

Quiz

Intro to Hadoop

1. What does IaaS provide?

1 / 1 point

- ☒ Hardware Only
- ☐ Software On-Demand
- ☐ Computing Environment

✓ **Correct**
See [this video](#) to review.

2. What does PaaS provide?

1 / 1 point

- ☐ Software On-Demand
- ☐ Hardware Only
- ☒ Computing Environment

✓ **Correct**
See [this video](#) to review.

3. What does SaaS provide?

1 / 1 point

- ☒ Software On-Demand
- ☐ Hardware Only
- ☐ Computing Environment

✓ **Correct**
See [this video](#) to review.

4. What are the two key components of HDFS and what are they used for?

1 / 1 point

- ☐ FASTA for genome sequence and Rasters for geospatial data.
- ☒ NameNode for metadata and DataNode for block storage.
- ☐ NameNode for block storage and Data Node for metadata.

✓ **Correct**
See [this video](#) to review.

5. What is the job of the NameNode?

1 / 1 point

- ☒ Coordinate operations and assigns tasks to Data Nodes
- ☐ Listens from DataNode for block creation, deletion, and replication.
- ☐ For gene sequencing calculations.

✓ **Correct**

See [this video](#) to review.

6. What is the order of the three steps to Map Reduce?

1 / 1 point

- ☐ Shuffle and Sort -> Map -> Reduce
- ☐ Shuffle and Sort -> Reduce -> Map
- ☒ Map -> Shuffle and Sort -> Reduce
- ☐ Map -> Reduce -> Shuffle and Sort

✓ **Correct**

See [this video](#) to review.

7. What is a benefit of using pre-built Hadoop images?

1 / 1 point

- ☐ Quick prototyping, deploying, and guaranteed bug free.
- ☐ Guaranteed hardware support.
- ☐ Less software choices to choose from.
- ☒ Quick prototyping, deploying, and validating of projects.

✓ **Correct**

See [this video](#) to review.

8. What is an example of open-source tools built for Hadoop and what does it do?

1 / 1 point

- ☐ Giraph, for SQL-like queries.
- ☐ Zookeeper, analyze social graphs.
- ☒ Giraph, for processing large-scale graphs.
- ☐ Pig, for real-time and in-memory processing of big data.

✓ **Correct**

See [this video](#) to review.

9. What is the difference between low level interfaces and high level interfaces?

1 / 1 point

- ☒ Low level deals with storage and scheduling while high level deals with interactivity.
- ☐ Low level deals with interactivity while high level deals with storage and scheduling.

✓ **Correct**
See [this video](#) to review.

10. Which of the following are problems to look out for when integrating your project with Hadoop?

1 / 1 point

☒ Task Level Parallelism

✓ **Correct**
See [this video](#) to review.

☒ Infrastructure Replacement

✓ **Correct**
See [this video](#) to review.

☒ Advanced Algorithms

✓ **Correct**
See [this video](#) to review.

☐ Data Level Parallelism

☒ Random Data Access

✓ **Correct**
See [this video](#) to review.

11. As covered in the slides, which of the following are the major goals of Hadoop?

1 / 1 point

☒ Enable Scalability

✓ **Correct**
See [this video](#) to review.

☒ Facilitate a Shared Environment

✓ **Correct**
See [this video](#) to review.

☒ Provide Value for Data

✓ **Correct**
See [this video](#) to review.

☐ Latency Sensitive Tasks

☒ Handle Fault Tolerance

✓ **Correct**
See [this video](#) to review.

☒ Optimized for a Variety of Data Types

✓ **Correct**
See [this video](#) to review.

12. What is the purpose of YARN?

1 / 1 point

- ☒ Allows various applications to run on the same Hadoop cluster.
- ☐ Enables large scale data across clusters.
- ☐ Implementation of Map Reduce.

✓ **Correct**
See [this video](#) to review.

13. What are the two main components for a data computation framework that were described in the slides?

1 / 1 point

- ☐ Node Manager and Container
- ☒ Resource Manager and Node Manager
- ☐ Node Manager and Applications Master
- ☐ Resource Manager and Container
- ☐ Applications Master and Container



Correct

See [this video](#) to review.