Simulation of BM@N data processing and some propositions

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 tasks.

The simulation software complex

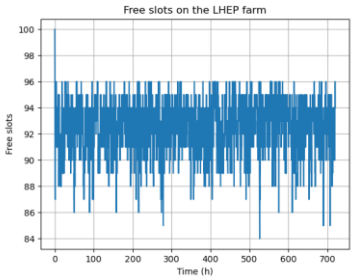
- equipment parameters
- list of jobs for processing

Database

- simulation results

Module for setting of equipment configurations

Stable core for transfer and processing data simulation

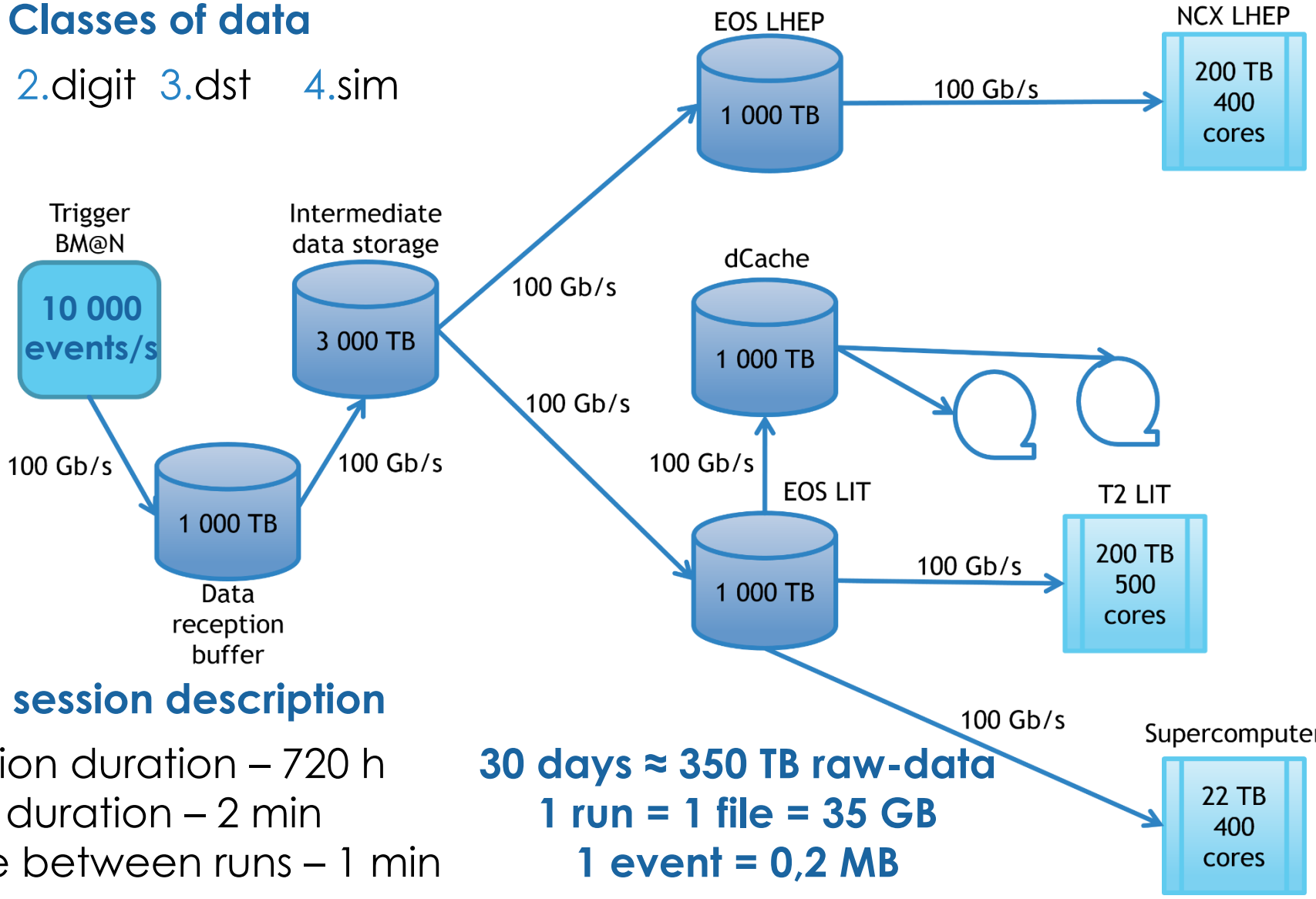


Module for presenting results

The simulated structure

Classes of data

