Index No:							

UNIVERSITY OF MORATUWA

Faculty of Engineering

Department of Computer Science & Engineering

B.Sc. Engineering
2013 Intake Semester 8 Examination

CS4262 Distributed Systems

Time allowed: 2 Hours December 2017

ADDITIONAL MATERIAL: *None*

INSTRUCTIONS TO CANDIDATES:

1. This paper contains **4** questions on **12** pages.
2. This examination accounts for 50% of the module assessment.
3. This is a closed book examination.
4. Answer **ALL** questions.
5. Answer the questions on the paper itself. **DO NOT** exceed the given space.
6. For MCQ and True/False questions, select the most appropriate answer. No penalty for wrong answers.
7. The maximum attainable mark for each question is given in brackets.
8. Only calculators approved by the Faculty of Engineering are permitted.
9. Assume reasonable values for any data not given in or with the examination paper. Clearly state such assumptions made on the script.
10. If you have any doubt as to the interpretation of the wording of a question, make your own decision, but clearly state it on the script.
11. Electronic/Communication devices are not permitted. Only equipment allowed is a calculator approved and labelled by the Faculty of Engineering.
12. This paper should be answered only in English.

Q1	Q2	Q3	Q4	Total

Question 1 **[25 marks]**

Circle the most appropriate answer. [2 × 10 marks]

(i) Which of the following statements are true about distributed systems?

(p) Distributed systems mostly rely on horizontal scaling compared to vertical scaling.
(q) In real implementations message pull is more common than push.
(r) Stateless servers/components are easier to scale.

a. *p* and *q* only
b. *p* and *r* only
c. *q* and *r* only
d. All three

(ii) Which of the following is **not** a scaling technique?

a. Caching
b. Distribution
c. Relocation
d. Replication

(iii) In _____ architecture components communicate through a common repository.

a. Data-centered
b. Event-based
c. Layered
d. Object-based

(iv) Which of the following statement is true about P2P networks?

a. Structured networks are scalable, easier to setup, and suitable for immutable objects
b. Unstructured networks are scalable, easier to setup, and suitable for mutable objects
c. Structured networks are more resilient to failures, harder to setup, and suitable for mutable objects
d. Unstructured networks are more resilient to failures, easier to setup, and suitable for mutable objects

(v) What is an advantage of multi-hop overlay communication?

a. Low latency
b. Low probability of message loss
c. High message propagation overhead
d. High anonymity

(vi) Which of the following is **not** an advantage of gossiping?

a. Easier to setup than deterministic walk
b. Lower overhead compared to flooding
c. Avoids implosion
d. Fast convergence

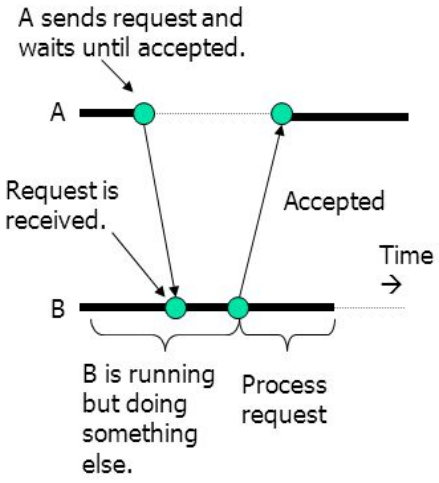
(vii) If a distributed system has 2 nodes each with availability of 99%, overall availability of the system is?

a. $(0.01 \times 0.01) \times 100$
b. $(1 - 0.01 \times 0.01) \times 100$
c. $(0.99 \times 0.99) \times 100$
d. $(2 \times 0.99) \times 100$

(viii) Which of the following is **not** an attribute of a good name?

a. Can be reused
b. Should be unique
c. Need to have 1-to-1 mapping
d. Should be easier to remember

(ix) Following diagram illustrates



a. Delivery-based, transient, synchronous communication

b. Receipt-based, transient, synchronous communication

c. Receipt-based, persistent, asynchronous communication

d. Response-based transient synchronous communication

(x) Which of the following is true about transactions?

a. Lost Update problem can be solved using Totally Ordered Multicast.

b. Inconsistent Retrievals problem can be solved using Serial Equivalence.

c. Dirty Read problem can be solved using Serial Equivalence.

d. Premature writes problem can be solved using Totally Ordered Multicast.

