



WLCG Tier-2 computing center at NRC "Kurchatov Institute" - IHEP: 20 years of operation

V. Gusev[†], V. Kotliar^{1*}, V. Kukhtenkov¹, V. Savin¹

NRC "Kurchatov institute" - IHEP, RU-142281, Protvino, Moscow region, Russia

E-mail: {Victor.Gusev, Viktor.Kotliar, kvi,savin}@ihep.ru

^{*} Corresponding author

WLCG Tier-2 computing center at NRC “Kurchatov Institute” - IHEP has been participating in the Worldwide LHC Computing Grid from very **beginning since 2003**. Over a twenty-year period it became one of the biggest WLCG Tier-2 centers in Russia. Ru-Protvino-IHEP Grid site provides computing resources for LHC experiments in high energy physics such as Atlas, Alice, CMS, LHCb and internal experiments at NRC “Kurchatov Institute” - IHEP such as OKA, BEC and other.

In this work the current status of the **computing capacities, networking and engineering infrastructure, used software** will be shown as well as **the evolution** of the computing center over last 20 years for stable and efficient operation.

1963 – Minsk-2, Minsk-22, M-220, BESM-4,
BESM-6, Minsk-32, EC-1040, EC-1045

1972 – ICL

1977 – DEC 10

1991 – mVAX-II – mail

1993 – Internet

2003 – Grid-cluster

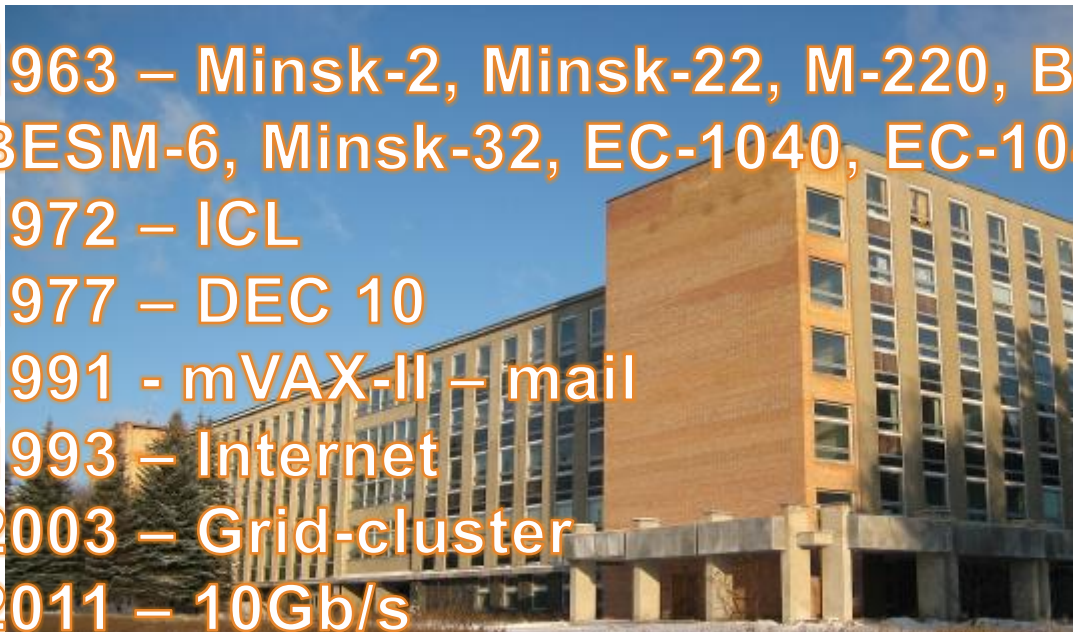
2011 – 10Gb/s

2013 – 2x10Gb/s

2015 – LHCONE

2018 – IPv6

2023 – support for mega projects





SU-Protvino-IHEP: history 2003

- GRID SW (LCG + EDG) installed at the end of **2003** (SU-Protvino-IHEP).
- In that time were installed and configured the first grid infrastructure services like CE, SE, WNs, UI on **16** two-core Pentium III 900MHz (still in use as cluster primary DNS!) 3x2Mb/s Internet.
- Starting from 2004 with EGEE project IHEP started to invest more resources to Linux cluster
 - NA2, NA3, NA4, SA1
 - RDIG (ROC functions)
 - site name RU-Protvino-IHEP
 - Create core team for cluster support (system administrating, networking, engineering)
 - Expertise for Beowulf clusters
 - Hardware investments
 - Moving IHEP experiments from Alpha servers (True64 Unix, OpenVMS) to Linux

History: 2004-2007



- Increasing network bandwidth to 100Mb/s
- buying a new hardware
- More Grid activities: SA2 for testing SW
- Collaboration with CERN to develop SW

History: 2007- 2010

- New hardware
- Migration to x86_64 from first 32bit cluster
- Migration to racks from commodity-grade hardware
- EGEE I,II, III projects



History: 2011- 2013



- Double CPU and Disks with RDIG support
- Setup new hardware – twin based (more power)
- End of EGEE era no more external support for grid activities
- Some sysadm members left the team
- 10Gbs channel to Moscow

History: 2013-2019

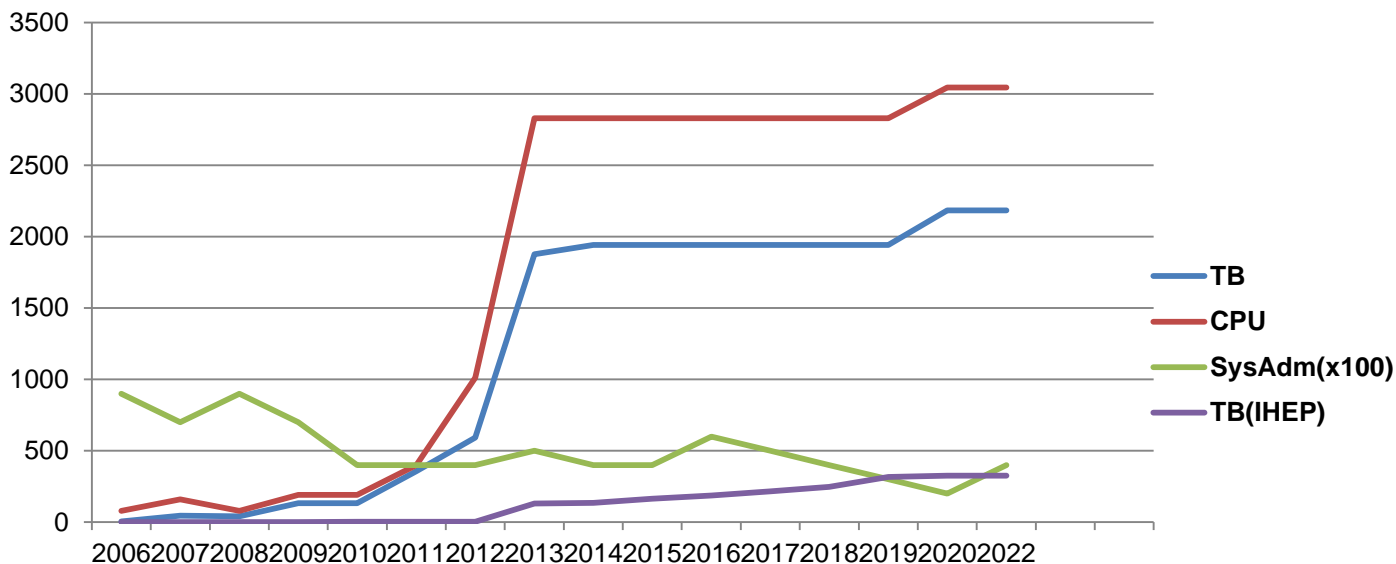
- Triple CPU and Disks with IHEP support (no more RDIG hardware)
- Renew power and cooling systems
- IHEP migration to NRC KI from Rosatom
- Double 10Gbs channel to Moscow
- Core cluster hardware for the present moment



- Stable operation
 - Investing human resources into the cluster management system
 - Research projects for additional monitoring
 - Create self-protection systems for cluster which are based on temperature, power, cooling status
 - we have BigRedButton for software cluster shutdown in case of major failure
- Some new hardware
 - Increase for local storage
 - Increase for LHCb storage (one of few Tier2 sites with storage for LHCb)
 - Minor increase for CPU (preparation for the hardware replacement from 2004-2013)
- Add IPv6 support for the cluster

