CST1
COMPUTER SCIENCE TRIPOS Part IB

Tuesday 8 June 2021 11:30 to 14:30 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 Design

(a) Moore's law and Dennard scaling both predict scaling properties of CMOS chips. What are the differences between these predictions and which predictions are valid today? [4 marks]

(b) How is the critical path in a clocked digital CMOS circuit determined and how does it impact the maximum clock frequency? [4 marks]

(c) What is a function calling convention and how does it impact the design of the RISC-V instruction set architecture (ISA)? [4 marks]

(d) Consider the following C function that computes the greatest common divisor, and the assembler produced by the compiler. The assembler has been split into segments. Describe what function each segment performs. [8 marks]

```
int gcd(int n1, int n2) {
  if (n2 == 0)
    return n1;
  else
    return gcd(n2, n1 % n2);
}
```

```
##### Segment A #####
gcd:
        bne     a1,zero,.L7
        jr      ra
##### Segment B #####
.L7:
        addi    sp,sp,-16
        sw      ra,12(sp)
##### Segment C #####
        mv      a5,a1
        rem     a1,a0,a1
        mv      a0,a5
        jalr    ra, gcd
##### Segment D #####
        lw      ra,12(sp)
        addi    sp,sp,16
        jr      ra
```

2 Computer Design

(*a*) How are forwarding paths used in processor pipelines? What pipeline hazards can they mitigate and what are the limitations of the mitigations? Provide one or more examples to illustrate the challenges. [8 marks]

(*b*) For a pipelined processor implementing the RISC-V instruction set, why might branches cause bubbles in the pipeline, why are they needed and what can be done to mitigate them? Provide one or more examples. [6 marks]

(*c*) In detail, describe the hardware mechanisms that are required to support virtual memory for a pipelined processor implementing the 32-bit RISC-V ISA. [6 marks]

3 Computer Design

A processor contains two cores, each with an L1 cache connected via a shared bus to an L2 cache, which is then connected to main memory. Each L1 is a direct-mapped, 4 KiB, write-back cache. The L2 is a 4-way set-associative 16 KiB cache with the least-recently-used replacement policy. All cache lines are 16 B long. The hierarchy is inclusive, runs the MSI cache coherence protocol and is initially empty.

(*a*) Considering this cache hierarchy,

(*i*) Explain whether it would be suitable for a system-on-chip that is only running single-threaded applications. [2 marks]

(*ii*) Explain whether it would be suitable for a system-on-chip where each core processes a small (e.g. $\leqslant$ 1 KiB) array of data at a time. [2 marks]

(*b*) Show the cache contents and coherence state of cache lines after each access in the following sequence of physical addresses, stating any assumptions you have made. All accesses are 4 bytes long and entirely complete before the next one starts.

```
Core 1:  Read   0xab18
Core 2:  Read   0xdb14
Core 1:  Write  0x1b10
Core 1:  Read   0xab14
Core 1:  Read   0xbb1c
Core 2:  Write  0xa010
Core 1:  Read   0x2b10
Core 2:  Read   0x1b10
Core 1:  Read   0xa018
Core 1:  Write  0x1b14
```

[8 marks]

(*c*) Describe the impact of each change below (in isolation) on the cache hierarchy.

(*i*) Increasing the L1 cache size.

(*ii*) Increasing the line size in all caches.

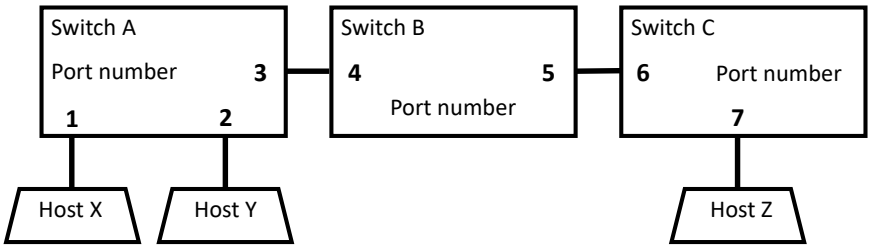
(*iii*) Increasing the associativity of the L1 caches.

