Data and task management systems in the experiments at the LHC

Fernando Barreiro Megino (University Texas at Arlington)

Second International Youth Scientific School-Conference on Distributed Heterogeneous Computing Infrastructures (The 2nd International School on Heterogeneous Computing Infrastructure
25-29 September 2017, Budva

UNIVERSITY OF TEXAS ARLINGTON

PanDA

About me

- SW Engineer (2004) and Telecommunications Engineer (2007), Universidad Autónoma de Madrid (Spain)
- Working on ATLAS Distributed Computing (ADC) since 2008
    - 2008-2012: Distributed Data Management developer
    - 2012-13: ATLAS Cloud Computing co-coordinator
    - 2015-now: Workload Management developer and co-coordinator since April 2016
- 2013-2014: JP Morgan Technology Division in Geneva

Contact: barreiro@uta.edu

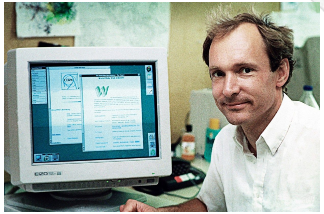
Outline

- CERN, LHC, ATLAS
- From collisions to papers
- The Worldwide LHC Computing Grid (WLCG)
- Central components
- Data management
- Workload management

CERN

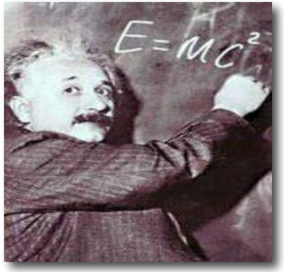
- European Organization for Nuclear Research in Geneva founded in 1954
- The world’s largest particle physics laboratory
- Particle accelerators and other infrastructure for high energy physics (HEP) research
- Worldwide community
  - 20 member states (+ 3 incoming members)
  - Observers: Turkey, Russia, Japan, USA, India
  - About 2300 staff
  - >10’000 users (about 5’000 on-site)
  - Budget (2015) ~1000 MCHF
- Birthplace of the World Wide Web

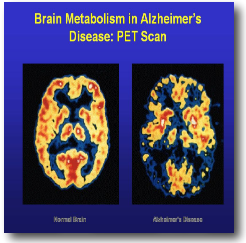
MEMBER STATES
ASSOCIATE MEMBER STATES
ASSOCIATE MEMBERS IN
THE PRE-STAGE TO MEMBERSHIP
OBSERVERS
OTHER STATES

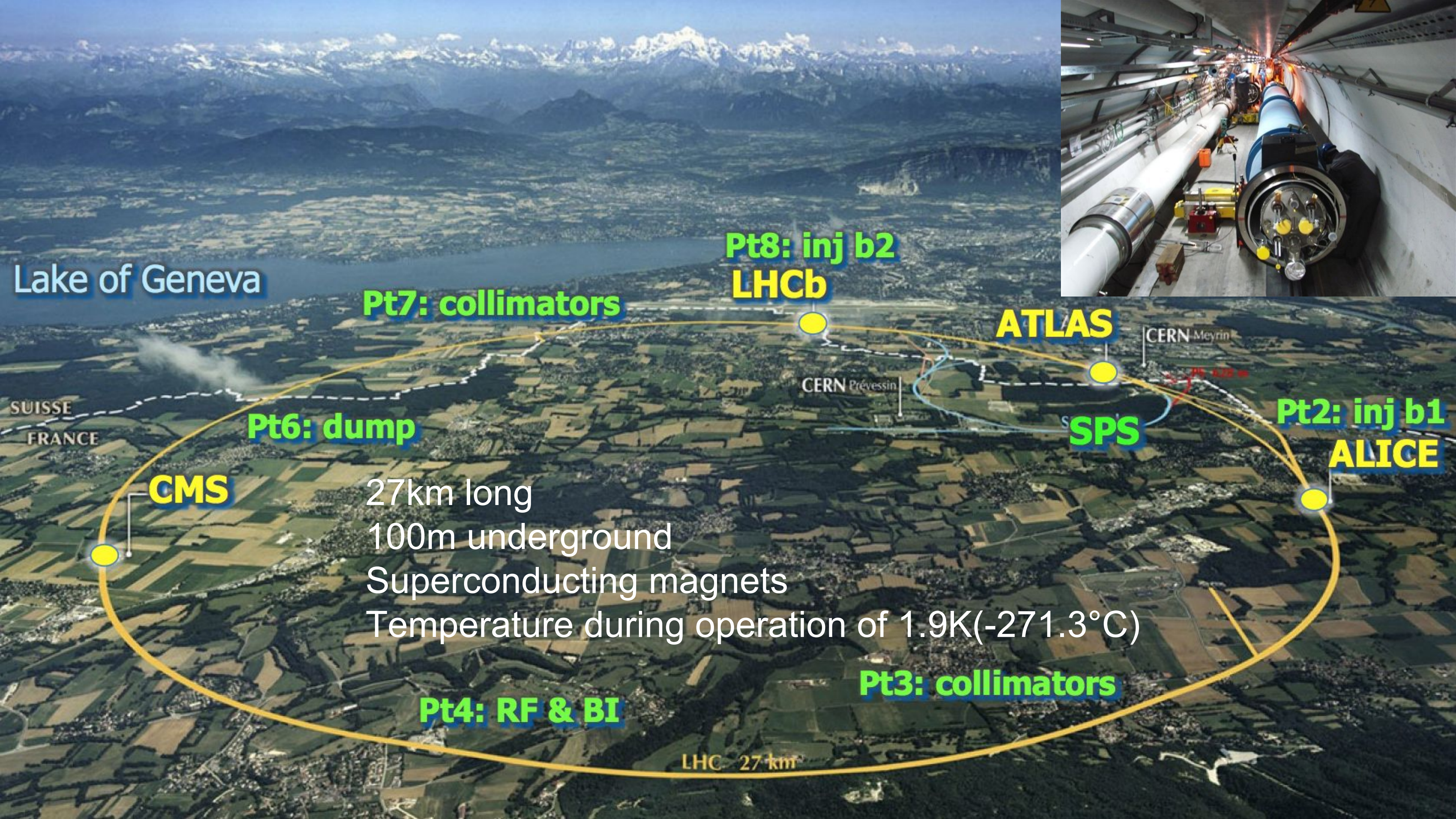


CERN’s mission

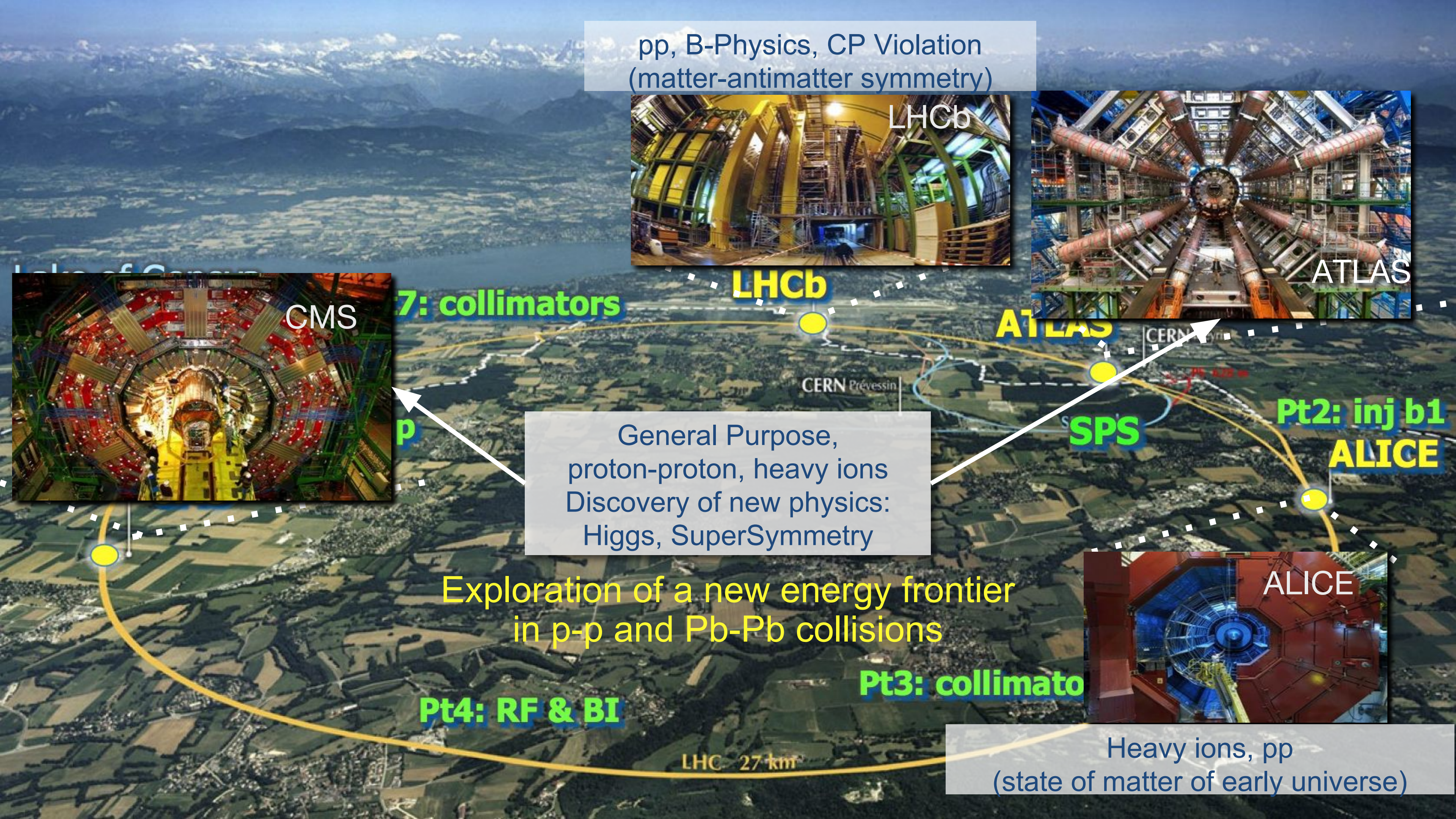
- Research
  - Seeking and finding answers to questions about the universe
  - The secrets of the Big Bang
  - Origin of mass
- Technology
  - Advancing the frontiers of technology
  - Information technology - the Web and the Grid
  - Medicine - diagnosis and therapy
- Collaborating
  - Bringing nations together through science
- Education
  - Training the scientists of tomorrow





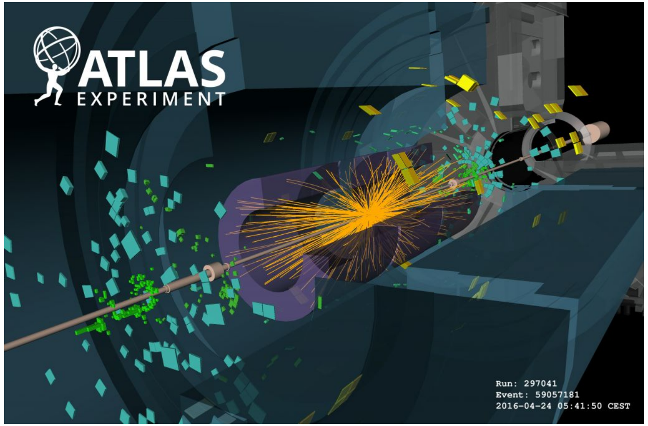


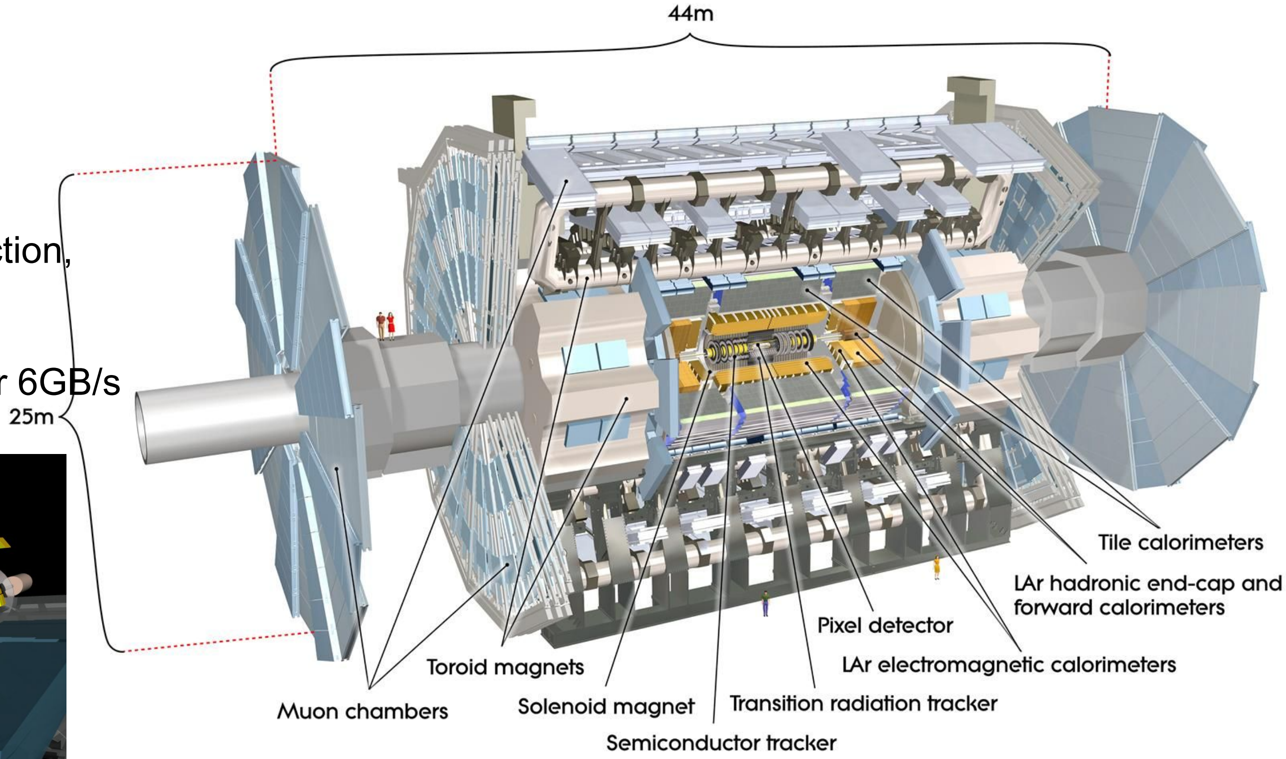
27km long
100m underground
Superconducting magnets
Temperature during operation of 1.9K(-271.3°C)


pp, B-Physics, CP Violation
(matter-antimatter symmetry)
LHCb
ATLAS
CMS
7: collimators
LHCb
CERN Prévessin
SPS
Pt2: inj b1
ALICE
General Purpose,
proton-proton, heavy ions
Discovery of new physics:
Higgs, SuperSymmetry
Exploration of a new energy frontier
in p-p and Pb-Pb collisions
ALICE
Pt3: collimato
Pt4: RF & BI
LHC 27 km
Heavy ions, pp
(state of matter of early universe)

ATLAS

- The largest detector
- Multiple components and sub-detectors
- 7000 tons
- 10 MW electric power
- 150M sensors measure direction, momentum and charge
- Collisions at 40MHz
    - Filtered to kHz or under 6GB/s

