JOINT INSTITUTE FOR NUCLEAR RESEARCH

JINR

DUBNA

6th MPD Collaboration meeting

DIRAC Interware for the MPD experiment

Speaker: Igor Pelevanyuk

29 October 2020

Authors

DIRAC:

Igor Pelevanyk, Andrey Tsaregorodtzev

MPD:

Andrey Moshkin, Oleg Rogachevskiy

Responsibles for resources:

Cloud: Nikolay Kutovskiy

dCache: Vladimir Trofimov

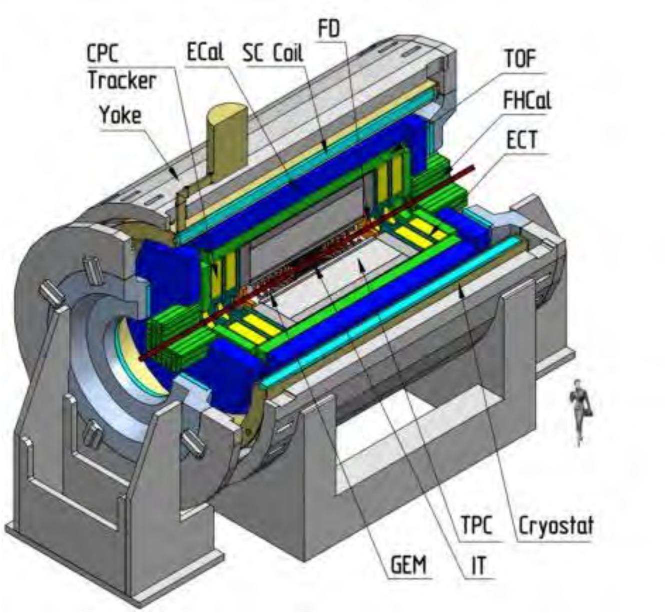
Govorun: Dmitry Podgainy

LHEP cluster: Boris Schinov

Tier-1,Tier-2, EOS: Valery Mitsyn

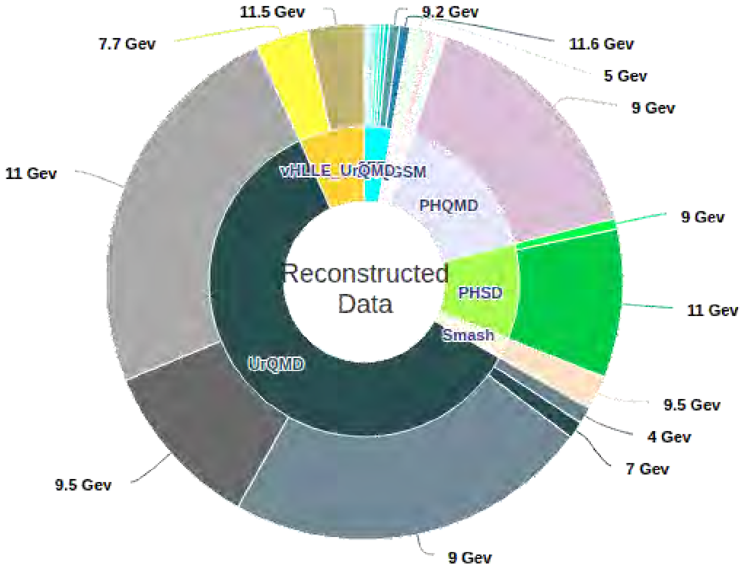
MPD MC generation

The MPD(Multi Purpose Detector) apparatus has been designed as a 4π spectrometer capable of detecting of charged hadrons, electrons and photons in heavy-ion collisions at high luminosity in the energy range of the NICA collider. To reach this goal, the detector will comprise a precise 3-D tracking system and a high-performance particle identification (PID) system based on the time-of-flight measurements and calorimetry.



MPD detector

Reconstructed Monte-Carlo events



MC generation math

In practice, to

generate 600M events

reconstruct 100M events

(size of reconstructed is around 1 MB)

You need to execute 500k jobs,

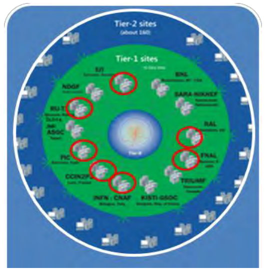
Each lasts for average 5.5 hours on one CPU core

(4-9 hours depending on the resource)

Notebook: ~80 years Server(24cores): ~13 years

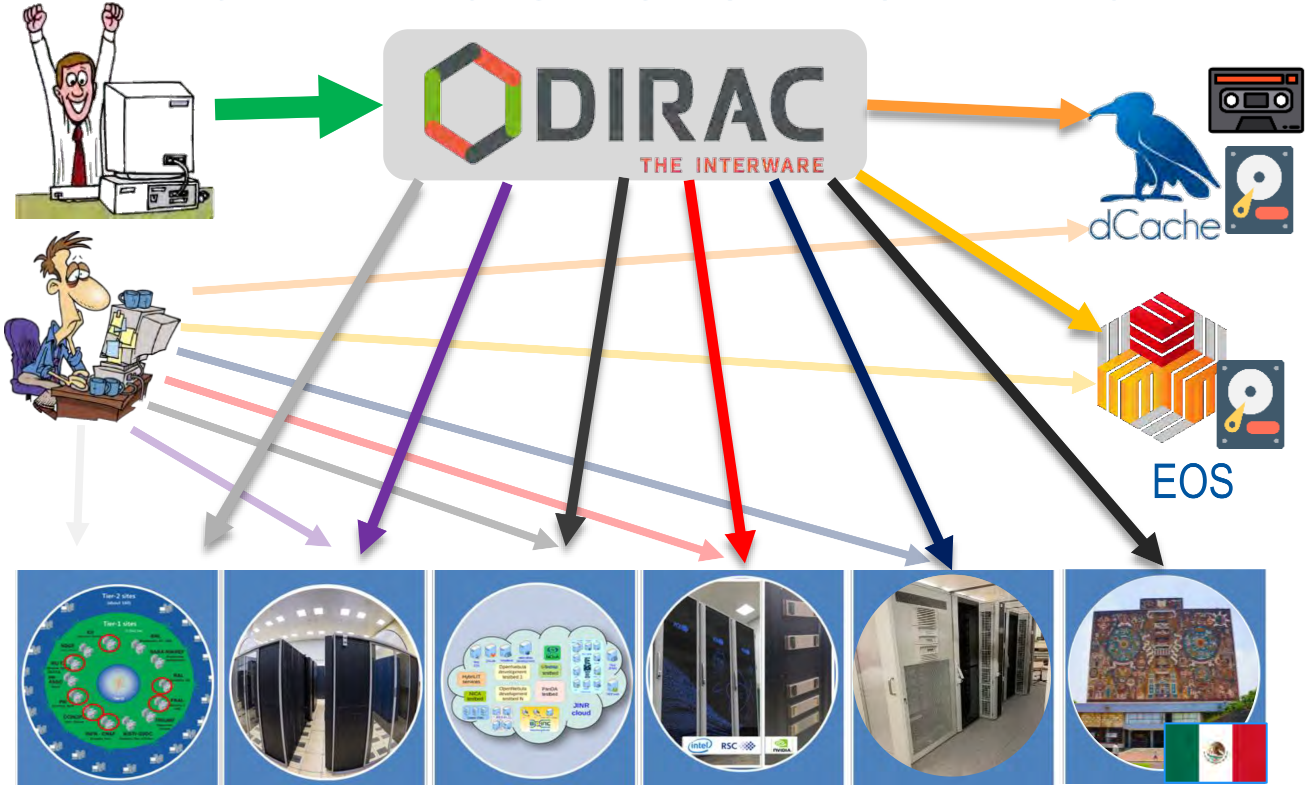
Cluster(10000 cores): ~ 11 days

MICC Resources

Storage	EOS		dCache Disk, Tape	ceph	lustre
Protocol	local, root		GridFTP, root	local	local
Auth Storage	Kerb. , x509		x509	ceph key	*HybriLIT*
Auth Jobs	Kerb.	x509	x509	SSO	*HybriLIT*
Job Submit.	Torque	Grid	Grid	OpenNebula	Slurm
Component				EOS OpenNebula development testbed 1 PanDA testbed NICA testbed DIRAC JINR cloud	
	Tier-2/CICC		**Tier-1**	**Cloud**	**Govorun/HybriLIT**

