

Technical aspects of the Spanish ATLAS Tier-2 Project @ IFIC

Santiago González de la Hoz and José Francisco Salt Cairols

Grupo de GRID y e-Ciencia del IFIC

Jornadas Técnicas del IFIC: 12 y 13 de Marzo de 2018

13 de Marzo de 2018

Contents:

- 1.- Introduction.**
- 2.- Evolution of the ATLAS Computing Model.**
- 3.- The Spanish ATLAS Tier-2**
- 4- Technical issues from the Tier-2 project**
- 5.- Perspectives and conclusions**

1.-INTRODUCTION

:1 year ago, we had the IFIC Scientific Days...



TIER-2 GRID del experimento ATLAS

(COMPUTACION en la Era del LHC)

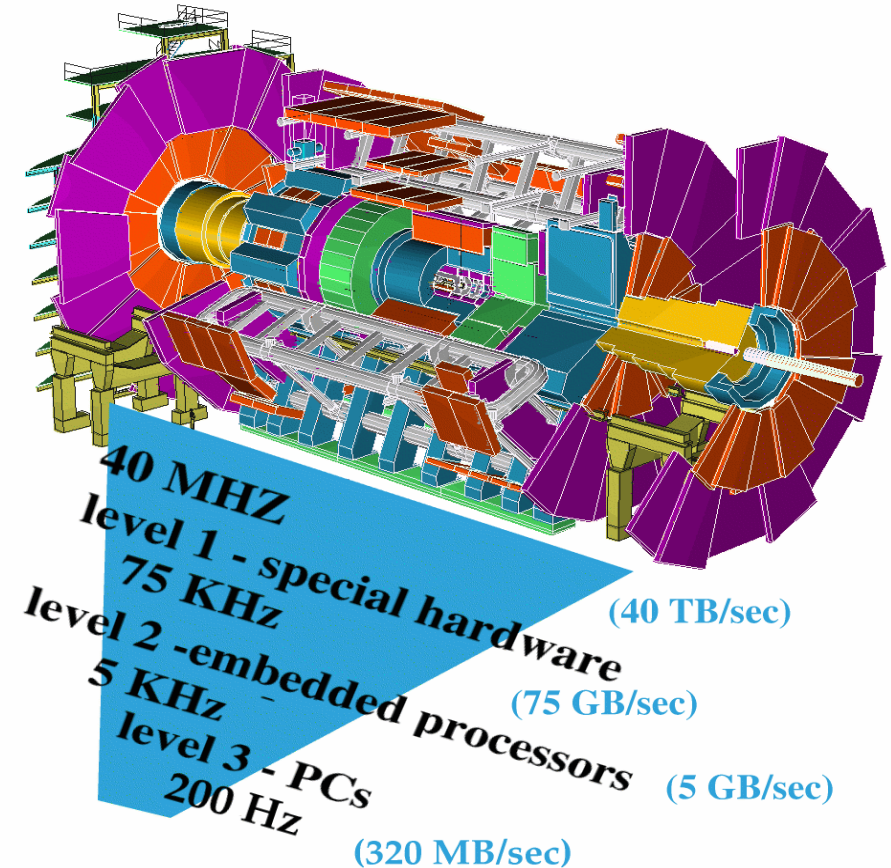
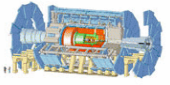
Presentado por: José F. Salt Cairols

Jornada Científica del IFIC: 19 y 20 de Enero de 2017

IFIC
INSTITUT DE FÍSICA
CORPUSCULAR



the ATLAS Experiment



- The Spanish ATLAS Tier-2 project is a it's an almost totally technological project
- So it fits well in these Technical Days: a lot of technical aspects can be mentioned

ATLAS Computing Challenges

The offline computing:

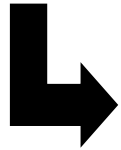
- Output event rate: 200 Hz ~ **10^9 events/year**
- Average event size (raw data): **1.6 MB/event**

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 Challenge (DC)

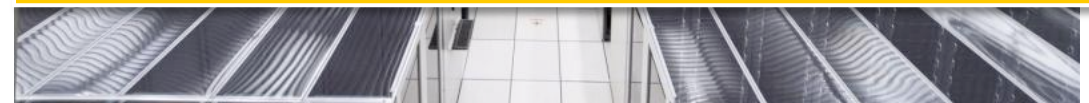
ATLAS Production System (ProdSys)

