

The XXIV International Scientific Conference of Young Scientists and Specialists

9 - 13 November 2020 JINR, Dubna

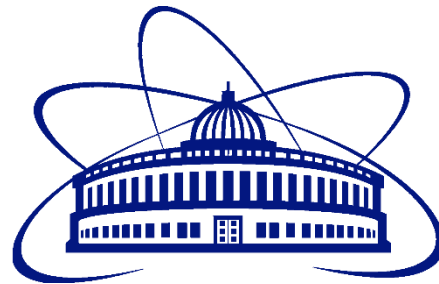
Data Center Simulation for the BM@N experiment of the NICA project

D. PRIAKHINA

V. TROFIMOV

G. OSOSKOV

K. GERTSENBERGER



Simulation of data storage and processing centers, both as from the BM@N detector, as for simulated particle collision events for comparison with the expected results of real storage and processing processes.

Probabilistic approach to simulate

- Representation of information processes as byte streams
- Using of probability distributions of significant data acquisition processes – the probabilities of loss of incoming information should be determined for different configurations of the data centers equipment

Simulation goal

Determine the hardware configuration that will ensure the operability of the data storage and processing system – takes into account hardware parameters and expected data flows and jobs.

The simulation software complex

- equipment parameters
- list of jobs for processing

Database

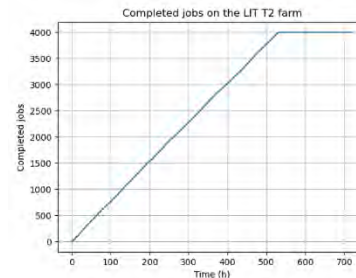
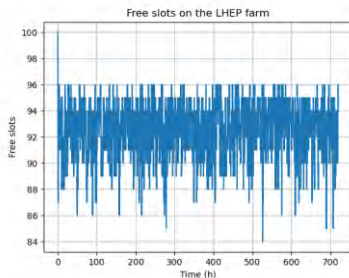
- simulation results

