

ITPassLeader

Pass Your Next Certification Exam Fast!

Select a vendor... Select an test... Your email address [Free Download Demo](#)



Instant Download



365 Days Free Updates



Money Back Guarantee



Security & Privacy

Choose the version that fits your needs

PDF Version

Desktop Test Engine

Online Test Engine

Latest and Up-to-Date exam dumps with real exam questions answers.



Get 12-Months free updates without any extra charges.



Experience same exam environment before appearing in the certification exam.



100% exam passing guarantee in the first attempt.



20% discount on more than one license and 30% discount on 5+ license purchases.



100% secure purchase on SSL.



Completely private purchase without sharing your personal info with anyone.



<http://www.itpassleader.com>

High-praise Exam Dumps Questions grant you success by high pass rate - ITPassLeader

Exam : **AWS-Big-Data-Specialty**

Title : AWS Certified Big Data -
Specialty

Vendor : Amazon

Version : DEMO

- NO.1** An organization currently runs a large Hadoop environment in their data center and is in the process of creating an alternative Hadoop environment on AWS, using Amazon EMR. They generate around 20 TB of data on a monthly basis. Also on a monthly basis, files need to be grouped and copied to Amazon S3 to be used for the Amazon EMR environment. They have multiple S3 buckets across AWS accounts to which data needs to be copied. There is a 10G AWS Direct Connect setup between their data center and AWS, and the network team has agreed to allocate
- A.** Use an offline copy method, such as an AWS Snowball device, to copy and transfer data to Amazon S3.
 - B.** Configure a multipart upload for Amazon S3 on AWS Java SDK to transfer data over AWS Direct Connect.
 - C.** Use Amazon S3 transfer acceleration capability to transfer data to Amazon S3 over AWS Direct Connect.
 - D.** Setup S3DistCop tool on the on-premises Hadoop environment to transfer data to Amazon S3 over AWS Direct Connect.

Answer: D

NO.2 You are managing the AWS account of a big organization. The organization has more than 1000+ employees and they want to provide access to the various services to most of the employees. Which of the below mentioned options is the best possible solution in this case?

- A.** The user should create IAM groups as per the organization's departments and add each user to the group for better access control
- B.** Attach an IAM role with the organization's authentication service to authorize each user for various AWS services
- C.** The user should create an IAM role and attach STS with the role. The user should attach that role to the EC2 instance and setup AWS authentication on that server
- D.** The user should create a separate IAM user for each employee and provide access to them as per the policy

Answer: B

NO.3 What does Amazon CloudFormation provide?

- A.** None of these.
- B.** The ability to setup Autoscaling for Amazon EC2 instances.
- C.** A template to map network resources for Amazon Web Services.
- D.** A templated resource creation for Amazon Web Services.

Answer: D

NO.4 An Operations team continuously monitors the number of visitors to a website to identify any potential system problems. The number of website visitors varies throughout the day. The site is more popular in the middle of the day and less popular at night.

Which type of dashboard display would be the MOST useful to allow staff to quickly and correctly identify system problems?

- A.** A single KPI metric showing the statistical variance between the current number of website visitors and the historical number of website visitors for the current time of day.
- B.** A vertical stacked bar chart showing today's website visitors and the historical average number of

website visitors.

C. A scatter plot showing today's website visitors on the X-axis and the historical average number of website visitors on the Y-axis.

D. An overlay line chart showing today's website visitors at one-minute intervals and also the historical average number of website visitors.

Answer: A

NO.5 A city has been collecting data on its public bicycle share program for the past three years. The SPB dataset currently on Amazon S3. The data contains the following data points:

- * Bicycle organization points
- * Bicycle destination points
- * Mileage between the points
- * Number of bicycle slots available at the station (which is variable based on the station location)
- * Number of slots available and taken at each station at a given time

The program has received additional funds to increase the number of bicycle stations, available. All data is regularly archived to Amazon Glacier.

