Subject : Cloud Computing   Unit : II   Class: BE

1. Which of the following is storage data interchange interface for stored data objects?
a) OCC
b) OCCI
c) OCMI
d) All of the mentioned

Answer: b
Explanation: OCCI stands for Open Cloud Computing Interface.

2. Point out the correct statement.
a) To determine whether your application will port successfully, you should perform a functionality mapping exercise
b) Cloud computing supports some application features better than others
c) Cloud bursting is an overflow solution that clones the application to the cloud and directs traffic to the cloud during times of high traffic
d) All of the mentioned

Answer: d
Explanation: Mapping process involves determining the critical application features and then matching them to the cloud provider's offering to see if those features can be supported.

3. _____ data represents more than 50 percent of the data created every day.
a) Shadow
b) Light
c) Dark
d) All of the mentioned

Answer: a
Explanation: Lots of shadow data does get retained, having never been touched by a human being.

4. Which of the following system does not provision storage to most users?
a) PaaS
b) IaaS
c) CaaS
d) SaaS

Answer: b
Explanation: Cloud storage is storage accessed by a Web service API.

5. The addition of a software package on top of a cloud storage volume makes most cloud storage offerings conform to a _____ as a Service model.
a) Software
b) Platform
c) Analytics
d) None of the mentioned

Answer: a
Explanation: Most users interact with cloud storage using backup, synchronization, archiving, staging, caching, or some other sort of software.

6. How many categories of storage devices broadly exist in cloud?
a) 1
b) 2
c) 3
d) None of the mentioned

Answer: b
Explanation: Storage devices may be broadly categorized as either block storage devices or file storage devices.

7. Which of the following storage devices exposes its storage to clients as Raw storage that can be partitioned to create volumes?
a) block
b) file
c) disk
d) all of the mentioned

Answer: a
Explanation: It is up to the operating system to create and manage the file system.

8. Which of the following impose additional overhead on clients and offer faster transfer?
a) Block storage
b) File Storage
c) File Server
d) All of the mentioned

Answer: a
Explanation: File-oriented storage devices are generally slower, but require less overhead from attached clients.

9. Which of the following is most important feature of cloud storage listed below ?

a)   Logon authentication
b)   Bare file
c)   Multiplatform support
d)   Adequate bandwidth

Answer: Option A

Explanation:

The interactive logon process is the first step in user authentication and authorization.

10. Which of the following is not a backup category ?
a)   Full system backup

b)   Half system backup

c)   Image backup

d)   All of the mentioned

Answer: Option B

11. Which of the following system does not provision storage to most users ?
a)   PaaS
b)   IaaS
c)   CaaS
d)   SaaS

Answer: Option B

Explanation:

Cloud storage is storage accessed by a Web service API.

12. Which of the following service is provided by Google for online storage ?
a)   Drive
b)   SkyDrive
c)   Dropbox
d)   All of the mentioned

Answer: Option A

Explanation:

Google Drive is a file storage and synchronization service created by Google.

13 A _____ is a logical unit that serves as the target for storage operations, such as the SCSI protocol's READs and WRITEs.
   a) GETs
   b) PUN
   c) LUN
   d) All of the mentioned

Answer: Option C

Explanation:

In traditional pooled storage deployments, storage partitions can be assigned and provide a device label called a Logical Unit Number

14. Which of the following backup create a cloned copy of your current data or drive ?
   a) Continuous Data Protection
   b) Open file backup
   c) Reverse Delta backup
   d) None of the mentioned

Answer: Option A

Explanation:

CDP is also referred to as mirroring.

15 .The source of HDFS architecture in Hadoop originated as
   a) Google distributed filesystem
   b) Yahoo distributed filesystem
   c) Facebook distributed filesystem
   d) Facebook distributed filesystem

Ans: a

16. What is HDFS?
   a) Storage layer
   b) Batch processing engine
   c) Resource management layer
   d) All of the above

Ans : a

17. Which among the following command is used to copy a directory from one node to another in HDFS?
   a) rcp
   b) distcp
   c) dcp
   d) drcp