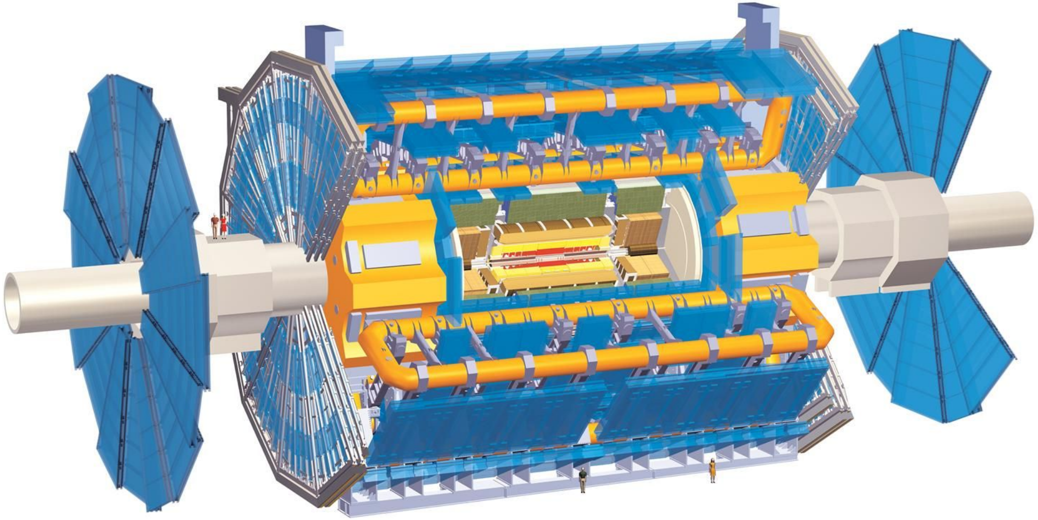




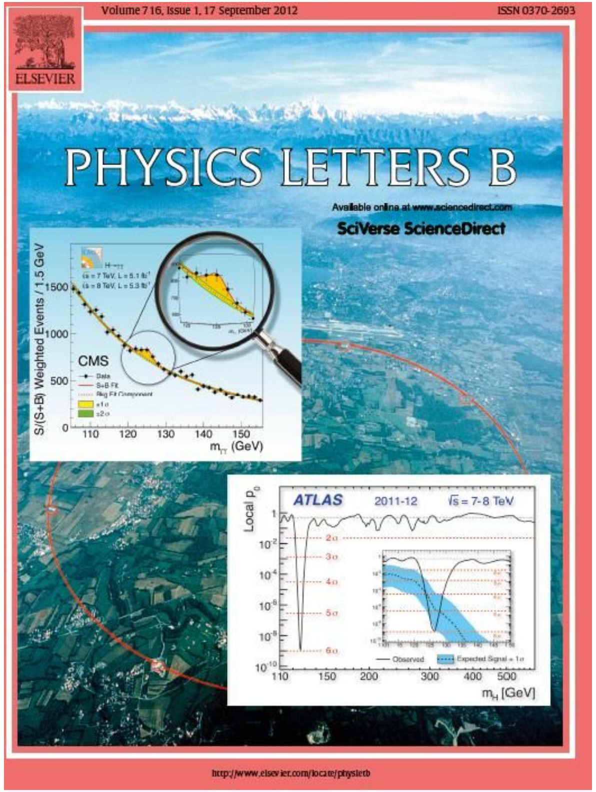
Outline

- CERN, LHC, ATLAS
- From collisions to papers
- The WLCG
- Central components
- Data management
- Workload management

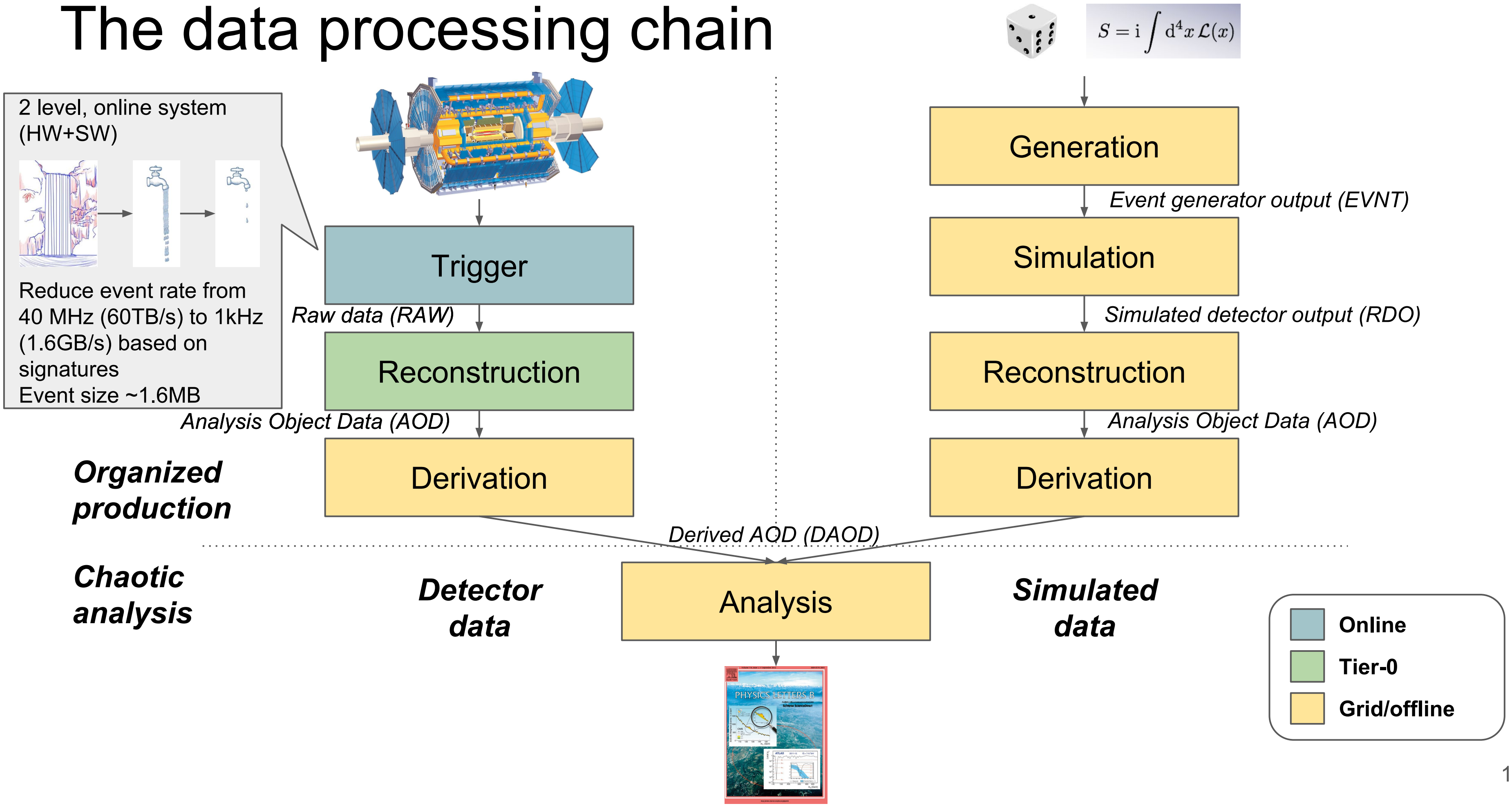
From collisions to papers in ATLAS



?



The data processing chain



What is the data: RAW data

Measurements of the sensors

```
0x01e84c10:  0x01e8 0x8848 0x01e8 0x83d8 0x6c73 0x6f72 0x7400 0x0000
0x01e84c20:  0x0000 0x0019 0x0000 0x0000 0x01e8 0x4d08 0x01e8 0x5b7c
0x01e84c30:  0x01e8 0x87e8 0x01e8 0x8458 0x7061 0x636b 0x6167 0x6500
0x01e84c40:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84c50:  0x01e8 0x8788 0x01e8 0x8498 0x7072 0x6f63 0x0000 0x0000
0x01e84c60:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84c70:  0x01e8 0x8824 0x01e8 0x84d8 0x7265 0x6765 0x7870 0x0000
0x01e84c80:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84c90:  0x01e8 0x8838 0x01e8 0x8518 0x7265 0x6773 0x7562 0x0000
0x01e84ca0:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84cb0:  0x01e8 0x8818 0x01e8 0x8558 0x7265 0x6e61 0x6d65 0x0000
0x01e84cc0:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84cd0:  0x01e8 0x8798 0x01e8 0x8598 0x7265 0x7475 0x726e 0x0000
0x01e84ce0:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84cf0:  0x01e8 0x87ec 0x01e8 0x85d8 0x7363 0x616e 0x0000 0x0000
0x01e84d00:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84d10:  0x01e8 0x87e8 0x01e8 0x8618 0x7365 0x7400 0x0000 0x0000
0x01e84d20:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84d30:  0x01e8 0x87a8 0x01e8 0x8658 0x7370 0x6c69 0x7400 0x0000
0x01e84d40:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84d50:  0x01e8 0x8854 0x01e8 0x8698 0x7374 0x7269 0x6e67 0x0000
0x01e84d60:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84d70:  0x01e8 0x875c 0x01e8 0x86d8 0x7375 0x6273 0x7400 0x0000
0x01e84d80:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84d90:  0x01e8 0x87c0 0x01e8 0x8718 0x7377 0x6974 0x6368 0x0000
```

Address: which detector element took the reading

Values: what the electronics wrote out

Fraction of a RAW event

- Was an detector element hit? How much energy and at what time?
- More than 300K such words in each event with full data from all detector elements
- Data size: ~1.6MB

What is the data: Reconstructed data

UNIVERSITY OF TEXAS ARLINGTON

PanDA

Event 1

Nch (charged tracks): 2

Pcha (Momentum of each track):
{{"-7.65698", "42.9725", "14.3404"},
{" 7.54101", "-42.1729", "-14.0108"}}
px py pz

Qcha (Charge of each track): {-1,1}

Event 2

Nch (charged tracks): 3

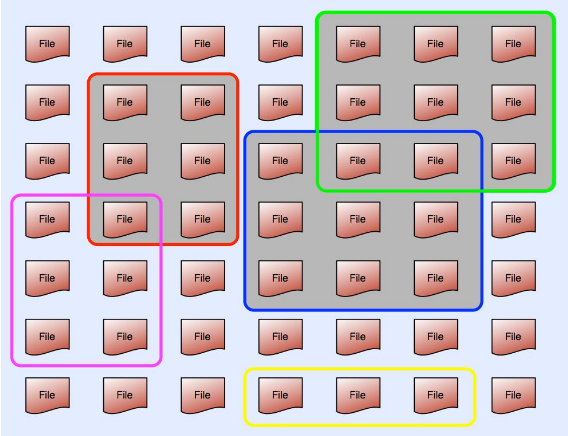
Pcha (Momentum of each track):
{{"-12.9305","12.2713","40.5615"},
{" 12.2469","-11.606","-38.7182"},
{"0.143435","-0.143435","-0.497444"}}
px py pz

Qcha (Charge of each track): {-1,1,-1}

Concepts

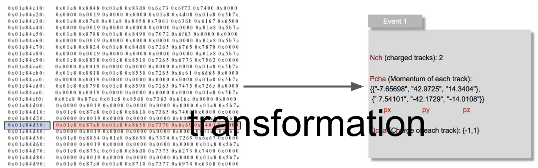
Data

- **Event:** Collision
- **Files**: Multiple event objects, parts of detector, etc.
- **Datasets**: logical grouping of files
  - Similar characteristics for improved data management
  - Flexible definition:
    - Related data: e.g. detector data from a given luminosity run
    - Auxiliary objects used for data movement purposes



Workload

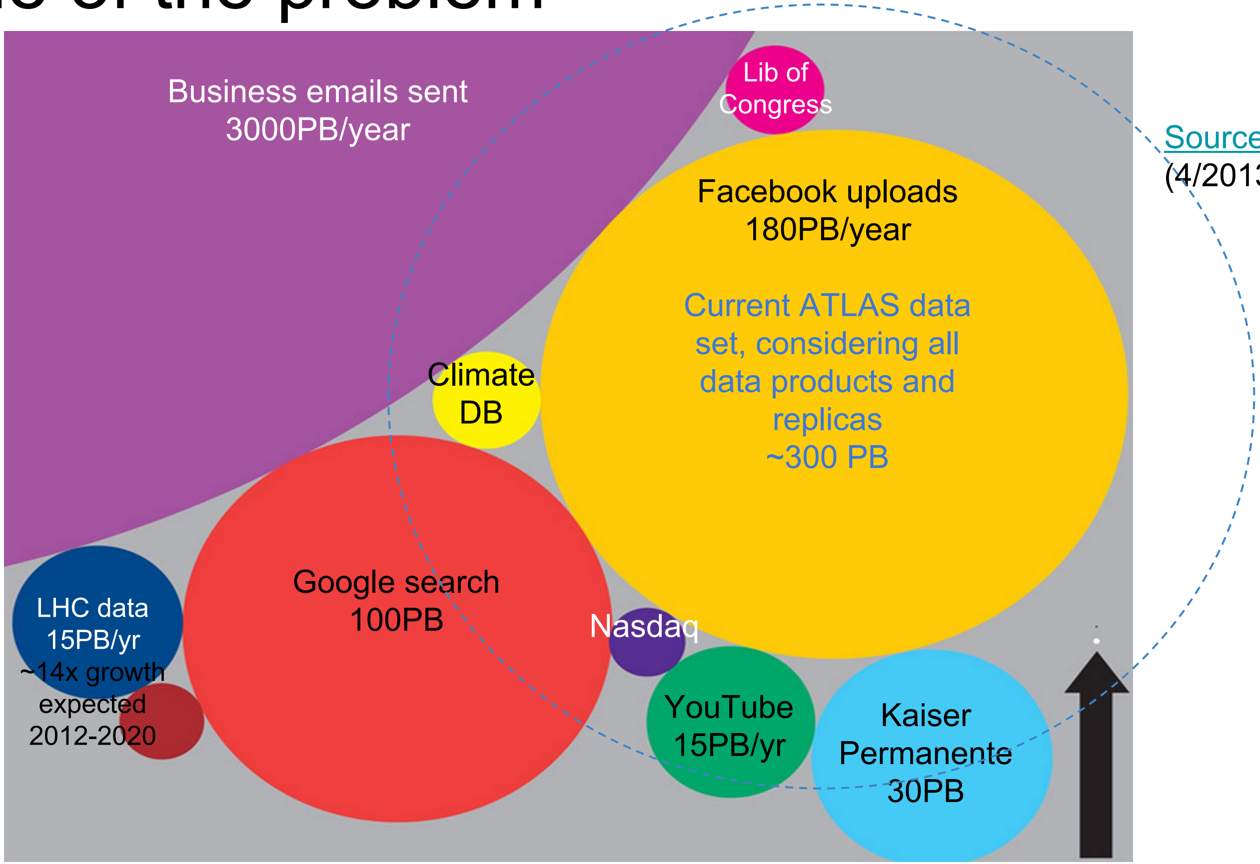
- **Jobs**: take input files, run transformation on them and produce output files
  - Transformations move data through various formats
- **Tasks:** logical grouping of jobs, usually executing over a dataset



Outline

- CERN, LHC, ATLAS
- From collisions to papers
- The Worldwide LHC Computing Grid (WLCG)
- Central components
- Data management
- Workload management

Magnitude of the problem



Source: Wired magazine
(4/2013 - a bit outdated)

Why not store and process everything at CERN?

