



Joint Institute for Nuclear Research
Meshcheryakov Laboratory of Information Technologies

GRID2023



3-7 July 2023



10th International Conference
“Distributed Computing and Grid Technologies in
Science and Education”

Prototype of a software complex for creating digital twins of distributed data acquisition, storage and processing centers

D. PRIAKHINA, V. KORENKOV, V. TROFIMOV

Meshcheryakov Laboratory of Information Technologies, JINR

Distributed data acquisition, storage and processing centers (DDPC)



- Geographically distributed infrastructure.
- Designed to work with extremely large amounts of data.
- Consists of various types of resources.
- Collective shared access to data storage and processing resources.

Distributed data acquisition, storage and processing centers (DDPC)



Important! The systems must guarantee high-quality and efficient operation.

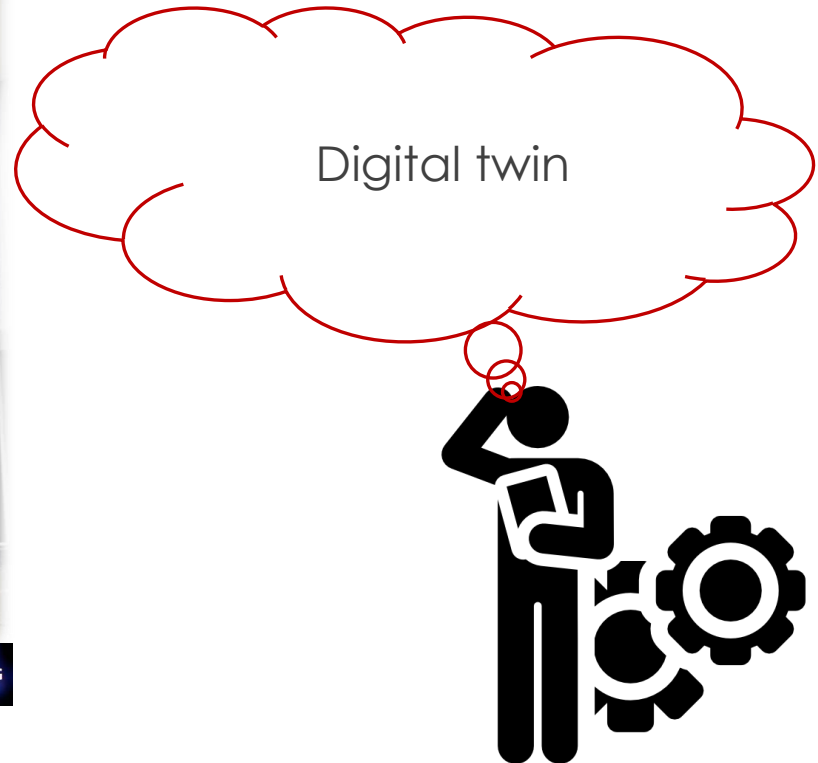
How to ensure continuous improvement and scaling of DDPC?



Distributed data acquisition, storage and processing centers (DDPC)



Important! The systems must guarantee high-quality and efficient operation.



Digital twin (DT)

Real-time operation throughout the entire DDPC life cycle.

COMPUTER MODEL



INPUT DATA

- Architecture and hardware parameters of DDPC.
- Characteristics of data flows and jobs flows.

FUNCTIONAL PURPOSE

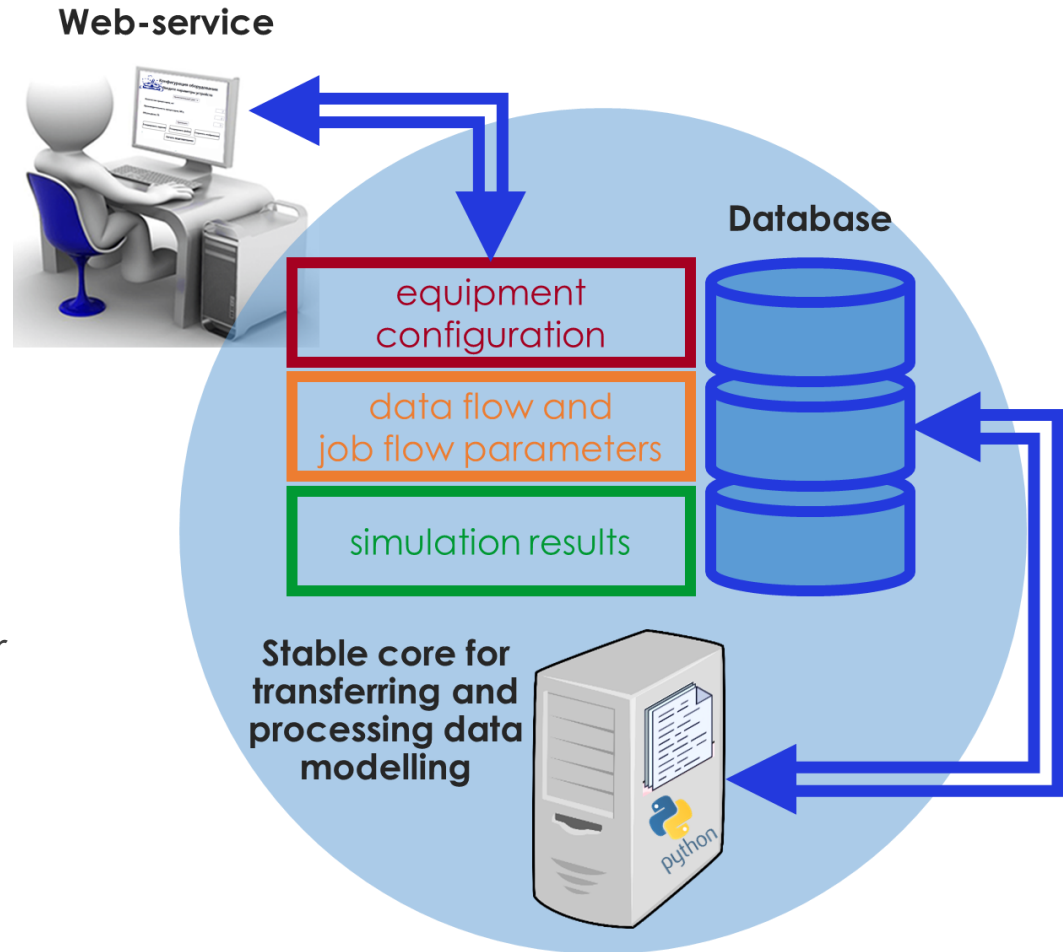
- Designing of DDPC.
- Analysis of the efficiency and reliability of DDPC.
- Testing scaling scenarios based on data flows and jobs flows requirements.
- Assessment of the required amount of resources for specific tasks.
- Checking jobs flows management strategies.

Software complex for creating digital twins of DDPC

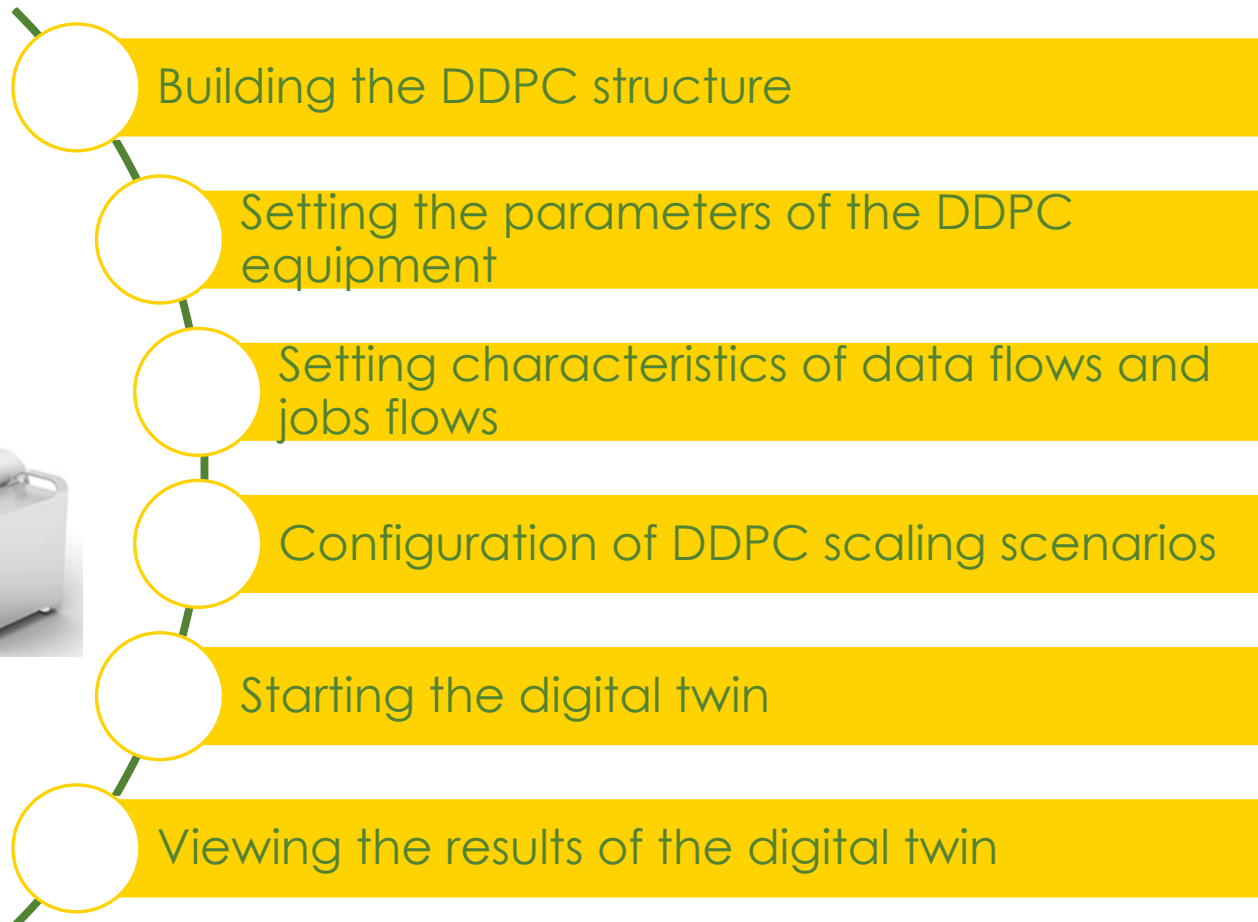


Modelling core

- Universal – applicable for modeling any data center without changing the program code.
- Probabilistic distributions are taken into account when forming data flows, jobs flows, and criteria for the functioning of equipment.
- Used for design tasks, data center scaling during operation, searching for problem areas when data flows and jobs flows change.



