CTA report

<u>Luisa Arrabito</u>, Johan Bregeon, Ricardo Graciani

**DIRAC User Workshop 26th – 28th May 2014, CERN**

CTA report

Outlook:

- CTA project
- Current activity: MC simulation
- DIRAC for CTA
- Conclusions and perspectives

CTA (Cherenkov Telescope Array) (I)

- CTA is the next generation instrument in the VHE gamma-ray astronomy
- 2 arrays of 50-100 Cherenkov Telescope (North and South hemisphere)
- 10x sensitivity (point sources) with respect to the current IACT experiments
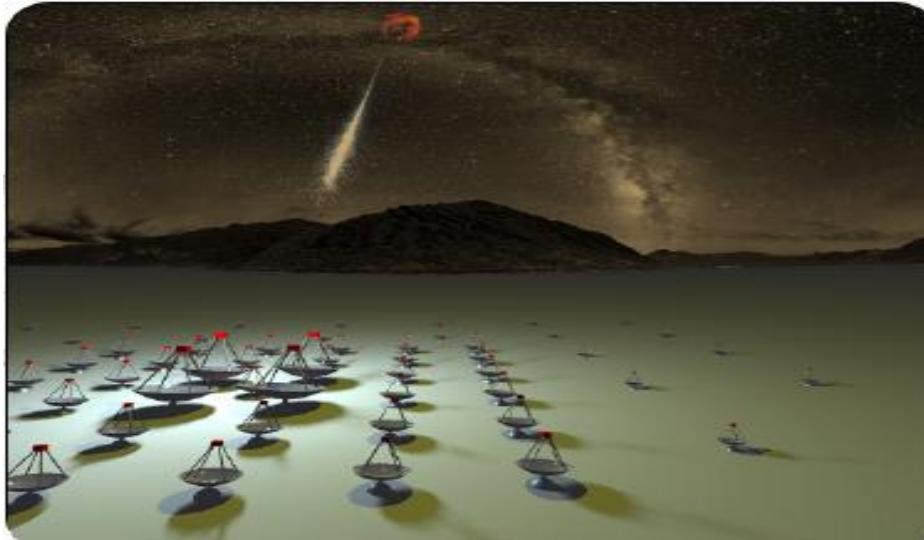
Scientific goals:

- Cosmic rays origins
- High Energy astrophysical phenomena
- Fundamental physics and cosmology

CTA (Cherenkov Telescope Array) (II)

CTA Project

- Consortium of > 1000 scientists in 25 countries
- Current preparatory phase (2011-2014):
- Find sites
- Build prototypes of telescopes
- Define optimal array layout
- *Prepare data pipeline and data distribution*
- The construction of the full array to be completed around 2018-2020
- Operate as an observatory

Current activity: MC Simulation (I)

Objectives:

- Site selection (North and South)
- Study of telescope and array design
- Algorithm optimisation
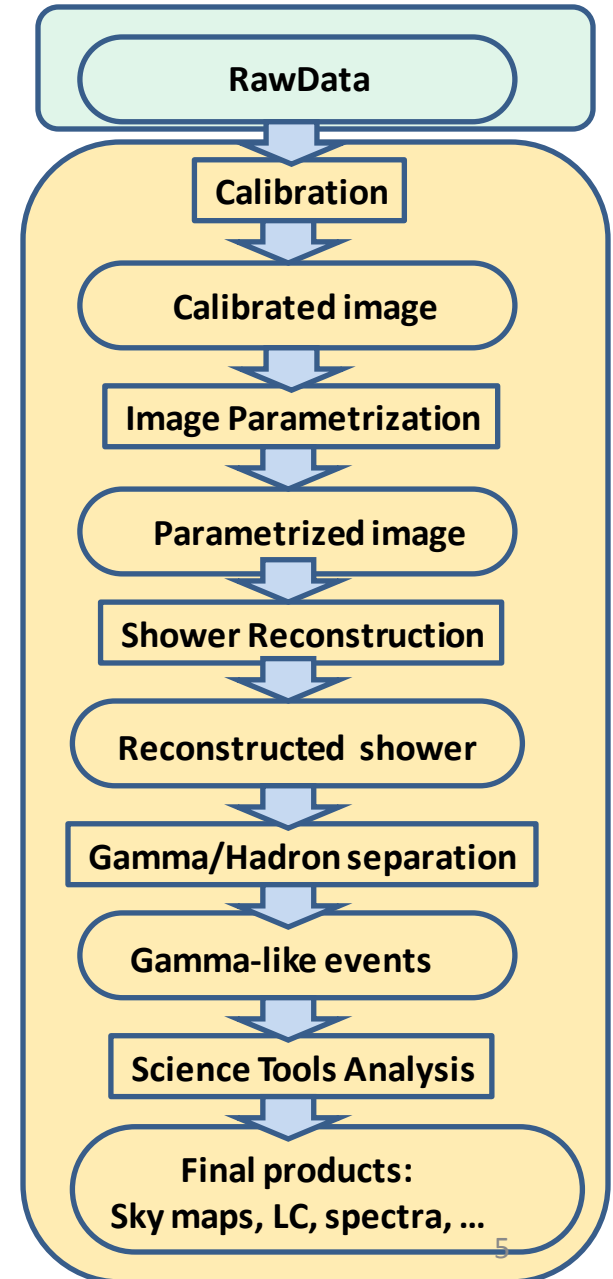- Studies of final instrument performances

High statistics needed:

- 10^9 - 10^{10} simulated events for each study

MC Production:

- Shower generation and propagation in the atmosphere
- Telescope response simulation

MC Analysis:

- Reconstruction and analysis

RawData

Calibration

Calibrated image

Image Parametrization

Parametrized image

Shower Reconstruction

Reconstructed shower

Gamma/Hadron separation

Gamma-like events

Science Tools Analysis

Final products:
Sky maps, LC, spectra, ...

Current activity: MC Simulation (II)

Computing needs:

- A typical production campaign:
- CPU: ~ 10 - 20 M HS06 h (about 2-4 weeks using 3000 cores in parallel)
- ~ 200k jobs
- Storage: 30 TB (600 TB, if rawdata are saved)
- ~ 600k files (1 M files, if rawdata are saved)
- About 10 campaigns per year

- A typical analysis:
- Processing of about 30 TB
- ~ 200k jobs

CTA VO grid resources

The CTA VO:

- Active since 2008
- Today: 21 sites in 7 countries
- About 100 members

Resources:

- Dedicated and opportunistic resources
- CPU:
  - 4000-5000 cores available on average
- Storage:
  - More than 1 PB dedicated in 6 sites

CTA VO EGI sites

DIRAC for CTA

- **DIRAC evaluation started in 2011**
  - First DIRAC instance dedicated to CTA installed at PIC:
  - 2 virtual servers having in total 6 cores, 6 GB of RAM and 1.5 TB of local disk
  - 1 server hosting the web portal
  - Now upgraded with:
  - 2 servers with 4 Cores: dcta-servers (8 GB) & dcta-db (8GB)
  - 2 servers with 2 Cores: dcta-agents (2 GB) & dcta-web (1GB)
  - First use-cases (see CTA talk at last DIRAC User Workshop)

- **Since 2013 DIRAC is used for MC production and analysis**
  - Simple CTA-DIRAC extension to configure and run 6 CTA applications
  - Use of DFC to save Meta-data information for MC production (particle type, zenith angle, sw version, etc.)
  - Use of the Request System for bulk file deletion
  - Use of parametric jobs for MC production
  - Preliminary tests using the Transformation System for MC production