*** This is a simplified schema to demonstrate complexity and variability of protocols and accesses approaches**

What was done in JINR

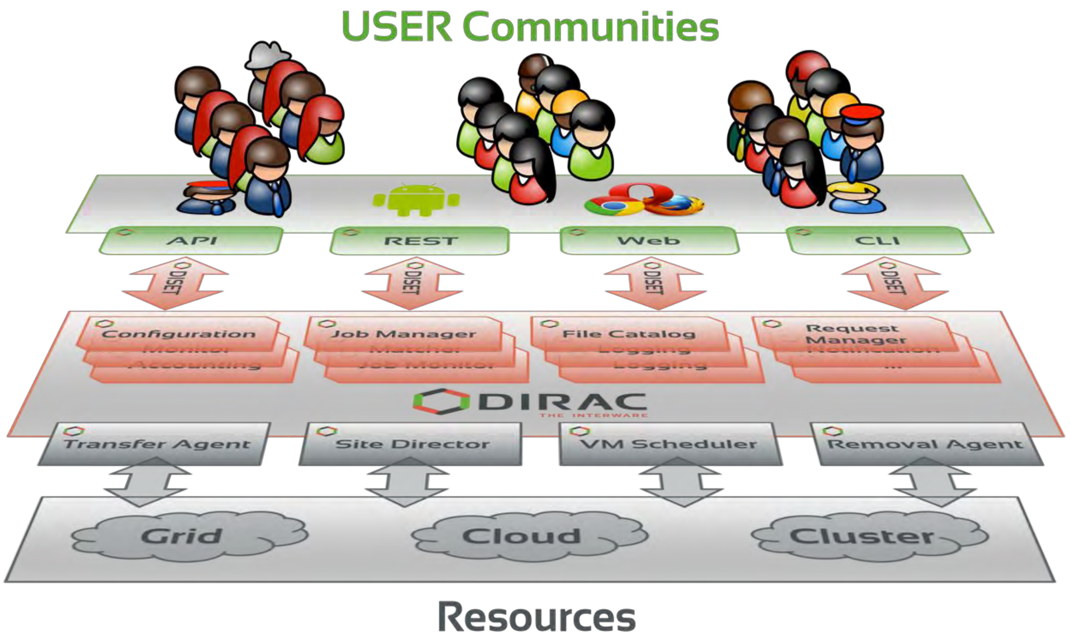


Tier-1	CICC/Tier-2	Clouds	Govorun	NICA Cluster	UNAM
Running	**Running**	**Running**	**Running**	**Running**	**Running**

The computing resources of the JINR Multifunctional Information and Computing Complex, clouds in JINR Member-States, cluster from Mexico University were combined using the DIRAC Interware.

What is DIRAC?

DIRAC provides all the necessary components to build ad-hoc grid infrastructures **interconnecting** computing resources of different types, allowing **interoperability** and simplifying **interfaces**. This allows to speak about the DIRAC *interware*.



Web

CLI

API

REST

Why DIRAC?

1. Single system for all aspects of computing

DIRAC
THE INTERWARE

User Interface

API

Central configuration

Workload management

Data management

Integration tools

File Catalog

Workflow management

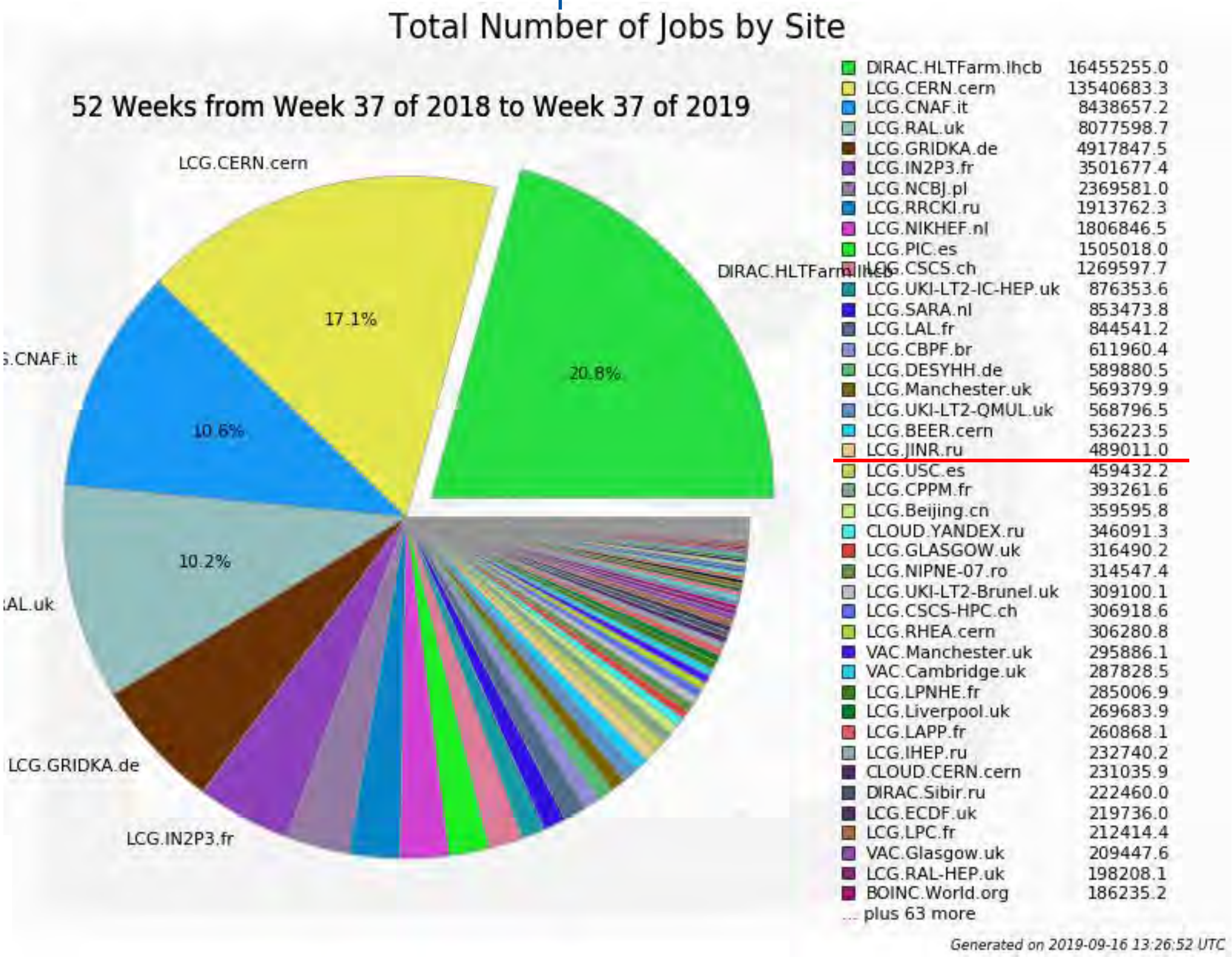
Metadata management

Accounting

Management

Why DIRAC?

2. Good performance



🟩	DIRAC.HLTFarm.lhcb	16455255.0
🟨	LCG.CERN.cern	13540683.3
🟦	LCG.CNAF.it	8438657.2
■	LCG.RAL.uk	8077598.7
🟫	LCG.GRIDKA.de	4917847.5
🟪	LCG.IN2P3.fr	3501677.4
■	LCG.NCBJ.pl	2369581.0
🟦	LCG.RRCKI.ru	1913762.3
🟪	LCG.NIKHEF.nl	1806846.5
🟩	LCG.PIC.es	1505018.0
■	LCG.CSCS.ch	1269597.7
■	LCG.UKI-LT2-IC-HEP.uk	876353.6
🟦	LCG.SARA.nl	853473.8
■	LCG.LAL.fr	844541.2
■	LCG.CBPF.br	611960.4
🟩	LCG.DESYHH.de	589880.5
🟫	LCG.Manchester.uk	569379.9
🟦	LCG.UKI-LT2-QMUL.uk	568796.5
🟦	LCG.BEER.cern	536223.5
🟧	LCG.JINR.ru	489011.0
🟨	LCG.USC.es	459432.2
■	LCG.CPPM.fr	393261.6
🟩	LCG.Beijing.cn	359595.8
🟦	CLOUD.YANDEX.ru	346091.3
🟥	LCG.GLASGOW.uk	316490.2
🟩	LCG.NIPNE-07.ro	314547.4
■	LCG.UKI-LT2-Brunel.uk	309100.1
🟦	LCG.CSCS-HPC.ch	306918.6
🟩	LCG.RHEA.cern	306280.8
🟦	VAC.Manchester.uk	295886.1
🟦	VAC.Cambridge.uk	287828.5
🟩	LCG.LPNHE.fr	285006.9
🟩	LCG.Liverpool.uk	269683.9
🟥	LCG.LAPP.fr	260868.1
■	LCG.IHEP.ru	232740.2
🟪	CLOUD.CERN.cern	231035.9
■	DIRAC.Sibir.ru	222460.0
🟪	LCG.ECDF.uk	219736.0
🟫	LCG.LPC.fr	212414.4
🟪	VAC.Glasgow.uk	209447.6
🟪	LCG.RAL-HEP.uk	198208.1
🟪	BOINC.World.org	186235.2
...	plus 63 more	