Module for setting of equipment configurations

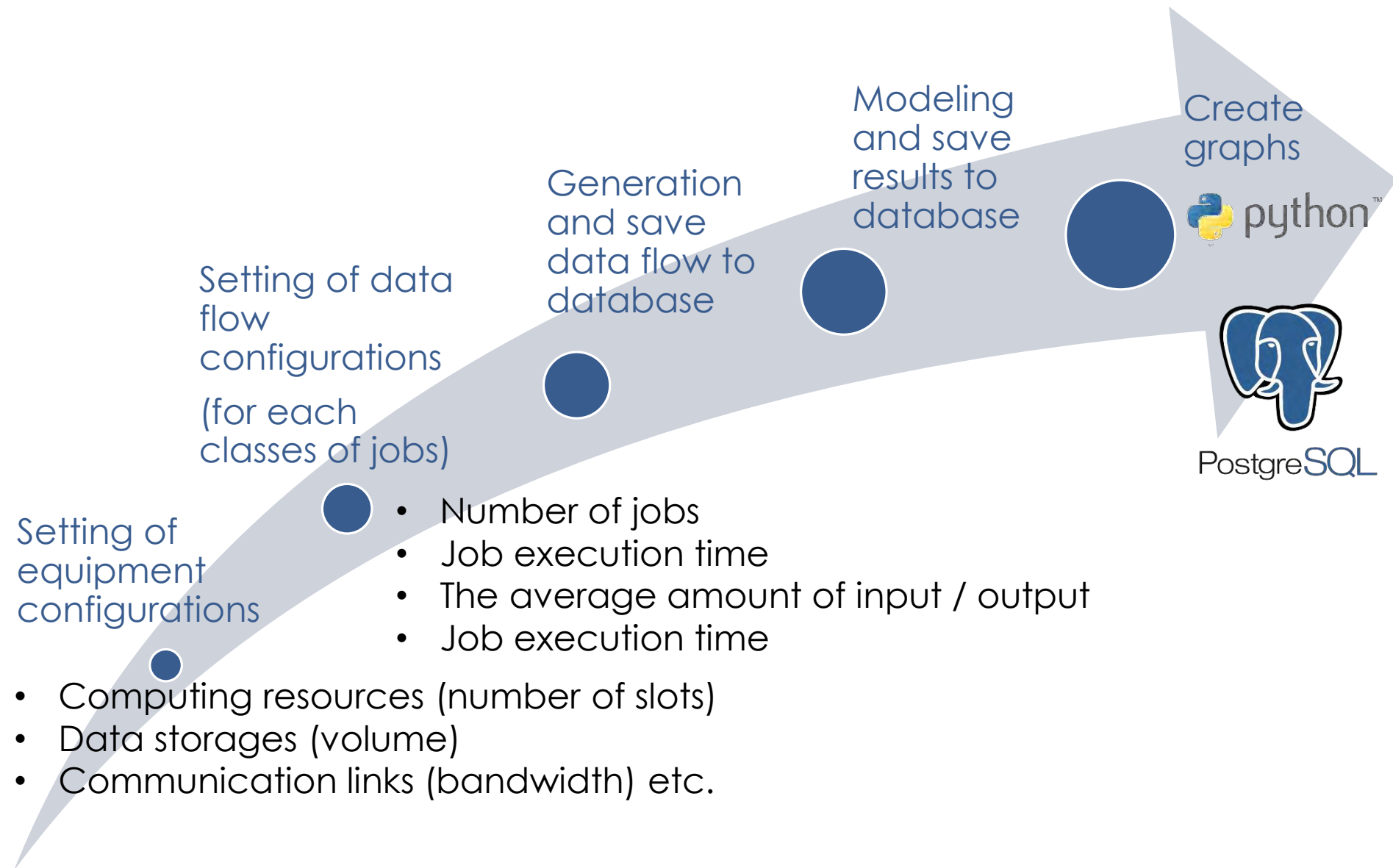
The software complex modules

Transfer and processing data simulation module

Module for presenting results



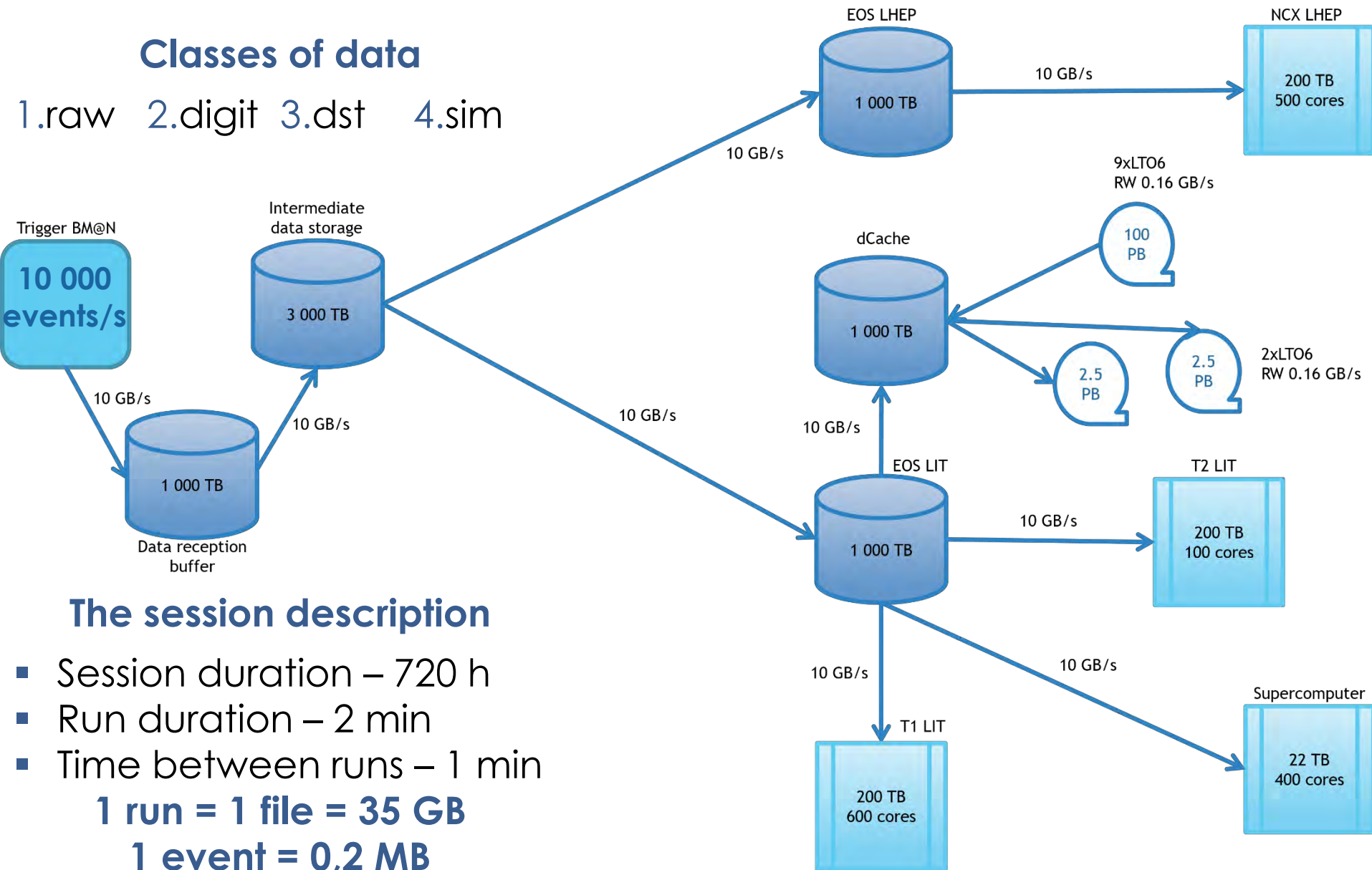
Stages of the software complex



The simulated structure

Classes of data

1.raw 2.digit 3.dst 4.sim



The session description

- Session duration – 720 h
- Run duration – 2 min
- Time between runs – 1 min

1 run = 1 file = 35 GB

1 event = 0,2 MB

Classes of jobs

No	Class	Event processing time on one processor (ms)	The average amount of input (GB)	Number of events in the file (1 file = 1 job)	Job execution time (s)	The average amount of output (GB)	Number of jobs
1	RawToDigit	150	35	175 000	26 250	1	10 000
2	DigitToDst	30	1	175 000	5 250	0,6	10 000
3	GenToSim	60	2	175 000	10 500	8	300
4	SimToDst	30	8	175 000	5 250	1	300
5	DstToAna	10	1	175 000	1 750	0,1	1 000

