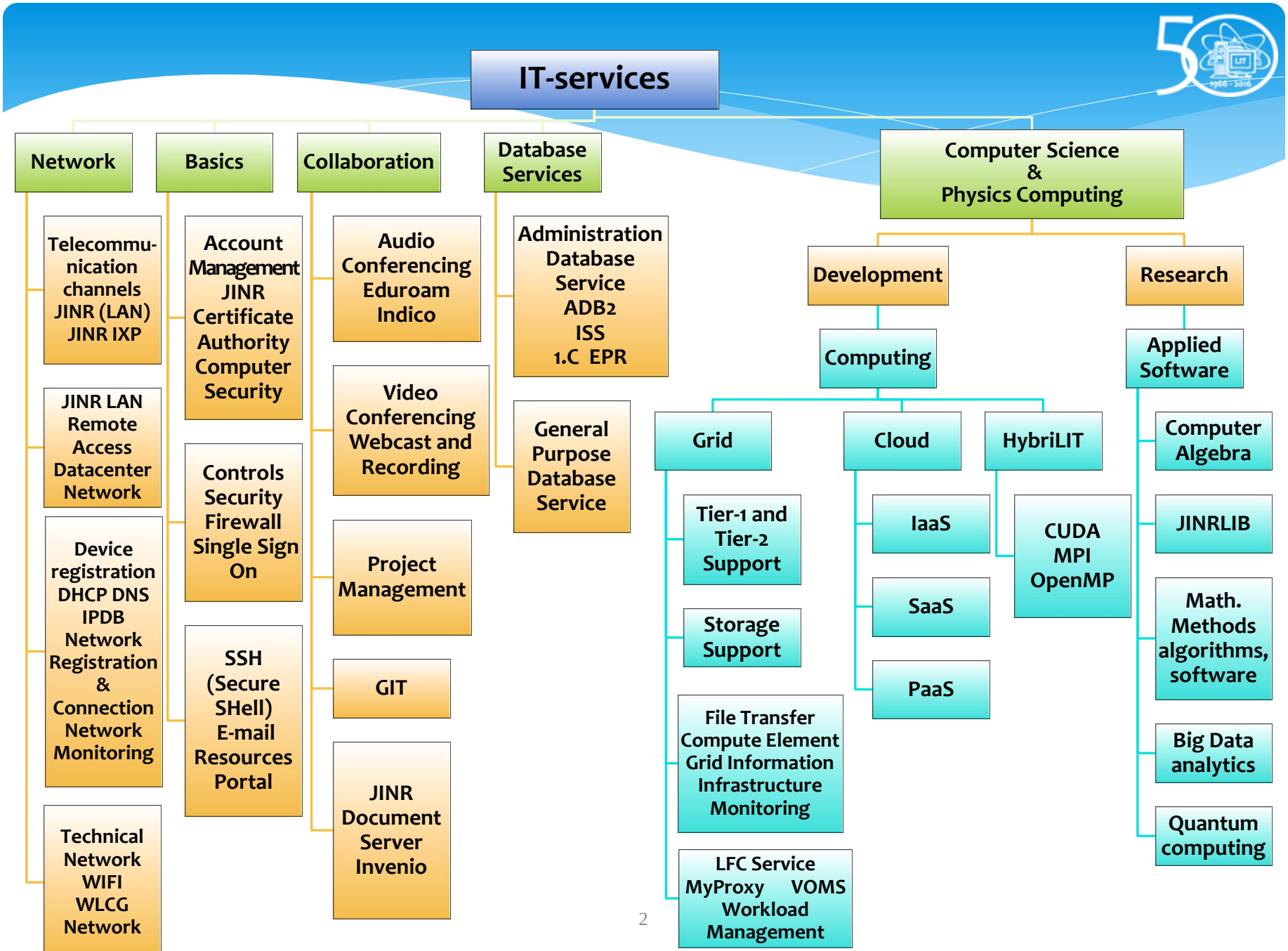
Distributed computing and Big data technologies at JINR

Korenkov Vladimir

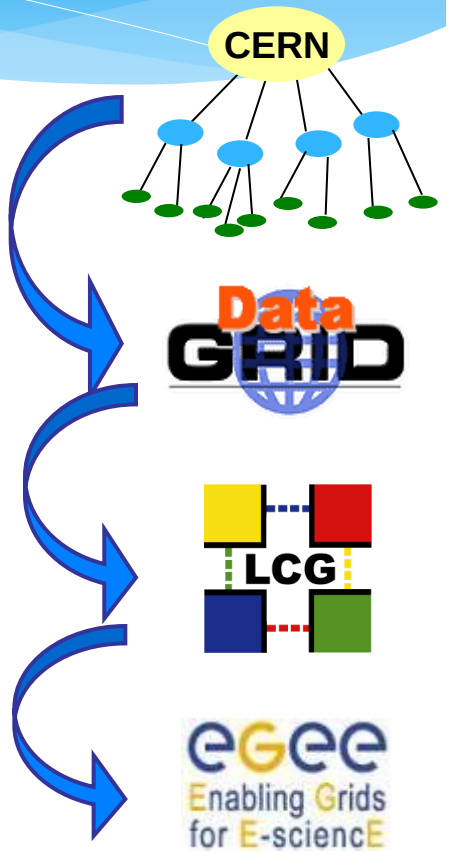
Director LIT JINR

CREMLIN Big data Workshop
Moscow, 16 February


IT-services
Network
Basics
Collaboration
Database Services
Computer Science & Physics Computing
Telecommu-nication channels JINR (LAN) JINR IXP
JINR LAN Remote Access Datacenter Network
Device registration DHCP DNS IPDB Network Registration & Connection Network Monitoring
Technical Network WIFI WLCG Network
Account Management JINR Certificate Authority Computer Security
Controls Security Firewall Single Sign On
SSH (Secure SHell) E-mail Resources Portal
Audio Conferencing Eduroam Indico
Video Conferencing Webcast and Recording
Project Management
GIT
JINR Document Server Invenio
Administration Database Service ADB2 ISS 1.C EPR
General Purpose Database Service
Development
Research
Computing
Applied Software
Grid
Cloud
HybriLIT
Tier-1 and Tier-2 Support
Storage Support
File Transfer Compute Element Grid Information Infrastructure Monitoring
LFC Service MyProxy VOMS Workload Management
IaaS
SaaS
PaaS
CUDA MPI OpenMP
Computer Algebra
JINRLIB
Math. Methods algorithms, software
Big Data analytics
Quantum computing

Some history

- 1999 – Monarc Project
  - Early discussions on how to organise distributed computing for LHC
- 2001-2003 - EU DataGrid project
  - middleware & testbed for an operational grid
- 2002-2005 – LHC Computing Grid – LCG
  - deploying the results of DataGrid to provide a production facility for LHC experiments
- 2004-2006 – EU EGEE project phase 1
  - starts from the LCG grid
  - shared production infrastructure
  - expanding to other communities and sciences
- 2006-2008 – EU EGEE-II
  - Building on phase 1
  - Expanding applications and communities …
- 2008-2010 – EU EGEE-III
- 2010-2014 - EGI-InSPIRE



