Batch Processing: Future Plans

Gareth Hughes

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

A long time ago …

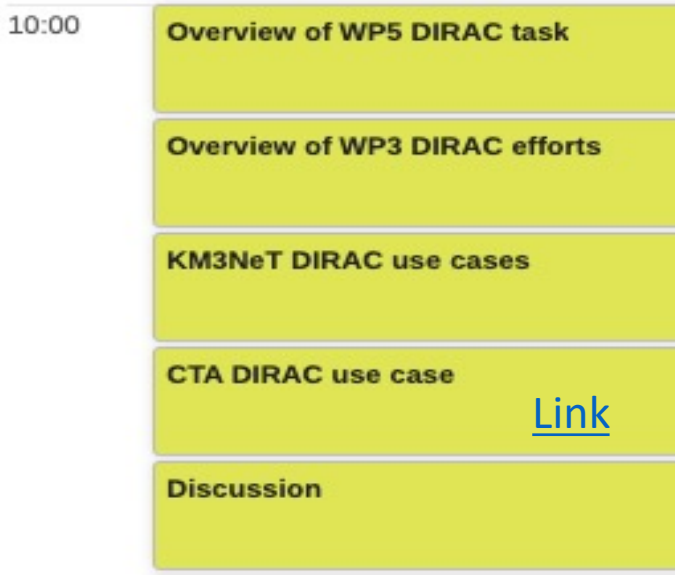
Deliverable 5.2: Project Plan

"Many existing workload management systems are proven to be capable of dealing with this, which makes the development of ESAP in terms of connecting to different computer infrastructures easier. **ESAP will investigate the existing workload management systems and select one to interface with**. In this way, batch job submission will be handled by the selected workload management system that in turn connects to the underlying HTC, HPC or Cloud platforms."

- Requirement R5:
  - Users should be able to execute a job on a given dataset, including but not limited to: batch or real-time queries & pipelines, depending on the capabilities of the facility, which need to be made clear to the user.
    - Use Cases: U-4, U-5, U-7
- Requirement R11:
  - Users should be able to select computing facilities on the basis of their capacity. E.g. She needs an HPC resource with a specific acceleration (GPU) because the software to be run requires it.
    - Use Cases: U-15, U-14, U-8
- Requirement R13
  - Users should be able to schedule computational tasks at regular intervals e.g. to periodically retrieve new classification data from a Citizen Science experiment.
    - Use Cases: U-18

Funded by the European Union's Horizon 2020 - Grant N° 824064

2020 Progress Meeting



- Step 1: Next steps
  - Finish collecting use cases
  - With WP3 integrate the DIRAC-CORSIKA container with the ESAP
- Step 2: Further integration and tech
  - Can we link the data from the shopping cart?
  - Do we need some testbed infrastructure (e.g. a server)?
- Step 3: Can we get some glue?
  - RUCIO
  - Interactive (jupyter etc …)

Funded by the European Union's Horizon 2020 - Grant N° 824064

Where is Workload Management useful?

Data Lake

Staff Scientist

Batch
WMS

JupyterHub
VO tools

Science
User

Data Level 0
Raw

DLX

DL3
Science

DLY

DL6
High

File size

Funded by the European Union's Horizon 2020 - Grant N° 824064

Batch Systems / Job schedulers

- Common at computing centres
- Head node that manages the schedule
- Worker nodes (never enough)
- Complex job weighting for queuing jobs
- e.g. slurm, condor, oracle grid engine, ….

