

README

December 6, 2023

1 Exercise 2 - Taxonomies and Definitions

The learning goal of this exercise is to understand **the most important definitions** that you need when programming parallel supercomputers. We will use these extensively in the coding exercises, so it is very important to master these concepts.

We undertake this exercise during the second lecture session in the format of group work; everyone participating will get full points. If you cannot make it, you can also write a learning diary and return it to [MyCourses](#). The diary will be graded by the course personnel. The group work outcomes will be distributed as the correct solution to this sheet.

If you decide to submit this exercise as a learning diary: the tasks have equal weight in the grading.

2 Your tasks

2.0.1 1. Inspect the list of [top 500 computers in the world](#)

Hints: use the same method (inspect evolution over time) as in sheet 1 and at least the same keywords. You are encouraged to inspect further keywords.

Guiding questions:

1. What kind of trends in terms of Flynn's taxonomy (SISD, MISD, SIMD, MIMD) can be seen, if any?
 2. What does that mean for the actual programming concepts?
-

2.0.2 2. Inspect the Mahti interconnect network either from CSC docs or lecture slides (slide 19).

Hint: you can also find a short description of it from CSC docs.

Guiding questions: Analyze it as per

1. Topology
 2. Connection type
 3. Bandwidth
-

2.0.3 3. You have a code of which 80% can be efficiently parallelized.

Hint: Remind yourself of Amdahl's (p. 30 and 32 of lecture slides) and Gustafsson's laws (p. 33-34).

1. Draw a schematic strong scaling plot, with ideal scaling and the expected one: $T(\text{serial})/T(p \text{ processors})$ as function of p .
 2. Draw a schematic weak scaling plot, when the proportion of work is kept fixed per processor count p (ideal and expected).
-

2.0.4 4. We have a 2D iterative stencil loop problem to program, see lecture slide 37.

Now we need to use a 4th order Moore stencil instead of the von Neumann one. If your 2D subdomain would have the dimension of 16 points, what would be the ratio of grid points that you can compute and communicate at the same time? Analyse the compute/communication performance according to the ACC model (slide 28).

2.0.5 5. Synthesis: why do you think these concepts are important to master?