25th Symposium on Nuclear Electronics and Computing - NEC'2015

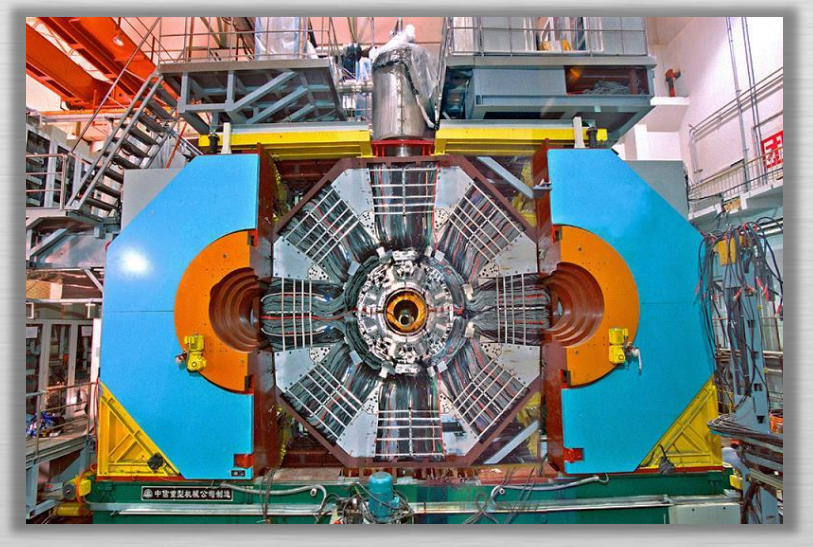
BES-III Distributed computing status

IHEP: Deng Z.Y., Li W.D., Lin T., MaZ.T., Nicholson C., Suo B., Yan T., Yan X.F., Zhang X.M., Zhao X.H.

JINR: Belov S., Korenkov V., Pelevanyuk I., Trofimov V., Uzhinskiy A., Zhemchugov A.

BESIII Introduction

- The BES-III experiment in Beijing is a world best facility to test Standard Model and QCD with high precision in taucharm domain.
- Uses grid computing since 2013
- Maximum data rate is about 40 MB/s.
- Amount of data: ~400 TB
- Around 500 000 jobs executed since this time in 2014.
- A lot of unique data collected



BES-III grid challenges

BES-III computing model before grid reminds traditional HEP computing model before year 2000.

Challenges:

- Lack of grid experience among communities
- SE is not affordable to all the sites
- Weak network connection
- Lack of manpower to maintain sites

Grid development was motivated to involve more computing resources from remote BES-III members

