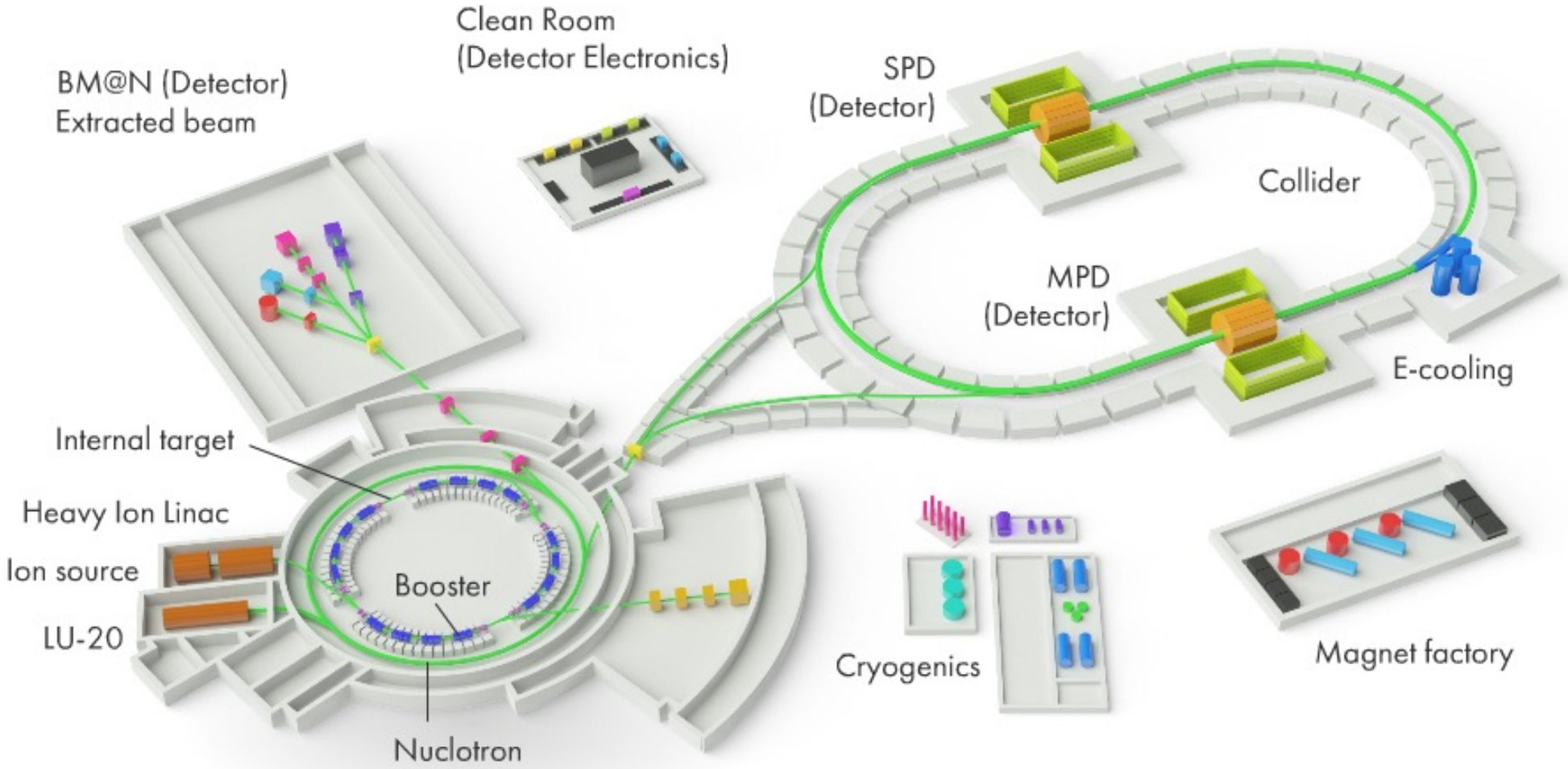
Nuclotron-based Ion Collider fAcility

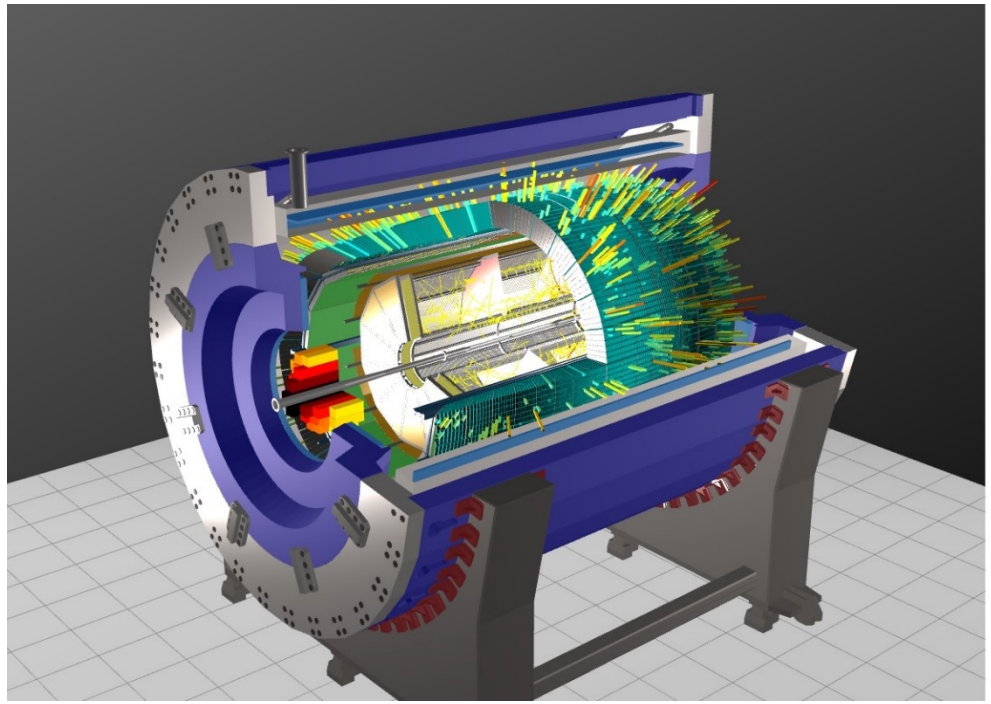


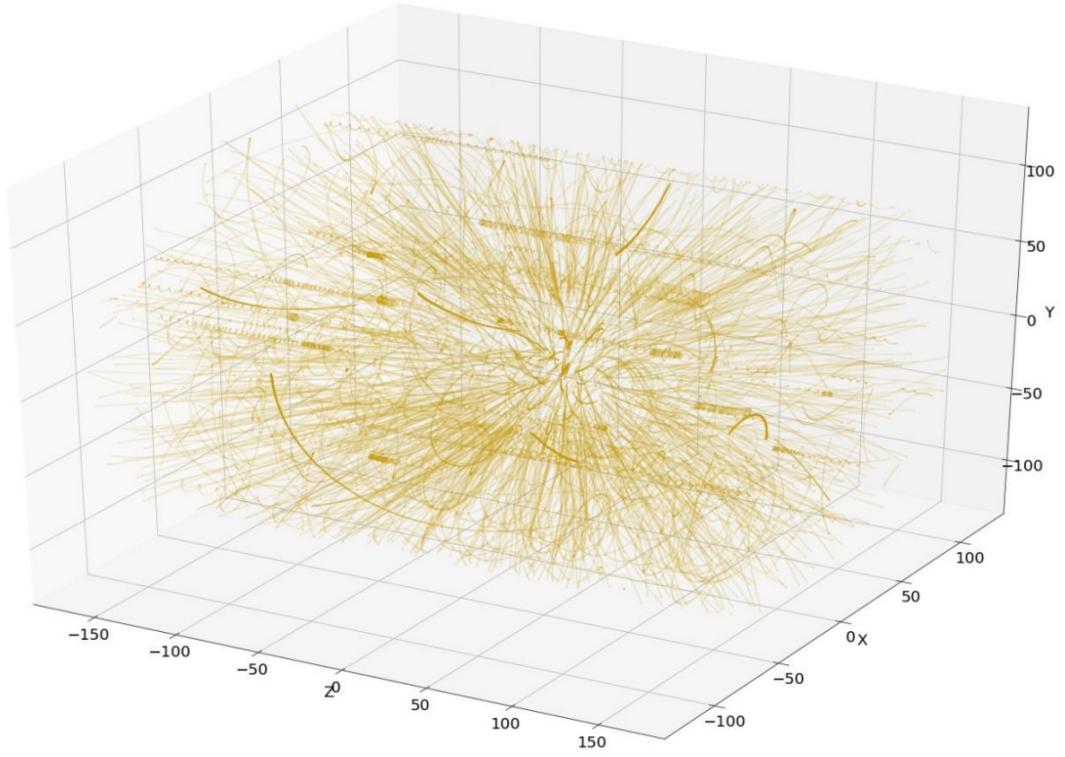
NICA is a new accelerator complex designed at the Joint Institute for Nuclear Research (Dubna, Russia) to study properties of dense baryonic matter.

After putting the **NICA** collider into operation JINR scientists will be able to create in the Laboratory a special state of matter in which our Universe stayed shortly after the Big Bang – the Quark-Gluon Plasma (QGP).

Computing for the NICA megaproject

A large number of tracks in events requires the development of approaches that have constant computational complexity regardless of the number of tracks in an event.

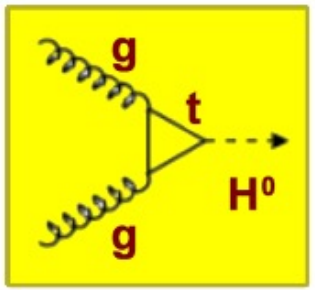




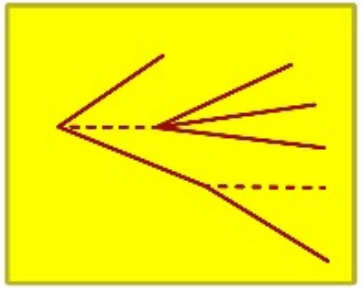
Data collection for NICA is due to start in 2022, but even preparation phase requires multi-petabyte storage with capability to rapidly process hundredterabyte datasets for collision events simulation and reconstruction. Data flow during experiment is expected from tens to hundreds of GB/sec and more with several PB for one experimental run. Meshcheryakov Laboratory of Information Technology (MLIT) at JINR should provide IT support for NICA, i.e. MLIT should provide data collection, data processing in real time and data processing in off-line mode.

