# Development of computing infrastructure for distributed data processing for SPD experiment

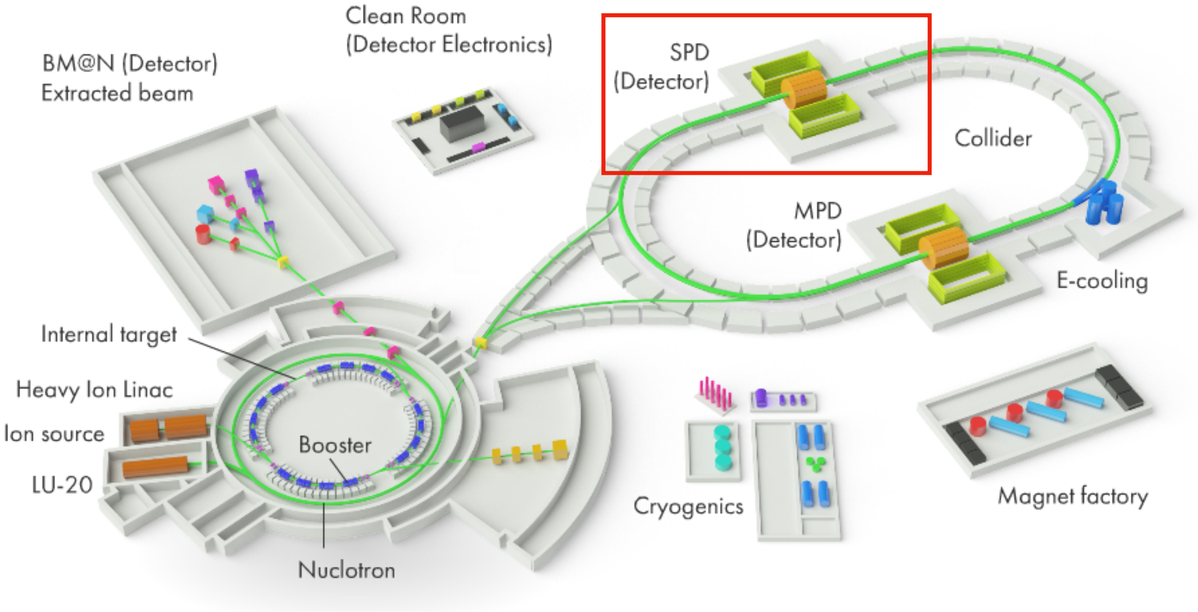
**Danila Oleynik, JINR LIT**

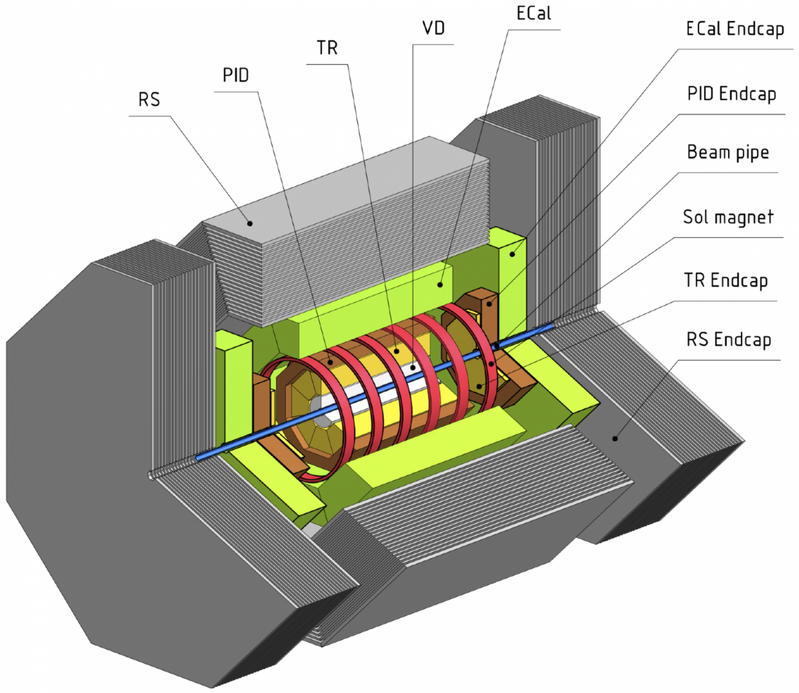
# Outline

- SPD experiment at NICA
- Data flow and projected amount of data
- Distributed heterogeneous computing infrastructure in JINR
- Services for data processing at distributed computing infrastructure
- Unified resources management system