- **At term, we are proposing DIRAC WMS for CTA operation phase**
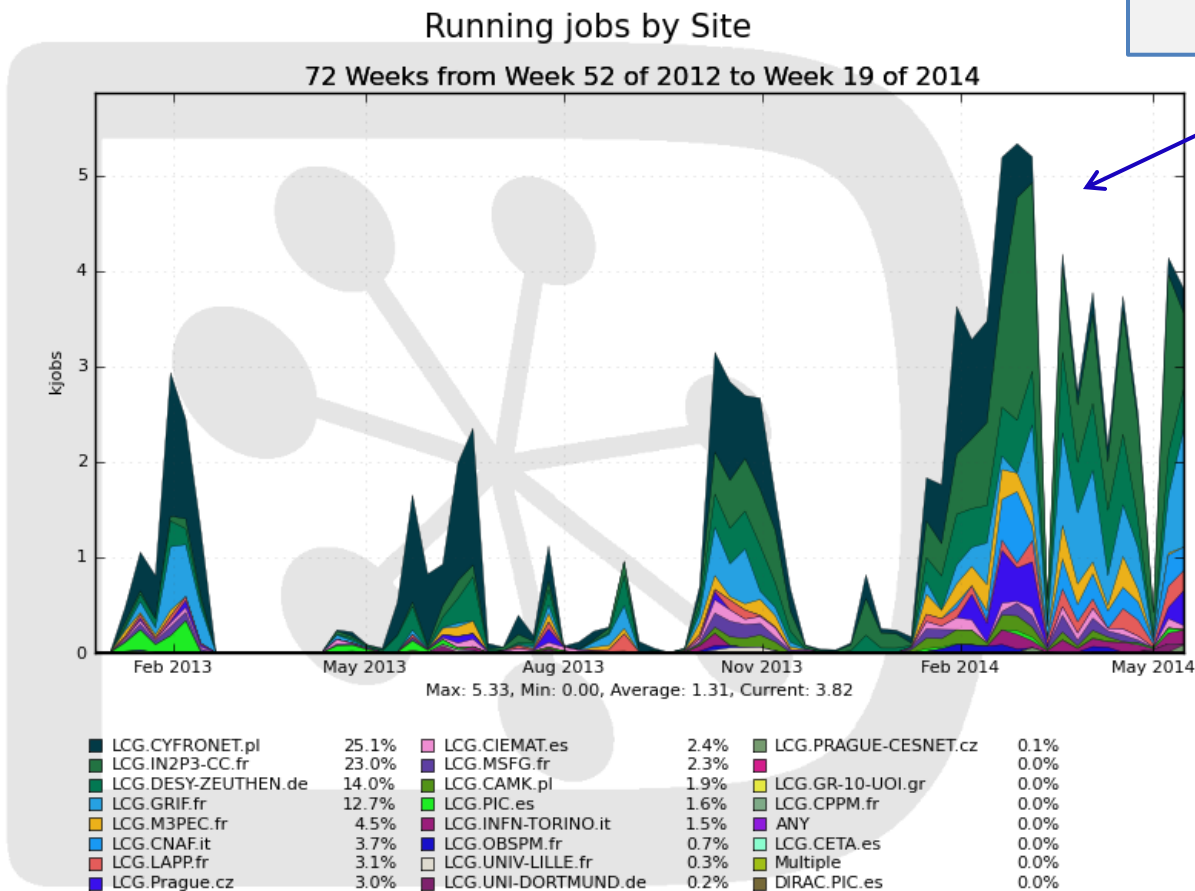 - Raw-data reconstruction, calibration and IRF production

