The JINR distributed information and computing environment: participants, features and challenges

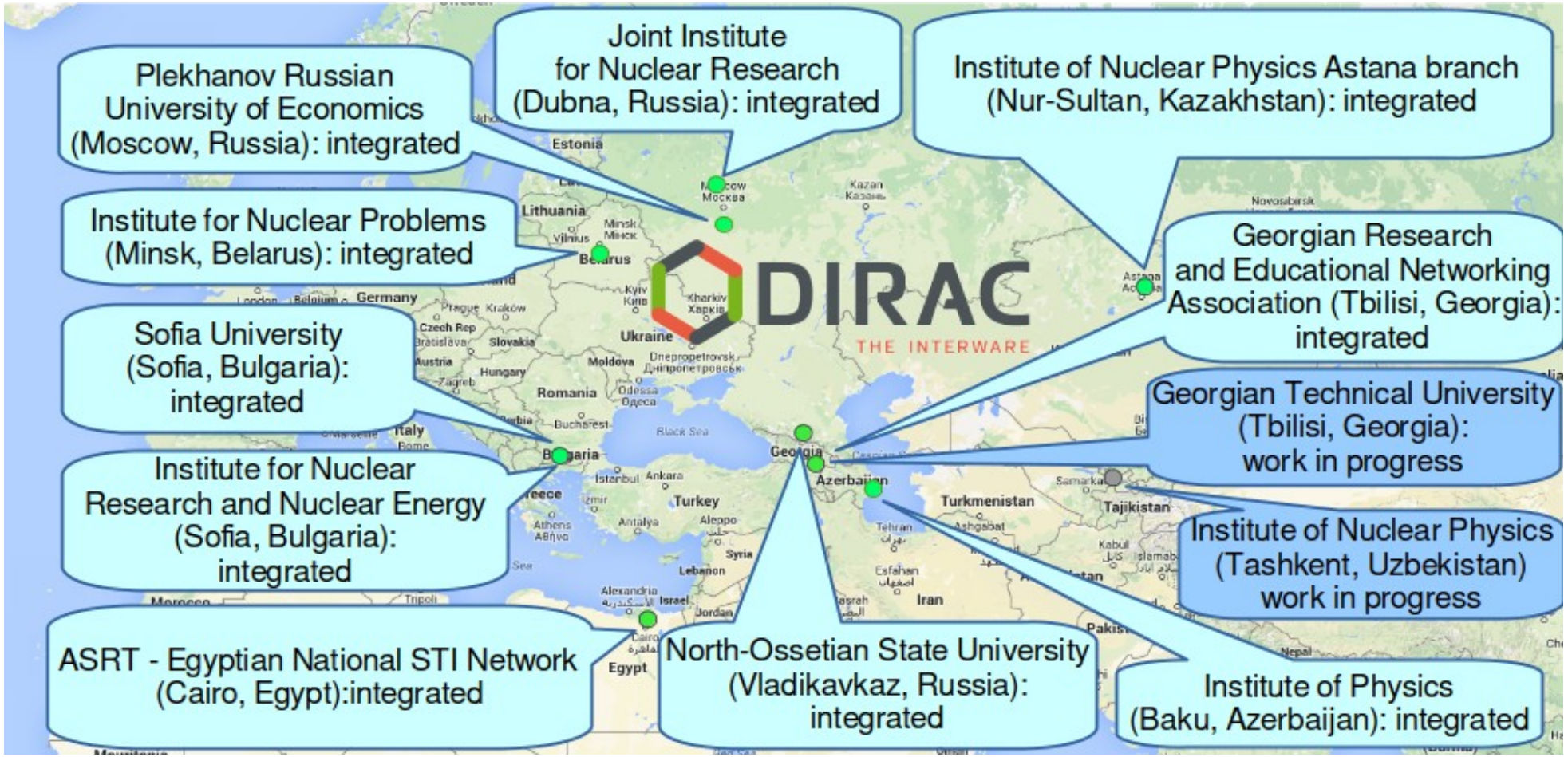
N.A. Balashov[1], I.S. Kuprikov[2], N.A. Kutovskiy[1], A.N. Makhalkin[1], Ye. Mazhitova[1,3], I.S. Pelevanyuk[1], R.N. Semenov[1,4]

[1] Laboratory of Information Technologies, Joint Institute for Nuclear Research
[2] Dubna State University, Dubna, Russia
[3] Institute of Nuclear Physics, Almaty, Kazakhstan
[4] Plekhanov Russian University of Economics, Moscow, Russia

The 9th International Conference "Distributed Computing and Grid-technologies in Science and Education" (GRID'2021)
July 5-9, 2021, JINR, Dubna, Russia

Clouds integration: participants

- To join resources for solving common tasks as well as to distribute a peak load across resources of partner organizations