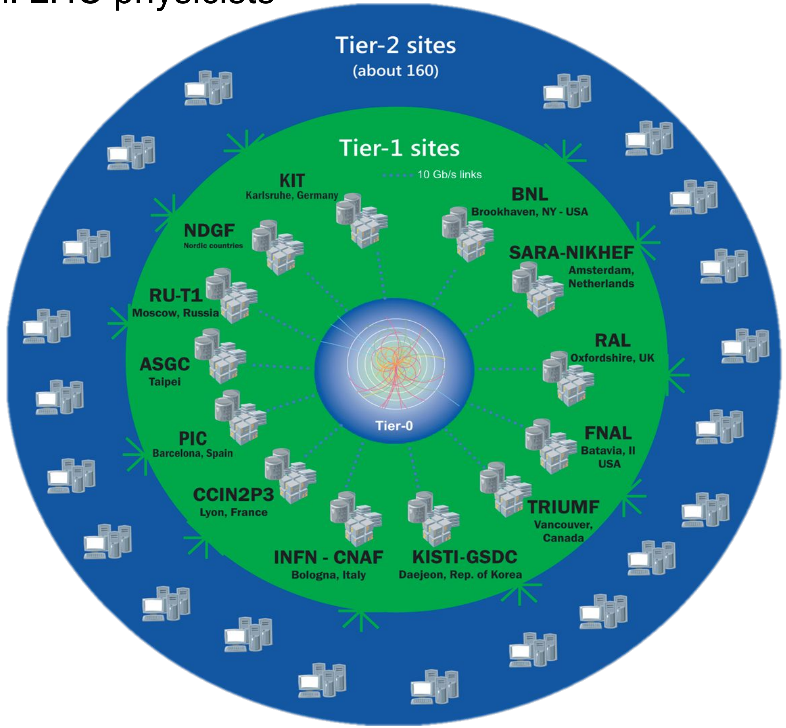
- Traditionally a single computing center at CERN could not physically provide all resources
- Data redundancy
  - LHC and ATLAS operation is expensive: we can't afford to lose any data
  - There are multiple copies, in particular of RAW detector data
- Funding reasons
  - ATLAS is an international collaborations with participants from 38 countries
  - Funding agencies prefer to invest and employ locally

http://ssl-computing.web.cern.ch/ssl-computing/default.htm

Worldwide LHC Computing Grid

- International collaboration to distribute and analyse LHC data
- Integrates computing centres worldwide that provide **computing** and **storage** resource into a single infrastructure accessible by all LHC physicists

- **Tier-0 (CERN)**: data recording and archival, prompt reconstruction, calibration and distribution
- **Tier-1s**: T0 overspilling, second tape copy of detector data, more intensive tasks
- **Tier-2s:** Processing centers, being the differences with T1s increasingly blurry



For **all** experiments:
- nearly 170 sites
- ~350k cores
- 200 PB of disk
- 10 Gb links and up

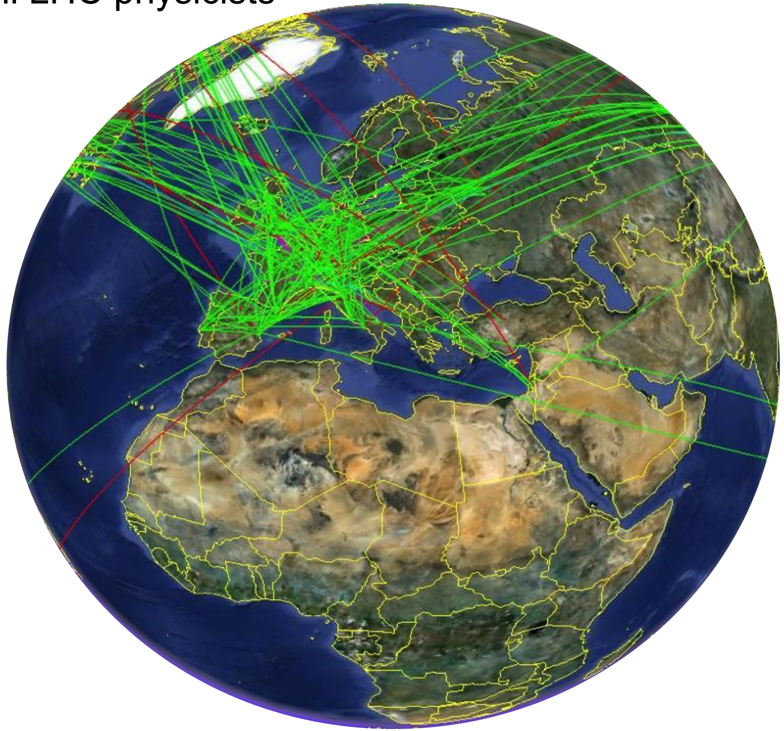
Worldwide LHC Computing Grid

- International collaboration to distribute and analyse LHC data
- Integrates computing centres worldwide that provide **computing** and **storage** resource into a single infrastructure accessible by all LHC physicists

- **Tier-0 (CERN)**: data recording and archival, prompt reconstruction and calibration and distribution

- **Tier-1s**: T0 overspilling, second tape copy of detector data, memory and CPU intensive tasks

- **Tier-2s:** Processing centers, being the differences with T1s increasingly blurry



For all experiments:

- nearly 170 sites
- ~350k cores
- 200 PB of disk
- 10 Gb links and up


BNL
New York
ASGC/Taipei
CCIN2P3/Lyon
TRIUMF
Vancouver
NIKHEF/SARA
Amsterdam
FNAL
Chicago
NRC-K, Moscow
RAL
Rutherford
JINR, Dubna
PIC
Barcelona
NDGF
Nordic countries
CNAF
Bologna
CERN
TIER2s
FZK
Karlsruhe

How does the CERN data center look nowadays?

- 2700 m2 surface for computing equipment
- 3.5 MW power
- Air and water cooling
- 10k servers, 110k cores

https://home.cern/about/computing

Main components of a WLCG site

CPU servers:

- CPU servers are grouped into **Batch systems** for processing of data

Disk servers:

- CPU server attached to several disks
- Disk servers are grouped into **Storage Elements**
- Data storage with fast access

Tape robots:

- Long term archival
- Slow access
- All raw data from the experiments has 2 tape copies

Grid middleware: the glue

- Heterogeneous resources are grouped and exposed in a uniform way
  - Computing Elements give access to CPUs
  - Storage Elements give access to data
  - Information systems describe the grid
  - Authentication is done via x509 public key infrastructure

Outline

- CERN, LHC, ATLAS
- From collisions to papers
- The WLCG
- Central components
- Data management
- Workload management

Paradigm shift in HEP Computing

Old paradigms	New ideas
• Distributed resources are **independent entities**	• Distributed resources are seamlessly **integrated worldwide** through a single submission system • **Hide middleware** while supporting diversity
• Groups of users utilize **specific resources** (whether locally or remotely)	• All users have access to **same resources**
• Fair shares, priorities and policies are managed **locally**, for each resource	• **Global** fair share, priorities and policies allow efficient management of resources
• **Uneven user experience** at different sites, based on local support and experience	• **Automation, error handling, and other features** improve user experience • Central support coordination
• Privileged users have access to **special resources**	• All users have access to **same resources**

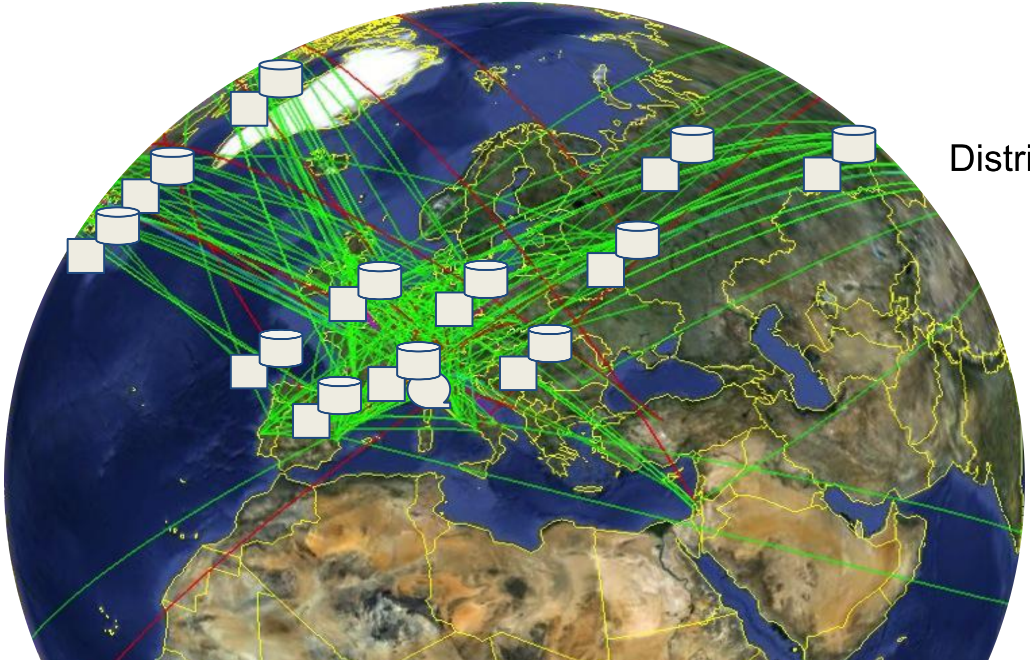
ATLAS orchestrators

Workload Management: submission and scheduling of jobs & tasks

PanDA

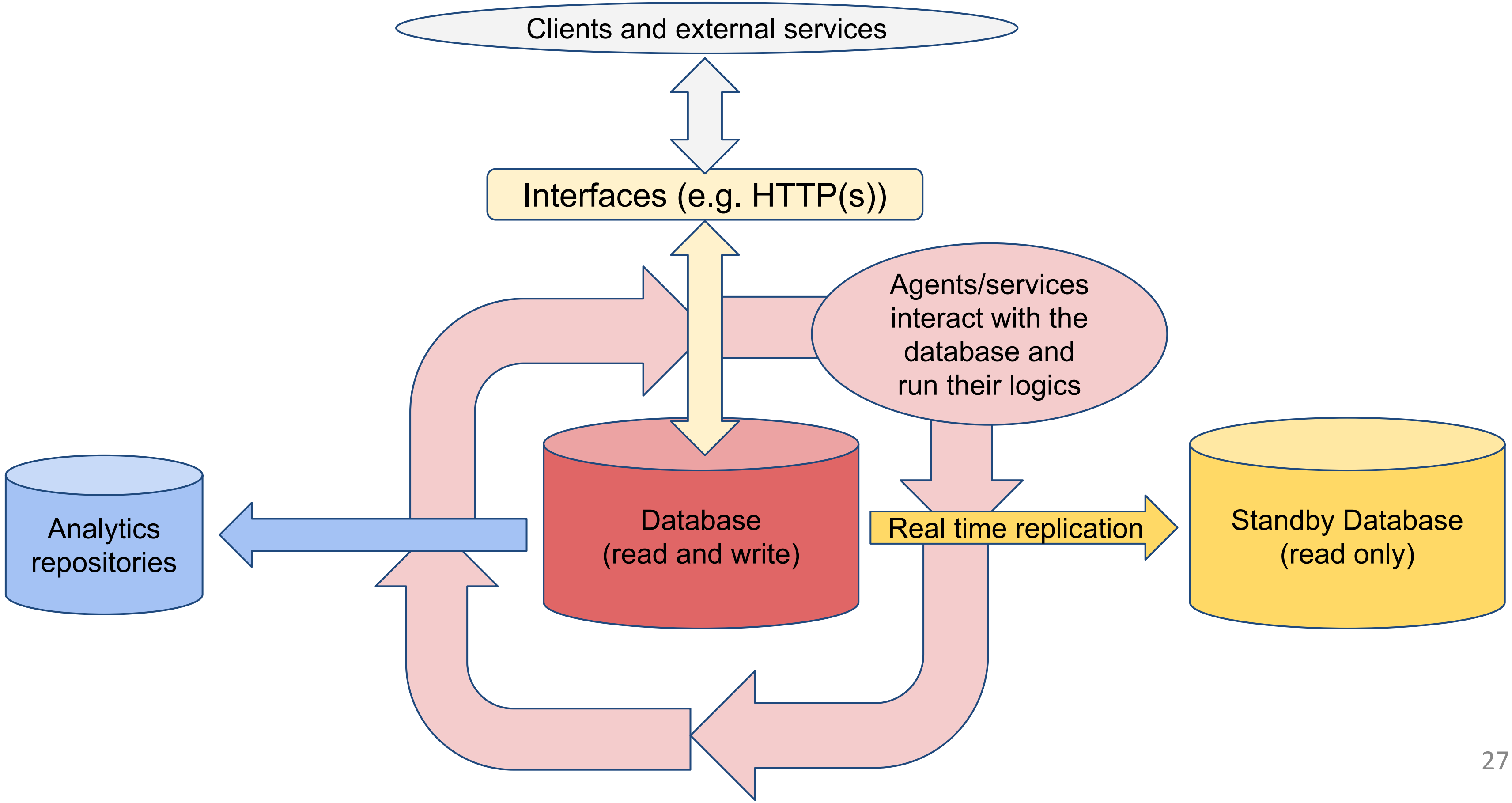
Central components at CERN

Data Management: bookkeeping and distribution of files & datasets



Distributed resources

General architecture of central components



Usual technology stack

Usually Open Source tools

- Linux: SLC6, CentOS7
- Python
  - HTTP: requests, urllib2, httplib...
  - Web services: django, flask…
  - DB access: cxOracle, SQLAlchemy...
- Apache httpd, ActiveMQ...
- Data analytics platforms:
  - Hadoop
  - Elastic Search

Advantages of Python

- Faster development
- Clean, straightforward syntax
- Developer does not take care of memory management
- High level native data types; no type declaration
- Duck typing
- Iterators, generators and comprehensions
- Huge standard library
- Great support for building web apps

NB: Python is also slower and uses more memory, but since the bottleneck in our programs is remote calls and DB queries, it's justified. We gain more by optimizing our SQL than our python code.

But there are use cases, where you need a lower level language

- C

```
#include <stdio.h>

int main(int argc, char ** argv)
{
   printf("Hello, World!\n");
}
```

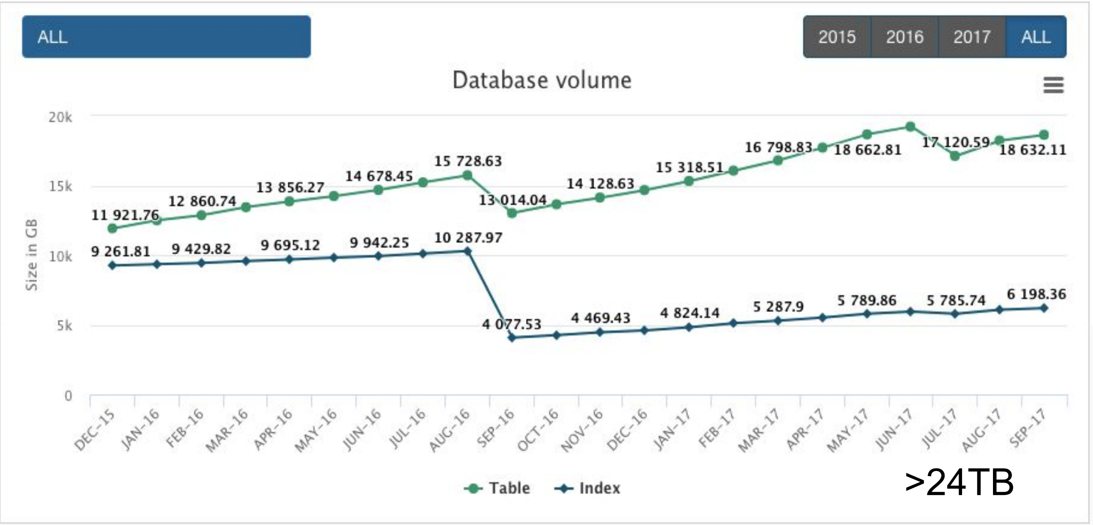
- Java

```
public class Hello
{
   public static void main(String argv[])
   {
      System.out.println("Hello, World!");
   }
}
```

- now in Python

```
print "Hello, World!"
```

Database growth over time



ATLAS has a very good DBA team that advises developers. Applications need to scale up

- Indexes
- Query optimisation
- Compression
- Time partitioning
- Archiving of data to separate schema

Outline

- CERN, LHC, ATLAS
- From collisions to papers
- The WLCG
- Central components
- Data management
- Workload management

Distributed Data Management in a nutshell

- Stores and manages all the experiment's data across a distributed environment following the computing model principles
  - Computing model determines the number and location of copies of different types of data
- High-level requirements:
  - **Data bookkeeping**
    - Location of files and datasets
    - Relationship between files and datasets
    - Owner, checksum and other metadata
  - **Data transfers**
  - **Data deletion**
  - **Data consistency**
    - Are the files really where we think they are?
- Each experiment has their own with slightly different features: we will focus on ATLAS Rucio (http://rucio.cern.ch/ developed by CERN and Univ. Oslo)

Rucio (ATLAS)

PhEDEx (CMS)

