INTERNATIONAL INTERGOVERNMENTAL ORGANIZATION
МЕЖДУНАРОДНАЯ МЕЖПРАВИТЕЛЬСТВЕННАЯ ОРГАНИЗАЦИЯ

JOINT INSTITUTE FOR NUCLEAR RESEARCH
ОБЪЕДИНЕННЫЙ ИНСТИТУТ ЯДЕРНЫХ ИССЛЕДОВАНИЙ

IT and LIT–strategy of development

Korenkov Vladimir

Director of the Laboratory of Information Technologies, JINR

30 September, 2019, NEC-2019, Montenegro

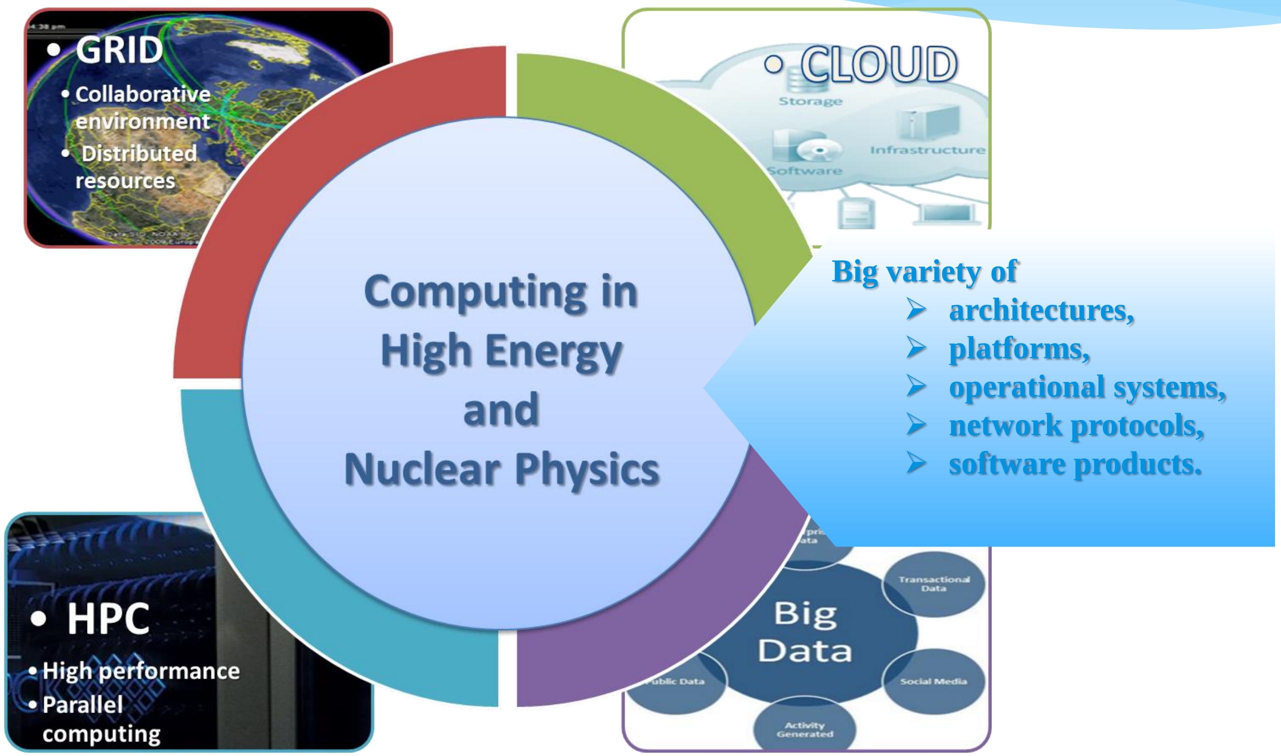
Grid technologies - a way to success

On a festivity dedicated to receiving the Nobel Prize for discovery of Higgs boson, CERN Director professor Rolf Dieter Heuer directly called the grid-technologies one of three pillars of success (alongside with the LHC accelerator and physical installations).

Without implementation of the grid-infrastructure on LHC it would be impossible to process and store enormous data coming from the collider and therefore to make discoveries.

Nowadays, every large-scale project will fail without using a distributed infrastructure for data processing.

Computing in High Energy and Nuclear Physics



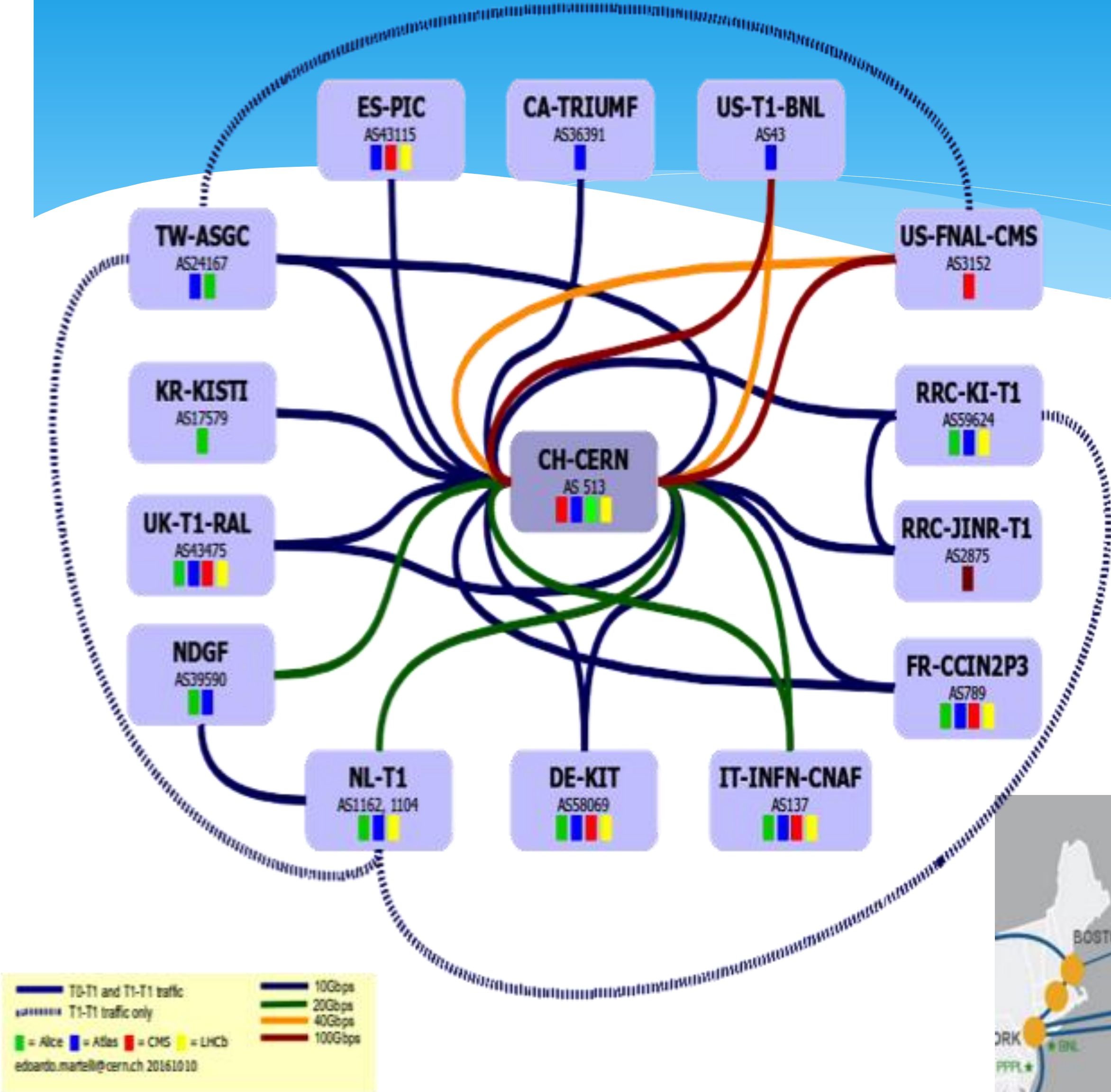
The interface of the uniform environment should provide a way for organization of collective development, solution of problems of various complexity and subject matter, management and processing of data of various volumes and structures, training and organization of scientific and research processes.

Networks

- HEP network data transfer volume continues to rise at roughly a factor of 10 every 4 years. Networking use will increase as we move toward Run3 and HL-LHC
- When discussing networking, the first parameter usually discussed is the bandwidth (capacity), either in the LAN or to the WAN
- WLCG sites usually have a pretty good handle on the LAN capacity required to allow their compute and storage to operate smoothly, but the WAN capacity is less well understood
  - For LHC Run2 we targeted 10Gbps as a minimum for Tier-2s. Large sites (Tier-1 and Tier-2) were encouraged to have multiple 40Gbps to 100Gbps connectivity
- For HL-LHC the communication speed should be more than 1 Tbps (LAN&WAN)

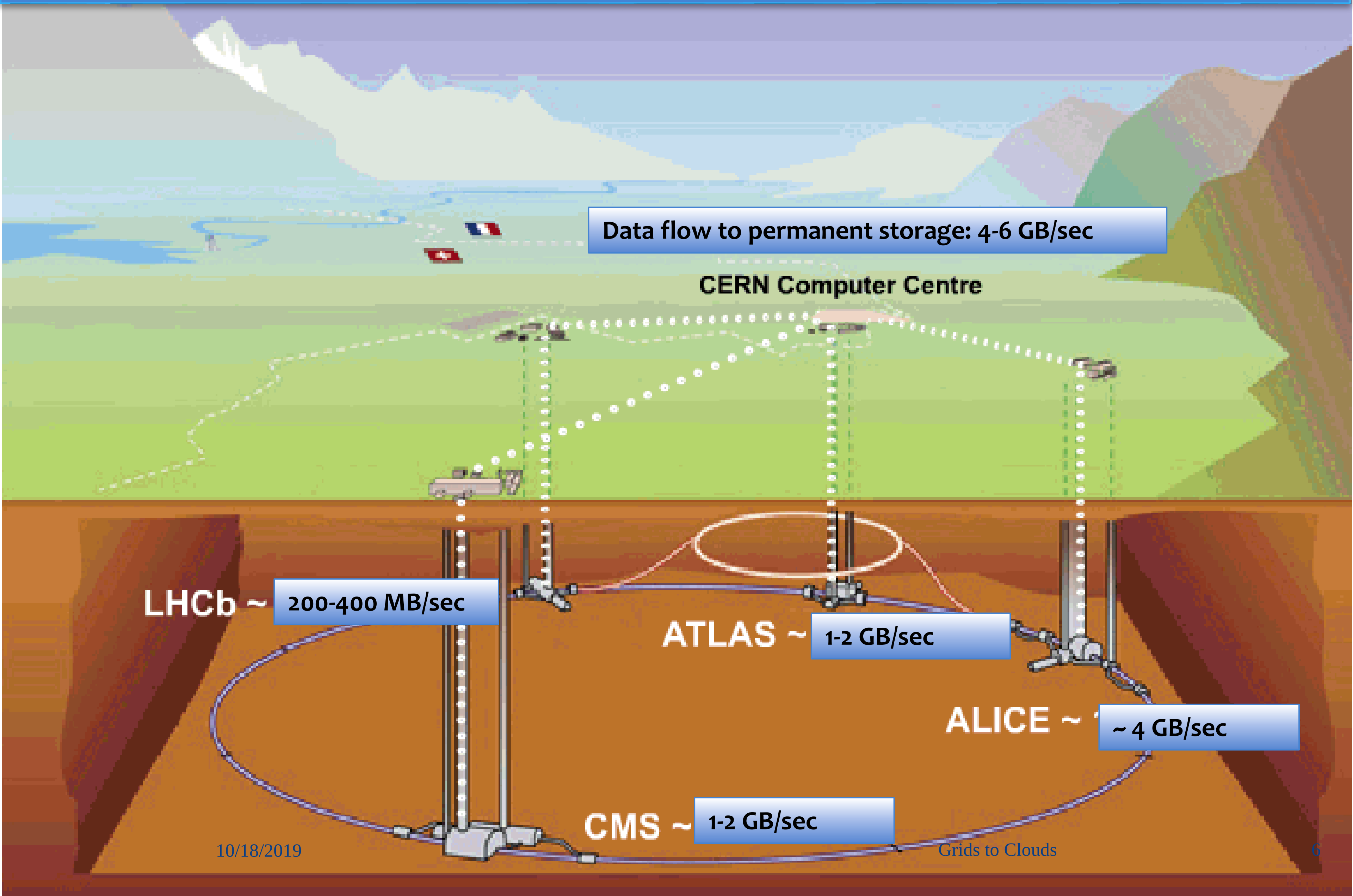
Networks

- Optical Private Network
- Support T0 – T1 transfers & T1 – T1 traffic
- Managed by LHC Tier 0 and Tier 1 sites



Up to 340 Gbps transatlantic

Data Collection and Archiving at CERN



JINR is a part of Worldwide LHC Computing Grid

WLCG:

An international collaboration to distribute and analyse LHC data integrates computer centres worldwide, which provide computing and storage resources, into a single infrastructure accessible by all LHC physicists.