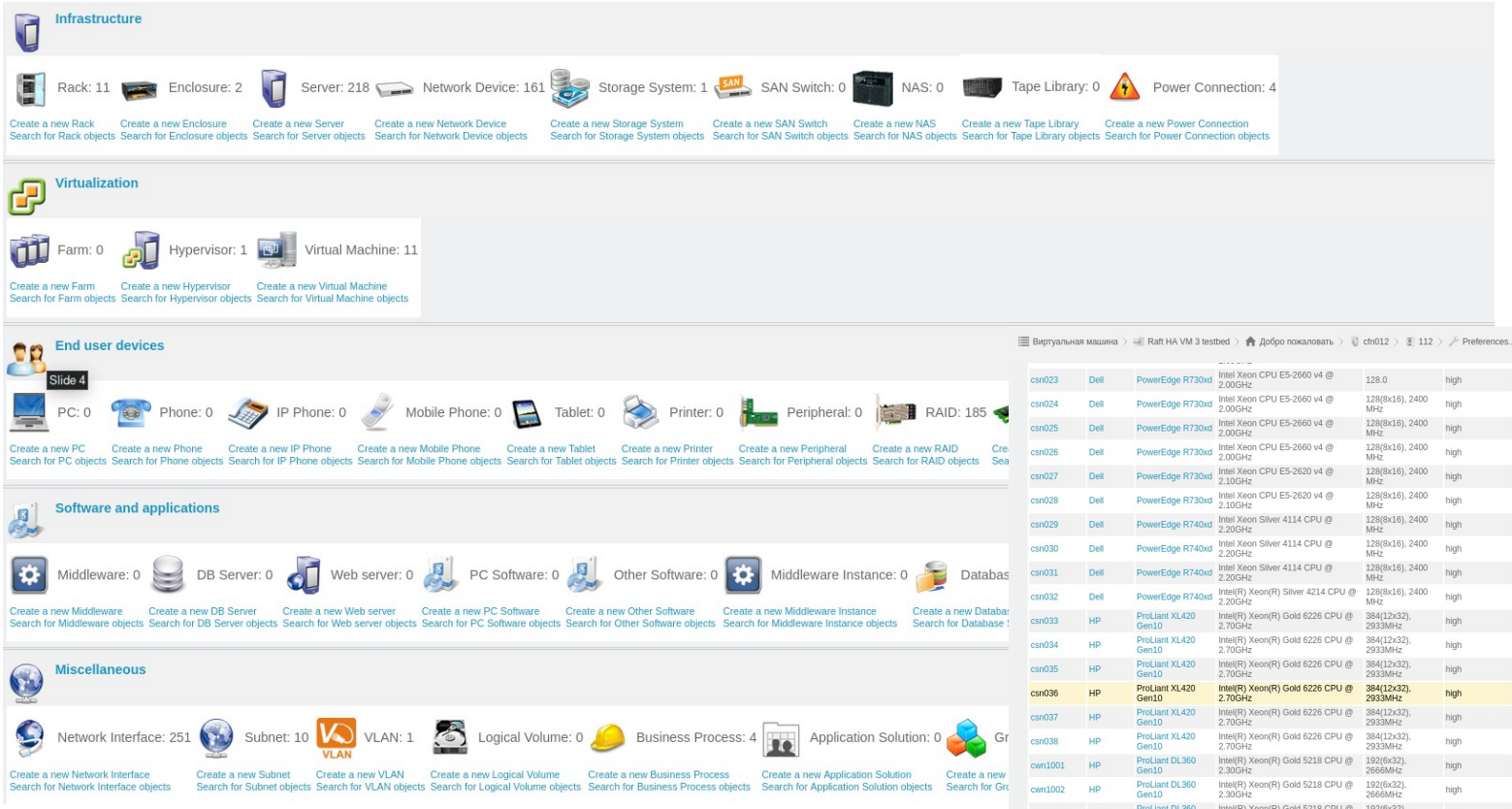
JINR distributed information and computing environment: resources

Organization	Country	Status	non-HT CPU cores	RAM, GB	Storage, TB
Plekhanov Russian Economic University	RU	integrated	132	608	51.1
Astana branch of the Institute of Nuclear Physics	KZ	integrated	84	840	6.8 (SSD)
Institute of Physics of the National Academy of Sciences of Azerbaijan	AZ	integrated	16	96	56
North Ossetian State University	RU	integrated	84	672	17
Academy of Scientific Research & Technology - Egyptian National STI Network	EG	integrated	98	704	13.8
Institute for Nuclear Research and Nuclear Energy	BG	integrated	20	64	4
St. Sophia University «St. Kliment Ohridski»	BG	integrated	48	250	4.7
Scientific Research Institute of Nuclear Problems of the Belarusian State University	BY	integrated	132	290	127
Institute of Nuclear Physics	UZ	in progress			
Georgian Technical University	GE	in progress	50	308	20
Total			664	3832	

