MSc in CSTE
High Performance Technical Computing

Assignment

Irene Moulitsas
Cranfield University

December 8, 2023

Hand in date: 05/02/2024 (FT), 19/02/2024 (PT), 09:30am

1 Introduction

In this assignment you are asked to examine the application of distributed-memory parallel programming techniques for the implementation of sparse-matrix fat-vector multiplication $M * v$

$$\begin{pmatrix} \times & \times & & & \times & & & \times \\ \times & \times & \times & & & & & \\ & & & & & \times & \times & \\ & \times & \times & \times & & & & \\ & & \times & \times & & & & \\ \times & & & & \times & \times & & \times \\ \times & & & & \times & & & \times \\ & & & & \times & \times & \times & \\ & & & & & \times & \times & \\ \times & & & & \times & & & \times \end{pmatrix} * \begin{pmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{pmatrix}$$

where M is a $m \times n$ sparse matrix and v is an $n \times k$ vector.

2 HPC Tasks

1. First devise a serial algorithm to implement the sparse-matrix fat-vector multiplication and then devise three different ways to parallelise it.
2. Analyse your parallel algorithms; present and describe the flow of the algorithms, significant data structures, problem decomposition, load balancing, and potential synchronisation points where applicable.

3. Implement your algorithms in a program written in a programming language of your choice, between C, C++, FORTRAN, and coupled with MPI.

4. Compare the solutions obtained by each one of the parallel algorithms with that of the serial program.

5. Measure the cost of communication and the cost of computations for each of the parallel solutions. How is the communication cost affected by the use of different communication patterns?

6. Measure the performance of your serial and parallel codes and discuss. How does the parallel program performance compare to the theoretical one? Is the performance of your parallel approaches the expected/theoretical one?

7. Replace your own methods with an appropriate one from an external mathematical library or repository. Study how the performance of this new different implementation. Is it better or worse compared to your implementations? Why?

8. Based on your results, reason on the problem sizes deemed necessary in order for MPI parallelisation to become efficient for the above algorithms.

9. Reflect upon the article of Irina Kupiainen, of the CSC-IT Center for Science in Finland where LUMI (#5 in the Nov 2023 TOP500 list) is based, and the historic information available at Green500. Discuss the effect of HPC on the environment. Use more sources as appropriate.

3 Source Code and Report Requirements

The source program will need to compile on Crescent 2 using the Intel compilers and Intel MPI. Your simulations for the above tasks should be performed on Crescent 2 using the queueing system and working nodes.

Write a report to present and discuss your findings. The report should be no less than 2,000 words and must not exceed 3,500 words. The report can contain any number of figures/tables, however all figures/tables should be numbered and discussed. The report should include a description of the design of your solution explaining your choices. The source code and a sample scheduler script should be included as Appendices to the report.

4 Assignment Submission

The source code and the sample scheduler script must be submitted electronically via the **Canvas Coding Files submission point** by 9:30am on the 5^{th} February (full-time students) or the 19^{th} February (part-time students).

The report should be submitted electronically via the **TurnItInUK Report submission point** by the prescribed deadline, for the assignment submission to be considered complete.

5 Marking

The assignment will be assessed based on the following marking scheme:

- 20% Introduction, methodology, conclusions
- 40% Source code, design
- 30% Discussion and analysis
- 10% Report structure, presentation, references