HEP experiments data flow

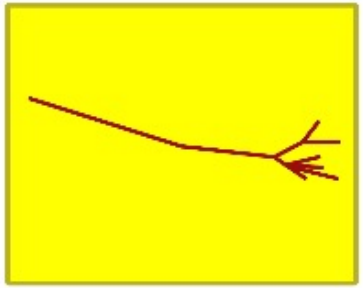
From Physics to raw data



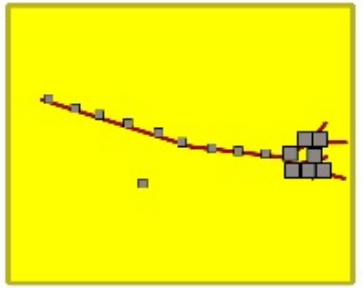
Basic physics



Fragmentation, Decay



Interaction with detector material
Multiple scattering, interactions



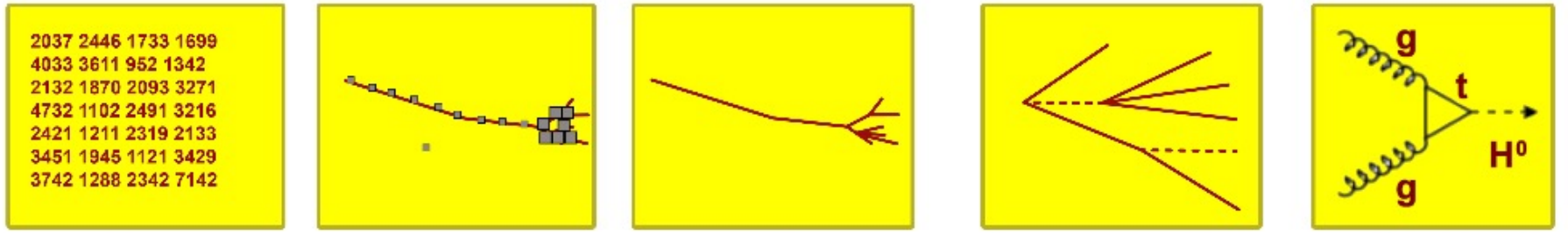
Detector response
Noise, pile-up, cross-talk, inefficiency, ambiguilty, resolution, response function, alignment

Raw data

Read-out addresses, ADC, TDC values, Bit patterns

HEP experiments data flow

From raw data to Physics



Raw data	Detector response	Interaction with detector material	Fragmentation, Decay	Basic physics
Convert to physics quantities	Apply calibration, alignment	Pattern, recognition, Particle identification	Physics analysis	Results

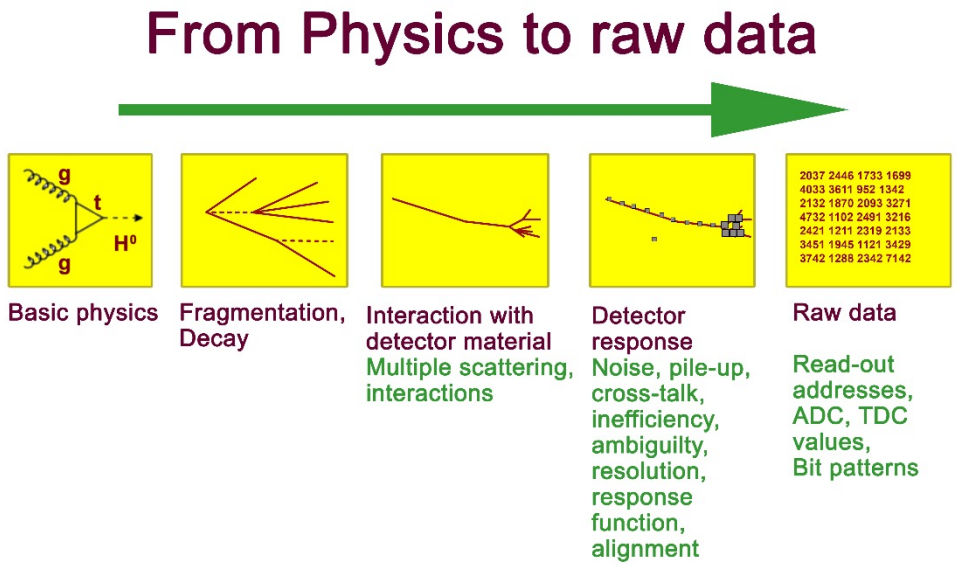
Reconstruction

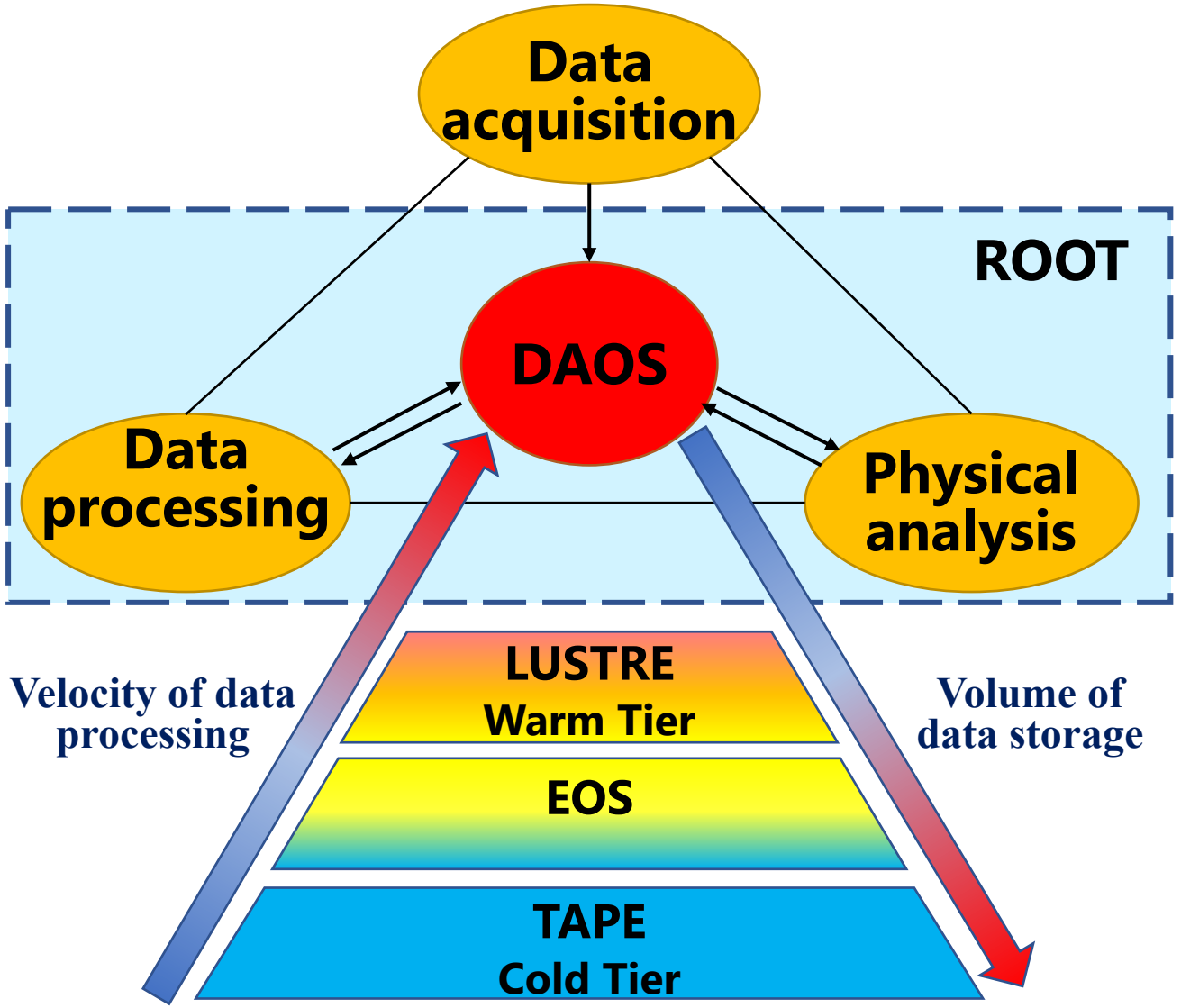
Analysis

Simulation (Monte-Carlo)

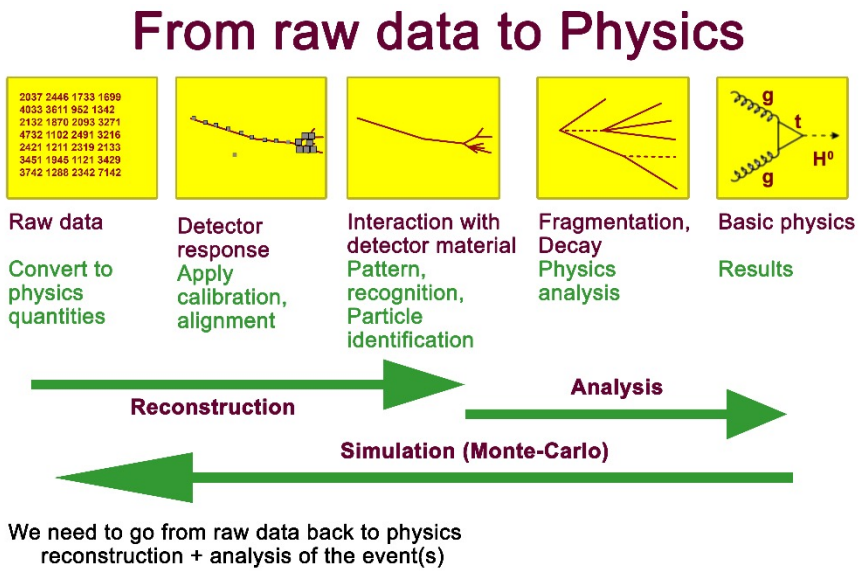
We need to go from raw data back to physics reconstruction + analysis of the event(s)