Tier-0 (CERN):
data recording, reconstruction and distribution
Tier-1:
permanent storage, re-processing, analysis
Tier-2:
simulation, end-user analysis
~170 sites
42 countries
1 000 000 cores
1 EB of storage
> 3 million jobs/day
10-100 Gb links

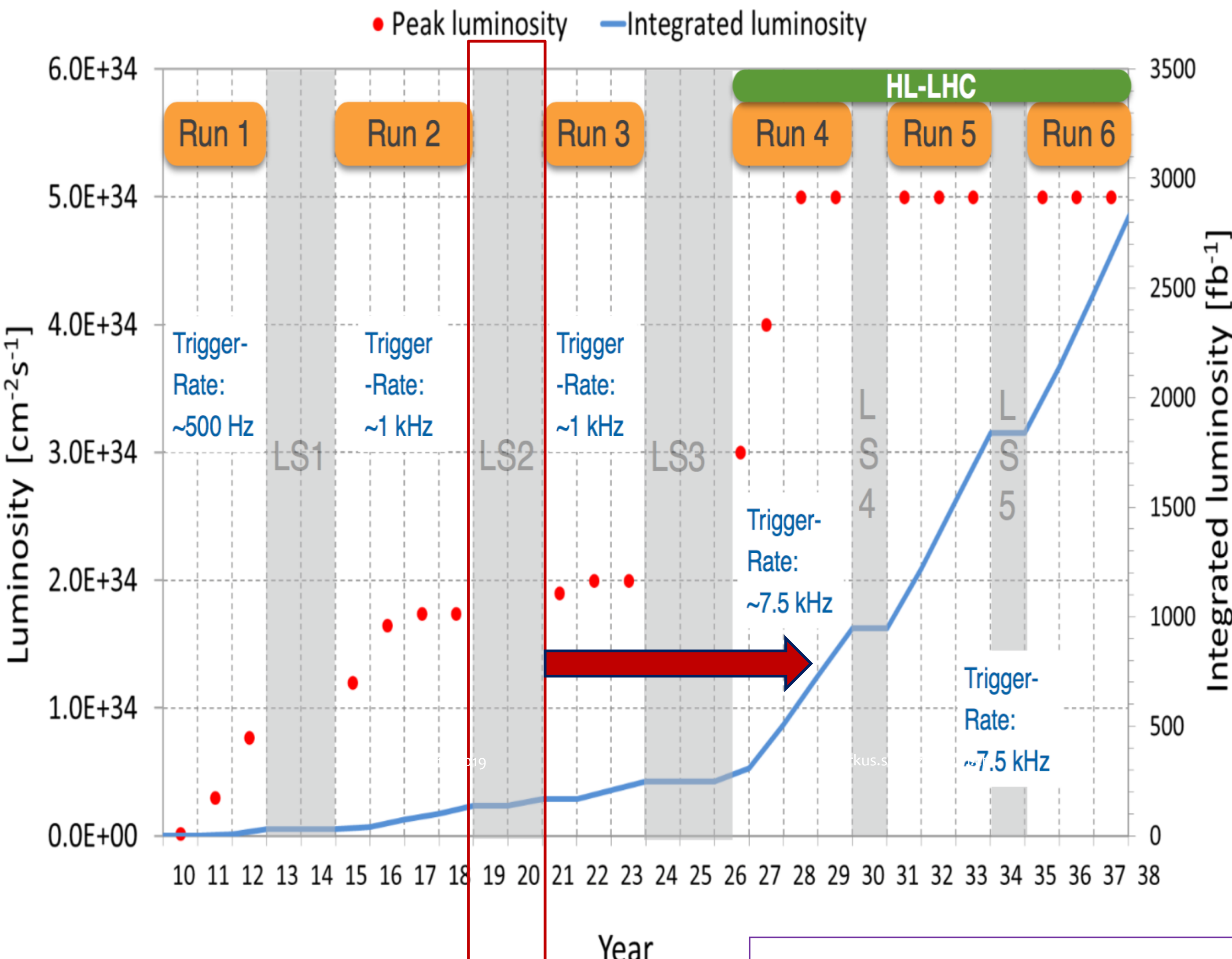
Worldwide LHC Computing Grid - 2019

Evolution of the computing models

- Must reduce the (distributed) provisioning layer of compute to something simple, we need a hybrid and be able to use:
  - Our-own resources
  - Commercial resources
  - Opportunistic use of clouds, grids, HPC, volunteer resources, etc.
- Move towards simpler site management
- Recognizing the need to re-engineer HEP software

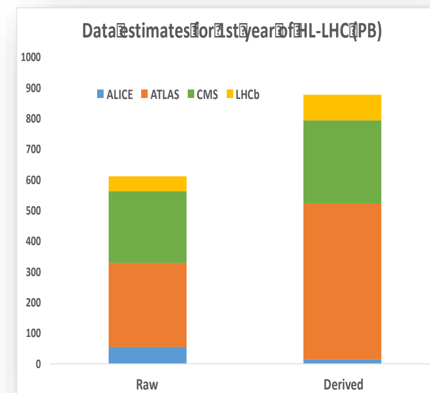
  - New architectures, parallelism everywhere, vectorisation, data structures, etc.
- Set up HEP Software Foundation (HSF)
- Optimization of the data processing and analysis
- The development of data management tools (Federation data storage, intelligent algorithms for data distribution, “data popularity”