Functionality of the web service



Цифровой двойник РЦОД ВМ@N

Построение инфраструктуры центра сбора, хранения и обработки данных



Генератор данных



Вычислительная компонента



Хранилище данных



Роботизированная библиотека

Сохранить изображение

Завершить построение инфраструктуры

Настроить каналы связи

Настроить потоки данных

Настроить потоки задач

Создать цифровой двойник

Цифровой двойник РЦОД ВМ@Н

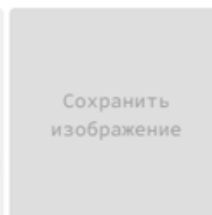
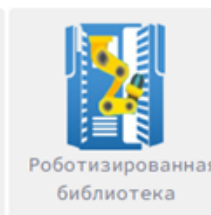
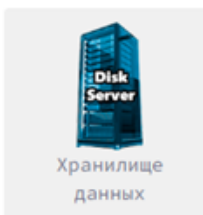
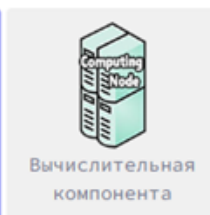
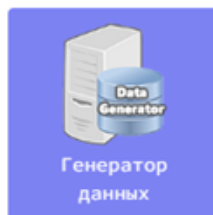
Построение инфраструктуры центра сбора, хранения и обработки данных

Настройка базовой конфигурации оборудования

* Обязательное поле для заполнения

- Название*:
- Описание:
- Средняя скорость генерации данных*: МБ/с

Добавить устройство



Input parameters for the **data generator**:

- name;
- description;
- average data generation rate (MB/sec).



Цифровой двойник РЦОД ВМ@Н

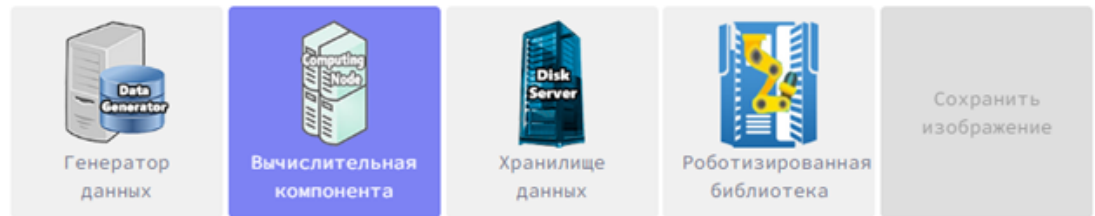
Построение инфраструктуры центра сбора, хранения и обработки данных

Настройка базовой конфигурации оборудования

* Обязательное поле для заполнения

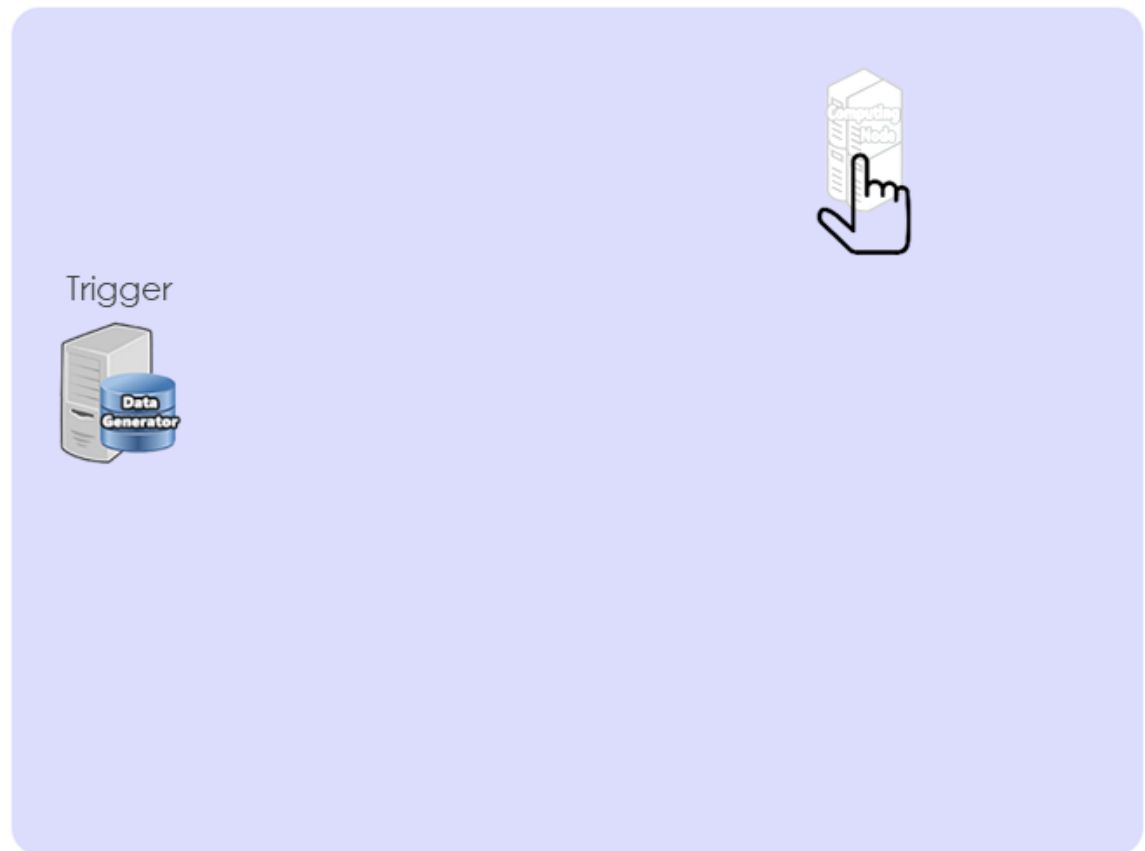
- Название*:
- Описание:
- Количество ядер:

Добавить устройство



Input parameters for the **computing component**:

- name;
- description;
- number of cores.



Цифровой двойник РЦОД ВМ@Н

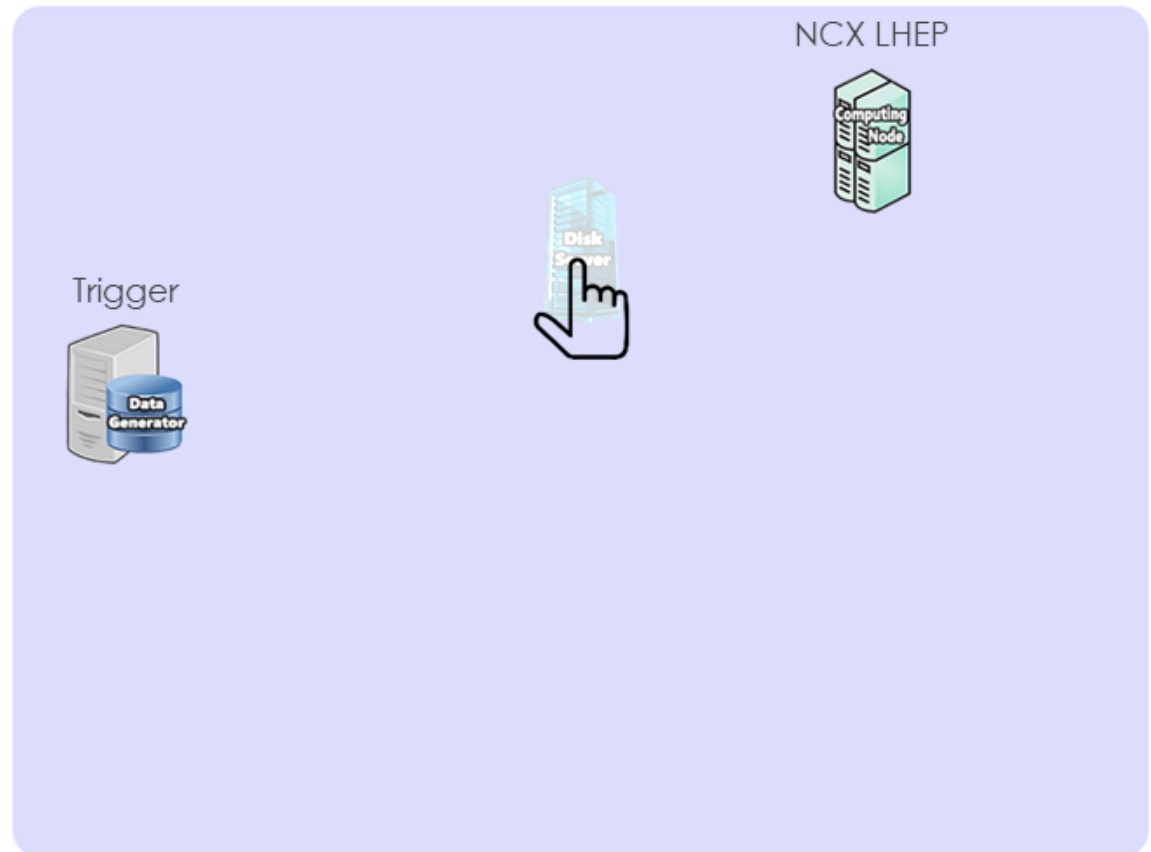
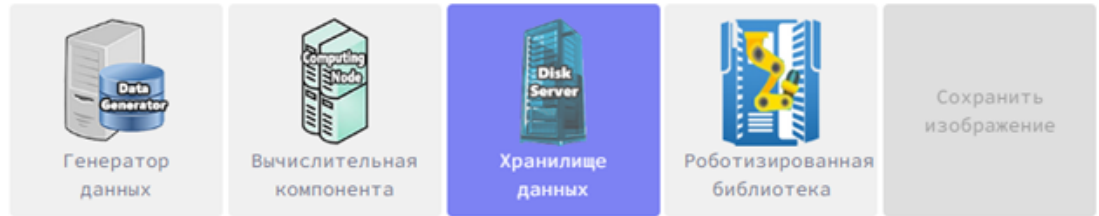
Построение инфраструктуры центра сбора, хранения и обработки данных

Настройка базовой конфигурации оборудования

* Обязательное поле для заполнения

- Название*:
- Описание:
- Объем: ТБ

Добавить устройство

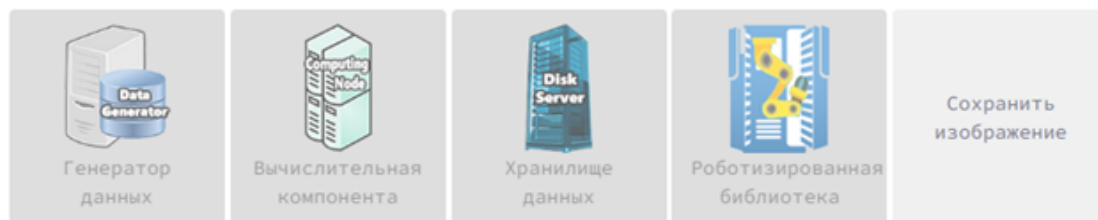


Input parameters for the **data storage**:

- name;
- description;
- volume (TB).

