JINR Cloud Service: Status and Perspectives

N. Balashov[1], A. Baranov[1], **V. Korenkov[1,2]**, N. Kutovskiy[1,2], A. Nechaevskiy[1], R. Semenov[1,2]

[1]Joint Institute for Nuclear Research, Russia, Dubna.
[2]Plekhanov Russian University of Economics, Russia, Moscow.

The conference «Cloud computing. Education. Research. Development»
Moscow, 3 December 2015

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.

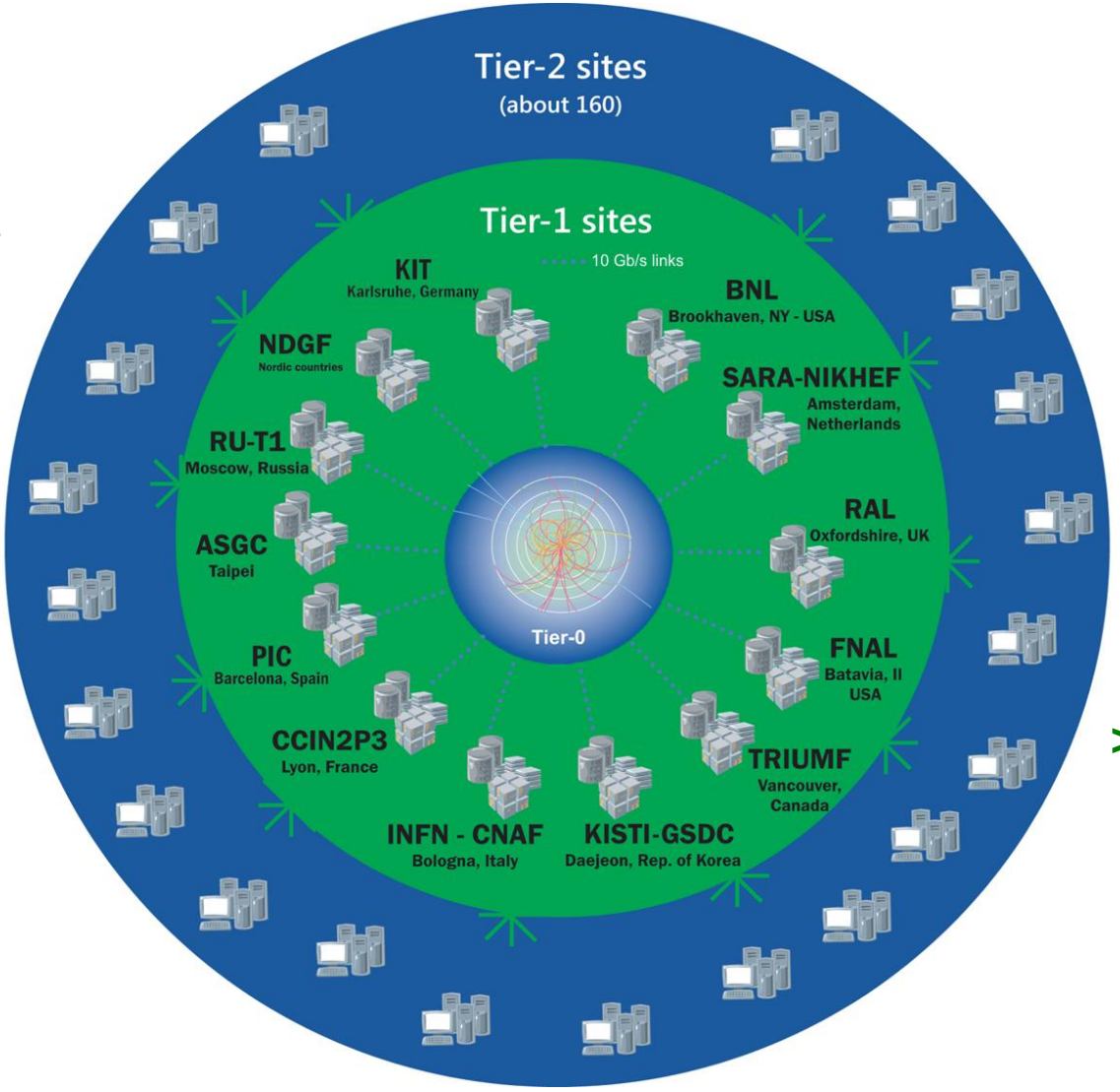
Nowadays, every large-scale project will fail without using a distributed infrastructure for data processing.

The Worldwide LHC Computing Grid

Tier-0 (CERN): data recording, reconstruction and distribution

Tier-1: permanent storage, re-processing, analysis

Tier-2: Simulation, end-user analysis



nearly 170 sites, 40 countries

~500'000 cores

500 PB of storage

> 2 million jobs/day

10-100 Gb links

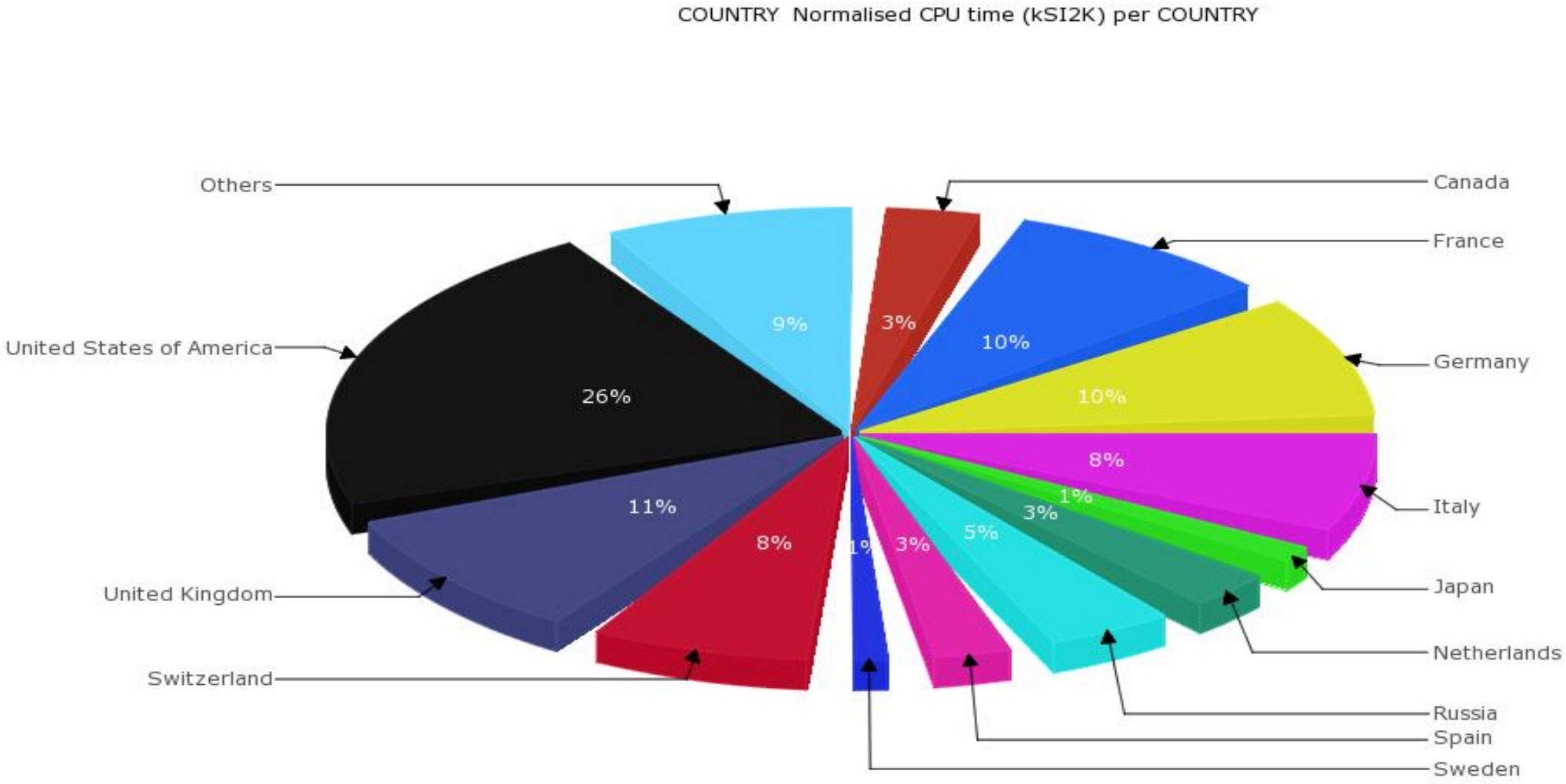
WLCG:
An International collaboration to distribute and analyse LHC data

Integrates computer centres worldwide that provide computing and storage resource into a single infrastructure accessible by all LHC physicists

CERN Computer Centre (Tier-0)

Country Normalized CPU time 2014-2015

A 'EGI View' : / normcpu / 2015:5-2015:10 / COUNTRY-VO / lhc (x) / GRBAR-LIN / l



All Country - 41,242,621,048 **Russia-** 1,981,877,204
Job 1,246,002,966 43,334,211

LIT JINR Fundamentals

User policies

Computing infrastructure

Mathematical and software support

Corporative information system

Training, education and user support

Multifunctional Information&Computing Complex

Tier-2 level grid-system to support LHC experiments (ATLAS, ALICE, CMS, LHCB), FAIR (PANDA) and other large-scale experiments and projects within the global grid-infrastructure

Tier1 CMS experiment on the Large Hadron Collider (LHC)

Heterogeneous computer complex for high-efficiency calculations

Cloud environment

Всего сотрудников – 324 из 18 стран мира, 30 –докторов наук, 65-кандидаты наук,
В 2014 году опубликовано: 270 статей в реферированных журналах, 2 монографии, 112 статей в трудах конференций

JINR Computing Complex (Tier2)

