Cloudera
CDP-2001
CDP Administrator - Private Cloud Base Exam

Questions&AnswersPDF

ForMoreInformation:
https://www.certswarrior.com/

Features:

- ➢ 90DaysFreeUpdates
- ➢ 30DaysMoneyBackGuarantee
- ➢ InstantDownloadOncePurchased
- ➢ 24/7OnlineChat Support
- ➢ ItsLatestVersion

Latest Version: 6.0

Question: 1

You are troubleshooting a performance issue in Cloudera Data Platform (CDP) Private Cloud Base and need to collect diagnostic bundles. Which command would you use to manually trigger the collection of diagnostic bundles from a specific CDP service, for example, Cloudera Manager?

A.
```
cm-agent -collectDiagnosticBundles
```
B.
```
cloudera-manager-agent -collectDiagnosticBundles
```
C.
```
cdp-cli diagnostics collect -s ClouderaManager
```
D.
```
cdp-cli diagnostics collect -t ClouderaManager
```
E.
```
cdp-cli diagnostics collect -a ClouderaManager
```

Answer: C

Explanation: The command `cdp-cli diagnostics collect -s ClouderaManager` is used to manually trigger the collection of diagnostic bundles from a specific CDP service. The -s option specifies the service name. You can replace 'ClouderaManager' with the name of the specific service you want to collect bundles from, such as 'CDH' or 'HDFS'.

Question: 2

Your Cloudera CDP Private Cloud cluster is experiencing frequent ZooKeeper connection failures. You suspect a network issue. Which Cloudera Manager API command would provide information about the ZooKeeper server's network connectivity status?

A.
```
curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/services/ZOOKEEPER/state/status"
```
B.
```
curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/services/ZOOKEEPER/metrics/network/connections"
```
C.
```
curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/services/ZOOKEEPER/metrics/server/connection/timeout"
```
D.
```
curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/services/ZOOKEEPER/configs/zookeeper.client.port/value"
```
E.
```
curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/services/ZOOKEEPER/roles/server/status"
```

Explanation:
The correct API commands for checking ZooKeeper server's network connectivity status are: 1. curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/services/ZOOKEEPER/metrics/network/connections": This command retrieves the number of current network connections to the ZooKeeper server. Low or fluctuating connection counts can indicate network issues. 2. curl -u admin:admin –X GET "http://cm server:7180/api/v18/clusters/clusterName/services/ZOOKEEPER/metrics/server/connection/timeout": This command retrieves the average connection timeout value for ZooKeeper connections. Frequent timeouts could point towards network problems. Other options are incorrect because: * Option A: This retrieves the overall status of the ZooKeeper service, not specifically network information. * Option D: This retrieves the value of the ZooKeeper client port configuration property, which is not directly related to network connectivity status. Option E: This retrieves the status of the ZooKeeper server role, not network connectivity details.

Question: 3

You are investigating a high disk space utilization on the Cloudera CDP Private Cloud cluster. Which Cloudera Manager API command would provide information about the current disk usage for each host in the cluster?

A.
```
curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/hosts/all/disk/usage/total"
```
B.
```
curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/hosts/all/disk/free/space"
```
C.
```
curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/services/all/metrics/disk/usage/percent"
```
D.
```
curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/services/all/metrics/disk/io/rate"
```
E.
```
curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/hosts/all/state/status"
```

Explanation:
The correct API command to get information about current disk usage for each host is: curl -u admin:admin -X GET "http://cm- server:7180/api/v18/clusters/clusterName/hosts/all/disk/usage/total" This command retrieves the total disk usage for all hosts in the cluster, enabling you to identify which hosts are experiencing high utilization. Other options are incorrect because: * Option B: This command retrieves the free disk space, not the total usage. While useful, it doesn't directly pinpoint hosts with high disk utilization. Option C: This command retrieves the disk usage percentage for all services, not individual hosts. Option D: This command retrieves the disk 1/0 rate for all services, not disk usage information. * Option E: This command retrieves the overall status of all hosts, not specific disk usage details.