Цифровой двойник РЦОД ВМ@N

Построение инфраструктуры центра сбора, хранения и обработки данных



Редактировать инфраструктуру

Настроить каналы связи

Настроить потоки данных

Настроить потоки задач

Создать цифровой двойник

Настройка каналов связи

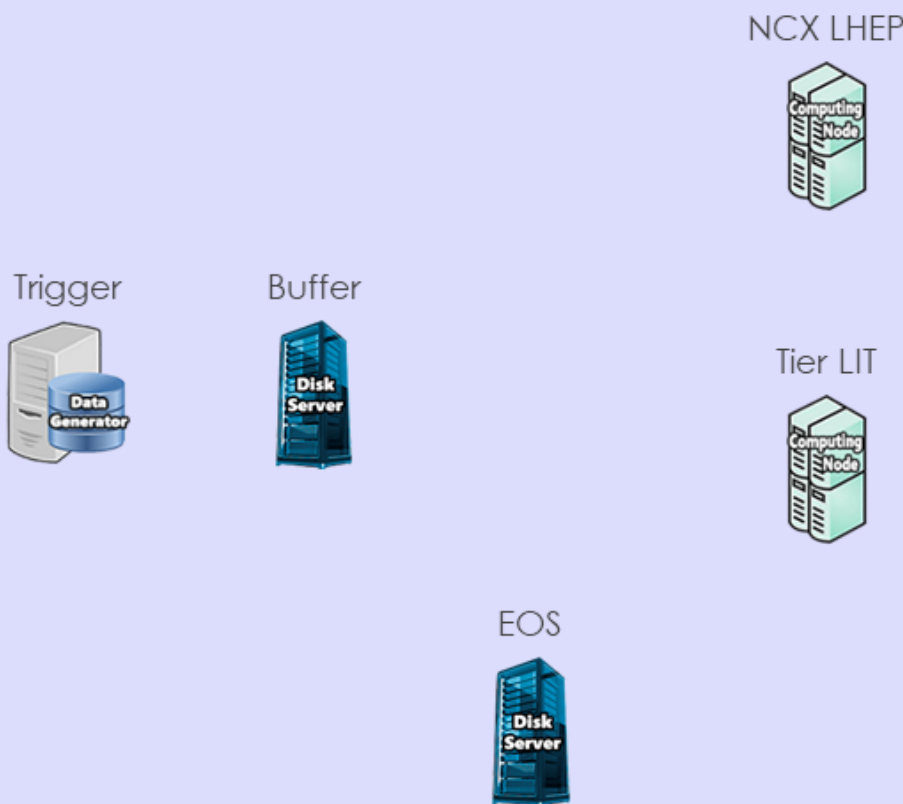
* Обязательное поле для заполнения

- Название*:
- Описание:
- Канал от*: до*:
- Пропускная способность: Гб/с

Добавить

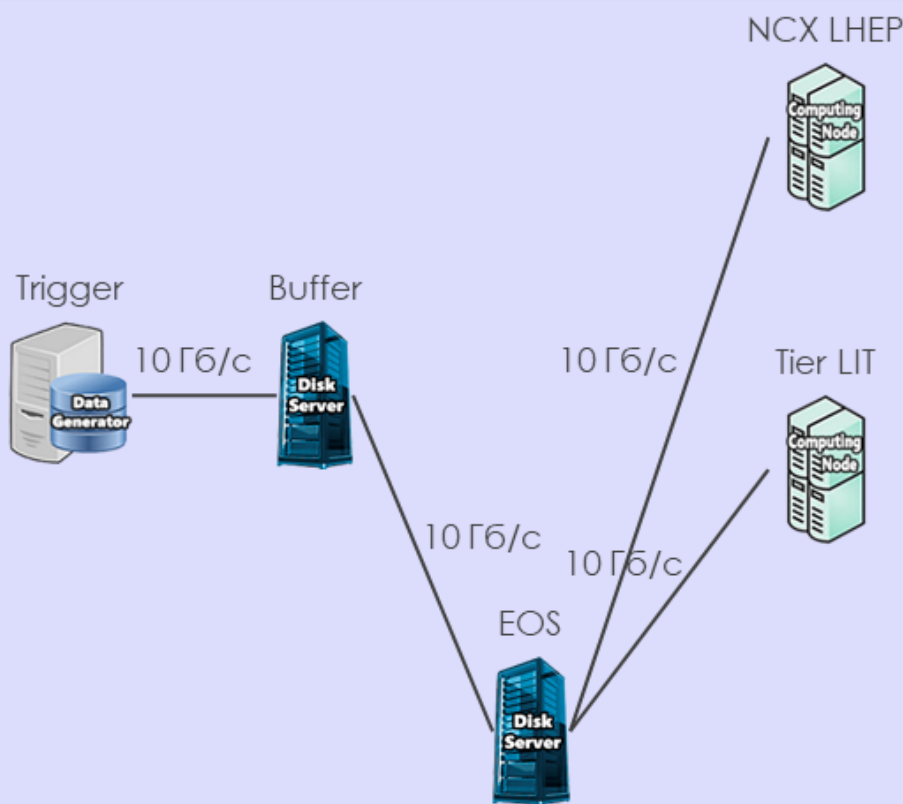
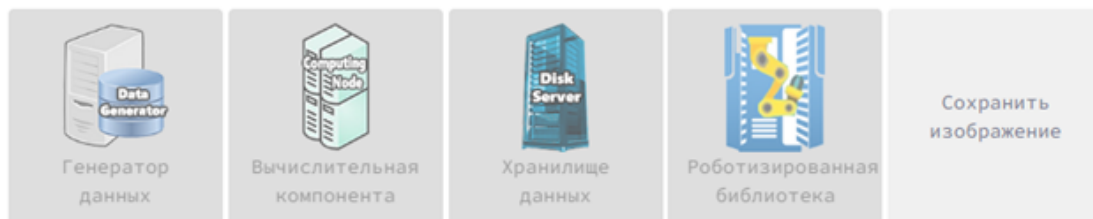
Input parameters for the **communication link**:

- name;
- description;
- objects to connect;
- bandwidth (Gbit/sec).



Цифровой двойник РЦОД ВМ@N

Построение инфраструктуры центра сбора, хранения и обработки данных



Редактировать инфраструктуру

Настроить каналы связи

Настроить потоки данных

Настроить потоки задач

Создать цифровой двойник

Настройка потоков данных

* Обязательное поле для заполнения

• Название*:

raw

• Описание:

data from trigger

• Канал передачи данных*:

link1

• Объем данных для передачи:

100 %

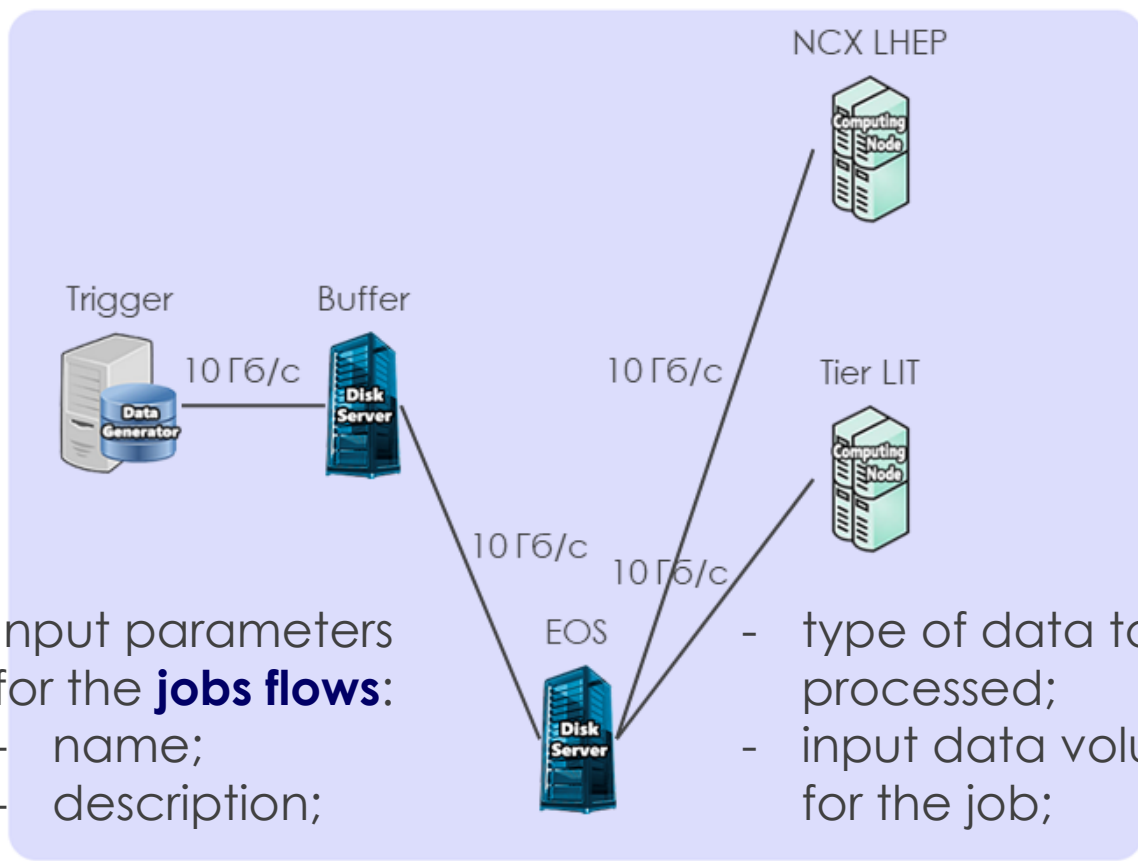
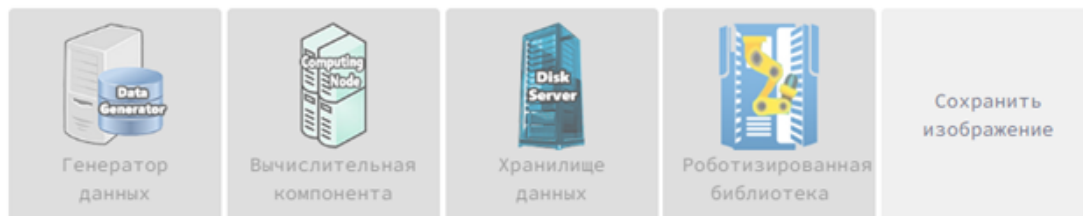
Добавить

Input parameters for the **data flows**:

- name;
- description;
- data transmission channels;
- the amount of data to transfer (%).