LHC Schedule



Run 3 — ALICE, LHCb upgrades

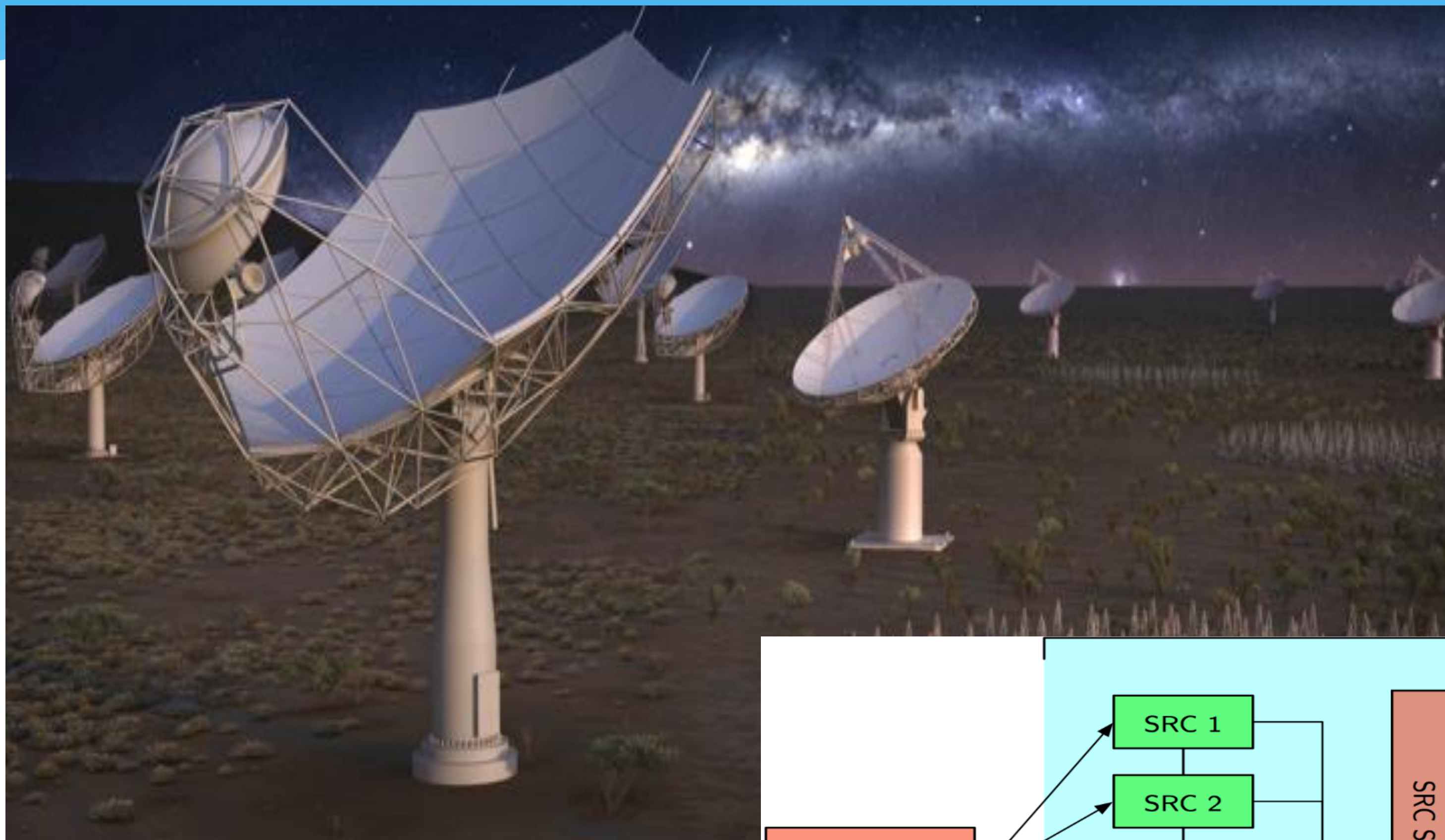
Run 4 — ATLAS, CMS upgrades

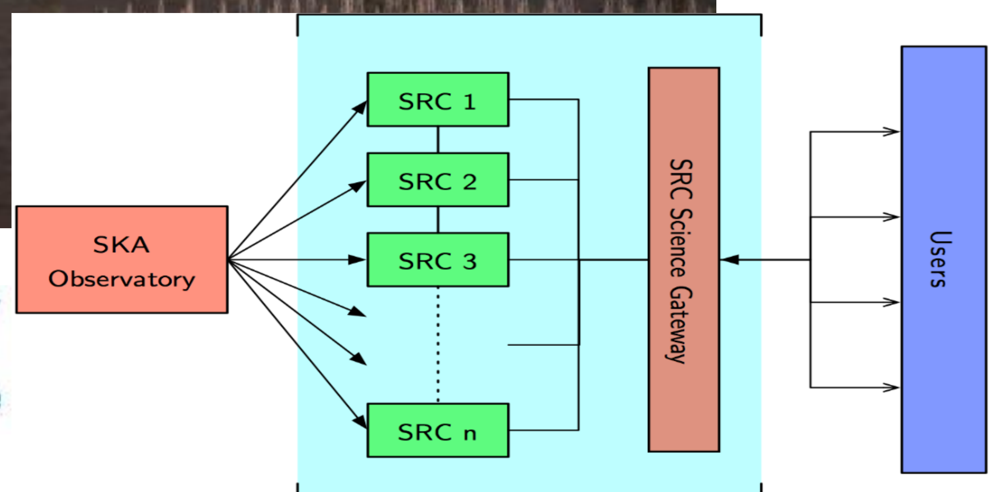


Data:
RAW → 2027: 600 PB
Derived (1 copy): → 2027: 900 PB

Square Kilometre Array

SKA Regional Centers





SQUARE KILOMETRE ARRAY

Exploring the Universe with the world's largest radio telescope

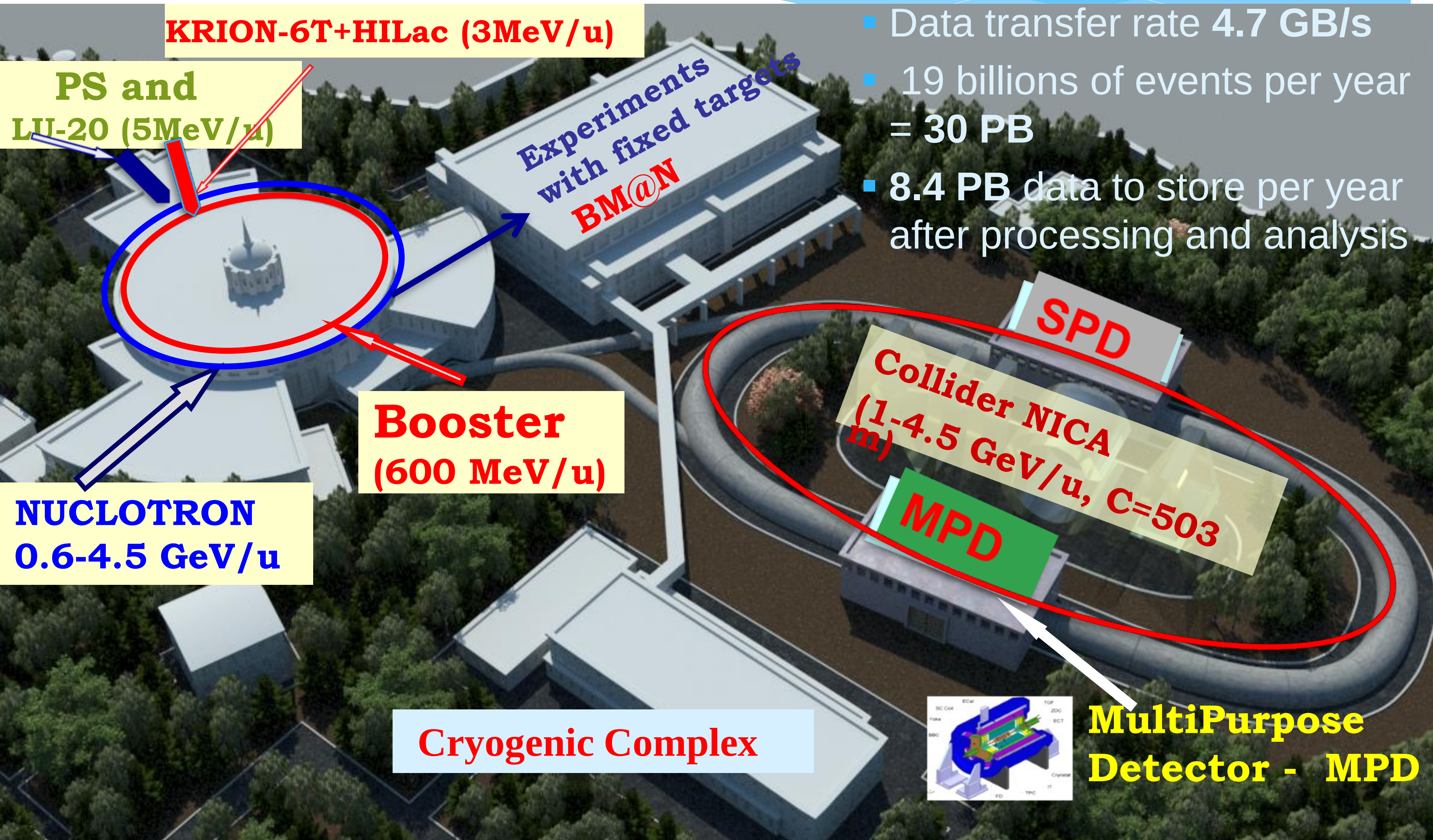
SKA > 1EB/Year

NICA Complex: *New era in the hot dense matter science*

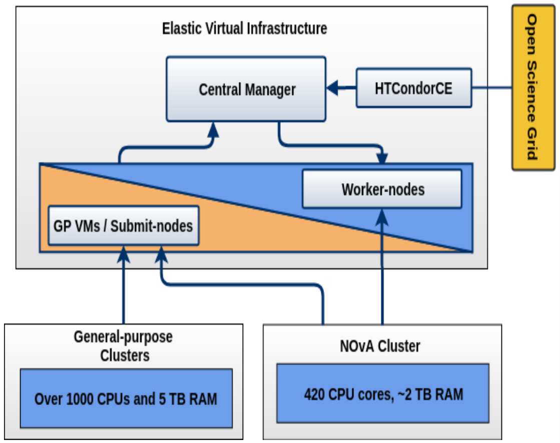
Collider basic parameters:

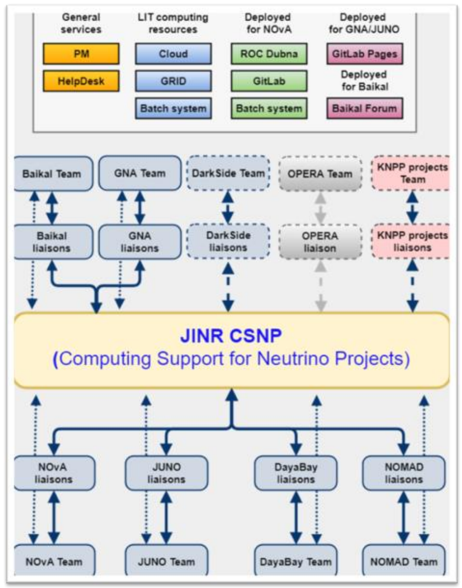
$\sqrt{S_{NN}}$ = 4-11 GeV; *beams: from* p to Au; L~10^{27} cm^{-2} c^{-1} (Au), ~10^{32} cm^{-2} c^{-1} (p)

- Data transfer rate **4.7 GB/s**
- 19 billions of events per year = **30 PB**
- **8.4 PB** data to store per year after processing and analysis

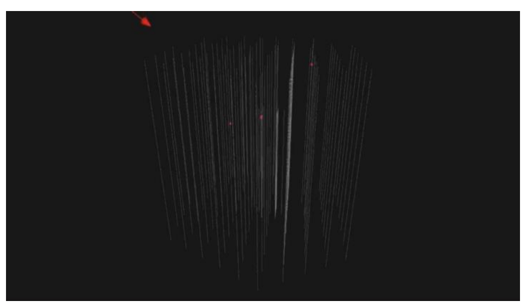


JINR NEUTRINO COMPUTING PLATFORM





- HTCondor - production and analysis, job processing;
- JupyterHub – development tool;
- Personal virtual machines - traditional interactive applications;
- GitLab CI - automated software testing and deployment.



The LIT and DLNP directorate agreed to establish a joint working group on writing a proposal about the dedicated project for developing computing facilities at JINR for neutrino experiments the Institute participates in.

LIT contribution:

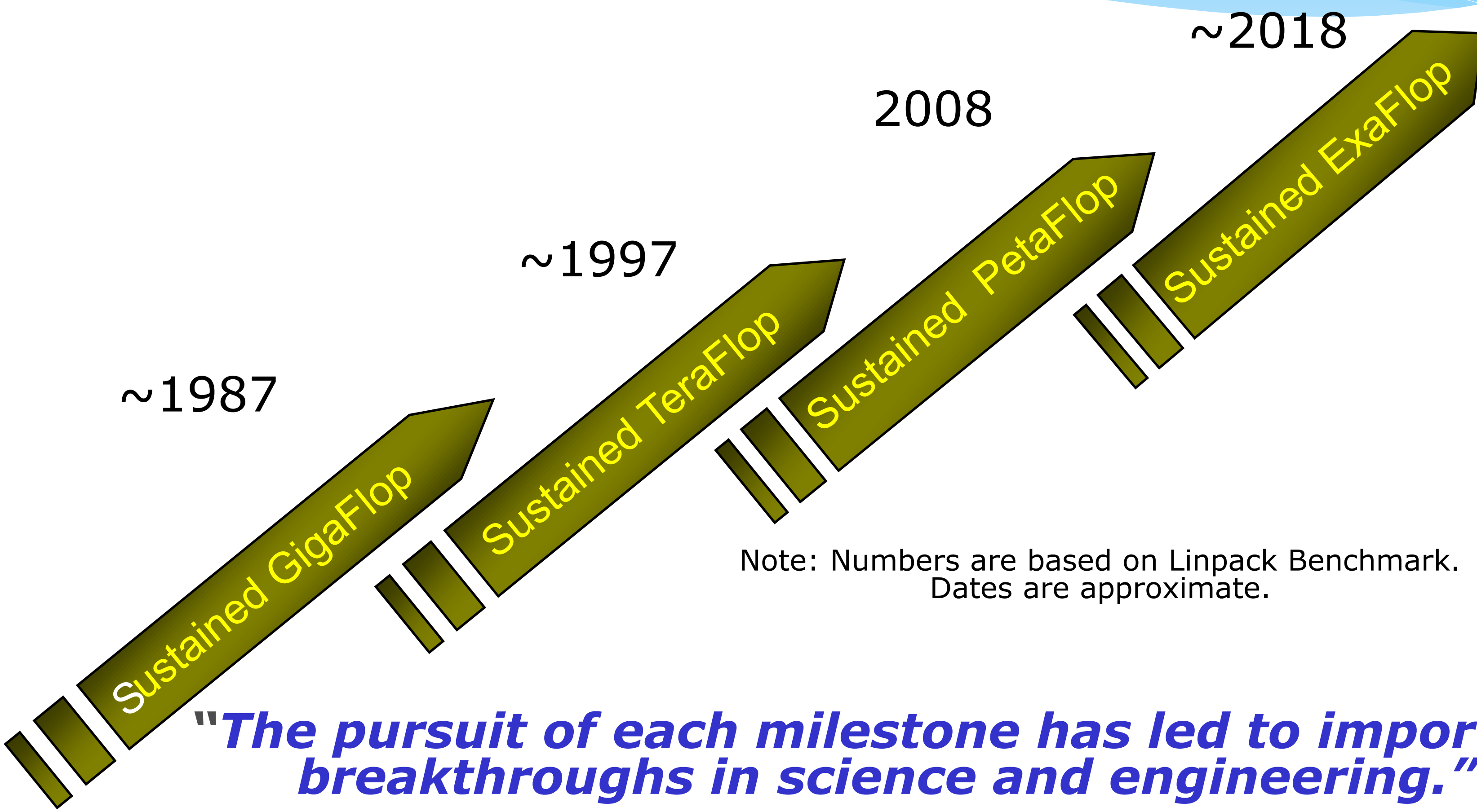
engineering infrastructure (electricity, UPS, cooling, network, racks, manpower)

DLNP contribution:

computing and storage resources (CPUs/GPUs&disks)

Reach Exascale by 2018

From GigFlops to ExaFlops



Note: Numbers are based on Linpack Benchmark.
Dates are approximate.

"The pursuit of each milestone has led to important breakthroughs in science and engineering."

