

# AWS BDS-C00

AWS BIG DATA SPECIALTY CERTIFICATION QUESTIONS & ANSWERS

---

Exam Summary – Syllabus – Questions

---

**BDS-C00**

[AWS Certified Big Data - Specialty](#)

65 Questions Exam – 750 / 1000 Cut Score – Duration of 170 minutes

[www.VMExam.com](http://www.VMExam.com)

## Table of Contents

Know Your BDS-C00 Certification Well: .....	2
AWS BDS-C00 Big Data Specialty Certification Details:	2
BDS-C00 Syllabus:.....	3
AWS BDS-C00 Sample Questions:.....	4
Study Guide to Crack AWS Big Data Specialty BDS-C00 Exam: .....	8

## Know Your BDS-C00 Certification Well:

The BDS-C00 is best suitable for candidates who want to gain knowledge in the AWS Specialty. Before you start your BDS-C00 preparation you may struggle to get all the crucial Big Data Specialty materials like BDS-C00 syllabus, sample questions, study guide.

But don't worry the BDS-C00 PDF is here to help you prepare in a stress free manner.

The PDF is a combination of all your queries like-

- What is in the BDS-C00 syllabus?
- How many questions are there in the BDS-C00 exam?
- Which Practice test would help me to pass the BDS-C00 exam at the first attempt?

Passing the BDS-C00 exam makes you AWS Certified Big Data - Specialty. Having the Big Data Specialty certification opens multiple opportunities for you. You can grab a new job, get a higher salary or simply get recognition within your current organization.

## AWS BDS-C00 Big Data Specialty Certification Details:

<b>Exam Name</b>	AWS Certified Big Data - Specialty (Big Data Specialty)
<b>Exam Code</b>	BDS-C00
<b>Exam Price</b>	\$300 USD
<b>Duration</b>	170 minutes
<b>Number of Questions</b>	65
<b>Passing Score</b>	750 / 1000
<b>Schedule Exam</b>	<a href="#">PEARSON VUE</a>
<b>Sample Questions</b>	<a href="#">AWS BDS-C00 Sample Questions</a>
<b>Recommended Practice</b>	<a href="#">AWS Certified Big Data - Specialty Practice Test</a>

## BDS-C00 Syllabus:

Section	Objectives	Weight
<b>Collection</b>	<ul style="list-style-type: none"> <li>- Determine the operational characteristics of the collection system</li> <li>- Select a collection system that handles the frequency of data change and type of data being ingested</li> <li>- Identify the properties that need to be enforced by the collection system: order, data structure, metadata, etc.</li> <li>- Explain the durability and availability characteristics for the collection approach</li> </ul>	<b>17%</b>
<b>Storage</b>	<ul style="list-style-type: none"> <li>- Determine and optimize the operational characteristics of the storage solution</li> <li>- Determine data access and retrieval patterns</li> <li>- Evaluate mechanisms for capture, update, and retrieval of catalog entries</li> <li>- Determine appropriate data structure and storage format</li> </ul>	<b>17%</b>
<b>Processing</b>	<ul style="list-style-type: none"> <li>- Identify the appropriate data processing technology for a given scenario</li> <li>- Determine how to design and architect the data processing solution</li> <li>- Determine the operational characteristics of the solution implemented</li> </ul>	<b>17%</b>
<b>Analysis</b>	<ul style="list-style-type: none"> <li>- Determine the tools and techniques required for analysis</li> <li>- Determine how to design and architect the analytical solution</li> <li>- Determine and optimize the operational characteristics of the Analysis</li> </ul>	<b>17%</b>
<b>Visualization</b>	<ul style="list-style-type: none"> <li>- Determine the appropriate techniques for delivering the results/output</li> <li>- Determine how to design and create the Visualization platform</li> <li>- Determine and optimize the operational characteristics of the Visualization system</li> </ul>	<b>12%</b>
<b>Data Security</b>	<ul style="list-style-type: none"> <li>- Determine encryption requirements and/or implementation technologies</li> <li>- Choose the appropriate technology to enforce data governance</li> <li>- Identify how to ensure data integrity</li> <li>- Evaluate regulatory requirements</li> </ul>	<b>20%</b>

## AWS BDS-C00 Sample Questions:

### Question: 1

An administrator decides to use the Amazon Machine Learning service to classify social media posts that mention your company into two categories:

posts that require a response and posts that do not. The training dataset of 10,000 posts contains the details of each post, including the timestamp, author, and full text of the post. You are missing the target labels that are required for training.

