Closer look at the site accounting. PIC

WLCG Accounting TF – CERN – 23/06/2016

J. Flix (PIC/CIEMAT) [remote]
for the PIC team

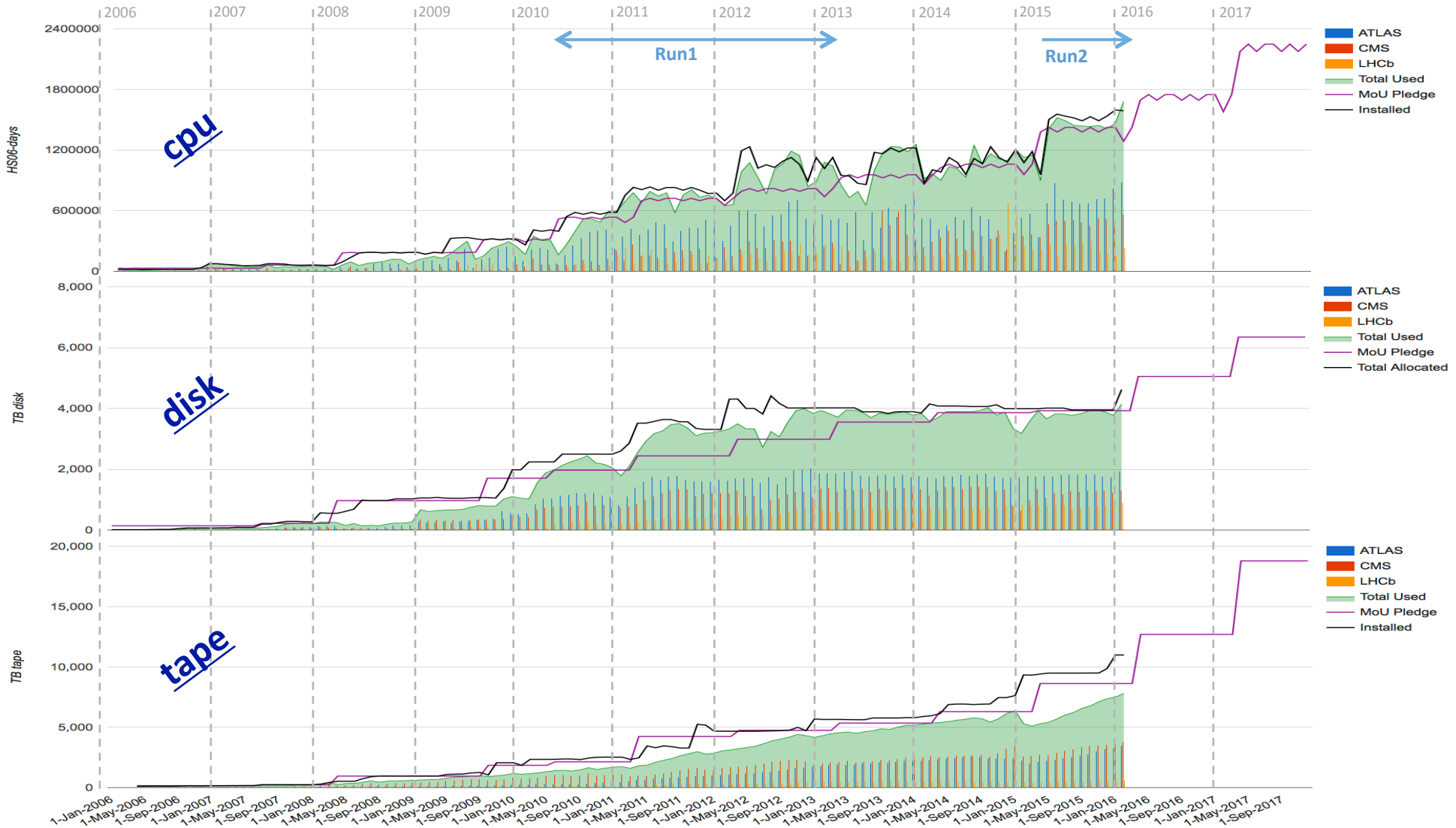
WLCG accounting

- Reliable and accurate accounting is a must
  - These values are extensively used by VOs, sites and funding agencies
  - Essential for scrutinizing the resource usage at the Tiers