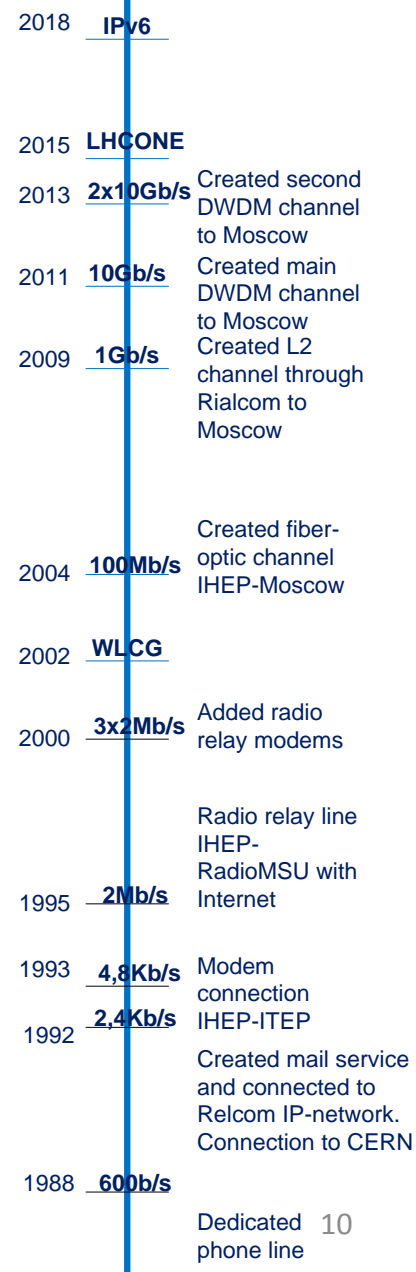


IHEP resources evolution



growth of the IHEP grid resources by year in TB and CPU

- At the moment our site serves for four LHC experiments (Atlas, Alice, CMS, LHCb) and many small experiments inside the Institute
- Implement shared CPU schema(mix Grid jobs 90% with non Grid jobs on nodes 10%) allows achieve 24x7x365 CPU 100% resource usage.
- Main goal to support available resources without degradation



IHEP resources: evolution



- Less humans more servers
 - Bare minimum for SysAdms, Engineers, NetAdms
- New expertise for running Linux clusters and distributed systems
- SysAdms gain unique Grid experience



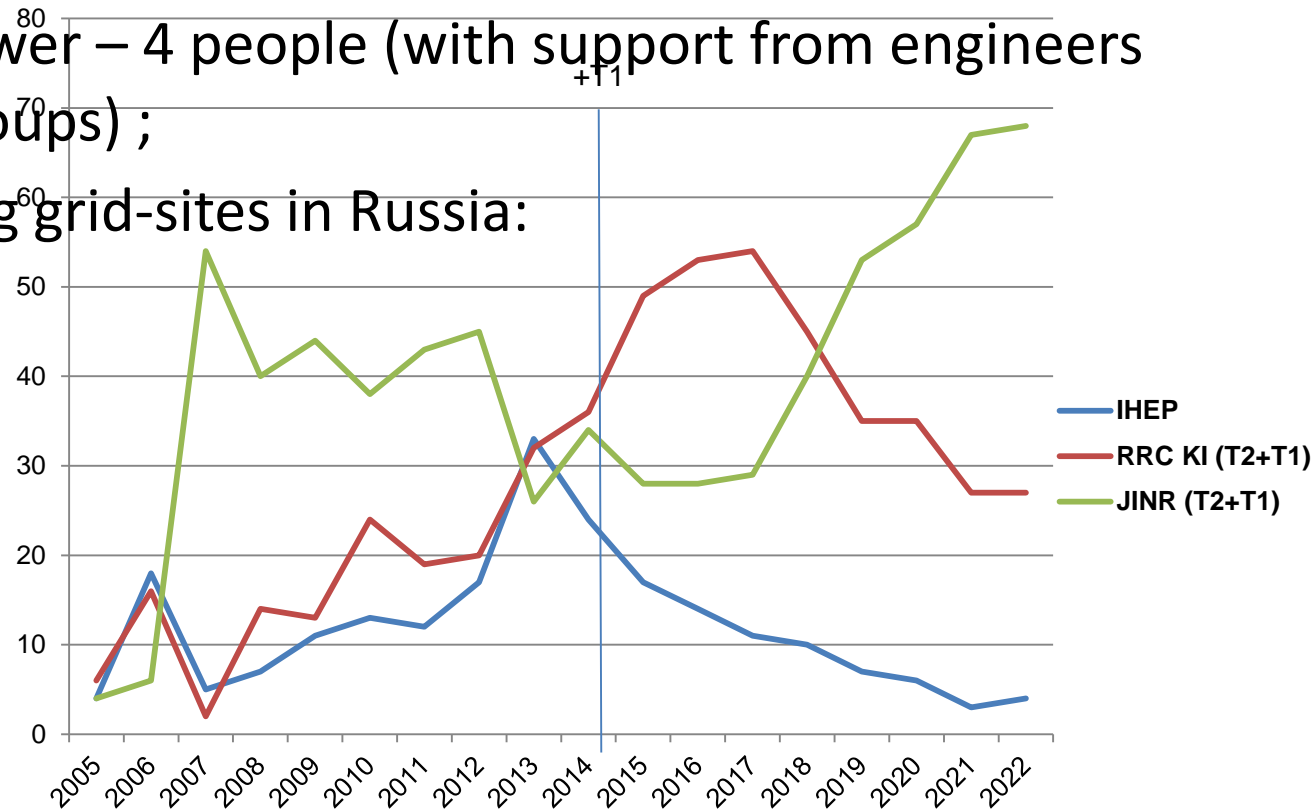
- Almost everything software defined
- Opensource software everywhere (constant unintruptable upgrades – new skill)
- Sys&Net Adm - > programmers

04.07.2023

Grid 2023

11

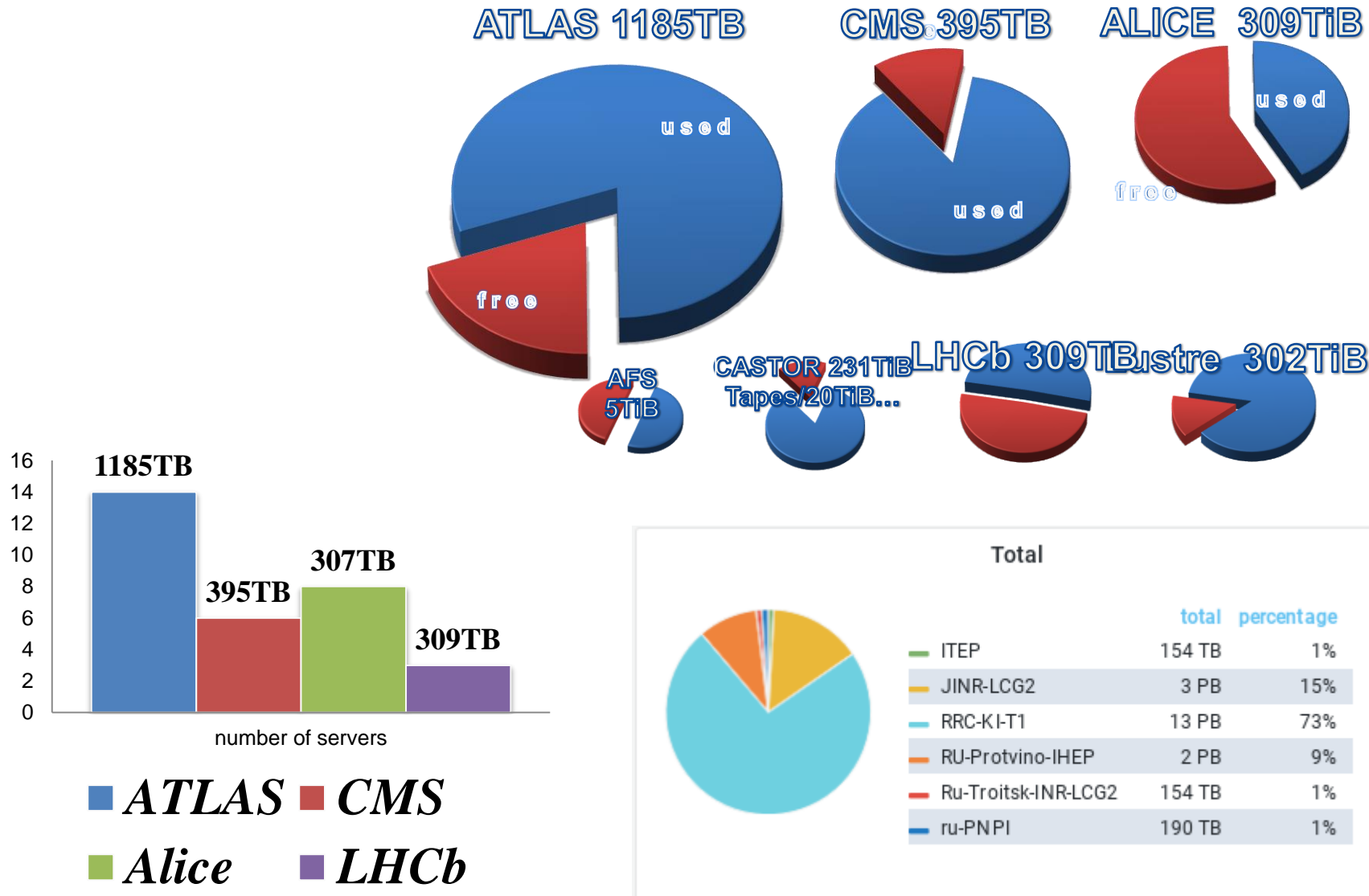
- 3044 CPU, 26875 HEP-SPEC06;
- 2183 TB: Atlas 1185, CMS 395, Alice 314 , LHCb 289 ;
- 2x10Gb/s Internet channels IPv6 dual stack;
- SysAdm manpower – 4 people (with support from engineers and network groups) ;
- One of three big grid-sites in Russia:



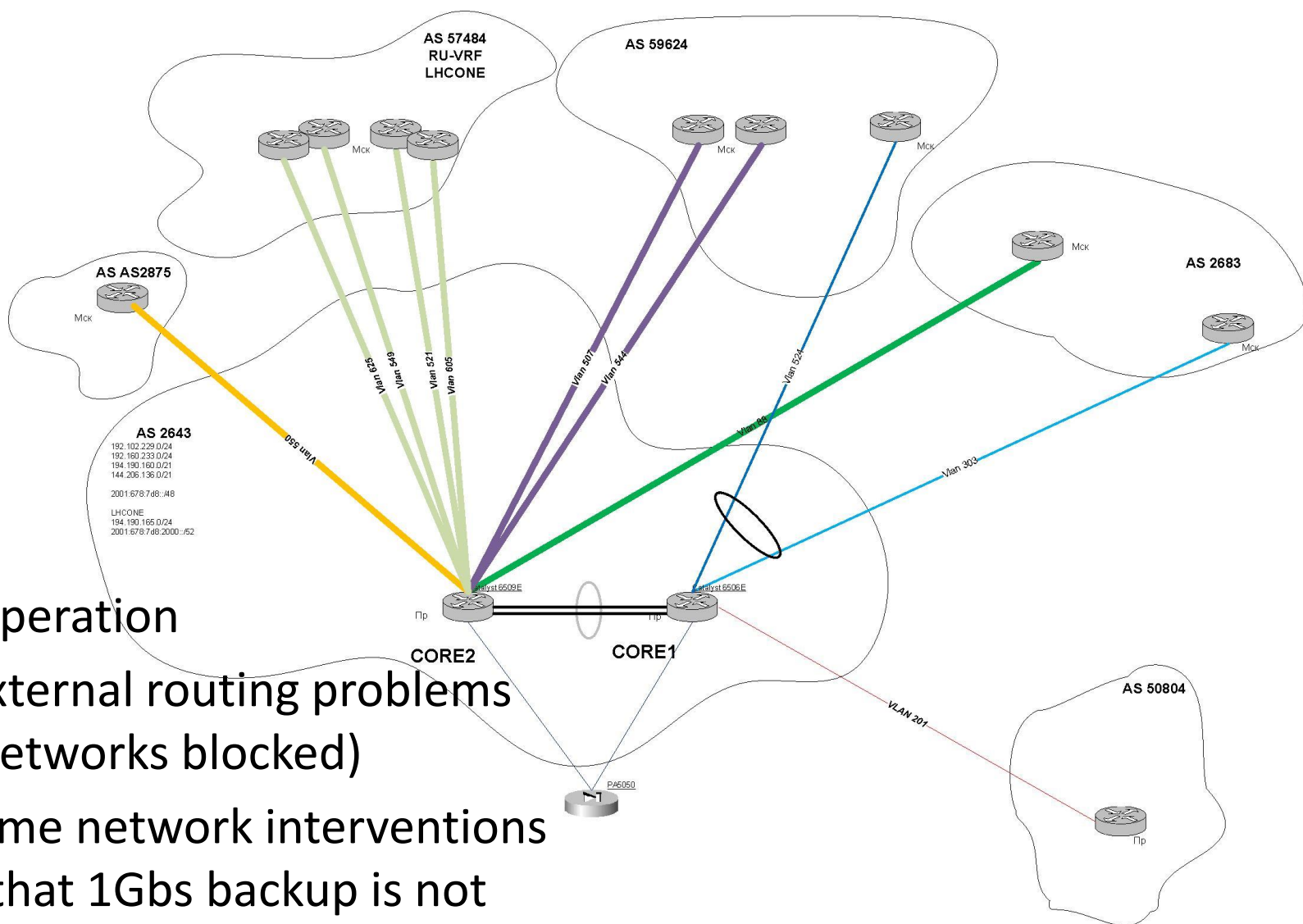
Contribution in % to RDIG by normalised CPU time

Grid 2023

Current status: storages

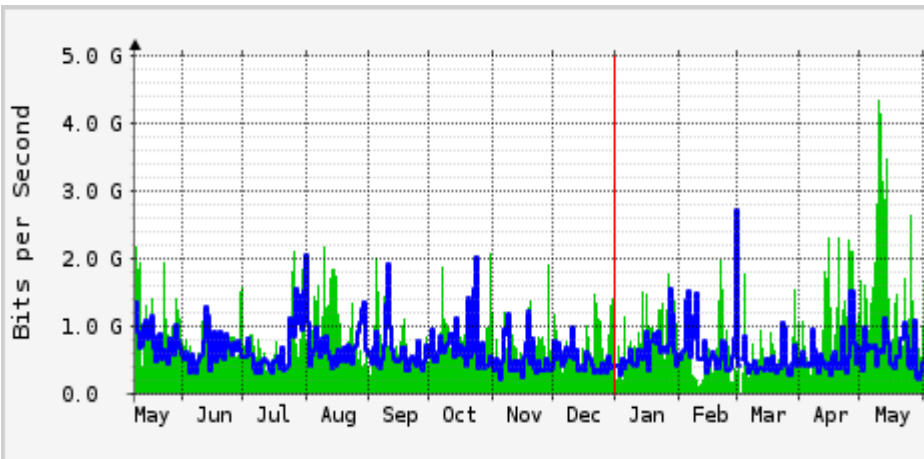


Current status: IHEP external network

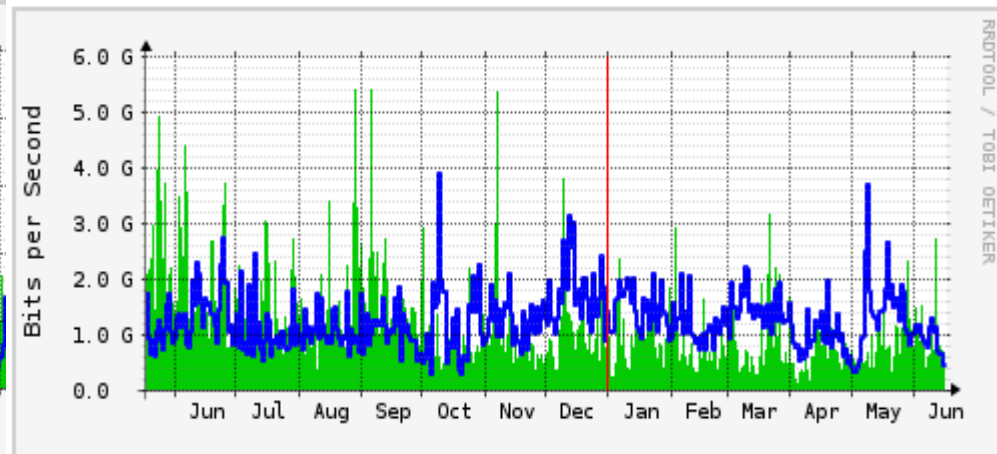


- Stable operation
- Some external routing problems (some networks blocked)
- After some network interventions we see that 1Gbps backup is not good enough