Цифровой двойник РЦОД ВМ@N

Построение инфраструктуры центра сбора, хранения и обработки данных



Редактировать инфраструктуру

Настроить каналы связи

Настроить потоки данных

Настроить потоки задач

Создать цифровой двойник

Настройка потоков задач

* Обязательное поле для заполнения

• Название*:

rtd

• Описание:

Raw to Digit

• Тип обрабатываемых данных*: raw

• Объем входных данных для задачи*: 40 Гб

• Тип результирующих данных*: digit

• Объем выходных данных*: 1 Гб

• Среднее время выполнения задачи*: 10 с

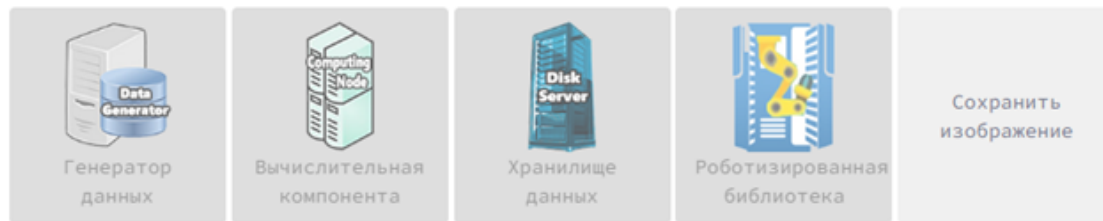
• Общее количество задач * 100

Добавить

- type of output data;
- output data volume;
- average job completion time;
- number of jobs.

Цифровой двойник РЦОД ВМ@N

Построение инфраструктуры центра сбора, хранения и обработки данных



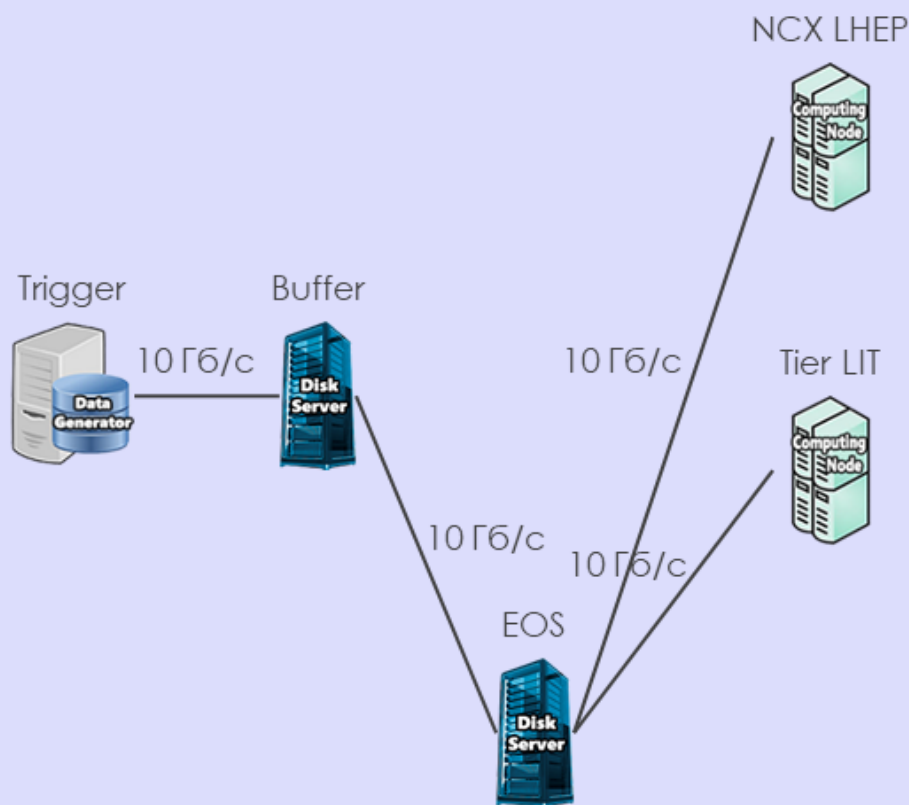
Редактировать инфраструктуру

Настроить каналы связи

Настроить потоки данных

Настроить потоки задач

Создать цифровой двойник



Эксперименты

Выберите существующий эксперимент или добавьте новый эксперимент для поиска оптимальной конфигурации оборудования

Добавить

To start the digital twin:

- add an experiment;
- configure parameters for DDPC modeling.

Цифровой двойник РЦОД ВМ@N

Добавление эксперимента

Заполните поля формы, чтобы добавить новый эксперимент для поиска оптимальной конфигурации оборудования

* Обязательное поле для заполнения

Название эксперимента *

Test 1

Описание эксперимента

Поиск оптимального количества ресурсов для хранения данных

Параметры моделирования

- Продолжительность работы моделируемой инфраструктуры – ч.
- Ускорение процесса моделирования в раз.

Параметры логирования

Выберите объекты и события, о которых необходимо сохранять информацию во время моделирования

• Объекты моделируемой инфраструктуры

- ☒ Хранилища данных
- ☐ Вычислительные компоненты
- ☐ Каналы связи

• События

- ☒ Генерация данных
- ☒ Потери данных
- ☒ Работа с файлами
- ☐ Генерация, запуск, выполнение задач

Parameters for modeling:

- experiment name;
- description;
- duration of DDPC work;
- speed up of modelling;
- objects and events for logging.

Добавить

Очистить

Отмена

Эксперименты

Выберите существующий эксперимент или добавьте новый эксперимент для поиска оптимальной конфигурации оборудования

Test 4

Test

Дата создания: 9 марта 2023 г. 15:04

Test 3

Исследование загрузки каналов связи

Дата создания: 7 февраля 2023 г. 10:42

Test 2

Поиск оптимального количества вычислительных ресурсов

Дата создания: 7 февраля 2023 г. 10:38

Test 1

Поиск оптимального количества ресурсов для хранения данных

Дата создания: 7 февраля 2023 г. 10:36

Добавить

Цифровой двойник РЦОД ВМ@N

Информация об эксперименте

Дата создания: 7 февраля 2023 г. 10:36

Название эксперимента

Test 1

Описание эксперимента

Поиск оптимального количества ресурсов для хранения данных

Параметры моделирования

- Продолжительность работы моделируемой инфраструктуры – 800 ч.
- Ускорение процесса моделирования в 1000 раз.

Параметры логирования

- Объекты моделируемой инфраструктуры

- ☒ Хранилища данных
- ☐ Вычислительные компоненты
- ☐ Каналы связи

- События

- ☒ Генерация данных
- ☒ Потери данных
- ☒ Работа с файлами
- ☐ Генерация, запуск, выполнение задач

Посмотреть результаты

Выбрать другой эксперимент

Базовая конфигурация

Хранилища данных

Название	Описание	Объем (ТБ)
trigger	Trigger ВМ@N	10000,0
buffer	Data reception buffer	5400,0
eoslhep	Main storage LHEP	1000,0
eoslit	Main storage LIT	1000,0
dcach	pp	1000,0

Вычислительные компоненты

Название	Описание	Количество ядер
t2lit	LIT T2 farm	500
ncxlhep	LHEP main farm	1200
super	Govoron	190

Каналы связи

Название	Описание	Пропускная способность (Гб/с)
raw0	trigger - buffer	100,0
raw1	buffer - lhep	10,0
raw2	buffer - lit	10,0
compute0	lhep - farm lhep	10,0
compute1	lit - Govoron	10,0
compute2	lit - farm lit	10,0
dataeoslhepLit	eoslhep - eoslit	10,0
dataeosLitLhep	eoslit - eoslhep	10,0

Добавить модификацию

Цифровой двойник РЦОД ВМ@N

Добавление модификации для эксперимента Test 1

Измените значения базовой конфигурации и сохраните модификацию

Хранилища данных

Название	Описание	Объем (ТБ)
trigger	Trigger ВМ@N	10000
buffer	Data receptipon buffer	5400
eoslhep	Main storage LHEP	1000
eoslit	Main storage LIT	1000
dcach	pp	1000

Вычислительные компоненты

Название	Описание	Количество ядер
t2lit	LIT T2 farm	500
ncxlhep	LHEP main farm	1200
super	Govorun	190

Каналы связи

Название	Описание	Пропускная способность (Гб/с)
raw0	trigger - buffer	100
raw1	buffer - lhep	10
raw2	buffer - lit	10
compute0	lhep - farm lhep	10
compute1	lit - Govorun	10
compute2	lit - farm lit	10
dataeosLhepLit	eoslhep - eoslit	10
dataeosLitLhep	eoslit - eoslhep	10

Сохранить

Очистить

Отмена

Цифровой двойник РЦОД ВМ@N

Информация об эксперименте

Дата создания: 7 февраля 2023 г. 10:36

Название эксперимента

Test 1

Описание эксперимента

Поиск оптимального количества ресурсов для хранения данных

Параметры моделирования

- Продолжительность работы моделируемой инфраструктуры – 800 ч.
- Ускорение процесса моделирования в 1000 раз.

