

# **Data-intensive scientific computing in National Research Center "Kurchatov Institute"**

V.E. Velikhov

Dubna, 2023.

National Research Center "Kurchatov Institute" unite a significant part of the country's scientific potential in the field of nuclear physics, energy, materials science, information technology, biology and genetics, microelectronics:

- NRC "Kurchatov Institute", Moscow
- NRC "Kurchatov Institute" - PNPI (St. Petersburg Institute of Nuclear Physics named after B.P. Konstantinov, Gatchina)
- NRC "Kurchatov Institute" - IHEP (Institute of High Energy Physics named after A.A.Logunov, Protvino)
- NRC "Kurchatov Institute" - CRISM «Prometey» (Central Research Institute of Structural Materials "Prometheus" named after I.V. Gorynin, St. Petersburg)
- NRC "Kurchatov Institute" - VIAM (All-Russian Research Institute of Aviation Materials, Moscow)
- NRC "Kurchatov Institute" - NII MP (Research Institute of Medical Primatology, Sochi)
- NRC "Kurchatov Institute" - RSIARAE (All-Russian Research Institute of Radiology and Agroecology, Obninsk)
- NRC "Kurchatov Institute" - FTIAN (Physico-Technological Institute named after K.A. Valiev, Moscow)
- NRC "Kurchatov Institute" - IDPM (Institute of Design Problems in Microelectronics, Zelenograd, Moscow)
- NRC "Kurchatov Institute" - ISVCH (Institute of Microwave Semiconductor Electronics named after V.G. Mokerov, Moscow)
- NRC "Kurchatov Institute" - NIISI (Federal Research Center "Scientific Research Institute for System Research", Moscow)
- NRC "Kurchatov Institute" - KiF (Federal Research Center "Crystallography and Photonics", Moscow)
- NRC "Kurchatov Institute" - ISC (Institute of Chemistry of Silicates named after I.V. Grebenshchikov, St. Petersburg)
- NRC "Kurchatov Institute" - IVS (Institute of Macromolecular Compounds, St. Petersburg)





NRC "Kurchatov Institute" is the leading scientific organisation in Integrated program "Development of engineering, technology and scientific research in the field of the use of atomic energy in the Russian Federation for the period up to 2030.

For all projects of NPPs with VVER and RBMK reactors, NRC "Kurchatov Institute" is the organisation authorised to carry out scientific supervision in the development of NPP and reactor plant designs for a specific site or basic NPP design.

Programs for mathematical modelling and precision calculations of reactors are being created. Neutron-physics, thermal physics, thermal-hydraulic, radiology, mechanics and other computing programs formed the basis for the safety justification of almost all reactor installations.

Industrial Detector for Reactor Antineutrino Monitoring (iDREAM) is dedicated for remote monitoring of PWR reactor operational modes by neutrino method in real-time.

NRC "Kurchatov Institute" has united the efforts of organisations to develop new types of steel for nuclear reactor vessels. Materials have been obtained that increase the power of reactors and their service life up to 100 years.





NRC "Kurchatov Institute" is the leading scientific organisation in several Federal programs:

- Federal Scientific and Technical Program for the Development of Synchrotron and Neutron Research and Research Infrastructure for 2019-2027;
- Federal Scientific and Technical Program for the Development of Genetic Technologies for 2019-2030;
- Federal scientific and technical program for the development of agriculture for 2017-2030.

NRC "Kurchatov Institute" participates, and in a number of projects provides scientific guidance on behalf of the Government, in International projects:

- International Thermonuclear Experimental Reactor ITER;
- European X-ray Free Electron Laser XFEL;
- Large Hadron Collider at CERN LHC;
- European Center for Synchrotron Radiation ESRF;
- FAIR European Center for Ion and Antiproton Research FAIR.

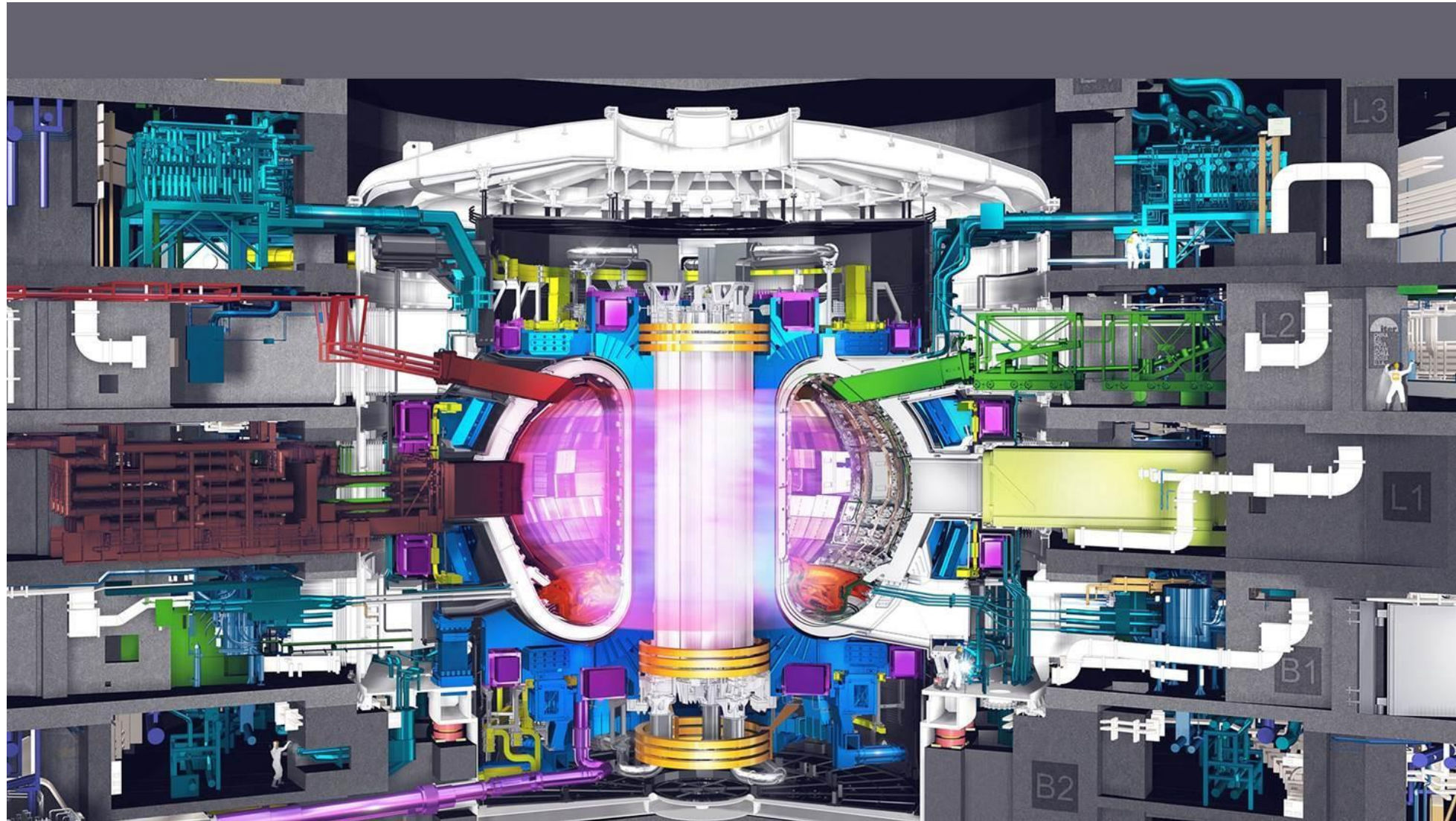
International experiment BOREXINO.

Russian-Italian project of an experimental thermonuclear reactor IGNITOR.