Source: IDC "In Pursuit of Petascale Computing: Initiatives Around the World," 2007

TOP500 List – June 2019

	Site	System	Cores	Rmax TFlop/s	Rpeak TFlop/s
1	DOE/SC/Oak Ridge NL US	**Summit** - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, IBM	2,414,592	**148,600.0**	200,794.9
2	DOE/NNSA/LLNL US	**Sierra** - IBM Power System S922LC, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, IBM / NVIDIA	1,572,480	**94,640.0**	125,712.0
3	NSCC in Wuxi China	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway NRCPC	10,649,600	**93,014.6**	125,435.9
4	NSCC in Guangzhou China	**Tianhe-2A** - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, NUDT	4,981,760	**61,444.5**	100,678.7
5	Texas ACC US	**Frontera** - Dell C6420, Xeon Platinum 8280 28C 2.7GHz, Mellanox Dell EMC	448,448	**23,516.4**	38,745.9
6	Swiss CSCS Switzerland	**Piz Daint** - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 Cray Inc.	387,872	**21,230.0**	27,154.3
7	DOE/NNSA/LANL/SNL US	**Trinity** - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Intel Xeon Phi 7250 68C 1.4GHz, Cray Inc.	979,072	**20,158.7**	41,461.2
8	AIST Japan	**AI Bridging Cloud Infrastructure (ABCI)** - PRIMERGY CX2570, 6148 2.4GHz, NVIDIA Tesla V100, Fujitsu	391,680	**19,880.0**	32,576.6
9	Leibniz RC Germany	**SuperMUC-NG** - ThinkSystem SD650, Xeon Platinum 8174 24C 3.1GHz, Intel Omni-Path Lenovo	305,856	**19,476.6**	26,873.9
10	DOE/NNSA/LLNL US	**Lassen** - IBM Power System S922LC, IBM POWER9 22C 3.1GHz, NVIDIA Tesla V100 IBM / NVIDIA / Mellanox	288,288	**18,200.0**	23,047.2

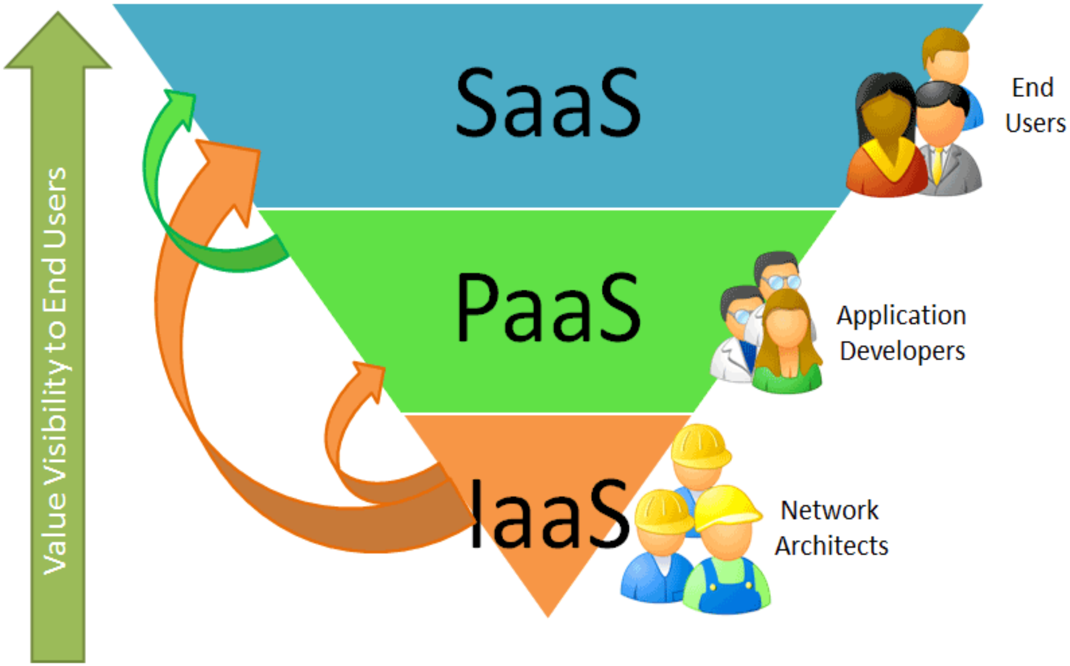
Supercomputers (Top 1,3,4)

Next Generation HPC Architectures

- Quantum Computing
- Photon Computers
- Special processors
- software platform for supercomputer modeling, digital testing and forecasting of complex technical systems
- Basic research in the development of physical and mathematical models and mathematical methods for exaflops computing (supercomputer doubles)
- Exaflops-class supercomputers (scientific and technical solutions, applied and system software, and others)
- Merging DA/ML/AI with Supercomputing

- In the next generation of supercomputers we see extensive use of accelerator technologies
  - Oak Ridge: **Summit** (2018)
    - 4608 IBM AC922 nodes *w/* 2x Power9 CPU
    - 3x NVIDIA Volta V100 + NVLink / CPU
  - LBL: NERSC-9 **"Perlmutter"** (2020)
    - AMD EPYC "Milan" x86 only nodes + mixed CPU / "next gen" NVidia GPU
  - Oak Ridge: **Frontier** (2021)
    - 1.5 exaflop
    - AMD EPYC CPU + 4x AMD "Instinct" GPU
  - Commercial clouds:
    - Brainwave / Azure FPGA
    - Google Cloud TPU
  - LLNL: **Sierra** (2018)
    - 4320 IBM AC922 nodes *w/* 2x Power9 CPU
    - 2x NVIDIA Volta V100 + NVLink / CPU
  - Argonne: **Aurora A21** (2021)
    - first exascale HPC (???)
    - Intel Xeon CPU + Intel X^e/gen12 GPU + Optane
  - Tsukuba: **Cygnus** (2020)
    - 2x Intel Xeon 6162+ 4x NVidia V100 GPU
    - 2x CPU + 4x GPU + 2x Intel Stratix FPGA
  - Japan: **Fugaku** (2021)
    - manycore ARM A64fx (48+2)
    - integrated "SVE" 512 bit GPU-like accelerator

Cloud computing



Gmail by Google

Dropbox

YouTube

Google App Engine

Microsoft Azure

Google Compute Engine

amazon

openstack

IBMSmartCloud Enterprise

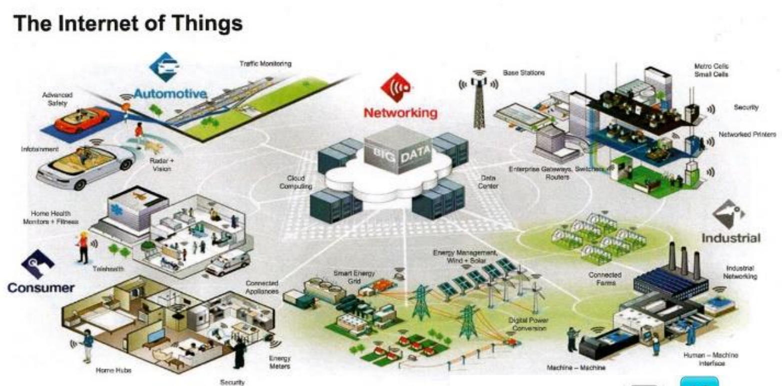
OpenNebula.org

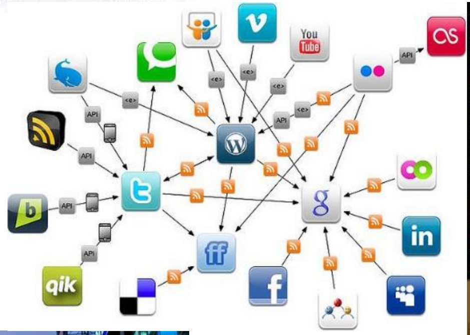
Big Data Analytics

Big data sources

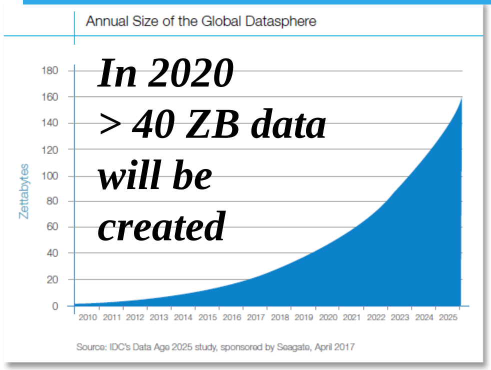
Where the data of such a volume is generated?...
In today's world - everywhere:

- ***Retail***
- ***Financial Institutions***
- ***Transport***
- ***WEB***
- ***Internet of Things (IoT)***
- ***Social Networks***
- ***Telecommunications***
- ***Energetics***
- ***Science***
- ***Industry***





Big Data + HPC (HPDA - High Performance Data Analysis)

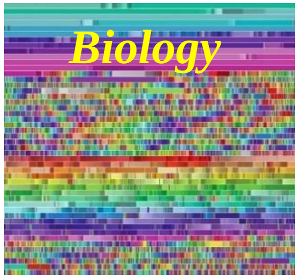


In 2020
> 40 ZB data
will be
created

Annual data production follows to exponential law.

High Energy Physics

Biology



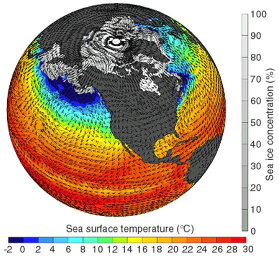
Nanotechnology

Science

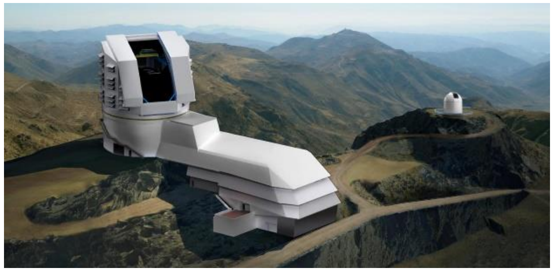
CERN Large Hadron Collider
HL > 600 Pb/Year

Astrophysics

Climate



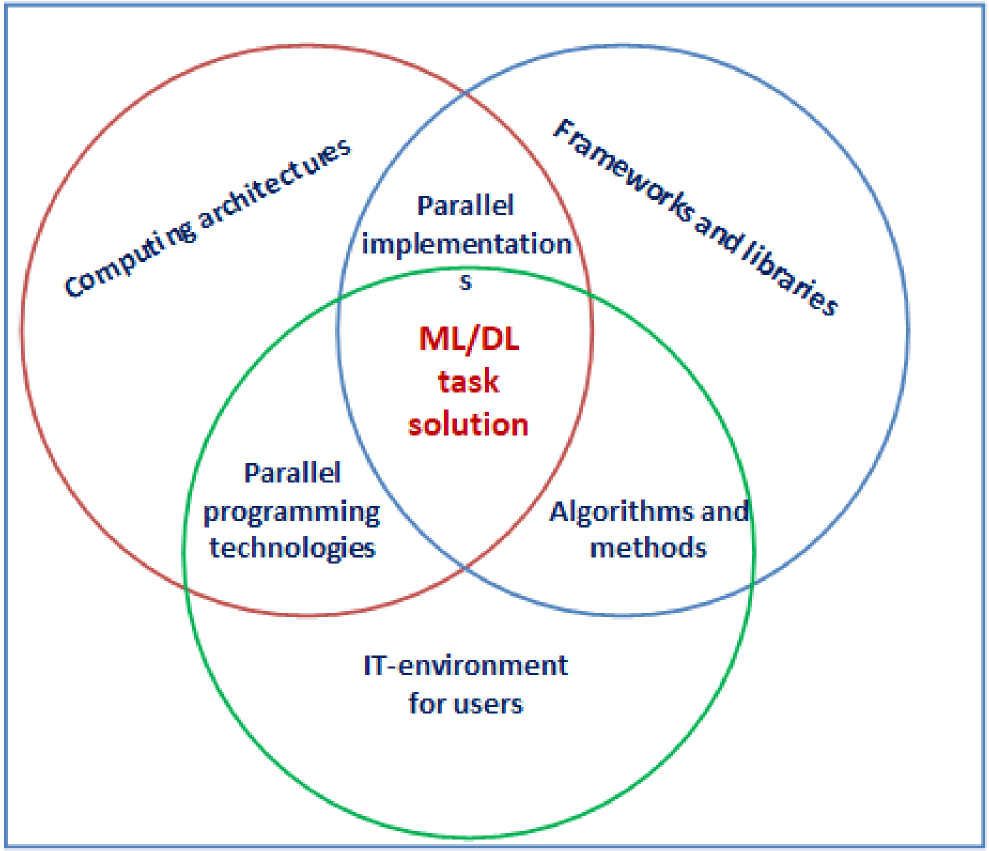
Square Kilometer Array radio telescope (SKA) > 1 Eb/Year



...et cetera

Large Synoptic Survey Telescope (LSST) > 10 Pb/Year (estimation)

Implementation of the neural network approach, methods and algorithms of ML/DL



Venn diagram on the implementation of the neural network approach, methods and algorithms of ML/DL to solve applied tasks.