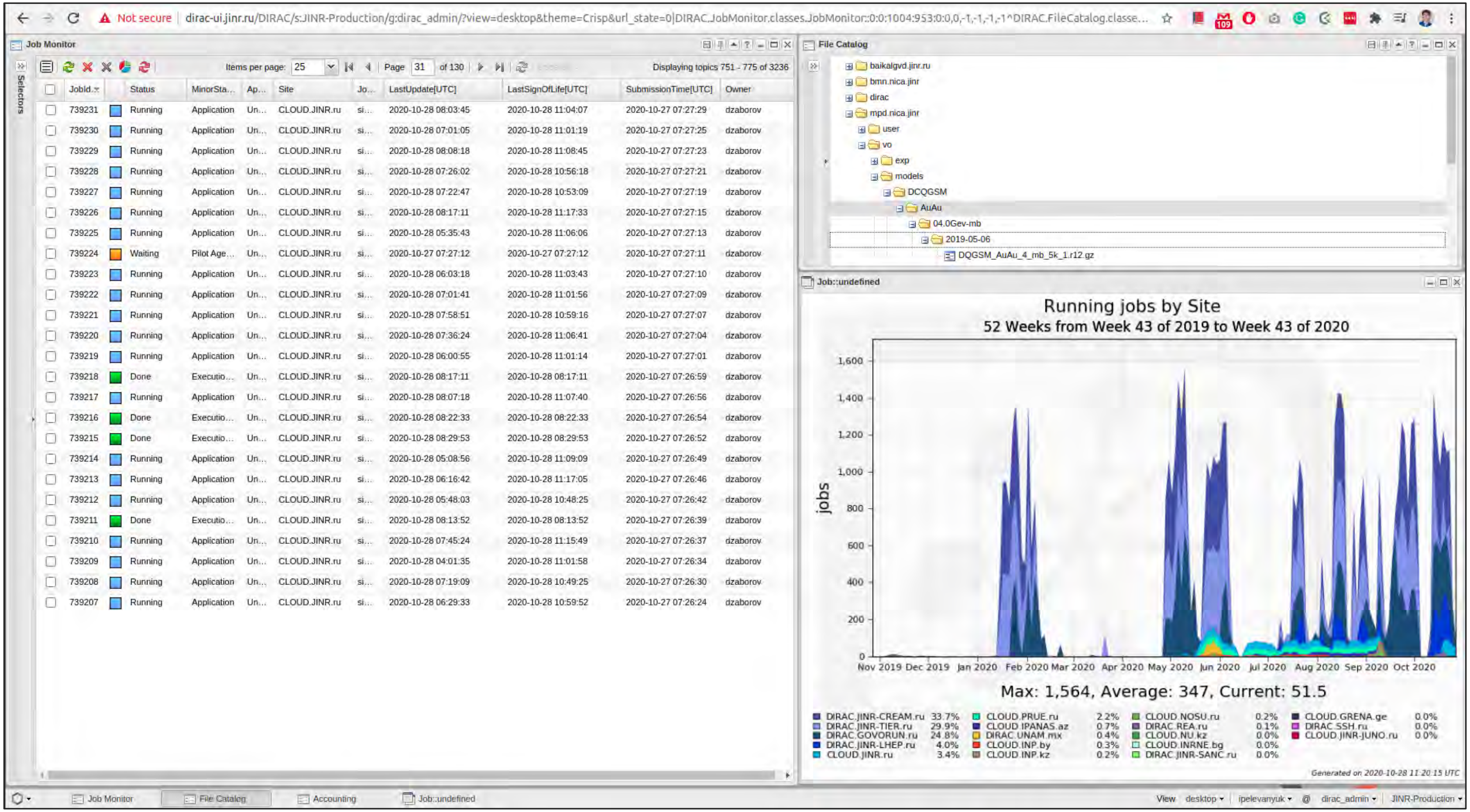
Generated on 2019-09-16 13:26:52 UTC

Why DIRAC?

3. Active users and developers community

- Dedicated installations
  - LHCb, Belle II, CTA
- Multi-community services
  - ILC, CALICE
  - IHEP: BES III, Juno, CEPC
  - FG-DIRAC
  - GridPP
  - DIRAC4EGI
  - PNNL
  - DIRAC@JINR
  - DIRAC@CNAF
- Several DIRAC evaluations are ongoing
  - Auger, ELI, NICA, Virgo, LSST, …

User Interface



Not secure | dirac-ui.jinr.ru/DIRAC/s:JINR-Production/g:dirac_admin/?view=desktop&theme=Crisp&url_state=0|DIRAC.JobMonitor.classes.JobMonitor::0:0:1004:953:0:0,0,-1,-1,-1,-1^DIRAC.FileCatalog.classe...

Job Monitor

Items per page: 25 | Page 31 of 130 | Displaying topics 751 - 775 of 3236

□	JobId		Status	MinorSta...	Ap...	Site	Jo...	LastUpdate[UTC]	LastSignOfLife[UTC]	SubmissionTime[UTC]	Owner
□	739231	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 08:03:45	2020-10-28 11:04:07	2020-10-27 07:27:29	dzaborov
□	739230	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 07:01:05	2020-10-28 11:01:19	2020-10-27 07:27:25	dzaborov
□	739229	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 08:08:18	2020-10-28 11:08:45	2020-10-27 07:27:23	dzaborov
□	739228	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 07:26:02	2020-10-28 10:56:18	2020-10-27 07:27:21	dzaborov
□	739227	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 07:22:47	2020-10-28 10:53:09	2020-10-27 07:27:19	dzaborov
□	739226	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 08:17:11	2020-10-28 11:17:33	2020-10-27 07:27:15	dzaborov
□	739225	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 05:35:43	2020-10-28 11:06:06	2020-10-27 07:27:13	dzaborov
□	739224	🟧	Waiting	Pilot Age...	Un...	CLOUD.JINR.ru	si...	2020-10-27 07:27:12	2020-10-27 07:27:12	2020-10-27 07:27:11	dzaborov
□	739223	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 06:03:18	2020-10-28 11:03:43	2020-10-27 07:27:10	dzaborov
□	739222	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 07:01:41	2020-10-28 11:01:56	2020-10-27 07:27:09	dzaborov
□	739221	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 07:58:51	2020-10-28 10:59:16	2020-10-27 07:27:07	dzaborov
□	739220	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 07:36:24	2020-10-28 11:06:41	2020-10-27 07:27:04	dzaborov
□	739219	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 06:00:55	2020-10-28 11:01:14	2020-10-27 07:27:01	dzaborov
□	739218	🟩	Done	Executio...	Un...	CLOUD.JINR.ru	si...	2020-10-28 08:17:11	2020-10-28 08:17:11	2020-10-27 07:26:59	dzaborov
□	739217	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 08:07:18	2020-10-28 11:07:40	2020-10-27 07:26:56	dzaborov
□	739216	🟩	Done	Executio...	Un...	CLOUD.JINR.ru	si...	2020-10-28 08:22:33	2020-10-28 08:22:33	2020-10-27 07:26:54	dzaborov
□	739215	🟩	Done	Executio...	Un...	CLOUD.JINR.ru	si...	2020-10-28 08:29:53	2020-10-28 08:29:53	2020-10-27 07:26:52	dzaborov
□	739214	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 05:08:56	2020-10-28 11:09:09	2020-10-27 07:26:49	dzaborov
□	739213	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 06:16:42	2020-10-28 11:17:05	2020-10-27 07:26:46	dzaborov
□	739212	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 05:48:03	2020-10-28 10:48:25	2020-10-27 07:26:42	dzaborov
□	739211	🟩	Done	Executio...	Un...	CLOUD.JINR.ru	si...	2020-10-28 08:13:52	2020-10-28 08:13:52	2020-10-27 07:26:39	dzaborov
□	739210	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 07:45:24	2020-10-28 11:15:49	2020-10-27 07:26:37	dzaborov
□	739209	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 04:01:35	2020-10-28 11:01:58	2020-10-27 07:26:34	dzaborov
□	739208	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 07:19:09	2020-10-28 10:49:25	2020-10-27 07:26:30	dzaborov
□	739207	🟦	Running	Application	Un...	CLOUD.JINR.ru	si...	2020-10-28 06:29:33	2020-10-28 10:59:52	2020-10-27 07:26:24	dzaborov

