Full-Scale Simulation of the Super C-Tau Factory Computing Infrastructure to Determine the Characteristics of the Necessary Hardware

D. WIENS, I. CHERNYKH, I. LOGASHENKO

BINP SB RAS, ICMMG SB RAS, Novosibirsk

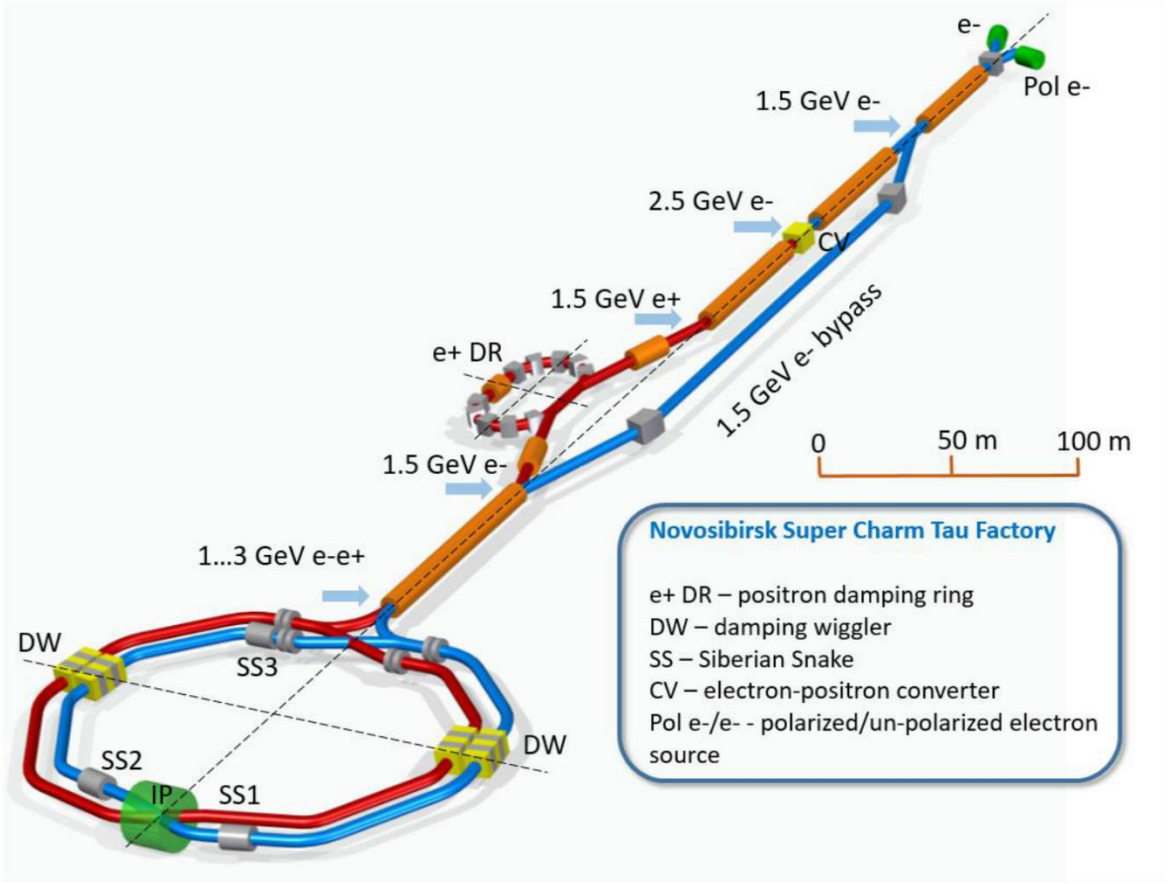
This work was supported by the Russian Science Foundation (Grant No. 19-72-20114)

Super Charm-Tau Factory Project

The «Super Charm-Tau Factory» project, which is a symmetric electron-positron collider of ultrahigh luminosity with a beam energy at the mass center system from 2 to 6 GeV, is developed at the BINP SB RAS.

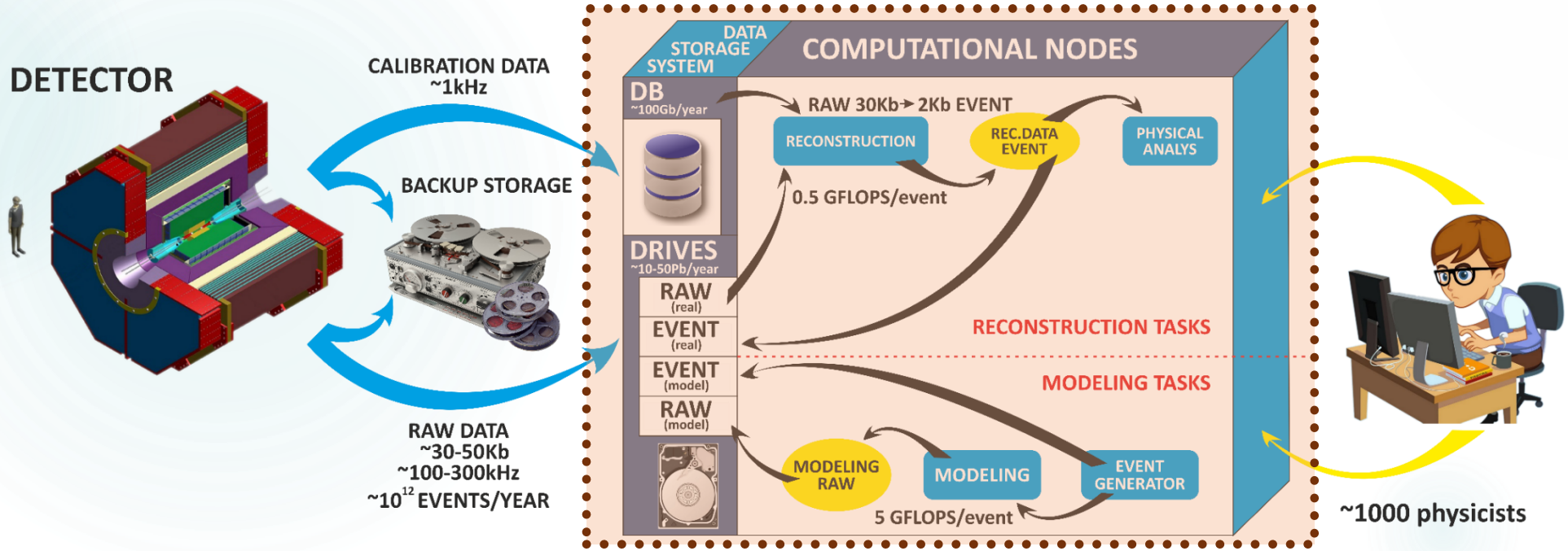
This project comprises a unique accelerating-storage complex with a luminosity of 10^{35} $cm^{-2}s^{-1}$ and a universal elementary particle detector.

The main goal of experiments carried out on the SCTF is to study the properties of tau lepton and charmed particles, subject the existing microworld theory and Standard Model to high-presicion verification, and to search for phenomena not described within the framework of this theory.

