# Big Data Systems

## Module Information

**2022.01, Approved**

## Summary Information

| Module Code | 7123COMP |
|---|---|
| Formal Module Title | Big Data Systems |
| Owning School | Computer Science and Mathematics |
| Career | Postgraduate Taught |
| Credits | 20 |
| Academic level | FHEQ Level 7 |
| Grading Schema | 50 |

**Teaching Responsibility**

| LJMU Schools involved in Delivery |
|---|
| Computer Science and Mathematics |

## Learning Methods

| Learning Method Type | Hours |
|---|---|
| Lecture | 11 |
| Practical | 11 |
| Tutorial | 11 |

## Module Offering(s)

| Display Name | Location | Start Month | Duration Number Duration Unit |
|---|---|---|---|
| JAN-CTY | CTY | January | 12 Weeks |

## Aims and Outcomes

| Aims | To allow students to develop new advanced cloud-based software development skillsTo combine existing and new networking and software development skills in a practical 'Big Data' contextTo use real-world Cloud-based and locally administered systems to apply their knowledge to 'Big Data' problems |
|---|---|

**After completing the module the student should be able to:**

**Learning Outcomes**

| Code | Number | Description |
|---|---|---|
| MLO1 | 1 | Evaluate new paradigms and techniques for data management, systems and programming |
| MLO2 | 2 | Design solutions using a big data paradigm that access and manipulate in a massively parallel fashion for offline stored data |
| MLO3 | 3 | Apply big data programming tools to build applications that access and manipulate in a massively parallel fashion for online streaming data |
| MLO4 | 4 | Interpret and categorise challenges involved in developing large scale applications with big data systems |

## Module Content

| Outline Syllabus | Distributed systems issues (fault tolerance, high performance, resource utilisation, caching, load balancing)Distributed file systemsData Management with data modelling, consistency, query processing, indexingFunctional language that underpins large scale parallelism, e.g. Map-ReduceParallel programming modelQuery languages and environments for big data systemsPublish/subscribe systems for large scale information dissemination |
|---|---|
| Module Overview | |
| Additional Information | The theoretical work will build on students existing knowledge of distributed system, refocussing on its application to building large scale cloud applications. The practical element will involve hands-on cloud application development using real-world cloud services. |

## Assessments

| Assignment Category | Assessment Name | Weight | Exam/Test Length (hours) | Module Learning Outcome Mapping |
|---|---|---|---|---|
| Report | Big Data System Development | 50 | 0 | MLO1, MLO2, MLO3 |
| Centralised Exam | Exam | 50 | 2 | MLO1, MLO4 |

## Module Contacts

**Module Leader**

| Contact Name | Applies to all offerings | Offerings |
|---|---|---|
| Paul Fergus | Yes | N/A |

**Partner Module Team**

| Contact Name | Applies to all offerings | Offerings |
|---|---|---|