The Spanish ATLAS Tier-2 and the perspectives in the LHC Computing

Santiago González de la Hoz and José Francisco Salt Cairols

Instituto de Física Corpuscular (Centro Mixto CSIC- Universitat de València)

2nd Red LHC Workshop
09 - 11 May 2018 @ CIEMAT (Madrid)

10th May 2018

EXCELENCIA SEVERO OCHOA

Contents:

1.- Introduction

2.- Evolution of the ATLAS Computing Model.

3.- The Spanish ATLAS Tier-2

4- The Tier-2 project towards the HL-LHC period

5.- Perspectives and conclusions

1.-INTRODUCTION

the ATLAS Experiment
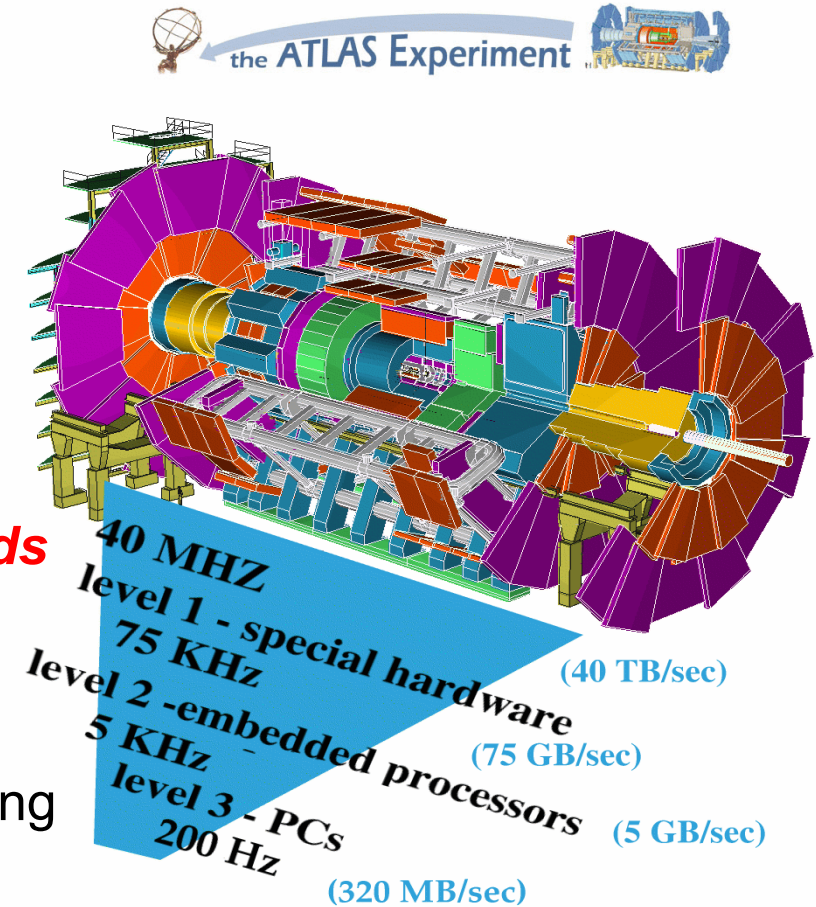
=> To take profit of these RED LHC Workshops to have *feedback from the Spanish Community participating in LHC* experiments (experimentalists & theorists)

⇒ Main argument is the discussion about Physics Analysis and the means to obtain the results is the Computing environment Provided by the LHC collaborations,

⇒ Evolution of the Computing Model is going very fast and the Computing teams of the LHC experiments need to know ***the trends and concerns*** of the physicits:

* new methods to be applied which need to revisit the present environment
* performance of the Computing Infrastructure in an fast changing period

- The ***Spanish ATLAS Tier-2 project*** is a it's an almost totally technological project…
- But it is a fundamental part of the IFIC participation in ATLAS experiment ***to ease the use of the computing infrastructures to undertake the physics Analysis***

40 MHZ
level 1 - special hardware (40 TB/sec)
75 KHz
level 2 - embedded processors (75 GB/sec)
5 KHz
level 3 - PCs (5 GB/sec)
200 Hz
(320 MB/sec)

WLCG
Worldwide LHC Computing Grid

ATLAS Computing in numbers
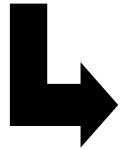
The offline computing:
- Output event rate: 200 Hz ~ 10^9 **events/year**
- Average event size (raw data): **1.6 MB/event**

Processing:
- order of 100k of today's fastest PCs

Storage:
- **Raw data recording rate 320 MB/sec**
- **Accumulating at 5-8 PB/year**

A solution: Grid technologies

Worldwide LHC Computing Grid (WLCG)

ATLAS Data Management

ATLAS MC Production

José Salt, 2nd RED LHC Workshop, 9-11 May 2018

eGee
Enabling Grids
for E-sciencE

Scheduled = 25002
Running = 12614

To analyze all this data, 100,000 current PCs and a huge storage capacacity are needed. And a excellent network