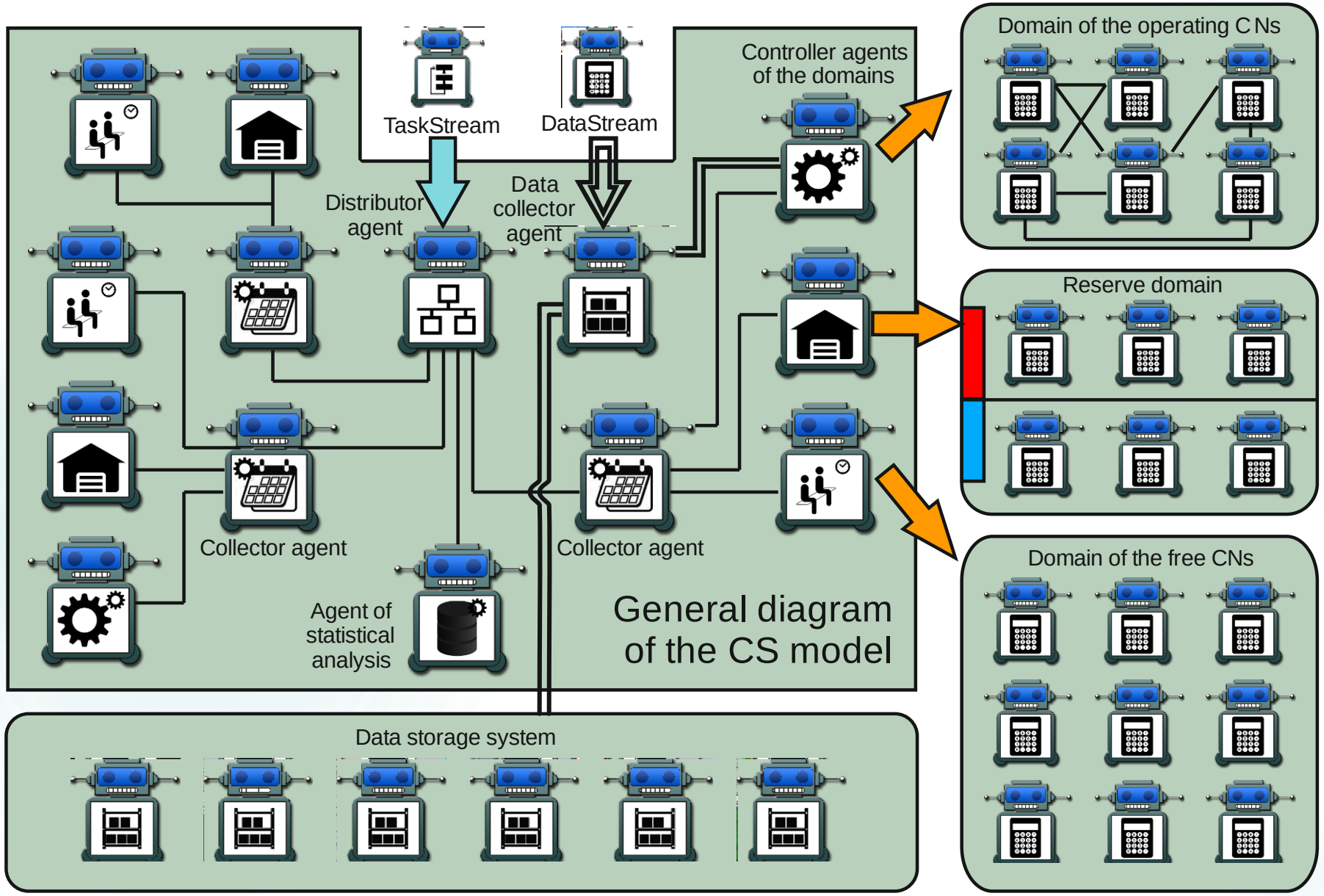


Super Charm-Tau Factory Project

In the course of the experiments, about 100 petabytes of "RAW" data is accumulated from the elementary particle detector of the SCTF. An important role in the project is played by the system for data processing and storage, whose tasks include the primary data processing, data transfer to long-term data storage system (decades), data extraction from the storage system for processing and processing using high-tech computing (HPC) systems. Specialized software should allow one to analyze the accumulated data by a collective of about 1000 physicists. The development of the data analysis algorithms and the optimization of the detector structure are carried out using modeling data generated via software for modeling of experiments.