DIRAC usage in 2013-2014 (I)
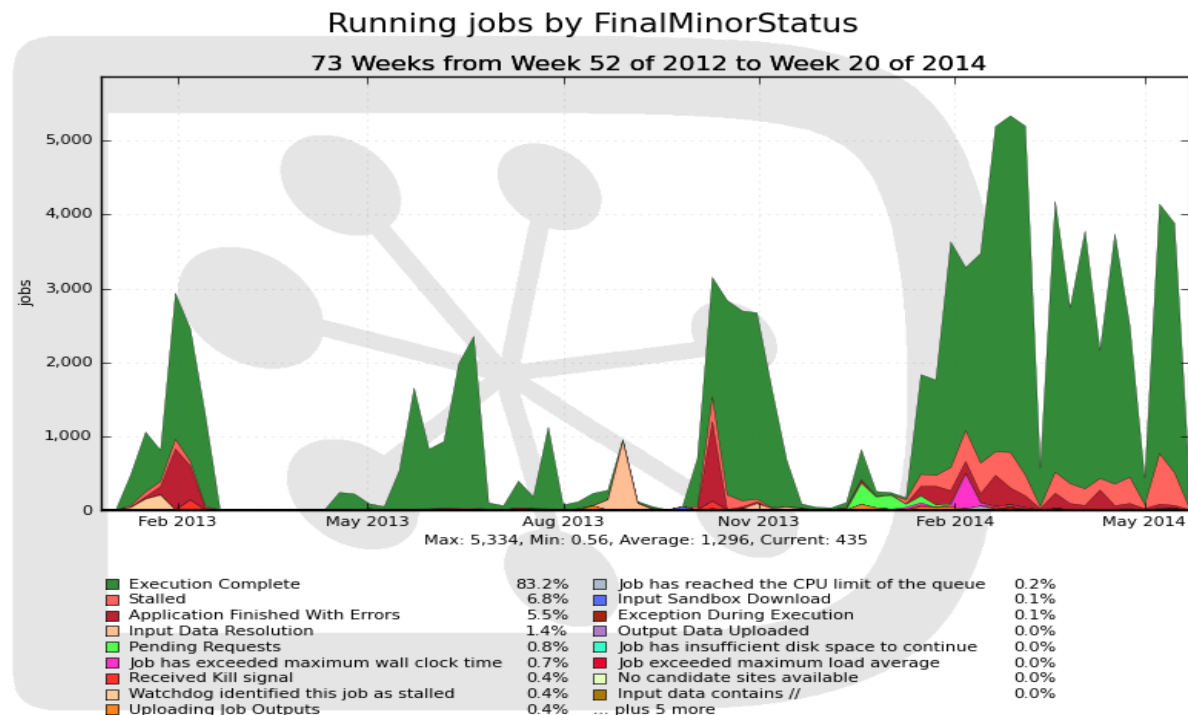
Running jobs:

- Stable regimes of 4000-5000 concurrent jobs
- About 4 M executed jobs
- 132 M HS06 CPU hours

Almost constant activity since Feb. 2014

Running jobs by Site

72 Weeks from Week 52 of 2012 to Week 19 of 2014

Max: 5.33, Min: 0.00, Average: 1.31, Current: 3.82

LCG.CYFRONET.pl	25.1%	LCG.CIEMAT.es	2.4%	LCG.PRAGUE-CESNET.cz	0.1%
LCG.IN2P3-CC.fr	23.0%	LCG.MSFG.fr	2.3%		0.0%
LCG.DESY-ZEUTHEN.de	14.0%	LCG.CAMK.pl	1.9%	LCG.GR-10-UOI.gr	0.0%
LCG.GRIF.fr	12.7%	LCG.PIC.es	1.6%	LCG.CPPM.fr	0.0%
LCG.M3PEC.fr	4.5%	LCG.INFN-TORINO.it	1.5%	ANY	0.0%
LCG.CNAF.it	3.7%	LCG.OBSPM.fr	0.7%	LCG.CETA.es	0.0%
LCG.LAPP.fr	3.1%	LCG.UNIV-LILLE.fr	0.3%	Multiple	0.0%
LCG.Prague.cz	3.0%	LCG.UNI-DORTMUND.de	0.2%	DIRAC.PIC.es	0.0%

