Grid - Hands-on

ATLAS-D Physics Meeting Göttingen 2021

Dr. Gen Kawamura
II.Physikalisches Institut, Universität Göttingen

Exercise overview (30mins)

- Your Grid environments
  - Lxplus at CERN, or NAF
  - Environments/Scripts for this tutorial
- Introduction to Grid computing
- PanDA (ATLAS job management system)
  - Hello world prun Grid job
  - Hello world pathena Grid job
  - Hello world PyRoot
  - Hello world PyRoot Grid job
- Rucio (ATLAS data management system)
  - Rucio Client CLI
- RucioUI (WebUI)
- ATLAS Metadata Interface (AMI)
  - AMI Client CLI
- Extra: Information Provider (of gLite)

Your Grid environments

Lxplus at CERN, or NAF

- Loging in lxplus (Scientific Linux 7)

```
## If you have a CERN account
ssh -Y YOUR_CERN_ACCOUNT@lxplus.cern.ch
```

- NAF
  - https://confluence.desy.de/display/IS/NAF+-+National+Analysis+Facility

```
## If you have a NAF account (yourusername = username, atlasXX = node)
ssh -XY yourusername@nafhh-atlasXX.desy.de
```

Environments/Scripts for this tutorial

- From GitHub

Cloning materials
git clone https://github.com/GenKawamura/ATLAS-D_2021_Computing_Tutorial
cd ATLAS-D_2021_Computing_Tutorial

Introduction to ATLAS Grid Computing

Setup CVMFS

- In LXPLUS, use setupATLAS command
- For example. write the command aliases in ~/.bashrc

```
## Alias to initialization of VOMS proxy
alias vinit='voms-proxy-init --voms atlas -hours 200 --valid 200:00'

## Alias to setupCVMFS
setupCVMFS(){
   export LCG_LOCATION=
   export ATLAS_LOCAL_ROOT_BASE=/cvmfs/atlas.cern.ch/repo/ATLASLocalRootBase
   source $ATLAS_LOCAL_ROOT_BASE/user/atlasLocalSetup.sh ""

   ## Using EMI LCG package
   source ${ATLAS_LOCAL_ROOT_BASE}/packageSetups/atlasLocalEmiSetup.sh --emiVersion ${emiVersionVal}
}

## Using CVMFS (with EMI LCG client tools)
setupCVMFS
```

Hands-on exercise user certificate - 1

- Getting your user certificate (if you have, skip)

```
## Check important environment variables for your certificate
env | grep X509

## Generating a proxy certificate
export X509_USER_CERT=~/.globus/usercert.pem
export X509_USER_KEY=~/.globus/userkey.pem

## Generate user certificate
## (usercert.p12 was already exported by your browser)
openssl pkcs12 -clcerts -nokeys -in usercert.p12 -out $X509_USER_CERT

## create a private certificate with passphrase
openssl pkcs12 -nocerts -in usercert.p12 -out $X509_USER_KEY

## Set permissions
chmod 644 $X509_USER_CERT
chmod 400 $X509_USER_KEY

## show enddate
openssl x509 -in $X509_USER_CERT -noout -enddate

## show if the certificate is valid
openssl verify -CApath $X509_CERT_DIR -purpose sslclient $X509_USER_CERT
```

Hands-on exercise user certificate - 2

- Checking your certificate and VO

Generating a proxy
voms-proxy-init
Enter GRID pass phrase for this identity:
Contacting voms2.cern.ch:15001 [/DC=ch/DC=cern/OU=computers/CN=voms2.cern.ch] "atlas"...
Remote VOMS server contacted succesfully.

voms2.cern.ch:15001: The validity of this VOMS AC in your proxy is shortened to 345600 seconds!

Generating a proxy certificate without VO
grid-proxy-init
voms-proxy-info -all

(it displays information without VO attributes)

Generating a proxy certificate with VO (a normal use)
voms-proxy-init --voms atlas -hours 200
voms-proxy-info -all

(it displays information with VO attributes)

Using another role (if you have another)
voms-proxy-init -voms atlas:/atlas/de/Role=production
voms-proxy-info -all

Hands-on exercise user certificate - 2

- Checking your certificate and VO

```
## Check context of your certificate
## The proxy certificate has 3 fields (Public Key, New Public Key, New Secret Key)
less /tmp/x509up_u$UID | grep '\-'

-----BEGIN CERTIFICATE-----
-----END CERTIFICATE-----
-----BEGIN RSA PRIVATE KEY-----
-----END RSA PRIVATE KEY-----
-----BEGIN CERTIFICATE-----
-----END CERTIFICATE-----

## Check X509 attribute
openssl x509 -in /tmp/x509up_u$UID -text | less

##  Using a different proxy certificate
## (switch them if you have several ones)
mv -v /tmp/x509up_u$UID /tmp/x509_different_cert
export X509_USER_PROXY=/tmp/x509_different_cert
voms-proxy-info -all
```

PanDA (ATLAS Job Management System)

Hello World prun Grid Job

Hands-on exercise
Using ATLAS client tools

- First “Hello world” job by PanDA client

PanDA client
lsetup panda

Make a Python script
cat hello_world.py

```
#!/usr/bin/python
print "Hello world!"
```

chmod 755 hello_world.py
./hello_world.py

```
Hello world!
```

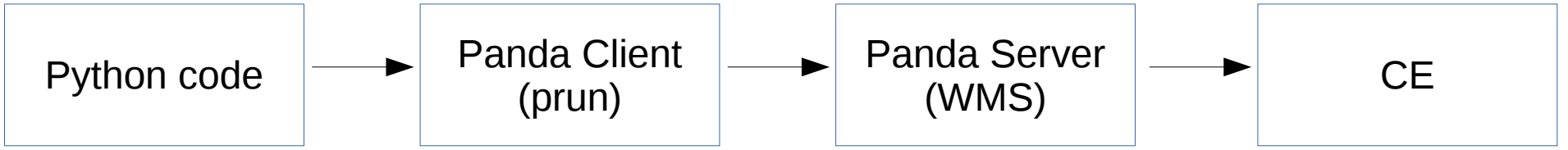
Submitting a prun job (--outDS <dataset_name> is an identifier of a PanDA job)
Please use your name (usually RUCIO_ACCOUNT name)
prun --outDS user.gkawamur.pruntest.$RANDOM --exec hello_world.py

```
INFO : gathering files under /home/gen/tmp/for_new_comer
INFO : upload source files
INFO : submit
INFO : succeeded. new jediTaskID=5107461
```

Submitting 5 prun jobs
prun --outDS user.gkawamur.pruntest.$RANDOM --exec hello_world.py --nJobs=5

How it works

- A Python code is serialized and is sent to the backends



What will happen?

- On PanDA web interface (bigpanda.cern.ch), find the jobs submitted

jobstatus (1)	finished (2)
minramcount (1)	1-2GB (1)
outputfiletype (2)	? (1) log (1)
priorityrange (2)	1000:1099 (1) 2000:2099 (1)
processingtype (1)	panda-client-0.5.72-jedi-athena (2)
prodsourcelabel (2)	panda (1) user (1)
produsername (1)	Gen Kawamura (2)
reqid (1)	94 (2)
specialhandling (1)	ddm:rucio (2)
transformation (2)	buildJob-00-00-03 (1) runAthena-00-00-12 (1)

Prodsys Jobs Handling

Job list Sort by PandaID, time since last state change, ascending mod time, descending mod time, priority, attemptnr, ascending duration, descending duration

PanDA ID Attempt#	Owner Group	Request Task ID	Transformation	Status	Created	Time to start d:h:m:s	Duration d:h:m:s	Mod	Cloud Site
3131853110 Attempt 1	Gen Kawamura	94 10262517	runAthena-00-00-12	finished	2016-12-19 14:51	0:0:10:54	0:0:01:14	12-19 15:09	DE ANALY_IEPSAS-Kosice online HC.Blacklist.set.online
	Job name: user.gkawamur.tutorial2016.12/.3131853110 #1								
	Datasets: **Out:** user.gkawamur.tutorial2016.12.log.112492285								
3131853105 Attempt 0	Gen Kawamura	94 10262517	buildJob-00-00-03	finished	2016-12-19 14:51	0:0:02:38	0:0:03:17	12-19 15:00	DE ANALY_IEPSAS-Kosice online HC.Blacklist.set.online
	Job name: user.gkawamur.tutorial2016.12/ #0								
	Datasets: **Out:** panda.1219145126.858464.lib._10262517								

Hello World pathena Grid Job

Hands-on exercise simple Athena job

- The xAOD Athena example (muon selection)

```
## xAOD Tutorial
$ cd xAOD_example

$ mkdir source build run
$ cd build
$ asetup AthAnalysis,21.2.90,here
$ mv CMakeLists.txt ../source/
$ cmake $TestArea/../source
$ source $TestArea/*/setup.sh

## Copying example files
$ cd $TestArea/../run/
$ cp -v ../myJobOptions.py .
```

- Or, simply

```
## xAOD Tutorial
$ cd xAOD_example
$ source set_athena_env.sh
```

Hands-on exercise simple Athena job

- The xAOD Athena example
  - The input files are defined by myJobOptions.py

```
import glob
fileInputs =
glob.glob('valid2.117050.PowhegPythia_P2011C_ttbar.digit.AOD.e2657_s1933_s1964_r5534/*')
svcMgr.EventSelector.InputCollections = fileInputs
```

- **“valid2.117050.PowhegPythia_P2011C_ttbar.digit.AOD.e2657_s1933_s1964_r5534” is a Rucio dataset stored in Grid storages**

```
## Get a sample file
$ export RUCIO_ACCOUNT=your_rucio_account_here
$ rucio download --nrandom 1 "valid2.117050.PowhegPythia_P2011C_ttbar.digit.AOD.e2657_s1933_s1964_r5534"

## Test Athena with the job option file
$ athena myJobOptions.py
```

Hands-on exercise simple pathena job

- Athena job using PanDA client
  - Pathena job (JOB ID = user.gkawamur.test.$RANDOM)

```
## Loading PanDA client
$ lsetup panda

## For example, you can seamlessly run Athena code on Grid
$ pathena --inDS valid2.117050.PowhegPythia_P2011C_ttbar.digit.AOD.e2657_s1933_s1964_r5534 --outDS
user.gkawamur.test.$RANDOM --nFilesPerJob=1 myJobOptions.py


....
INFO : checking symbolic links
INFO : uploading source/jobO files
INFO : submit
INFO : succeeded. new jediTaskID=15301491
```