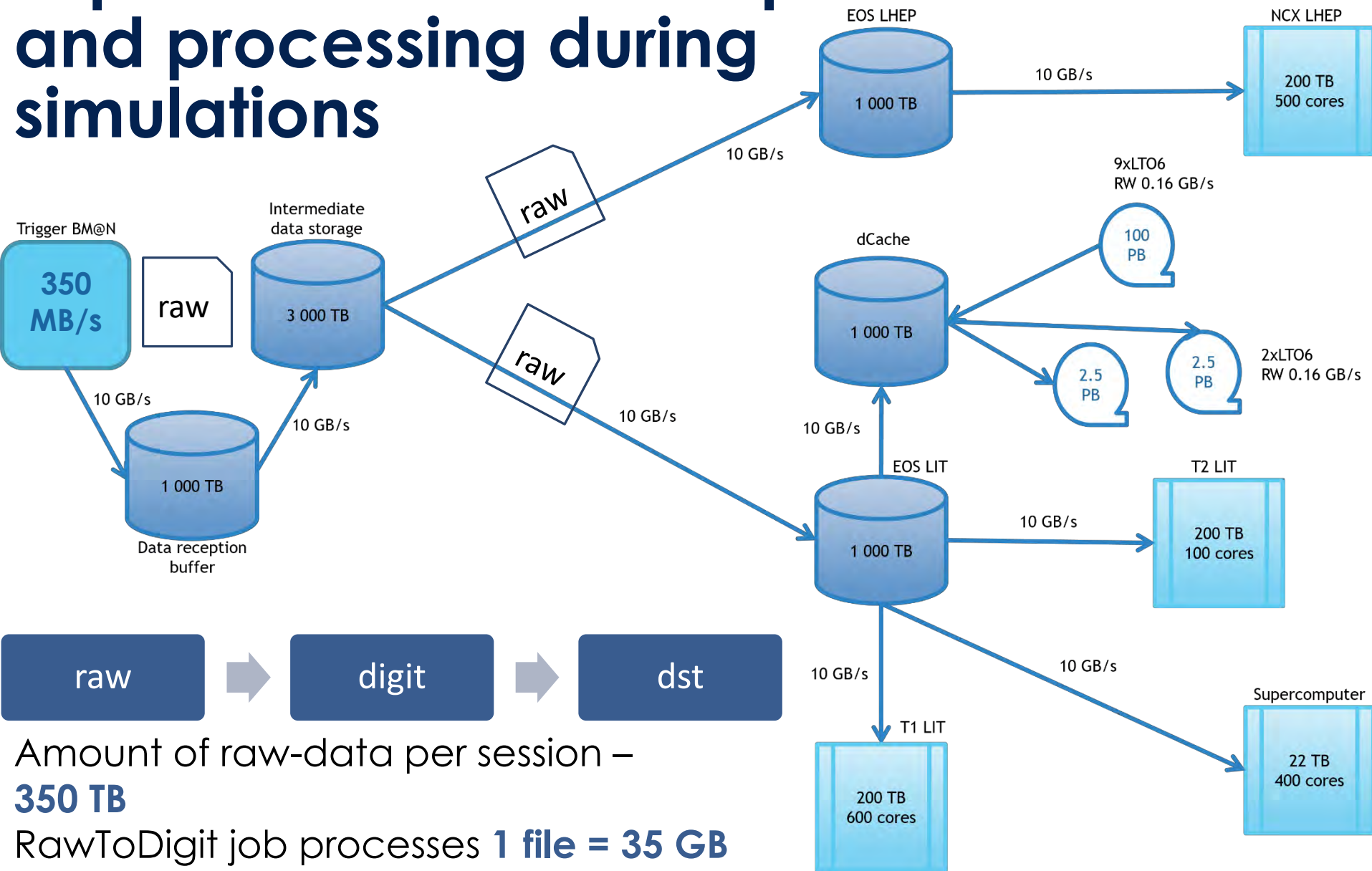
Experimental data processing



Model data processing



Experimental data acquisition and processing during simulations



Scenarios for executing jobs

№	Class	Location of the executing jobs / % of jobs	
		Scenario 1	Scenario 2
1	RawToDigit	NCX LHEP / 40% T2 LIT / 45% Supercomputer / 15%	NCX LHEP / 50% T2 LIT / 15% Supercomputer / 35%
2	DigitToDst	NCX LHEP / 40% T2 LIT / 45% Supercomputer / 15%	NCX LHEP / 50% T2 LIT / 15% Supercomputer / 35%