# SPD Spin Physics Detector

Study of the nucleon spin structure and spin-related phenomena in polarized *p-p*, *d-d* and *p-d* collisions



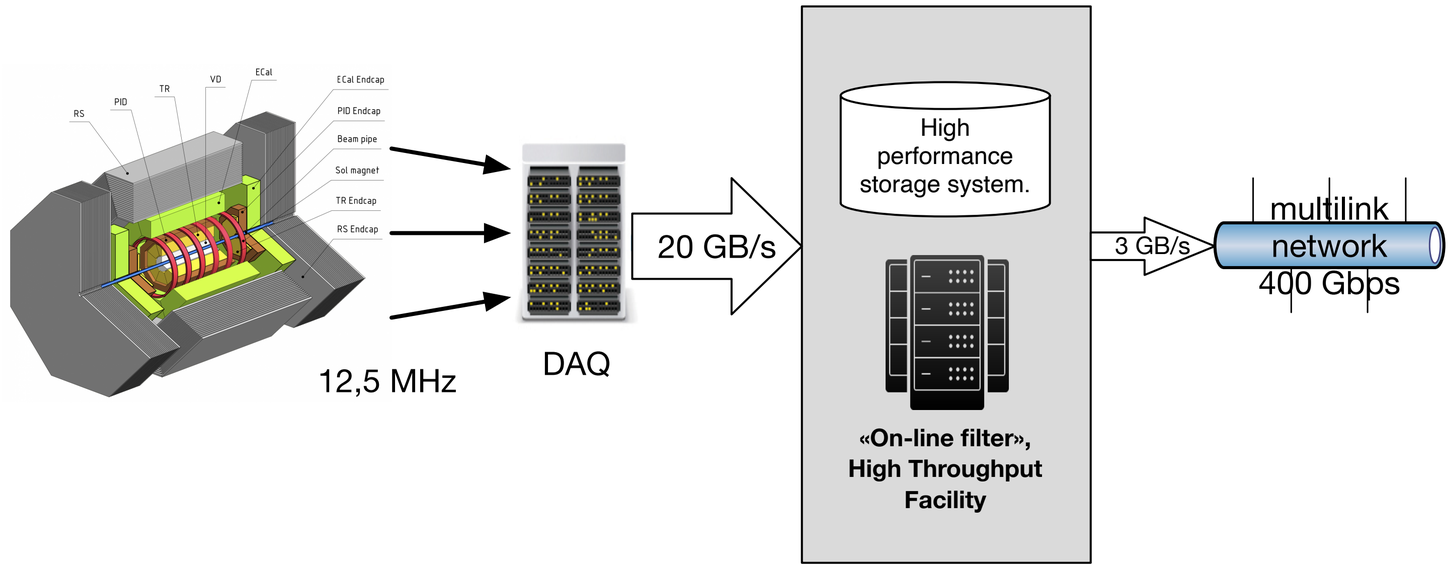

SPD - a universal facility for comprehensive study of gluon content in proton and deuteron

# SPD as data source

- Bunch crossing every 80 ns = crossing rate 12.5 MHz
- ~ 3 MHz event rate (at $10^{32}$ $cm^{-2}s^{-1}$ design luminosity) = **pileups**
- **20 GB/s** (or **200 PB/year** "raw" data, $3*10^{13}$ events/year)
  - High requirements for longterm storage
  - Selection of physics signal requires momentum and vertex reconstruction → no simple trigger is possible
  - Comparable amount of simulated data

