



Verification of the Modelling Program for BM@N Computing Infrastructure based on Run 8 mass production

D. PRIAKHINA

V. KORENKOV

V. TROFIMOV

K. GERTSENBERGER



Introduction

The software complex for simulation of distributed data processing systems is being developed at the MLIT.

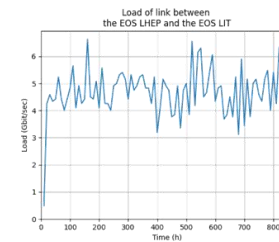
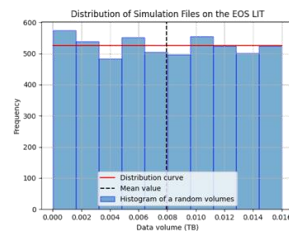
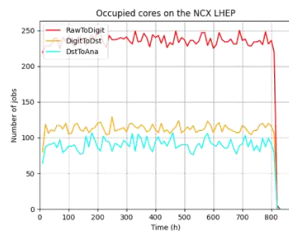
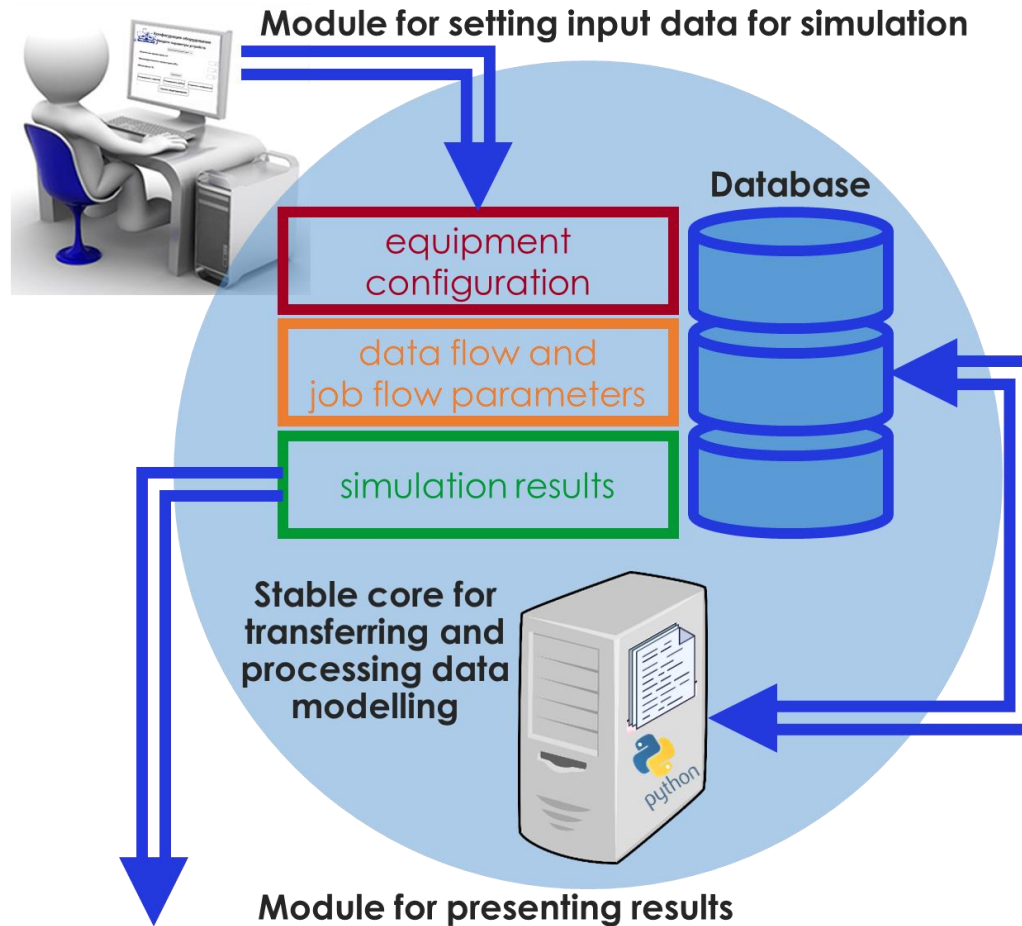
The important task

The data processing simulation of the BM@N experiment.

Simulation goal

- to find out how the data storage and processing system will work with the available computing power;
- to calculate the load on computing farms and communication links with the specified parameters of data flows and jobs flows.

The simulation software complex



Completed works

(from 9th BM@N meeting)

Goal: verification of the modelling program (core of software complex).

1. Upgrade the program and speed up the running time.
 2. Modelling the BM@N Computing Infrastructure with parameters of equipment, data flows and jobs flows based on Run 8 mass production.
 3. Comparison of modelling results with results of BM@N Mass Data Production on distributed infrastructure for Run 8 using DIRAC.
- Thanks to Igor Pelevanyuk*

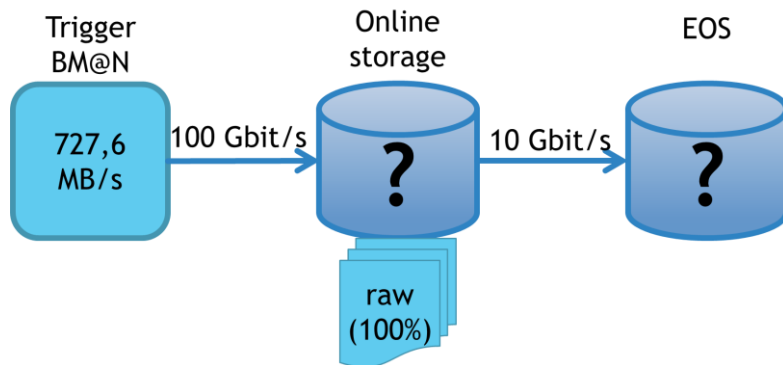
Computations were held on the basis of the Hybrid heterogeneous computing platform.