Current status: IHEP external network usage



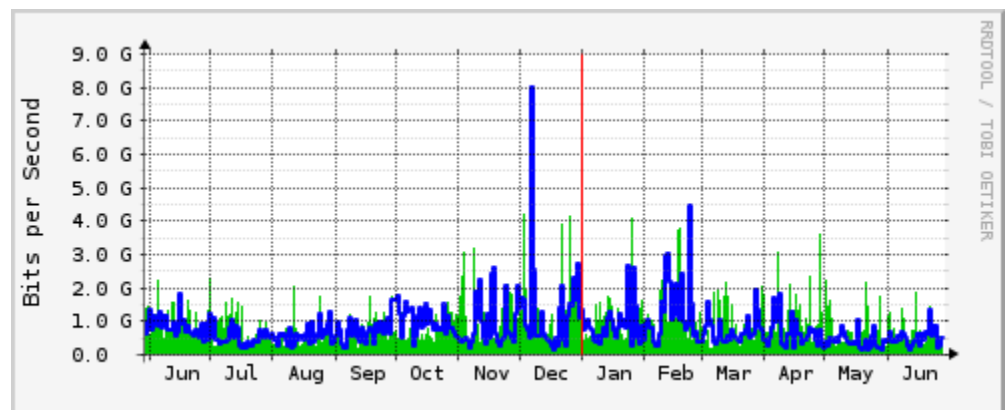
2016



2021

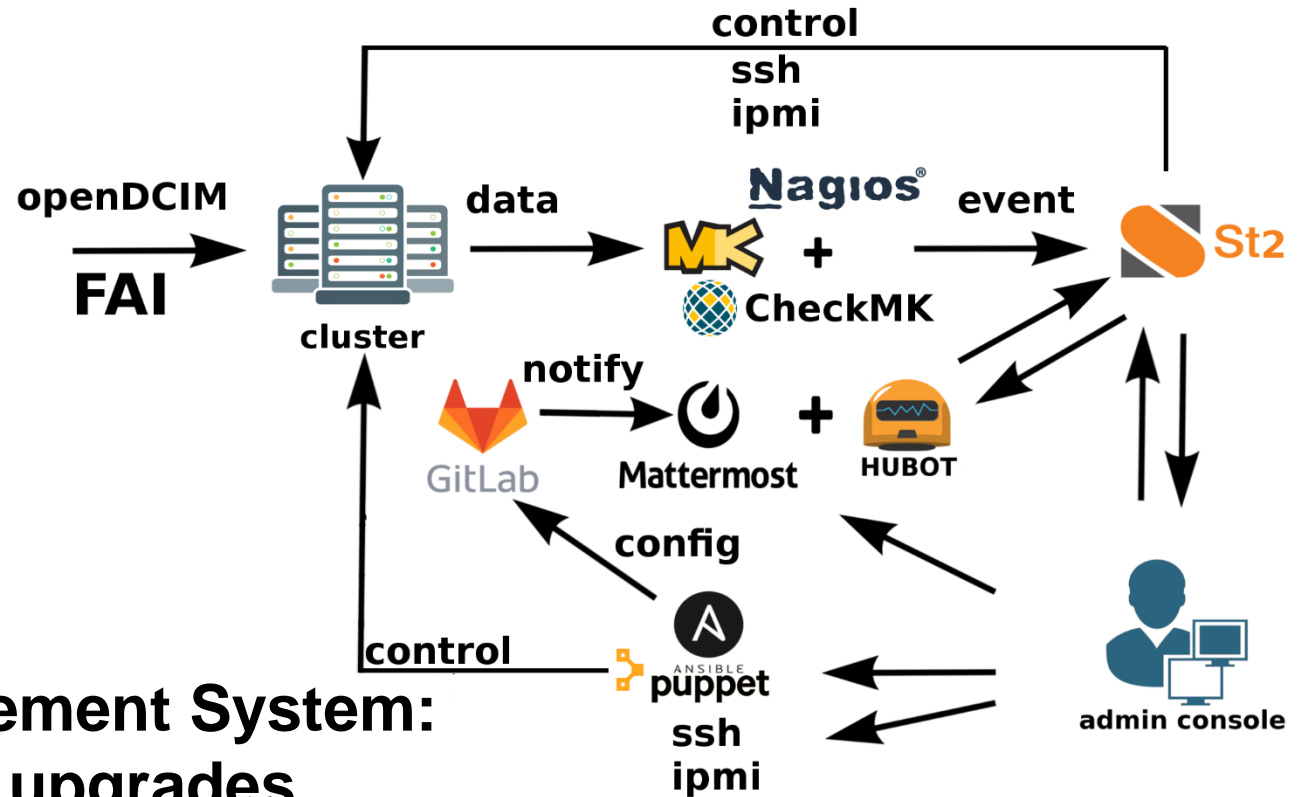
No traffic increases

- No major changes in resources
- No major changes in computing paradigm for Grid jobs
 - Maybe even less traffic for multicore jobs



2023

Current status: SW usage



Cluster Management System:

- Easy system upgrades
- Easy deployment of new systems
- Ansible with Puppet combo
- Easy cluster management (2-3 people)
- Monitoring in push mode

Current status: SW usage

Four aspects of self-management

1. Self-configuration

- Configure themselves automatically
- High-level policies (what is desired, not how)

2. Self-optimization

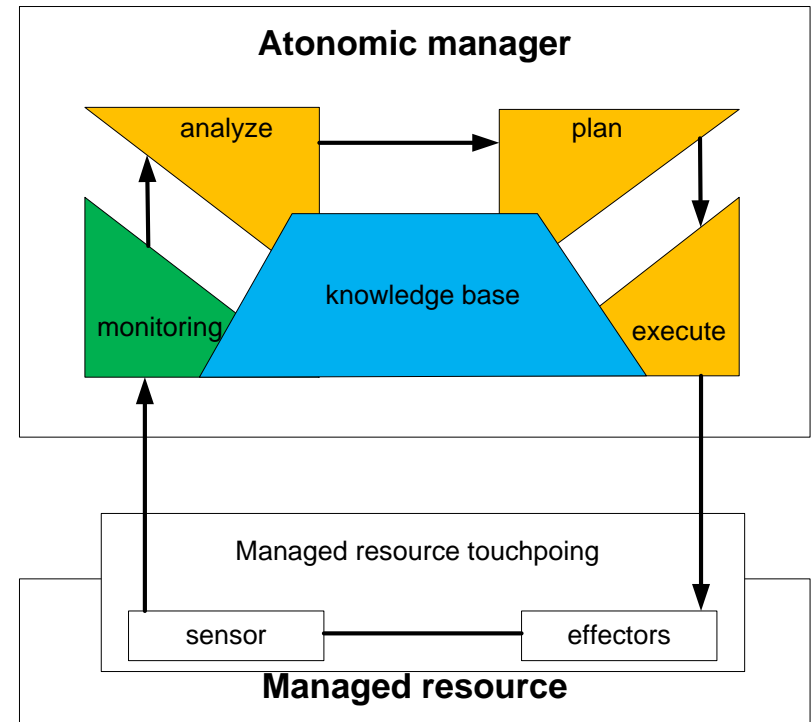
- Hundreds of tunable parameters
- Continually seek ways to improve their operation

3. Self-healing

- Analyze information from log files and monitors

4. Self-protection

- Malicious attacks
- Cascading failures





Current status: GRID networking

New conditions since 2022 (Network black holes GEANT + NORDU NET)

...

I wanted to make you aware that at their meeting of the 9th of June 2022, the GÉANT Executive team agreed to block the AS numbers of institutions found on the EUs 6. These are the Institute of High Energy Physics (IHEP) and the Institute of Theoretical and Experimental Physics (ITEP) - AS2148 and AS2643.