# SPD as data source

## On-line facility

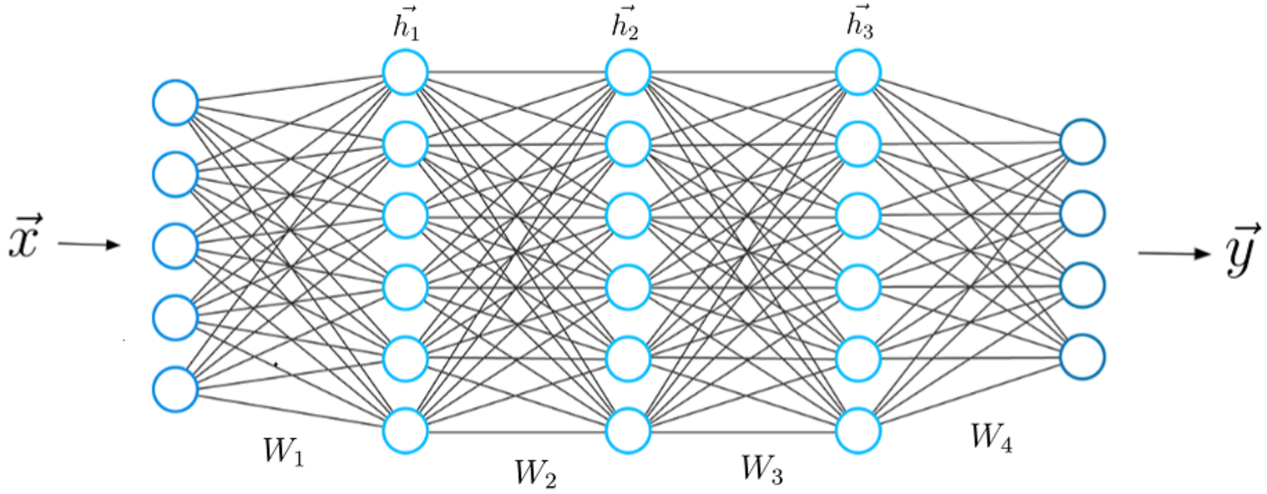



- “On-line filter” - dedicated high throughput computing facility with integrated high performance storage system for:
  - Intelligent data reduction
  - Initial data organization

# On-line filter

## details

- Partial reconstruction
  - Fast tracking
  - Fast ECAL clustering
- Event unscrambling
- Software trigger
  - several data streams
- Monitoring and Data quality assessment
- Local polarimetry

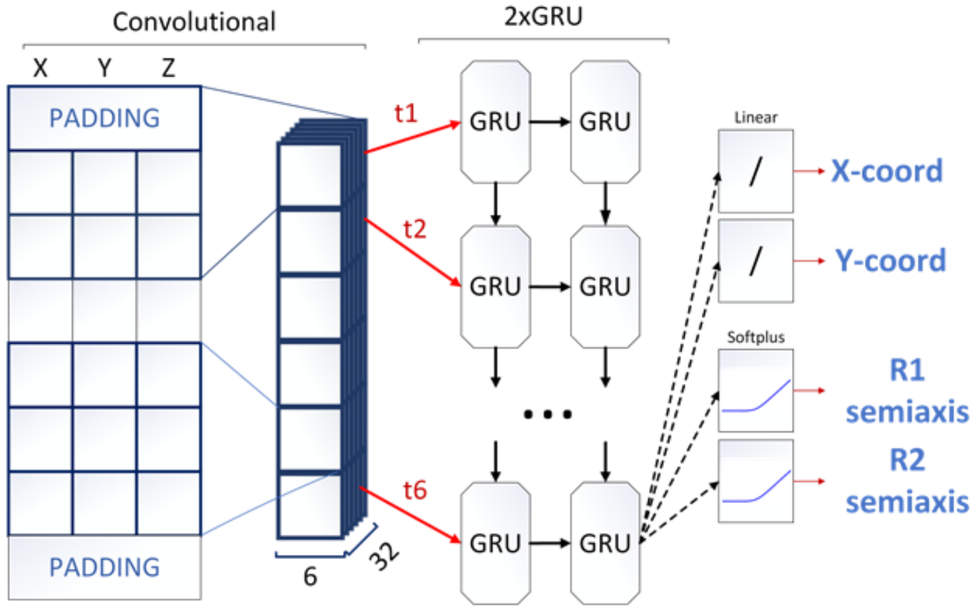



Machine learning is a promising technology

# Machine learning for SPD

## Under research: TrackNETv2

- works like learnable version of the Kalman filter
- for the starting part of a track predicts an elliptical area at the next station where to search for the continuation
- if there is not continuation candidate track is thrown away
- Results (Based on BM@N experiment data):
  - 12K tracks/sec on Intel Core i3-4005U @1.70 Ghz
  - 96% of tracks were reconstructed without any mistake



