11th International Conference "Distributed Computing and Grid Technologies in Science and Education"

Software Complex for Distributed Data Processing during Run 9 of the BM@N experiment

Konstantin Gertsenberger
BM@N Software Coordinator
V. Veksler and A. Baldin Laboratory of High Energy Physics
Joint Institute for Nuclear Research

July 8, 2025

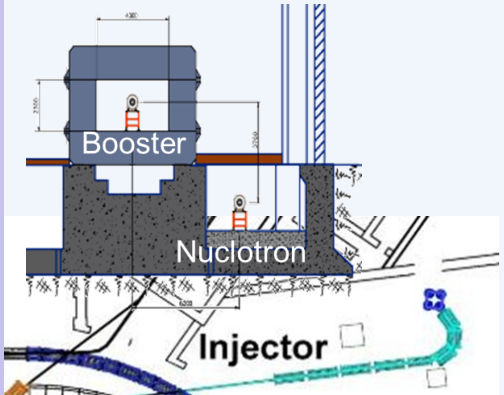
NICA Accelerator Complex

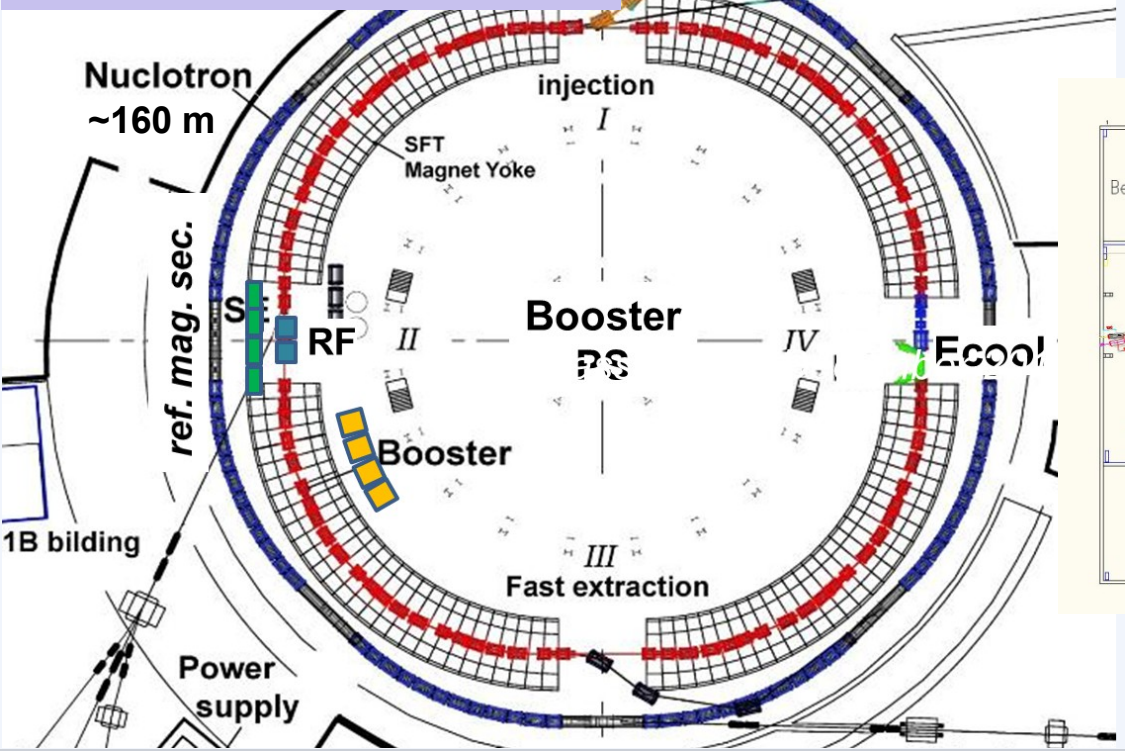
- Beams: from $\boldsymbol{p^{\uparrow}, d^{\uparrow}}$ to $\boldsymbol{Bi}$
- Luminosity: $\boldsymbol{10^{27}}$ (Bi), $\boldsymbol{10^{31}}$ $(p^{\uparrow})$ $cm^{-2}s^{-1}$
- Collision energy: $\sqrt{S_{NN_{BI}}}$ = ***4 - 11*** *GeV* $\sqrt{S_{NN_P}}$ = ***12 - 27*** *GeV*
- Fixed target: BM@N (2018) E_{lab} = *1 - 4.5 AGeV*
- 2 interaction points: MPD (2025end) & SPD (2029)
- Official site: *bmn.jinr.int, nica.jinr.ru*

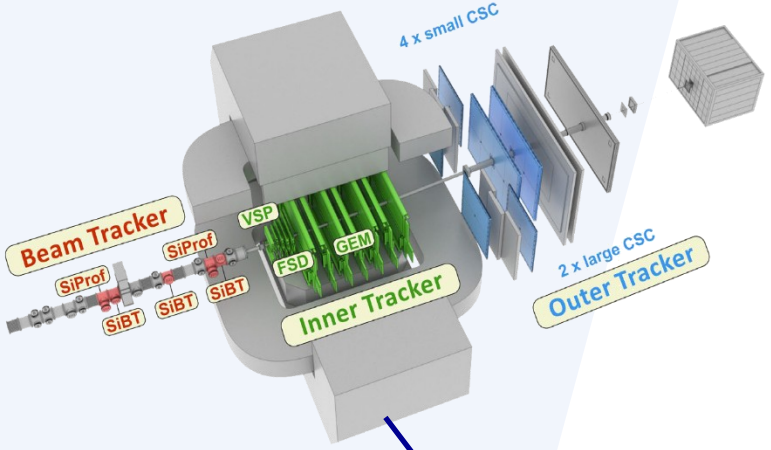
Baryonic Matter @ Nuclotron

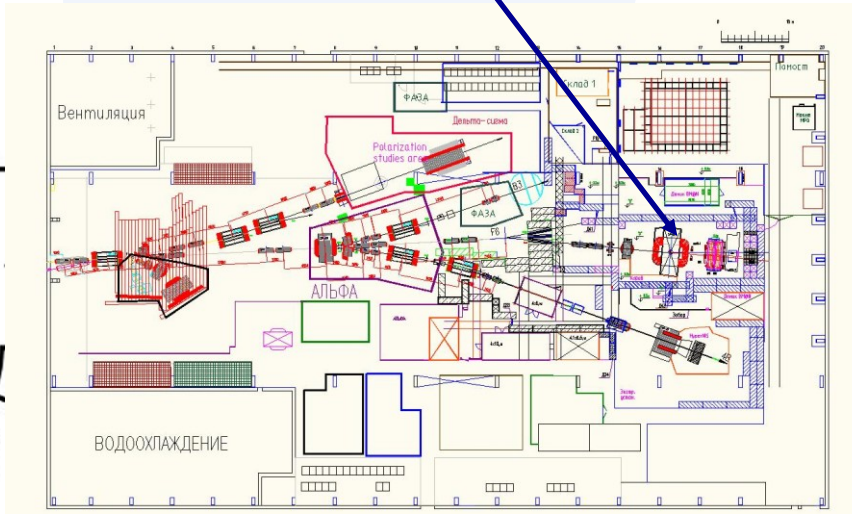
BM@N Physics Program:

- ✓ *strange / multi-strange hyperon and hypernuclei production at the threshold*
- ✓ *in-medium modifications of strange & vector mesons in dense nuclear matter*
- ✓ *hadron femtoscopy*
- ✓ *short range correlations*
- ✓ *event-by event fluctuations*
- ✓ *electromagnetic probes, states decaying into γ, e*



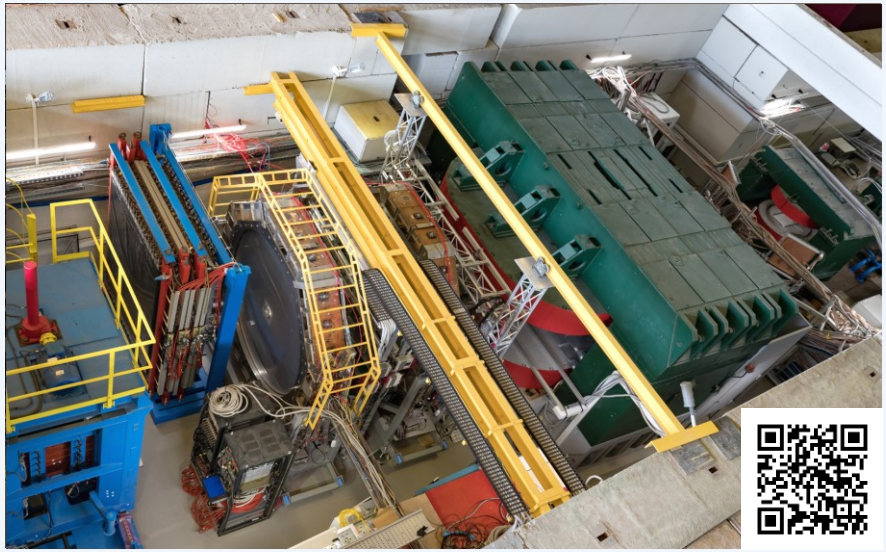






BM@N in Nuclotron Runs (2015 – 2025)

❖ **Nuclotron Run 51** (d,C)		*Feb. 22 – Mar. 15,* ***2015***
❖ **Nuclotron Run 52** (d)		*June 29 – June 30,* ***2016***
❖ **Nuclotron Run 53** (d, d↑)	***Technical*** ***event count: ~300 MEvents***	*Dec. 09 – Dec. 23,* ***2016***
❖ **Nuclotron Run 54** (C)		*Mar. 07 – Mar. 18,* ***2017***
❖ **Nuclotron Run 55** (Ar,Kr)		*Mar. 03 – Apr. 05,* ***2018***
❖ **Nuclotron Run 57** (Xe)	***Physics*** ***event count: ~650 MEvents***	*Dec. 12 – Feb. 02,* ***2023***
❖ **NICA Run 1** (Xe)	***exp. event count: 1000 MEvents***	*August – October,* ***2025***



- Beam: *Xe* (3.8, 3.0 → 2.0 AGeV),
previous runs: *Kr* (2.3, 2.6, 3.0 AGeV), *Ar* (3.2 AGeV), C^{12} (3.5–4.5 AGeV), *d* (4, 4.6 AGeV)
- Target: *CsI* or empty
previous runs: *Pb*, *Sn*, *Cu*, *Al*, C_2H_4, *C*
- Integrated DAQ, T_0 and Trigger systems
- Detectors: VSP, FSD, GEM, CSC, ToF-400&700, DCH 1&2, FHCal, ECal, profilometers…
- Detect min bias beam-target interactions to reconstruct hyperons, strange particles, identify charged particles and nucleus fragments…

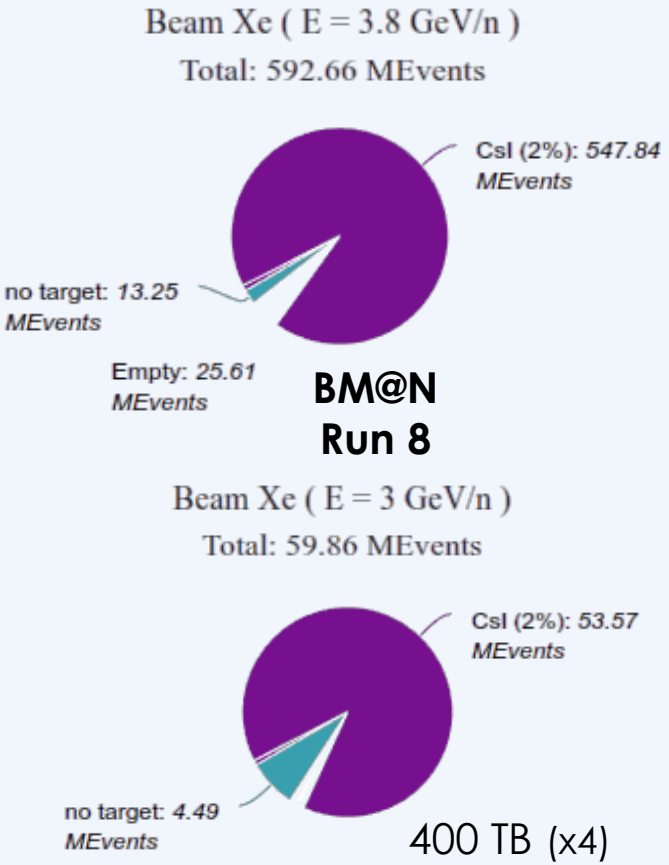
Data Production in BM@N Run 8

1st Physics BM@N Run (2022/2023y)