CICC comprises 2800 Cores
Disk storage capacity 2500 TB

Availability and Reliability = 99%

30 million Jobs and
840 millions normalized CPU time
were executed in 2010-2015

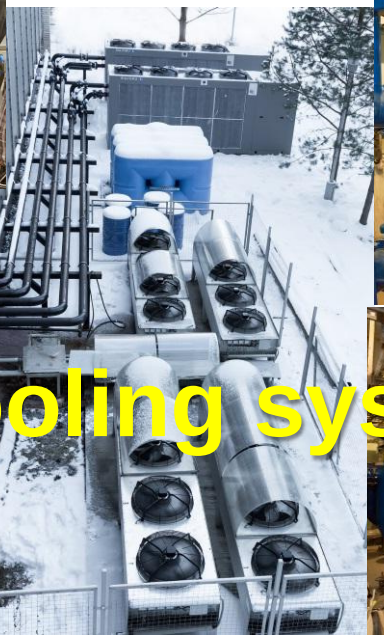
Year 2015 numbers of jobs and CPU time run:

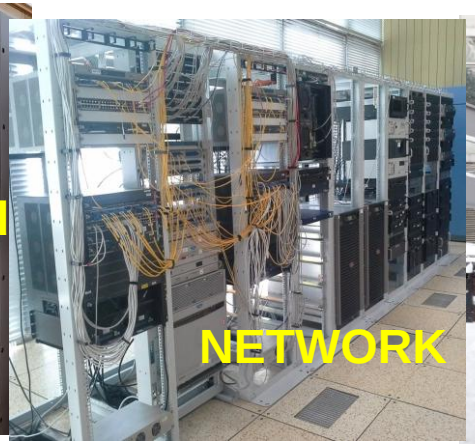
Tier2 total -	**4,829,032**	**181,000,104**
ATLAS -	3,263,835	63,560,916
CMS -	918,688	37,069,196
ALICE -	492,459	55,805,316

Local Users - 940 Jobs – 252,344

JINR Tier1 Center

March 2015 – CMS Tier1 Inauguration
LHCOPN – 10Gbps, 3000 cores (~ 50 kHS06),
5 PB tapes (IBM TS3500), 3.4 PB disk
Close-coupled, chilled water cooling InRow
Hot and cold air containment system
MGE Galaxy 7000 – 2x300 kW energy efficient solutions 3Ph power protection with high adaptability
Annual increase: 15 kHS06, 2 PB tapes, 1 PB disk



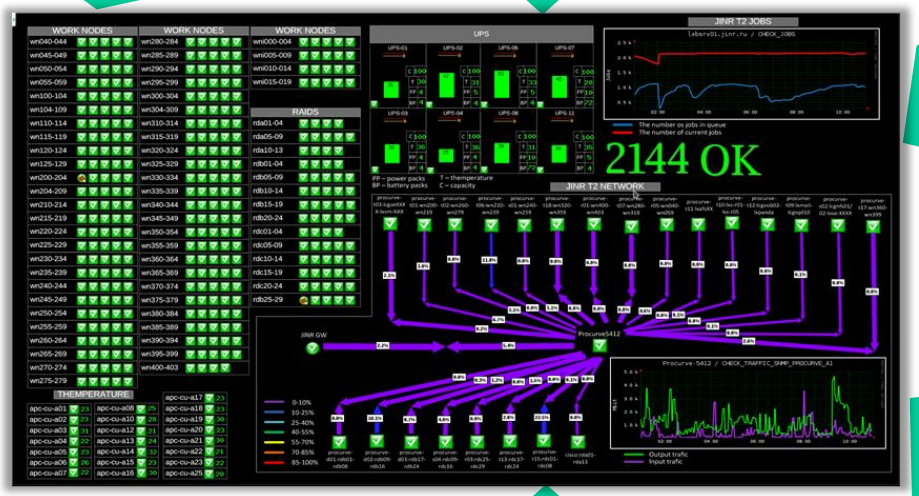


Monitoring System

The monitoring system of the JINR Computing Complex has been developed and put into exploitation.
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.