15:38:56 UTC

GridPP
UK Computing for Particle Physics

- CERN provides 30% of the computing power for LHC: 30,000 CPUs.
- 5 million Gbits of disk storage.
- 16 million Gbits of magnetic tape storage.
- Connected to other centers at 10Gbits / s.
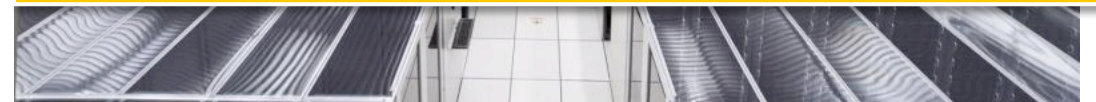
2.- Evolution of the ATLAS Computing Model.

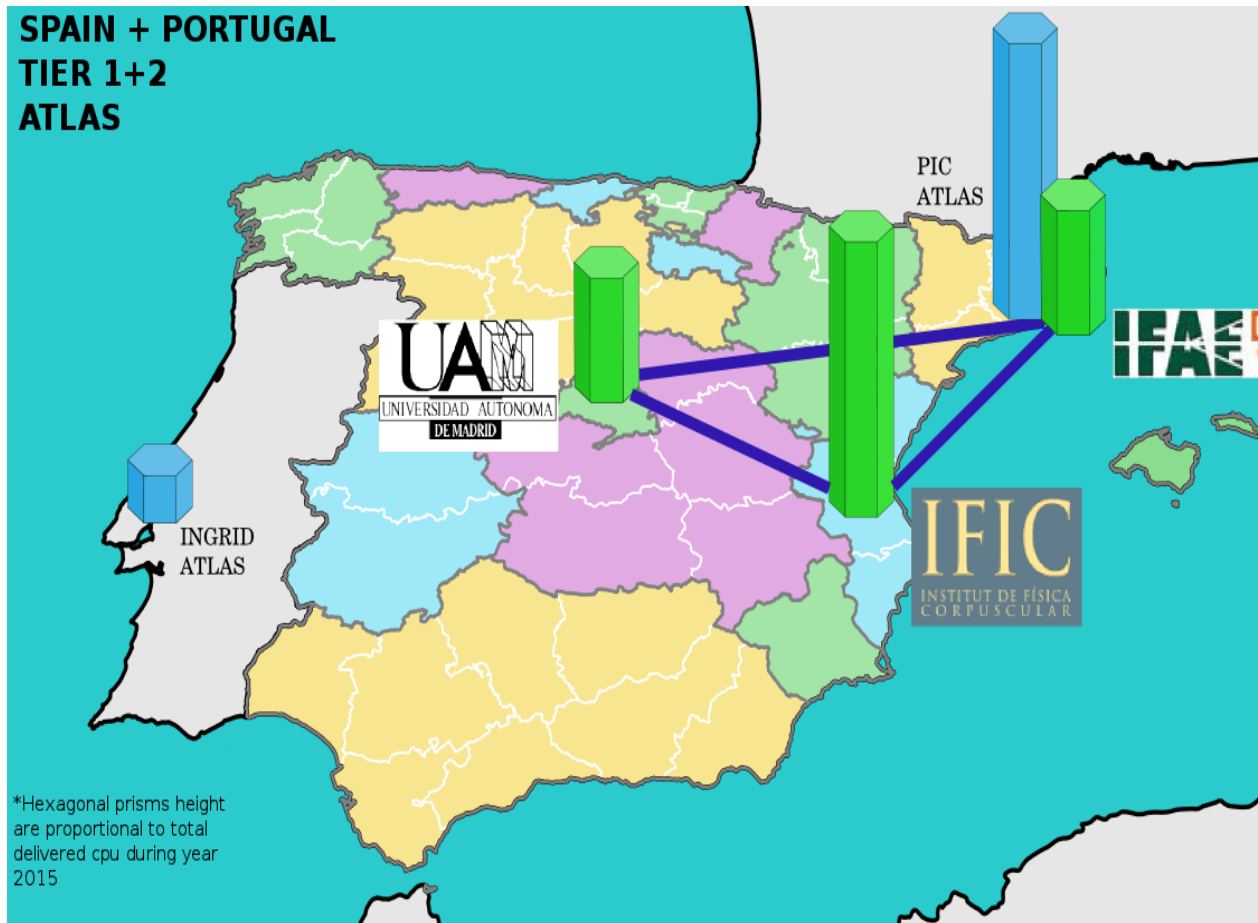Run-I (first LHC operating period): March 2010 - January 2013

