Please write clearly. Make your answers concise, but complete.

1. (10 points) The figure below shows a binary trie implementation of an IP route lookup data structure. Draw an equivalent multibit trie with stride of 2.
2. (10 points) The state transition diagram shown below is for two port switch element with a shared buffer capable of holding up to two cells. A state label \((i,j)\) signifies that in the given state there are \(i\) cells in the buffer for output 0 and \(j\) cells for output 1. Label all the bold arrows with the correct transition probability, assuming that the probability of a cell arrival at either input is \(p\) and that arriving cells are equally likely to be addressed to either of the outputs. Assume that there is no inter-stage flow control, so there is no constraint on arriving cells and also no constraint on departing cells.
3. (15 points) Consider a variant of a bus-based cell switch with a sub-divided bus and \( n \) inputs and outputs. In this version, there is one knockout concentrator for every group of four output links. The \( k \) outputs of this concentrator go to a shared memory used by all four outputs. The four outputs take turns reading from this memory in order to get the cells they need to send on their links. Give an expression for the total memory bandwidth required by each port, assuming that the bandwidth of the external links is \( R \) bits per second.

\[(k+4)R \text{ b/s}.\]

Give an expression for the width of the memory bus, assuming that the clock rate to the memory is \( M \) MHz (meaning that each wire on the memory bus can transfer one bit every \( 1/M \) microseconds).

\[W = (k+4)R \times 10^6 / M, \text{ so } W = (k+4)R / 10^6 M.\]

Give an expression for the probability that during one operational cycle of the switch, we receive more cells at a given knockout concentrator than we can transfer into the memory. Assume that the inputs system are fully loaded and that the traffic is uniformly distributed across all the outputs.

\[
\sum_{i=k+1}^{n} \binom{n}{i} (4/n)^i (1-4/n)^{n-i}
\]

How do you think the total memory bandwidth for this system compares to the original system, in which each output has its own knockout concentrator and its own memory? Assume that in both systems, the number of outputs of the knockout concentrator is chosen to give the same low loss probability. You needn’t quantify the difference in memory bandwidth. Just state which of the two would require more memory bandwidth and explain your reasoning.

The system with the shared concentrator will have smaller memory bandwidth. In order for the total memory bandwidths to be the same, its knockout concentrator would have to have 4 times as many outputs as are required in the other case. But the probability that all four of them are busy enough to lose cells in this case, is much less likely than that one output is busy enough to lose cells in the original system. The outputs in the group are very unlikely to be busy at the same time.
4. (15 points) Show that $B_{d^k,d} \approx D_{d^{k-1},d} \otimes X_{d,d} \otimes D_{d^{k-1},d}$ using the associativity properties of the serial and parallel construction operators and induction on $k$.

The basis for the induction is $k=2$. In this case

$$B_{d^2,d} = X_{d,d} \otimes X_{d,d} \otimes X_{d,d} = D_{d,d} \otimes X_{d,d} \otimes D_{d,d}$$

by definition of $B$ and $D$.

The induction step, for $k>2$ is

$$B_{d^k,d} = X_{d,d} \otimes B_{d^{k-1},d} \otimes X_{d,d}$$

$$= X_{d,d} \otimes (D_{d^{k-2},d} \otimes X_{d,d} \otimes D_{d^{k-2},d}) \otimes X_{d,d}$$

$$= (X_{d,d} \times D_{d^{k-2},d}) \otimes X_{d,d} \otimes (D_{d^{k-2},d} \times X_{d,d})$$

$$= D_{d^{k-1},d} \otimes X_{d,d} \otimes D_{d^{k-1},d}$$

The first step above is just the definition of $B$. The second line uses the induction hypothesis. The third line uses the associativity of the parallel operation. The last line uses the definition of $D$ and the general property of the delta network that says that any delta network is strongly isomorphic to the series combination of two smaller delta networks.
5. (15 points) Consider a buffered, multistage switch in which each switch element (SE) has two inputs and outputs, and each input has its own FIFO buffer that can store up to \( k \) cells. At the start of a cell cycle, an SE gives a grant to each of its upstream neighbors if the corresponding input buffer is not full. It then attempts to forward the first cell in each of its input buffers, assuming that the outputs needed by those cells have sent a grant. Let \( \pi(i,j) \) be the steady-state probability that a buffer in a stage \( i \) SE contains exactly \( j \) cells. Give an expression for the probability that a downstream neighbor of a stage \( i \) SE sends it a grant (this expression will be a function of \( \pi(i+1,j) \)).