Two beam energy available for *Xe*-beam

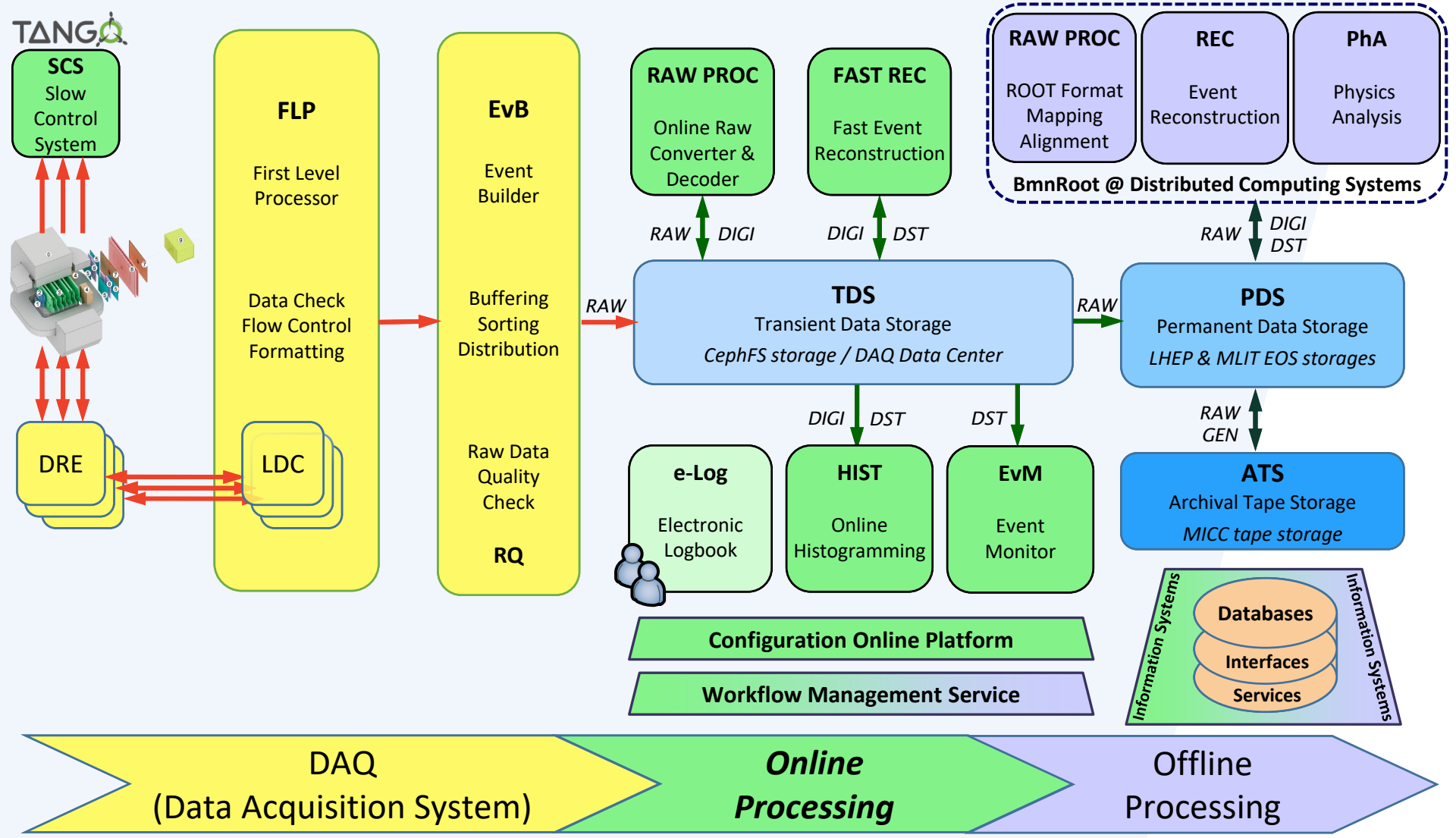
CsI target is used as more similar to *Xe*

More than 600M events were collected



Parameter	Value (approx.)
Data acquisition time	**720 hours**
Average run duration	20 minutes
Average run time break	2.5 minutes
Beam intensity (3.8 AGeV)	up to 900k/2.2 Xe^+/sec up to 900k/12 Xe^+/sec
Trigger rate	**8 000/2.2 event/sec**
Average event size	**0,57 МБ**
Data rate	**up to 2 GB/sec**
Raw file size	**15 GB**
Event count per file (total)	25 000
Total event count (+test, calibration, pedestal)	**645 M**
Total (complete) file count	25 800
Total run count	1 920
Total raw data size	**400 TB**
Total replicated raw data	1.6 PB
Avg digit file size	**1.2 GB**
Avg DST file size	**2.2 GB**

BM@N Data Processing Flow



Electronic Logbook Platform

Online Information System

"I'm trying to liven up my entries. What's another word for 'slept?'"

e-Log Platform

create a new run

advanced search

current day records

user cabinet (event subs)

work with dictionaries

records per page

file attachments

username

page

fast search

BM@N Electronic Logbook

bmn-elog.jinr.ru

Logged in as shift

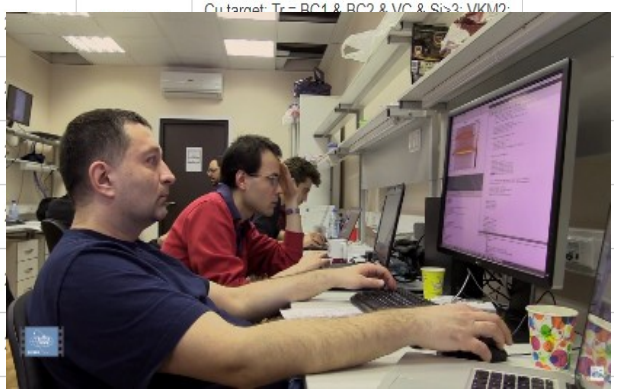
Home New Find Last day Account Reference Book

Page: 1 of 282

Number of items per page: 10

Logout

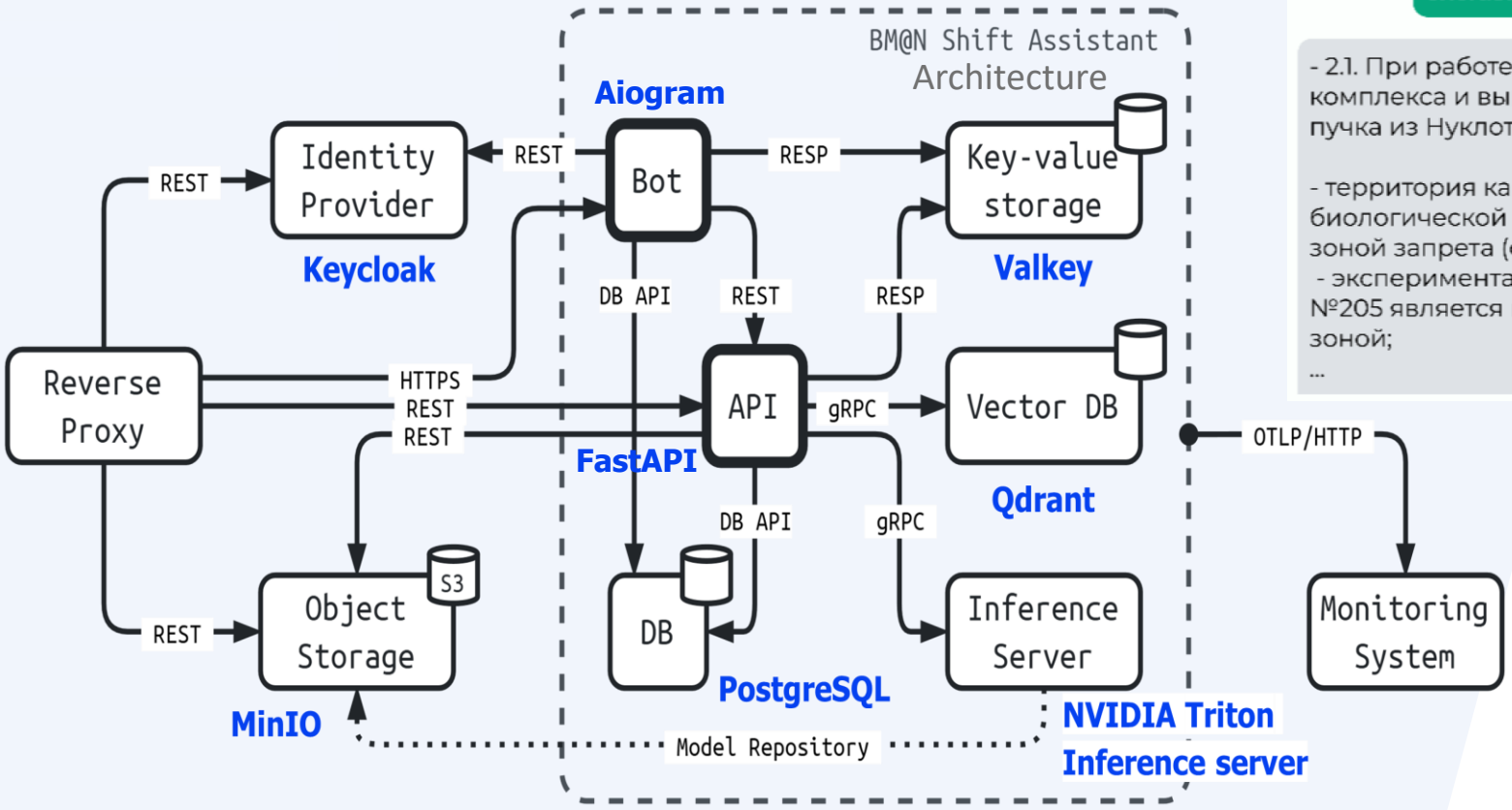
Date	Shift Leader	Type	№ Run	Trigger	DAQ Status	SP-41, A	SP-57, A	VKM2, A	Beam	Energy, GeV	Target	Comment	Attachment
2018-04-05 11:47:06	Rumyantsev	Inform All	5185 per.7	Special Trigger	All	0	0	0	Kr	2.94	Cu (2 mm)	End of the RUN7	
2018-04-05 11:09:20	Rumyantsev	New Run	5184 per.7	Beam Trigger + Si >3	All	1250	50	125	Kr	2.94	Cu (2 mm)	Cu target; Tr.= BC1 & BC2 & VC & Si>3 VKM2: I=125A, SP-57=50A, SP41=1250A; 100 k	
2018-04-05 08:12:35	Rumyantsev	New Run	5183 per.7	Beam Trigger + Si >3	All	1250	50	125	Kr	2.94	Cu (2 mm)	Cu target; Tr.= BC1 & BC2 & VC & Si>2 VKM2: I=125A, SP-57=50A, SP41=1250A; 120 k	
2018-04-05 07:46:35	Babkin	New Run	5182 per.7	Beam Trigger + Si >3	All	1250	50	125	Kr	2.94	Cu (2 mm)	Cu target; Tr.= BC1 & BC2 & VC & Si>3 VKM2: I=125A, SP-57=50A, SP41=1250A; 208 kev	
2018-04-05 07:41:29	Babkin	New Run	5180 per.7	Beam Trigger + Si >3	All	1250	50	125	Kr				
2018-04-05 07:25:08	Babkin	New Run	5179 per.7	Beam Trigger + Si >3	All	1250	50	125	Kr				
2018-04-05 06:01:07	Babkin	New Run	5178 per.7	Beam Trigger + Si >3	All	1250	50	125	Kr				
2018-04-05 05:27:39	Babkin	New Run	5177 per.7	Beam Trigger + Si >3	All	1250	50	125	Kr				
2018-04-05 05:27:06	Babkin	New Run	5176 per.7	Beam Trigger + BD>3	All	1250	50	125	Kr				
2018-04-05 04:47:27	Babkin	New Run	5174 per.7	Beam Trigger + BD>3	All	1250	50	125	Kr				

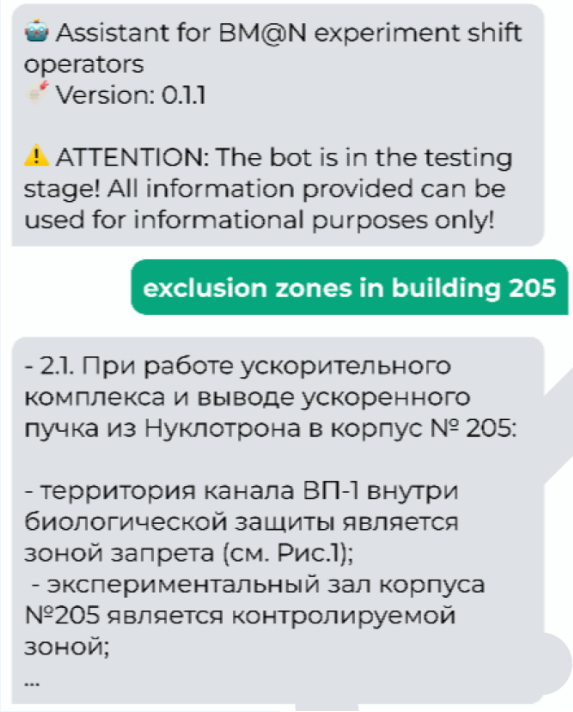


2020 - software team (contact e-mail: gertsen@jinr.ru)

BM@N Bot Assistant for Shift Operators

BGE-M3 is a state-of-the-art text retrieval model:
- large context window (8192 tokens)
- multilingual and cross-lingual
- simultaneous dense and sparse retrieval



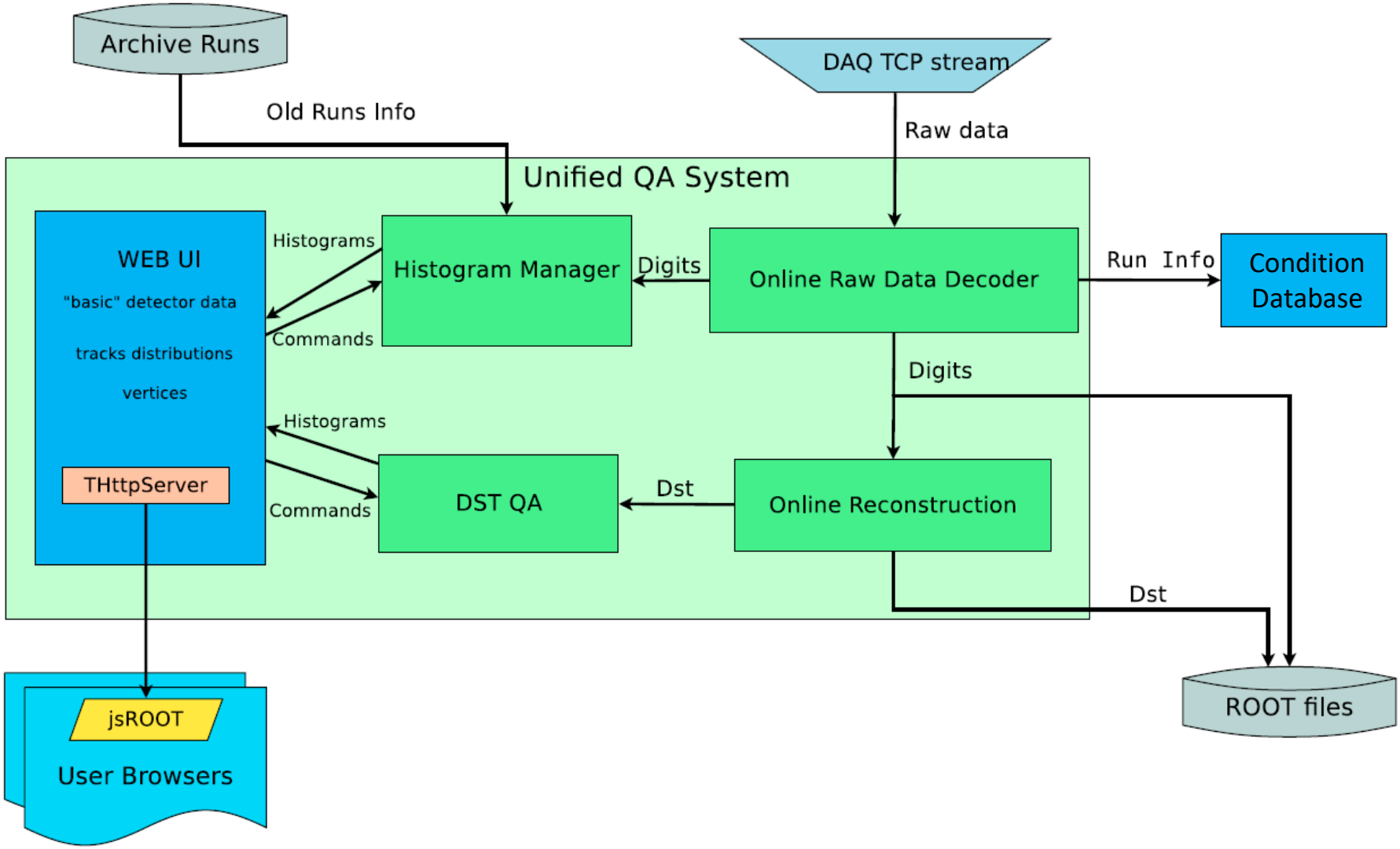


Online Monitoring System

Online Histogramming + Event Display

She says she's from Quality Control. We've failed the furniture inspection.