File Catalog

```
⊞ baikalgvd.jinr.ru
⊞ bmn.nica.jinr
⊞ dirac
⊟ mpd.nica.jinr
   ⊞ user
   ⊟ vo
      ⊞ exp
      ⊟ models
         ⊟ DCQGSM
            ⊟ AuAu
               ⊟ 04.0Gev-mb
                  ⊟ 2019-05-06
                     DQGSM_AuAu_4_mb_5k_1.r12.gz
```

Job::undefined

Running jobs by Site

52 Weeks from Week 43 of 2019 to Week 43 of 2020

Max: 1,564, Average: 347, Current: 51.5

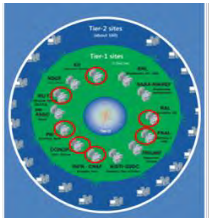
DIRAC.JINR-CREAM.ru	33.7%	CLOUD.PRUE.ru	2.2%	CLOUD.NOSU.ru	0.2%	CLOUD.GRENA.ge	0.0%
DIRAC.JINR-TIER.ru	29.9%	CLOUD.IPANAS.az	0.7%	DIRAC.REA.ru	0.1%	DIRAC.SSH.ru	0.0%
DIRAC.GOVORUN.ru	24.8%	DIRAC.UNAM.mx	0.4%	CLOUD.NU.kz	0.0%	CLOUD.JINR-JUNO.ru	0.0%
DIRAC.JINR-LHEP.ru	4.0%	CLOUD.INP.by	0.3%	CLOUD.INRNE.bg	0.0%		
CLOUD.JINR.ru	3.4%	CLOUD.INP.kz	0.2%	DIRAC.JINR-SANC.ru	0.0%		