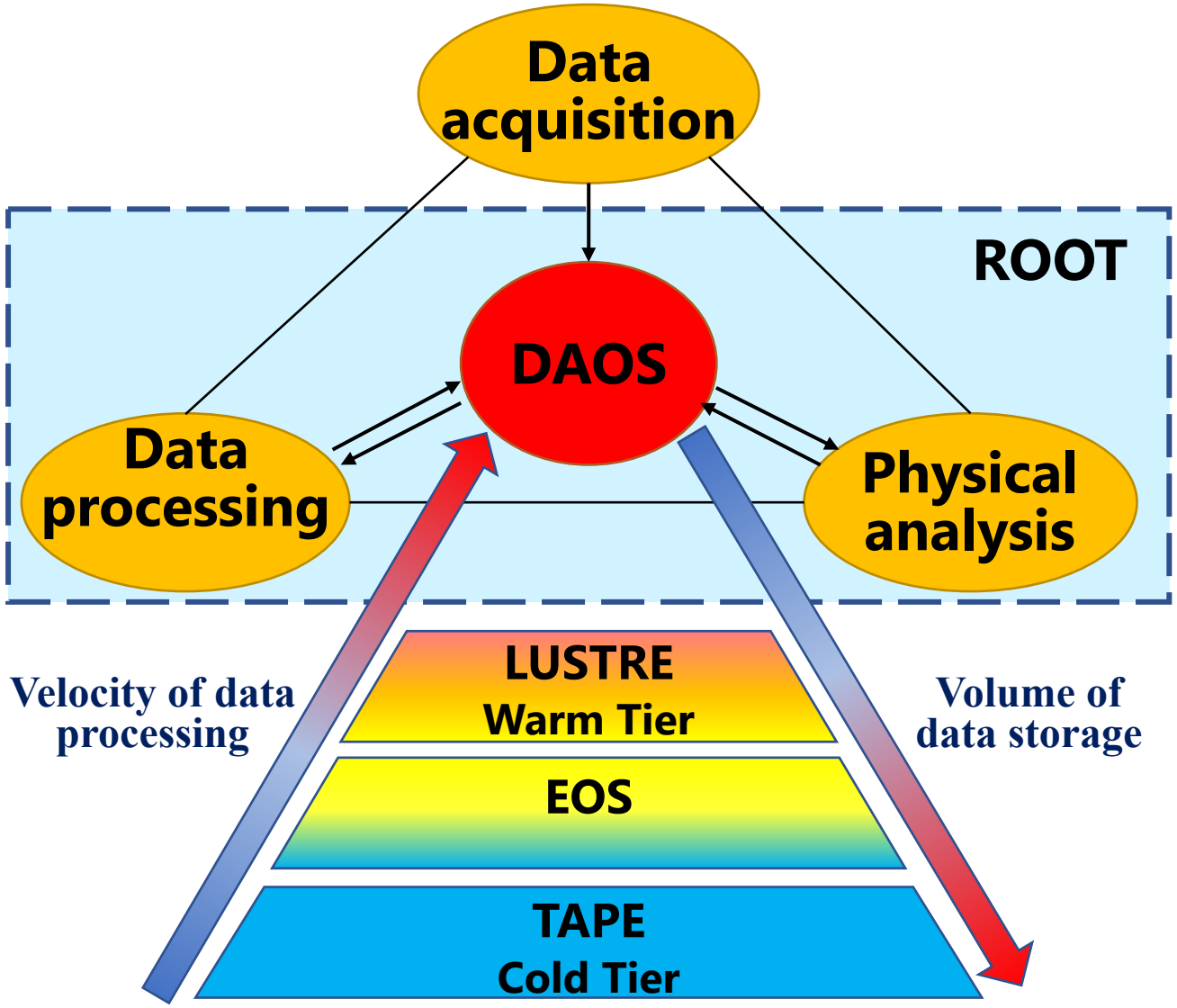
HEP experiments data flow





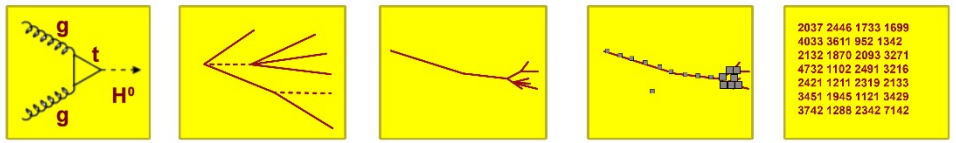
HEP experiments data flow





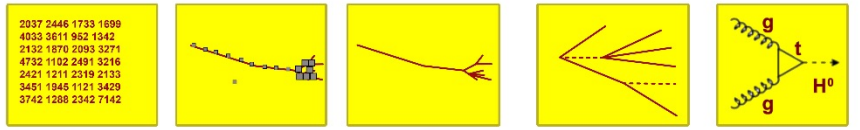
HEP experiments data flow

From Physics to raw data



Basic physics | Fragmentation, Decay | Interaction with detector material: Multiple scattering, interactions | Detector response: Noise, pile-up, cross-talk, inefficiency, ambiguity, resolution, response function, alignment | Raw data: Read-out addresses, ADC, TDC values, Bit patterns

From raw data to Physics



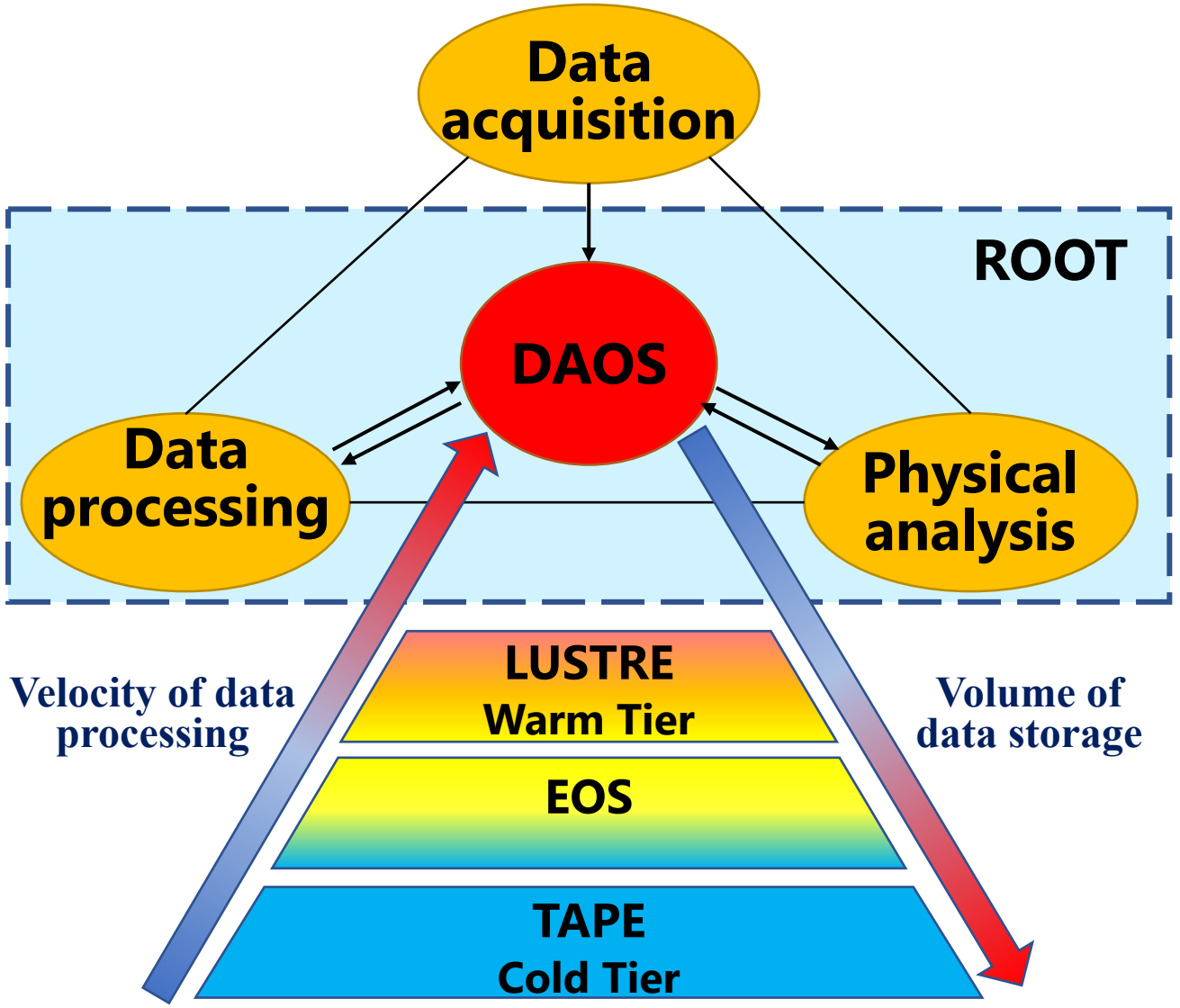
Raw data: Convert to physics quantities | Detector response: Apply calibration, alignment | Interaction with detector material: Pattern, recognition, Particle identification | Fragmentation, Decay: Physics analysis | Basic physics: Results

Reconstruction

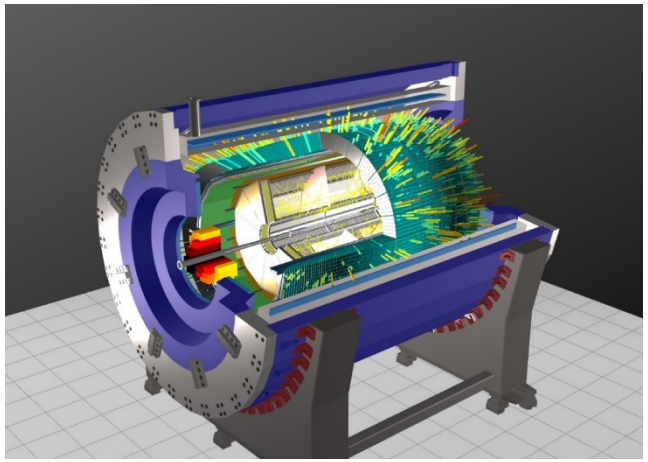
Analysis

Simulation (Monte-Carlo)

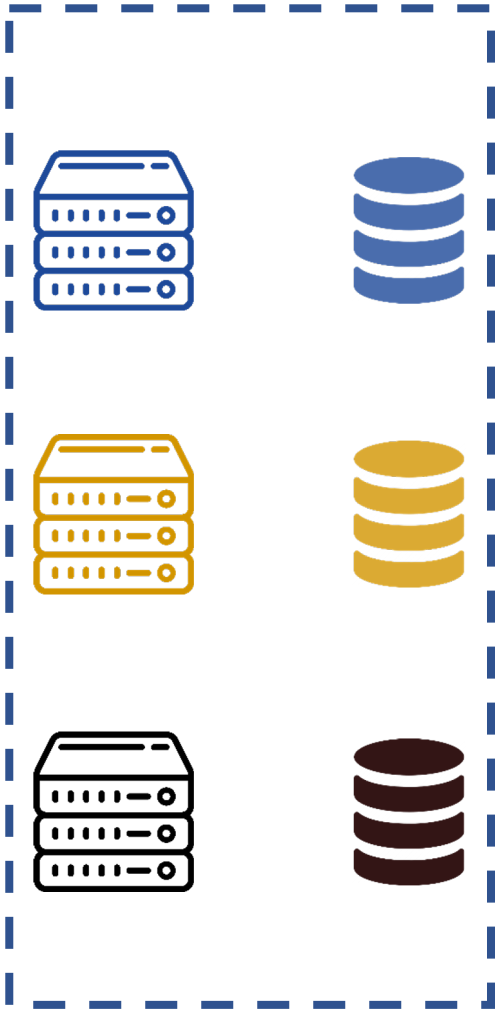
We need to go from raw data back to physics reconstruction + analysis of the event(s)