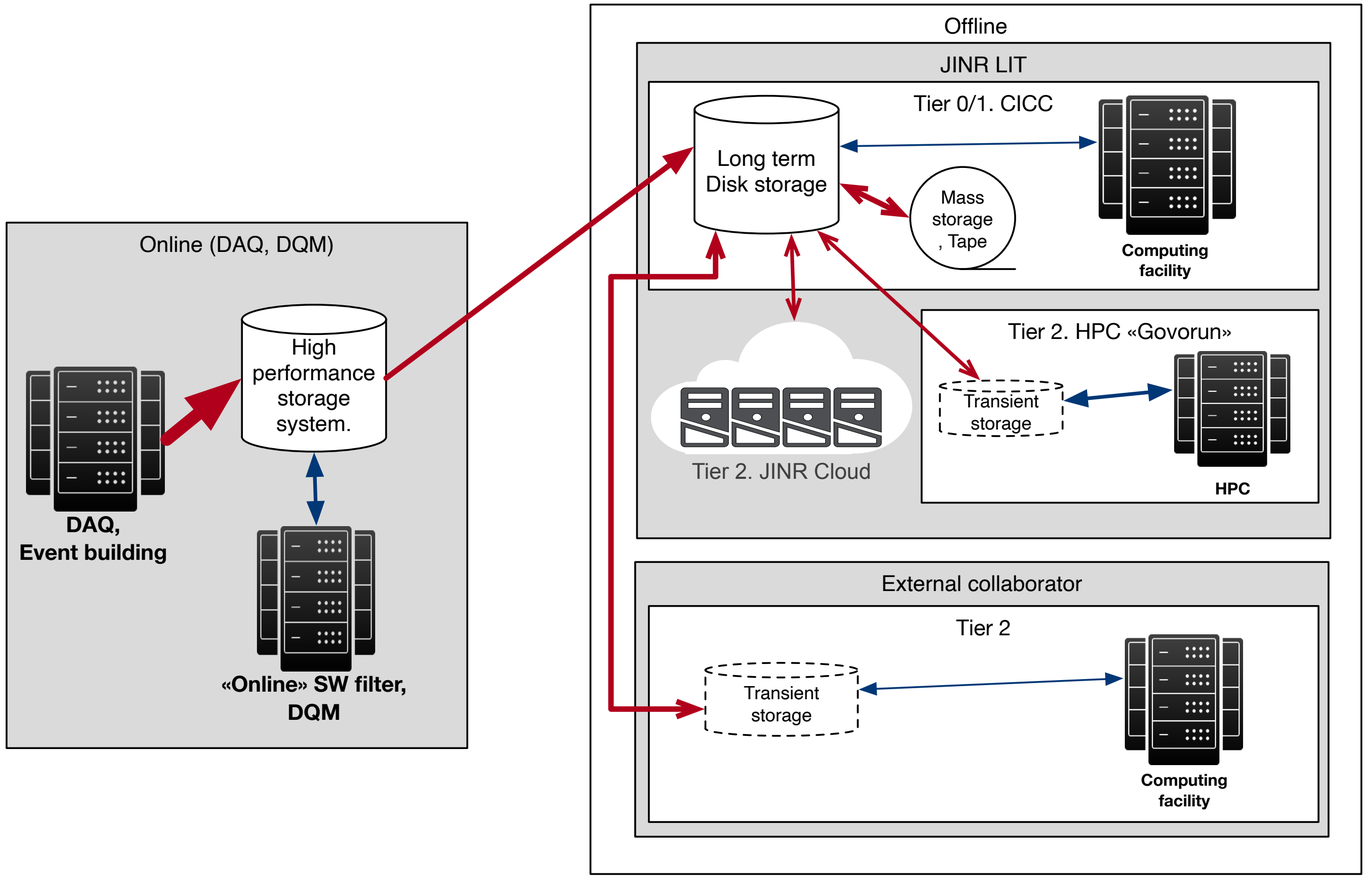

*P.Goncharov, G. Ososkov, D. Baranov*
*AIP Conf 2163, 040003 (2019)*

*Work supported by the RFBR-NFSC project No. 19-57-53002*