Raw Data Decoder in the Online Histogramming



Online Histogramming System

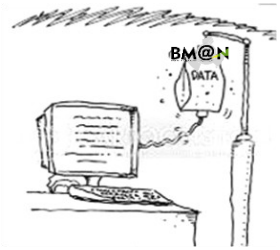
jsROOT (Javascript ROOT) server provides control histograms via the Web

Welcome to the BM@N Experiment | Triggers | GEM | Silicon | ToF400 | ToF700 | DCH | MWPC | ZDC | ECAL | SRC Triggers | LAND | MSC

Run: 4147

Event: 20000

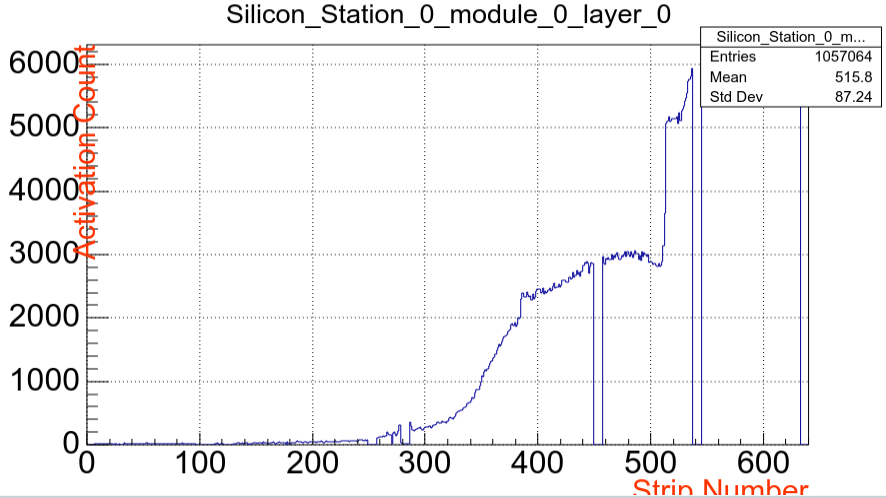
Run Type: beam

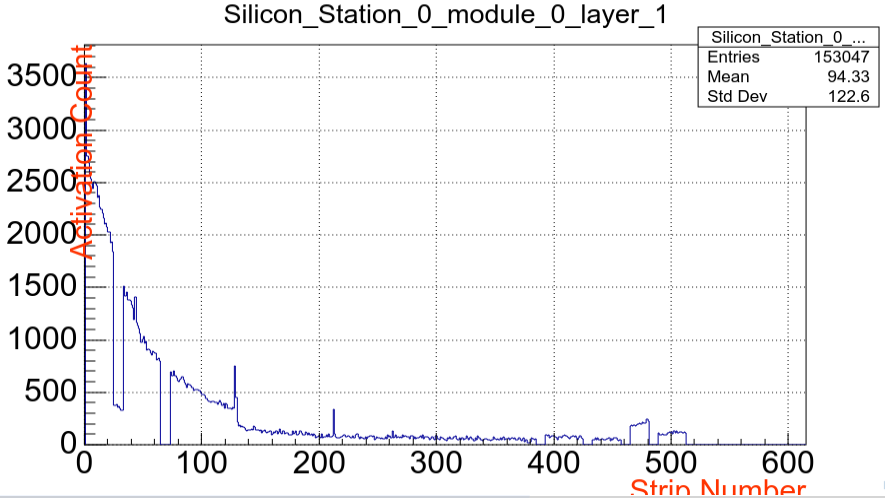


BM@N Silicons

Energy: 3.20
Beam: Ar
Target: C
Field Voltage: 77.60

Reset | Select Reference Run | Run 3946, beam Ar, energy 3.2, target C, Voltage 77.597222

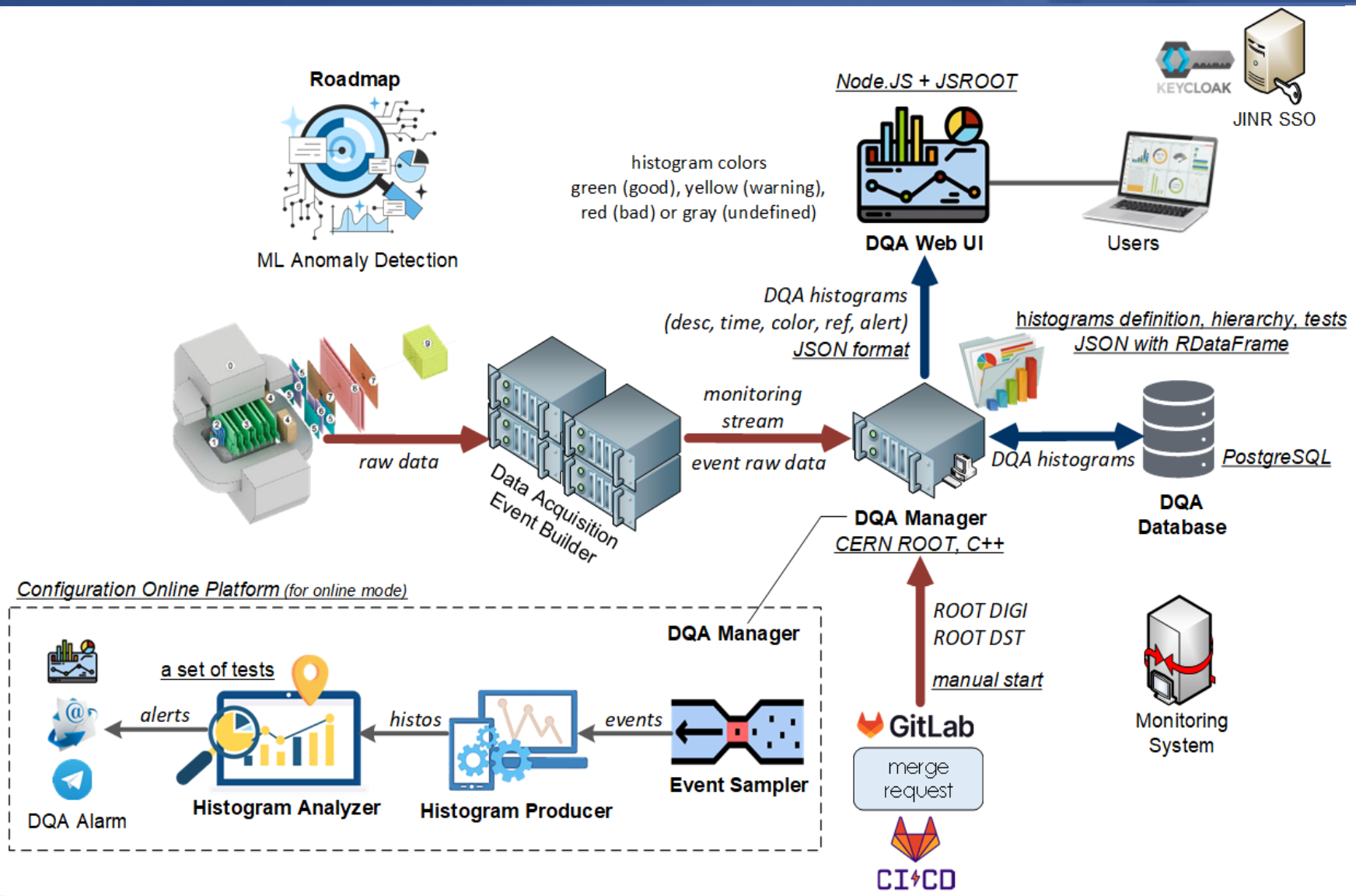




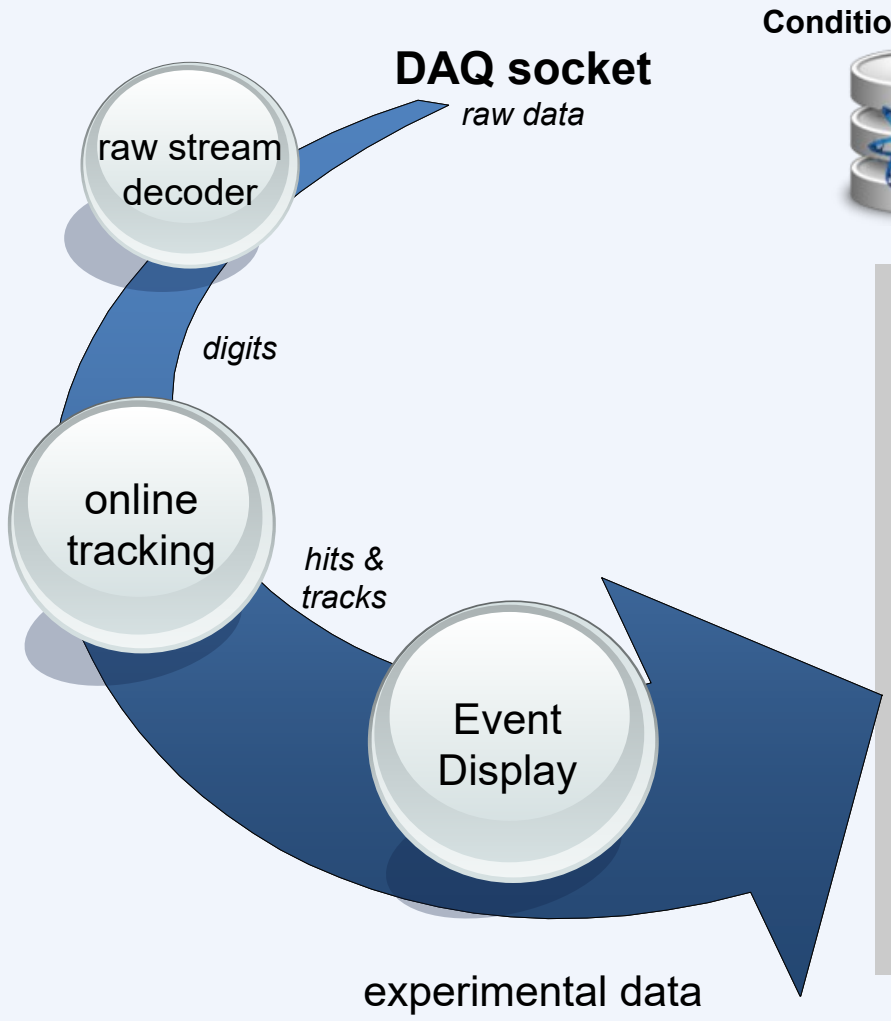
Data Quality Assurance (DQA) system for BM@N

- Single DQA system for the BM@N experiment is being developed to provide control histograms in the same way in 3 modes:
  - for online decoded and reconstructed data (online data monitoring)
  - for qualitative assessment of new BmnRoot versions (after MRs)
  - for manual run to check user versions of the software
- The BM@N DQA architecture provides predefined checks and graphical representation of control histograms on a central Web service, which receives the displayed data distributions from the DQA manager (sampler → histogram producer → histogram analyzer)
- The DQA system ensures the ability to easily add new control histograms with predefined checks and alerts in a user-friendly format (suitable for physicists and detector team, who are not developers) using *JSON*-description