Параметры логирования

- Объекты моделируемой инфраструктуры
 - ☒ Хранилища данных
 - ☐ Вычислительные компоненты
 - ☐ Каналы связи
- События
 - ☒ Генерация данных
 - ☒ Потери данных
 - ☒ Работа с файлами
 - ☐ Генерация, запуск, выполнение задач

Посмотреть результаты

Выбрать другой эксперимент

Базовая конфигурация

Хранилища данных

Название	Описание	Объем (ТБ)
trigger	Trigger ВМ@N	10000,0
buffer	Data reception buffer	5400,0
eosllep	Main storage LHEP	1000,0
eoslit	Main storage LIT	1000,0
dcach	pp	1000,0

Вычислительные компоненты

Название	Описание	Количество ядер
t2lit	LIT T2 farm	500
ncxllep	LHEP main farm	1200
super	Govoron	190

Каналы связи

Название	Описание	Пропускная способность (Гб/с)
raw0	trigger - buffer	100,0
raw1	buffer - llep	10,0
raw2	buffer - lit	10,0
compute0	llep - farm llep	10,0
compute1	lit - Govoron	10,0
compute2	lit - farm lit	10,0
dataeosllepLit	eosllep - eoslit	10,0
dataeosLitllep	eoslit - eosllep	10,0

Добавить модификацию

Список модификаций

№	Статус	Дата обновления	
16	NEW	9 марта 2023 г. 14:52	Просмотр Запуск Результаты
15	DONE	10 марта 2023 г. 10:18	Просмотр Запуск Результаты

Simultaneous run of all modifications is possible

Цифровой двойник РЦОД ВМ@N

Результаты эксперимента Test 1

Выберите вкладку для просмотра результатов

Хранилища данных

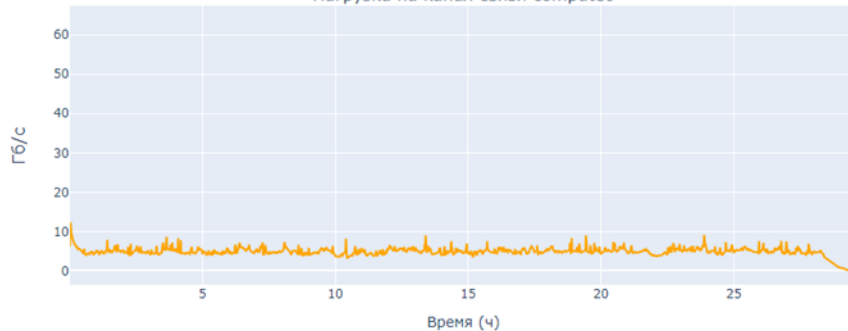
Вычислительные компоненты

Каналы связи

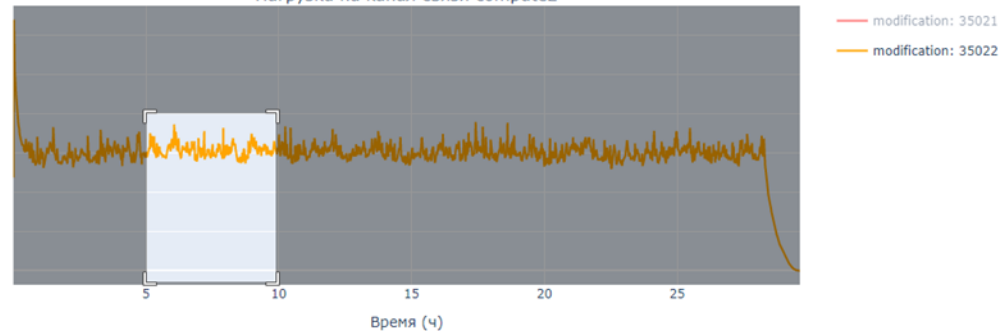
Очереди задач

Распределения файлов

Нагрузка на канал связи compute0



Нагрузка на канал связи compute2



Available for viewing:

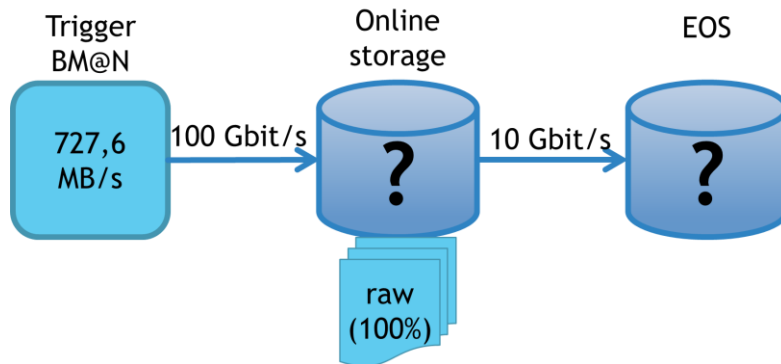
- data storage load volume;
- using cores on computing components;
- load on communication links;
- job queues, the number of completed jobs;
- distribution of files in storages.

Example of software complex application

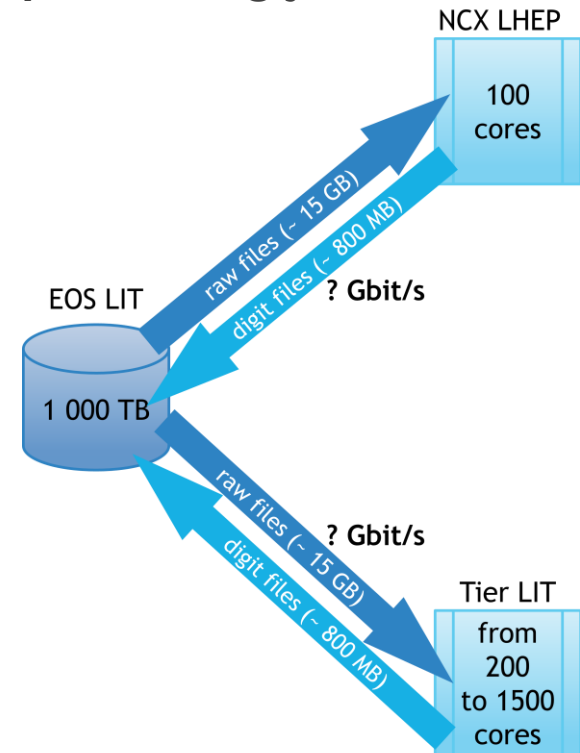


BM@N computing infrastructure (Run 8)

Experimental data acquisition
and storage



Running experimental data
processing jobs

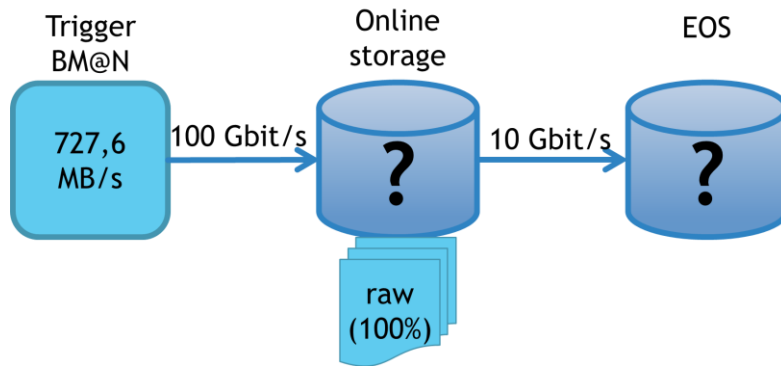


Example of software complex application



BM@N computing infrastructure (Run 8)

Experimental data acquisition and storage



Experiment description

- Duration ≈ 750 hours
- 1 raw file ≈ 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.

Example of software complex application



BM@N computing infrastructure (Run 8)

Experimental data acquisition and storage

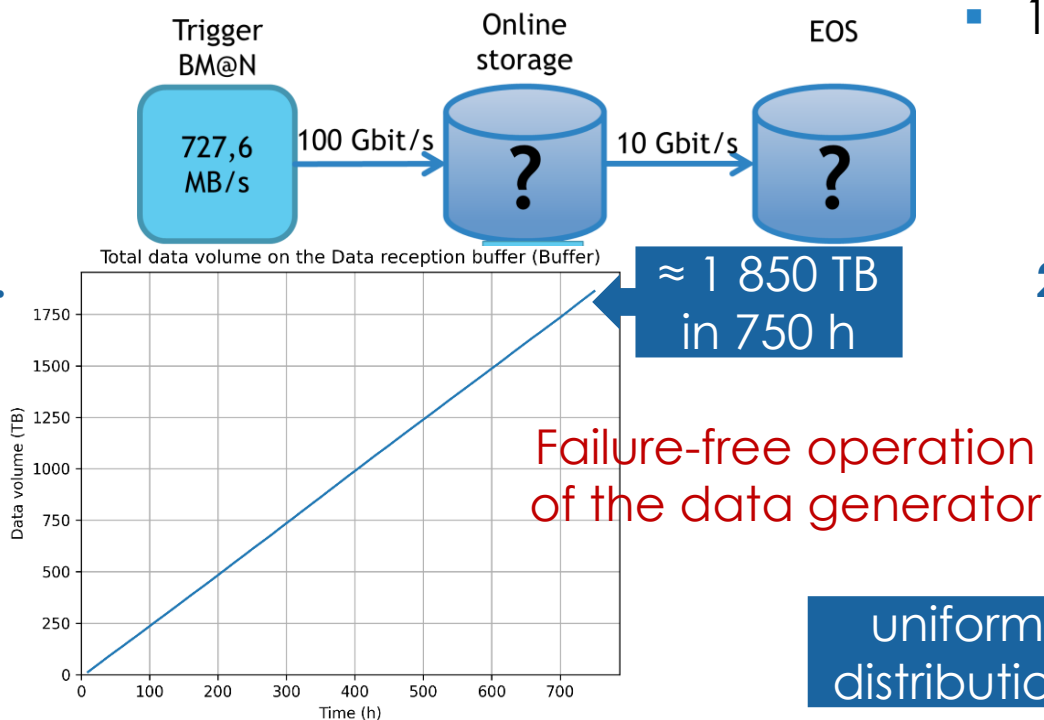
Experiment description

- Duration ≈ 750 hours
- 1 raw file ≈ 15 GB

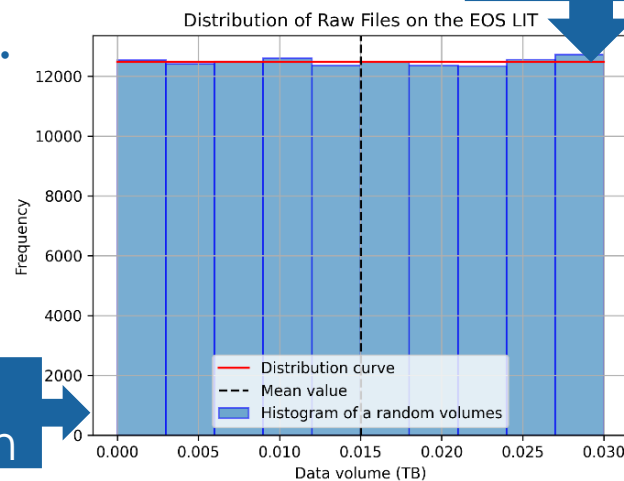
Results

$\approx 125\,000$ raw files

1.