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

15:38:56 UTC



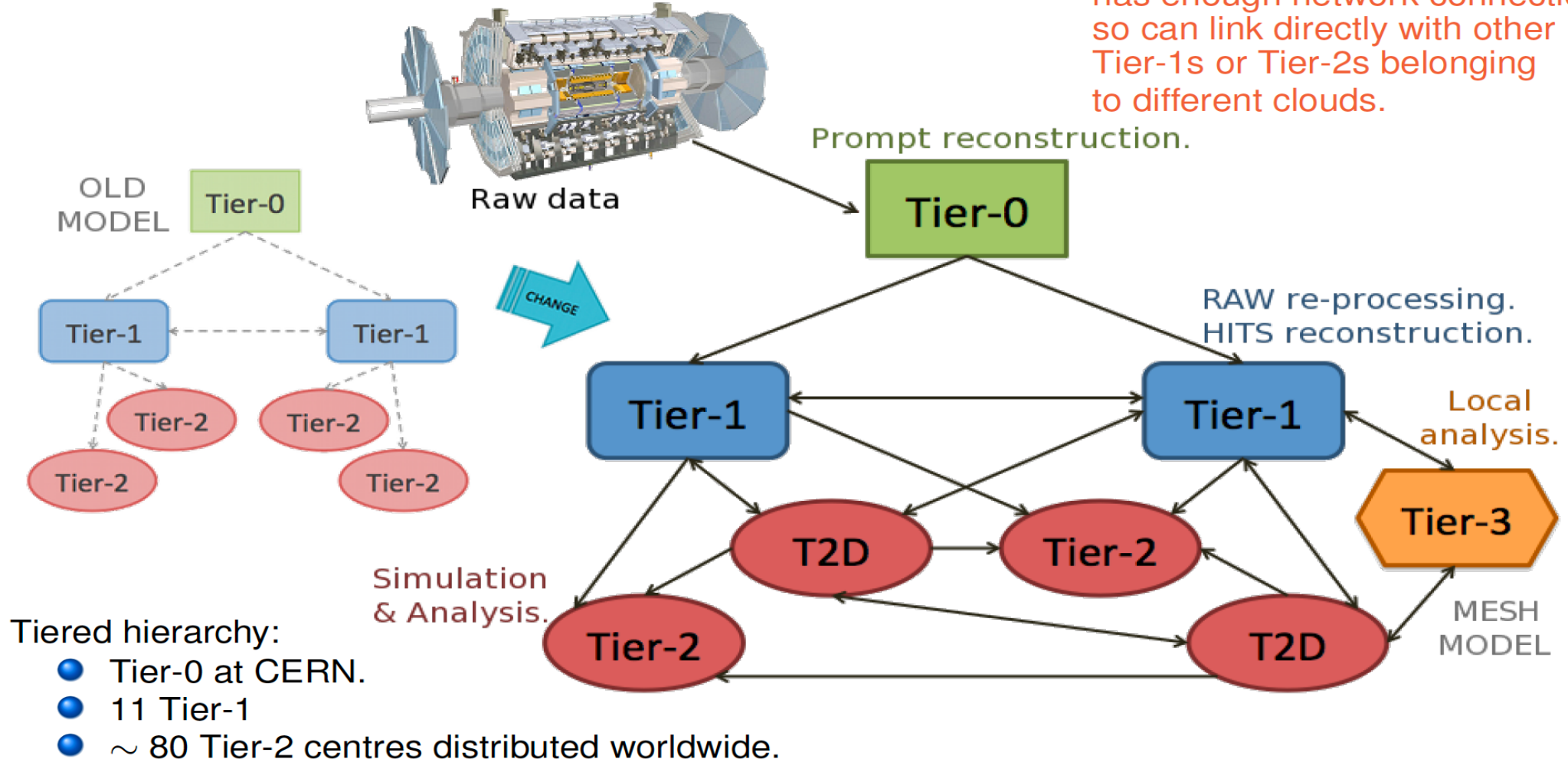
- CERN provides 30% of the computing power for LHC: 30,000 CPUs.
- 5 million Gbits of disk storage.
- 16 million Gbits of magnetic tape storage.
- Connected to other centers at 10Gbits / s.