Hardware inventory

Organizations:

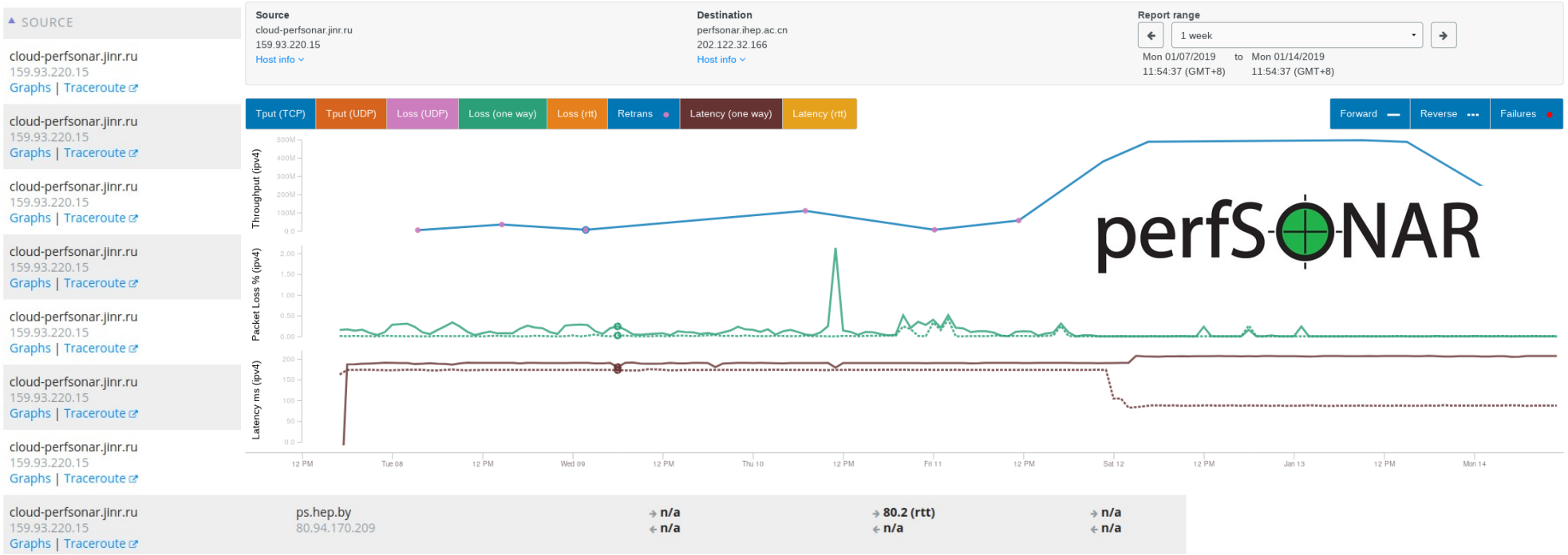
- ASRT
- GTU
- INP
- JINR
- NOSU



Virtual Machine	Organization	Status	Business criticity	IP	Move to production date	Name
NAT	АНИТ	production	high	195.43.16.35	2020-12-18	NAT
one-fn	ИЯФ	production	high	10.0.2.2	2020-07-02	one-fn
one-fn	СОГУ	production	high	10.55.0.10	2020-03-24	one-fn
one-fn	АНИТ	production	high	195.43.16.34	2020-12-14	one-fn
perfSONAR	АНИТ	production	high	195.43.16.36	2020-12-21	perfSONAR
Raft HA VM 1	ОИЯИ	production	high	10.220.0.101		Raft HA VM 1
Raft HA VM 1 testbed	ОИЯИ	production	low	10.220.31.243		Raft HA VM 1 testbed
Raft HA VM 2	ОИЯИ	production	high	10.220.0.102		Raft HA VM 2
Raft HA VM 2 testbed	ОИЯИ	production	low	10.220.31.244		Raft HA VM 2 testbed
Raft HA VM 3	ОИЯИ	production	high	10.220.0.103		Raft HA VM 3
Raft HA VM 3 testbed	ОИЯИ	production	low	10.220.31.245		Raft HA VM 3 testbed

