

21st International Conference on Computing in High Energy and Nuclear Physics (CHEP2015)



Contribution ID: 527

Type: oral presentation

Agile Research - Strengthening Reproducibility in Collaborative Data Analysis Projects

Monday, April 13, 2015 5:45 PM (15 minutes)

Reproducibility of results is a fundamental quality of scientific research. However, as data analyses become more and more complex and research is increasingly carried out by larger and larger teams, it becomes a challenge to keep up this standard. The decomposition of complex problems into tasks that can be effectively distributed over a team in a reproducible manner becomes nontrivial.

Overcoming these obstacles requires a shift in both management methodology as well as supporting technology.

The LHCb collaboration is experimenting with different methods and technologies to attack such challenges. In this talk we present a language and thinking framework for laying out data analysis projects. We show how this framework can be supported by specific tools and services that allow teams of researchers to achieve high quality results in a distributed environment. Those methodologies are based on so called agile development approaches that have been adopted very successfully in industry. We show how the approach has been adapted to HEP data analysis projects and report on experiences gathered on a pilot project.

Primary authors: USTYUZHANIN, Andrey (ITEP Institute for Theoretical and Experimental Physics (RU)); LINN, Christian Peter (CERN); NEUBERT, Sebastian (CERN); KARBACH, Till Moritz (CERN)

Presenter: NEUBERT, Sebastian (CERN)

Session Classification: Track 4 Session

Track Classification: Track4: Middleware, software development and tools, experiment frameworks, tools for distributed computing