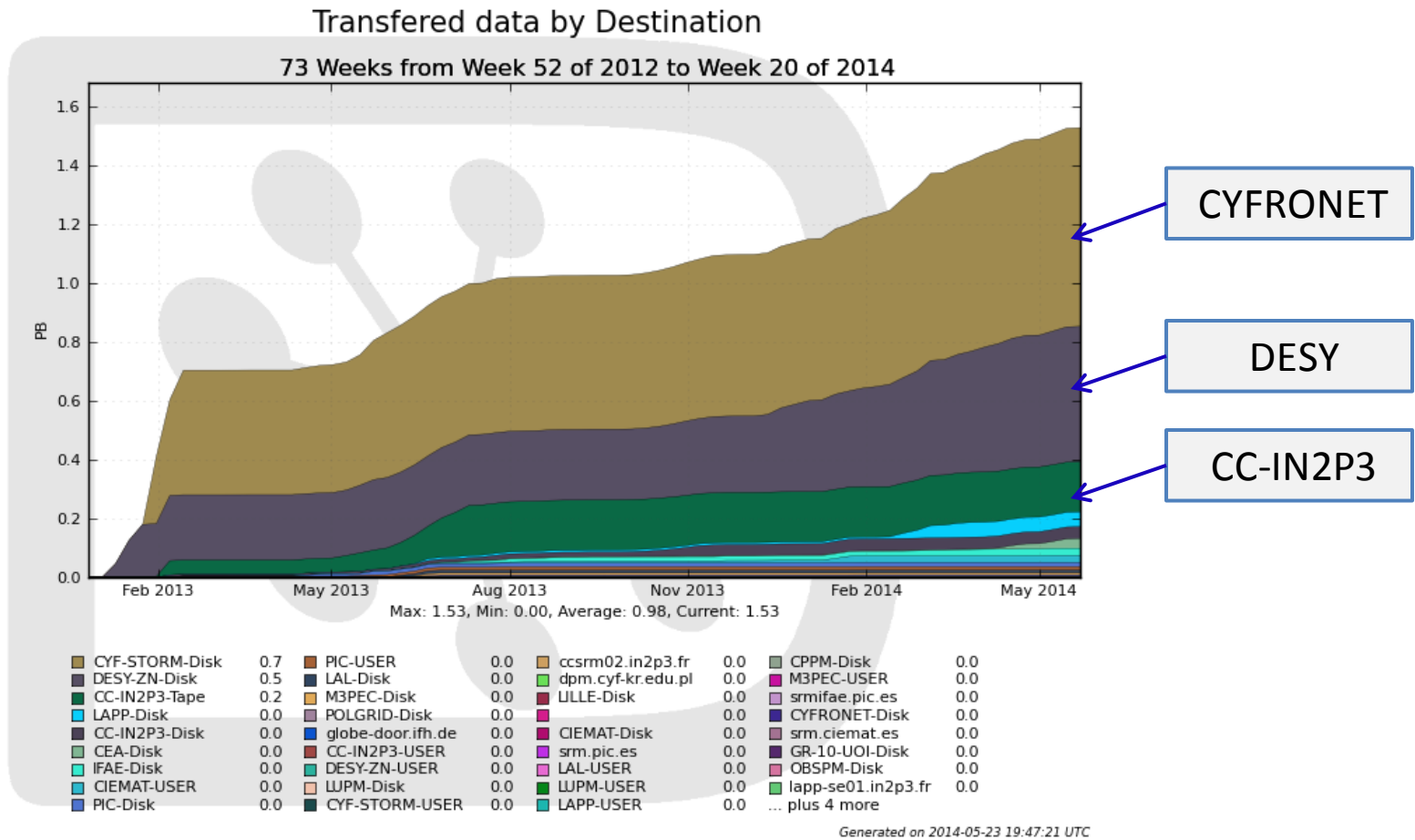Generated on 2014-05-22 11:27:30 UTC

DIRAC usage in 2013-2014 (II)

- Overall success rate:
- 83% , not so good
- Main reasons for failures:
- Stalled jobs (proxy renewal, CPU limit exceeded, etc.)
- Timeouts while connecting to DFC
- Access to Input Data
- Pending requests
- Some limitations are due to the hw of the CTA-DIRAC servers