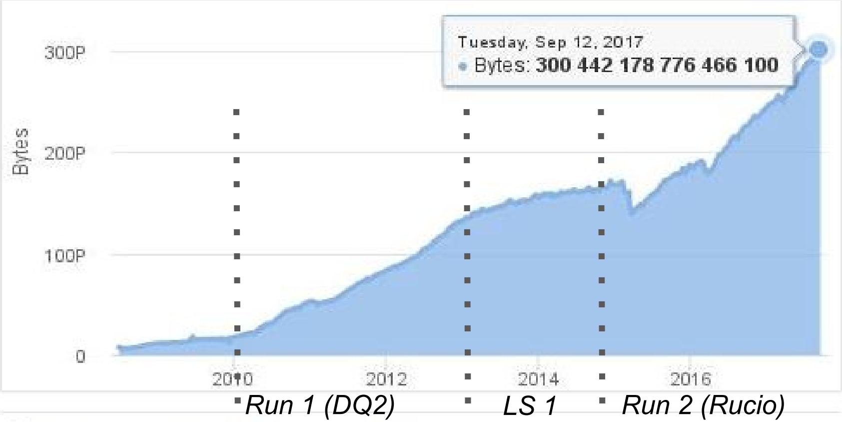
AliEn (Alice)

DIRAC (LHCb)

ATLAS Distributed Data Management: Rucio

Mario Lassnig is at CERN.
4 hrs · Meyrin ·

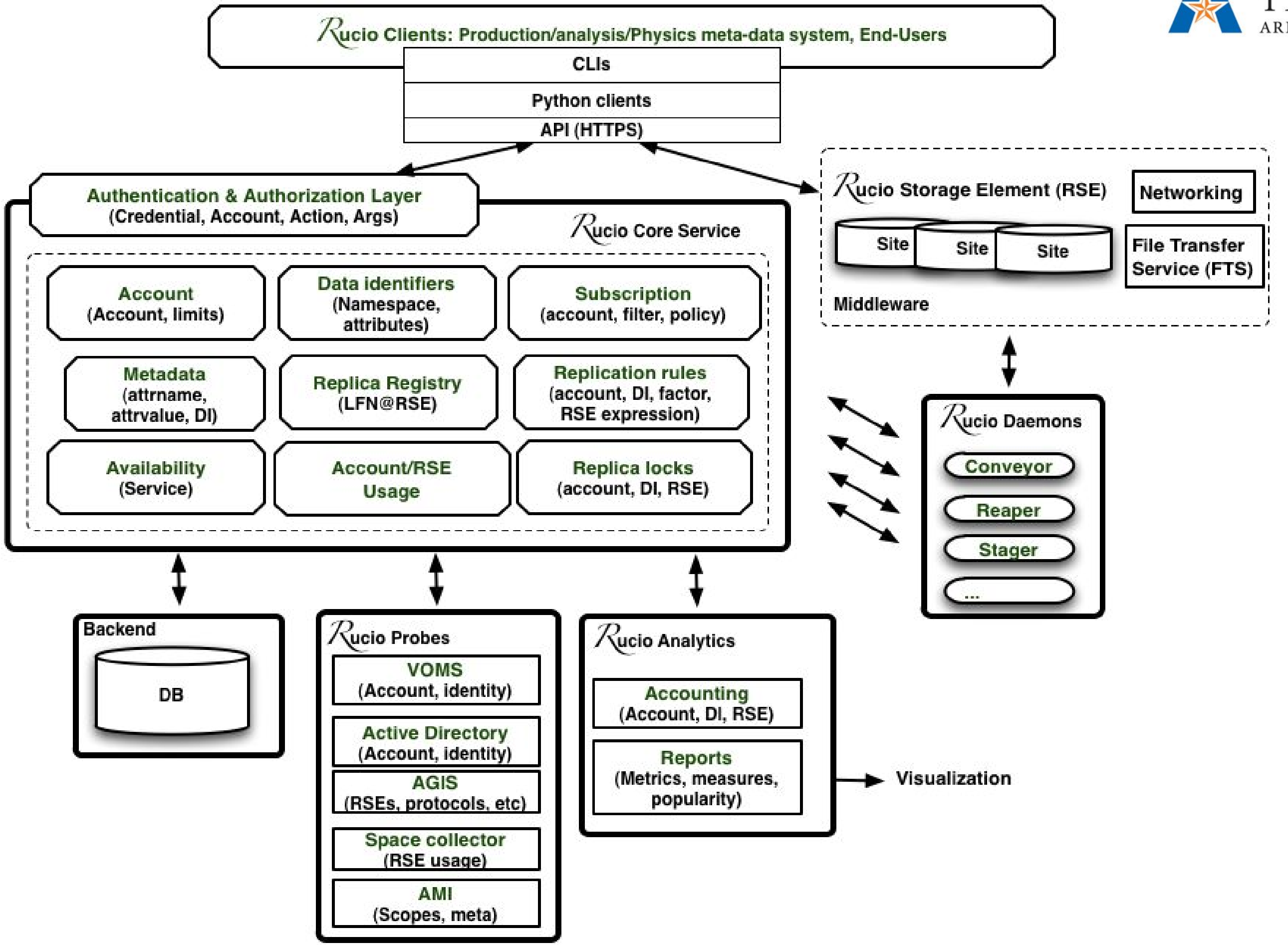
We're proud that our data management system Rucio has breached 300 Petabytes of scientific data for the ATLAS Experiment. Over 1 billion data products, 150 data centres worldwide, multi petabyte transfers and deletions per day amounting to more than 1 Exabyte in the last year. Extremely scalable, stable, and robust. Congratulations to everyone involved, we are now moving to a factor 10 increase in usage for LHC Runs 3 and 4.



Like Comment Share

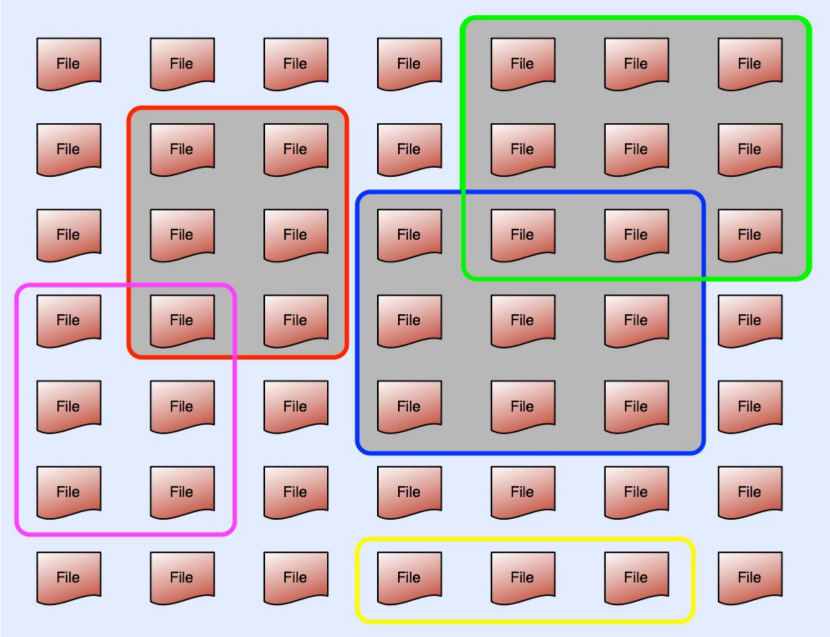
Dan van der Ster, Xavier Espinal and 27 others

Data Management: Rucio architecture



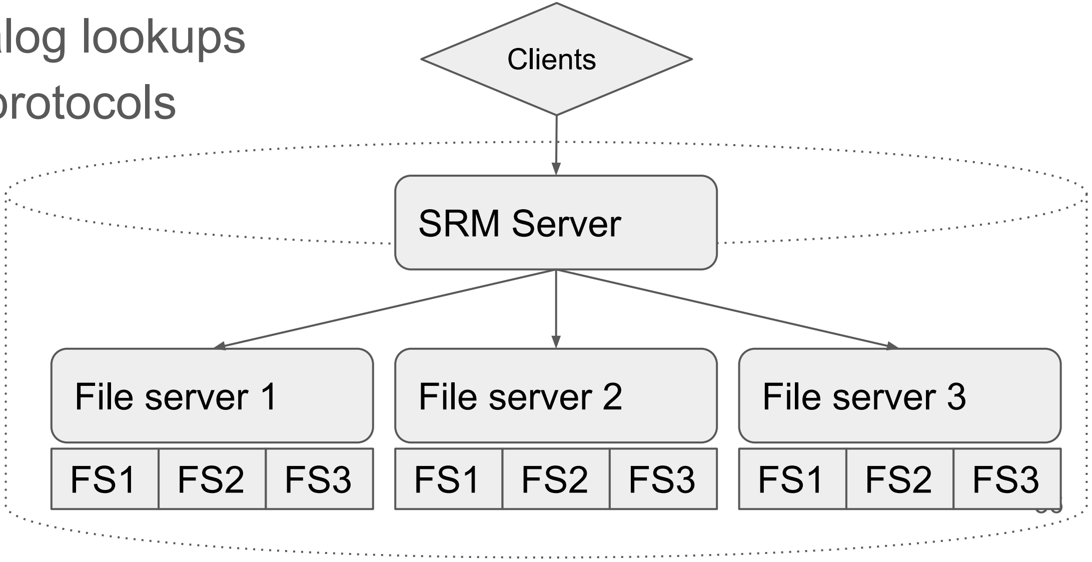
Rucio data concepts

- Events: collisions
- Files: Collections of events (e.g. C++ objects)
- Datasets: logical grouping of files
  - Units of replication



Rucio Storage Element (RSE)

- Software abstraction for a storage end-point
  - E.g. CERN_DATADISK, JINR_DATADISK,...
- A deterministic mapping between the logical file name and its path can be used to remove file catalog lookups
- RSEs support multiple protocols
  - GridFTP, HTTP, S3, etc.

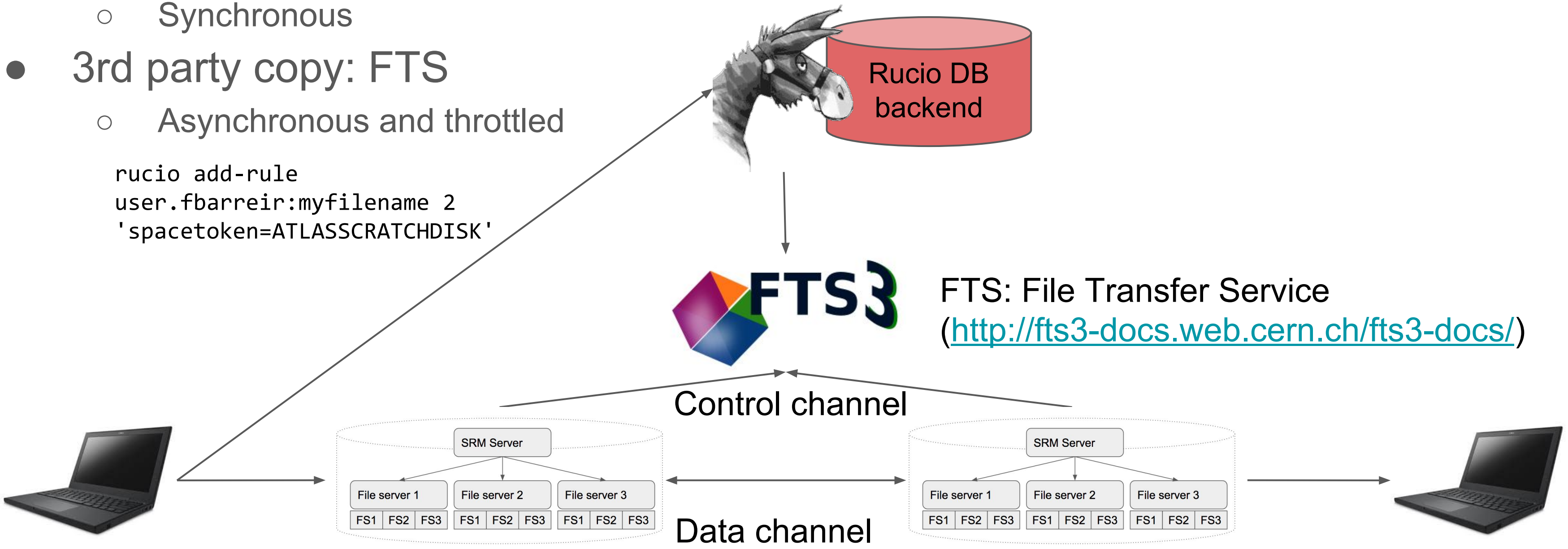


SRM: Storage Resource Manager

Interaction with the data

- Upload and download
  - Synchronous
- 3rd party copy: FTS
  - Asynchronous and throttled

```
rucio add-rule
user.fbarreir:myfilename 2
'spacetoken=ATLASSCRATCHDISK'
```



FTS: File Transfer Service
(http://fts3-docs.web.cern.ch/fts3-docs/)

```
rucio upload --rse MOCK
--scope fbarreir
--files myfilename
--dsn mydataset
```

```
rucio download --dir=/tmp/
fbarreir:myfilename
```

Rucio hides all the complicated details (paths, protocols, hosts) from the users!

Listing, copying and removing files.

```
[ui03] > edg-gridftp-ls --verbose gsiftp://i2g-se01.lip.pt/flatfiles/itut
total 4
drwxrwxr-x    3 itut         4096 Nov  8 15:06 generated
drwxrwxr-x    2 itut         4096 Nov  8 18:32 tut-14-11-07
```

```
[ui03] > globus-url-copy -vb file:/home/tutorial/user01/data-manag/dm-file-user01
gsiftp://i2g-se01.lip.pt/flatfiles/itut/tut-14-11-07/dm-file-user01
      1048576 bytes       329.49 KB/sec avg       329.49 KB/sec inst
```

```
[ui03] > edg-gridftp-ls --verbose gsiftp://i2g-
se01.lip.pt/flatfiles/itut/tut-14-11-07
total 9412
-rw-rw-r--    1 itut      9621413 Nov  8 18:33 dm-file-user01
```

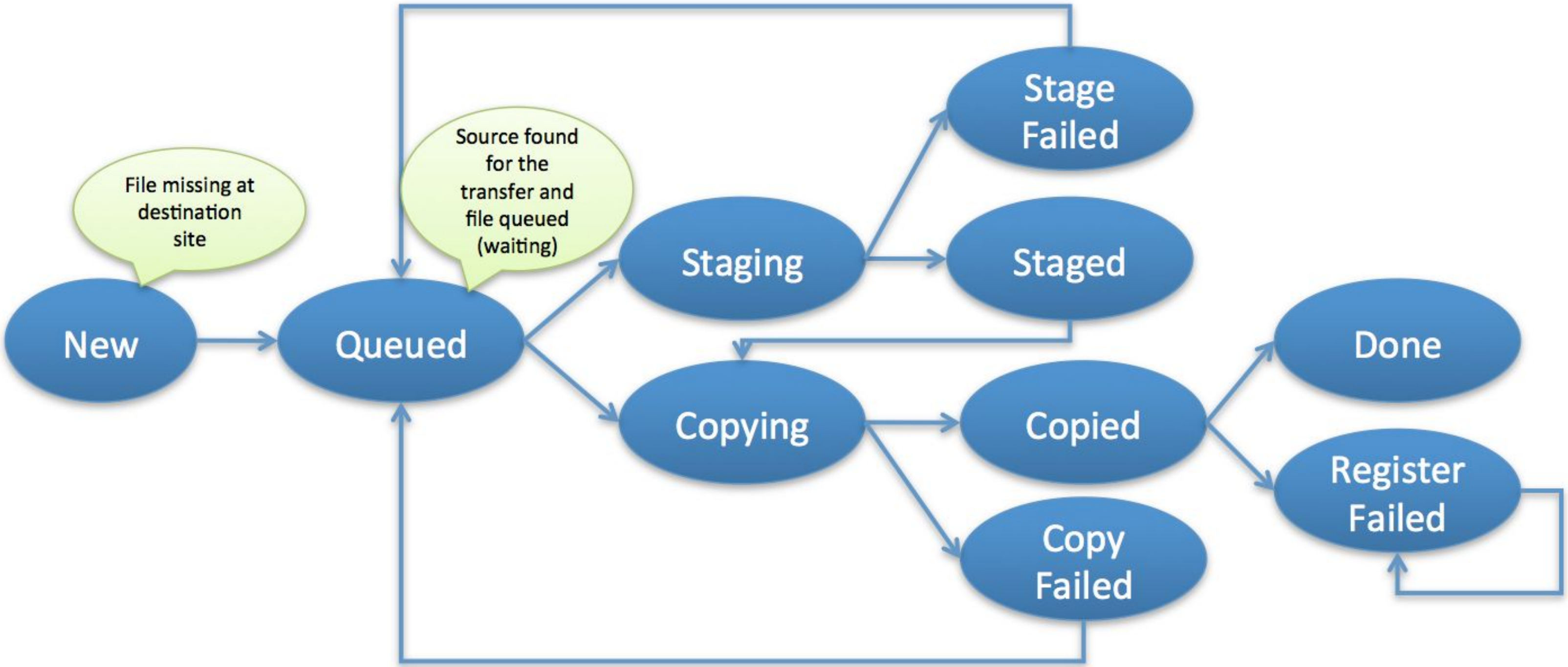
```
[ui03] > edg-gridftp-rm gsiftp://i2g-
se01.lip.pt/flatfiles/itut/tut-14-11-07/dm-file-user01
```