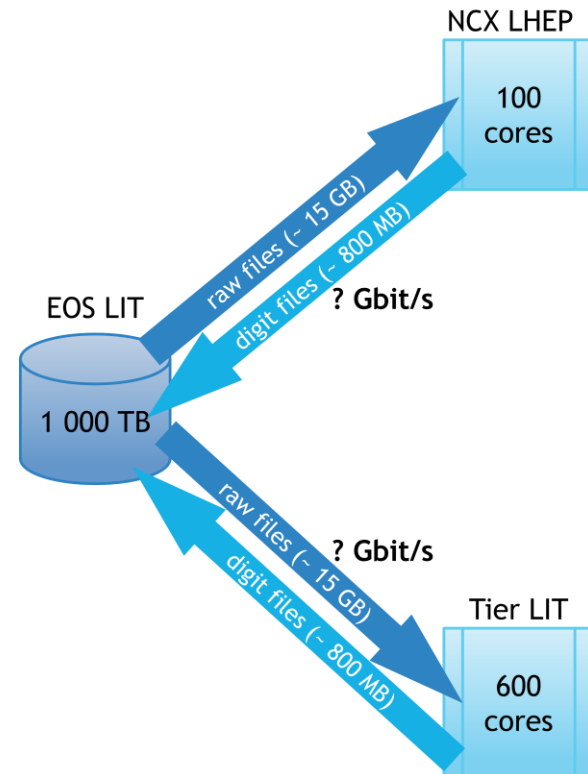
This work is supported by JINR grant for young scientists
No. 23-602-03.

Stages of modelling

DATA ACQUISITION AND STORAGE



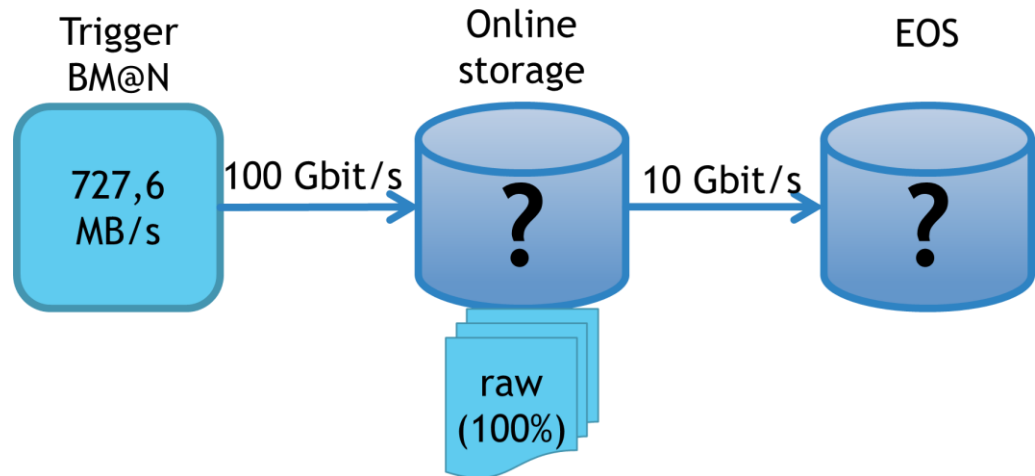
RUNNING AND EXECUTING JOBS



Modelling of data acquisition and storage processes

The Run 8 description (expected)

- Duration \approx 750 h (Dec-22 – Feb-23)
- 1 raw file \approx 15 GB



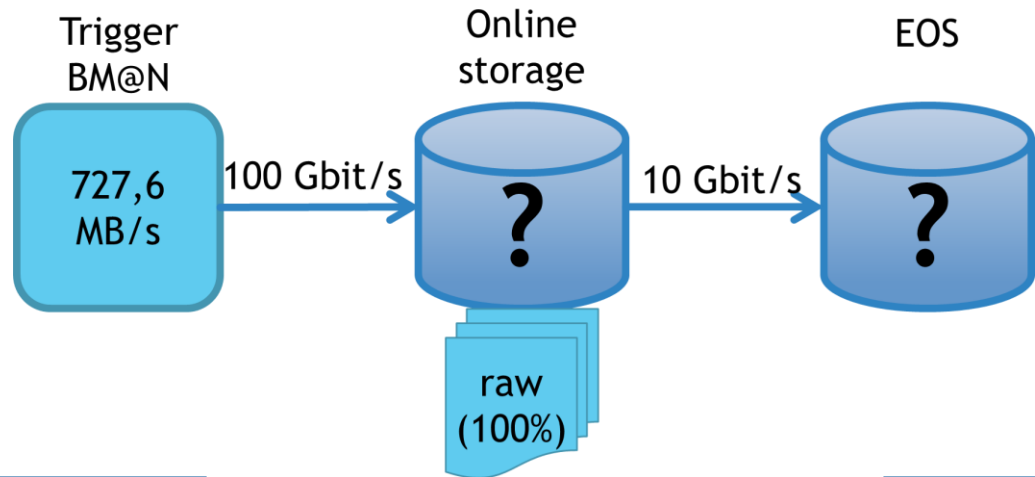
Tasks

1. Find the amount of resources that are needed to store all raw-data on the Online storage.
2. Find the number of raw files in the EOS storage.

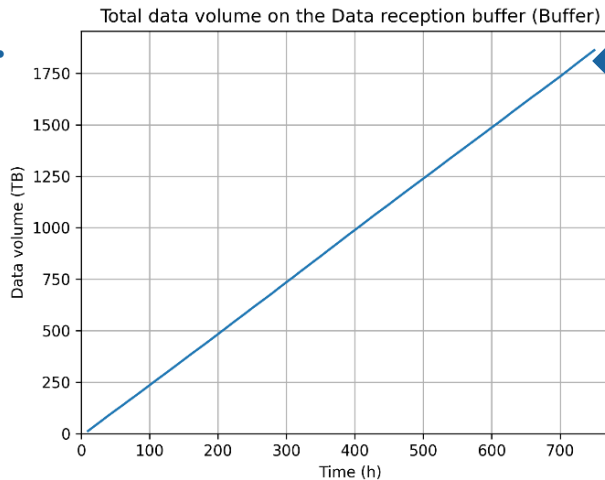
Modelling of data acquisition and storage processes

The Run 8 description (expected)

- Duration ≈ 750 h (Dec-22 – Feb-23)
- 1 raw file ≈ 15 GB

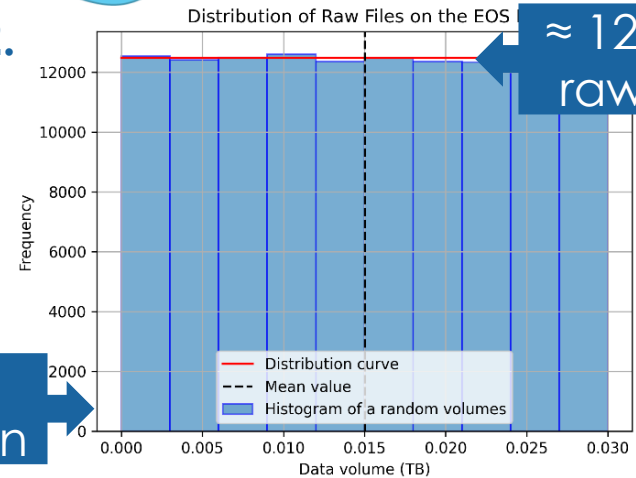


Results



$\approx 1\,850$ TB
in 750 h

2.



