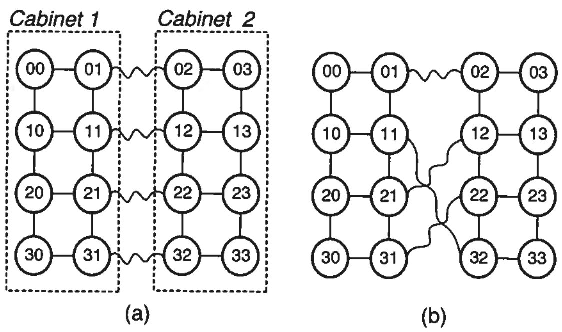

Figure 3.14 The packaging of a 16-node network across 2 cabinets (a) using coaxial cables to create a mesh topology, and (b) with a random permutation of the cables between cabinets.

Coaxial cables (wavy lines in the graph) connect corresponding nodes between the two cabinets.

(a) What is the diameter of this mesh network?

(b) How does the diameter of this network change if the cable connections are randomly permuted? (Figure 3.14[b] shows one such permutation.) What are the minimum and maximum diameters of the network over all permutations? Give a permutation that realizes the minimum diameter.

3.10 *Performance of a fat tree network.* Figure 3.15 shows a 16-node, radix-2 fat tree topology.

(a) Assuming all channel bandwidths equal the injection and ejection rates of the terminal nodes, what is the capacity of this network?

(b) Consider a randomized approach for routing on this network: for each packet, route first from the source "up" to the top of the tree (these are the 8 small, center nodes in the figure), along the way randomly choosing between the two possible "up" channels, and then finish by routing along the "down" channels to the destination. What is the maximum channel load you can reach by using this routing algorithm for *any* traffic pattern?

3.11 *Performance and packaging of a cube-connected cycles topology.* The topology in Figure 3.16 is called the 3^{rd}-order cube-connected cycles.

(a) What is the channel bisection B_C of this topology?

(b) If minimal routing is used, what is the maximum hop count H_{max}? What is the average hop count H_{min}?

(c) Now we want to package this topology under a constraint of $W_n = 128$ signals per node and $W_s = 180$ across the backplane. Assume a packet size of $L = 200$ bits and a signaling frequency of 800 MHz, and also ignore wire length. What is maximum channel width w under these constraints? Is this network

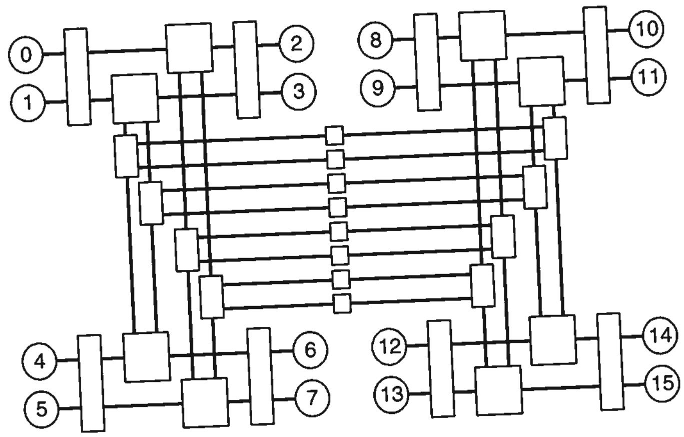

Figure 3.15 A 16-node, radix-2 fat tree. All rectangular nodes are switch nodes.

pin or bisection bandwidth limited? What does the router latency t_r need to be to ensure a zero-load latency of 75 ns?

3.12 *Physical limits of performance.* Using simple ideas, it is possible to compute realizable bounds on the diameter of a network and the physical distance between nodes once it has been packaged. First, if radix k switches (out degree k) are used, what is the smallest possible diameter of an N node network? Now assume that each node has a volume V. What packaging shape in three-dimensions gives the smallest maximum distance between nodes? What is this distance, assuming a large number of nodes?

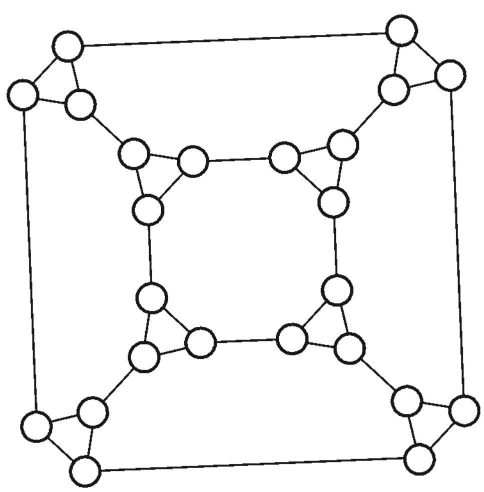

Figure 3.16 3^{rd}-order cube-connected cycles.

[119, 165, 33]. You will have the opportunity to experiment with such topologies in Exercise 5.6.