Hands-on exercise simple pathena job

- Athena job using PanDA client
  - Pathena job (JOB ID = user.gkawamur.test.$RANDOM)

```
## Loading PanDA client
$ lsetup panda

## For example, you can seamlessly run Athena code on Grid
$ pathena --inDS valid2.117050.PowhegPythia_P2011C_ttbar.digit.AOD.e2657_s1933_s1964_r5534 --outDS
user.gkawamur.test.$RANDOM --nFilesPerJob=1 myJobOptions.py


....
INFO : checking symbolic links
INFO : uploading source/jobO files
INFO : submit
INFO : succeeded. new jediTaskID=15301491
```

1 Grid subjob processes only
1 file in the input dataset

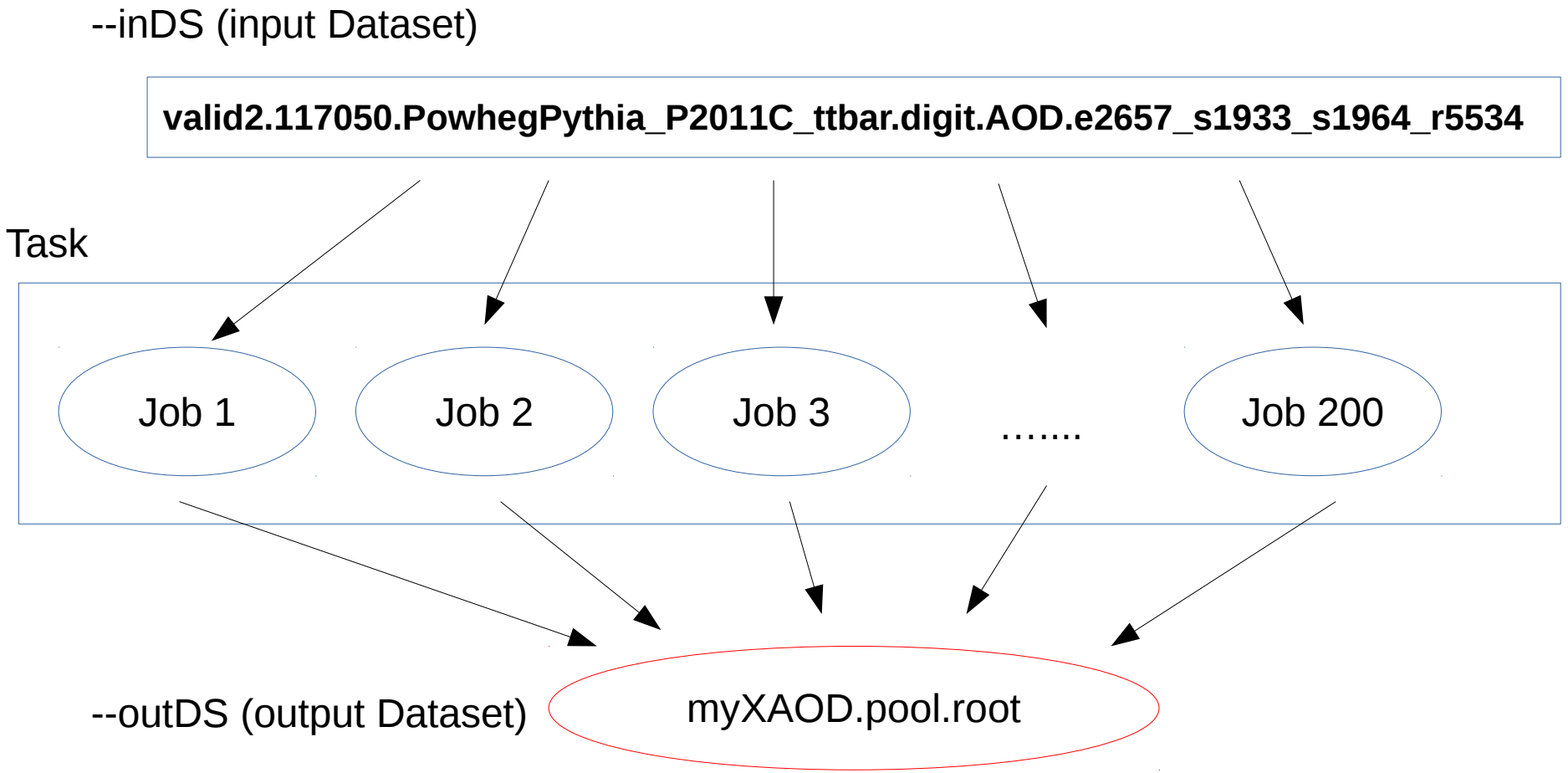
Input dataset containing xAOD files

Output dataset also used for job identifier

Final output dataset which will contain results

How it works

- Executing a task processing processing events per job (per file)



Hello World PyRoot

Hands-on exercise
PyRoot example

- **ATLAS xAOD EDM tutorial using pyRoot**
  - **https://twiki.cern.ch/twiki/bin/viewauth/AtlasComputing/SoftwareTutorialxAODEDM**
  - **We will run it on Grid later**

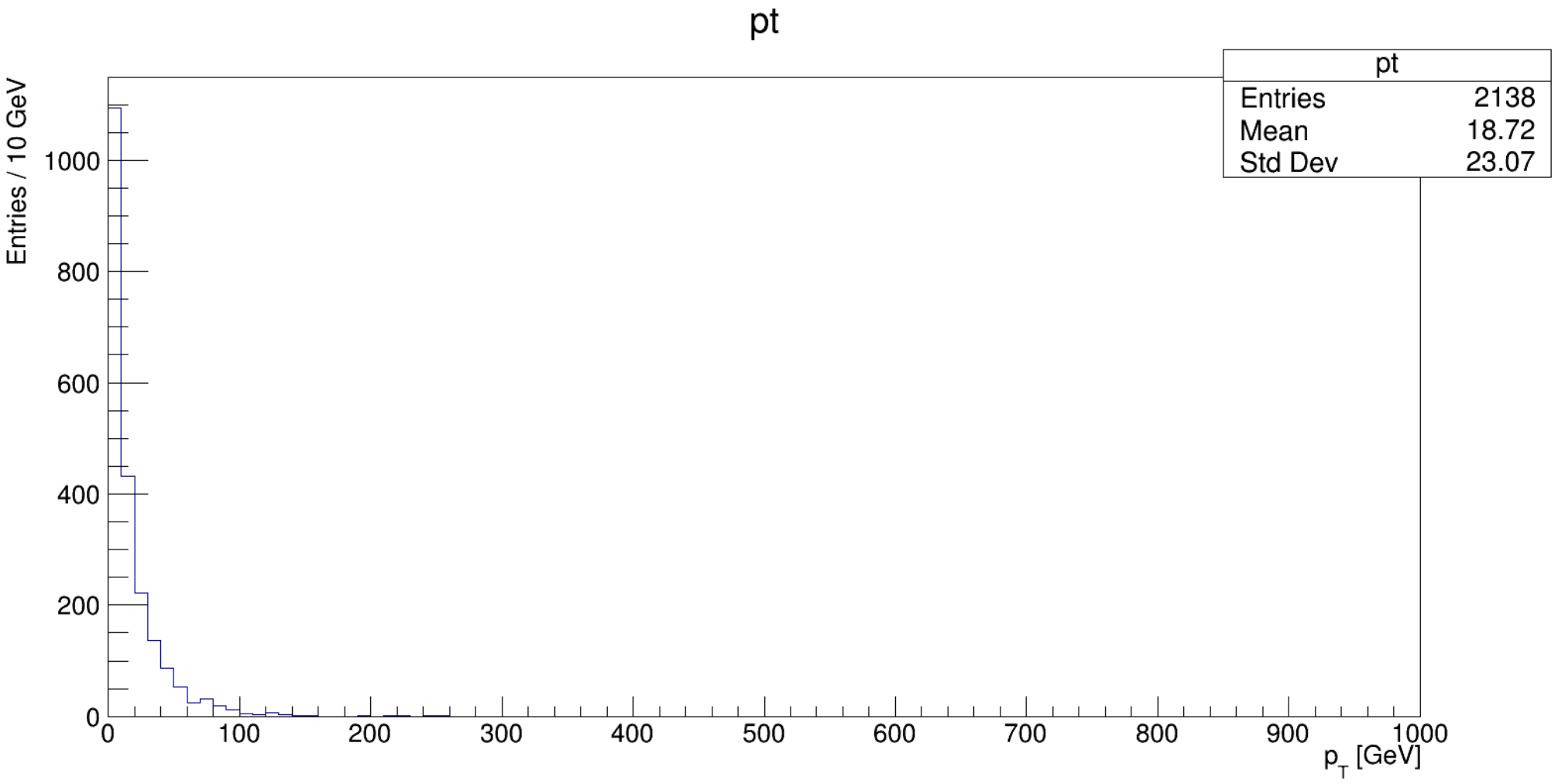
```
## Making PyRoot environments (using RootCore)
$ cd pyroot
$ source pyroot_env.sh

## Getting a sample file (set your rucio account)
$ export RUCIO_ACCOUNT=your_rucio_account_here
$ rucio download --nrandom 1 "valid2.117050.PowhegPythia_P2011C_ttbar.digit.AOD.e2657_s1933_s1964_r5534"
$ ls valid2.117050.PowhegPythia_P2011C_ttbar.digit.AOD.e2657_s1933_s1964_r5534/* > input.txt

## Extracting and counting electron energy
$ less xAOD_electron_hist_example.py
$ ./xAOD_electron_hist_example.py -i input.txt -o hist.root

## Plotting electron Pt distribution (on X-window TBrowser).
$ root hist.root
root [1] TBrowser t
```

Plot of electron Pt distribution



How it works - 1

- Just looping entries (events) in a Root tree and counting electron Pt in histogram object