The block will be enforced at 13.00 BST, Wednesday 29th June.

...

No more *.egi.eu services since NET block

- Site switched to direct VO support without EGI support and monitoring
- GGUS based or E-mail based

End of 2022

- **WLCG MB** did not object to the plan according to which the suspension will become effective as of 01 February 2023.



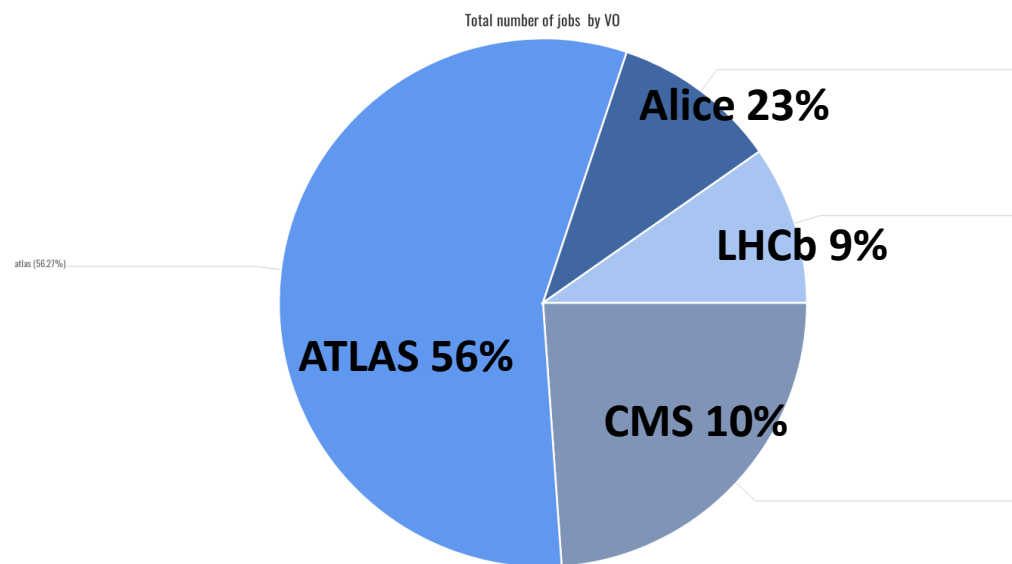
Current status: GRID services

Suspended since February 2023

- Measure: suspension of IHEP and IHEP in GOCDB to 'remove' the concerned data centres from the production infrastructure. The sites will then need to be recertified once conditions for suspensions are removed.
- Consequences:
 - the VOs supported by the site are no longer able to discover and use resources (e.g. through configuration information sources such as GOCDB and BDII). The site information is no longer collected in the top-level BDII.
 - the accounting data of the site is no longer accepted by the APEL accounting repository
 - the site is no longer monitored by the ARGUS (Monitoring) and Security monitoring tools
 - the site information is no longer collected in the top-level BDII
 - the site can no longer receive tickets through the EGI Helpdesk
 - the site is no longer discoverable in AppDB (only relevant to data centres in the EGI Federated Cloud)
- Since that RU-Protvino-IHEP started to work like independent GRID site with direct communications to VO (thanks to distributed GRID nature it is possible)
- Our experience allows us to run site without external monitoring and security tools, we can use VO specific accounting if needed

Current fair share setup (no changes for many years):

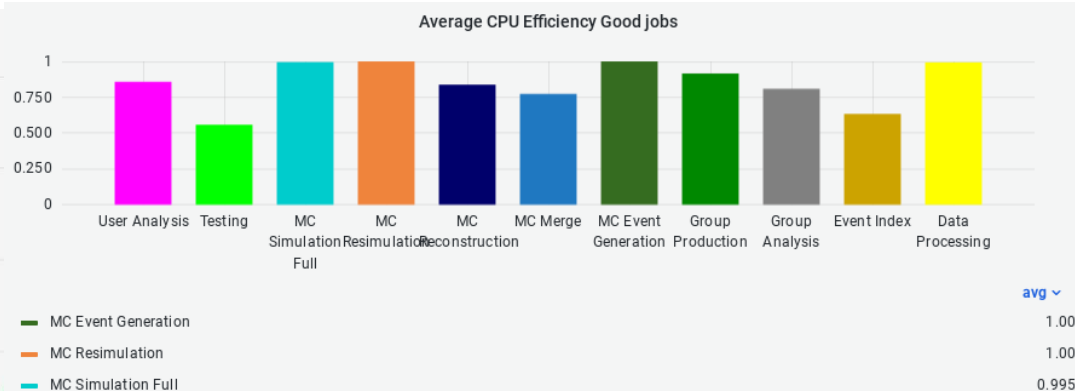
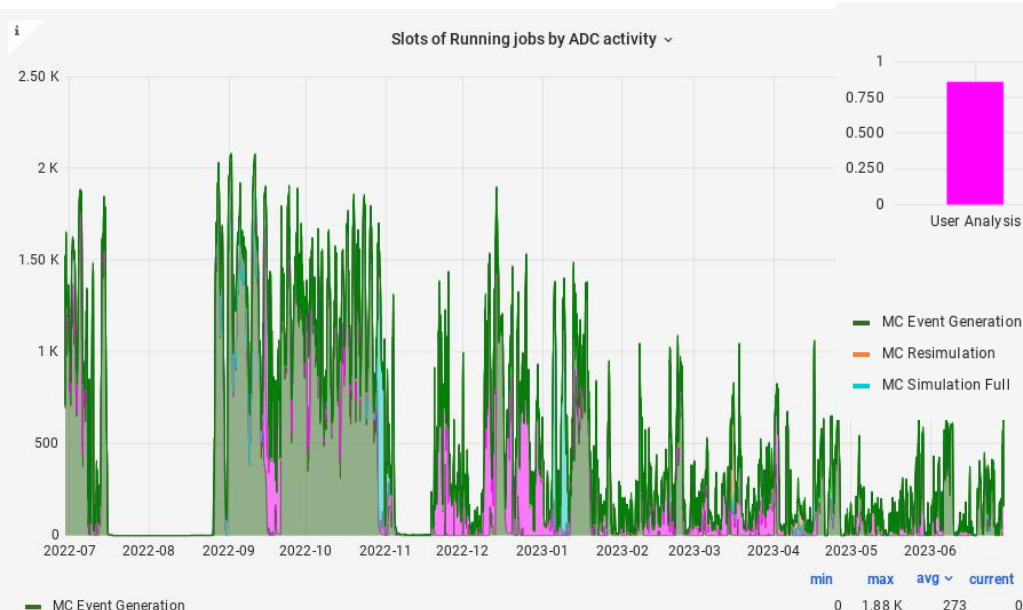
- ATLAS 52%
- CMS 30%
- ALICE 12%
- LHCb 6%