Ans B

18. Which utility is used for checking the health of an HDFS file system?
   a) fsck
   b) fchk
   c) fsch

d) fcks

Ans:a

20. Which among the following is the correct statement
    a) Datanode manage file system namespace
    b) Namenode stores metadata
    c) NameNode stores actual data
    d) All of the above
        Ans B

21. What is default replication factor?
    a) 1
    b) 2
    c) 3
    d) 5

        Ans c

22. The namenode knows that the data node is active using a mechanism known as
    a) Active pulse
    b) Data pulse
    c) Heartbeats
    d) h-signal

Ans C

23. What is the default size of HDFS Data Block?
    a) 16MB
    b) 32MB
    c) 64MB
    d) 128MB

Ans d

24. What is HDFS Block in Hadoop?
    a) It is the logical representation of data
    b) It is the physical representation of data
    c) Both the above
    d) None of the above

Ans b

25. Which of the following is the correct statement?
    a) DataNode is the slave/worker node and holds the user data in the form of Data Blocks
    b) Each incoming file is broken into 32 MB by default
    c) NameNode stores user data in the form of Data Blocks
    d) None of the above

Ans A

26. The need for data replication can arise in various scenarios like
    a) Replication Factor is changed

b) DataNode goes down

c) Data Blocks get corrupted

d) All of the above

Ans d

27. A file in HDFS that is smaller than a single block size

   a) Cannot be stored in HDFS

   b) Occupies the full block's size.

   c) Occupies only the size it needs and not the full block

   d) Can span over multiple blocks

Ans c

28. Which among the following are the duties of the NameNodes

   a) Manage file system namespace

   b) It is responsible for storing actual data

   c) Perform read-write operation as per request for the clients

   d) None of the above

Ans a

29. If the IP address or hostname of a data node changes

   a) The namenode updates the mapping between file name and block name

   b) The data in that data node is lost forever

   c) The namenode need not update mapping between file name and block name

   d) There namenode has to be restarted

Ans c

30, For the frequently accessed HDFS files the blocks are cached in

   a) The memory of the data node

   b) In the memory of the namenode

   c) Both the above

Ans a

31. Which scenario demands highest bandwidth for data transfer between nodes

   a) Different nodes on the same rack

   b) Nodes on different racks in the same data center.

   c) Nodes in different data centers

   d) Data on the same node

Ans c

32. When a client contacts the namenode for accessing a file, the namenode responds with

   a) Size of the file requested

   b) Block ID of the file requested

   c) Block ID and hostname of all the data nodes containing that block

Ans c

33. In HDFS the files cannot be

   a) Read

   b) Deleted

   c) Executed

   d) Archived

Ans c

34. Which among the following is the duties of the Data Nodes
    a) Manage file system namespace
    b) Stores meta-data
    c) Regulates client's access to files
    d) Perform read-write operation as per request for the clients

Ans d

35. NameNode and DataNode do communicate using
    a) Active pulse
    b) Heartbeats
    c) h-signal

ans b

36. HDFS is inspired by which of following Google project
    a) BigTable
    b) GFS
    c) MapReduce
    d)

Ans b

37. In HDFS, data node sends frequent heartbeats to name node
    a) True
    b) False

Ans a

38. Clients connect to _____ for I/O
    a) NameNode
    b) DataNode

 Ans b

39. For reading/writing data to/from HDFS, clients first connect to
    a) NameNode
    b) Checkpoint Node
    c) DataNode

Ans a

40. The namenode loses its only copy of fsimage file. We can recover this from which of the following?
    a) Secondary Namenode
    b) It can not be recovered
    c) Datanodes
    d) Checkpoint Node

Ans b

41. The HDFS command to create the copy of a file from a local system is which of the following?

a) copyFromLocal

b) CopyFromLocal

c) CopyLocal

d) Copyfromlocal

42. HDFS provides a command line interface called _____ used to interact with HDFS

a) HDFS Shell

b) FS Shell

c) DFS Shell

Ans b

43. Which of the following is the daemon of HDFS?

a) Secondary namenode

b) Node manager

c) Resource manager

Ans a

44. Which of the following stores metadata?

a) DataNode

b) NameNode

c) Secondary Data Node

Ans b

45. Which of the following statement is true about Secondary NameNode

a) It store the modified FsImage into persistent storage

b) It stores the merged FsImage with EditLogs back to active namenode.