Generated on 2020-10-28 11:20:15 UTC

Workload management

Submit thousand of jobs to DIRAC Job Queue

DIRAC
THE INTERWARE


Tier-1

CICC/Tier-2

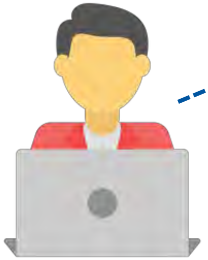
Clouds

Govorun

NICA Cluster

UNAM

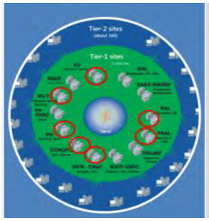
Workload management



Submit thousand of jobs to DIRAC Job Queue

User
User
User
User
User
User
User
User
User Job

DIRAC
THE INTERWARE



Tier-1

CICC/Tier-2

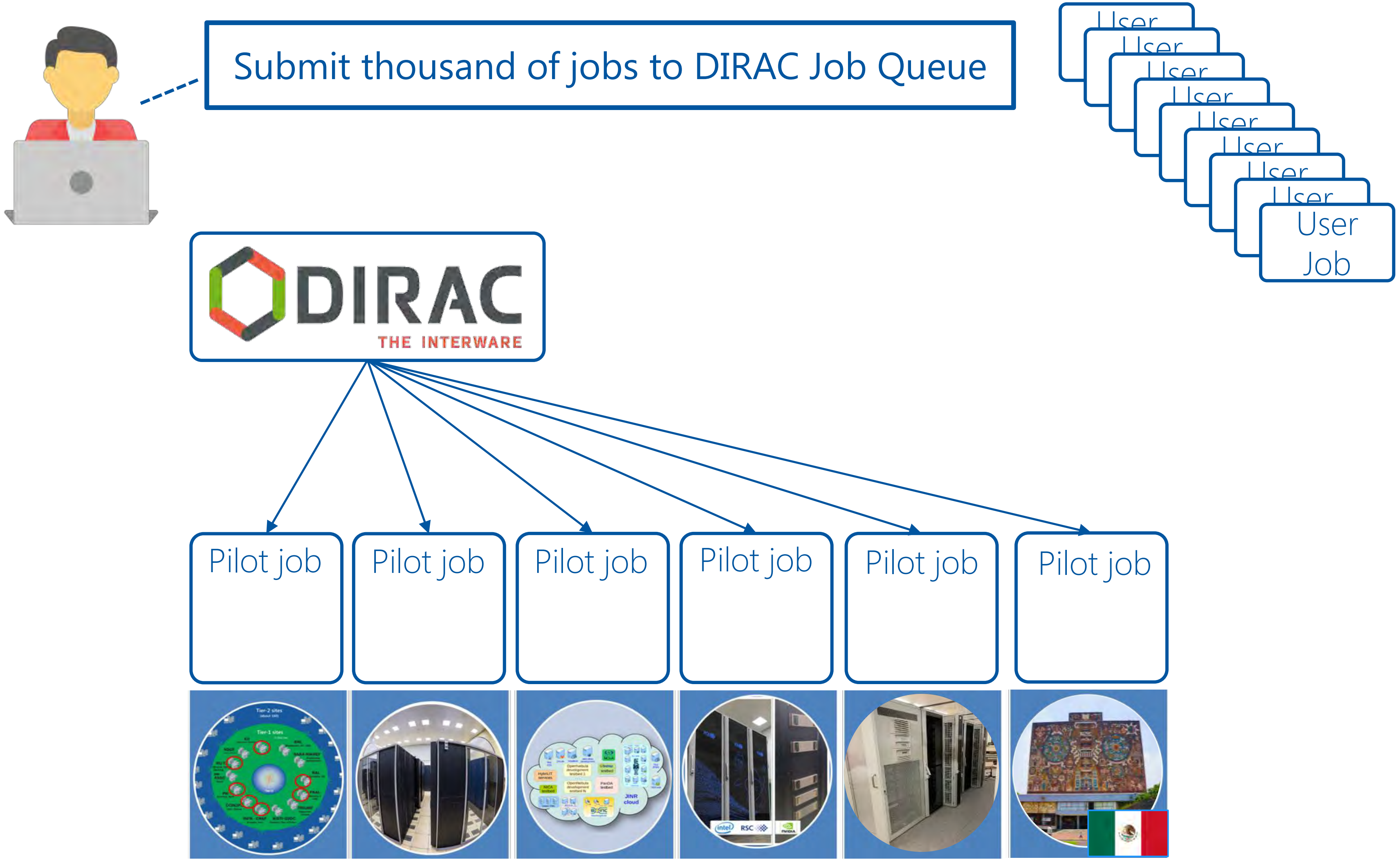
Clouds

Govorun

NICA Cluster

UNAM

Workload management



Submit thousand of jobs to DIRAC Job Queue

User Job

User Job

User Job

User Job

User Job

User Job

User Job

User Job

User Job

Pilot job

Pilot job

Pilot job

Pilot job

Pilot job

Pilot job

Tier-1

CICC/Tier-2

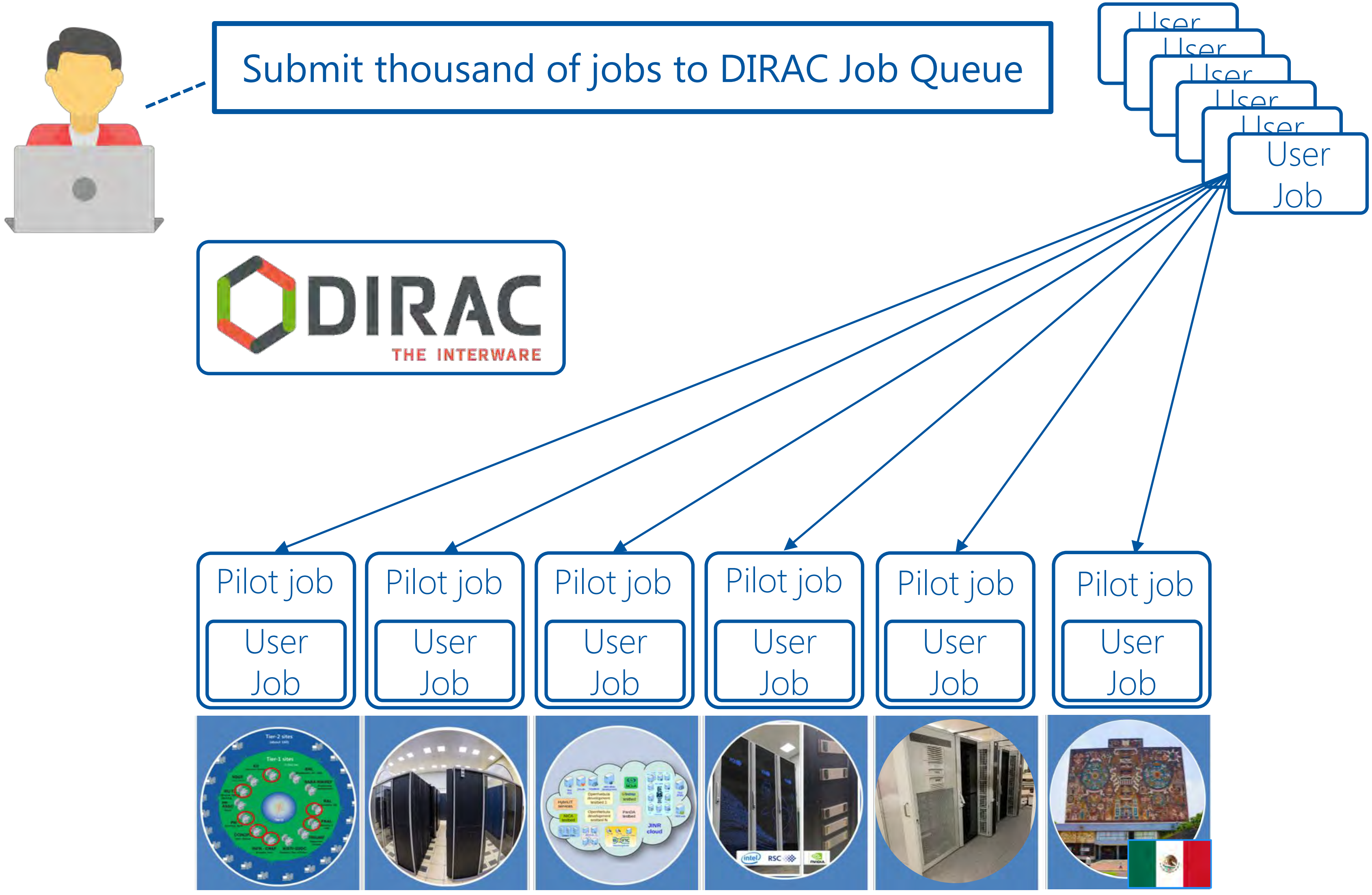
Clouds

Govorun

NICA Cluster

UNAM

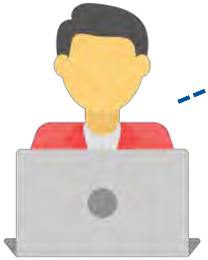
Workload management

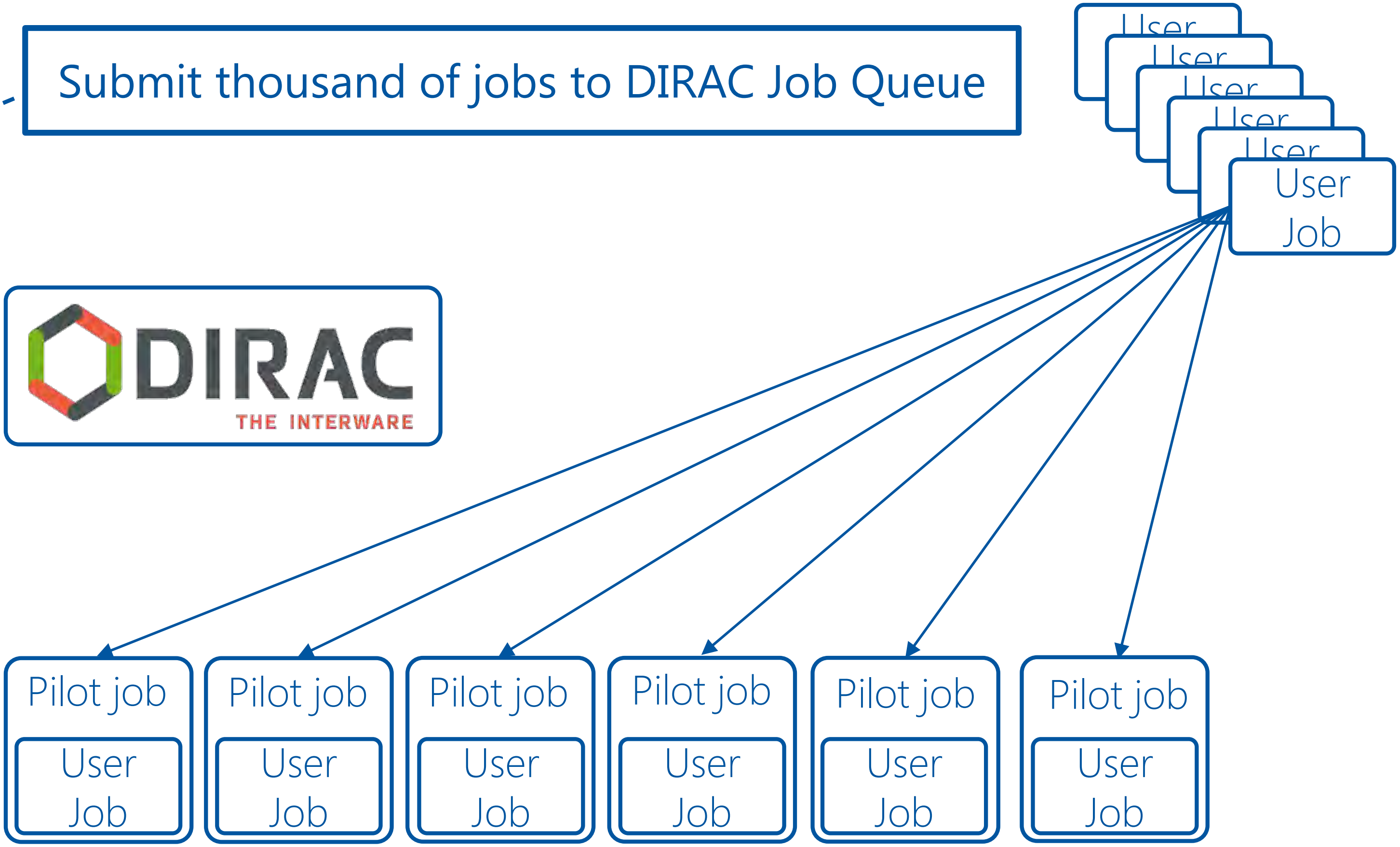


Submit thousand of jobs to DIRAC Job Queue

Tier-1 CICC/Tier-2 Clouds Govorun NICA Cluster UNAM

Workload management

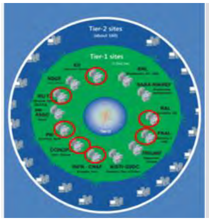




Submit thousand of jobs to DIRAC Job Queue

User
Job

Pilot job	Pilot job	Pilot job	Pilot job	Pilot job	Pilot job
User Job	User Job	User Job	User Job	User Job	User Job



Tier-1

CICC/Tier-2

Clouds

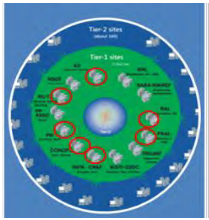
Govorun

NICA Cluster

UNAM

MPD Computing resources

JINR



Tier-1
595 slots
Quota increased

CICC/Tier-2
500 slots

Clouds
100 slots

Govorun
184-470 slots

NICA Cluster
250 slots
(single user quota)

MPD collaboration

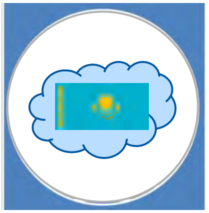
UNAM
100 slots

Member-states clouds

IPANAS
20 slots

INP
10 slots

INRNE
3 slots



INP
50 slots

REA Plehanova
40 slots

NOSU
60 slots

Quotas in different resources may be increased in case of successful and effective usage.

Workload management

1. Initial configuration

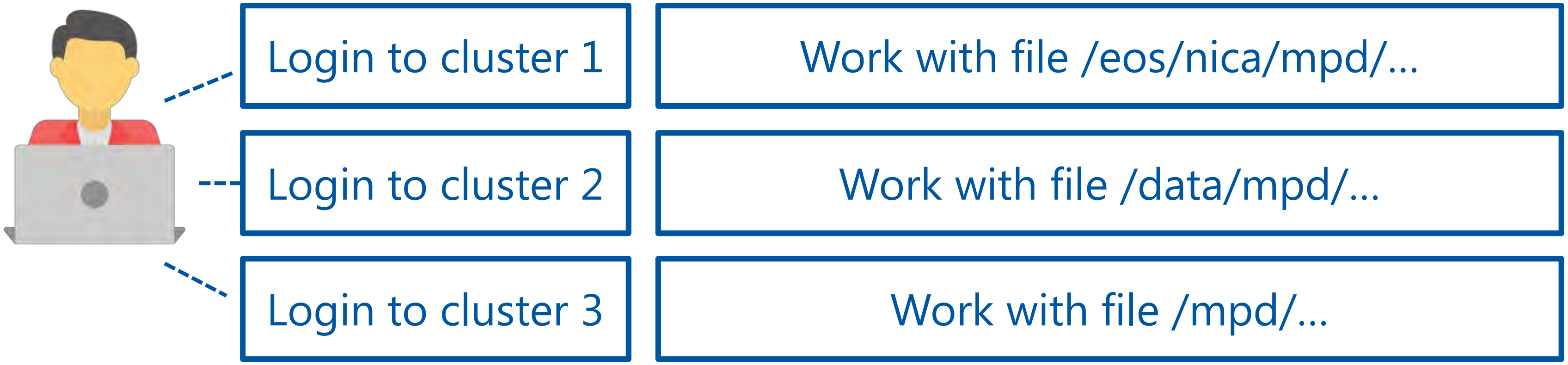
2. Input data download

3. Processing

4. Output data upload

5.Finalization

Data access



Login to cluster 1 — Work with file /eos/nica/mpd/…

Login to cluster 2 — Work with file /data/mpd/…

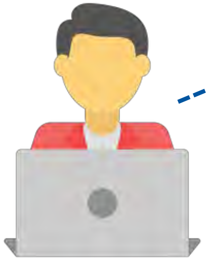
Login to cluster 3 — Work with file /mpd/…

Issues:

1. Different path to files on different clusters.
2. User need to remember that path names.
3. And keep track where different files exist.