To provide all the possibilities both for developing mathematical models and algorithms and carrying out resource-intensive calculations including graphics accelerators, which significantly reduce the calculation time, an ecosystem for tasks of machine learning and deep learning (ML/DL) and data analysis has been created and is actively developing for HybriLIT platform users.

Platforms of Big Data

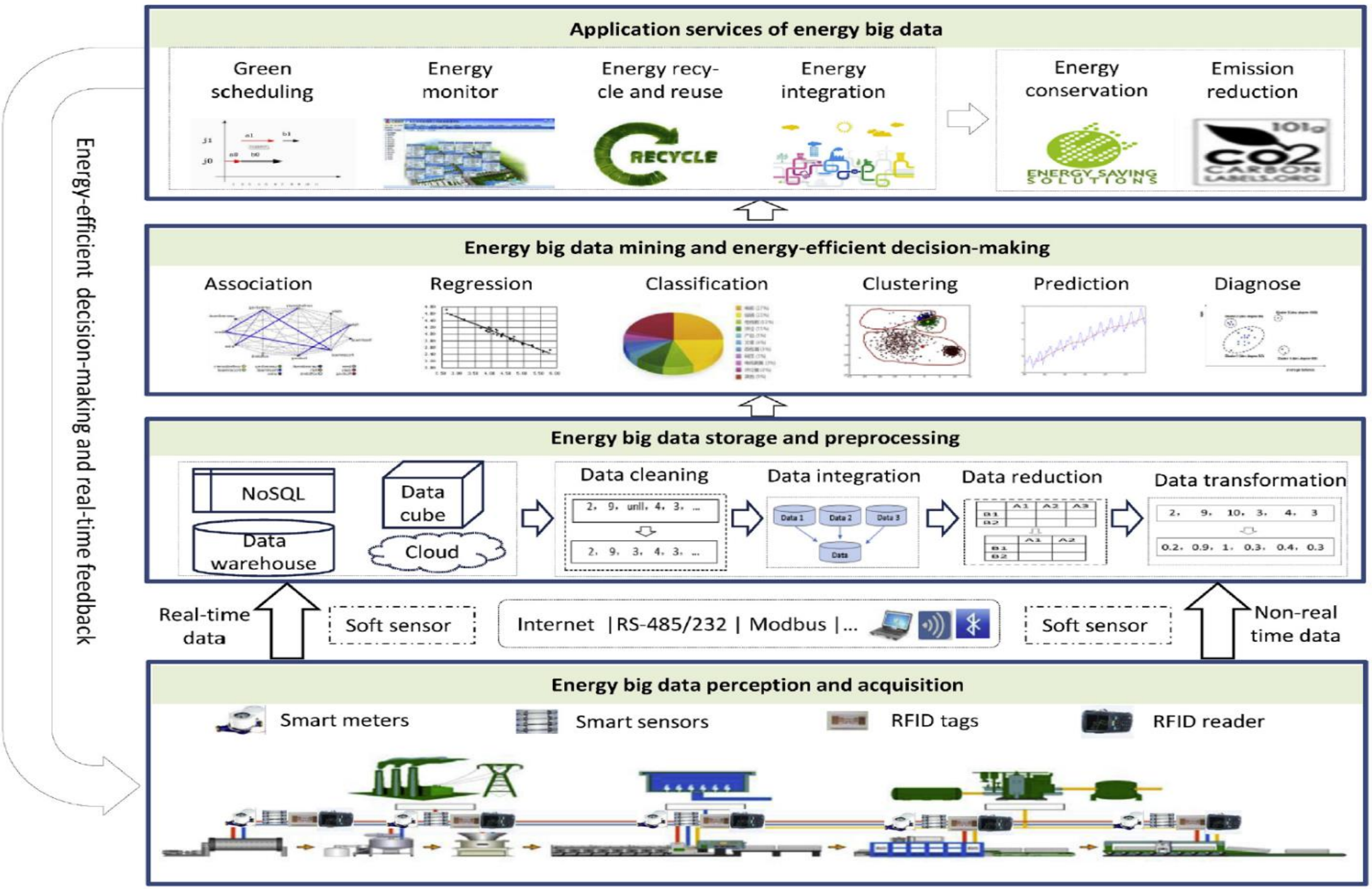
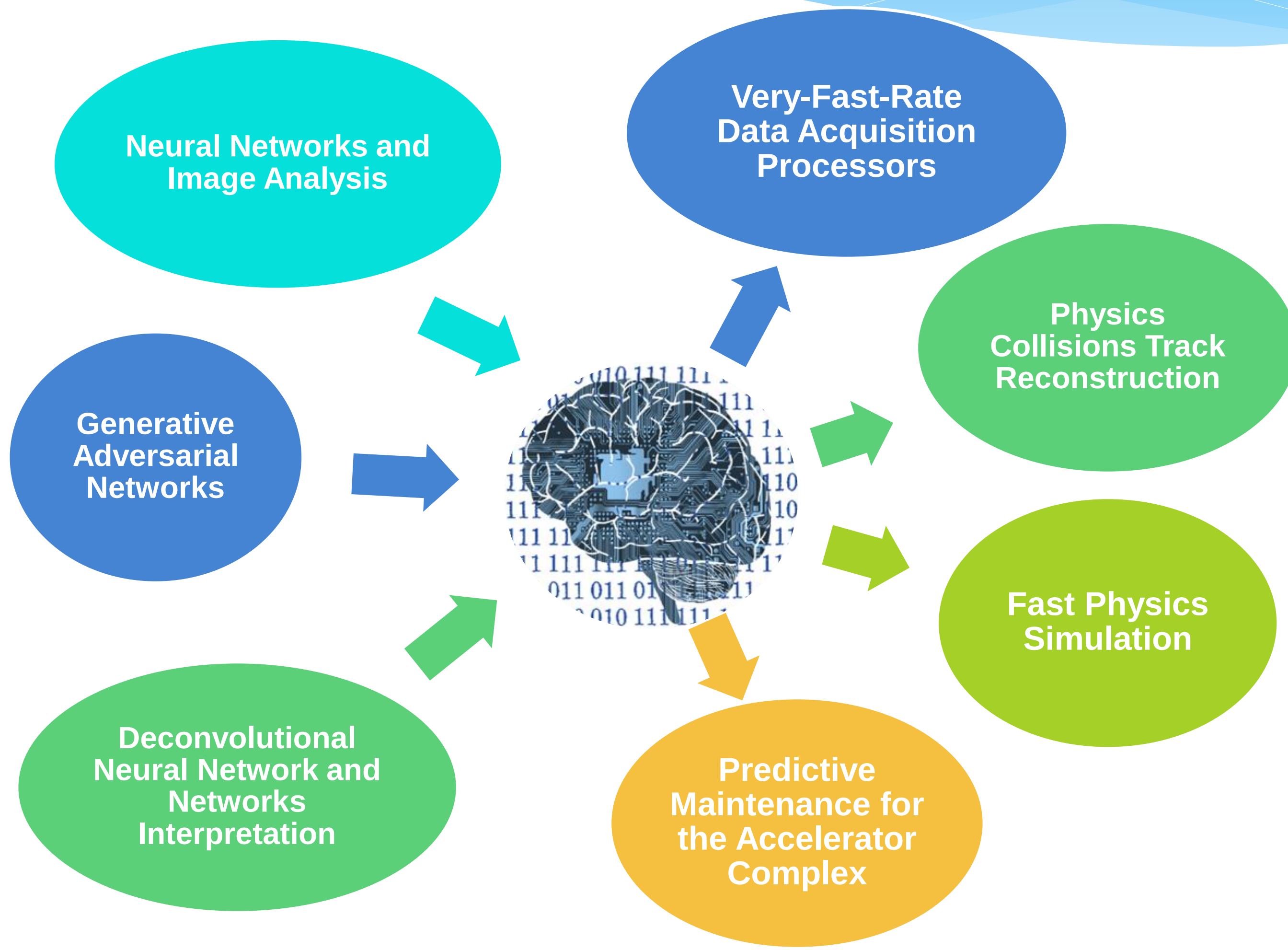
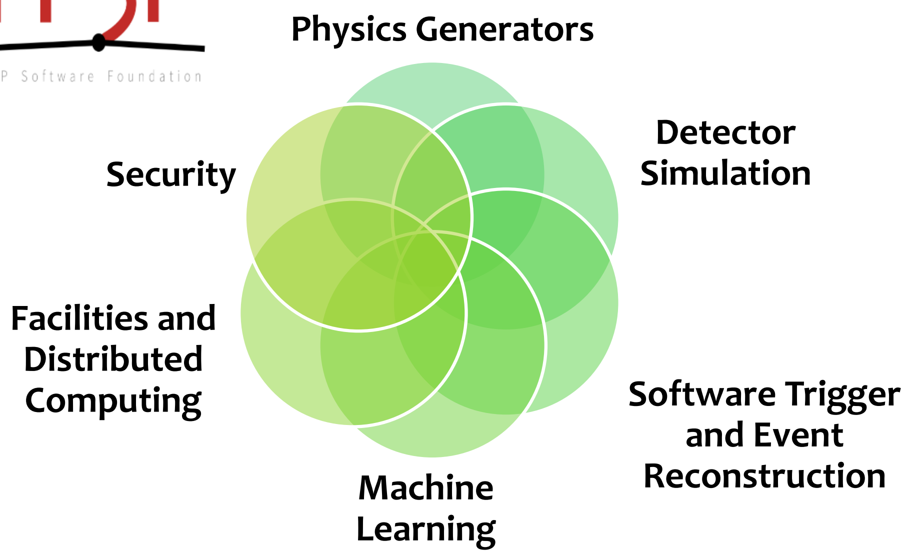


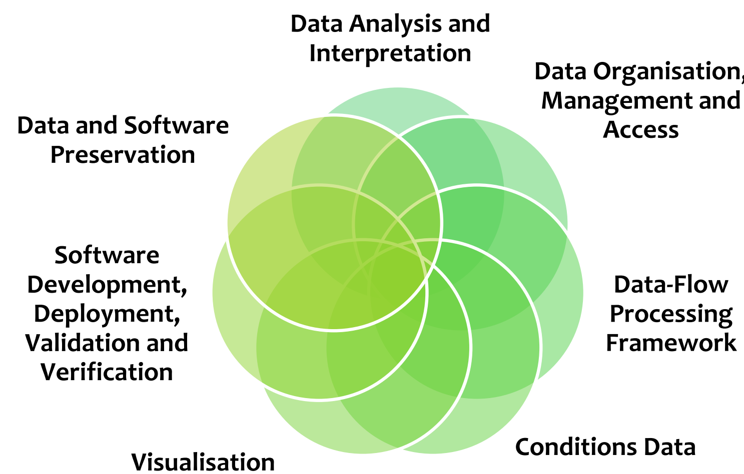
Fig. 3. A big data driven analytical framework for energy-intensive manufacturing industries (based on the example of ceramics industry).

