

Big Data SoSe 2025

Module Overview

Lecture Materials

Exercise Materials

Session 1: Introduction

Session 2: Storage Infras

Apache Spark™ Programming

Quiz 2: Storage Infrastru

Session 3: Column Stores

Quiz 3: Column Stores

Session 4: Distributed Da

Data Ingestion with Delta

Quiz 4: Distributed Data

Forum

Quiz 3: Column Stores

Tutorial on 15.05.2025

Performance summary

Assessed

Success status



Undefined

Score



39 of 40 points

Attempts



1

Results

Course	Big Data SoSe 2025 ID: 4853531345 / 107711829655247
Test	Test 3 Big Data 2025 ID: 4903141552

This are your test results

Duration	0h 14m 17s	5/10/2025, 2:33 PM - 5/10/2025, 2:48 PM
Answered	8 of 8 questions (100%)	
Your score	39 of 40 points (98%)	



Recall

Relational vs non-Relational DB

Status

Answered

Your score

6 / 6

100%

Response

Which of the following statements are true for RDBMS and which for NoSQL? Please note that some statements might be true for both.

	RDBMS	NoSQL
Column based storage	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Single data record represented as Tuple	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Partition tolerance	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Allow to retrieve, update, and delete stored data	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Support more unstructured data such as JSON files	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Vertically scalable	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Aggregate-oriented databases

Status

Answered

Your score

4 / 4

100%

Response

Which of the following statements are true for Aggregate-oriented databases?

Unanswered	Right	Wrong	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	In Aggregate-oriented databases, it is easier to manage data storage over clusters
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Aggregate-oriented databases store data in the form of Tuples
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	There are ACID transactions that span multiple aggregates
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Allow you to manipulate any combination of rows from any table in a single transaction

BigTable

Status

Answered

Your score

4 / 4

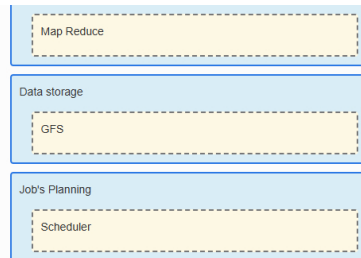
100%

Response

master election, location bootstrapping

Lock service

simplified large-scale data processing



HBase and Hive

Status

Answered

Your score

5 / 6

83%

Response

Which of the following statements are True?

Unanswered	Right	Wrong	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	HBase allows random write and update
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	HBase is not suitable fore individual record look up
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	HBase uses Column storage instead of tables
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Hive is built on Hadoop
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Hive is a relational database
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	HBase is build on HDFS

STF Scheduling

Status

Answered

Your score

10 / 10

100%

Response

Task	Length	Arrival
T1	8	0
T2	2	5
T3	1	4
T4	5	2

Take a look at the table above, where 4 tasks with their length and arrival times are shown.
Calculate the Average completion time in case of STF Scheduling (Shortest Task First)
Round your answer to the nearest hundredth

8.25

Knowledge Questions

Bigtable

Status

Answered

Your score

5 / 5

100%

Response

Please read the paper "Bigtable: A Distributed Storage System for Structured Data" and answer the following

Unanswered	Right	Wrong	
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	The Bigtable master is a potential bottleneck since all read/write operations go through it.
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Bigtable's SSTables are immutable after creation.
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Compression in Bigtable is applied across entire tables for maximum efficiency.
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Bigtable supports general-purpose distributed transactions across multiple rows.
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Major compactions in Bigtable are used to remove deleted data and obsolete versions.

Facebook's Use of HBase

Status

Answered

Your score

4 / 4

100%

Response

Please answer based on the key concepts and components discussed in the Facebook Messages HBase infrastructure paper

Used as a task queue for distributed coordination (e.g., log splitting)

Zookeeper

Increases cache efficiency by reducing redundancy in data blocks

Data Block Encoding

Improves read performance by filtering non-matching HFiles

Bloom Filters

Tests new systems using real-world traffic without affecting production

Shadow Testing

⦿ Data Modeling in the NoSQL World

Status

Answered

Your score

1 / 1

100%

Response

☐ To standardize SQL queries for NoSQL databases

☐ To implement a new type of relational database

☐ To replace existing NoSQL systems with a new database engine

☒ To provide a system-independent data model to support NoSQL database design

Test execution

Information

⌚ Availability: Expired at 5/15/2025, 4:00 PM

🔄 Max. attempts: Unlimited

👁 Results of this test are visible to administrators and tutors of this course.

Start test

▶ Change log

[^ Go to top](#)

Logged in as Ravi Himmatbhai Ramani (5452 People are online)

Imprint
Datenschutzerklärung

OpenOlat 19.1.11