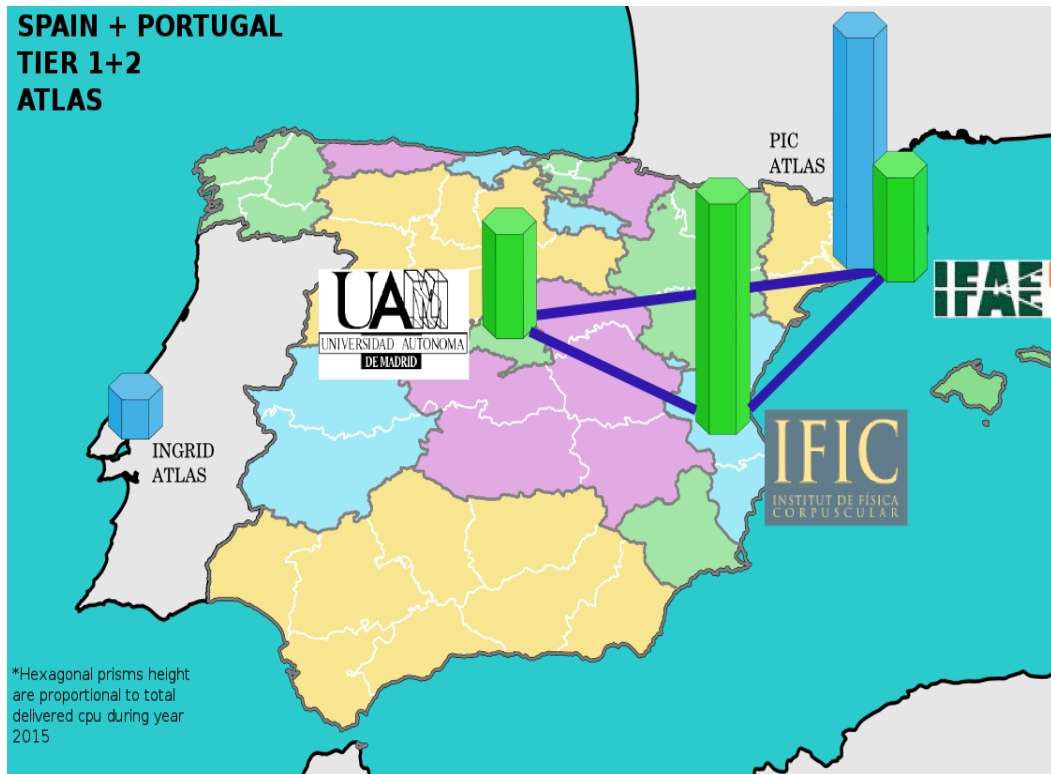
2.- Evolution of the ATLAS Computing Model

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.



3.- The Spanish ATLAS Tier-2

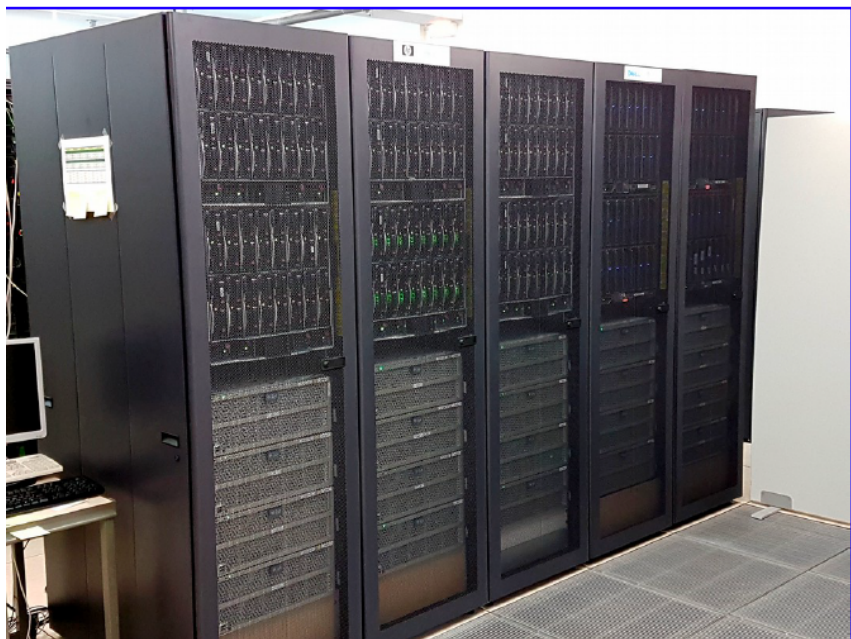


- Federation of 3 centers:
 - IFIC (Valencia), 60%
 - IFAE(Barcelona), 25%
 - UAM(Madrid),15%
- Spanish ATLAS Tier-2 Infrastructure, started in 2005
Associated to Iberian Cloud (PIC as Tier-1)
 - 4% of all ATLAS Tier-2

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)