High Energy Physics workflow before DAOS



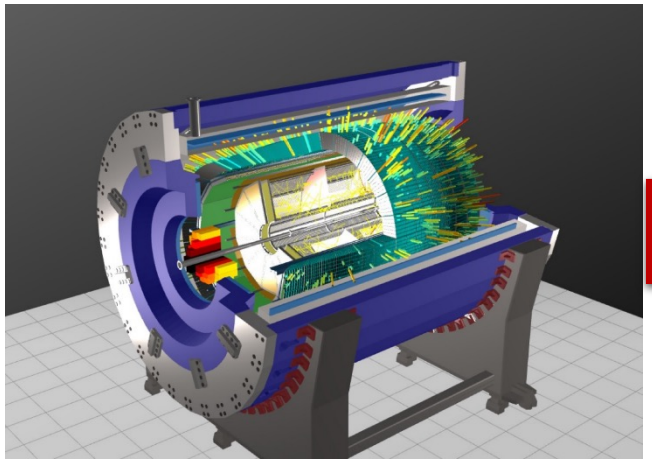
GRID



Scientists

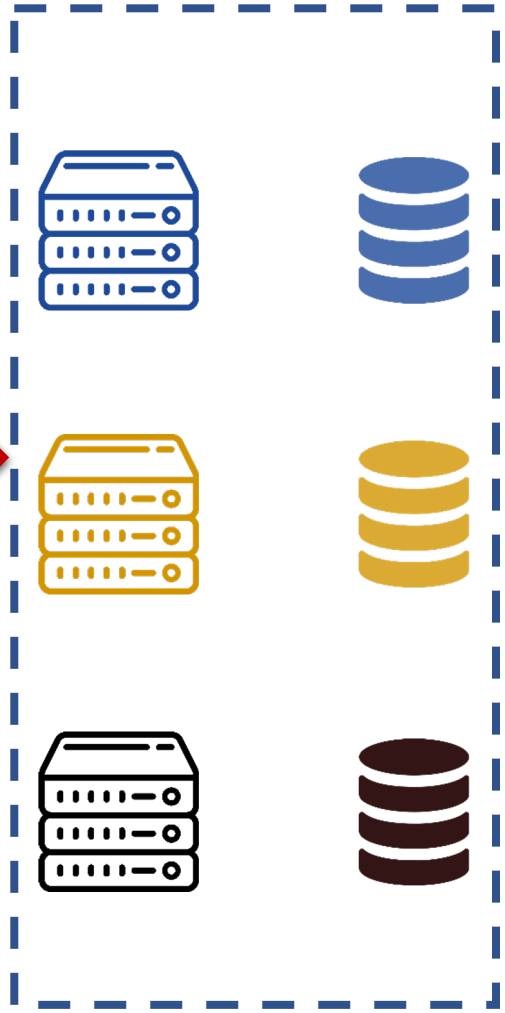
We load GRID infrastructure at all stages

High Energy Physics workflow before DAOS



RAW data

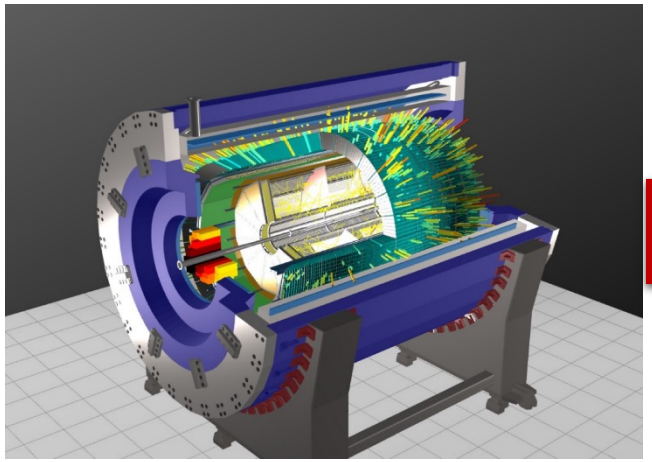
GRID



Scientists

We load GRID infrastructure at all stages

High Energy Physics workflow before DAOS



RAW data

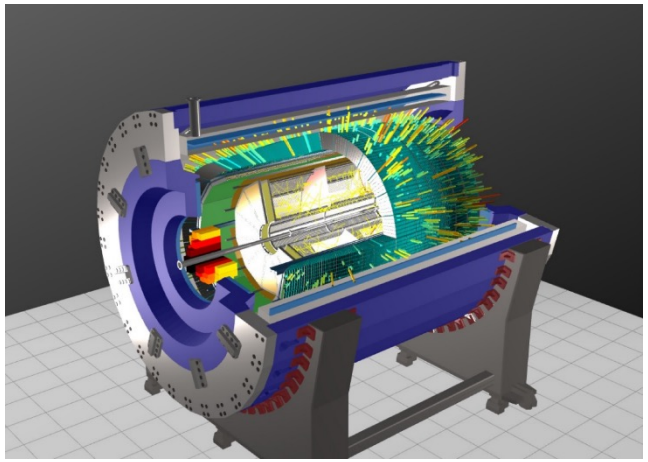
GRID

Processing

Scientists

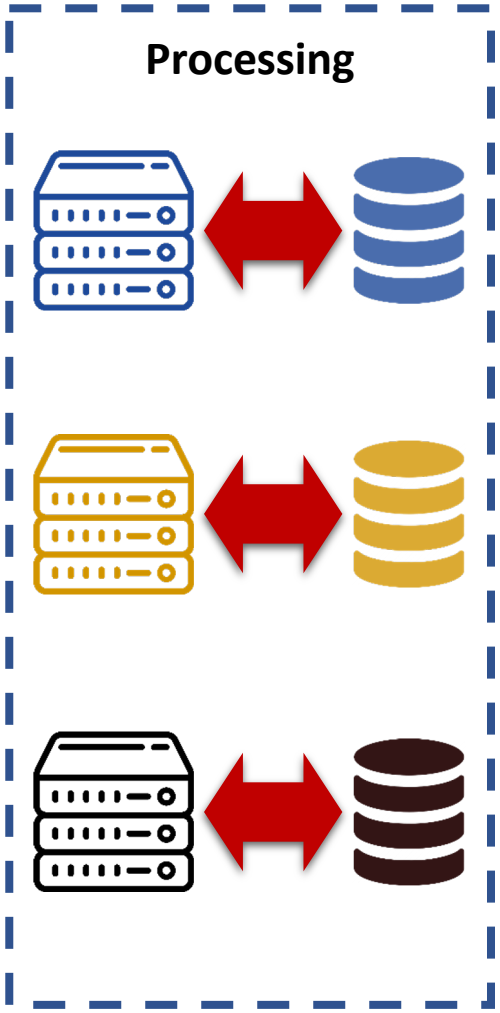
We load GRID infrastructure at all stages

High Energy Physics workflow before DAOS

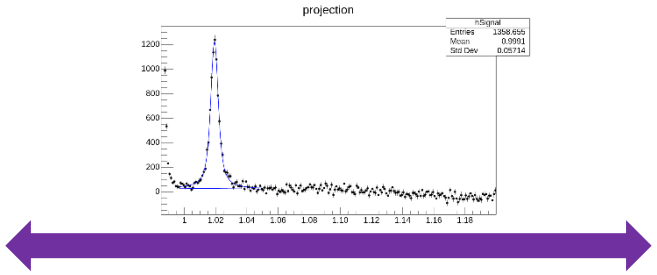


GRID

Processing



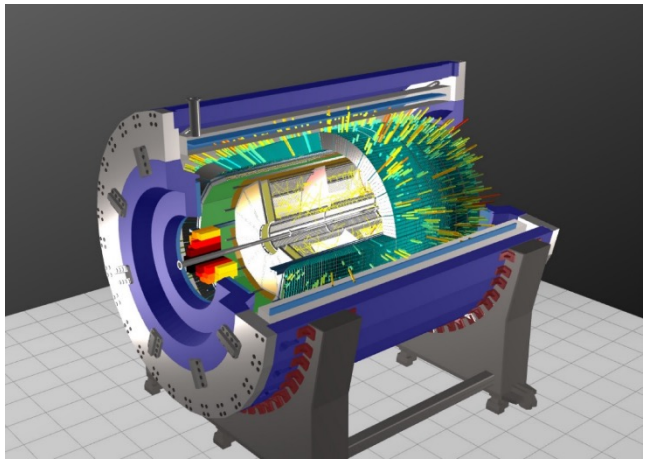
Analysis requests to **GRID**



Scientists

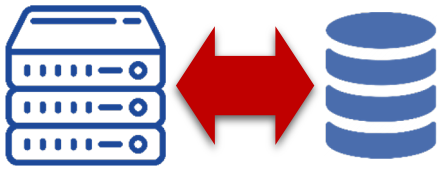
We load GRID infrastructure at all stages

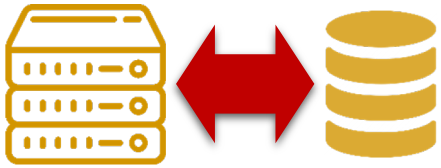
High Energy Physics workflow before DAOS