Information Technology 2016

LIT IT-infrastructure is the one of JINR basic facilities

- LAN: 10 Gbps
- WAN: 100 Gbps
- Tier-1: 3600 core, 4PB disk, 5.4PB tape
- CICC/Tier-2: 3500 core, 2PB disk
- HybriLIT: 252 CPU, 77184 GPU cores, 182 PHI-cores, 2.4 TB RAM, 57.6 TB HDD, 142 Tflops
- Cloud: 330 CPU, 840GB RAM

Project MICC: main components

JINR grid sites of WLCG/EGI: Tier-1 for CMS
Tier-2 for ALICE, ATLAS, CMS, STAR, LHCb,
BES, biomed, fermilab

Cloud infrastructure

Heterogeneous(CPU + GPU) computing cluster HybriLIT

Off-line cluster and storage system for BM@N, MPD, SPD Storage and computing facilities for local users

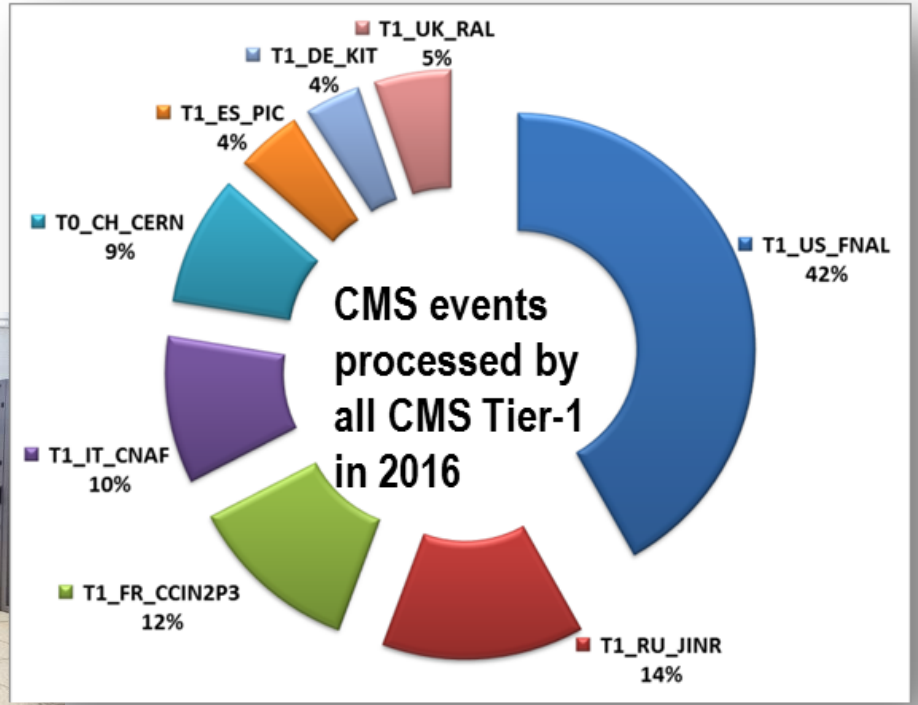
Network infrastructure

Engineering infrastructure

Tier1 center JINR for the CMS experiment

Current configuration

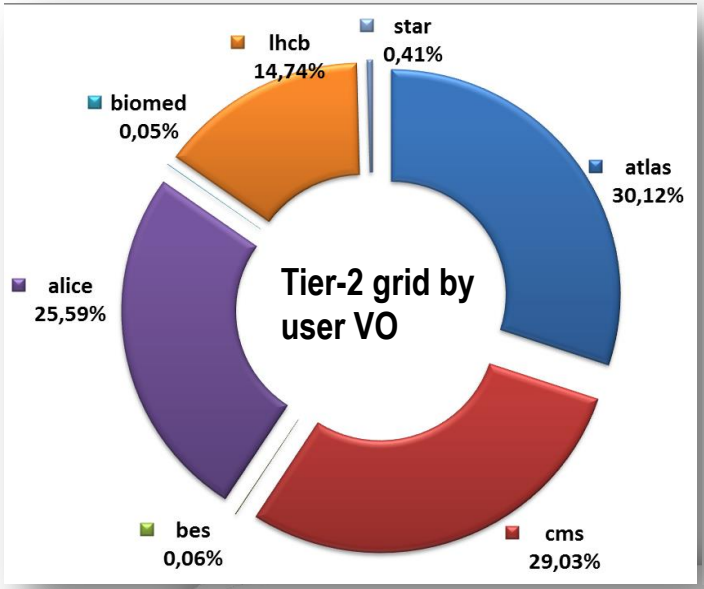
- **3600 cores**
- **5.4 PB tapes**
- **4.5 PB disks**
- **100% R/A**
- **70 TB per day**

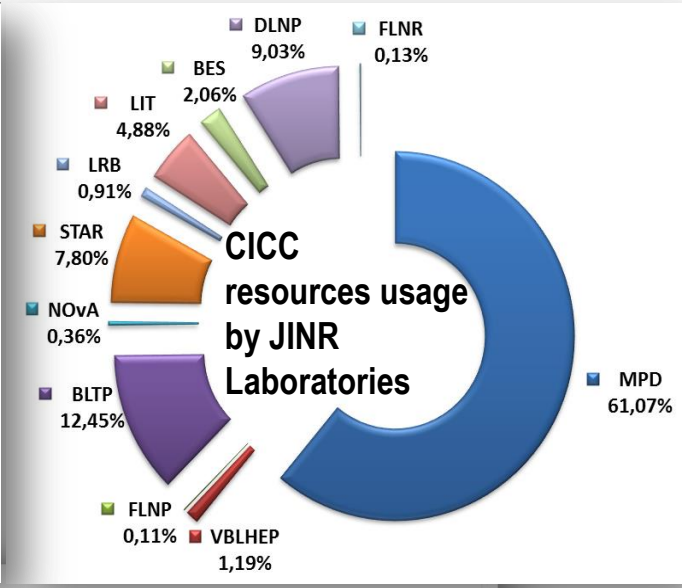


Tier2 & CICC

Experiments

BM@N,
MPD,
CMS,
ATLAS,
ALICE,
LHCb,
COMPASS,
PANDA,
CBM,
STAR,
NOvA,
BESIII,
DIRAC,
OPERA
NEMO
Mu2e,
NUCLON,
HONE,
FUSION
BIOMED

