CST1
COMPUTER SCIENCE TRIPOS Part IB

Tuesday 7 June 2022 11:00 to 14:00 BST

COMPUTER SCIENCE Paper 5

*Answer **five** questions.*

*Submit each question answer in a **separate** PDF. As the file name, use your candidate number, paper and question number (e.g., `1234A-p5-q6.pdf`). Also write your candidate number, paper and question number at the start of each PDF.*

You must follow the official form and conduct instructions for this online examination

1 Computer Networking

Consider a satellite network consisting of a constellation of S satellites placed at an altitude of 1,000km above the Earth (Earth's diameter is ≈12,750km) and G ground stations connecting both customers and Internet exchange points. Each orbit of the constellation consists of 50 satellites (20 orbits total). Each satellite is equipped with five communications laser-receivers; one to communicate with the ground and four to communicate with the nearest neighbour satellites. The satellites are not geostationary. You may assume each satellite only communicates with its nearest neighbours at any time. Each node—satellite, or ground station—has a unique identifier of 16 bytes which it knows.

Design a topology-discovery protocol than can identify the shortest path among any two ground stations using the 1,000 satellites in orbit.

Symmetric paths may be presumed.

(a) Outline a protocol (including message formats) for a node to learn about its immediate neighbours. [3 marks]

(b) Design a protocol (including message formats) for distributing this information across the network. [7 marks]

(c) Give a bound on the total amount of non-redundant information which is transmitted to ensure that every node acquires complete topology information. [5 marks]

(d) The channel bandwidth is given as B, an approximation of 3×10^8 meters per second may be used for the speed of light, and per node packet processing time may be considered zero.

(i) Make an estimate of the worst-case total amount of time the exchange of this information will take to propagate across the network. [2 marks]

(ii) Outline one method to improve the time-period. Speculate on the improved upper bound of total time for the information exchange. [3 marks]

2 Computer Networking

A networking enthusiast says, "TCP does not perform well for the very long, the very short, the very fast, or the very slow."

(*a*) When and why are they correct? Explain your answer clearly. [8 marks]

(*b*) QUIC is a new transport protocol that offers many advantages over TCP, it may not solve the issues the enthusiast of the last question discusses. Prototyping new protocols to replace TCP has proven difficult in the past. Discuss the QUIC approach to this particular challenge. [2 marks]

(*c*) Considering the capacity and limits of end-system components (e.g., CPU(s), cache(s), main memory, peripheral interconnects, network adapters), propose a practical strategy for end-systems to deliver the next generation of Ethernet speeds (100Gbit/s and 400Gbit/s).

Alongside your strategy you may wish to also consider the particular challenges invoked by improved security in networking such as the wider use of IP header encryption (IPsec), the increasing use of full application-payload encryption, and so-forth.

Justify your approach by stating your assumptions throughout.

Hint: your approach need not be universal. It may help to consider aspects of Computer Architecture, Algorithms and Computer Networking in you approach. [10 marks]

3 Computer Networking

(*a*) Using an explanation of the difference between flow-control and congestion-control, discuss the impact of a stable end-to-end latency. [5 marks]

(*b*) Mobile end-points present a challenge for any connection-based application on an IP network. Using an example, discuss how this challenge might be resolved without a complete reworking of the end-system application. [5 marks]

(*c*) A reverse proxy such as Varnish makes static snapshots of the content of a dynamic webpage. Outline a use-case for using such a reverse proxy, paying particular attention to a case when an otherwise useful system might need disabling. [5 marks]

(*d*) In 2021 a large social network had to reboot all its systems after systemic failure. Discussing the extensive use of image caching, outline the challenges that face an image serving system in the social media space (that is, one that uses popularity, hashtags, and so-forth). Discuss in particular why, for weeks after the reboot, many users complained of poor performance, and of being served irrelevant or old images.

Hint: A distinguishing feature of social networks has been a heavy-tailed distribution of interest among images and users.

For your information: a large image sharing site has 500×10^6 active daily users, 500×10^6 active posts (groups of images and interactions) and 100×10^6 new images each day. Sizes varies considerably but an average image size approaches 100 Kbytes. [5 marks]

4 Concurrent and Distributed Systems

(*a*) A shared resource is updated by one writer at a time but must support concurrent readers. One or more locks are used to manage exclusion.

(*i*) What design choices exist regarding fairness? [2 marks]

(*ii*) Compare having two simple locks, respectively used for read and write exclusion, with having a single, specialised locking primitive. What additional state transition(s) could the specialised lock support? [4 marks]

(*iii*) If deadlock is to be avoided using a locking order, is a complete or partial order needed? [2 marks]

(*iv*) How should the various transitions possible with the specialised lock mentioned above be integrated into a lock ordering policy? [3 marks]

(*b*) A mutex is acquired at the start of the body of a function.

(*i*) What can go wrong if the function is recursive? [1 mark]

(*ii*) To avoid the problem, a friend suggests that threading systems should allow lock acquisitions to proceed if the lock is already held by the same piece of code. Could something like this work and is it worthwhile? [3 marks]

(*c*) The behaviours provided by a shared, global variable are to be emulated within a message-passing system.

