The European Science Cluster of Astronomy & Particle Physics ESFRI research infrastructures

Xavier Espinal (CERN) with contributions from:

R. Bolton, A. Bruzzese, S. Campana, A. Ceccanti, M. Delfino, D. Ciangottini, R. Di Maria, R. Dona, A. Fkiaras, Y. Grange, S. Jezequel, G. Merino, P. Millar, G. Rahal, A. Sarwar, D. Spiga

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.

Science Projects

HiLumi HL-LHC PROJECT
SKA SQUARE KILOMETRE ARRAY
VIRGO
FAIR
KM3NeT
JIVE Joint Institute for VLBI ERIC
cta cherenkov telescope array
LSST
EST european solar telescope
ESO
CERN

ESCAPE

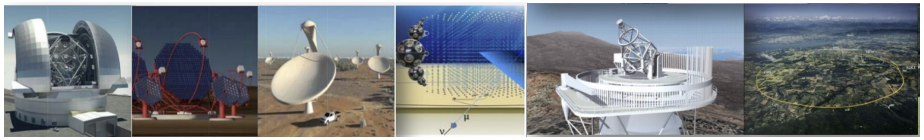
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

EUROPEAN OPEN SCIENCE CLOUD

Horizon 2020 funded project

Project Goals

- Prototype an infrastructure adapted to Exabyte-scale needs of large science projects
- Ensure the sciences **drive** the development of the EOSC
- Address FAIR data management principles



Data centres

CERN
INFN
GSI
DESY
Nikhef
SURF SARA
rijksuniversiteit groningen
CCIN2P3
cnrs
PIC port d'informació científica
LAPP Laboratoire d'Annecy de Physique des Particules
INAF ISTITUTO NAZIONALE DI ASTROFISICA NATIONAL INSTITUTE FOR ASTROPHYSICS

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

The ESCAPE Project Work Packages

- **ESCAPE Data Lake (DIOS)**: a scalable federated data infrastructure as the basis of an open science for the ESFRI projects within ESCAPE
- **ESCAPE Software Repository (OSSR)**: the repository of scientific software services of the research infrastructures concerned by the ESCAPE
- **ESCAPE Virtual Observatory (EVO)**: astronomical high-level products archive and related services
- **ESCAPE Science Platform (ESAP)**: a flexible science platform for the analysis of open access data
- **ESCAPE Citizen Science (CS)**: an open gateway dedicated to the public through Citizen Science and communication actions

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

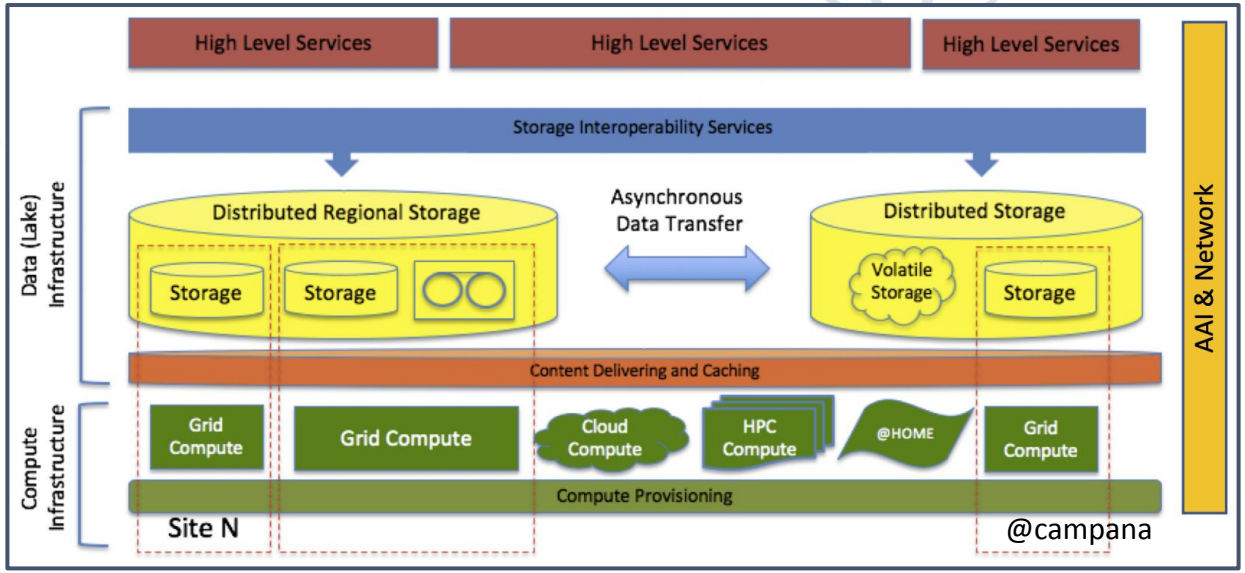
ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

ESCAPE DIOS | Data Infrastructure for Open Science

- The DIOS work package aims at **delivering a Data Infrastructure for Open Science**. This infrastructure is a **non HEP specific implementation of the Data Lake concept** elaborated in the HSF Community White Paper and endorsed in the WLCG Strategy Document for HL-LHC
- The science projects in ESCAPE are in different phases of evolution, some of them are **defining now** their computing models. Special interest on data storage, organisation, management and access.
- The backbone of the ESCAPE Data Lake consists of services operated by the partner institutes and connected through reliable networks. Leveraging the existing expertise in LHC/WLCG:
  - Data management and organisation orchestrated with **Rucio**. Data Transfers steered by **FTS**. Layer of caching and latency hiding services based on **XCache**. **CRIC** is used as global Data Lake Information System. Token based authentication and authorisation system implemented by Indigo **IAM** (with legacy x509 support)
  - Supporting various access protocols (**http**, **xrootd** and *gridftp*) to serve the data to heterogeneous facilities, from conventional **Grid** sites to **HPC** centres and **Cloud** providers