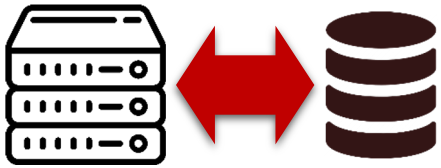


GRID

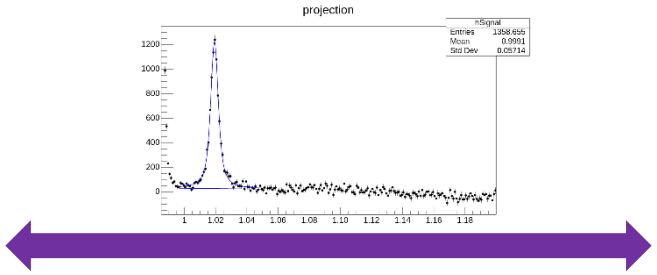
Processing

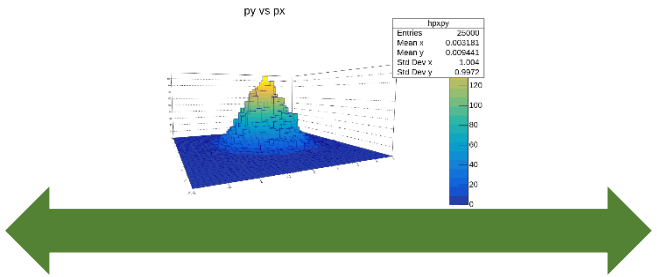






Analysis requests to **GRID**

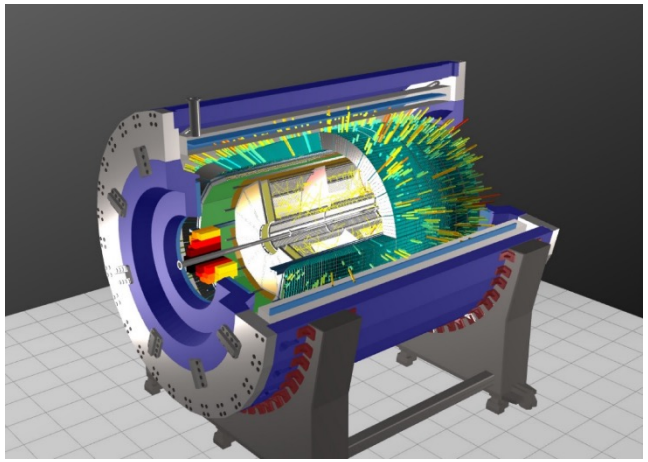




Scientists

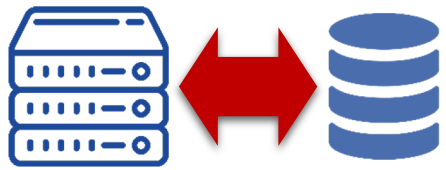
We load GRID infrastructure at all stages

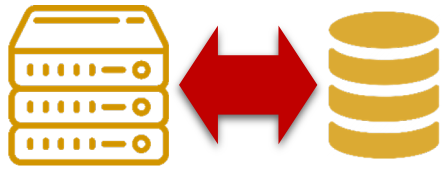
High Energy Physics workflow before DAOS

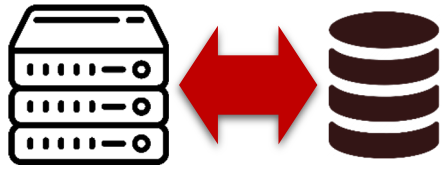


GRID

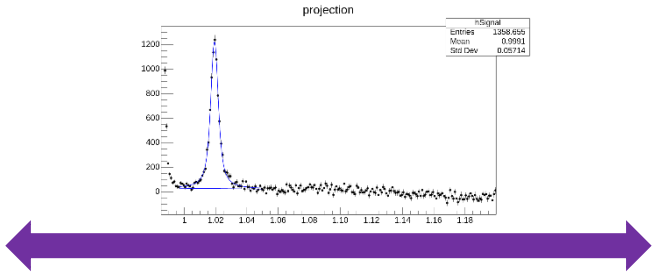
Processing

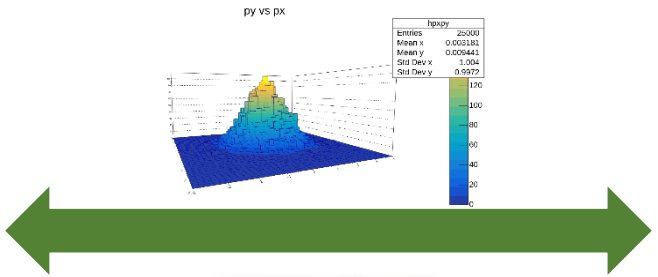






Analysis requests to **GRID**





Scientists

We load GRID infrastructure at all stages

“Govorun” supercomputer

8 nodes: 4 PMEM and 2 SSD per one node
Total: 8 TB PMEM and 32 TB SSD

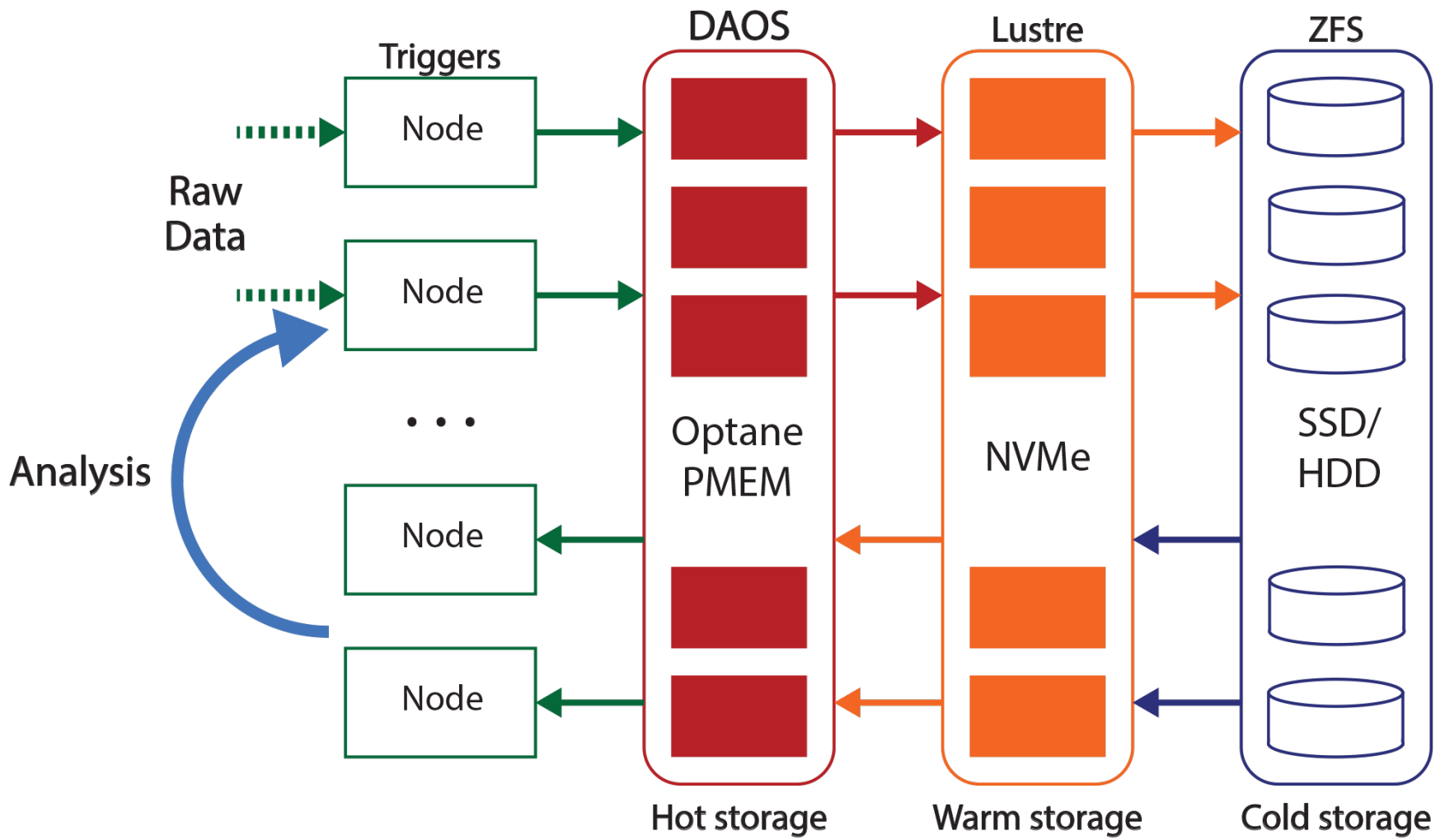
Upcoming 8 PB

The “Govorun” supercomputer is a hyper-converged software-defined system and occupied **26th** and **31st** in the current edition of the **IO500** list (July 2021). For the high-speed data storage system, RSC Group has received the prestigious Russian DC Awards 2020 in “The Best IT Solution for Data Centers” nomination at the awards ceremony held on 10 December 2020 in Moscow.

CPU-component based on the newest Intel architectures:
Intel Xeon Phi gen.2 and Intel CascadeLake CPUs

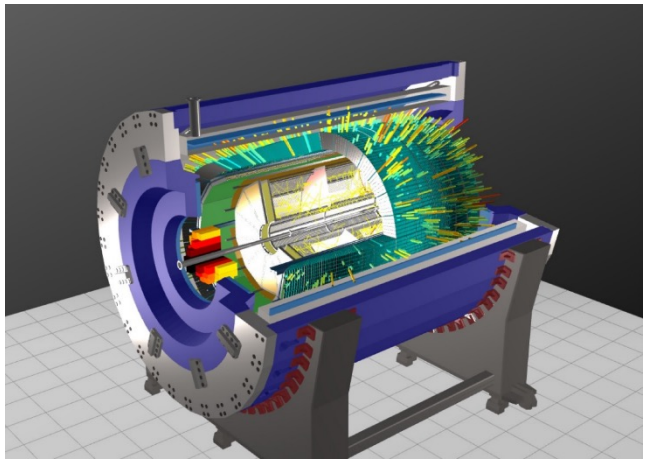
Total peak performance:
1.7 PFLOPS SP
860 TFLOPS DP
300 Gb/s Data IO rate

Hierarchical system for processing and storing data on the “Govorun” supercomputer

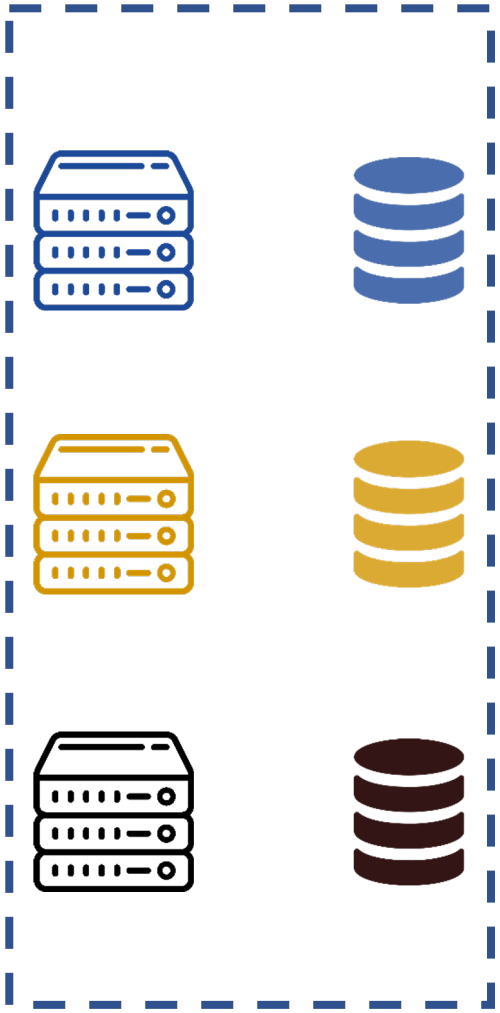


The storage and memory hierarchy is differentiated by the type and temperature of the data. The data workflow is significantly more advanced than just managing multi tier storage layout. It uses a storage tier which is appropriate by the current state of the data processing. RSC Basis Orchestration software is designed for that purpose.