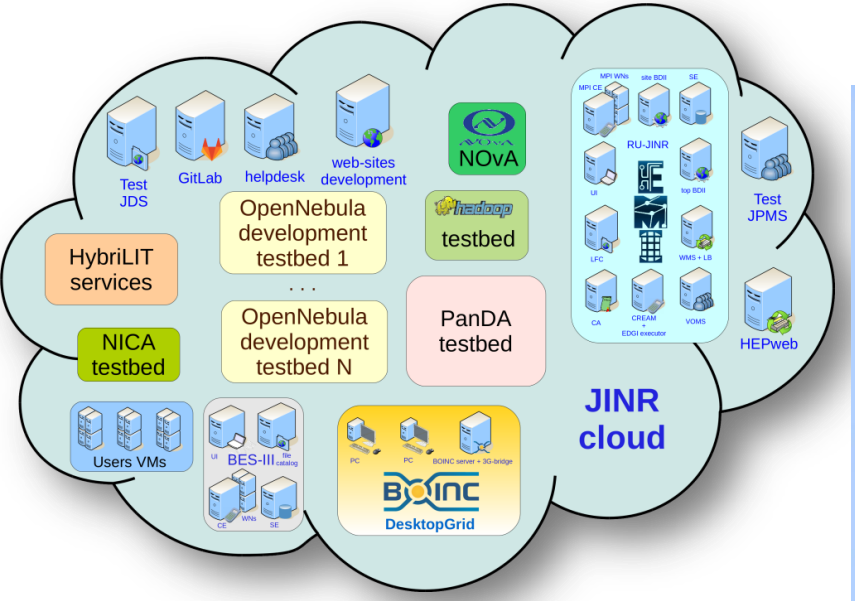




Tier-1 and Tier-2:
JINR in Worldwide LHC Computing Grid

JINR Cloud

Services and testbeds currently deployed in JINR cloud



Geographical locations of the partner cloud infrastructures from JINR Member States which provide part of their computational resources being integrated into the JINR cloud

A spectrum of the tasks solved with the JINR in 2016:

- PanDA testbed was deployed for PanDA software validation and extensions development for ATLAS and COMPASS experiments;
- DIRAC-based testbed for BESIII experiment;
- a set of VMs of NOvA experiment users for analysis and software development;
- NICA testbed for grid middleware evaluation for NICA computing model development;
- a standalone Spark instance for Machine Learning and BigData analysis.

Heterogeneous computing cluster HybriLIT

•**Purpose:** The HybriLIT can be characterized as a modular heterogeneous High Performance Computing (HPC) complex answering to three basic tasks:

●●Task 1: Design and implementation of parallel software for computing intensive research;

●●Task 2: Porting to the cluster open software packages, numerical libraries, and programs which are already tuned for hybrid architectures;

●●Task 3: Development of new mathematical methods and parallel algorithms adapted to heterogeneous architectures.

In total, the cluster includes:

— 252 CPU-cores,
— 77184 CUDA-cores;
— 182 PHI-cores;
— 1920 GB RAM;
— 60 TB HDD.

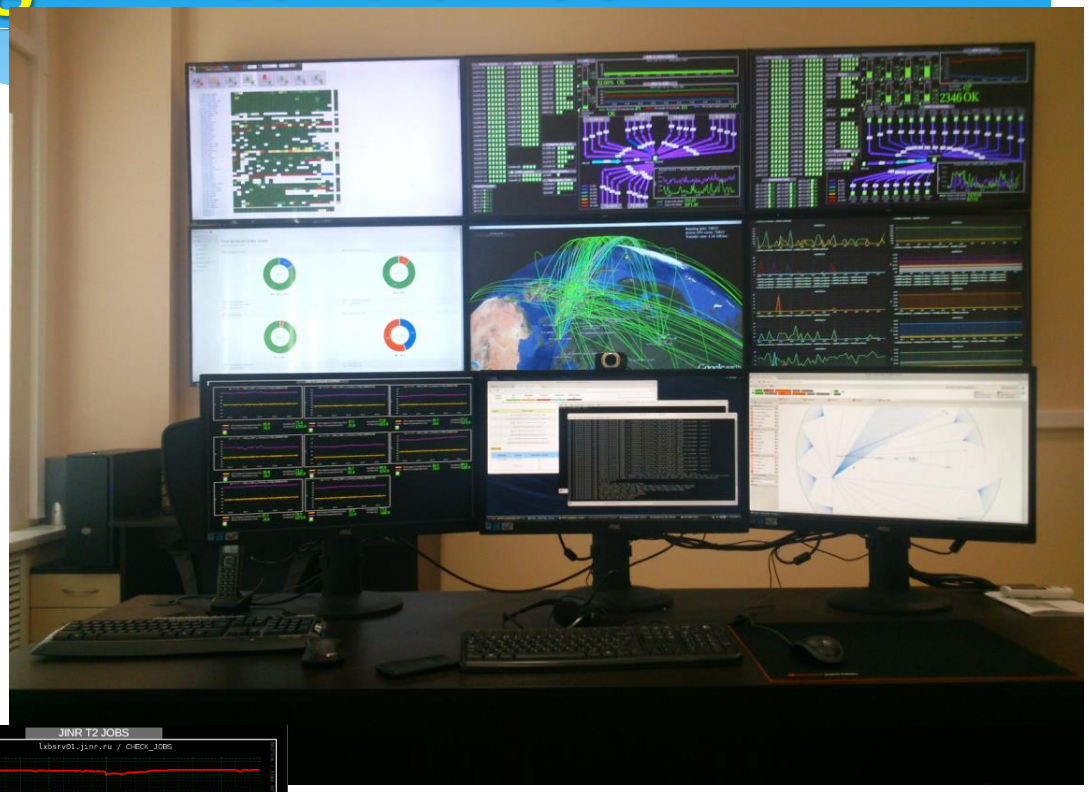
Total performance:
142 TFlops for operations with single precision;
50 TFlops for operations with double precision.

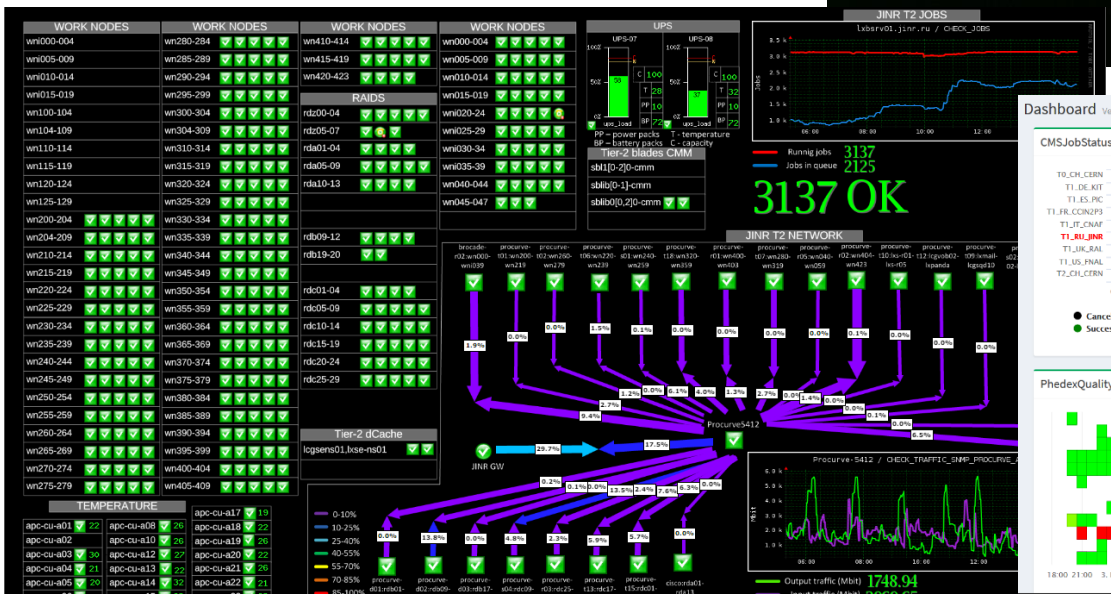
At present, the cluster is used by:

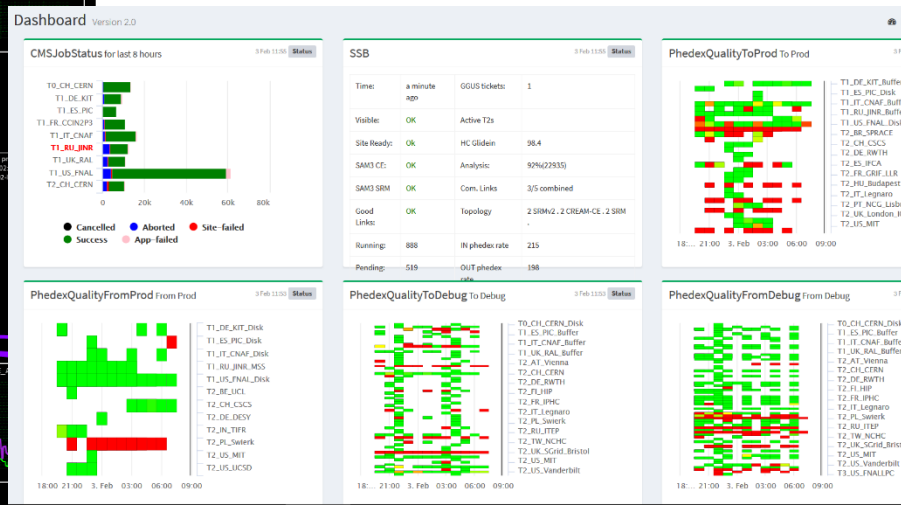
120 persons, including **26** from JINR Member States and **19** – Russia Universities.

Monitoring and Control room

For a robust performance of a complex it is necessary to monitor the state of all nodes and services- from the supply system to the robotized tape library. System allows one, in a real time mode, to observe the whole computing complex state and send the system alerts to users via e-mail, sms, etc. **690 elements are under observation 3497 checks in real time**







Worldwide LHC Computing Grid Project (WLCG)

The primary goal of the WLCG project is to create a global infrastructure of regional centers for processing, storage and analysis of data of the LHC physical experiments.
The grid-technologies are a basis for constructing this infrastructure.

A protocol between CERN, Russia and JINR on participation in the LCG project was signed in 2003. MoU about participation in the WLCG project was signed in 2007.

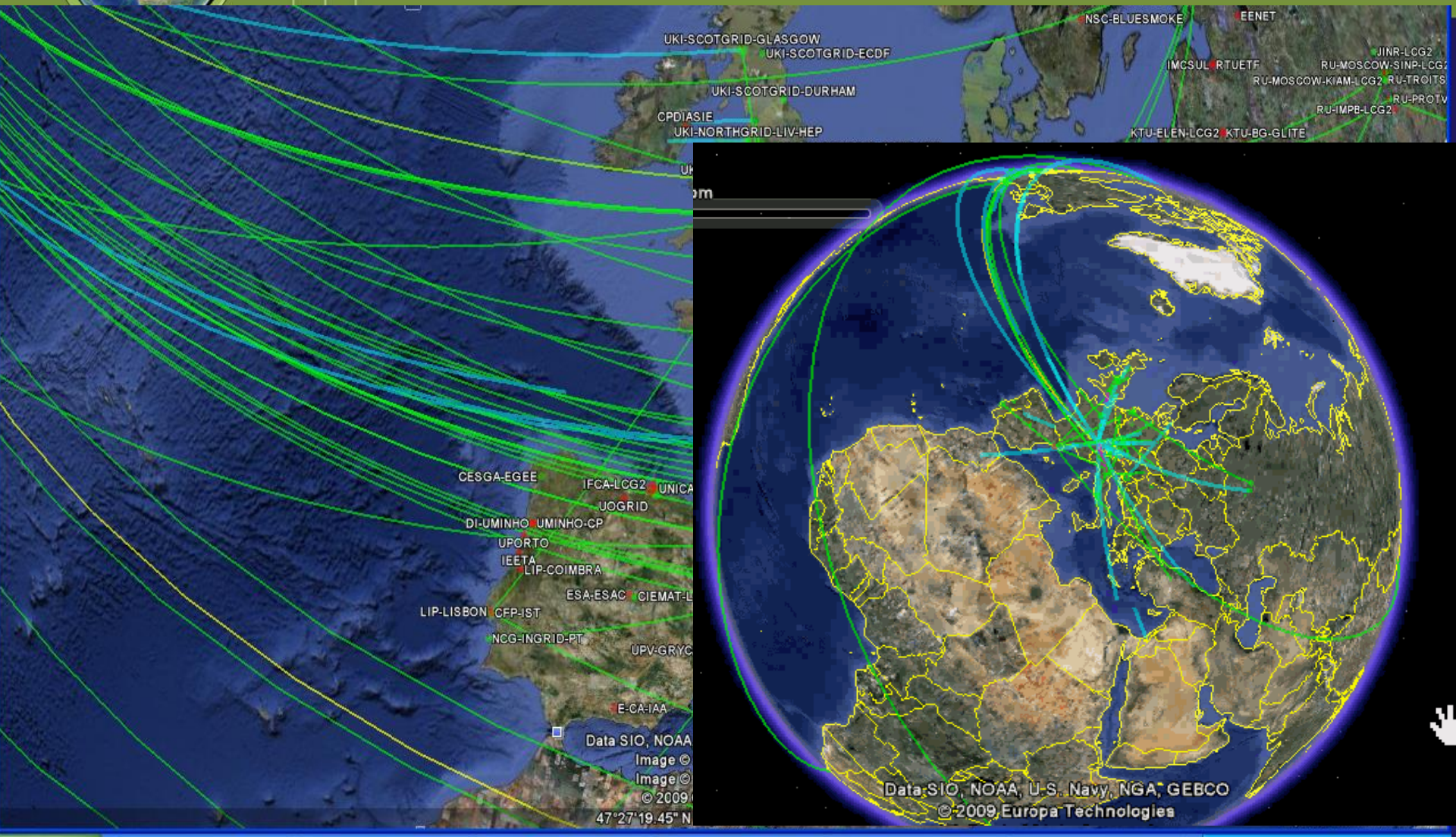
Tasks of the Russian centers and JINR within WLCG :