- Send your job to the farm
- Multiple cores / parallel analysis
- Jobs run: few mins or hours

- Current Gamma-ray astronomy:
  - You prepose an observation
  - Data is taken (hopefully not so many clouds)
  - You download raw (or slightly processed) data
  - You install your experiments software
  - Run workflow (hopefully using a batch system)
  - Create high-level plots
  - Publish

- With a Science Platform
  - data discovery
  - software workflow discovery
  - Access to computing facilities

Funded by the European Union's Horizon 2020 - Grant N° 824064

Use Cases

Name	2. Data Reprocessing
ID	CTA002
Goal/Aim	The ability to reprocess all raw dat
Workflow	• Raw (DL0) data is identified on t via **metadata** e.g. using getMetaD https://tinyurl.com/myyvab5 • Data volume is calculated. • Data is staged from tape (COLD

Name	3. Generation of Instrument Response Function (IRF)
ID	CTA003
Goal/Aim	Generation of Instrument response function which enables the processing of data to higher level data products

me period, model, systematic uncertainty)
lection of further information from DB).
s not already exist.
l events.
es.
IRAC).

th appropriate metadata

Name	4b Analysis of (simulated) CTA science data by a Principal Investigator (PI) [Batch]
ID	CTA004b
Goal/Aim	A project PI is able to login to the ESAP and find and analyse the data from their proposal
Workflow	1. User logs in to the ESAP and is identified as a CTA project PI 2. Search for Data in the datalake a. Search for (simulated) CTA DL3 level data by project ID. b. Select data from search results or select all 3. Search for corresponding IRF (instrument response function) for the data selected 4. Search for corresponding metadata, log files etc .... 5. The data can now be analysed in interactive mode, in batch mode or downloaded 6. Batch mode a. Search for compute resources b. Data is transferred to resource cache c. Submit job(s) d. Monitor job(s) e. Save data to temporary quicklook location or download final results

Redmine: Use Cases

Funded by the European Union's Horizon 2020 - Grant N° 824064

Link to WP2 & DAC21

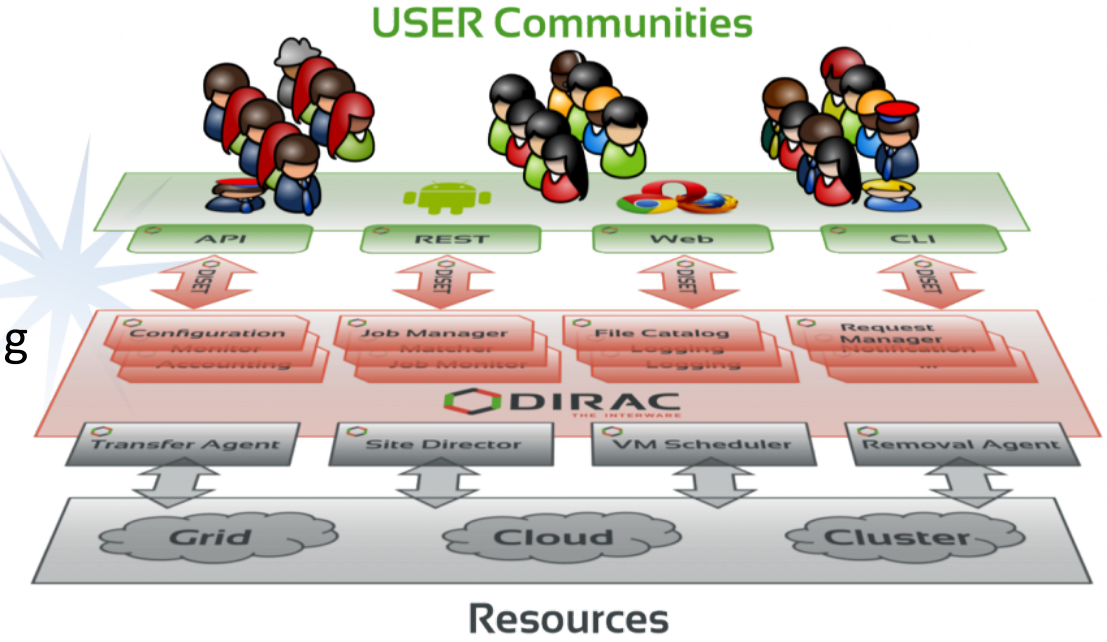
Name	2. Data Reprocessing
ID	CTA002
Goal/Aim	The ability to reprocess all raw data (DL0) to higher (DL3) level
Workflow	• Raw (DL0) data is identified on tape (obsid or time range) via **metadata** e.g. using getMetaData method https://tinyurl.com/myyvab5 • Data volume is calculated. • Data is staged from tape (COLD) storage to temporary disk (HOT). • Data is reprocessed using CTA pipeline software via the workload management system (WMS, based on DIRAC) using a cache area for on-the fly, transient data products • Final data products (DL3) are verified. • Cache and temporary data is cleared. • Ingest the resulting new DL3 data into the datalake. • Update the corresponding metadata.