PerfSONAR

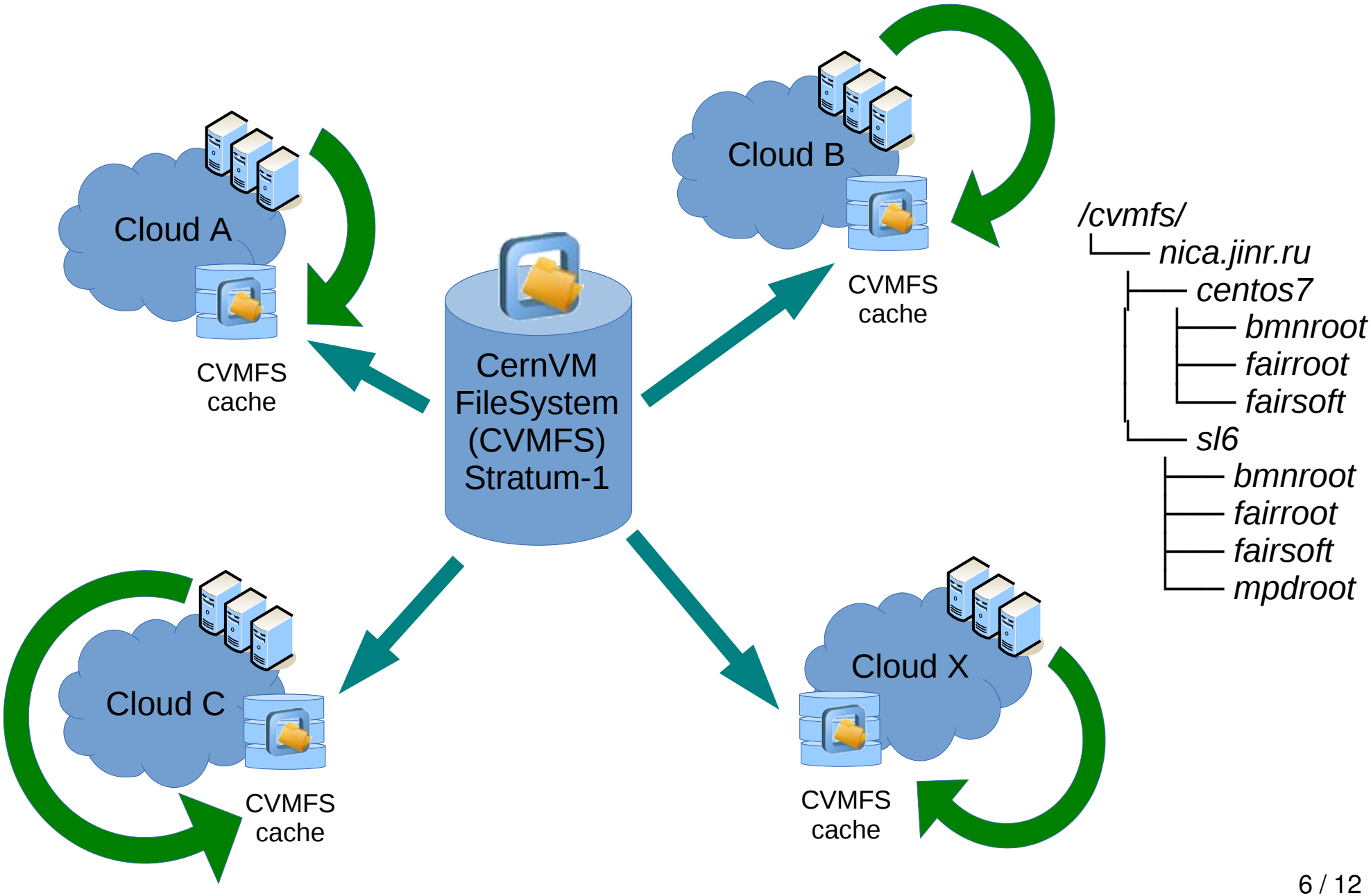
- To monitor network connectivity of participants
  - http://cloud-perfsonar.jinr.ru