LIP

gLite and INT.EU.GRID training for end-users

5

State machines: file transfer



Dark data and consistency checks

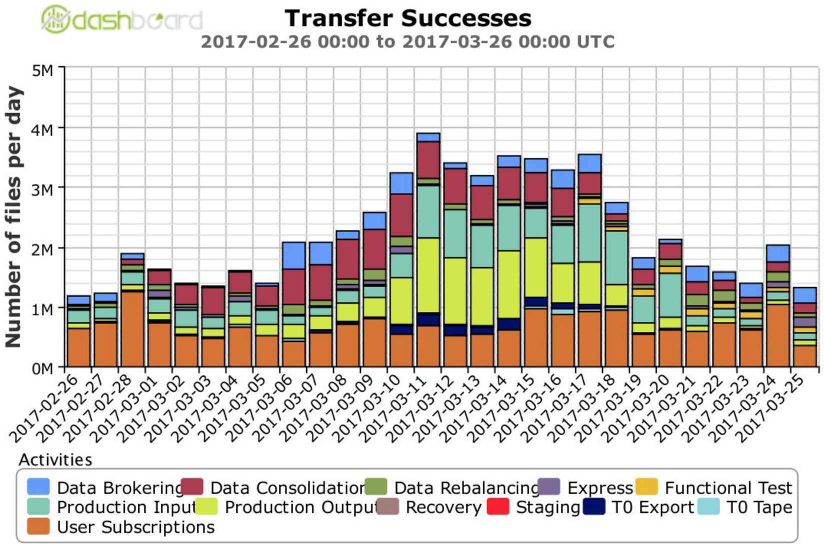
- Consistency between the Rucio Storage Elements and the Rucio DB
  - **Lost Files:** Files in the catalog(s) but not physically on the SE
  - **Dark Data:** Files on the SE but not registered in the catalog(s)
- Automatic consistency check is based on comparison of information dumps
  - Each site provides storage elements dumps on a regular basis (monthly or quarterly)
  - Rucio dumps of expected file replicas generated every day
- Comparison times scale from **few seconds** for small sites to **few hours** for the biggest one (70M files)
  - **Dark Data** is automatically collected and deleted by another daemon
  - **Lost Files** are reported to site support for investigation
    - Confirmed Lost Files are
      - Copied from other SEs if other copies exist
      - Notified to the owner and deleted from the dataset

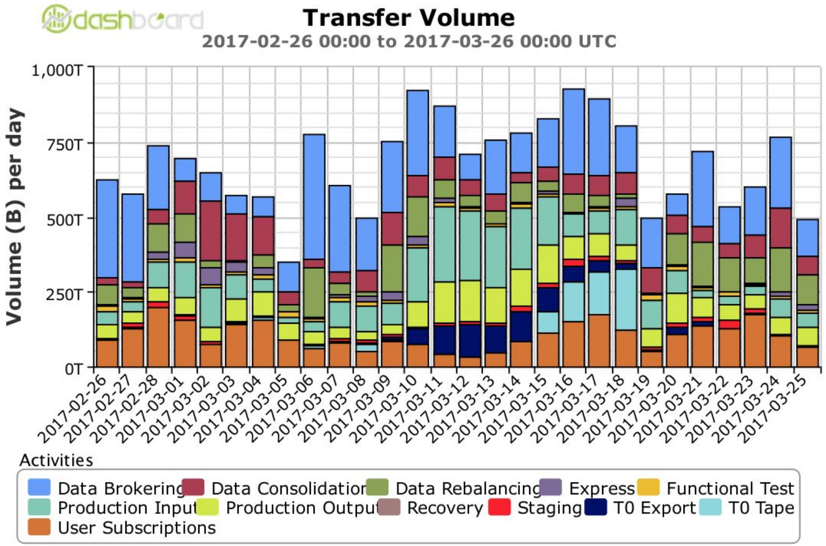
Traces, data popularity and analytics

- Common questions
  - Which files/datasets are popular in the system?
  - Which files/datasets are not used at all?
  - Statistics on transfers times, deletion times, etc.
- Traces: each event is sent to HDFS (Hadoop File System)
  - Important information: event type, file, dataset, source/destination, user, size, transfer time
  - 6M json dictionaries per day (~5GB)
- **Data reduction:** redundant, unused copies of old data can be removed
- **Data pre-placement:** popular data can be replicated to facilitate usage
- **Network map:** measure current bandwidth between sites

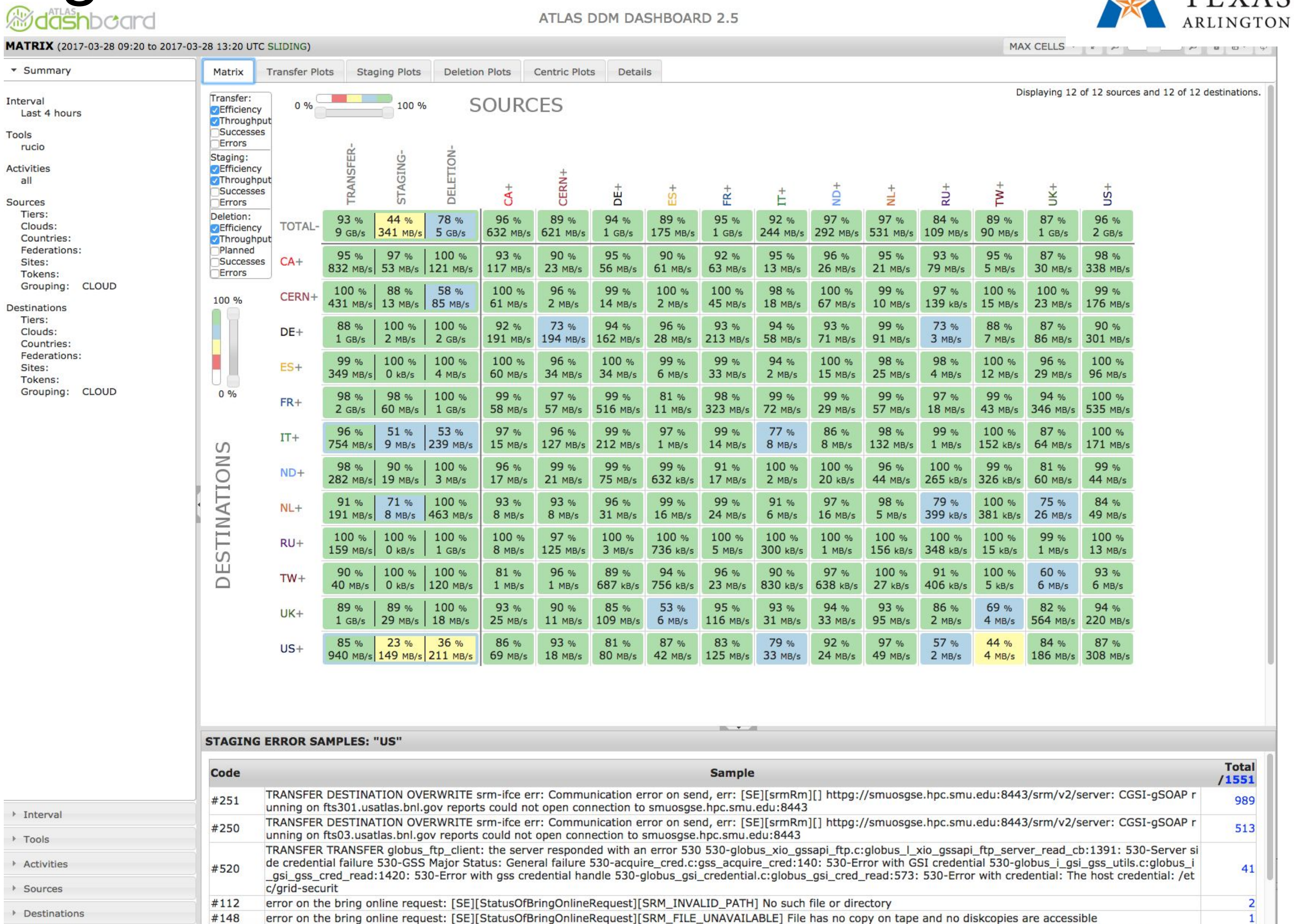
Data Management: some metrics

- Transfers
  - >40M files/month
  - Up to 40 PB/month
- Download
  - 150M files/month
  - 50 PB/month
- Deletion
  - 100M files/month
  - 40 PB/month





Monitoring: DDM Dashboard

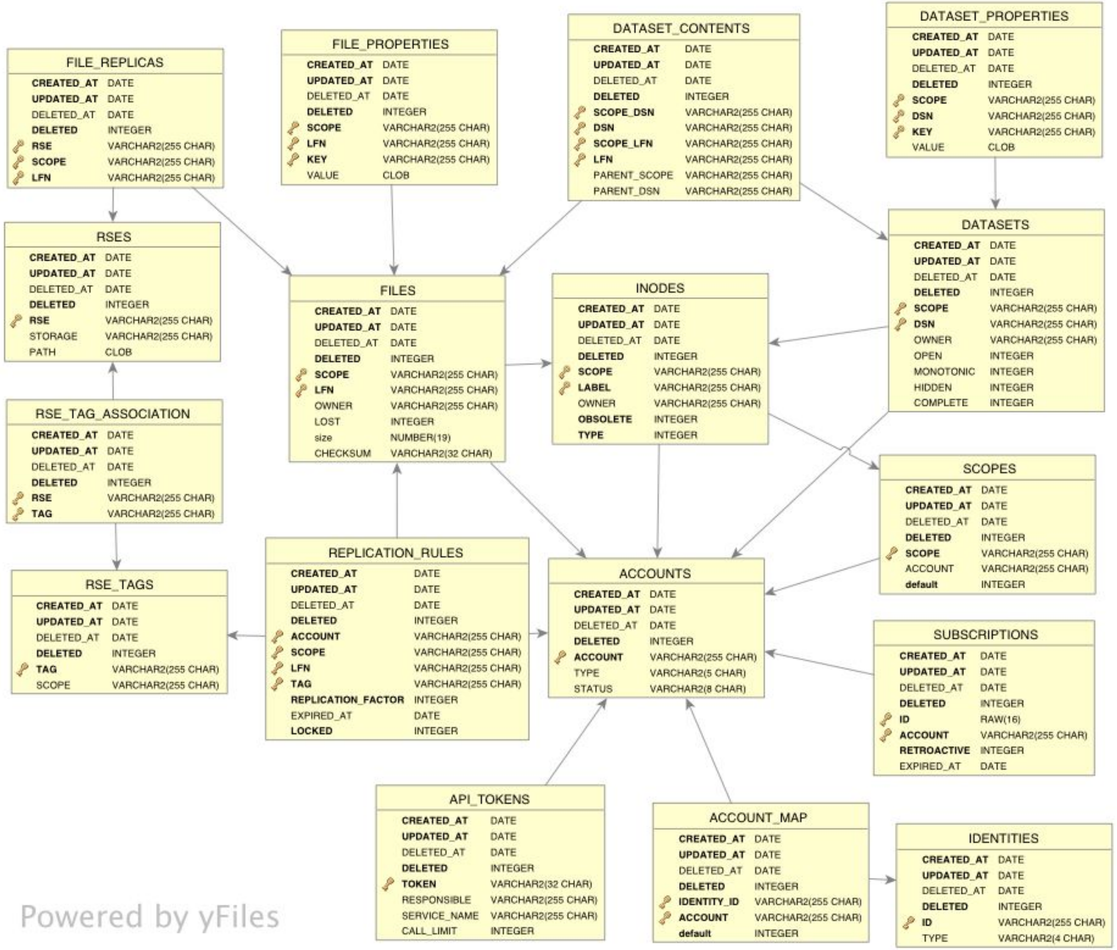


	TRANSFER-	STAGING-	DELETION-	CA+	CERN+	DE+	ES+	FR+	IT+	ND+	NL+	RU+	TW+	UK+	US+
TOTAL-	93 % 9 GB/s	44 % 341 MB/s	78 % 5 GB/s	96 % 632 MB/s	89 % 621 MB/s	94 % 1 GB/s	89 % 175 MB/s	95 % 1 GB/s	92 % 244 MB/s	97 % 292 MB/s	97 % 531 MB/s	84 % 109 MB/s	89 % 90 MB/s	87 % 1 GB/s	96 % 2 GB/s
CA+	95 % 832 MB/s	97 % 53 MB/s	100 % 121 MB/s	93 % 117 MB/s	90 % 23 MB/s	95 % 56 MB/s	90 % 61 MB/s	92 % 63 MB/s	95 % 13 MB/s	96 % 26 MB/s	95 % 21 MB/s	93 % 79 MB/s	95 % 5 MB/s	87 % 30 MB/s	98 % 338 MB/s
CERN+	100 % 431 MB/s	88 % 13 MB/s	58 % 85 MB/s	100 % 61 MB/s	96 % 2 MB/s	99 % 14 MB/s	100 % 2 MB/s	100 % 45 MB/s	98 % 18 MB/s	100 % 67 MB/s	99 % 10 MB/s	97 % 139 kB/s	100 % 15 MB/s	100 % 23 MB/s	99 % 176 MB/s
DE+	88 % 1 GB/s	100 % 2 MB/s	100 % 2 GB/s	92 % 191 MB/s	73 % 194 MB/s	94 % 162 MB/s	96 % 28 MB/s	93 % 213 MB/s	94 % 58 MB/s	93 % 71 MB/s	99 % 91 MB/s	73 % 3 MB/s	88 % 7 MB/s	87 % 86 MB/s	90 % 301 MB/s
ES+	99 % 349 MB/s	100 % 0 kB/s	100 % 4 MB/s	100 % 60 MB/s	96 % 34 MB/s	100 % 34 MB/s	99 % 6 MB/s	99 % 33 MB/s	94 % 2 MB/s	100 % 15 MB/s	98 % 25 MB/s	98 % 4 MB/s	100 % 12 MB/s	96 % 29 MB/s	100 % 96 MB/s
FR+	98 % 2 GB/s	98 % 60 MB/s	100 % 1 GB/s	99 % 58 MB/s	97 % 57 MB/s	99 % 516 MB/s	81 % 11 MB/s	98 % 323 MB/s	99 % 72 MB/s	99 % 29 MB/s	99 % 57 MB/s	97 % 18 MB/s	99 % 43 MB/s	94 % 346 MB/s	100 % 535 MB/s
IT+	96 % 754 MB/s	51 % 9 MB/s	53 % 239 MB/s	97 % 15 MB/s	96 % 127 MB/s	99 % 212 MB/s	97 % 1 MB/s	99 % 14 MB/s	77 % 8 MB/s	86 % 8 MB/s	98 % 132 MB/s	99 % 1 MB/s	100 % 152 kB/s	87 % 64 MB/s	100 % 171 MB/s
ND+	98 % 282 MB/s	90 % 19 MB/s	100 % 3 MB/s	96 % 17 MB/s	99 % 21 MB/s	99 % 75 MB/s	99 % 632 kB/s	91 % 17 MB/s	100 % 2 MB/s	100 % 20 kB/s	96 % 44 MB/s	100 % 265 kB/s	99 % 326 kB/s	81 % 60 MB/s	99 % 44 MB/s
NL+	91 % 191 MB/s	71 % 8 MB/s	100 % 463 MB/s	93 % 8 MB/s	93 % 8 MB/s	96 % 31 MB/s	99 % 16 MB/s	99 % 24 MB/s	91 % 6 MB/s	97 % 16 MB/s	98 % 5 MB/s	79 % 399 kB/s	100 % 381 kB/s	75 % 26 MB/s	84 % 49 MB/s
RU+	100 % 159 MB/s	100 % 0 kB/s	100 % 1 GB/s	100 % 8 MB/s	97 % 125 MB/s	100 % 3 MB/s	100 % 736 kB/s	100 % 5 MB/s	100 % 300 kB/s	100 % 1 MB/s	100 % 156 kB/s	100 % 348 kB/s	100 % 15 kB/s	99 % 1 MB/s	100 % 13 MB/s
TW+	90 % 40 MB/s	100 % 0 kB/s	100 % 120 MB/s	81 % 1 MB/s	96 % 1 MB/s	89 % 687 kB/s	94 % 756 kB/s	96 % 23 MB/s	90 % 830 kB/s	97 % 638 kB/s	100 % 27 kB/s	91 % 406 kB/s	100 % 5 kB/s	60 % 6 MB/s	93 % 6 MB/s
UK+	89 % 1 GB/s	89 % 29 MB/s	100 % 18 MB/s	93 % 25 MB/s	90 % 11 MB/s	85 % 109 MB/s	53 % 6 MB/s	95 % 116 MB/s	93 % 31 MB/s	94 % 33 MB/s	93 % 95 MB/s	86 % 2 MB/s	69 % 4 MB/s	82 % 564 MB/s	94 % 220 MB/s
US+	85 % 940 MB/s	23 % 149 MB/s	36 % 211 MB/s	86 % 69 MB/s	93 % 18 MB/s	81 % 80 MB/s	87 % 42 MB/s	83 % 125 MB/s	79 % 33 MB/s	92 % 24 MB/s	97 % 49 MB/s	57 % 2 MB/s	44 % 4 MB/s	84 % 186 MB/s	87 % 308 MB/s