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

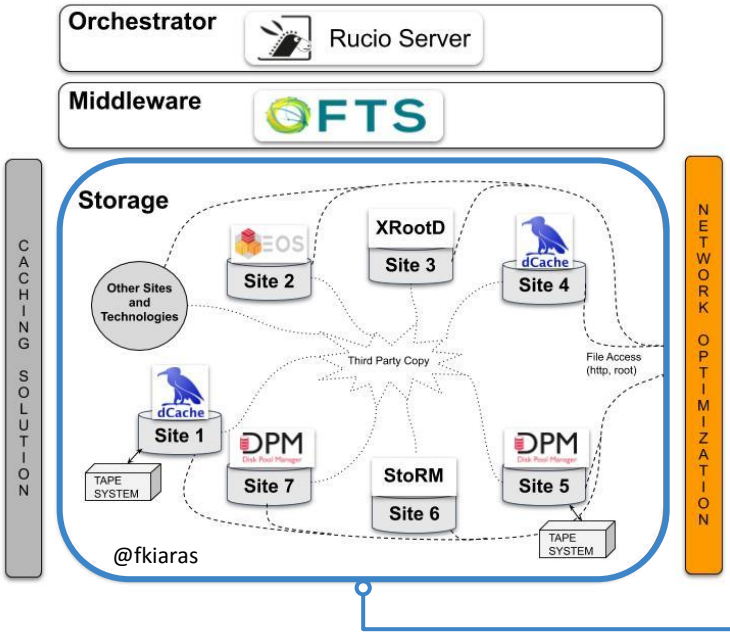
The ESCAPE Data Infrastructure for Open Science

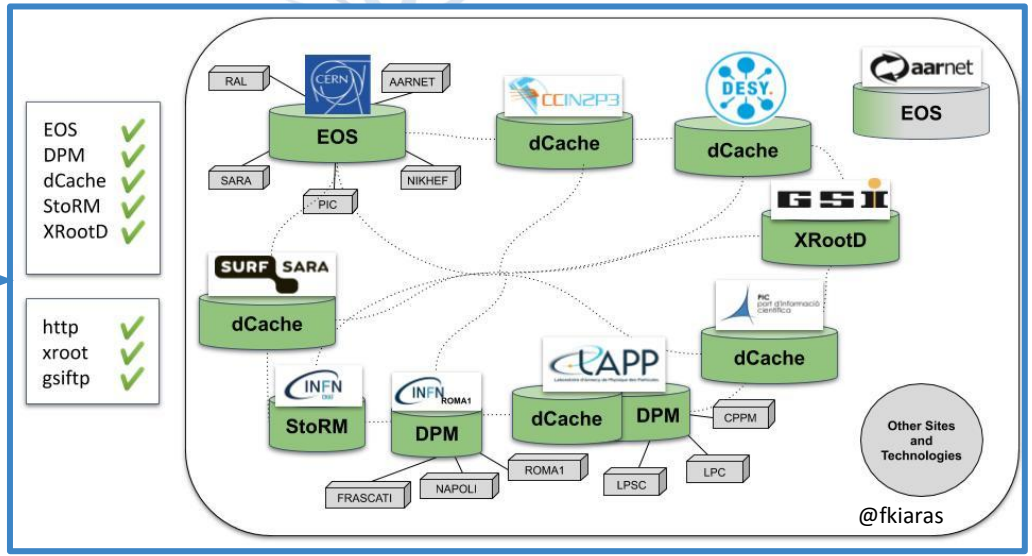


- Define, integrate and commission an ecosystem of tools and services to build a data lake, in **synergy with the WLCG DOMA project**
- Science **projects to drive** the services requirements most suitable to their needs
- Contributes to deliver **Open Access and FAIR data services**: trustable data repositories; enable data management policies; transparent data access layer

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

The ESCAPE Data Lake





- Hiding complexity and providing transparent access to data
- Heterogeneous federated storage and operations model
- Some centers joining even if not funded by ESCAPE

Further info: https://wiki.escape2020.de/index.php/WP2_-_DIOS#Datalake_Status

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

Monitoring and IS (1/2)

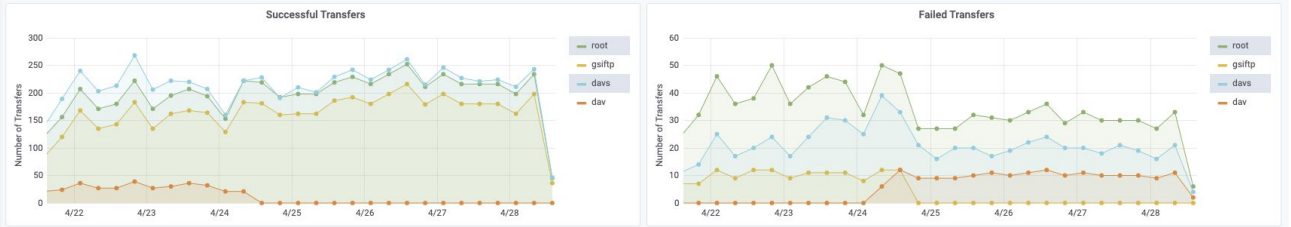
filter by Experiment site | filter by Storage Unit | ACTIVE