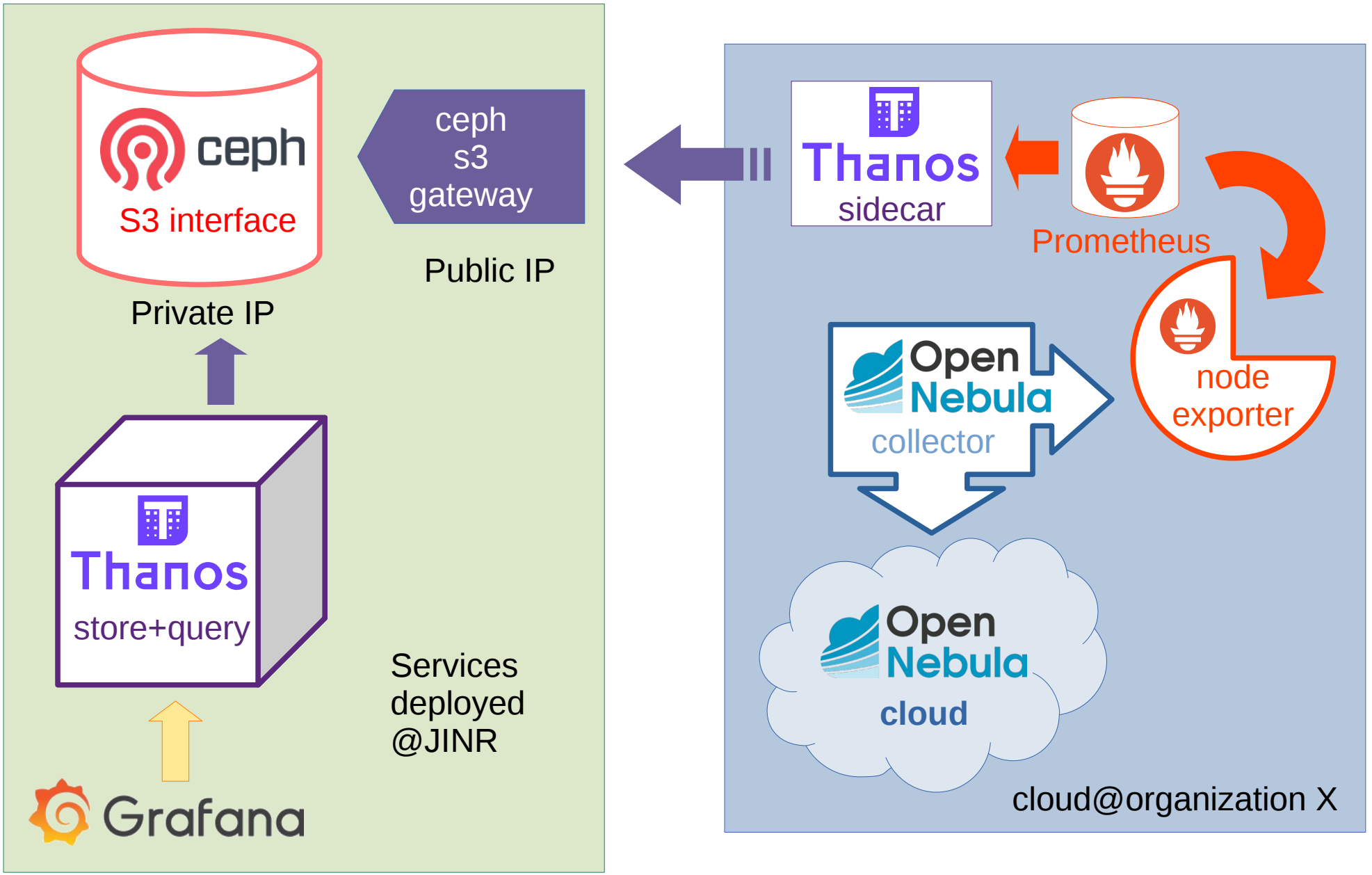
There is a challenge to deploy PS instance at some sites because all cloud VMs are behind NAT

Low external network **bandwidth** (e.g. 100 Mbps shared with the whole organization) is the main contributor into high CPU wall time of jobs
Most **suitable** type of jobs for such kind of resources is **MC with negligible input data**