T2D (Tier-2 Directly): has enough network connection so can link directly with other Tier-1s or Tier-2s belonging to different clouds.
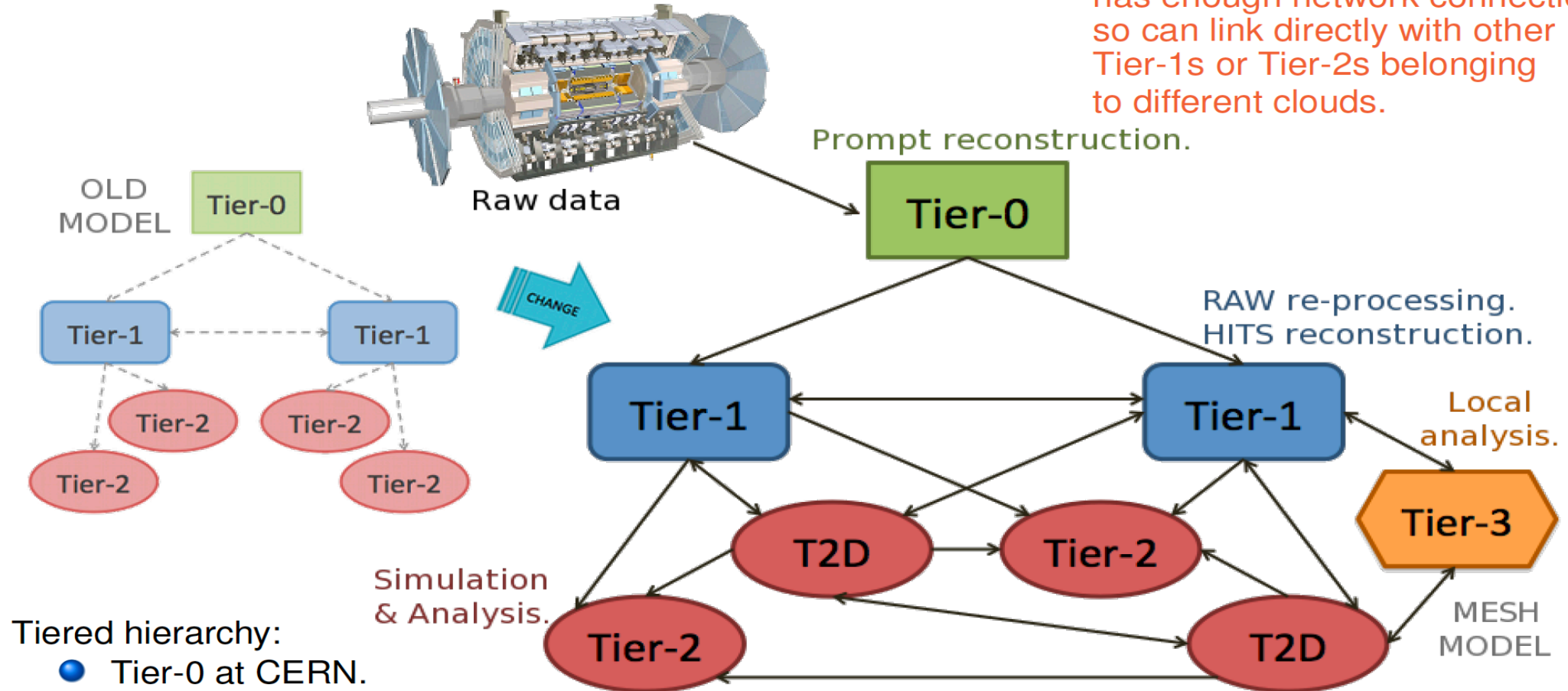
Tiered hierarchy:
- Tier-0 at CERN.
- 11 Tier-1
- ~ 80 Tier-2 centres distributed worldwide.

3.- The Spanish ATLAS Tier-2

SPAIN + PORTUGAL
TIER 1+2
ATLAS

PIC ATLAS

INGRID ATLAS

*Hexagonal prisms height are proportional to total delivered cpu during year 2015

- Federation of 3 centers:
  - IFIC (Valencia), 60%
  - IFAE(Barcelona), 25%
  - UAM(Madrid),15%
- Spanish ATLAS Tier-2 Infrastructure, started in 2005 Associated to Iberian Cloud (PIC as Tier-1)
  - 4% of all ATLAS Tier-2