High Energy Physics workflow with DAOS

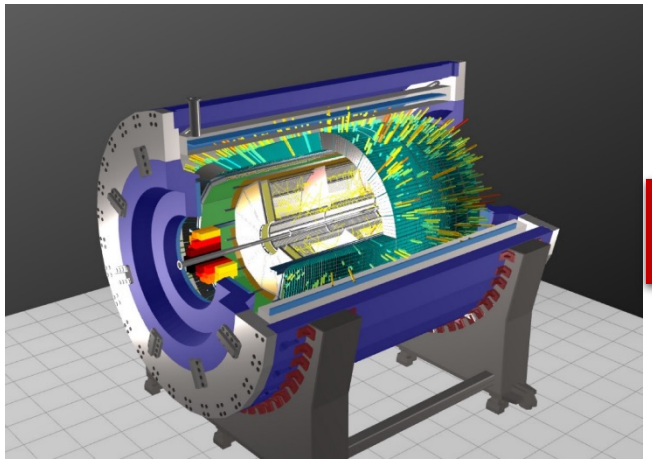


GRID



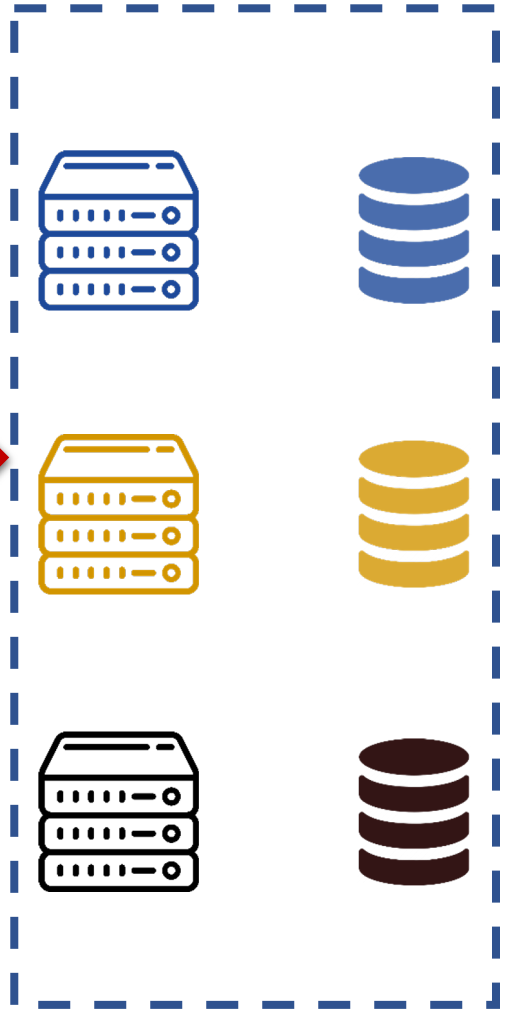
Scientists

High Energy Physics workflow with DAOS



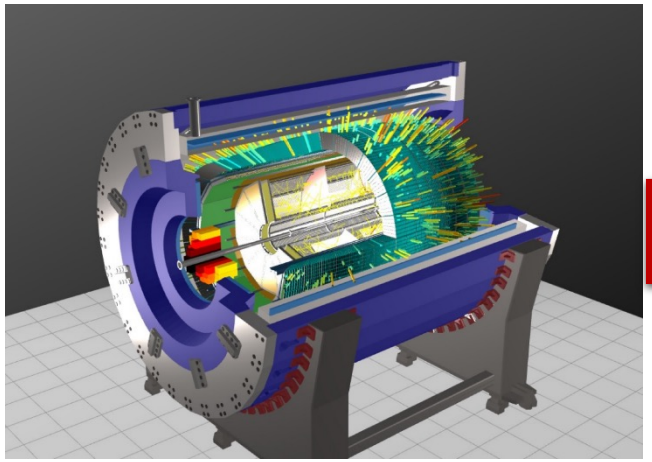
RAW data

GRID



Scientists

High Energy Physics workflow with DAOS



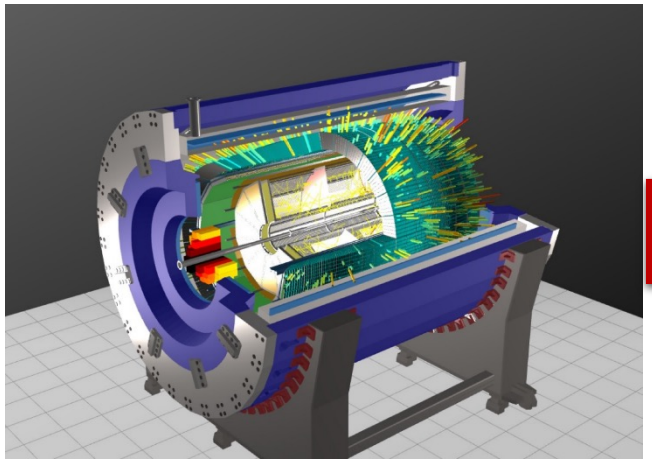
RAW data

GRID

Processing

Scientists

High Energy Physics workflow with DAOS

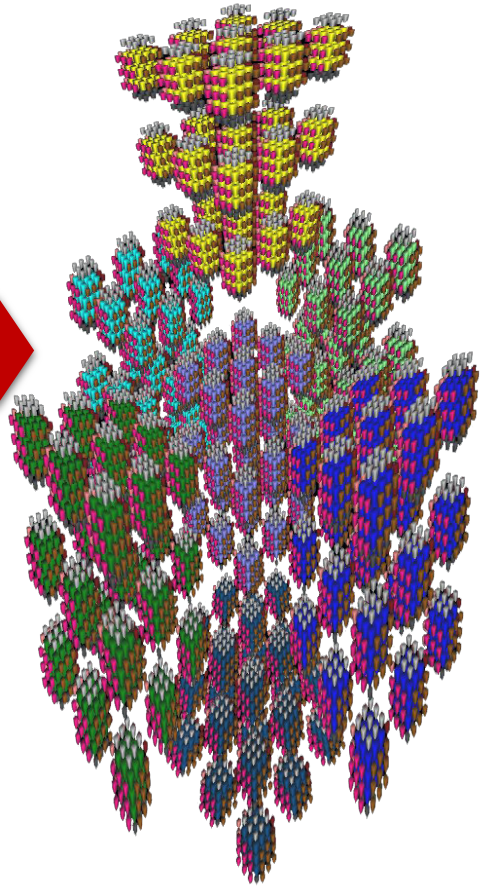


RAW data

GRID

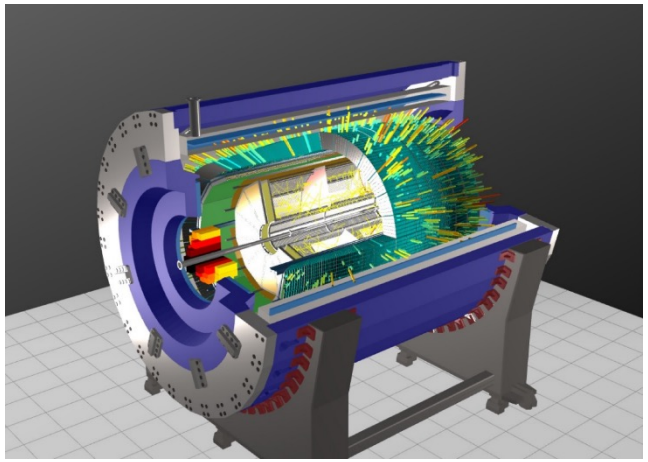
Processing

Presentation layer with DAOS

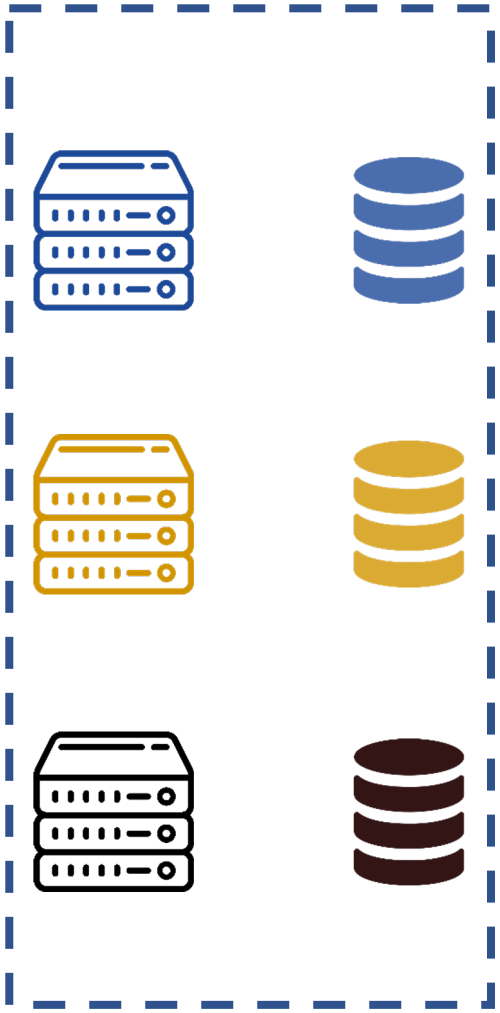


Scientists

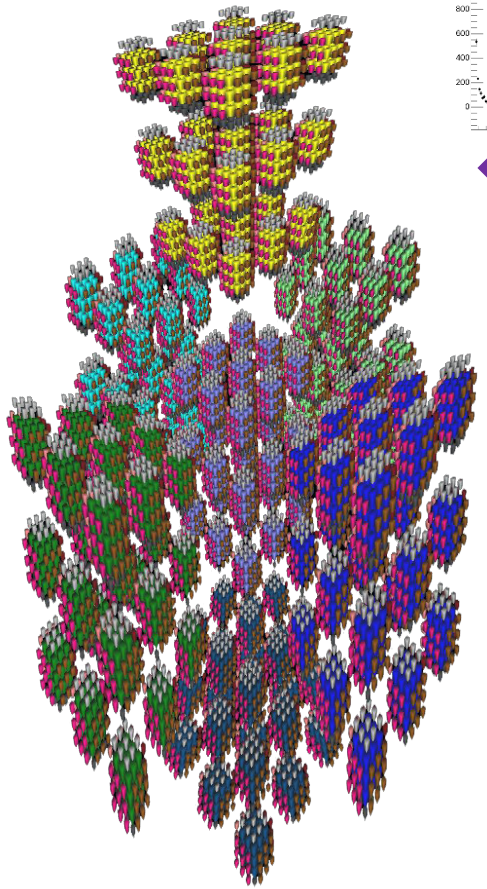
High Energy Physics workflow with DAOS



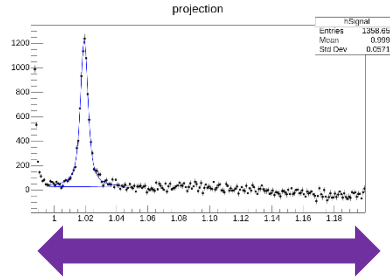
GRID



Presentation layer with DAOS

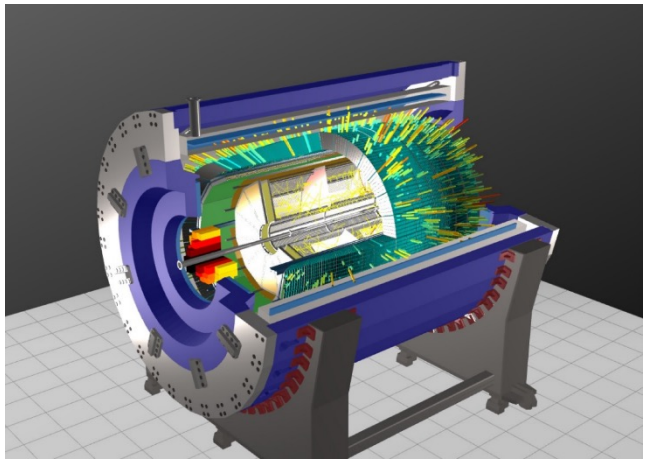


Analysis requests to **DAOS** only



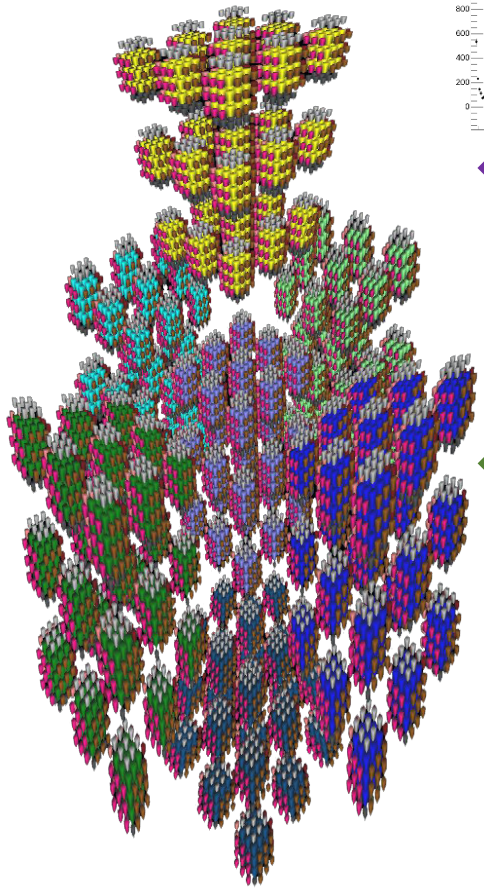
Scientists

High Energy Physics workflow with DAOS

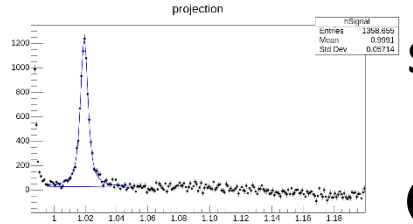


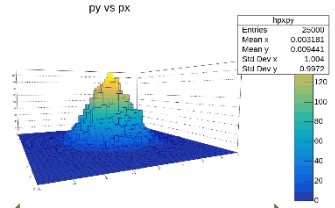
GRID

Presentation layer with DAOS



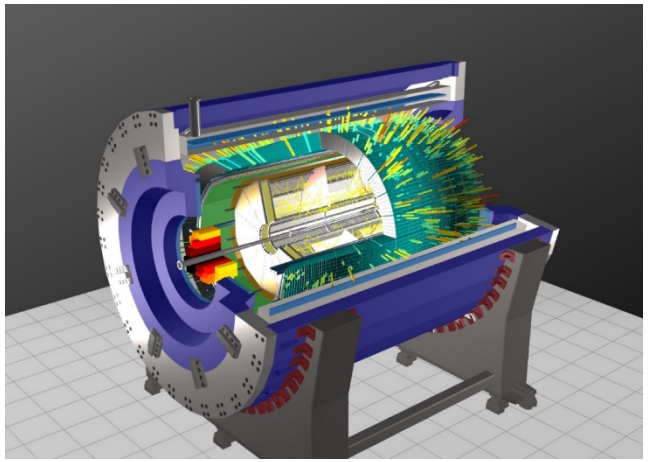
Analysis requests to **DAOS** only





Scientists

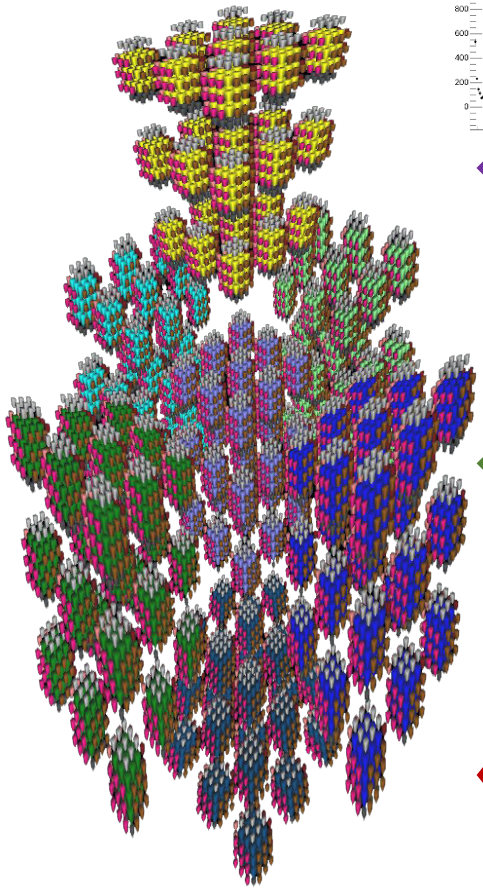
High Energy Physics workflow with DAOS

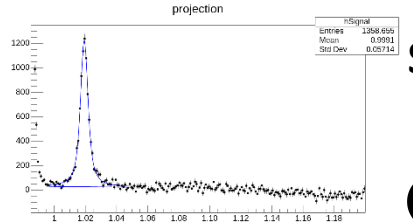


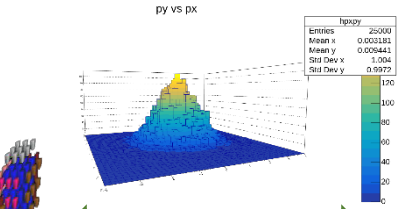
GRID