Results of Scenario 1

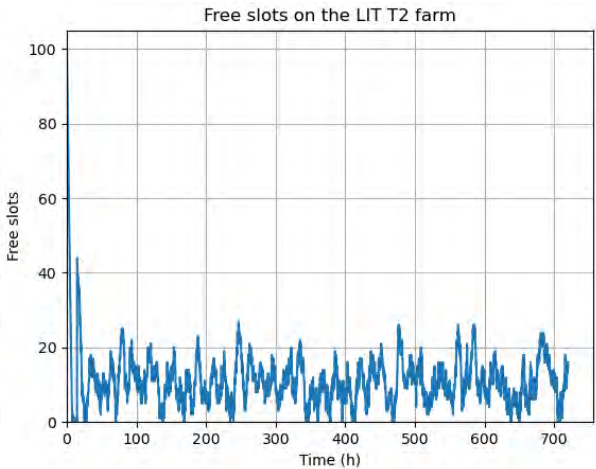
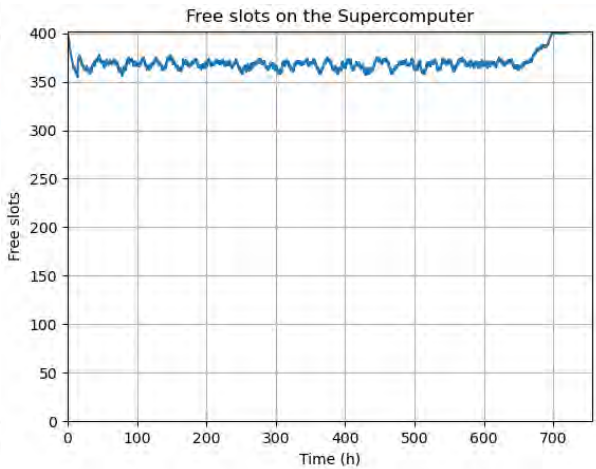
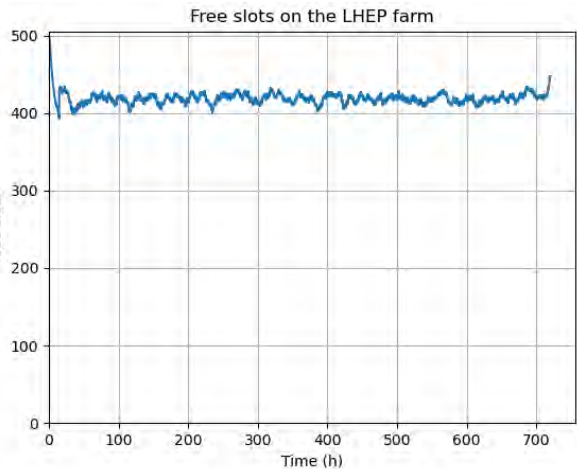
Total number

RawToDigit jobs – 10 000
DigitToDst jobs – 10 000

LHEP farm: 500 slots
RawToDigit – 4 000 (40%)
DigitToDst – 4 000 (40%)

Supercomputer: 400 slots
RawToDigit – 1 500 (15%)
DigitToDst – 1 500 (15%)

T2 LIT farm: 100 slots
RawToDigit – 4 500 (45%)
DigitToDst – 4 500 (45%)



- **400 slots are free**
- There are not jobs queues
- The farm is not fully loaded

- **350 slots are free**
- There are not jobs queues
- The Supercomputer is not fully loaded