Personnel and resources of the ATLAS TIER-2 Valencia
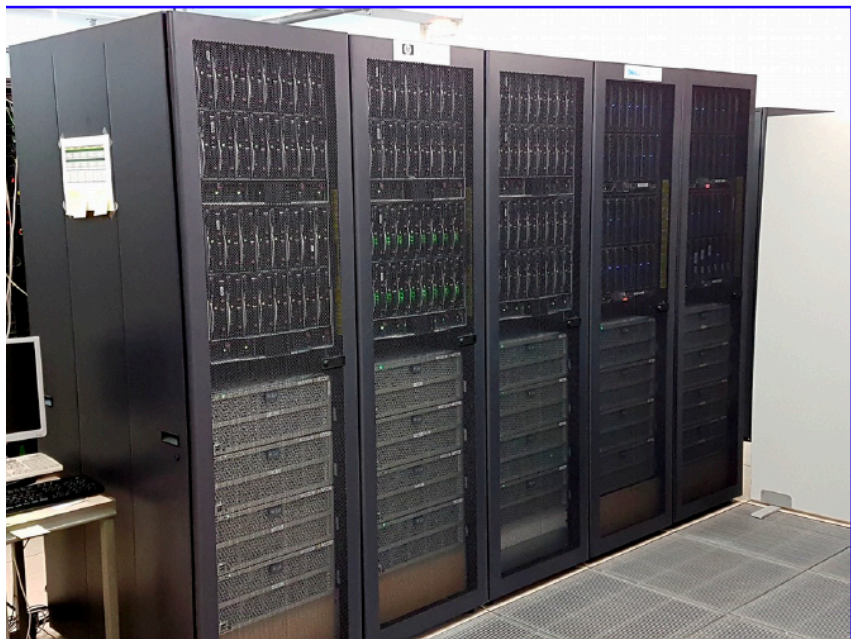
| STAFF	A. Fernández Casani S. González de la Hoz J. F. Salt Cairols J. Sánchez Martínez
STUDENTS POSTDOCS ENGINEERS	E. Fullana Torregrosa (postdoc S8A) J. Lozano (Contratado proyecto) F. Fassi (Dr. Vinculada) C. Garcia Montoro (PTA) J. Aparisi (Contratado proyecto)
COLAB. EXTERNOS	G. Amorós (AEMET) M. Vilaplana (postdoc en INFN-Milán)

	nodos	CPU	RAM	DISCO	cores	compra	>5y
ATLAS T2	32	2 x Intel Xeon E5472 @ 3.00 GHz (4 cores)	16 GB	2x300 GB SAS 15krpm	256	jun-2008	✎
	32	2 x Intel Xeon L5520 @ 2.26 GHz (4 cores)	24 GB	2x146 GB SAS 10krpm	256	mar-2010	✎
	16	2 x Intel Xeon E5-2660 @ 2.20 GHz (8-cores)	64 GB	2x300 GB SAS 15krpm	256	jul-2012	
	8	2 x Intel Xeon E5-2690 v2 @ 3.00 GHz (10-cores)	128 GB	2x300 GB SAS 15krpm	160	oct-2014	
	6	2 x Intel Xeon E5-2690 v2 @ 3.00 GHz (10-cores)	128 GB	2x300 GB SAS 15krpm	120	jul-2015	
	13	2 x Intel Xeon E5-2690 v3 @ 2.60 GHz (12 cores)	128 GB	2x300 GB SAS 15krpm	312	dic-2016	
	107				1360		