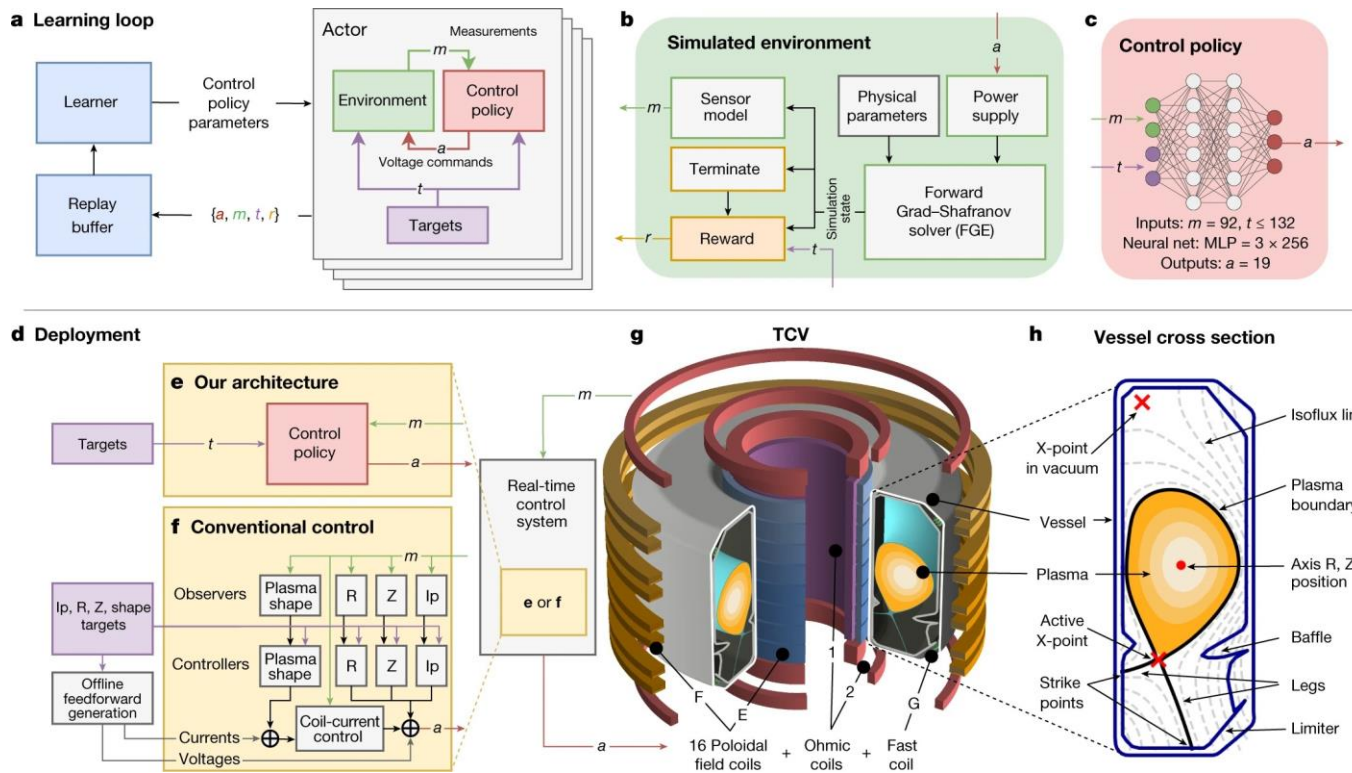
The PIK complex will become the basis of the world-class International Center for Neutron Research.

The NRC "Kurchatov Institute" is participating in an international project to create a Russian proton and heavy ion collider NICA in Joint Institute for Nuclear Research.





|              |                                           |
|--------------|-------------------------------------------|
| Astrophysics | Radiotelescopes<br>Neutrino detectors     |
| HEP          | Colliders                                 |
| Fusion       | TOKAMAKs                                  |
| SNR          | Synchrotrones<br>FELs<br>Neutron reactors |