2.



Example of software complex application



BM@N computing infrastructure (Run 8)

Experimental data acquisition and storage

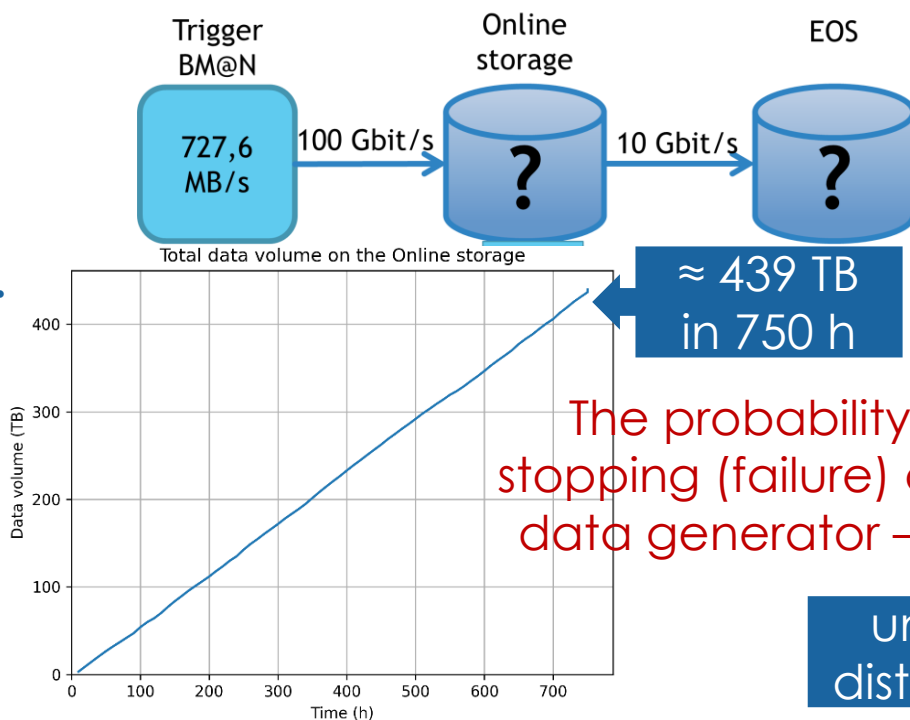
Experiment description

- Duration ≈ 750 hours
- 1 raw file ≈ 15 GB

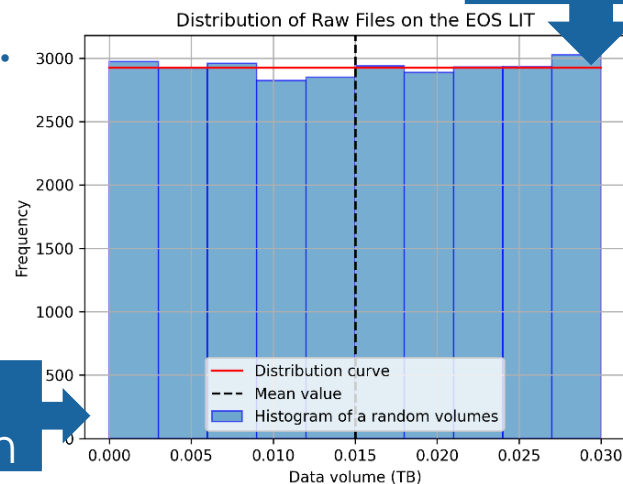
Results

$\approx 29\,241$
raw files

1.



2.



Example of software complex application



BM@N computing infrastructure (Run 8)

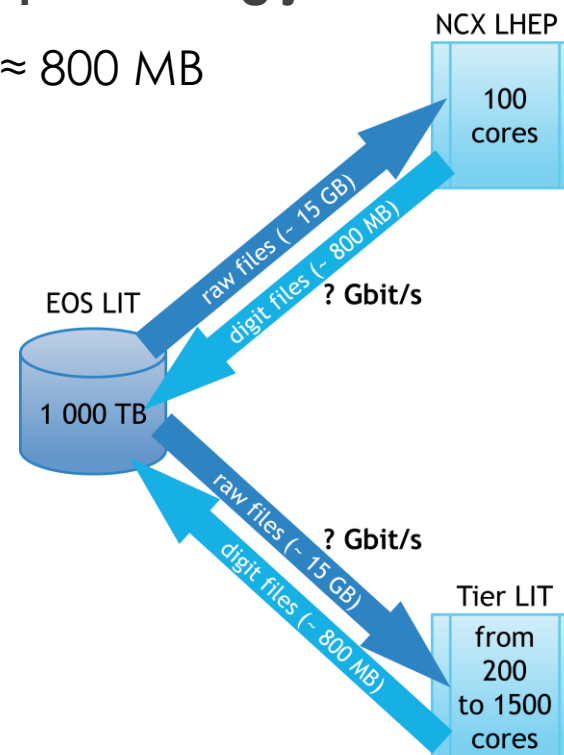
Experiment description

- Number of jobs – 31 306
- 1 raw (input) file \approx 15 GB; 1 digit (output) file \approx 800 MB
- Time to complete a 1 job \approx 2 500 cek

Running experimental data processing jobs

Tasks

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



Example of software complex application



BM@N computing infrastructure (Run 8)

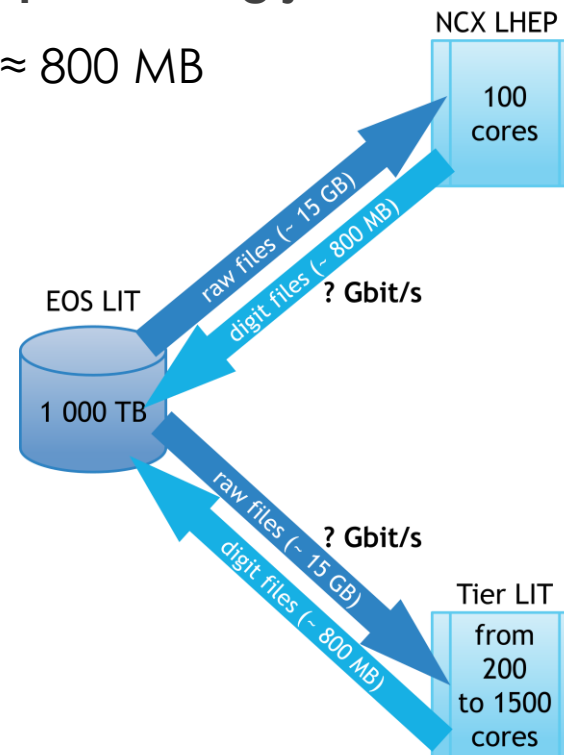
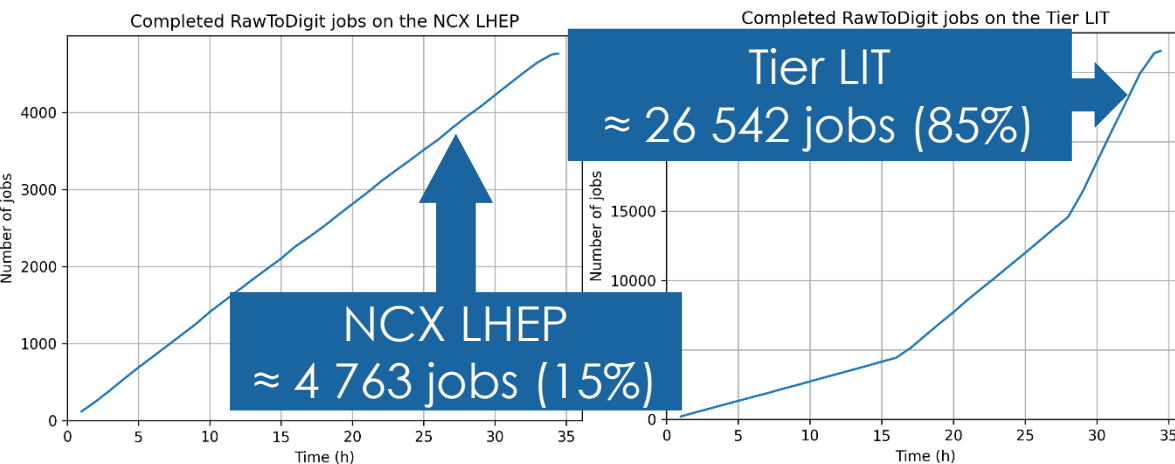
Experiment description

- Number of jobs – 31 306
- 1 raw (input) file \approx 15 GB; 1 digit (output) file \approx 800 MB
- Time to complete a 1 job \approx 2 500 cek

Running experimental data processing jobs

Results

1. Processing time of all jobs \approx 35 hours



Example of software complex application



BM@N computing infrastructure (Run 8)