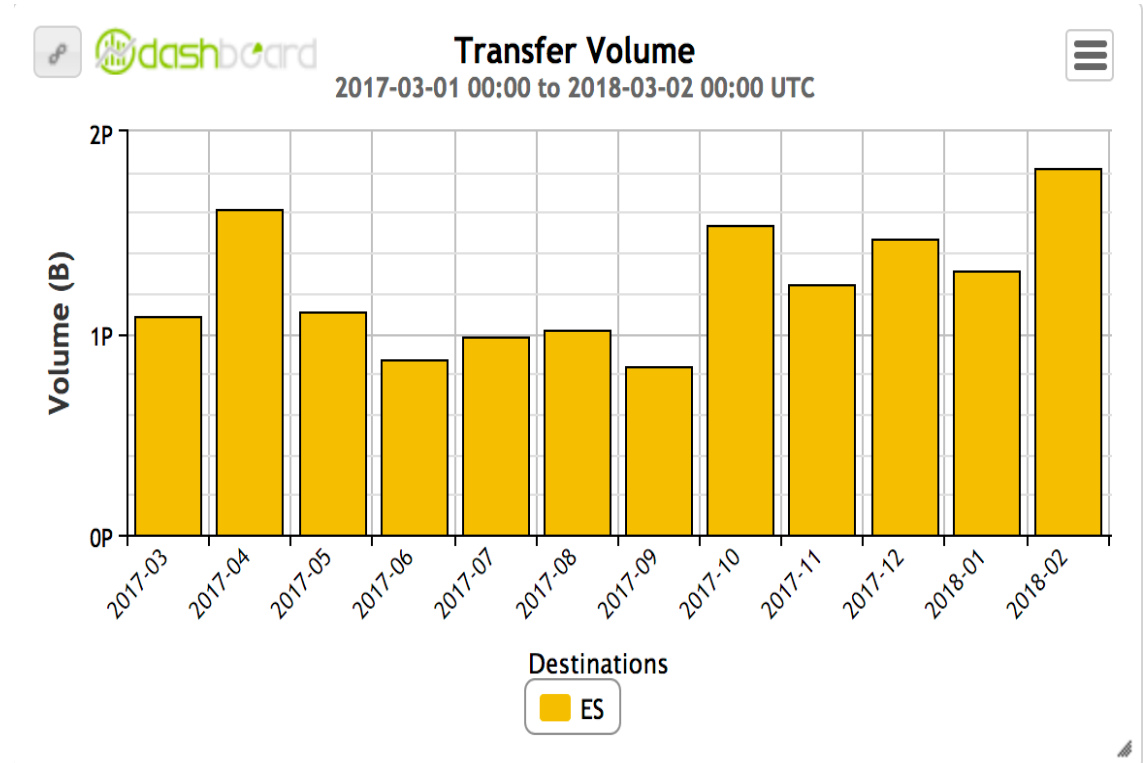
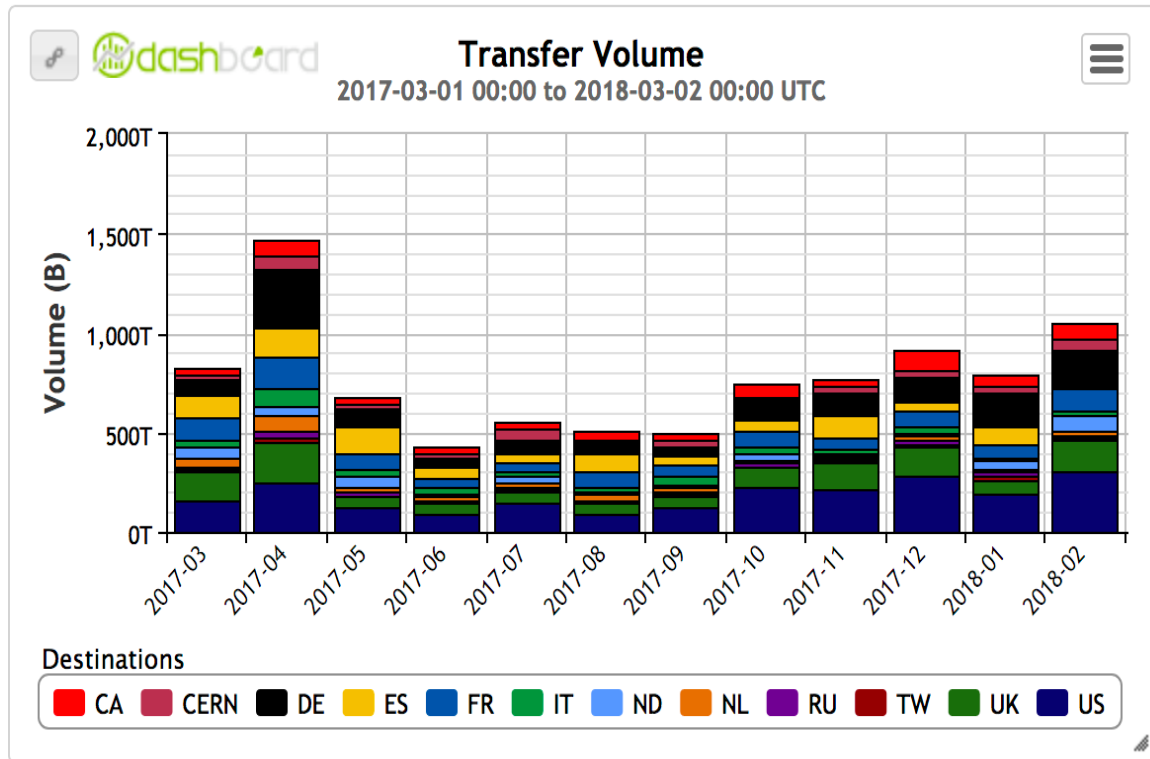
	nodos	CPU	RAM	DISCO	cores	compra	>5y
ATLAS T2	32	2 x Intel Xeon E5472 @ 3.00 GHz (4 cores)	16 GB	2x300 GB SAS 15krpm	256	jun-2008	👁
	32	2 x Intel Xeon L5520 @ 2.26 GHz (4 cores)	24 GB	2x146 GB SAS 10krpm	256	mar-2010	👁
	16	2 x Intel Xeon E5-2660 @ 2.20 GHz (8-cores)	64 GB	2x300 GB SAS 15krpm	256	jul-2012	
	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	7 2244