Frederic Gillardo
Luisa Arrabito

Link

Funded by the European Union's Horizon 2020 - Grant N° 824064

What is DIRAC



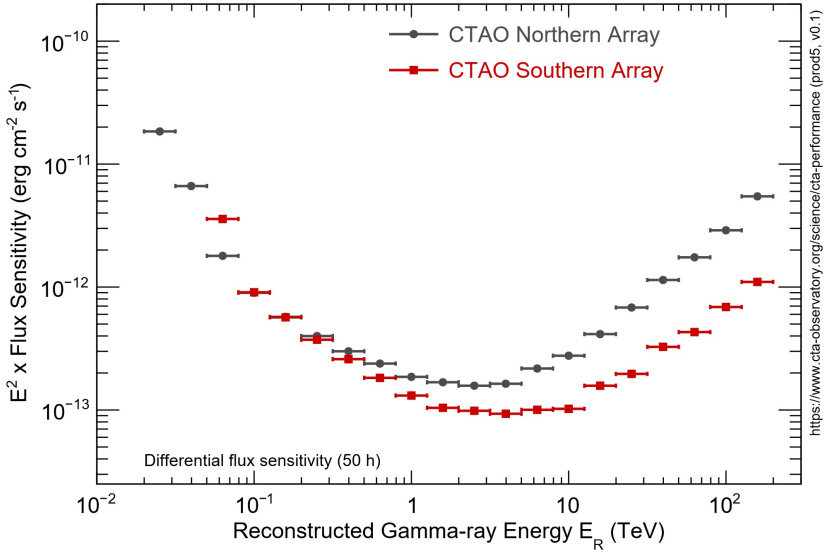
- DIRAC: **D**istributed **I**nfrastructure with **R**emote **A**gent **C**ontrol
- An open source **software framework** for distributed computing
  - http://diracgrid.org/
- Started at CERN, LHCb
  - Used by a large number of high energy and astronomy experiments
- Systems include:
  - workload management
  - data management (ESCAPE uses RUCIO)
  - Job management API
  - accounting (provenance)
  - +++ …. Much more
- DIRAC Workload Management System (WMS)
  - uses novel approach of pilot jobs
  - allows for detailed job monitoring
  - can submit to a wide range of computing centers (HPC/HTC/Cloud)
  - able to create workflows: full automatization of multi-step workflow execution
- DIRAC is not just a WMS

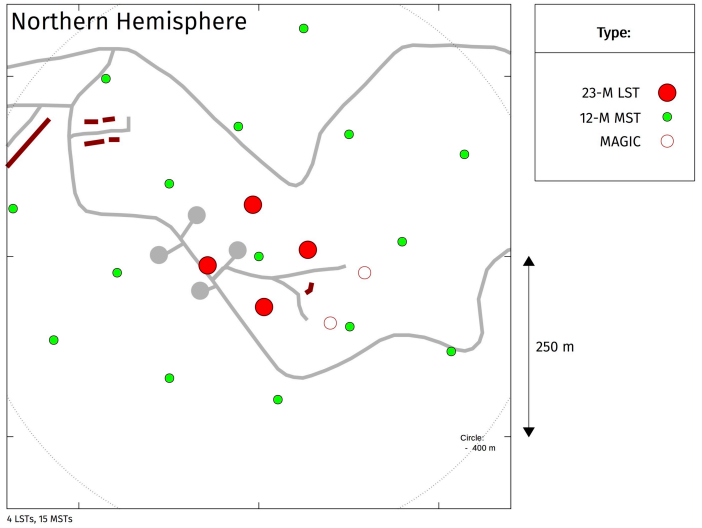
Funded by the European Union's Horizon 2020 - Grant N° 824064