(*iv*) Changing to an exclusive cache hierarchy.

[2 marks each]

4 Computer Networking

(a) Consider the layer-2 switch topology shown below.



Assuming all switches start with empty switch-forwarding tables; Host X (with physical address X) sends a packet destined for Host Z. Enumerate in the style below, all packets sent across the network until the message arrives at Host Z. You may assume packet-processing, latency and transmission time are negligible. Additionally, indicate packets transmitted simultaneously. [5 marks]

Time	Sent		Frame	
Step	by Device	on Link	Source	Destination
0	X	X-A1	X	Z
⋮	⋮	⋮	⋮	⋮
?	C	C7-Z	X	Z

(b) Enumerate in the style below, the forwarding table of Switch B at the end of Part (a).

Destination	Port
⋮	⋮

[2 marks]

(c) Consider a layer-2 network consisting of S+1 switches, S directly attached to H hosts. Each host runs V virtual machines, each with a single address. S switches are connected using a star topology with a single switch C at the centre. Each host exchanges data with a selected and stable subset of other hosts.

Estimate the worst case number of entries in the forwarding table for any typical switch S and the number of entries in the forwarding table for switch C.
[3 marks]

(d) Users of this cluster of machines complain of occasional misbehaviour attributed to network timeouts and network slowdown.

Outline two plausible chains of events causing the problem. Describe two appropriate, cost-effective, strategies for overcoming the issues faced by the users. You may assume the switches are state-of-the-art and buying more hosts is not the answer. [10 marks]

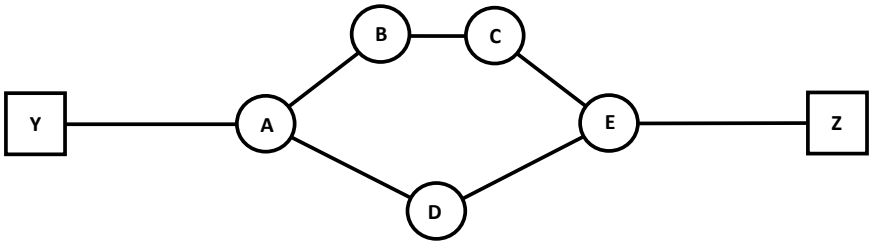
5 Computer Networking

(*a*) A user complains that their web application `Times Out` after successfully connecting to a remote host, yet, an investigation with `ping` indicates the remote host is alive.

Propose a cause of this fault and outline a test-strategy that could be used automatically to detect such a fault. [4 marks]

(*b*) A user would like to infer the path between two hosts; they suggest using the utility `ping`. Make your argument for why `traceroute` is the more appropriate utility, comparing and contrasting their operation. [4 marks]

(*c*) Consider the test network below; Y and Z are two unix IPv4 hosts, while nodes A through E represent fully-conformant IPv4 routers providing the only connectivity between the two hosts.



Per-packet load balancing by A and E means packets may be sent on any valid path. No firewalls or packet filtering is used anywhere in the network.

(*i*) Using `traceroute` infers the following paths:
(Y-A-B-C-E-Z), (Y-A-D-E-Z), (Y-A-B-D-E-Z), and (Y-A-B-E-E-Z)

Explain (with diagrams as appropriate) why these results have occurred and identify the two true and two false results; [8 marks]

(*ii*) A firmware upgrade for routers A and E means they now do per-flow and not per-packet load balancing. Despite the functions of the flow-based load balancing, `traceroute` results have not changed; two false paths are inferred along with two true paths.

Explain this outcome and a strategy to identify only true paths.
[4 marks]

6 Computer Networking

(*a*) For IPv4 ISPs, each domestic installation typically gets a /32 network. You have a complicated configuration requiring NAT and multiple IPv4 subnets.

(*i*) Why would an IPv6 based provider allocate four /64 networks for your premises when each /64 represents 2^{64} addresses? [2 marks]

(*ii*) A colleague has IPv6 with another provider; they only allow one /64 for each domestic installation. In the past your colleague has used a NAT and many IPv4 private address blocks, but keenly adopted IPv6 permitting them to upgrade their home network. They are now using blocks of the allocated /64 and a router in their home to interconnect the subnets.

