

Knowledge Check 10: Data Management in MapReduce Systems

Due No due date **Points** 0 **Questions** 10
Time Limit None **Allowed Attempts** Unlimited

Take the Quiz Again

Attempt History

	Attempt	Time	Score
LATEST	<u>Attempt 1</u>	7 minutes	0 out of 0

Submitted Feb 13 at 8:13pm

Question 1

0 / 0 pts

How does Hadoop and MapReduce relate to each other?

You Answered

- Hadoop is a server-side application while MapReduce runs on client computers.

This is not correct. Hadoop MapReduce system is a distributed computing application, not a client-server application.

- ☐ Hadoop is a computer operating system while MapReduce is a software application.

Correct Answer

- ☐ MapReduce is the framework used by Hadoop software.
- ☐ Hadoop is the framework used by MapReduce software.

Question 2**0 / 0 pts**

Which of the following is the correct order of functions in the typical processing of big data?

Correct!

- ☒ The Map function has to finish before the Reduce function starts.

Correct! All of the mapping processes in the system must be finished before the Reduce function starts.

- ☐ The Reduce function has to finish before the Map function starts.
- ☐ Map and Reduce functions can run simultaneously.
- ☐ Map and Reduce functions can run in parallel.

Question 3**0 / 0 pts**

What is the name of the transitional phase between the Map and Reduce phases in a big data process?

Correct!

- ☐ Data Scrubbing

- ☒ Data shuffling

Correct! Data shuffling is an intermediate step between the Map and Reduce phases.

- ☐ Data Mapping
- ☐ Data Mining

Question 4**0 / 0 pts**

What happens during the data shuffling phase in a typical big data process?

☐

Data generated during the reduce phase is routed to different nodes in the cluster.

☐

Data generated during the map phase is encrypted to make it secure.

☐

Data generated during the reduce phase is encrypted to make it secure.

Correct!☒

Data generated during the map phase is routed to different nodes in the cluster.

Correct! Data shuffling is an intermediate phase where mapping data is routed to different nodes in the cluster.

Question 5**0 / 0 pts**

Which of the following phases in a typical Hadoop process provide full programming control to users?

☐

Reduce and Shuffling

☐

Compress and Shuffling

☐

Map and Reduce

☒

Map and Shuffling

This is not correct. The shuffling phase is not fully controlled by users.

Correct Answer**You Answered**

Question 6**0 / 0 pts**

How many copies of a piece of data are generated by the Hadoop File System (HDFS) in order to allow for fault tolerance?

☐ 8☒ 3

Correct! The HDFS replicates data three times for fault tolerance.

☐ 10☐ 64**Correct!****Question 7****0 / 0 pts**

In which programming language are Map and Reduce functions written?

☒ Java

Map and Reduce functions are written in Java.

☐ Python☐ C++☐ HTML**Correct!****Question 8****0 / 0 pts**

What is the size of each data block in the Hadoop file system?

You Answered

☐ 1 GB☒ 100 MB

This is not correct. Hadoop file system data blocks are 128 MB each.

☐ 1 MB

Correct Answer

☐ 128 MB**Question 9****0 / 0 pts**

What does each node correspond to in a Hadoop cluster?

You Answered

☒ Data Center

This is not correct. Data center is not called a node.

☐ A data block

Correct Answer

☐ A computing machine☐ A data cloud**Question 10****0 / 0 pts**

What is the name of the special node in a Hadoop cluster that stores metadata of the entire cluster?

☐ Meta node**Correct!**☒ Name node

Correct! The name node stores metadata of a Hadoop cluster.

☐ Master node

☐ Hub node