Storage serves

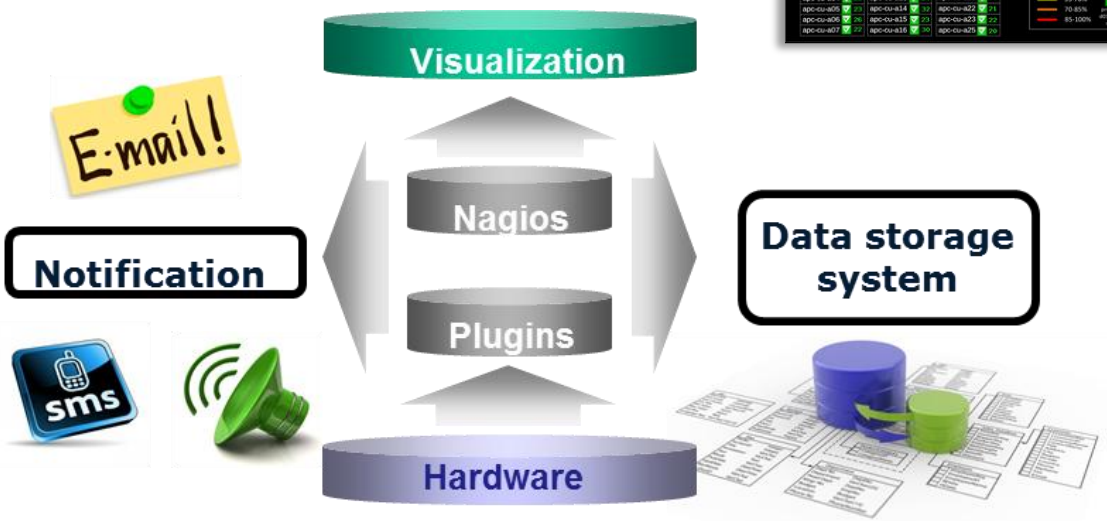
Computing servers



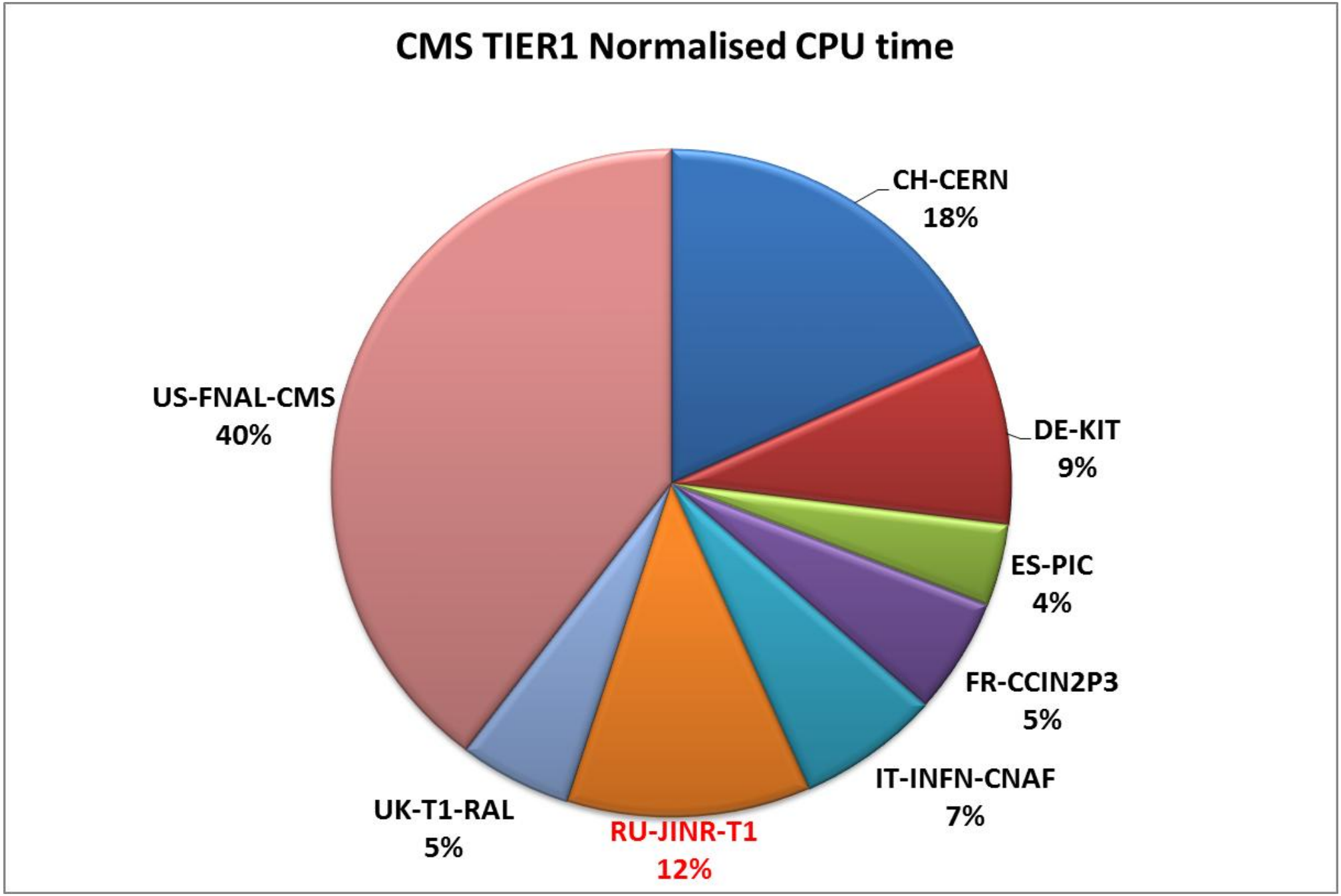
UPS

The cooling system

The tape library



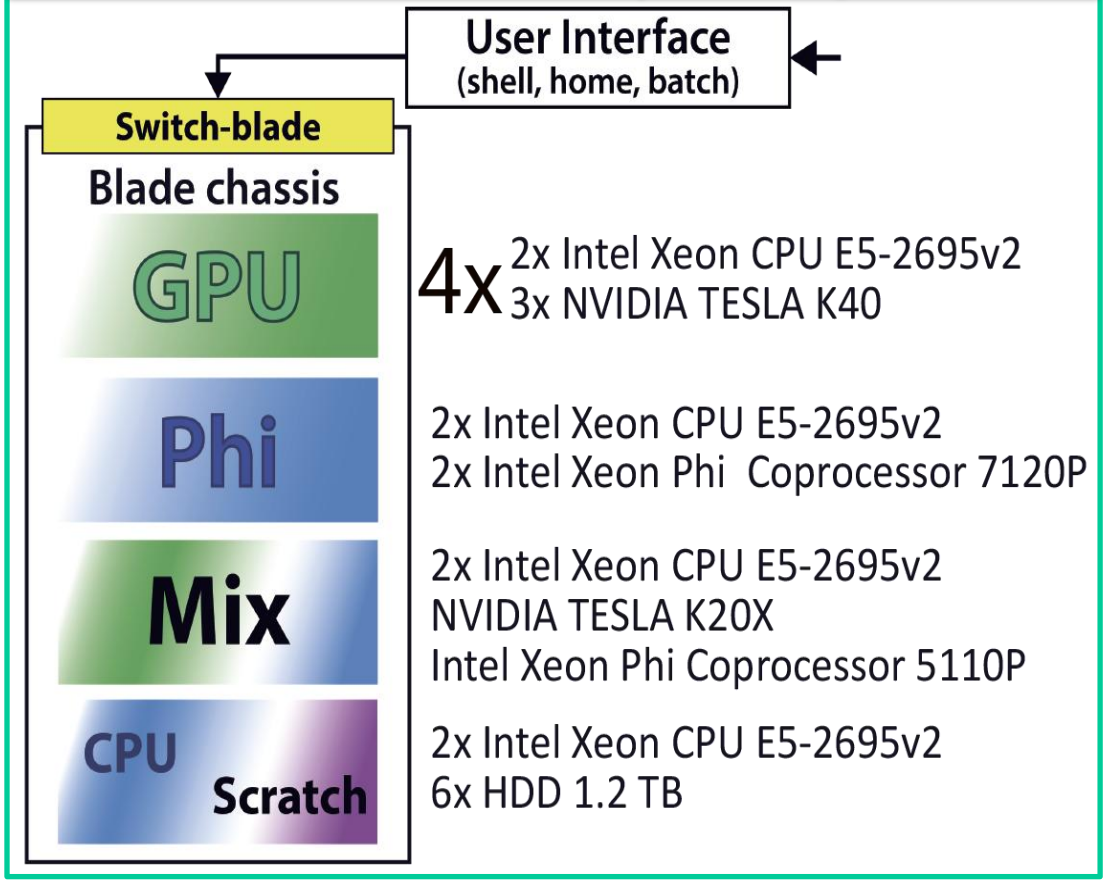
CMS Tier1 Normalized CPU time Sep-Nov 2015



HybriLIT: heterogeneous computation cluster

Current Status 2015:

Total performance: single precision 75 TFLOPS
double precision 30 TFLOPS



Nearest Plans 2015:

Total performance: +35 TFLOPS

- **Hardware upgrades:**

2 X GPU

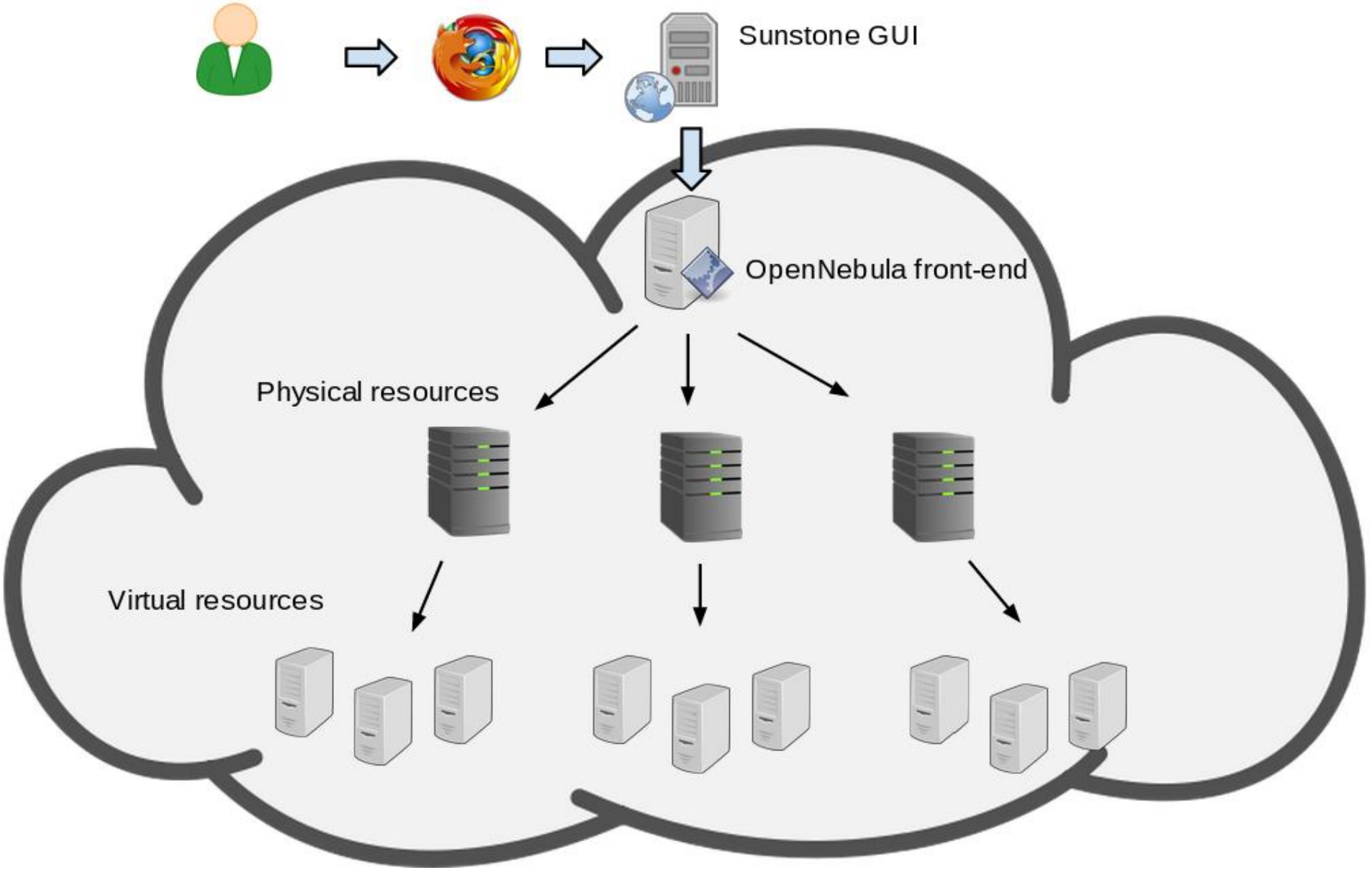
Blades with 4 NVIDIA TESLA K80

- **Creating of polygon for training courses on Parallel Computing Technologies**

Currently the **total number** of users comprise **95** people:
26 are from JINR member-countries: *Armenia, Bulgaria, Mongolia, Romania, Slovakia, Ukraine*, etc.
19 people are from the following universities of Russia: *MSU, SPSU, PFUR and "Dubna" University.*