$\approx 125\,000$
raw files

uniform
distribution

Run 8 mass production (monitoring by I. Pelevanyuk)

Total raw size ~ 436 TB

Data per day

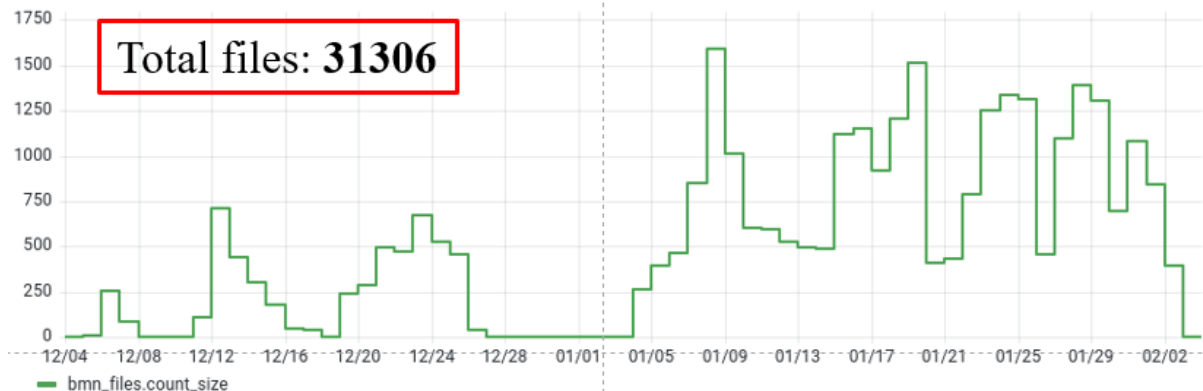


- Non uniform distribution.
- No data is received at all by some periods.

Average data
generation rate
 $\approx 169,3$ MB/s

Total files: 31306

Files per day

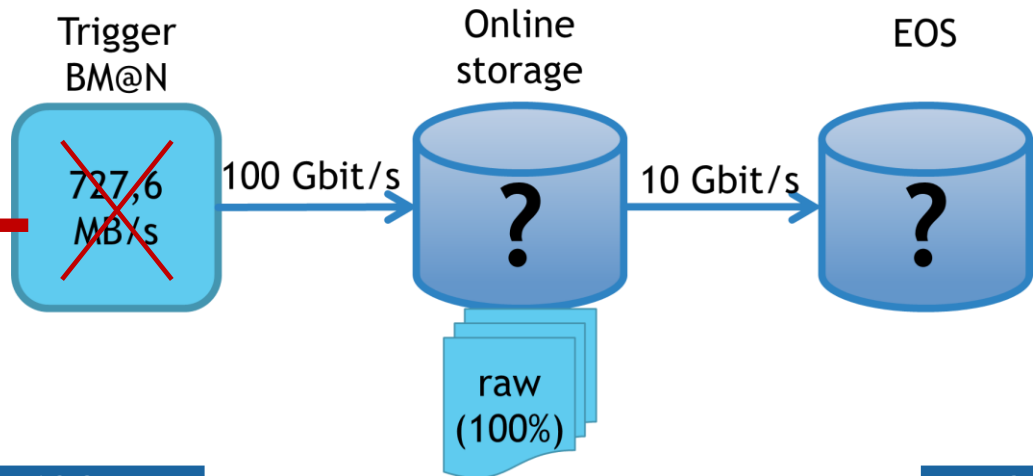


Modelling of data acquisition and storage processes

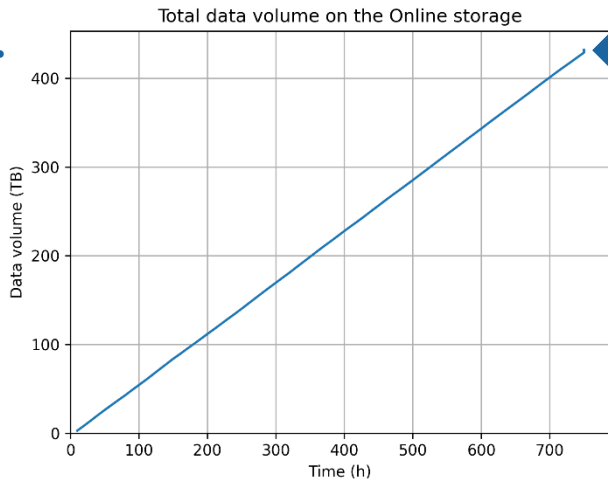
The Run 8 description (by monitoring)

- Data generation rate \approx **169,3 MB/s**
- Duration \approx 750 h (Dec-22 – Feb-23)
- 1 raw file \approx 15 GB

Results

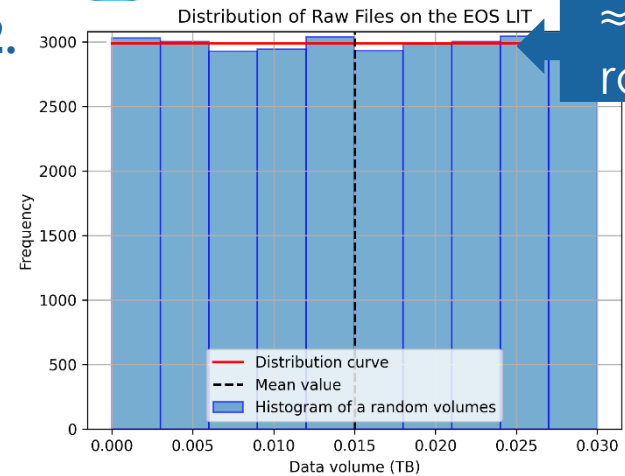


1.



\approx 432 TB
in 750 h

2.



\approx 30 000
raw files

Run 8 mass production (monitoring by I. Pelevanyuk)

Only RawToDigit jobs!

