Hide menu

Monitoring & Tuning

Course Final Exam

Reading: Instructions for the Final Exam
2 min

Quiz: Final Exam
20 questions

Course Wrap-Up

Final Exam

Quiz • 1h 40m

Submit your assignment

Due Aug 27, 11:59 PM EDT **Attempts** 3 every 24 hours

Start assignment

Receive grade

To Pass 80% or higher

Your grade

-

Like **Dislike** **Report an issue**

1. How does Apache Spark solve read/write problems encountered by other tools? **1 point**

- (•) By keeping much of the required data in-memory.
- ( ) By leveraging redundancy.
- ( ) By only using certain processors in the distributed group.
- ( ) By using special proprietary APIs.

48:55

Expand

2. One component of Spark architecture is the executor. Which of the following is true? **1 point**

- ( ) Executors control the function of drivers.
- ( ) Only one executor can function at a time.
- ( ) Executors work on only one worker node.
- (•) Executors complete single tasks.

48:55

Expand

3. Which of the following is NOT a way to create a dataset? **1 point**

- ( ) From a text file using an explicit schema declaration and the "String" data type
- ( ) Using the toDS function in Scala
- (•) Combine DataFrames within a dataset.
- ( ) From a JSON file and custom classes

Expand

4. Which of the following is true of Tungsten? **1 point**

- ( ) Does rely on the JVM object model.
- ( ) Does not support on-demand JVM byte code generation.
- (•) Does not generate virtual function dispatches.
- ( ) Does not enable computation of algorithms using STRIDE-based memory access.

Expand

5. Which tool mentioned in the course creates production-ready environments for AI and machine learning? **1 point**

- ( ) IBM Analytics Engine
- (•) IBM Watson
- ( ) Spark
- ( ) Apache Hadoop Cluster

Expand

6. What is the name of the Spark unified interface? **1 point**

- (•) SUI
- ( ) spark-default
- ( ) YARN
- ( ) spark-submit

Expand

7. Why does Spark queue tasks and wait for available cores? **1 point**

- ( ) To keep the number of cores in the pool low
- ( ) To use more cores
- (•) To maximize parallel processing
- ( ) To start as many tasks as possible

Expand

8. If a task fails due to a dependency problem, what is the best way to identify the issue? **1 point**

- ( ) Cataloging the libraries on the system
- ( ) Checking required data files for corruption
- ( ) Checking APIs
- (•) Examining the event log for stack trace errors

Expand

9. Which of the following was mentioned in the course as a common application of Big Data? **1 point**

- ( ) Writing new video games
- ( ) Optimizing streaming video services
- (•) Recommendation engines on websites like Amazon and Google
- ( ) Running automotive assembly lines

Expand

10. Which of the following is true of parallel processing? **1 point**

- (•) It's the best technique for processing Big Data.
- ( ) It's difficult to scale.
- ( ) It can be inefficient and time-consuming.
- ( ) It's not particularly flexible.

Expand

11. What is Data Scaling? **1 point**

- (•) Data scaling is a technique to manage, store, and process the overflow of data.
- ( ) Data scaling divides workloads to run in parallel.
- ( ) Data scaling is only applicable within cloud environments.
- ( ) Data scaling is the process of transforming data values for end use.

Expand

12. Semi-structured data _____. **1 point**

- (•) Includes some metadata that identifies certain characteristics.
- ( ) Has a pre-structured data model.
- ( ) Includes sensor data from Internet of Things devices.
- ( ) Includes databases and spreadsheets.

Expand

13. Which of these is a good situation for using Hadoop? **1 point**

- ( ) Processing many small files
- ( ) Processing data with dependencies
- (•) Processing enormous data sets
- ( ) Processing transactions

Expand

14. A driver program is a process that _____. **1 point**

- (•) creates work and sends it to the cluster
- ( ) acts in parallel to do work
- ( ) runs multiple threads
- ( ) is one of several similar processes in an application

Expand

15. Which configuration method enables you to adjust settings on a per-machine basis? **1 point**

- ( ) Manual
- (•) Environment variables
- ( ) Properties
- ( ) Logging

Expand

16. Kubernetes runs containerized applications on a cluster. What else is true about it? **1 point**

- ( ) It only runs in the cloud.
- ( ) It cannot be run on a single machine.
- (•) It's portable.
- ( ) It cannot be deployed automatically.

Expand

17. The biggest component of Big Data is _____. **1 point**

- ( ) Kubernetes
- ( ) Apache Spark
- (•) Hadoop
- ( ) HDP

Expand

18. Which of the following happens in the map task of MapReduce? **1 point**

- ( ) Give consistent names to pieces of data
- ( ) Aggregate a set of results
- ( ) Produce a final report
- (•) Process data into key value pairs

Expand

19. What are the three components of Hive architecture? **1 point**

- ( ) Clients, Services, Execution
- (•) Services, Metastore, Database
- ( ) Clients, Services, Storage and Computing
- ( ) Storage, Computing, Command Line Interface

Expand

20. Which of the following is a benefit of using the Spark Application UI to monitor applications? **1 point**

- ( ) You can pinpoint and remove corrupted data.
- ( ) You can limit usage if the system is overloaded.
- (•) You can quickly identify failed jobs and locate the root cause of failure.
- ( ) You can install new applications when necessary.

Expand

Upgrade to submit