EECE5640
High Performance Computing
Homework 4
*Submit your work on Turnitin on Blackboard

1. (40) On the Discovery Cluster, develop an MPI program that computes pi using the dartboard method. You can find a discussion of this method in the MPI tutorial provided on Blackboard. Vary the number of darts and the number of tasks. Report on the execution time of your program.

   *Answers to this question should be included in your homework write-up in pdf format, and submitted through Turnitin. You should include a C/C++ program submitted through Turnitin.

2. (50) Develop a parallel histogramming program using C/C++ and MPI. A histogram is used to summarize the distribution of data values in a data set. The most common form of histogramming splits the data range into equal-sized bins called classes. For each class, the number of data values in the data set that fall into that class are totaled. Your input to this program will be integers in the range 1-1000 (use a random number generator). Your input data set should contain N integers, where N can be varied.

   Perform binning on each node using MPI, and then perform a reduction on node 0 to combine your partial results. Run this on 1, 2 and 4 nodes on Discovery. Your program should print out the values that fall into each class, in ascending class order.

   a. For your binning, use each node to produce a histogram for (1/# classes)th of the input data values. Experiment with varying the number of classes and different values for N. Plot your performance results.
   b. For your binning, use each node to only produce values for a single class of the input data values. Experiment with varying the number of classes and different values for N. Plot your performance results.
   c. Compare the performance results for a and b. Attempt to explain any differences.
   *Answers to this question should be included in your homework write-up in pdf format, should include two C programs submitted on Blackboard, and submitted through Turnitin.

3. (10) Performance analysis of an MPI application has been an active area of research. There have been many performance tools developed. Please identify two of these frameworks and compare and contrast the capabilities of the toolsets you have selected. Make sure to cite all your resources.
   *Please do not copy text out of user guides for the frameworks.

4. (20) Read the paper on Blackboard titled "MPI on Millions of Cores." Given that this paper was published in 2010 (9 years ago), can you comment on what changes have occurred since 2010 that could positively and/or negatively impact our ability to fully exploit parallelism on millions of cores? Many of the papers today discuss exascale computing. Select a recent paper on exascale-computing and compare/contrast the barriers identified in the two papers that impact our ability to achieve these milestones.
   This problem is extra credit for the undergraduates in the class, though is required by graduate students.

   *Answers to this question should be included in your homework write-up in pdf format, and submitted through Turnitin.