Which two options will create valid target label data?

- a) Ask the social media handling team to review each post and provide the label.
- b) Use the sentiment analysis NLP library to determine whether a post requires a response.
- c) Use the Amazon Mechanical Turk web service to publish Human Intelligence Tasks that ask Turk workers to label the posts.
- d) Using the a priori probability distribution of the two classes, use Monte-Carlo simulation to generate the labels.

**Answer: a, c**

### Question: 2

A company logs data from its application in large files and runs regular analytics of these logs to support internal reporting for three months after the logs are generated.

After three months, the logs are infrequently accessed for up to a year. The company also has a regulatory control requirement to store application logs for seven years.

Which course of action should the company take to achieve these requirements in the most cost-efficient way?

- a) Store the files in S3 Glacier with a Deny Delete vault lock policy for archives less than seven years old and a vault access policy that restricts read access to the analytics IAM group and write access to the log writer service role.
- b) Store the files in S3 Standard with a lifecycle policy to transition the storage class to Standard - IA after three months. After a year, transition the files to Glacier and add a Deny Delete vault lock policy for archives less than seven years old.
- c) Store the files in S3 Standard with lifecycle policies to transition the storage class to Standard – IA after three months and delete them after a year. Simultaneously store the files in Amazon Glacier with a Deny Delete vault lock policy for archives less than seven years old.
- d) Store the files in S3 Standard with a lifecycle policy to remove them after a year. Simultaneously store the files in Amazon S3 Glacier with a Deny Delete vault lock policy for archives less than seven years old.

**Answer: c**

**Question: 3**

A data engineer needs to architect a data warehouse for an online retail company to store historic purchases. The data engineer needs to use Amazon Redshift.

To comply with PCI:DSS and meet corporate data protection standards, the data engineer must ensure that data is encrypted at rest and that the keys are managed by a corporate on-premises HSM.

Which approach meets these requirements in the most cost-effective manner?

- a) Create a VPC, and then establish a VPN connection between the VPC and the on-premises network. Launch the Amazon Redshift cluster in the VPC, and configure it to use your corporate HSM.
- b) Use the AWS CloudHSM service to establish a trust relationship between the CloudHSM and the corporate HSM over a Direct Connect connection. Configure Amazon Redshift to use the CloudHSM device.
- c) Configure the AWS Key Management Service to point to the corporate HSM device, and then launch the Amazon Redshift cluster with the KMS managing the encryption keys.
- d) Use AWS Import/Export to import the corporate HSM device into the AWS Region where the Amazon Redshift cluster will launch, and configure Redshift to use the imported HSM.

**Answer: a**

**Question: 4**

A company needs to deploy a data lake solution for their data scientists in which all company data is accessible and stored in a central S3 bucket.

The company segregates the data by business unit, using specific prefixes. Scientists can only access the data from their own business unit.

The company needs a single sign-on identity and management solution based on Microsoft Active Directory (AD) to manage access to the data in Amazon S3.

Which method meets these requirements?

- a) Use AWS IAM Federation functions and specify the associated role based on the users' groups in AD.
- b) Create bucket policies that only allow access to the authorized prefixes based on the users' group name in Active Directory.
- c) Deploy the AD Synchronization service to create AWS IAM users and groups based on AD information.
- d) Use Amazon S3 API integration with AD to impersonate the users on access in a transparent manner.

**Answer: a**

**Question: 5**

A customer needs to load a 550-GB data file into an Amazon Redshift cluster from Amazon S3, using the COPY command.

The input file has both known and unknown issues that will probably cause the load process to fail. The customer needs the most efficient way to detect load errors without performing any cleanup if the load process fails.

Which technique should the customer use?

- a) Compress the input file before running COPY.
- b) Write a script to delete the data from the tables in case of errors.
- c) Split the input file into 50-GB blocks and load them separately.
- d) Use COPY with NOLOAD parameter.