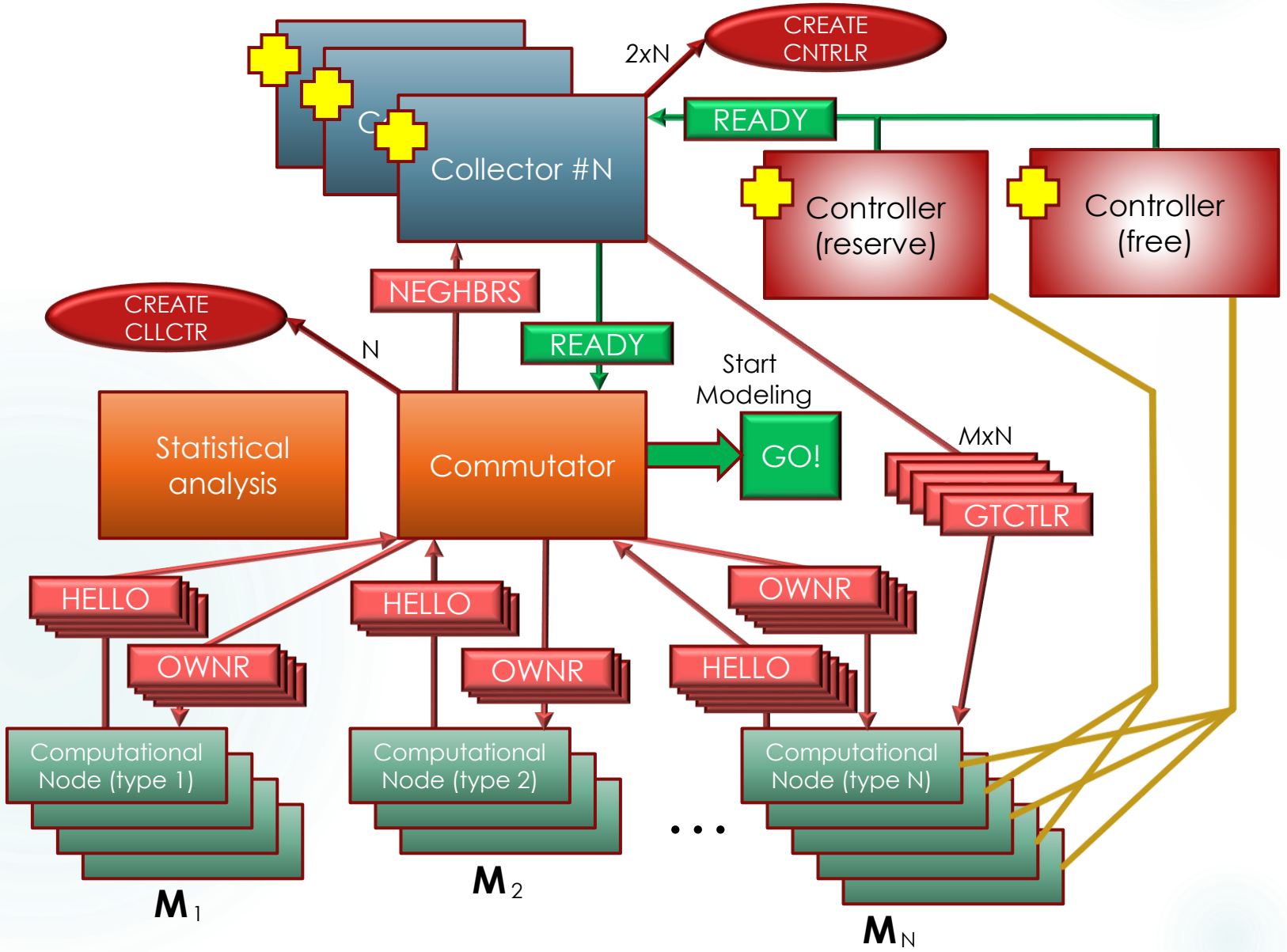


Model of HTC System. General scheme

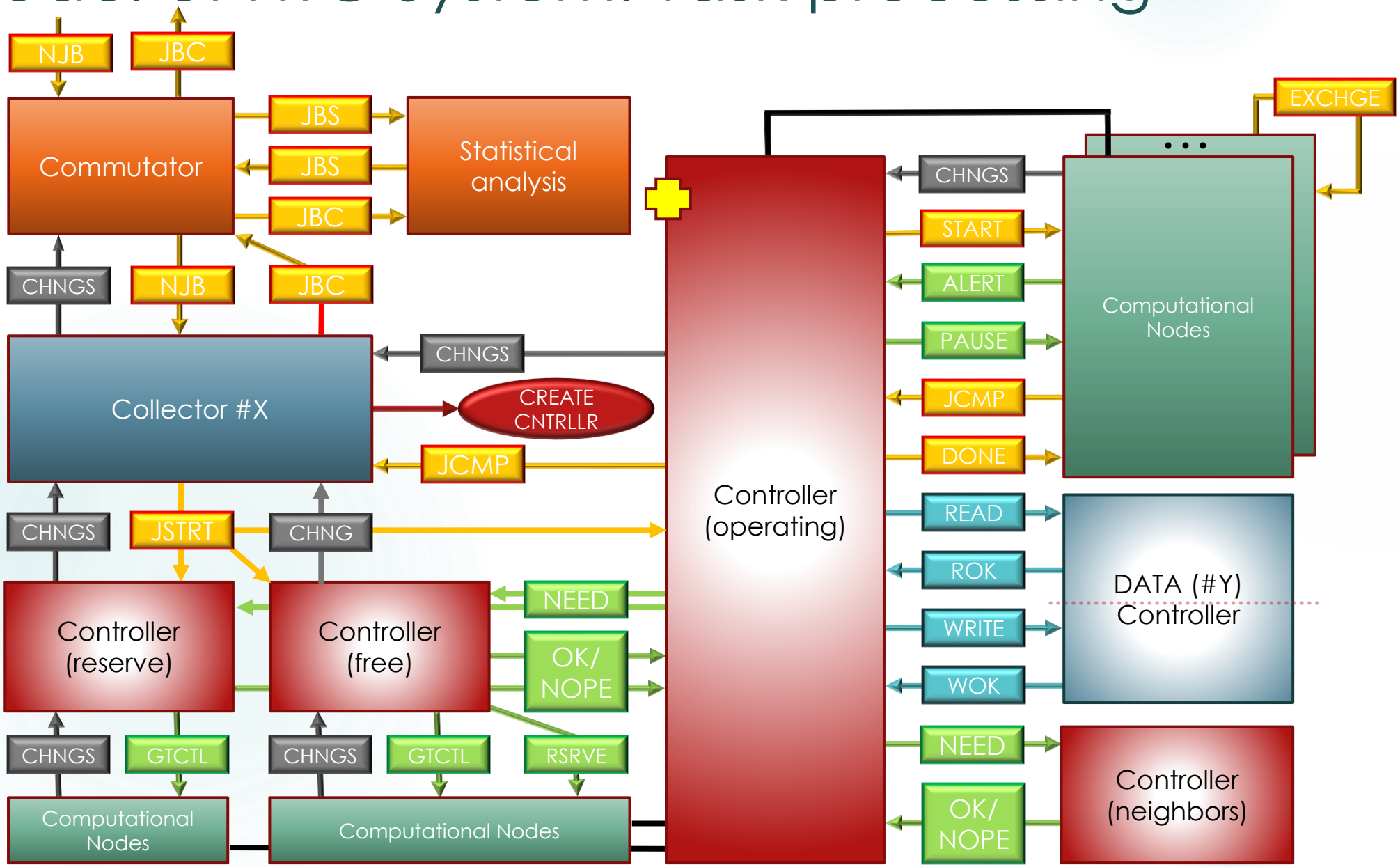


*Created in AGNES (Jade+JM)

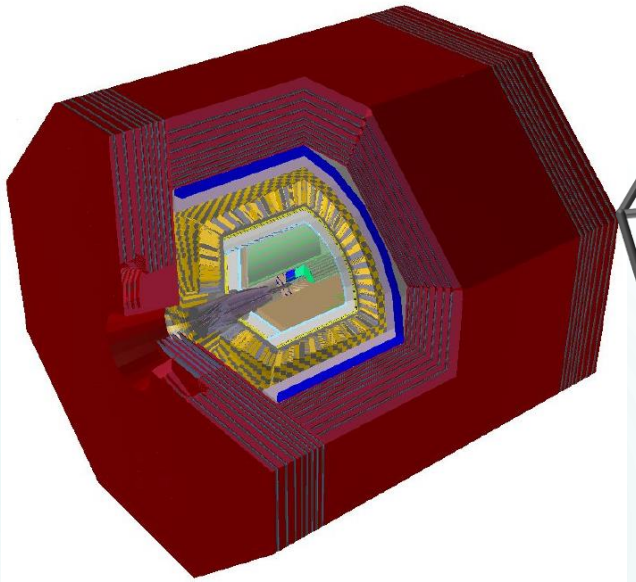
Model of HTC System. Initialization.



Model of HTC System. Task processing



Model of Storage Data System



DataCollector (DB)

DataCollector IEVENT real)

DataCollector (RAW real)

DataCollector (RAW modell)

DataCollector IEVENT model)

DataCollector (Backup)

Read ops

Write ops

Read ops

Read ops

Write ops

Read ops

Reconstruction task

Analysis task

Modeling task

Full-scale modeling. Initial data

DATA	Calibration (KALIB)	Raw * (RAW)	Reconstructed (RECO)	Modeling raw (MRAW)
Size	1kB/sec	30kB/event	2kB/event	30kB/event