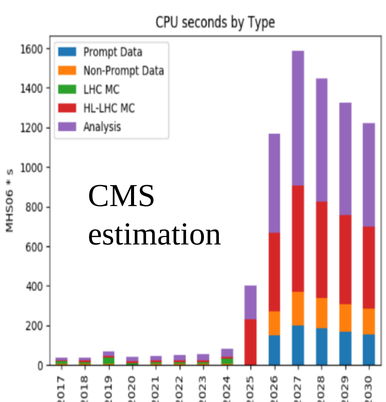
Artificial Intelligence (ML/DL)

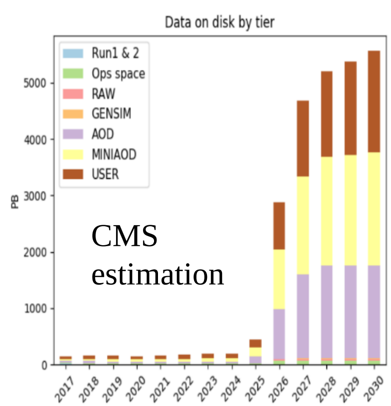


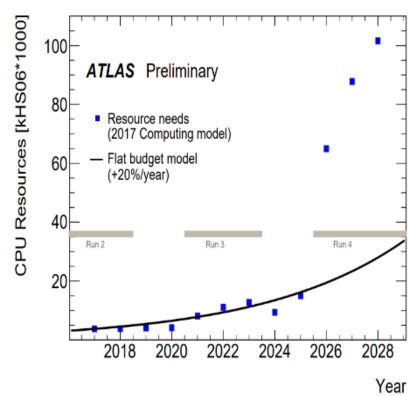
CHALLENGE: R&D of software to acquire, manage, process, and analyse the big amounts of data to be recorded

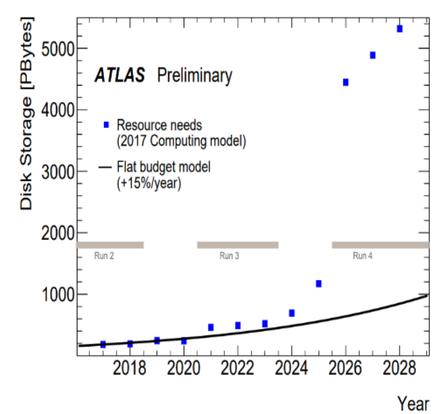












CHALLENGES: distributed data storage evolution: DATALAKES

GOAL:

•to provide a computing infrastructure to the experiments and the community to store and analyze data,

•to achieve storage consolidation where geographically distributed storage centers (potentially deploying different storage technologies) are operated and accessed as a single entity.

EOS - a CERN open-source storage software solution to manage multi PB storage.

XRootD - core of the implementation framework providing a feature-rich remote access protocol.

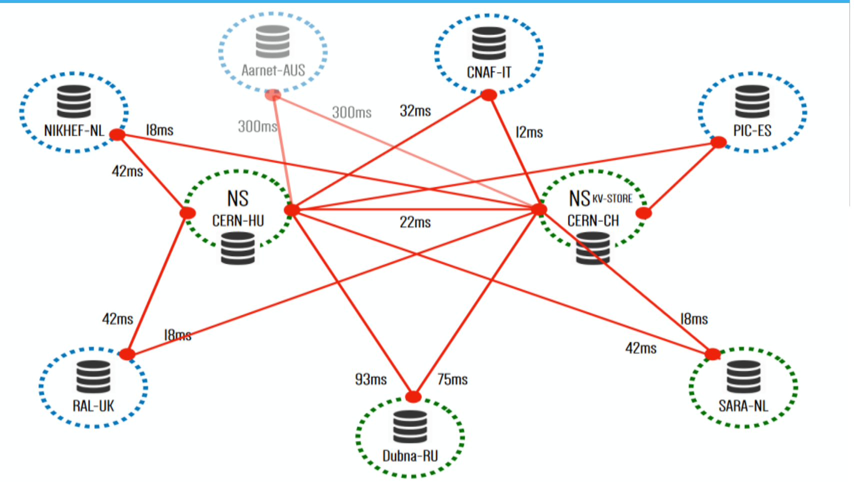
Improvement of already existing production quality Data Management services.

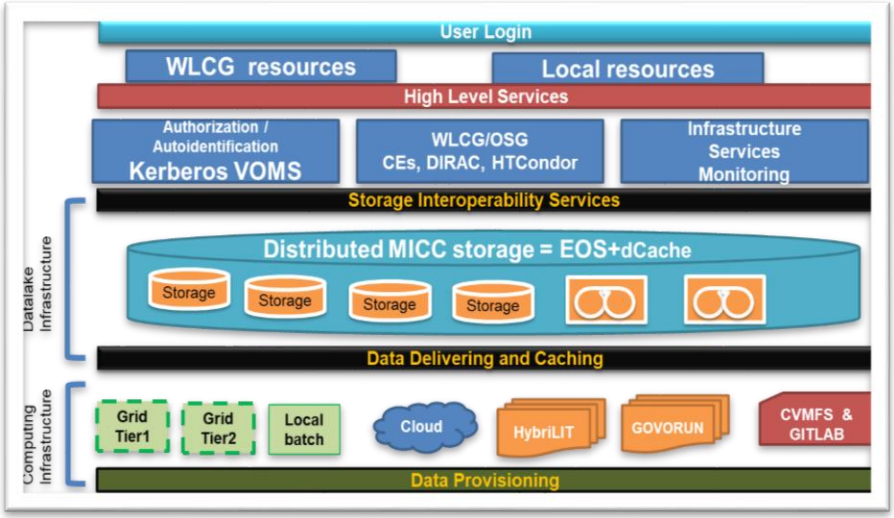
Scalable technologies for federating storage resources and managing data in highly distributed computing environments.

JINR DATALAKE

* logical separation of computing infrastructure and data storage;
* the presence of high-level services (control of job flows, loading of data) that interact with all elements of the infrastructure and control the use of resources
* the presence of a hierarchical, geographically distributed structure from regional data lakes of various sizes with a specific network topology, internal balancing mechanisms;
* the presence of "smart" services for transferring data between all components of the infrastructure, as well as services for determining and predicting the amount of computing resources needed to complete task flows.

EU DataLake



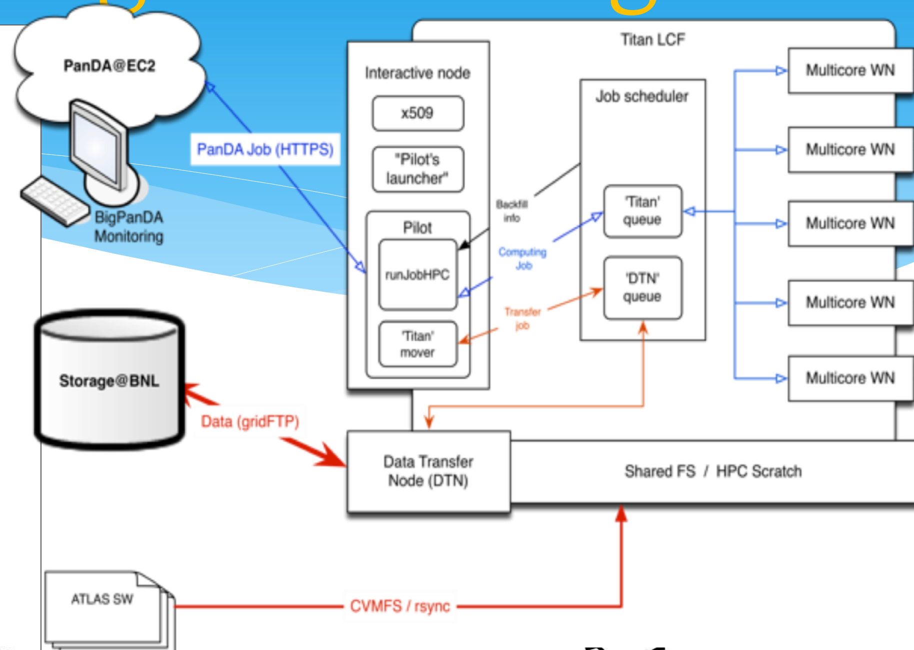


- DataLake deployed at JINR
- The new storage system successfully integrated into the MICC structure
- It shows great performance for storing and accessing big arrays of information.
- It can be applied for all the steps of data handling.

Integrating Distributed Computing Technologies for Big Data Management

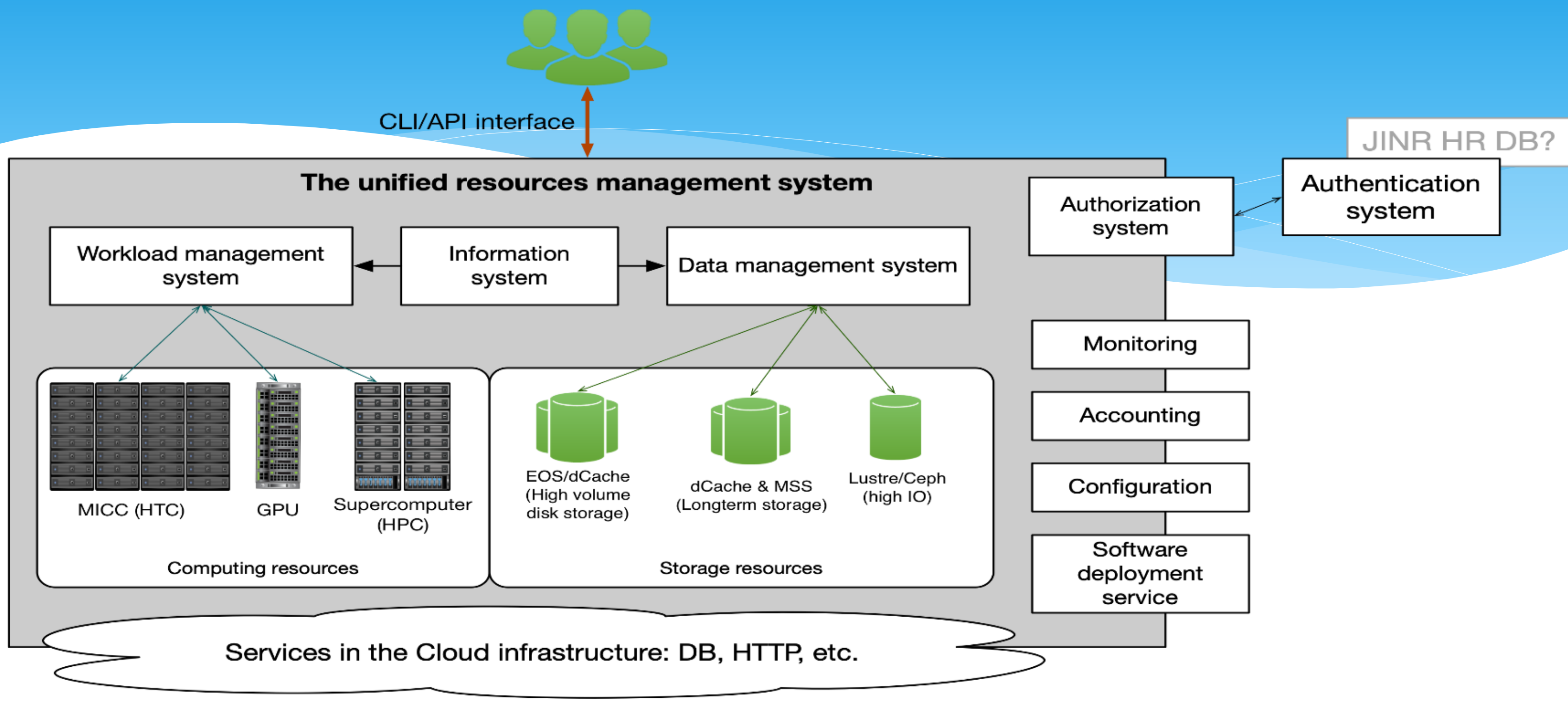
Titan System (Cray XK7)			
Peak Performance	27.1 PF 18,688 compute nodes	24.5 PF GPU	2.6 PF CPU
System memory	710 TB total memory		
Interconnect	Gemini High Speed Interconnect	3D Torus	
Storage	Lustre Filesystem	32 PB	
Archive	High-Performance Storage System (HPSS)	29 PB	
I/O Nodes	512 Service and I/O nodes		

12 OLCF 20 OAK RIDGE National Laboratory



Development of software platforms that provide transparency in the process of storing, processing and managing data for applications with large data streams and massive calculations. Integration of various architectures of distributed and parallel computing (grid, cloud, clusters, data centers, supercomputers) in order to create a universal platform for large-scale big data management projects.