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

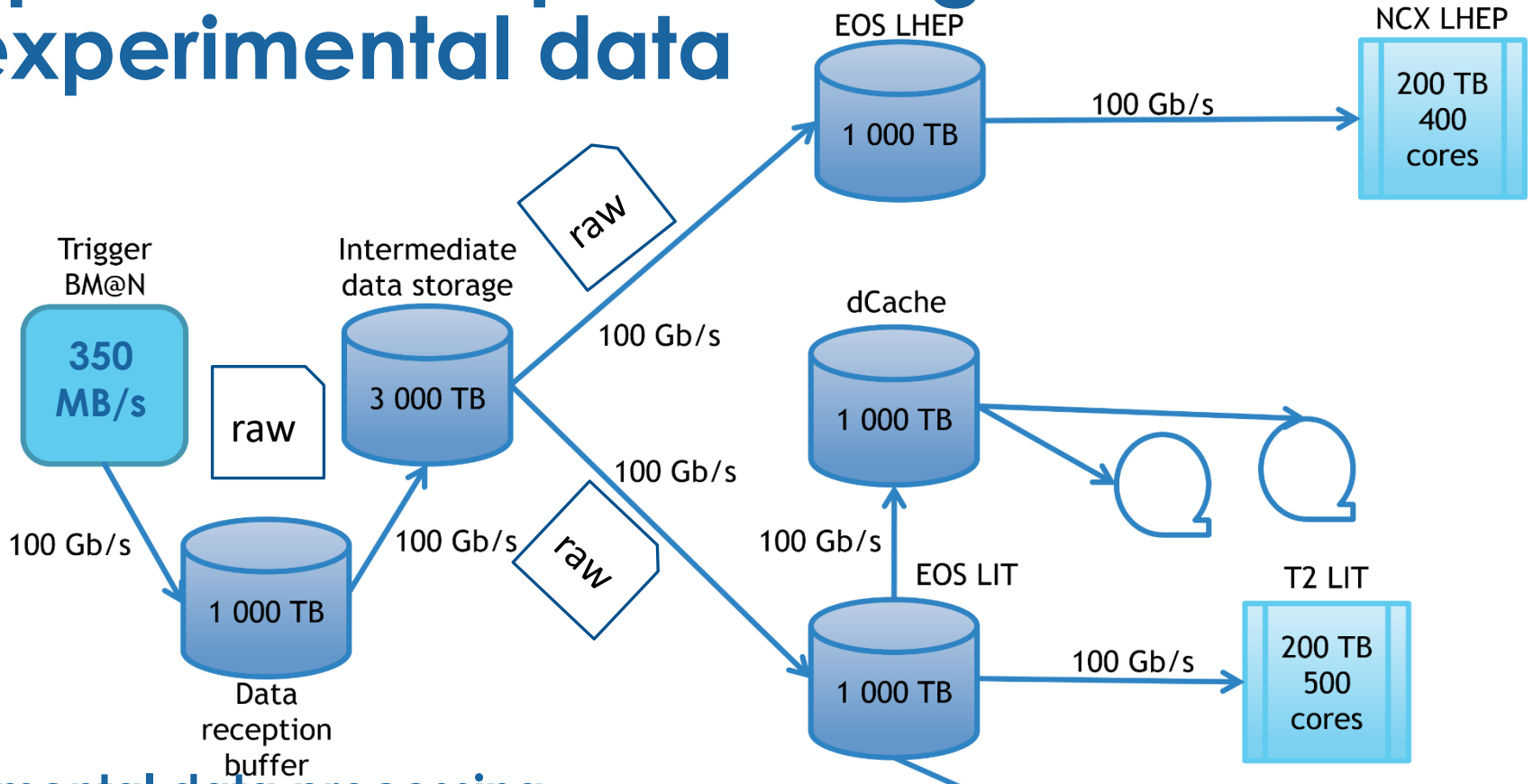


The session description

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

30 days ≈ 350 TB raw-data
1 run = 1 file = 35 GB
1 event = 0,2 MB

Acquisition and processing of experimental data



Experimental data processing

raw → digit → dst → analysis

Model data processing

gen → sim → dst → analysis

Amount of raw-data per session – **350 TB**
RawToDigit job processes **1 file = 35 GB**