c) It does not store the modified FsImage into persistent storage

Ans a

46. Which statement is true about NameNode

a) It is the slave node that stores actual data

b) It is the Master node that stores actual data

c) It is the slave node that stores metadata

d) It is the Mater node that stores metadata

Ans d

47. Which of the following is true about metadata

a) Metadata shows the structure of HDFS directories/files

b) Metadata contain information like number of blocks, their location, replicas

c) FsImage & EditLogs are metadata files

d) All of the above

Ans d

48. What is the major advantages of storing data in block size 128MB

a) It saves disk seek time

b) It saves disk processing time

c) It saves disk access time

d) It saves disk latency time

Ans a

49. HDFS performs replication, although it results in data redundancy?

a) True

b) False

Ans a

50.Which of the following is the Single Point of Failure

a) NameNode

b) Secondary NameNode

c) DataNode

Ans a

51.Which of the following are the features of HDFS

a) Fault Tolerance

b) High Availability

c) Replication

d) All of the above

Ans d

52. HBase is a distributed _____ database built on top of the Hadoop file system.

a) Column-oriented

b) Row-oriented

c) Tuple-oriented

d) None of the mentioned

Answer: a

Explanation: HBase is a data model that is similar to Google's big table designed to provide quick random access to huge amounts of structured data.

53. Point out the correct statement.

a) HDFS provides low latency access to single rows from billions of records (Random access)

b) HBase sits on top of the Hadoop File System and provides read and write access

c) HBase is a distributed file system suitable for storing large files

d) None of the mentioned

Answer: b

Explanation: One can store the data in HDFS either directly or through HBase. Data consumer reads/accesses the data in HDFS randomly using HBase.

54. HBase is _____ defines only column families.

a) Row Oriented

b) Schema-less

c) Fixed Schema

d) All of the mentioned

Answer: b

Explanation: HBase doesn't have the concept of fixed columns schema.

55. Apache HBase is a non-relational database modeled after Google's _____

a) BigTop

b) Bigtable

c) Scanner

d) FoundationDB

Answer: b

Explanation: Bigtable acts up on Google File System, likewise Apache HBase works on top of Hadoop and HDFS.

56. Point out the wrong statement.

a) HBase provides only sequential access to data

b) HBase provides high latency batch processing

c) HBase internally provides serialized access

d) All of the mentioned

Answer: c

Explanation: HBase internally uses Hash tables and provides random access.

57. The _____ Server assigns regions to the region servers and takes the help of Apache ZooKeeper for this task.

a) Region

b) Master

c) Zookeeper

d) All of the mentioned

Answer: b

Explanation: Master Server maintains the state of the cluster by negotiating the load balancing.

58. Which of the following command provides information about the user?

a) status

b) version

c) whoami

d) user

Answer: c

Explanation: status command provides the status of HBase, for example, the number of servers.

59. Which of the following command does not operate on tables?

a) enabled

b) disabled

c) drop

d) all of the mentioned

Answer: b

Explanation: is_disabled command verifies whether a table is disabled.

60. _____ command fetches the contents of a row or a cell.
a) select
b) get
c) put
d) none of the mentioned

Answer: b

Explanation: put command puts a cell value at a specified column in a specified row in a particular table.

61. Point out the correct statement.
a) MapReduce tries to place the data and the compute as close as possible
b) Map Task in MapReduce is performed using the Mapper() function
c) Reduce Task in MapReduce is performed using the Map() function
d) All of the mentioned

Answer: a

Explanation: This feature of MapReduce is "Data Locality".

62. _____ part of the MapReduce is responsible for processing one or more chunks of data and producing the output results.
a) Maptask
b) Mapper
c) Task execution
d) All of the mentioned

Answer: a

Explanation: Map Task in MapReduce is performed using the Map() function.

63. _____ function is responsible for consolidating the results produced by each of the Map() functions/tasks.
a) Reduce
b) Map
c) Reducer
d) All of the mentioned

Answer: a

Explanation: Reduce function collates the work and resolves the results.