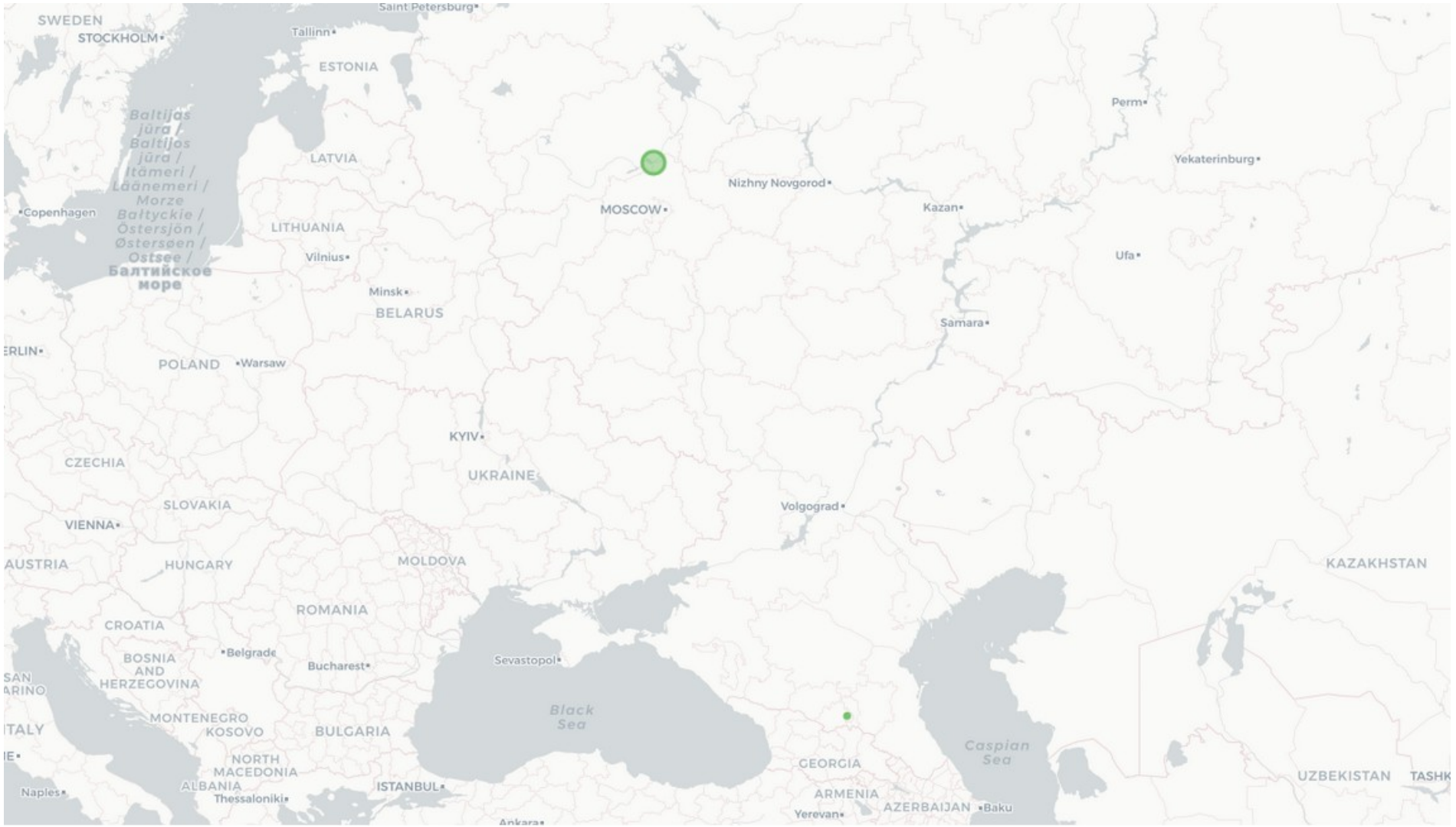
Experiments software distribution model



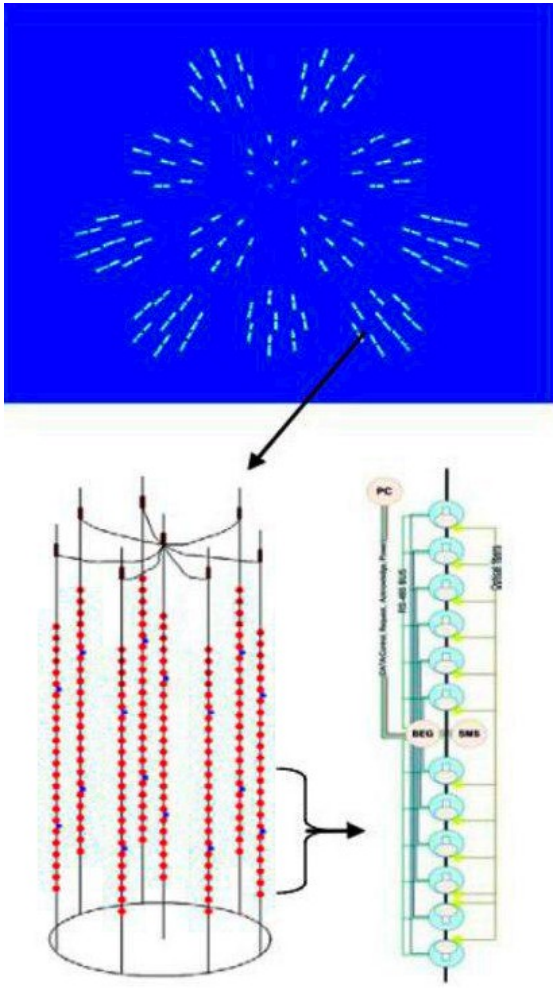
Metrics aggregation

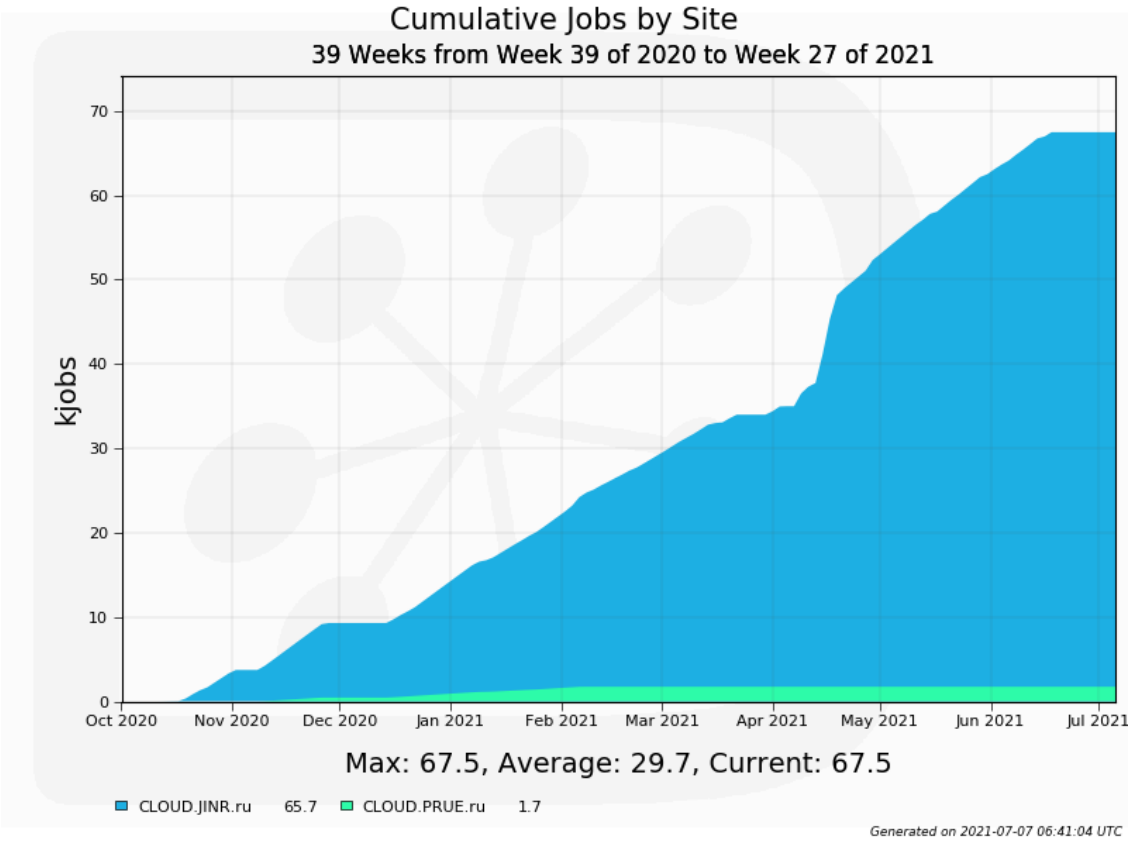


Grafana World Map plugin



