Management of the Monte Carlo simulation process in the MPD experiment on the NICA cluster *

Łukasz Sawicki[1,†], Jakub Zieliński[1,], Krzysztof Dygnarowicz[2,]

[1]Faculty of Physics Warsaw University of Technology
[2]Faculty of Electronics and Information Technology Warsaw University of Technology

This paper shows the development of software aimed at automatization of the process of Monte-Carlo simulation of event and particle generation and propagation in Au-Au collisions in the MPD detector. It contains information about the tools used and a description of the software functioning.

PACS numbers: 07.05.Kf

1. Introduction

During an internship at the Joint Institute for Nuclear Research [1] at the Laboratory of High Energy Physics we created software for the need of management of the generated data. This data was simulations of heavy-ion collisions inside the MPD (Multi-Purpose Detector) located at the NICA [2] (Nuclotron-based Ion Collider fAcility).
Major problem during the practice was to make process of event generation and simulation of particles as automated as possible. We were running simulations and event generations for Au-Au collision in NICA collider range of energy. It meant having to solve problems with marking, storing and using all the generated data.
This problem can be divided into four separate steps. Firstly, event generation, secondly event reconstruction using macros, then checking correctness of received files and future data analysis.

* Presented at the IV NICA Days 2019 Conference associated with the V Slow Control Warsaw 2019, Warsaw, Poland, October 21-25, 2019
† Corresponding author: l.saw99@outlook.com

2. Software

For this purpose we had NICA computing cluster with about 7PB disc space to use at our disposal. Responsible for event generation was UrQMD 3.4 [3] (Ultra relativistic Quantum Molecular Dynamics) Monte Carlo simulation package. Using it we generated most of our gathered data. In order to compare simulation models we generated a small amount of data by using vHLLE package (A 3+1 dimensional viscous hydrodynamic code for relativistic heavy ion collisions). Event generations were executed for Au-Au collision at energy from 4 GeV to 11 GeV. On the cluster we had MPDRoot framework available which contains MPD detector models. By using this framework we were able to simulate transition of particles created in the collision through the detector and we were able to see detector's response to this event.

2.1. Software operating principle

For the purpose of realisation of each step we made a few BASH scripts and programs written in C++. NICA cluster shares a folder named "data", its size is about 7PB without quota limit. Thanks to the access to this folder we were able to easily collect data coming from the event generation and detector response into one place. The full path to this folder is */eos/nica/mpd/data*. In this location every simulation model and each energy have their separate folder. Simulations for UrQMD model were carried out for energies at center of mass frame $\sqrt{s_{NN}}$ = 4 Gev, 7 GeV, 9 GeV and 11 GeV. For vHLLE model we were making simulations only for 7.7 GeV collisions, due to it being the only possible energy from NICA collider's energy range we can choose from. Event generation and detector response simulation process was working according to this algorithm:

- In program's config file set number of events in one generation
- Run program with an argument of energy and number of jobs to run (in vHLLE model case we don't pass energy as a argument, because it is set in config file)
- Program makes in data folder new folder with set energy and simulation model (if this name already exist, this step is skipped)
- Inside that a folder for every running job, program is creating folder with name of running job. It is unique JOBID number given by queue system working on the cluster
- Necessary macros and folders are placed in the created folder

- Program runs event generation
- Output file from every generation is passing through runMC.C macro which is simulating detector's response to generated event and through reco.C macro which can reconstruct event from information given by previous macro. Final file ready for analysis is mpddst.root

User can only set energy and number of events for one simulation in the config file. The rest of the parameters is the same in every simulation. User can modify them by changing source code, but this need to change names of making folders in code too. This is because it was our first attempt to create program for running simulations in one command. We used a lot of same parameters on purpose in order to collect as much data as possible for analysis. Source code and installation instruction are available on GitHub. Queue system on NICA cluster allows to run 200 jobs parallelly per user, which limited the possibility of producing data.

2.2. Data correctness checking

Next stage is checking correctness of generated files. Due to the complexity of the content of each file, file validation has been limited to checking if a file with the correct name exists and if its size on the disk is large enough. Correct size for every energy and number of events in one simulation is calculated by checking the size of a few generated files with various energies and number of events in one simulation. Based on this, we can determine what the right size of each file should be. Program knows number of events in each simulations, because it is set in config file(inputfile). It turns out that the file size is practically constant, and damaged files are at least a few precent smaller than the correct size. Following to this, the program is able to determine which simulations ended correctly, and in the case of incorrect ones, it is able to determine in which place the error occured. Paths to valid files are placed in a txt file, which is used in further collision analysis.

3. Summary

During the 2-month internships, most of the time was devoted to learning how to use the cluster. First weeks were spent on writing the program and fixing its errors. The last 3 weeks of practice were devoted to data generation and analysis. During this time, we successfully reconstructed and analysed 2,120,000 events. Together they had a size of 60 TB. Numbers of the reconstructed events for each energy are shown in the Table 1.

A typical job is 200 events for one simulation, this value was chosen because, by trial and error, it was noticed that a larger number of events

UrQMD				vHLLE
4GeV	7GeV	9GeV	11GeV	7.7GeV
reconstructed events				
203600	448346	593851	721652	151393

Table 1. Number of reconstructed events for each energy

results in more simulations ending with errors, and a smaller number of events reduces the amount of data produced during the day. 200 events were accepted as the most optimal number. Generation of so many events in one job and their reconstruction took about 24h. Over 2 million events in the cluster were generated in 3 weeks. Using your own laptop for this would take a little over 10 years.

All data and images come from the presentation given at the conference. You can download the software from GitHub https://github.com/Lukaszz99/work

REFERENCES

[1] JINR, "http://www.jinr.ru/about-en/." Accesed: 28-08-2019.

[2] A. Sissakian, A. Sorin, N. collaboration, *et al.*, "The nuclotron-based ion collider facility (nica) at jinr: new prospects for heavy ion collisions and spin physics," *Journal of Physics G: Nuclear and Particle Physics*, vol. 36, no. 6, p. 064069, 2009.

[3] S. A. Bass, M. Belkacem, M. Bleicher, M. Brandstetter, L. Bravina, C. Ernst, L. Gerland, M. Hofmann, S. Hofmann, J. Konopka, *et al.*, "Microscopic models for ultrarelativistic heavy ion collisions," *Progress in Particle and Nuclear Physics*, vol. 41, pp. 255–369, 1998.