Distributed computing workflow

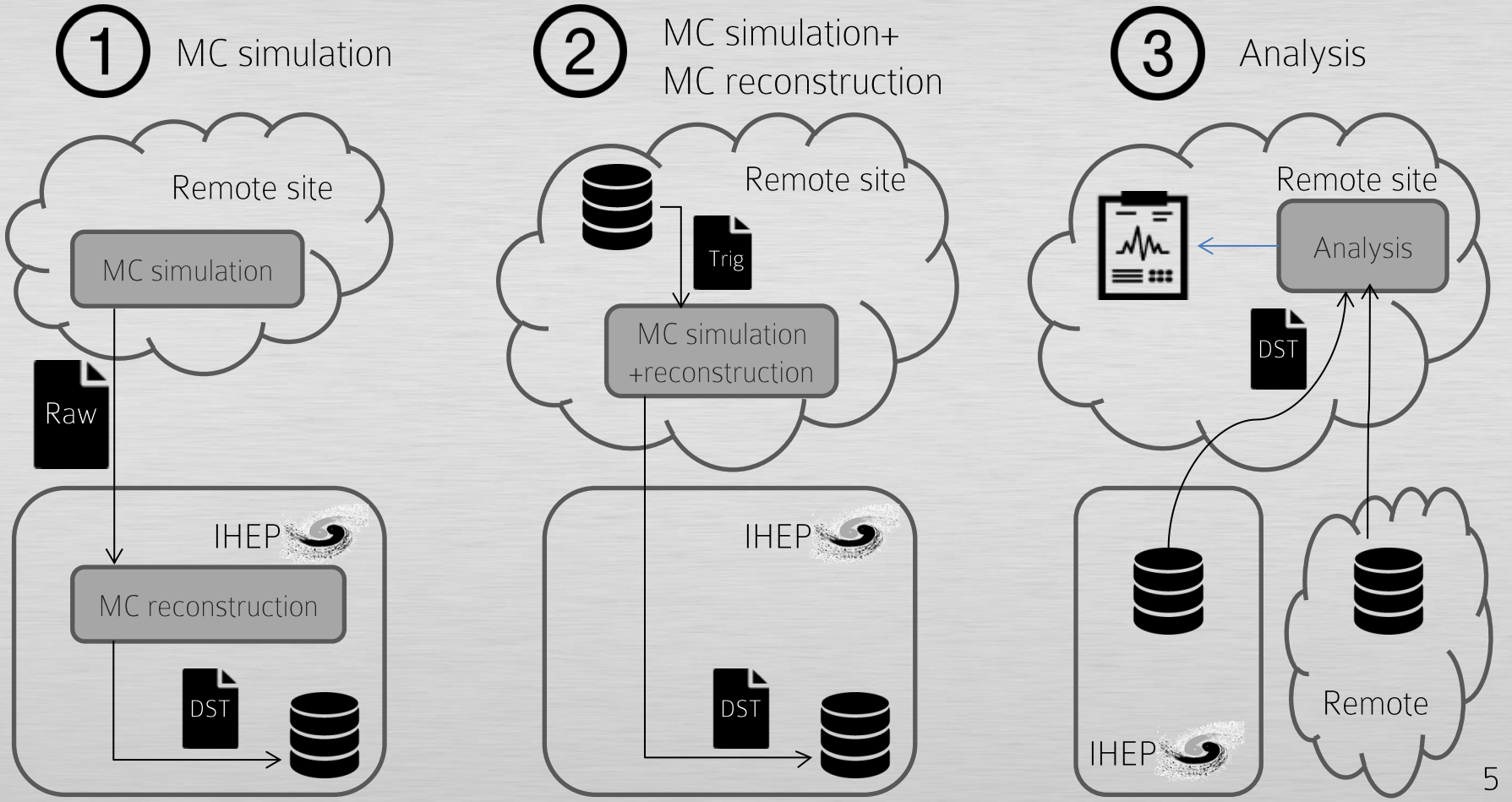
Remote sites participate only in MC production and physics analysis, while all reconstruction of experimental and simulated data is done at IHEP.

Raw experimental data stored on the tape managed by CASTOR.

DSTs stored in a disk pool managed by Lustre + dCache/STORM

Distributed computing workflow

With the growth of resources the workflow model expanded



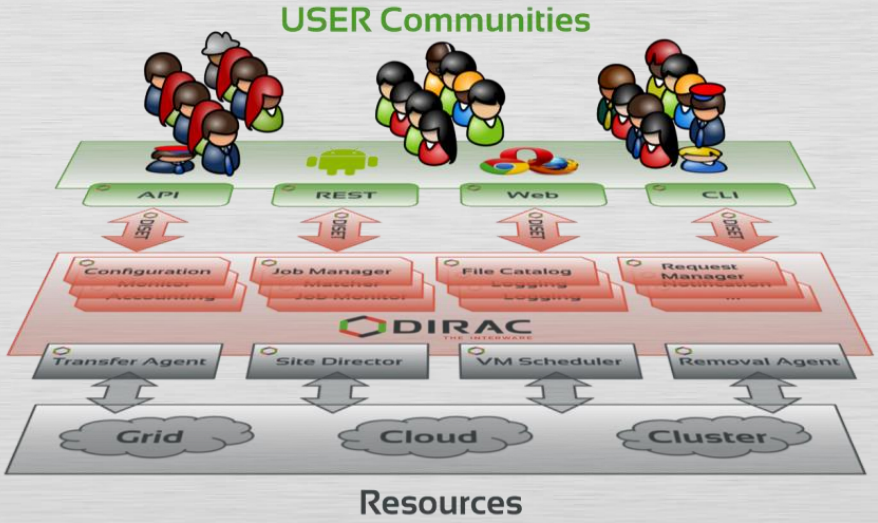
BES-III grid solution

DIRAC
THE INTERWARE

- DIRAC provide a complete and grid solution for both workload and data management system, and it is design to minimize the effort of local sites and system maintainer.