	nodos		discos	capacidad (TB)	cap tot (TB)
tier2	14	Supermicro	972	2244	2244

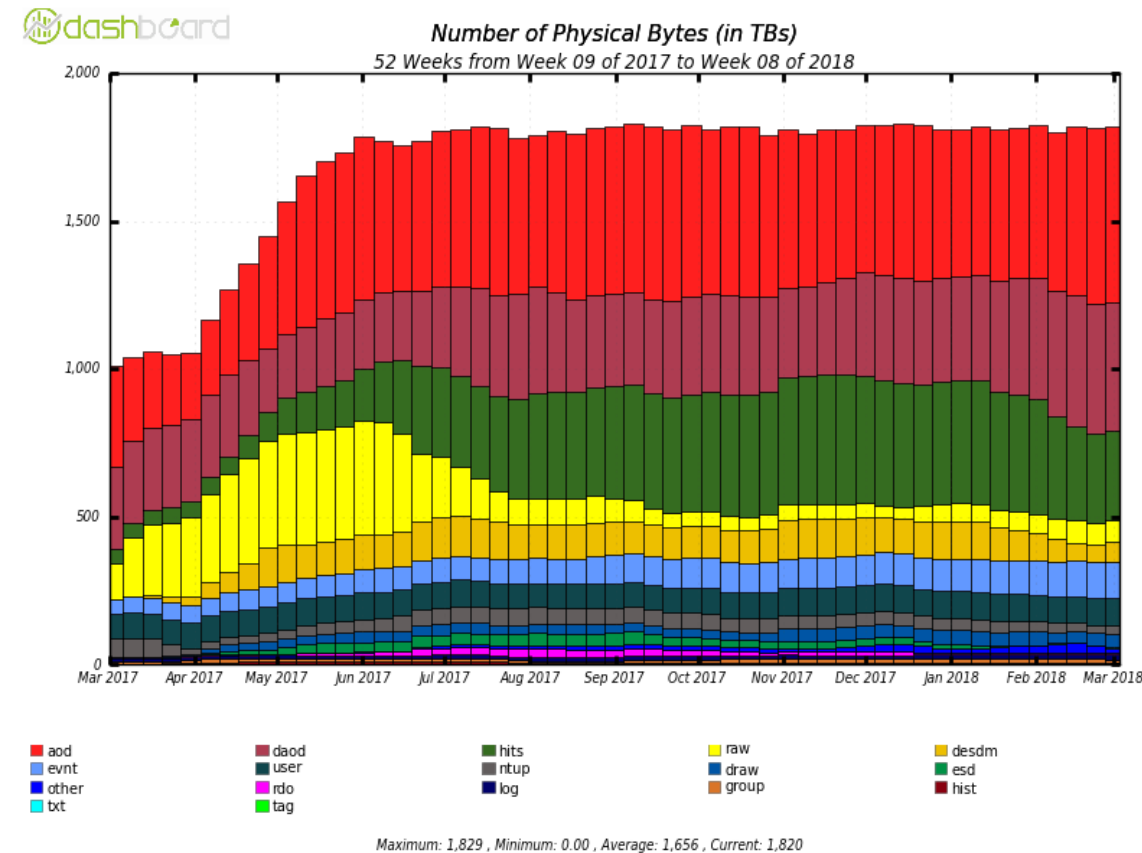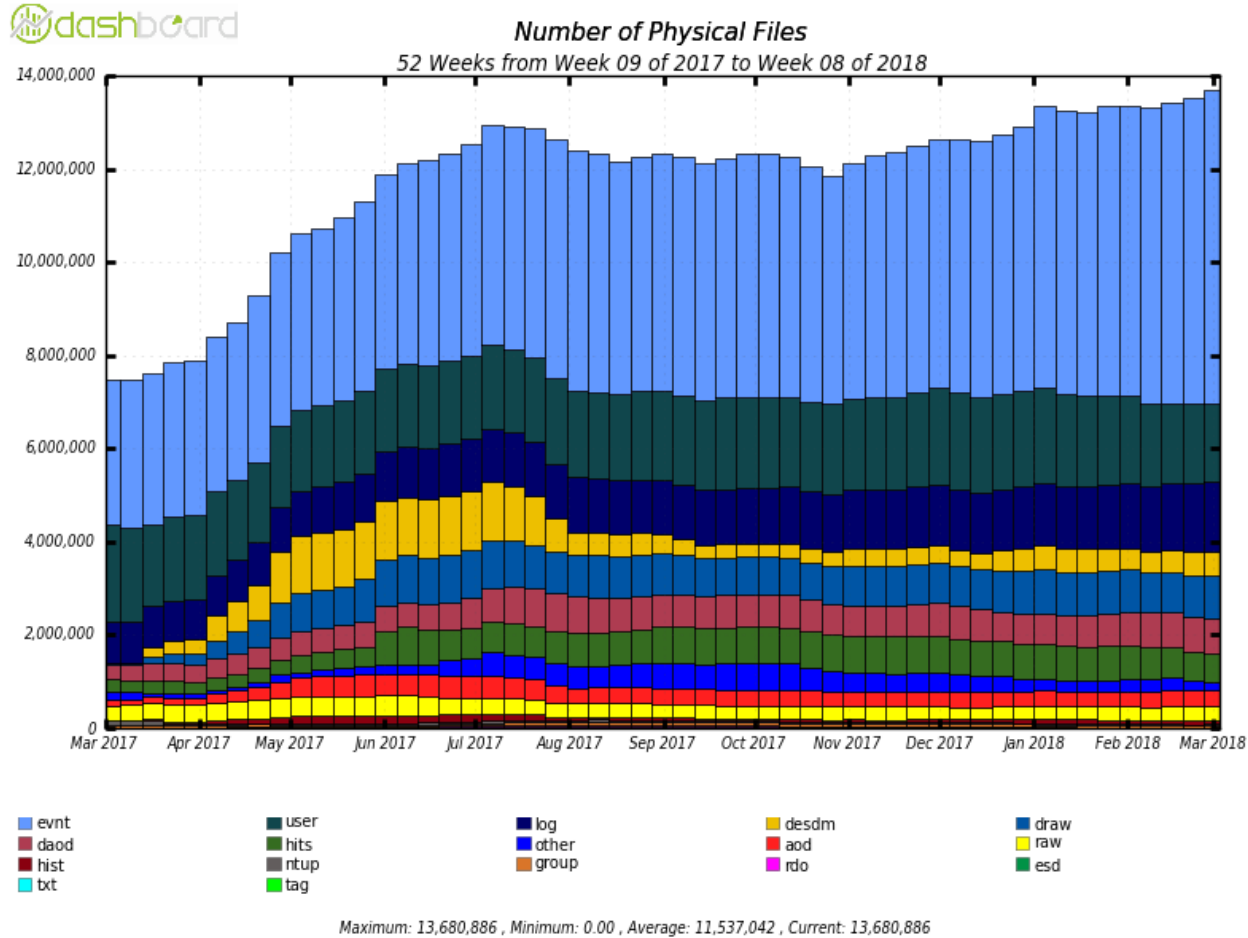# *Maintenance & Operation of a complex Infrastructure and continouosly updated*

