February 7th, 2018

Exam of Switching technologies for data centers (2017/18)

Rules for the exam. It is **forbidden** to use notes, books or calculators. Use only draft paper provided by the professor. When needed, use approximations. The answers must be provided in correct English. Any notation must be defined.
Time available: 70 minutes.

Problem A

Consider an $N \times M$ input queued switch supporting multicast traffic, with optimal queueing structure MC-VOQ for multicast.

1. Define the MC-VOQ queueing and compute the total number of queues.
2. Write the pseudocode of a maximal scheduler allowing fanout-splitting (i.e. a multicast packet can be sent to a subset of destinations). Comment the code. Use the data structures reported below. Any data structure must be properly defined and eventually initialized.
3. Comment about the fairness of the above code.
4. In which application scenarios multicast traffic is relevant for high speed networking?

As notation, assume that `Q[i][k]` is the occupancy of queue `k` at input `i` and `FS(k)` returns the fanout set (i.e. the set of destination outputs) of the packets stored in queue `k`. The crossbar configuration is described by two vectors: let `X` be a vector of size M defined as follows: `X[j]=i` whenever input `i` sends a copy of the packet to output `j`; let `Y` be a vector of size N describing the queues to serve, defined as `Y[i]=k` whenever queue `k` is served at input `i`.

Problem B

Consider the design of Jupiter data center at Google, based on a basic building block implemented with a chipset with 16 ports at 40 Gbps. The adopted oversubscription ratio is 3:1. Each server is equipped with a single port running at 10 Gbps. Draw the architecture and compute the total number of servers and basic building blocks (i.e., chipsets) for each of the following scenarios:

1. 2-layers topology;
2. 2-layers POD;
3. 3-layers topology;
4. 3-layers POD;
5. 4-layers topology.

Finally, describe for the 4-layers topology all the possible ways to interconnect the data center to the Internet.

Problem C

Consider the adoption of the fingerprinting scheme in data storage system.

1. Describe how fingerprinting works.
2. Describe the problem of false positives.
3. Describe how deletion is supported.
4. Compute analytically the minimum number of bit per fingerprint to achieve a probability of false positive equal to ϵ. Define *every* notation adopted in the proof and explain all the steps.
5. Assume to store 1000 elements in a bit array with fingerprinting. What is the size of the storage to achieve a probability of false positive equal to 10^{-6}?

Hints for the solution

Problem A

1. In the MC-VOQ structure, each input has one dedicated queue for each possible fanout set, thus a total of $N \times (2^M - 1)$ queues are present in the switch.

```
// initialize data structures
for j=1...M // for each output port
    X[j]=-1 // output j non connected
for i=1...N // for each input port
    Y[i]=-1 // the queue has not be chosen so far
// take scheduling decision
for i=1...N     // for each input port
     for k=1...(2^M-1) // for each queue
          if (Q[i][k]>0) // check if the queue is not empty
                for each j in FS(k) // for each output in the fanout set
                       if (X[j]==-1) // check if the output is still available
                          // matching the input with the output
                          Y[i]=k // store the queue to serve
                          X[j]=i // match input i with output j
          if (Y[i]>0) // check if an input has been found
              break // consider a new input
```

2. The above code is unfair since the queues belonging to the first inputs and the ones identified by smaller values of `k` will be served at higher priority than the other queues.
3. Example of applications for which multicast is important is: broadcast in VLAN scenarios or virtualized networks, duplication of data in data center, financial transactions (high-frequency trading).

Problem B

This is a verbatim copy of the class notes.

Problem C

This is a verbatim copy of the class notes.
As a result of the proof, the total number of bits per fingerprint is $b = \log_2(m/\epsilon)$ where $m = 10^3$ and $\epsilon = 10^{-6}$, thus $b = \log_2(10^9) = 3\log_2(1000) \approx 30$ bits per fingerprint. Thus, the total memory size for the bit array is 10^9 bits, equivalent to 125 MB.