

Liverpool John Moores University

Title: APPLICATIONS IN DISTRIBUTED DATABASES
Status: Definitive
Code: **7125COMP** (121348)
Version Start Date: 01-08-2021

Owning School/Faculty: Computer Science and Mathematics
Teaching School/Faculty: Computer Science and Mathematics

Team	Leader
Andrew Laws	Y

Academic Level: FHEQ7
Credit Value: 20
Total Delivered Hours: 55
Total Learning Hours: 200
Private Study: 145

Delivery Options

Course typically offered: Semester 2

Component	Contact Hours
Lecture	22
Practical	33

Grading Basis: 50 %

Assessment Details

Category	Short Description	Description	Weighting (%)	Exam Duration
Artefacts	AS1	Database setup and population	40	
Artefacts	AS2	Database application development	60	

Aims

The aim of the module is to provide students with hands-on experience of setting up, populating and querying large scale distributed databases using de-facto industry standard tools.

Learning Outcomes

After completing the module the student should be able to:

- 1 Formulate, design and construct robust example distributed databases.
- 2 Formulate effective query mechanisms for distributed databases using de facto industry standard tools.
- 3 Critically evaluate the effectiveness of the solutions developed.

Learning Outcomes of Assessments

The assessment item list is assessed via the learning outcomes listed:

Database setup and population	1	
Database application develop	2	3

Outline Syllabus

Distributed databases, tools and approaches

Review Mapreduce and Hadoop

Cluster setup tools e.g. Ambari

Populating distributed databases

Re-configuring SQL database data for distributed datases e.g. Sqoop

Hadoop Distributed File system (HDFS)

Fault tolerance in distributed file systems

Google's Bigtable, HBase and GUI Clients e.g. HareDB

Data extraction tools e.g. Hive

Distributed database application development e.g. Pig

Large cluster management tools e.g. Zookeeper

Emerging trends in applications of large scale databases

Learning Activities

Lectures will explore the practical aspects of large scale data storage approaches
Practical sessions explore and practice those ideas to provide hands-on experience in building and managing large scale data storage solutions.

Notes

This module provides both theoretical and practical experience of large scale data storage considerations and the use of tools to support the processing of that data.