Usage: Baikal-GVD

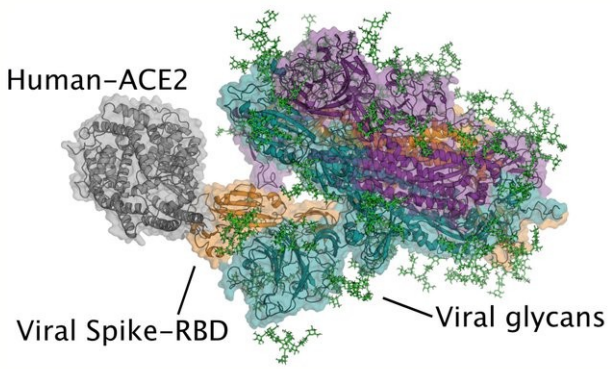


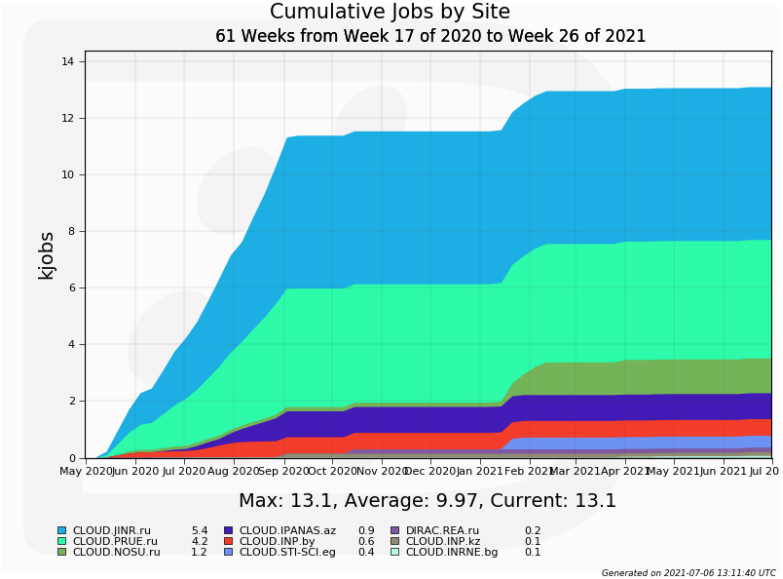


Main issue: 2.4 GB input file size for each job
For sites with low external network speed it's a problem.