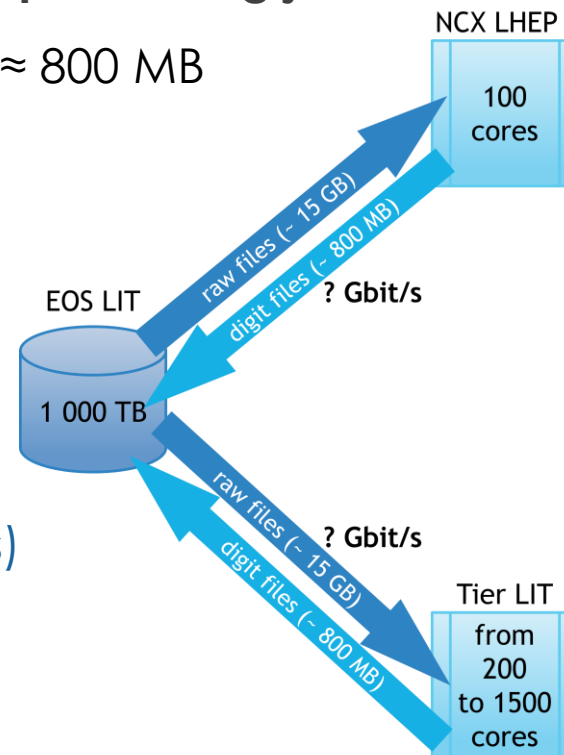
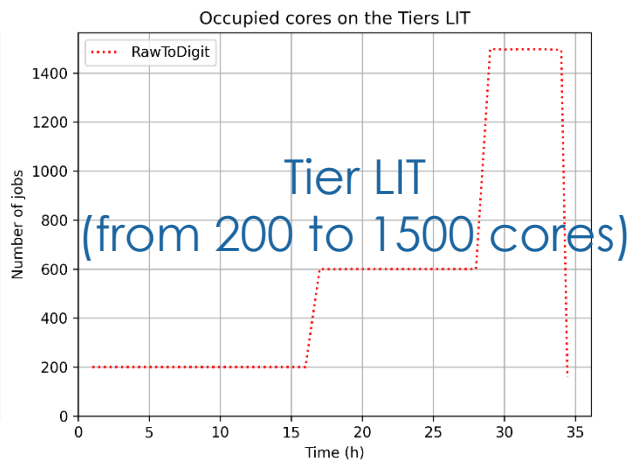
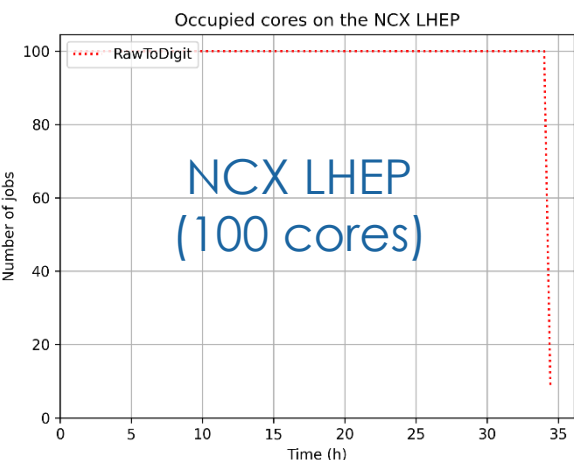
Experiment description

- Number of jobs – 31 306
- 1 raw (input) file \approx 15 GB; 1 digit (output) file \approx 800 MB
- Time to complete a 1 job \approx 2 500 cek

Running experimental data processing jobs

Results

2. Use of computing resources



Example of software complex application



BM@N computing infrastructure (Run 8)

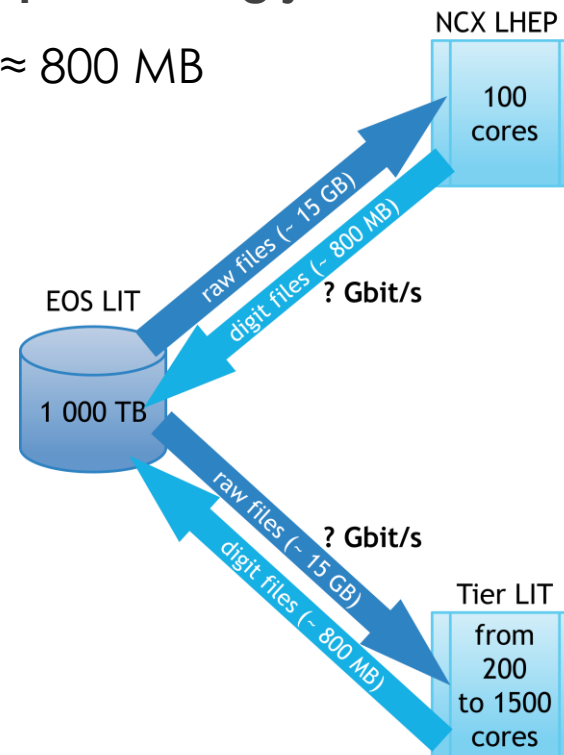
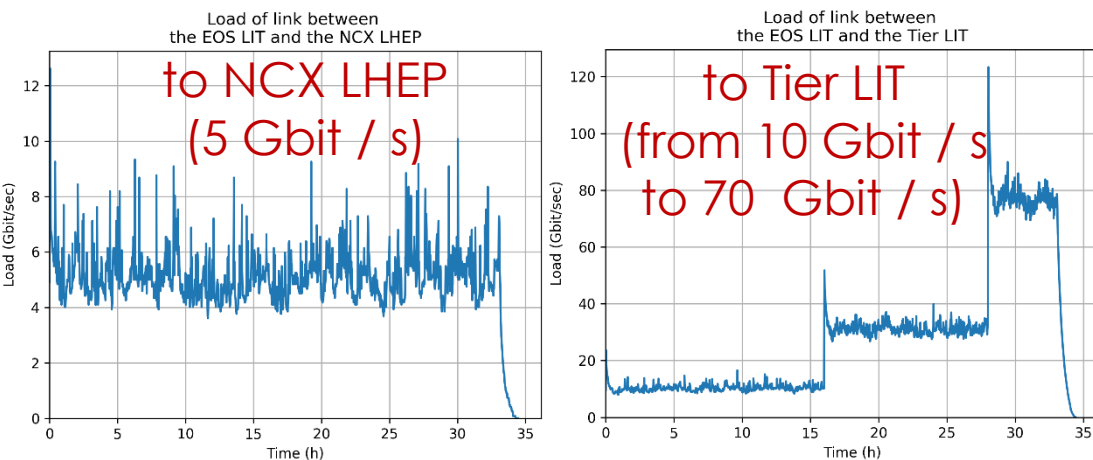
Experiment description

- Number of jobs – 31 306
- 1 raw (input) file \approx 15 GB; 1 digit (output) file \approx 800 MB
- Time to complete a 1 job \approx 2 500 cek

Running experimental data processing jobs

Results

3. Load of communication links



Conclusions

- A prototype of a software complex has been developed to create digital twins of distributed data acquisition, storage and processing centers:
 - database;
 - modelling program (**successful approbation**);
 - web-service (prototype: building DT, setting configurations, starting DT, viewing results).
- Approbation during the construction of the distributed computing infrastructure of the BM@N experiment of the NICA project.
- The modeling program takes into account:
 - the probability of failures and changes in equipment parameters;
 - requirements for stored data flows;
 - requirements for data processing jobs flows.
- Planning to use other experiments of the NICA project for the design of computing infrastructures.

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



Thank you for the attention!

