**Blue Waters Petascale Semester Curriculum v1.0**

**Unit 2: Parallel Computing Concepts**

**Lesson 1: Types of Parallel Work - Data and Task Parallelism**

**Exercise Instructions for Students**

*Developed by Peter J. Hawrylak for the Shodor Education Foundation, Inc.*



*Except where otherwise noted, this work by The Shodor Education Foundation, Inc. is licensed under CC BY-SA 4.0. To view a copy of this license, visit*[*https://creativecommons.org/licenses/by-sa/4.0*](https://creativecommons.org/licenses/by-sa/4.0)

*Browse and search the full curriculum at*[*http://shodor.org/petascale/materials/semester-curriculum*](http://shodor.org/petascale/materials/semester-curriculum)

*We welcome your improvements! You can submit your proposed changes to this material and the rest of the curriculum in our GitHub repository at*[*https://github.com/shodor-education/petascale-semester-curriculum*](https://github.com/shodor-education/petascale-semester-curriculum)

*We want to hear from you! Please let us know your experiences using this material by sending email to* [*petascale@shodor.org*](mailto:petascale@shodor.org)

Problem 1: Identify data and task parallelism aspects of finding the area under the curve of a polynomial on a range of [x1, x2]. Include a justification to explain why a part or the entire problem is data or task parallel. For each aspect of parallelism include a discussion of why it is not the other type of parallelism too.

Problem 2: Recall that Gaussian elimination can be used to solve a set of linear simultaneous equations. A simple implementation is to start at the top row and first column and convert the matrix into the identity matrix and update the extra solution row with each step. The process starts with the leftmost column and works to the right of the matrix. Once the left part is the identity matrix the extra solution row will contain the values for each of the *n* variables in the set that solve the problem. Assume that there are 5 linear equations with 5 unknowns yielding a 5x6 matrix (the 6th column (rightmost column) is the solution set column and is not converted into the identity matrix). Devise a schedule of operations and identify if each operation is data or task parallel. Include rational as to why each operation is data or task parallel.

Problem 3: Recall that Gaussian elimination can be used to solve a set of linear simultaneous equations. One problem with this approach is that as the size of the set of equations grows the inaccuracies of floating point mathematical operations make the final solution less accurate. There are several ways to address or minimize this. For this problem assume that you have a routine, MINIMIZE\_ERROR(), that can minimize error of the matrix but that it needs to work on the entire matrix. There is a routine GAUSSIAN\_STEP() that performs one step of Gaussian elimination. Also, assume that you need 20 steps of Gaussian elimination to find the solution and that the maximum allowable error will occur if MINIMIZE\_ERROR() is not called at least once every 5 Gaussian elimination steps. Devise a schedule of operations and identify if each operation is data or task parallel. Include rational as to why each operation is data or task parallel.