- Total 31 306 jobs
- Each job processes 1 file
- Each file is processed 1 time
- 1 raw file \approx 15 GB
- 1 digit file \approx 800 MB
- Processing time of all jobs \approx 35 h
- Computing resources:
NCX LHEP & Tier LIT

NCX LHEP (100 cores)

Total number of jobs **5451**



- Uniform using of resources
- \approx 100 jobs / h
- 20% jobs of the total number

Tier LIT (avg. 600 cores)

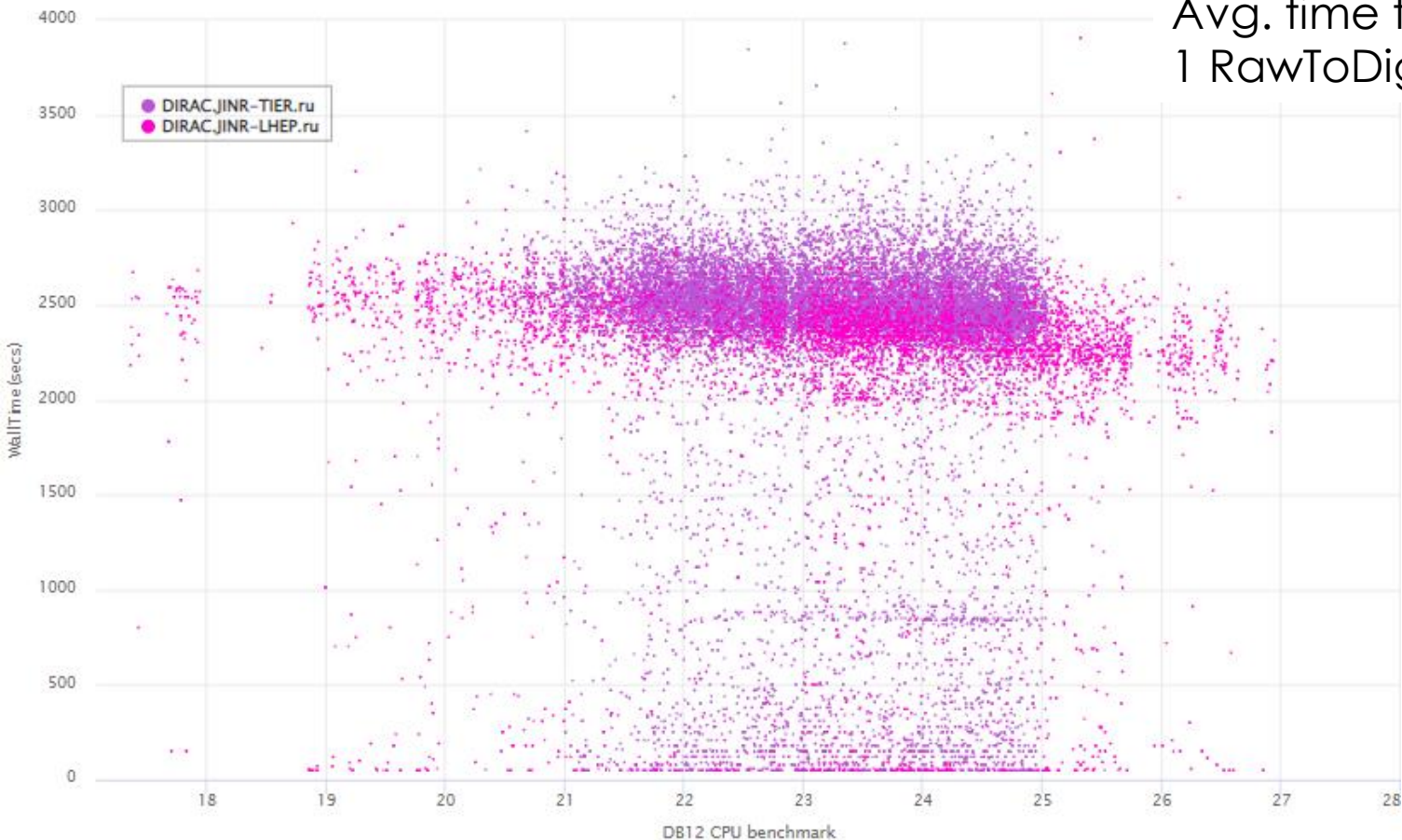
Total number of jobs **25855**



- Non uniform using of resources
- 200 – 1 500 jobs / h
- 80% jobs of the total number

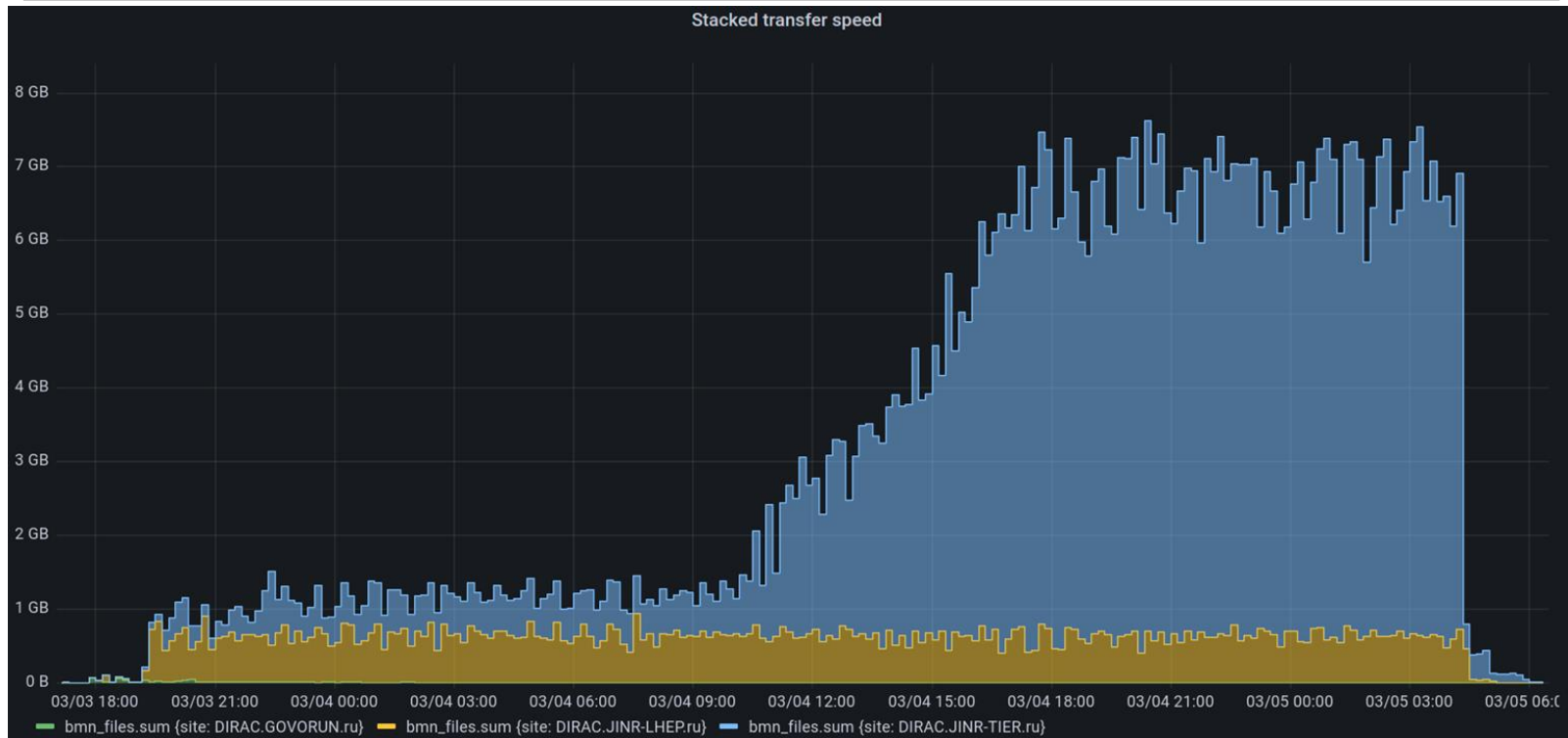
Run 8 mass production (monitoring by I. Pelevanyuk)