# Off-line computing

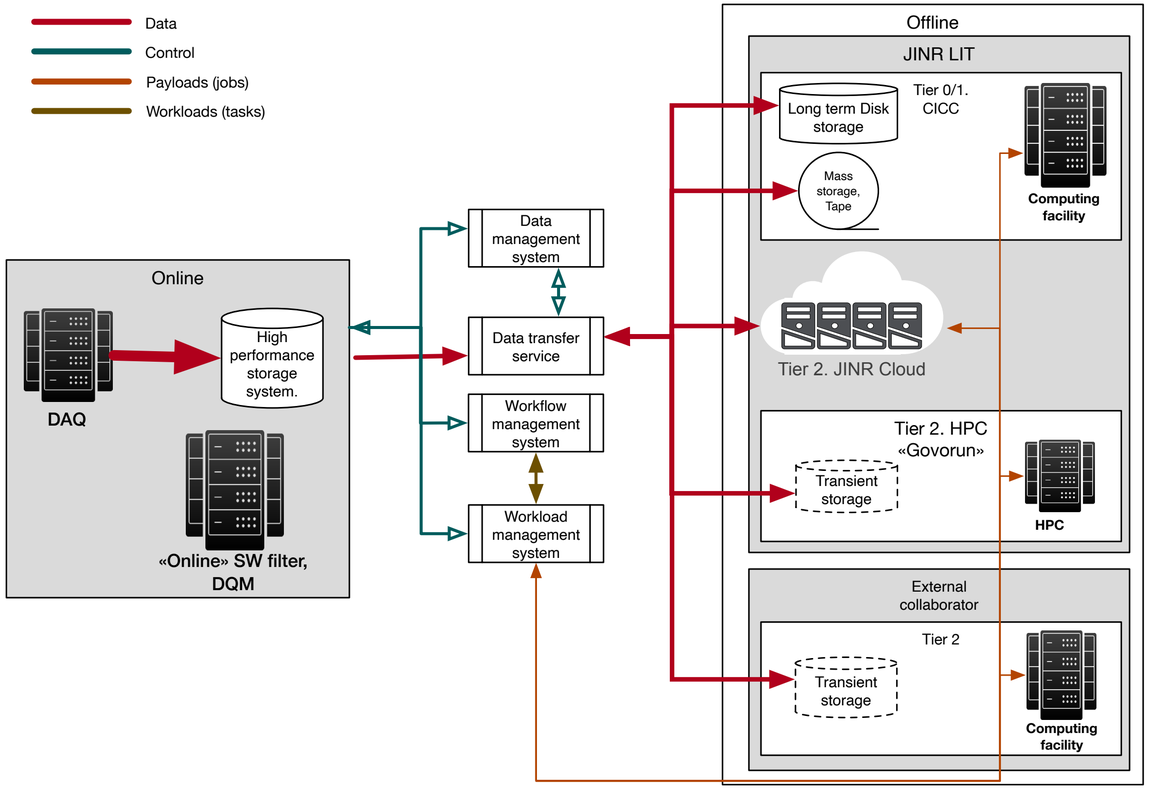
## Infrastructure

- Tier 0
  - Longterm data storage
  - Final reconstruction
- Tier 1
  - Reprocessing
  - Simulation
- Tier 2
  - Simulation
  - Analysis