Maintenance & Operation of a complex Infrastructure and continuously updated

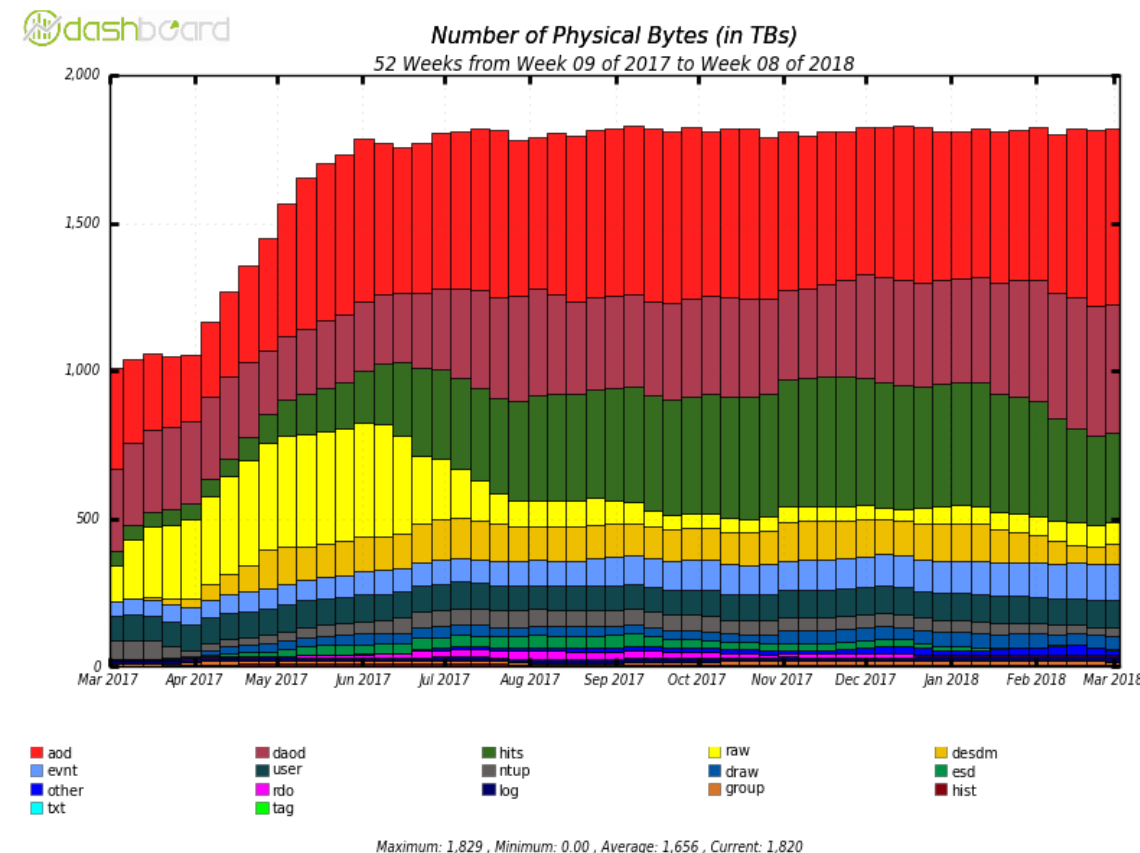
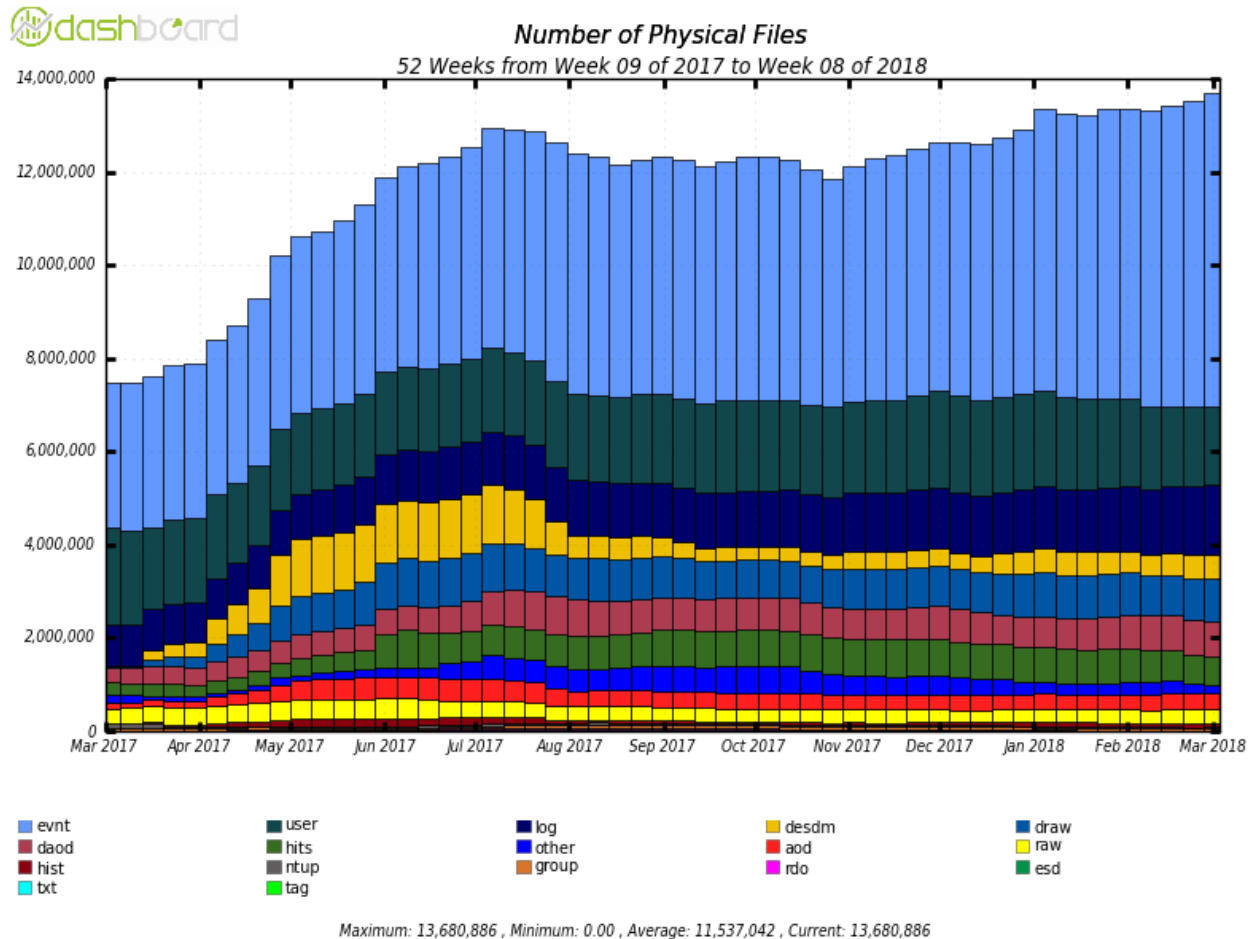
- Transfers in the last year from the rest of Tier2 to the rest of the World