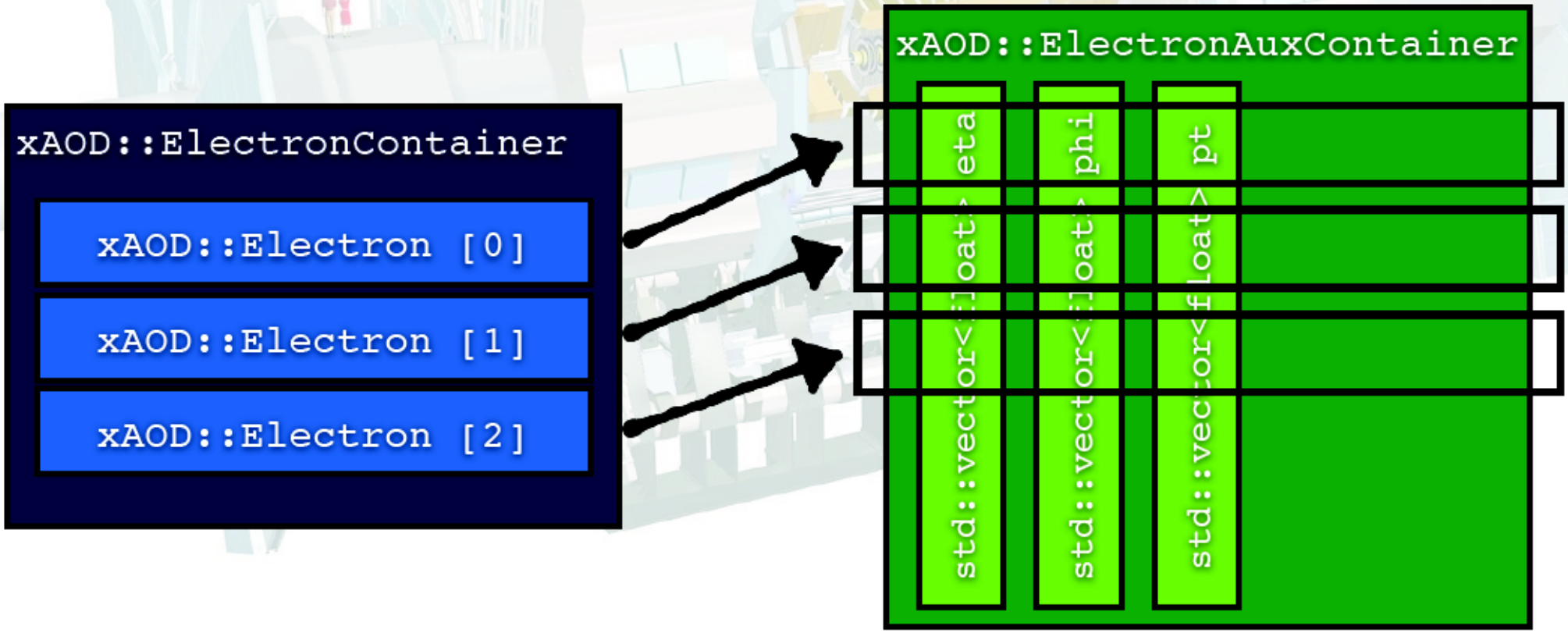
*** xAOD_electron_hist_example.py**

```
# Make the "transient tree":
t = ROOT.xAOD.MakeTransientTree( f, treeName)

print( "Number of input events: %s" % t.GetEntries() )
for entry in xrange( t.GetEntries() ):
   t.GetEntry( entry )
   print( "Processing run #%i, event #%i" % ( t.EventInfo.runNumber(), t.EventInfo.eventNumber() ) )
   print( "Number of electrons: %i" % len( t.ElectronCollection ) )
   # loop over electron collection
   for el in t.ElectronCollection:
      pthist.Fill(el.pt()/1000.)
      pass # end for loop over electron collection
   pass # end loop over entries
f.Close()
pass
```

How it works - 2

- Is technically quite smart code…
  - Provides an "array of structs" interface to data held as "struct of arrays" in memory
  - This "struct of arrays" layout allows us to write files that can be browsed similar to D3PD files



Hello World PyRoot Grid Job

Hands-on exercise
PyRoot with Grid

- First “Hello world” PyRoot job by PanDA client

```
## Making PyRoot environments
 $ inDS="valid2.117050.PowhegPythia_P2011C_ttbar.digit.AOD.e2657_s1933_s1964_r5534"
 $ outDS="user.gkawamur.DStutorial.pyroot.xAOD.v0.1_$$"
 $ infile="input.txt"
 $ outfile="hist.root"
 $ prun --useRootCore --inDS=$inDS --forceStaged \
--outDS=$outDS --outputs=$outfile --nFiles=100 --nFilesPerJob=1 \
--exec="echo %IN > $infile; xAOD_electron_hist_example.py -i $infile -o $outfile"
```

```
## or, use a simple submitter script
 $ bash submit.sh
```

How it works

- Executing a task processing processing events per job (per file)

--inDS (input Dataset)

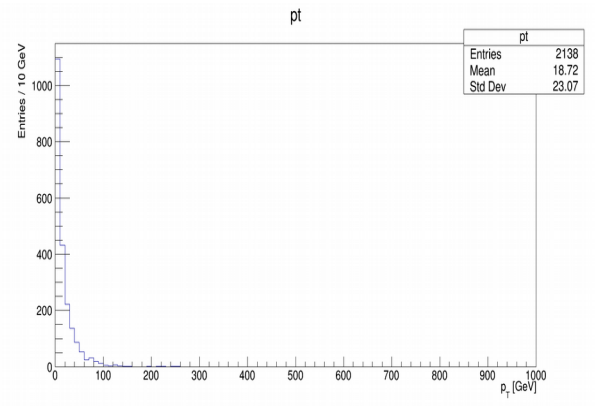
valid2.117050.PowhegPythia_P2011C_ttbar.digit.AOD.e2657_s1933_s1964_r5534

Task

Job 1 | Job 2 | Job 3 | ..... | Job 100

--outDS (output Dataset)

hist.root



Rucio (ATLAS data management system)

Setup Rucio

- Initializing Rucio client

Loading Rucio client
lsetup rucio

```
**********************************************************************
Requested:  rucio ...
 Setting up emi 3.17.1-1_v2.sl6 ...
  Skipping: grid middleware already setup (from UI)
 Setting up rucio 1.7.3 ...
Info: Setting compatibility to slc6
Info: Set RUCIO_AUTH_TYPE to x509_proxy
Info: Set RUCIO_ACCOUNT to gkawamur
>>>>>>>>>>>>>>>>>>>>>>>>> Information for user <<<<<<<<<<<<<<<<<<<<<<<<
 emi:
   Your proxy has 95h:54m:0s remaining
**********************************************************************
```

Rucio command
rucio

```
usage: rucio [-h] [--version] [--verbose] [-H ADDRESS] [--auth-host ADDRESS]
             [-a ACCOUNT] [-S AUTH_STRATEGY] [-T TIMEOUT] [--robot]
             [--user-agent USER_AGENT] [-u USERNAME] [-pwd PASSWORD]
             [--certificate CERTIFICATE] [--ca-certificate CA_CERTIFICATE]
```

In Rucio, check which account you use
rucio whoami

Hands-on exercise
list scopes

- Each user account has one default scope, e.g., user.rucio_user for account rucio_user

Listing sopes
rucio list-scopes

user.gkawamur
...
group.phys-sm
group.phys-susy
...
data15_13TeV
data15_1beam
data15_900GeV
...
mc15_5TeV
mc15_8TeV
mc15_900GeV

Your user scope
rucio list-scopes | grep user.gkawamur

user.gkawamur

Hands-on exercise
list DIDs

- A data Identifier is found by name and scope

To list all DIDs within a scope

rucio list-dids user.gkawamur:*

```
+-------------------------------------------------------------+--------------+
| SCOPE:NAME                                                  | [DID TYPE]   |
|-------------------------------------------------------------+--------------|
| user.gkawamur:user.gkawamur.pruntest_7168.log               | CONTAINER    |
| user.gkawamur:user.gkawamur.tut.helloworldathena.test.log   | CONTAINER    |
| user.gkawamur:user.gkawamur.20160323144306.1.log            | CONTAINER    |
| user.gkawamur:user.gkawamur.tutorial.grid.deriv.test1.log   | CONTAINER    |
| user.gkawamur:user.gkawamur.pruntest_14222.log              | CONTAINER    |
```

Using filters (search)

rucio list-dids data16_13TeV:* --filter type=DATASET,datatype=AOD

```
+-------------------------------------------------------------------------------------------------------------------------+--------------+
| SCOPE:NAME                                                                                                              | [DID TYPE]   |
|-------------------------------------------------------------------------------------------------------------------------+--------------|
| data16_13TeV:data16_13TeV.00293572.physics_CosmicCalo.merge.AOD.x387_m1554                                              | DATASET      |
| data16_13TeV:data16_13TeV.00297447.physics_Standby.merge.AOD.x416_m1583                                                 | DATASET      |
| data16_13TeV:data16_13TeV.00297041.physics_Main.merge.AOD.f686_m1583                                                    | DATASET      |
| data16_13TeV:data16_13TeV.00297041.physics_CosmicCalo.merge.AOD.f686_m1583                                              | DATASET      |
| data16_13TeV:data16_13TeV.00297041.physics_ZeroBias.merge.AOD.f686_m1583                                                | DATASET      |
```

Hands-on exercise list contents

List file contents of dataset or container

rucio list-files data16_13TeV:data16_13TeV.00303819.physics_Main.merge.AOD.f716_m1620

```
+-----------------------------------------------------------------------------------------+-------------------------------------+-------------+------------+----------+
| SCOPE:NAME                                                                              | GUID                                | ADLER32     | FILESIZE   |   EVENTS |
|-----------------------------------------------------------------------------------------+-------------------------------------+-------------+------------+----------|
| data16_13TeV:data16_13TeV.00303819.physics_Main.merge.AOD.f716_m1620._lb0004._0001.1 | 98325FCF-9E4E-B442-A4B2-2A64E271C697 | ad:130a9a7c | 182.9 MB   |      887 |
| data16_13TeV:data16_13TeV.00303819.physics_Main.merge.AOD.f716_m1620._lb0005._0001.1 | 85303413-544F-EC4A-B494-494091163A92 | ad:d78501da | 220.4 MB   |      836 |
| data16_13TeV:data16_13TeV.00303819.physics_Main.merge.AOD.f716_m1620._lb0006._0001.1 | 970BCEFD-CF3D-C541-8507-471F54E3167A | ad:abc20882 | 2.7 GB     |     9097 |
```

List file contents of dataset or container

rucio list-content data16_13TeV:data16_13TeV.00303819.physics_Main.merge.AOD.f716_m1620

```
+-----------------------------------------------------------------------------------------+--------------+
| SCOPE:NAME                                                                              | [DID TYPE]   |
|-----------------------------------------------------------------------------------------+--------------|
| data16_13TeV:data16_13TeV.00303819.physics_Main.merge.AOD.f716_m1620._lb0004._0001.1 | FILE         |
| data16_13TeV:data16_13TeV.00303819.physics_Main.merge.AOD.f716_m1620._lb0005._0001.1 | FILE         |
| data16_13TeV:data16_13TeV.00303819.physics_Main.merge.AOD.f716_m1620._lb0006._0001.1 | FILE         |
```

