

# I/O Access Pattern Detection of HPC Workloads at Runtime

Jean Luca Bez<sup>1</sup>, Gessica F. M. Azevedo<sup>1</sup>, Francieli Zanon Boito<sup>2</sup>, Ramon Nou<sup>3</sup>, Alberto Miranda<sup>3</sup>,  
Toni Cortes<sup>3,4</sup>, and Philippe O. A. Navaux<sup>1</sup>

<sup>1</sup>Institute of Informatics, Federal University of Rio Grande do Sul (UFRGS) — Porto Alegre, Brazil

<sup>2</sup>Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LIG, 38000 Grenoble, France

<sup>3</sup>Barcelona Supercomputing Center (BSC) — Barcelona, Spain

<sup>4</sup>Universitat Politècnica de Catalunya — Barcelona, Spain

{jean.bez,gessica.azevedo,navaux}@inf.ufrgs.br,  
francieli.zanon-boito@inria.fr,{ramon.nou,alberto.miranda}@bsc.es,toni@ac.upc.edu

## Abstract

*In this paper, we seek to guide optimization and adaptation techniques according to the applications' I/O access pattern classified at runtime. We evaluate three machine learning techniques to detect the I/O access pattern of HPC applications: decision trees, random forests, and neural networks. We seek to detect the access pattern by using file-level metrics as seen by the I/O nodes. We applied these three detection strategies in a case study in which the correct detection of the current access pattern is paramount to adjust a parameter of an I/O scheduling algorithm. We show that such approaches accurately classify the access pattern, regarding file layout and spatiality of accesses with up to 99% precision. The access patterns selected as classes are the most common ones used by the HPC I/O community and by I/O benchmarking tools. Finally, after applying these approaches to the TWINS case study, we observe the tuning the mechanism achieve 99% performance for an Oracle solution.*