MDPI
sciforum

MOL2NET, International Conference Series on Multidisciplinary Sciences

A proposal for distributed processing for large scale virtual screening using Python
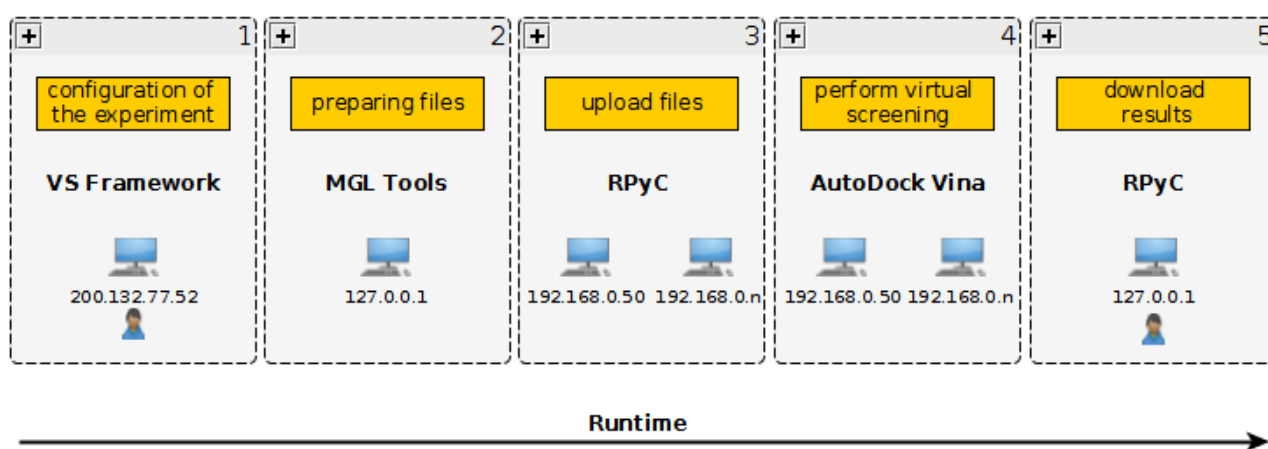
Alex Dias Camargo (alexcamargoweb@gmail.com) [a],

Vinícius Rosa Seus (viniciusseus@gmail.com) [b]

[a] Federal University of Pelotas - UFPEL

[b] Federal University of Rio Grande - FURG

Graphical Abstract

Abstract

The distributed virtual screening architecture proposed in this work is defined in 5 steps: Step 1 (configuration of the experiment), the VS Framework is used for the initial configuration of the experiment, with the other steps are automated by a Python script; Step 2 (preparing files), all virtual screening entries are converted to the default formats of the rest of the run, for which AutoDock Tools is used; Step 3 (upload files), the converted files are submitted to the computers that integrate the distributed execution; Step 4 (perform virtual screening) and Step 5 (download results), the virtual screening is processed and the results are sent to the computer that started the experiment. The results were listed among four combinations of simplified distributed computing infrastructures, divided by 5 virtual machines.

Introduction

Bioinformatics can be defined as the use of computational tools in the study of problems and biological issues, covering applications related to several areas of knowledge, such as biochemistry and computer science [Verli et al. 2014]. Thus, virtual screening acts in a standard step before laboratory experiments, for the discovery of new drugs. This technique involves calculating the estimated affinities (molecular docking) and the plausible binding modes of many drug candidates [Zhu et al. 2013]. The main focus of this work is the use of AutoDock Vina [Trott and Olson 2010], a

tool for molecular docking and virtual screening, as well as a distributed implementation using the Python programming language.

Materials and Methods

For the configuration of the experiments VS Framework [Seus et al. 2016] was used, this is a tool of preparation of virtual screening in which it allows the user to configure in a simplified way an experiment through a web platform. RPyC (Remote Python Call) [Filiba 2018], a Python library for symmetric remote procedure calls was used to execute jobs. Basically, RPyC uses the Remote Procedure Call (RPC) protocol for a client-server implementation.

Results and Discussion

Table. Results with Exhaustiveness = 8

IP list	Receptor	Ligand	Runtime
192.168.0.51 (1 VM)	1YT9.pdb	ZINC_results_AZT.mol2	7,60 min
192.168.0.51-52 (2 VMs)	1YT9.pdb	ZINC_results_AZT.mol2	5,12 min
192.168.0.51-53 (3 VMs)	1YT9.pdb	ZINC_results_AZT.mol2	4,05 min
192.168.0.51-54 (4 VMs)	1YT9.pdb	ZINC_results_AZT.mol2	3,97 min
192.168.0.51-55 (5 VMs)	**1YT9.pdb**	**ZINC_results_AZT.mol2**	**3,85 min**

The processing of each receptor-ligand in a distributed way had a better performance with the increase of computers in the structure. Finally, the analysis showed that these platforms are suitable for virtual screening executions with different receptor/ligand sizes. These considerations may guide scientists in choosing the best computing framework for their large-scale experiments.

References

Filiba, T. (2018). Rpyc - transparent, symmetric distributed computing. Available in: http://rpyc.readthedocs.io/en/latest/index.html.

Seus, V. R., Silva Jr, L., Gomes, J., da Silva, P. E., Werhli, A. V., and Machado, K. S. (2016). A framework for virtual screening. In Proceedings of the 31st Annual ACM Symposium on Applied Computing, pages 31–36. ACM.

Trott, O. and Olson, A. J. (2010). Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of computational chemistry, 31(2):455–461.

Verli, H. et al. (2014). Bioinformática – Da Biologia à Flexibilidade Molecular. Porto Alegre.

Zhu, T., Cao, S., Su, P.-C., Patel, R., Shah, D., Chokshi, H. B., Szukala, R., Johnson, M. E., and Hevener, K. E. (2013). Hit identification and optimization in virtual screening: practical recommendations based on a critical literature analysis: miniperspective. Journal of medicinal chemistry, 56(17):6560–6572.