- PIC Tier-1 takes the resources accounting seriously:
  - Internal accounting system deployed since many years
  - Monthly x-checks prior reporting to WLCG about resources usage
  - Pro-active in detecting/reporting accounting issues
  - Helping in the integration/debug of new accounting values (such as MultiCore accounting)

- I will (mostly) concentrate in this talk in the **CPU Accounting**

PIC Tier-1 resource accounting 2006-2016



PIC Accounting | Tape | dCache | NFS | Computing | Users | Reports | Admin

Daily
Monthly

Back

Computing Yearly Report

CPU / Wall Time

CPU | Wall | Wall N

2015

VO	Jan 2015	Feb 2015	Mar 2015	Apr 2015	May 2015	Jun 2015	Jul 2015	Aug 2015	Sep 2015	Oct 2015	Nov 2015	Dec 2015	TOTAL
atlas-T1	333552.74	457705.16	476637.27	282703.79	580286.20	738485.38	609705.88	588270.59	612893.98	612753.15	603903.47	521100.55	6417998.17
atlas-T2	174057.36	149710.75	130477.59	72133.48	176417.40	248218.25	154761.53	168612.48	170832.80	168883.99	143828.49	157022.41	1914956.52
atlas-T3	8713.91	8881.49	6519.70	8796.68	11979.67	1383.75	5159.77	5858.66	3281.78	38441.52	28376.85	17978.36	145372.13
cms	185565.86	288459.02	223494.77	202898.82	235948.89	269320.43	293463.55	216276.85	288074.24	298104.88	413735.09	359670.34	3275012.73
cta	0.84	0.83	1.73	2.18	1.96	1.81	0.80	1.91	0.60	0.60	0.93	1.70	15.91
ctaifae	10908.47	1862.09	635.99	1155.22	1137.96	1054.21	5711.89	23771.67	9032.89	3738.41	11851.05	23023.43	93883.27
dteam	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
euclid	2.79	8.91	0.00	1271.07	25.31	875.44	903.06	187.62	217.46	847.66	1007.43	791.97	6138.73
fmat2	0.00	0.00	1883.34	3898.72	1660.93	8368.70	5187.39	3816.46	1340.65	0.00	0.00	0.00	26156.19

CPU Accounting

- All of the WNs CPU usage at PIC are scaled to a reference CPU
  - GlueCECapability: CPUScalingReferenceSI00=3050

Online Worker Nodes

CPU Type	Number of Nodes	Numbre of Slots	HS06	HS06/slot
E5-2640v3 (1.1487)	48	1144	16032.1761	14.01
E5-2640v3 (1.2212)	36	864	12872.4249	14.90
E5-2650-2.00 (0.8586)	11	264	2765.3788	10.47
E5-2650v2-2.60 (1.1011)	32	768	10316.8665	13.43
E5-2650v2-2.60 (1.4848)	16	256	4637.3273	18.11
E5645-2.40 (0.8889)	120	1928	20908.3502	10.84
L5530-2.40 (1.0902)	0	0	0	-
X5650-2.67 (1.1864)	56	672	9726.5817	14.47
TOTAL	**319**	**5896**	**77259.1055**	**13.10**

HT on

We re-run HS06 benchmarks adapting to the offered slots

factor w.r.t reference CPU ($HS06/core_{measured}$ / $HS06/core_{reference}$)

CPU Accounting

- All of the WNs CPU usage at PIC are scaled to a reference CPU
  - GlueCECapability: CPUScalingReferenceSI00=3050
  - We change the reference CPU every ~4/5 years (so it matches the average)