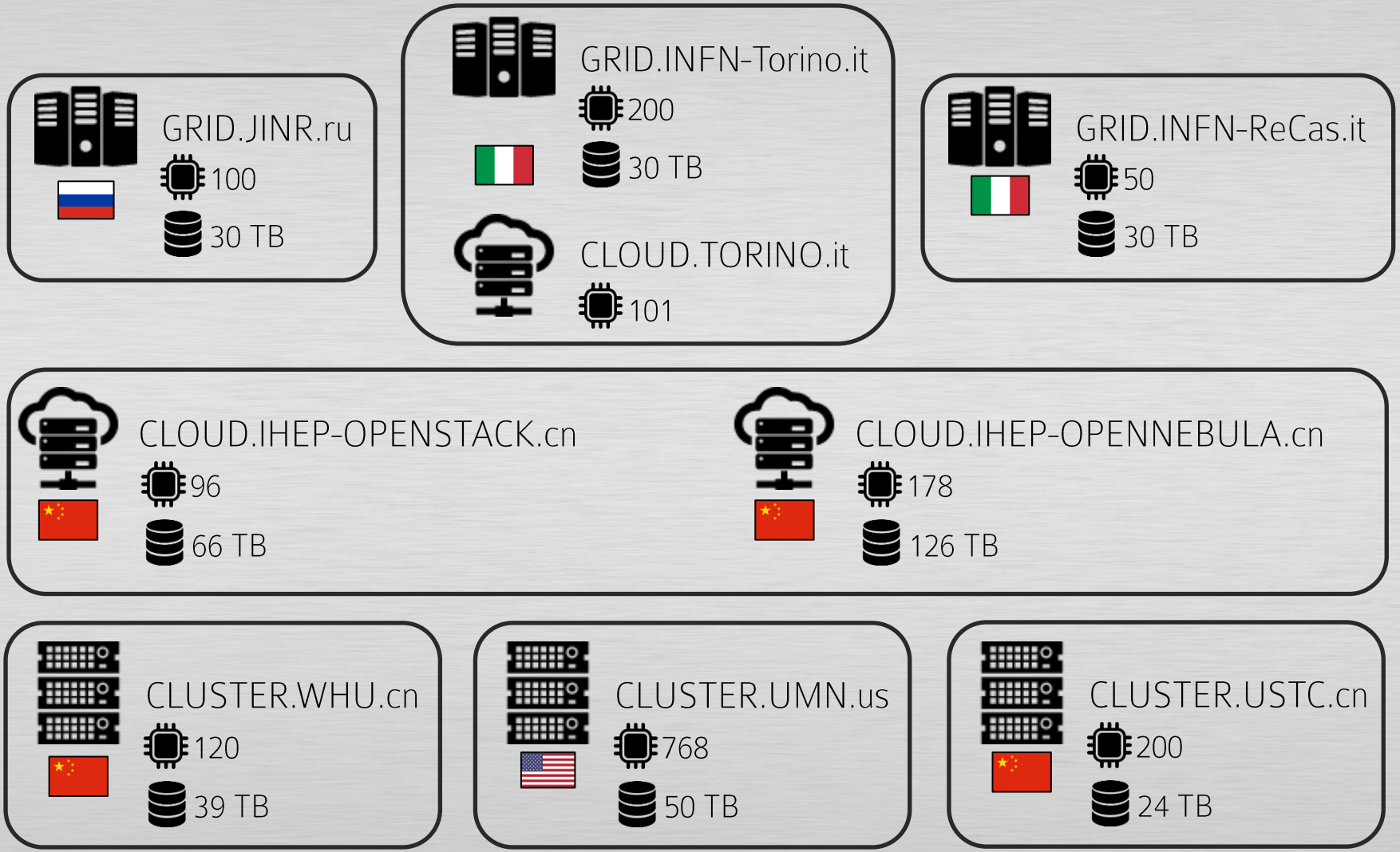
Technologies:

- Python – main language
- JavaScript (ExtJS) – for web
- MySQL – DIRAC database
- GitHub – controls code versions



It was greatly improved for the purposes of BES-III.

Grid resources



Clouds

Cloud resources continue to work well: ⅓ of all jobs done on clouds

There are 6 cloud sites in total. Most use KVM.

Including Amazon EC2.