Not everything is working as they hoped; for example, sometimes IoT devices can't connect to the Internet to update and your colleague can not connect to their front-door camera when at work.

Explain what sort of problems your colleague may face along with methods by which they could verify the root cause. [6 marks]

(*iii*) Explain to your colleague why you might not be able to *lend them* one of your /64 allocations, even though the /64 blocks (provided to you by your ISP) are each globally routable addresses. [2 marks]

(*b*) A local area network may carry several different LANs simultaneously; such a network would be designated for known sets of HomePlug devices.

Describe a physical line coding approach for the HomePlug devices that: allows two or more simultaneous virtual local area networks to fairly share the same physical channel, but does not permit trivial interception of network traffic.

Outline your approach along with its benefits and drawbacks, comparing it with the simplest use of VLAN tags in Ethernet. [10 marks]

7 Concurrent and Distributed Systems

(*a*) Explain what happens to the state space, the possible behaviours and the reachable state space when two automata are coupled. [3 marks]

(*b*) The Banker's Algorithm can be viewed as a predicate over shared state. What state does it operate over and does this include the program counter of the participating threads? When does it return true or false? [6 marks]

(*c*) What is the difference between strict and non-strict isolation in a transaction processing system? What do both approaches ensure? Which can lead to a transaction abort being forced by the system and why? [3 marks]

(*d*) "Increment and decrement operations are freely commutable" — what two assumptions are required for this statement to hold? Is it true that the effects of transactions containing increment and decrement operations are always serialisable? [3 marks]

(*e*) Customers interact with a transaction processing system over a web interface but confirmations are also sent by email, such as 'please collect from your local branch'.

Should the email message be generated before, during, or after the process of committing the order transaction? What are the advantages and disadvantages of different approaches? Fully justify at least two design decisions in terms of system complexity and durability over a system crash.

[5 marks]

8 Concurrent and Distributed Systems

You are developing a distributed system in which you want some task to be assigned to exactly one node at any given time. If that node crashes, the task must be automatically reassigned to a different node, but you also want to ensure that there are never two or more nodes executing the task at the same time.

This is known as a *lease*. A lease is a concurrency primitive similar to a lock (only one node may hold a lease at any time); the difference is that a lease times out if it is not renewed for some time. After timing out, another node can acquire the lease.

(*a*) Briefly summarise how leader election works in the Raft consensus algorithm, and discuss the commonalities and differences between leader election and a lease. (Focus only on leader election, and ignore the rest of the Raft algorithm. Include the role of the term number in your explanation.) [5 marks]

(*b*) In a partially-synchronous system with crash-recovery failures, is it possible to guarantee that a lease is always held by exactly one node? Justify your answer. [5 marks]

(*c*) You are asked to design a lease algorithm for a system in which the set of nodes is not known in advance, and may change over time. Can the Raft leader election algorithm be used here? Why/why not? [2 marks]

(*d*) A colleague proposes the following lease algorithm:

- There are three servers, each storing a value that is initially null.
- Assume every client has a unique ID $clientId \neq \mathsf{null}$. Every 10 seconds, each client that wants to acquire the lease, or currently holds the lease, sends a request $(\mathsf{acquire}, clientId)$ to all three servers.
- When a server with current stored value v receives $(\mathsf{acquire}, n)$:
  If $v = \mathsf{null} \ \vee \ v = n$, or if its value was last set more than 30 seconds ago, then it sets its value to n and replies true. This counts as "setting the value", even if the value does not change.
  If its value was last set to $v \neq n$ less than 30 seconds ago, it leaves its value unchanged and replies false.
- If a client receives two or more true responses from the servers, it now holds the lease, otherwise it does not hold the lease.

Discuss the strengths and weaknesses of this algorithm. What faults does it tolerate? What assumptions does it make for its correctness? How might the algorithm be improved to avoid some assumptions or weaknesses? [8 marks]

END OF PAPER