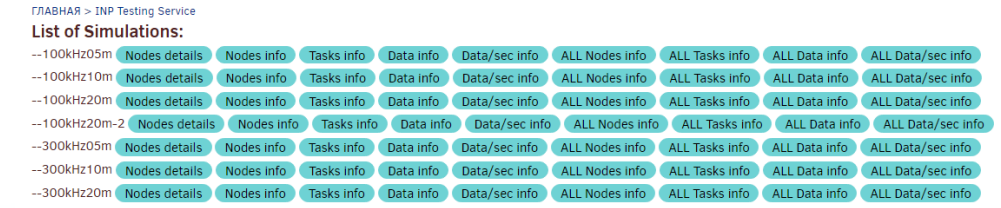
*Raw data events frequency 100/300kHz

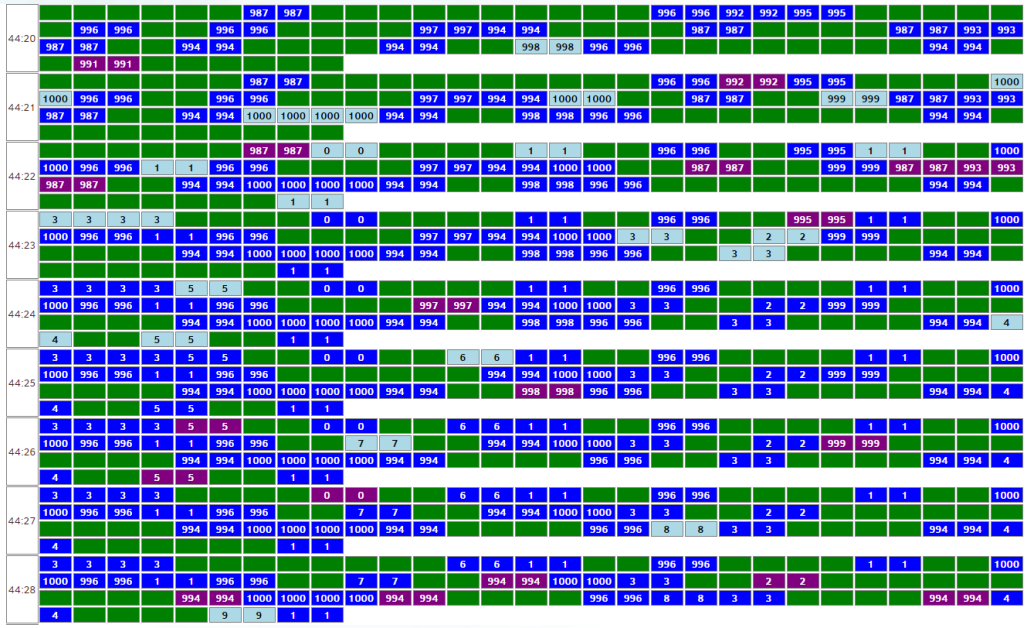
TASKS	Reconstruction (RECO)	Modeling (MOD)	Compare (COMP)
Start	auto	Pmod = 5/10/20%	Pcomp = 2%
Complexity	0.5GFLOPs/event	5GFLOPs/event	insignificantly
Input	100K RAW	none	Needed amount of RAW/MRAW
Output	100K RECO	100K MRAW	none

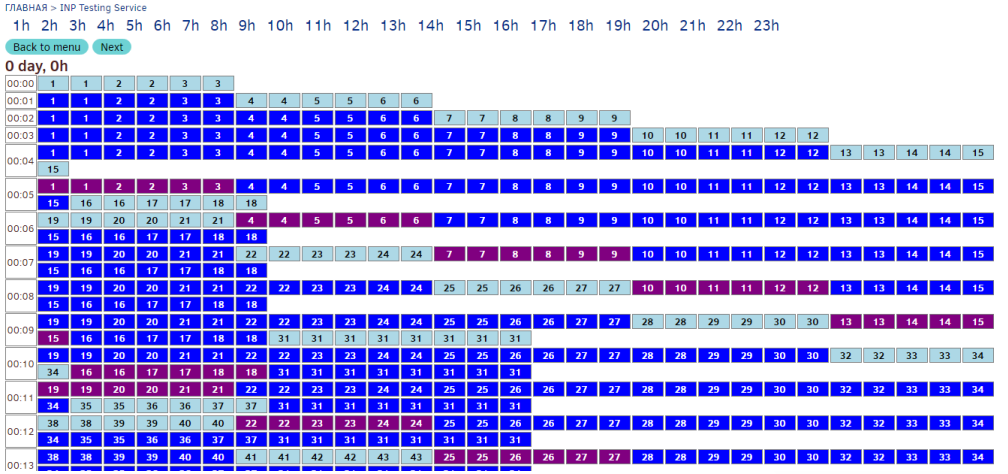
NODE	CPU	RAM	ROM
6.3TFLOPs	2x Intel Xeon 8368Q (2.6GHz, 38 cores)	512Gb – 4Tb	infinitely

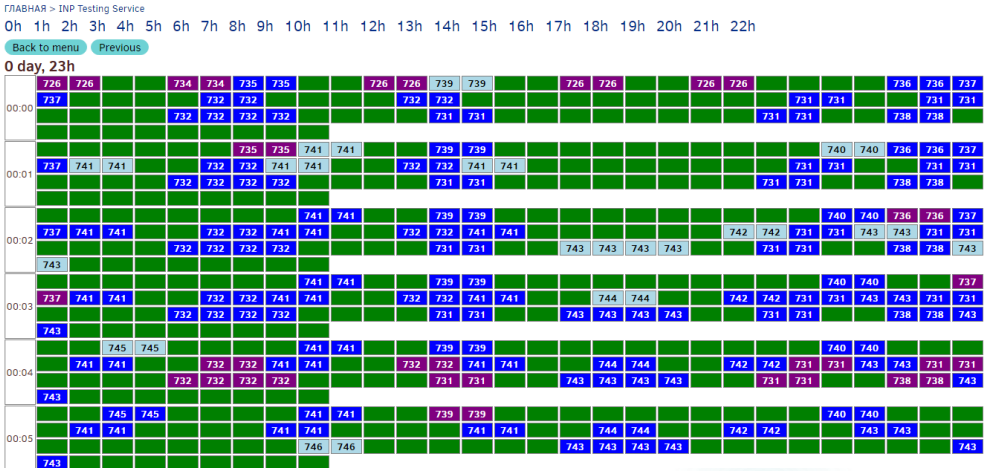
Testing modeling. Agent of statistical analysis

A module for interactive viewing logs from agent of statistical analysis was created based on the SQL+JS+PHP.





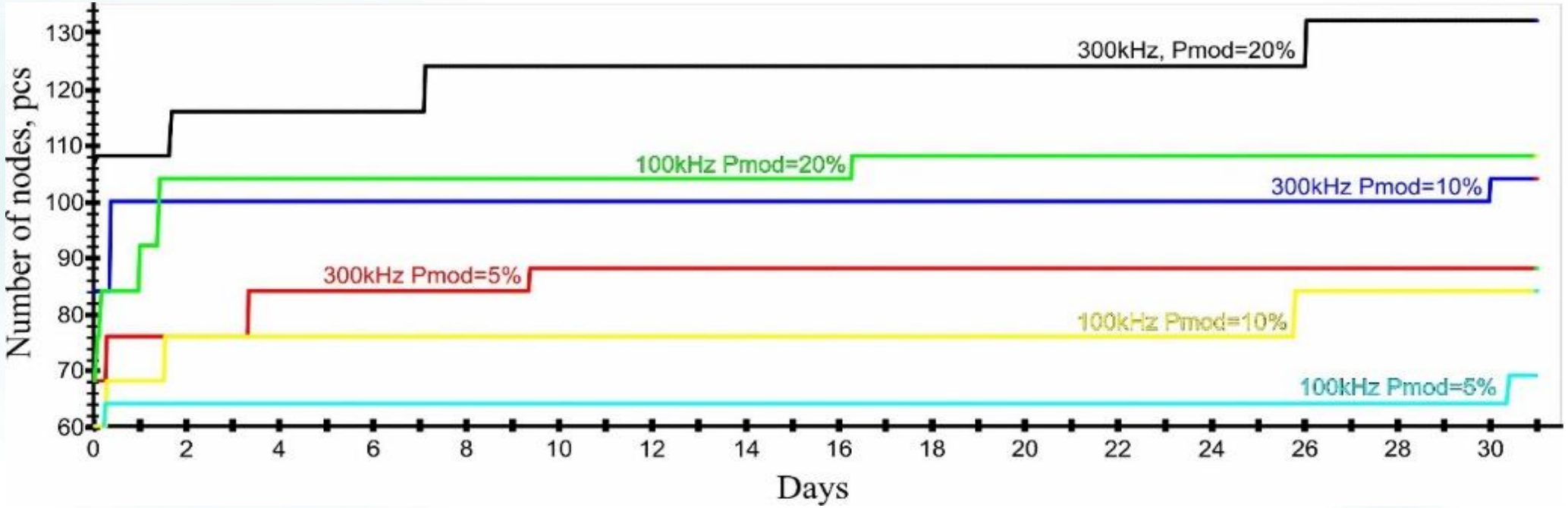




Full-scale modeling. Data flow.