Tier1/2	EOS MICC	Govorun	EOS HybriLIT	EOS LHEP	Cloud
dCache Tapes Disks	EOS	lustre	EOS	EOS	ceph

DIRAC File catalog

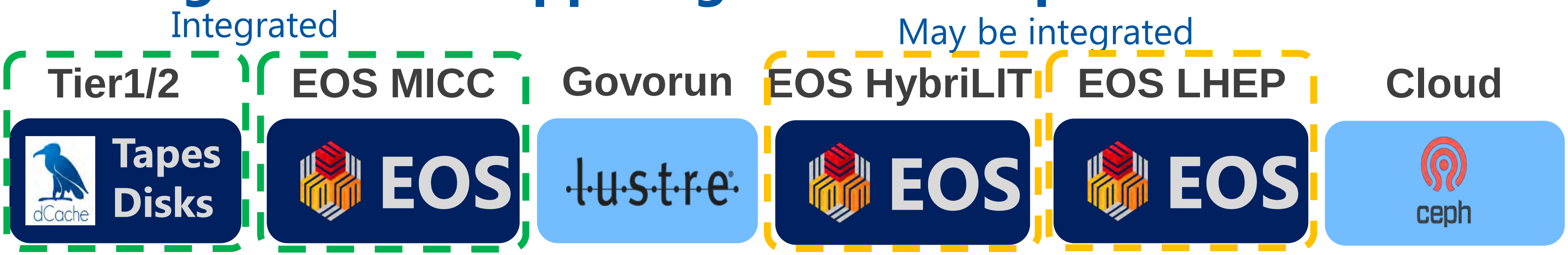


File catalog, give me file /mpd/...

Logical name	Physical name
/mpd/file1	root://eos.jinr.ru:1094:/eos/nica/mpd/file1
/mpd/file2	root://eos.jinr.ru:1094:/eos/nica/mpd/file2
/mpd/file2	srm://lxse-dc01.jinr.ru:8443/pnfs/jinr.ru/data/file2
/mpd/file3	srm://lxse-dc01.jinr.ru:8443/pnfs/jinr.ru/data/file3

Same file

DIRAC File Catalog provide single namespace for all files and replicas across different storage systems. To be used storages should support grid transfer protocols.



Integrated

May be integrated

Tier1/2

Tapes
Disks

EOS MICC

Govorun

EOS HybriLIT

EOS LHEP

Cloud

Metadata Management + File catalog

Logical name	Physical name
/mpd/file1	root://eos.jinr.ru:1094:/eos/nica/mpd/model/DQGSM/v4_3 /Au/Au/7GeV/2020-05-06/2k_001.root

Traditionally, a lot of information about data coded in file names. It is not straight forward how to work with this data, especially in case of complex searches and filter requests.

Logical name	Metadata name	Metadata value
/mpd/file1	type	model
/mpd/file1	generator	DQGSM
/mpd/file1	version	4.3
/mpd/file1	beam	Au
/mpd/file1	target	Au
/mpd/file1	energy	7.0
/mpd/file1	events	2000

```
dirac-find /mpd/model LastAccess < 01-10-2020 \\
             GaussVersion=v1,v2 SE=EOS-MPD Name=*.root
```

The use of metadata provide tool for efficient search and filtering of the data.

Workflow Management

DIRAC provide tools for automatization of different processes

1. Create metadata selector for finding files with right metadata

```
Path=/mpd/raw type=raw reconstructed=false Name=*.raw
```

2. Create job template

```
                    process_raw.sh <job_args*>
          (1. Run reconstruction script for data from job_args
    2. In case of success change “reconstructed” metadata to “true”)
```

3. Add files to file catalog and attach metadata

```
dirac-dms-add-file /mpd/raw/run7/6TeV_102.raw 6Tev_102.raw JINR-EOS-MPD
dirac-add-metadata /mpd/raw/run7/6TeV_102.raw reconstructed=false
                                              recoVersion=7.6
```

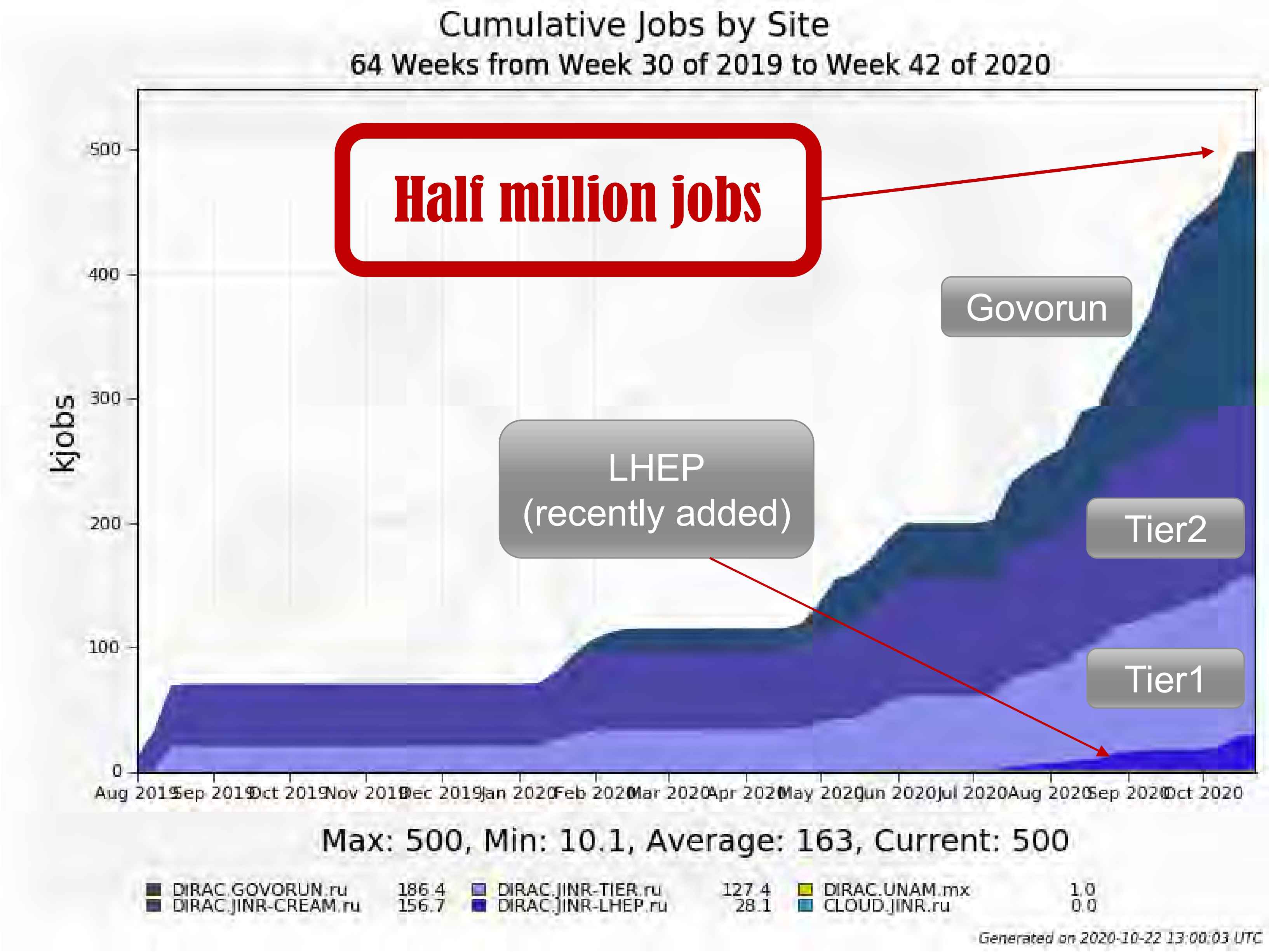
- DIRAC will automatically notice new data satisfying query from point 1.
- For every new file with satisfying query submit job:

```
process_raw.sh <filename> <software version> <other args…>
```

Disclaimer for further slides

- Further statistics applied only for MPD **centralized mass-production** submitted only **via DIRA**C in JINR.
- Difference between resources mostly due to the fact that some resources integrated by DIRAC for longer time.
- Computing power of different components mostly determined by quotas on the resources.