The unified resource management system



The main objectives of a unified resource management system are:

–to provide the ability to process large amounts of data;

–to provide the possibility to organize massive computing tasks;

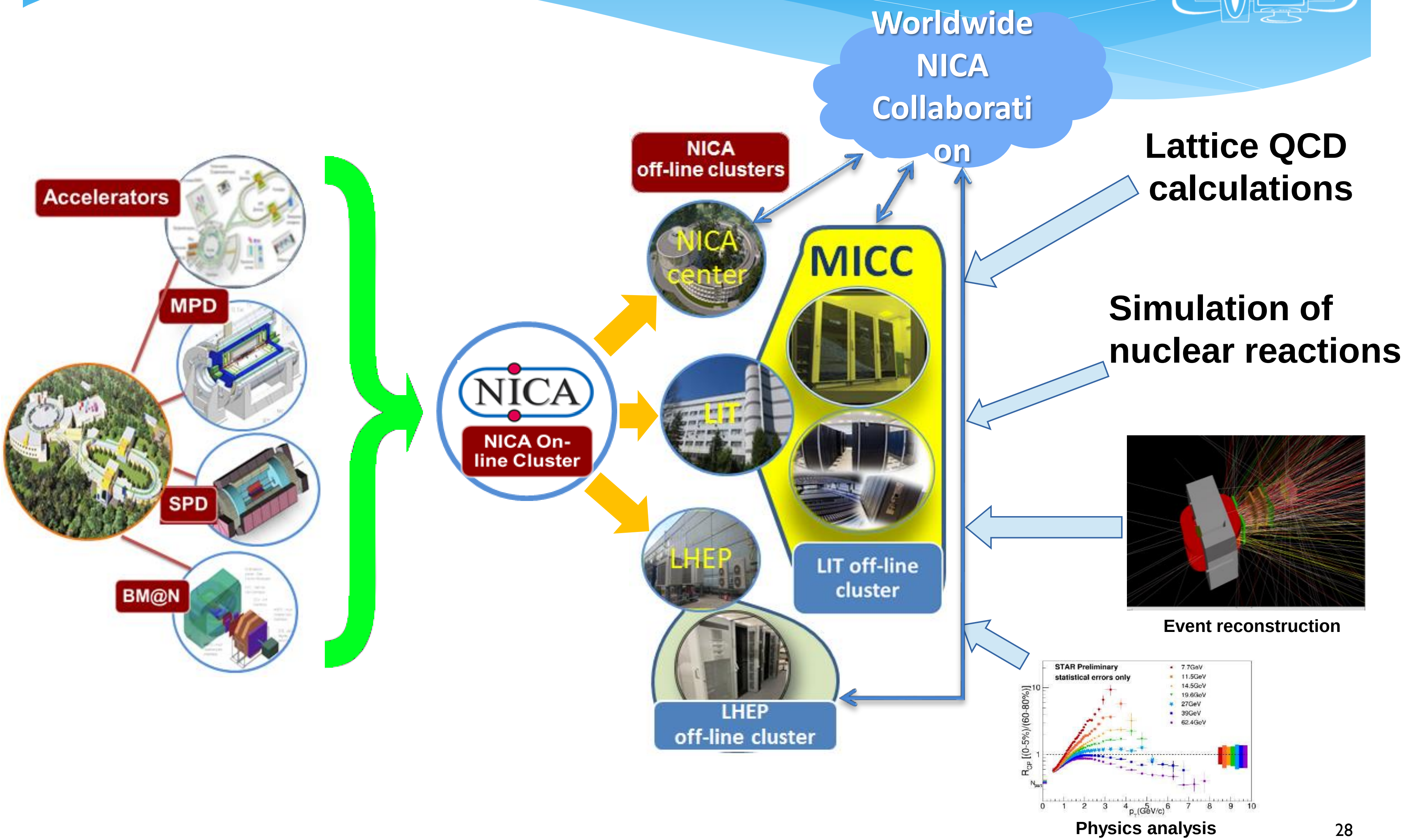
–to optimize the efficiency of the use of computing and storage resources;

–to effectively monitor the resource loading;

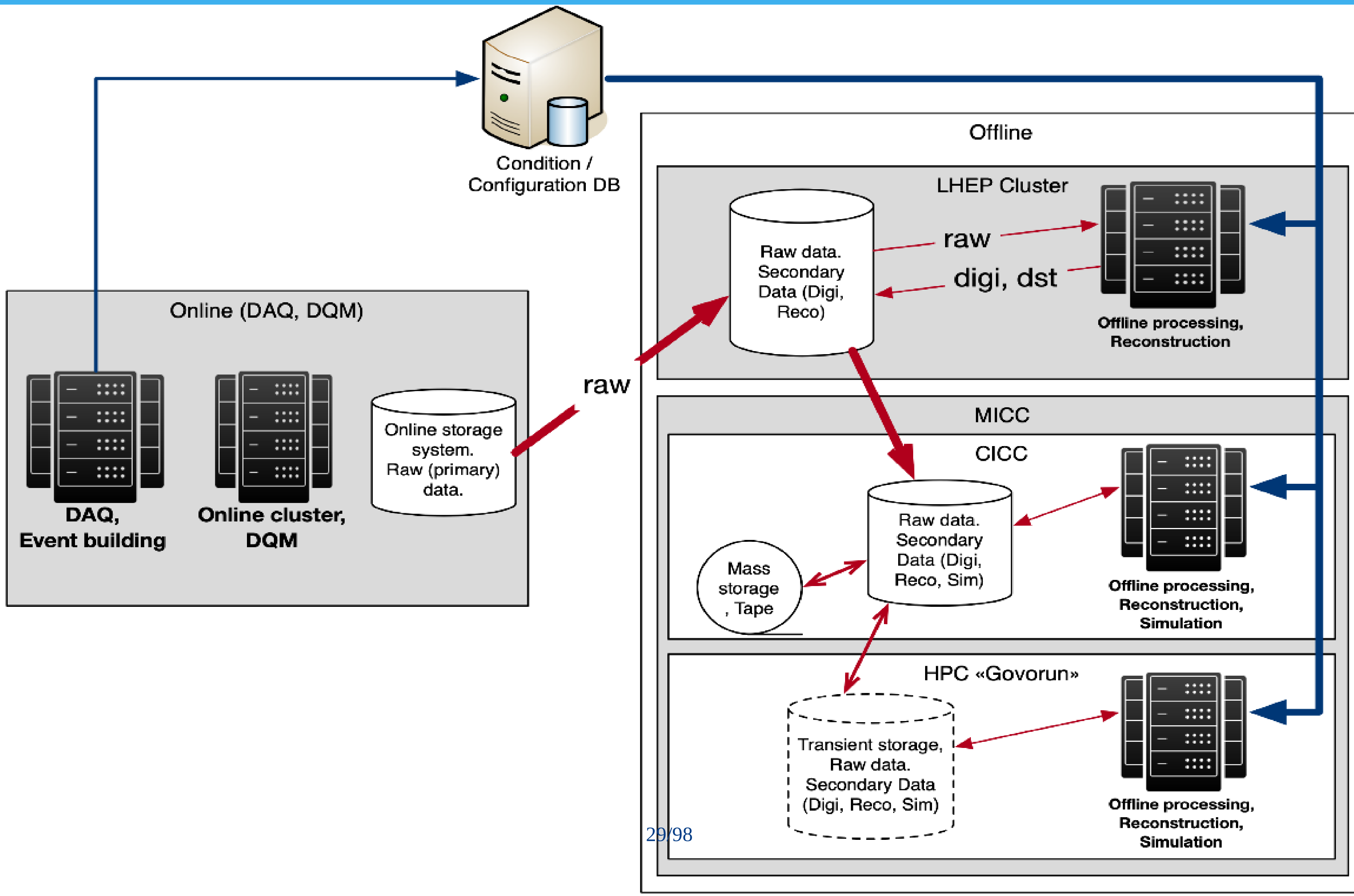
–to consolidate the accounting for the use of resources;

–to provide a unified interface of access to resources.

NICA Computing Concept & Challenges

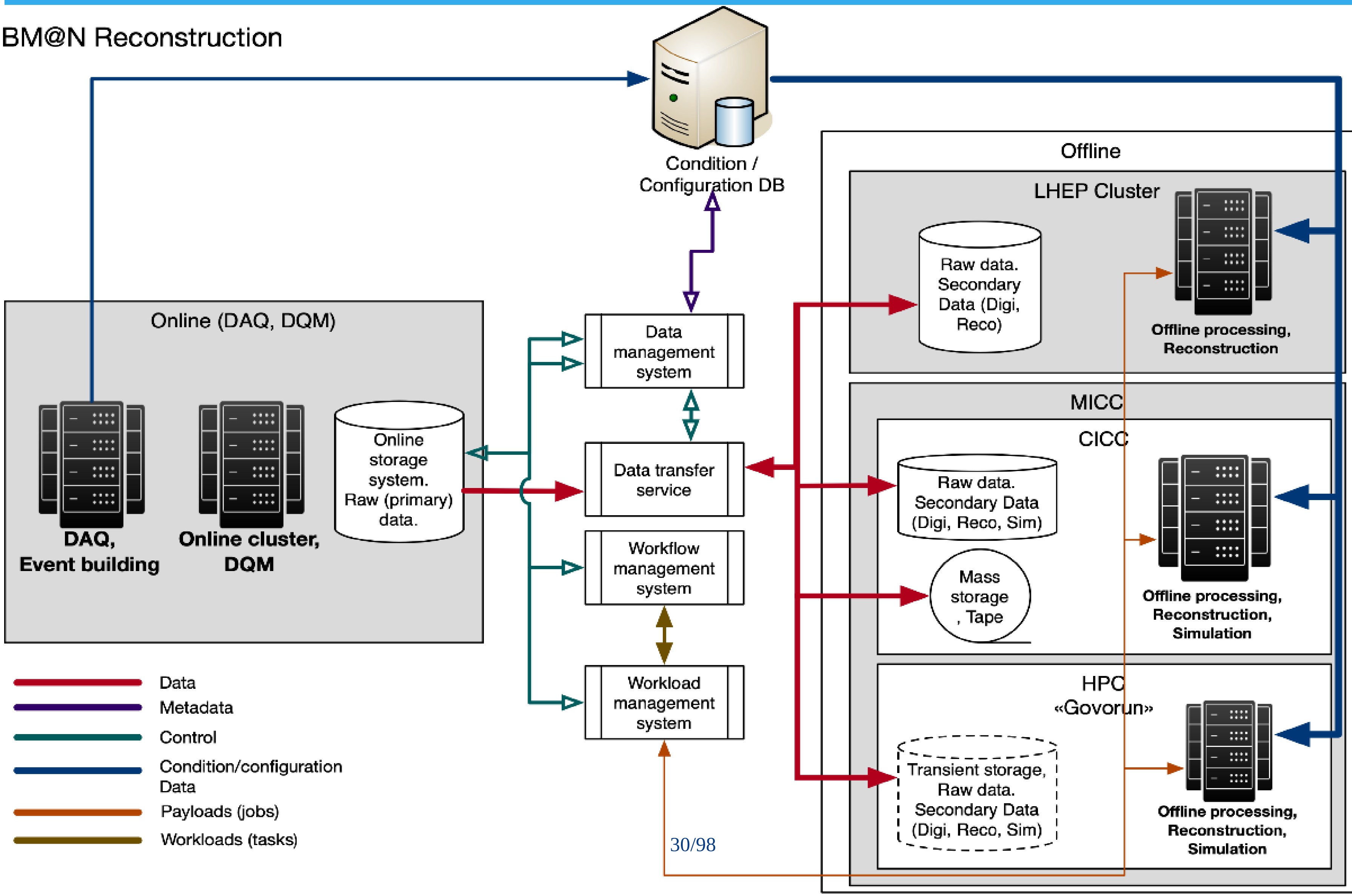


Data flow (for BM@N)



Services (for BM@N)

BM@N Reconstruction



International IT-School “Data Science”

The main goal of the International School of Information Technologies “Data Science” is to prepare highly qualified IT-specialists in Data Science who are able to define and solve theoretical and practical problems with the help of Big Data analytics.

Among major School directions special attention will be paid to the development of computing models, software platforms of the system for acquisition, storage, processing and analysis of experimental data from the megascience installations (NICA, PIC, LHC, FAIR, SKA, etc.).