The new bicycle station must be located to provide the most riders access to bicycles. How should this task be performed?

A. Move the data from Amazon S3 into Amazon EBS-backed volumes and EC2 Hadoop with spot instances to run a Spark job that performs a stochastic gradient descent optimization.

B. Persist the data on Amazon S3 and use a transits EMR cluster with spot instances to run a Spark streaming job that will move the data into Amazon Kinesis.

C. Keep the data on Amazon S3 and use an Amazon EMR based Hadoop cluster with spot instances to run a spark job that perform a stochastic gradient descent optimization over EMBFS.

D. Use the Amazon Redshift COPY command to move the data from Amazon S3 into RedShift and platform a SQL query that outputs the most popular bicycle stations.

Answer: C

NO.6 An organization is setting up a data catalog and metadata management environment for their numerous data stores currently running on AWS. The data catalog will be used to determine the structure and other attributes of data in the data stores. The data stores are composed of Amazon RDS databases, Amazon Redshift, and CSV files residing on Amazon S3. The catalog should be populated on a scheduled basis, and minimal administration is required to manage the catalog. How can this be accomplished?

A. Use AWS Glue Data Catalog as the data catalog and schedule crawlers that connect to data sources to populate the database.

B. Set up Amazon DynamoDB as the data catalog and run a scheduled AWS Lambda function that connects to data sources to populate the database.

C. Use an Amazon database as the data catalog and run a scheduled AWS Lambda function that connects to data sources to populate the database.

D. Set up Apache Hive metastore on an Amazon EC2 instance and run a scheduled bash script that connects to data sources to populate the metastore.

Answer: A

Explanation

<https://docs.aws.amazon.com/glue/latest/dg/populate-data-catalog.html>

NO.7 Are you able to integrate a multi-factor token service with the AWS Platform?

- A.** Yes, you can integrate private multi-factor token devices to authenticate users to the AWS platform.
- B.** Yes, using the AWS multi-factor token devices to authenticate users on the AWS platform.
- C.** No, you cannot integrate multi-factor token devices with the AWS platform.

Answer: B

NO.8 What does Amazon ELB stand for?

- A.** Elastic Linux Box.
- B.** Elastic Load Balancing.
- C.** Encrypted Load Balancing.
- D.** Encrypted Linux Box.

Answer: B

NO.9 A sys admin is planning to subscribe to the RDS event notifications. For which of the below mentioned source categories the subscription cannot be configured?

- A.** DB security group
- B.** DB parameter group
- C.** DB snapshot
- D.** DB options group

Answer: D

NO.10 A telecommunications company needs to predict customer churn (i.e. customers who decide to switch a computer). The company has historic records of each customer, including monthly consumption patterns, calls to customer service, and whether the customer ultimately quit the service. All of this data is stored in Amazon S3. The company needs to know which customers are likely going to churn soon so that they can win back their loyalty.

What is the optimal approach to meet these requirements?

- A.** Use the Amazon Machine Learning service to build the binary classification model based on the dataset stored in Amazon S3. The model will be used regularly to predict churn attribute for existing customers
- B.** Use EMR to run the Hive queries to build a profile of a churning customer. Apply the profile to existing customers to determine the likelihood of churn
- C.** Use AWS QuickSight to connect it to data stored in Amazon S3 to obtain the necessary business insight. Plot the churn trend graph to extrapolate churn likelihood for existing customer
- D.** Use a Redshift cluster to COPY the data from Amazon S3. Create a user Define Function in Redshift that computes the likelihood of churn

Answer: A

Explanation

<https://aws.amazon.com/blogs/machine-learning/predicting-customer-churn-with-amazon-machine-learning/>

NO.11 To help you manage your Amazon EC2 instances, images, and other Amazon EC2 resources, you can assign your own metadata to each resource in the form of _____

- A. functions
- B. special filters
- C. wildcards
- D. tags

Answer: D

NO.12 True or False: When you add a rule to a DB security group, you do not need to specify port number or protocol.