List all contents recursively

rucio list-dids --recursive data16_13TeV:data16_13TeV.periodA3.physics_Main.PhysCont.AOD.t0pro20_v01

```
+-----------------------------------------------------------------------------------------+--------------+
| SCOPE:NAME                                                                              | [DID TYPE]   |
|-----------------------------------------------------------------------------------------+--------------|
| data16_13TeV:data16_13TeV.00297730.physics_Main.merge.AOD.f694_m1583                    | DATASET      |
| data16_13TeV:data16_13TeV.00297730.physics_Main.merge.AOD.f694_m1583._lb0108._0001.1 | FILE         |
| data16_13TeV:data16_13TeV.00297730.physics_Main.merge.AOD.f694_m1583._lb0108._0002.1 | FILE         |
| data16_13TeV:data16_13TeV.00297730.physics_Main.merge.AOD.f694_m1583._lb0108._0003.1 | FILE         |
| data16_13TeV:data16_13TeV.00297730.physics_Main.merge.AOD.f694_m1583._lb0109._0001.1 | FILE         |
| data16_13TeV:data16_13TeV.periodA3.physics_Main.PhysCont.AOD.t0pro20_v01             | CONTAINER    |
```

Hands-on exercise
show metadata

Showing metadata of a dataset
rucio get-metadata data16_13TeV:data16_13TeV.00303819.physics_Main.merge.AOD.f716_m1620

purge_replicas: None
campaign: None
is_new: None
is_open: False
closed_at: 2016-07-18 07:39:44
deleted_at: None
availability: AVAILABLE
eol_at: None
guid: None
panda_id: None
provenance: None
accessed_at: 2016-09-19 22:00:22
version: f716_m1620
scope: data16_13TeV
hidden: False
md5: None
events: 25475721
adler32: None
complete: None
lumiblocknr: None
monotonic: False
updated_at: 2016-09-19 22:00:42
obsolete: False
transient: None
did_type: DATASET
suppressed: True
expired_at: None
stream_name: physics_Main
account: tzero
run_number: 303819
name: data16_13TeV.00303819.physics_Main.merge.AOD.f716_m1620
task_id: None
datatype: AOD
created_at: 2016-07-17 04:17:53
bytes: 6462688464808
project: data16_13TeV
length: 2595
prod_step: merge
phys_group: None

Hands-on exercise list replicas

Listing dataset replicas

rucio list-dataset-replicas data16_13TeV:data16_13TeV.00303819.physics_Main.merge.AOD.f716_m1620

```
DATASET: data16_13TeV:data16_13TeV.00303819.physics_Main.merge.AOD.f716_m1620
+--------------------+---------+---------+
| RSE                |   FOUND |   TOTAL |
|--------------------+---------+---------|
| CERN-PROD_TZDISK   |    2567 |    2595 |
| GRIF-LPNHE_DATADISK |    2595 |    2595 |
| IN2P3-CC_DATADISK  |    2595 |    2595 |
| CERN-PROD_DERIVED  |    2595 |    2595 |
+--------------------+---------+---------+
```

Listing file replicas

rucio list-file-replicas data16_13TeV:data16_13TeV.00303819.physics_Main.merge.AOD.f716_m1620

```
+-------------+-------------------------------------------------------------------------+------------+----------
+-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------+
| SCOPE       | NAME                                                                    | FILESIZE   | ADLER32   | RSE: REPLICA

                                                |
|-------------+-------------------------------------------------------------------------+------------+----------
+-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------|
| data16_13TeV | data16_13TeV.00303819.physics_Main.merge.AOD.f716_m1620._lb0004._0001.1 | 182.9 MB   |
130a9a7c  | GRIF-LPNHE_DATADISK: srm://lpnse1.in2p3.fr:8446/srm/managerv2?
SFN=/dpm/in2p3.fr/home/atlas/atlasdatadisk/rucio/data16_13TeV/39/bc/data16_13TeV.00303819.physics_Main.merge.A
OD.f716_m1620._lb0004._0001.1                                                                                                                                                  |
| data16_13TeV | data16_13TeV.00303819.physics_Main.merge.AOD.f716_m1620._lb0004._0001.1 | 182.9 MB   |
130a9a7c  | TAIWAN-LCG2_DATADISK: https://f-
dpm000.grid.sinica.edu.tw:443/dpm/grid.sinica.edu.tw/home/atlas/atlasdatadisk/rucio/data16_13TeV/39/bc/data16_13TeV
.00303819.physics_Main.merge.AOD.f716_m1620._lb0004._0001.1
```

Hands-on exercise download data

- Make a local copy from Rucio RSEs

Downloading a dataset to local disk (select an output dataset of your test job)
rucio download user.gkawamur:user.gkawamura.test1

2019-02-01 20:26:57,534 INFO [Starting download for user.gkawamur:user.gkawamura.test1 with 0 files]

Download summary

DID user.gkawamur:user.gkawamura.test1
Total files : 0
Downloaded files : 0
Files already found locally : 0
Files that cannot be downloaded : 0

Hands-on exercise request a replica

- Request a replica to a RSE space
  - The same action can be performed by R2D2

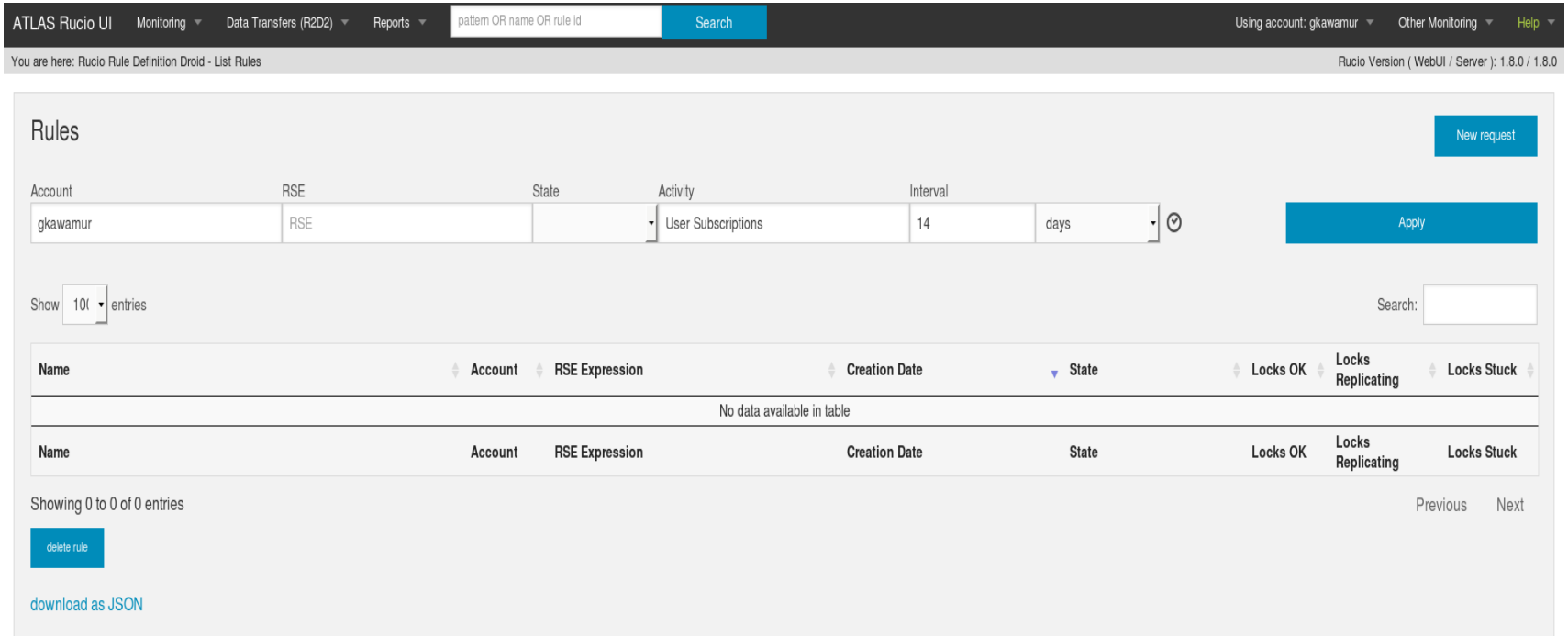
Request a replication
rucio add-rule user.gkawamur:user.gkawamura.test1 --grouping DATASET 1 "DESY-HH_SCRATCHDISK"
…
<Transfer ID>

Check the transfer status (or by RucioUI https://rucio-ui.cern.ch/list_rules)
rucio rule-info <Transfer ID>

RucioUI (WebUI)
- Request replicas -

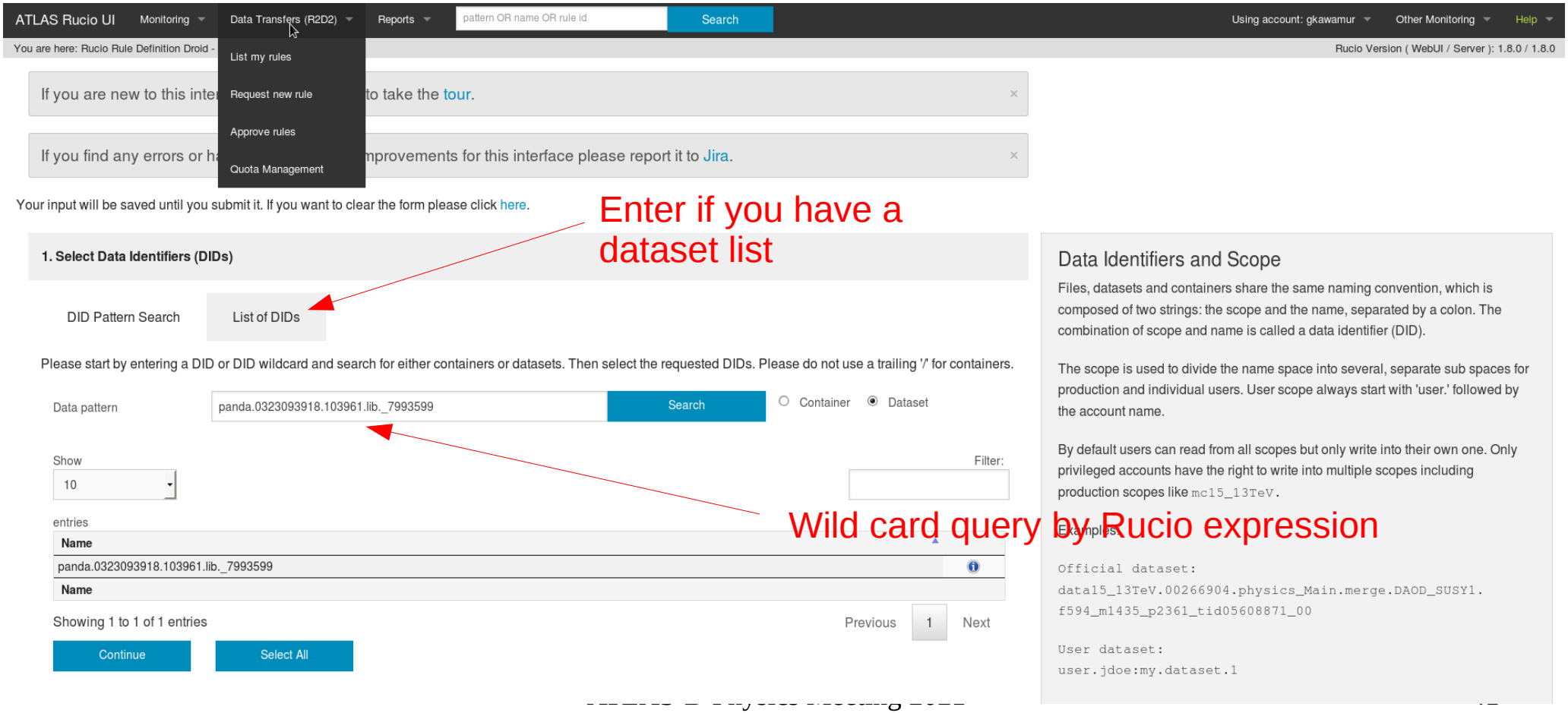
RucioUI

- Move data between RSEs, use Rucio replication rules
- Such request can be generated by CLI and UI
- The UI tool is *the Rucio Rule Definition Droid (R2D2)* *https://rucio-ui.cern.ch/r2d2*
- Basic need: X509 certificate in your browse