Question: 4

You need to identify the specific YARN application running on your Cloudera CDP Private Cloud cluster that is consuming the most resources. Which Cloudera Manager API command would provide information about the current resource utilization (memory, CPU) of individual YARN applications?

A.
 curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/services/YARN/metrics/resource/total/used"
B.
 curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/services/YARN/metrics/nodemanager/memory/used"
C.
 curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/services/YARN/applications/running/memory/usage"
D.
 curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/services/YARN/metrics/resourcemanager/cpu/usage"
E.
 curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/services/YARN/configs/yarn.nodemanager.resource.memory-mb/value"

Answer: C

Explanation:
The correct API command for retrieving the resource utilization of individual YARN applications is: curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/services/YARN/applications/running/memory/usage" This command provides the current memory usage of each running YARN application, allowing you to identify the application consuming the most resources. Other options are incorrect because: * Option A: This command retrieves the total YARN resources used, not individual application metrics. * Option B: This command retrieves the memory used by NodeManagers, not specific applications. Option D: This command retrieves the CPU usage of the ResourceManager, not individual applications. * Option E: This command retrieves the configuration property for the NodeManager's memory limit, not actual resource usage.

Question: 5

You are troubleshooting a Cloudera CDP Private Cloud cluster with a large number of data nodes. You want to quickly identify the data nodes with the highest CPU utilization. Which Cloudera Manager API command would provide this information?

A.
 curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/services/DATANODES/metrics/cpu/usage/average"
B.
 curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/hosts/all/metrics/cpu/usage/percent"
C.
 curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/services/all/metrics/cpu/usage/max"
D.
 curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/services/DATANODES/metrics/disk/io/rate"

E.

```
curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/services/DATANODES/state/status"
```

<div style="text-align: right">

Answer: B

</div>

Explanation:
The correct API command to identify data nodes with the highest CPU utilization is: curl -u admin:admin -X GET "http://cm-server:7180/api/v18/clusters/clusterName/hosts/all/metrics/cpu/usage/percent" This command retrieves the CPU usage percentage for all hosts in the cluster, allowing you to quickly identify data nodes with the highest utilization. Other options are incorrect because: * Option A: This command retrieves the average CPU usage for all DataNodes, not individual host-specific utilization. Option C: This command retrieves the maximum CPU usage across all services, not specifically data node metrics. * Option D: This command retrieves the disk 1/0 rate for DataNodes, not CPU usage information. * Option E: This command retrieves the overall status of the DataNodes service, not specific CPU usage details.

Question: 6

Which of the following factors influence the size and contents of a diagnostic bundle collected from a CDP Private Cloud Base environment?

A. The specific CDP services selected for bundle collection
B. The configuration of the 'Diagnostic Collection' policy
C. The time period for which logs and metrics are collected
D. The available disk space on the nodes where bundles are collected
E. The current state and health of the CDP services

<div style="text-align: right">

Answer: A,B,C,D,E

</div>

Explanation:
All of the listed factors influence the size and contents of a diagnostic bundle. The specific services selected for bundle collection determine the components included. The 'Diagnostic Collection' policy defines the level of detail (e.g., logs, metrics, configuration files) to be collected. The time period for which data is collected affects the size of the bundle. Available disk space limits the maximum bundle size, and the state and health of services influence the information included in the bundle (e.g., error logs, performance metrics).

Question: 7

You have manually triggered the collection of diagnostic bundles from your CDP Private Cloud Base environment. Where are the collected bundles stored by default?

A. In a dedicated directory within the Cloudera Manager server's file system
B. In a shared network file system accessible by all nodes in the cluster
C. In a dedicated S3 bucket within the AWS account associated with CDP
D. In a directory on each individual node where the bundles were collected

E. In a temporary location on the Cloudera Manager server, which is automatically deleted after transfer to Cloudera Support

Answer: D

Explanation:
By default, diagnostic bundles collected from CDP Private Cloud Base are stored in a dedicated directory on each individual node where the bundles were collected. This allows for easy access and retrieval, even if the Cloudera Manager server is unavailable. You can configure the bundle storage location to an alternate location, like an S3 bucket, but the default location is the local node.

Question: 8

How do you transfer diagnostic bundles collected from your CDP Private Cloud Base environment to Cloudera Support for analysis?

