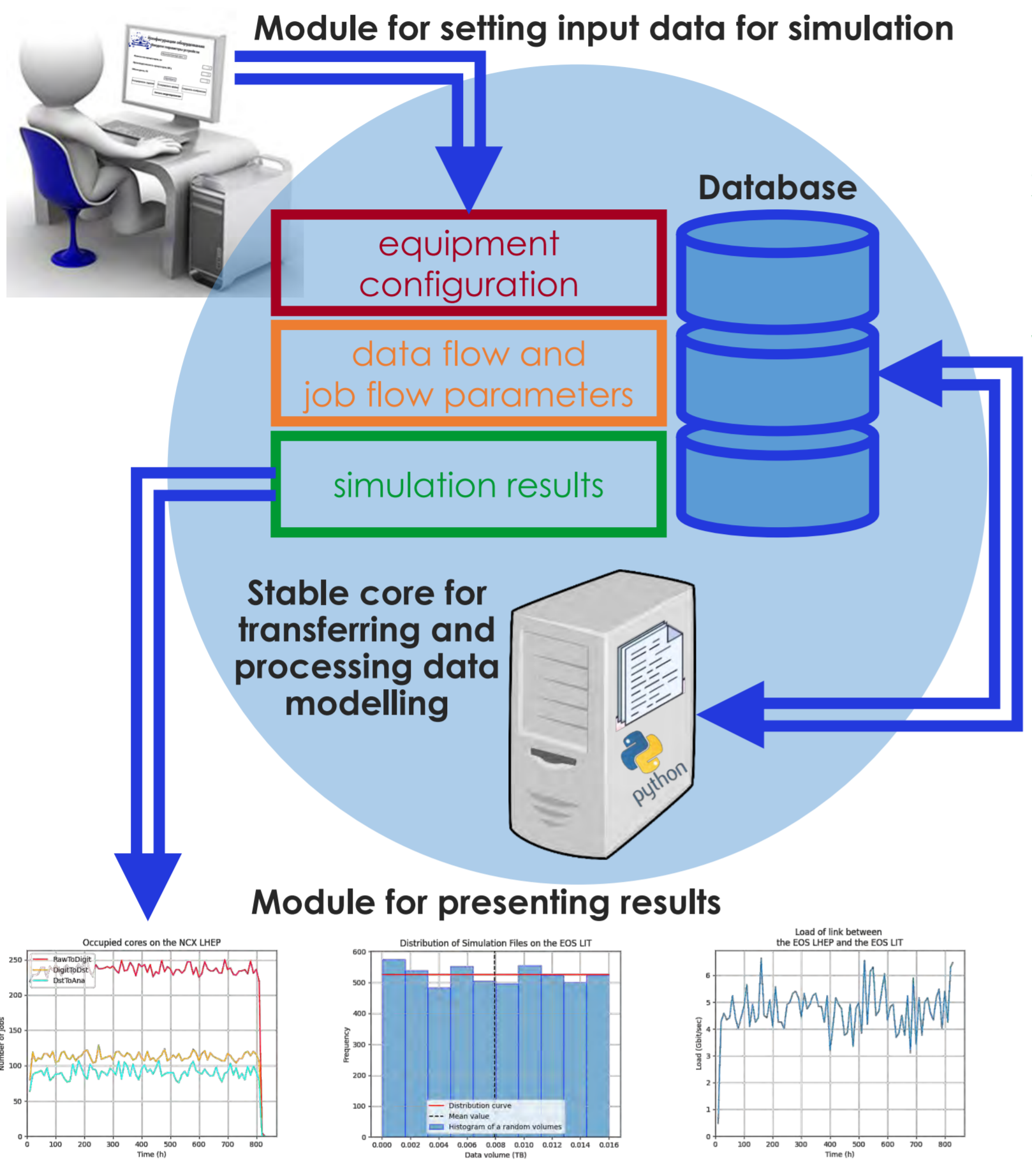


# Simulation of computing infrastructures for data acquisition, storage and processing within the NICA experiments

D. Priakhina, V. Korenkov, V. Trofimov, K. Gertsenberger \*  
 pryahinad@jinr.ru

\* Collaboration with Veksler and Baldin Laboratory of High Energy Physics

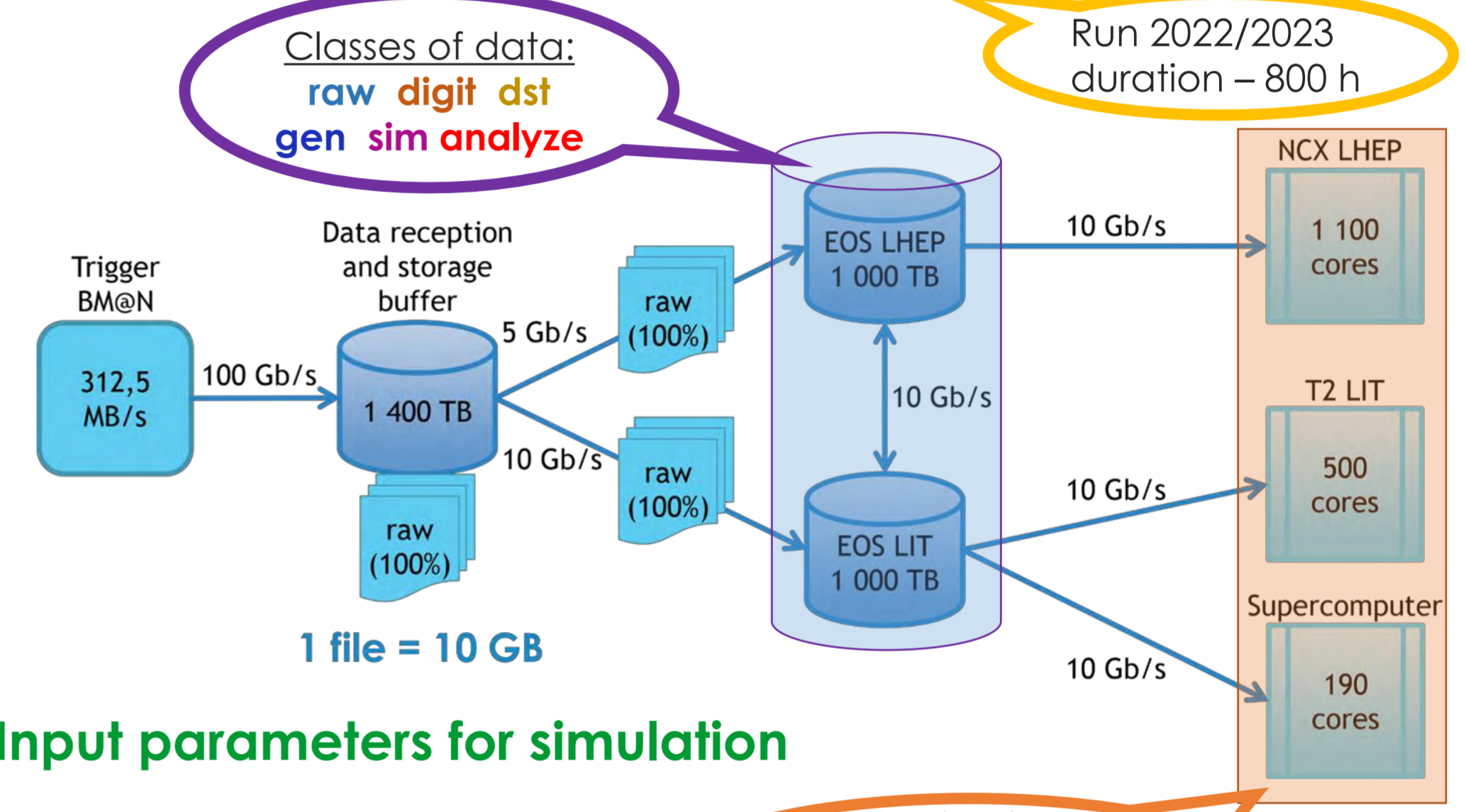
## Simulation software complex



## Goals of simulating computing infrastructures of the NICA experiments

1. Assessment of the current requirements for data acquisition, storage and processing.
2. Obtaining predictive values for necessary resources within the perspective of the development of the BM@N computing infrastructure.
3. Searching for optimal equipment parameters that will ensure data processing depending on the specified requirements.