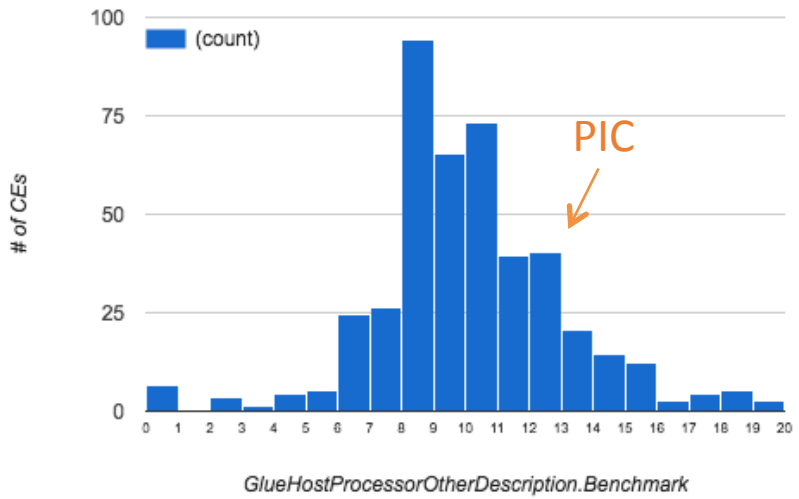
    a) GlueCECapability: CPUScalingReferenceSI00=3050

    b) GlueHostBenchmarkSI00: 3030 (=12.1205*250)

    250 SI00 is 1 HS06

    c) GlueHostProcessorOtherDescription: Cores=8.91,Benchmark=12.1205-HEP-SPEC06

    - We don't treat a), b) or c) as dynamic info



About 5% of the CE's in BDii have Benchmarks > 20 (suspicious)

CPU Accounting

- All of the WNs CPU usage at PIC are scaled to a reference CPU
  - GlueCECapability: CPUScalingReferenceSI00=3050
  - We change the reference CPU every ~4/5 years (so it matches the average)
    a) GlueCECapability: CPUScalingReferenceSI00=3050
    b) GlueHostBenchmarkSI00: 3030 (=12.1205*250)
    c) GlueHostProcessorOtherDescription: Cores=8.91,Benchmark=12.1205-HEP-SPEC06
    - We don't treat a), b) or c) as dynamic info
  - PBS manages all of PIC WNs referenced to this ReferenceCPU
  - These referenced CPU/Wall Clock times are parsed by APEL → sent to EGI accounting portal

CPU Accounting

- The “Sum CPU Time” or “Sum Elapsed CPU time” in EGI Accounting portal shows time in hours, for each site… but the sites have different ReferenceCPUs
  - Good for intra-site checks…
  - … but, these ‘times’ cannot be added, or relative % be evaluated

The following table shows the distribution of Sum CPU time grouped by TIER1 and VO (only information about **LHC VOs** is returned).

Sum CPU time by TIER1 and VO						
TIER1	alice	atlas	cms	lhcb	Total	%
CA-TRIUMF	0	2,754,550	0	0	2,754,550	2.79%
CH-CERN	6,299,187	1,488,417	1,515,481	1,867,570	11,170,655	11.33%
DE-KIT	1,717,676	3,226,270	610,284	1,877,121	7,431,352	7.54%
ES-PIC	0	1,250,735	742,818	589,017	2,582,570	2.62%
FR-CCIN2P3	1,943,360	4,182,375	1,672,661	2,813,140	10,611,536	10.77%
IT-INFN-CNAF	2,078,560	2,429,679	2,274,949	2,593,821	9,377,010	9.51%
KR-KISTI-GSDC	1,892,382	0	0	0	1,892,382	1.92%
NDGF	661,612	1,390,294	0	0	2,051,906	2.08%
NL-T1	1,241,653	2,540,314	0	1,261,807	5,043,774	5.12%
NRC-KI-T1	1,978,024	1,882,311	0	1,146,558	5,006,893	5.08%
RU-JINR-T1	0	0	1,193,619	0	1,193,619	1.21%
TW-ASGC	0	1,651,324	0	0	1,651,324	1.68%
UK-T1-RAL	3,093,349	10,984,268	2,337,671	9,817,572	26,232,861	26.62%
US-FNAL-CMS	0	0	5,037,856	0	5,037,856	5.11%
US-T1-BNL	0	6,517,383	13	0	6,517,396	6.61%
Total	20,905,804	40,297,921	15,385,352	21,966,607	98,555,684	
Percentage	21.21%	40.89%	15.61%	22.29%		
Click here for a CSV dump of this table						
Click here for a Extended CSV dump of this table						
Click here for XML encoded data						

CPU Accounting

- The “Sum CPU Time” or “Sum Elapsed CPU time” in EGI Accounting portal shows time in hours, for each site… but the sites have different ReferenceCPUs
  - Good for intra-site checks…
  - … but, these ‘times’ cannot be added, or relative % be evaluated

The following table shows the distribution of Sum CPU time grouped by TIER1 and VO (only information about **LHC VOs** is returned).