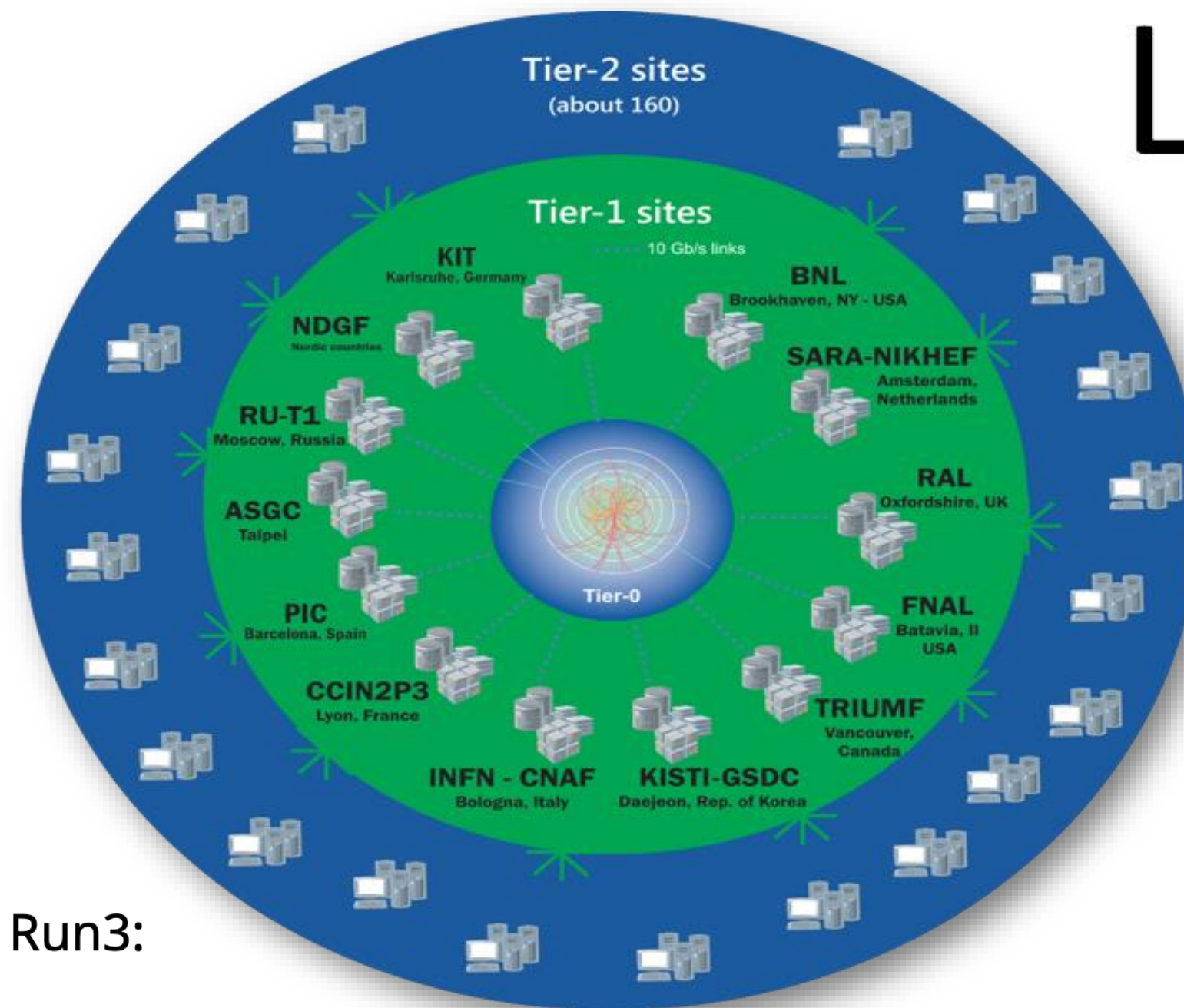


Physical models

Surogate models



# LHCOPN

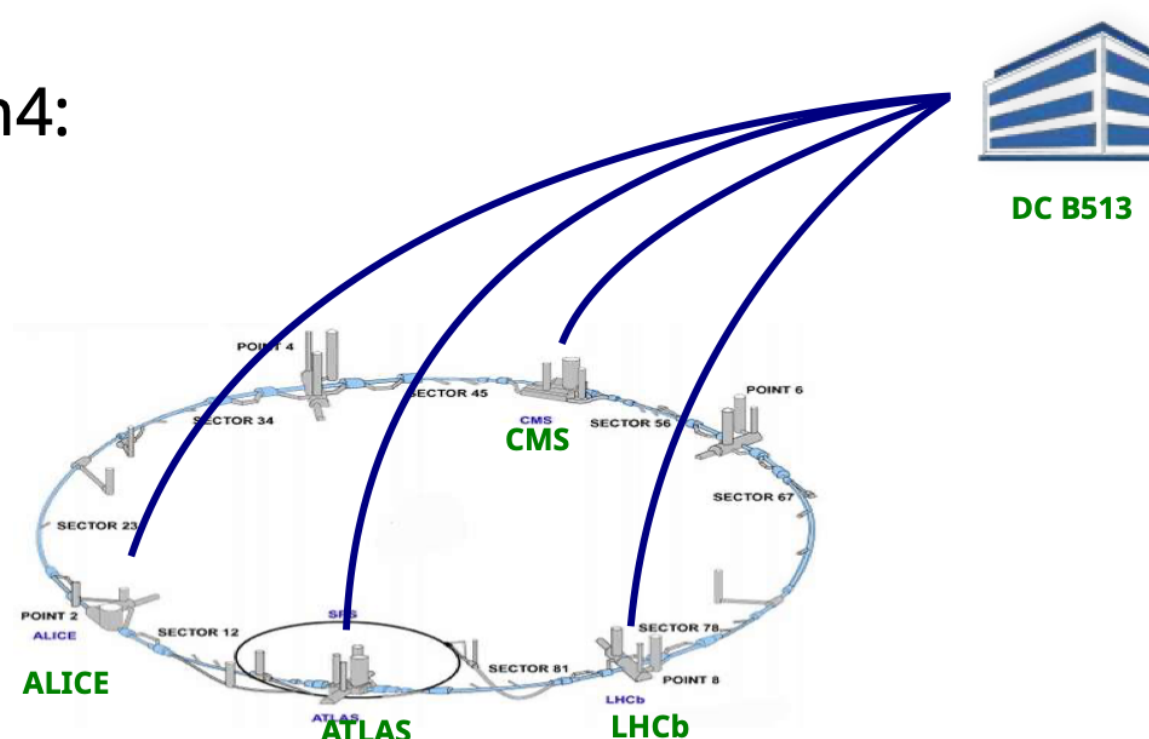
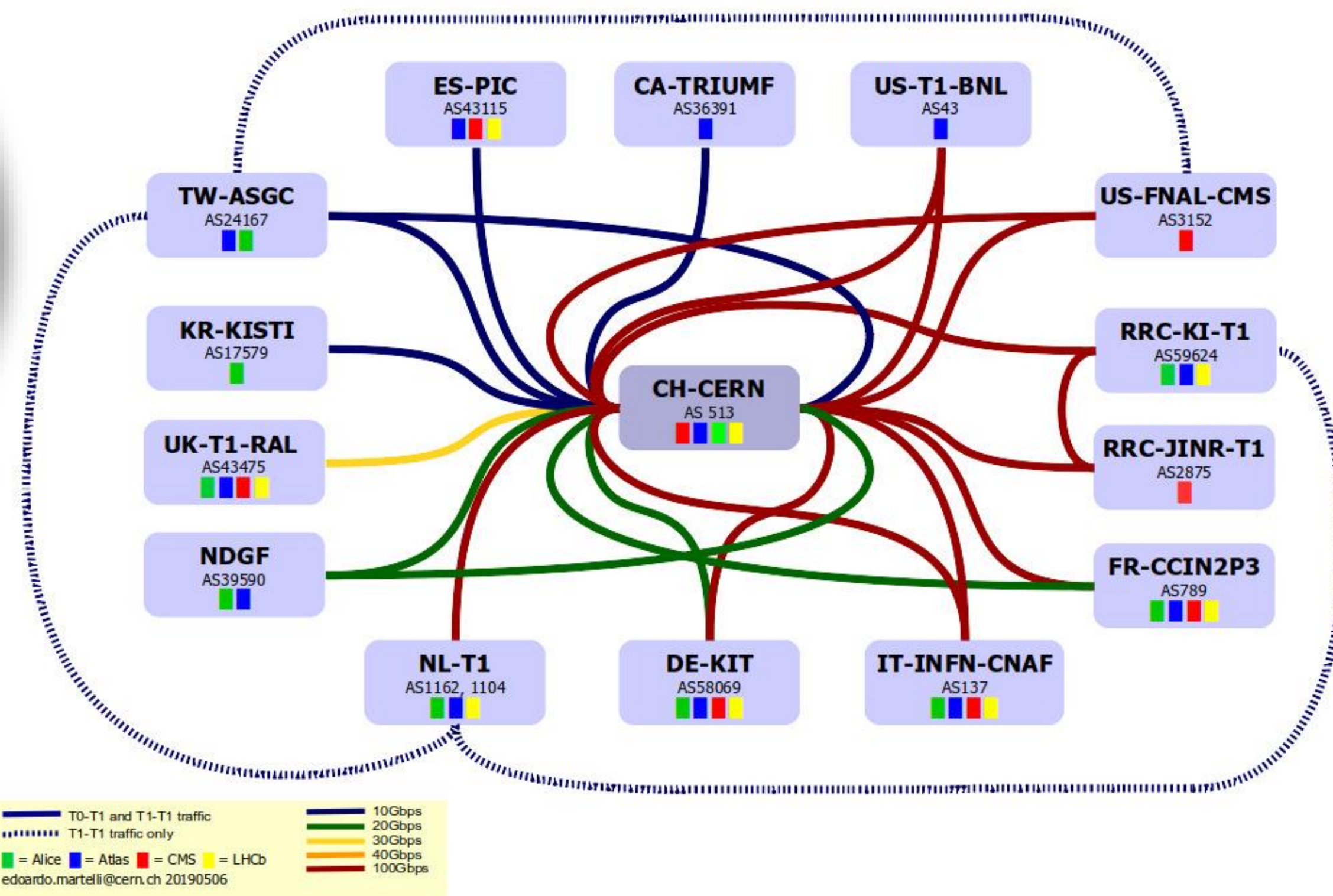


Received requirements for Run3:

- ALICE: 2Tbps
- LHCb: 1Tbps
- CMS: 400Gbps
- ATLAS: 200Gbps

Estimated requirements for Run4:

- ATLAS: 4Tbps
- CMS: 4Tbps





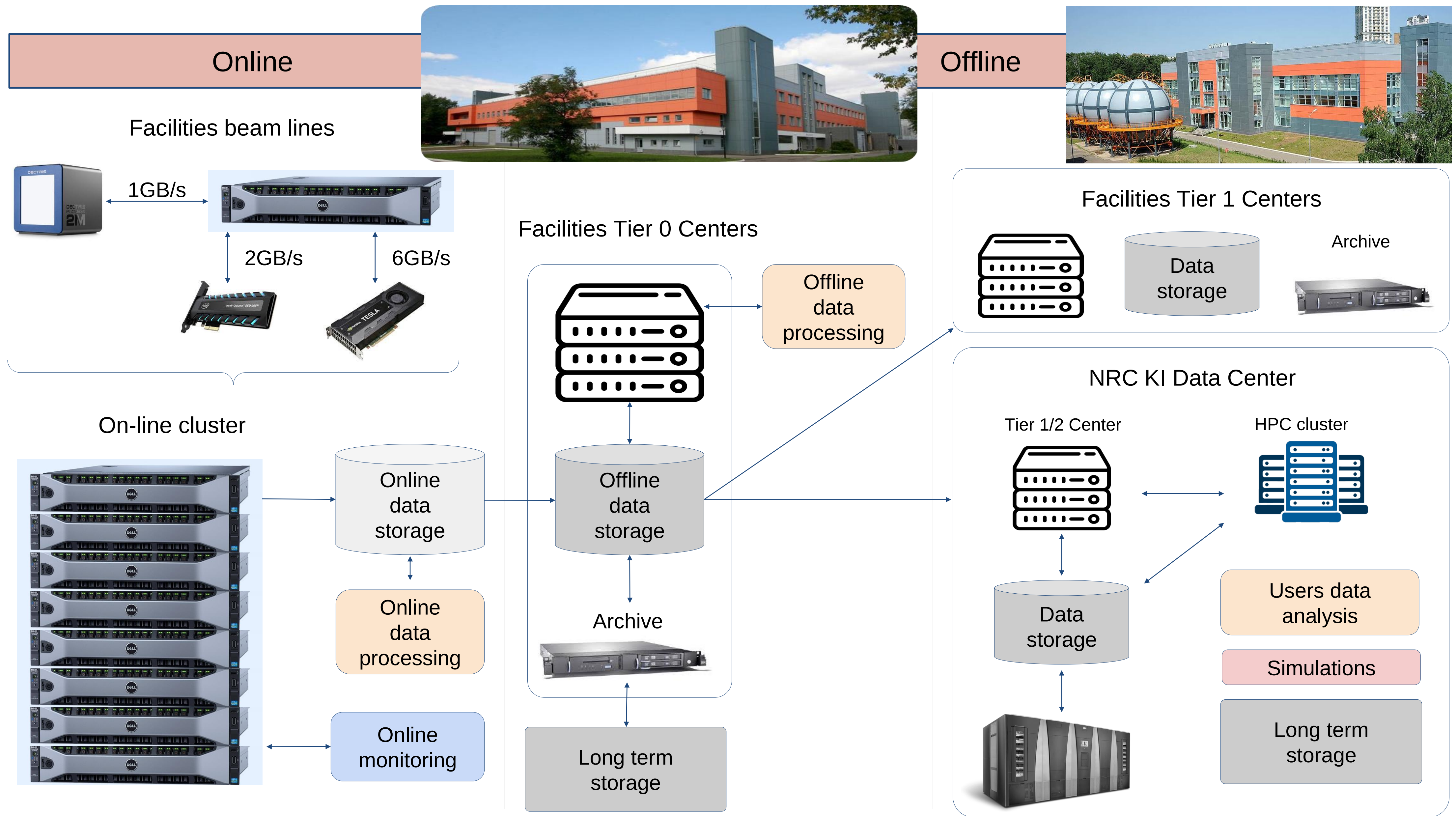
The Strategy for the Scientific and Technological Development of the Russian Federation, approved by Decree of the President of the Russian Federation, defines the state policy of a priority nature to support the creation and development of unique scientific installations of the class “mega science” as a part of large research infrastructures on the territory of the Russian Federation, which requires an effective solution of research problems of the mega science class based on domestic scientific, technical and human resources.