- **storage and data management for WLCG**
- **Introduction of WLCG services for experiments**
- **Development of WLCG monitoring systems**
- **Development of simulation packages for experiments**
- **Creation of a Tier1 center in Russia**

JINR activity at WLCG project

* Participation in development of software for ATLAS, ALICE, CMS
* Development WLCG Dashboard (WLCG Google Earth monitoring, Global data transfer monitoring, Local and global Monitoring of Tier3 centers, xROOTd Monitoring,…)
* NOSQL storage testing (Hadoop, ElasticSearch,…)
* Development of DDM (Deletion service), AGIS for ATLAS
* BigPanda development (TITAN, IT4, COMPASS,…)
* GENSER & MCDB
* Tier1 center for CMS

WLCG Google Earth Dashboard

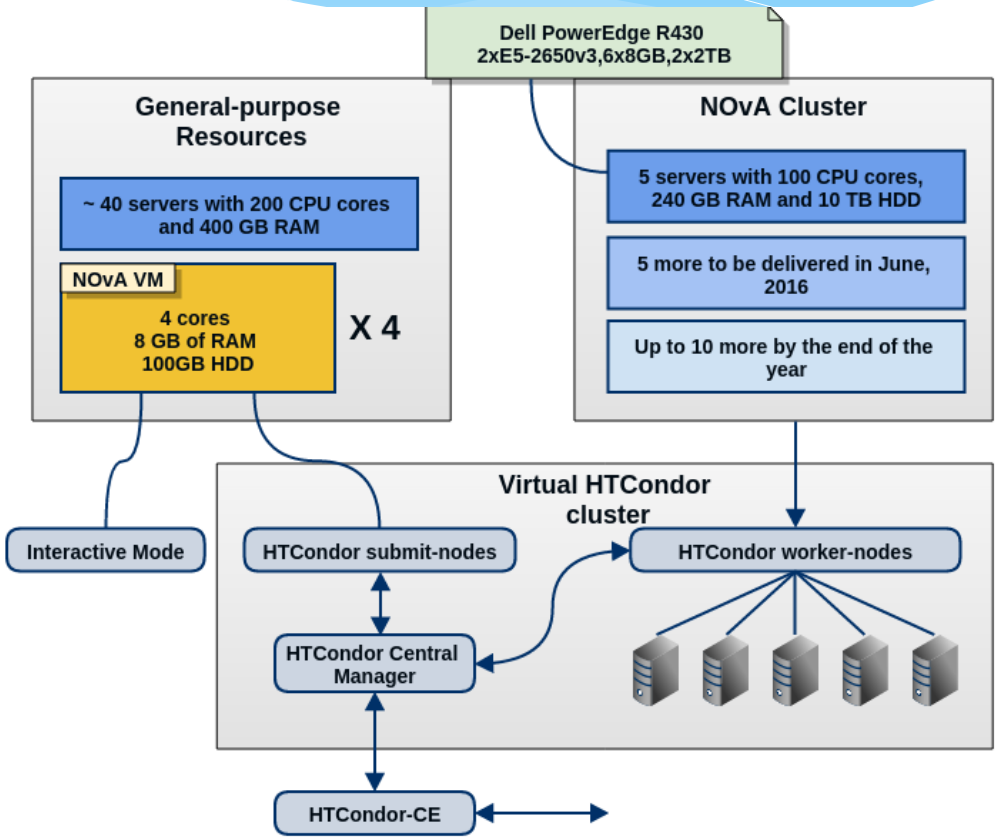


Computing Support for Neutrino Projects

NOvA (Fermilab, USA) is the first neutrino experiment actively using JINR Cloud:

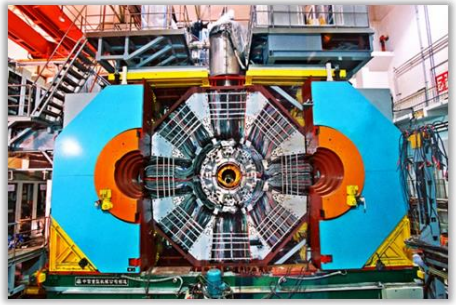
- ✓ Computing support team was formed including physicists and IT specialists

These resources may also be used by other future experiments at Fermilab, such as DUNE and mu2e.



Reactor neutrino experiments Daya Bay and JUNO also showed its interest in using JINR cloud resources. At the moment the experiments' tasks and required computing capacities are being discussed.

BES-III Distributed Computing



GRID.JINR.ru
100
30 TB
CLOUD.JINR.ru
5

GRID.INFN-Torino.it
200
30 TB
CLOUD.TORINO.it
101

GRID.INFN-ReCas.it
50
30 TB

CLOUD.IHEP-OPENSTACK.cn
96
66 TB

CLOUD.IHEP-OPENNEBULA.cn
178
126 TB

CLUSTER.WHU.c
120
39 TB

CLUSTER.UMN.u
768
50 TB

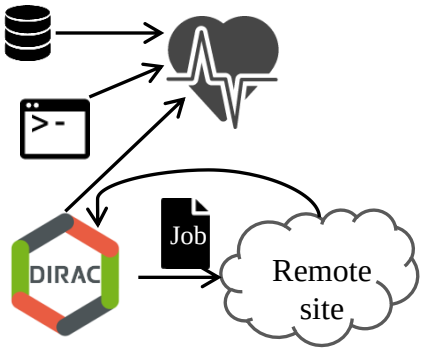
CLUSTER.USTC.c
200
24 TB

What have been done in computing:

- Grid monitoring system developed from scratch
- JINR cloud was integrated in BES-III infrastructure
- 6 % of all jobs was done in JINR during the past year

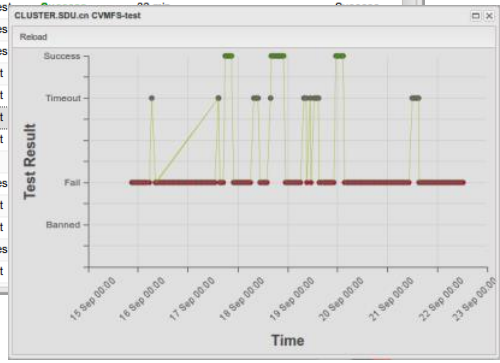
Planning to continue participate in BES-III experiment by:

- Improving monitoring
- Research on clouds in grid
- Providing storage and CPU cores



