February 22nd, 2016

Exam of Packet switch architectures (2015/16)

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.
Time available: 70 minutes.

Problem A

Consider a $N \times M$ input queued switch with Virtual Output Queueing and supporting Strict Priority Queueing based on C classes. For each (input-i,output-j) pair there exist C queues: $[VOQ_{ij}^c]_{c=1}^C$, with decreasing priority level, where $c = 1$ refers to the highest priority traffic and $c = C$ refers to the lowest priority traffic (i.e., best effort traffic). Therefore, $C \times N \times M$ queues are present in total. We assume that an input traffic classifier sends the incoming packets to the correct queue.

1. Write in pseudocode a greedy algorithm to schedule the transmissions across the switching fabric, to be maximal and to maximize the number of higher priority packets that are selected at each timeslot. Define all the data structures adopted in your code.

2. Does the scheduler obtain always the maximum throughput under admissible Bernoulli i.i.d. traffic? Motivate your answer.

3. Is it possible that low priority traffic will be starved indefinitely by high priority traffic? Motivate your answer with an example.

Problem B

Design a 8×8 self-routing multistage network, adopting basic modules of size 2×2.

1. Draw the whole network with all the interconnections and modules.
2. Identify all the modules, input ports, output ports and all the links in order to support self-routing.
3. How does the self-routing work?
4. What are the properties of the network in terms of blocking?
5. Using the Lee method, find the blocking probability in function of the overall traffic λ (measured in Erlang) feeding the whole network.
6. Is the blocking probability obtained with Lee compatible with the properties of the networks highlighted in question 4?

Problem C

Consider the design of a large data center based only on Ethernet switches with P ports.

1. Describe the advantages and disadvantages to adopt a layer-2 addressing and routing scheme.
2. Describe the advantages and disadvantages to adopt a layer-3 addressing and routing scheme.
3. For each of the following design problems, draw at the high-level the network, compute the required number of P-port switches and the maximum number of supported servers,
   (a) Design the largest possible two level (leaf-and-spine) data center.
   (b) Design the largest possible two level (leaf-and-spine) POD.
   (c) Design the largest possible three level (POD-and-spine) data center.
4. Consider a data center with 400 servers each of them equipped with one port at 1 Gbps. Each rack hosts 40 servers. All the adopted switches have $P = 80$ ports running at 1 Gbps each.
   (a) Draw the network.
   (b) Compute the number of required 80-port switches.

Hints for the solution

Problem A

The problem is a variant of Ex.82.

Problem B

The problem is a variant to Ex.32.
5. The Lee graph is just a sequence of two edges in series. The probability of busy for each edge is $\lambda/8$, thus the blocking probability is $1-(1-\lambda/8)^2$.

Problem C

3.(a) The design uses P switches in the leaf level and $P/2$ switches in the spine level. Thus, the number of servers is $P^2/2$ and the number of switches is $1.5P$.
3.(b) The design uses $P/2$ switches in the leaf level of the POD and $P/2$ switches in the spine level of the POD. Thus, the total number of server ports is $P^4/4$, the number of ports to the spine is $P^2/4$, and the number of switches is P.
3.(c) The design uses P PODs in the leaf level and $P^2/4$ switches in the spine. Thus, the total number of servers is $P^3/4$ and the number of switches is $(5/4) \times P^2$.
4.(a) The leaf level is based on 10 switches whereas the spine level is based on 5 switches. The interconnection is full between leaf and spine level, with a parallelism of 8 Gbps.
4.(b) The total number of switches is 15.