The Federal Scientific and Technical Program for the Development of Synchrotron and Neutron Research and Research Infrastructure for 2019-2027 includes the development and creation of a unified digital platform for storing, processing and analyzing experimental data.

The most important part of these projects is the creation and development of computer systems (including systems with extramassive parallelism) for processing, storing and analysing experimental data, the development and implementation of effective methods, algorithms and software for modelling physical systems, search and access data, development of algorithms in the field of machine and deep learning, methods of computer algebra.







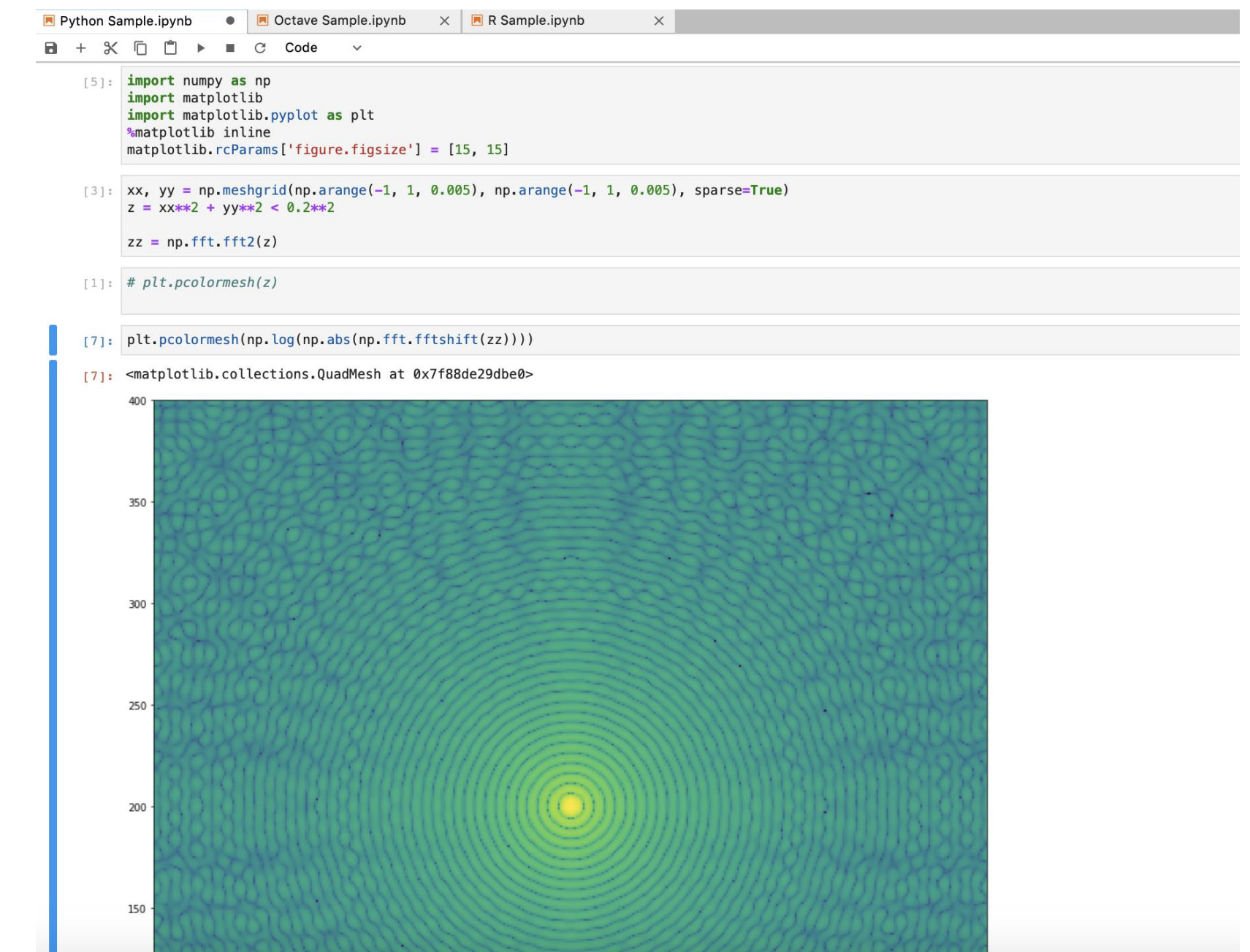


A high-performance platform has been created to automate the processing of experimental data through the use of big data processing and analysis technologies and AI methods.

The composite task for data processing includes all stages, from the primary data of the experimental setup to the reconstruction of the spatial structure of the object.

In 2020, the platform was used to reconstruct the virus (bacteriophage) PR772 with a resolution of 6.9 nm, based on an experiment at LCLS (Stanford). Work is underway on the encephalitis virus.

- Web interface for working with popular environments for programming and data analysis: R, Octave, Julia, Python and others
- Integration with 1PB storage (CEPHfs) and Kubernetes cloud (HPC4, 130 TFLOPS, Nvidia K80 GPU accelerators)
- Big data environment and machine learning (ML & AI) libraries Tensorflow, Keras, PyTorch, MXNet
- Ability to integrate a wide range of application software: Biopython and Bioconductor for bioinformatics tasks, SimNet for physical modeling
- Data processing of the first SPI experiment on EuXFEL - filtering and classification of experimental images (Sobolev, Egor, et al. "Megahertz single-particle imaging at the European XFEL." Communications Physics 3.1 (2020): 1-11.)





Purpose: To create a global competitive infrastructure for storage and high-performance processing and analysis of genetic information.

The technical design of the IAS NBGI was completed.

A high-performance complex for processing and analyzing genomic data (model of the IAS NBGI) was created and put into experimental operation. On its basis, the genomic center is implementing new projects.

More than 50 databases are loaded into the layout, including:

Mycancergenome

NCBI

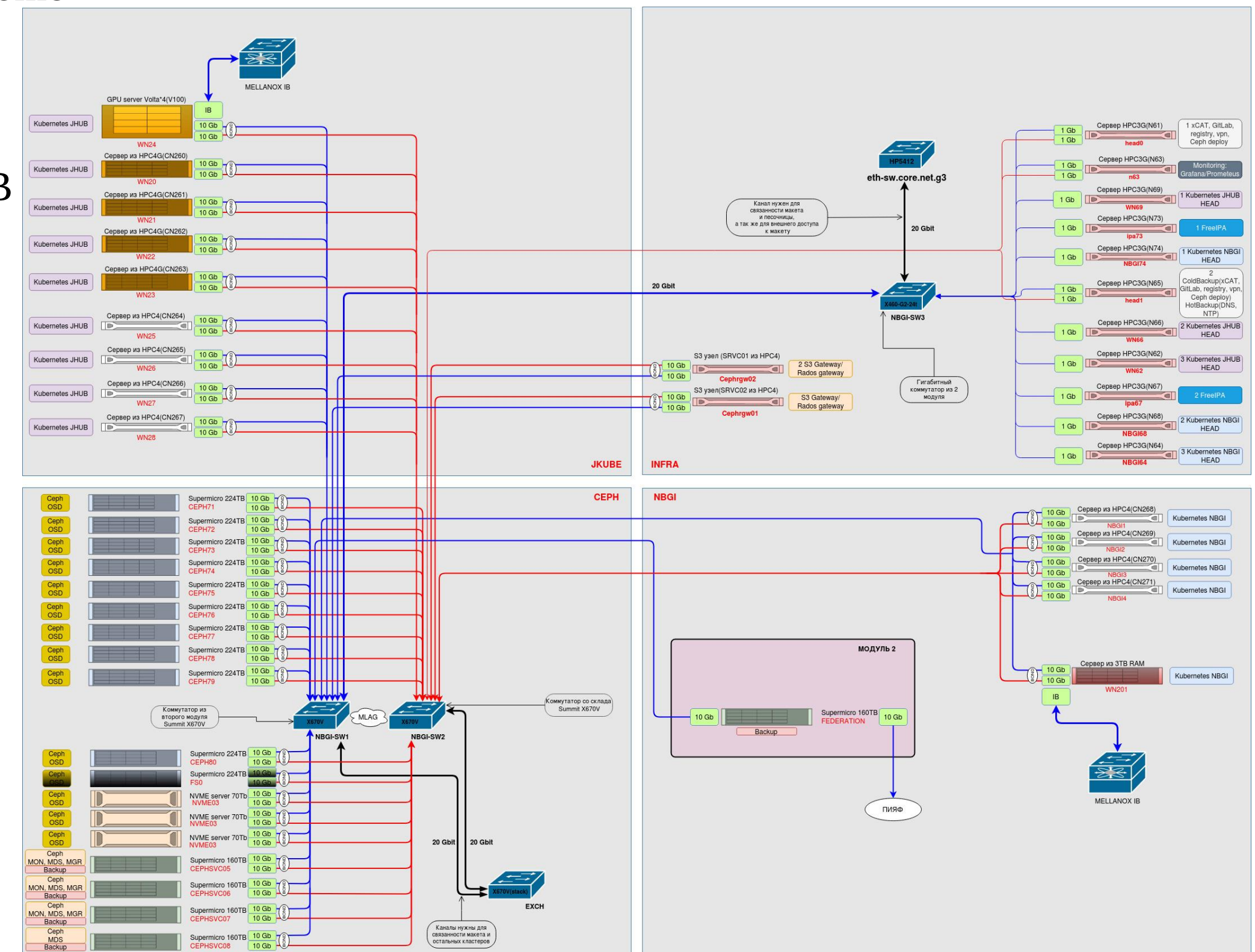
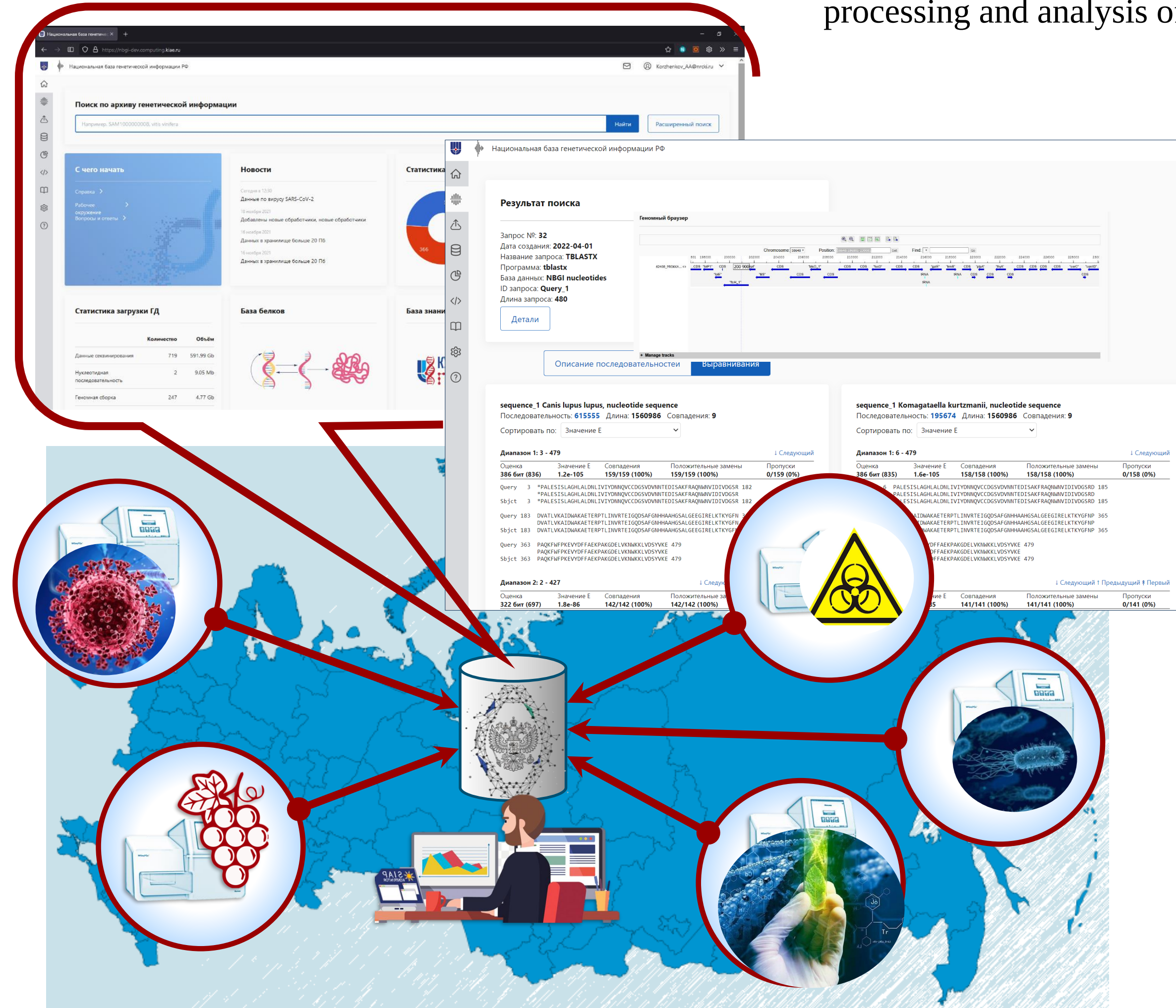
ENA

Ensembl.org

PDB + EMDB

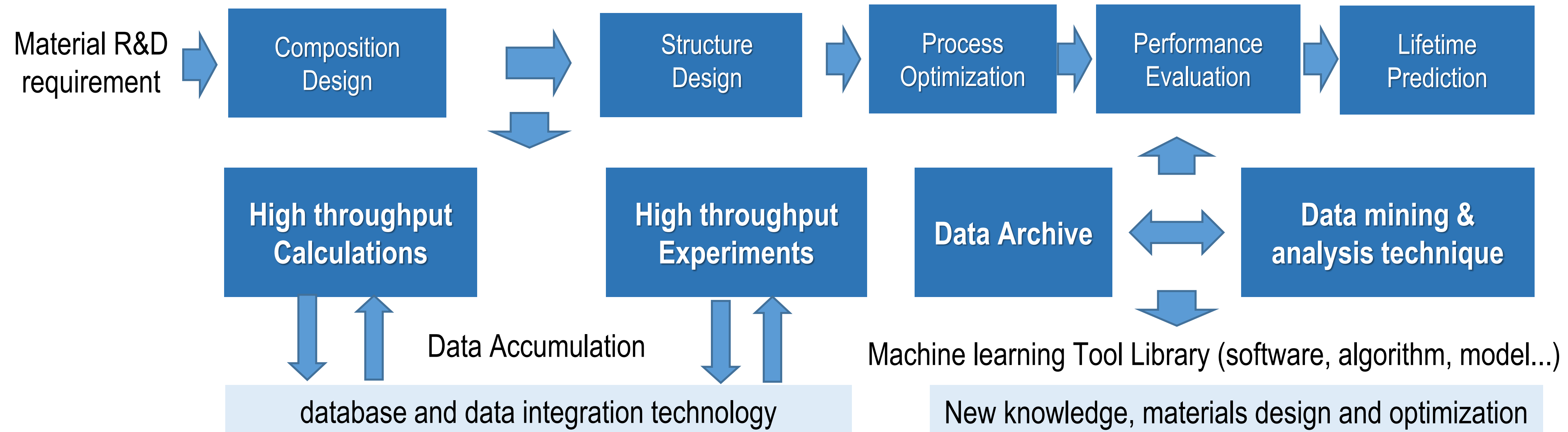
Addgene

Sasbdb.