hYBRI

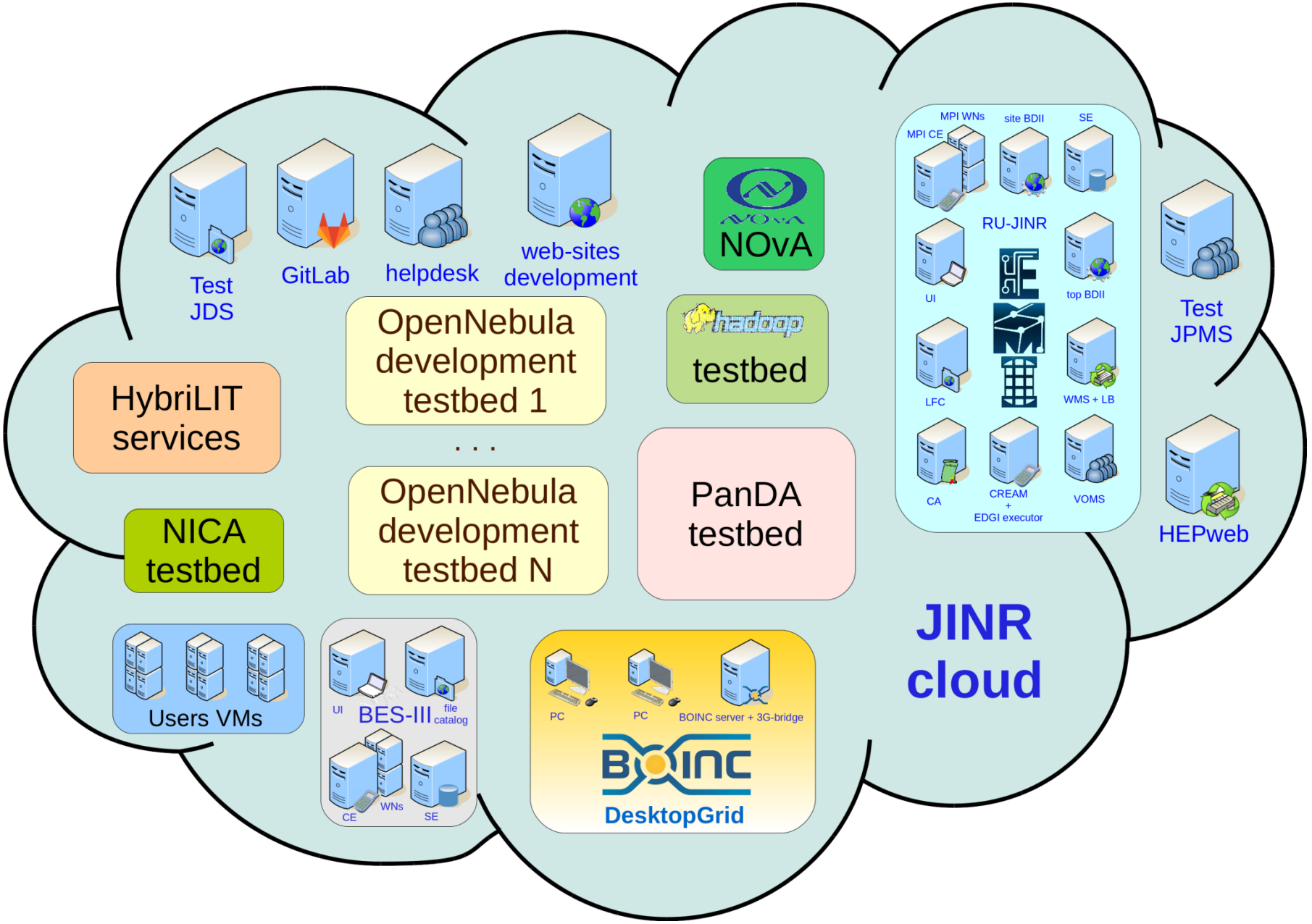
JINR Cloud Service



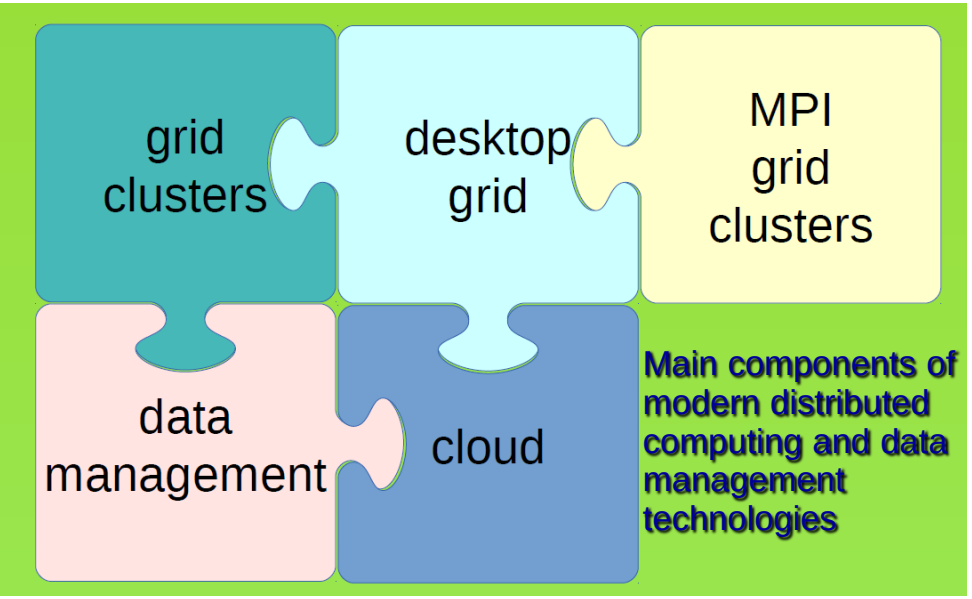
Activities

- Developers
  - development, testing and debugging various apps in various environments.
- System administrators
  - testing and studying specifics of installation and operation of new apps or testing updates
- Users
  - Physicists – major workload source

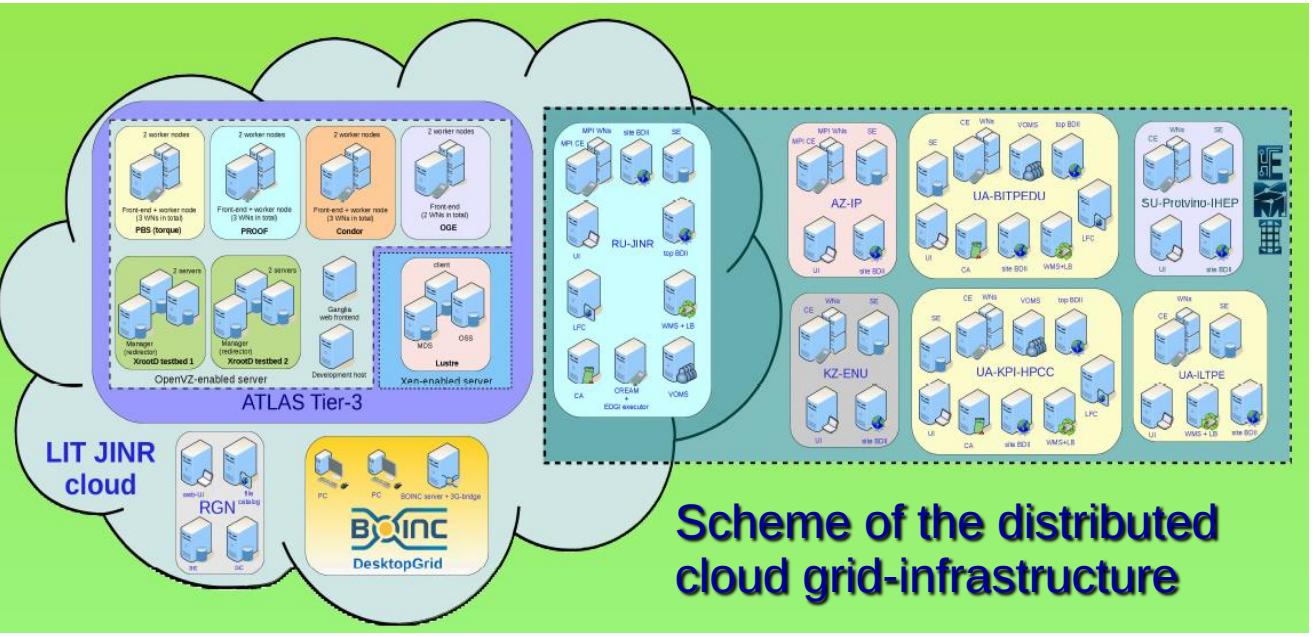
JINR Cloud Utilization



JINR distributed cloud grid-infrastructure for training and research

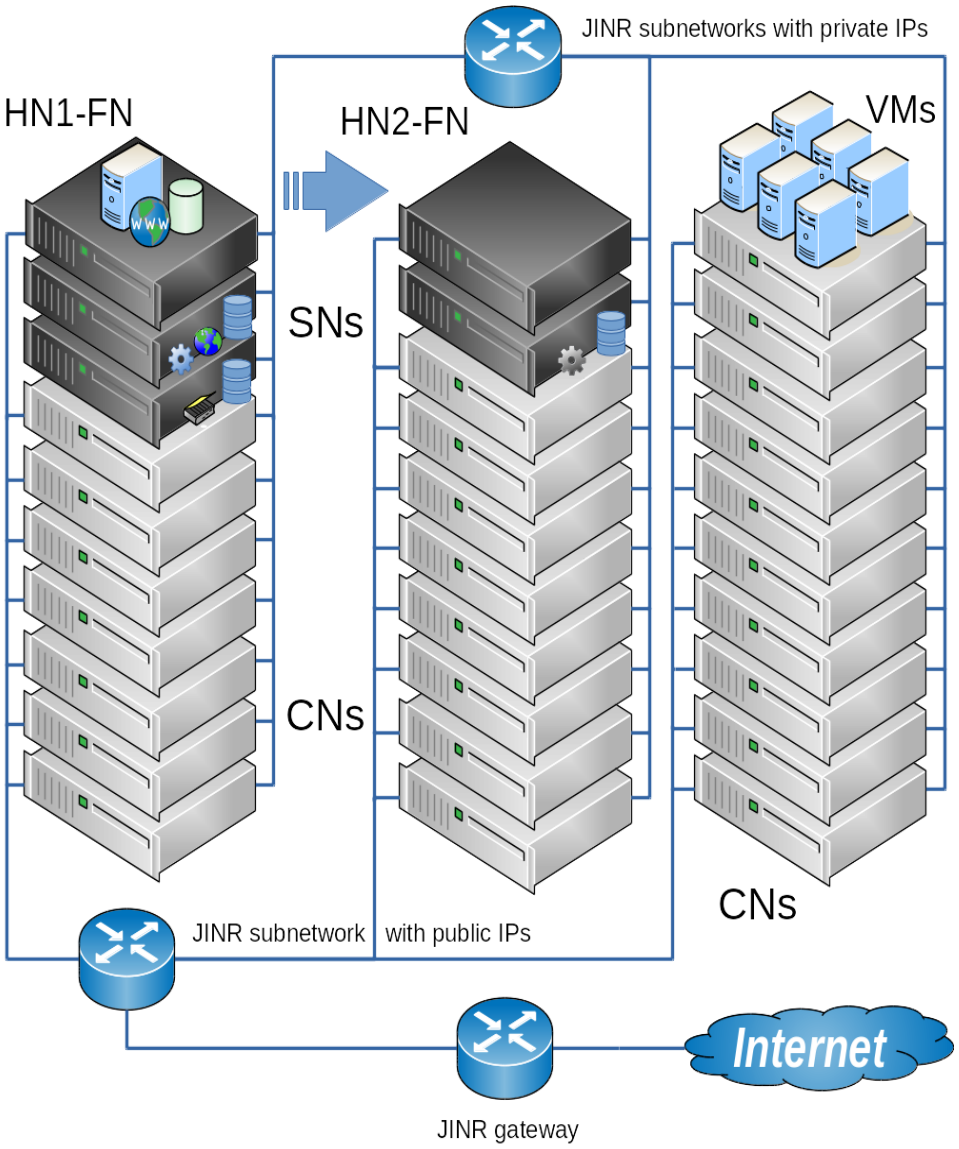


There is a demand in special infrastructure what could become a platform for training, research, development, tests and evaluation of modern technologies in distributed computing and data management.
Such infrastructure was set up at LIT integrating the JINR cloud and educational grid infrastructure of the sites located at the following organizations:



- Institute of High-Energy Physics (Protvino, Moscow region),
- Bogolyubov Institute for Theoretical Physics (Kiev, Ukraine),
- National Technical University of Ukraine "Kyiv Polytechnic Institute" (Kiev, Ukraine),
- L.N. Gumilyov Eurasian National University (Astana, Kazakhstan),
- B.Verkin Institute for Low Temperature Physics and Engineering of the National Academy of Sciences of Ukraine (Kharkov,Ukraine),
- Institute of Physics of Azerbaijan National Academy of Sciences (Baku, Azerbaijan)

Architecture: HA-setup



- Two physical servers (HN1-FN, HN2-FN) with DRBD and Heartbeat
- HN1-FN and HN2-FN are connected to different UPSes
- OpenNebula front-end (FN) deployed on OpenVZ CT running on DRBD partition
- All VM images and data are stored on distributed network file system based on LizardFS with automatic data replication
- CNs are connected to pub and priv nets, SNs - priv

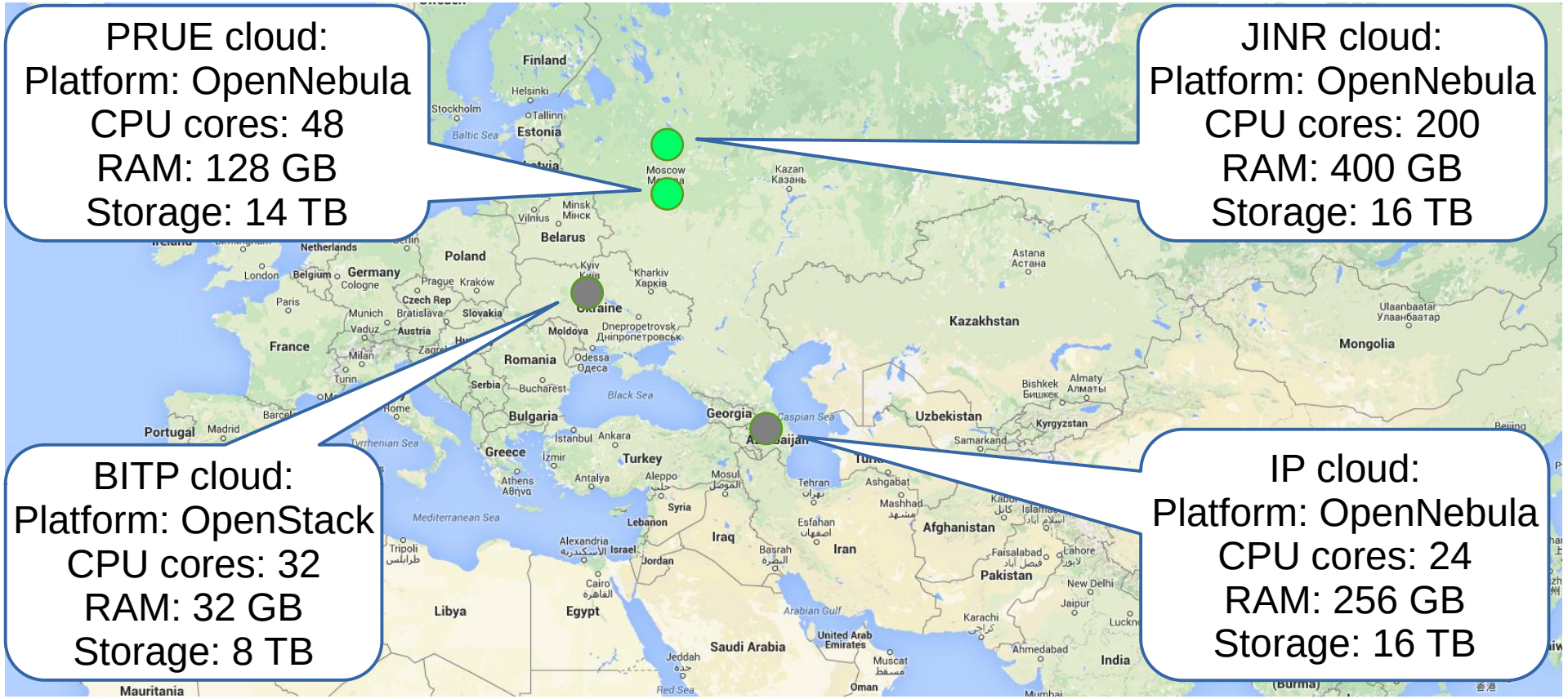
Cloud Storage

- Distributed
- Scalable
- Fault-tolerant
- Highly available

VMs Images Storage

Access, sync and share files

Clouds Integration



- To join resources for solving common tasks as well as to distribute a peak load across resources of partner organizations
- JINR cloud integration with clouds of partner organizations:
  - Institute of Physics of Azerbaijan National Academy of Sciences – IP (Baku, Azerbaijan)
  - Bogolyubov Institute for Theoretical Physics of the National Academy of Sciences of Ukraine – BITP (Kiev, Ukraine)
  - Plekhanov Russian University of Economics – PRUE (Moscow, Russia)
  - EGI Federated cloud

HNSciCloud H2020 PCP Project

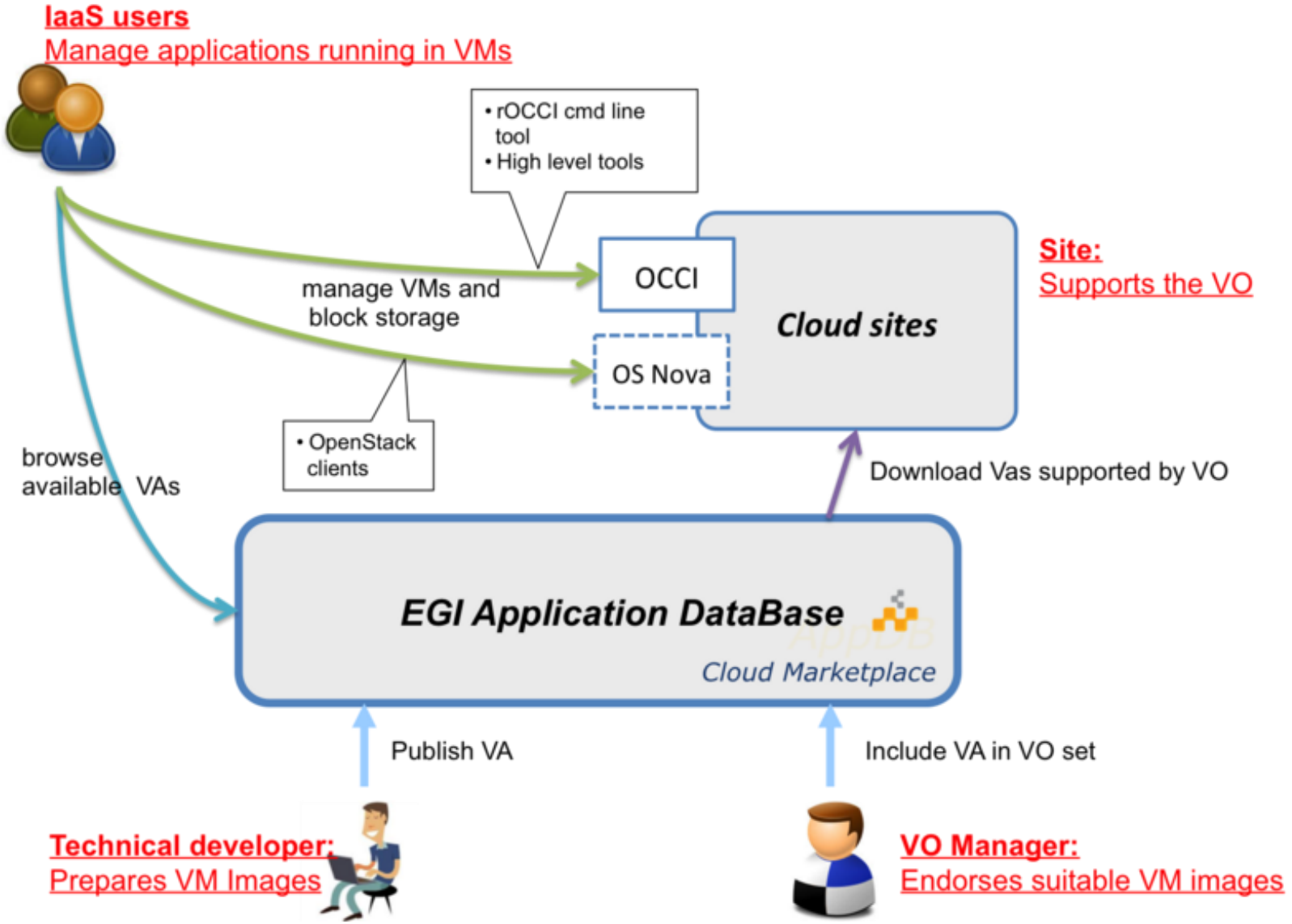
To procure innovative IaaS cloud services integrated into a hybrid cloud model