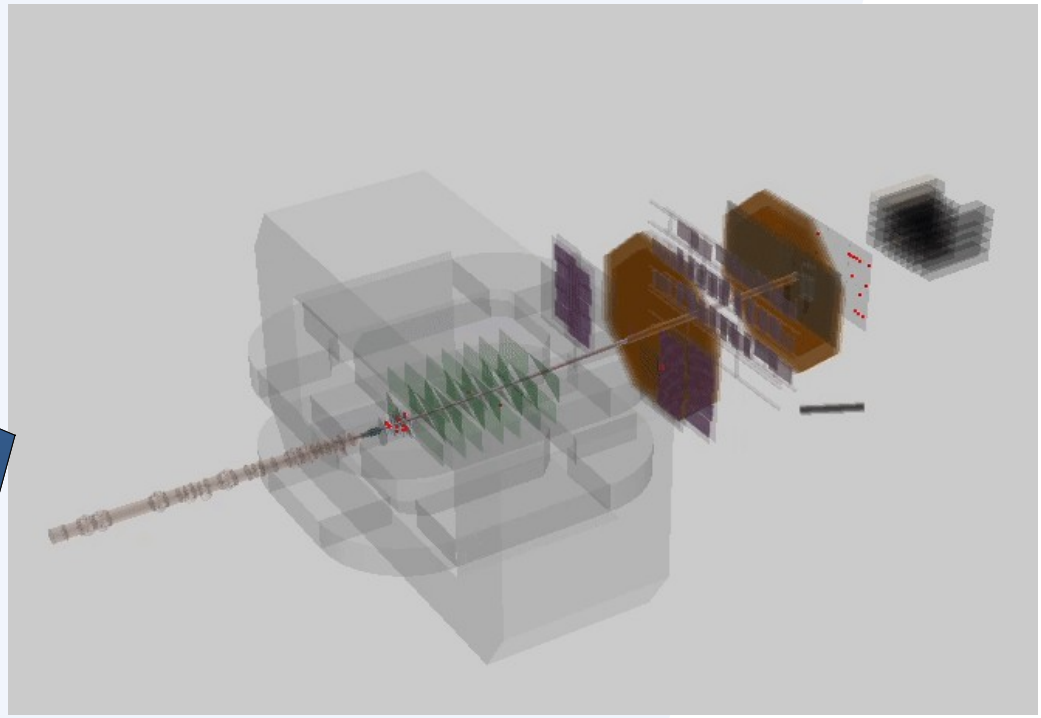
Design of the BM@N DQA system



BM@N Event Display for Online Monitoring

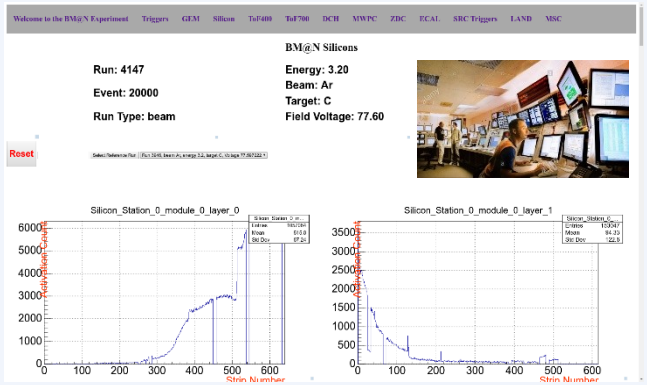


*based on the **ROOT EVE package***



Online Event Display
BM@N hits and tracks

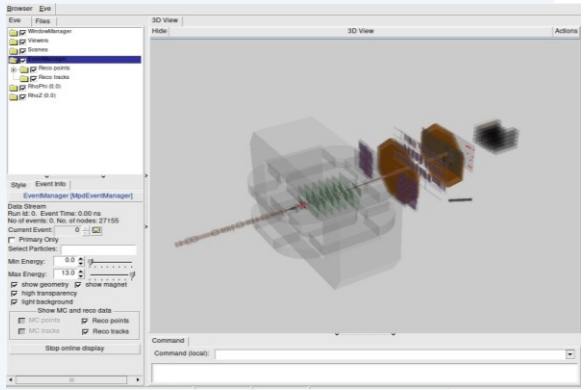
Online Metadata for BM@N Data Processing



DAQ system

EvB

Raw Event Builder



Online Raw

ROOT Format Digit Decoder

HIST

Online Histogramming

EvM

Event Monitor

raw file metadata

period, run
start & end time, event count, file path, file size

e-Log platform

run metadata

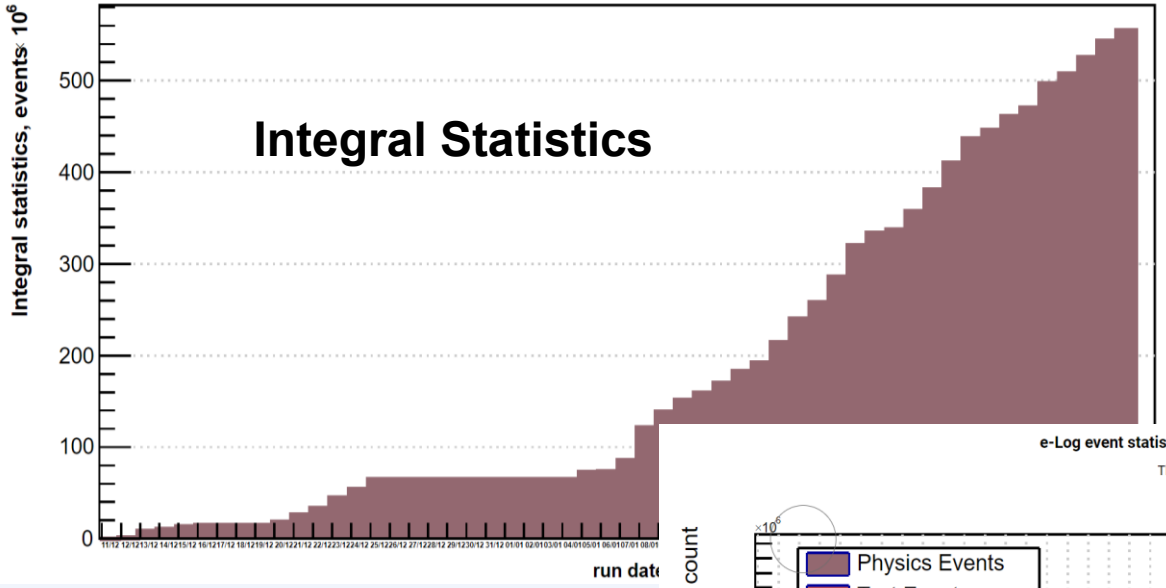
period, run
beam, energy, target, magnetic field

Condition Database

online & offline data processing

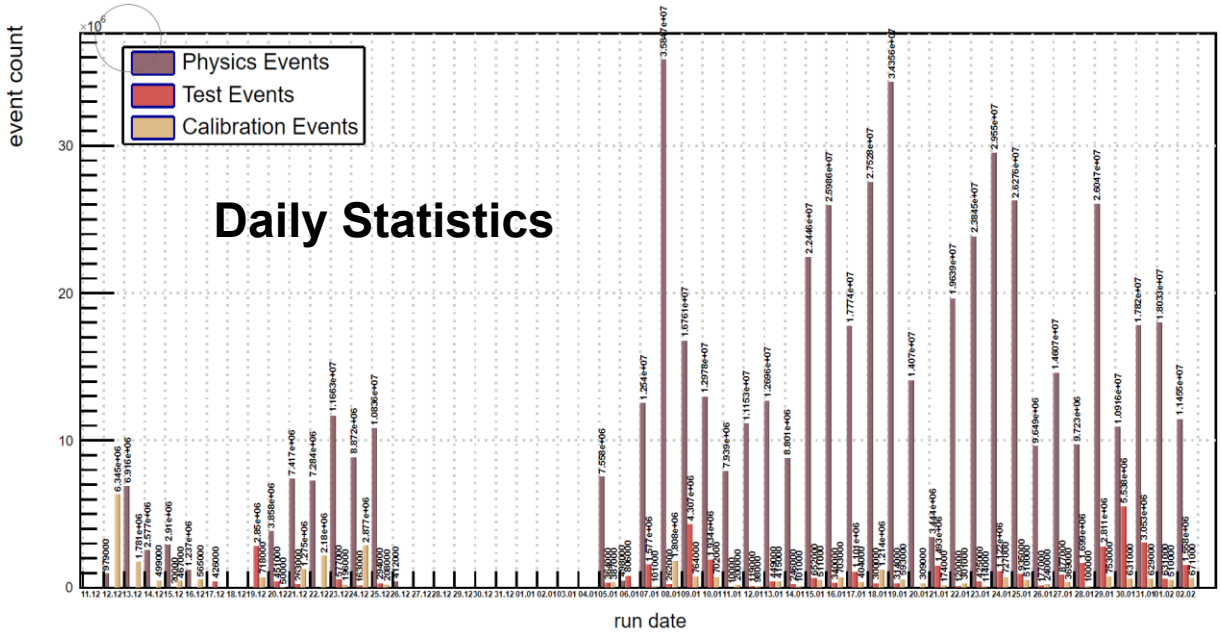
Event & Run Statistics

Online Event Statistics during BM@N Runs

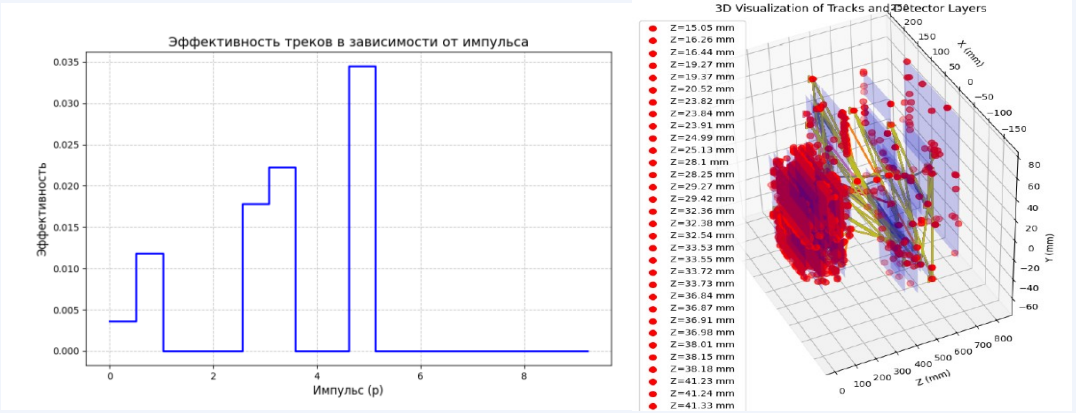


e-Log event statistics (Run 8): statistics by event type / integral statistics of physics events

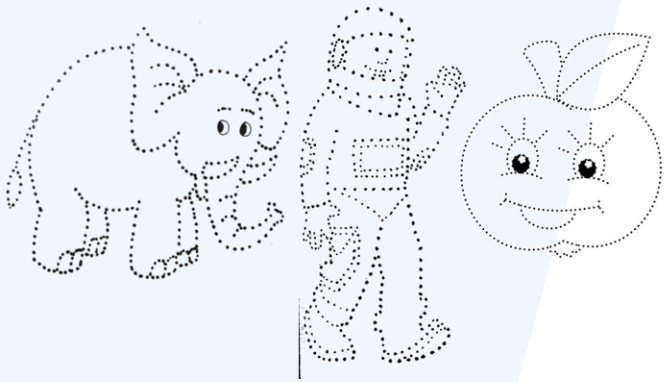
The information is current as of February 20 2023 23:59.

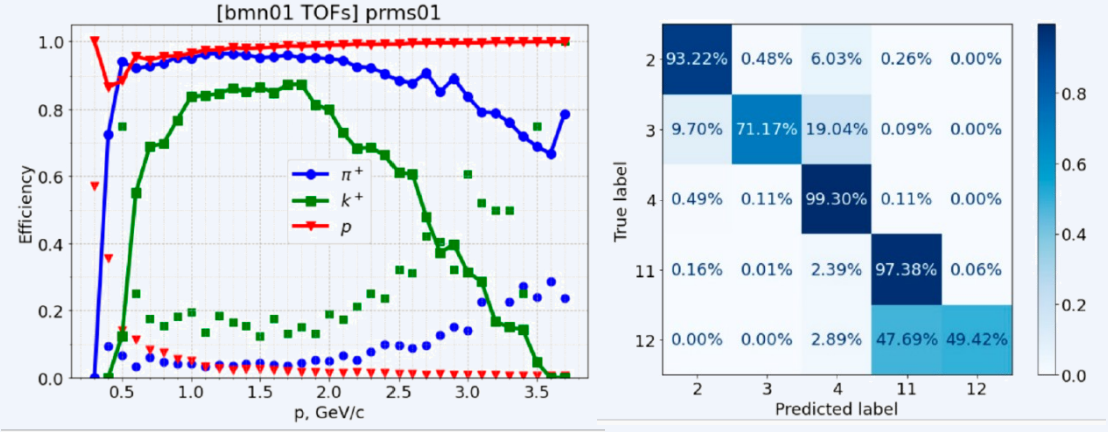


Machine Learning for Fast Tracking and PID

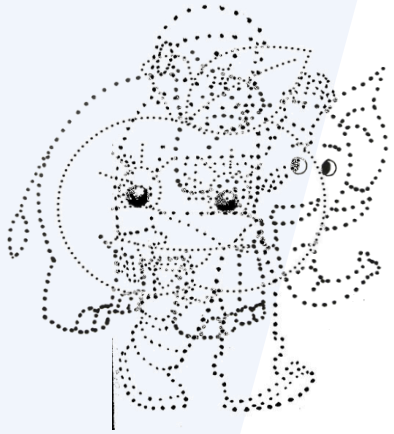


Anatoly ALEXANDROV, Sergey A. NEMNYUGIN
Application of machine learning for particle tracking in BM@N





Alexander AYRIYAN, Vladimir PAPOYAN
Implementation of Particle IDentification based on Machine Learning