Classes of jobs

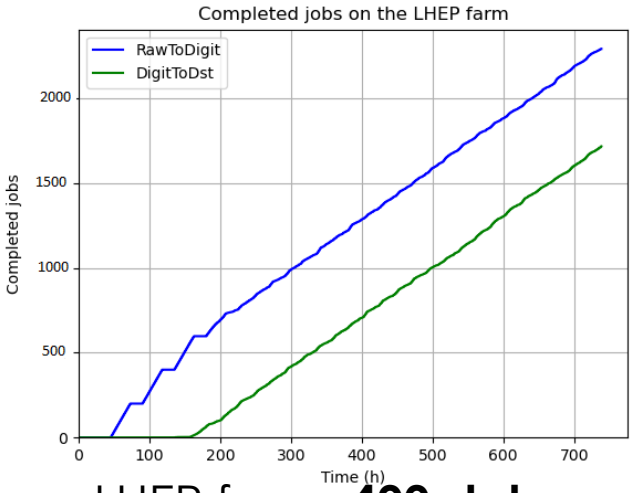
№	Class	Event processing time on one processor (ms)	The average amount of input (GB)	Number of events in the file (1 file = 1 job)	Job execution time (s)	The average amount of output (GB)	Number of jobs
1	RawToDigit	350 (HPC) 1 000 (NCX)	35	175 000	61 250 (HPC) 175 000 (NCX)	1	15 552
2	DigitToDst	150 (HPC) 430 (NCX)	1	175 000	26 250 (HPC) 75 250 (NCX)	1	15 552
3	GenToSim	60	0,6	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

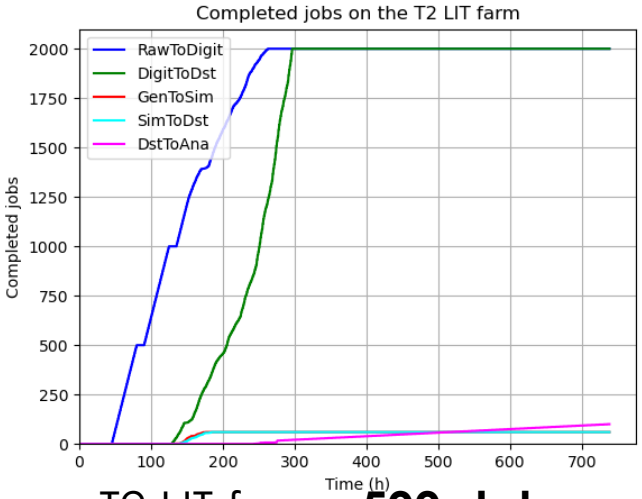
Scenarios for executing jobs

Distribution of data processing jobs (in %) to computing nodes

№	Class	Scenario 1			Scenario 2			Scenario 3		
		NCX LHEP	T2 LIT	Super-comp.	NCX LHEP	T2 LIT	Super-comp.	NCX LHEP	T2 LIT	Super-comp.
1	RawToDigit	50	15	35	80	20	-	-	10	90
2	DigitToDst	50	15	35	80	20	-	-	10	90
3	GenToSim	-	20	80	-	-	100	100	-	-
4	SimToDst	-	20	80	-	-	100	100	-	-
5	DstToAna	-	20	80	70	10	20	80	20	-

Results of Scenario 1





LHEP farm: **400 slots**
RawToDigit jobs – **7 776**
DigitToDst jobs – **7 776**

T2 LIT farm: **500 slots**
RawToDigit jobs – **2 333**
DigitToDst jobs – **2 333**
GenToSim jobs – **90**
SimToDst jobs – **90**
DatToAna jobs – **300**

Supercomputer: **200 slots**
RawToDigit jobs – **5 443**
DigitToDst jobs – **5 443**
GenToSim jobs – **210**
SimToDst jobs – **210**
DatToAna jobs – **700**

- Only 30% raw data will be converted to reconstruction data (during session – 30 days).
- 60% of simulation data will be converted to reconstruction data by 720 h.
- We will have to wait several more months until the end of processing all the raw data after the end of the session.
- There are not enough resources for data analysis.

