

## Quiz 4

**Due** Mar 17 at 5pm**Points** 20**Questions** 13**Available** Mar 13 at 7pm - Mar 17 at 5pm**Time Limit** 30 Minutes

## Instructions

You have 30 minutes to complete this quiz about Storage Virtualization, Cloud Databases and Big Data.

The quiz has 13 questions worth 20 points.

**PLEASE DO NOT CONSULT OTHER STUDENTS.**

## Attempt History

	Attempt	Time	Score
<b>LATEST</b>	<u><a href="#">Attempt 1</a></u>	30 minutes	16 out of 20 *

\* Some questions not yet graded

 **Correct answers are hidden.**

Score for this quiz: **16** out of 20 \*

Submitted Mar 16 at 6:45pm

This attempt took 30 minutes.

### Question 1

2 / 2 pts

What is not a drawback of host-based storage virtualization?

- ☐ Harder to scale to many hosts
- ☐ Must manage virtualization software on every host
- ☒ Must retrieve metadata for each host

- ☐ Uses extra resources on the host

**Question 2****2 / 2 pts**

What is not a drawback of storage device-based storage virtualization?

- ☐ Hard to mix and match storage targets
- ☐ Hard to visualize the virtualization logic
- ☐ Each storage target is likely to cost more
- ☒ Generates network bottlenecks due to metadata transfer

**Question 3****Not yet graded / 2 pts**

Please provide an example when network-based storage virtualization makes more sense than host-based storage virtualization.

Your Answer:

- Network-based storage virtualization is more suitable when there is a need to provide storage to multiple servers.
- Host-based storage virtualization is better suited for smaller environments with fewer hosts and storage resources.
- When many servers in a big company need to access the same storage, it is better to use network-based storage virtualization. This means that all the servers can share a single storage pool over the network.
- It can make better use of resources and make it easier to manage storage.

**Question 4****1 / 1 pts**

True or False

In a cloud environment where the cloud provider implemented storage virtualization, it is impossible to recover data from a failed implementation of the storage virtualization software.

☐ True☒ False**Question 5****1 / 1 pts**

True or False

Out-of-band storage virtualization is more robust to failures.

☒ True☐ False**Question 6****2 / 2 pts**

File storage is good for...

☐ Unstructured data storage☒ Image sharing☐ Database storage☐ Storage of large number of files

**Question 7****2 / 2 pts**

What is a good use case of block storage?

- ☐ Storing audio files
- ☒ Storing qcow images
- ☐ Storing COEN 241 HW Instructions
- ☐ Storing social security numbers in txt files

**Question 8****1 / 1 pts**

True or False

Object storage does not allow you to have "a notion of directories" as every object is stored at the same level.

- ☐ True
- ☒ False

**Question 9****1 / 1 pts**

True or False

You cannot use file storage for data analytics, because it is way too slow.

- ☐ True

☒ False

### Question 10

1 / 1 pts

True or False

Assuming you have infinite amount of read and writes to the cloud storage, file storage is the cheapest.

☒ True

☐ False

### Question 11

2 / 2 pts

Which option is not considered Big Data?

☐ Data arriving at a continuous stream from multiple sensors

☒ Relational database containing grades for COEN 241 students

☐ User behavior data that is hard to fit in a single machine

☐ Cloud object storage consisting of logs, images and audio files over 10 years

### Question 12

1 / 1 pts

True or False

MapReduce is useful for complex algorithms requiring amount of resources with tight dependency between each resource.

☐ True

☒ False

### Question 13

Not yet graded / 2 pts

Provide an example where hadoop would be more beneficial than spark.

Your Answer:

- Hadoop would be more beneficial than Spark in situations where there is a need to process and store large volumes of data in a distributed manner.
- Hadoop can split up the log data into smaller chunks and process them in parallel across multiple computers, making it faster and more efficient.
- Each node can process its own chunk of data in parallel, which can result in faster processing times.
- The MapReduce framework consists of two main functions: Map and Reduce. The Map function takes the input data and converts it into key-value pairs. The Reduce function then takes these key-value pairs and performs a summary operation on them to generate the final output.
- The MapReduce framework also includes a built-in fault-tolerance mechanism that ensures that if a node fails during processing, the task is automatically reassigned to another node in the cluster. This makes the system more reliable and ensures that processing can continue even if some nodes fail.
- 

Quiz Score: **16** out of 20