

## CSE 5441 Spring'22

### Assignment #1 – Basic Concepts and Principles of Parallel Computing

1. What is parallel overhead and provide an equation to compute parallel overhead?
2. What do you call a parallel program that does little or no communication?
3. Define Scalability of a parallel system and describe what is weak scaling and strong scaling?
4. Assume 19% of the runtime of a program is not parallelizable. This program is run on 128 cores of a state-of-the-art 128-core AMD Milan system. Under the assumption that the program runs at the same speed on all those cores, and there are no additional overheads, what is the parallel speedup?
5. Given an algorithm where the fraction that the program is serial is 0.35, what is the efficiency on 56, 64, 128 and an infinite number of processors. You may ignore communication costs and the problem size is fixed.

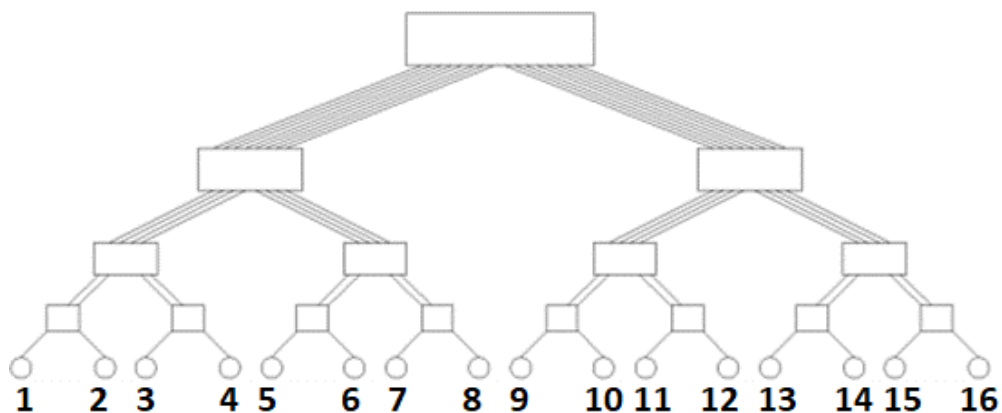
<b>p</b>	<b>56</b>	<b>64</b>	<b>128</b>	<b><math>\infty</math></b>
<b>Efficiency</b>				

6. What are the main differences between a Data Parallel Model and Distributed Memory Model?

7. Why is CPU utilization an important metric for high-performance interconnects?

8. What are the benefits and deficiencies of fine-grained parallelism when compared to coarse grained parallelism?

9. You are the designer of a high-performance communication stack and are asked to design an efficient implementation of one-to-all broadcast on a full fat tree utilizing cut through routing as shown below. The time complexity of the algorithm should be better than or equal to  $\log(n)$  where 'n' is the number of processes (in our case, assume  $n = 16$ ). Please show the communication operations involved in your algorithm during each step assuming that the cost per hop is trivial and can be ignored?



Hint: Please write the communication steps in the following fashion. Note that the communication steps shown here are just examples and are not the actual answer.

Time Step – 1

1 -> 2

Time Step – 2

1-> 4

2 -> 3

...

Time Step – n

10. How would the above communication steps differ if you assume that the cost per hop is non-trivial and cannot be ignored? Please show the communication steps as in the previous question. Refer slides on “Cut-Through Routing” for equation on total communication time.