CE availability monitoring

Site	Test	Result	Received ago	Description
CLUSTER.SDU.cn	WMS-test	Success	43 min	Remote call
GRID.INFN-Torino.it	WMS-test	Success	8 min	Remote call
CLOUD.IHEP-OPENSTACK.cn	WMS-test	Success	86 min	Remote call
GRID.INFN-ReCas.it	WMS-test	Success	40 min	Remote call
CLOUD.CNIC.cn	WMS-test	Success	44 min	Remote call
CLUSTER.UMN.us	CVMFS-test	Success	37 min	Success
GRID.INFN-Torino.it	CVMFS-test	Success	8 min	Success
CLOUD.IHEP-OPENSTACK.cn	CVMFS-test	Success	30 min	Success
CLOUD.IHEP-OPENNEBULA.cn	CVMFS-tes			
GRID.INFN-ReCas.it	CVMFS-tes			
CLOUD.CNIC.cn	CVMFS-tes			
CLUSTER.UMN.us	BOSS-test			
GRID.INFN-Torino.it	BOSS-test			
GRID.INFN-ReCas.it	BOSS-test			
CLOUD.CNIC.cn	BOSS-test			
CLOUD.AWS.cn	WMS-test			
CLOUD.AWS.cn	CVMFS-tes			
BOINC.IHEP.cn	BOSS-test			
CLOUD.AWS.cn	BOSS-test			
CLUSTER.SDU.cn	CVMFS-tes			
CLUSTER.SDU.cn	BOSS-test			

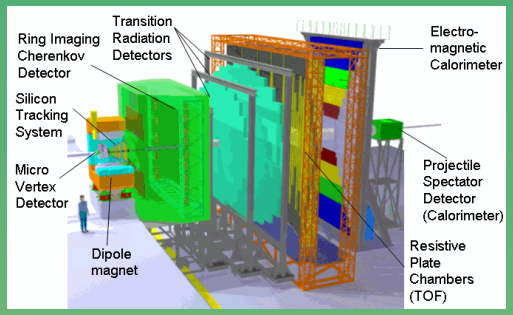


J/ψ→e⁺e⁻ decay selection criteria for Au+Au at 10 AGeV in the CBM experiment

The study of charmonium production is one of the key objectives of the CBM experiment.

Main difficulty of J/ψ→e+e- :

- extremely low yield of *J*/ψ mesons
- low probability (about 6%) of decay
- intense hadron background



The investigation goal is fast and efficient selection of the signal events for J/ψ→e+e- decays reconstruction in the real time experiment

J/ψ reconstruction technique

These criteria are applied successively to a sample formed as a result of applying the preceding criterion.

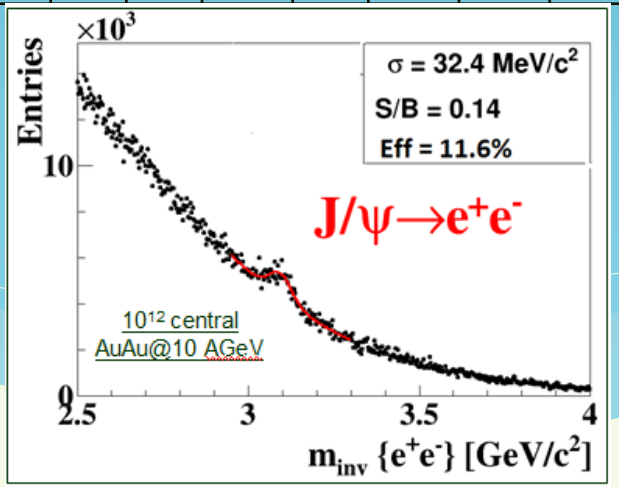
1. Transverse momentum p_t > 1 GeV/c
2. Number of hits in the TRD-track > 2
3. Electron identification in TRD with ANN
4. Limit on the track deviation in the TRD
5. Electron identification in TOF
6. The presence of the Cherenkov ring in the RICH

J/ψ reconstruction technique allows to almost completely suppress the background and accumulate considerable J/ψ meson statistics upon its decay into the dielectron channel using the CBM setup.

Results

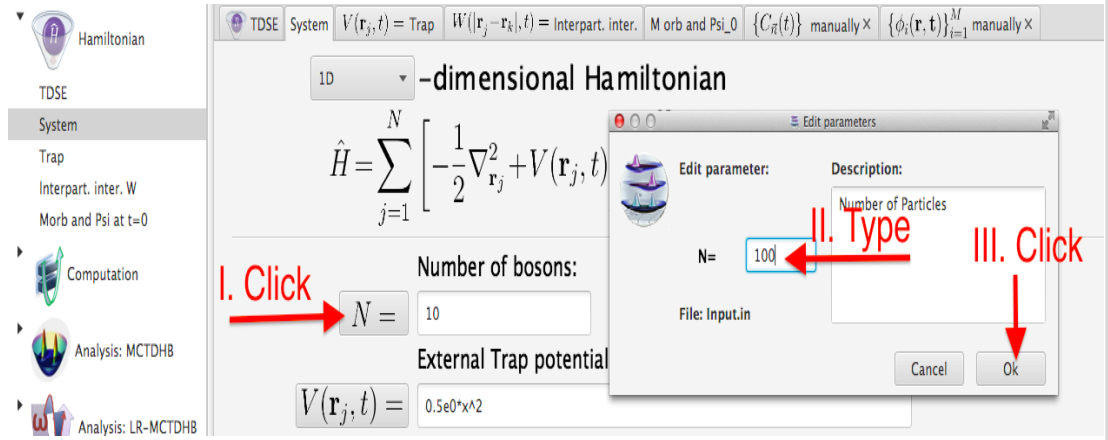
The loss of signal events and the corresponding suppression background in each stage independently

	1	2	3	4	5	6	Total
signal loss (%)	16	18	28.5	7.5	1.1	1.6	56
Bg supp(%)	92	26	99.5	16	54	61.5	99.9



In 2015 year was presented MCTDHB-Lab package – solver of the many-boson Time-Dependent Schrödinger Equation. This solver can be used for theoretical investigations of the highly-non-equilibrium quantum dynamics realized in trapped systems of ultra-cold atoms and molecules.

MCTDHB-Lab is available as a FREE-for-download, cross-platform software with a mouse-click interface



