Probabilistic macroeconomic approach to optimization of distributed data storage systems for physical experiments

LIT

D. Priakhina, A. Nechaevskiy, G. Ososkov, V. Trofimov

Introduction

Problem: the optimal choice of necessary equipment configuration during data processing and storage system design.

Solution: to develop and research data movement model of the HEP experiment.

Disadvantages of the imitative simulation approach (from own simulation experience and accessible literature*):

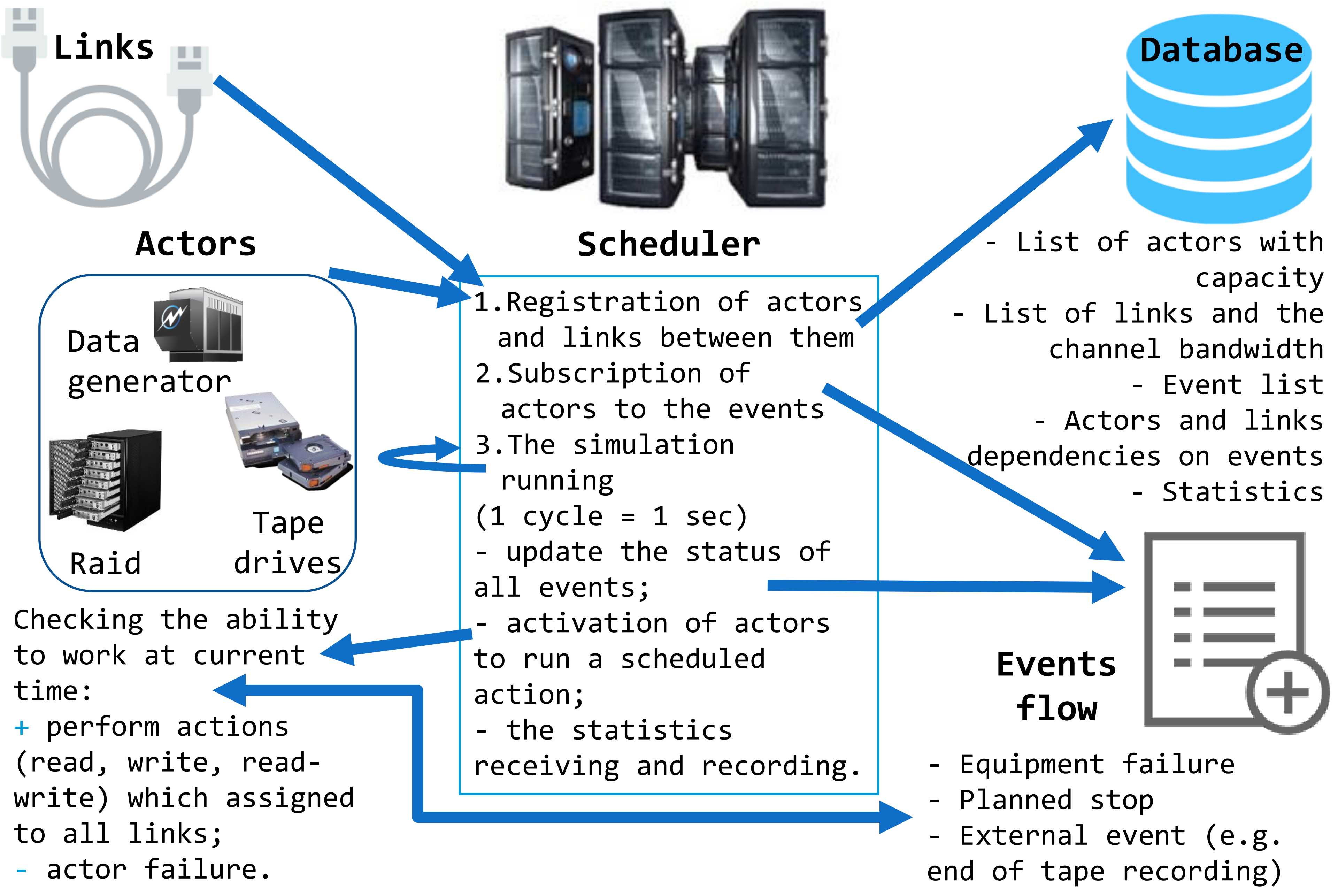
1. the data flow detailed up to packages/files level
   - → difficult program structure;
   - → high computational cost;
2. the equipment parameters detailed description
   - → difficulties in taking decision for the choice optimal equipment configuration;
3. the economic aspects of modeling are not taken into account

A simulation scheme is proposed and implemented:

- ✓ data flow is a static bytes flow (individual flow parts are not analyzed);
- ✓ equipment price is considered;
- ✓ **probabilistic approach** is used to estimate different equipment configurations:
  - data loss probability is determined for each configuration,
  - choice configuration with minimal cost (data loss probability should not exceed the threshold value).

* https://indico.cern.ch/event/587955/contributions/2937260/
https://indico.cern.ch/event/587955/contributions/2936249/

The simulation program structure



Simulation of data acquisition and storage system of the BM@N experiment*

Task: find the equipment configuration which will ensure the working capacity with the defined quality and the minimal cost.

The approach to the searching configuration is based on the definition of statistical dependencies of the system quality on the investments in the storage and data transmission equipment.

