

The DataForge framework for data acquisition and analysis

Alexander Nozik

INR RAS

Abstract

Recently, the software solutions became a major problem in particle physics. Most software packages for it were developed in 1990's or even 1980's, solving some specific problems for specific theoretical or experimental tasks. While mathematical solutions themselves do not strongly change over time and could not be improved very much, the software engineering made a great progress since that time. Nowadays particle physicists waste a lot of effort due to inefficient software frameworks and lack of automation.

The DataForge introduces some trends from modern software development into scientific software for particle physics research and physics in general:

1. Data analysis as a metadata processing. Everything aside from data itself is treated as a metadata. The analysis process is constructed depending on this metadata automatically.
2. Context encapsulation. Environment variables replaced by independent context variables. This allows easily making any analysis process parallel.
3. Cross-platform implementation. The framework is written in Java language and could be easily installed on any operating system supporting JVM.
4. Modular design. The plugin system allows using only the modules needed for precise task and nothing extra.

The prototype for DataForge framework is being successfully tested at "Troitsk nu-mass" experiment in search for neutrino mass.