STAGING ERROR SAMPLES: "US"

Code	Sample	Total /1551
#251	TRANSFER DESTINATION OVERWRITE srm-ifce err: Communication error on send, err: [SE][srmRm][] httpg://smuosgse.hpc.smu.edu:8443/srm/v2/server: CGSI-gSOAP running on fts301.usatlas.bnl.gov reports could not open connection to smuosgse.hpc.smu.edu:8443	989
#250	TRANSFER DESTINATION OVERWRITE srm-ifce err: Communication error on send, err: [SE][srmRm][] httpg://smuosgse.hpc.smu.edu:8443/srm/v2/server: CGSI-gSOAP running on fts03.usatlas.bnl.gov reports could not open connection to smuosgse.hpc.smu.edu:8443	513
#520	TRANSFER TRANSFER globus_ftp_client: the server responded with an error 530 530-globus_xio_gssapi_ftp.c:globus_l_xio_gssapi_ftp_server_read_cb:1391: 530-Server side credential failure 530-GSS Major Status: General failure 530-acquire_cred.c:gss_acquire_cred:140: 530-Error with GSI credential 530-globus_i_gsi_gss_utils.c:globus_i_gsi_gss_cred_read:1420: 530-Error with gss credential handle 530-globus_gsi_credential.c:globus_gsi_cred_read:573: 530-Error with credential: The host credential: /etc/grid-securit	41
#112	error on the bring online request: [SE][StatusOfBringOnlineRequest][SRM_INVALID_PATH] No such file or directory	2
#148	error on the bring online request: [SE][StatusOfBringOnlineRequest][SRM_FILE_UNAVAILABLE] File has no copy on tape and no diskcopies are accessible	1

Database schema



UNIVERSITY OF TEXAS ARLINGTON

PanDA

http://rucio.cern.ch/

Outline

- CERN, LHC, ATLAS
- From collisions to papers
- The WLCG
- Central components
- Data management
- Workload management

Workload Management in a nutshell

- High level requirements:
  - Bookkeeping of all active and past jobs
  - Centralized job queue
  - Management of priorities and shares
  - Brokerage: matchmaking of jobs, data locality and resource capacities
  - Management of complex tasks
  - Job and task progress monitoring
  - Error handling and recovery
- Each experiment has its own framework with slightly different features: we will focus on ATLAS PanDA (http://news.pandawms.org/ developed mainly by BNL and UTA together with many international partners)

(ATLAS)

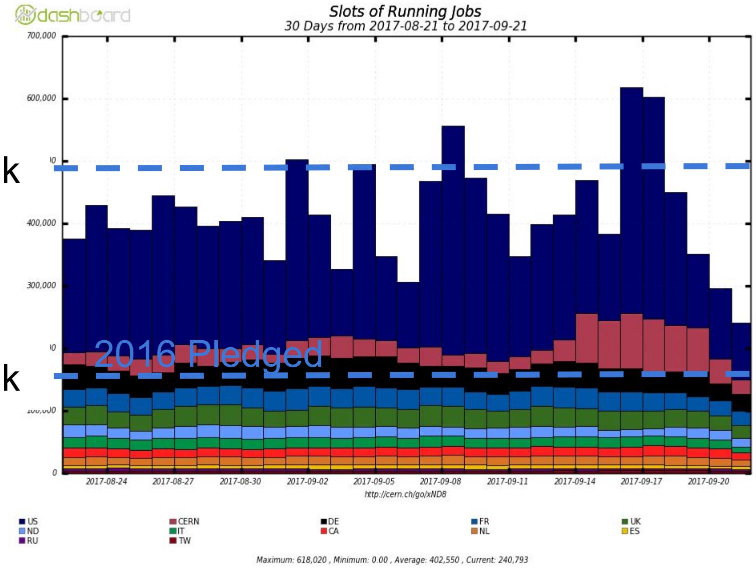
DIRAC (LHCb)

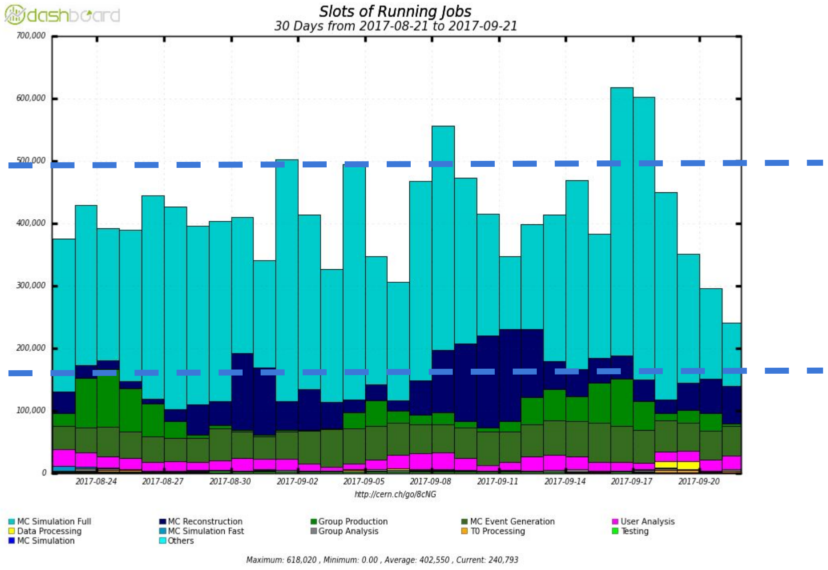
AliEn (Alice)

ATLAS Distributed Workload Management: PanDA

500k

150k



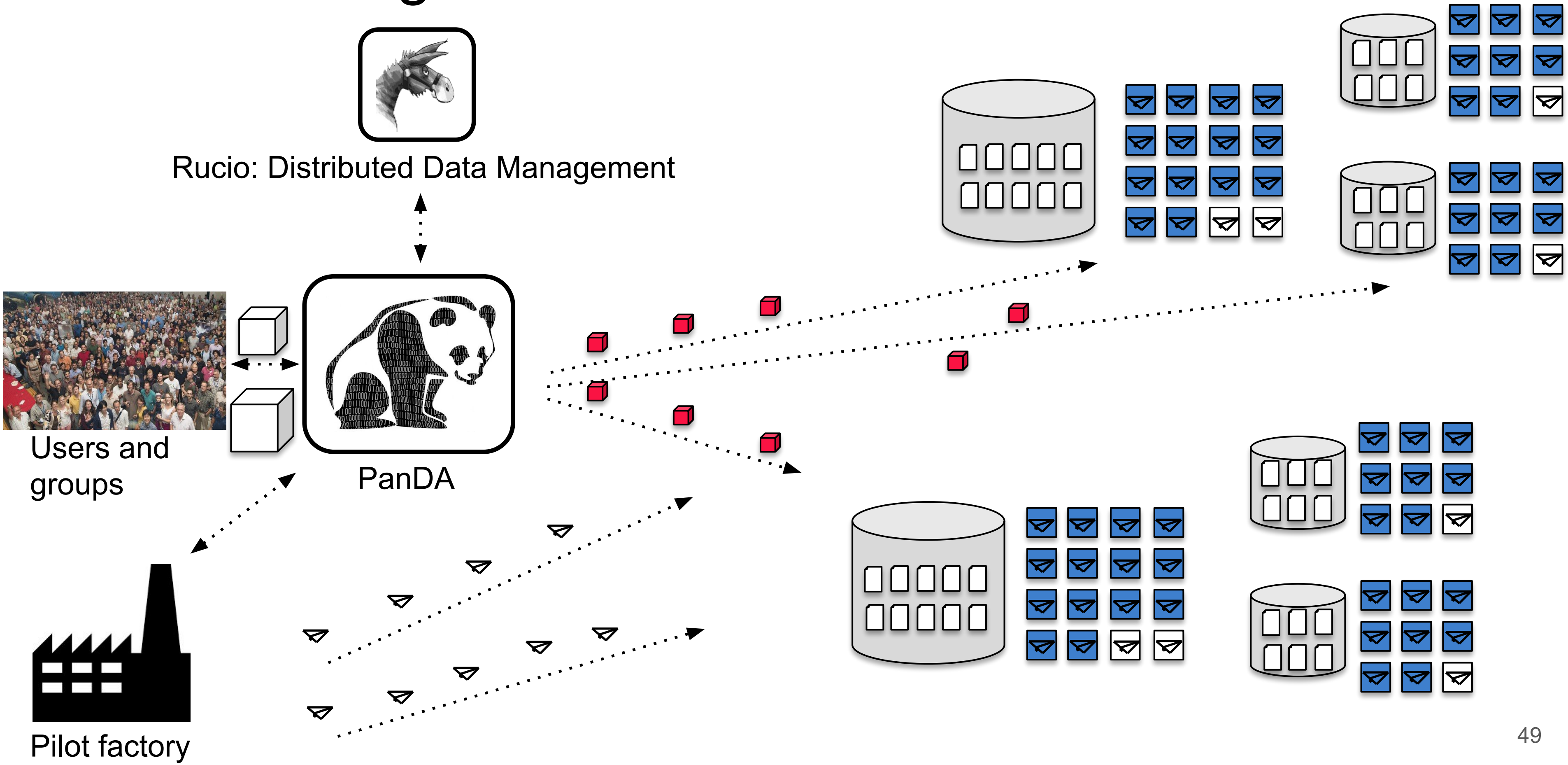


- Full grid utilization
- > 1M successful jobs per day
- Resources on T1 and T2 sites are exploited beyond pledge (200% for T2s)
- Various types of resources: grid, cloud and HPCs

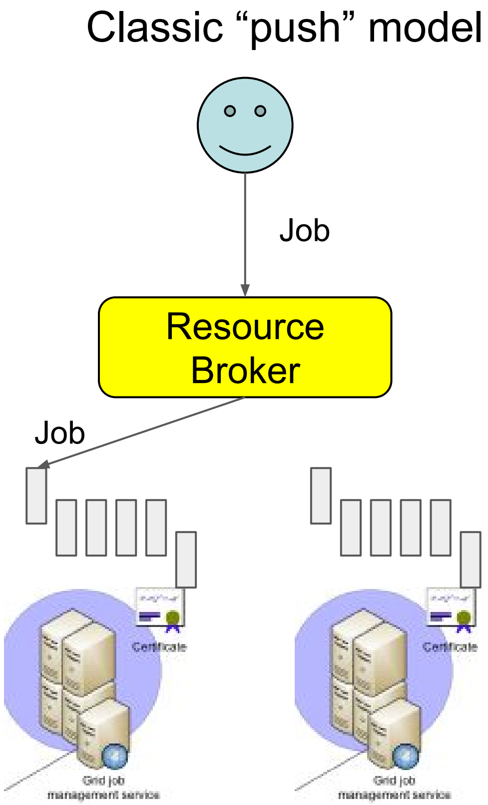
Core ideas in PanDA

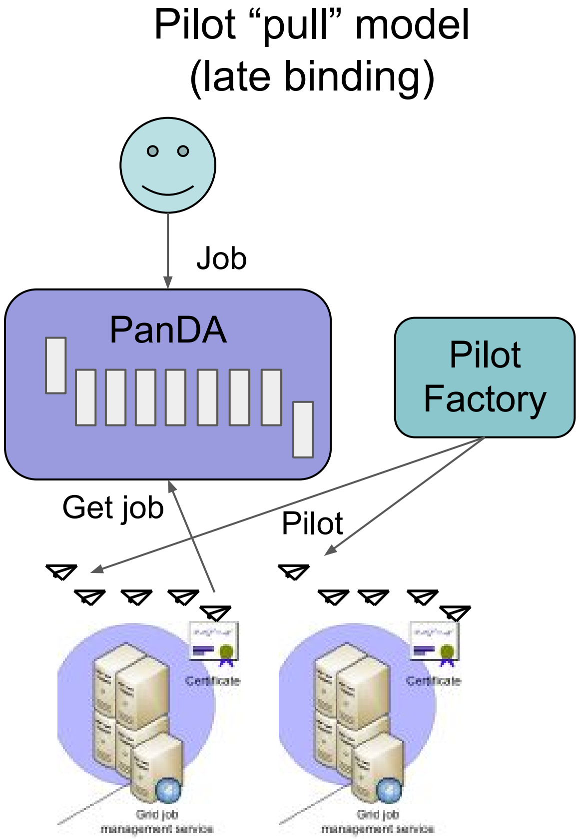
- Single entry point to the WLCG
  - Provide a central queue for users – similar to local batch systems
  - Make hundreds of distributed sites appear as local
  - Reduce site related errors and reduce latency
- Build a pilot job system – late transfer of user payloads
  - Crucial for distributed infrastructure maintained by local experts
- Hide middleware while supporting diversity and evolution
  - PanDA interacts with middleware – users see high level workflow
  - Hide variations in infrastructure
- PanDA presents uniform 'job' slots to user (with minimal sub-types)
  - Easy to integrate grid sites, cloud computing resources, HPC sites …
  - Same set of distributed resources available to all users
  - Highly flexible system, giving full control of priorities to experiment

