

OLLSCOIL NA hÉIREANN, CORCAIGH
THE NATIONAL UNIVERSITY OF IRELAND, CORK

COLÁISTE NA hOLLSCOILE, CORCAIGH
UNIVERSITY COLLEGE, CORK

End-Semester Examination
Semester 1 - Winter 2018

CS6321 Model-based Software Development
2018-2019

M.Sc. Computer Science

Answer all questions

Total marks: 40

Time: 60 minutes

PLEASE DO NOT TURN THIS PAGE UNTIL INSTRUCTED TO DO SO
THEN ENSURE THAT YOU HAVE THE CORRECT EXAM PAPER

1. Compare containerization and virtualization technologies. Explain why containerization technology cannot completely replace virtualization technology? Use an example to support your answer.
(4 marks)
2. When designing a cloud application, what are the three most important quality attributes a developer needs to consider? Use an example to show why a cloud application developer should pay particular attention to these attributes.
(4 marks)
3. There are two common approaches for doing autoscaling, namely scaling-out/in and scaling-up/down. What is the main difference between these two autoscaling approach?
(2 marks)
4. In cloud storage offerings, storage can be operated in either strict consistency or eventual consistency mode. State the main difference between the two storage operation modes. Give a real-world application where eventual consistency mode storage is most appropriate.
(3 marks)
5. What is the Idempotent processor in cloud processing offering? Give two arguments on why the use of idempotent processor is important for cloud applications?
(3 marks)
6. Compare distributed file systems (e.g., HDFS) and relational database systems under each of the following headings: processing paradigms, operation model, data schema, and scalability.
(2 marks)
7. Monitoring activity on twitter for a certain set of corporate users generates a large data file containing the number of tweets per day over an extended period, each line representing the number of tweets for a user on a particular day, e.g. @BigBroCorp 22.
Hadoop is a popular big data analytics platform, consisting of a distributed file system (HDFS) and a data processing paradigm (MapReduce). Give an overview, with diagram(s), of a MapReduce architecture to determine the maximum daily tweets for each user during the monitored period.
(4 marks)
8. A software system supporting an online tv and film streaming service is being redesigned as a microservices architecture. Describe two main characteristics of a microservices architecture.
(1 mark)

9. In the context of the Apache Spark framework, what is the Resilient Distributed Dataset (RDD) Lineage?
(2 marks)
10. In a large-scale computing environment, ensuring data high-availability is vitally important. Discuss how Hadoop and how Spark ensure high-availability of data.
(4 marks)
11. Discuss how Apache Spark deals with streaming data. (Use a diagram to support your answer)
(4 marks)
12. Explain the PageRank Algorithm.
(4 marks)
13. Modularity is central for designing a scalable and extensible software system. Controlling software class visibility is one of the main tasks when designing modules. Explain how class visibility control is achieved in the OSGi framework.
(1 mark)
14. Serverless computing is about dynamically allocating resources for event-driven function execution. State two advantages and two disadvantages of using serverless computing.
(2 marks)