- Transfers in the last year from the rest of the World to Tier2
=> More than 2 PetaBytes of transferred data

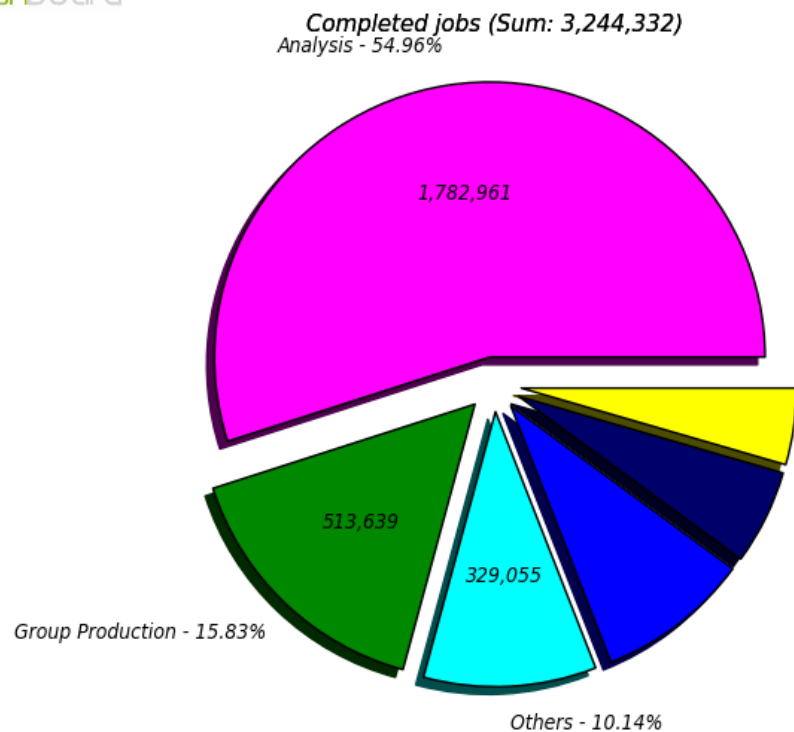


Tier-2 IFIC performance:

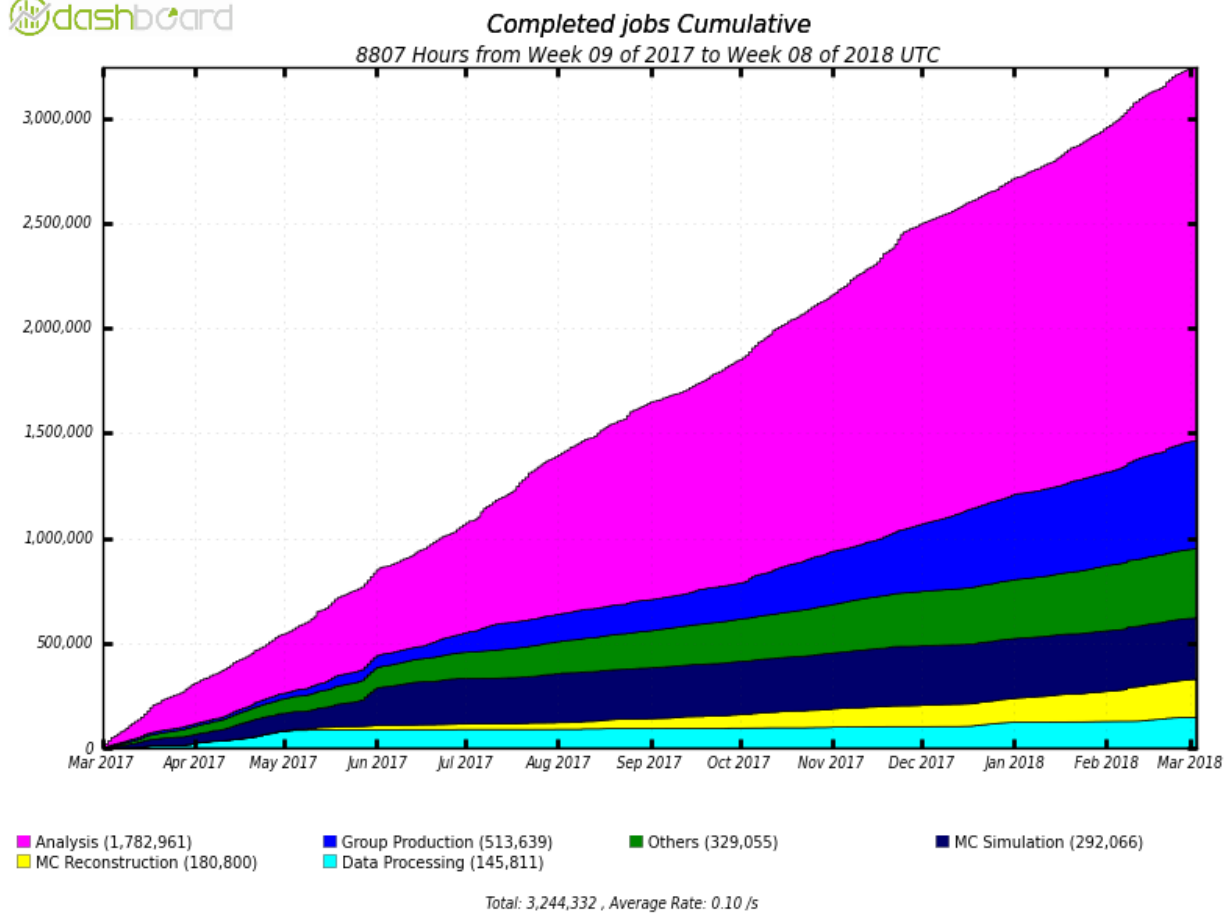
- Data produced at Tier-2/IFIC in the last year (More than 14M files and more than 2 PBs)



- Executed jobs in the last year @ Tier-2 (more than 3 M)



Analysis - 54.96% (1,782,961) Group Production - 15.83% (513,639) Others - 10.14% (329,055)
 MC Simulation - 9.00% (292,066) MC Reconstruction - 5.57% (180,800) Data Processing - 4.49% (145,811)



4.- Technical issues from Tier-2 project

Proyectos/tareas de Investigación relacionados con el Tier2 de IFIC: the computing challenge for HL-LHC

- ATLAS and CMS will need x20 more resources 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
- 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.**

