Spanish and Portuguese contributions to CMS Computing Grid activities

E.Acción[1], N.Almeida[4], G.Bernabeu[1], G.Borges[4], A.Bria[1], I.Cabrillo[3], N.Colino[2], M.David[4], M.Delfino[1], A.Delgado-Peris[2], J.Flix[1], J.Gomes[4], I.González Caballero[3], J.Hernandez[2], A.López[3], R.Marco[3], F.Martinez[1], J.P.Martins[4], F.Matorras[3], G.Merino[1], M.Oliveira[5], P.Orviz[3], F.J.Rodríguez-Calonge[2]

[1] Port d'Informació Científica, PIC (CIEMAT - IFAE - UAB), Bellaterra, Spain
[2] Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas, CIEMAT, Madrid, Spain
[3] Instituto de Física de Cantabria, IFCA (U. Cantabria - CSIC), Santander, Spain
[4] Laboratório de Instrumentação e Física Experimental de Partículas (LIP), Lisboa, Portugal
[5] Laboratório de Instrumentação e Física Experimental de Partículas (LIP), Coimbra, Portugal

Josep Flix - CMS Tier-1 contact at PIC (CIEMAT)

MINISTERIO DE EDUCACIÓN Y CIENCIA

Ciemat
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

PIC
port d'informació científica

The Large Hadron Collider [LHC] I

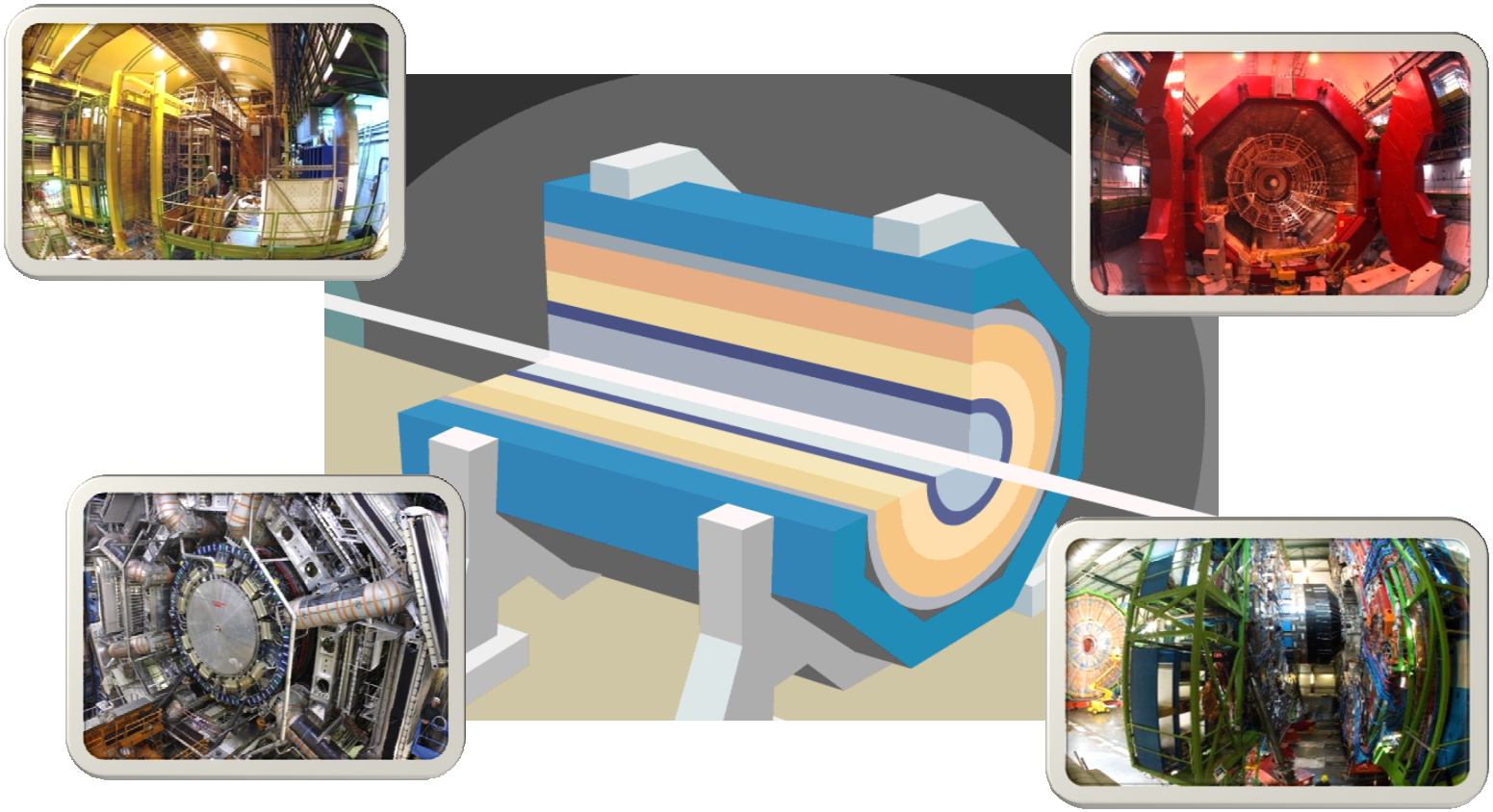
Large Hadron Collider & CERN area

- Built in a 100 meter underground 27 Kms circumference tunnel.
- p-p collisions at 14 TeV @ 25 ns.
- Starts operations in summer 2008.

Large Hadron Collider – Tunnel view (2008)

The Large Hadron Collider [LHC] II

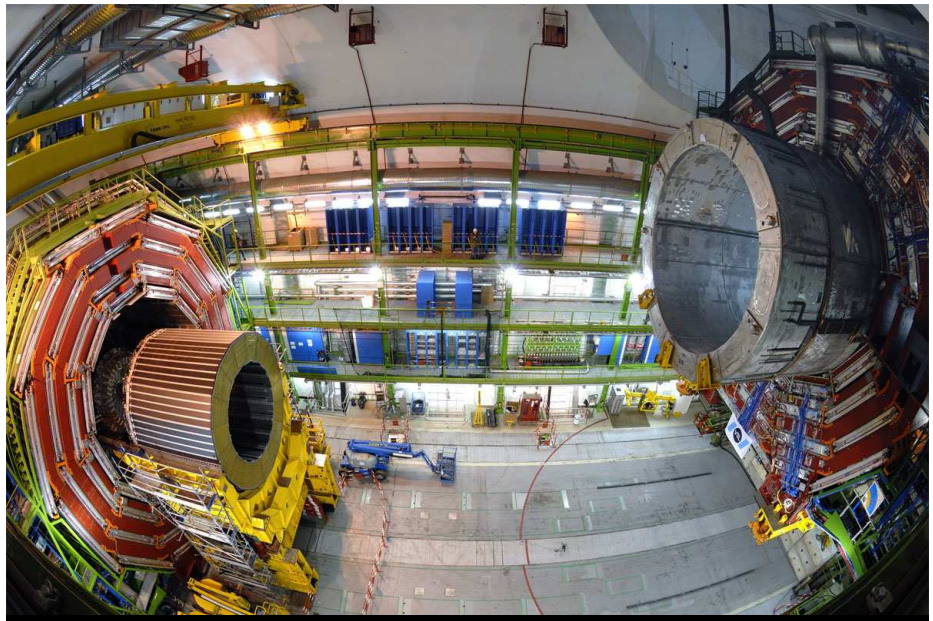
- Four particle detectors responsible to detect results of p-p collisions:



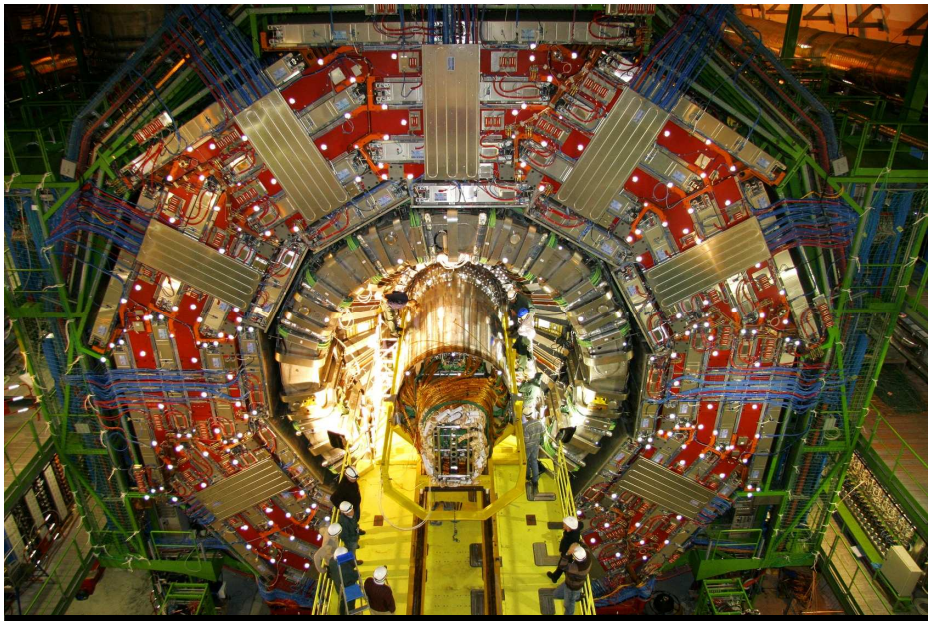
- A plethora of new particles produced per collission → Their tracks are examined for evidences of new physics, to explore the fundamental nature of matter, and the basic forces that shape our universe.

The Compact Muon Solenoid [CMS]

- CMS is a large worldwide collaborative effort:
  - 2900 physical scientists (of a total of 3400 collaborators)
  - 180 Universities and Research Laboratories (40 countries)



CMS central part (YB0) landing at the pit (Feb. 2007)

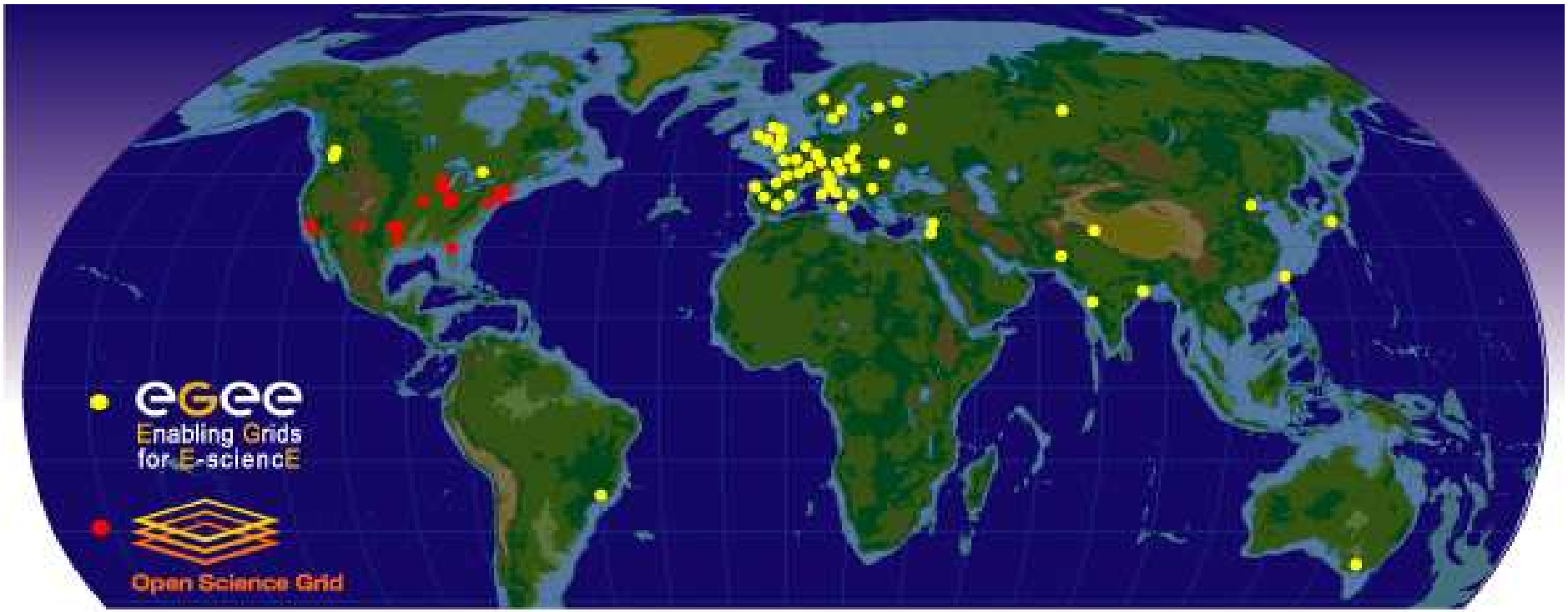


CMS inner tracker installation (Dec. 2007)

- CMS sub-detectors almost **commissioned**.
- When the LHC starts operations in 2008, the CMS DAQ rate ~ 300 Hz will provide data from the detector at **~450 MB/s → ~PBs/year**.

WLCG computing architecture I

- **WLCG**: Grid infrastructure to store and analise all the LHC data.
- Formed by hundreds of computing centres in more than 20 countries.
- Operated by two large Grid projects: **EGEE** and **OSG**.

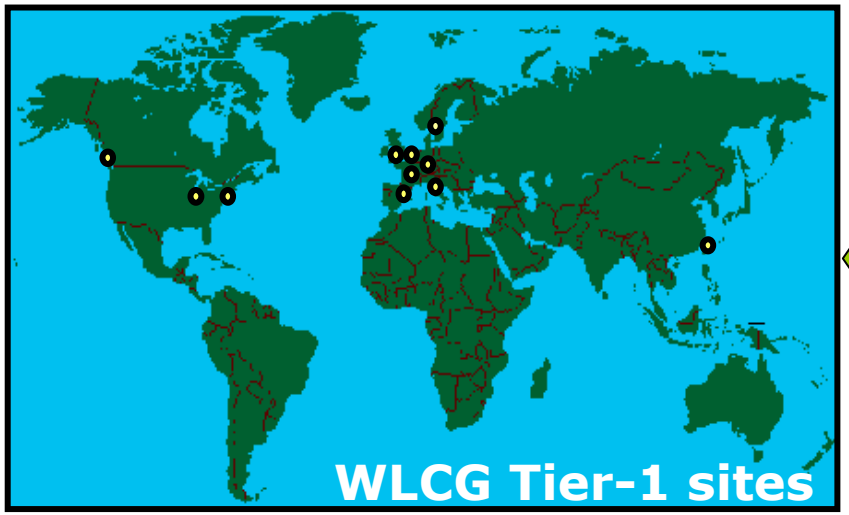


A map of the worldwide LCG infrastructure operated by EGEE and OSG.

