

HPC Code Review Club @BSC

Objectives

Peer code review is a good practice in code development. Here at BSC, each department creates some code to run jobs in a High-Performance Computing (HPC) setting. The HPC code review club will meet monthly to review HPC related code created at BSC. In each meeting, someone from one of the departments will present a code connected to a project. The presentation and conversation will include a description of the application and how it works, nodes being used, how the code is measured for optimization, and questions the person or team might have on what approach to follow or how to improve their code.

Abstract: Disciplines like bioinformatics, which address the computational needs of life sciences, depend on software that is not typically designed with HPC in mind but that often require extensive calculations. Another characteristic of bioinformatics work is that it typically involves recipes of steps involving software tools of quite diverse nature. The focus of bioinformatics developers, beyond solving specific scientific questions, has been on reproducibility/reusability; thus, there has been a lot of interest in workflows and workflow managers, with almost every self-respecting bioinformatics lab developing their own workflow manager at some point. Here I will present Rbbt (ruby bioinformatics toolkit), a framework for bioinformatics that features its own workflow manager as well as many tools for the bioinformatics developer. I will focus on the aspects regarding workflow HPC integration as well as internal tools for parallelization or relevant for performance.

Short bio: Miguel Vazquez holds a PhD on bioinformatics and has worked in this field for almost 20 years, with a special focus on cancer genomics and text-mining. He arrived at bioinformatics after having his mind blown away by a data mining course he took in the University of Texas at Austing back in 2001 and later being told that biology had plenty of data problems to work on. He very early got interested in reproducibility/reusability which led him to start hoarding code into his own bag of tricks which later became Rbbt. Bioinformatics had a need for better software design than anything else, he soon found out, and this has occupied him more than the data mining itself--such is life.

Speakers

Speaker: Miguel Vazquez, Life Sciences, BSC

Host: Mercè Crosas, Head of Computational Social Sciences, BSC

Source URL (retrieved on 3 Dic 2023 - 08:22): <https://www.bsc.es/es/news/events/hpc-code-review-club-bsc-0>