- Commercial cloud services
- European e-Infrastructures
- In-house IT resources

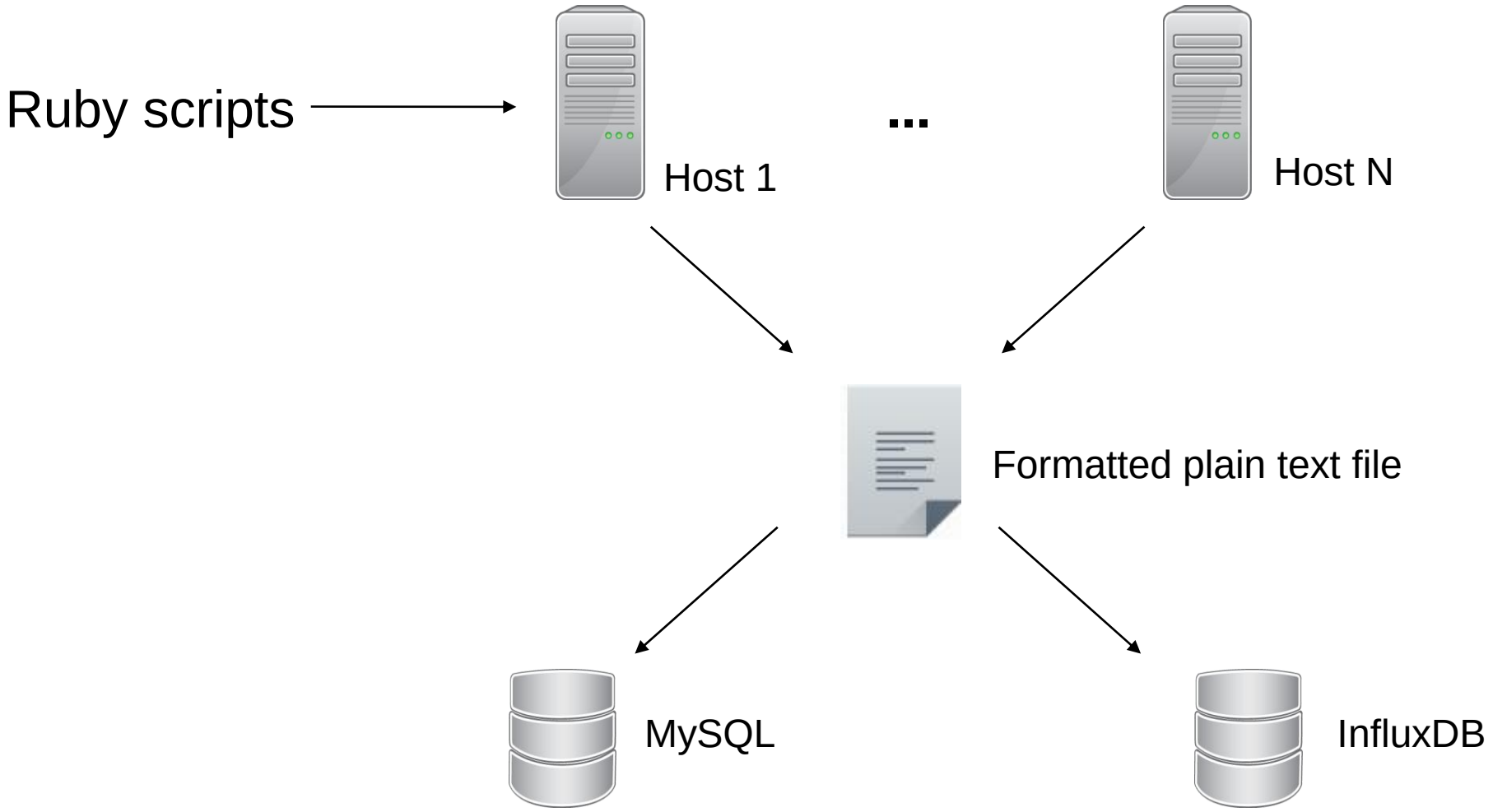
Procured services will be made available to end-users from many research communities

EGI Federated Cloud

"A grid of clouds"



Monitoring System

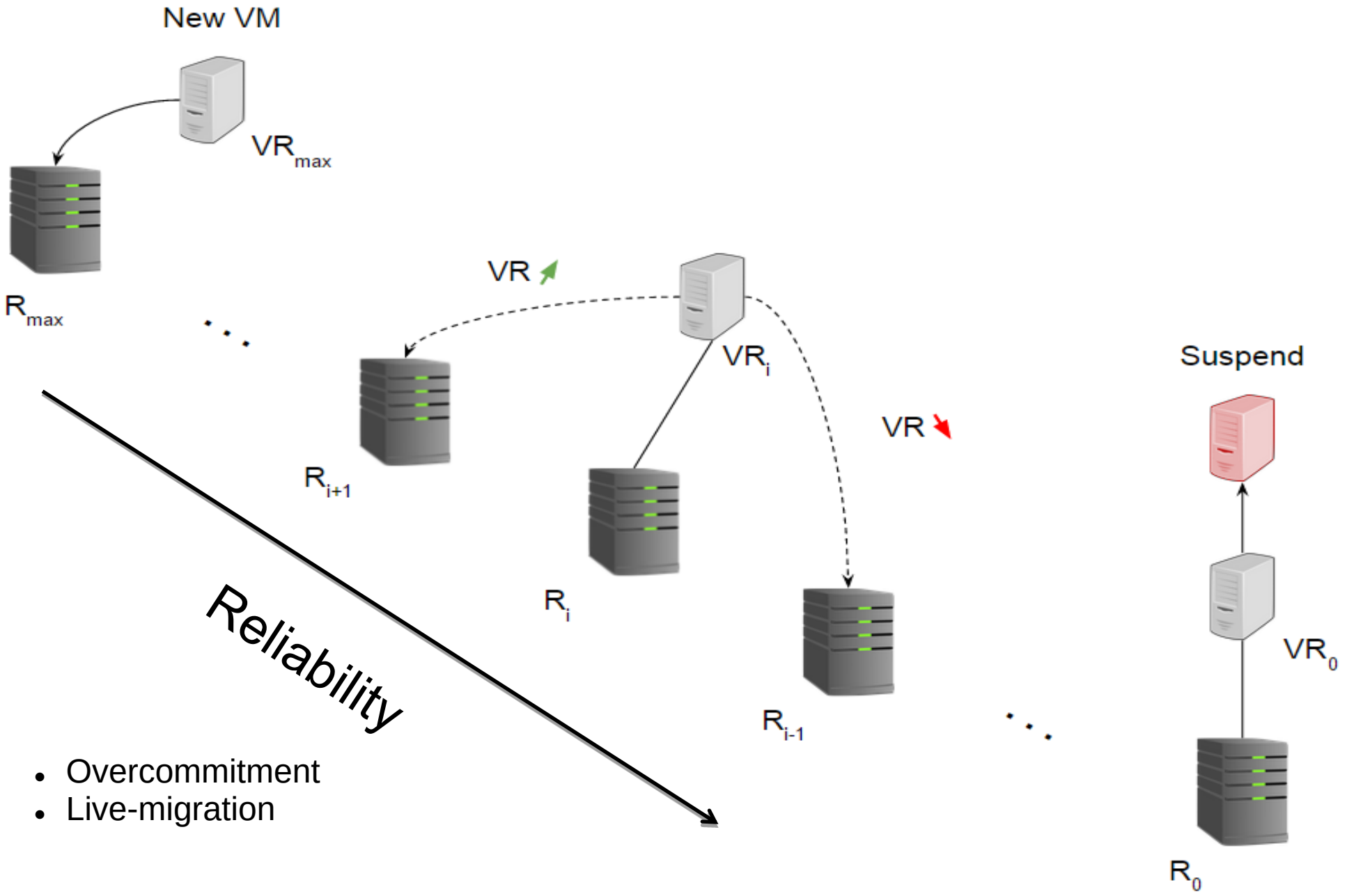


Modeling and analysis

Visualization

Grafana

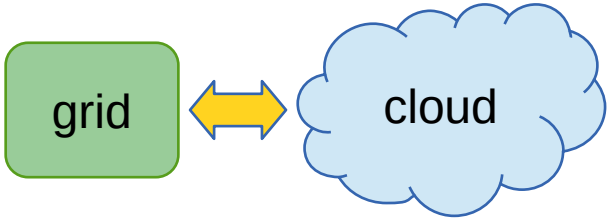
Smart Cloud Scheduler



- Overcommitment
- Live-migration

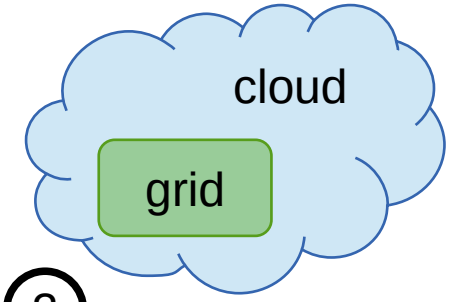
Clouds and grids

- Both ones **share** two main attributes:
  - they provide access to remote computing resources; and
  - they provide a service.
- But they are based on different paradigms:
  - clouds are being based on **virtualization of resources**,
  - grids being based on the **sharing of resources across boundaries**.
- Modern trend is a synthesis of these two technologies:



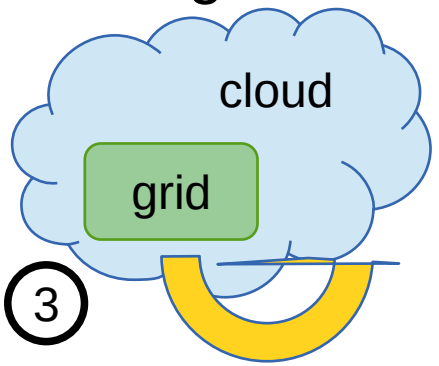
1

Cloud resources suppliments grid ones (e.g. during peak load to provide required QoS)