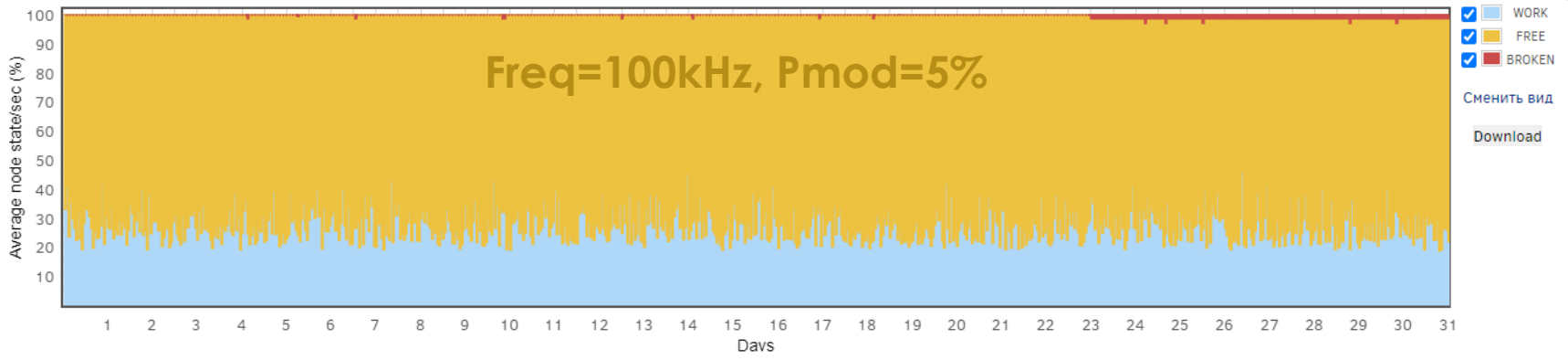
Several model experiments simulating a month of the system operation were performed with different startup parameters. Two modes of operation of the detector were considered: with a frequency of 100 kHz and 300 kHz with different probabilities of tasks.

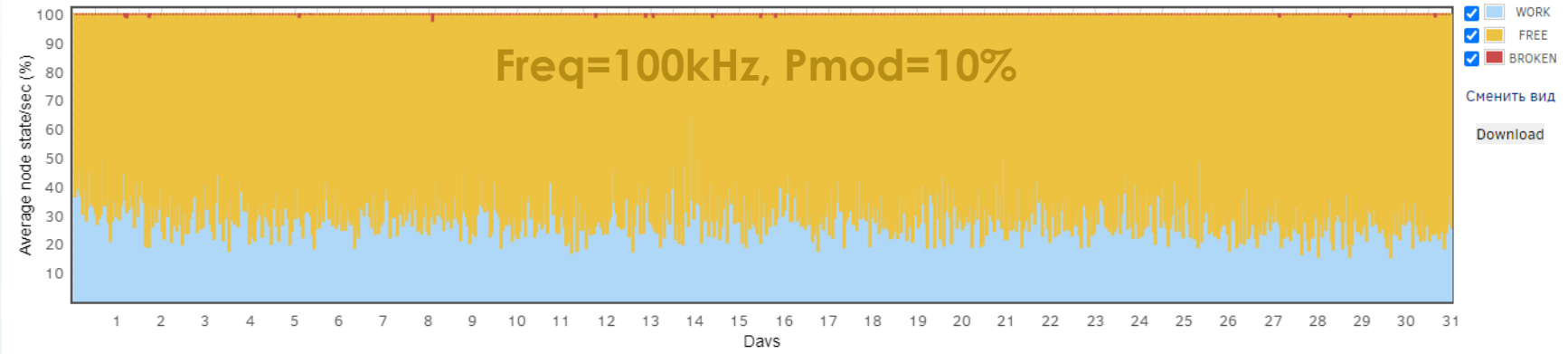
As can be seen on the graph, the required number of nodes directly depends on both the operating mode and the number of incoming modeling tasks. The maximum performance of the HPC system of SCTF at the same time is ~600 TFLOPs.

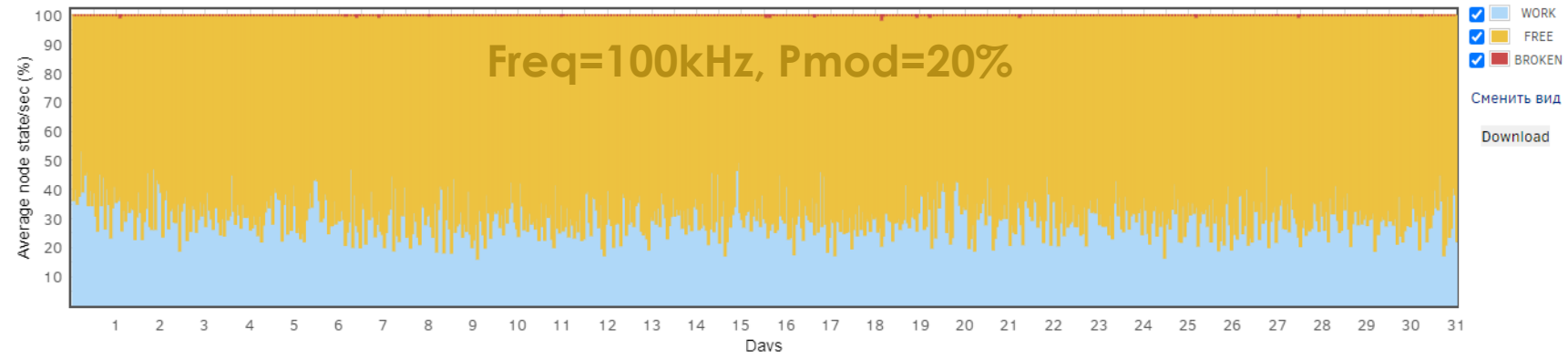


Full-scale modeling. Data flow.

However, as can be seen from the graphs of the average relative number of nodes in different states about 70% (in 100 kHz mode) …