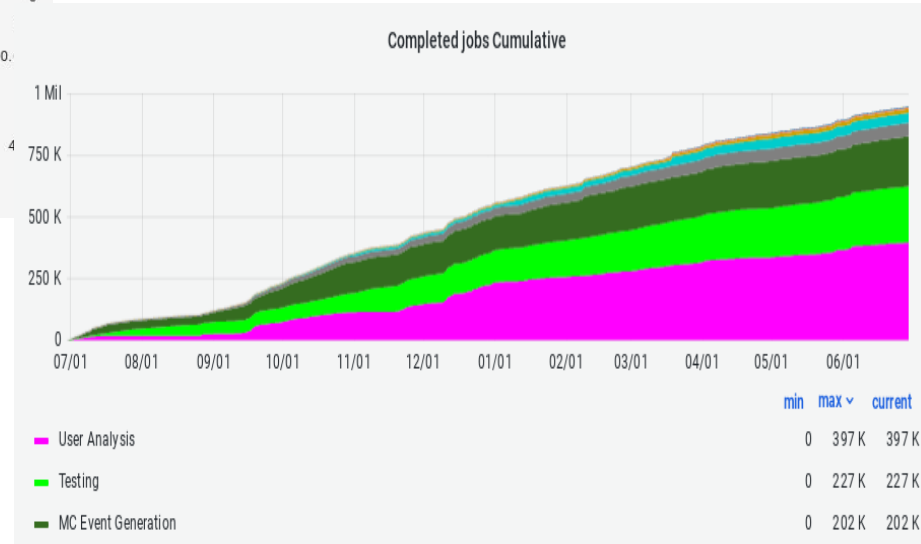
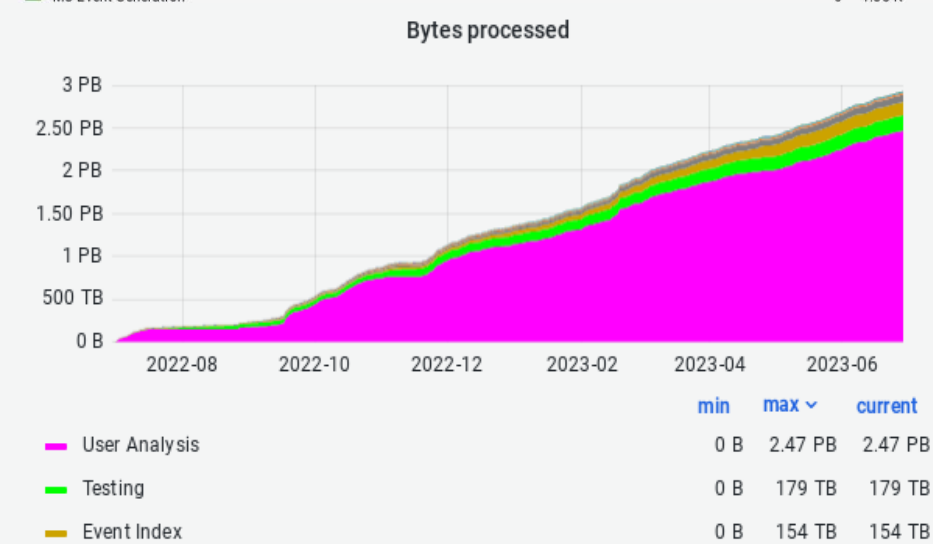
Grid site usage by VO for 2022



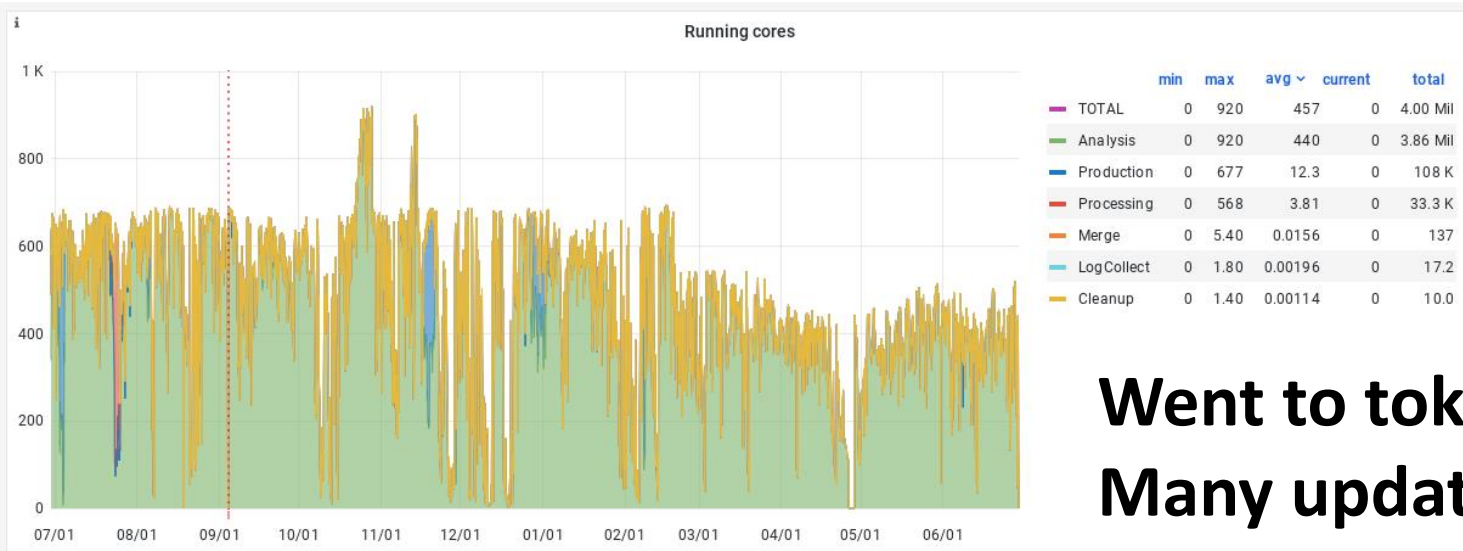
Sites usage by ATLAS last year



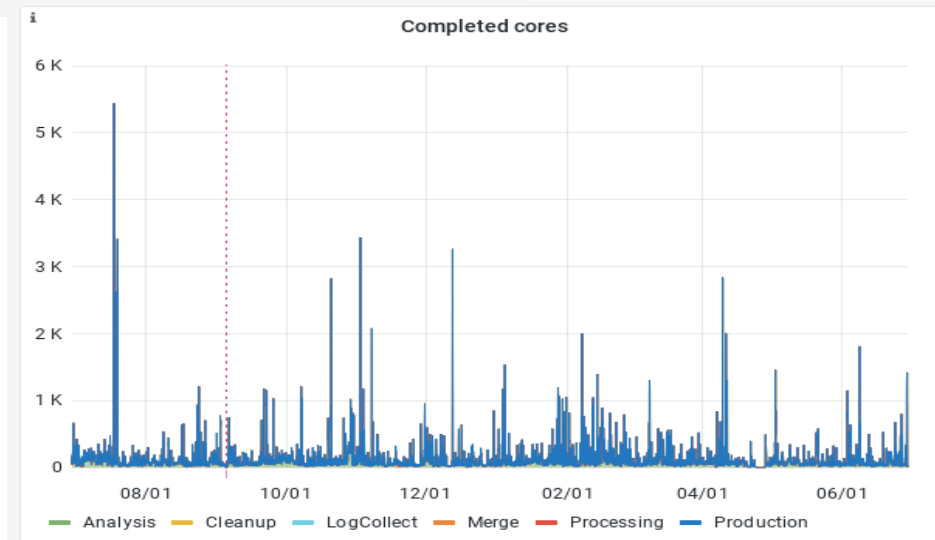
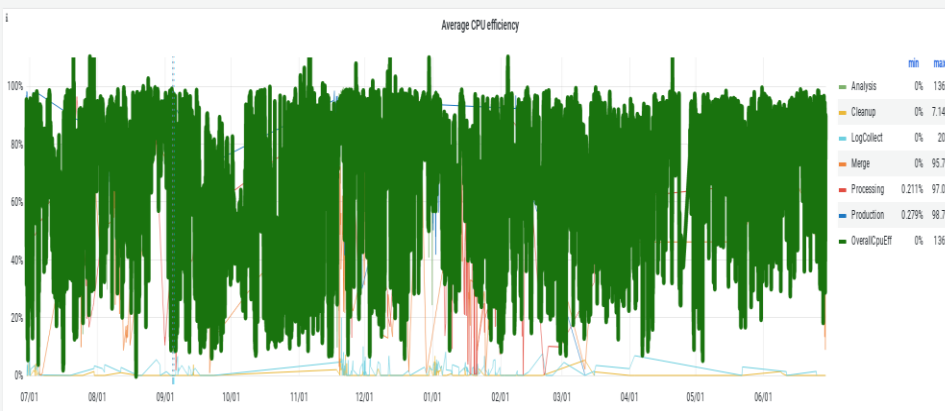
No changes in computing



