

# Modernized supercomputer "Govorun" as a computing environment for BM@N data processing

The studies in the given direction were supported by the RFBR grant ("Megascience - NICA") № 18-02-40101.

## Zuev M.

on behalf of HybriLIT team

Joint Institute for Nuclear Research, Dubna, Russia

5th Collaboration Meeting of the BM@N Experiment at the NICA Facility 20-21 April 2020

### **Off-Line Computer Complex for the NICA Project**



Technologies for storing, processing and analyzing experimental data in megasience projects



DIRAC

### Modernization of the supercomputer "Govorun"



First modification, 2018:Total peak performance:1 PFlops for single precision500 TFlops for double precision







Number of computing cores: 1440 4224

> Capacity of UDSS: 120 TB 288 TB

Data input/output rate 56 GB/s 300 GB/s



#10 in Top50

Second modification, 2019:Total peak performance:1.7 PFlops for single precision860 TFlops for double precision

## Modernized supercomputer "Govorun" for the MPD experiment

The hierarchical structure of working with data on the "Govorun" supercomputer.



Events of the MPD experiment are simulated and reconstructed on ultrafast data storage system under the **FS Lustre** management with a subsequent transfer to semi-cold storages (**FS ZFS, EOS**) and to the tape library for long-term storage. **1.5 million events** were generated for the MPD experiment using the hierarchical structure of working with data. The acceleration of calculations on the upgraded supercomputer in comparison with the previous configuration was **1.45 times**.





## LIT infrastructure for the NICA megaproject



## JINR computing resources integration for the MPD experiment



The computing resources of the JINR Multifunctional Information and Computing Complex were combined using DIRAC Interware: **Tier1/Tier2**, supercomputer **"Govorun"** and **storage** resources. JINR and Member-States **cloud resources** were tested and are ready to accept jobs. **NICA Cluster** is the next on the list. More than **120 000** jobs were performed on the Tier1/Tier2 components and the supercomputer "Govorun" using the DIRAC platform in the framework of Monte-Carlo data simulation for the MPD experiment.



## Information systems for online and offline data processing on the NICA complex

- 1. Information system for accounting for the geometry of detectors (geometric database): storage and processing of information about the composition and structure of detectors.
- 2. Database of system states: storage, processing and the possibility of using various operating parameters and modes of devices and experiment detectors in algorithms for reconstruction and analysis of particle collision events.
- 3. Online logging system: recording data on events, the status of systems and the operating conditions of the detectors.
- 4. Configuration information system: storage and provision of data about the configuration of the subsystems of the experiment when collecting data from detectors in online mode.
- 5. Metadata system of physical events: managing information about a unique number of events, saving links to events occurring in an experiment, triggers triggered by online processing, a list of recovered particles, and other information.

#### **The studies in the given direction were supported by the RFBR grant ("Megascience - NICA") № 18-02-40125.** Gertsenberger K. et al, Development of Information Systems for Online and Offline Data Processing for the Experimental Setups of the NICA Complex

## ML Tracks Reconstruction in BM@N Experiment



Machine learning algorithms bring a lot of potential to the tracks reconstruction problem due to their capability to learn effective representations of high-dimensional data through training, and to parallelize on HPC architectures.



#### Simulation data



Input data for the first step algorithm were simulated by GEANT in MPDRoot framework for the real BM@N configuration.

G.Ososkov et al, Deep learning for tracking

Statistics of using all components of the supercomputer for the NICA megaproject

Other computations [ 34% ] Computations for the NICA megaproject [ 66% ] 72% MPD

28% BM@N

## BM@N collaboration on supercomputer "Govorun"

NICA

BM@N

- 19 users from JINR and other organizations
- an individual group on supercomputer "Govorun"
- an individual computational queue (384 CPU, ~800 Gb RAM) with the possibility of enlargement
- access to experiment data at the EOS repository at JINR
- access to FS **ZFS** (home directory, 2 TB), Luster (computations, 256 TB) and EOS (storage, 300 TB)

		Registration Form		
hYBRI	Platform "HybriLIT" ~ SU	5		Q
		*Full name		and the
1 1 1 2		*Laboratory	Select laboratory ~	
5		*Country		No.
	education	*E-mail address		POLID
0		Phone number		h 3
and the	Colores No.	*Funding source	Select funding source	
A A A A A A A A A A A A A A A A A A A			You can find the Full names list of Funding source here	
ß	Ţ	Project from the funding source		
REGISTRATION	SERVICES	*Login (latin)		
			Max 8 symbols	
Heterogeneous platform "Hybr		*Name of the scientific project		
"HybriLIT" Heterogeneous Platform is a part of the Multi			÷	
Technologies of JINR. Heterogeneous platform consists of "		*Summary of the scientific project		
"Govorun" Supercomputer is a two-component system tha				
Skylake processors), and GPU-component based on NVIDIA				
Education and testing polygon is based on heterogeneous				100
computation using various computing architectures such a				
also provides a possibility to hold tutorials in parallel progra				
		*Choose resource categories which are planned to be used		
		🗆 GPU		
		□ Intel coprocessorWrite down a list of software packages that are planned to be used on the cluster		



#### Software coordinator