Performace monitoring



Avg. time to complete a
1 RawToDigit job \approx 2 500 s

Run 8 mass production (monitoring by I. Pelevanyuk)



- Avg. data transfer rate to NCX LHEP ≈ 0.5 GB / s
- Avg. data transfer rate to Tier LIT ≈ 4 GB / s

Modelling of running and executing jobs processes

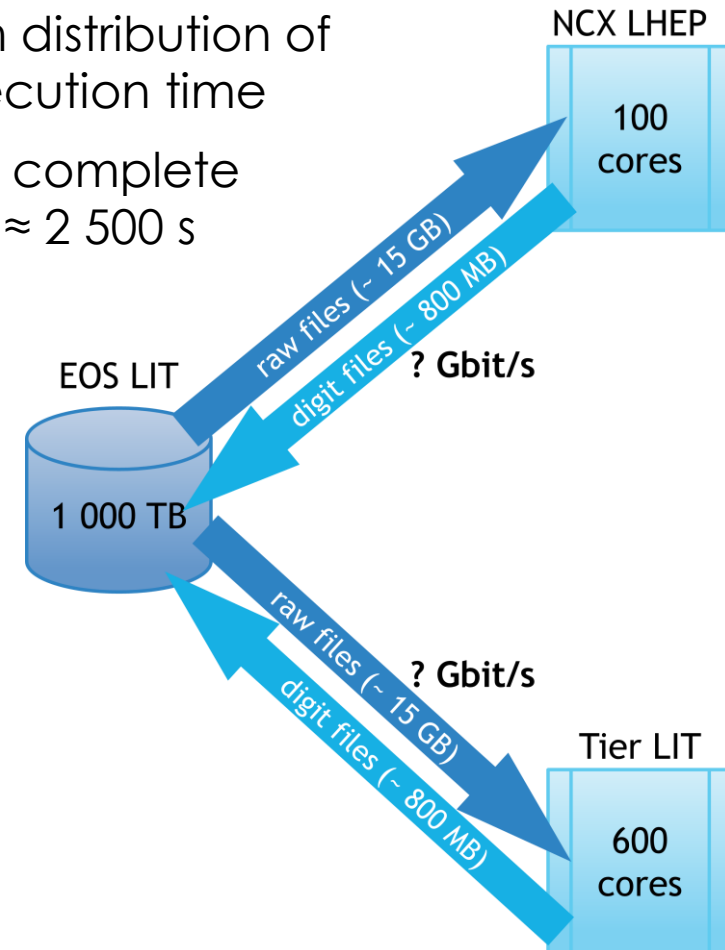
Input data for model

- Total 31 306 RawToDigit jobs
- Each job processes 1 file
- Each file is processed 1 time
- 1 raw file \approx 15 GB
- 1 digit file \approx 800 MB

Tasks

1. Find the total execution time of all RawToDigit jobs.
2. Calculate the load of computing resources during the execution of RawToDigit jobs.
3. Calculate the load of communication links.

- Uniform distribution of job execution time
- Time to complete a 1 job \approx 2 500 s



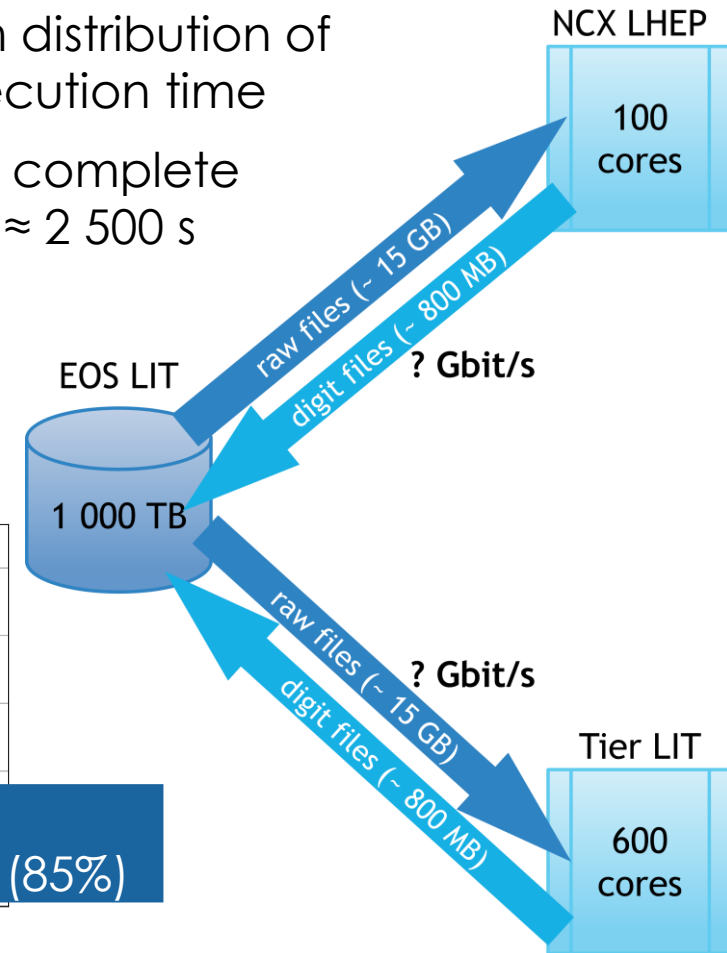
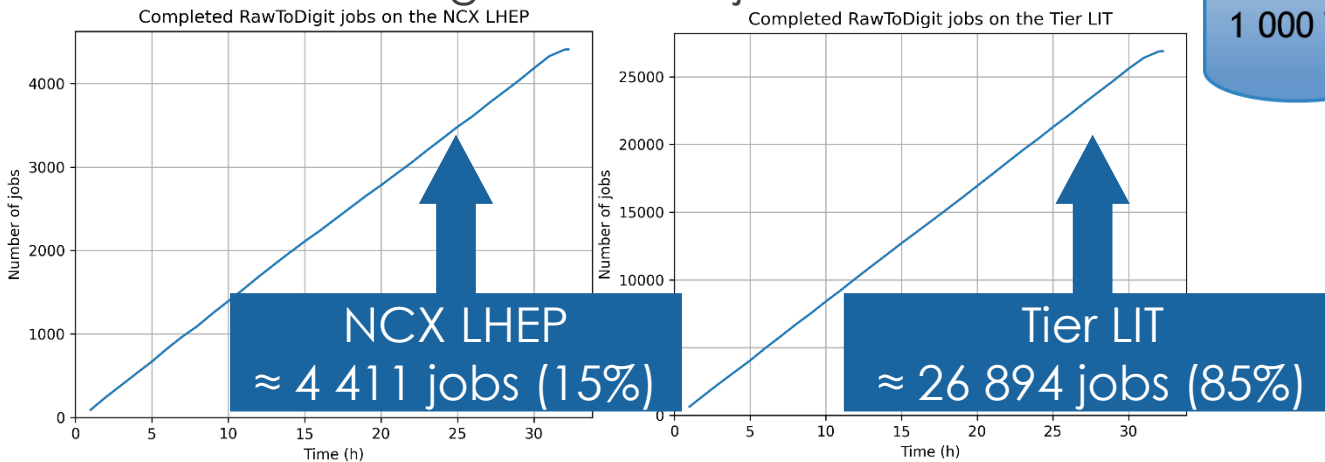
Modelling of running and executing jobs processes