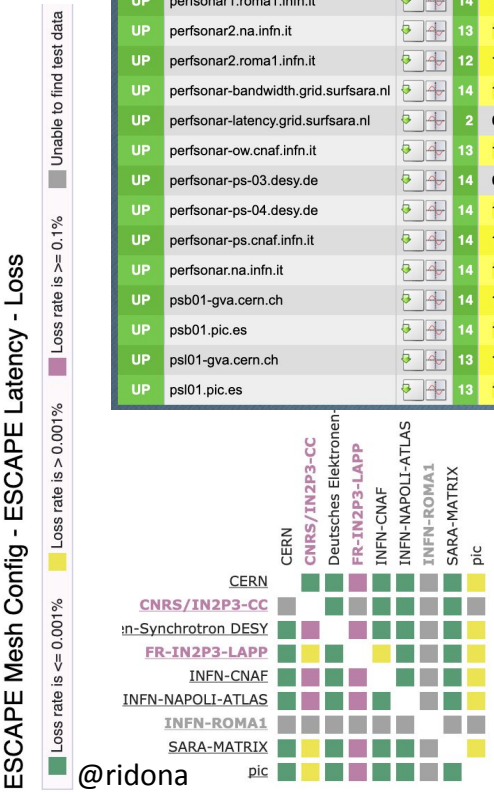
Experiment site	Storage Unit	State
CNAF-STORM-ES	CNAF-STORM	ACT
CNAF_CMS_TEMP	CNAF_CMS_TEMP	AC
DESY-DCACHE	DESY-DCACHE	ACTIVE
EULAKE-1	EULAKE-1	ACTIVE
EULAKE-2	EULAKE-2	ACTIVE
IN2P3-CC-DCACHE	IN2P3-CC-DCACHE	ACTIVE
INFN-NAPOLI-DPM	INFN-NAPOLI-DPM	ACTIVE
LAPP-DCACHE	LAPP-DCACHE	ACTIVE
LAPP-DPM	LAPP-DPM	ACTIVE
PIC-DCACHE	PIC-DCACHE	ACTIVE
PIC-DCACHE	PIC-DCACHE	ACTIVE
SARA-DCACHE	SARA-DCACHE	ACTIVE

perfSONAR

	Host	Icons	OK	Wa	Un	Cr	Pd
	perfsonar1.in2p3.fr		14	1	0	0	0
	ccperfsonar2.in2p3.fr		13	1	0	0	0
UP	cmsrm-perfsonar1.roma1.infn.it		13	0	3	0	0
UP	lapp-ps01.in2p3.fr		14	1	0	0	0
UP	lapp-ps02.in2p3.fr		13	1	0	0	0
UP	perfsonar1.roma1.infn.it		14	1	0	0	0
UP	perfsonar2.na.infn.it		13	1	0	0	0
UP	perfsonar2.roma1.infn.it		12	1	0	1	0
UP	perfsonar-bandwidth.grid.surfsara.nl		14	1	0	0	0
UP	perfsonar-latency.grid.surfsara.nl		2	0	9	3	0
UP	perfsonar-ow.cnaf.infn.it		13	1	0	0	0
UP	perfsonar-ps-03.desy.de		14	0	0	0	0
UP	perfsonar-ps-04.desy.de		14	1	0	0	0
UP	perfsonar-ps.cnaf.infn.it		14	1	0	0	0
UP	perfsonar.na.infn.it		14	1	0	0	0
UP	psb01-gva.cern.ch		14	1	0	0	0
UP	psb01.pic.es		14	1	0	0	0
UP	psl01-gva.cern.ch		13	1	0	0	0
UP	psl01.pic.es		13	1	0	0	0

FTS





@ridona

Monitoring and IS (2/2)

WebUI Portal for Core Resources

DOMA RucioStorageElement PIC-DCACHE

General Information

RSE Name	PIC-DCACHE
Storage Unit	PIC-DCACHE
Last modification date	2019-11-20 16:51:25.774627

State

Object state	ACTIVE
State comment	

Attributes

Space Usage URL	
Deterministic	True
Volatile	False
LFN to PFN Algorithm	hash
Credentials	
RSE Type	
Relation to FTS	https://fts3-pilot.cern.ch:8446

Edit

DELETE LAN | DELETE WAN | READ LAN | READ WAN | Third party copy | WRITE LAN | WRITE WAN

Priority	Rucio Priority	Endpoint	Protocol	Activity	Edit protocol
1	2	door05.pic.es:8452	davs	delete_lan	
2	2	xrootd.pic.es:1094	root	delete_lan	

Re-arrange protocols | Use new Protocol | Create new Protocol

DOMA RucioStorageElement EULAKE-1

General Information

RSE Name	EULAKE-1
Storage Unit	EULAKE-1
Last modification date	2019-11-20 16:50:53.233567

State

Object state	ACTIVE
State comment	

Attributes

Space Usage URL	
Deterministic	True
Volatile	False
LFN to PFN Algorithm	hash
Credentials	
RSE Type	
Relation to FTS	https://fts3-pilot.cern.ch:8446

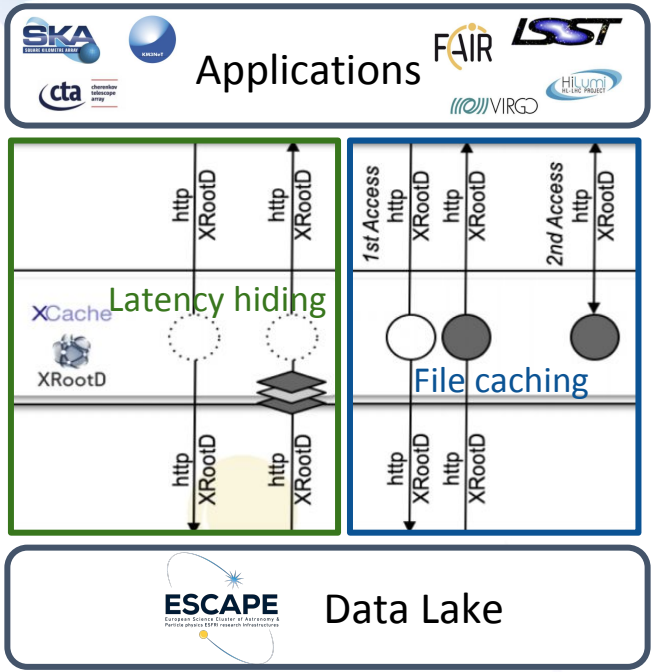
Edit

DELETE LAN | DELETE WAN | READ LAN | READ WAN | Third party copy | WRITE LAN | WRITE WAN