WLCG computing architecture II

Tier-0: the accelerator centre

- Data acquisition & initial processing
- Long-term mass data storage
- Distribution of data → Tier-1 centres



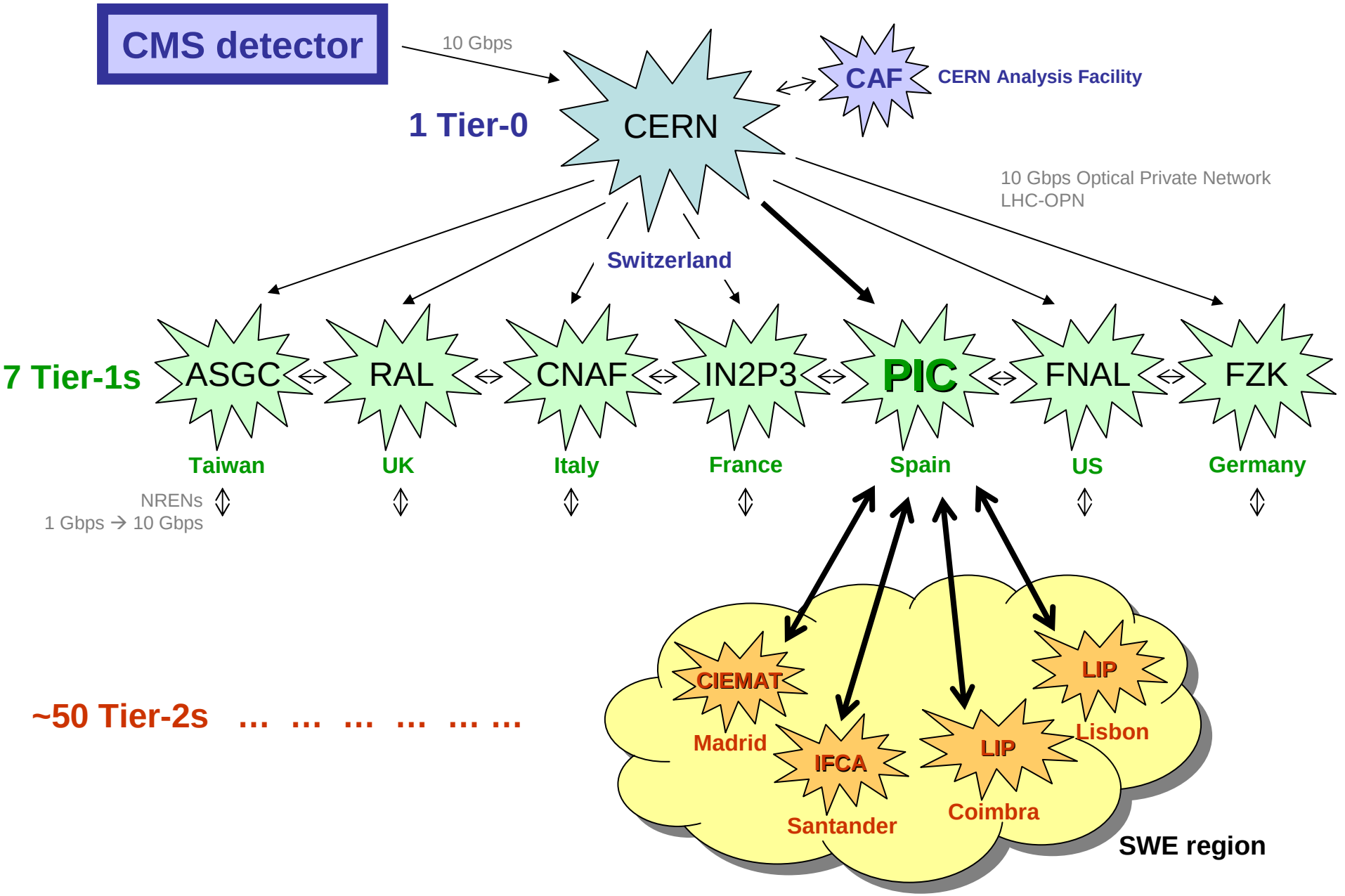
Tier-1: "online" to the DAQ

- High availability centers
- Custodial mass storage of share of data
- Data reconstruction and reprocessing
- Distribute analysis data → Tier-2s

Tier-2: ~100 centres in ~20 countries

- MonteCarlo Simulation → Tier-1
- End-user analysis

CMS WLCG computing architecture



CMS SWE region: Tier-1 & Tier-2s

- Tier-1 SWE region → **PIC**
- Tier-2 Spanish federation → **CIEMAT**, **IFCA**
- Tier-2 Portuguese federation → **LISBON**, **COIMBRA**

- **Commitments**:

  - Provide resources equivalent to **5% of the total CMS needs** for Tier-1 and Tier-2.
  - Provide the corresponding services with a **good quality**.
  - Provide **computing environment** to the Spanish and Portuguese CMS HEP community for **an appropriate profit of LHC physics potential**.
  - ...

- **Main goal**: be prepared in mid-2008 for first physics LHC run.

Tier-1: Port d'Informació Científica I

SWE region: Tier-2s associated to Tier-1 PIC

CMS: CIEMAT&IFCA – LIP_Lisbon&LIP_Coimbra

ATLAS: IFAE&IFIC&UAM – LIP_Lisbon&LIP_Coimbra

LHCb: UB&USC

Port d'Informació Científica* [PIC]

Generalitat de Catalunya, CIEMAT, UAB, IFAE

- Provider of computing services for scientific research in collaborative environments with lots of data. (TB-PB)
- **Tier-1 for ATLAS, CMS and LHCb.**
- Services provided for other disciplines: Astrophysics, medical imaging…

(*) 150 m² machine room - UPS of 200 KVA - Diesel generator 500 KVA - 1000 rack units - 300 KW air conditioned – Tape mass storageSTK-5500+IBM-3584.

Tier-1: Port d'Informació Científica II

- **Reliability** is an important issue for a T1 (24x7):

  - Typical PIC monthly WLCG site reliability has shown to be above 95%.

- PIC computing contribution for storage and processing:

Spain PIC	2007	2008	2009	2010
Tier-1 CPU [kSI2K]	501	1509	2591	5109
Tier-1 Disk [TBs]	218	967	1702	3009
Tier-1 Tape [TBs]	243	953	1844	3402

CMS PIC	2007	2008	2009	2010
Tier-1 CPU [kSI2K]	294	477	1058	2516
Tier-1 Disk [TBs]	81	358	630	1113
Tier-1 Tape [TBs]	142	487	974	1677

- As of **May 2008** (ramping up to 2008 deliverables):

  - Total CPU capacity of 1500 kSI2k (~600 job slots).
  - Disk storage service of ~ 600TB (120 TB for CMS) → to grow to 1 PB by mid-2008.