# Automation of data processing
## in heterogenous distributed computing system

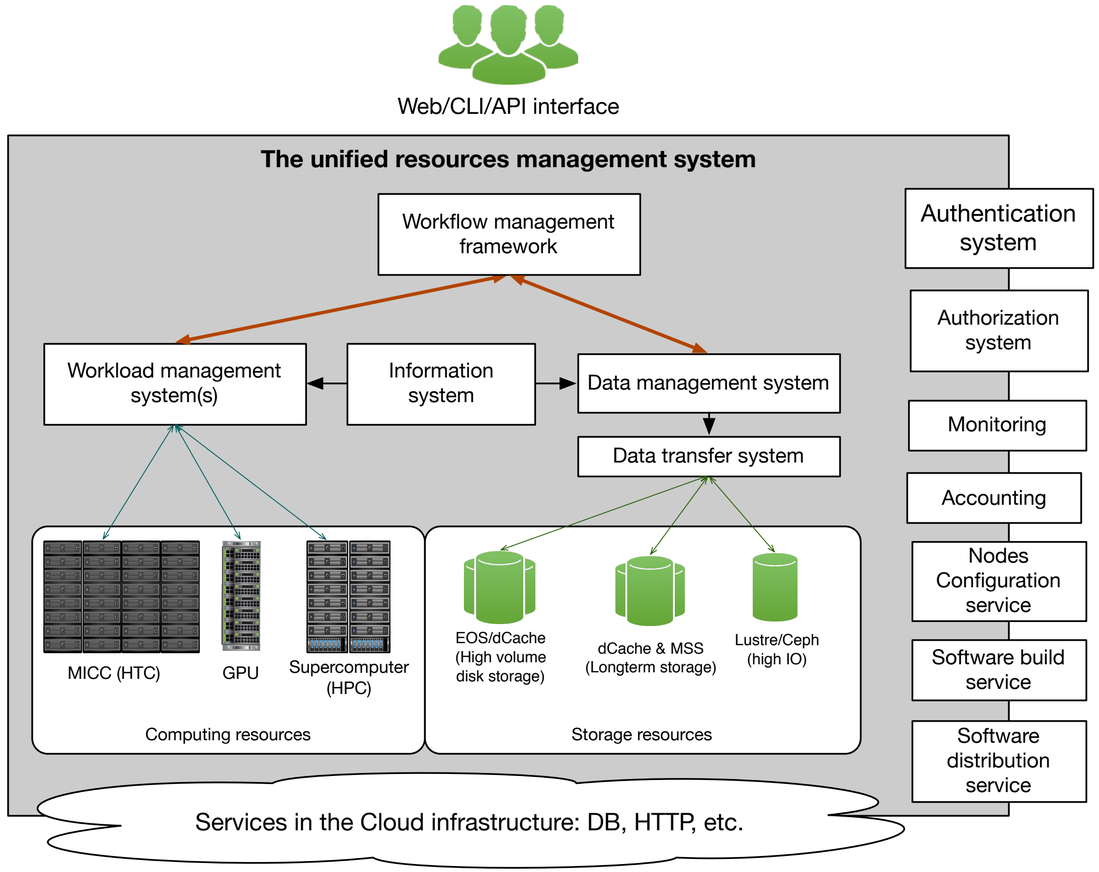
- Automation of data processing means the sequence of transformations of source data to the data in the format which is used for final analysis
- Key components required for automation:
  - **Workflow management system** - control the process of processing of data on each step of processing. Produce tasks, which required for processing of certain amount of data, manages of tasks execution.
  - **Workload management system** - processes tasks execution by the splitting of the task to the small jobs, where each job process a small amount of data. Manage the distribution of jobs across the set of computing resources. Takes care about generation of a proper number of jobs till task will not be completed (or failed)
  - **Data management system** - responsible for distribution of all data across computing facilities, managing of data (storing, replicating, deleting etc.)
  - **Data transfer service**: takes care about major data transfers. Allow asynchronous bulk data transfers.