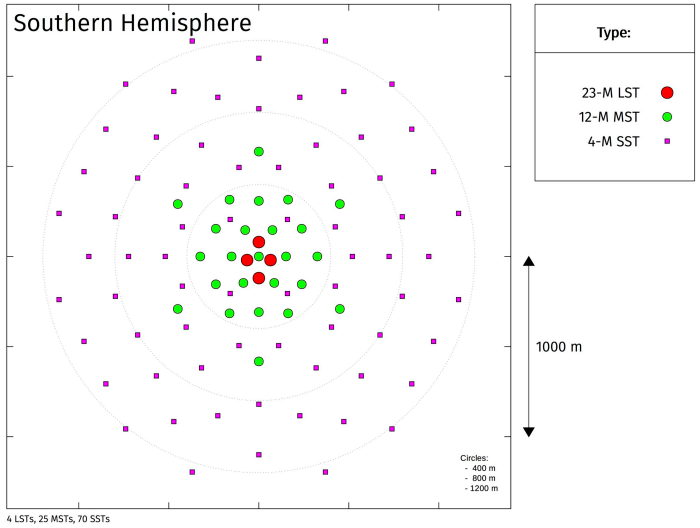
CTA and DIRAC

https://www.cta-observatory.org/science/ctao-performance/

- >5 Monte Carlo Productions (calculate sensitivity, array layout studies, etc...)
- ~18 million of jobs
- ~100 million of CPU HS06 hours per year
- ~4 PB per year between the different comp facilities
- Currently, about 4.4 PB of MC and users data are permanently stored on disk, distributed among 7 grid Storage Elements







Funded by the European Union's Horizon 2020 - Grant N° 824064

DIRAC for Developers

- BILD meetings: every 2nd week: https://indico.cern.ch/category/4205/
- Certification hackathons: every (other) 2nd week
- Official docs: http://dirac.readthedocs.io/en/latest/index.html
- GitHubWiki: https://github.com/DIRACGrid/DIRAC/wiki
  - e.g. for operational changes for new releases
- Google forum, for operations and assistance
- github issues, for… filing issues

- ~6 FTE as core developers, a dozen contributing developers

Funded by the European Union's Horizon 2020 - Grant N° 824064

DIRAC developments

- Virtual DIRAC Users' Workshop
- DIRAC needs an X509 certificate: tokens being investigated
- Staging the data, there is a RUCIO plug-in from Belle2
  - Constraints maybe placed on RUCIO storage setup: being investigated
- Python API
- ~~REST API~~ Deprecated (based on Tornado using M2Crypto)
- HTTP RPC
- Jupyterhub integration (commandline)

Funded by the European Union's Horizon 2020 - Grant N° 824064

What did we try

- My own server in a container: no certificate
- REST API
- Fake Rest API ok (but deprecated)
- Started looking at local RPC