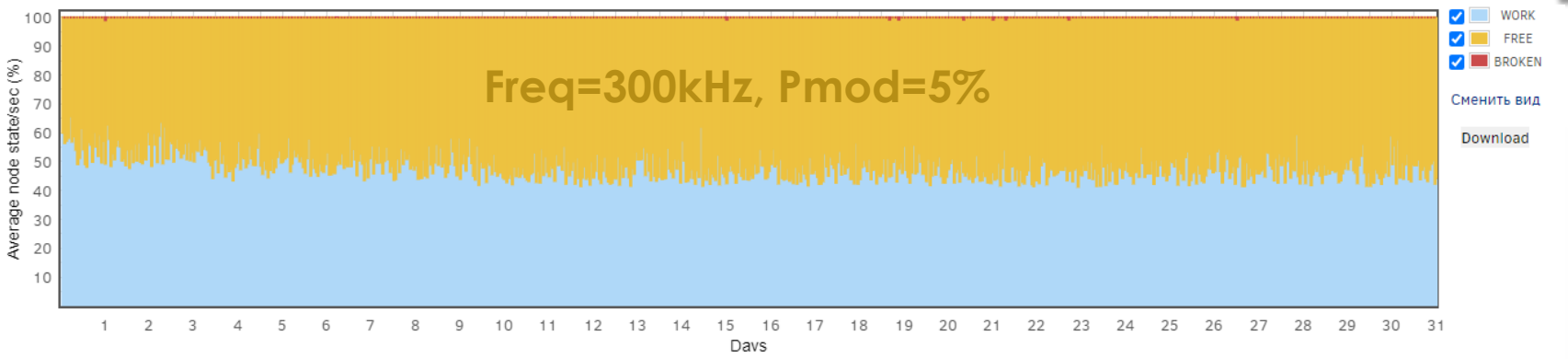


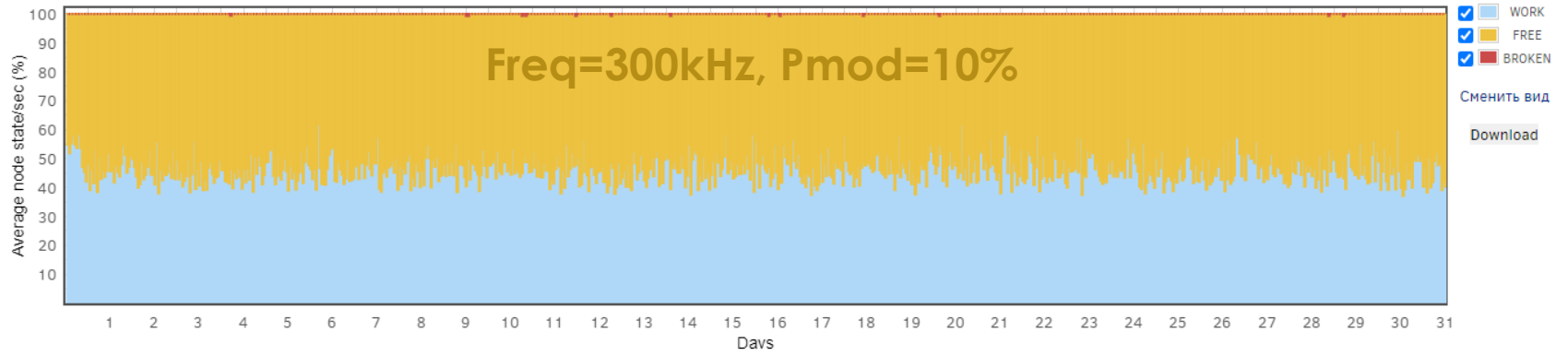


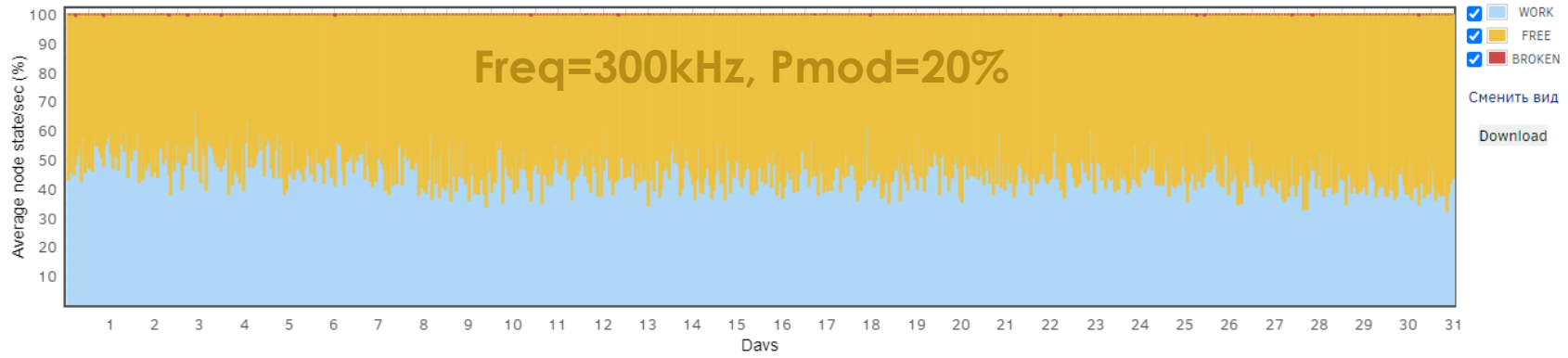
Full-scale modeling. Data flow.

or half (in 300 kHz mode) the nodes are simply idle.

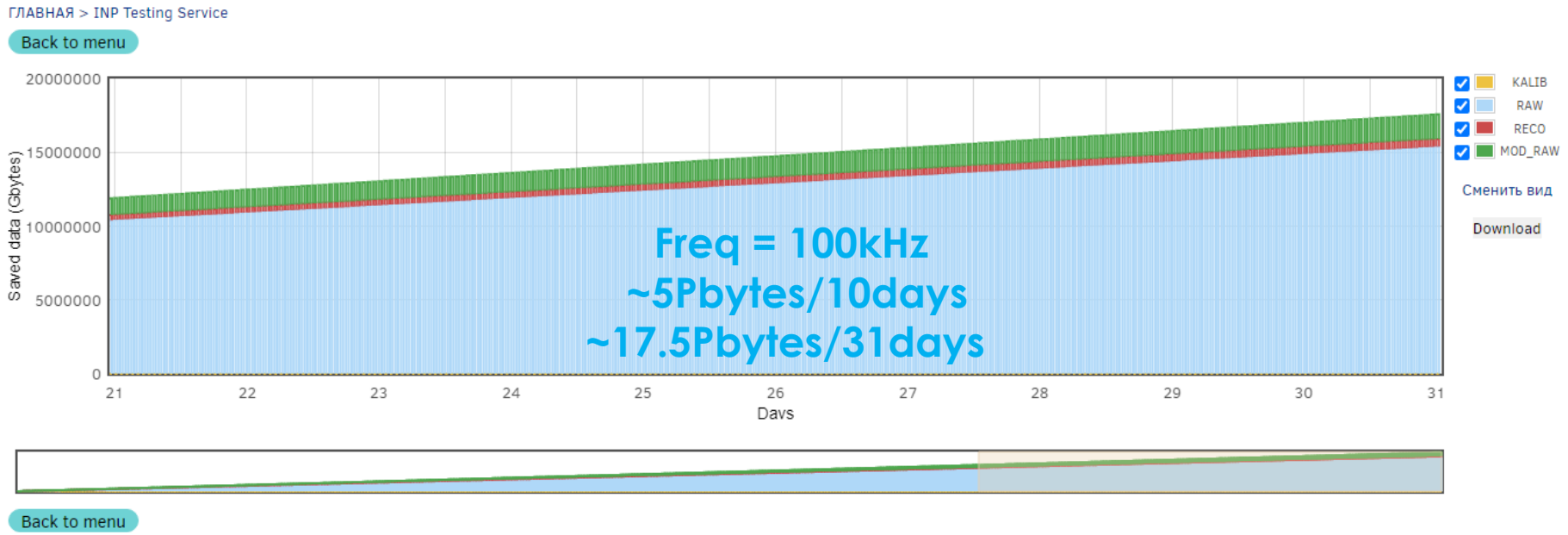
Therefore, it will be necessary to conduct additional model experiments with limited resources in a known range (70-110 nodes).