\[
1 - \pi_{i+1}(k)
\]

Give an expression for the probability that an upstream neighbor of a stage \( i \) SE has a cell at the front of at least one in one of its input buffers that is addressed to the given stage \( i \) SE (this expression should be a function of \( \pi(i-1,j) \)).

\[
1 - (1 - (1/2)(1 - \pi_{i-1}(0)))^2
\]

Give an expression for the probability that a cell at the front of a buffer in a stage \( i \) SE has a competing cell at the front of the other buffer in that SE. That is, the other buffer has a cell and it is addressed to the same output.

\[
(1/2)(1 - \pi_i(0))
\]

Let \( g_i, a_i \) and \( x_i \) be the probabilities in the previous three parts of the problem. Give a balance equation for \( \pi(i,j) \) in terms of \( a_i, g_i, x_i, \pi(i-1,j) \) and \( \pi(i,j+1) \). You may assume that \( 0 < j < k \).

\[
\pi_i(j) = \pi_i(j-1)(a_i((1-g_i)+(g_i,x_i/2)) + \pi_i(j+1)((1-a_i)g_i,(1-x_i/2))
+ \pi_i(j)(a_i,g_i,(1-x_i/2) + (1-a_i)((1-g_i)+(g_i,x_i/2)))
\]

\[
\pi_i(j) = \frac{\pi_i(j-1)(a_i((1-g_i)+(g_i,x_i/2)) + \pi_i(j+1)((1-a_i)g_i,(1-x_i/2))}{1-(a_i,g_i,(1-x_i/2) + (1-a_i)((1-g_i)+(g_i,x_i/2)))}
\]
6. (15 points) Consider a system with 600 Mb/s links and a speedup of 1.5 that implements the basic distributed queueing algorithm described in class. Suppose we perform a stress test on this system with 2 inputs and 5 phases and that the first phase lasts for 10 ms. How big will the backlogs be in each of the VOQs at the end of the first phase?

Each input sends to the first output at 450 Mb/s, while receiving at 600 Mb/s. So, it builds a backlog at rate 150 Mb/s, meaning that it will have 1.5 Mbits in its queue after 10 ms.

How big will the backlogs be in each of the output side queues at the end of the first phase?

The output is receiving at 900 Mb/s and sending at 600 Mb/s, so it will have 300 Mbits in its queue after 10 ms.

When will the second phase end, assuming that the second phase ends when the input side backlog from input 1 to output 1 is equal to the input side backlog from input 1 to output 2?

During the second phase each input sends at 450 Mb/s to both of the outputs. This means that it clears its backlog going to the first output while building a backlog for the second at 150 Mb/s. If \( t \) is the amount of time from the start of the second phase, when the input side backlogs at input 1 are equal, then

\[
150 \cdot 10^6 t = 1.5 \cdot 10^6 - 450 \cdot 10^6 t
\]

\[
t = 1.5 / 600
\]

seconds or 2.5 ms. So, the second phase ends at 12.5 ms.
7. (15 points) The diagram below shows the consecutive copy+route multicast architecture. In this architecture, the copy network delivers copies of an original multicast cell to consecutive outputs (wrapping around from the last output to the first, if necessary). Show how the multicast routing table entries should be initialized to support a multicast session from input 3 to outputs 0, 4 and 5. Assume there is just one fanout class.