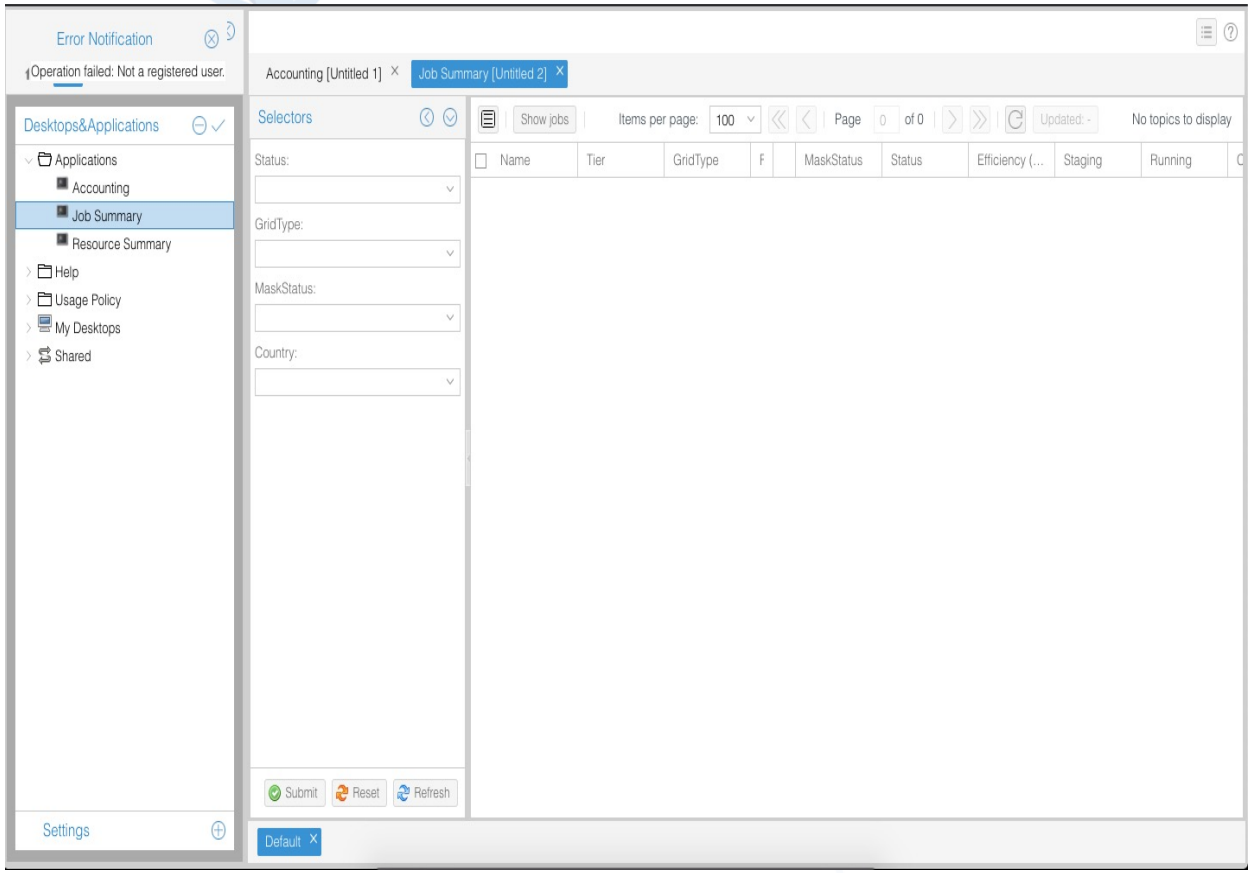
- Multiple talks with CTA-DIRAC people
- Andrei Tsaregorodtsev (several times - all roads lead back here)
- WP2-5 common meeting: some interest and volunteers

Funded by the European Union's Horizon 2020 - Grant N° 824064

Resource: EGI

- Compute Platform of the European Open Science Cloud

- Combined service inherited all the communities
  - 43 registered VOs
  - ~700 registered users
  - Accessible via the same endpoint: WebAPP: https://dirac.egi.eu/DIRAC/