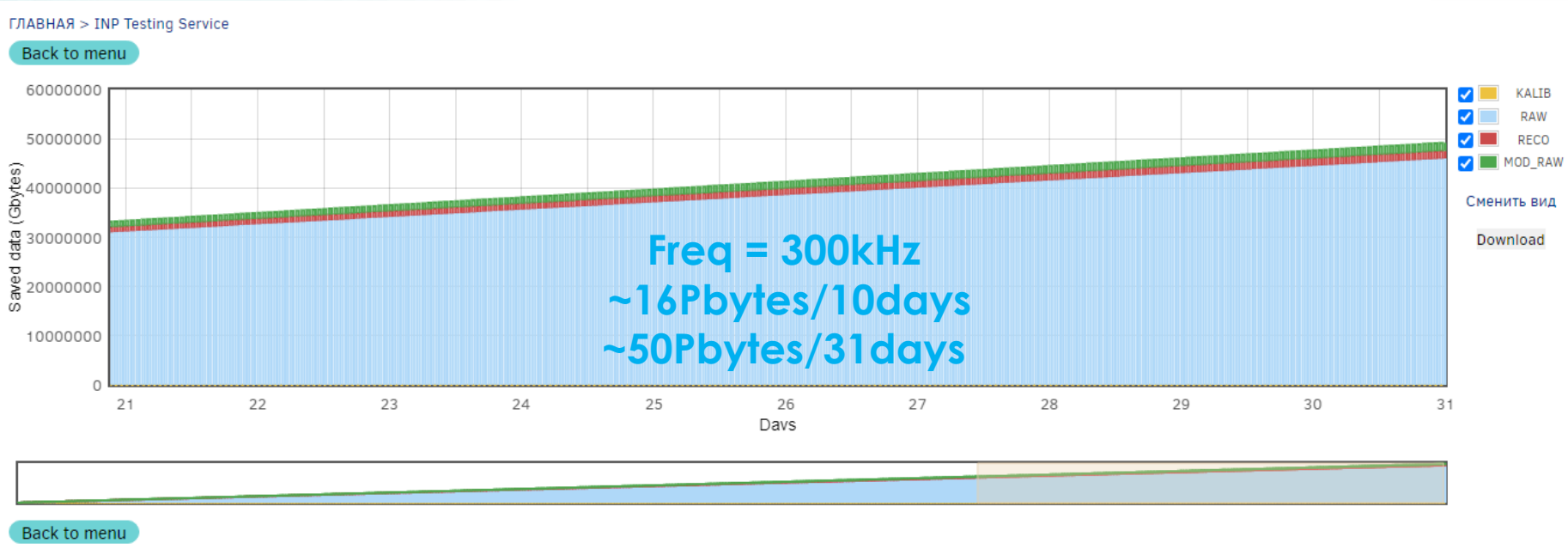






Full-scale modeling. Data capacity.

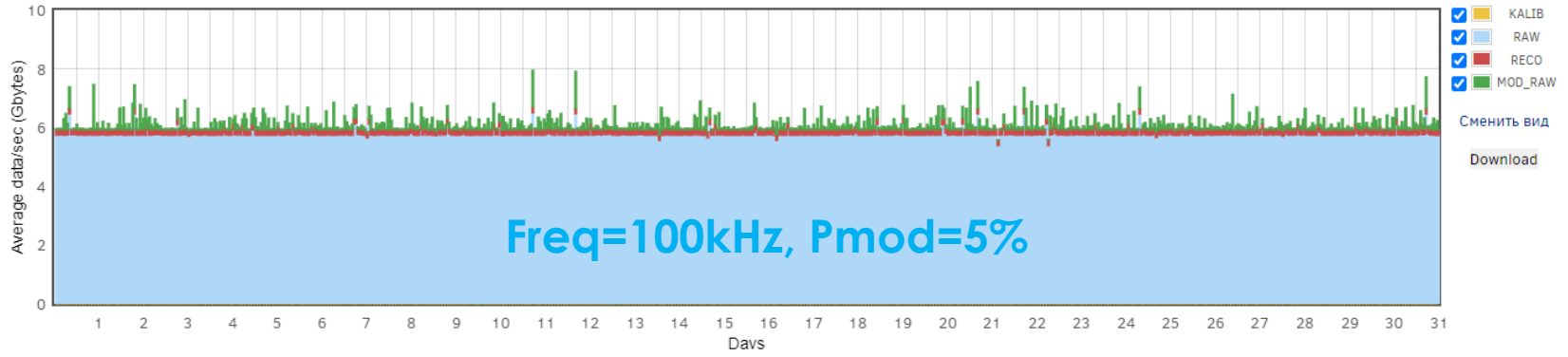


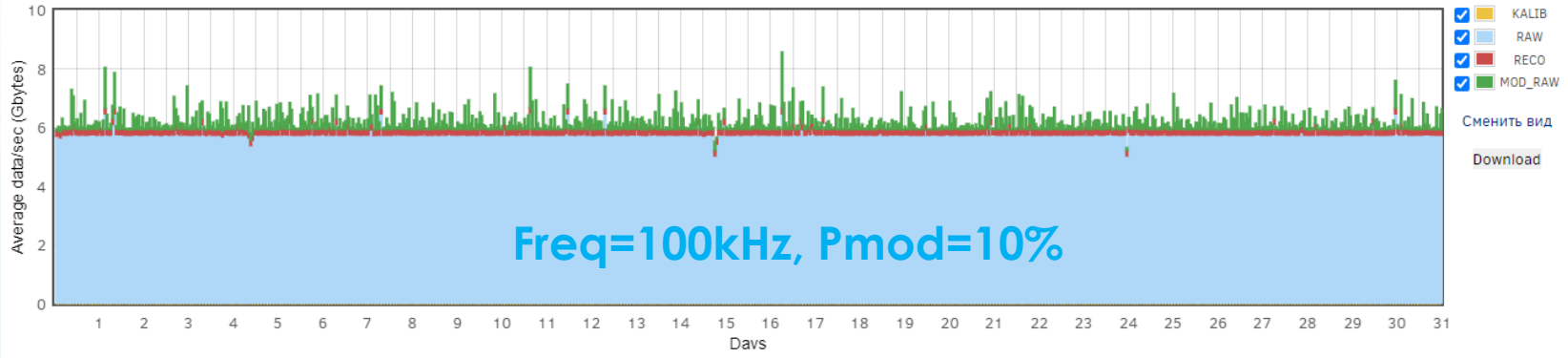


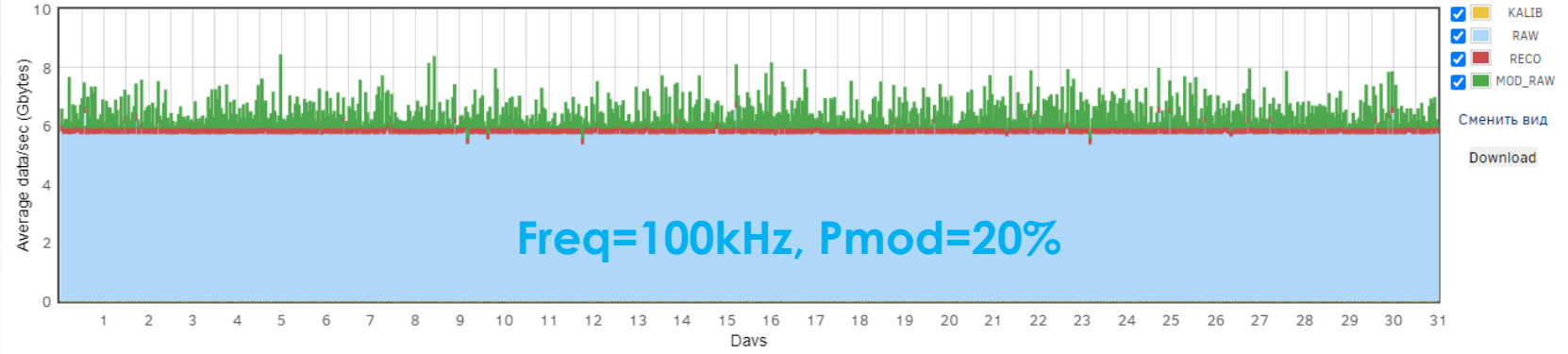
Full-scale modeling. Data flow params.

