SWE 507 PARALLEL PROGRAMMING PROJECT STUDY 2 (2024)

Due date: 19 April 2024, in class time presentation. Do not send any document.

In this project, you are expected to carry out experiments on distributed parallel programming environment using MPI and AWS cloud cluster. You should measure the performance (execution time) versus process (CPU count) count.

The problem is matrix vector multiplication with different number of processes.

$$\mathbf{A}_{1000x1000} \times \mathbf{b}_{1000x1} = \mathbf{c}_{1000x1}$$

A is a square matrix with size of 1000x1000, **b** and **c** are vectors, all matrix elements are double.

In the first phase you should develop and test your program on your local computer, and then performance (timing) experiments will be done on both local computer and AWS cloud. Check the following links for how to create a MPI cluster on Amazon Cloud for experiments, you can chose one of them (don't forget to terminate your cluster when you are done):
1- https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/efa-start.html
2 - https://mpitutorial.com/tutorials/launching-an-amazon-ec2-mpi-cluster/

Project steps:

- In master process create the matrices, and initialize **A** and **b** with random real values between [0-100].
- Scatter the matrix **A** rows to all nodes equally, broadcast the vector **b** to all nodes.
- Do the calculations on the nodes including the master.
- Do the experiments 1, 2, 4 and 8 processes on local computer. Check the correctness with a smaller size of matrices. Get timing as you did on your first project.
- Create a EC2 MPI cluster on AWS, move your source code to AWS. You can use GitHub or simply sftp.
- Do your experiments on AWS cluster get the timing.
- Draw a graph using spreadsheet that shows thread count (x-axis) versus execution time (y-axis)
- Compare local runs with ASW cluster runs and comment the results.

You should justify that your program works correctly. Hybrid parallel approach (MPI+Pthread) is not expected in this project. Send me email if you have any questions.

Grading:

No	Task	Grade
1	Application is implemented locally, and it runs correctly on local computer.	50
2	Local time measurements are done, graph/comments are OK.	30
3	EC2 MPI cluster experiments are done, and performance study is OK.	20

PS: All experiments will be done in class presentation, be prepared before coming.