- PIC is connected to Tier-0 and Tier-1s through **10 Gbps LHC-OPN**.

- Connectivity to Tier-2 centers is routed through the Spanish NREN RedIRIS. Currently limited to 2 Gbps → upgrade to 10Gbps soon.

Spanish and Portuguese Tier-2s

- A total of four Spanish Institutes actively contribute to CMS activities: **CIEMAT** and **UAM** in Madrid, **IFCA** in Santander, and **Universidad de Oviedo** in Asturias → ~ 80 scientists.

CIEMAT & IFCA:

CMS Tier-2 Spain	2007	2008	2009	2010
Tier-2 CPU [kSI2K]	380	760	1280	2260
Tier-2 Disk [TBs]	65	210	420	665

- In Portugal, about 25 Portuguese physicists and engineers, from five institutes (**LIP**, **INESC**, **IST**, **INEGI** & **Univ. Porto**) collaborate in CMS.

LISBON & COIMBRA:

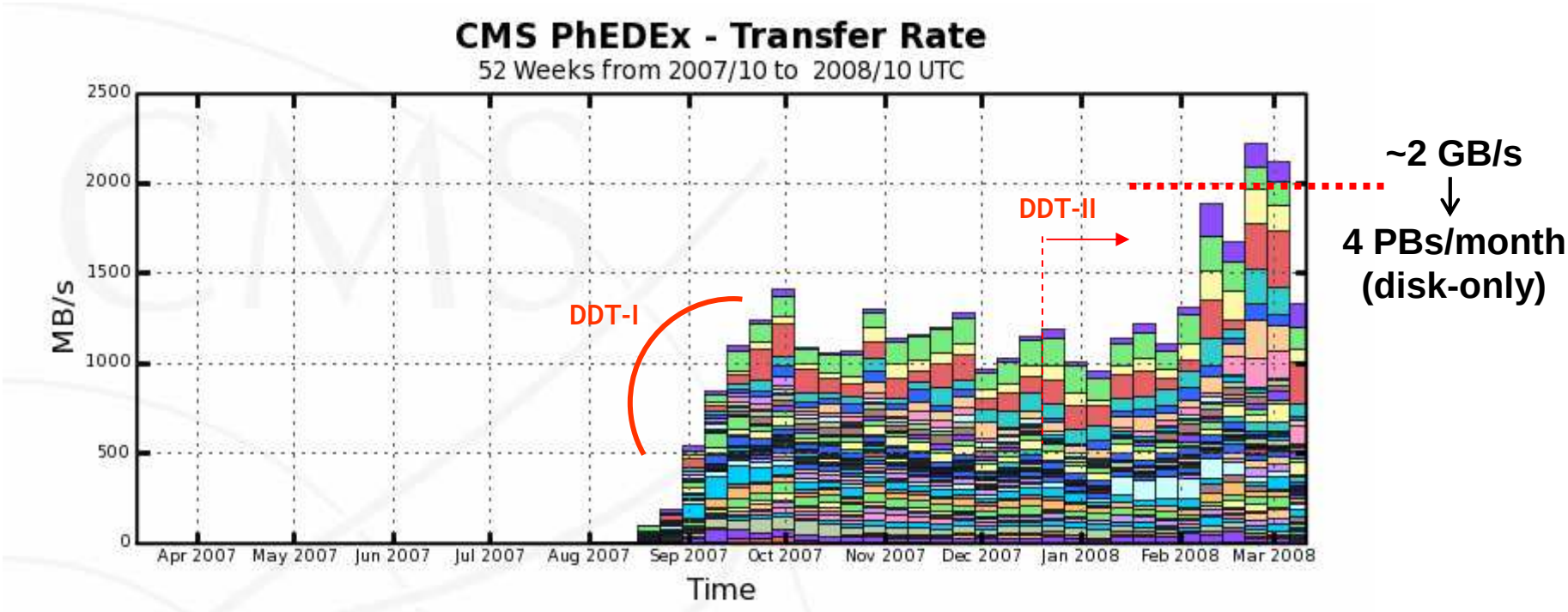
CMS Tier-2 Portugal	2007	2008
Tier-2 CPU [kSI2K]	500	750
Tier-2 Disk [TBs]	80	160

Recent/Ongoing CMS computing challenges

- CMS undertakes **periodic computing challenges** of increasing scale and complexity to test its computing model and Grid systems:
  - **Debugging Data Transfers (DDT):** CMS program intended to deliver fully debugged and operational end-to-end links between CMS Tier centres.
  - **Computing, Software and Analysis challenge 2007 (CSA07):** data challenge in end of 2007 designed to test the system at 50% of the design goal for 2008.
  - **CSA08:** challenge to test the system at 100% full capacity, ongoing atm...
  - **Common Computing Readiness Challenge of 2008 (CCRC08):** an LHC-wide computing challenge in which all LHC experiments test their computing activities in parallel. Preparatory Phase-I series of tests was scheduled for February 2008. Main Phase-II is ongoing atm…
  - **CMS Global Runs:** Data taking with Cosmic Muons. A Global Run ongoing atm…
- Performance values are measured, problems are identified and feedback into the design, integration and operation of the computing system is provided.

Debugging Data Transfers

- In Sept. 2007, all 'fake' LoadTest traffic was moved to independent non-production data management environment.
- Only links satisfying certain conditions were enabled or commissioned and be ready for CMS production activities. **Metric:** ~20 MB/s sustained for 1 day.



- The program has **delivered** a fully debugged and **operational end-to-end links** to the CMS Data Operations team → **~300 certified links**.

Debugging Data Transfers @ SWE I

- The DDT has been a useful tool to establish **robust** to improve the **connectivity of PIC and SWE Tier-2** sites to many other centes.

- **2007 results:**

150 TBs/month → ~7% of CMS Tier-1s traffic

Link	LoadTest (TBs)	DDT-I (TBs)	Total (TBs)	% (CMS Total)	comm. links
PIC T1 Import	648	402	1050	10.5	16
PIC T1 Export	435	415	850	4.7	11
CIEMAT T2 Import	215	180	395	2.4	5
CIEMAT T2 Export	90	50	140	2.7	2
IFCA T2 Import	165	160	325	2.0	7
IFCA T2 Export	50	75	125	2.4	3
LISBON T2 Import	13	70	83	0.5	2
LISBON T2 Export	4	73	77	1.5	2
COIMBRA T2 Import	0.05	80	80.05	0.5	2
COIMBRA T2 Export	0	82	82	1.6	2

