Lessons learned from the ATLAS performance studies of the Iberian Cloud for the first LHC running period

V. Sánchez-Martínez(1), G. Borges(2), C. Borrego(3), J. del Peso(4), M. Delfino(5,6), J. Gomes(2), S. González de la Hoz(1), A. Pacheco Pages(3,5), J. Salt(1), A. Sedov(5,3), M. Villaplana(1), H. Wolters(7)
for the ATLAS Collaboration

(1) Instituto de Física Corpuscular (IFIC), University of Valencia and CSIC, Valencia, Spain
(2) Laboratório de Instrumentação e Física Experimental de Partículas - LIP, Lisboa, Portugal.
(3) Institut de Física d'Altes Energies, Universitat Autònoma de Barcelona, Spain.
(4) Departamento de Fisica Teorica C-15, Universidad Autonoma de Madrid, Madrid, Spain
(5) Port d'Informació Científica (PIC), Campus UAB, Bellaterra, Spain.
(6) Departament de Física, Universitat Autònoma de Barcelona, Barcelona, Spain
(7) Laboratório de Instrumentação e Física Experimental de Partículas, Coimbra, Portugal.

GOBIERNO DE ESPAÑA · MINISTERIO DE ECONOMIA Y COMPETITIVIDAD

First Running Period (Run I) ⇒ March 2010 to Jan 2013
Change Computing Model ⇒ Summer 2011
Long Shut Down (LS1) ⇒ 14 Feb 2013 to Jan 2015

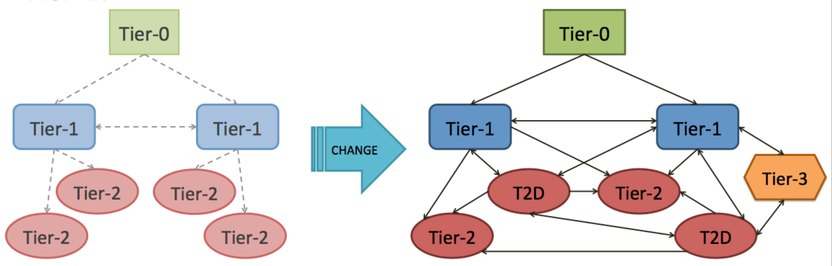
ATLAS Computing Model (ACM)

Facilities:

- Tier-0 at CERN.
- 10 Tier-1s & 78 Tier-2s distributed world wide.
- End-user private analysis facility (Tier-3).

Old ACM:
It was hierarchical, where tier-2s/3s only received data from the corresponding Tier-1.

Current ACM:
Selected Tier-2s are connected with other Tier-1 directly (T2D).



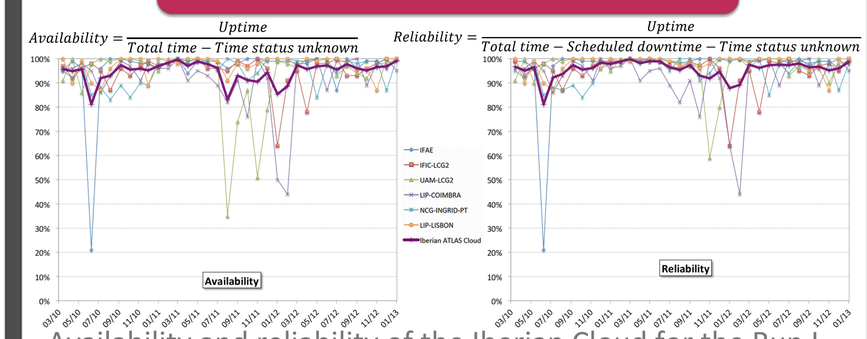
Some advantages about the new ACM are:

- It improves the usability and it reduces the dependency on Tier-1 sites.
- Sites with good availability, reliability and connectivity enter the category T2D and are candidates to receive more dataset replicas, more analysis and production jobs.
- T2D can directly exchange data with any Tier-1 and other Tier-2.

Availability & Reliability

$$Availability = \frac{Uptime}{Total\ time - Time\ status\ unknown}$$

$$Reliability = \frac{Uptime}{Total\ time - Scheduled\ downtime - Time\ status\ unknown}$$



Availability and reliability of the Iberian Cloud for the Run I. Both have always been, on average, inside the interval 90-100%.