(xi) Fill in the blanks using one of the following keywords: [1 × 5 marks]

`Active, Decouple, Drift, Heartbeat, Integrate, Jitter, Latency, Passive, Skew, Traceroute`

a. In distributed systems link and node failure is detected using _____ .

b. Pub/Sub solutions _____ event producers from event consumers.

c. Due to _____ it is difficult to determine how much to buffer in video streaming.

d. _____ caching is becoming less effective due to the use of HTTPS.

e. Due to clock _____ even nodes that are synchronized goes out of sync with time.

Question 2 **[25 marks]**

(i) Socket Programming and Message Queuing (MQ) systems provide mechanisms to communicate with remote entities. List 2 advantages and 1 disadvantage of MQ systems when compared to Socket Programming. [3 marks]

Figure 1 illustrates an IoT solution deployed to detect movements of elephants in a jungle. A ZigBee-based (i.e., low power and low bandwidth wireless technology with a range of ~100m) collar is attached to an elephant. When an elephant comes near a ZigBee receiver a message is triggered and send to the Cloud-based backend using a 3G connection. Typical event includes (*receiver ID*, *elephant ID*) pair. Location of receivers are known and configured on the backend.

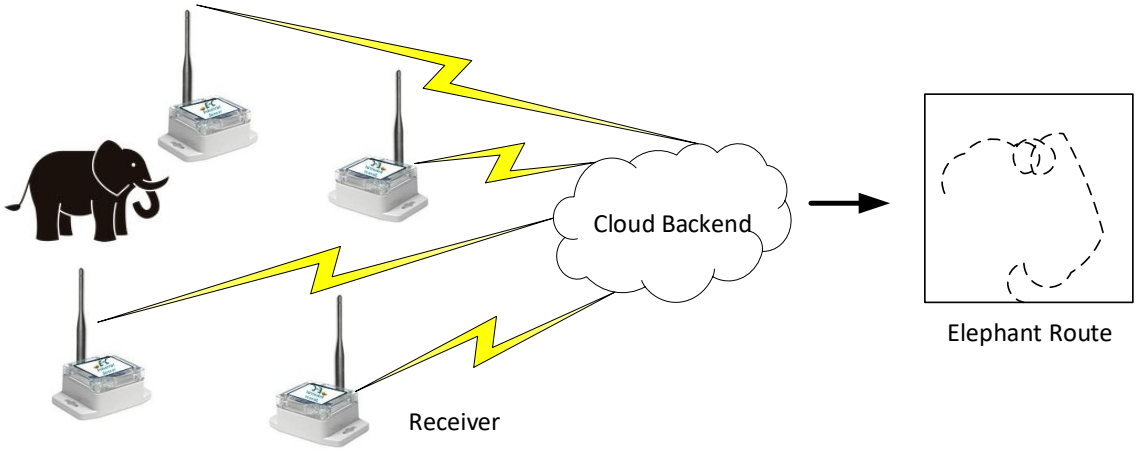


Figure 1 – Overview of elephant movement tracking solution.

(ii) Following 3 options are suggested to order the elephant detection events such that the route taken by an elephant can be estimated.

(a) Receiver's clock is set to physical time at the time of deployment, and is used to timestamp each event.
(b) Events are to be timestamped only at the server based on arrival time.
(c) Each event is to be tagged with an event ID.

Recommend the most suitable solution while considering pros and cons of each of the solution. [7 marks]

(iii) Can Totally Ordered Multicast among receivers be used to order the elephant detection events? Discuss. [3 marks]

(iv) Would you recommend persistent-asynchronous or transient-asynchronous communication for the data transfer between receivers and the Cloud-backend? Briefly Discuss. [3 marks]

(v) Cloud-based server maintains an in-memory table with the list of received events. Is this server *stateful* or *stateless*? Briefly discuss. [3 marks]

(vi) Propose a protocol for the messages exchanged between a receiver and Cloud-based server. Your proposal should contain message formats, order of messages, data types, and lengths. [6 marks]

Question 3 **[25 marks]**

Spotify is a music, podcast, and video streaming service. Spotify is a freemium service; basic features are free with advertisements or limitations, while additional features such as improved quality and music downloads are offered via paid subscriptions. Figure 2 illustrates the Spotify architecture where Spotify services are partitioned based on features, e.g., music, podcast, and video streaming. If one feature fails, other features are independent and will continue to work. When there is a weak dependency between features, failure of one feature may sometimes lead to degradation of service of another feature. Feature partitioning gives scalability, reliability, and an efficient way of focusing on feature improvements. Spotify runs several server instances across the World to achieve geographic scalability.