5.7 Exercises

5.1 *Comparing butterfly and torus topologies.* Compare 1,024-node butterfly and torus networks where the node bandwidth is 300 Gbits/s and the bisection bandwidth is 4 Tbits/s. For both networks, choose the smallest values of k and n so that the bisection bandwidth is saturated (for simplicity, consider only combinations of k and n where $k^n = 1{,}024$). What is the serialization latency of these networks for a packet length of $L = 1{,}280$? What is the average hop count H_{min}? Ignoring wire latency, with a per-hop latency of $t_r = 12$ ns, what is the zero-load latency of both networks?

5.2 *Tradeoffs in a 4,096-node torus.* Examine the tradeoff between k and n for a 4,096-node torus. For each combination of k and n where $k^n = 4,096$, what is the ideal throughput and average zero-load message latency? Assume each node has 120 signal pins, the bisection width of the system is 1,500 signals, the signalling frequency is $f = 2.5$ GHz, the packet length is $L = 512$ bits, and the router hop delay is 20 ns. Ignore wire latency ($T_w = 0$).

5.3 *A three-level packaging hierarchy.* A 256-node torus needs to be packaged in a three-level hierarchy under the following constraints: each node has 384 signal pins, 1,200 signals may go off a board, and 6,000 signals can cross the midsection of the backplane. Determine k and n that maximize the network's bandwidth. If several values of k and n achieve this goal, choose the one with the minimum zero-load latency. Explain how you would package this network by using nodes, boards, and a backplane while minimizing the number of boards used.

5.4 *Channel load in unidirectional tori.* Calculate the average channel load in unidirectional tori under uniform traffic. Approximately how many times greater is this load than in bidirectional tori? Explain the sources of this increase.

5.5 *Number of slightly non-minimal routes.* As a function of the minimal hop counts in the x and y dimensions, Δ_x and Δ_y respectively, how many routes are there in a 2-D torus if routes that are at most 1 hop longer than minimal are allowed? How many routes are there that are at most 2 hops longer than minimal? Assume k is odd.

5.6 *Doubly twisted tori.* The topology shown in Figure 5.13 is a doubly-twisted torus [165]. Assuming minimal routing, compare the average hop count H_{min} of this topology with that of a standard 4-ary 2-cube. Why does twisting change the hop count?

5.7 *Wire length in the layout of a torus.* Calculate the average wire length traversed by a message in a k-ary 6-mesh network laid out in a plane. Assume that all nodes are aligned on a 2-D grid and the inter-node spacing is 10 cm. For simplicity, measure distances between node centers. How does the average distance change if the network is packaged in three dimensions and the inter-node spacing in third dimension is also 10 cm? Assume k is even in both cases.

and router node is packaged in a separate chip and each chip can have a total pin bandwidth of up to 100 Gbits/s. Which topology offers the lowest pin bandwidth? Which offers the lowest pin bandwidth without incurring additional serialization latency?

7.2 *Distributing traffic from a line card.* Using distributors, suggest how you might connect 64 40-Gbits/s line cards using a torus network composed of channels that do not exceed 10 Gbits/s. The underlying network should support worst-case traffic in which all nodes send across the bisection. Does this arrangement change the bisection bandwidth compared to a network in which the channels could operate at 40 Gbits/s?

7.3 *Slicing a butterfly.* Consider the partitioning of a radix-4 butterfly node with $w = 2$-bit-wide channels into two modules using bit slicing, dimension (port) slicing, and channel slicing. Flits are 64-bits, 16-bits of which are header information. Sketch each partitioning, labeling all channel widths and the number of signals required per chip, and qualitatively compare the latency in each case.

7.4 *Sub-signal slicing.* It it possible to choose a channel slicing factor such that the resulting channels are less than one signal wide. To implement these channels, several could be multiplexed onto a single physical signal. For example, if the channels are sliced to one-half a signal, two channels would share one physical signal. Can this level of slicing ever reduce zero-load latency? Explain why or why not.

7.5 *Channel slicing a Clos network.* Find the channel slicing factor x that gives the minimum latency for an $N = 256$ rearrangeable ($m = n$) Clos network built from switches with 32 signals. Each node requires 8 Gbits/s of bandwidth, $L = 128$ bits, t_r=20 ns, and f=1 Gbit/s.

7.6 *Bit slicing a 4,096 node butterfly.* Consider a network that has the parameters of the network of the first row of Table 7.1. Find the *bit* slicing of this network that gives minimum zero-load latency. Assume that a 32-bit header is repeated in each bit slice. (No control signals are required between slices.) Also, the full header must be received before any of the router latency t_r is experienced.

7.7 *Header distribution in the Tiny Tera.* In the bit-sliced Tiny Tera switch, crossbar configurations are computed by the centralized scheduler and then redistributed to the input ports. This header information is then duplicated and attached to each outgoing packet slice. However, for a port A, the packet leaving A does not contain the destination port, but rather the address of the port writing to A (that is, configurations are not described by where to write a packet, but rather from where to read one). Why does this result in a more efficient encoding? Hint: The Tiny Tera supports multicast traffic.