

MPD DATA LAB: TOWARDS THE MODERN DATA ANALYSIS FRAMEWORK FOR THE MPD EXPERIMENT

J. Buša, Jr., A. Bychkov, S. Hnatič, A. Krylov, V. Krylov,
O. Rogachevsky*

Joint Institute for Nuclear Research, Dubna

MPDRoot is an off-line software framework for simulation, reconstruction, and physical analysis of the simulated or experimental data for the MPD experiment at the NICA collider. The experiment is projected to run for a few decades and to obtain $\sim 10^8$ events of heavy-ion collisions, collecting the data for physics analysis at the 100 PB scale. For overall experiment success, it is imperative to have state-of-the-art data analysis software, which integrates the best of the latest technologies available, while adhering to time-proven, most effective development methodologies.

We introduce the MPD Data Lab — the technological integration of Acceptance Test Driven Development and Rapid Development concepts into the MPDRoot framework. At the beginning, we standardized the existing codebase by designing and writing API. This was a necessary step to be able to plug in the external diagnostic software entities and to make the in-depth comparison of different realizations of the reconstruction modules possible. The logic of the diagnostics is encapsulated into the separate controller — the QA Engine, while its visualization is provided by JupyterLab framework. We show how full integration of MPDRoot’s libraries into JupyterLab enables us to use the power of rapid development provided by JupyterLab technology to enhance productivity by fast prototyping of MPDRoot’s algorithms. The combination of these technologies together with the existing development environment forms a software complex, providing means to accomplish the long-term strategic objectives: competent software development with reliable quality control and algorithm innovation.

MPDRoot — это офлайн-фреймворк для моделирования, реконструкции и физического анализа моделированных или экспериментальных данных эксперимента MPD на коллайдере NICA. Эксперимент продлится несколько десятилетий и позволит получить $\sim 10^8$ событий столкновений тяжелых ионов с масштабом данных 100 ПБ для физического анализа. В целях успешного эксперимента крайне важно иметь современное программное обеспечение для анализа данных, которое интегрирует в себя лучшие из новейших технологий, его разработка осуществляется наиболее эффективными и временем проверенными методологиями.

Представлена MPD Data Lab — технологическая интеграция концепций разработки через приемочное тестирование и быстрое прототипирование в фрейм-

* E-mail: hnatics@jinr.ru

ворк MPDRoot. Вначале была стандартизирована имеющаяся кодовая база путем проектирования и написания API. Это был необходимый шаг для возможности подключения внешних диагностических программ и детального сравнения различных реализаций модулей реконструкции. Логика диагностики инкапсулирована в отдельный контроллер — QA Engine, а ее визуализация обеспечивается фреймворком JupyterLab. Показано, как полная интеграция библиотек MPDRoot в JupyterLab позволяет использовать возможности быстрой разработки, обеспечиваемые технологией JupyterLab, для повышения производительности за счет быстрого прототипирования алгоритмов MPDRoot. Комбинация этих технологий вместе с существующей средой разработки образует программный комплекс, обеспечивающий средства для достижения долгосрочных стратегических целей — компетентной разработки программного обеспечения с надежным контролем качества и инноваций в алгоритмах.

PACS: 07.05.Rm; 12.38.Mh