Priority	Rucio Priority	Endpoint	Protocol	Activity	Edit prot
1	1	eoseulakehttp.cern.ch:443	davs	delete_lan	
2	2	eulake.cern.ch:1094	root	delete_lan	
3	3	eulakeftp.cern.ch:2811	gsiftp	delete_lan	

Content delivery and caching (1/2)

- Leveraging the know-how being acquired in DOMA/WLCG with XCache investigations
- Effort made towards a vanilla installation (experiment-unbiased) caching service. Easy deployable by the partners
  - Installations at CERN, CNAF and CC-IN2P3
- Main use-cases:
  - Latency hiding and file re-usability
    - benchmarking multi-caching layers between client and origin
    - http and tokens aware
  - Facilitate ingress/egress with Commercial Clouds and HPC
- Main goal:
  - Investigate and understand whether caching can help on non-event based files e.g. images, data-cubes,...



@dimaria

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

Content delivery and caching (2/2)

The objective: produce a working prototype for CMS use case in ESCAPE data-lake.

- Analysis data on a lake endpoint + cache layer + computing facility for data processing
  - we start little with a single storage/lake endpoint at CNAF
    - to then extend at least to CERN

Embargoed data

- Integrating **capability based** auth
  - we used ESCAPE IAM to self mana
- we uses NANOAOD based analysi
  - as an interesting use case for futu

Open data

- dataset imported in the data lake

An XCache server serving data to the analysis facility

- both xrootd and https flow
- keeping ACLs in sync with the lake endpoint for both used protocols

XRootD

The XCache server points to the origin above

IS) can access data
oxy to contact origin
should have a super-user proxy
On hold, waiting for a version to test

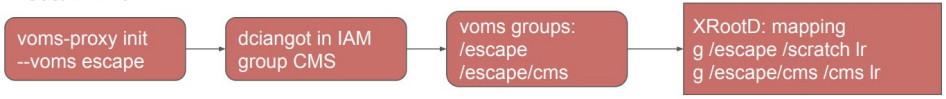
/cms scope can access it through https

Tests summary

- Checked the correct ACLs management
- Tried RUCIO download of a registered embargoed data
- Submitted CMS condor jobs reading through the cache
  - XRootD protocol
  - via WebDAV with DAVIX
- Already visible some latency hiding effect in this simple test setup

@spiga,@ciangottini

X509/IAM flow



Full JWT scope authz

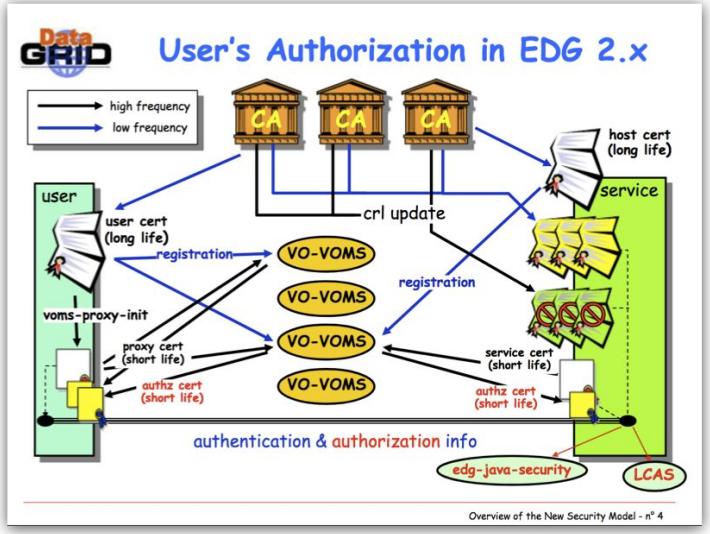
Still work to be done for full token integration, e.g. token and scope based authz, multi-VO/group support

More info: https://indico.in2p3.fr/event/19942/

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

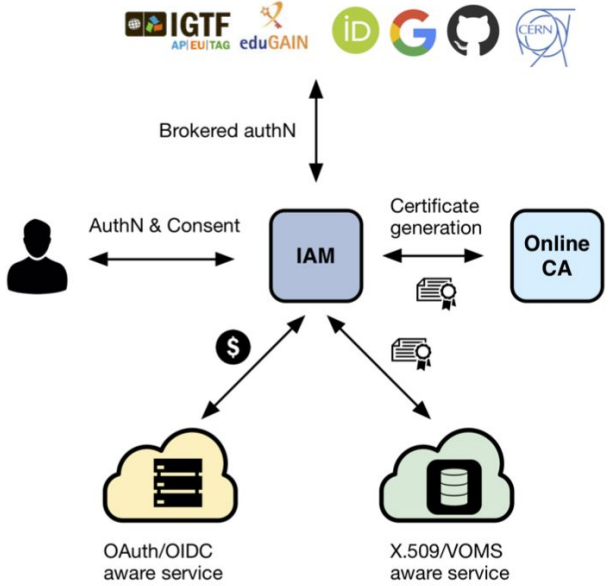
AuthN/Z in the ESCAPE Data-lake testbed (1/2)