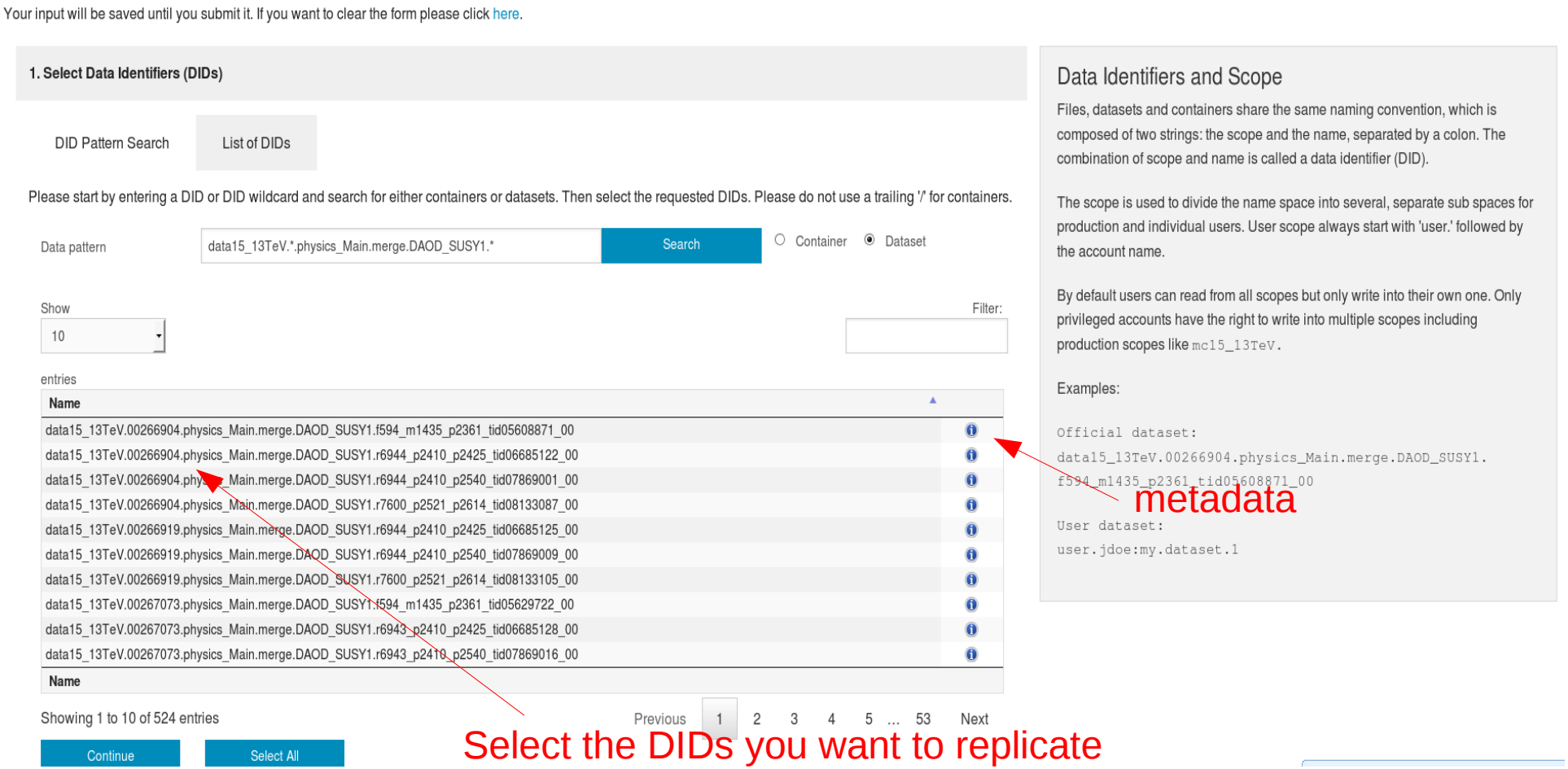
RucioUI: Select DIDs - 1

- Data transfers (R2D2) → Request new rule



RucioUI: Select DIDs - 2

- E.g. a wild card pattern
  - “data15_13TeV.*.physics_Main.merge.DAOD_SUSY1.*”



RucioUI: Select DIDs - 3

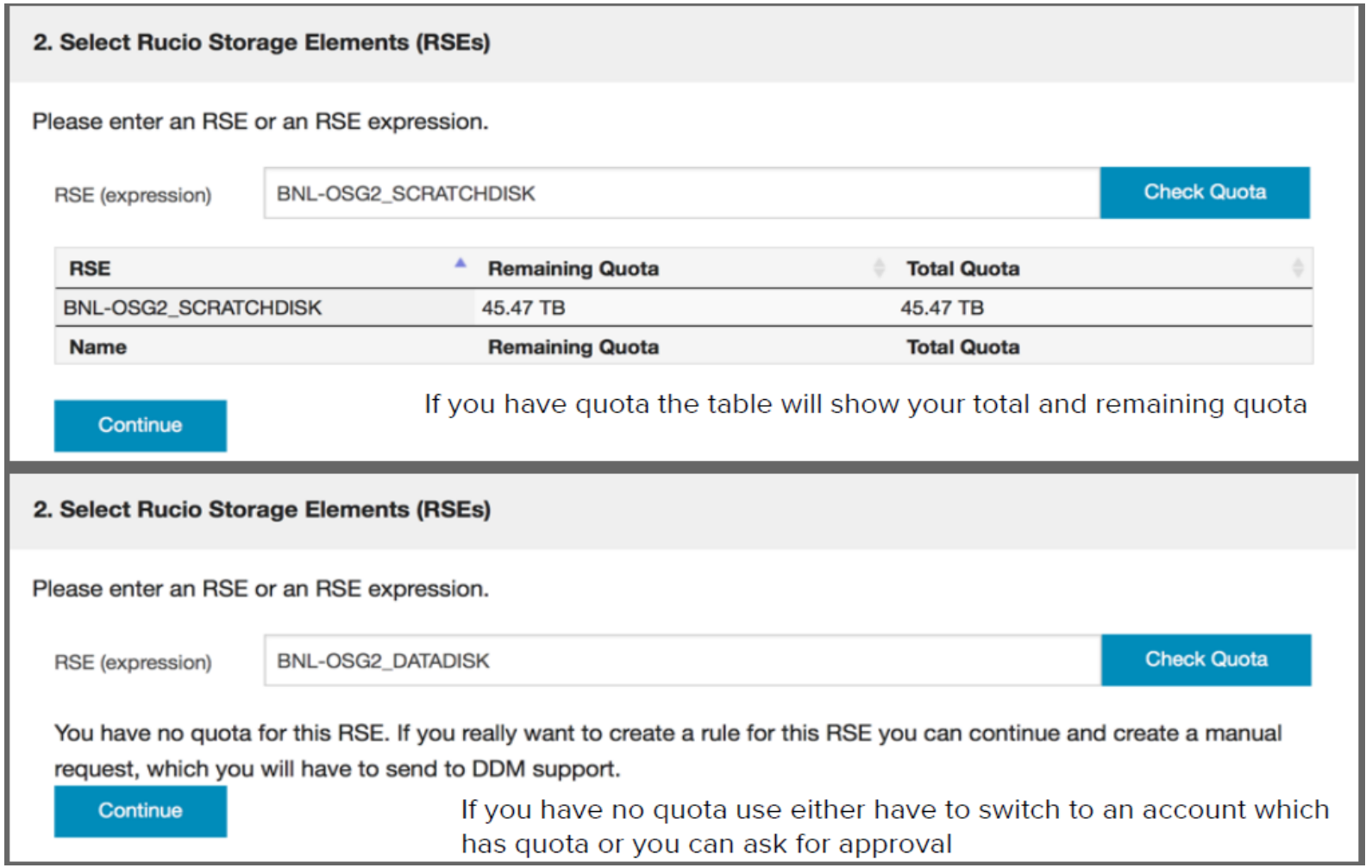
- Metadata

data15_13TeV.00266904.physics_Main.merge.DAOD_SUSY1.f594_m1435_p2361_tid05608871_00

accessed_at	Mon, 14 Sep 2015 04:34:53 UTC
account	panda
availability	AVAILABLE
closed_at	Mon, 24 Aug 2015 14:33:30 UTC
created_at	Sun, 07 Jun 2015 10:09:59 UTC
datatype	DAOD_SUSY1
did_type	DATASET
events	10593
filesize	237.94 MB
hidden	false
is_open	false
length	7
monotonic	false
name	data15_13TeV.00266904.physics_Main.merge.DAOD_SUSY1.f594_m1435_p2361_tid05608871_00
obsolete	false
prod_step	merge
project	data15_13TeV
run_number	266904
scope	data15_13TeV
stream_name	physics_Main
suppressed	true
task_id	5608871
updated_at	Mon, 14 Sep 2015 06:35:06 UTC
version	f594_m1435_p2361