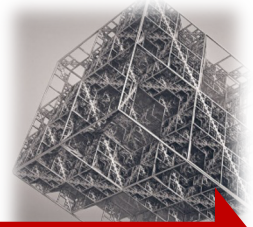
Presentation layer with DAOS

Analysis requests to **DAOS** only



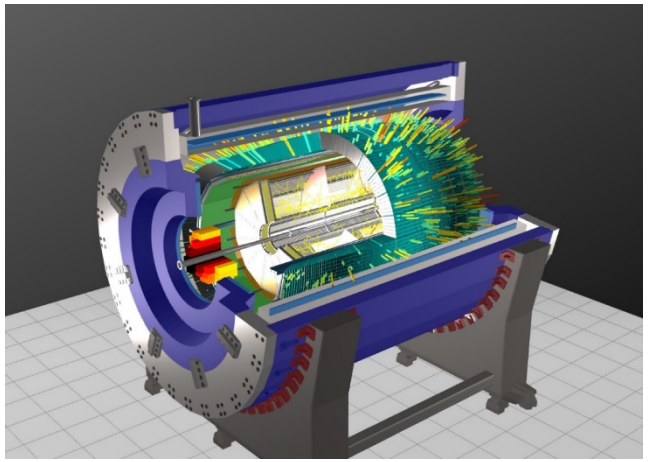






Scientists

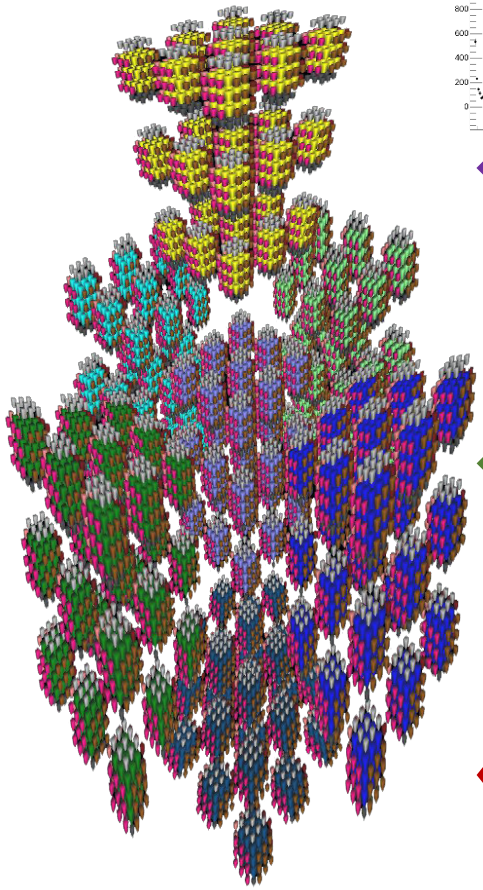
High Energy Physics workflow with DAOS

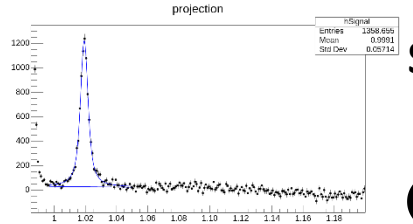


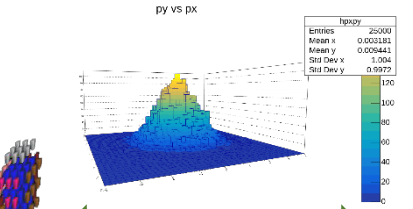
GRID

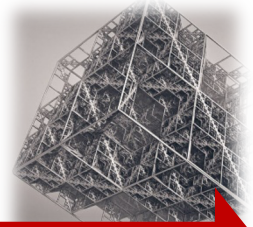
Presentation layer with DAOS

Analysis requests to **DAOS** only









Scientists

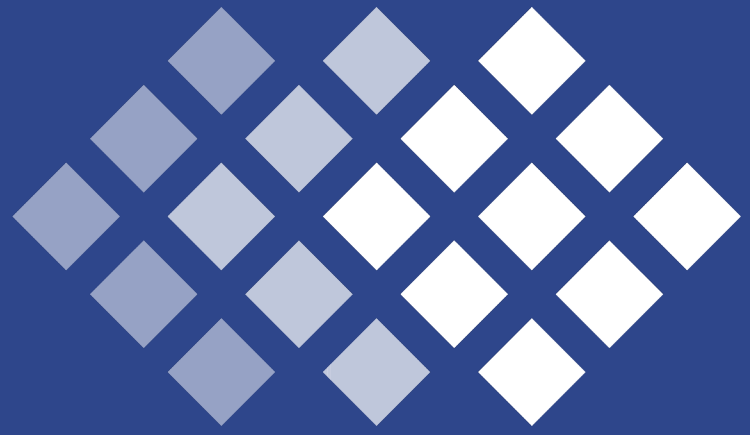
DAOS eliminates GRID usage at Analysis stage

Results

Of DAOS usage in High Energy Physics workflow:

- Allows us to store and read multidimensional data structure at TB scale in shared address space
- Allows to create multiuser presentation layer for scientific results analysis by researchers
- Many hundreds times cheaper than DDR memory usage
- Reduce GRID infrastructure usage (compute/storage/network) for analysis stage
- Easily integrates with other hot/warm storages

RSC Group



Leading Solutions for HPC and Data Centers

16th place in the ISC21 10 node challenge!

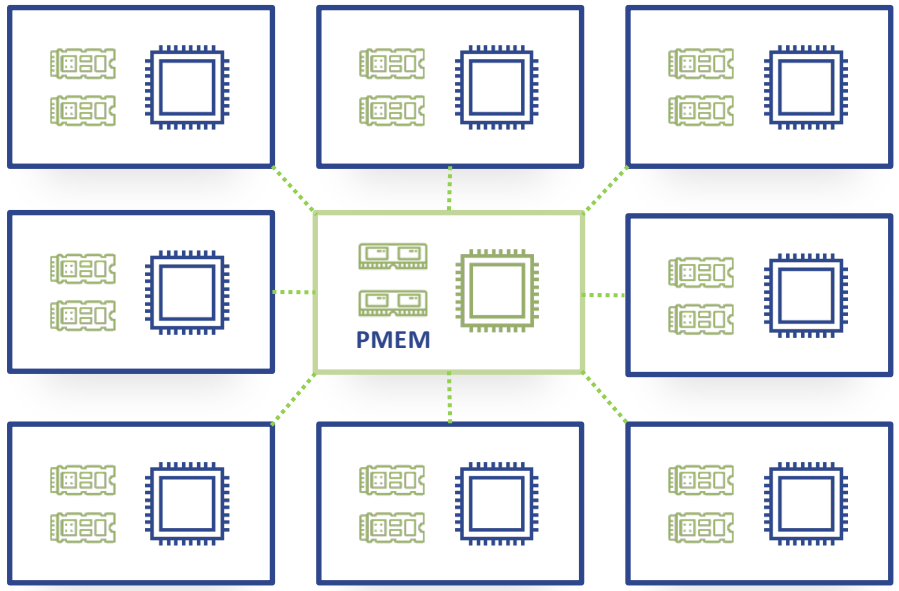
JINR

RSC

14	SC19	WekaIO	WekaIO	WekaIO	WekaIO Matrix	10	2,610	156.51	56.22	435.76
15	ISC21	University of Tsukuba	Cygnus	OSS	CHFS	10	240	148.69	30.39	727.61
16	ISC21	Joint Institute of Nuclear Research	Govorun	RSC	DAOS	10	160	132.06	20.19	863.69
