

Status since last meeting - Operations

➔ **Generally troubled operations across the board** (ce01, ce02)

- Ailing hardware, reduced capacity, cooling instabilities, many issues with ARC, constantly catching up with issues, delays in deployment of new hardware, etc.
- ATLAS quiet until March, then startup was quite hectic; took a while to tune the task definitions (mem, cpu, wt) => large numbers of job failures. ARC troubles did the rest.
- Re-gained stability from July

➔ **ARC upgraded from 4.2.0-1.el6 to 5.0.2-1.el6** (via 5.0.0-2.el6) (smooth upgrades on live system)

- 5.0.0-2.el6 (mid-April): major upgrade also introduced major bugs.
- Hard to reproduce, sometimes depending on the load on the filesystem or other obscure factors
- Skipped 5.0.1-1.el6 (still bugged), eventually 5.0.2-1.el6 (beginning of July) settled the issues

➔ **Cluster re-installation with ROCKS 6.2** (released beginning of July)

- Support for newest hardware
- Full chain prototyped: ARC CE, mds, oss, wn, IB, lustre.
- All node definitions and wn image ready. Only slurm setup to be done

➔ **ce03 idle since Feb 2015** (targeting Todi at CSCS)

HammerCloud Gangarobot

History Legend

■ offline
 ■ brokeroff
 ■ online
 ■ NoQueue
 ■ test

Historic view for "panda_queues_all" from 00:00 01.01.2015 to 00:00 18.08.2015

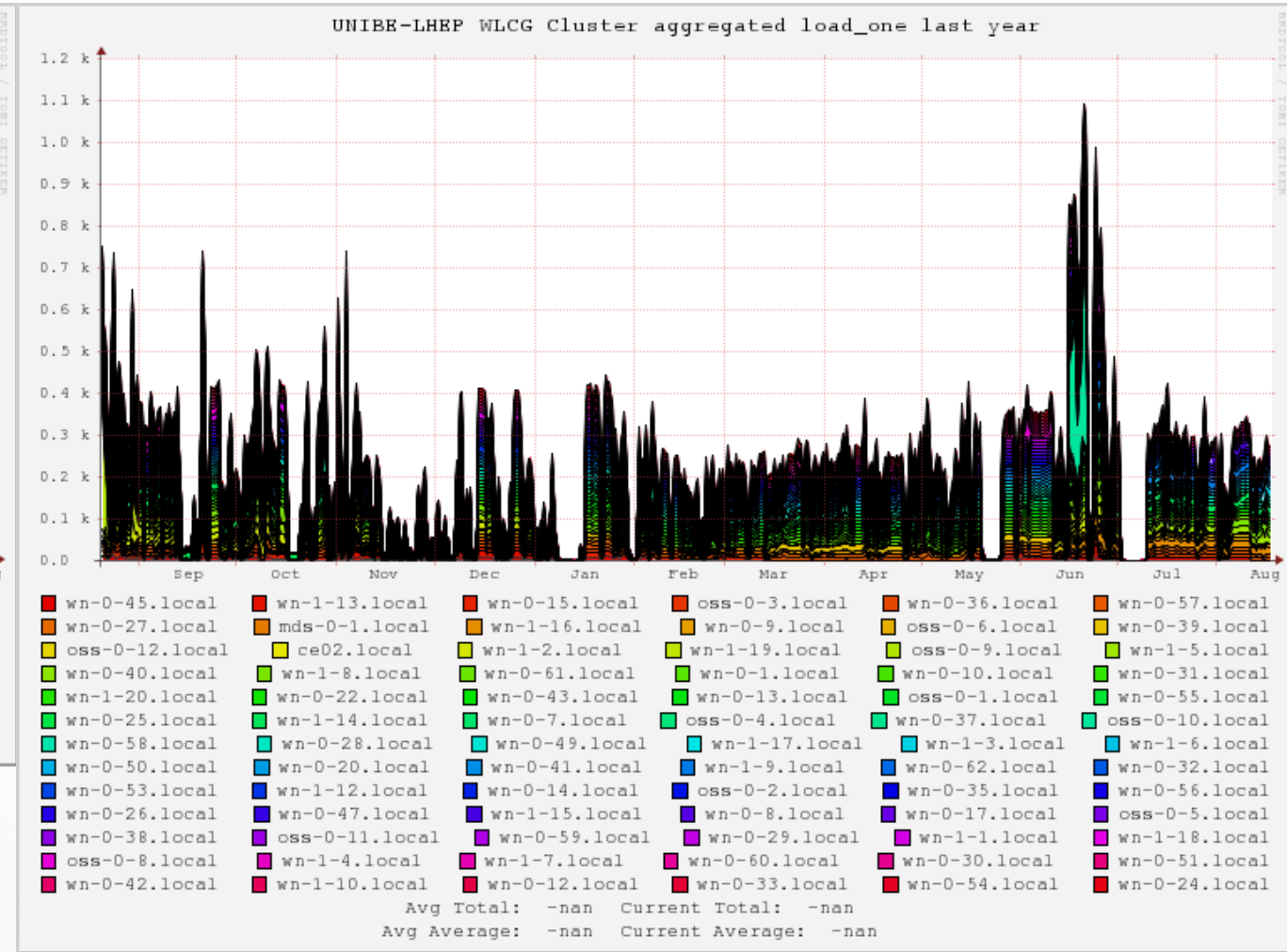
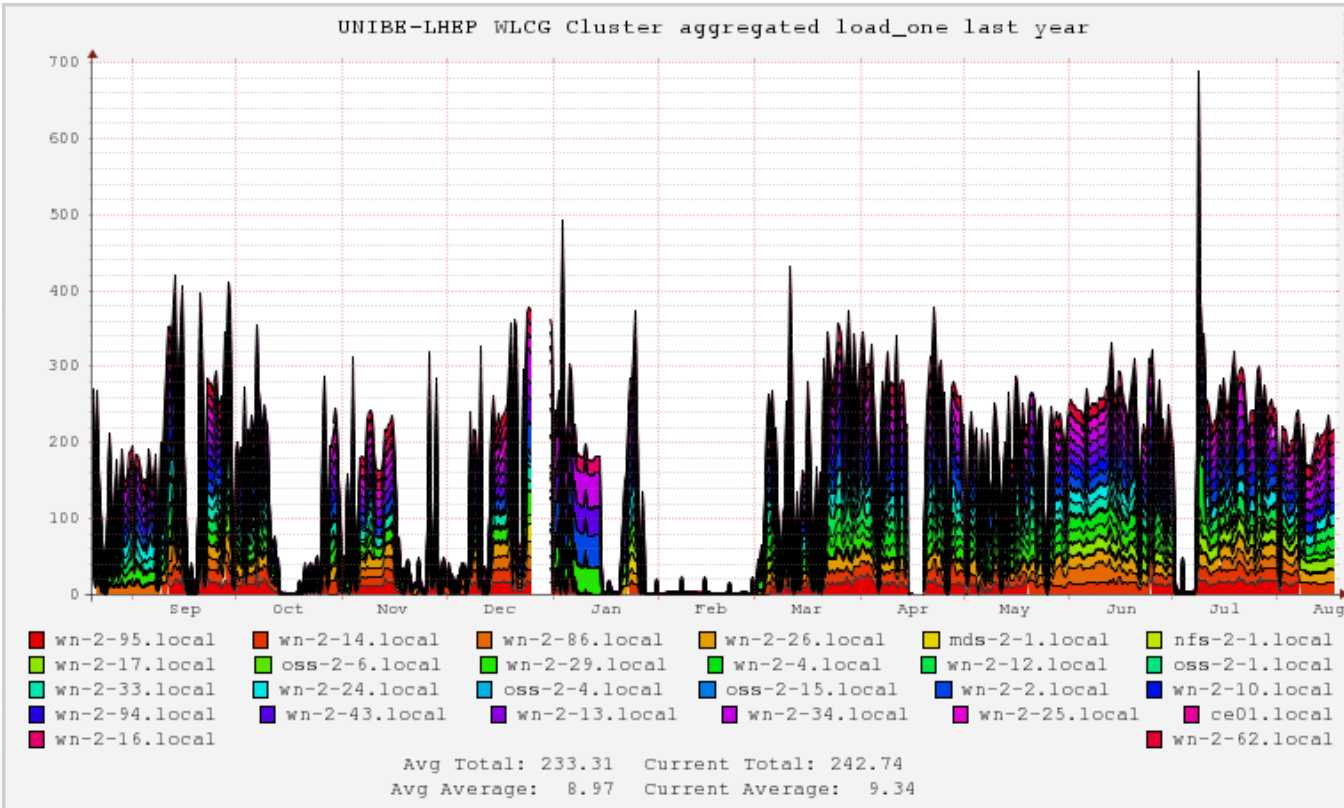
Show entries Search:

PANDA queue	SITE Name	TIER	CLOUD	History plot time bin = 458 hours	offline		brokeroff		online		NoQueue		test	
					%	count	%	count	%	count	%	count	%	count
ANALY_CSCS	CSCS-LCG2	T2D	DE		0.93	4	0.15	2	94.88	54	0	0	3.59	41
ANALY_CSCS_GLEXEC	CSCS-LCG2	T2D	DE		0	0	0	0	0	0	0	0	99.55	16
CSCS-LCG2-all-prod-CEs	CSCS-LCG2	T2D	DE		1.23	4	0	0	94.87	53	0	0	3.46	39
CSCS-LCG2-all-prod-CEs_MCORE	CSCS-LCG2	T2D	DE		1.23	4	0	0	94.87	53	0	0	3.46	39
UNIBE-LHEP	UNIBE-LHEP	T2	ND		2.21	2	0	0	79.93	27	0	0	17.4	37
UNIBE-LHEP-UBELIX	UNIBE-LHEP	T2	ND		0.05	1	0	0	97.94	30	0	0	1.56	15
UNIBE-LHEP-UBELIX_MCORE	UNIBE-LHEP	T2	ND		0.05	1	0	0	66.58	15	0	0	1.24	13
UNIBE-LHEP_MCORE	UNIBE-LHEP	T2	ND		0.05	1	0	0	61.62	15	0	0	6.19	13

Showing 1 to 8 of 8 entries First Previous 1 Next Last

ce01

ce02



Issues and mitigations (1/3)

➔ A range of new issues introduced by ARC 5

- Job priority out of range for gridengine in submit-sge-jobs
- CE dropping out of infosys regularly (all services running): add cron every 4h
- a-rex crashes (these were actually introduced by ARC 4)
- Bug (?) in a-rex causing the a-rex infoprovider to stop updating the bdii (cluster drops out of the GIIS). Manual workaroud: crons to restart a-rex, then nordugrid-arc-ldap-infosys
- grid-manager.log exploding: full partitions, recovery very fiddly/time-consuming (bug in a-rex)
- controldir stuffed with over 100k files (0k) and as many directories in "joblinks".

```
[root@ce01 joblinks]# for i in `find /grid/lustre/cache/joblinks/ -maxdepth 1 -type d -mtime +14 -print`; \
do echo -e "Deleting directory $i";rm -rf $i; done
```

- cache-clean <defunct> processes(also arched and gridftpd): all night to recover
- obscure corruption in control dir (prevents a-rex from starting)

```
[root@ce01 ~]# rm -rf /var/spool/nordugrid/jobstatus/gm.fifo
```


Issues and mitigations (2/3)

➔ Problems with multicore job scheduling:

- Multicore jobs waiting in the queue and eventually cancelled by Panda (not 100%)
- Probably due to the introduction of some longer reconstruction jobs (to improve cpu/wt efficiency)
 - => cron to `qalter -R y` for the mcore jobs in the queue
 - => allow 20 reservations: `# qconf -ssconf|grep reservation`
`max_reservation 20`

Hardware issues

LAN network upsets on both ce01 and ce02 (unrelated)

- 28 Jan (ce02) and 24 Apr (ce02): resolved by power-cycling the switches

Glitch on the MDS (ce02)

- `sd 6:2:0:0: rejecting I/O to offline device` (resolved by power-cycling)

GB NIC failure on ce01 (Friday issue :-)

- Recovery:
 - => swap to unused network interface
 - => register network and interface changes in ROCKS
 - => redeploy Lustre from scratch
 - => power-cycle and re-install stuck nodes

Increasing failure rate of flash cards on SunBlades

- Obtained 22 spares from UniZH

Issues and mitigations (3/3)