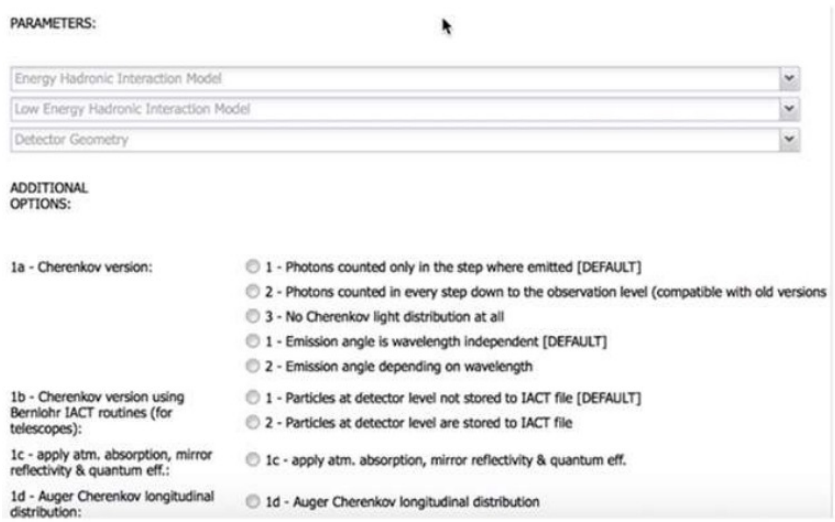


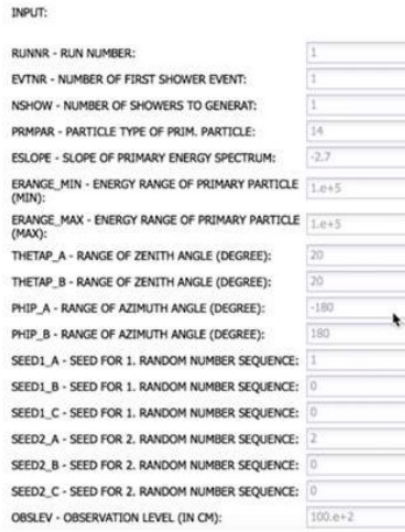
https://dirac.egi.eu/DIRAC/

Funded by the European Union's Horizon 2020 - Grant N° 824064

CONCORDIA (WP3 Task 3.3)

- ConCORDIA (Container Corskia DIRAC)
- Provide a common simulation tool between experiments (CTA & KM3Net)
  - Singularity containers to be deployed on the GRID
- Integrated as a DIRAC WebApp
  - Access to GRID job submissions and existing resources utilities
- EGI resources for DIRAC access
  - First developments in the EGI DIRAC-client docker
  - First tests on EGI-connected resources
- Containers can be tweaked on-demand
  - GUI for container creation ✅
  - Tune the simulation parameters according to needs ✅
  - CORSIKA setup and runcards ✅
- Running the containers:
  - Scripting access and management ✅
  - GUI access and management ✅
- They would like to integrate with ESAP
  - Certificates / IAM