Availability & Reliability around 90-100%

HAMMER CLOUD (HC) TESTS:

Sites are tested with typical analysis and MC production jobs. Problematic sites are identified online and automatically blacklisted. Last year Iberian Cloud had more than 90% availability according to HC.

PANDA queue	SITE Name	TIER	CLOUD	History plot From September 2011 to September 2013	offline %	offline count	brokeroff %	brokeroff count	online %	online count	NoQueue %	NoQueue count	test %	test count
ANALY_IFAE	ifae	T2D	ES		1.26	15	0.47	19	96.1	133	0	0	2.11	49
ANALY_IFIC	IFIC-LCG2	T2D	ES		1.3	12	1.35	26	93.39	208	0	0	3.91	95
ANALY_LIP-Coimbra	LIP-Coimbra	T2	ES		9.08	24	3.51	61	84.66	135	0	0	2.69	69
ANALY_LIP-Lisbon	LIP-Lisbon	T2D	ES		2.37	14	1.02	19	63.24	139	0	0	2.89	39
ANALY_NCG-INGRID-PT	NCG-INGRID-PT	T2D	ES		1.87	13	1.36	21	94.24	158	0	0	2.47	49
ANALY_PIC	pic	T1	ES		2.6	14	0.85	23	89.39	139	0	0	1.56	47
ANALY_PIC_SL6	pic	T1	ES		0.4	1	0	0	7.66	4	0	0	0.23	3
ANALY_UAM	UAM-LCG2	T2D	ES		4.87	58	10.91	49	77.57	176	0	0	6.34	97

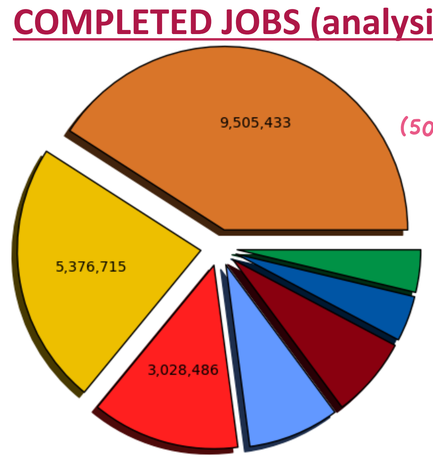
Good connectivity implies: Five Tier-2s tagged as T2D!

- Direct transfers from/to Tier-0 and all ATLAS Tier-1s.
- Direct transfers to Tier-2 from different clouds.
- Process data from many Tier-1s.

Analysis and Production Jobs

COMPLETED JOBS (analysis + production):

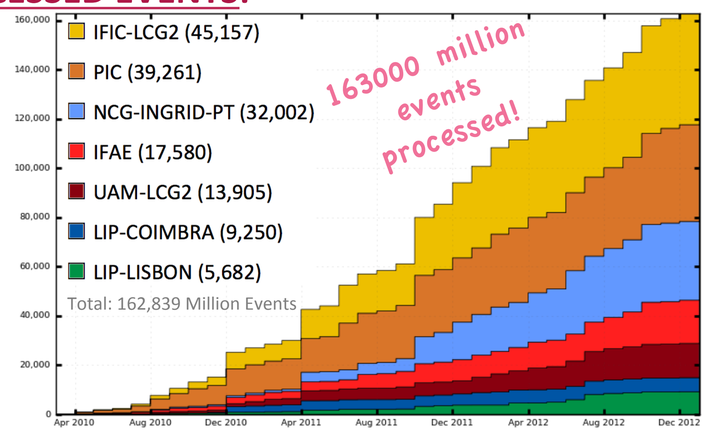
More than 23 million jobs completed! (50% Production and 50% Analysis)



- PIC – 40.91% (9,505,433)
- IFIC-LCG2 – 23.14% (5,376,715)
- IFAE – 13.04% (3,028,486)
- NCG-INGRID-PT – 8.04% (1,876,592)
- UAM-LCG2 – 7.13% (1,655,608)
- LIP-COIMBRA – 4.01% (931,066)
- LIP-LISBON – 3.74% (868,573)

Total: 23,233,473 Jobs

PROCESSED EVENTS:

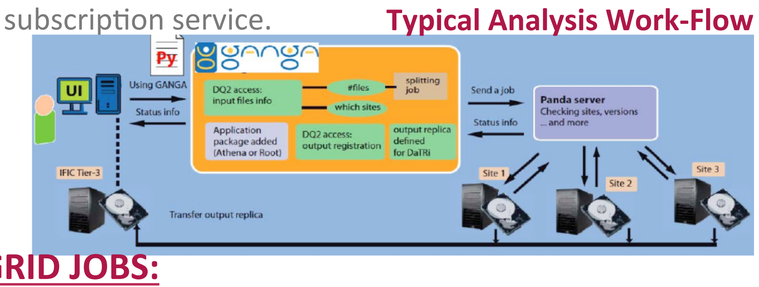


TOOLS

FOR DATA MANAGEMENT:

★ **dq2** (Don Quijote 2): to download and obtain information about data and to register files on GRID.
★ **AMI** (ATLAS Metadata Interface): web page to monitor datasets, releases, number of events…
★ **DaTRi** (Data Transfer Request): end-user dataset subscription service.

Typical Analysis Work-Flow



FOR GRID JOBS:

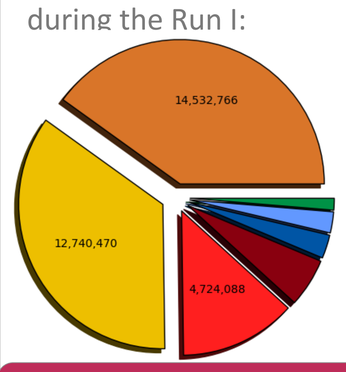
❖ **ganga** (Gaudi/Athena and Grid Alliance): job definition management tool for local, batch system and the GRID.
❖ **PanDA client** (Production and Distributed Analysis): analysis job submission tool.
◆ **pathena** (Panda Athena): client tool to submit user-defined jobs to distributed analysis systems.
◆ **prun** (Panda Run): panda-client software to submit general jobs to panda.
◆ **pbook** (Panda Bookkeeping): bookkeeping application.

SPACE TOKENS:

In ATLAS, the storage is organized using **SRM Space Tokens**, which are controlled through the ATLAS Distributed Data Management (DDM) system and they are associated to a path to a GRID Storage Element.

IBERIAN CLOUD Tier-1 AND Tier-2s:

The following graph shows the number of bytes processed during the Run I:



- PIC – 40.08% (14,532,767 GB)
- IFIC-LCG2 – 35.14% (12,740,471 GB)
- IFAE – 13.03% (4,724,088 GB)
- UAM-LCG2 – 5.44% (1,972,627 GB)
- LIP-COIMBRA – 2.64% (955,582 GB)
- NCG-INGRID-PT – 2.36% (855,760 GB)
- LIP-LISBON – 1.32% (476,967 GB)

Total: 36,258,261 GB

36 PB Processed!

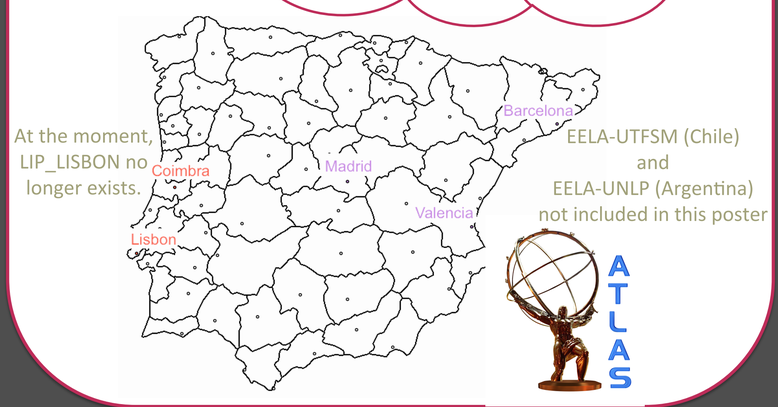
Data distribution

Iberian ATLAS Cloud

ES-PIC (Tier1)
PIC (Barcelona)