Sum CPU time by TIER1 and VO

TIER1	alice	atlas	cms	lhcb	Total	%
CA-TRIUMF	0	2,754,550	0	0	2,754,550	2.79%
CH-CERN	6,299,187	1,488,417	1,515,481	1,867,570	11,170,655	11.33%
DE-KIT	1,717,676	3,226,270	610,284	1,877,121	7,431,352	7.54%
ES-PIC	0	1,250,735	742,818	589,017	2,582,570	2.62%
FR-CCIN2P3	1,943,360	4,182,375	1,672,661	2,813,140	10,611,536	10.77%
IT-INFN-CNAF	2,078,560	2,429,679	2,274,949	2,593,821	9,377,010	9.51%
KR-KISTI-GSDC	1,892,382	0	0	0	1,892,382	1.92%
NDGF	661,612	1,390,294	0	0	2,051,906	2.08%
NL-T1	1,241,653	2,540,314	0	1,261,807	5,043,774	5.12%
NRC-KI-T1	1,978,024	1,882,311	0	1,146,558	5,006,893	5.08%
RU-JINR-T1	0	0	1,193,619	0	1,193,619	1.21%
TW-ASGC	0	1,651,324	0	0	1,651,324	1.68%
UK-T1-RAL	3,093,349	10,984,268	2,337,671	9,817,572	26,232,861	26.62%
US-FNAL-CMS	0	0	5,037,856	0	5,037,856	5.11%
US-T1-BNL	0	6,517,383	13	0	6,517,396	6.61%
Total	20,905,804	40,297,921	15,385,352	21,966,607	98,555,684	
Percentage	21.21%	40.89%	15.61%	22.29%		

Click here for a CSV dump of this table

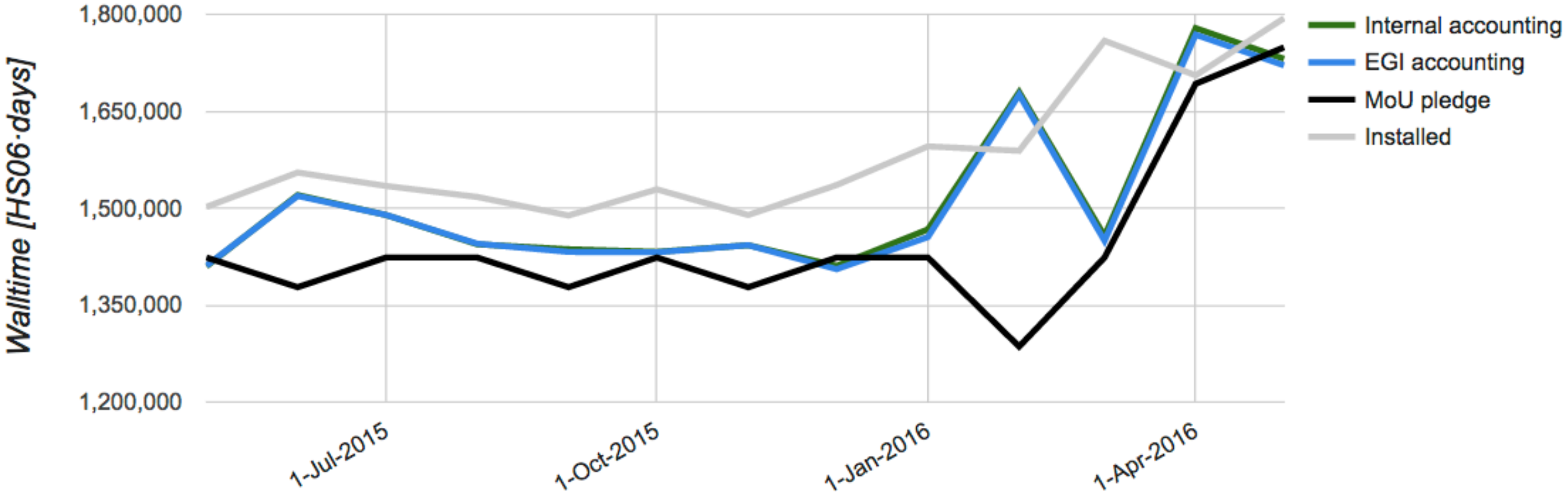
Click here for a Extended CSV dump of this table

