Parallel Processing

Fall 2018

Assignment No. 1

Total Points: 100

Multithreading

1. There are three types of implementations of multithreading:

    - Fine-grain multithreading
    - Coarse-grain multithreading
    - Simultaneous multithreading

    What is the exact difference between Simultaneous multithreading and the other approaches? Does it mean that the fine-grain and coarse-grain multithreading are only suitable for single core processor?

2. Consider the enumeration sort algorithm presented below:

```
procedure ENUM SORT (n)
begin
   for each process P1,j do
       C[j] :=0;
   for each process Pi,j do
       if (A[i] < A[j]) or ( A[i]= A[j] and i < j) then
           C[j] := 1;
       else
           C[j] := 0;
   for each process P1,j do
       A[C[j]] := A[j];
end ENUM_SORT
```

The basic idea behind enumeration sort is to determine the rank of each element. The rank of an element a_i is the number of elements smaller than a_i in the sequence to be sorted. The rank of a_i can be used to place it in its correct position in the sorted sequence.

Assume that concurrent writes to the same memory location of the PRAM result in the sum of all the values written being stored at that location. Consider the n^2 processes as being arranged in a two-dimensional grid. This algorithm consists of two steps. During the first step, each column j of processes computes the number of elements smaller than a_j. During the second step, each process $P_{1,j}$ of the first row places a_j in its proper position as determined by its rank. It uses an auxiliary array $C[1...n]$ to store the rank of each element. The crucial steps of this algorithm are lines 7 and 9. There, each process $P_{i,j}$ writes 1 in C [j] if the element A[i] is smaller than A[j] and writes 0 otherwise. Analyze the performance of each of the following:

- CRCW PRAM
- CREW PRAM
- EREW PRAM

Evaluation Metrics

3. The bisection bandwidth of a network is the maximum transmission performance over the bisection line, i.e. the sum of all single channel bandwidths from all edges that are "cut" via the bisection line.
    a. Given is a 3-dimensional torus with size 8×8×16 (i.e. 1024 nodes). Furthermore, every edge (i.e. cable) of this network has a bandwidth of 400 MBps. Every message sent through this network – assuming there is no congestion – has a latency of 0.5 ns to be passed from one node to another.
        i. Calculate the bisection bandwidth of this network.
        ii. Calculate the diameter of this network.
        iii. Calculate the maximum delay of this network.

    b. Given is a 10-dimensional hypercube (i.e. 1024 nodes). Furthermore, every edge (i.e. cable) of this network has a bandwidth of 120 MBps. Every message sent through this network – assuming there is no congestion – has a latency of 0.7 ns to be passed from one node to another.
        i. Calculate the bisection bandwidth of this network.
        ii. Calculate the diameter of this network.
        iii. Calculate the maximum delay of this network.

    c. Calculate the cost of both networks described in a. and b. in order to make a decision which topology to choose for a 1024 processor machine concerning bisection bandwidth, delay, and cost. Discuss your decision!

Routing Mechanisms

4. Can the switching network below, built of 2×2 exchange elements perform all permutations from inputs to outputs? Prove or disprove as simply as possible.