- Transfers in the last year from the Tier2 to the rest of the World

- Transfers in the last year from the rest of the World to Tier2

  => More than 2 PetaBytes of transferred data
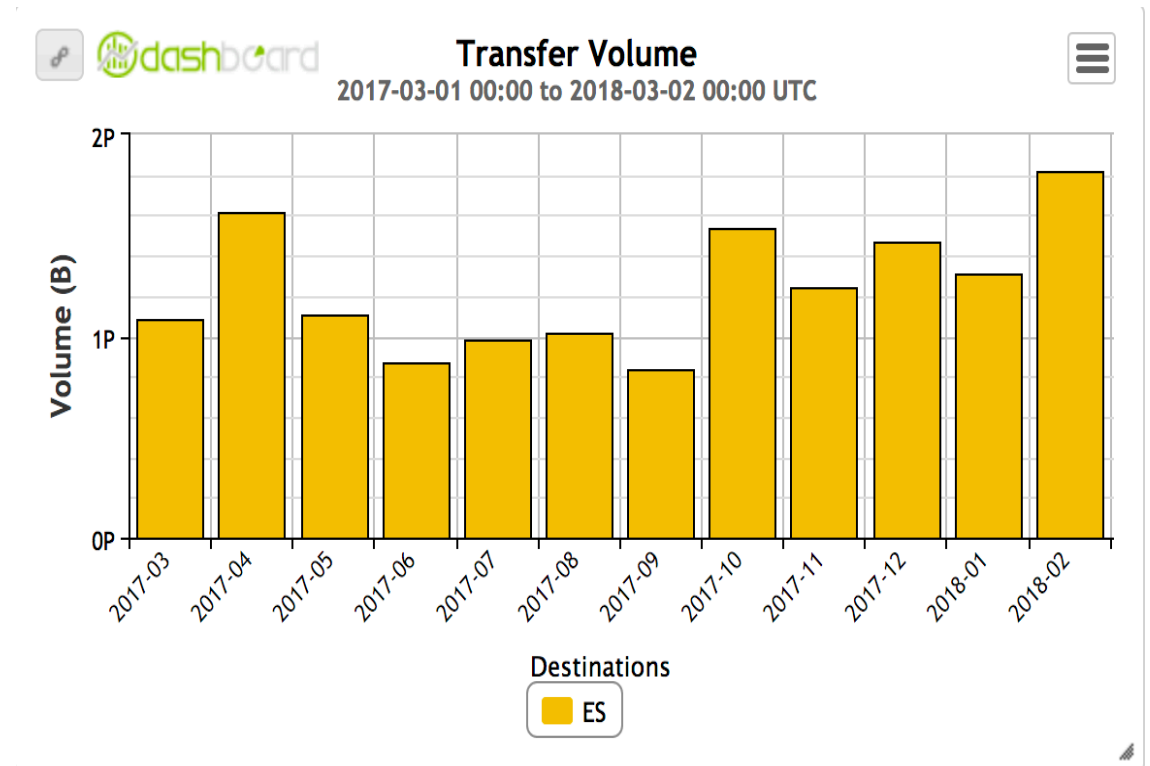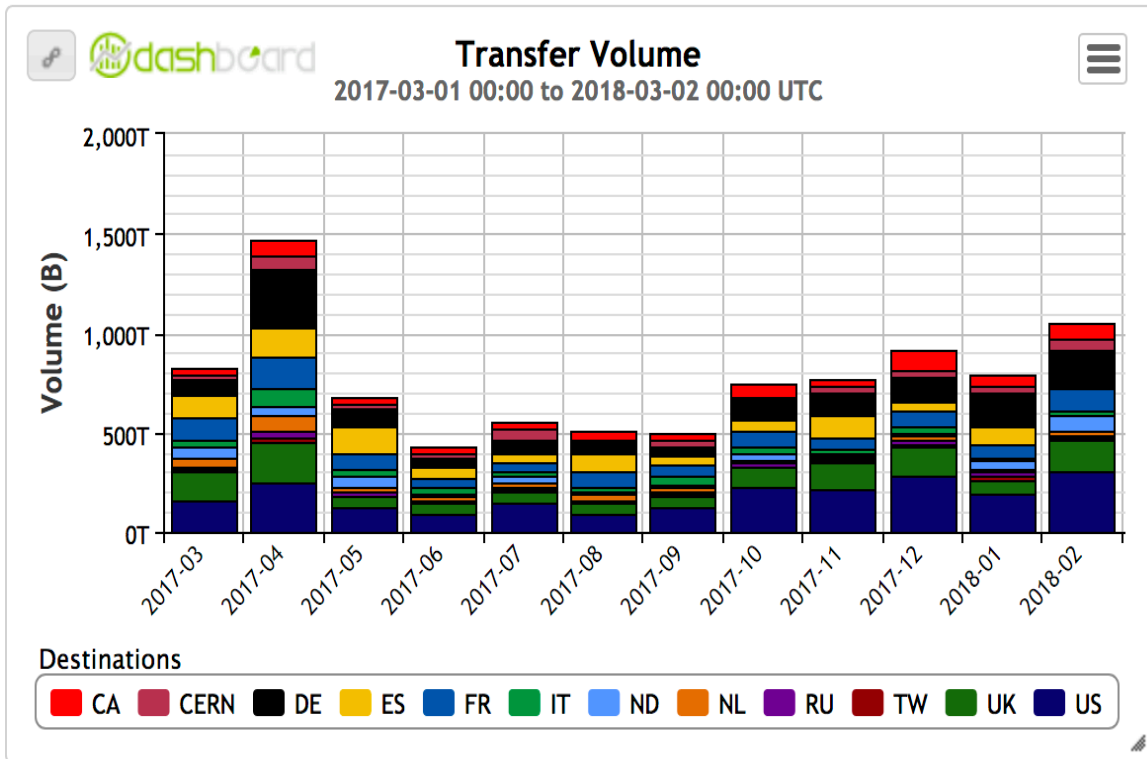
Tier-2 IFIC performance:

- Data produced at Tier-2/IFIC in the last year (More than 14M files and more than 2 PBs)

4.-The Tier-2 project towards the HL-LHC period

- Project of the National Program of Particle Physics: FPA2016, 3 years project (2017-18-19)

  - **Towards a genuine TIER-2 Spanish Federation of ATLAS to face the challenge of the management and processing of the Big Data of the LHC.**

  - ATLAS and CMS will need x20 more ressources at HL-LHC with respect to today

  - Flat budget and +20%/year from technology evolution fills part of this gap but there is still a factor x5. Storage looks like the main challenge to address

On top of the Maintenance & Operation of the Tier2 infrastructure, the project has another research activities to face with the HL-LHC challenge:
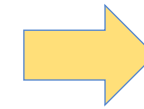
- **Distributed Computing**
- **Data Bases**
  - **=> Event index**
- **Big Data /Machine Learning**
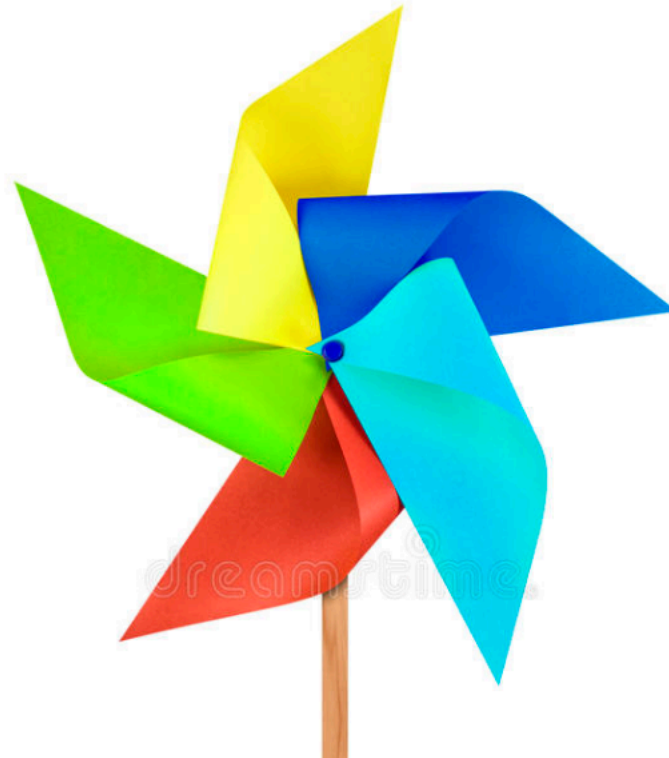- **Opportunistic Resources**
....