Results of Scenarios 2 & 3

The results obtained were similar to the results of the first scenario.

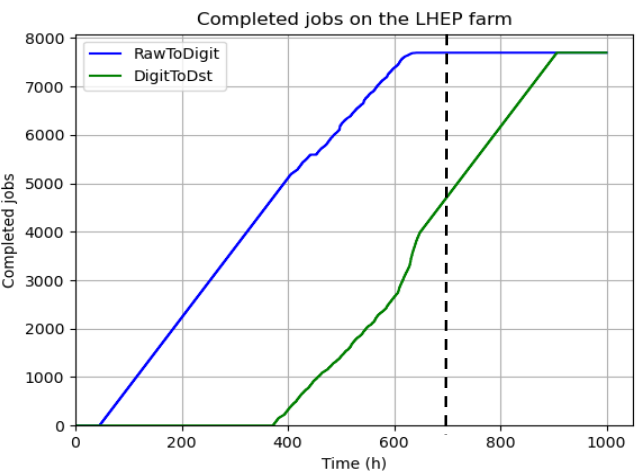
Scenario 2	Scenario 3
10%	**15%**
of all jobs session can be processed by 720 h	
1.5%	**1%**
of raw data will be converted to reconstruction data by 720 h	
100%	**100%**
of simulation data will be converted to reconstruction data	
LHEP farm & T2 LIT farm	**LHEP farm & Supercomputer**
all slots are occupied	

Problem: the result is unsatisfactory…

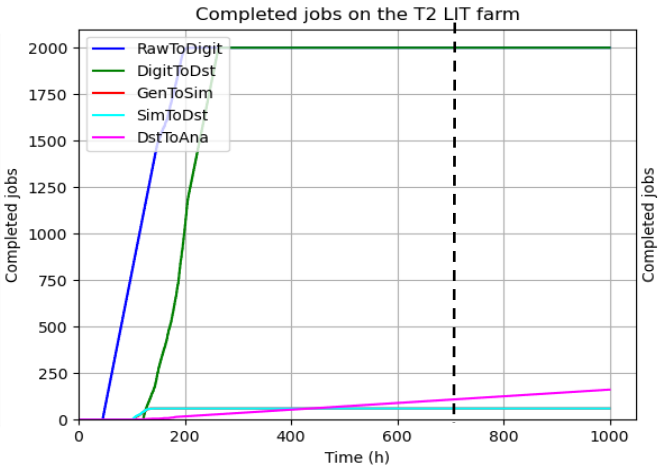
Solving the problem

- To increase the number of cores on computing nodes:
  - LHEP farm – 700 cores
  - T2 LIT farm – 800 cores
  - Supercomputer – 400 cores
- Do not occupy computing resources with other jobs until the jobs of primary data processing (RawToDigit) begin to free up the cores.

Improvement results



LHEP farm: **700 slots**



T2 LIT farm: **800 slots**

Supercomputer: **400 slots**

By end of the Run (30 days)

- 100% raw data will be converted to digit data
- 90% of raw data will be converted to reconstruction data

We will have to wait after the end of the session

- 1 week until the end of processing all the raw data to reconstruction data.

Conclusions and Outlook

- Developed a tool for modeling the data processing.
- Based on the simulation results, we can predict problems that may appear during the experiment and data processing.

- 3 scenarios for executing jobs are modeled. Some problems were found: a small amount of experimental data can be processed by the end of the session.
- Need to increase the number of cores on computing nodes (LHEP – 700 cores, T2 LIT – 800 cores, Supercomputer – 400 cores) and adjust the start time of jobs. Result: 90% of all raw data will be processed by 720 h.
- **Next steps:**
  - developing module like pilot for starting jobs;
  - conducting computational experiments taking into account the fact that the equipment does not have absolute reliability (calculating probability of equipment failure and recovery times);
  - find the optimal number of cores to perform all jobs, taking into account their updated parameters.

Thank you for the attention!

D. PRIAKHINA

V. TROFIMOV

G. OSOSKOV

K. GERTSENBERGER