D. PRIAKHINA, V. KORENKOV, V. TROFIMOV

Meshcheryakov Laboratory of Information Technologies, JINR

04.07.2023

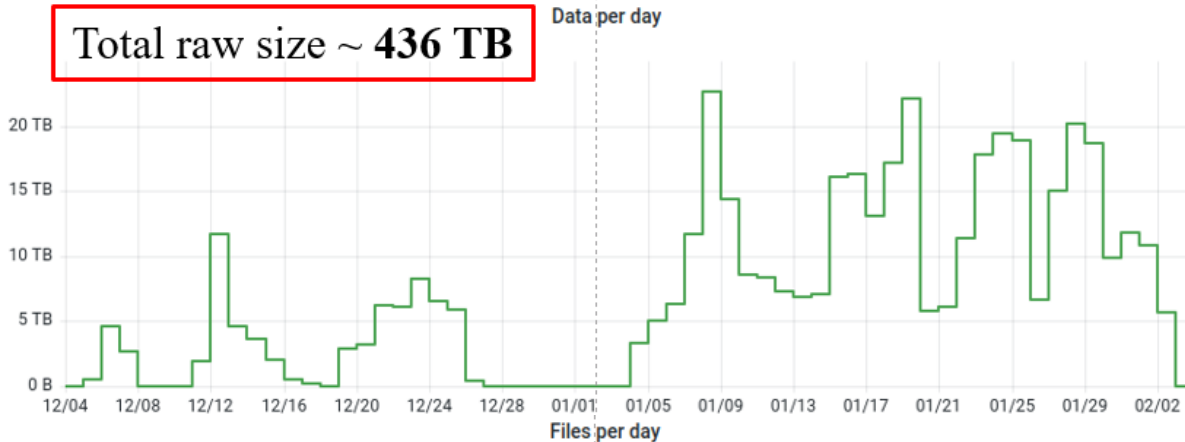
Implementers a web service prototype



Monitoring VS Digital Twin

The probability of
stopping (failure) of the
data generator – 80%

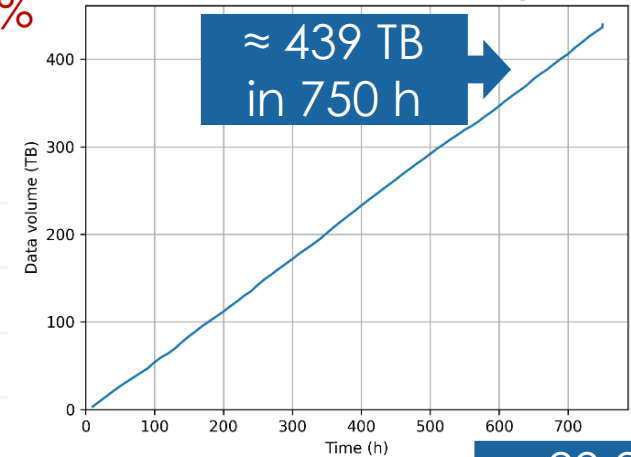
Total raw size ~ **436 TB**



Total files: **31306**

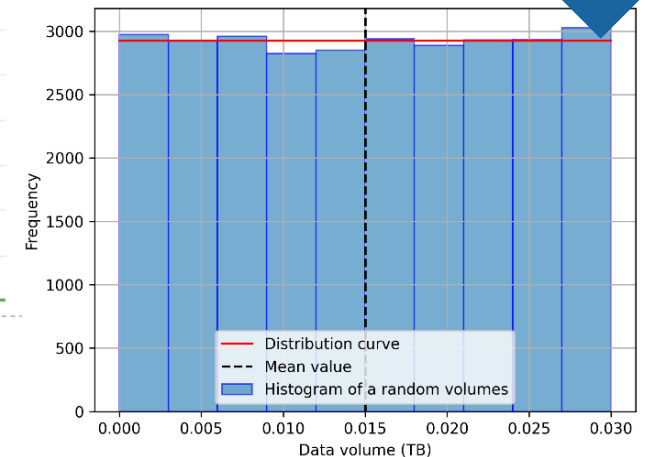


Total data volume on the Online storage



≈ 29 241
raw files

Distribution of Raw Files on the EOS LIT

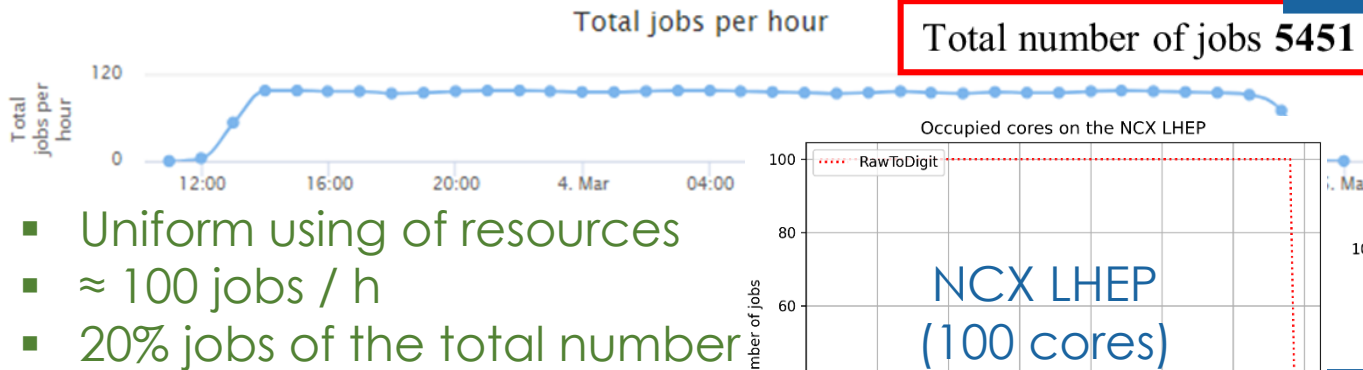


- Non uniform distribution.
- No data is received at all by some periods.

Monitoring VS Digital Twin

NCX LHEP (100 cores) Processing time of all jobs ≈ 35 h

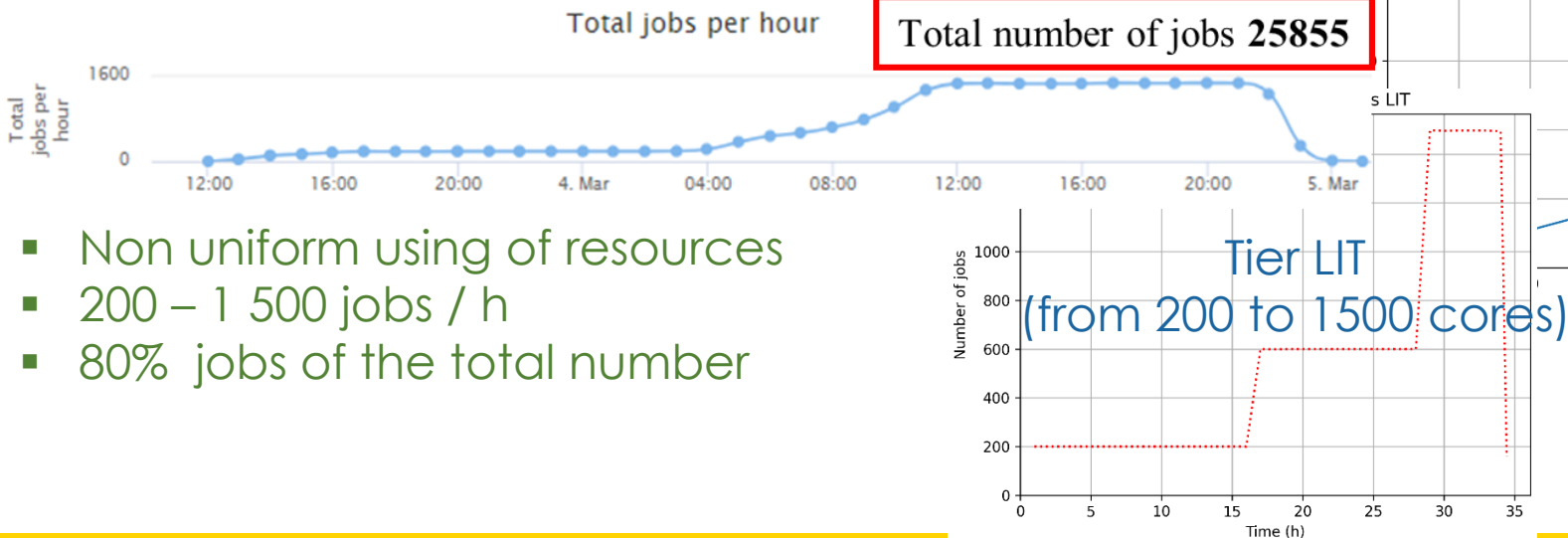
NCX LHEP
 $\approx 4\,763$ jobs (15%)



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

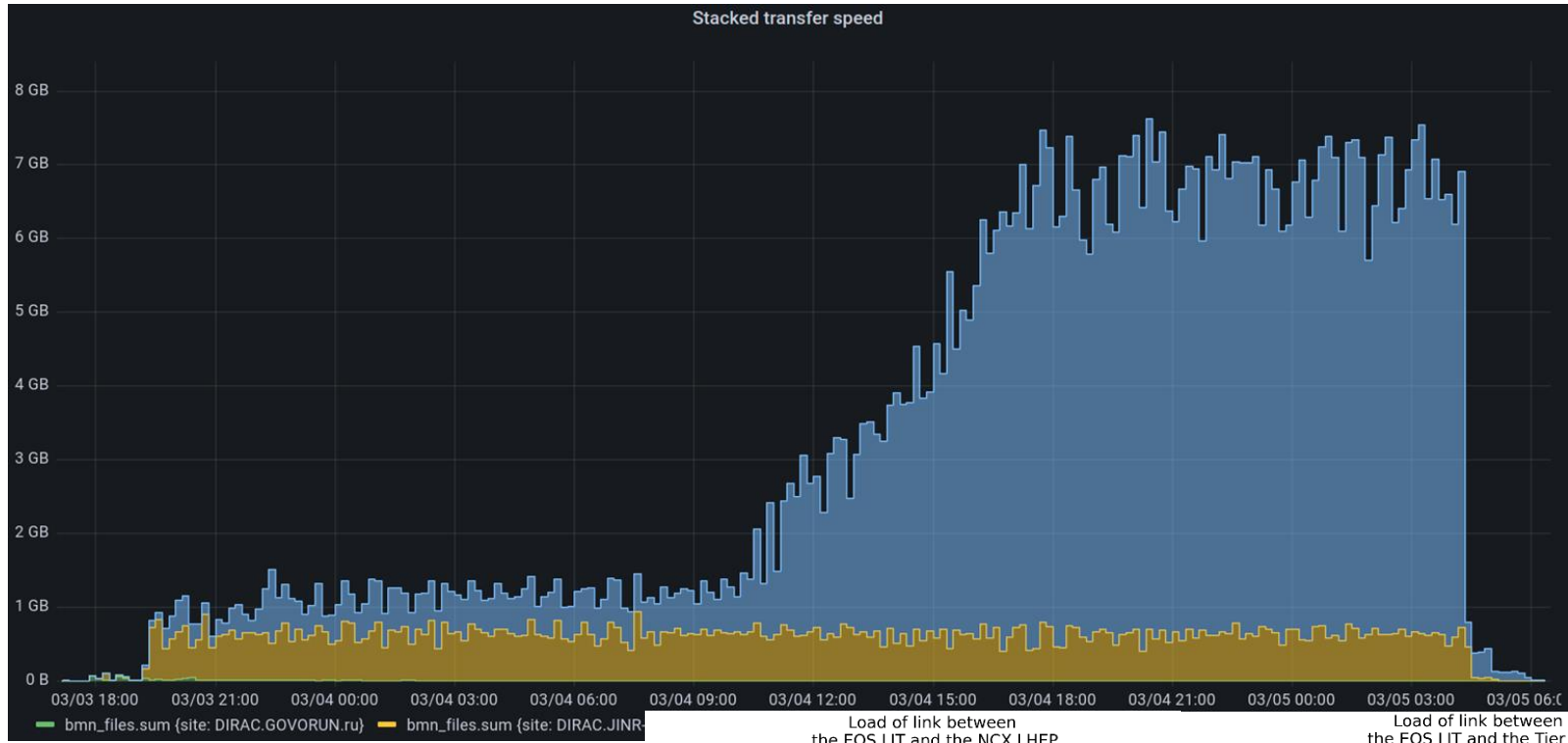
Tier LIT (from 200 to 1500 cores)

Tier LIT
 $\approx 26\,542$ jobs (85%)



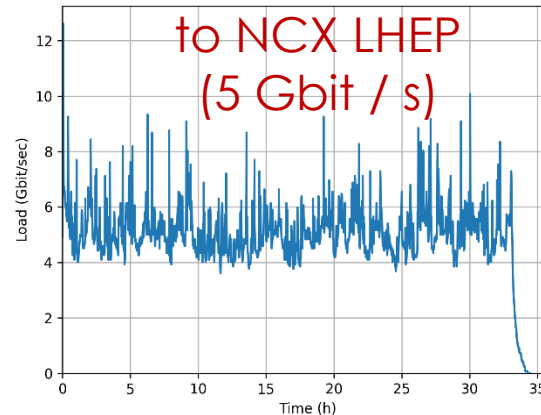
- Non uniform using of resources
- 200 – 1 500 jobs / h
- 80% jobs of the total number

Monitoring VS Digital Twin



- Avg. data transfer rate to NCX LHEP ≈ 4 Gbit / s
- Avg. data transfer rate to Tier LIT from 8 Gbit / s to 64 Gbit / s

Load of link between the EOS LIT and the NCX LHEP



Load of link between the EOS LIT and the Tier LIT