17	SC20	TACC	Frontera	DDN	IME	10	280	109.91	176.23	68.55
18	ISC21	Japan Agency for Marine-Earth Science and Technology	Earth Simulator	DDN	Lustre	10	320	101.88	48.19	215.38

daos

RSC

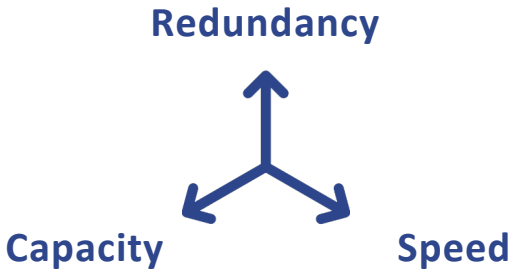


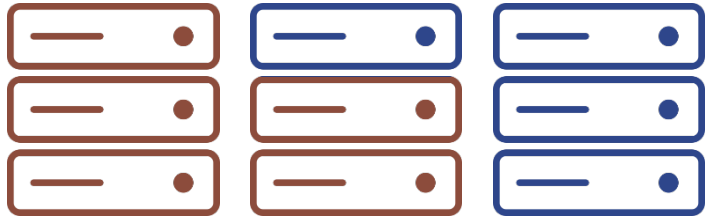
DAOS System with NVMeOF drives

Problems solved by RSC Storage on-Demand

How to aggregate underutilized cluster resources in order to more efficiently use the resources of the data center and increase its performance by $?

Different users have different storage requirements. How can all of their needs be met?



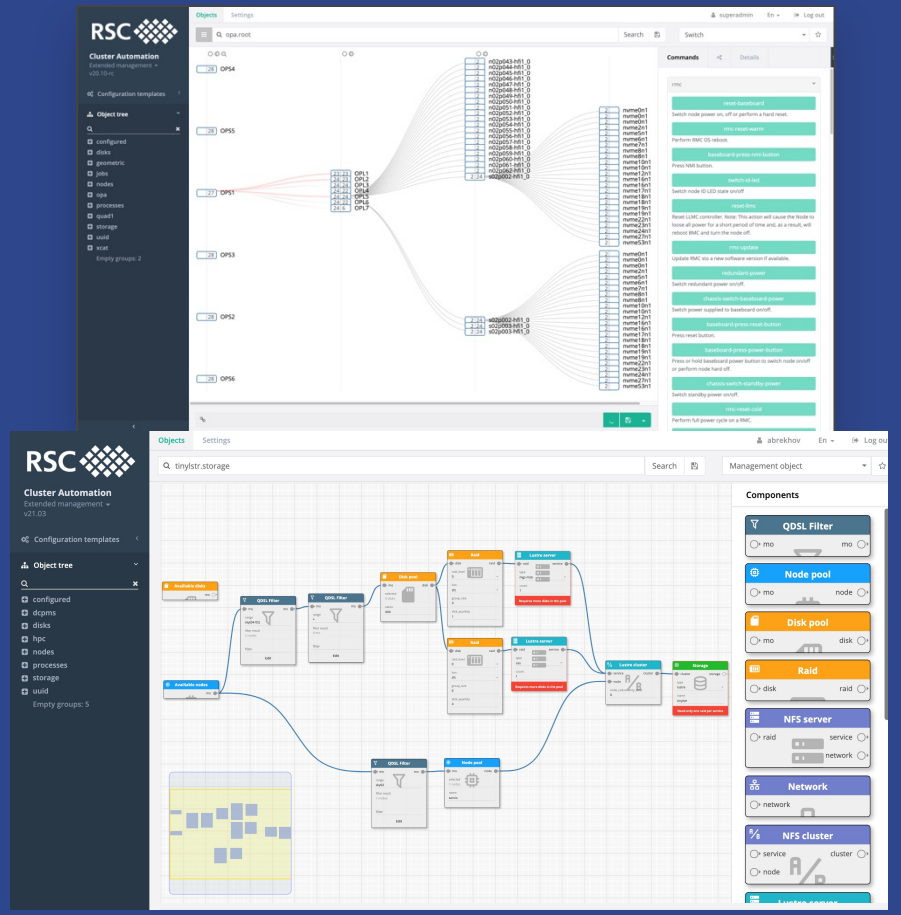


Is it possible to eliminate the overhead of security and quotas and isolate resources quickly and easily?

Data Center Storage Orchestrator

RSC

RSC BasIS Storage on-Demand



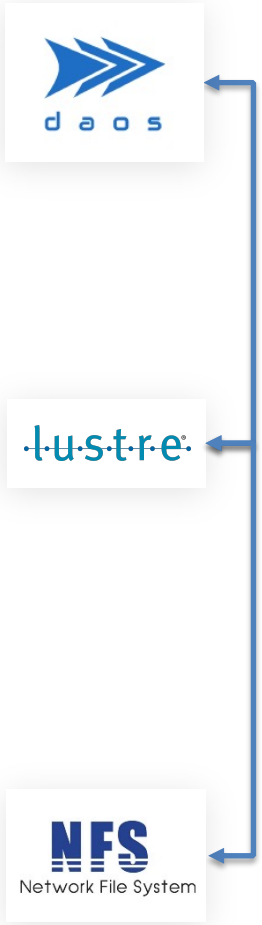
RSC BasIS Storage on-Demand is resource orchestrator for composable and disaggregated infrastructure.

It consists of two components:

- **cluster management platform which handle all the information about available cluster resources (BasIS Stack)**

- **scheduler that efficiently manages storage resources and creates storage systems with varying lifespan and on demand (BasIS Storage on-demand)**

It’s a lot of storages right now



There’s a need to automate data moving and standardize their presentation