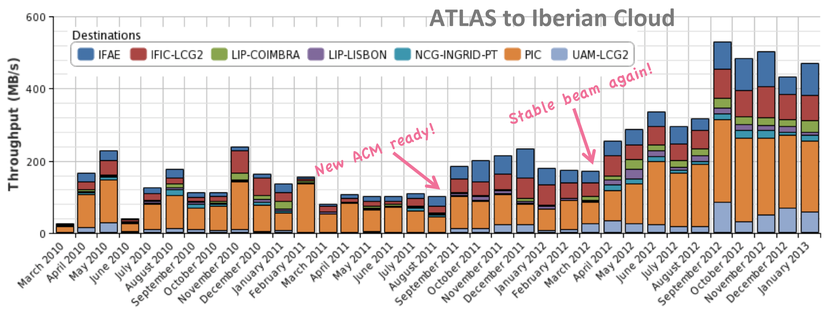
ES-ATLAS-T2
(T2 Federation)
50% IFIC (Valencia)
25% IFAE (Barcelona)
25% UAM (Madrid)

PT-LIP-LCG-Tier2
(T2 Federation)
50% LIP-Coimbra (Coimbra)
50% NCG-Ingrid (Lisbon)



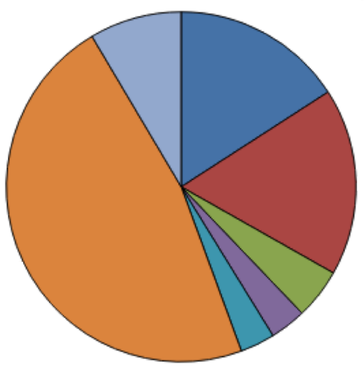
TRANSFER THROUGHPUT:

During the Run I, the transfer throughput reached 550 MB/s in September 2012



DATA TRANSFER VOLUME:

The data is transferred over the tiers after the reprocessing and after the ATLAS official production.



20 PB of data transferred over the Run I!

- PIC – 47% (9339.497 TB)
- IFIC-LCG2 – 17% (3435.064 TB)
- IFAE – 16% (3150.623 TB)
- UAM-LCG2 – 9% (1690.460 TB)
- LIP-COIMBRA – 5% (925.598 TB)
- NCG-INGRID-PT – 3% (634.874 TB)
- LIP-LISBON – 3% (661.595 TB)

Total: 19,837.7 TB

Files transferred

RESOURCES

The installed capacities at the Iberian Cloud sites and their corresponding pledges are summarized in this table. Currently, the hardware provided by the sites is roughly the pledges.

Federation	CPU (HEP-SPEC06)		DISK (TB)	
	Pledges 2013	Current	Pledges 2013	Current
ES-ATLAS-T2	18000	17800	2800	2558,3
ES-PIC	16269	16269	1785	1812,0
PT-LIP-LCG-Tier2	3200	3200	220	183,0

Current = September 2013

The evolution of the pledges in CPU, Disk and Tape for the Iberian Cloud Tier-1 and Tier-2 is shown in the following tables:

Iberian Cloud's Tier-1	2010	2011	2012	2013	2014	2015
CPU (HEP-SPEC06)	9145	11774	13209	16269	18105	24378
DISK (TB)	1043	1292	1377	1785	1683	2397
TAPE (TB)	681	1568	1836	2193	2244	3774

Iberian Cloud's Tier-2	2010	2011	2012	2013	2014	2015
CPU (HEP-SPEC06)	13508	17100	16500	21200	23800	26100
DISK (TB)	1307	2100	2570	2470	3020	3250

Iberian ATLAS Cloud plans during LS1:

☑ **Deployment of the FAX federation:**
FAX (Federated ATLAS Xrootd) will be used to data access via a single entrance (using Xrootd's redirection tech), to read a dataset directly from WAN, to bring data to local Tier 3 Xrootd disk (storage cache) and to Users sharing non-DDM data between sites.

☑ **Development of the EventIndex subproject:**
It is a complete catalogue of ATLAS events. Its contents will be, for instance: event identifiers, online trigger pattern & hit counts and references (pointers) to the events at each processing stage (RAW, ESD, AOD, NTUP) in all permanent les on storage. The main motivations for this new project were that EventTAG was designed and developed long time ago and the implementation in Oracle is an intensive labour and expensive. For that reason database technologies based on NoSQL seem well adapted to this type of application.

☑ **Operation of the data reduction framework:**
Requirements: to develop a simple mechanism to control the addition of "user data" to the new persistent format, to prepare the repository and operational procedures for collection and operation of reduction-framework tools and to prepare the reduction framework for inclusion of "smart slimming"

Long Shutdown 1