MPD Jobs Total



Half million jobs

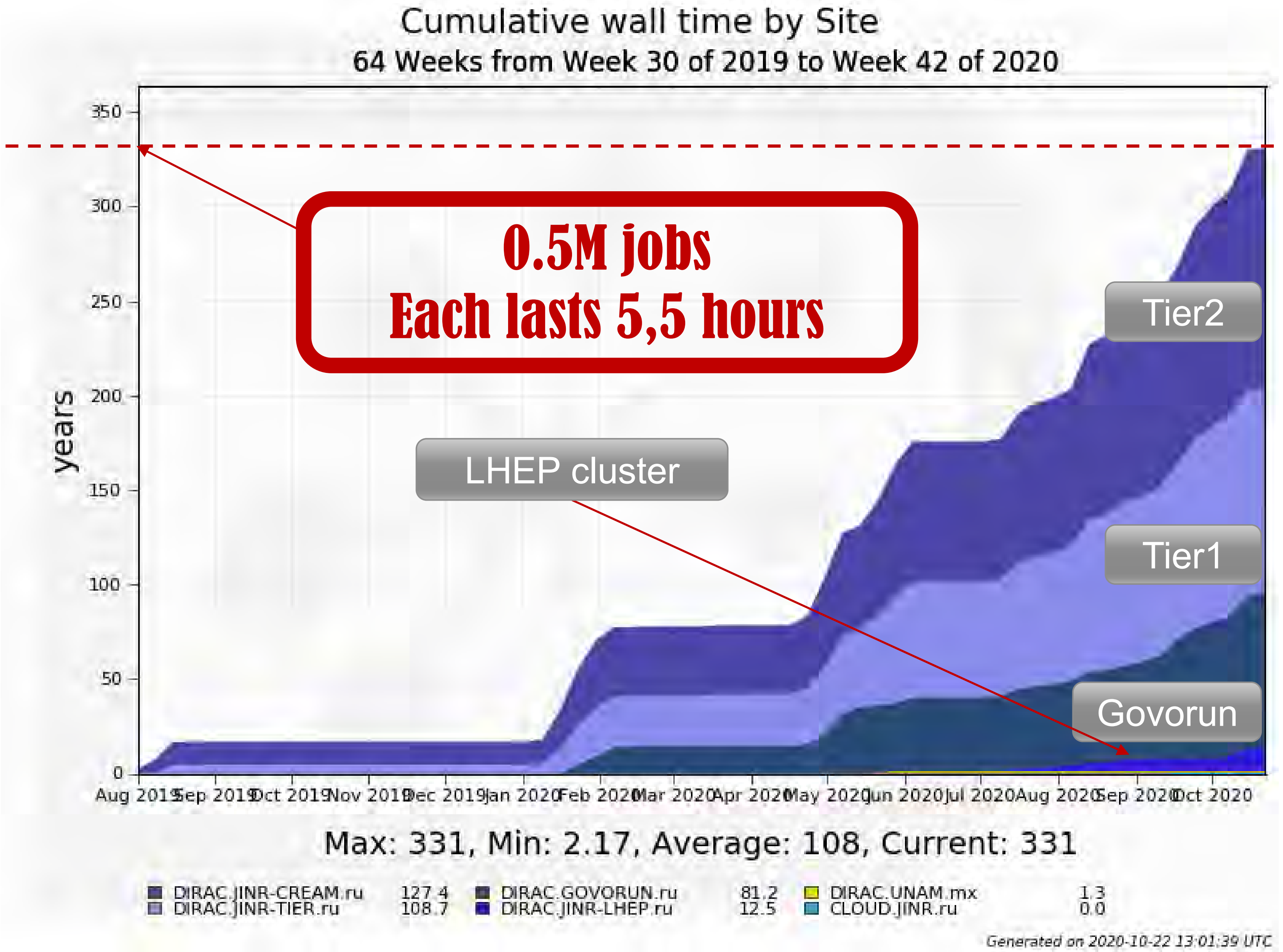
Govorun

LHEP
(recently added)

