Exercise 2: Collective communications

12/06/2019

1. Write a modification of the solution to Exercise 1 that uses collective, instead of point-to-point communications.

2. When computing average of the result of N results of some quantity A in a simulation that is subject to correlations between successive iterations, one might see the effect of these correlations in the variance of the data. In order to remove this unwanted effect (as the values should behave as independent random variables so that the central limit theorem can be used), it is necessary to partition the data into n_b blocks of size τ_b and take the average

$$\langle A \rangle_b = \frac{1}{\tau_b} \sum_{i=1}^{\tau_b} A_i \tag{1}$$

of the quantity over each block, which one expects to behave like independent random variables.

As τ_b is increased, block averages are expected to be uncorrelated. The *statistical efficiency* due to correlations is

$$s = \lim_{\tau_b \to \infty} \frac{\tau_b \ \sigma_b^2(A)}{\sigma^2(A)}, \tag{2}$$

where

$$\langle A \rangle = \frac{1}{N} \sum_{i=1}^{N} A_i = \frac{1}{n_b} \sum_{b=1}^{n_b} \langle A \rangle_b \tag{3}$$

and

$$\sigma^2(A) = \frac{1}{N}\sum_{i=1}^{N}(A_i - \langle A\rangle)^2 \tag{4}$$

are respectively the average and variance of the complete set of N measurements, while

$$\sigma_b^2(A) = \frac{1}{n_b}\sum_{b=1}^{n_b}(\langle A\rangle_b - \langle A\rangle)^2 \tag{5}$$

is the variance of the block averages.

Finally, the corrected standard deviation of the data is estimated from the simulation as

$$\sigma_{est}(A) = \sigma(A)\sqrt{\frac{s}{n_b}} \tag{6}$$

Write a program that reads a list of data and, using MPI, computes the corrected variance for a given number of blocks n_b, given that N is a multiple of n_b.

(a) MPI_Scatter the data among all the processes in the communicator

(b) Make each process to compute the average of its block of data $\langle A\rangle_b$ and then use MPI_Allreduce to compute the total average $\langle A\rangle$, which is needed by all processes

(c) Within each block compute the partial sum $\sum_i(A_i - \langle A\rangle)^2$ and then use MPI_Reduce to compute the total variance $\sigma^2(A)$

(d) Make process 0 to MPI_Gather the $\langle A\rangle_b$ values and compute the block average variance $\sigma_b^2(A)$

(e) Finally, make process 0 compute the statistical efficiency s and the corrected estimation of $\sigma_{est}(A)$

The file `block_average_ser.cpp` has a serial implementation of the algorithm, which can be used as a base for the MPI code.