarantees in a cloud service model.
17. Analyze the role of virtualization in optimizing cloud storage.
18. Compare server virtualization and application virtualization with examples.
19. Examine the architecture of storage clouds and their role in cloud computing.
20. Distinguish between network virtualization and storage virtualization.
21. Analyze how hypervisors contribute to VM creation and resource allocation.
22. Compare different types of virtualization used in cloud platforms.
23. Assess the performance of 5 virtual machines running on a 16-core CPU with 64 GB RAM, each VM allocated 4 cores and 12 GB RAM.
24. Examine the importance of machine images in deploying virtual machines.
25. Assess the performance of 6 VMs on a host with 24-core CPU and 96 GB RAM, where each VM requires 6 cores and 20 GB RAM.
26. Evaluate whether a 32-core, 128 GB RAM server can efficiently host 10 VMs if each VM needs 2 cores and 16 GB RAM for an analytics application.
27. Justify deploying 12 lightweight VMs (1 core, 2 GB RAM each) on a hypervisor host with 8 cores and 24 GB RAM.
28. Assess the impact of allocating 8 VMs with 8 GB RAM each on a host with 48 GB RAM.
29. Evaluate the efficiency of running 4 database VMs each requiring 6 vCPUs on a dual-socket 8-core (16-core total) machine.
30. Validate if a server with 20 TB of virtualized storage can support 100 VMs each needing 150 GB persistent disk.
31. Recommend VM configurations (CPU and memory) for a video rendering farm of 5 machines with 24-core CPUs and 128 GB RAM each.
32. Evaluate how resource virtualization enhances scalability in cloud infrastructure.
33. Critique security concerns associated with virtualization in public cloud.
34. Outline the fundamentals of parallel and distributed programming in cloud platforms.
35. Show how big data spreadsheets support data analysis in cloud environments.
36. Identify storage models used for handling large-scale cloud data.
37. Choose the most suitable storage architecture for unstructured data in cloud.
38. Analyze how cloud security differs across network level, host level, and application level.
39. Differentiate between access control and authentication in the context of cloud security.
40. Examine the structure of a mobile cloud architecture and identify potential security threats.
41. Interpret how application-level vulnerabilities can lead to data breaches in cloud platforms.

42. Compare traditional security models with cloud-specific access control mechanisms.
43. Categorize various types of threat agents in cloud computing environments.
44. Analyze how encryption techniques address cloud data privacy issues.
45. Evaluate the effectiveness of multi-factor authentication (MFA) in securing cloud applications.
46. Justify the use of network segmentation as a method for improving cloud security posture.
47. Recommend security best practices for managing sensitive data in media cloud platforms.
48. Defend the use of host-based intrusion detection systems (HIDS) in virtualized cloud environments.
49. Assess the risks and mitigation strategies for insider threats in cloud infrastructure.
50. Argue the importance of data privacy policies in managing cloud compliance and trust.
51. Prioritize cloud security controls for a healthcare organization migrating to cloud infrastructure.
52. Critique current cloud authentication mechanisms in terms of scalability and resilience.
53. Recommend suitable access control models for managing a multi-tenant SaaS platform.
54. Evaluate how cloud providers' shared responsibility model affects application security management.