Running jobs by FinalMinorStatus
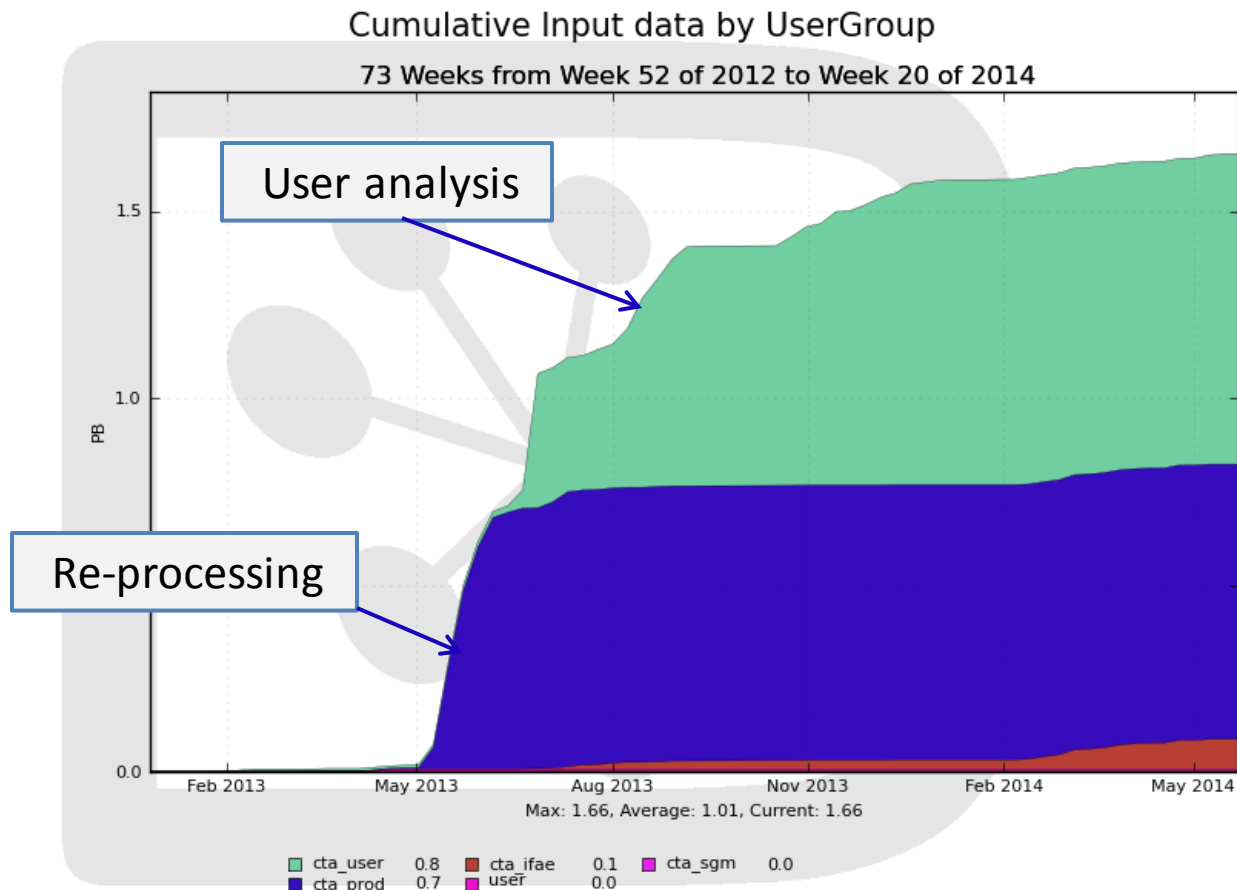#### 73 Weeks from Week 52 of 2012 to Week 20 of 2014