Click here for XML encoded data

CPU Accounting

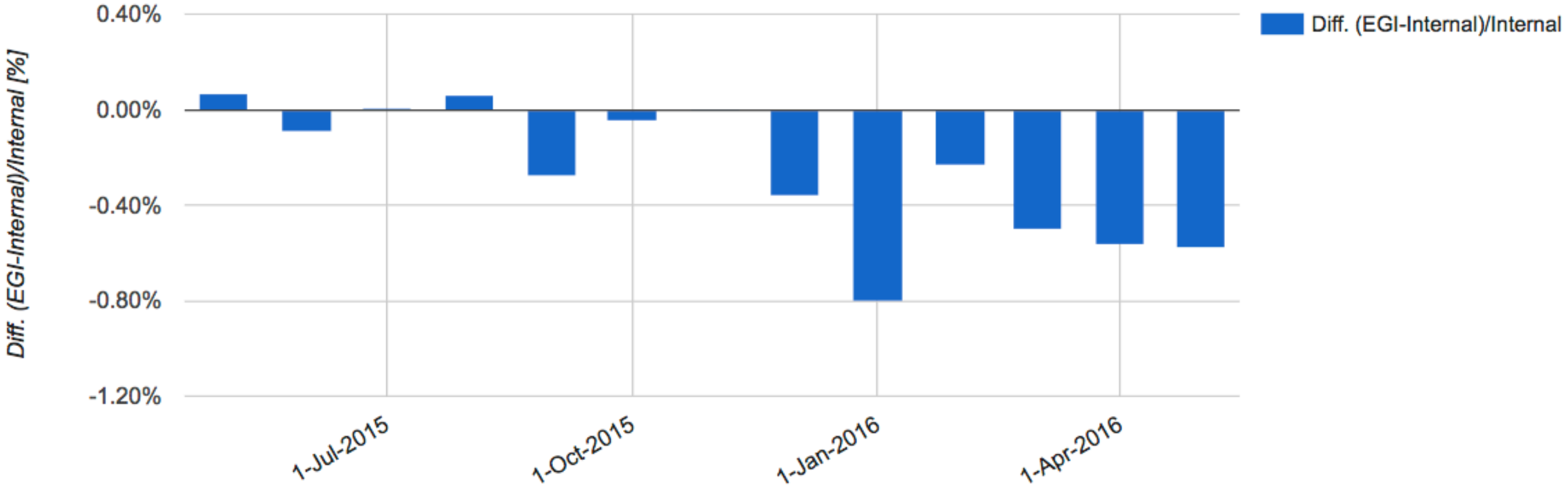
- GlueCECapability: CPUScalingReferenceSI00 is used in the EGI accounting portal to calculate the HS06·hours for all of the sites
  - Then, things are in the same units, can be added, relative % can be evaluated…

PIC Tier-1 CPU usage (WallTime)



Rather good agreement between PIC internal accounting and values from the EGI portal

PIC Tier-1 CPU usage (WallTime)