Current, X.509 based AAI



Move beyond X.509

Future, token-based AAI



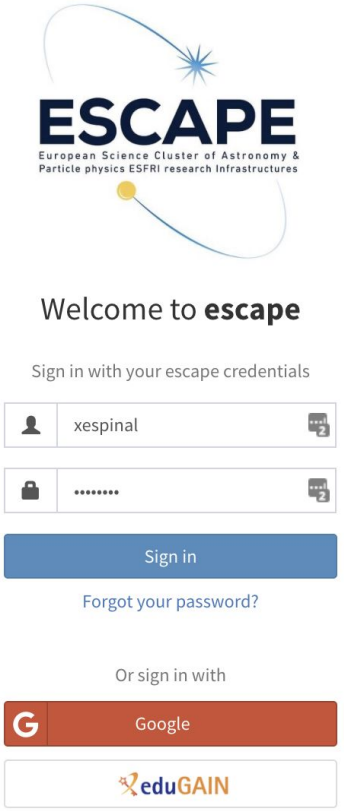
Approach: leverage and build upon the WLCG experience

@ceccanti

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

AuthN/Z in the ESCAPE Data-lake testbed (2/2)

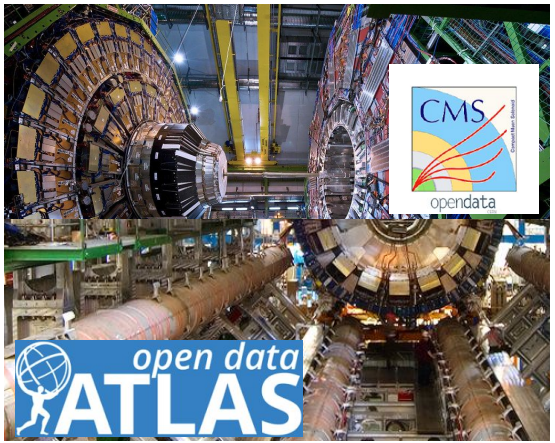
- Start with “traditional” Grid AuthN/Z approach
  - GSI X.509 authN + VOMS authorization, Coarse-grained VO-level authorization, Fine-grained group/role-based authorization
- Demonstrate Token-based AuthN/Z approach
  - Flexible AuthN (e.g., EduGAIN) + OAuth-based authorization, Coarse-grained VO-level authorization, Fine-grained, group or scope-based authorization
- ESCAPE IAM instance deployed and integrated with EduGAIN, supporting grid and token approaches
  - GSI/VOMS AuthN/Z supported by all data management services
  - Token-based AuthN/Z supported by most data management services
- Next steps:
  - Demonstrate interoperability/integration for communities already having their own AAI solution/user database
  - Showcase fine-grained AuthZ in support of use cases with embargoed data
  - Support other WPs integration efforts with training events/f2f hackathons



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

Data and Data access in the ESCAPE Data Lake (1/3)

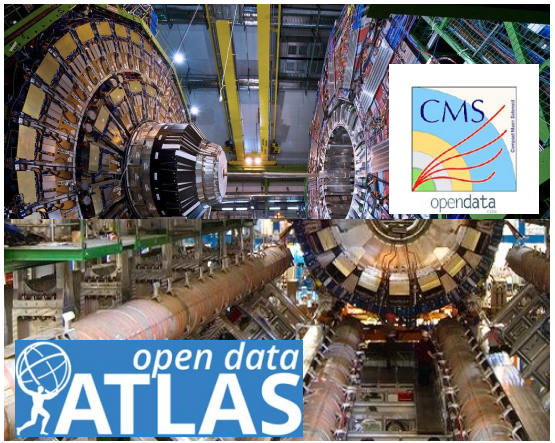
- **LOFAR** datasets, data injection from SURF/SARA
  - 354 tar files, file size between 3GB and **450**GB for a total of 32TB
- **CMS** Open and Embargoed data (300 files, 1TB), data injection from CNAF
- **ATLAS** Open Data (few files uploaded for a PoC), data injection from LAPP
- **LSST and CTA** 3.5k files for a total of 700GB, data injection from CC-IN2P3 and LAPP
- **SKA** - few test files, data injection from SKAO



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

Data and Data access in the ESCAPE Data Lake (2/3)

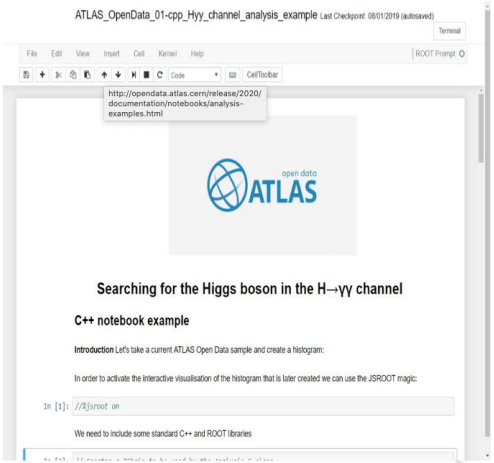
- Setting up data access use cases for LOFAR, CMS, ATLAS, LSST, CTA and SKA (KM3net joining soon)
- Show how the use case fits the current model for Radio Astronomy. Making sure we have a definition of what the "parameter space is" to assess the representativity of the use case (data type, access patterns)