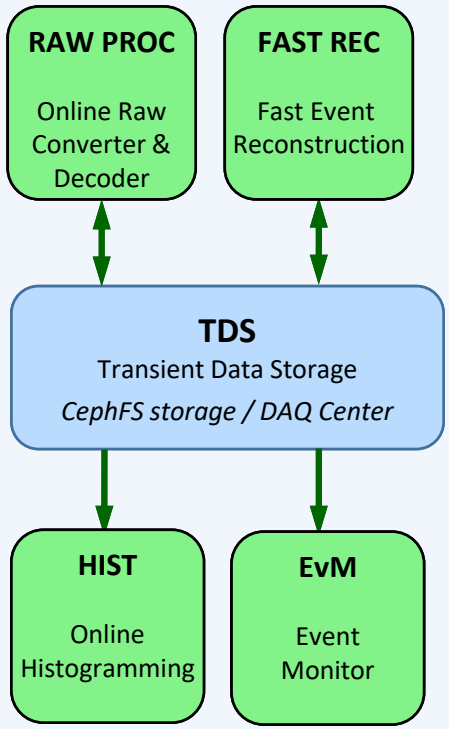


Nikolay ERSHOV
Own implementation of Fast Event Reconstruction based on ML/ NN

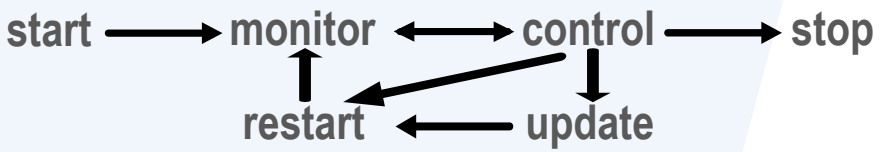
Configuration Online Platform

Online Data Processing

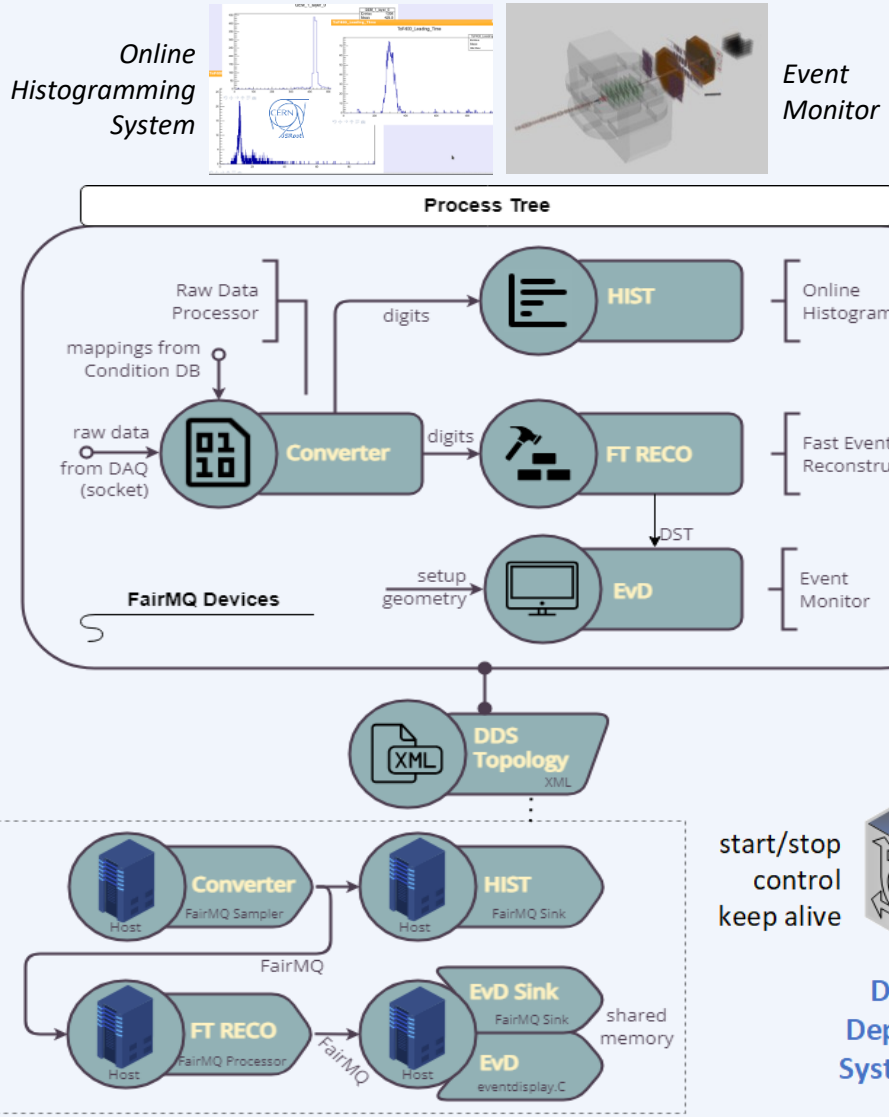
Online Process Control



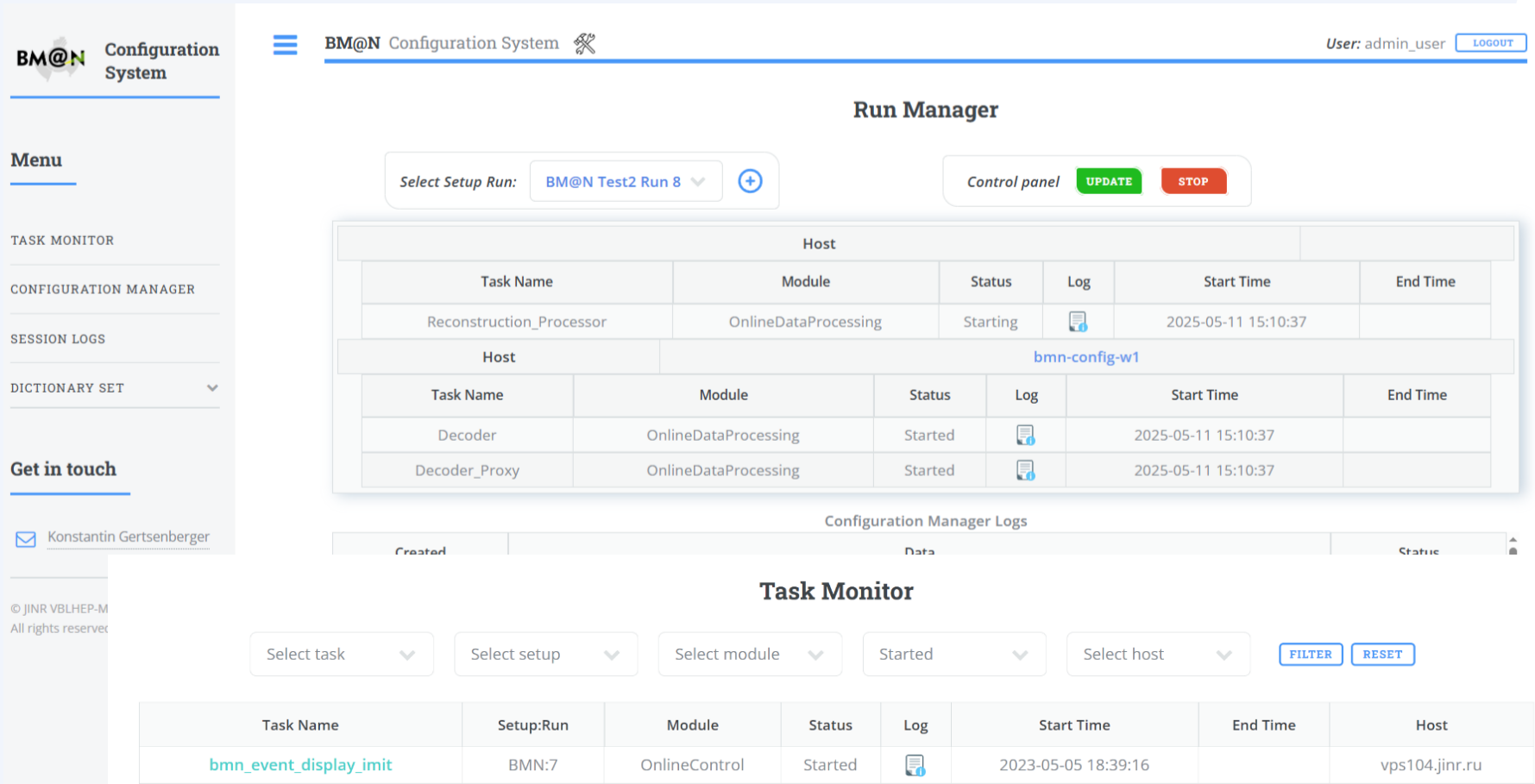
Online Processing



BM@N Configuration Online Platform



Configuration Designer & Monitor

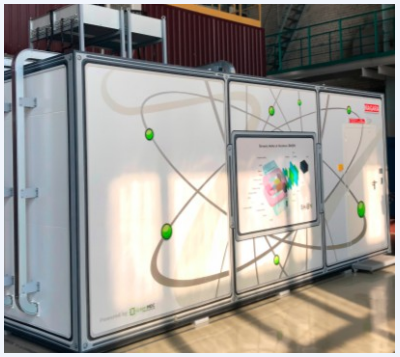


Distributed Data Processing during BM@N Runs

"Let's shrink Big Data into Small Data ...
and hope it magically becomes Great Data."

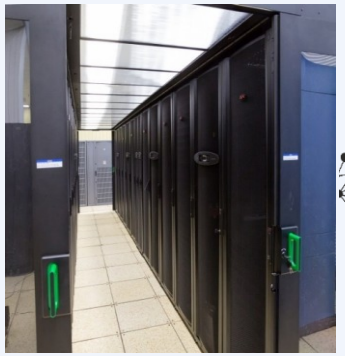
BM@N Computing Platforms

BM@N Online Cluster
(DAQ Data Center)
ddc.jinr.ru
(LHEP, b.205)



NICA Cluster
ncx[101-106].jinr.ru
(LHEP, b.216)

GRID Tier1&2 Centres
(CICC complex)
lxui.jinr.ru
(MLIT, b.134)



SC «Govorun»
(HybriLIT platform)
hydra.jinr.ru
(MLIT, b.134)

OS: AlmaLinux 9.5

OS: CentOS / Scientific Linux 7.9 (EOL on June 30, 2024)

Central Software Repository based on ***CVMFS*** for the experiment

CEPH: 2.8 PB (*replica*)
CEPH (hot): 100 TB$_{ssd}$
SLURM: ≈**700** cores

EOS: 1.3 PB (*replica*)
NFS: **300 TB** (*for NICA*)
SLURM: ≈**1800** cores
(*for all NICA users*)

EOS: 1.9 PB (*replica*)
EOS CTA: **500 TB$_{tapes}$**
SLURM: 2500 cores
(*for all NICA users*)

ZFS: 200 TB
Lustre: **300 TB$_{ssd}$** (*for NICA*)
SLURM: *bmn* – **192** cores

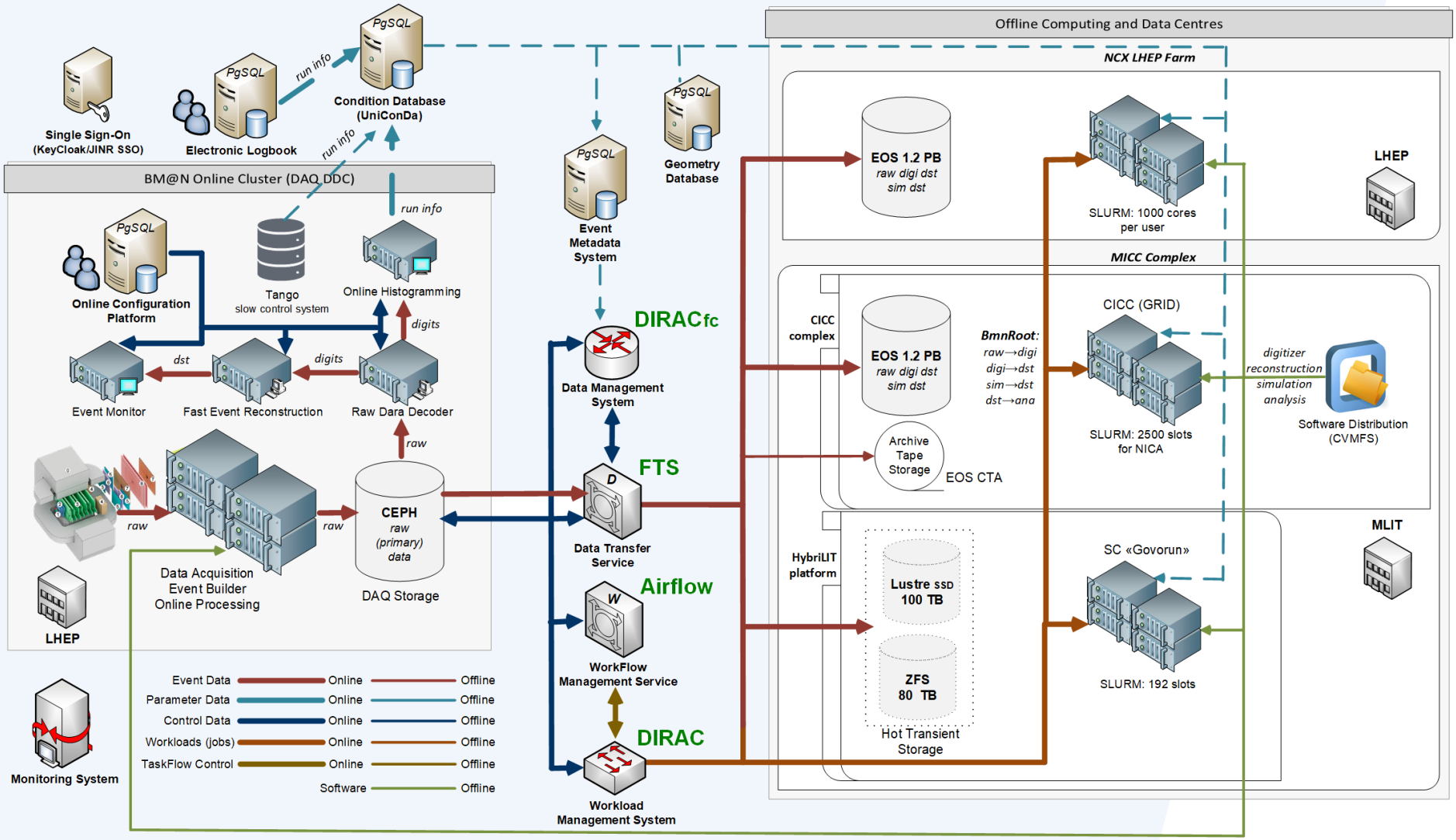
BM@N software is deployed on *JINR CVMFS* for Centos 7 and AlmaLinux 9

Automatic software deployment of the BmnRoot package on CVMFS with *GitLab CI*

BmnRoot Containers for Distributed Computing

- Dockers for testing BmnRoot before MR in GitLab CD/CI
  - simplify *CI*-infrastructure using BmnRoot Pipelines
  - quickly add any OS environments to *CI* pipelines (CentOS 7 / AlmaLinux 9 / Ubuntu 22.04)
- User Docker Containers with BmnRoot software
  - base image = OS + FairSoft + FairRoot
  - users do not need to install software – just run the BmnRoot container
  - hosting computer can potentially run any operating system
  - great for short-period students and fast analysis
  - BmnRoot & its containers are automatically built and published with *GitLab CI*
- Apptainers for using BmnRoot Containers
  - 2 containers (on AlmaLinux 9 & Ubuntu 22.04) with full local installation
  - 2 containers with CernVM-FS client to the central JINR CVMFS repository
  - new “container” execution mode in NICA-Scheduler to launch distributed jobs inside an Apptainer container (with CVMFS client)