2

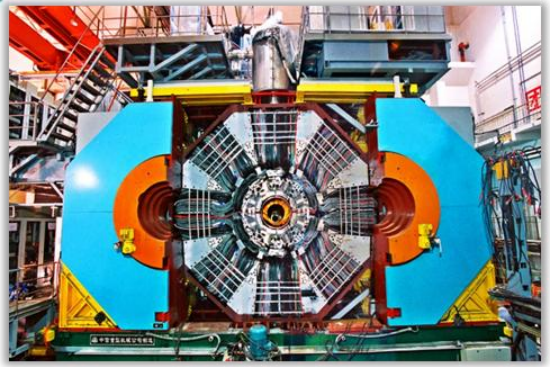
Services of grid site are deployed on cloud VMs (to increase hardware utilization efficiency and simplify admins' work)

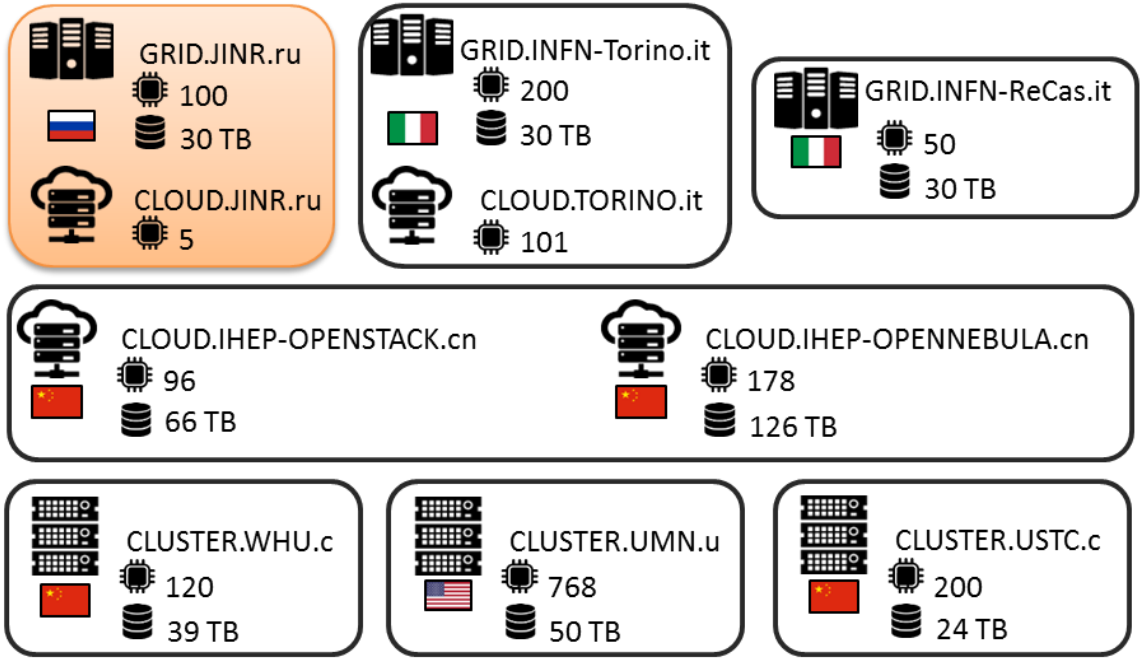


3

Synergy of 1st and 2nd approach

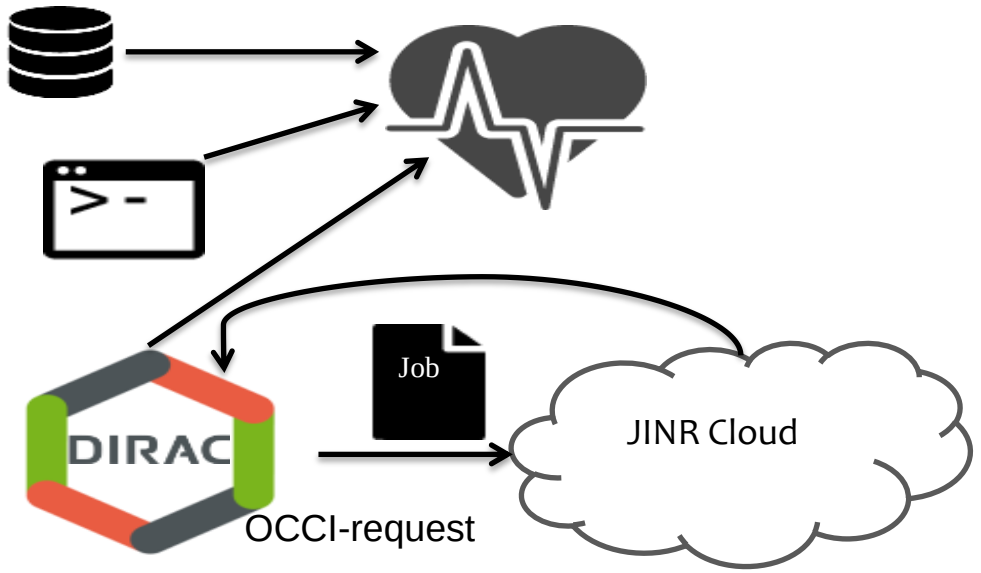
BES-III Distributed Computing and JINR Cloud





What have been done in computing:

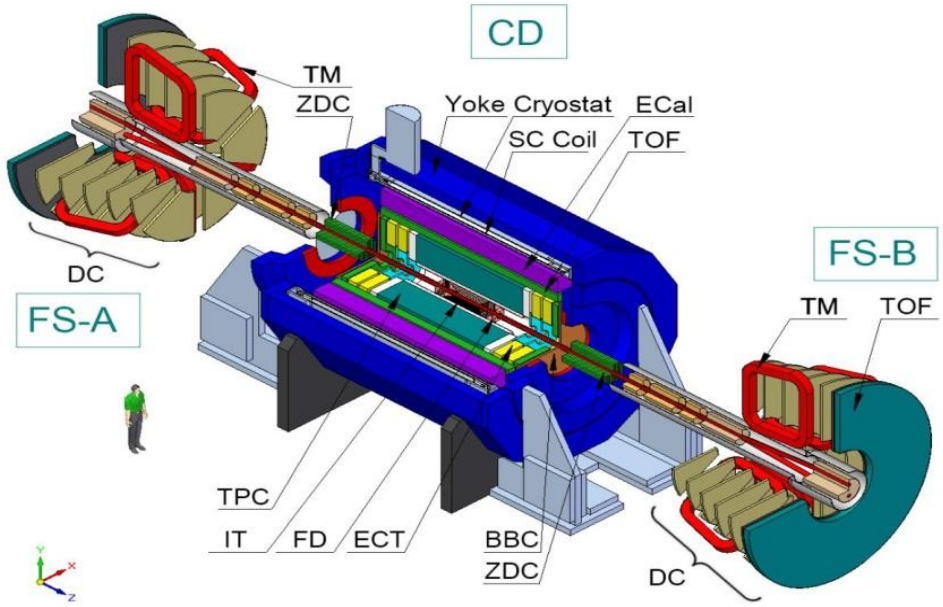
- Grid monitoring system developed from scratch
- JINR cloud was integrated in BES-III infrastructure
- 12 % of all jobs was done in JINR during the past year



Planning to continue participate in BES-III experiment by:

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

NICA Accelerator Complex



For the NICA project the data stream has the following parameters:

- high speed of the event set (up to 6 kHz),
- in central Au-Au collision at the NICA energies, about 1000 charged particles are generated,
- predicted event quantity - 19 billion;
- the total amount of initial data can be valued as 30 PB annually or 8.4 PB after processing.

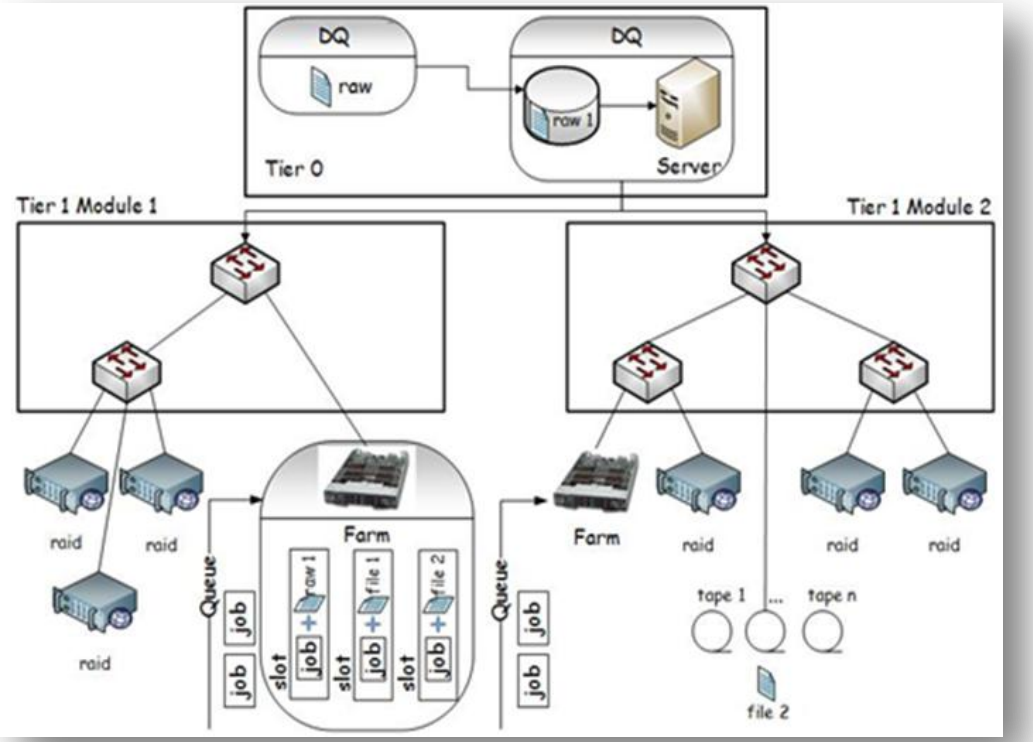
Simulation of NICA-MPD-SPD Tier0-Tier1 computing facilities

