IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2015

MEng Honours Degree in Mathematics and Computer Science Part IV
MEng Honours Degrees in Computing Part IV
MSc in Computing Science (Specialist)
MRes in High Performance Embedded and Distributed Systems
for Internal Students of the Imperial College of Science, Technology and Medicine

This paper is also taken for the relevant examinations for the Associateship of the City and Guilds of London Institute

PAPER C410H

SCALABLE DISTRIBUTED SYSTEMS DESIGN

Wednesday 25 March 2015, 10:00
Duration: 60 minutes

Answer TWO questions

Paper contains 3 questions
Calculators not required

1 a State and explain the *CAP theorem* that governs the design of distributed systems.

b Explain how each of the following systems complies with the CAP theorem. For each, give the reasons why a particular trade-off was chosen.

i) Amazon Dynamo

ii) Google Bigtable

iii) Apache Zookeeper

iv) Apache Spark

c The CAP theorem has been criticised for the fact that not all of its choices are sensible. Discuss this statement, and suggest a better version of the CAP theorem.

The three parts carry, respectively, 20% 60%, and 20% of the marks.

2a Briefly explain each of the following concepts.

i) Data partitioning

ii) Data-parallel processing

iii) Data replication

iv) Eventual consistency

b In distributed dataflow frameworks, workers execute compute tasks on a cluster of machines.

i) Explain the importance of fault-tolerance in such frameworks.

ii) Describe **two** different techniques for making computation in a distributed dataflow framework fault-tolerant, and give examples of frameworks adopting each technique.

iii) Discuss the trade-offs between the two fault-tolerance techniques that you identified under ii).

iv) Suggest a hybrid fault-tolerance technique that combines the benefits of the two techniques that you identified under ii).

The two parts carry, respectively, 40% and 60% of the marks.

3 You work as a software engineer for YOURAREA, a start-up company that wants to design a distributed system that can push dynamic, geographically-local content, such as shopping promotions and news bulletins, to mobile devices. As soon as new content becomes available from external sources, it should be sent to potentially millions of devices.

The YOURAREA system should allow mobile devices to provide *subscriptions* that define the *content messages* that devices want to receive. Subscriptions include filter expressions in order to decide if particular content messages are relevant for a given device, e.g. by defining a geographic area. Subscriptions change over time, e.g. as devices move between locations. Content messages contain attributes such as the specification of geographic relevance that can be matched against the filters in subscriptions.

The design of the YOURAREA system should be (i) incrementally-scalable, (ii) fault-tolerant and (iii) maintain low latency (on the order of seconds) when pushing content messages to mobile devices. The system should be deployed in a single data centre.

(You should make justified decisions about any other requirements that are left unspecified.)

a Draw a high-level diagram of the *system design*, clearly labelling all distributed components. Explain the operation of each component, and justify its function.

b Explain how your design is *scalable* in that it can handle an increasing number of concurrent subscriptions and content messages.

c Explain how your design is *fault-tolerant* in that it can handle the failure of individual machines in the data centre.

The three parts carry, respectively, 35% 35%, and 30% of the marks.