The system quality criterion: the relative data loss during the simulation period:

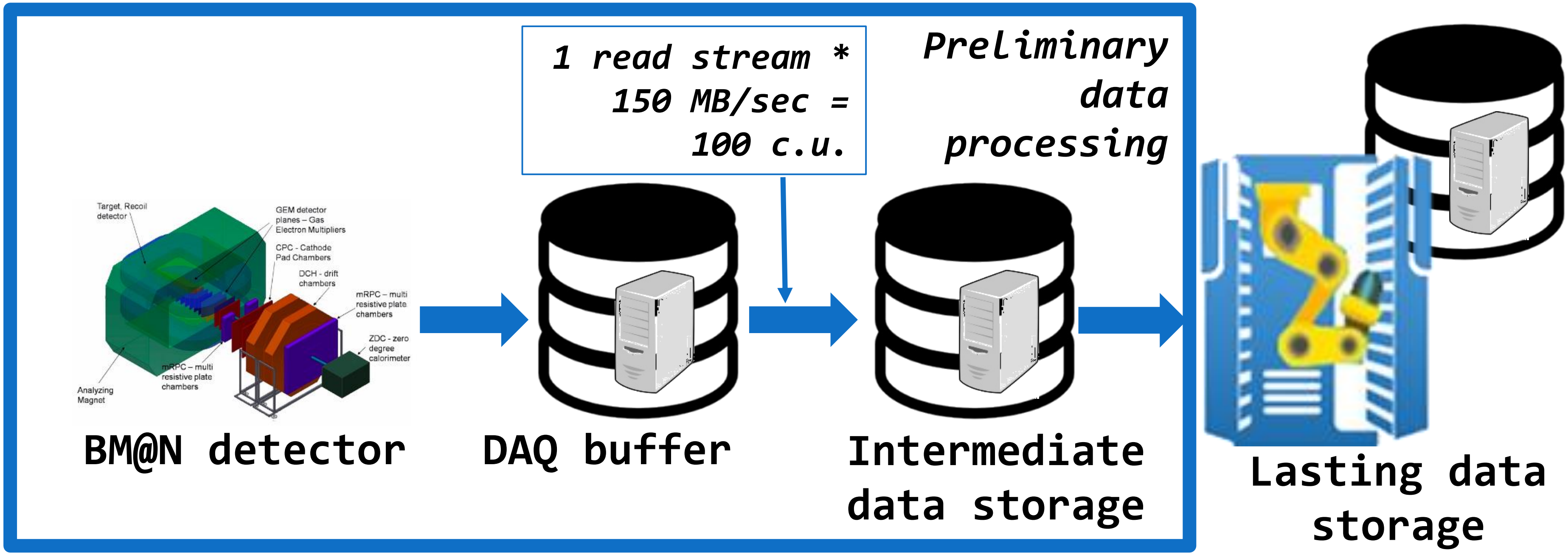
$$1 - \frac{\textit{value of the ratio of the volume stored data}}{\textit{data produced by detector}}$$

If the relative data loss don’t exceed the threshold value at the simulation end, the system of this configuration is considered as workable.

All equipment configurations are characterized by their price (parameter which includes factors such as the cost of equipment, the cost of operation, personnel requirements, etc.)

* Baryonic Matter at Nuclotron
http://nica.jinr.ru/projects/bman.php

Simulation of data acquisition and storage system of the BM@N experiment



Task: to determine the dependence of the probability of system workable on the investments to the intermediate data storage system.

The trigger frequency - 3 kHz
The buffer size - 0.5 MB

The probability of equipment failure at any time - 0.005
The recovery equipment time – 100 sec

The simulation period – 48 h

Simulation algorithm

1. **Setting input parameters:**
   - system configuration parameters;
   - number of computational experiments in each configuration;
   - the time interval for one experiment (seconds).
2. **Running the simulation program:**
   - generation of random variables (e.g. equipment shutdown or failure time);
   - simulation of the data transmission process from the trigger to the storage system for each computational experiment;
   - determining the percentage of total losses for the experiment and comparing it with a threshold (exceeding the threshold is considered as a failure, the calculations are stopped, configuration is not workable with these random values).
3. **Calculate the probability of failure** for each configuration
4. **Choice optimal configuration with minimal cost** (data loss probability does not exceed the threshold value).

Simulation results of data acquisition and storage system of the BM@N experiment

The probability of workable system at different number of read streams (150 MB / sec) and different thresholds of possible losses

Number of streams	The average value of losses	The dispersion of the average values	Prob. factor 0,005	Prob. factor 0,01	Prob. factor 0,02
33	0,045532217	0,011470388	0	0	0
34	0,02328578	0,015501326	0,1	0,2	0,4
35	0,011962853	0,014654211	0,333333	0,44444	0,888889
36	0	0	1	1	1

Initial number of read streams – 33

1 read stream = 100 c.u.

Result: to achieve the probability of workable system >0,5:

- **if 1% loss is acceptable - add 3 read streams (invest 300 c.u.)**
- **if 2% loss is acceptable - add 2 read streams (invest 200 c.u.)**

Conclusions

- **Data acquisition and storage system** of the BM@N experiment can be **more reliable** (to reduce the loss probability) by **adding 2 or 3 read streams** with a 150 MB/sec bandwidth which will cost 200 or 300 c.u., respectively.

- The simulation results stimulated meaningful and well-reasoned discussions with the designers of DAQ and triggers systems on the data flow parameters that facilitate the adoption of motivated solutions.
- The program is focused so far only on the analysis of losses for B@MN + MPD experiments.
- **Development prospects:**

- create classes to change the topology of the storage system;

- testing a parallel program realization since it promises to analyze a significant number of alternative configurations (tens of thousands).

Thank you for the attention!

Probabilistic and statistical approach to optimization of distributed data storage systems for physical experiments

D. Priakhina, A. Nechaevskiy, G. Ososkov, V. Trofimov