**Answer: d**

**Question: 6**

An administrator has a 500-GB file in Amazon S3. The administrator runs a nightly COPY command into a 10-node Amazon Redshift cluster. The administrator wants to prepare the data to optimize performance of the COPY command.

How should the administrator prepare the data?

- a) Compress the file using gz compression.
- b) Split the file into 500 smaller files.
- c) Convert the file format to AVRO.
- d) Split the file into 10 files of equal size.

**Answer: b**

**Question: 7**

A data engineer needs to collect data from multiple Amazon Redshift clusters within a business and consolidate the data into a single central data warehouse. Data must be encrypted at all times while at rest or in flight.

What is the most scalable way to build this data collection process?

- a) Run an ETL process that connects to the source clusters using SSL to issue a SELECT query for new data, and then write to the target data warehouse using an INSERT command over another SSL secured connection.
- b) Use AWS KMS data key to run an UNLOAD ENCRYPTED command that stores the data in an unencrypted S3 bucket; run a COPY command to move the data into the target cluster.
- c) Run an UNLOAD command that stores the data in an S3 bucket encrypted with an AWS KMS data key; run a COPY command to move the data into the target cluster.
- d) Connect to the source cluster over an SSL client connection, and write data records to Amazon Kinesis Firehose to load into your target data warehouse.

**Answer: b**

**Question: 8**

An organization needs a data store to handle the following data types and access patterns:

- Key-value access pattern
- Complex SQL queries and transactions
- Consistent reads
- Fixed schema

Which data store should the organization choose?

- Amazon S3
- Amazon Kinesis
- Amazon DynamoDB
- Amazon RDS

**Answer: d**

**Question: 9**

A mobile application collects data that must be stored in multiple Availability Zones within five minutes of being captured in the app.

What architecture securely meets these requirements?

- The mobile app should write to an S3 bucket that allows anonymous PutObject calls.
- The mobile app should authenticate with an Amazon Cognito identity that is authorized to write to an Amazon Kinesis Firehose with an Amazon S3 destination.
- The mobile app should authenticate with an embedded IAM access key that is authorized to write to an Amazon Kinesis Firehose with an Amazon S3 destination.
- The mobile app should call a REST-based service that stores data on Amazon EBS. Deploy the service on multiple EC2 instances across two Availability Zones.

**Answer: b**

**Question: 10**

A web application emits multiple types of events to Amazon Kinesis Streams for operational reporting. Critical events must be captured immediately before processing can continue, but informational events do not need to delay processing.

What is the most appropriate solution to record these different types of events?

- Log all events using the Kinesis Producer Library.
- Log critical events using the Kinesis Producer Library, and log informational events using the PutRecords API method.
- Log critical events using the PutRecords API method, and log informational events using the Kinesis Producer Library.
- Log all events using the PutRecords API method.

**Answer: c**



# Study Guide to Crack AWS Big Data Specialty BDS-C00 Exam:

- Getting details of the BDS-C00 syllabus, is the first step of a study plan. This pdf is going to be of ultimate help. Completion of the syllabus is must to pass the BDS-C00 exam.
- Making a schedule is vital. A structured method of preparation leads to success. A candidate must plan his schedule and follow it rigorously to attain success.
- Joining the AWS provided training for BDS-C00 exam could be of much help. If there is specific training for the exam, you can discover it from the link above.
- Read from the BDS-C00 sample questions to gain your idea about the actual exam questions. In this PDF useful sample questions are provided to make your exam preparation easy.
- Practicing on BDS-C00 practice tests is must. Continuous practice will make you an expert in all syllabus areas.

## Reliable Online Practice Test for BDS-C00 Certification

Make VMExam.com your best friend during your AWS Certified Big Data - Specialty exam preparation. We provide authentic practice tests for the BDS-C00 exam. Experts design these online practice tests, so we can offer you an exclusive experience of taking the actual BDS-C00 exam. We guarantee you 100% success in your first exam attempt if you continue practicing regularly. Don't bother if you don't get 100% marks in initial practice exam attempts. Just utilize the result section to know your strengths and weaknesses and prepare according to that until you get 100% with our practice tests. Our evaluation makes you confident, and you can score high in the BDS-C00 exam.

**Start Online practice of BDS-C00 Exam by visiting URL**

<https://www.vmexam.com/aws/bds-c00-aws-certified-big-data-specialty>