- A. TRUE
- B. FALSE
- C. Depends on the RDMS used

Answer: A

NO.13 An organization's data warehouse contains sales data for reporting purposes. data governance policies prohibit staff from accessing the customers' credit card numbers. How can these policies be adhered to and still allow a Data Scientist to group transactions that use the same credit card number?

- A. Store a cryptographic hash of the credit card number.
- B. Mask the credit card numbers to only show the last four digits of the credit card number.
- C. Encrypt the credit card number with a symmetric encryption key, and give the key only to the authorized Data Scientist.
- D. Encrypt the credit card number with an asymmetric encryption key and give the decryption key only to the authorized Data Scientist.

Answer: A

NO.14 Does Amazon RDS allow direct host access via Telnet, Secure Shell (SSH), or Windows Remote Desktop Connection?

- A. No
- B. Yes
- C. Depends on if it is in VPC or not

Answer: A

NO.15 As part of your continuous deployment process, your application undergoes an I/O load performance test before it is deployed to production using new AMIs. The application uses one Amazon Elastic Block Store (EBS) PIOPS volume per instance and requires consistent I/O performance.

Which of the following must be carried out to ensure that I/O load performance tests yield the correct results in a repeatable manner?

- A. Ensure that the Amazon EBS volume has been pre-warmed by creating a snapshot of the volume before the test
- B. Ensure that snapshots of the Amazon EBS volumes are created as a backup
- C. Ensure that the I/O block sizes for the test are randomly selected

- D. Ensure that the Amazon EBS volume is encrypted
- E. Ensure that the Amazon EBS volumes have been pre-warmed by reading all the blocks before the test

Answer: E

NO.16 Which of the following notification endpoints or clients are supported by Amazon Simple Notification Service? Choose 2 answers

- A. File Transfer Protocol
- B. Simple Network Management Protocol
- C. Email
- D. CloudFront distribution
- E. Short Message Service

Answer: A,D

NO.17 Can I initiate a "forced failover" for my Oracle Multi-AZ DB Instance deployment?

- A. Only in VPC
- B. Yes
- C. No
- D. Only in certain regions

Answer: B

NO.18 You have an ASP.NET web application running in Amazon Elastic BeanStalk. Your next version of the application requires a third-party Windows installer package to be installed on the instance on first boot and before the application launches.

Which options are possible? Choose 2 answer

- A. In the source bundle's .ebextensions folder, create a file with a .config extension. In the file, under the "packages" section and "msi" package manager, include the package's URL
- B. In the environment's configuration, edit the instances configuration and add the package's URL to the "Packages" section
- C. Launch a new Amazon EC2 instance from the AMI used by the environment. Log into the instance, install the package and run sysprep. Create a new AMI. Configure the environment to use the new AMI
- D. In the application's Global.asax file, run msixexec.exe to install the package using Process.Start() in the Application_Start event handler
- E. In the source bundle's .ebextensions folder, create a "Packages" folder. Place the package in the folder

Answer: A,C

NO.19 What is the maximum response time for a Business level Premium Support case?

- A. 30 minutes
- B. 12 hours

C. 10 minutes

D. 1 hour

Answer: D

NO.20 Managers in a company need access to the human resources database that runs on Amazon Redshift, to run reports about their employees. Managers must only see information about their direct reports.

Which technique should be used to address this requirement with Amazon Redshift?

A. Define a view that uses the employee's manager name to filter the records based on current user names.

B. Use Amazon Redshift snapshot to create one cluster per manager. Allow the managers to access only their designated clusters.

C. Define a key for each manager in AWS KMS and encrypt the data for their employees with their private keys.

D. Define an IAM group for each employee as an IAM user in that group and use that to limit the access.

Answer: A

Explanation

<https://www.intermix.io/blog/iam-to-generate-temporary-amazon-redshift-passwords/>

<https://docs.aws.amazon.com/redshift/latest/mgmt/generating-iam-credentials-steps.html>