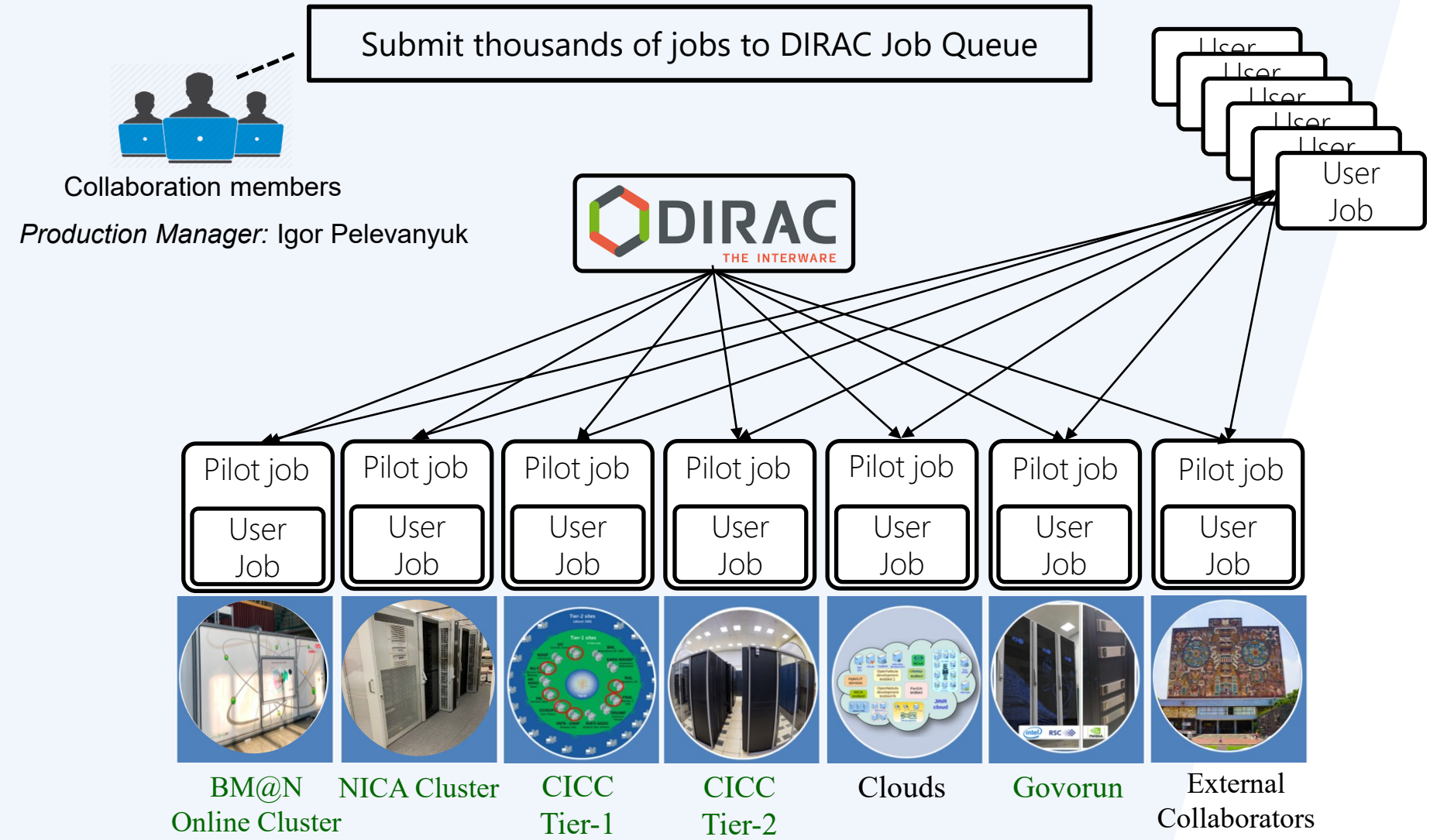
BM@N Computing Software Architecture



BM@N Workload Manager

DIRAC Interware

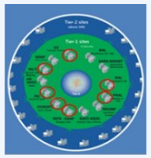
DIRAC Workload Manager for BM@N



DIRAC jobs on BM@N Computing Resources

BM@N
Online Cluster

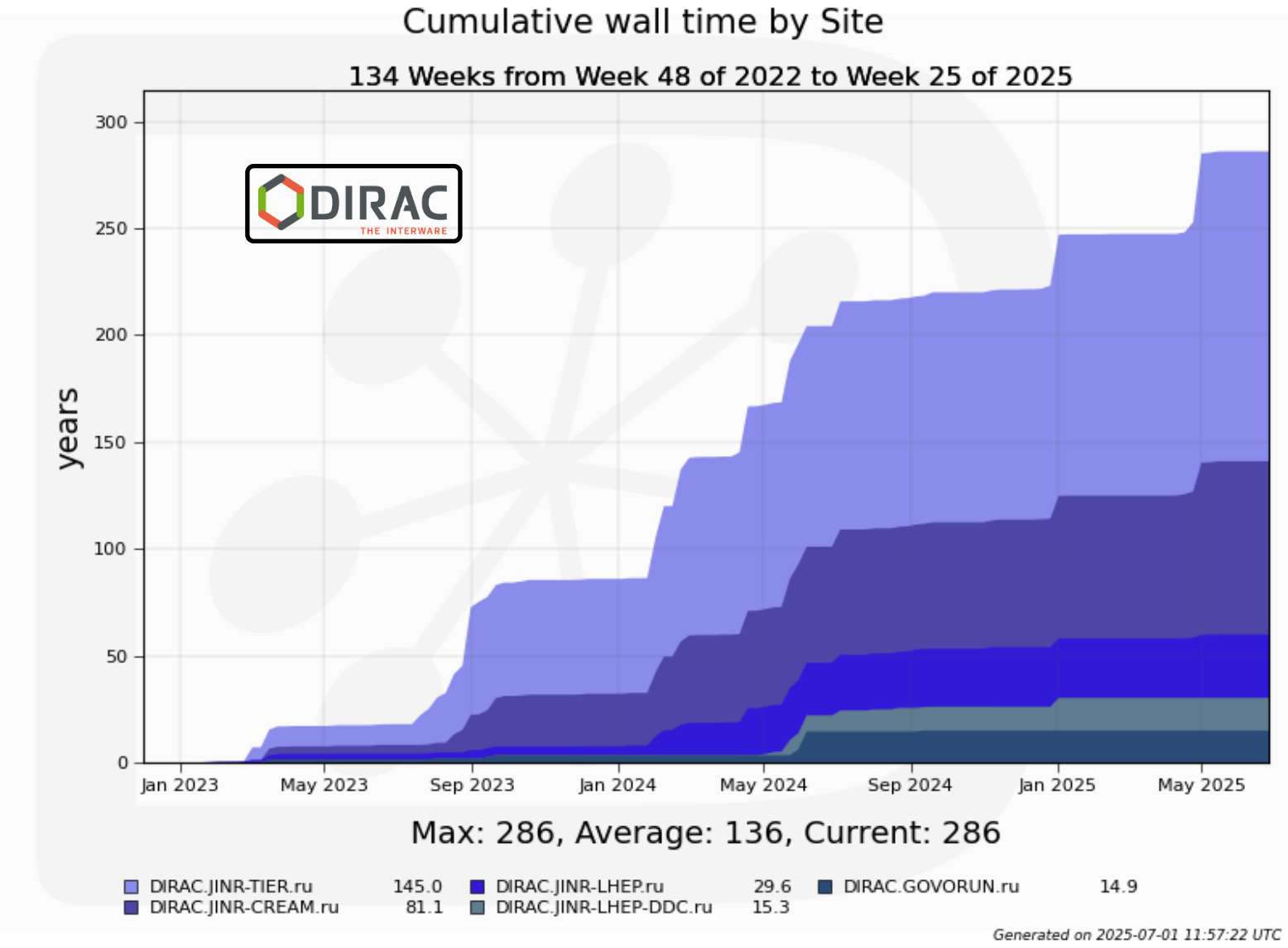
NICA Cluster



CICC
Tier-1

CICC
Tier-2

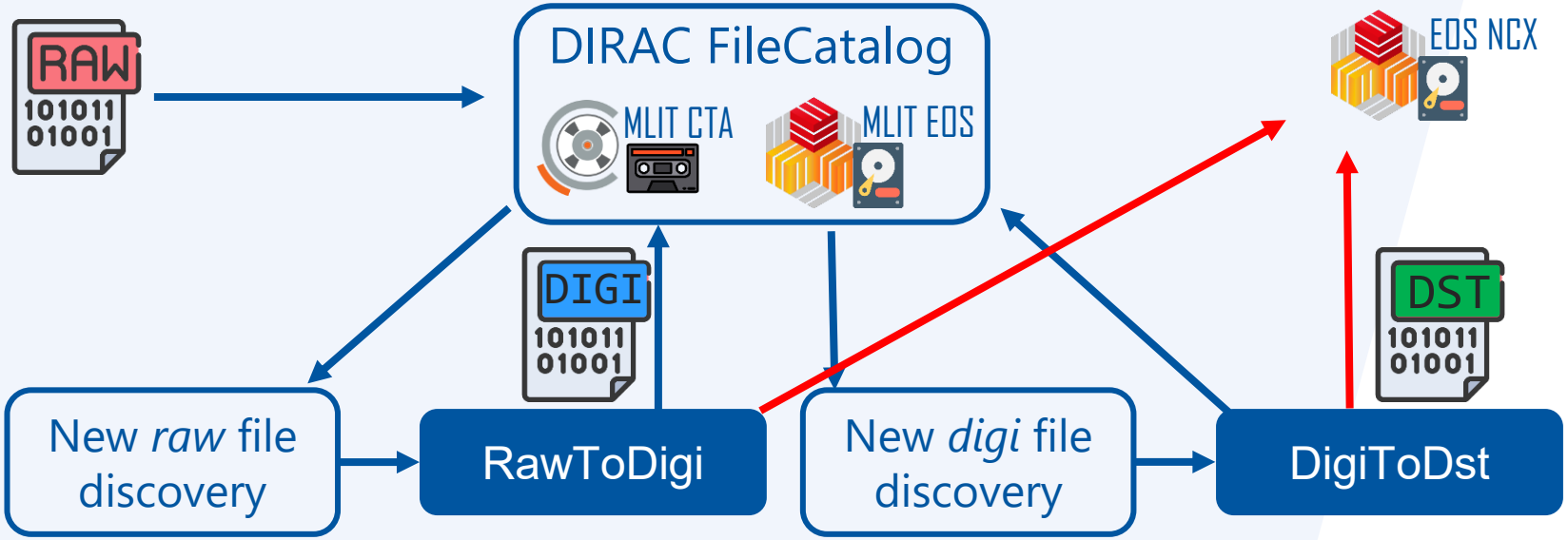
Govorun



Automation Scheme with DIRAC

To automate job processing the following tasks should be completed:

1. **Upload** new *raw* file to DIRAC file catalog.
2. **Detect** new *raw* file in DIRAC file catalog and **submit** RawToDigi job for each new *raw* file.
3. **Detect** new *digi* file in DIRAC file catalog and **submit** DigiToDst job for each new *digi* file.
4. **Upload** *digi* and *dst* files to NCX cluster.



BM@N Data Management

DIRAC File Catalog

Data Storages for BM@N

ceph

for BM@N online (2.8 PB)

build on HDD with SSD buffer

NICA cluster

EOS

for BM@N offline (1.3 PB)

build on HDD

sync | data

MLIT MICC

EOS

for BM@N offline (1.9 PB)

