Assignment 5: File compressor

Student name: Pham Tran Huong Giang

Matricola: 606579

Implement

Emitter: Walk through the directory and check for the size of each file. If file size is smaller than BIGFILE_LOW_THRESHOLD the emitter will chunk the file into smaller parts. The size of each part is not greater than BIGFILE_LOW_THRESHOLD. After having all the files smaller than BIGFILE_LOW_THRESHOLD, the emitter sends the file to the workers to compress them.

Worker: Compresses and writes files.

Collector: The collector merges the parts of the same file together and also deletes all the intermediate files including file *".part*"* and file *".part*.zip".*

I set the BIGFILE_LOW_THRESHOLD = 10000000 (10MB)

Performance

Compression

The program is tested on the "silesia" folder. This folder consists of 13 files with 4 files larger than BIGFILE_LOW_THRESHOLD, before compression, the file of the folder is 212MB. After compression, the size of the folder is 69.5MB.

Running time

For the running time, the program is tested on the XEON PHI machine. For each number of workers, I run the program 5 times and take the average running time. The performance is as figure below: