EE/CSCI 451
Fall 2020
Homework 6

Assigned: October 2, 2020
Due: October 9, 2020, AOE
Total Points: 100

1 [10 points]

Explain the following terms:

- All-to-all broadcast primitive
- Prefix sum primitive
- Gather primitive
- Permutation communication primitive
- All-to-all personalized communication primitive

2 [25 points]

In the class, we discussed a recursive doubling algorithm for one-to-all broadcast. Note that when the algorithm is executed on a linear array, with nodes represented in binary, a message is first sent along the processors with different most significant bits, then with different second most significant bits and so on (in the algorithm discussed in the slides, j goes down to 0). For example, in a 8 node linear array, a message is sent from processor 0 to processor 4, then in the next iteration messages are sent from processor 0 to processor 2 and processor 4 to processor 6 and so on.

Let's now change the algorithm so that the message is sent along the least significant bit first (i.e., in the algorithm discussed in the slides, j goes up to $logp - 1$). So in the first time step, processor 0 will communicate with processor 1; in the second time step, processors 0 and 1 will communicate with 2 and 3, respectively; and so on (**Note:** the if condition needs to be modified accordingly so that number of processors sending data in time step j is 2^j).

1. Write an algorithm using recursive doubling technique for this revised algorithm for one-to-all broadcast?

2. What is the run time of this revised algorithm on a linear array network?

For these derivations, if k messages have to traverse the same link at the same time, then assume that the effective per-word-transfer time for these messages is kt_w.

3 [25 points]

Given a balanced binary tree as shown in Figure 5, describe a procedure to perform all-to-all broadcast that takes time $(t_s + t_w mp/2)\log p$ for m-word messages on p nodes. Assume that only the leaves of the tree contain nodes, and that an exchange of two m-word messages between any two nodes connected by bidirectional channels takes time $t_s + t_w mk$ if the communication channel (or a part of it) is shared by k simultaneous messages.

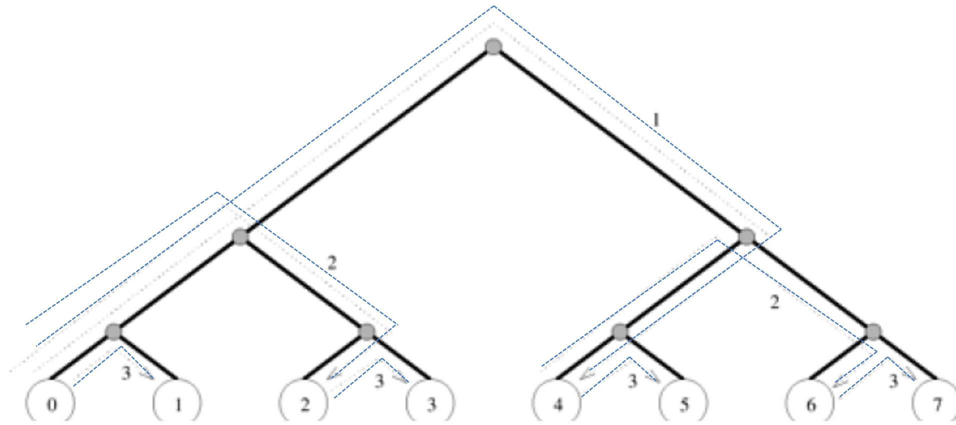


Figure 1: Figure 4.7 of textbook

4 [20 points]

Consider the all-reduce operation in which each processor starts with an array of m words, and needs to get the global sum of the respective words in the array at each processor. This operation can be implemented on a ring using one of the following two alternatives:

1. All-to-all broadcast of all the arrays followed by a local computation of the sum of the respective elements of the array.
2. Single node accumulation of the elements of the array, followed by a one-to-all broadcast of the result array.

For each of the above cases, compute the run time in terms of m, t_s, and t_w.

5 [20 points]

- Consider the figure below which has 3 processors. Each processor has an array of 3 elements.

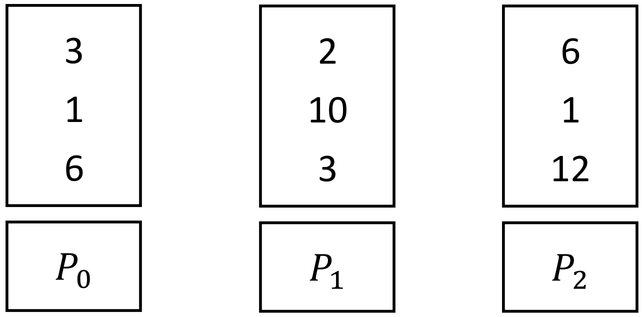
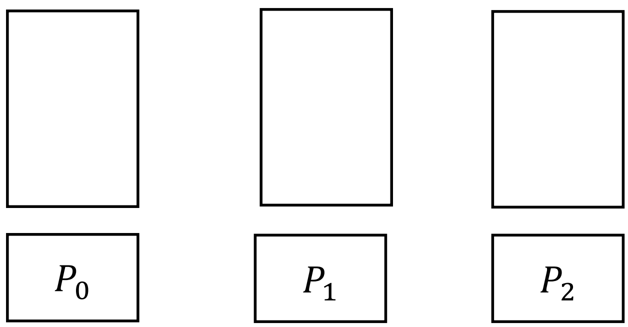


Figure 2: Problem 5

Perform all-reduce using sum reduction operation. Show the results after all-reduce.



- Assuming all-reduce is implemented as all-to-one reduction followed by one-to-all broadcast. Both all-to-one reduction and one-to-all broadcast use recursive doubling algorithm. Assuming there are p processors, each having an array of m elements. What is the total amount of communicated data in terms of p and m when performing all-reduce using sum reduction operation?