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

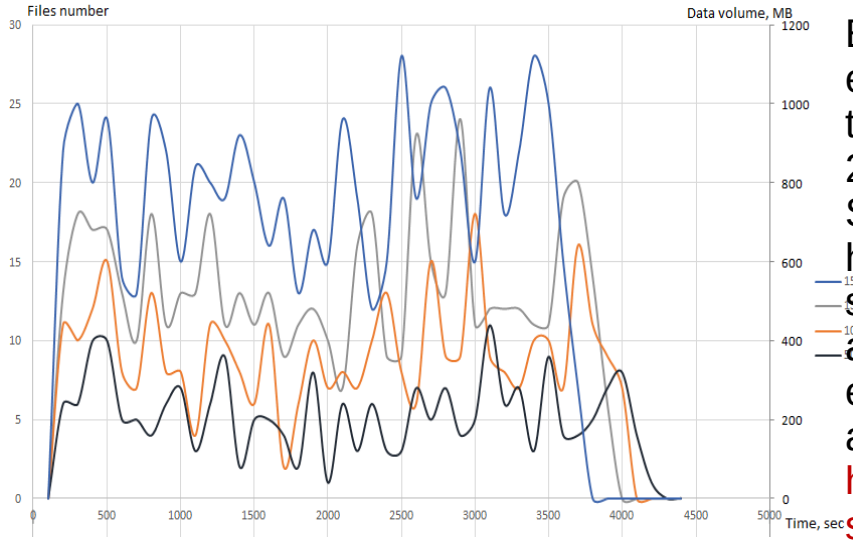
- large increase of CPU and network resources;
- combined grid and cloud access;
- Intelligent dynamic data placement
- distributed parallel computing;
- renewal most of simulation and analysis software codes.

The program SyMSim (Synthesis of Monitoring and SIMulation) for simulation of grid-cloud structures is developed.

The **originality** consists in **combining a simulation program with a real monitoring system** of the grid/cloud service in frame of the same program.



Data storage and processing scheme of Tier0-Tier1 level



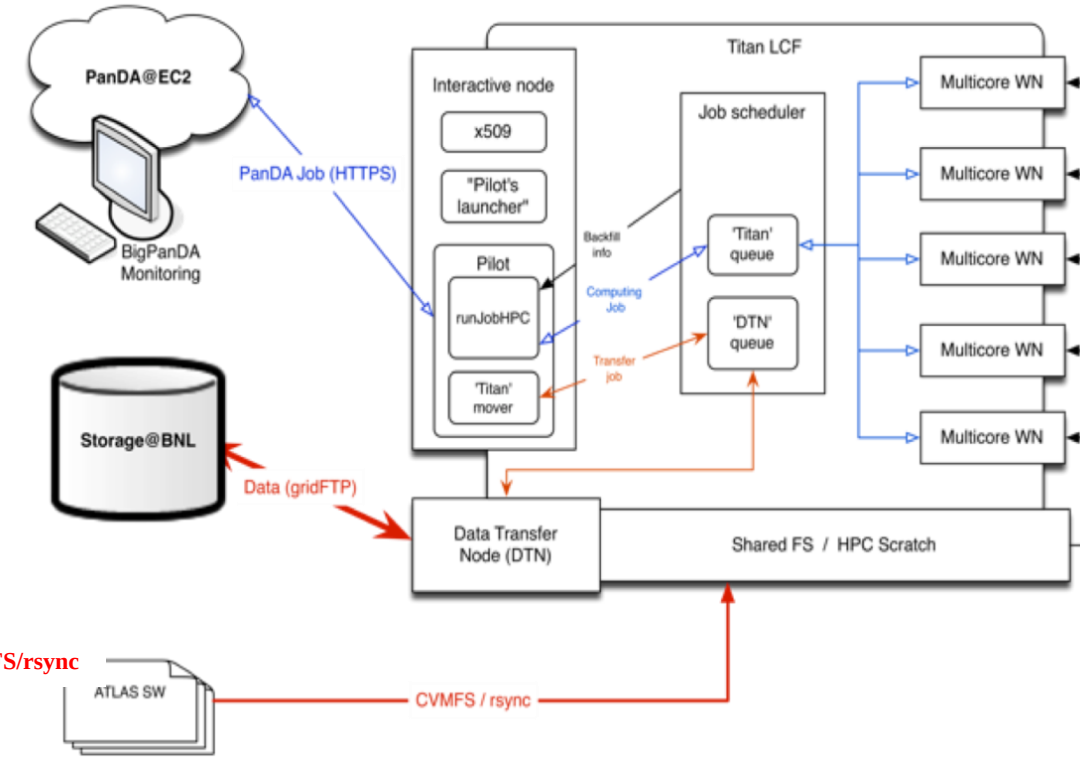
Number of DAQ data files stored on output disk buffer for growing data volumes

Estimated rate of NICA-MPD experimental data to be transferred to Tier 1 is about 24 PB by one month. Simulation result shows what happened in the grid/cloud system if the data volumes are grow up to 1,5 times for example. Simulation result allows one to understand how the intensity of the input stream determines the reserves of the system capacity

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		

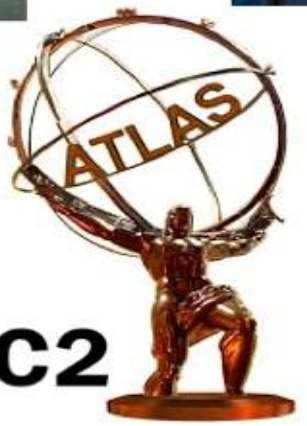
12 OLCF | 20

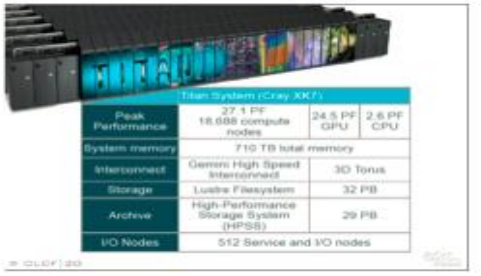


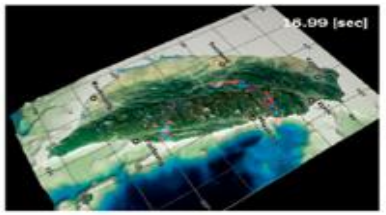
ATLAS (BNL, UTA), OLCF, ALICE (CERN,LBNL,UTK), LIT JINR:

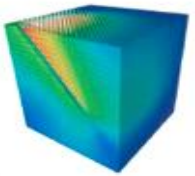
- adapt PanDA for OLCF (Titan)
- reuse existing PanDA components and workflow as much as possible.
- PanDA connection layer runs on front-end nodes in user space. There is a predefined host to communicate with CERN from OLCF, connections are initiated from the front-end nodes
- SAGA (a Simple API for 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

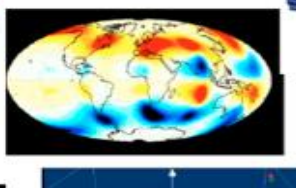
BigPanDA

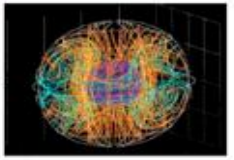


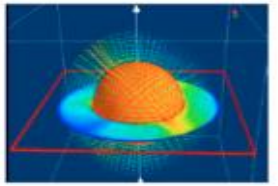


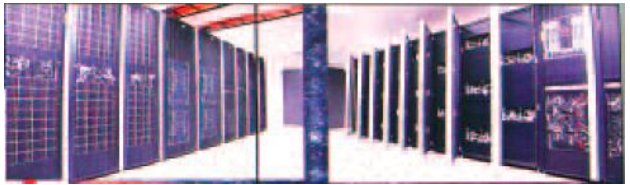












LIT traditional conferences and schools in 2015

MATHEMATICAL MODELING AND COMPUTATIONAL PHYSICS 2015

Stará Lesná, High Tatra Mountains, Slovakia

July 13 — 17, 2015

XXV INTERNATIONAL SYMPOSIUM ON NUCLEAR ELECTRONICS & COMPUTING

Montenegro, Budva, 28 september - 02 october

and

IT – Student School

JINR/CERN/MEPHI

GRID AND ADVANCED INFORMATION SYSTEMS

SCHOOL ON JINR/CERN GRID AND ADVANCED INFORMATION SYSTEMS

November 02-06, 2015

In LIT holds regular tutorial courses and traineeship of young scientists and students from the JINR Member States

Projects of LIT in distributed computing

- Worldwide LHC Computing Grid (WLCG)
- EGI-InSPIRE
- RDIG Development
- Project BNL, ANL, UTA "Next Generation Workload Management and Analysis System for BigData"
- Tier1 Center in Russia (NRC KI, LIT JINR)
- 6 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" was supported in frames of JINR-South Africa cooperation agreement;
- Development of grid segment at Cairo University and its integration to the JINR GridEdu infrastructure
- 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 of Tier-2 level at LIT JINR and Bogolyubov Institute for Theoretical Physics for data processing from ALICE experiment"
- JINR-Romania cooperation Hulubei - Meshcheryakov programme
- JINR-Moldova cooperation (MD-GRID, RENAM)
- JINR-Mongolia cooperation (Mongol-Grid)
- JINR-China cooperation (BES-III)
- Cooperation with Belarus, Slovakia, Poland, Bulgaria, Kazakhstan, Armenia, Georgia, Azerbaijan

Conclusions

- Cloud technologies are highly suitable for modern scientific activities
- There are many different ways in cloud evolution
- There are ways to optimize cloud utilization efficiency

Acknowledgments

The activities are partially supported by the RFBR grants 15-29-07027 and 14-07-90405

In 2016 JINR will celebrate its 60th anniversary.
You all are welcome to take part in this remarkable event !

Thanks for your attention!

V.V. Korenkov
LIT director
E-mail: korenkov@jinr.ru