Possible solution: to put it in CVMFS repository. In that case each job (apart from the first one) would download it with a speed of local network from CVMFS caching node.

Usage: SARS-CoV-2 research via F@H





FOLDING @HOME | STATISTICS — Home | Team | Donor | Project | OS

Team Statistics

Joint Institute for Nuclear Research

Rank 7,349

Top 10K Ranked Team

Team was founded by Igor Pelevanyuk and has earned 31,898,062 points by contibuting 13,418 work units.

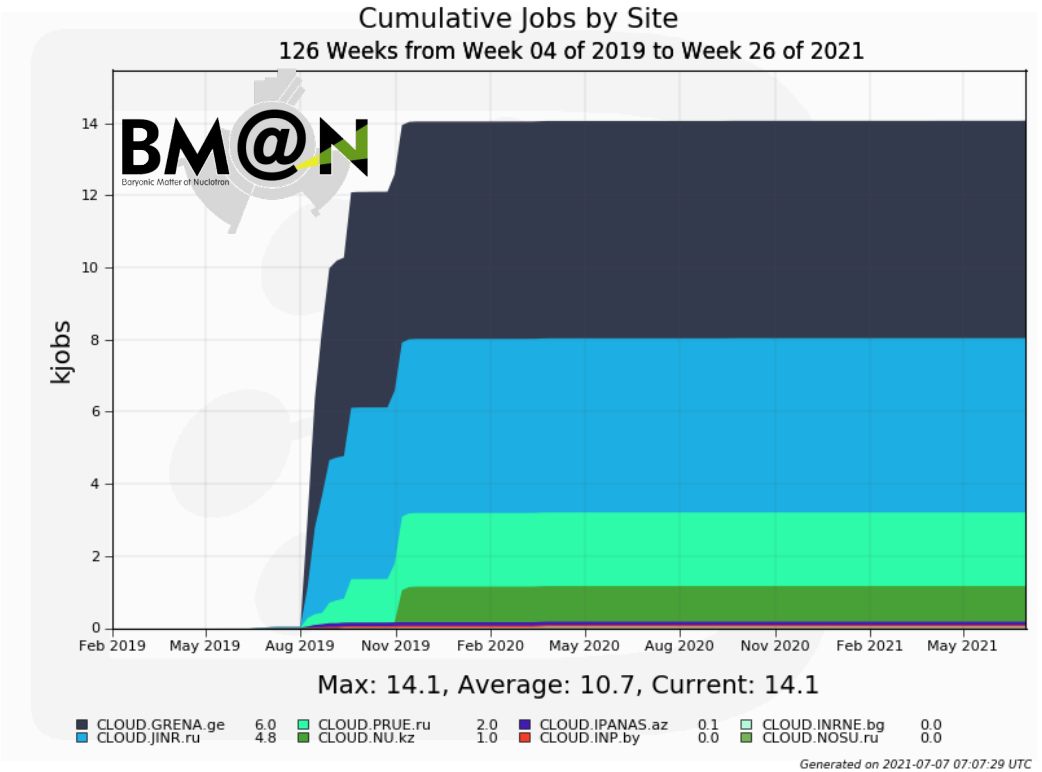
Team Website | ID: 265602

Team Donors

Rank	Donor Name		Score	WUs
89,114	CLOUD.JINR.ru	174385822	12,645,224	5,355
105,315	CLOUD.PRUE.ru	176451176	9,453,851	4,175
158,398	CLOUD.NOSU.ru	191767895	4,154,811	1,352
238,270	CLOUD.IPANAS.az	240838025	1,542,618	910
242,831	CLOUD.INP.by	176598058	1,465,167	599
269,612	CLOUD.INP.kz	336298929	1,101,339	433
298,521	CLOUD.STI-SCI.eg	451198000	817,728	381
356,701	DIRAC.REA-Parallel.ru	377146365	471,543	155
434,117	CLOUD.INRNE.bg	191781076	245,781	58

https://stats.foldingathome.org/team/265602

Usage: BM@N and SPD



BM@N workflow with simulation jobs was tested successfully.

Simulation and reconstruction SPD jobs was tested successfully as well but not yet on production storage (EOS)

Conclusion

- Number of JINR DICE participants and an amount of its resources is growing as well as number of users
- Only MC jobs with negligible input data are suitable for resources with low external network bandwidth (100 Mbps)
- Technical implementation of OpenNebula metrics aggregation is done. One needs to disseminate that experience to other JINR DICE clouds
- Migration from hand-drawn JINR DICE map to grafana World Map plugin is in progress