![Diagram of consecutive copy+route multicast architecture]

If this multicast session has a bandwidth of 100 Mb/s, how much bandwidth does it use on each of the copy network outputs?

100 Mb/s = 100 \cdot \frac{3}{8} = 37.5 \text{ b/s}

Suppose we simplify the architecture by eliminating the wrap-around feature. In this case, how much bandwidth does the example multicast session place on each of the copy network outputs?

Outputs 0 and 7 get \( \frac{100}{6} \approx 16.7 \text{ M b/s} \), outputs 1 and 6 get \( \frac{100}{3} \approx 33.3 \text{ M b/s} \) and outputs 2, 3, 4 and 5 get \( \frac{100}{2} = 50 \text{ M b/s} \).

In general, what is the minimum load that a multicast session with fanout \( F \) and input bandwidth \( B \) places on any copy network output in a system with \( n \) ports? What is the maximum?

The minimum is \( \frac{B}{n - F - 1} \) and the maximum is \( \frac{FB}{n - (F - 1)} \), so the load on some outputs is \( F \times \) the load on others. Also, the most heavily loaded outputs can have a total load which is \( 1 + (F - 1)/(n - (F - 1)) \) times larger than the load they can have, if the wrap-around feature is included. This is a strong argument in favor of including the wrap-around feature, when \( F \) can be a large fraction of \( n \). On the other hand, if \( F \) is limited to no more than say \( n/10 \), it may be reasonable to omit the wrap-around feature, since the effect on the required bandwidth is relatively small in that case.
8. (15 points) Consider a switching system that uses a static routing network with topology $C_{312,32,16}^3$. If the external link rate is 1 Gb/s and the fastest individual session has a bandwidth of 50 Mb/s, what bandwidth is needed on the switch’s internal data paths to make the system strictly nonblocking for unicast traffic?

The inequality needed for this is $r > 2 \left[ \frac{d - B / \beta}{1 / \beta - B / \beta} - 2 \right]$. Substituting for the given parameters, we have $16 > 2 \left[ \frac{32 - 0.05}{1 / \beta - 0.05} - 2 \right]$. The smallest value of $\beta$ that makes this true satisfies $8 = \frac{32 - 0.05}{1 / \beta - 0.05}$, so $\beta = 4.04$, so the required bandwidth on the internal datapaths is 4.04 Gb/s.

Consider a switching system that uses a static routing network with topology $C_{312,32,r}^3$. If the external link rate is 1 Gb/s, the internal link rate is 8 Gb/s and the fastest individual session has a bandwidth of 50 Mb/s, what is the smallest value of $r$ that makes the system reroutably nonblocking for multicast traffic, if the maximum fanout in the first stage switches is limited to 4?

The inequality needed is

$$r > \left[ \frac{d - B / \beta}{1 / \beta - B / \beta} - 1 \right] + \left[ \frac{d - B / \beta}{1 / \beta - B / \beta} - 1 \right] (n / d)^{1/f}.$$ 

Substituting for the given parameters, we have

$$r > \left[ \frac{4 \times 32 - 0.05}{8 - 0.05} - 1 \right] + \left[ \frac{32 - 0.05}{8 - 0.05} - 1 \right] (16)^{1/4} = 16 + 4 \cdot 2 = 24$$

So, we need $r = 25$. 

8
9. (15 points) Give an expression the blocking probability for the following networks using Lee’s method. Assume each external link and internal link can carry just one virtual circuit at a time. Let $p$ denote the probability that each network input is busy.

\[
C^2_{128,16,24} \left[ 1 - (1 - 2p/3)^3 \right]^{24}
\]

\[
X_{2,4} \otimes (X_{2,4} \times X_{2,4}) \otimes X_{4,8} \left[ 1 - (1 - p/2)(1 - p/4)(1 - p/8) \right]^4
\]

\[
B_{64,4} \left\{ 1 - (1 - p)^2 \left[ 1 - (1 - (1 - p)^2)^4 \right] \right\}^4
\]