Tier2

Tier1

MPD Wall time



0.5M jobs
Each lasts 5,5 hours

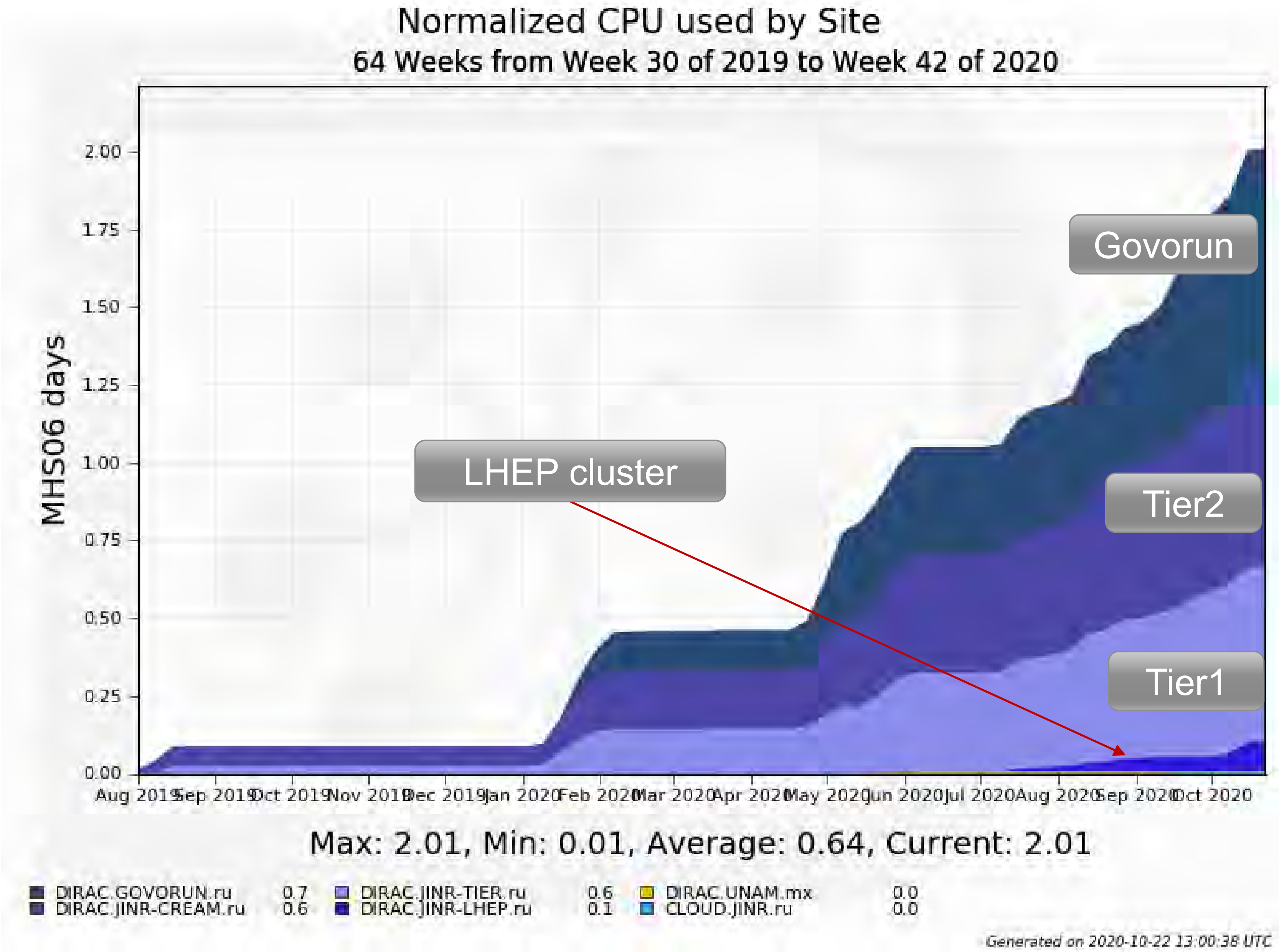
Tier2

LHEP cluster

Tier1

Govorun

MPD Normalized time



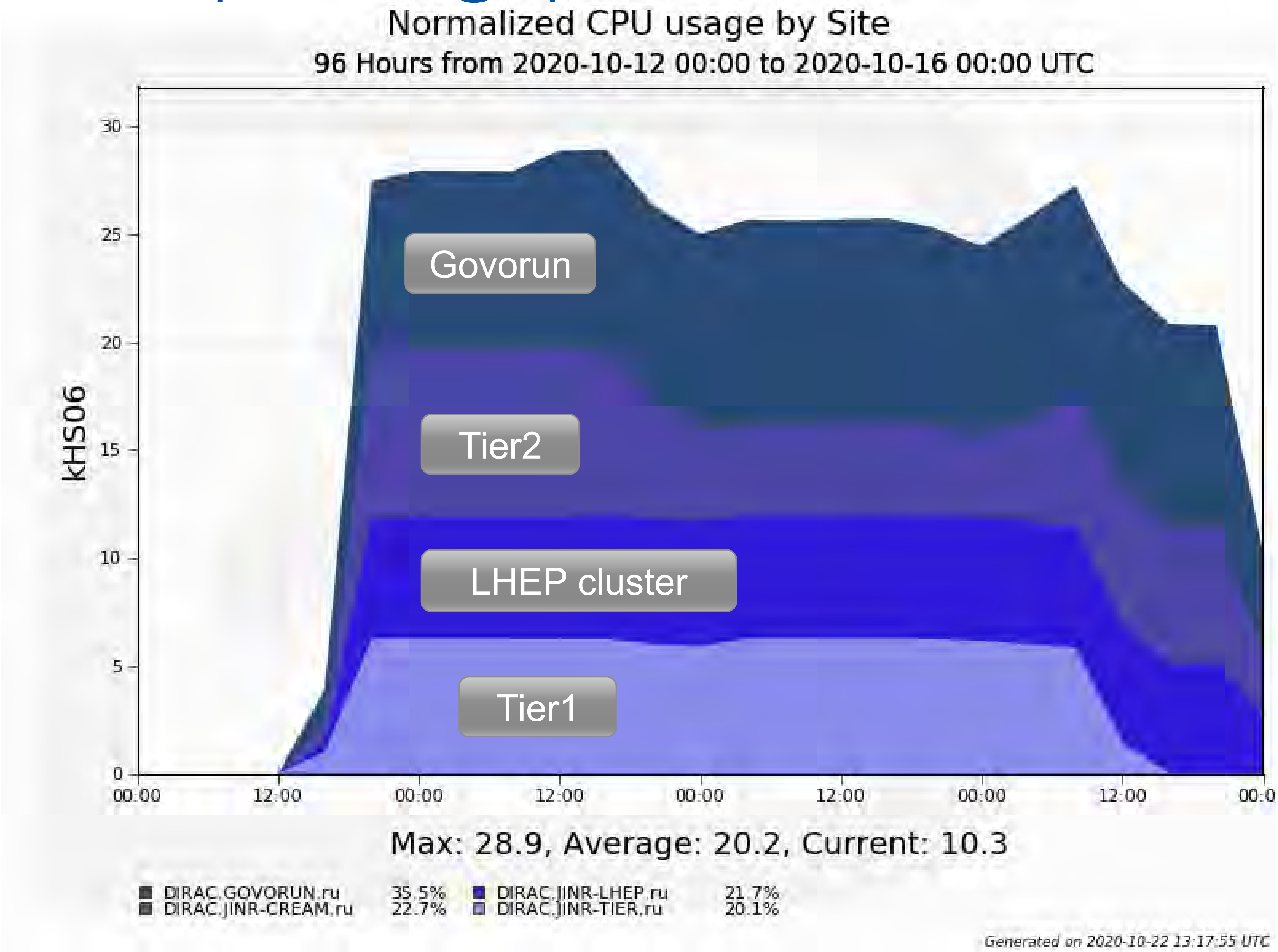
Govorun

LHEP cluster

Tier2

Tier1

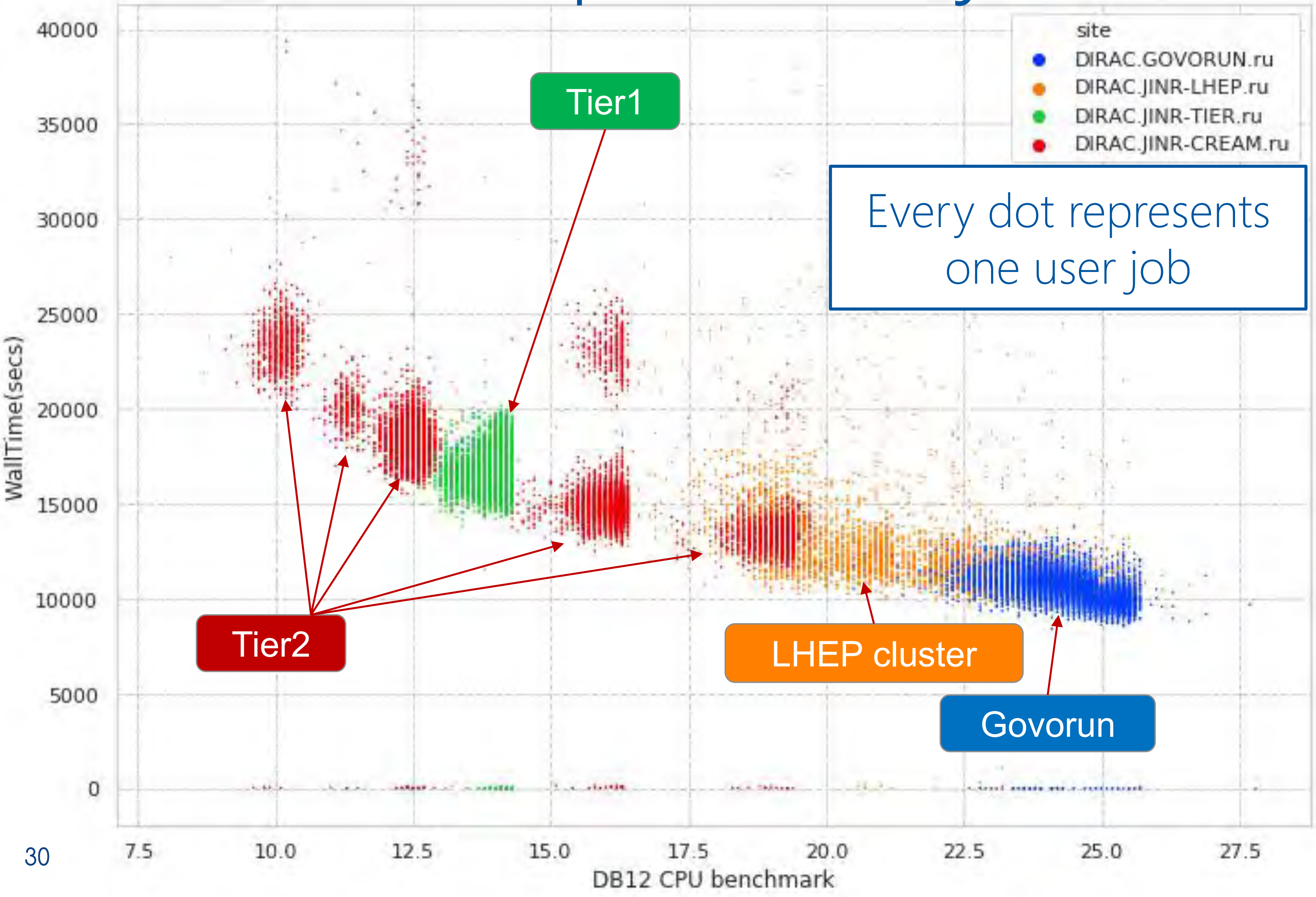
Computing power in DIRAC



Individual CPU core performance study

- Centralized job management gives possibility for centralized and unified performance study of different computing resources.
- Before running user jobs DIRAC Pilots execute benchmark for CPU core they are running on.
- Benchmark is DiracBenchmark2012 or DB12. It evaluate just CPU core performance. Disk I/O, RAM speed, Network, CPU caches and other highly important aspects of performance are **neglected by DB12**.

MPD plot - July



Tier1

Tier2

LHEP cluster

Govorun

Every dot represents one user job

Conclusion on MPD+DIRAC

- Cooperation is the key.
- >500k jobs successfully done
- >130TB data written to EOS disks(all registered in DIRAC FileCatalog)
- Some resources are not presented in accounting:
  - JINR Cloud and other clouds were not actively used up to now.
  - UNAM Cluster: 1000 jobs completed as an experiment. Network is week point. Using local storage will solve the issue.
  - dCache Tapes access over DIRAC is successfully tested. Mostly needed for RAW data from detector.
- DIRAC accounting provide normalized accounting across all resources.

JOINT INSTITUTE FOR NUCLEAR RESEARCH

JINR

DUBNA