IT-school “Data Science” groups created in PRUE, Dubna University…

Events:

- IT-School for young scientists “Modern IT-technologies for solving scientific problems” on the basis of North Ossetian State University in Vladikavkaz, May 2019
- IT-school “Machine Learning, Parallel and Hybrid Computations & Big Data Analytics” in framework of the International Conference “MATHEMATICAL MODELING AND COMPUTATIONAL PHYSICS”, High Tatra Mountains, Slovakia, July, 2019
- Summer Computer School "Big Data Analytics Dubna-2019“, July, 2019
- Parallel programming school for Czech and Slovak students (LIT JINR, Dubna, July 2019)
- International students school “Big Data mining and distributed systems“ in framework of the XXVII International Symposium on Nuclear Electronics and Computing (NEC-2019), Budva, Montenegro, September - October 2019.

Development of the system for training and retraining IT-specialists

Training courses, tutorials and lectures

Parallel programming technologies

OpenMP
MPI
OpenCL
NVIDIA CUDA

Tools for debugging and profiling of parallel applications

Intel Cluster Studio

Work with packages of applied software

COMSOL MULTIPHYSICS
ROOT
MATLAB
Wolfram Mathematica
GEANT4
Maple

Frameworks and tools for ML/DL tasks

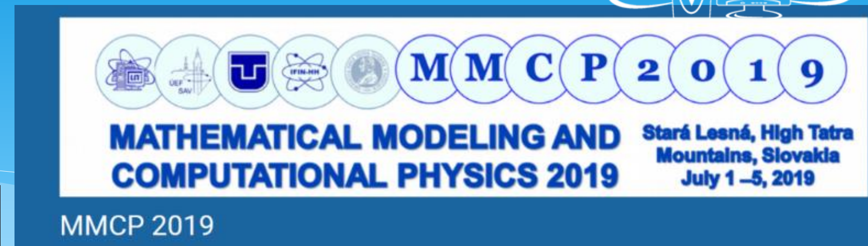
jupyterhub
TensorFlow
learn
NumPy

Training of the Institute staff, students and young scientists from the JINR Member States is carried out within :

- activities organized by the JINR University Centre;
- conferences and schools organized by JINR;
- international cooperation programs at JINR Member States institutes.

Traditional LIT conferences

International Conference “Distributed Computing and Grid-technologies in Science and Education”



27th Symposium on Nuclear Electronics and Computing (NEC’2019): Montenegro, Budva, September 30 – October 4, 2019

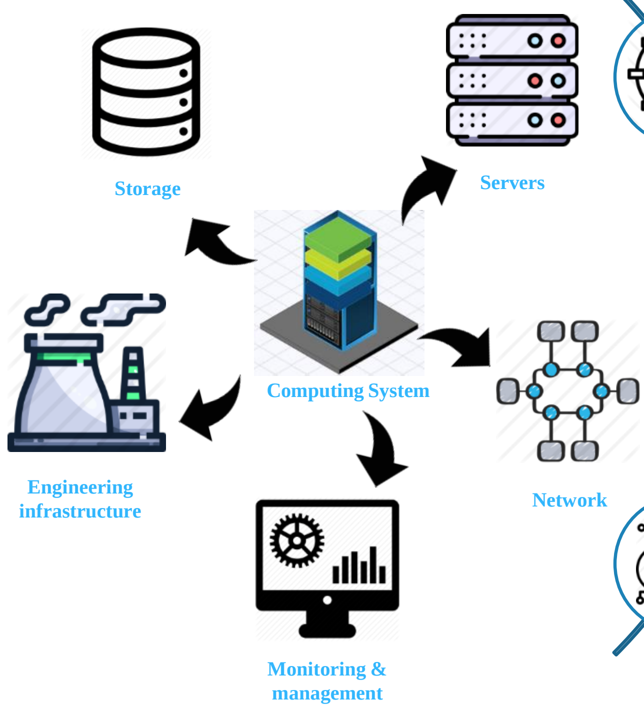
Long-Term Concept of a Scientific IT-ecosystem at JINR

Telecommunication technol

- terabit networks
- on-demand networks, networks of networks, software-defined networks, etc.
- dynamic resource allocation network systems
- networks enabling the large amount of data transfers
- networks allowing distributed processing of the exabyte data volume.

Long-Term Concept of a Scientific IT-ecosystem at JINR

Computing systems



- distributed software-configured platform
- supercomputer (heterogeneous)
- grid and cloud technologies
- effective use of novel computing architectures

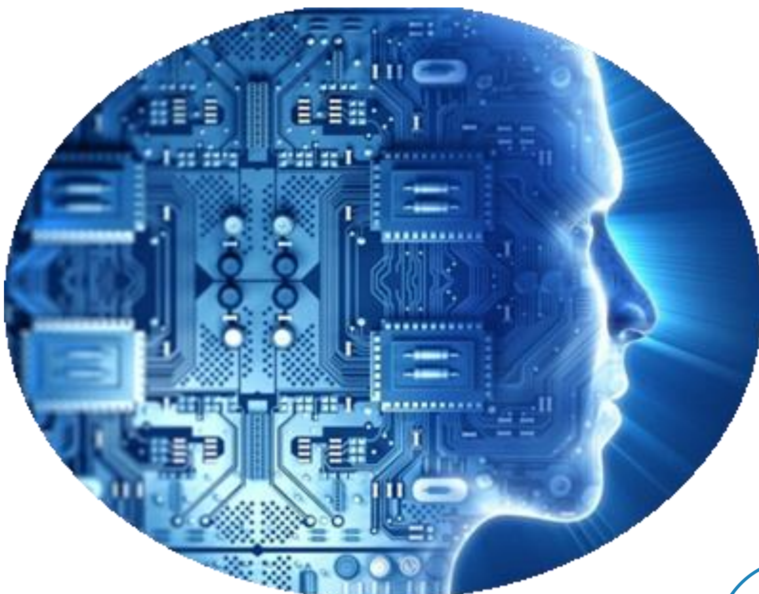
Long-Term Concept of a Scientific IT-ecosystem at JINR

Algorithms and software

- complex computing problems
- machine learning
- data mining
- quantum computing

Long-Term Concept of a Scientific IT-ecosystem at JINR

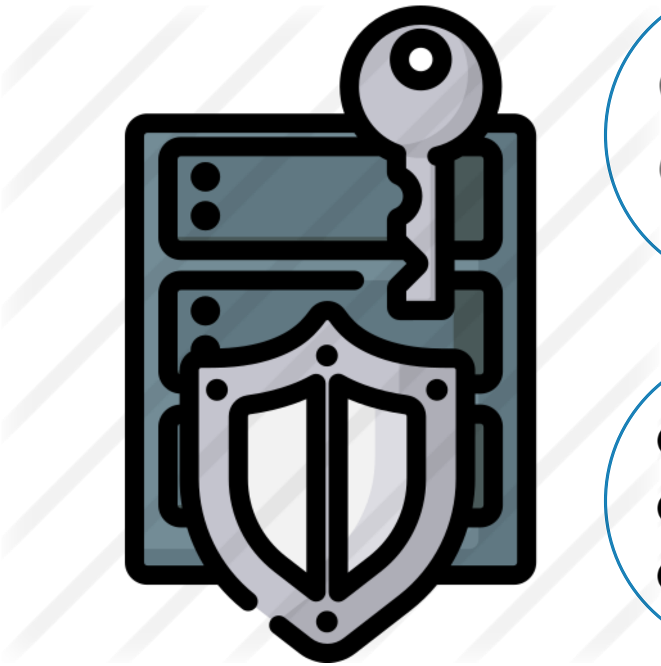
Technologies of data processing and analysis



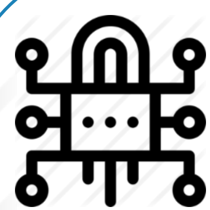
- computing with the use of supercomputers
- processing of multi-dimensional and hierarchical data of Exabyte volumes
- Big Data analysis and processing
- Artificial Intelligence and robotics
- SyMSim — modeling of computing architectures and workflows

Long-Term Concept of a Scientific IT-ecosystem at JINR

Information security

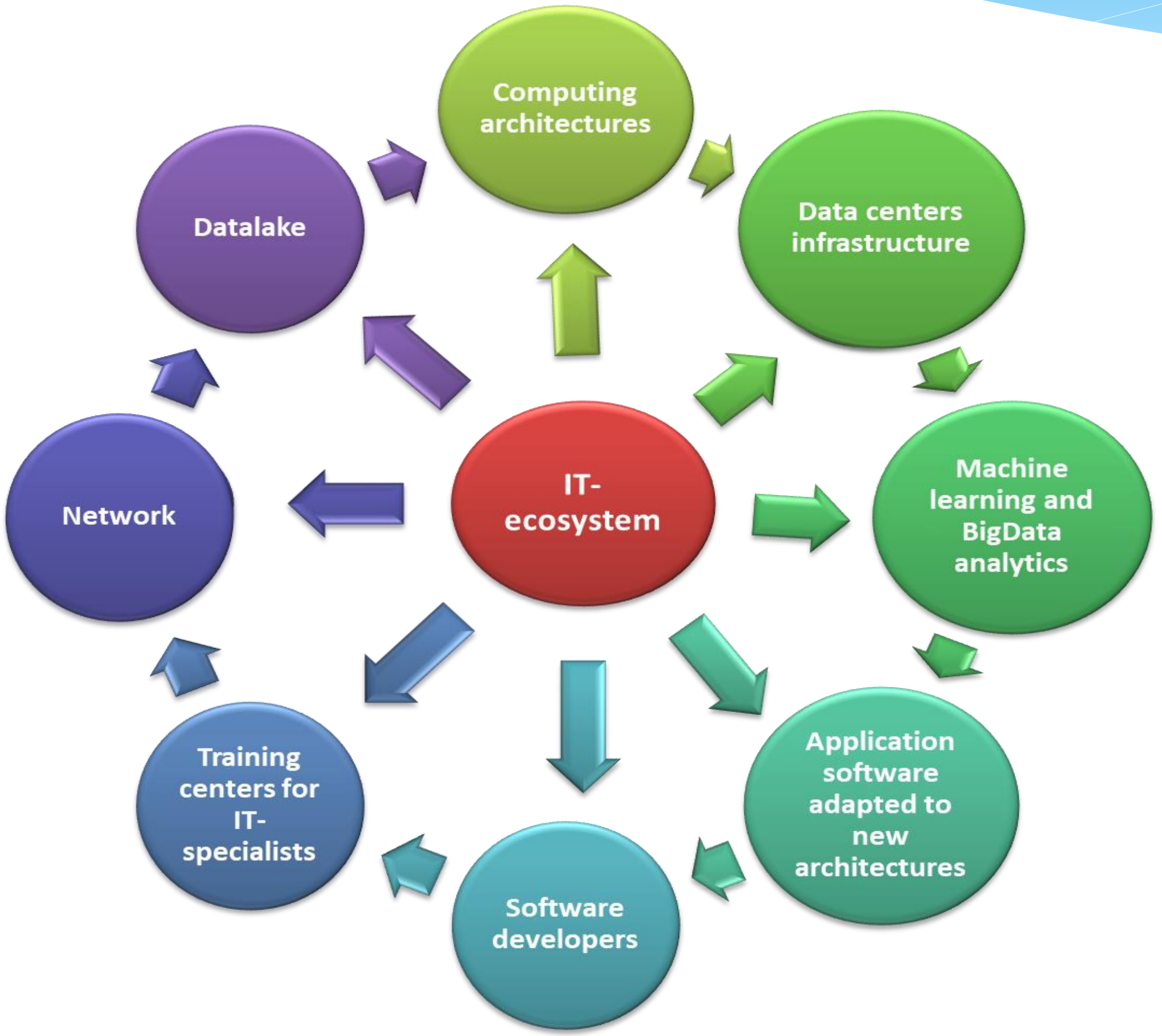


new security paradigms including quantum cryptography and neurocognitive principles



novel and up-to-date tools and software systems for data and information protection

STRATEGIC LONG-TERM PLAN



AIM

Expandable worldwide dynamicaly evolving IT-ecosystem that combines a variety of technological solutions, state-of-art computing concepts and methodologies.

PURPOSES

Significantly reduce the time spent on the implementation of projects that require computing resources and IT expertise

BENEFICIARIES

JINR, its Member States and international collaborators

Big lake with large flow