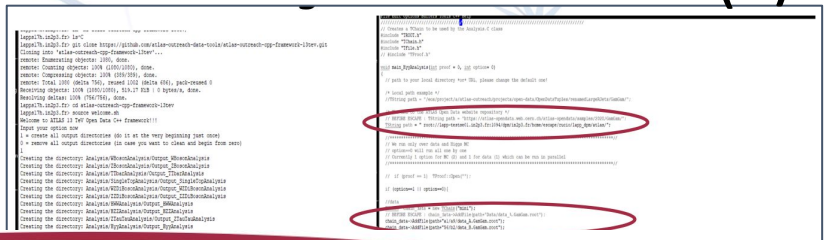
Funded by the European Union's Horizon 2020 - Grant N° 824064

Data and Data access in the ESCAPE Data Lake (3/3)

ATLAS analysis demo



@jezequel



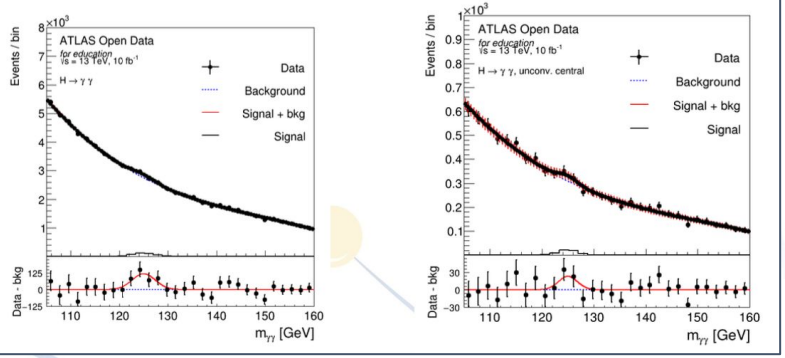
```
// BEFORE ESCAPE : TString path = "https://atlas-opendata.web.cern.ch/atlas-opendata/samples/2020/GamGam/";
TString path = " root://lapp-testse01.in2p3.fr:1094/dpm/in2p3.fr/home/escape/rucio/lapp_dpm/atlas/";
```

- C++ code
- Produce H → yy plot

- ESCAPE exercise :
  - Upload files to ESCAPE datalake with rucio client
  - Adapt file access to ESCAPE datalake
  - Produce plots

- Path for input files :
  - Can be automatised decoded by just providing dataset and location
  - Pattern can nbe built to include xcache in the path

- Post processing root code to make plots



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

Storage Quality of Service - QoS (1/2)

- Prototyping and testbed
  - *Deploying* storage endpoints with different QoS
  - Providing demos that showcase QoS possibilities
  - Building up practical experience within ESFRI communities on using QoS
- Scientific engagement
  - Build contacts within ESFRI communities
  - Conduct interviews to understand how QoS may match their workflows (or anticipated workflows)
    - Initial interview completed with ATLAS, will be used as a template for other ESFRI interviews
  - Update testbed (and potentially software) to match desired QoS usage
    - Interview process with SKA and CTA has started, in the preparatory phase for FAIR
- Software development
  - Build architecture / design paper
    - Design paper has been written during ESCAPE QoS mini-workshop (CERN, 27 Feb)
  - Identify limitations of current approach
  - Implement missing functionality (based on feedback)

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

Storage Quality of Service - QoS (2/2)

Demo (continued) 1

- Create RSE -- done once

```
$ rucio-admin rse add QOS-A-PIC
Added new deterministic RSE: QOS-A-PIC

QOS-A-PIC
QOS-B-SARA
```

- Add Protocol about RSE

```
$ rucio-admin rse add-protocol --hostname xrootd.pic.es --scheme root --prefix '/pnfs/pic.es/tape/' --port 1094 QOS-A-PIC
```

- RSE Properties

```
$ rucio-admin rse info PIC-DCACHE
```


Demo (continued) 2

- Add attribute on RSE -- likely done only once

```
$ rucio-admin rse set-attribute --rse QOS-A-PIC --key QOS --value A
Added new RSE attribute for QOS-A-PIC: QOS-A
```

- List RSE's by QoS Label

```
$ rucio list-rses --expression 'QOS=B'
PIC-DCACHE
CNAF-STORM
```

Demo (continued) 3

- File Upload -- done once per file

```
$ rucio upload test_qos --scope testing --rse DESY-DCACHE

2020-04-20 16:01:52,856  INFO Preparing upload for file test_qos
2020-04-20 16:01:53,037  INFO Successfully added replica in Rucio catalogue at DESY-DCACHE
2020-04-20 16:01:53,178  INFO Successfully added replication rule at DESY-DCACHE
2020-04-20 16:01:53,799  INFO Trying upload with davs to DESY-DCACHE
2020-04-20 16:01:54,821  INFO Successfully uploaded file test_qos
```

@aleem@afkiaras https://indico.in2p3.fr/event/21221/

Demo (continued) 4

- File availability on a QOS=B Site

```
$ rucio add-rule testing:test_qos 1 'QOS=B'
1f87409a72934e0bab2e9168ae3f5d58
```

- Requiring particular QoS -- done each time desired QoS changes.
- Check file transfer status