## Next challenge - Computational materials design



High throughput calculation and experiment:

- Atomic potentials database construction
- Typical material thermodynamics and dynamics database construction
- High throughput DFT driven engine construction
- Capture and analysis of high throughput testing data

Database Construction for Material design:

- First-principles calculation database
- Composition-microstructure-property database
- Data mining and application technology

Intelligent software platform for data mining:

- 3D microstructure reconstruction technology
- Application research of materials data

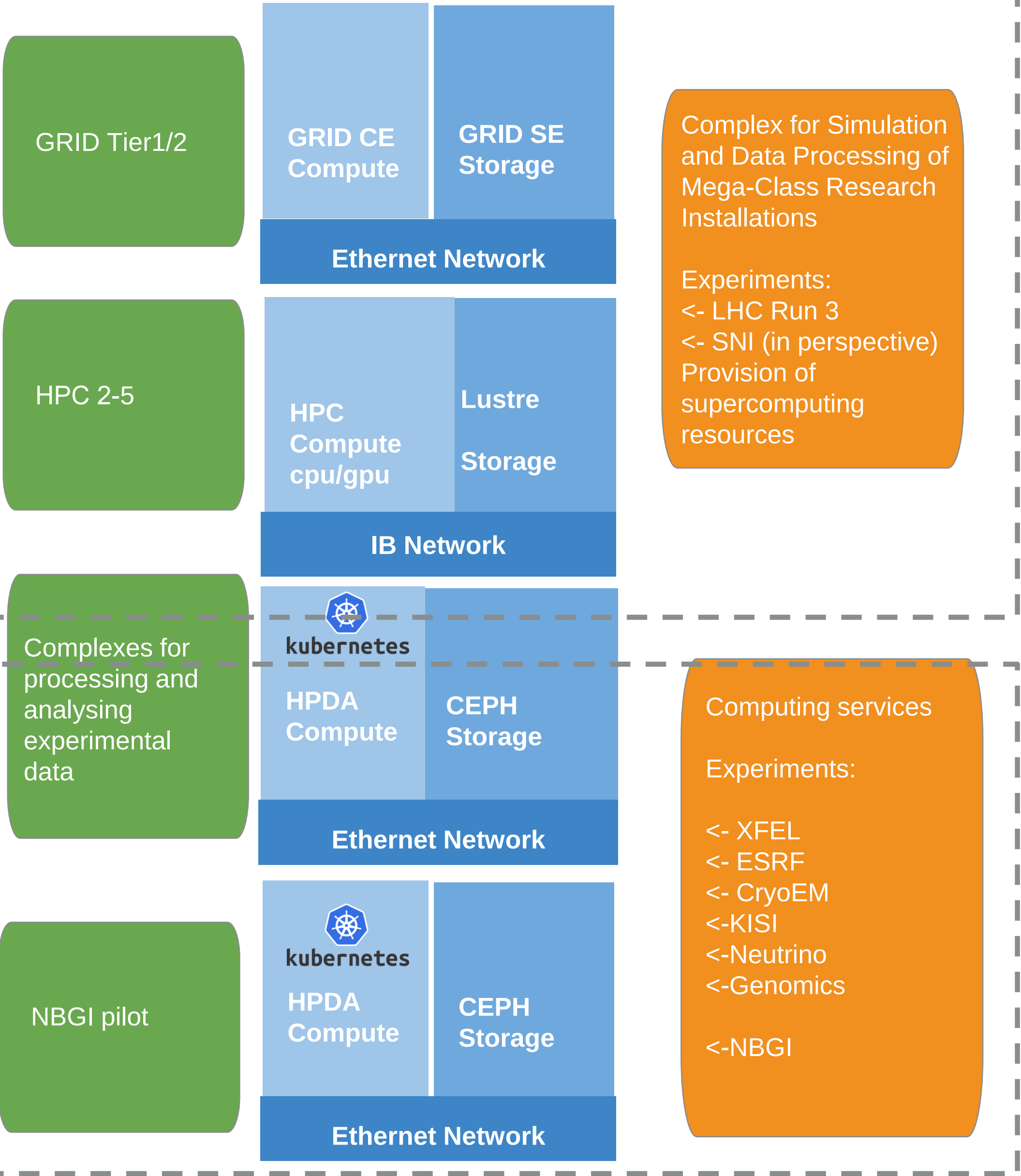
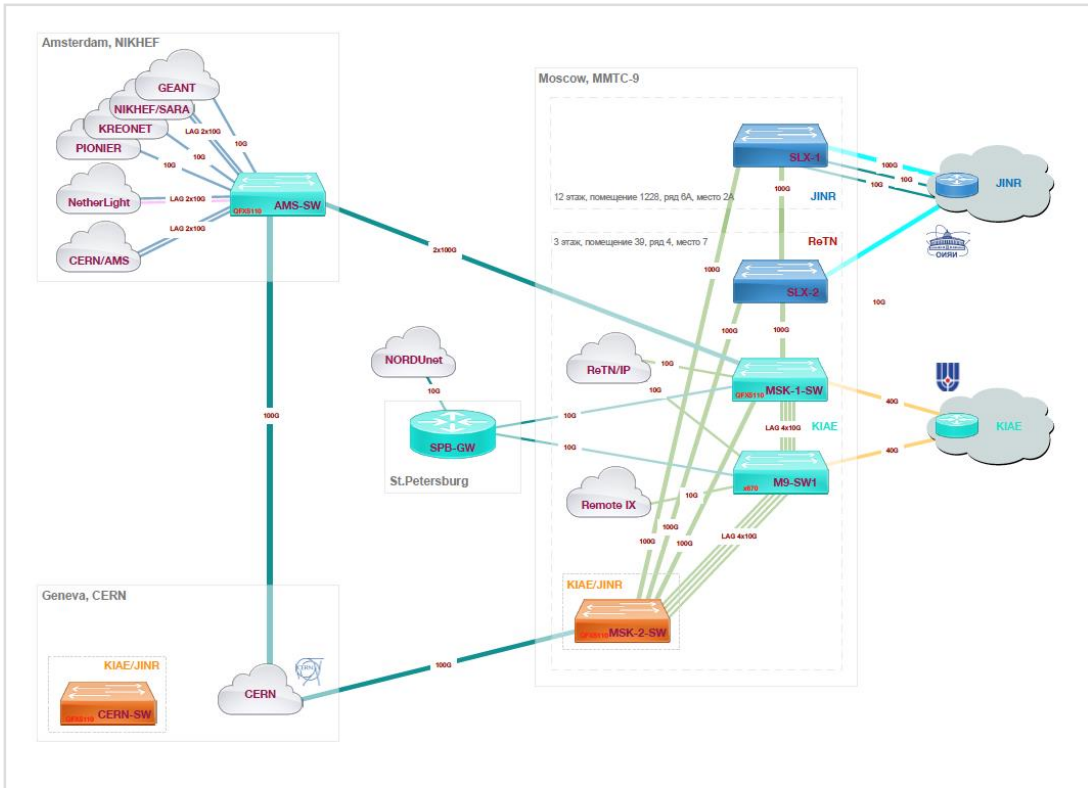
**Big data + Machine Learning : new research and development mode**