(*i*) Give pseudocode that provides the emulation. [2 marks]

(*ii*) Discuss whether emulation of locking primitives, such as semaphores, is possible or would be worthwhile? [3 marks]

5 Concurrent and Distributed Systems

You are designing *Warbler*, a social networking service in which users can post public messages called *chirps*. A chirp may or may not be a reply to another chirp. Moreover, one user can *follow* another user, which means subscribing to the messages they post.

(*a*) Write pseudocode for three operations: (1) user u_1 following user u_2; (2) u_1 unfollowing (ceasing to follow) u_2; and (3) obtaining the current set of followers of some user u. These operations must be fault-tolerant: you should store the data on five servers, and the operations must be able to complete successfully as long as any three or more servers are available. Your algorithm should ensure read-after-write consistency.

Assume an asynchronous system model with fair-loss network links and crash-recovery faults. Write your algorithm in terms of messages sent over point-to-point (unicast) links. You may assume that the user is already authenticated (e.g. their password has been checked), and you may assume that users' client software can generate Lamport timestamps.

Start by describing the nodes in the system, the structure of the data stored at each node, and the form and purpose of messages exchanged by nodes. Then give the pseudocode for the operations listed above. [12 marks]

(*b*) When a user posts a chirp, it needs to be sent to all of that user's followers. Write pseudocode to do this, using the operation for getting a user's followers from part (*a*). Chirps must be delivered to a given recipient in the following order: chirps posted by the same user should be delivered in the order they were posted; and when one chirp is a reply to another, the reply should be delivered after the chirp it is replying to. Apart from these rules, the chirps may be delivered in any order. [8 marks]

In both parts, in addition to the pseudocode, please also briefly explain and justify the design of your algorithm, and specify any assumptions you are making.

6 Introduction to Computer Architecture

The Thrupenny Bit FPGA company produces FPGAs with a sea of logic elements (LEs) depicted below (Fig 1) containing 3-bit LUTs (LookUp Tables) and a D flip-flop (DFF). Each LUT can be programmed with any Boolean function of three inputs and one output. The output of the LUT can be sent over the programmable wiring or to the input of the DFF. The DFF has data (D), asynchronous level-sensitive clear and edge triggered clock inputs, and the Q output. The clear input can be used to reset the DFF to zero.

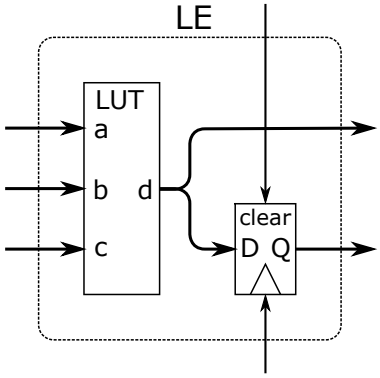

Fig. 1: Thrupenny Bit LE

```
module counter(input  logic       clk,
               input  logic       rst,
               input  logic       enable,
               output logic [2:0] count);

   always_ff @(posedge clk or posedge rst)
     if(rst)
       count <= 0;
     else
       if(enable)
         count <= count+3'd1;
endmodule
```

Fig. 2: SystemVerilog code

(a) What is the minimum number of LEs required to implement the `counter` module in Fig. 2 on a Thruppenny Bit FPGA? Provide a detailed rationale. [8 marks]

(b) Produce a circuit diagram of your implementation showing how the LEs are wired together. [6 marks]

(c) Tabulate the contents of the LUTs for each LE used (i.e. provide a state transition table for each LUT).

[6 marks]

7 Introduction to Computer Architecture

(*a*) The following code is written in C, where elements within the same row are stored contiguously. Assume each element of the two dimensional arrays A and B is a 64-bit integer. Assume that each cache line is 16-bytes.

```
for(i=0; i<1024; i++)
  for(j=0; j<1024; j++)
    A[i][j] = B[i][0] + A[j][i];
```

(*i*) Which variable references exhibit temporal locality? Explain your answer including a definition of temporal locality. [4 marks]

(*ii*) Which variable references exhibit spatial locality? Explain your answer including a definition of spatial locality. [4 marks]

(*b*) Imagine we have a tiny 256-byte direct-mapped data cache with 16-byte cache lines. The cache is initially empty. Below is a sequence of 32-bit memory load accesses. For each load identify the tag, index, offset and hit/miss status.

`0x003, 0x0b4, 0x001, 0x102, 0x001, 0x2c2, 0x004, 0x2c0`

[4 marks]

(*c*) Data hazards can impact the performance of a pipelined processor. What is a data hazard and how can load and arithmetic instructions be reordered to minimise data hazards arising? [4 marks]

(*d*) For a write-back L1 cache that is full of dirty cache lines, how is a write-miss handled? [4 marks]

8 Introduction to Computer Architecture

(*a*) Why are modern systems-on-chip (SoC) heterogeneous (i.e. contain a range of different processor cores)? [4 marks]

(*b*) What is the *von Neumann bottleneck* and to what extent do modern SoCs suffer from it? [4 marks]

(*c*) How is virtual memory used to provide isolation between applications? [4 marks]

(*d*) How do GPUs hide memory access latency? [4 marks]

(*e*) Why is conditional program flow control managed differently on GPUs and CPUs? [4 marks]

END OF PAPER