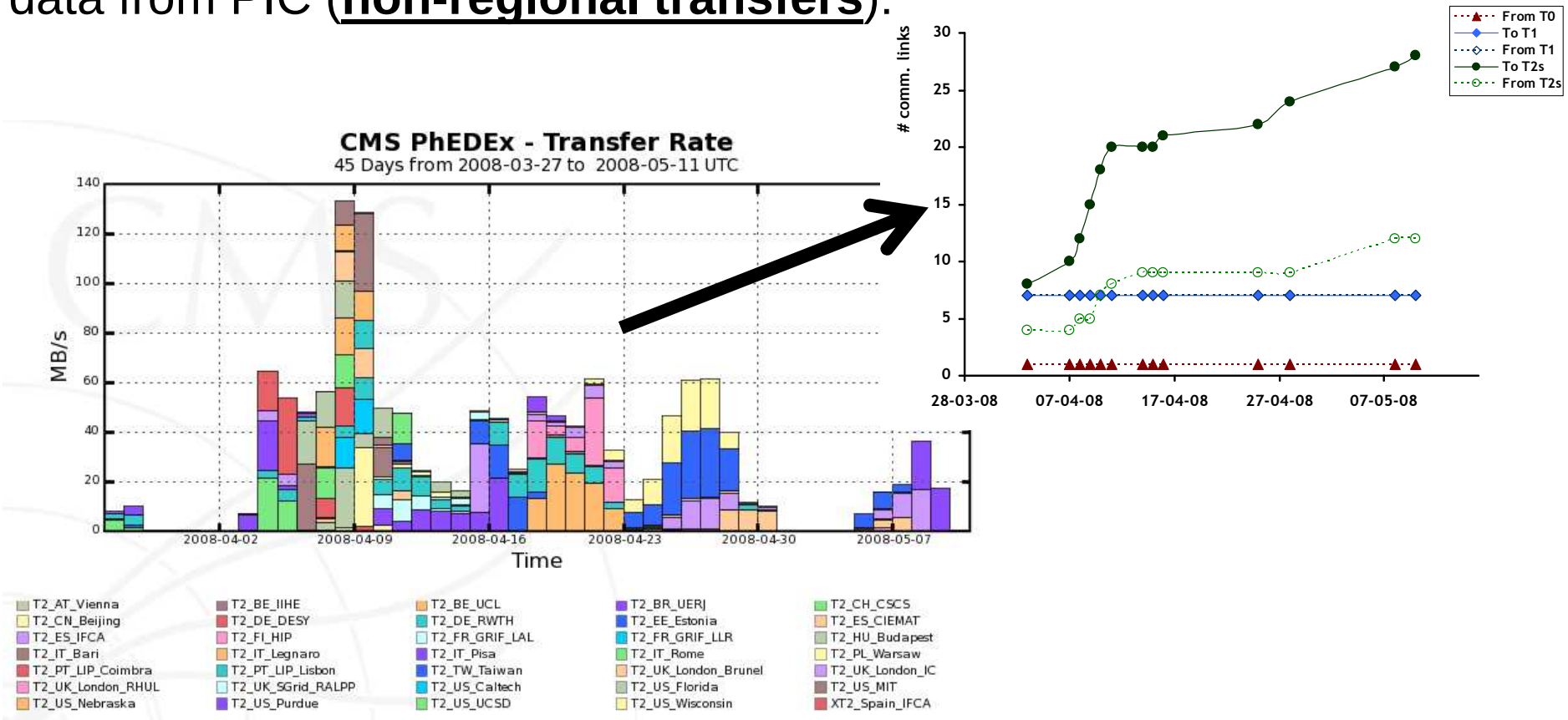
~9% of total CMS

- **2008 DDT results**:

- PIC import/exports → up to **350 TB/month** →15(30) TBs export(import) per day.
- Up to **120 TBs/month** obtained for T2-Spain → 3(10) TBs export(import) per day.
- Up to **60 TBs/month** for T2-Portugal →1(4) TBs export(import) per day.

Debugging Data Transfers @ SWE II

- As May 2008, about 30 Tier-2 centers (worldwide) are able to export data from PIC (**non-regional transfers**):



- This accounts for **~85% of the "active" CMS Tier-2 centers**.
- PIC is the 2nd Tier-1 better connected to Tier-0/1s&Tier-2s (1st FNAL).

The CSA07 CMS challenge

- This challenge was designed to be a **50% capacity test** of what is required for operations in 2008 data taking operations:

  - T0 reconstruction at **100 Hz**.
  - Distribution of raw and reconstructed data from T0 to T1 sites at **300 MB/s**.
  - Transfers between T1s at a rate of **100 MB/s**, mimicking AOD replication.
  - Data skimming jobs at T1 sites (**25K jobs/day**) and the resulted data propagated to T2 sites (**20-200 MB/s**).
  - Data analysis at the T2 sites on the skimmed data (**75K jobs/day**).
  - MC simulation in T2 sites at a rate of **50 Mevts/month**.

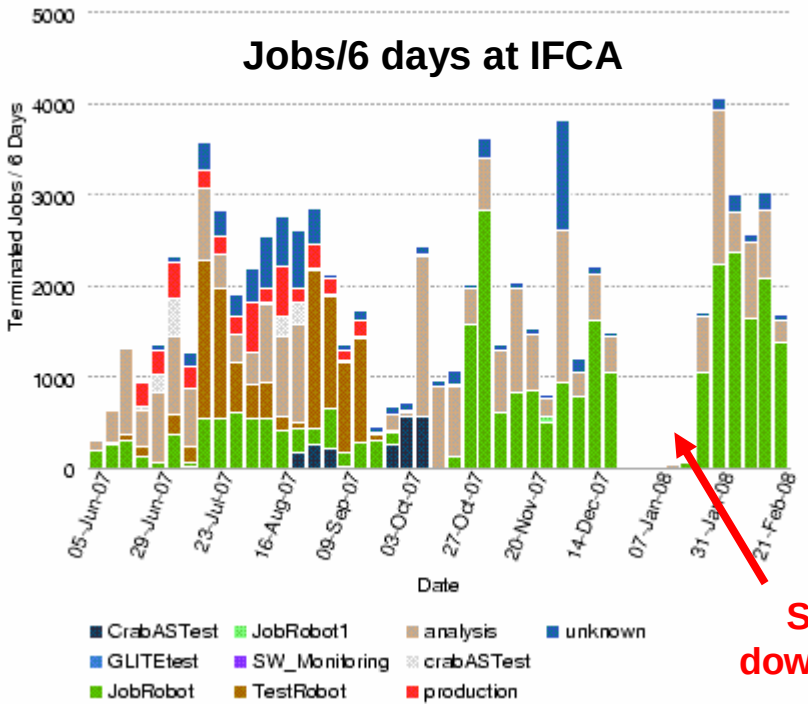
- CSA07 started in October 2007 and finished by February 2008:

  - Up to 160M MC events were produced at all CMS T2 sites.
  - Total of ~920M produced events → 1.9 PB of data (not including replicas).
  - Many of the CSA07 goals were fulfilled in all sites.
  - Transfers T1-T1 were marginally tested and job submission to T2s did not reach the goal (20K jobs/day).