NRC KI  
Data Center



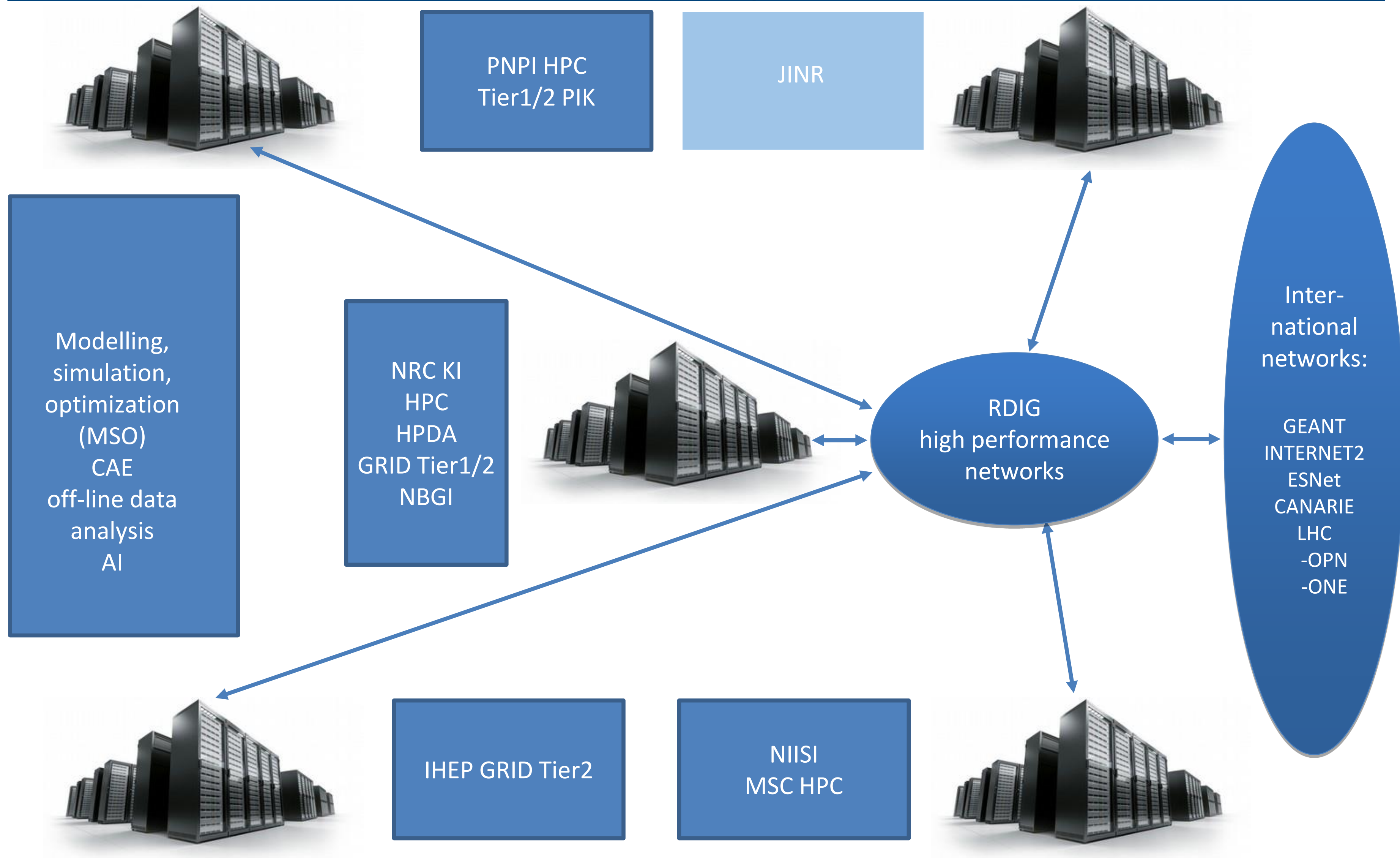
RDIG



| S&E fields                            |                                                          |
|---------------------------------------|----------------------------------------------------------|
| HEP, Astrophysics                     |                                                          |
| SNR                                   |                                                          |
| Computational materials science       |                                                          |
| Computational biology & genomics      |                                                          |
| Nuclear power plants                  |                                                          |
| Nuclear medicine                      |                                                          |
| Microelectronics simulations & design |                                                          |
| Geophysics                            |                                                          |
| Computing services                    | SW                                                       |
| Neutron physics                       | MCU                                                      |
| MSO & CAE                             | Orange Synchrotron Suite Ansys OCTAVE                    |
| CFD                                   | Star CCM+ FlowVision Ansys OpenFOAM                      |
| Data processing                       | Apache Spark Kafka Flume Hadoop R Ceph Lustre dCache EOS |
| ML frameworks                         | TensorFlow PyTorch Simnet Deepspeed                      |



## Scientific computing infrastructure in NRC “Kurchatov Institute”





# Russian Grid for Data Intensive Operations

NRC "Kurchatov Institute" operates high-performance computing complexes located in Moscow, Protvino, and Gatchina, including the “Complex for Simulation and Data Processing of Mega-Class Research Installations”.

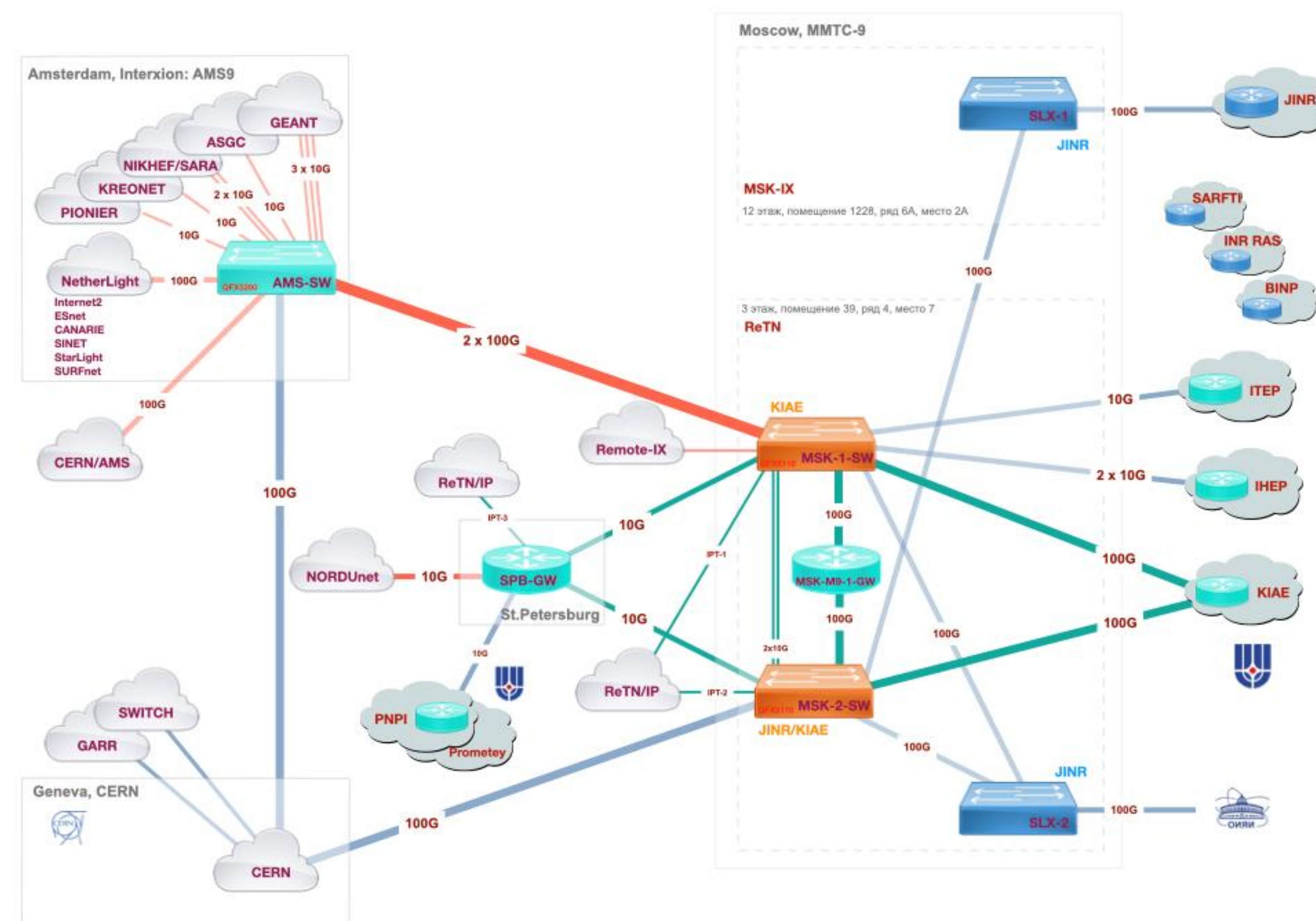
Computing systems are connected to the national scientific grid infrastructure “Russian Grid for Data Intensive Operations” (RDIG) and integrated into the European Grid infrastructure EGI and the global grid infrastructure of the Large Hadron Collider WLCG with high-speed (100 Gbit/s) data transmission channels.

JINR, MEPhI, SarFTI, INR RAS, INP SB RAS, St. Petersburg State University are also connected to this scientific GRID infrastructure. Interaction with other scientific and educational organizations is carried out through the MMTS-9 communication center using networks for science and education NIKS.

The National Certification Center RDIG at the NRC "Kurchatov Institute" provides authentication of Russian researchers in international projects.

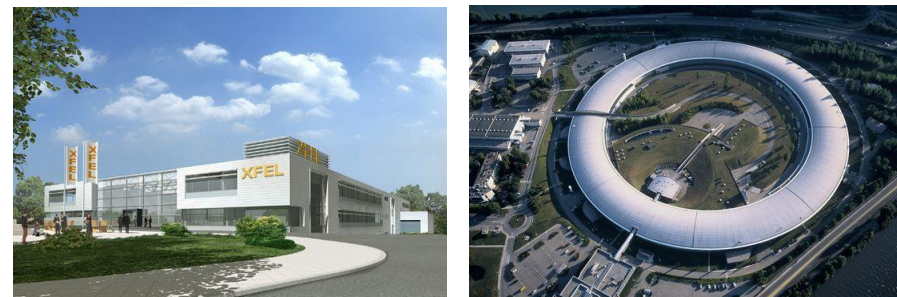
The functioning of this infrastructure requires the widespread use of supercomputer, network, GRID and cloud technologies, as well as technologies for federated storage, intelligent processing and analysis of super-large arrays and streams of scientific data, search algorithms and access to them.

