FAST DATA STRUCTURES (MERGESORT) IN NON-C SYSTEMS LANGUAGES

Tofig Bakhshiyev[1]

[1]Institute of Computer Science, University of Tartu

Introduction

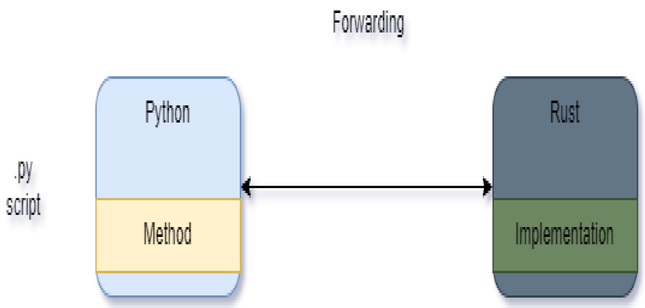
It is common idea to build programs quickly using high-level languages like Python and Ruby but implement data structures in C/C++ but the C language is believed to be error-prone, which can also cause security issues because of that in this project, The main goal is that to implement Merge sort algorithm in Rust language in parallel way and bind to Python language in order to get performance.

Merge sort

In computer science, merge sort (also commonly spelled as mergesort) is an efficient, general-purpose, and comparison-based sorting algorithm. Most implementations produce a stable sort, which means that the order of equal elements is the same in the input and output and time complexity of this algorithm is O(n*logn). The Merge sort algorithm is very scalable through its high parallelization capability, which allows the use of many processors [2].

About binding

Binding generally refers to a mapping of one thing to another. In the context of software libraries, bindings are wrapper libraries that bridge two programming languages, so that a library written for one language can be used in another language [2]



Results

In terms of time, with single thread python and binding library performed same in sorting, 12 second and 13 second respectively. On the other hand, with eight thread binding-Rust library performed significantly better than Python parallel merge sort, 5min and 7.7 second. Figure 1. But in memory test, with single thread and multiple thread binding Rust library performed very well. As as result, with this library we gained speedup and less memory footprint Figure 2.

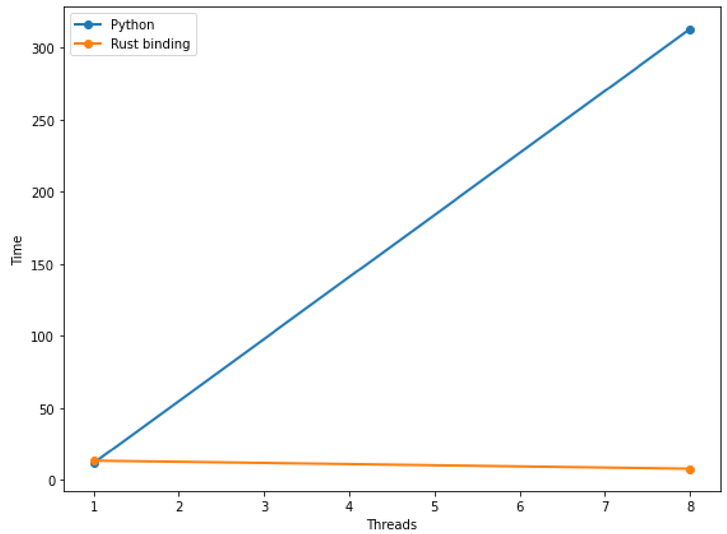


Fig. 1: Figure 1 Time.

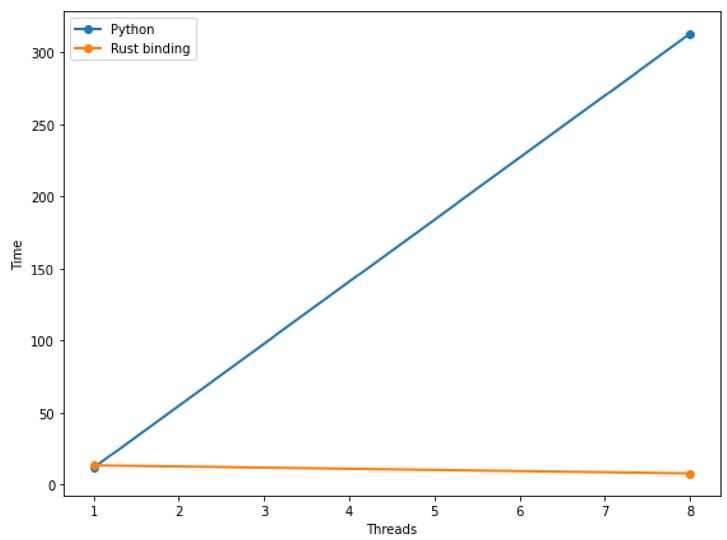


Fig. 2: Figure 2 Memory usage.

References

[1]https://github.com/TofigBakhshiyev/AlgoProject_parallelsort-rust

[2]https://en.wikipedia.org/wiki/Language_binding

[3]https://github.com/PyO3/maturin

[4]https://gist.github.com/stephenmcd/39ded69946155930c347