Input data for model

- Total 31 306 RawToDigit jobs
 - Each job processes 1 file
 - Each file is processed 1 time
 - 1 raw file \approx 15 GB
 - 1 digit file \approx 800 MB
- Uniform distribution of job execution time
 - Time to complete a 1 job \approx 2 500 s

Results

1. Processing time of all jobs \approx 33 h



Modelling of running and executing jobs processes

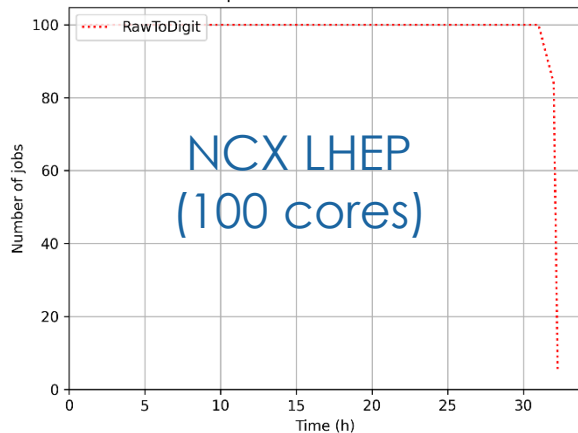
Input data for model

- Total 31 306 RawToDigit jobs
 - Each job processes 1 file
 - Each file is processed 1 time
 - 1 raw file \approx 15 GB
 - 1 digit file \approx 800 MB
- Uniform distribution of job execution time
 - Time to complete a 1 job \approx 2 500 s

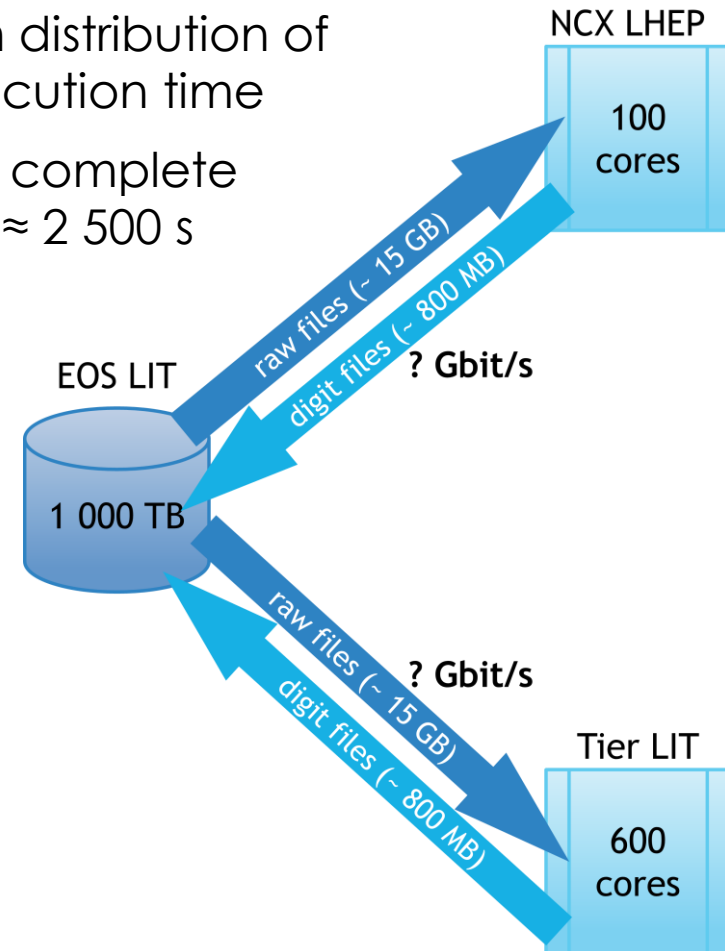
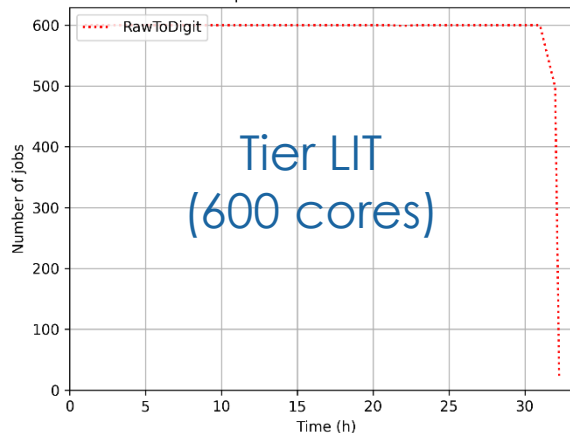
Results