build on HDD

NICA cluster

NFS

fast on NVMe SSD

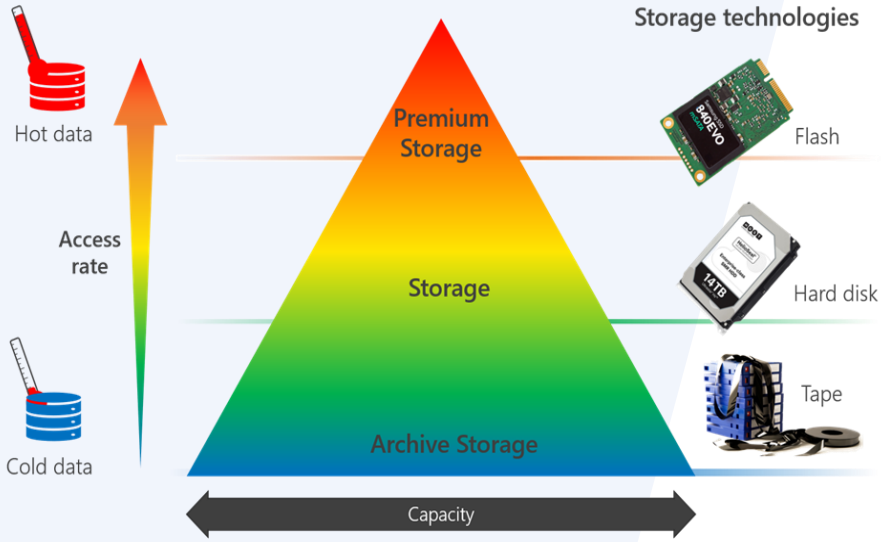
MLIT CICC

nope

HybriLIT

lustre

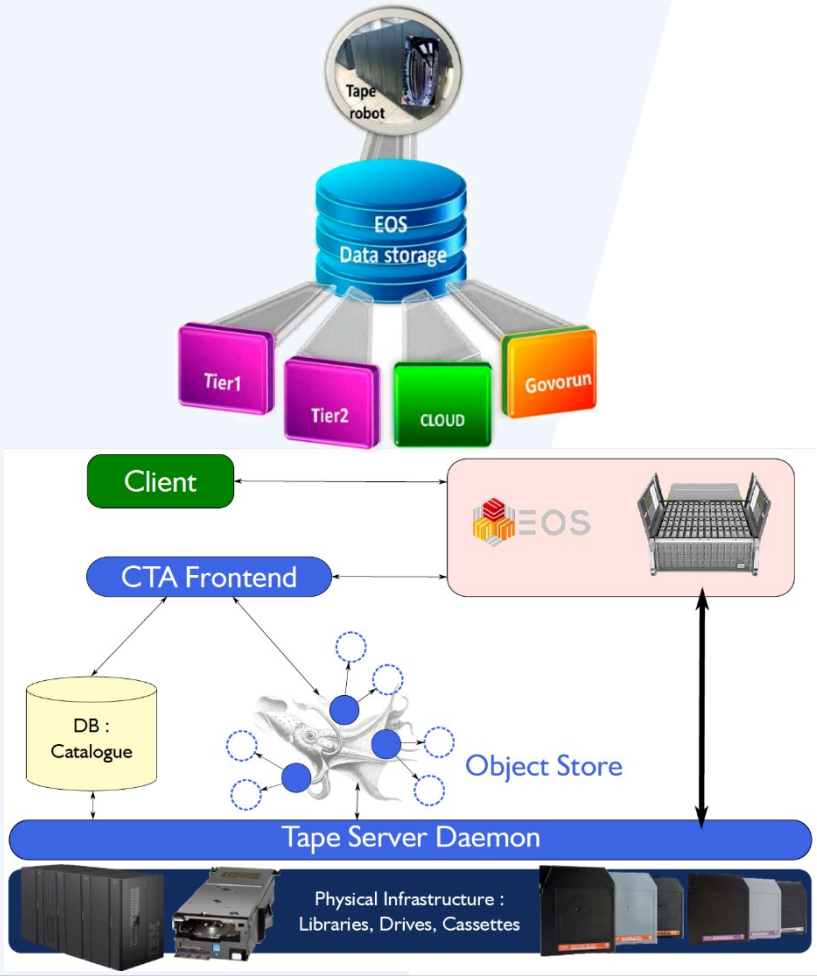
fast on NVMe SSD



Archival Tape Storage for BM@N

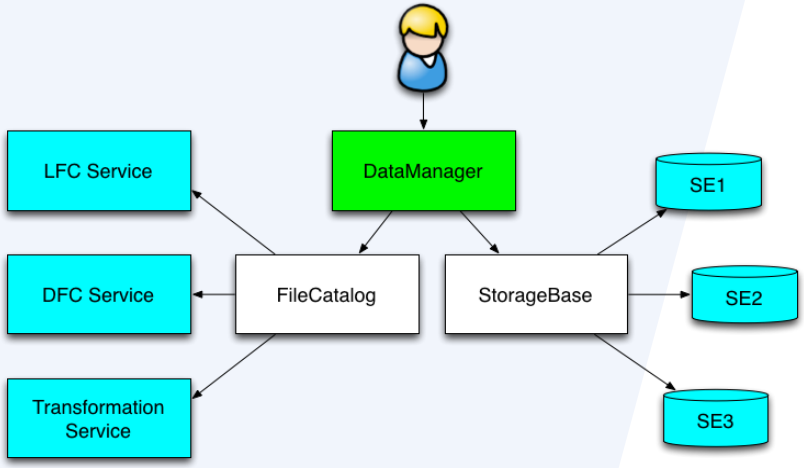
EOS CTA Integration in MLIT

- CTA tape is a new archive solution developed at CERN to replace Castor
- Extends MLIT EOS with tape backend functionality
- Tape “bringonline” exposed via EOS, and XRootD protocols, Gfal2 support
- Can be handled transparently by FTS
- Advantages: long lifespan, cost of use, energy efficiency, security
- Tape robotic systems – a long-term storage for BM@N, stores ***raw*** and ***gen*** data, ***online raw data backup*** to tapes
- Stores Run8 raw data already



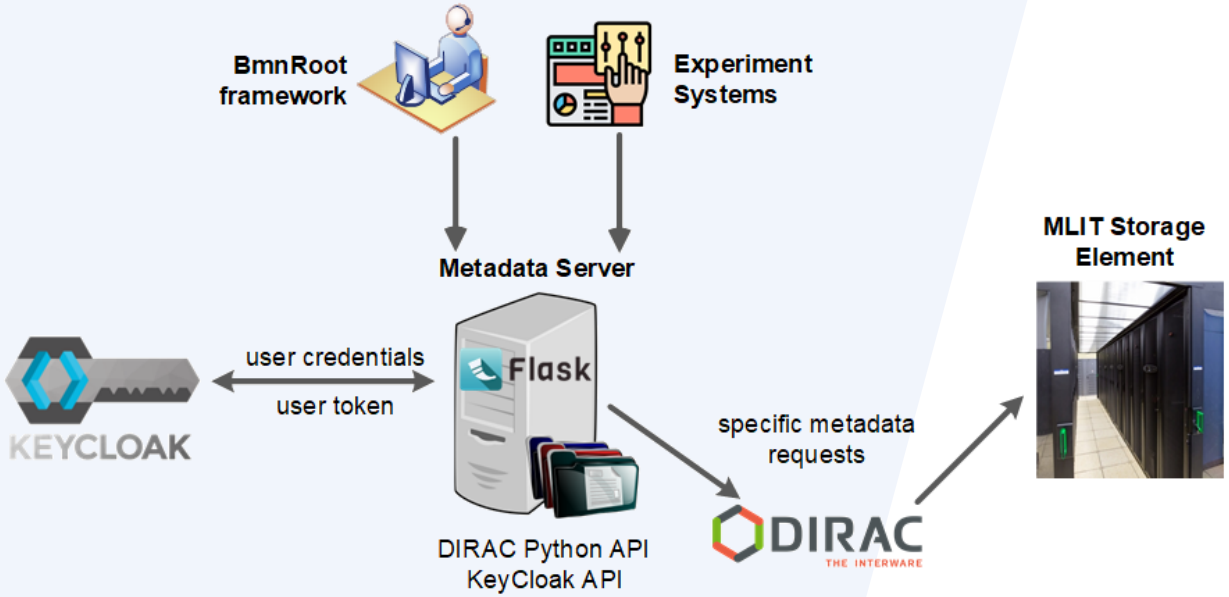
Data Management System for BM@N

- DIRAC File Catalog (DFC) is maintaining a single global logical name space
- A user sees it as a single catalogue with additional features
- DataManager is a single client interface for logical data operations
- DFC also hosts Metadata



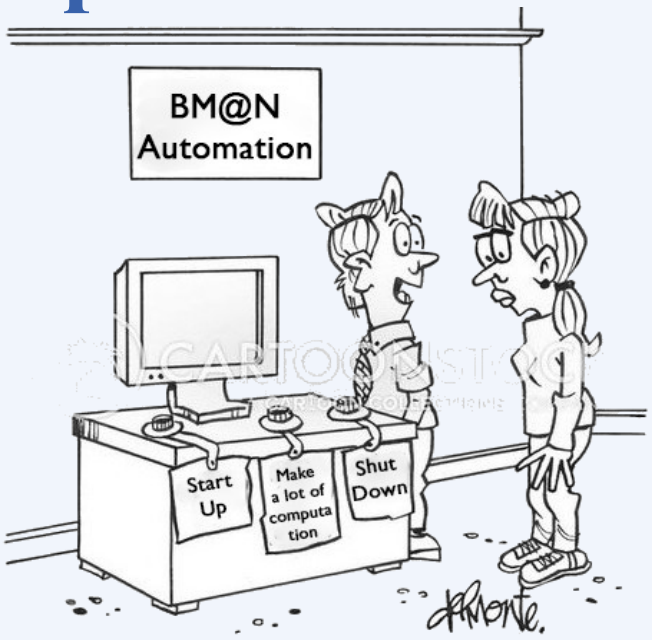
BM@N DFC Metadata:

- **period and run number**
- **start and end datetime**
- **beam and target particles**
- **run type**
- **energy**
- **magnetic field**
- **start and end event number**
- **event count**
- **file size**



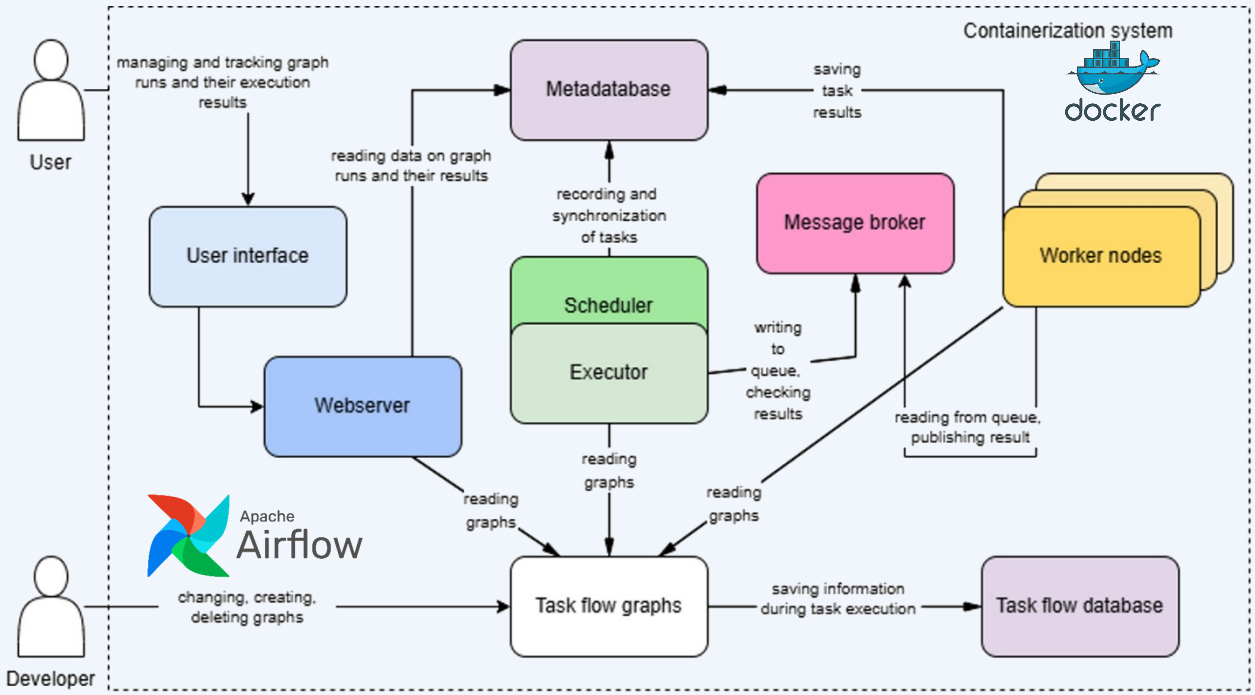
Workflow Management Service

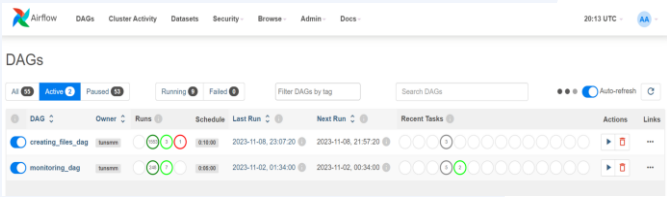
Apache Airflow



"I've simplified our work process!

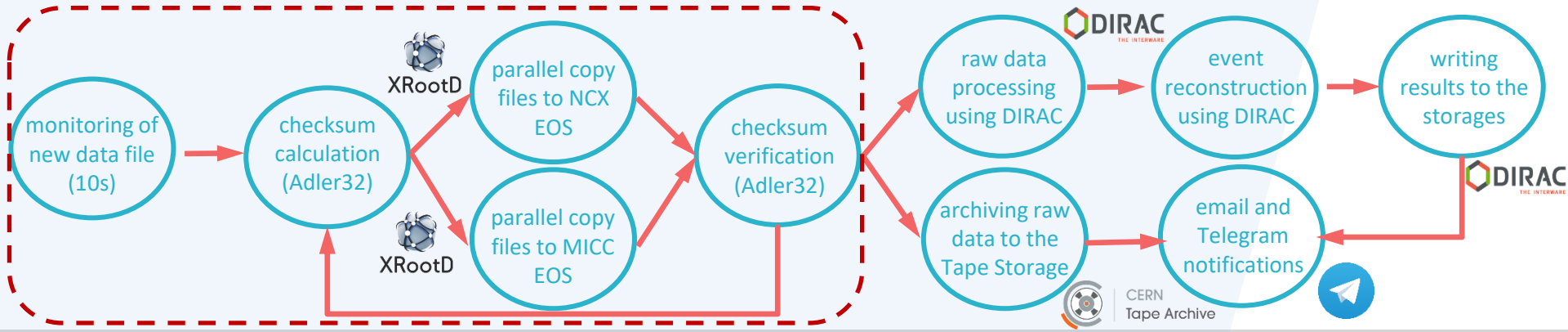
BM@N Orchestration with Workflow Manager



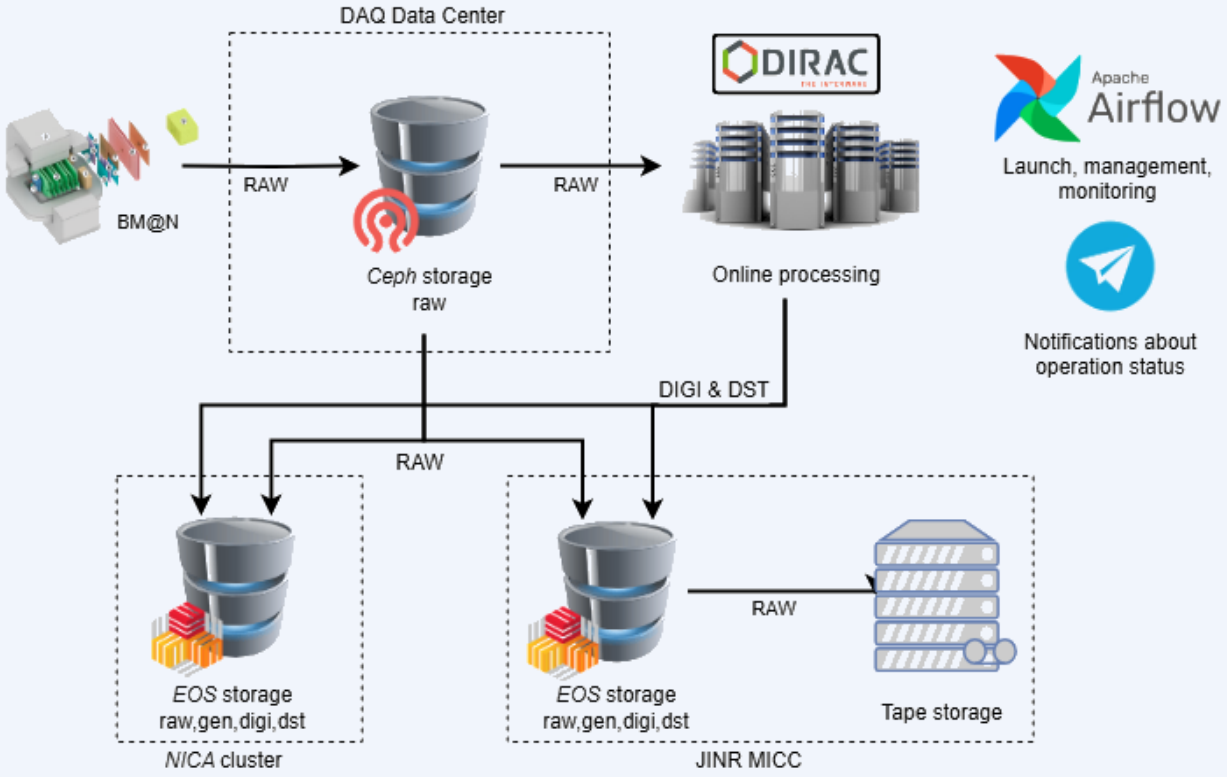


Airflow DAGs can be used for repeating data processing of simulation and experimental events