Average stream:

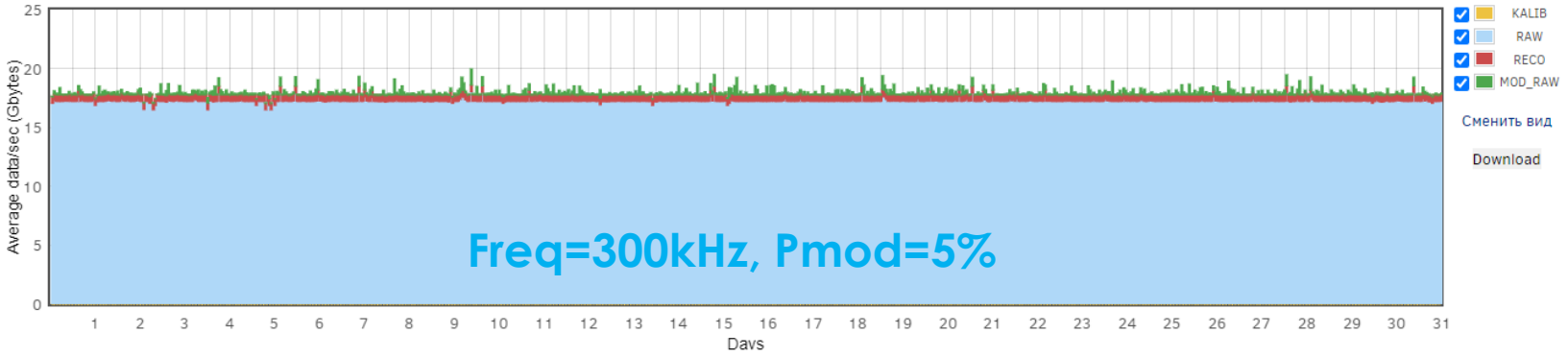
KALIB : 0.01Gb/sec
RAW : 5.8Gb/sec
RECO : 0.2Gb/sec
MRAW : 0.35Gb/sec

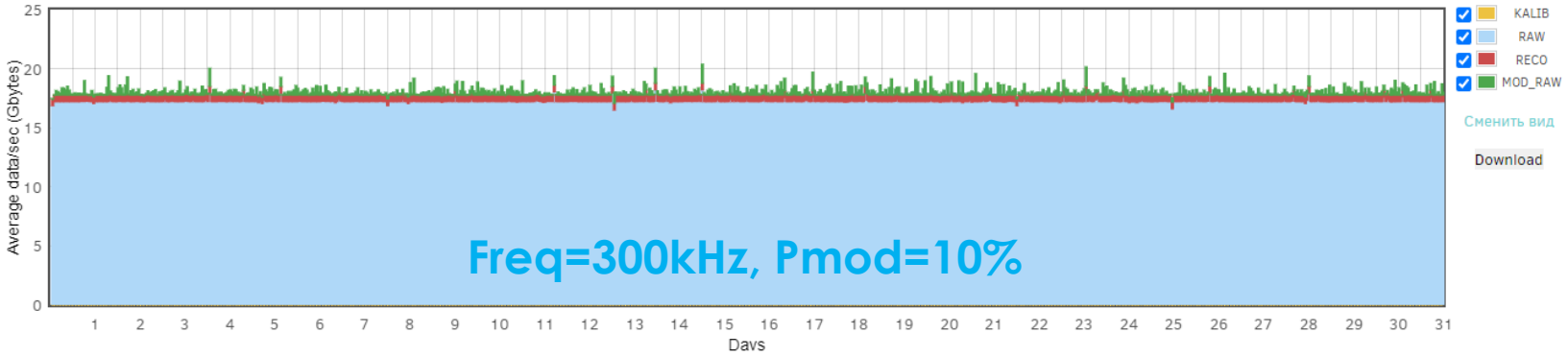


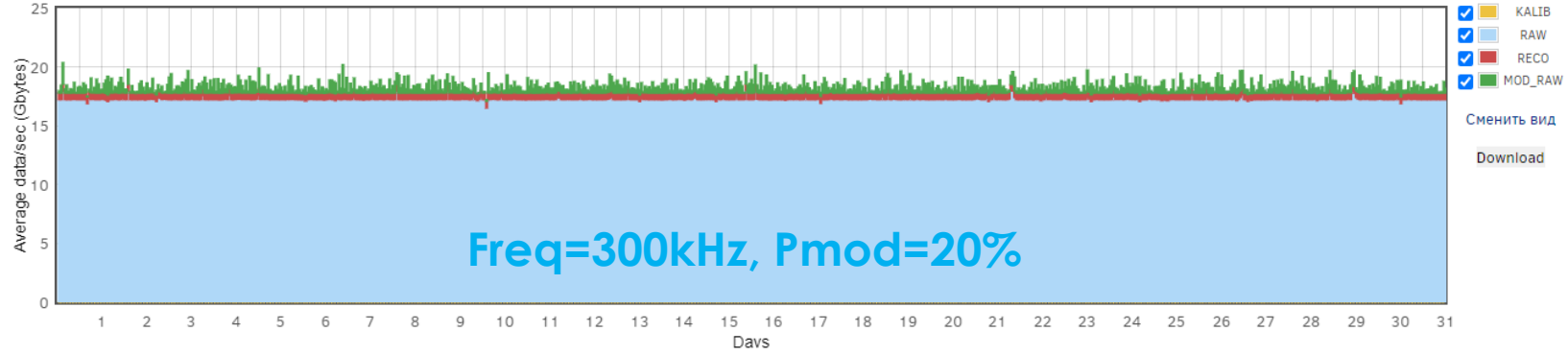




Full-scale modeling. Data flow params.







Average stream:

KALIB : 0.01Gb/sec

RAW : 17.5Gb/sec

RECO : 0.6Gb/sec

MRAW : 0.35Gb/sec

Conclusion

To determine the maximum amount of computing resources and the necessary characteristics of a data storage system for a high-performance computing system serving SCTF, a full-scale simulation of such a system with unlimited resources was carried out. In this simulation, the parameters of data flows and tasks received by the system, as well as the characteristics of modern hardware, were considered.

Model experiments have shown that from 70 to 100 modern computing nodes with performance of about 600TFLOPs are needed for the full operation of the system. However, with this approach, the system's performance is consumed inefficiently, because there is a large percentage of idle nodes. Therefore, in continuation of this work, additional modeling with a limited amount of resources is needed to determine the optimal number of computing nodes (with an acceptable execution latency).

The parameters for the data storage system were also defined. It is necessary to use solutions that allows to organize the storage of hundreds of petabytes of data and a bandwidth from 5 to 20 Gb/s. It is also reasonable to use a combination of different distributed data storage systems for the most efficient use of it.

Thank you for attention