

EECE5640
High Performance Computing
Homework 5

***Submit your work on Turnitin on Blackboard**

1. (40) In this problem we will be working with a graph-based data set. We are interested in characterizing relationships between authors that have co-authored publications that have appeared in the DBLP computer science bibliography. We are going to use the co-authorship network of the DBLP computer science bibliography (see <http://konect.uni-koblenz.de/networks/com-dblp>). Nodes in the graph are authors, with undirected edges between two nodes present if the corresponding authors have published at least one paper together. Each entry in the data set indicates that the first author has co-authored a publication with the second author. The data has been deidentified – numbers are used to identify unique authors. Your job is to develop an MPI program (use at least 2 nodes) that can produce the following features about this graph-based dataset:
 - a. Find the list of authors that have the most co-authors (this list may contain 1 or multiple authors).
 - b. Plot the distribution showing the number of authors (on the y-axis) that have published with exactly d authors (on the x-axis).

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

2. (40) Develop parallel CUDA programs to implement both parts a. and b. described in Problem 1. Run on a single GPU. Compare the performance of your CUDA implementation with your MPI implementation.

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

3. (20) Read the Pascal whitepaper provided, and then identify the key features that were introduced in the Pascal P100 architecture, as compared against the Volta V100 architecture (make sure to identify the source for the information you obtained on the V100). Please do not just repeat what you read in the Pascal whitepaper, go into more detail on each of the features you identify.

*Answers to this question should be included in your homework write-up in pdf format.

4. (30 points extra credit for both UG and Grad) Extending Problems 1 and 2, develop a parallel CUDA program that computes the number authors (but not co-authors) that share a common co-author.

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