Cloud Computing

Computación en la nube

Supercomputing

Use of extra resources: from MereNostrum and Lusitania (RES) together we have got 300 Kh since September 2017
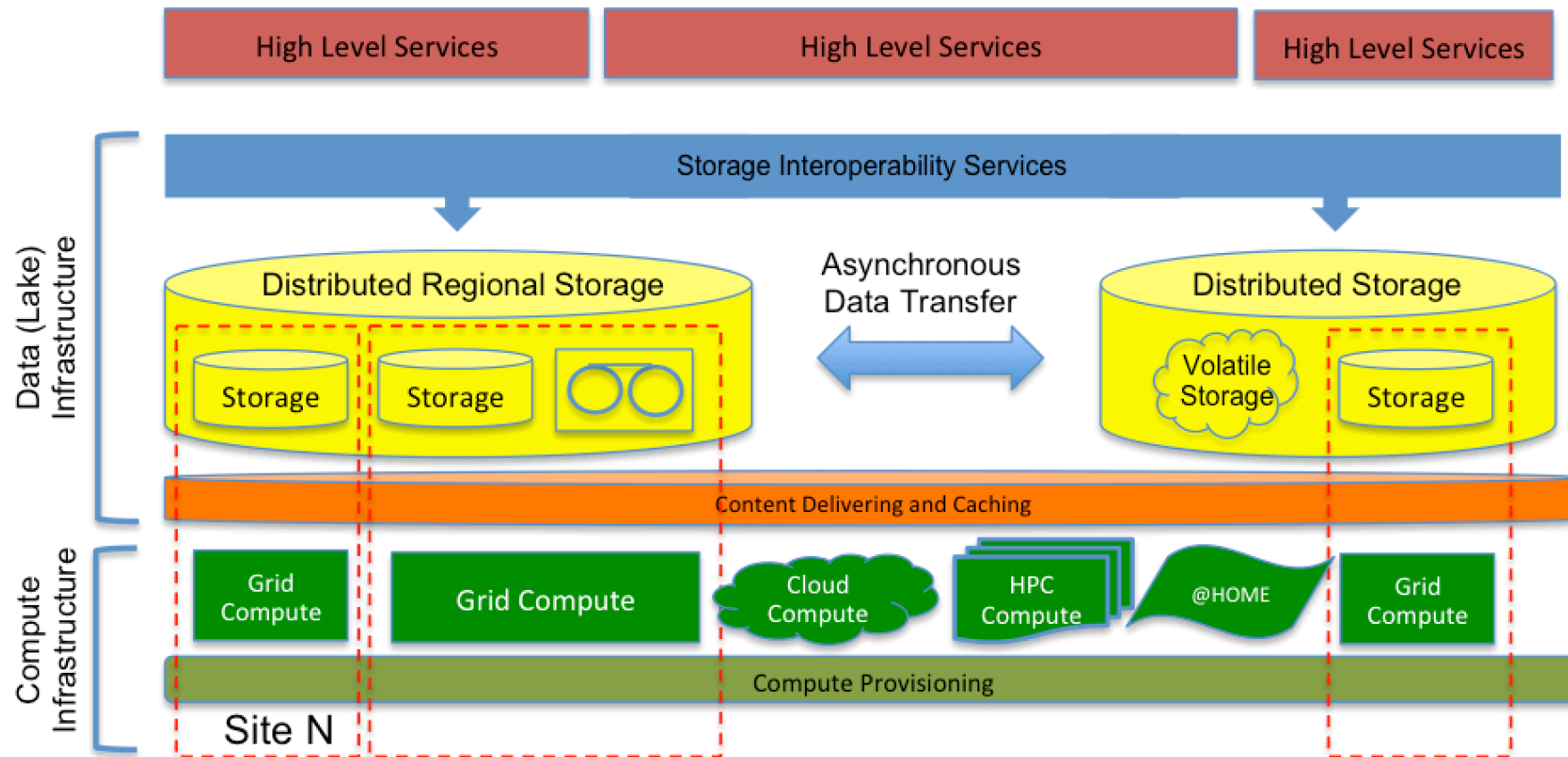
LHC @ Home Volunteer Computing

GRID /WLCG

The Resources Windmill

5.-Perspectives and conclusions

Quo Vadis, Computing?:::Towards: Data & Compute Infrastructure (also known as 'Data Lake')

From Ian Bird's presentation @ S&C Technical Meeting ATLAS

Conclusions:

- The computing(&data) *'ecosystem' is more complex* than 10 years ago:
  - The GRID groups have got a high level experience
  - Scenario: Cloud computing, HPC, GRID resources, volunteer (@home),…
- The Spanish LHC Computing (GRID) community: to establish the main guidelines for the next years (HL-LHC):
  - *Update of the LHC Computing @ Spain*
  - Extension to another FPA sectors to take profit of experience
- Moving towards *common implementations* (where realistic); joint efforts => i.e. In Data Management, Data lake(s), Resource Federations, Machine/Deep Learning Facilities
- *Training of 'Data Scientists'* (physicists & technologists): staff well valued at the experiment
- *Strategy*: Confluence of interests, synergy between projects, collaborative environments,…
- Global LHC Community meets computing teams of the groups *common issues – new techniques- workflows*
  - Tier-2 sites are placed in the centers involved in LHC experiments (Detectors, Physics Analysis,…)

- **THANKS FOR YOUR ATTENTION**
- **QUESTIONS**