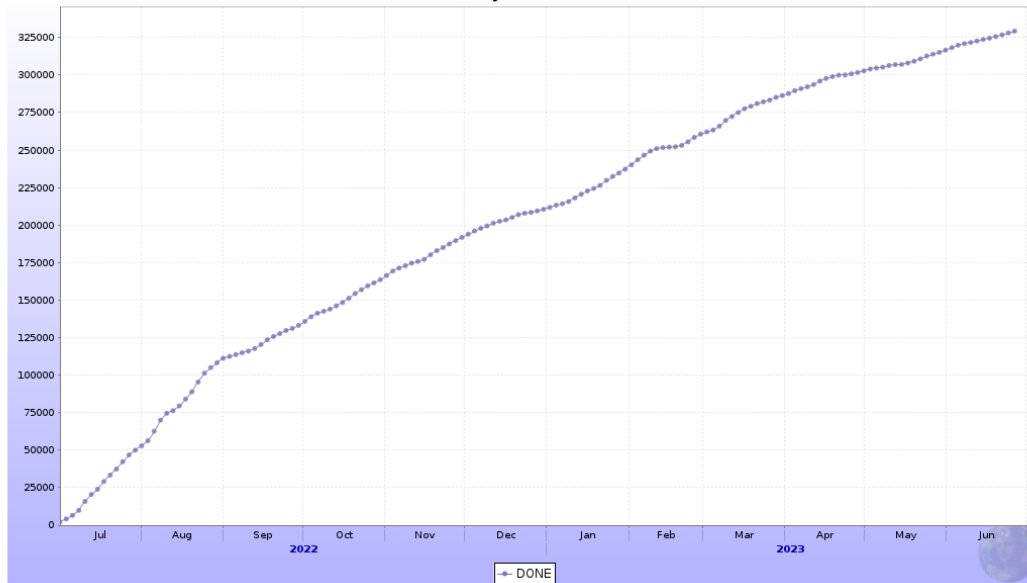
T2 sites usage by CMS last year



Went to tokens
Many updates for ARC CE

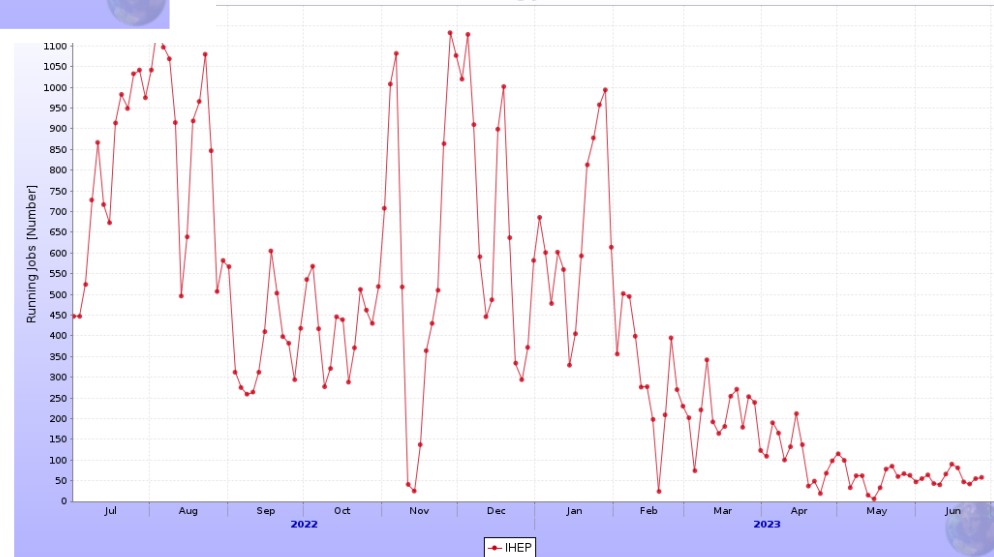


Done jobs in IHEP

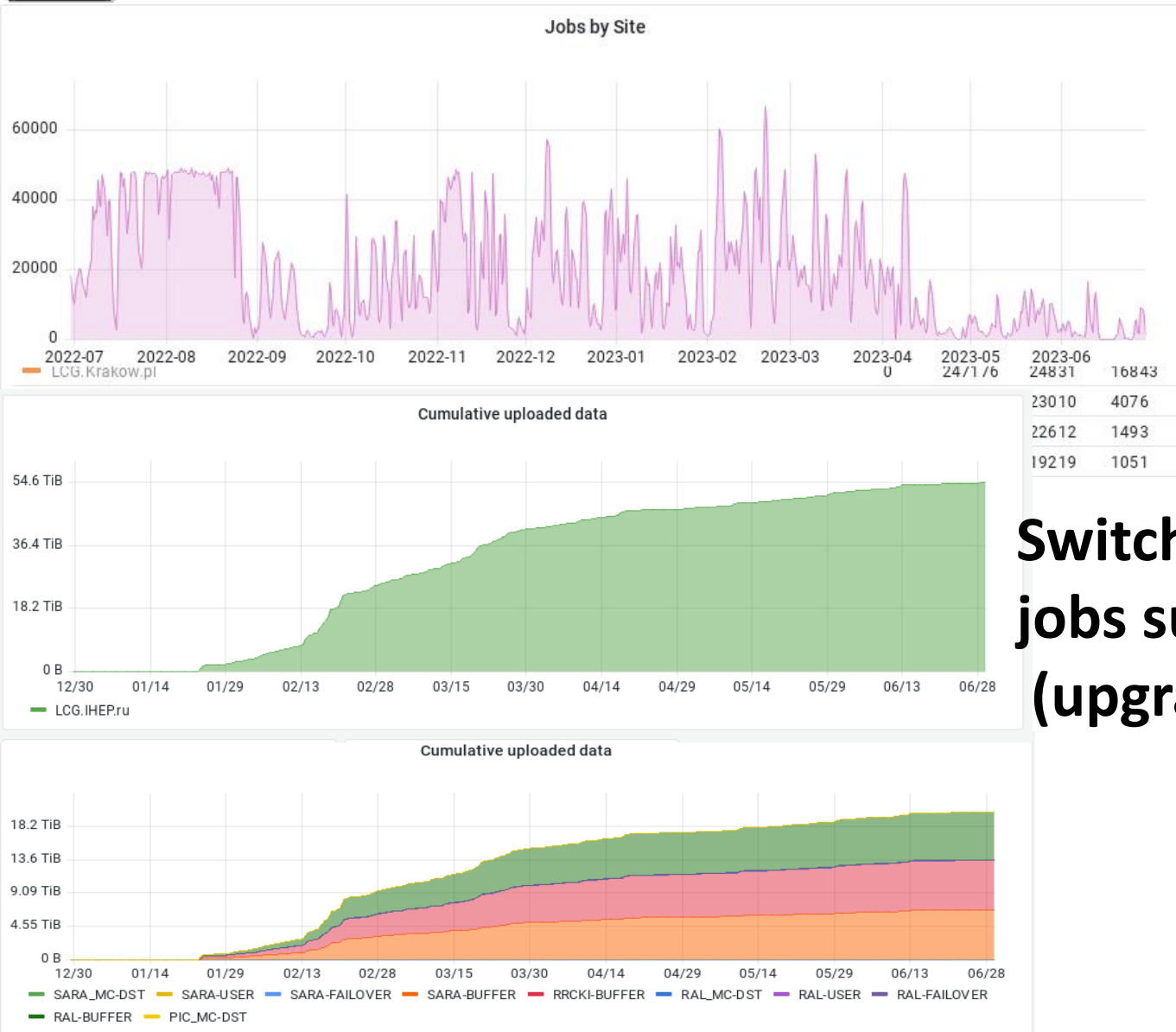


Went to Multicore jobs

Running jobs



T2 sites usage by LHCb last year

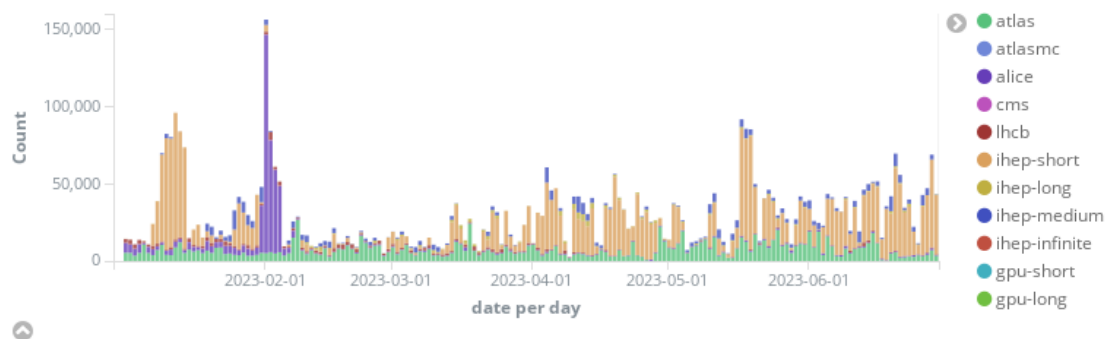


**Switch to AREX
jobs submission
(upgrade ARC CE)**



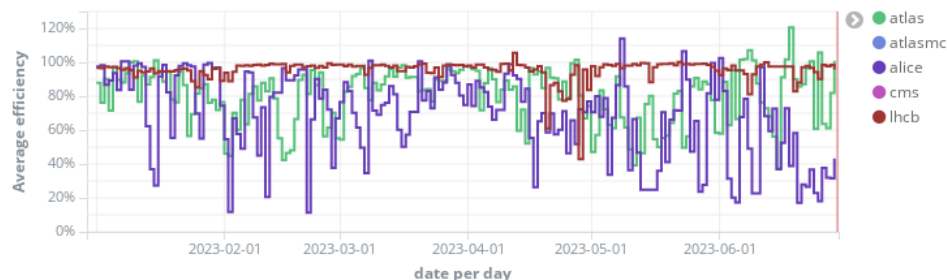
IHEP 24x7 cluster with high reliability and availability and efficiency