Center for Quantum Dynamics

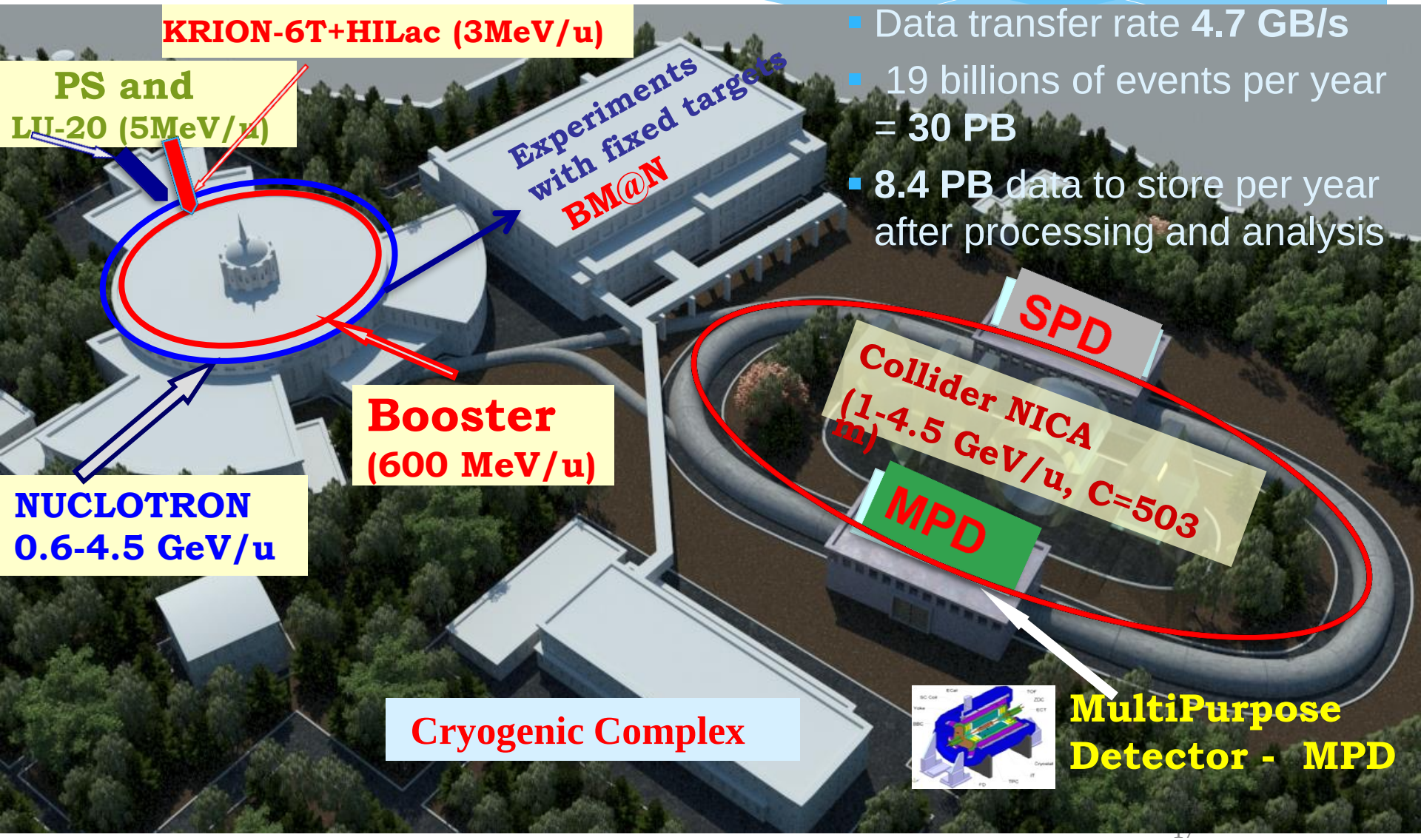
Many-body theory of bosons group Heidelberg, Germany

Heterogeneous Computations **HybriLIT** - team

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

Collider basic parameters:

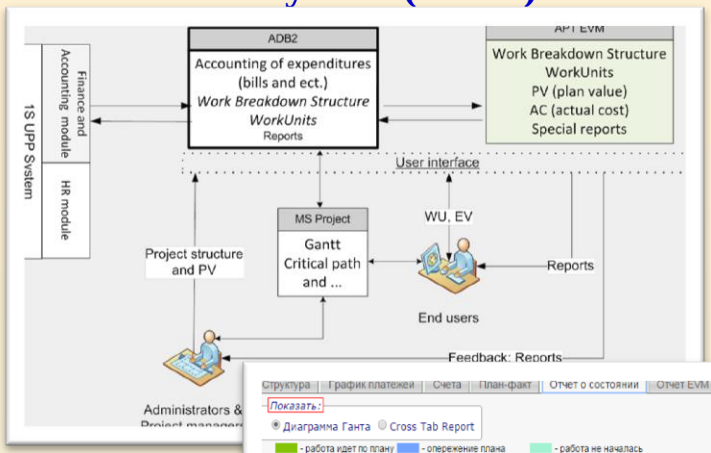
$\sqrt{s_{NN}}$ = 4-11 GeV; *beams: from* p to Au; L~10^{27} cm^{-2} c^{-1} (Au), ~10^{32} cm^{-2} c^{-1} (p)



- 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

Computing for NICA

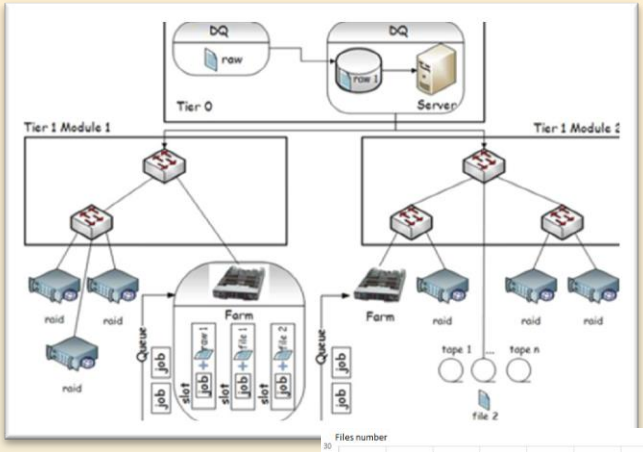
NICA Project Management Information System (PMIS)



NICA PMIS has the following features:

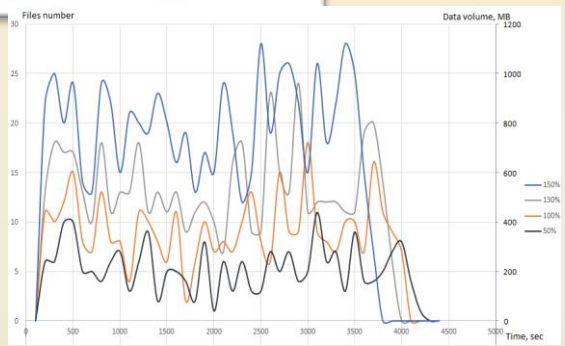
- ❑Control over the project structure (WBS);
- ❑Planning and replanning the project work;
- ❑Versions of the project plans (baselines);
- ❑Monitoring of the implementation of the project in terms of AC (actual payments) and EV (earned value);
- ❑System alerts users via e-mail (for the timely report on the progress of the work);
- ❑Charts by the method of EVM (PV, AC, EV).