Full processing of 1 B J/psi would cost ~ 30 000 $.

Process	Loss(No opt)	Loss(Opt)
Simulation	11.12%	2.79%
Reconstruction	5.98%	1.89%
Analysis	12.91%	1.75%

Problems:

- need improvements in track of failure and control of VMs to reduce failure rate and save resources
- cloud specific site monitoring is necessary

Task Management System

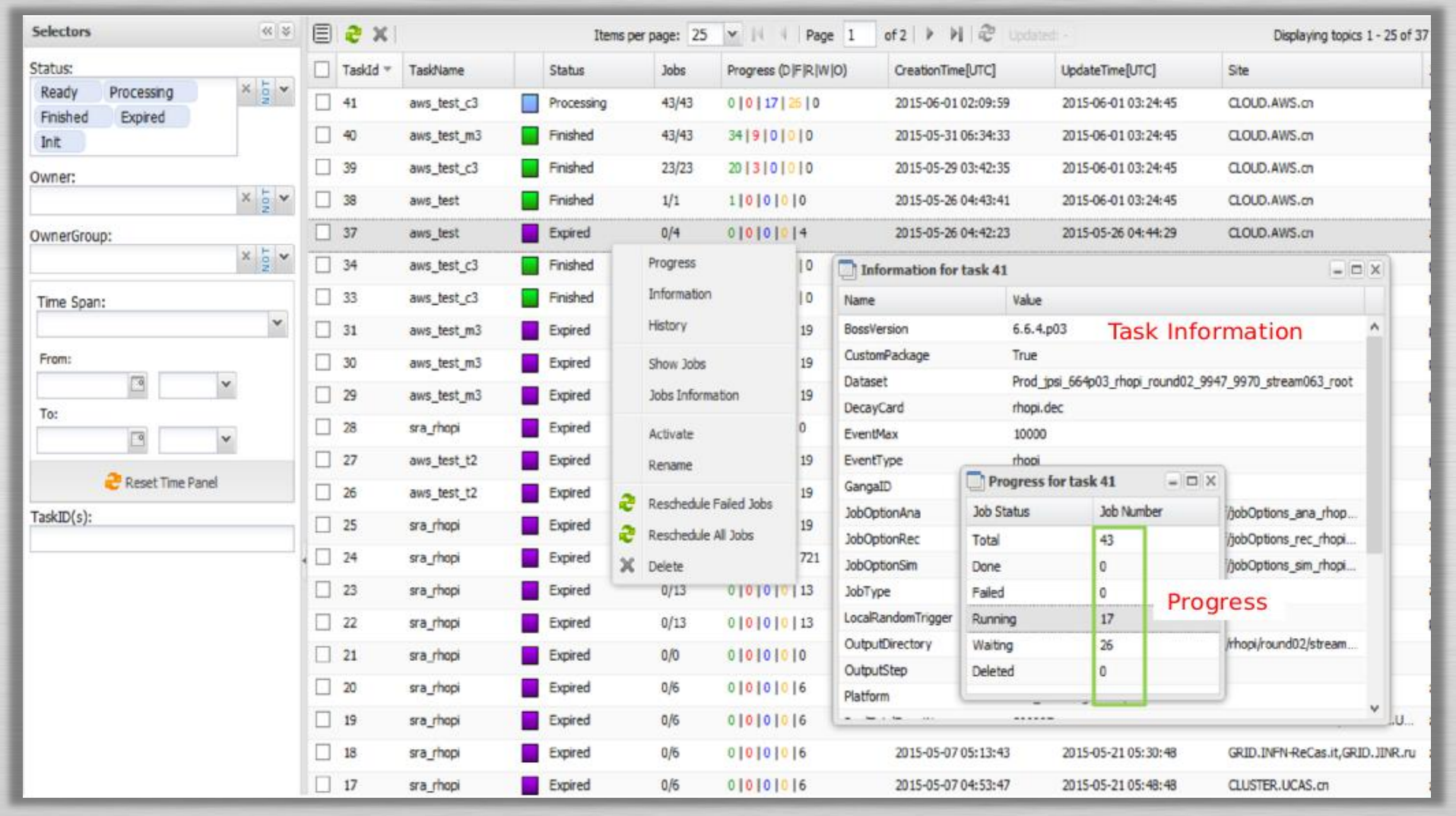
Usually production consists of MANY jobs:

- It is convenient combine all jobs of specific physics production into a logical unit - a task
- common monitoring
- common control (restart, kill etc)

The task management system has been developed with the following functions:

- Get status and progress of all jobs in the task
- Reschedule all or failed jobs of the task
- Delete all the jobs in the task

Task Management System



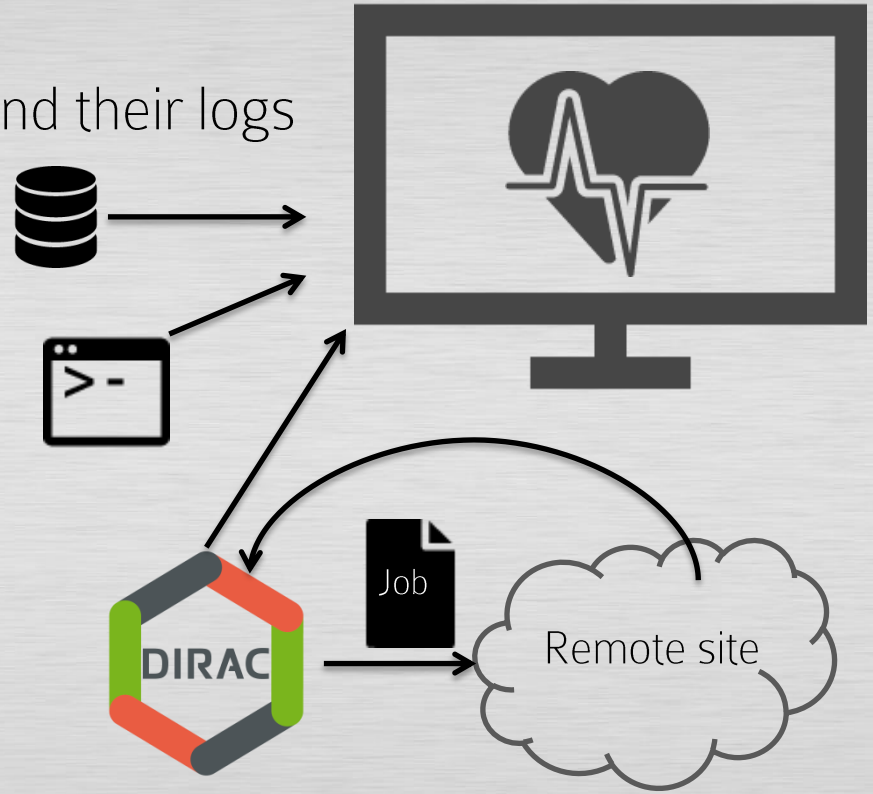
Monitoring System

Sources of information:

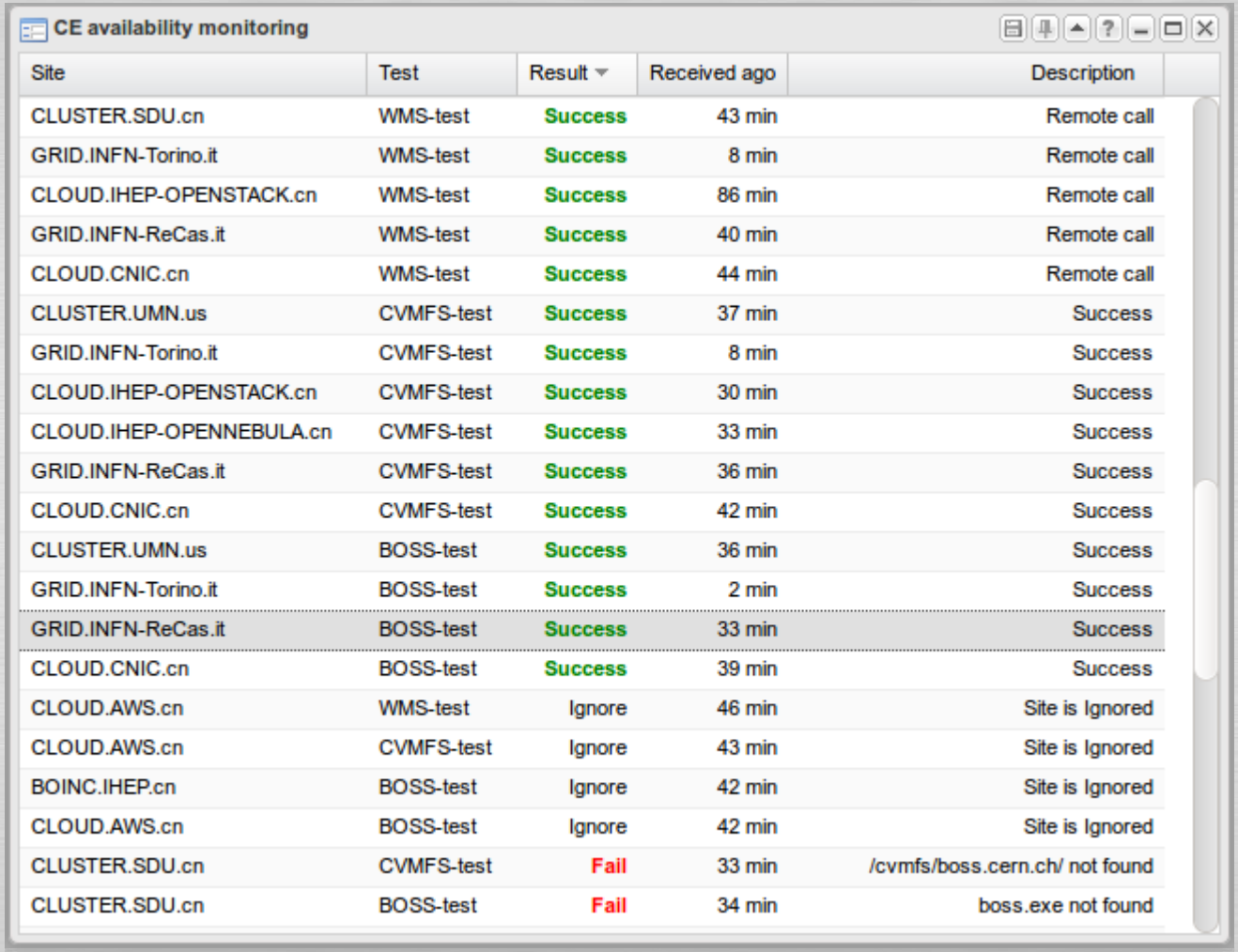
- Periodical functional(SAM) tests and their logs
- Inner DIRAC data base
- Auxiliary scripts

What to show to users(admins):