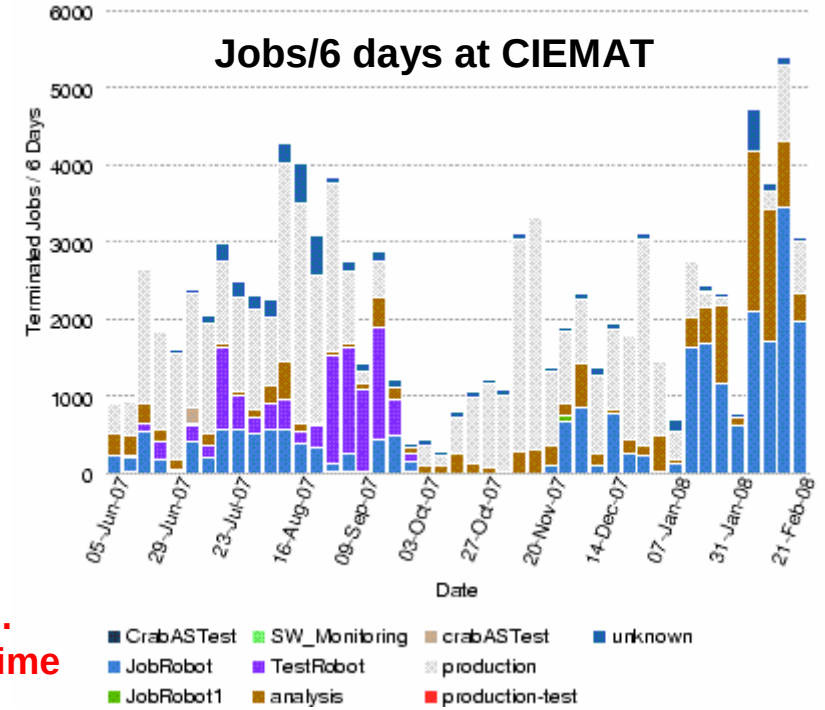
CSA07 challenge @ SWE I

- Successful participation of SWE region in CSA07 activities.
- MC production from Spanish and Portuguese sites in CSA07, w/PIC, was of about **<u>24 Mevt</u>** (around **<u>7% of total production</u>**).

  (*) CIEMAT contributed with 22 Mevt → 10% of total T2s production.



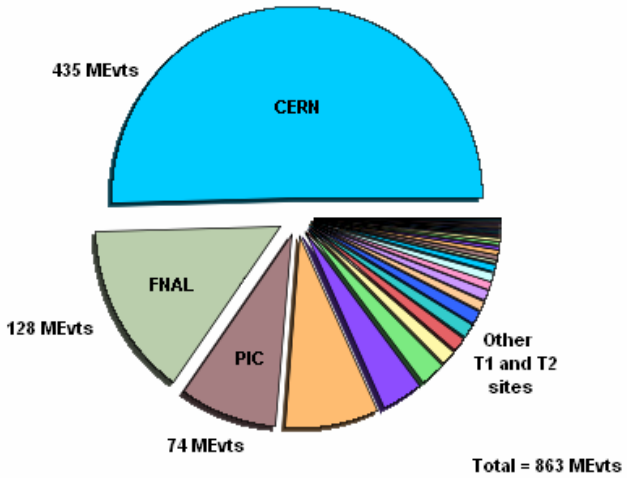
Sch. downtime

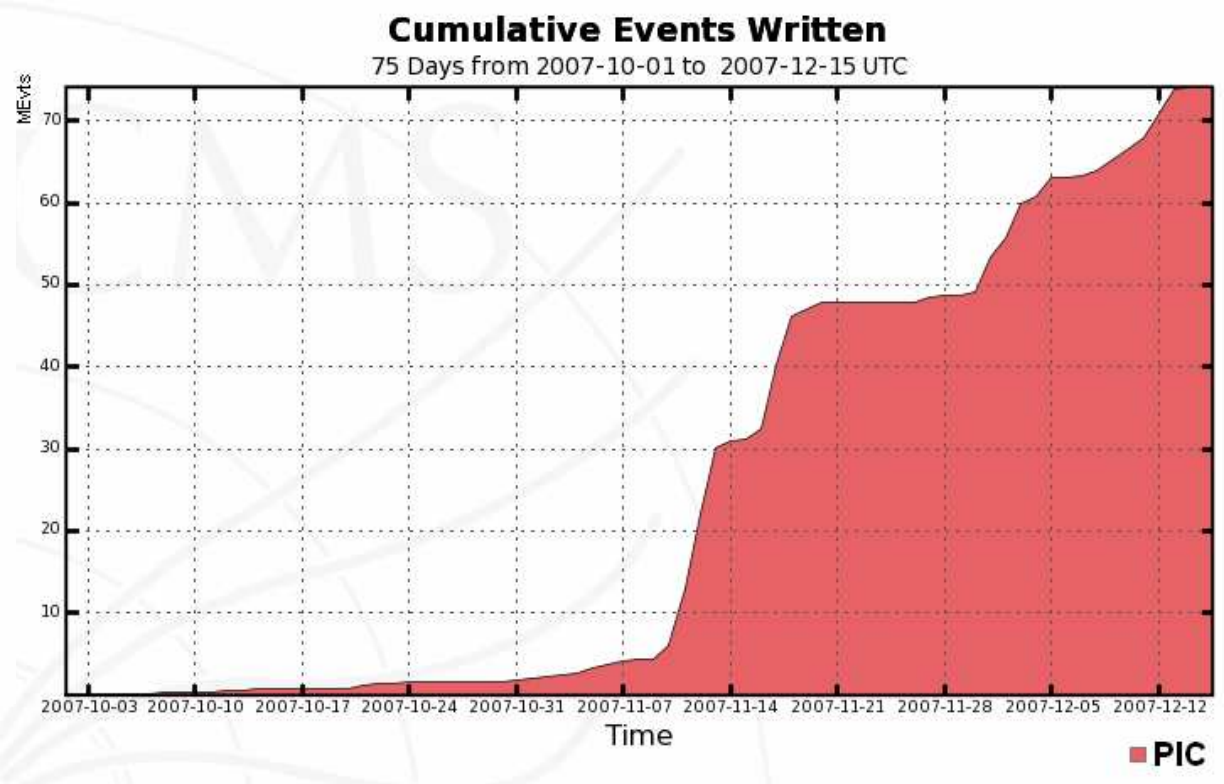


7000 jobs/week in 2008

CSA07 challenge @ SWE II

- Significant skimming contribution was made by PIC Tier-1
  → 3rd processing place, after CERN and FNAL at end 2007.





- The challenge showed that **SWE region is in good shape** to carry out CMS workflows.

The CCRC'08 challenge

- CCRC08: **wide computing challenge**, involving all LHC experiments to drive their computing workflows at once.
- The main goal is to ramp-up the load on the Tier infrastructure towards nominal values and test the stability and performance.
- First tests scheduled in Feb. 2008. Main CMS goals (transfers):

