

Introduction

PRIMAGE plans to deliver a platform to support decision making in the clinical management of malignant solid tumours. Among other elements, this task will be facilitated by development and execution, on available medical cohort data, a series of multi-scale simulations, working on the level of molecules, tissue and the whole tumour. These *in silico* techniques will be assisted with a set of image-based approaches to identify biomarkers, AI models for clinical predictions, and advanced visualisation for simulations. To efficiently execute these codes on both retrospective clinical data, and for prospective cohorts of two clinical trials planned by PRIMAGE, the project will need to apply High Performance Computing (HPC).

This deliverable presents the HPC resources made available for PRIMAGE consortium members, in order to achieve the project's goals. Description of the type of resources is given, followed by a detailed explanation on the procedure to register and obtain access to them. The document also provides description of several access modes which may be used by PRIMAGE users to utilize the computational power and the storage space provided to them in the scope of WP2.

When elaborating methods and tools for an efficient usage of different compute and storage resources for the PRIMAGE project, we will take into account typical requirements for large scale computing¹ as well as the experience gained during our extensive, long collaboration with scientists involved in medical computing on distributed e-infrastructures^{2,3}. Data will be processed according to the FAIR principle⁴. Furthermore, we will consider suggestions presented during the workshop organised by the Health Directorate of the European Commission on 12 March 2019⁵. Finally, we will take into consideration the innovative vision of a unified approach to big data and extreme-scale computing⁶.

¹ W. C. Chew and L. J. Jiang: *Overview of Large-Scale Computing: The Past, the Present, and the Future*, in Proceedings of the IEEE, vol. 101, no. 2, pp. 227-241, Feb. 2013.

doi: 10.1109/JPROC.2012.2220311

² M. Kasztelnik, E. Coto, M. Bubak, M. Malawski, P. Nowakowski, J. Arenas, A. Saglimbeni, D. Testi, A. F. Frangi: *Support for Taverna Workflows in the VPH-Share Cloud Platform*, Computer Methods and Programs in Biomedicine, 146, 37-46, 2017

³ P. Nowakowski, M. Bubak, T. Bartyński, T. Gubała, D. Harężlak, M. Kasztelnik, M. Malawski, J. Meizner: *Cloud computing infrastructure for the VPH community*, Journal of Computational Science, 24, 169-179, 2018

⁴ Marc D. Wilkinson et al.: *The FAIR Guiding Principles for scientific data management and stewardship*, Scientific Data, <https://doi.org/10.1038/sdata.2016.18>

⁵ A. Albeyatti, J. Armitage, L. Augello, et al.: *Towards a European health research and innovation cloud (HRIC)*, to be published in Genome Medicine.

⁶ M. Asch, Terry Moore, Rosa M. Badia, et al.: *Big data and extreme-scale computing: Pathways to Convergence - Toward a shaping strategy for a future software and data ecosystem for scientific inquiry*, The International Journal of High Performance Computing Applications- IJHPCA 32(4): 435-479, 2018, doi: 10.1177/1094342018778123

5. CONCLUSION

This deliverable presents in details the High Performance Computing (HPC) resources available to members of the PRIMAGE consortium. Leveraging such resources is inevitable in order to achieve the project's main goals, especially in terms of reaching adequate spatial and temporal scales/resolutions for multi-scale simulations envisaged in the scope of the project.

The document presents the Prometheus cluster, a supercomputer managed by ACC Cyfronet AGH (Kraków, Poland) in the framework of the PLGrid Infrastructure for science. Subsequent chapters present the procedure of obtaining a PLGrid account, required to perform computations on Prometheus, and store data there, and also how PRIMAGE is going to request an adequate data/CPU grant from PLGrid to be able to perform the required large-scale computations.

The remaining chapters present methods of using Prometheus and associated software (tools, services) in order to efficiently perform simulations and other computations in this fair-share, multi-user environment. Several access options are presented and PRIMAGE users will have many opportunities to engage with Cyfronet team in order to put Prometheus' resources to their use.

The Cyfronet team is open to any form of collaboration with PRIMAGE scientists, willing to use HPC resources to achieve the project's goals or wishing to integrate their own IT solutions with PLGrid's services in order to provide HPC power to their users.