# Unified Resource Management System

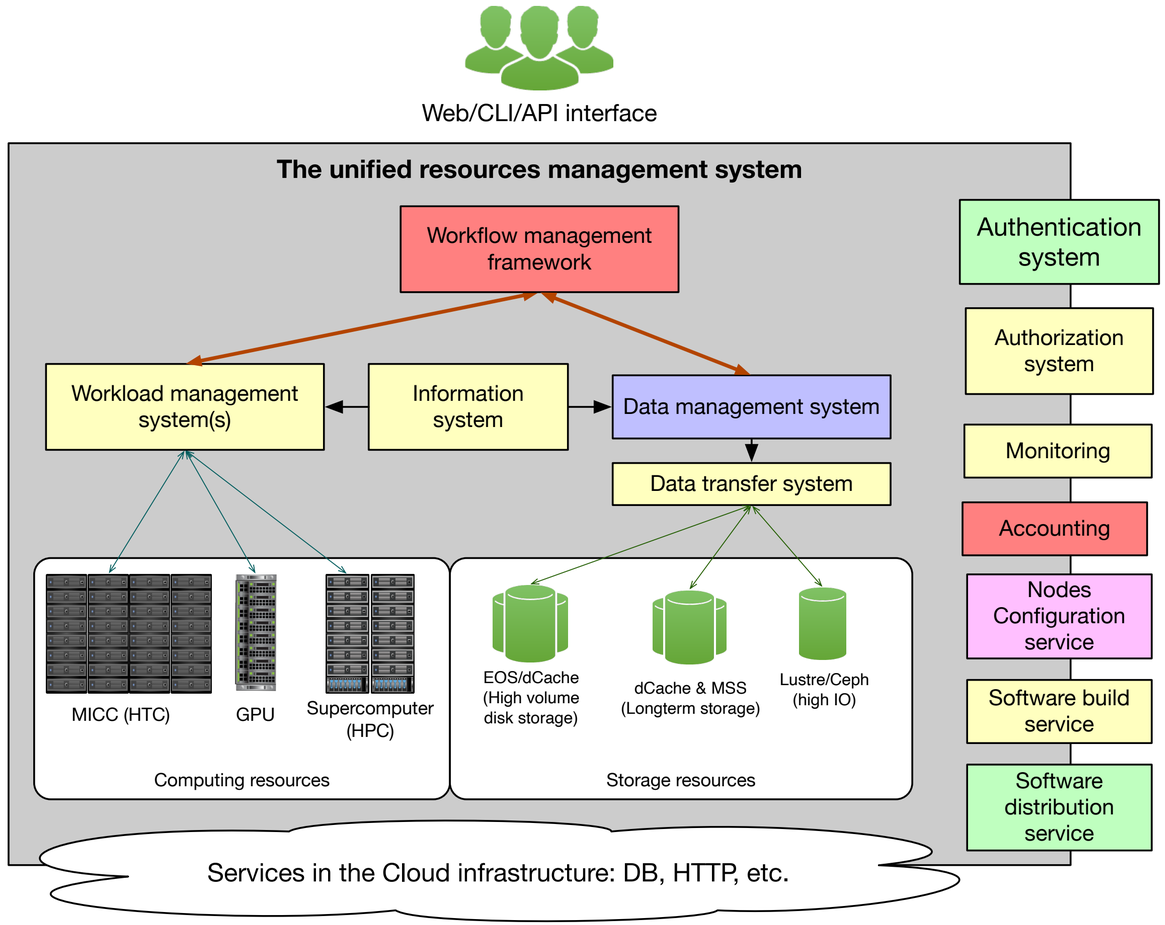
## Under development in JINR LIT

- The **Unified Resource Management System** is a IT ecosystem composed from the set of subsystem and services which should:
  - Unify of access to the data and compute resources in a heterogeneous distributed environment
  - Automate most of the operations related to massive data processing
  - Avoid duplication of basic functionality, through sharing of systems across different users (if it possible)
  - As a result - reduce operational cost, increase the efficiency of usage of resources,
  - Transparent accounting of usage of resources
- Mostly based on already existed components and technologies
  - approved during LHC data processing



# Unified Resource Management System

## in progress




- Some core subsystem already exist in JINR
  - Authentication system (Kerberos based, with SSO supporting for Web applications)
  - CVMFS as Software distribution service
- In progress:
  - integration of FTS as the core of Data transfer system
  - CRIC based Information system already deployed, but a lot of integration work expected
  - We already have some infrastructure monitoring
  - A lot of research in WFMS and WMS fields, we may declare a list of requirements:
    - We should avoid limitations by scale as much as possible.
    - Advanced monitoring system
    - WMS with MultiVO support
    - Priority and share management
    - Task-based job management
  - Looks like that Rucio will be natural choice  as cross experiment Data Management System
  - SSO + VOMS as authorization system
  - Software build service: a prototype already existed, but not fully unified across experiments
- Work to be done:
  - Accounting - we have nothing for the moment. It causes problems when we try to analyse our facilities.
  - Nodes configuration - does not look like a service for the moment. Homemade scripts, and human-oriented instructions