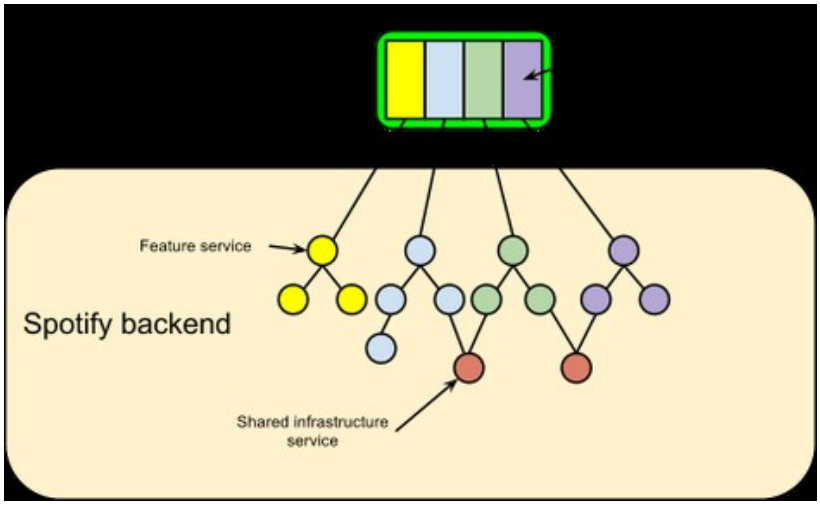


Figure 2 – High-level architecture of Spotify. Source: https://labs.spotify.com

(i) Tick True or False. [1 × 5 marks]

		True	False
(a)	Partitioning is one of the scaling techniques.		
(b)	Degradation of service due to failure of another component is an example of Failure Transparency.		
(c)	When bandwidth between Spotify client and backend drops, one solution is to degrade quality of the stream, e.g., downgrade from MP3 to MP4.		
(d)	Dynamic load balancing among backend feature services is more suitable for Spotify.		
(e)	Pub/Sub solution could be used to notify Spotify users about new songs.		

(ii) What do you recommend, RPC, RMI, REST API, or Web services to communicate between Spotify clients and backend? Justify. [4 marks]

(iii) Briefly explain the CAP theorem in the context of geographically distributed Spotify servers. [6 marks]

(iv) What active CDN technology would you recommend to scale Spotify over having their own servers? Discuss. [4 marks]

(v) Given that Spotify uses geographically distributed set of servers (with or without CDN), how can Spotify ensure transactions related to paid subscribers will satisfy ACID properties? Discuss. [6 marks]

Question 4 **[25 marks]**

With the advent of Social Media and the decreasing cost of data storages, the amount of data created per second increases exponentially. It is claimed that by the year 2020, about 1.7 MB of new data will be created every second for every human being on the planet.

Researchers utilize parallelism to build machine-learning models using large volumes of data. There are 3 different ways to develop parallelism in building machine-learning models. They are Data, Model, and Task Parallelism.

(i) List 2 challenges of using traditional machine-learning algorithms for building models out of large volume of data. [2 marks]

(ii) Define Data Parallelism. You may use a diagram to illustrate. [3 marks]

(iii) Define Model Parallelism. You may use a diagram to illustrate. [3 marks]

(iv) Explain the key challenge in using MapReduce for implementing iterative machine learning algorithms. [2 marks]

(v) Tick True or False. [1 × 5 marks]

		True	False
(a)	Task Queuing is the suitable load balancing technique for balancing requests between web server replicas.		
(b)	Paravirtualization creates dependency on Operating Systems.		
(c)	In distributed file systems, location transparency means if stored copy moved, name do not have to change.		
(d)	NFS protocol is implemented using a stateful servers.		
(e)	Discovery services prevent devices getting disconnected from network unpredictable manner.		

(vi) 'Next word suggester' tool suggests the highly likely next word to be typed given the previous word. Next word suggester is a useful tool for many applications including information retrieval, grammar checkers, and style checkers. These applications benefit by estimating the likelihood of a word following another word. Frequency of a sequence 'A B' is seen in a large corpus is directly correlated to the likelihood of word 'B' occurring after word 'A'.

Given a huge text corpus composed of a large set of files, where each file consists of one sentence per line. Design a solution using MapReduce to estimate the sequence counts of various word pairs in order. Clearly state all your assumptions and show pseudo code of mappers and reducers. [10 marks]

---------- END OF THE PAPER ----------