Research tasks included in the Tier-2 Project

Talks

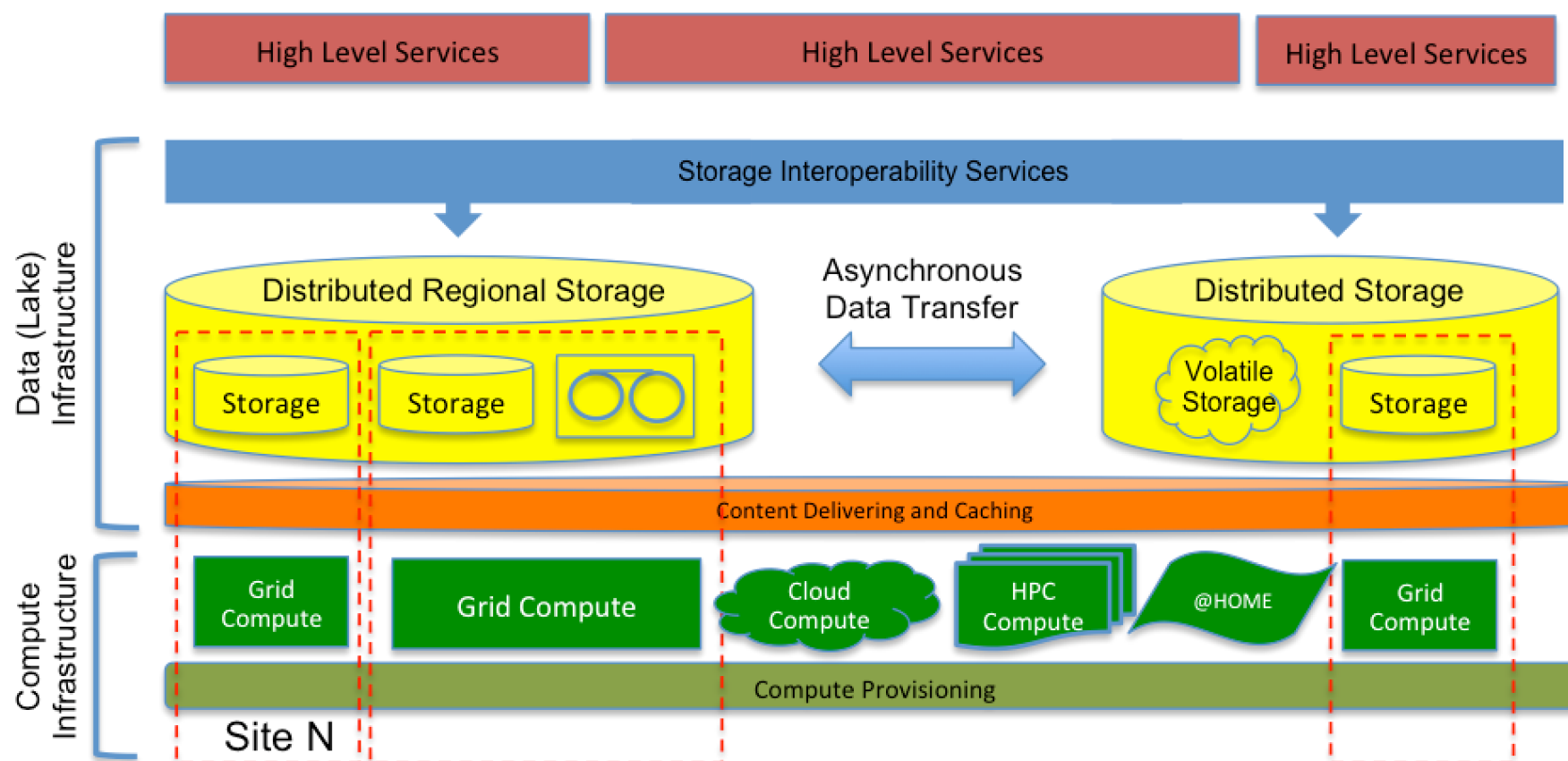
- Monitorización de los servidores Frontier de ATLAS (ver charla de LOZANO BAHILLO, Julio)
- Monitoring de las transferencias del TIER2 con ELK (ver charla de APARISI POZO, Javier)
- ATLAS EventIndex y EventWhiteboard. Un sistema de indexado y catalogado de eventos para experimentos con grandes cantidades de datos (ver charla de FERNÁNDEZ, Álvaro)
- ATLAS EventIndex Trigger Counter, una herramienta de análisis del trigger que emplea MapReduce (ver charla de GARCÍA MONTORO, Carlos)
- Aplicación de Machine Learning al análisis de física en ATLAS (ver charla de CAMPO, Samuel)

Tasks

- Infrastructure Management & Operations
- Distributed Computing: GRID, Cloud, HPC
- Data Bases: NoSQL,
- Big Data
- Data Storage
- Machine Learning techniques
- ...

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

José Salt, Jornadas Técnicas IFIC, IFIC 12-13 de Marzo 2018



Conclusions:

- The computing(&data) *'ecosystem' is more complex* than 10 years ago:
 - The GRID IFIC group has 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)
- *Training of 'Data Scientists'* (physicists & technologists): staff well valued at the experiment
- *Strategy @ IFIC* : Confluence of interests, synergy between projects, collaborative environments,...

- **GRACIAS POR VUESTRA ATENCION**
- **PREGUNTAS**