- Status/Availability
- SAM tests results
- «General state» of grid

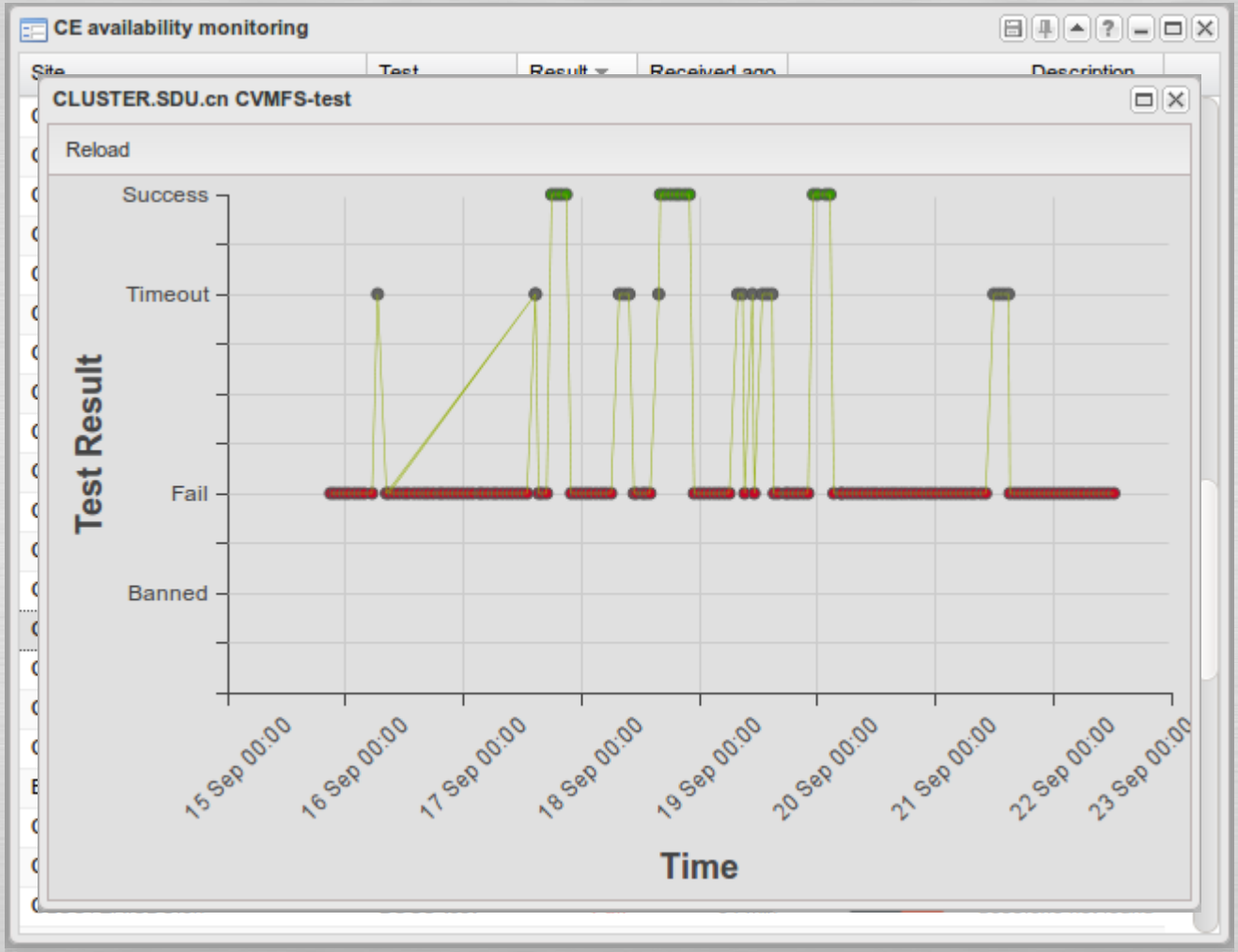


Monitoring System



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-test	Success	33 min	Success
GRID.INFN-ReCas.it	CVMFS-test	Success	36 min	Success
CLOUD.CNIC.cn	CVMFS-test	Success	42 min	Success
CLUSTER.UMN.us	BOSS-test	Success	36 min	Success
GRID.INFN-Torino.it	BOSS-test	Success	2 min	Success
GRID.INFN-ReCas.it	BOSS-test	Success	33 min	Success
CLOUD.CNIC.cn	BOSS-test	Success	39 min	Success
CLOUD.AWS.cn	WMS-test	Ignore	46 min	Site is Ignored
CLOUD.AWS.cn	CVMFS-test	Ignore	43 min	Site is Ignored
BOINC.IHEP.cn	BOSS-test	Ignore	42 min	Site is Ignored
CLOUD.AWS.cn	BOSS-test	Ignore	42 min	Site is Ignored
CLUSTER.SDU.cn	CVMFS-test	Fail	33 min	/cvmfs/boss.cern.ch/ not found
CLUSTER.SDU.cn	BOSS-test	Fail	34 min	boss.exe not found

Monitoring System



Multy VO support

Using DIRAC is easy. Installing, configuring and supporting – not so easy.

CEPC* and JUNO** have joined to BES-III computing infrastructure.

*Circular Electron Positron Collider

**Jiangmen Underground Neutrino Observatory

Resource utilization 

Support needs 

Feature development 

Multy VO support

Using DIRAC is easy. Installing, configuring and supporting – not so easy.

CEPC* and JUNO** have joined to BES-III computing infrastructure.

*Circular Electron Positron Collider

**Jiangmen Underground Neutrino Observatory

Resource utilization 

Support needs 

Feature development 

BES-III Future plans

A number of new developments is planned:

- Data Management System optimization
- Increase effectiveness of cloud resources use
- Storage Accounting System development
- Action-based monitoring
- Job Submission Tool for multi-vo support
- The StoRM-based central storage system is planned to ease sharing experiment data among sites
- Reduce impact of weak storage on site performance

Thank you for attention

References

Most of the icons in this presentation were taken from:

http://www.flaticon.com/

http://www.freepik.com/