2. Uniform using of resources (100%)

Occupied cores on the NCX LHEP



Occupied cores on the Tiers LIT



Modelling of running and executing jobs processes

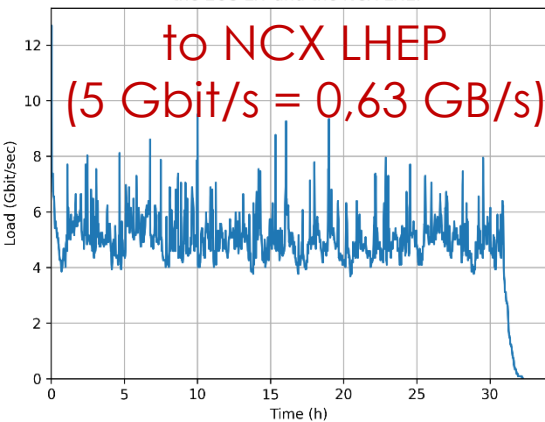
Input data for model

- Total 31 306 RawToDigit jobs
 - Each job processes 1 file
 - Each file is processed 1 time
 - 1 raw file \approx 15 GB
 - 1 digit file \approx 800 MB
- Uniform distribution of job execution time
 - Time to complete a 1 job \approx 2 500 s

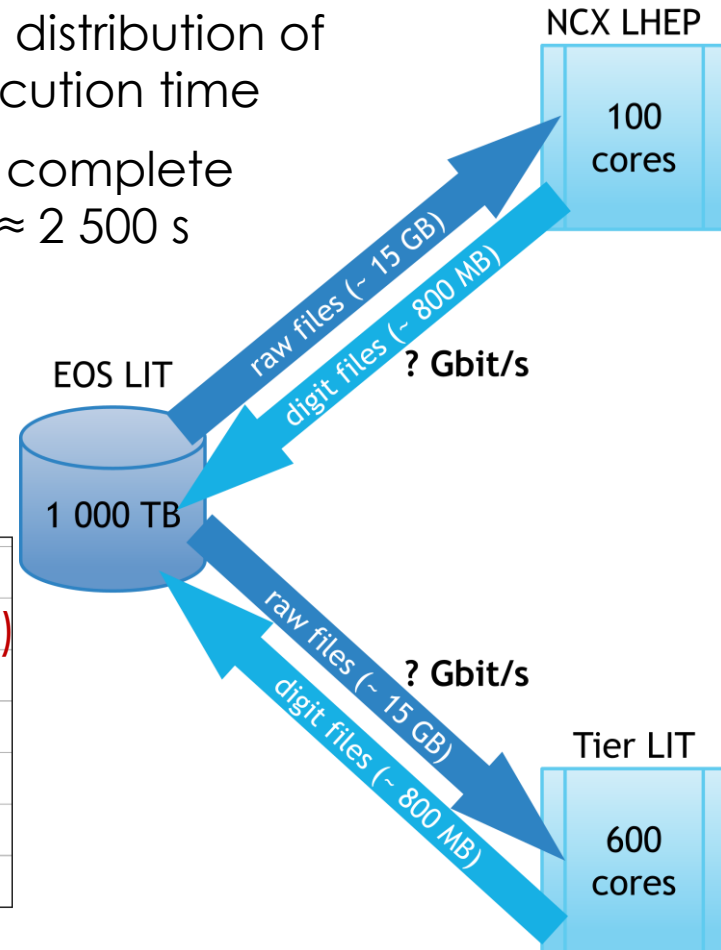
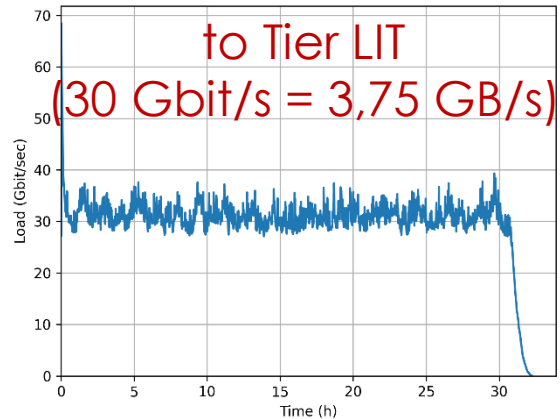
Results

3. Average load of communication links

Load of link between the EOS LIT and the NCX LHEP



Load of link between the EOS LIT and the Tier LIT



Modelling of running and executing jobs processes

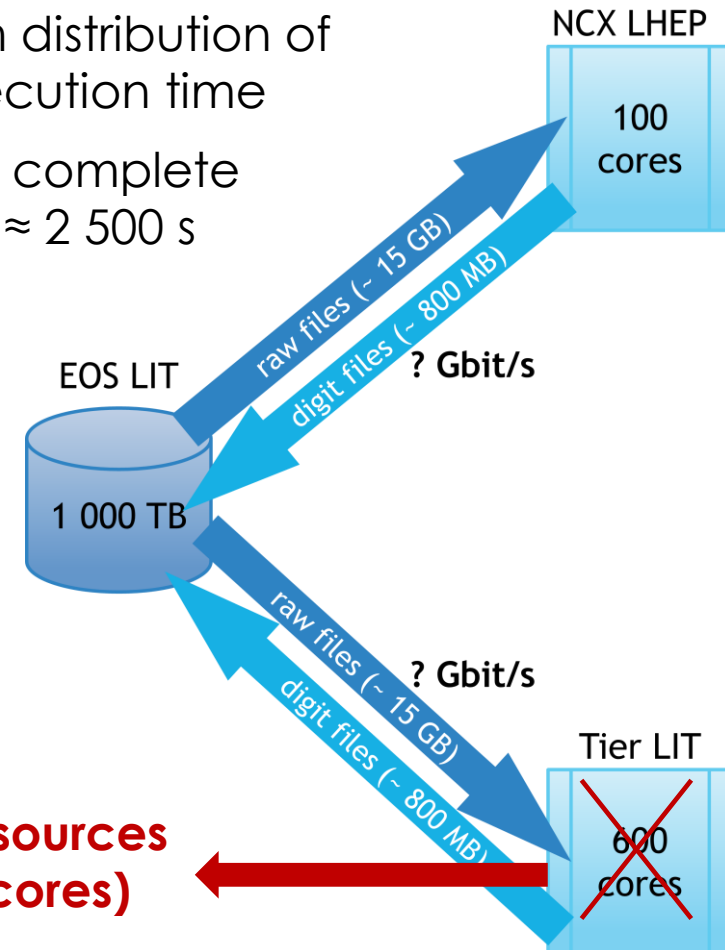
Input data for model

- Total 31 306 RawToDigit jobs
- Each job processes 1 file
- Each file is processed 1 time
- 1 raw file \approx 15 GB
- 1 digit file \approx 800 MB

- Uniform distribution of job execution time
- Time to complete a 1 job \approx 2 500 s

What would happen if...

