GEORGIA INSTITUTE OF TECHNOLOGY

College of Computing

CS6290/CS4290 — High-Performance Computer Architecture

Fall 2001

CS6290/CS4290	Issued: November 9, 2001
Homework 6	Due: November 16, 2001

Purpose: This homework covers predication and introduces some issues in network performance.

Reading: H&P Sections 4.6, 7.1 and 7.2
[Kubi93] paper

Problems:
1. Effective Bandwidth.
2. Topology.
3. Reading: [Kubi93].

Collaboration: *(As in the syllabus)* collaboration on projects and homework in **pairs** is encouraged. If you work in a pair, turn in one write-up with the names of both collaborators. You're welcome to discuss high-level concepts with other groups, but all homework solutions must be worked out and written up separately.

Problem 1: Effective Bandwidth

A: Problem 7.1 in the book. Also, what message size is required to achieve half of the boilerplate (maximum) bandwidth of the network (half of 9Mbits/sec, according to the problem)?

B: Problem 7.2 in the book.

Problem 2: Topology

Assume you have to connect 64 nodes, each with and input and an output link. You also have available 16x16 switches that each have 16 inputs and 16 outputs (as in Figure 7.35 in the book). Assume that the latency through a switch (or node) is one unit of time and that the bandwith of a link is one unit.

> **A:** The lowest-cost network has no switches – instead you connect the 64 nodes into a ring. What is the bisection bandwidth of the network? What is the average message latency between randomly selected pairs of nodes?
>
> **B:** Using 16x16 switches, design a switch topology that has the minimum latency through the switches. Sketch the topology. How many switches are required? What is minimum latency? What is the bisection bandwidth of your topology?

Problem 3: Reading

Read the [Kubi93] paper on the Alewife message passing mechanism.

> **A:** Figure 8 shows the instruction sequence for composing and launching a message with a one-word payload of data where the data is written directly out of a register. Show the sequence for sending a message with an additional payload of 256 bytes from location `foo` in memory.
>
> **B:** Figure 8 also shows the detailed pipeline timing for launching a message: three cycles per `stio` plus one cycle for the `ipilaunch` plus two cycles of overhead. The processor runs at 33MHz; the network is byte-wide and runs at 66MHz (528Mbit/S). The instruction sequence shown launches one 32-bit word of payload in nine cycles. What bandwidth (in Mbits/S) is achieved if we repeat this sequence continuously? What is the maximum bandwidth if we send longer messages via the same mechanism (more `stio`s from registers)? Note that the network is sufficiently fast that the processor is always the bottleneck. What length of message is required to achieve half the boilerplate bandwidth of the network (half of 528Mbit/S)?