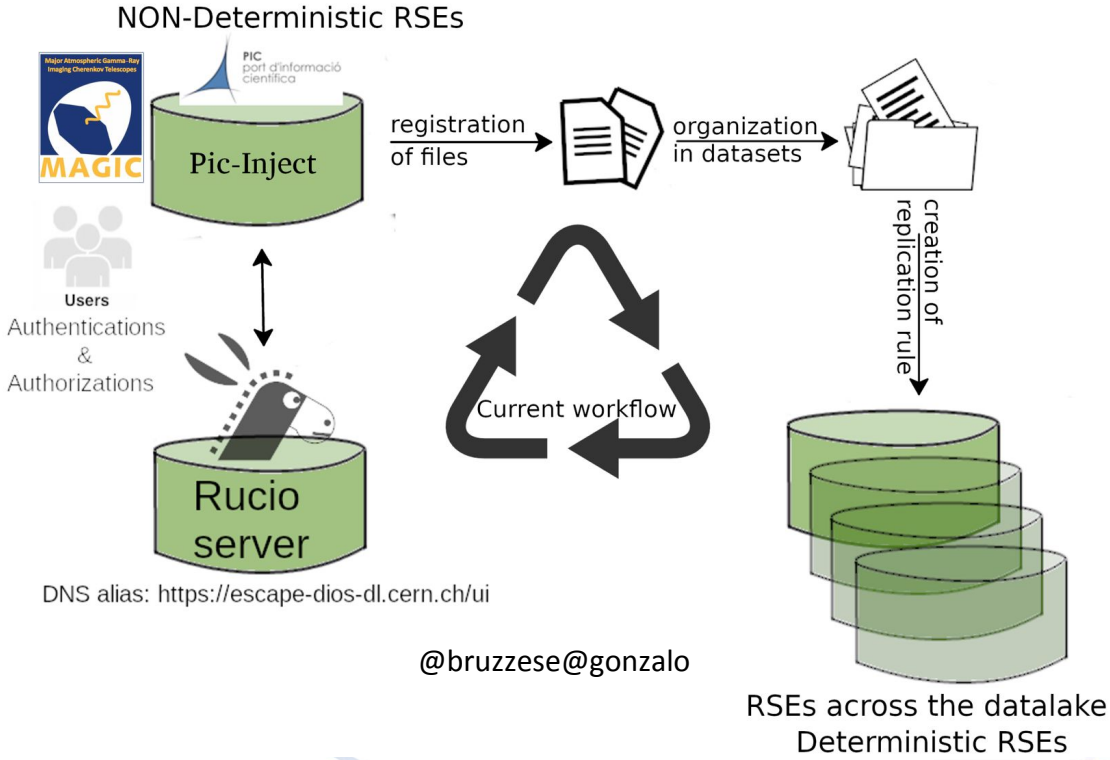
```
$ rucio list-rules testing:test_qos

ID                                ACCOUNT  SCOPE:NAME        STATE[OK/REPL/STUCK]  RSE_EXPRESSION  COPIES  EXPIRES (UTC)          CREATED (UTC)
--------------------------------  -------  ----------------  --------------------  --------------  ------  ---------------------  ---------------------
1a97227b69034c5d9c3528d542831bde  root     testing:test_qos  OK[1/0/0]             DESY-DCACHE     1       2020-04-20 14:01:53    2020-04-20 13:01:53
1f87409a72934e0bab2e9168ae3f5d58  root     testing:test_qos  REPLICATING[0/1/0]    QOS=B           1       2020-04-20 14:06:47    2020-04-20 13:06:47
```

17

Demonstrator: PIC γ-ray telescope data injector (1/2)

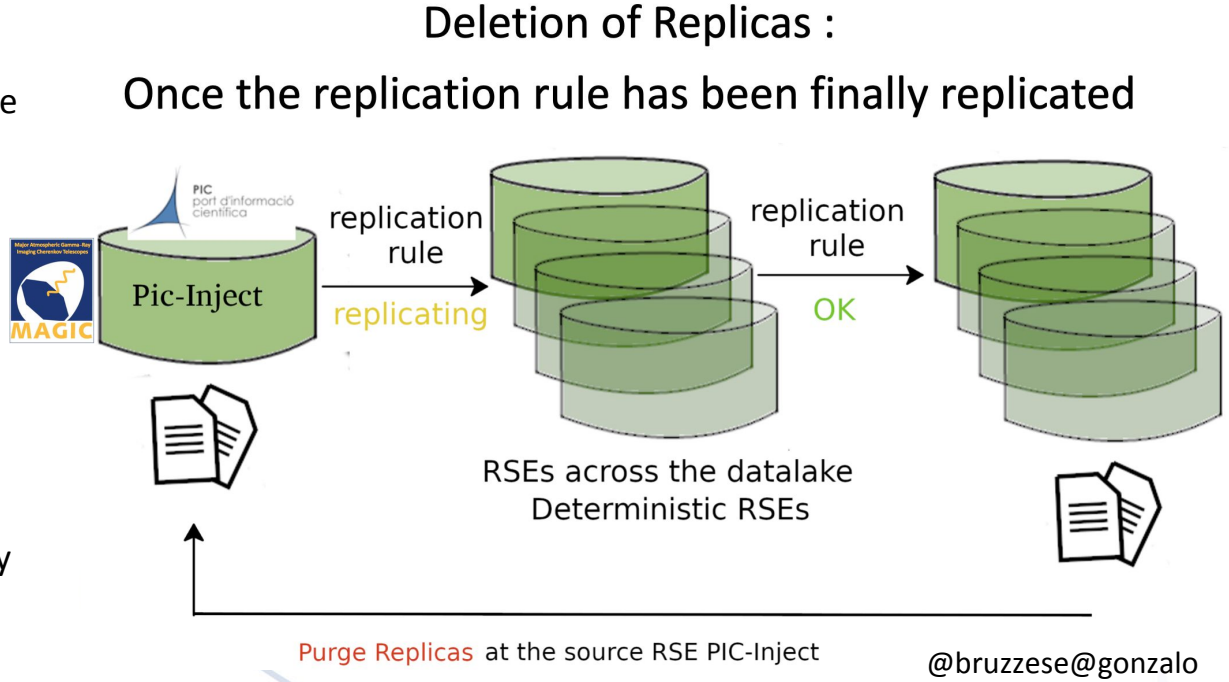
- Data acquisition demonstrator
- Files streamed from the telescope to the Data Lake for permanent storage and access
- Data is injected from the storage system located at the source
  - The origin RSE is a non-deterministic RSE allowing to register files with their original path in the detector
- Mock data used by now. Plan to use MAGIC telescope real data by next month.
  - Distribution and processing tests.