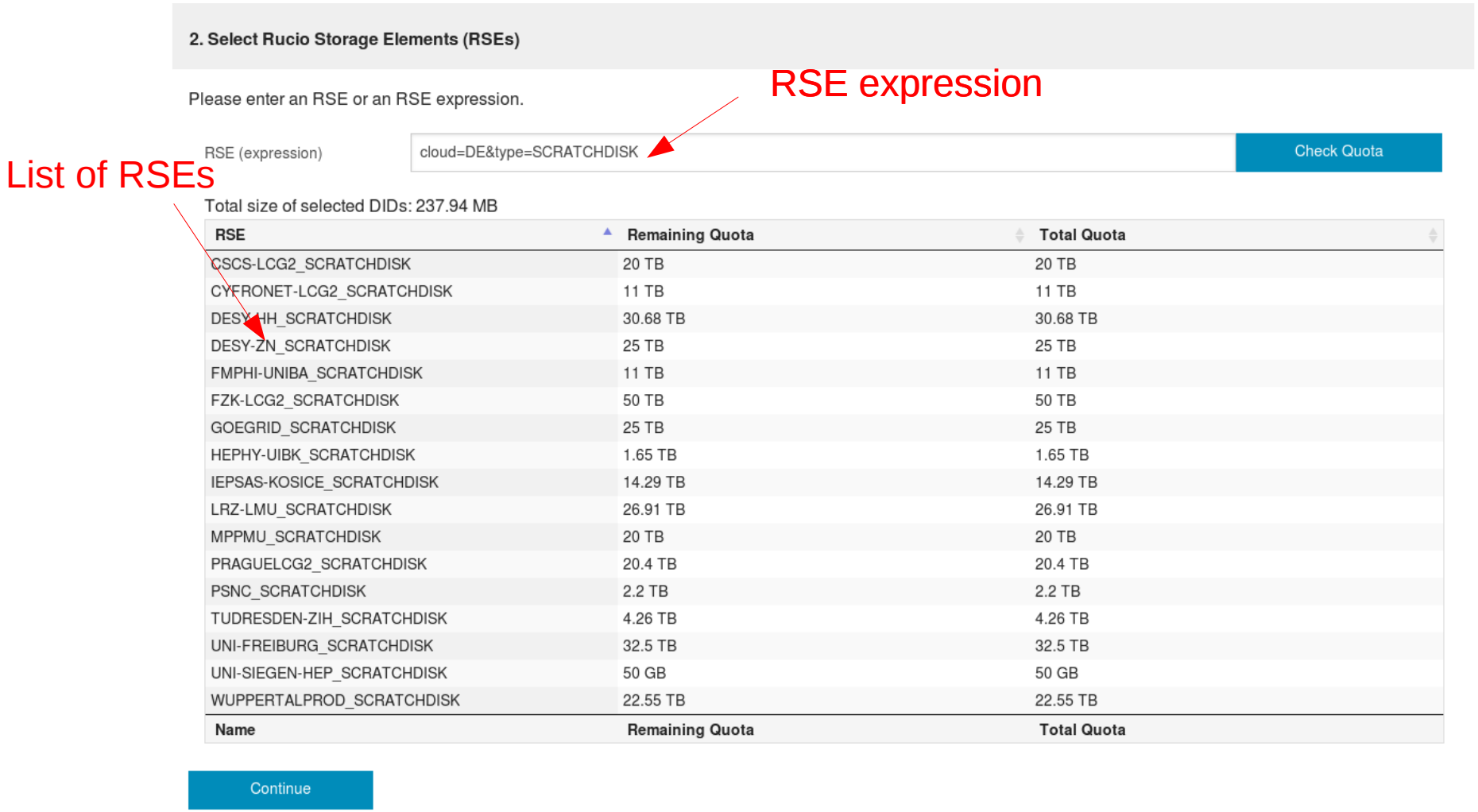
RucioUI: select a RSE - 1

- Find your destination space

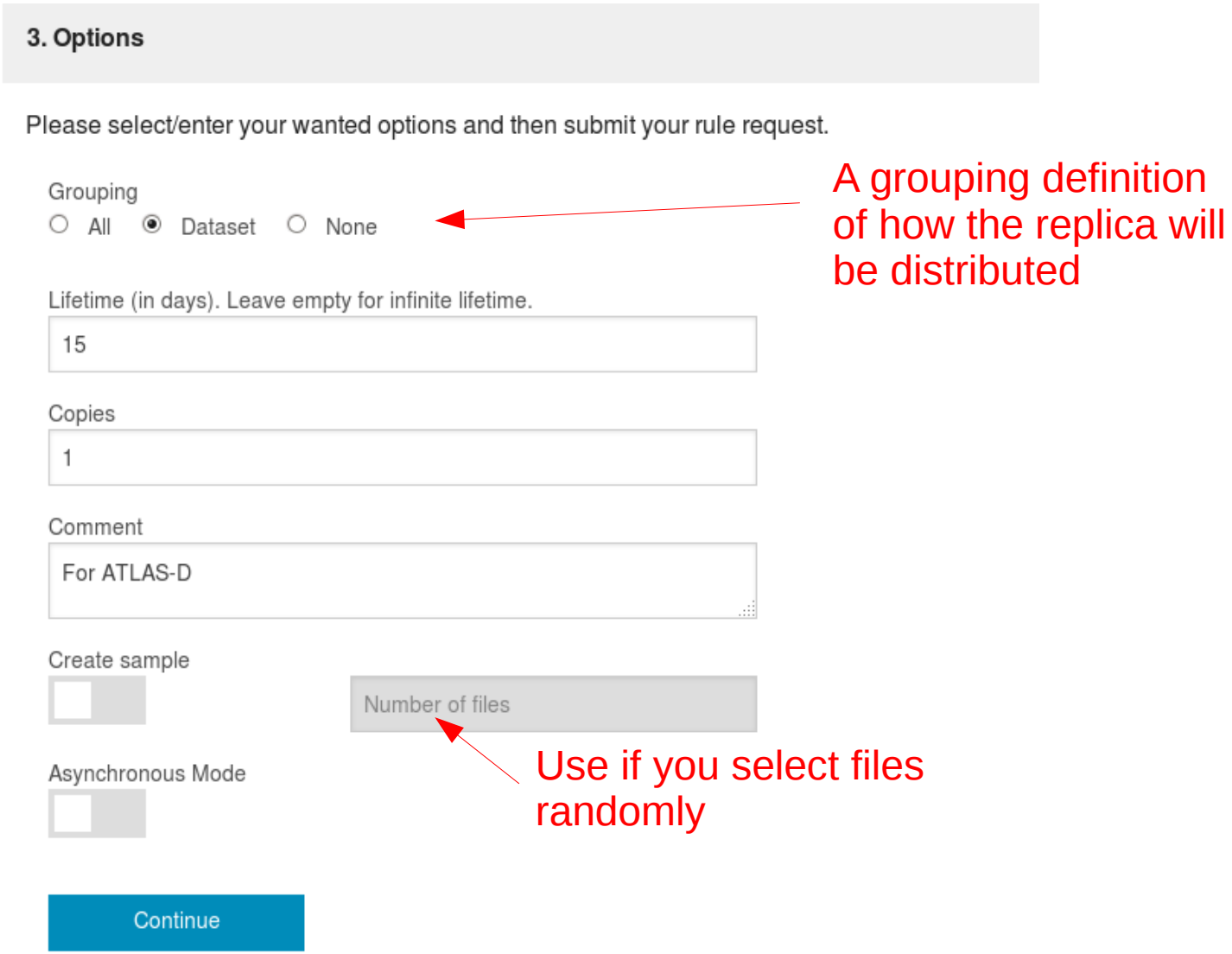


RucioUI: select a RSE - 2

- SCRATCHDISKs in Germany
  - cloud=DE&type=SCRATCHDISK



RucioUI: select options



Rucio UI: summary

- Before submission check rules

4. Summary

This request will create rules for the following DIDs:

DID	Copies	Files	Size	Requested Size
data15_13TeV:data15_13TeV.00266904.physics_Main.merge.DAOD_SUSY1.f594_m1435_p2361_tid05608871_00	1	7	237.94 MB	237.94 MB
data15_13TeV:data15_13TeV.00266904.physics_Main.merge.DAOD_SUSY1.r6944_p2410_p2425_tid06685122_00	1	14	2.96 GB	2.96 GB
data15_13TeV:data15_13TeV.00266904.physics_Main.merge.DAOD_SUSY1.r6944_p2410_p2540_tid07869001_00	1	5	3.4 GB	3.4 GB
data15_13TeV:data15_13TeV.00266904.physics_Main.merge.DAOD_SUSY1.r7600_p2521_p2614_tid08133087_00	1	10	2.62 GB	2.62 GB
Total	**4**	**36**	**8.59 GB**	**8.59 GB**

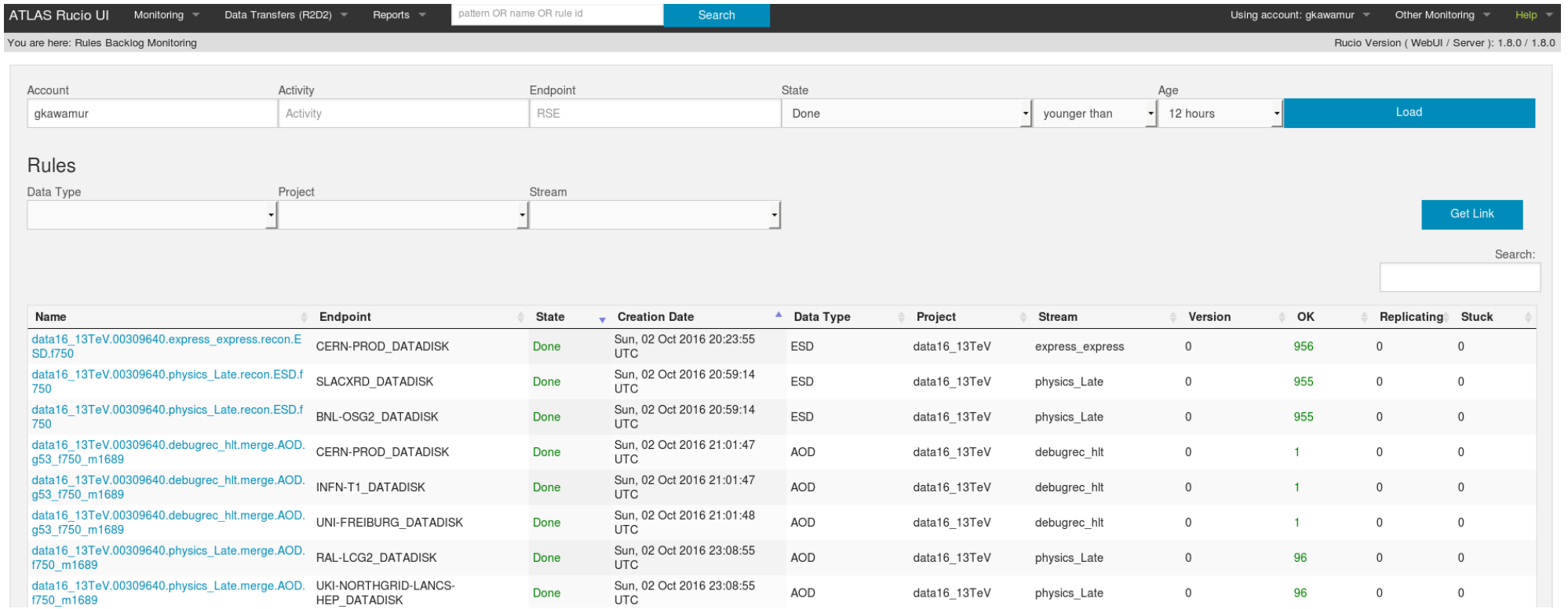
Check quota limit carefully!

