Garlic experiments

Rodrigo Arias Mallo

Barcelona Supercomputing Center

January 14, 2021

Approach 1

This was the approach proposed for hybrids PM

- Perform a granularity experiment with a *reasonable* problem size.
- Take the best blocksize
- Analyze strong and weak scaling with that blocksize.
- Plot speedup and efficiency comparing multiple PM.

The main problem is that it may lead to **bogus comparisons**. Additionally, there is no guarantee that the best blocksize is the one that performs better with more resources.

Approach 2

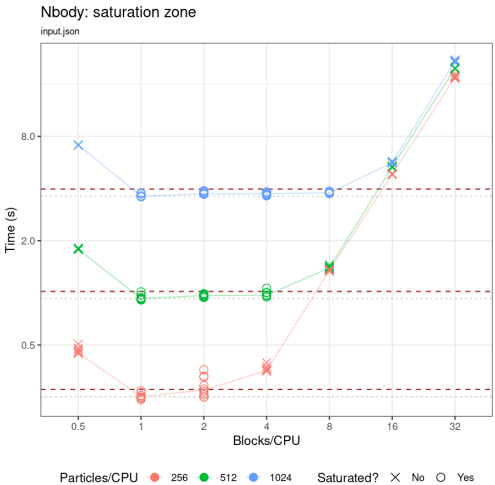
We want to measure scalability of the application **only**, not mixed with runtime overhead or lack of parallelism.

We define **saturation** as the state of an execution that allows a program to potentially use all the resources (the name comes from the transistor state, when current flows freely).

Design a new experiment which tests multiple blocksizes and multiple input sizes to find these states: **the saturation experiment**.

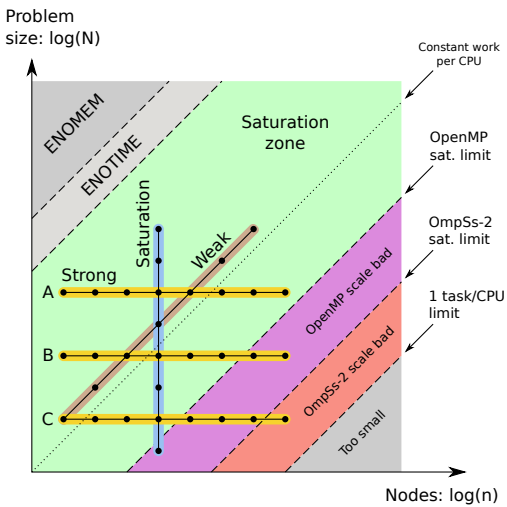
Begin with small problems and increase the size, so you get to the answer quickly.

Saturation experiment



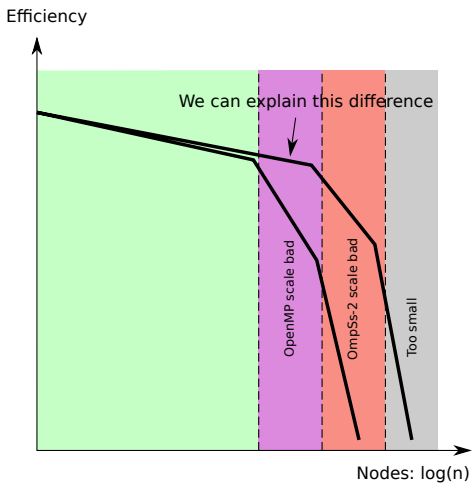
- The objetive is to find the minimum input size that allows us to get meaningful scalability results.
- More precisely, a unit is in **saturation state** if the median time is below the **saturation time limit**, currently set to 110% the minimum median time (red dashed lines).
- An input size is in **saturation zone** if it allows at least K=3 consecutive points in the saturation state.
- With less than 512 particles/CPU (green line) we cannot be sure that the performance is not impacted by the runtime overhead or lack of parallelism.

Experiment space



- **Saturation limit**: small tasks cannot be solved without overhead from the runtime, no matter the blocksize.
- Different limits for OmpSs-2 and OpenMP.
- Experiment A will show the scaling of the app while in the saturation zone.
- Experiment B will show that OpenMP scales bad in the last 2 points.
- Experiment C will show that at some point both OpenMP and OmpSs-2 scale bad.

Experiment space: experiment C



- The experiment C will show a difference in performance when approached to the saturation limit.
- We could say that OmpSs-2 introduces less overhead, therefore allows better scalability.

Reproducibility

How easy can we get the same results? Three properties R0 < R1 < R2 (no common nomenclature yet!):

- R0: **Same** humans on the **same** machine obtain the same result
- R1: **Different** humans on the **same** machine obtain the same result
- R2: **Different** humans on a **different** machine obtain same result

Garlic provides 2 types of properties: for software and for experimental results:

- Software is R2: you can get the exact same software by any one, in any machine
- Experimental results are R1: you cannot change the machine MN4 (yet)

Same experimental result means that the mean of your results is in the confidence interval of our results **and the relative std is < 1%**.