- **The T2 LIT farm is fully loaded**
- There are jobs queues

We can process more jobs on the LHEP farm and Supercomputer

Solution: to redistribute the number of jobs across compute nodes of data center

Results of Scenario 2

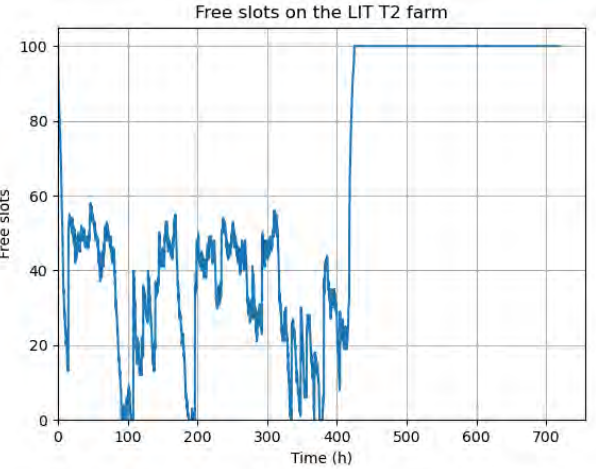
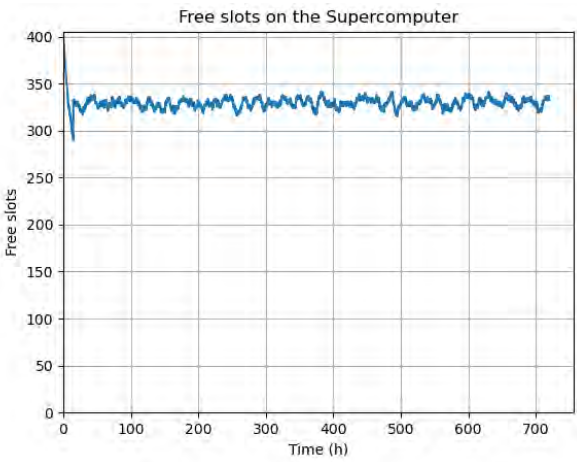
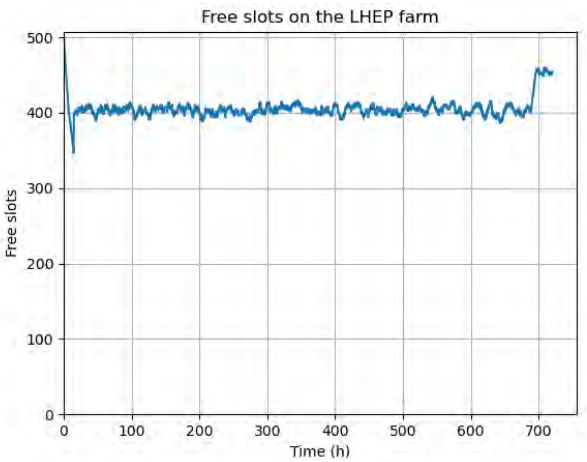
Total number

RawToDigit jobs – 10 000
DigitToDst jobs – 10 000

LHEP farm: 500 slots
RawToDigit – 5 000 (50%)
DigitToDst – 5 000 (50%)

Supercomputer: 400 slots
RawToDigit – 3 500 (35%)
DigitToDst – 3 500 (35%)

T2 LIT farm: 100 slots
RawToDigit – 4 500 (45%)
DigitToDst – 4 500 (45%)



- **350 slots are free**
- There are not jobs queues
- The farm is not fully loaded

- **250 slots are free**
- There are not jobs queues
- The Supercomputer is not fully loaded

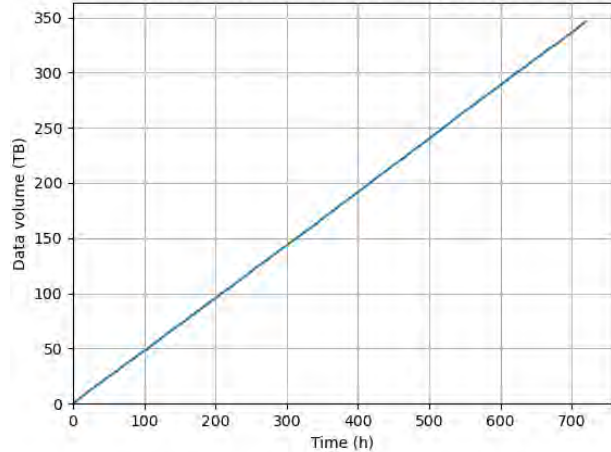
- **All jobs were processed in 400 hours**
- There are jobs queues