Russian Data-Intensive Grid (RDIG) :: Магистральная инфраструктура





EuXFEL



ESRF



TNK

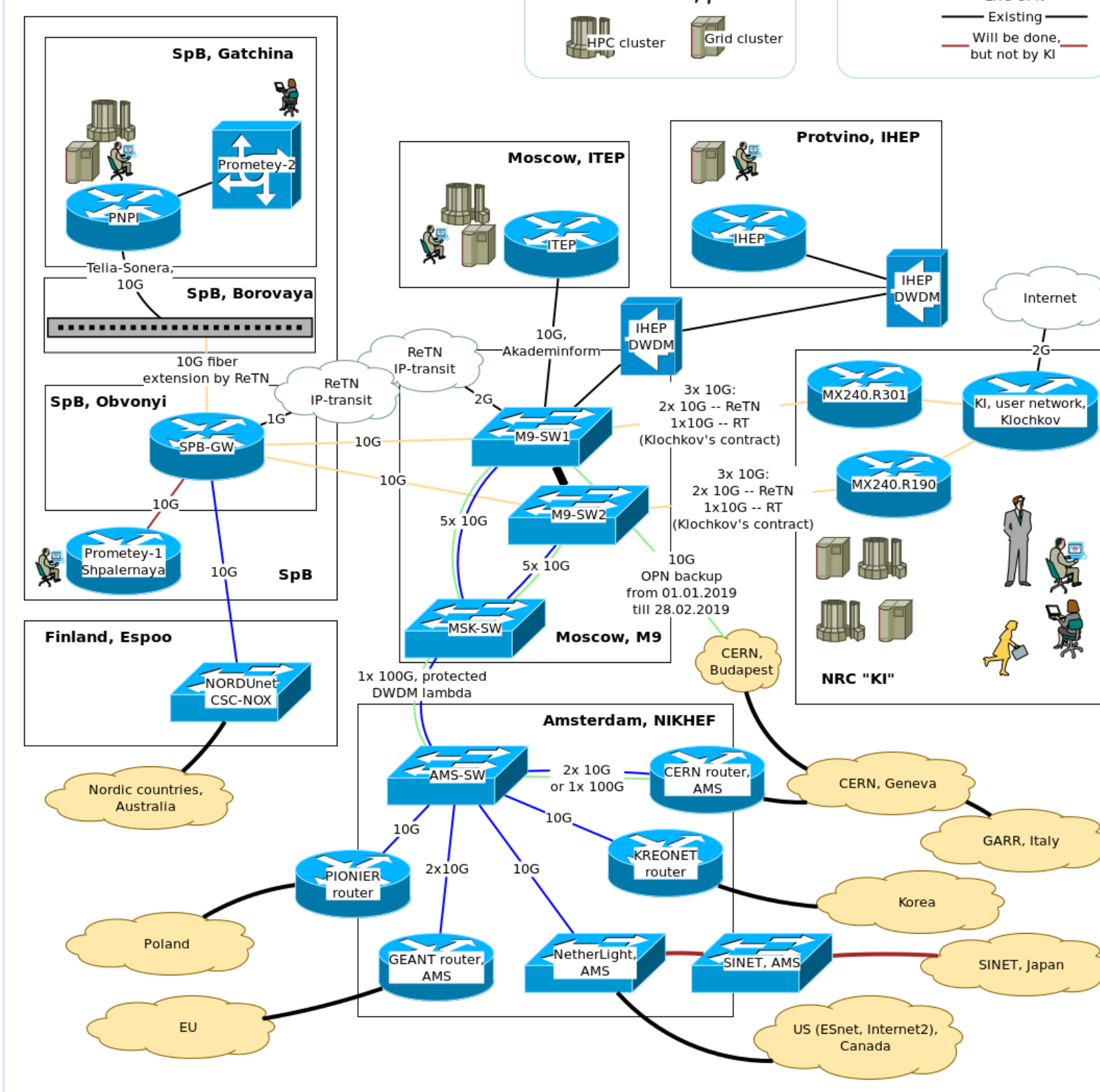


RIF



SKIF

KI WAN infrastructure, 2019, extended view,  
minimal variant.  
Version 2.3



NRC KI



OMEGA



KISI  
Kurchatov



PIK



SILA





## FAIR

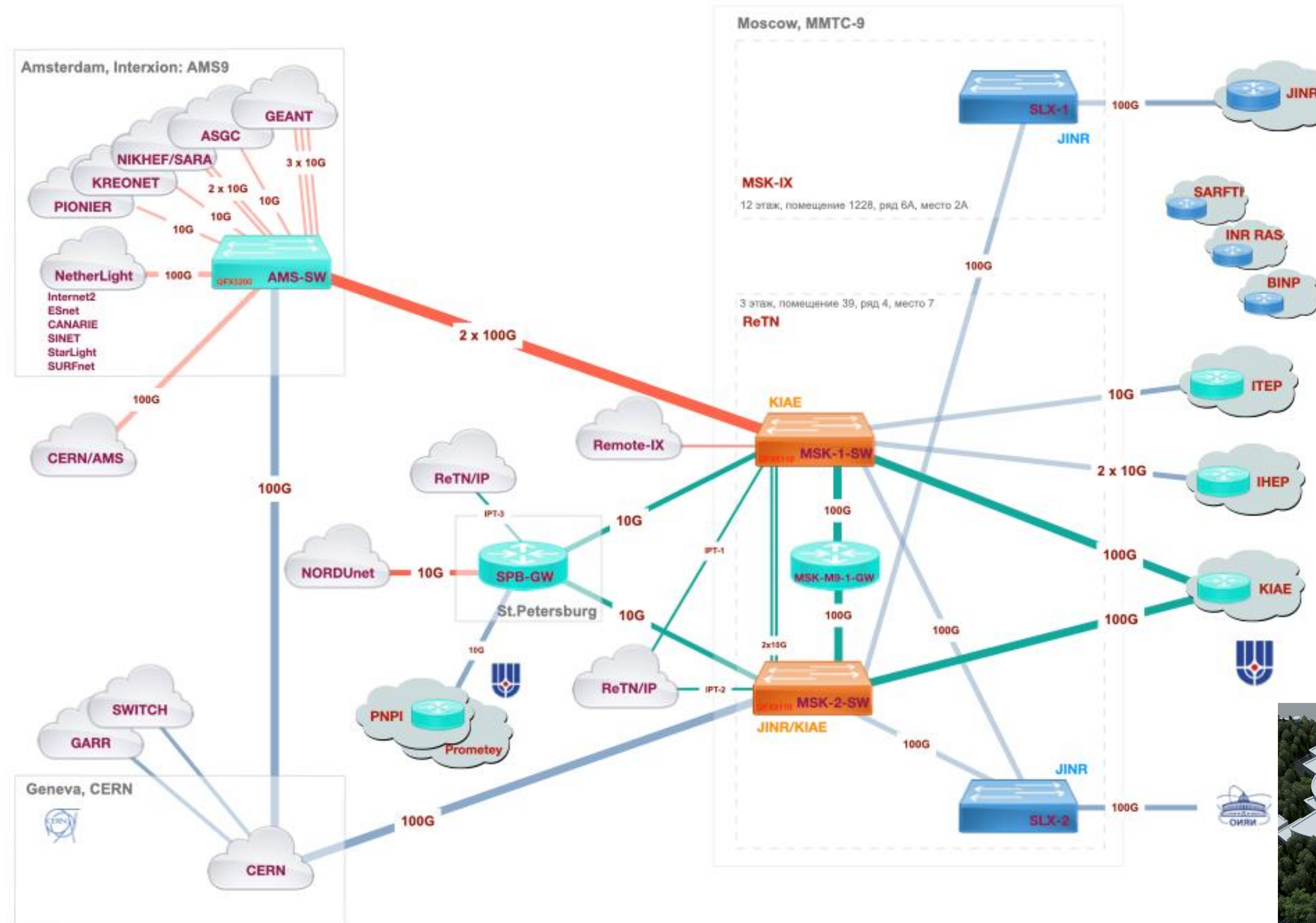


AMS-IX - 200 Gbps  
CERN - 100 Gbps  
GEANT - 30 Gbps  
NORDUnet - 10 Gbps

## LHC



Russian Data-Intensive Grid (RDIG) :: Магистральная инфраструктура



Tier 1/2 JINR



Tier 1/2 NRC KI

## NICA





Any questions?



Благодарю за внимание!