PanDA at a glance

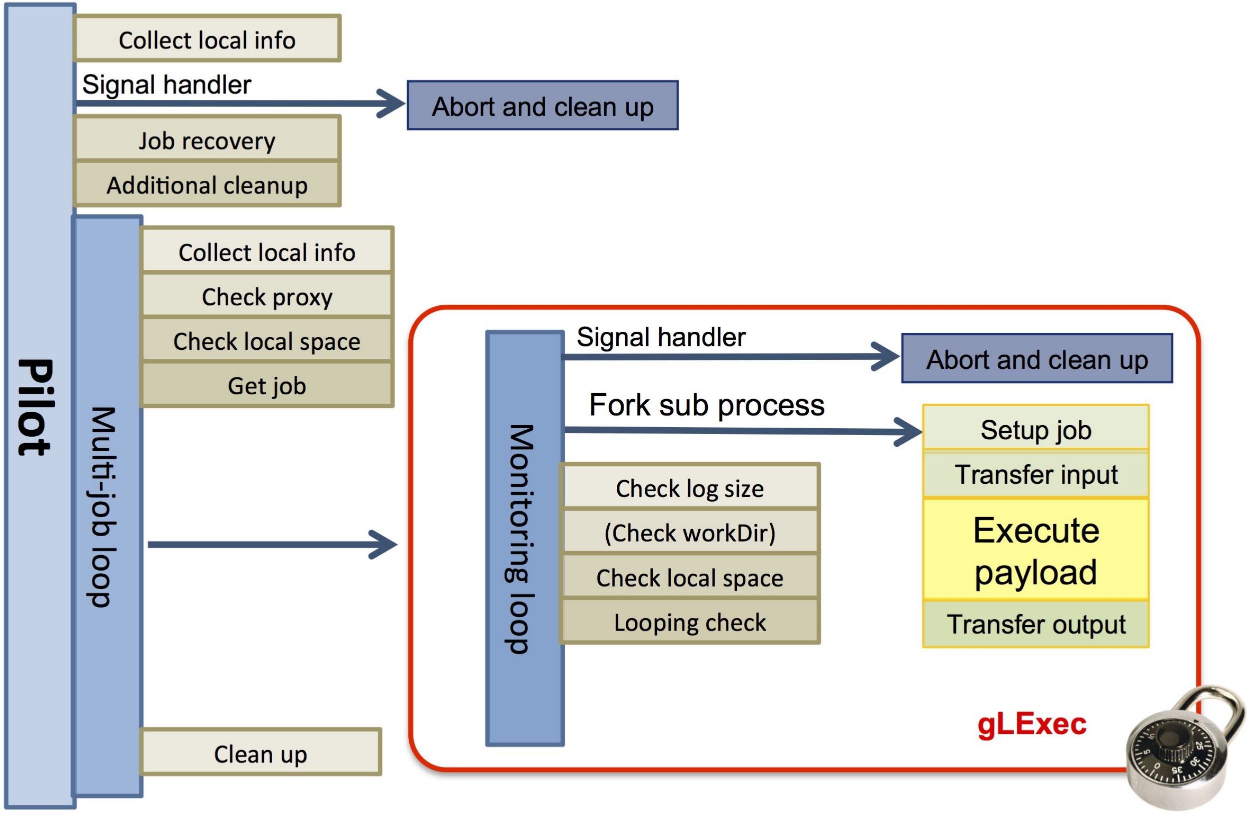


Grid job management





On the worker node: PanDA pilot

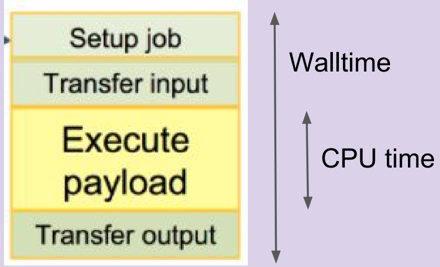


Important job concepts and metrics

- RAM consumption: max and avg
  - Swap
  - RSS
  - VMEM

- I/O intensity: input bytes/walltime

- Timings
  - Walltime
  - Setup
  - Stage in
  - Stage out

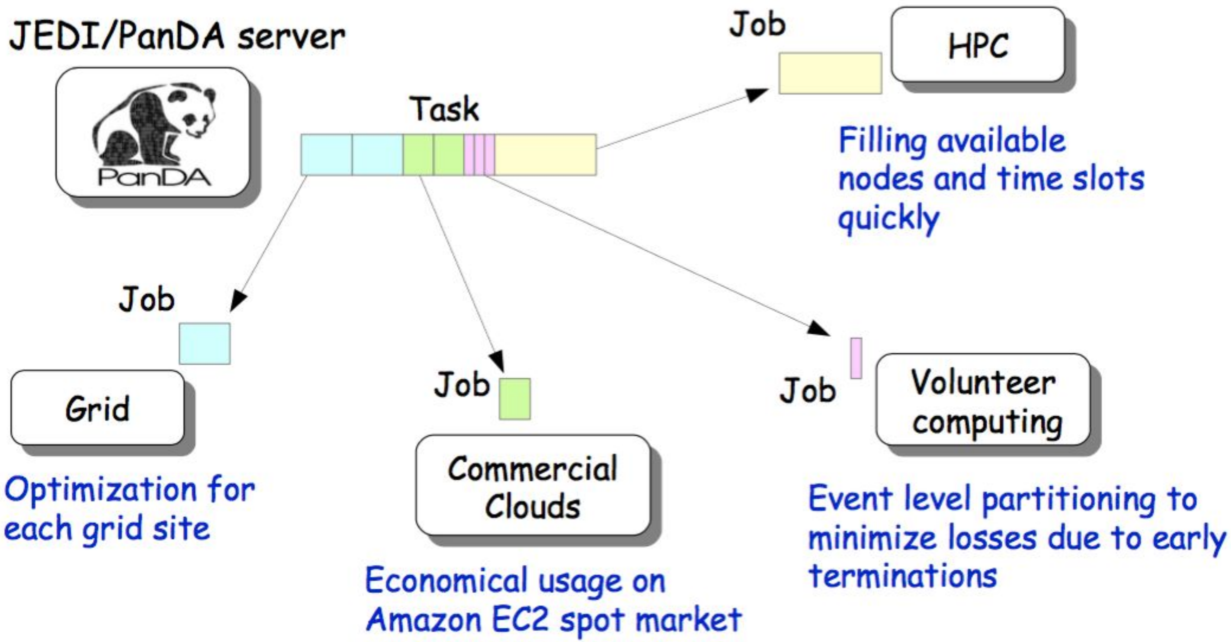


- Efficiency: CPU time / Walltime

- CPU:
  - SCORE
  - MCORE: 1, 2, 4, 8, 16, 32…
  - Corepower

Dynamic job definition

- Dynamically split workload for optimal usage of resources
- Manages workload at task, job, file and event level



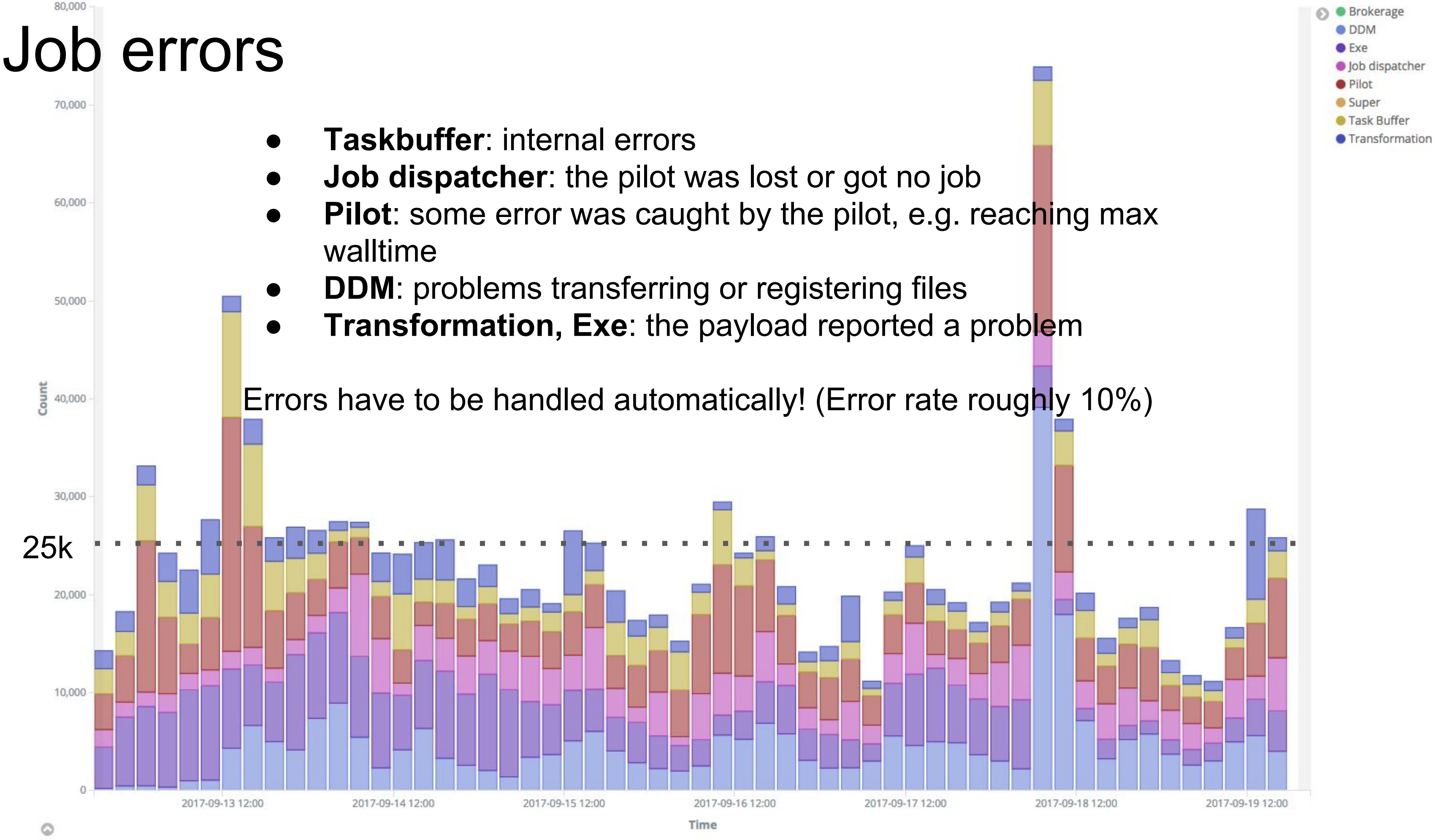
Job brokerage

- Matchmaking between the requirements of a job and what the worker node can handle
  - Max walltime: maximum time the worker node can be occupied
  - Memory usage: maximum memory the job can use
  - Disk: available storage on the worker node and the storage element
- Traditionally jobs go to data
  - It is expensive to move the input of each job around, particularly for jobs with high IO intensity
  - However data has to be replicated sometimes to avoid empty CPUs
- All sites need to be kept permanently full
  - A good ratio is to have `#queued = 2 * #running`
  - Bursty sites are more difficult to handle

Job errors

- **Taskbuffer**: internal errors
- **Job dispatcher**: the pilot was lost or got no job
- **Pilot**: some error was caught by the pilot, e.g. reaching max walltime
- **DDM**: problems transferring or registering files
- **Transformation, Exe**: the payload reported a problem

Errors have to be handled automatically! (Error rate roughly 10%)



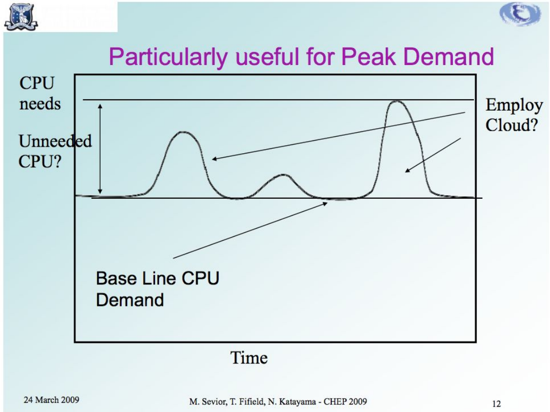
Task and job parameter auto-tuning

- Task and job parameters are tuned automatically
- **Scout jobs** collect real job metrics like memory and walltime
  - 10 scout jobs are generated at the beginning of each task
  - Parameters for successive jobs in the task are optimized based on these metrics
- **Retrial module** acts on failed jobs
  - Extending memory and walltime requirements for related types of errors
  - Preventing jobs with irrecoverable errors - don’t waste CPU time retrying jobs that will never succeed
  - Rules for error codes and actions are configurable through ProdSys User Interface

PanDA and upcoming computing paradigms

It is not about replacing the WLCG, but about integrating additional computing resources

Overspilling into the **CLOUD**



Future Grid

HELIX NEBULA THESCIENCECLOUD

Google Compute Engine

rackspace

amazon

Backfilling and allocations on **HPC**

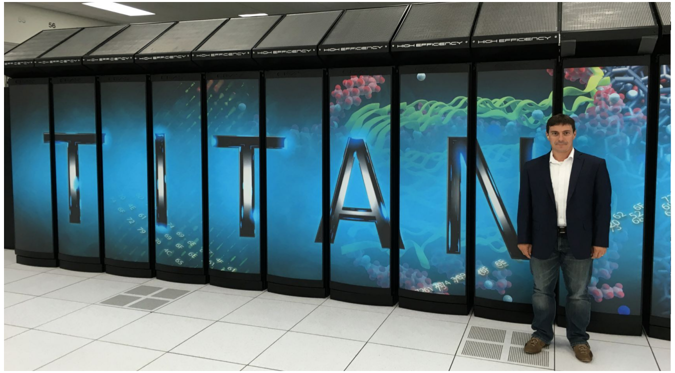
Oak Ridge National Laboratory

NERSC

Monte Carlo jobs as ideal candidates for external compute

Opportunistic resources

- Centers willing to contribute to ATLAS, but not part of WLCG
  - HPC centers
  - Shared academic clusters
  - Academic and commercial Clouds
  - Volunteer computing
- Reconfiguration of ATLAS online cluster
- Some of these centers have more computing power than the WLCG altogether
  - Even a backfill of leftover cycles (no dedicated allocation) is extremely interesting for us
- Need to adapt our systems to be able to fully exploit these offers