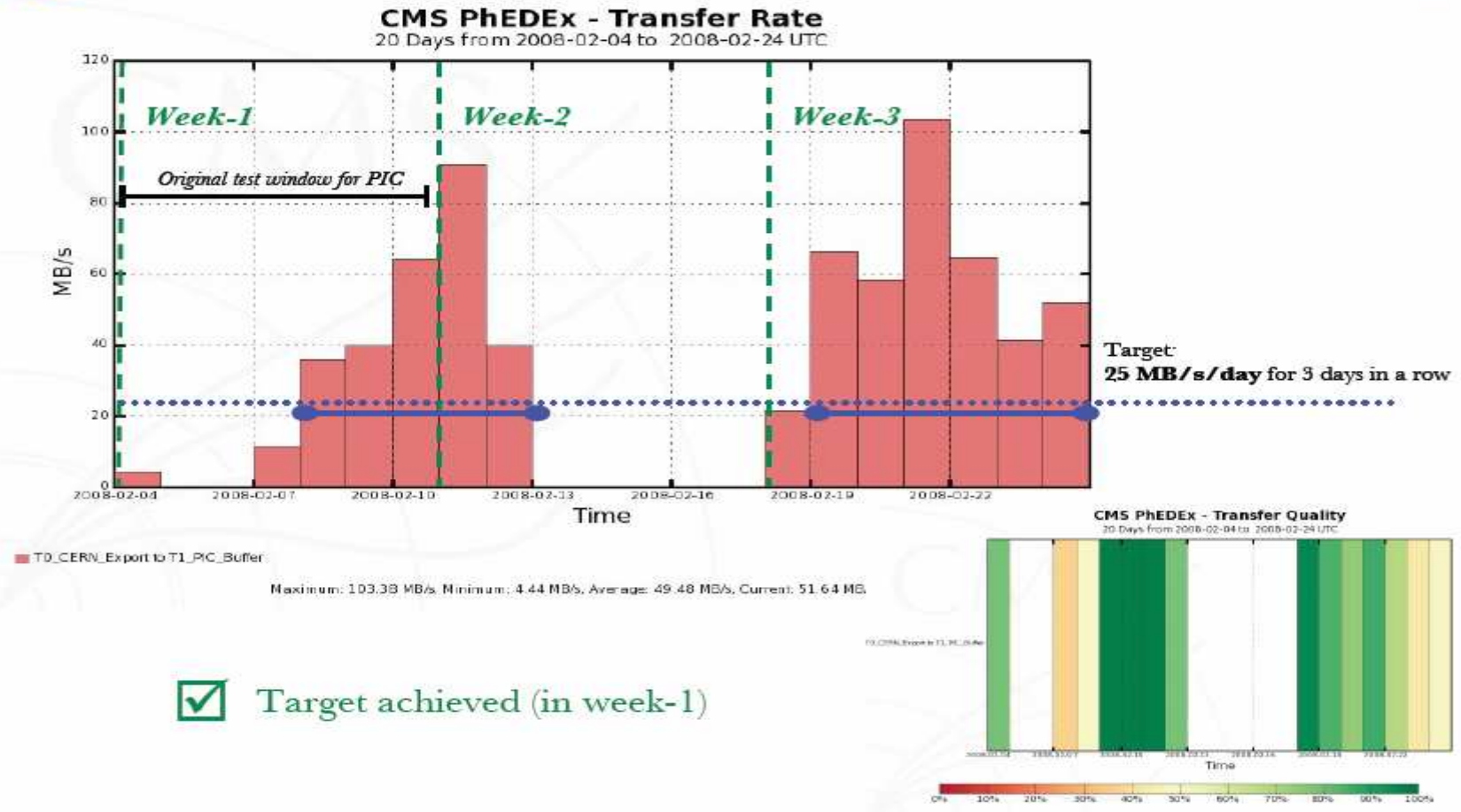
1) **T0→T1 targets**: to disk at 40% of nominal rate (25 MB/s for PIC) sustained for 3 days, and data migration to tape within 1 week with stable pattern.
2) **T1→T1 targets**: 50% of the overall 2008 nominal outbound rate to T1s (4 MB/s to each T1 from PIC). Exchange of data with at least 3 T1s, with at least 1 T1 not placed in the same continent (FNAL or ASGC in this case).
3) **T1→T2 targets** (regional T2s only): for PIC the goal was to achieve 20 MB/s to T2-Spain (CIEMAT, IFCA) and 3 MB/s toT2-Portugal (LIP-Lisbon, LIP-Coimbra).
4) **T2→T1 targets**: Traffic only from regional T2s according to the computing model. For PIC the goal was to get 1.5 MB/s from T2-Spain and 0.5 MB/s from T2-Portugal.

The CCRC'08 challenge @ SWE I

T0-T1 transfers

from:	to:	ASGC	CNAF	FNAL	FZK	IN2P3	PIC	RAL
T0		13	29	149	46	39	25	40
# days at target rate		*7*	*10*	*3*	*4*	*3*	*6*	*5*
highest rate [MB/s] satisfying metric		**75**	**70**	**149**	**58**	**51**	**58**	**53**
... corresponding to: (wrt 2008 rates)		231%	97%	40%	50%	52%	93%	53%

T0 → PIC



☑ Target achieved (in week-1)

The CCRC'08 challenge @ SWE II

Summary of T1→T1 tests

T1-T1 transfers

CCRC metric: T1 export at rates below (see matrix) demonstrated for at least 3 days in a row to at least 3 T1s, among which 1 must be in a different continent wrt source T1

FULL T1 MATRIX

from: / to:	ASGC	CNAF	FNAL	FZK	IN2P3	PIC	RAL	>=3 T1s?	other cont?	achieved?
ASGC	-	4	4	4	4	4	4	yes	yes	yes
CNAF	6	-	6	6	6	6	6	yes	yes	yes
FNAL	46	46	-	46	46	46	46	no	no	no (*)
FZK	11	11	11	-	11	11	11	yes	yes	yes
IN2P3	10	10	10	10	-	10	10	yes	yes	yes
PIC	4	4	4	4	4	-	4	yes	yes	yes
RAL	9	9	9	9	9	9	-	yes	yes	yes

	ASGC	CNAF	FNAL	FZK	IN2P3	PIC	RAL
# days at aggregate T1-outbound rate	*10*	*7*	*-*	*6*	*8*	*12*	*11*
highest rate [MB/s] (peak) achieved	**68**	**105**	**115**	**205**	**118**	**130**	**137**
... corresponding to: (wrt 2008 rates)	142%	146%	21%	155%	98%	271%	127%