The rules will replicate to one of the following RSEs:

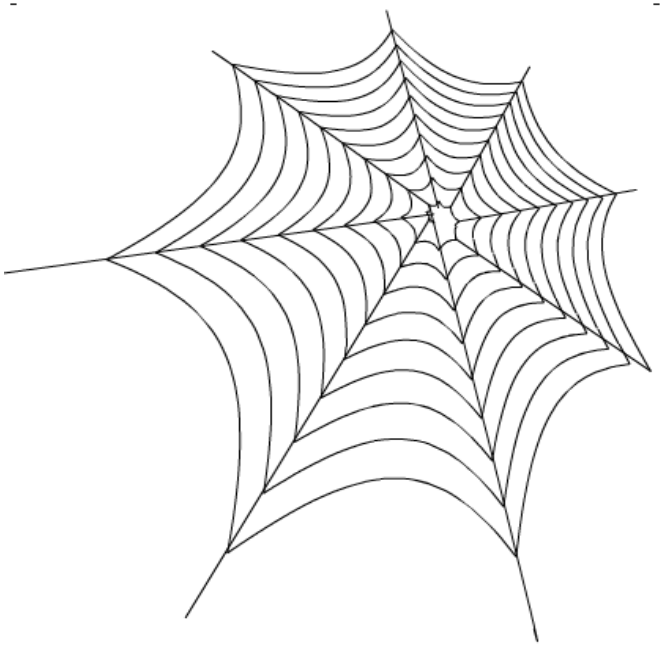
RSE	Remaining Quota	Total Quota
CSCS-LCG2_SCRATCHDISK	20 TB	20 TB
CYFRONET-LCG2_SCRATCHDISK	11 TB	11 TB
DESY-HH_SCRATCHDISK	30.68 TB	30.68 TB
DESY-ZN_SCRATCHDISK	25 TB	25 TB
FMPHI-UNIBA_SCRATCHDISK	11 TB	11 TB
FZK-LCG2_SCRATCHDISK	50 TB	50 TB
GOEGRID_SCRATCHDISK	25 TB	25 TB
HEPHY-UIBK_SCRATCHDISK	1.65 TB	1.65 TB
IEPSAS-KOSICE_SCRATCHDISK	14.29 TB	14.29 TB
LRZ-LMU_SCRATCHDISK	26.91 TB	26.91 TB
MPPMU_SCRATCHDISK	20 TB	20 TB
PRAGUELCG2_SCRATCHDISK	20.4 TB	20.4 TB
PSNC_SCRATCHDISK	2.2 TB	2.2 TB
TUDRESDEN-ZIH_SCRATCHDISK	4.26 TB	4.26 TB

Rucio UI: Transfer status

- Go to top page and see your transfer activity
  - *https://rucio-ui.cern.ch/r2d2*



ATLAS Metadata Interface (AMI)
If no time, please skip this part



Hands-on exercise pyAMI Interface

- AMI CLI interface

```
##  Loading the pyAMI environment
$ lsetup pyami

##  Search data of 2016 and period A1
$ ami list datasets data16_13TeV%periodA1.%
    data16_13TeV.periodA1.physics_Main.PhysCont.AOD.t0pro20_v01
    data16_13TeV.periodA1.physics_Main.PhysCont.DAOD_STDM2.grp16_v01_p2623
    data16_13TeV.periodA1.physics_Main.PhysCont.DAOD_STDM4.grp16_v01_p2623
    data16_13TeV.periodA1.physics_Main.PhysCont.DAOD_STDM5.grp16_v01_p2623
    data16_13TeV.periodA1.physics_Main.PhysCont.DAOD_STDM7.grp16_v01_p2623
```

Hands-on exercise
check metadata by pyAMI

Show metadata of a dataset

```
$ ami show dataset info data16_13TeV.00284285.physics_Main.merge.AOD.f662_m1453_r8067_p2645
   logicalDatasetName : data16_13TeV.00284285.physics_Main.merge.AOD.f662_m1453_r8067_p2645
   nFiles : 0
   totalEvents : 0
   totalSize : NULL
   runNumber : 284285
   period : J6
   prodsysStatus : NO EVENTS YET
   dataType : AOD
   beamType : NULL
   conditionsTag : NULL
   geometryVersion : NULL
   streamName : physics_Main
   version : f662_m1453_r8067_p2645
   lastModified : 2016-06-09 18:35:05
   amiStatus : VALID
   created : 2016-06-09 18:35:04
   inContainer : 0
   added_comment : NULL
   keyword : NULL
   prodsysIdentifier_0: 8650873
   taskStatus_0 : UNKNOWN:METADATA ERROR
   TIDState_0 : added
   task_lastModified_0: 2016-06-10 09:24:25
```

Hands-on exercise
check metadata by pyAMI

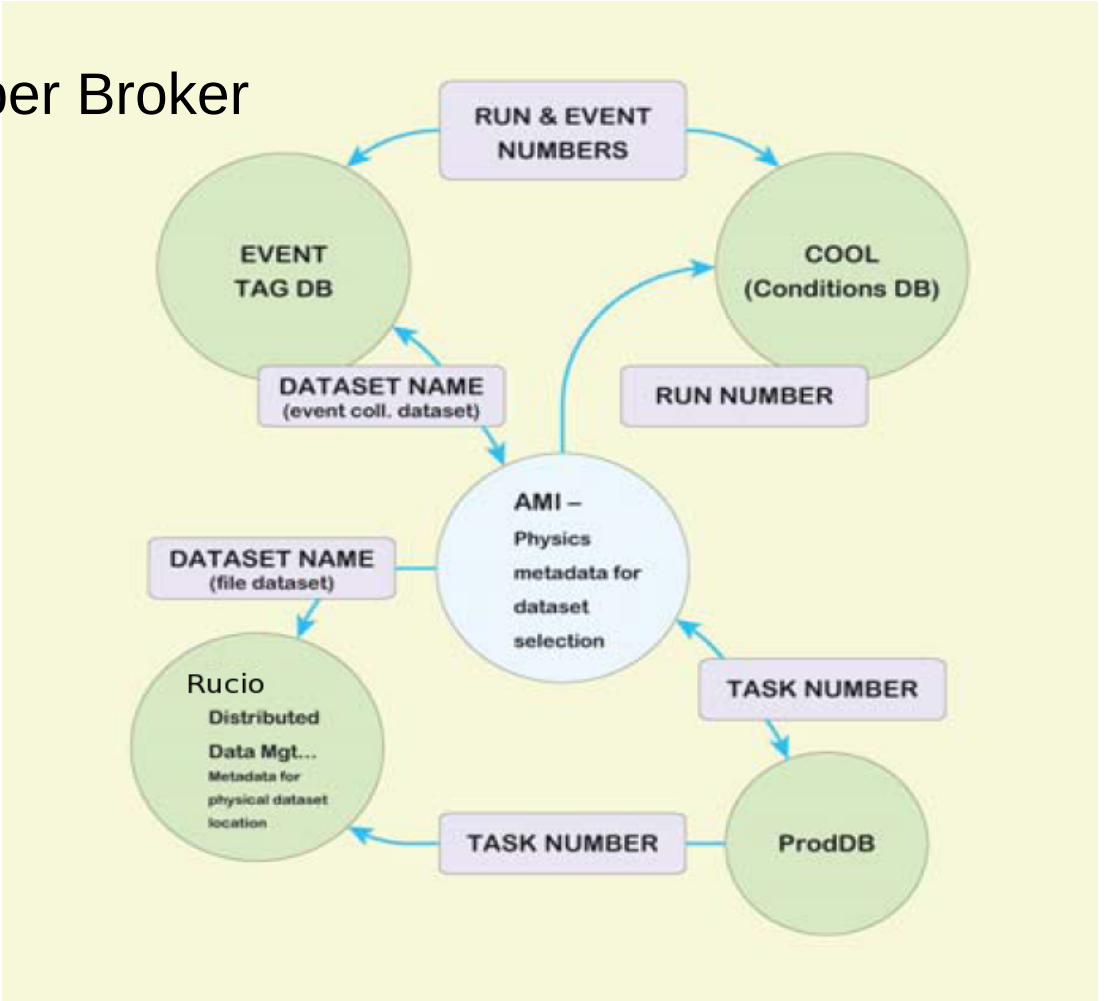
Show RAWs

$ ami show dataset prov data16_13TeV.00284285.physics_Main.merge.AOD.f662_m1453_r8067_p2645

...