Simulation of NICA-MPD-SPD Tiero-Tier1 computing



Data storage and processing scheme of Tier0-Tier1 level

Number of DAQ data files stored on output disk buffer for growing data volumes

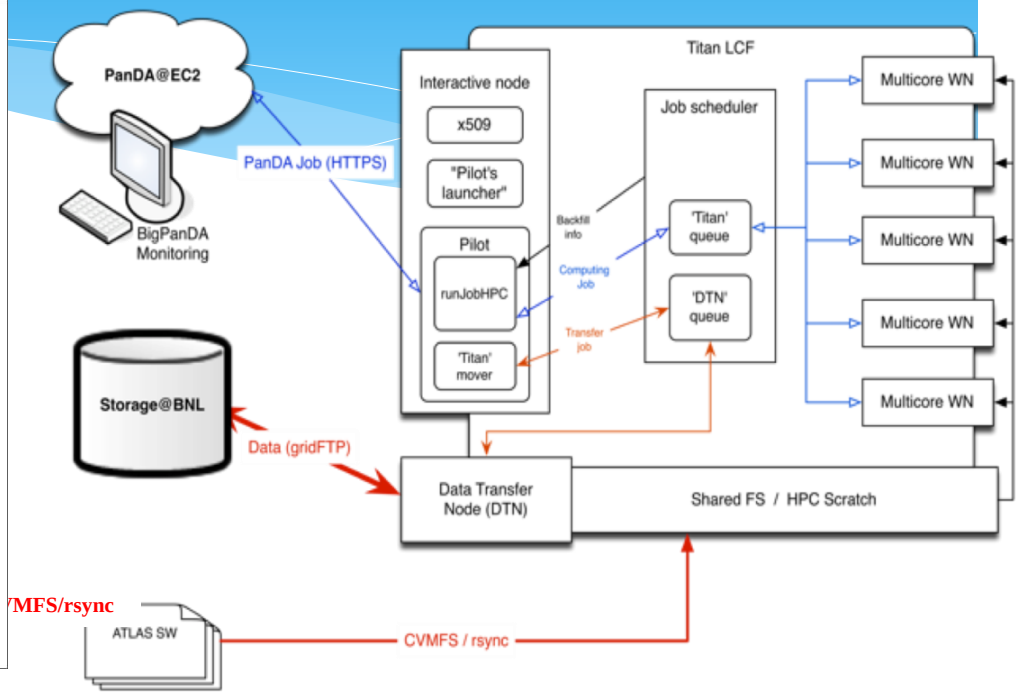


Working at TB scale the NICA MPD-SPD experiments will face with great challenges in distributed computing:

- ❑large increase of CPU and network resources;
- ❑combined grid and cloud access;
- ❑Intelligent dynamic data placement
- ❑distributed parallel computing;
- ❑renewal most of simulation and analysis software codes.

Evolving PanDA for Advanced Scientific Computing

	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		



ATLAS (BNL, UTA), OLCF, ALICE (CERN,LBNL,UTK), LIT JINR:

- adapt PanDA for OLCF (Titan)
- reuse existing PanDA components and workflow as much as possible.
- PanDA connection layer runs on front-end nodes in user space. There is a predefined host to communicate with CERN from OLCF, connections are initiated from the front-end nodes
- SAGA (a Simple API for Grid Applications) framework as a local batch interface.
- Pilot (payload submission) is running on HPC interactive node and communicating with local batch scheduler to manage jobs on Titan.
- Outputs are transferred to BNL T1 or to local storage

BigPanDA. Generalizing PanDA. Beyond Grid and ATLAS

PanDA

COMPASS

Google Cloud Platform

Alpha Magnetic Spectrometer AMS-02 Europe • Asia North America

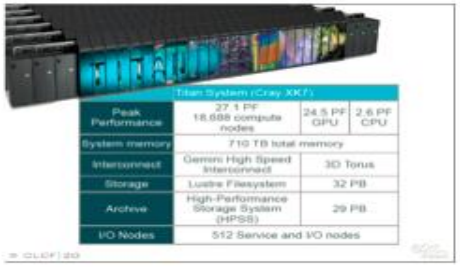
ALICE

ATLAS

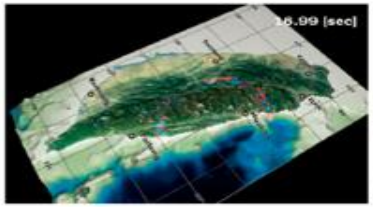
NICA

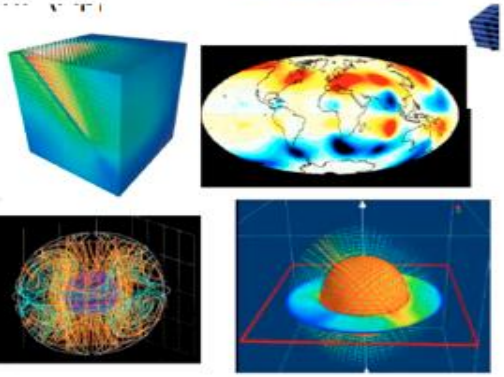
nectar

amazon web services | EC2



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	

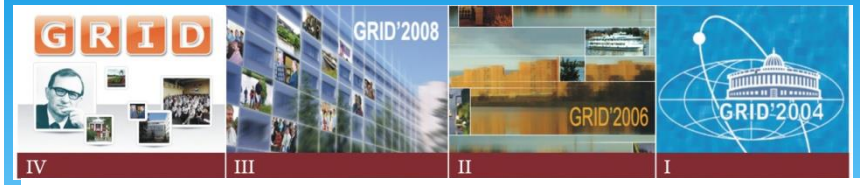




Projects in framework Distributed computing

- Worldwide LHC Computing Grid (WLCG)
- RDIG Development
- Tier1 Center in Russia (NRC KI, LIT JINR)
- Projects at CERN
- CERN-RFBR project “Global data transfer monitoring system for WLCG infrastructure”
- BMBF grant “Development of the grid-infrastructure and tools to provide joint investigations performed with participation of JINR and German research centers”
- Development of grid segment for the LHC experiments” with South Africa;
- Development of grid segment at Cairo University and its integration to the JINR GridEdu
- JINR - FZU AS Czech Republic Project “The grid for the physics experiments”
- NASU-RFBR project “Development and implementation of cloud computing technologies on grid-sites at LIT JINR and BITP for ALICE experiment”
- JINR-Romania cooperation Hulubei-Meshcheryakov programme
- JINR-China cooperation (BES-III)
- Cooperation with Armenia, Azerbaijan, Belarus, Bulgaria, Georgia, Kazakhstan, Moldova, Mongolia, Poland, Slovakia,

LIT traditional conferences



Distributed Computing and Grid-technologies in Science and Education

The International Conference
Mathematical Modeling and Computational Physics, 2015

Mathematics. Computing. Education

LIT schools

JINR / CERN

GRID AND ADVANCED INFORMATION SYSTEMS

International Conference-School for Young Scientists
"Modern Problems of Applied Mathematics & Computer Science"
August 22 - 27, 2012, Dubna, Russia

MPAMCS 2012

INFORMATION

DIGITAL LIBRARIES:
ADVANCED METHODS AND TECHNOLOGIES
DIGITAL COLLECTIONS

NEC'2017

26 th International Symposium on Nuclear Electronics & Computing

25-29 September 2017

Montenegro, Budva, Becici

NEC2017.jinr.ru

50 Years – LIT JINR