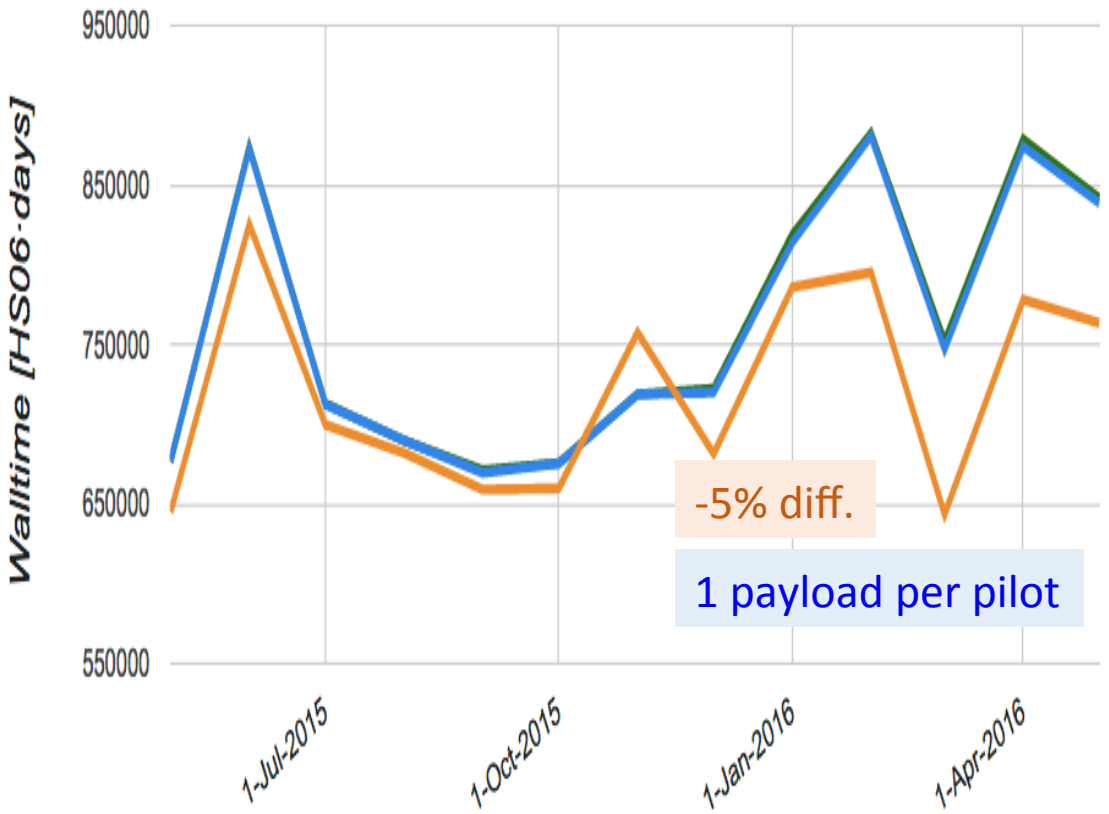
Generally, the differences are < 0.5%

- Same CREAM-ID assigned for different PBS jobs
  → error of the CREAM system, maybe in high load cases, not so often
- APEL works in UTC, our local monitoring in CEST