How it works

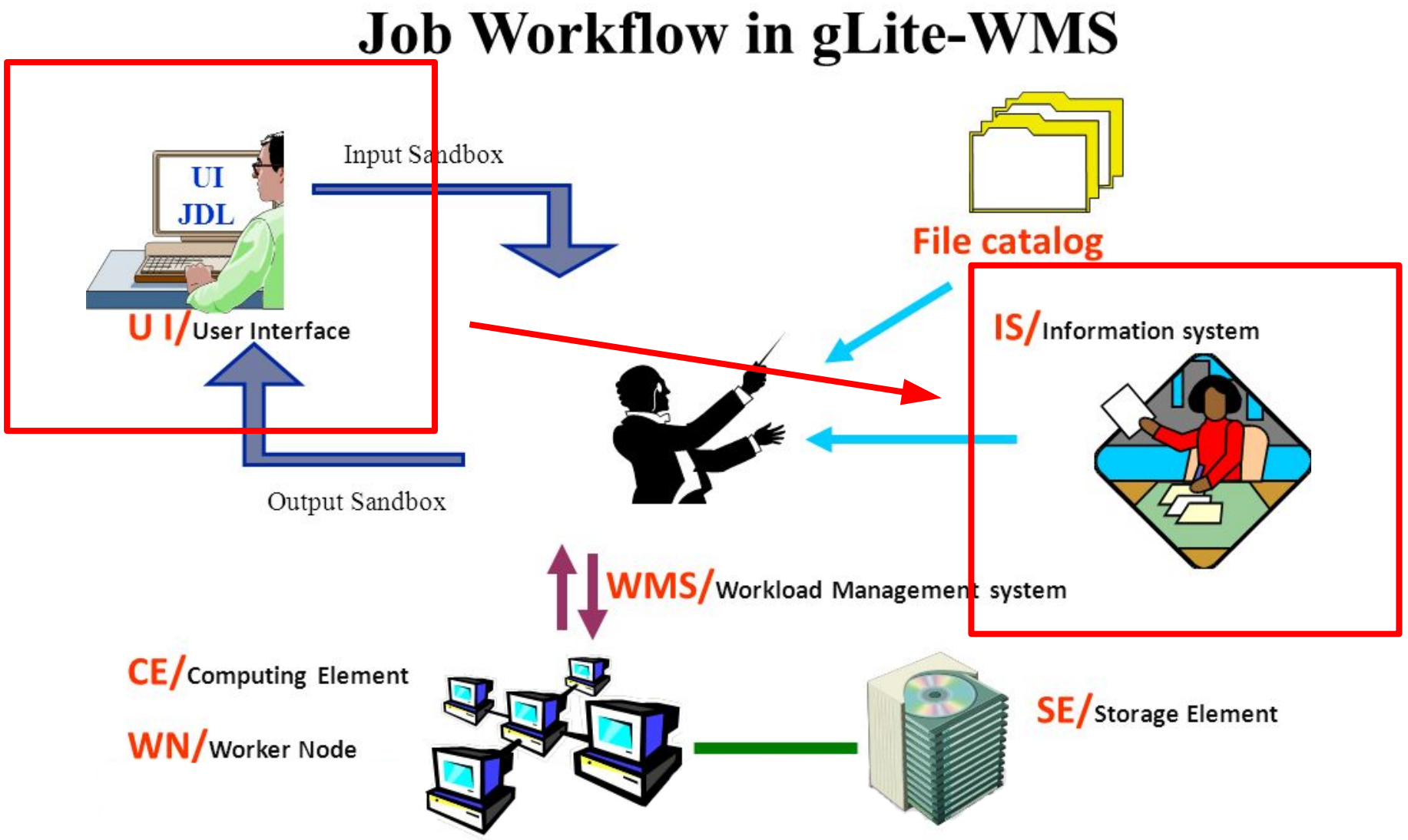
- Applications
  - The Monte-Carlo Dataset Number Broker
  - The ATLAS Metadata directory
  - Tag collector
- ProdDB
  - For Monte-Carlo simulation



S. Albrand, T. Doherty, J. Fulachier, F. Lambert. The ATLAS Metadata Interface. Interna-tional Conference on Computing in High Energy and Nuclear Physics (CHEP-07), Sep 2007, Victoria, Canada. IOP Publishing, 120, pp.072003, 2008, <10.1088/1742-6596/120/7/072003>. <in2p3-00192624>

Extra: Information Provider (of gLite)

Information Provider



Hands-on exercise Information Provider - 1

- Searching resources

EMI (European Middleware Initiative) LCG (LHC Computing Grid) tools
$ setupCVMFS

See usage
$ lcg-infosites

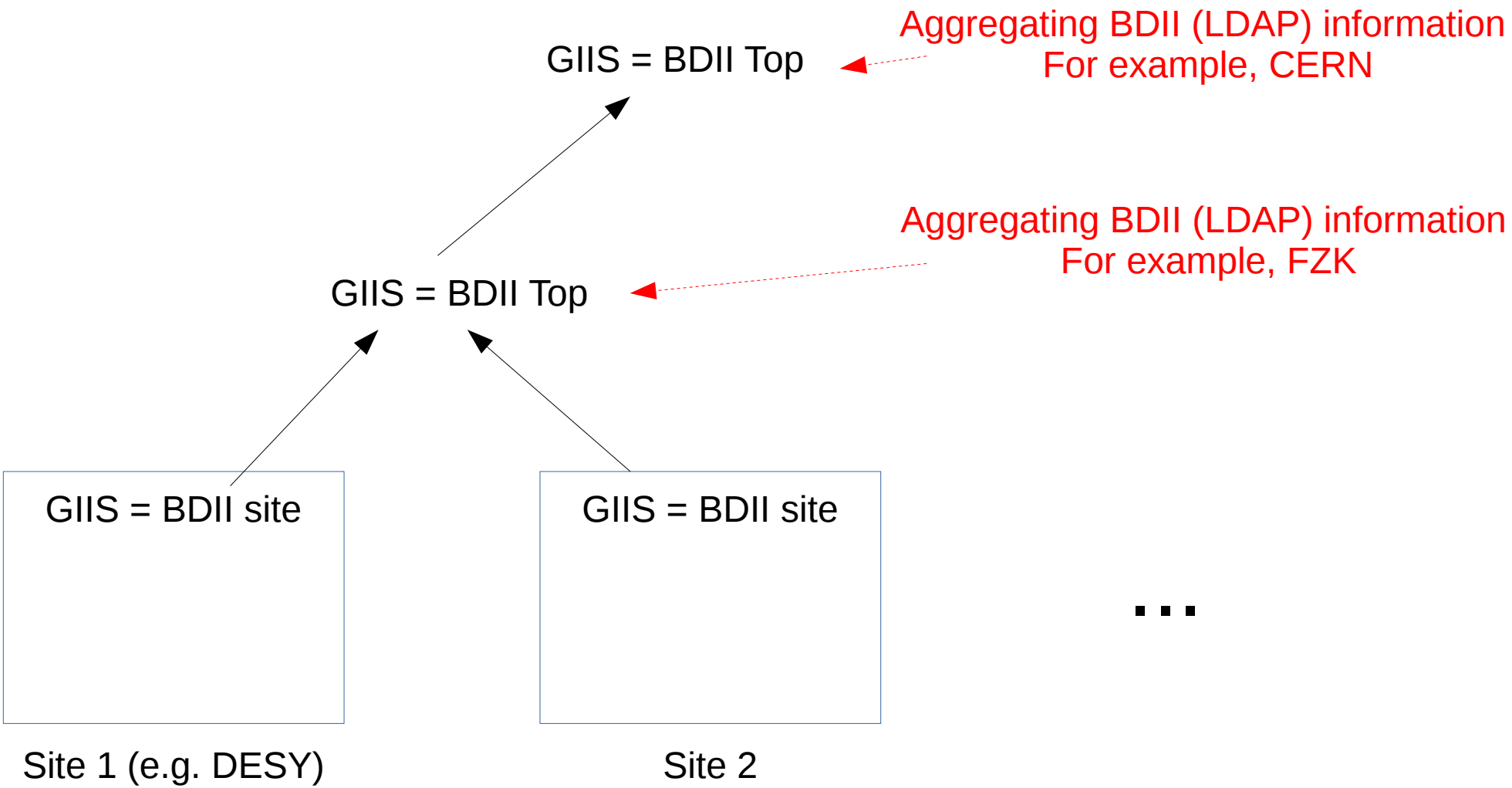
```
Usage: lcg-infosites [options] selection(s)

Selections:
   all          dli         lfc          tag
   bdii_site    dliLocal    lfcLocal     vobox
   bdii_top     fts         myproxy      voms
   ce           gridice     se           voms-admin
   closeSE      lb          sitenames    voview
   cream        lcg-ce      space        wms
```

Searching for storage element at DESY-HH
$ lcg-infosites --vo atlas se -f DESY-HH

```
 Avail Space(kB)  Used Space(kB)  Type         SE
------------------------------------------------
            n.a            n.a  SRM    dcache-se-atlas.desy.de
   144686862585   1155357848246  SRM    dcache-se-atlas.desy.de
   989417595283   2662137428541  SRM    dcache-se-atlas.desy.de
     1991791084            n.a  SRM    dcache-se-atlas.desy.de
    33956785021    30907581359  SRM    dcache-se-atlas.desy.de
     3972844749            n.a  SRM    dcache-se-atlas.desy.de
            n.a            n.a  SRM    prometheus.desy.de
      963849833        2517808  SRM    prometheus.desy.de
```

Structure of Information Providers among sites



GlueSchema & LDAP

- https://www.slac.stanford.edu/grp/eg/minos/dist/dist_aux2/packages/GridTools/HEAD/docs/glue_schema.html
- https://www.centos.org/docs/5/html/CDS/ag/8.0/Finding_Directory_Entries-LDAP_Search_Filters.html

Hands-on exercise Information Provider - 2

- Using 'ldapsearch' command

Checking site-level information provider (GIIS)
$ lcg-infosites --vo atlas bdii_site -f DESY-HH
ldap://grid-giis0.desy.de:2170/mds-vo-name=DESY-HH,o=grid
ldap://grid-giis1.desy.de:2170/mds-vo-name=DESY-HH,o=grid

Getting GridFTP endpoints from GIIS at DESY-HH
$ SE=dcache-se-atlas.desy.de
$ ldapsearch -xLLL -b 'o=grid' “(GlueChunkKey=GlueSEUniqueID=$SE)” -p 2170 -h grid-giis0.desy.de | grep gsiftp
GlueSEAccessProtocolType: gsiftp

Getting SRM endpoints from GIIS at DESY-HH
$ ldapsearch -xLLL -b 'o=grid' “(GlueChunkKey=GlueSEUniqueID=$SE)” -p 2170 -h grid-giis0.desy.de | grep httpg

ATLASDATADISK (using AND), the unit of size is “GigaByte”
$ **ldapsearch -xLLL -b 'o=grid' "(&(GlueChunkKey=GlueSEUniqueID=$SE)(GlueSALocalID=atlas:ATLASDATADISK))" **
-p 2170 -h grid-giis0.desy.de

Only OnlineSize (=GlueSATotalOnlineSize)
$ **ldapsearch -xLLL -b 'o=grid' "(&(GlueChunkKey=GlueSEUniqueID=$SE)(GlueSALocalID=atlas:ATLASDATADISK))" **
GlueSATotalOnlineSize -p 2170 -h grid-giis0.desy.de

Hands-on exercise Information Provider - 3

- Getting SE information given by a site BDII

Connecting to DESY-HH SE via SRM protocol
srmls srm://dcache-se-atlas.desy.de
512 /
512 /upload/
512 /admin/
512 /usr/
512 /pnfs/

SRM space token (which is used by Rucio Storage Endpoint)
srm-get-space-tokens -space_desc=ATLASLOCALGROUPDISK srm://dcache-se-atlas.desy.de
Space Reservation Tokens:
540002

Checking site-level information provider
lcg-infosites --vo atlas bdii_site -f DESY-HH
ldap://grid-giis0.desy.de:2170/mds-vo-name=DESY-HH,o=grid
ldap://grid-giis1.desy.de:2170/mds-vo-name=DESY-HH,o=grid

Getting GridFTP endpoints
SE=dcache-se-atlas.desy.de
ldapsearch -xLLL -b 'o=grid' '(GlueChunkKey=GlueSEUniqueID=$SE)' -p 2170 -h grid-giis0.desy.de | grep gsiftp
GlueSEAccessProtocolType: gsiftp

Connecting to DESY-HH SE via GridFTP protocol
uberftp -ls gsiftp://dcache-door-atlas12.desy.de/