pbs-queue

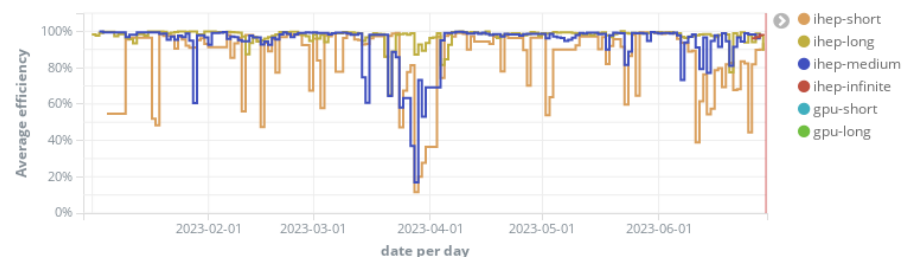


Year	GGUS tickets solved	Disks replaced	Availability and Reliability
2015	30	25	99.9
2016	34	56	96.67
2017	22	35	98.04
2018	26	50	94.79
2019	36	78	85.81
2020	39	60	89.34
2021	33	48	81.63
2022	33	51	90.2

pbs-efficiency-grid



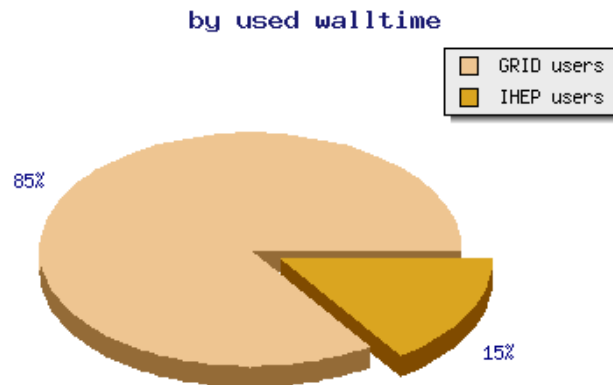
pbs-efficiency-ihep



Use of the cluster by local experiments

Cluster usage by local experiments increased

- Migration data from external resources (CERN, BNL)
- No access to other clusters anymore from computations
- Developing (simulation, preparations) for new mega projects



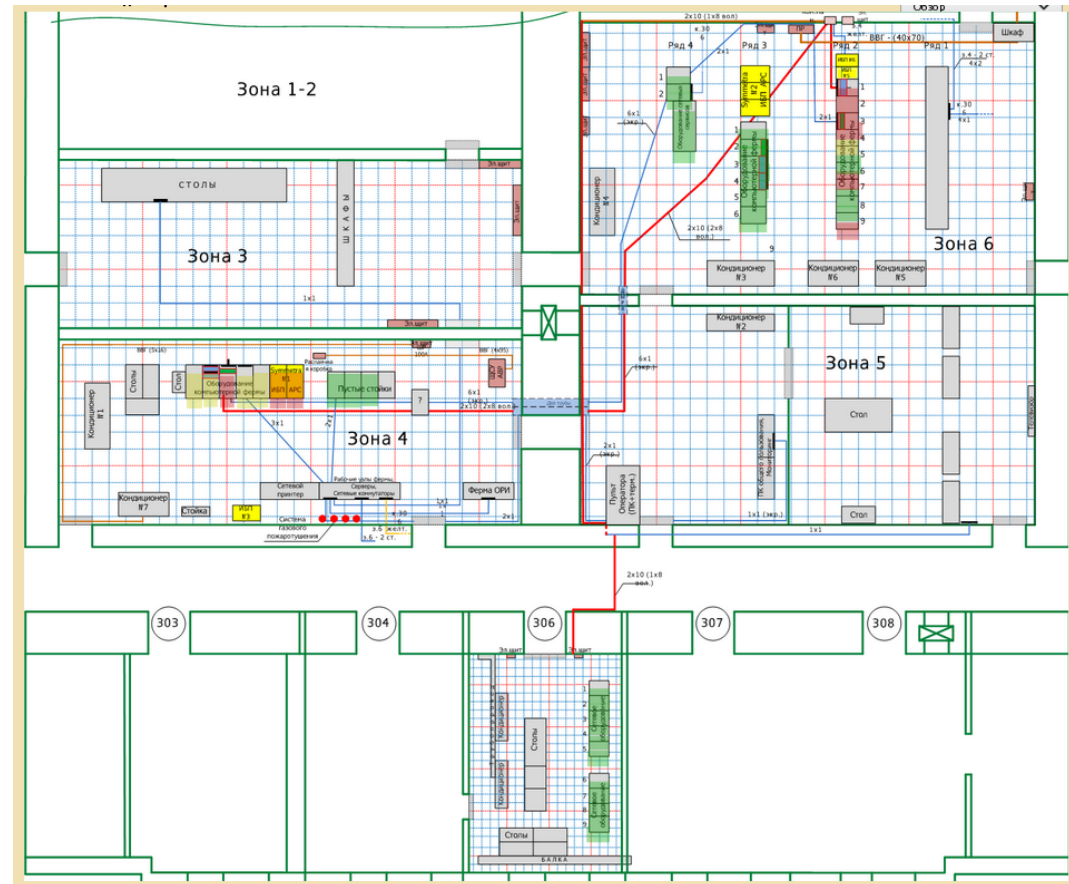
```
1 ivanilov
50 semakart
500 knovikov
500 sumaneev
543 minaev
611 moiseev
2188 gotman
```

server: ce0001-int.m45.ihep.su

Queue	Memory	CPU Time	Walltime	Node	Run	Que	Lm	State
-----	-----	-----	-----	----	---	---	---	-----
cmsmc	--	768:00:0	96:00:00	--	46	0	26	E R
cms	--	60:00:00	72:00:00	--	0	0	10	D R
gpu-long	--	720:00:0	288:00:0	--	0	0	13	E R
ihep-infinite	--	720:00:0	1440:00:	--	51	0	12	E R
ihep-long	--	144:00:0	160:00:0	--	256	355	20	E R
alice	--	256:00:0	26:00:00	--	12	0	16	E R
ihep-short	--	02:00:00	06:00:00	--	355	1564	20	E R
ihep-medium	--	16:00:00	24:00:00	--	321	1096	10	E R
lhcb	--	60:00:00	72:00:00	--	25	14	14	E R
ihep-serial	--	144:00:0	160:00:0	--	0	0	20	E R
ops	--	01:00:00	01:00:00	--	0	0	5	E R
gpu-short	--	16:00:00	18:00:00	--	0	0	12	E R
ihep-special	--	00:10:00	00:20:00	--	0	0	20	E R
atlas	--	84:00:00	96:00:00	--	0	112	25	E R
					1066	3141		

- Over many years of operation IHEP gain expertise in operating Linux cluster with hundreds of nodes
- Our cluster architecture and chosen stack of software technologies show good
 $\text{Effectiveness} = \text{Availability} * \text{Reliability} * \text{Maintainability} * \text{Capability}$
- WLCG Grid infrastructure plays important role in the cluster development and maintenance and allows to use it for 100% in 24x7x365 mode

- More support for local experiments
- Hope for new experiments from mega projects based on IHEP infrastructure
- No drops for Grid support
 - It becomes a part of the Cluster
 - It allows us to be in sync with HEP software and technologies
- Stable operation with high availability and reliability without decreasing of supported resources
- No big investments in to the current infrastructure for near future



Thank you!

Any questions?