Dashboard accounting

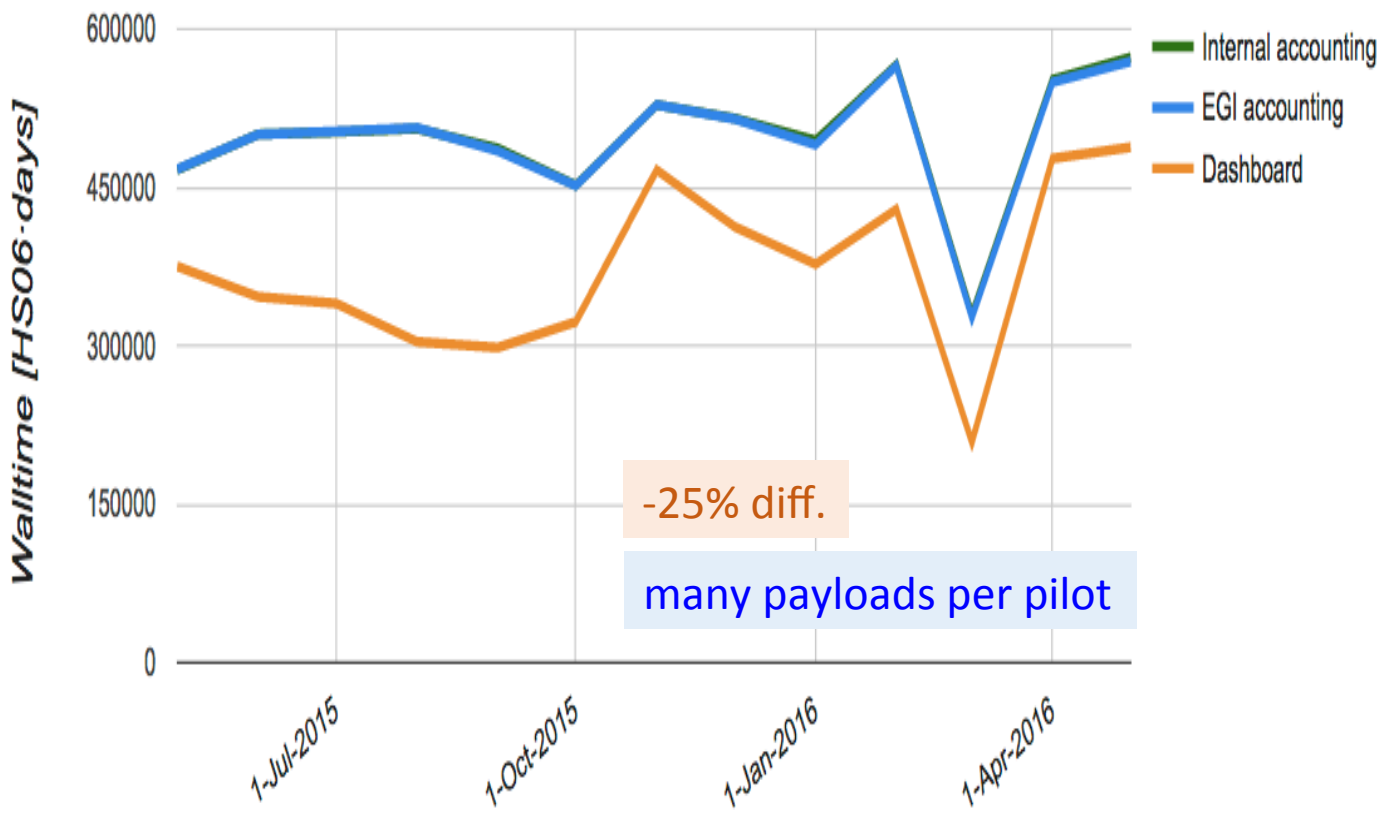
- For those VOs using payloads, it holds payloads information (not pilots)
- The CPU and Wall Times are not scaled
- The average HS06 reported by the farms in BDii are used to get HS06·hours
  - This is not very accurate
  - payloads (dashboard) cannot be directly compared to pilots (EGI accounting portal)
    - ATLAS and CMS models are very different

ATLAS CPU usage at PIC Tier-1 (WallTime)



ATLAS Wallclock consumption ALL jobs (days)	775288
ATLAS Wallclock HS06·days	9374827
ratio	12.09

CMS CPU usage at PIC Tier-1 (WallTime)



CMS Wallclock consumption ALL jobs (days)	389868
CMS Wallclock HS06·days	4845227
ratio	12.43

PIC GlueHostProcessorOtherDescription: Cores=8.91,Benchmark=12.1205-HEP-SPEC06 ??

Conclusions

- PIC internal accounting matches well with EGI accounting
  - If not, GGUS tickets are opened and addressed
- We should have ways to automatically check if a site is reporting wrong values in the EGI accounting portal
  - **We should eliminate any portal view that is controversial** (not all of the people know the story of scaled times and that sites cannot be compared in EGI portal)
    - If something cannot be compared, we shouldn't show it!
  - Still some views are wrong in the EGI accounting portal
    - Reports > WLCG Tier1 / Tier2 don't report Elapsed times * number of processors
- The dashboard procedure to evaluate HS06·hours is **not accurate**
  - Sites should deploy the MJF, then the payloads can directly report accurate HS06·hours values to the dashboard