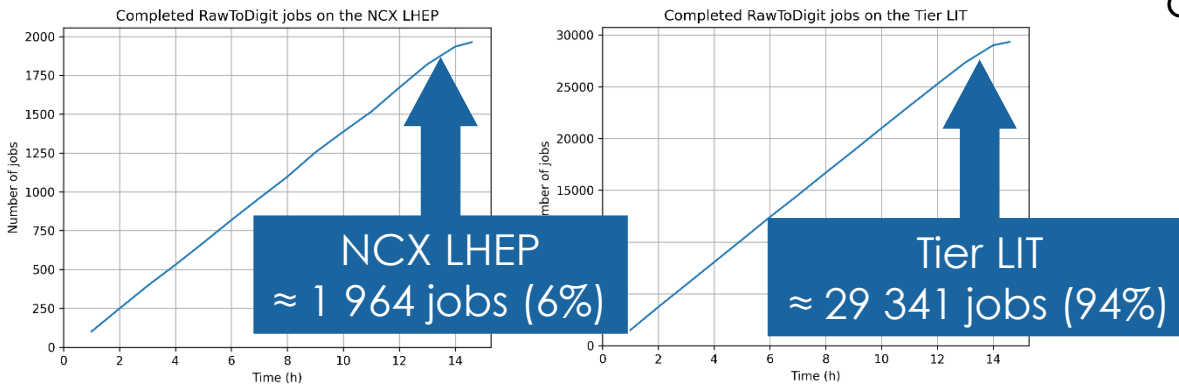
Uniform using of resources on Tier LIT (1 500 cores)



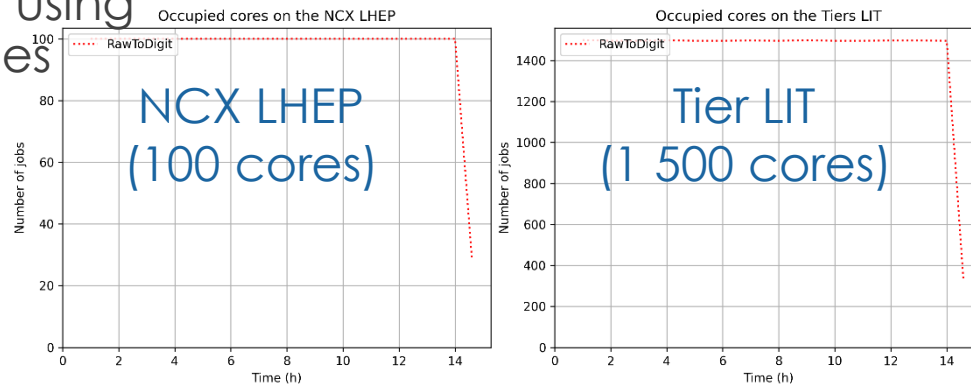
Modelling of running and executing jobs processes

Opportunity results

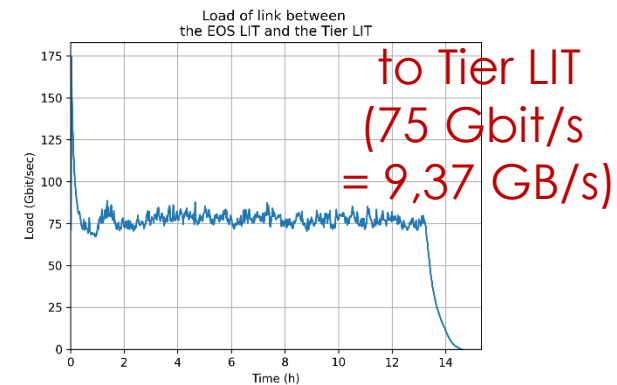
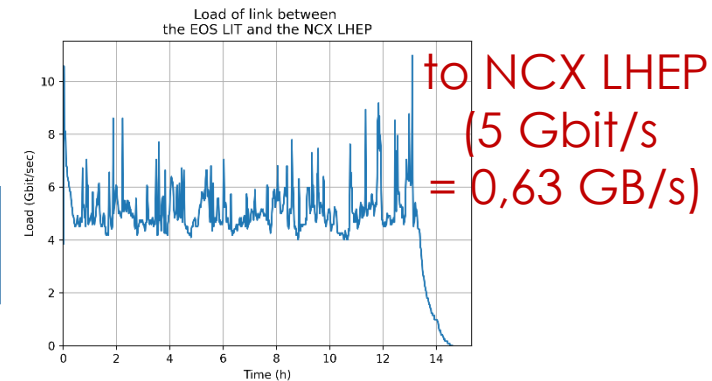
1. Processing time of all jobs ≈ 15 h



2. Uniform using of resources (100%)



3. Average load of communication links



Conclusions

Comparison of monitoring and modelling results

	Monitoring	Modelling
The amount of raw data for 750 h (TB)	436	432
Number of raw files	31 306	30 000
Total execution time of all RawToDigit jobs (h)	35	33
Number of completed jobs on NCX LHEP / % of the total number	5 451 / 20%	4 411 / 15%
Number of completed jobs on Tier LIT / % of the total number	25 855 / 80%	26 894 / 85%
Data transfer rate up to NCX LHEP (GB/s)	0,5	0,63
Data transfer rate up to Tier LIT (GB/s)	4	3,75

The verification results proved the correct and accurate of the modelling program!

The available errors are related to the non uniform real use of resources on the Tier LIT and the averaging of the number of allocated cores during modeling.

Conclusions

Future plans

- Add to the modelling program the ability to change the probabilities of various events in the system:
 - changing the rate of data generation;
 - increase/decrease the number of allocated resources for jobs;
 - equipment failure;
 - equipment recovery times;
 - incorrect completion of jobs ...
- Get statistics on the results of monitoring the process of executing other types of jobs for processing experimental and model data.
- Make a full verification of the simulation program.

Conclusions

Opportunity results

- Modelling of data acquisition and storage processes:
 - 1 850 TB raw-data on the Online storage in 750 h;
 - 125 000 raw files (15 GB) in the EOS LIT storage.
- Modelling of running and executing 31 306 RawToDigit jobs processes (100 cores on NCX LHEP; 1500 cores on Tier LIT):
 - \approx 15 h - processing time of all jobs;
 - 5 Gbit/s - average load of communication links to NCX LHEP;
 - 75 Gbit/s - average load of communication links to Tier LIT.



Thank you for the attention!

D. PRIAKHINA

V. KORENKOV

V. TROFIMOV

K. GERTSENBERGER