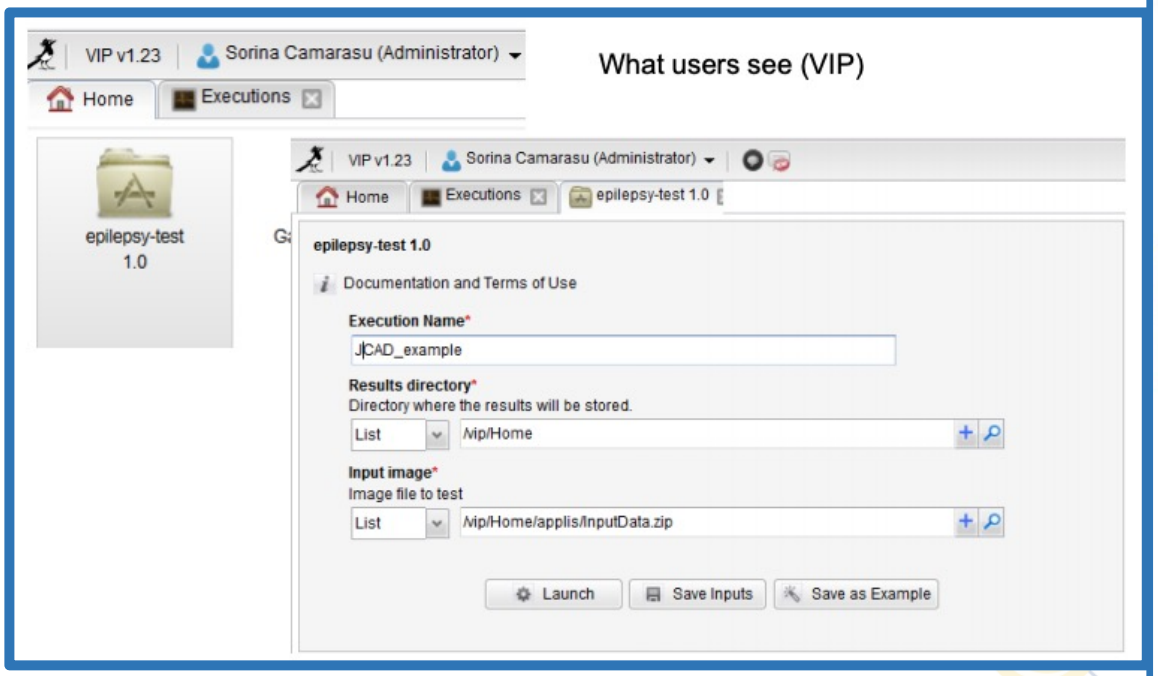


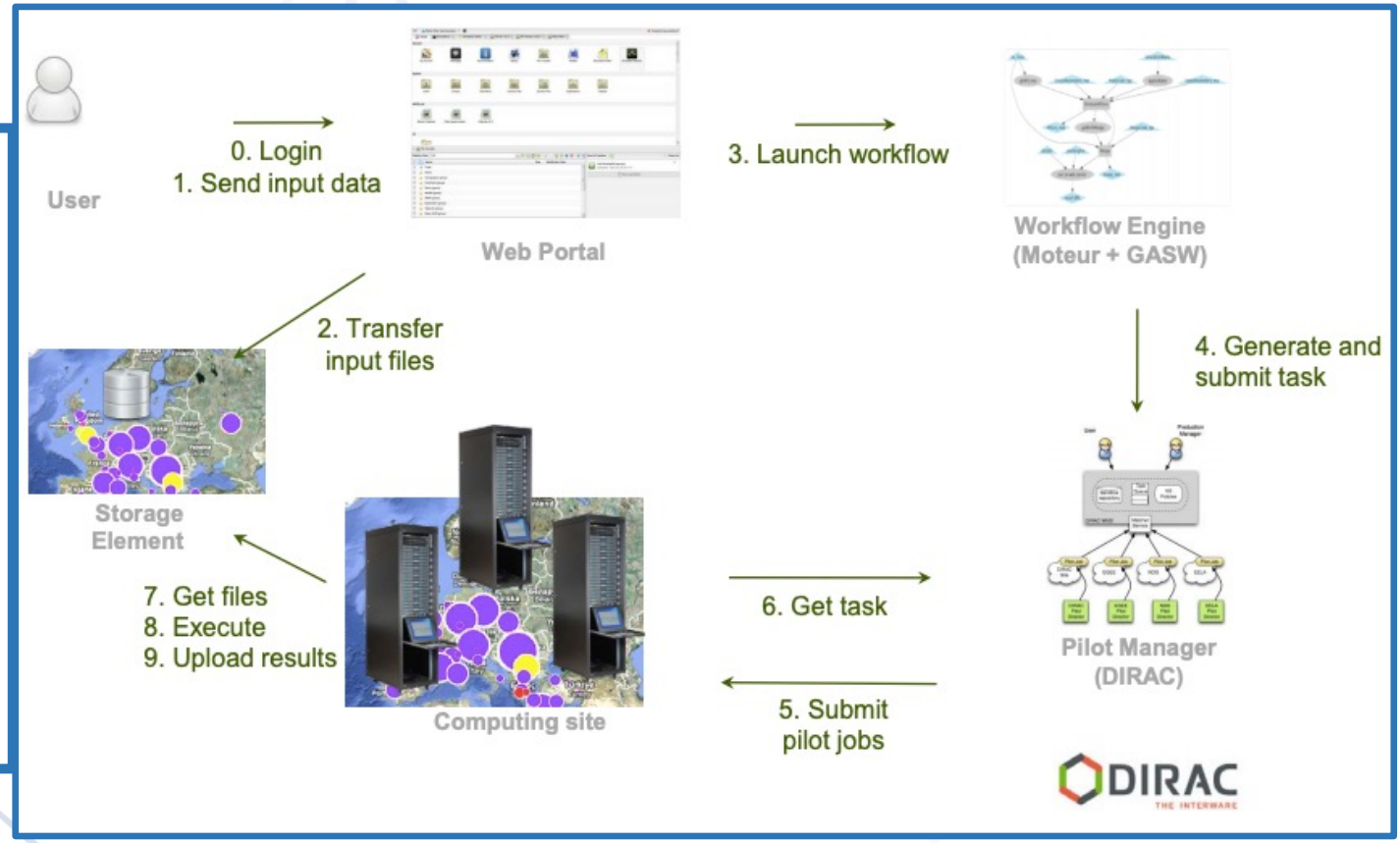


Funded by the European Union's Horizon 2020 - Grant N° 824064

Other examples

- Lyon Medical website: https://vip.creatis.insa-lyon.fr/
- https://www.creatis.insa-lyon.fr/site7/sites/www.creatis.insa-lyon.fr/files/Exploiting_GPUs_with_VIP_and_Dirac.pdf





Funded by the European Union's Horizon 2020 - Grant N° 824064

Discussion

- Is there an interest in integrating a batch system on ESAP?
- Is there interest in DIRAC?
- Who would be willing to join regular meeting?

- Challenges / Open Questions:
  - Authentication (would be true RE all external services)
  - Datalake / shopping basket (what if the data is not near the compute resources)
  - user ID / IAM (getting the x509, do you have to reinstall initiate the client)
  - Identifying workflows / containers and from OSSR
  - Identifying the resources needed to run these
  - …

Funded by the European Union's Horizon 2020 - Grant N° 824064

Backup Slides

Funded by the European Union's Horizon 2020 - Grant N° 824064

RPC

- The xmlrpc.server module provides a basic server framework for XML-RPC servers written in Python

```
import xmlrpc.client

s = xmlrpc.client.ServerProxy('http://localhost:8000')
print(s.pow(2,3))  # Returns 2**3 = 8
print(s.add(2,3))  # Returns 5
print(s.mul(5,2))  # Returns 5*2 = 10

# Print list of available methods
print(s.system.listMethods())
```


Funded by the European Union's Horizon 2020 - Grant N° 824064

Hello world?

- First steps:
  - Compute Resources (LAPP CTA DIRAC, EGI, other instances ….)
  - X509 Grid Cert
  - Access to specific VO (Multi VO should be available)
  - Local DIRAC client installation
  - Run ”Hello world” (on the command line)
  - Create “Archive” for DIRAC on in ESAP (started)
  - Simple one click launch of “Hello world”
  - Add parameter (box/drop down) run again

- Other possible steps:
  - Compute Resources (LAPP CTA DIRAC, EGI, other instances ….)
  - X509 Grid Cert
  - Access to specific VO (Multi VO should be available)
  - Create “Archive” for DIRAC on in ESAP (started)
  - Use the http RPC
  - Simple one click launch of “Hello world”
  - Add parameter (box/drop down) run again

Funded by the European Union’s Horizon 2020 - Grant N° 824064