

USAGE OF APACHE CASSANDRA FOR PROTOTYPING THE EVENT METADATA SYSTEM OF THE NICA EXPERIMENTS

A. Degtyarev^a, K. Gertsenberger^b, P. Klimai^{a, c, 1}

^a Moscow Institute of Physics and Technology (National Research University),
Dolgoprudny, Russia

^b Joint Institute for Nuclear Research, Dubna

^c Institute for Nuclear Research, Russian Academy of Sciences, Moscow

Nuclotron-based Ion Collider Facility (NICA) experiments at the Joint Institute for Nuclear Research have already generated substantial volumes of simulated and experimental (BM@N facility) data, and it is expected that the overall number of the collected events will increase from the current value of hundreds of millions to several billion events per year. An information system based on the event database (Event Catalogue) is required to store and index information on the events obtained at the NICA facilities, enabling quickly searching and using only necessary physics events based on various criteria for further processing and analysis. The paper describes the development of the Event Catalogue based on Apache Cassandra database. It is argued that Cassandra provides acceptable performance and is a possible candidate for creating the Event Catalogue in NICA and similar experiments. Principles of creating the database objects, in particular, selecting the tables' primary keys, are formulated.

Эксперименты на ионном коллайдерном комплексе на базе нуклотрона (NICA) в Объединенном институте ядерных исследований уже произвели значительный объем смоделированных и экспериментальных (установка BM@N) данных, и ожидается, что общее количество полученных событий увеличится с текущего уровня порядка сотен миллионов до нескольких миллиардов событий в год. Информационная система на основе базы данных событий (каталога событий) необходима для хранения и индексирования информации о событиях, полученных на установках комплекса NICA, что позволит осуществлять быстрый поиск и использовать только необходимые физические события, отбирая их по различным критериям, для дальнейшей обработки и анализа. В статье описывается разработка каталога событий на основе базы данных Apache Cassandra. Показано, что Cassandra обеспечивает приемлемую производительность и является возможным кандидатом для создания каталога событий в экспериментах на NICA и других подобных экспериментах. Сформулированы принципы создания объектов представленной базы данных, в частности выбор первичных ключей таблиц.

PACS: 07.05.-t; 07.05.Bx; 07.05.Kf

Received on January 26, 2022.

¹E-mail: pklimai@gmail.com