We can process more jobs on the LHEP farm and Supercomputer

The results require additional research

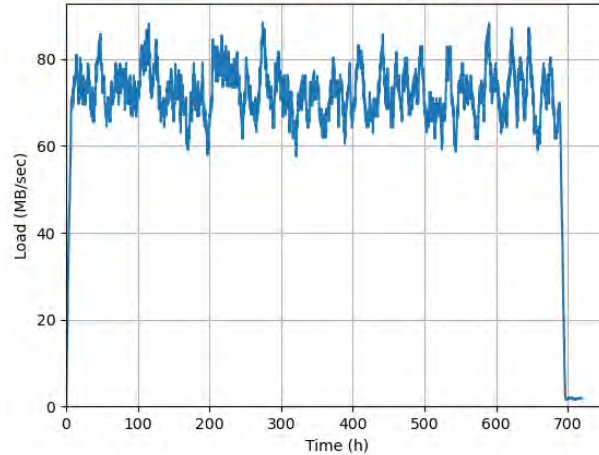
Total results

Data volume on the Intermediate data storage / EOS LHEP / EOS LIT



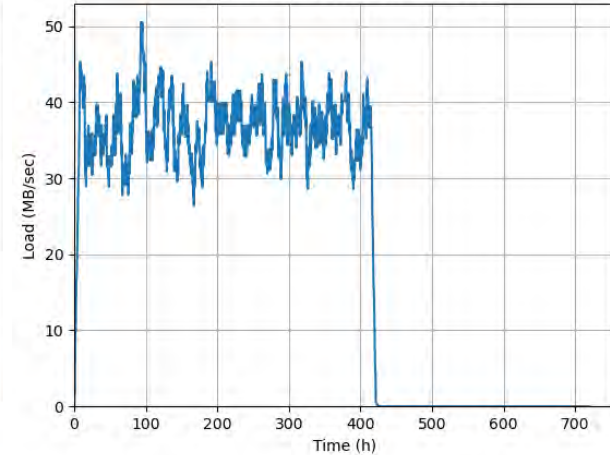
- Amount of raw-data per session – 350 TB

Load of link to the LHEP farm



- Maximum load of link to the LHEP farm – 90 MB / sec

Load of link to the LIT T2 farm



- Maximum load of link to the LIT farm – 50 MB / sec

Conclusions and Outlook

- Developed a tool for modeling the process of data acquisition and processing.

- Based on the simulation results, we can predict the load of farms, data pools and communication links.
- Modeling of 2 primary processing scenarios (executing **RawToDigit** and **DigitToDst** jobs).
- **Next steps:**
 - including other types of jobs (GenToSim, SimToDst, DstToAna) in the described scenarios;
 - modeling other possible scenarios for executing jobs;
 - adding % of jobs completion (now 100% completion);
 - adding probability of equipment failure and recovery time;
 - the probabilities of loss of incoming information.

The XXIV International Scientific Conference of Young Scientists and Specialists

9 - 13 November 2020 JINR, Dubna

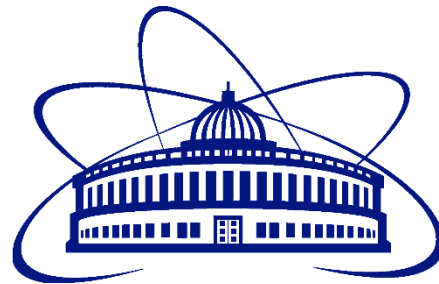
Thank you for the attention!