Airflow DAGs has been developed for online data processing, transferring to the EOS storages, archiving to the Tape Storage



Online Processing Pipeline for BM@N Run 9



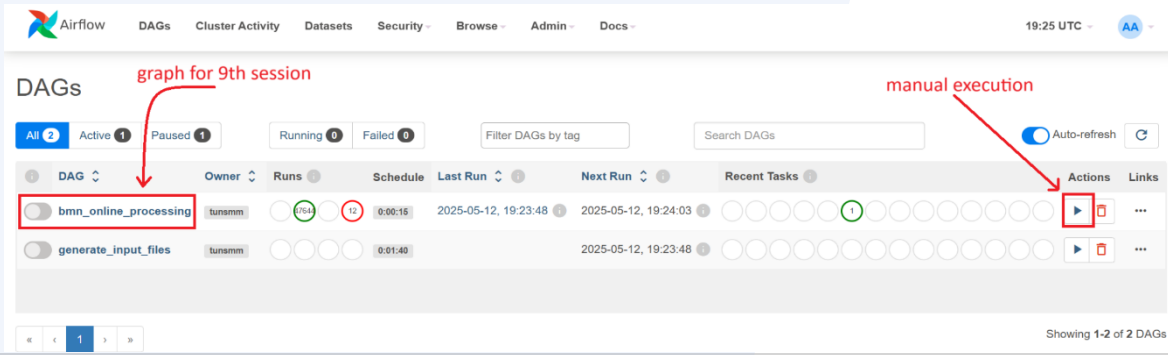
copying of the raw data to the storage systems of the NICA and MICC clusters

"online" processing of collected data on the provided distributed computing resources

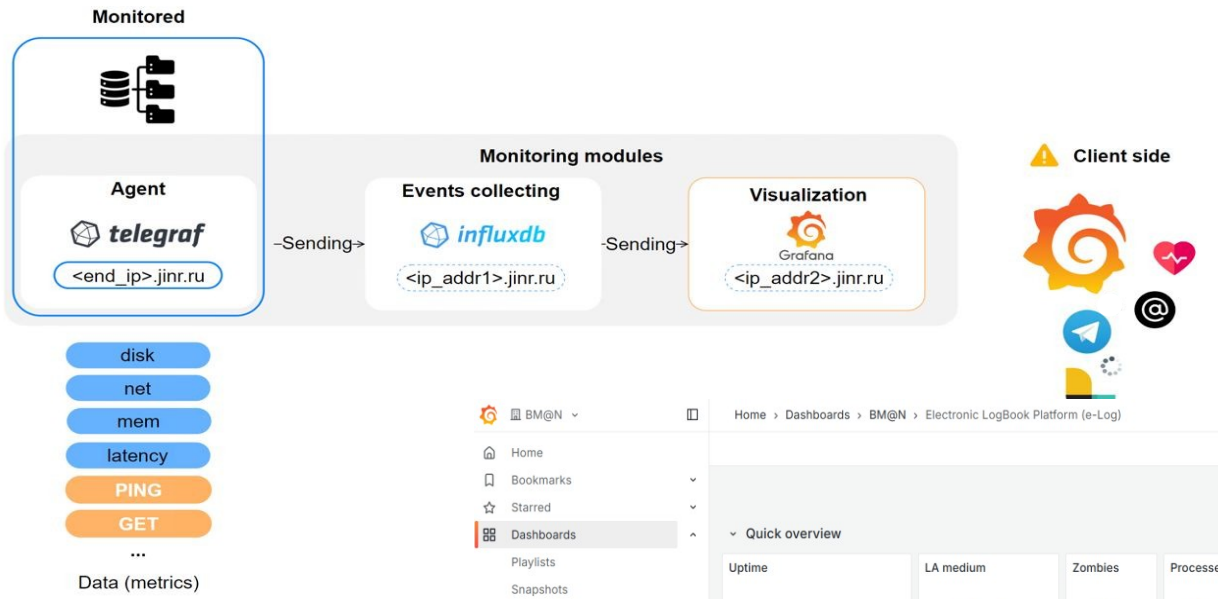
backup to the NICA Tape Storage

Apache Airflow

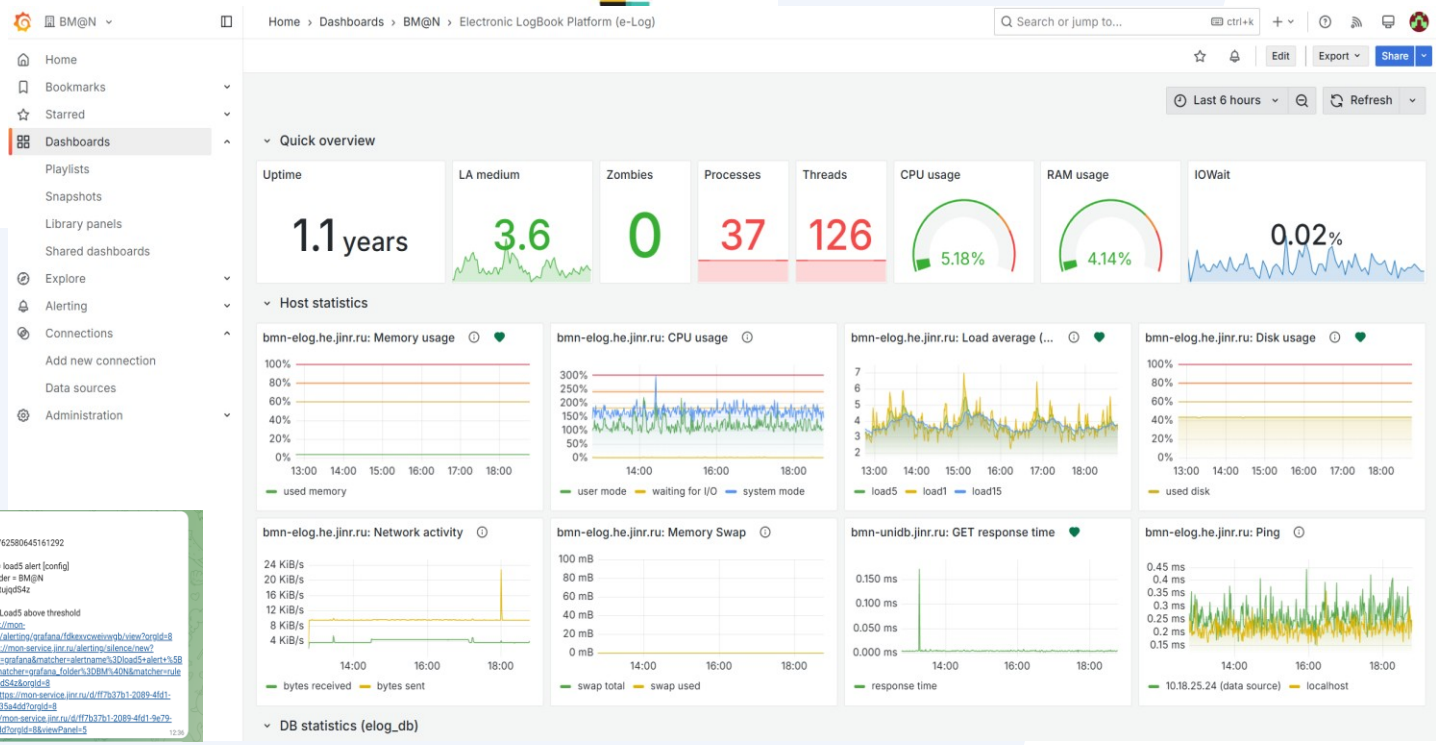
Online and Offline TaskFlows for mass data processing



Monitoring System for the Software Complex

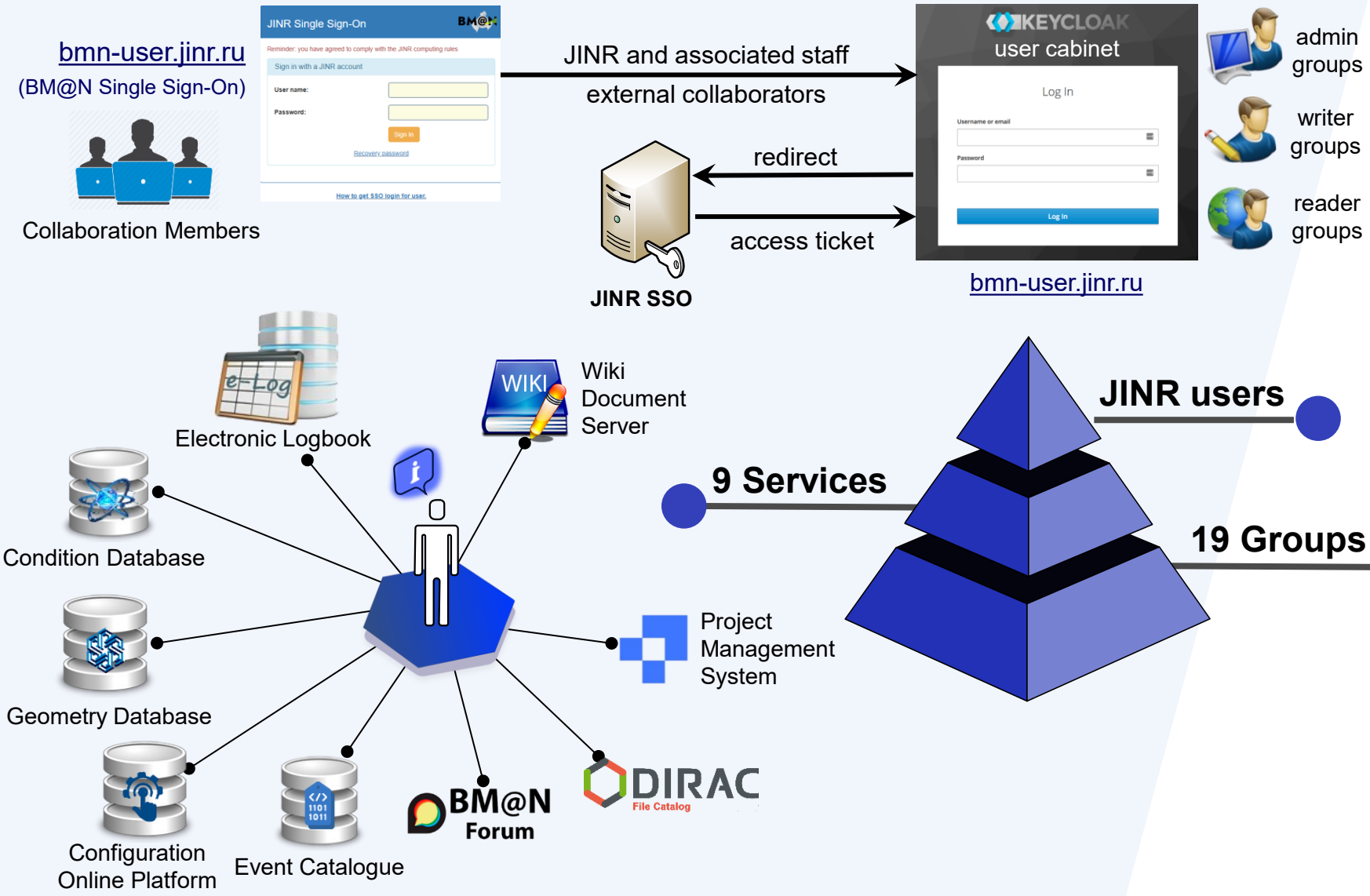


- Ping servers, HTTP request and SQL latency to check current status
- Monitor server parameters, such as *Disk*, *CPU*, *Memory*, etc. (Telegraf)
- Response time stored in InfluxDB
- Use JINR Grafana service for visualization and failure alerting



- Module architecture
- Configurable via **JSON** file
- Deploying with **Ansible playbooks**
- Email and **Telegram** notifications

BM@N SSO System



JINR MLIT Contribution to BM@N

***Director*: S. V. SHMATOV. *Scientific Leader*: V. V. KORENKOV**

JINR MLIT Contribution to BM@N

BM@N is open for cooperation and new people!

Igor ALEXANDROV, Evgeniy ALEXANDROV, Irina FILOZOVA, et alia
Geometry Database, Configuration Online Platform, DQA System

Nikita BALASHOV
CVMFS Deployment, GitLab Services, Docker Containers

Igor PELEVANYUK
DIRAC workload management system and BM@N mass production

Andrey DOLBILOV
NICA Computing Management

Dmitriy PODGAYNY, Oksana STRELTSOVA, Maksim ZUEV
HybriLIT and SC Govorun support

Zarif SHARIPOV, Zafar TUKHLIEV
Automation of BM@N Alignment

Alexander AYRIYAN, Vladimir PAPOYAN
Implementation of Particle IDentification based on ML

thanks to the DDC, CICC, NCX & HybriLIT teams for computing support

Thank you for your attention

contact email: *gertsen@jinr.ru*

BM@N Software Ecosystem