A. Using the cdp-cli diagnostics upload command, specifying the location of the bundles and the Cloudera Support case ID
B. By transferring the bundle files manually to a dedicated Cloudera Support server via FTP or SFTP
C. By uploading the bundles to a shared network file system, which is accessible to Cloudera Support
D. Through the Cloudera Manager web IJI, where you can select the collected bundles and upload them directly to Cloudera Support
E. Using a third-party file transfer tool like WinSCP or Filezilla to connect to the Cloudera Manager server and transfer the bundles

Answer: A

Explanation: You can transfer diagnostic bundles collected from CDP Private Cloud Base to Cloudera Support for analysis using the cdp-cli command. This command allows you to specify the location of the bundles and the Cloudera Support case ID. This ensures diagnostics upload that the bundles are securely uploaded to the correct location for efficient analysis.

Question: 9

Your CDP Private Cloud Base environment is experiencing intermittent performance issues. You suspect a specific service, for example, Apache Spark, might be the cause. What steps would you take to collect diagnostic bundles for this service and analyze them to identify the root cause?

A. 1. Identify the specific service (Spark) experiencing issues. 2. Use the cdp-cli diagnostics collect -s spark command to manually trigger bundle collection. 3. After bundle collection, examine log files within the bundle for error messages or warning signals. 4. Analyze performance metrics within the bundle to identify any unusual trends or spikes.
B. 1. Trigger automatic bundle collection using the 'Diagnostic Collection' policy. 2. Review the collected bundles, focusing on the configuration files for the suspect service (Spark). 3. Identify any misconfigured settings or missing dependencies that might be contributing to performance issues.

C. 1. Analyze performance metrics through Cloudera Manager's web IJI for the service in question (Spark). 2. If anomalies are detected, use the cdp-cli diagnostics collect -s spark command to collect diagnostic bundles. 3. Examine logs and configuration files within the bundle to confirm potential issues.
D. 1. Access the Cloudera Manager server and manually collect the diagnostic bundles from the service directory for Spark. 2. Analyze the bundle content, including logs and system information, to determine the root cause.
E. 1. Use the 'Diagnostic Collection' policy to collect bundles from all CDP services. 2. Filter the collected bundles to focus on the logs and metrics specific to the suspect service (Spark) for analysis.

Answer: A,C

Explanation:
The most effective approaches involve focusing on the specific service in question. Options A and C provide a focused and efficient method for collecting and analyzing diagnostic bundles. Option A directly targets the suspect service, while Option C utilizes monitoring tools to identify potential issues before manually collecting the bundle. Option B, D, and E are less efficient, as they involve collecting a broader range of data, making it more difficult to isolate the issue quickly.

Question: 10

You are monitoring a Cloudera Data Platform (CDP) cluster and notice a significant spike in the number of failed tasks on a specific Impala service. Your investigation reveals that the Impala daemon is experiencing high CPU utilization and memory pressure. Which of the following actions are the most appropriate steps to address this issue?

A. Increase the number of Impala daemons on the cluster nodes.
B. Review the Impala queries being executed and identify any inefficient queries that might be contributing to the resource consumption.
C. Check for any hardware failures or resource contention on the cluster nodes.
D. Adjust the Impala configuration settings, such as the number of concurrent queries or the maximum memory allocation.
E. Scale down the cluster resources to reduce the overall workload on the Impala service.

Answer: B,C,D

Explanation: The most appropriate steps to address high CPU utilization and memory pressure on an Impala service include: Reviewing Impala queries: Identifying and optimizing inefficient queries can significantly reduce the workload on the Impala daemon. * Checking for hardware failures: Hardware issues like faulty disks or insufficient RAM can contribute to resource contention and impact Impala performance. Adjusting Impala configuration: Adjusting configuration settings such as the number of concurrent queries or the maximum memory allocation can help manage resource consumption and improve performance. Increasing the number of Impala daemons (Option A) might not be the most effective solution, as it could lead to further resource contention if the underlying issue is not addressed. Scaling down the cluster resources (Option E) would only reduce the overall workload, potentially hindering other services and not directly addressing the Impala performance problem.