64. Point out the wrong statement.
a) A MapReduce job usually splits the input data-set into independent chunks which are processed by the map tasks in a completely parallel manner
b) The MapReduce framework operates exclusively on <key, value> pairs
c) Applications typically implement the Mapper and Reducer interfaces to provide the map and reduce methods
d) None of the mentioned

Answer: d

Explanation: The MapReduce framework takes care of scheduling tasks, monitoring them and re-executes the failed tasks.

65. Although the Hadoop framework is implemented in Java, MapReduce applications need not be written in _____

a) Java

b) C

c) C#

d) None of the mentioned

Answer: a

Explanation: Hadoop Pipes is a SWIG- compatible C++ API to implement MapReduce applications (non JNITM based).

66. _____ maps input key/value pairs to a set of intermediate key/value pairs.

a) Mapper

b) Reducer

c) Both Mapper and Reducer

d) None of the mentioned

Answer: a

Explanation: Maps are the individual tasks that transform input records into intermediate records.

67. HBase is a distributed _____ database built on top of the Hadoop file system.

a) Column-oriented

b) Row-oriented

c) Tuple-oriented

d) None of the mentioned

Answer: a

Explanation: HBase is a data model that is similar to Google's big table designed to provide quick random access to huge amounts of structured data.

68. Point out the correct statement.

a) HDFS provides low latency access to single rows from billions of records (Random access)

b) HBase sits on top of the Hadoop File System and provides read and write access

c) HBase is a distributed file system suitable for storing large files

d) None of the mentioned

Answer: b

Explanation: One can store the data in HDFS either directly or through HBase. Data consumer reads/accesses the data in HDFS randomly using HBase.

69. HBase is _____ defines only column families.

a) Row Oriented

b) Schema-less

c) Fixed Schema

d) All of the mentioned

Answer: b

Explanation: HBase doesn't have the concept of fixed columns schema.

70. Point out the wrong statement.
a) HBase provides only sequential access to data
b) HBase provides high latency batch processing
c) HBase internally provides serialized access
d) All of the mentioned

Answer: c

Explanation: HBase internally uses Hash tables and provides random access.

71. HDFS is implemented in _____ programming language.
a) C++
b) Java
c) Scala
d) None of the mentioned

Answer: b

Explanation: HDFS is implemented in Java and any computer which can run Java can host a NameNode/DataNode on it.

72. Which of the following service provider provides the least amount of built in security?
a) SaaS
b) PaaS
c) IaaS
d) All of the mentioned

Answer: c

Explanation: You get the least amount of built in security with an Infrastructure as a Service provider, and the most with a Software as a Service provider.

73. Which of the following services that need to be negotiated in Service Level Agreements?
a) Logging
b) Auditing
c) Regulatory compliance
d) All of the mentioned

Answer: d

Explanation: Logging, auditing, and regulatory compliance are all features that require planning in cloud computing systems.

74. Which of the following allows you to create instances of the MySQL database to support your Web sites?
a) Amazon Elastic Compute Cloud
b) Amazon Simple Queue Service
c) Amazon Relational Database Service
d) Amazon Simple Storage System

Answer: c

Explanation: RDS provides features such as automated software patching, database backups, and automated database scaling via an API call.

75. Point out the correct statement.

a) Amazon Elastic Cloud is a system for creating virtual disks(volume)

b) SimpleDB interoperates with both Amazon EC2 and Amazon S3

c) EC3 is an Analytics as a Service provider

d) None of the mentioned

Answer: b

Explanation: Amazon SimpleDB stores data in "buckets" and without requiring the creation of a database schema

76. Which of the following is a structured data store that supports indexing and data queries to both EC2 and S3?

a) CloudWatch

b) Amazon SimpleDB

c) Amazon Cloudfront

d) All of the mentioned

Answer: b

Explanation: SimpleDB isn't a full database implementation.

77. Which of the following is the most important area of concern in cloud computing?

a) Security

b) Storage

c) Scalability

d) All of the mentioned

Answer: a

Explanation: When your data travels over and rests on systems that are no longer under your control, you have increased risk due to the interception and malfeasance of others.

78. You can't count on a cloud provider maintaining your _____ in the face of government actions.

a) scalability

b) reliability

c) privacy

d) none of the mentioned

Answer: c

Explanation: Cloud computing industry continues to address security concerns.