D. PRIAKHINA

V. TROFIMOV

G. OSOSKOV

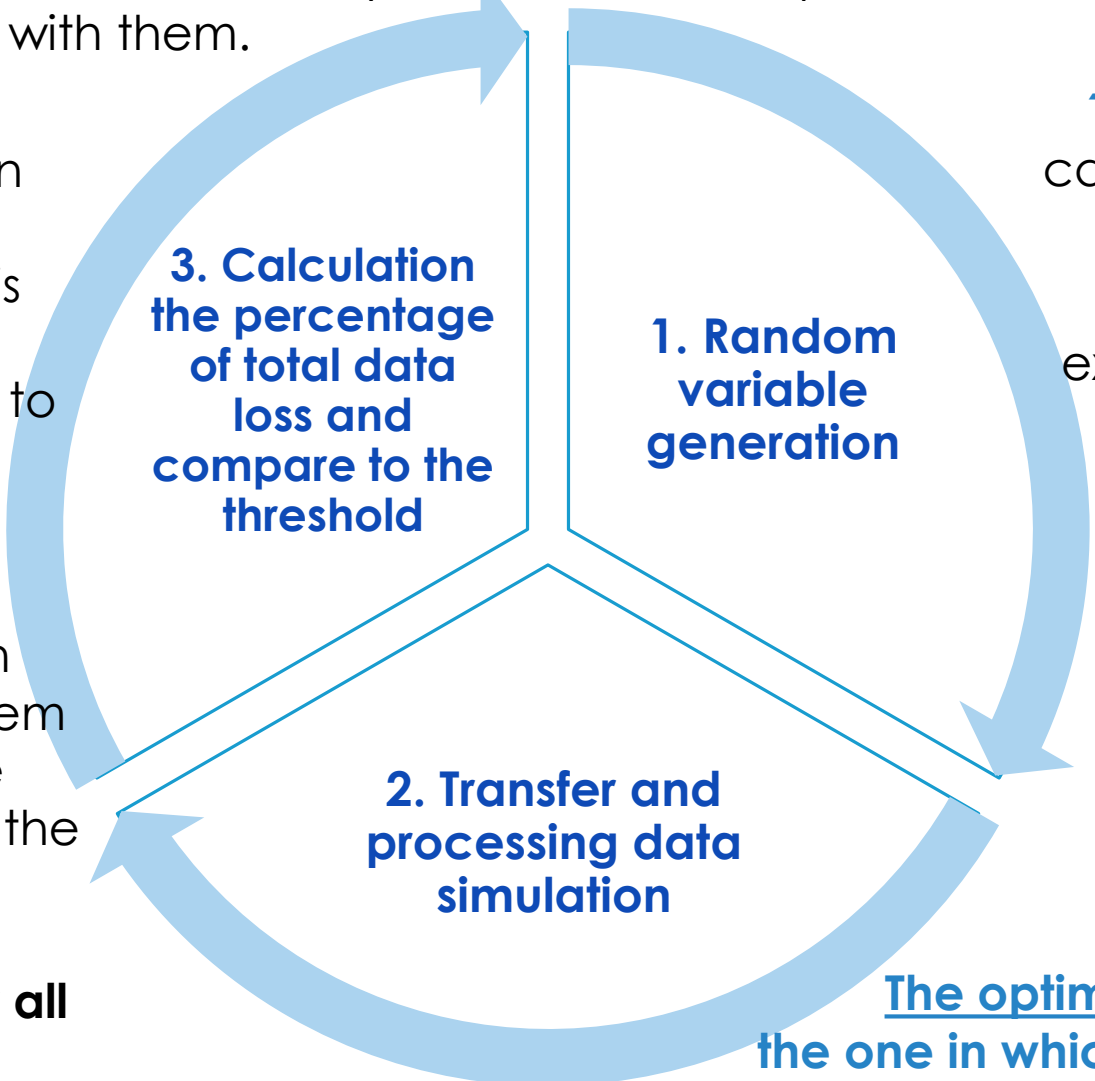
K. GERTSENBERGER



One equipment configuration simulation

- ✓ **Exceeding the threshold = a failure.**
- ✓ Suppose we generate a set of random parameters, but the system does not work with them.
- ✓ However that does not mean the given configuration is inoperable.
- ✓ Then we have to generate new set of random parameters.
- ✓ The conclusion about the system quality can be made only on the basis of **failure probability calculation for all experiments.**

A large number (**N**) of computational experiments with one configuration are required to determine the failure probability:



1. Data losses are calculated in each experiment.
2. The number of experiments (**M**) in which the data losses did not exceed the threshold is calculated.
3. **M/N estimates probability of system quality.**

The optimal configuration:
the one in which the probability doesn't exceed the threshold, and the price is minimal.