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

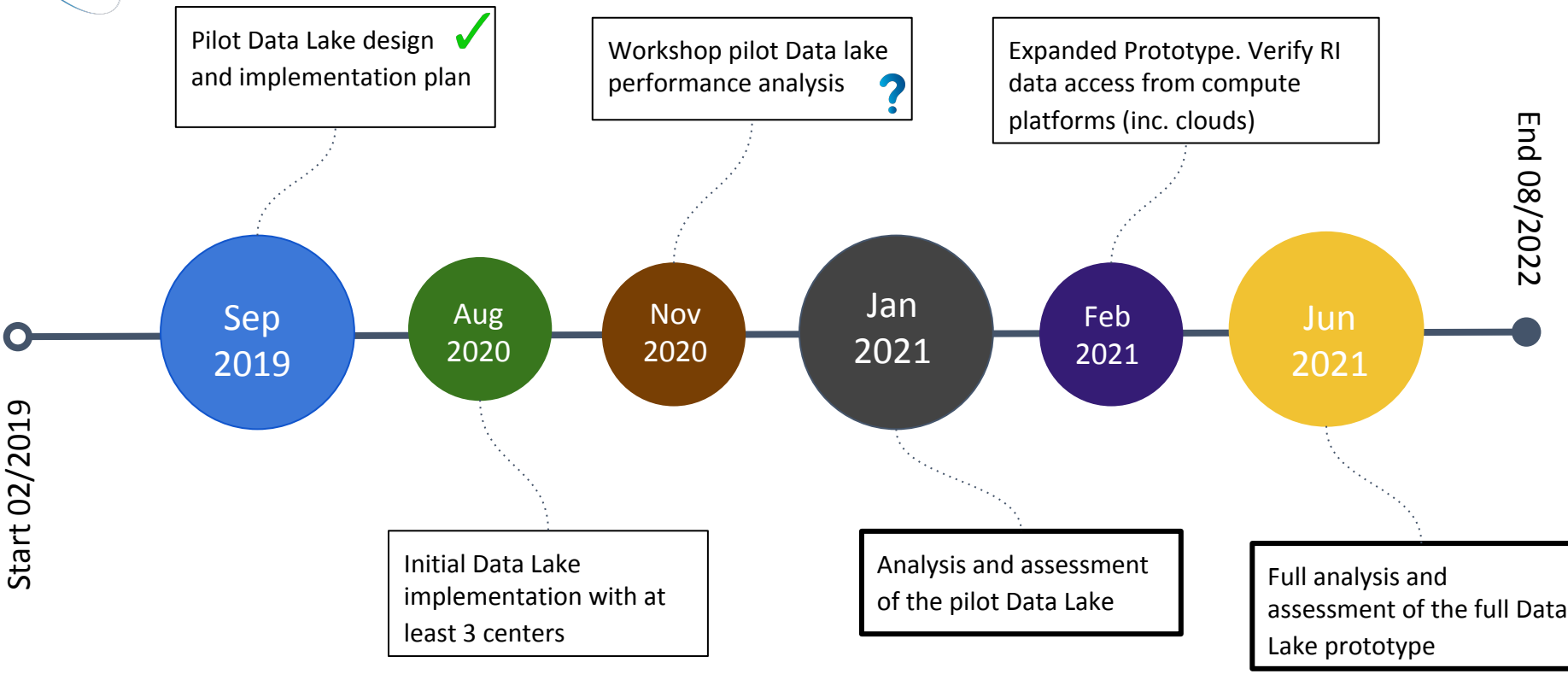
Demonstrator: PIC γ-ray telescope data injector (1/2)

- Data acquisition demonstrator
- Files streamed from the telescope to the Data Lake for permanent storage and access
- Data is injected from the storage system located at the source
  - The origin RSE is a non-deterministic RSE allowing to register files with their original path in the detector
- Mock data used by now. Plan to use MAGIC telescope real data by next month.
  - Distribution and processing tests.



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

Some important milestones



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

Next steps (not an exhaustive list)

- 2020: Pilot Data Lake Functional Data Transfer Tests Machinery
  - Data Transfer Tests Machinery to demonstrate stable and sizeable data movement across sites in the datalake
    - Moving bulk data from A to B (to C), changing QoS within a site and across sites (on demand, by policy)
    - Datalake data can be accessed by clients (e.g. simple workloads)
  - *Performance* monitoring in place (e.g. transfer matrix):
    - Show transfer metrics: bandwidth, number of files, success/errors, perfsonar,…
      - Allowing us to debug the infrastructure: network, storage, data management tools,...
    - Basis for an operations and deployment model
- 2021: Extending the pilot to a full Data Lake prototype
  - Automated infrastructure testing (based on HC) ready to run realistic research infrastructure workloads
  - Real data distribution and analysis for several non-HEP RI and HL-LHC reference workloads
  - Ability to plug comercial cloud resources into the datalake infrastructure
  - Caching mechanism deployed, vanilla software suite ready and deployed on several sites
  - Bonus: PoC integration of SA and AUS sites into the datalake infrastructure (SKA/LOFAR data)
- Continue the **synergy and feedback** with the WLCG DOMA R&D projects

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

ESCAPE

European Science Cluster of Astronomy &
Particle physics ESFRI research Infrastructures

Thanks for listening!

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