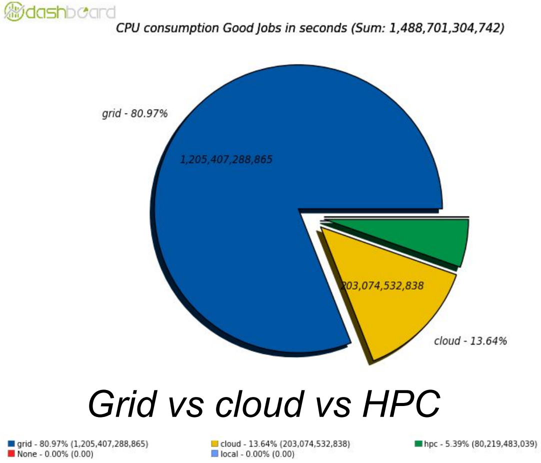


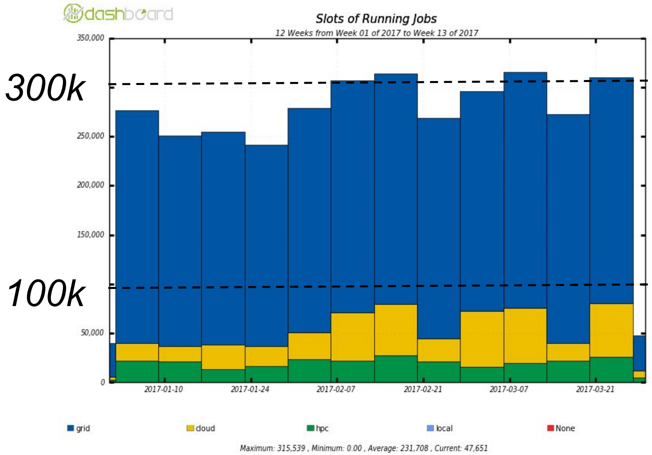
Google Compute Engine

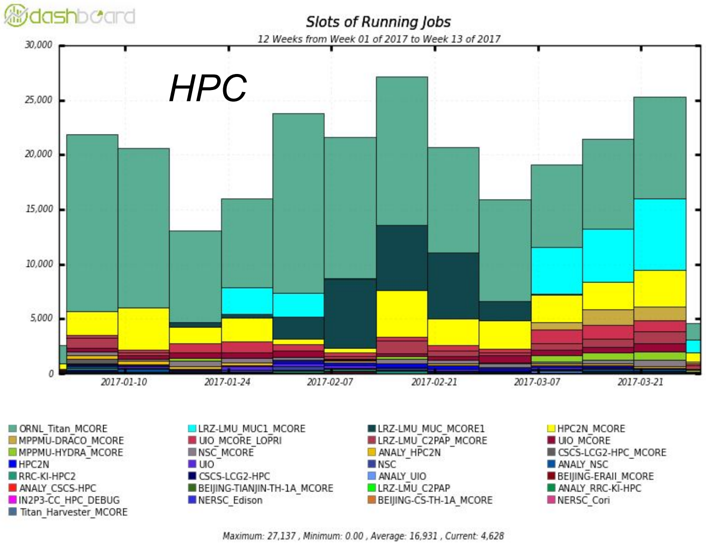
amazon web services EC2

openstack

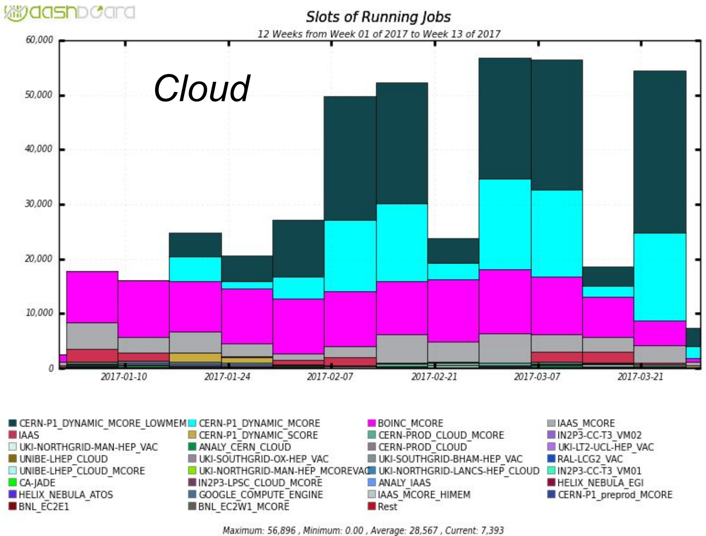
Opportunistic resources







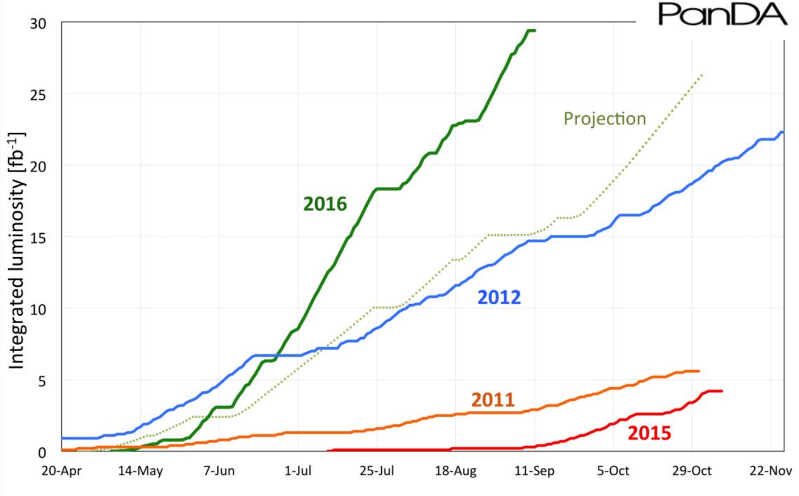
Major HPC contributor is Titan running on purely backfill mode. Constraints on tasks it can run and still a lot of backfill to exploit further



Beautiful example of how online farm is re-configured to run Grid jobs when idle. Also important, steady contribution from ATLAS@Home

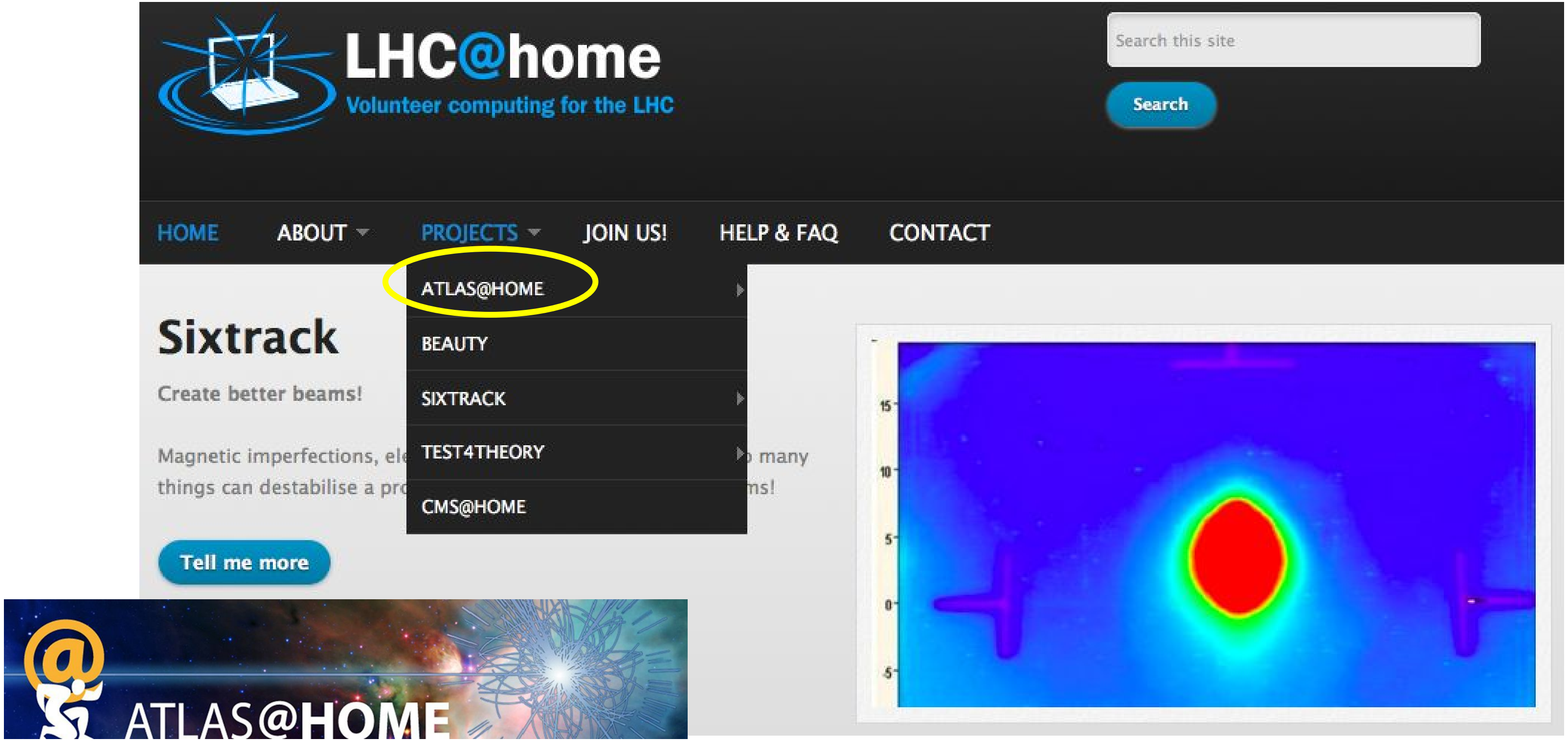
Tier-0 processing

- Tier-0 facility is a powerful cluster designed to cope with the data processing needs
  - Powerful worker nodes: SSD, 4GB/core
- Switch from T0 data processing to grid workflows during periods without data



BOINC: volunteer computing

- How YOU can help LHC experiments including ATLAS! http://lhcathome.web.cern.ch/
- Run simulation of collisions inside the ATLAS detector at home

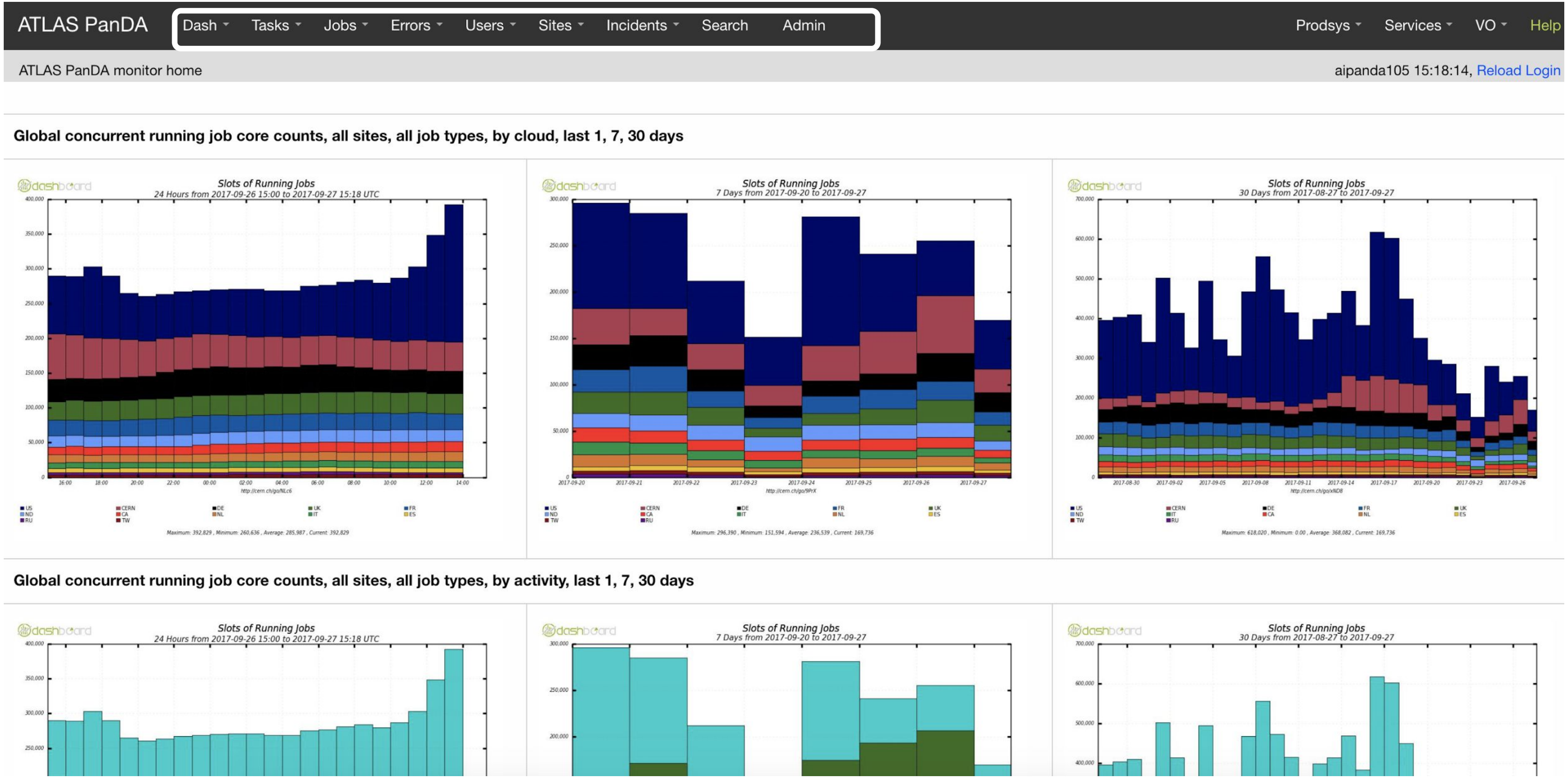


Who are the volunteers and why participate?

Rank	Name	Recent average credit	Total credit	Country
1	hartmut	727,050	269,722,681	International
2	Wenjing Wu	417,780	18,740,187	China
3	dthonon	203,274	33,547,098	France
4	MPI für Physik	171,104	89,747,772	Germany
5	Gunde	165,803	40,793,783	Sweden
6	Toby Broom	113,094	78,883,547	Switzerland
7	USTL-FIL (Lille Fr)	98,828	72,062,655	France
8	WLCG Performance-Test Cluster	98,301	54,282,843	Switzerland
9	grcpool.com-3	80,462	1,643,439	International
10	jaibenz	74,708	7,230,398	International

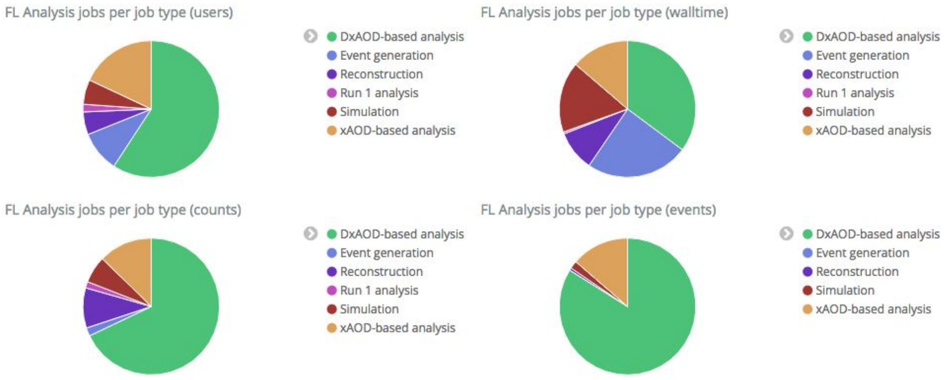
- It's fun
- Strong community, many volunteers and credits given for each processed event
  - No monetary value, but personal satisfaction in contributing to leading science
- 1% of the compute resources

Monitoring: BigPanDA



PanDA Analytics

- Job data and logs is streamed to ElasticSearch
  - Facilitates analytics and easy aggregation and filters
- Example: Identify incoherent user behaviour, such as individual users occupying non-negligible amounts of resources, can be easily identified



produsername: Descending	Sum of Walltime times core	Count	Sum of inputfilebytes	Sum of outputfilebytes
[redacted]	109 years	89,555	4.6TB	5.7TB
[redacted]	81 years	3,109,997	6.6PB	14.6TB
[redacted]	44 years	1,258,790	6TB	518.6GB
[redacted]	31 years	156,144	458.9TB	5TB
[redacted]	25 years	9,907	24.1TB	242.6GB
[redacted]	23 years	47,278	20.8TB	4TB
[redacted]	19 years	603,890	1.7PB	2TB
[redacted]	19 years	160,732	431TB	3.7TB
[redacted]	17 years	148,551	84.6TB	611.9GB
[redacted]	16 years	452,578	701.4TB	6.7TB

Wrapping up

Conclusions

- LHC experiments are data and compute intensive
- Compute is done on distributed computing resources: the WLCG
- Resources are managed centrally: data and workload management
  - We’ve learned many important concepts about both
- LHC needs keep increasing and resources are spare! We need to be creative and extend the WLCG to any opportunistic resources we can get: Supercomputers, Cloud, Volunteer computing

Questions?

Reference material

- ATLAS
  - J. Catmore: From collisions to papers
  - ATLAS Resource Request for 2014 and 2015
- CERN Computing Center
  - B. Panzer: Introduction to CERN Computing Services
- ATLAS Distributed Computing (ADC)
  - T. Wenaus: Computing Overview
  - A. Filipcic: ATLAS Distributed Computing Experience and Performance During the LHC Run-2
  - C. Serfon: ATLAS Distributed Computing
- ADC Data Management
  - V. Garonne: Experiences with the new ATLAS Distributed Data Management System
  - F. Lopez: Rucio Auditor
- ADC Workload Management
  - T. Maeno: The Future of PanDA in ATLAS Distributed Computing