## Computing infrastructure of the BM@N experiment



## Input parameters for simulation

Running jobs:  
 RawToDigit DigitToDst  
 GenToSim SimToDst  
 DstToAna

Characteristics of job classes

Class	Description	Avg. input file size (GB)	Avg. output file size (GB)	Job execution time (s)	Number of jobs	Job start frequency (s)
RawToDigit	Converting raw data to digit data	10	0,4	20 000 (NCX, T2)	32 000	90
DigitToDst	Converting digit data to reconstructed data	0,4	0,4	8 600 (NCX, T2)	32 000	90
GenToSim	Converting event data received by the event generator into simulation data	0,084	8	5 000 (Sup.) 15 000 (NCX, T2)	5 250	549
SimToDst	Converting simulation data to reconstructed data	8	0,4	12 000 (Sup.) 35 000 (NCX, T2)	5 250	549
DstToAna	Physical analysis of reconstructed data	0,4	0,05	3 000 (Sup.) 10 000 (NCX, T2)	37 250	77

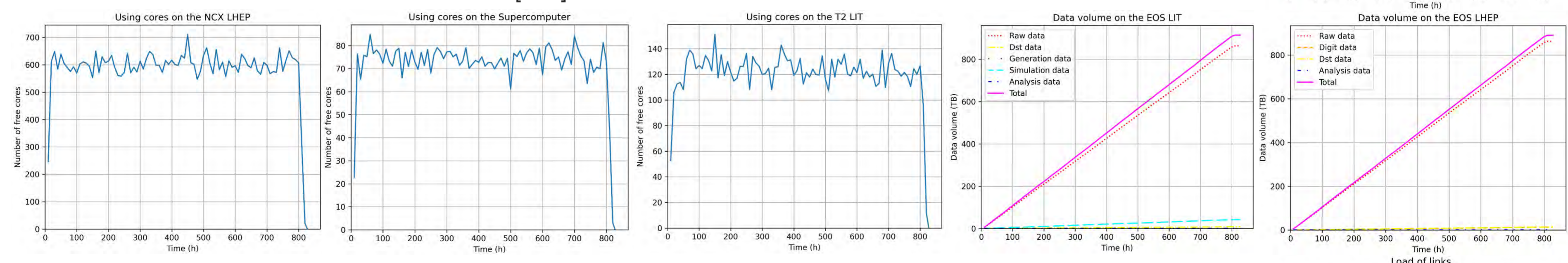
Planned distribution of computing power

Class	Number of cores that are allocated on computing components		
	NCX LHEP	T2 LIT	Sup.
RawToDigit	350	-	-
DigitToDst	150	-	-
GenToSim	-	200	90
SimToDst	-	200	90
DstToAna	600	100	10

## Simulation results for the BM@N computing infrastructure (winter Run 2022/2023)

1. About 850 TB of raw data (60% of online storage) will be accumulated.
2. Approximately 98% of all jobs will be completed during the experiment.
3. All jobs will be finally completed 30 hours after the end of the experiment.
4. The allocated computing resources are sufficient to execute 99% of jobs.
5. Most of the computing resources are not used.
6. The EOSes will be filled by 90%, it will not be enough with the increase in data flows and jobs.
7. The available bandwidth of communication links is sufficient.

More detailed information is available in [1-3].



## Future plans

1. Verifying the software complex based on the results of BM@N Run 2022/2023.
2. Finding optimal equipment parameters that will ensure data processing.
3. Using probabilistic distributions of significant data acquisition processes.
4. Obtaining predictive values of necessary resources within the perspective of the development of the BM@N computing infrastructure for 2023-2030.
5. Using complex for the other NICA experiments.

## References

1. Priakhina D., Korenkov V., Trofimov V., Gertsenberger K. Simulation results of BM@N computing infrastructure // Physics of Particles and Nuclei Letters (accepted).
2. Priakhina D., Korenkov V., Gertsenberger K., Trofimov V. Simulation of Data Processing for the BM@N Experiment of the NICA Complex // CEUR Workshop Proceedings. 2021. V. 3041. P. 483.
3. Priakhina D., Trofimov V., Ososkov G., Gertsenberger K., Data center simulation for the BM@N experiment of the NICA project // AIP Conference Proceeding. 2021. V. 2377. P. 040007.

The computations were performed on the basis of the HybriLIT heterogeneous computing platform (MLIT, JINR, <http://hlit.jinr.ru/>).



MESHCHERYAKOV LABORATORY OF INFORMATION TECHNOLOGIES