Max: 5,334, Min: 0.56, Average: 1,296, Current: 435

Execution Complete	83.2%	Job has reached the CPU limit of the queue	0.2%
Stalled	6.8%	Input Sandbox Download	0.1%
Application Finished With Errors	5.5%	Exception During Execution	0.1%
Input Data Resolution	1.4%	Output Data Uploaded	0.0%
Pending Requests	0.8%	Job has insufficient disk space to continue	0.0%
Job has exceeded maximum wall clock time	0.7%	Job exceeded maximum load average	0.0%
Received Kill signal	0.4%	No candidate sites available	0.0%
Watchdog identified this job as stalled	0.4%	Input data contains //	0.0%
Uploading Job Outputs	0.4%	... plus 5 more	

Generated on 2014-05-23 20:05:06 UTC

DIRAC usage in 2013-2014 (III)

- About 1.5 PB of transferred data:
- to 6 SE: CYFRONET, DESY, CC-IN2P3, GRIF, LAPP, CNAF

Transfered data by Destination

73 Weeks from Week 52 of 2012 to Week 20 of 2014

Max: 1.53, Min: 0.00, Average: 0.98, Current: 1.53

CYF-STORM-Disk	0.7	PIC-USER	0.0	ccsrm02.in2p3.fr	0.0	CPPM-Disk	0.0
DESY-ZN-Disk	0.5	LAL-Disk	0.0	dpm.cyf-kr.edu.pl	0.0	M3PEC-USER	0.0
CC-IN2P3-Tape	0.2	M3PEC-Disk	0.0	LILLE-Disk	0.0	srmifae.pic.es	0.0
LAPP-Disk	0.0	POLGRID-Disk	0.0		0.0	CYFRONET-Disk	0.0
CC-IN2P3-Disk	0.0	globe-door.ifh.de	0.0	CIEMAT-Disk	0.0	srm.ciemat.es	0.0
CEA-Disk	0.0	CC-IN2P3-USER	0.0	srm.pic.es	0.0	GR-10-UOI-Disk	0.0
IFAE-Disk	0.0	DESY-ZN-USER	0.0	LAL-USER	0.0	OBSPM-Disk	0.0
CIEMAT-USER	0.0	LUPM-Disk	0.0	LUPM-USER	0.0	lapp-se01.in2p3.fr	0.0
PIC-Disk	0.0	CYF-STORM-USER	0.0	LAPP-USER	0.0	... plus 4 more	

Generated on 2014-05-23 19:47:21 UTC

CYFRONET

DESY

CC-IN2P3

DIRAC usage in 2013-2014 (IV)

- About 1.6 PB of processed data:
- Re-processing
- User analysis

Cumulative Input data by UserGroup

73 Weeks from Week 52 of 2012 to Week 20 of 2014

User analysis

Re-processing

Max: 1.66, Average: 1.01, Current: 1.66

| cta_user | 0.8 | cta_ifae | 0.1 | cta_sgm | 0.0 |
| cta_prod | 0.7 | user | 0.0 | | |

Generated on 2014-05-23 20:57:55 UTC

12

DIRAC usage in 2013-2014 (V)

DIRAC File Catalog:

- Used for MC production
- 11.7 M replicas
- Typical user queries select tens of thousands of files

Browsing the Catalog

Meta-data selection

Query result

Conclusions and perspectives

- CTA will be producing *several PB/year*, but the Computing Model is not yet defined

- DIRAC WMS identified as a promising solution for data processing during CTA operation

- *CTA-DIRAC prototype is running since 3 years:*
 - Using several components: WMS, DMS, DFC, RS, TS, WebApp, REST
 - First CTA-DIRAC tutorials held in Montpellier in Feb. 2014

- Need to further upgrade the hw of the CTA-DIRAC instance

- In future, we will certainly need to connect the CTA-DIRAC to non-grid resources

- Work in progress at CYFRONET to fully interface the CTA Gateway prototype to DIRAC (REST interface)