79. Data stored in _____ domains doesn't require maintenance of a schema.

- A. SimpleDB
- B. SQL Server

- C. Oracle
- D. RDS

   Answer: Option A

   Explanation:

   To create a high performance "simple" database, the data store created is flat; that is, it is non-relational and joins are not supported.

80. Benefits of using Dynamo DB
   a) It is a managed service where there is no need to hire the experts or worry about installation, setup, cluster etc.,
   b) It is scalable.
   c) It provides the users high throughput at very low latency.
   d) All   the above

Ans d

   81. What is a row called in DynamoDB
   a)   row
   b)   tuple
   c)   item
   d)   none of these

ans b

82. Do items needs to share the same number of attributes?
   a)   True
   b)   False
      ans b

83.There are no columns in DynamoDB
      a) True
      b) False

Ans a

84. the maximum size of an item of Dynamo DB
      a) 400K
      b) 100k
      c) 128 k
      d)64k

ans a

85. Dynamo DB API is used to perform DDL operations?
      a) true
      b) False

ans a

86. dynamo DB support SQL

a) True

b) False

ans b

87. How would you scale horizontally on a RDMS?

    a) scaling horizontal

    b) shrading

    c) Vertical scale

    d) none of these

ans b

88. DynamoDB is a low latency NoSQL database service (Non-Relational Database)

    a) true

    b) False

ans a

89. Data stored in _____ domains doesn't require maintenance of a schema.

    a) SimpleDB

    b) SQL Server

    c) Oracle

    d) RDS

 Ans a

90. Which of the following is relational database service provided by Amazon?

    a) SimpleDB

    b) SQL Server

    c) Oracle

    d) RDS

 Ans d

91. Amazon Relational Database Service is a variant of the _____ 5.1 database system.

    a) Oracle

    b) MySQL

    c) SQL Server

Ans b

92. Which of the following database should be used for a solution that has a very high availability?

     a)   SimpleDB

     b)   RDS

     c)   Amazon EC2

Ans a

93. Database marketplace based on SQL Azure Database is code-named _____

     a)   Akamai

     b)   Dallas

     c)   Denali

Ans b

94. **names of some large cloud providers and databases?**

     a)   Google bigtable

     b)   Amazon simpleDB

     c)   Cloud based SQL

     d)   all of these

Ans d

95. open source cloud computing platform databases?

     a)   MongoDB,

     b)   CouchDB,

     c)   LucidDB

     d)   All of these

Ans d

96. example of large cloud provider and databases?

     a)   Google bigtable

     b)   Amazon simpleDB

     c)   Cloud based SQL

     d)   All the above

Ans d

97. DynamoDB supports four scalar data types

     a)   Number Set

     b)   String Set

     c)   Binary Set

     d)   String

Ans d

98. Which of the following is a standard protocol for network monitoring and discovery?

a) SNMP

b) CMDB

c) WMI

d) All of the mentioned

Answer: d
Explanation: WMI stands for Windows Management Instrumentation (WMI).

99. Which of the following factors are deemed to be important considerations for evaluating cloud solutions?
a) Seamless maintenance and upgrades
b) Performance characteristics
c) Client self-service
d) All of the mentioned

Answer: d
Explanation: In addition to assigning security-provisioned cloud storage to a client or service requester, a cloud storage service provider must be able to deliver a measured level of service.

100. How many categories of backup exist in cloud storage?
a) 3
b) 4
c) 5
d) None of the mentioned

Answer: d
Explanation: Seven categories of backup exist for cloud storage.

101. Which of the following is not a backup category?
a) Full system backup
b) Half system backup
c) Image backup
d) All of the mentioned

Answer: b
Explanation: An image backup creates a complete copy of a volume.

102 Which of the following backup lets you restore your data to a point in time and saves multiple copies of any file that has been changed?
a) Point-in-time
b) Differential
c) Reverse backup
d) None of the mentioned

Answer: a
Explanation: At least 10 to 30 copies of previous versions of files should be saved.

103. The amount of time needed to backup a system is referred to as its _____
a) backup time
b) backup sheet

c) backup window
d) all of the mentioned

Answer: c
Explanation: The first backup is quite slow over an Internet connection, but the incremental backup can be relatively fast.