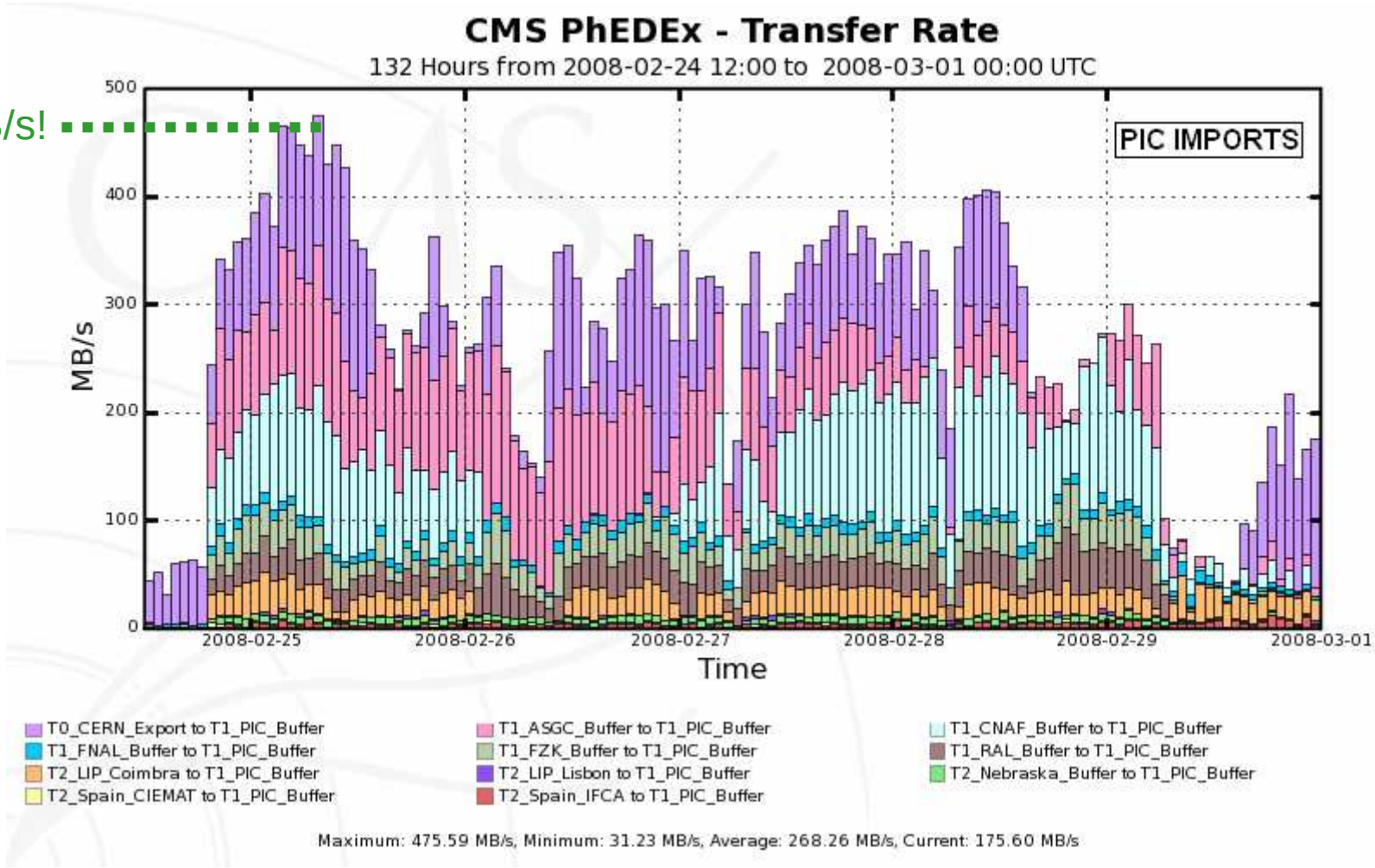
(*) the overall AOD replication scheme may be reviewed and better balanced

Goals met by the end of Feb. 2008 (next slide)

- The average **PIC export rate to T1** centres was of about 50 MB/s, with a daily averaged maximum of 130 MB/s.
- Data **imports to PIC from T1s** got an average rate of 57 MB/s, with a daily averaged maximum of 240 MB/s.

The CCRC'08 challenge @ SWE III

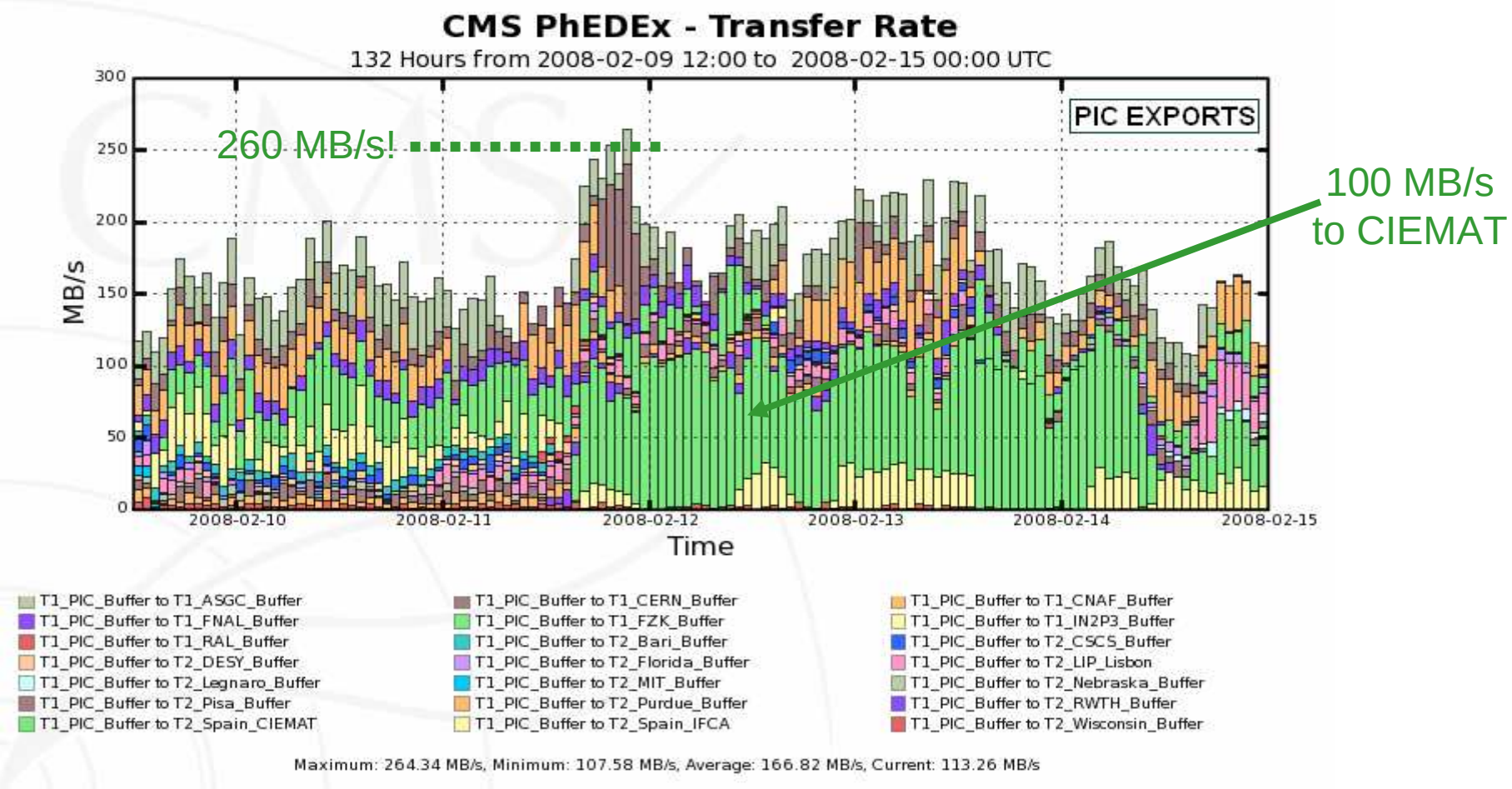
475 MB/s!



- Regional T2s (CIEMAT, IFCA, LIP-Coimbra and LIP-Lisbon) to PIC achieved 23 MB/s, with a daily averaged maximum of 34 MB/s.

The CCRC'08 challenge @ SWE IV

- **PIC to regional T2s** achieved 32 MB/s ,with a daily averaged maximum of 104 MB/s.



- PIC&T2-Spain/Portugal sites met all the CCRC08 Phase-I goals by far.

Summary and Conclusions

- The Iberian community is demonstrating to poses **functional sites** to run a complex computing system at significant scales.

- Results **above the expectations**… Overlaping tests ongoing to prove the system at higher capacities. **(CSA08, CCRC'08,GlobalRun…)**

- **Active participation** on software activities and coordination and operations of important CMS computing activities:

  - Integration Program coordination.
  - CMS sites commissioning program coordination.
  - Active participation on Facilities Operations.

- Very **valuable experience** has been gained by participating on all these activities.

- 2008 resources almost ready → **looking forward for collisions!**