➔ **Cooling instabilities:**

- **Several incidents over the summer**

- Cooling water is normally supplied to the room at temperature of 7.5 degree

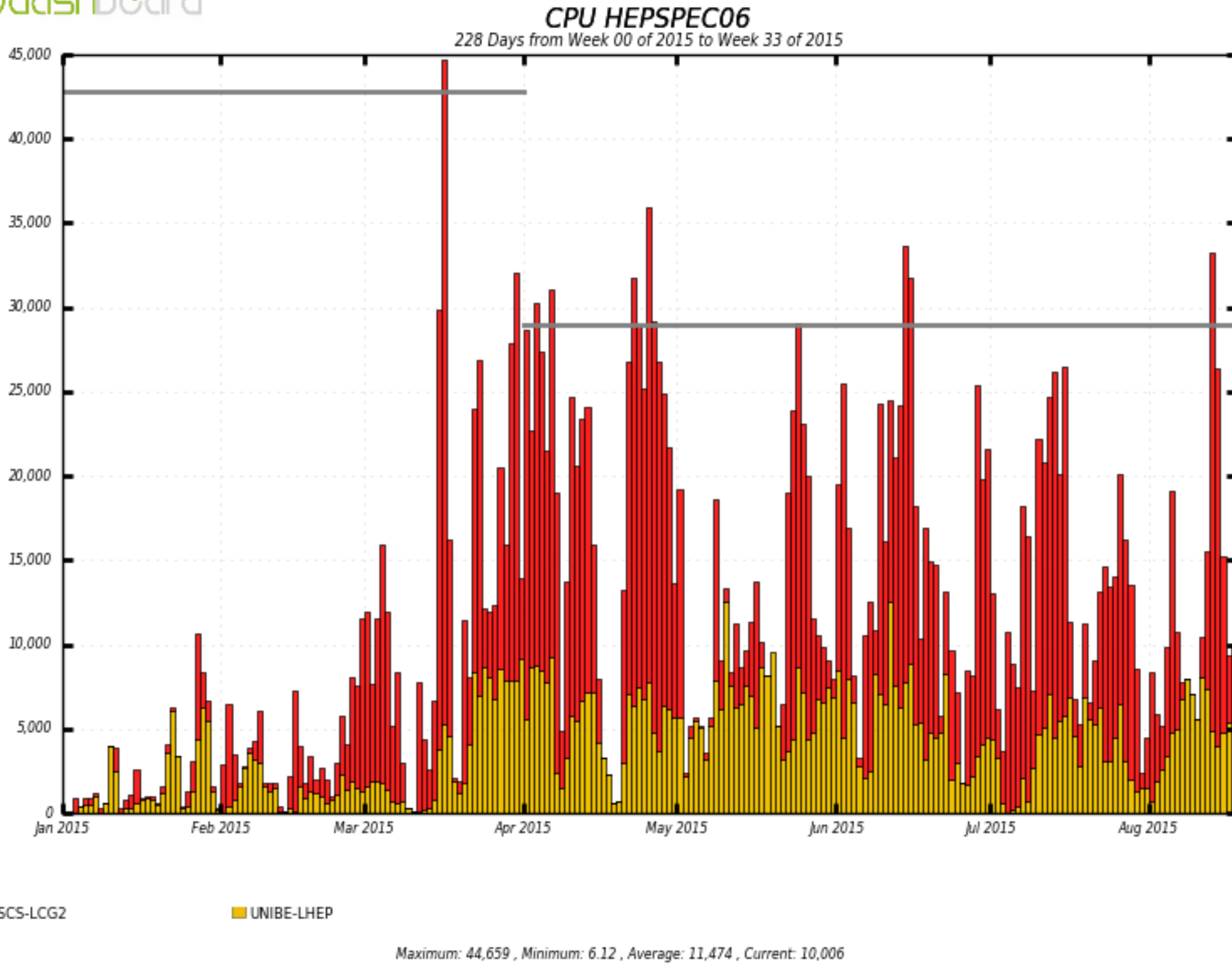
- When supply is at higher temperature, cooling degrades and water cooled racks are impacted first:

- => nodes shutdown and/or power trips

- Water supply quality is out of our control and we receive almost no communication from the responsible(s) in house

- Looking at improving this

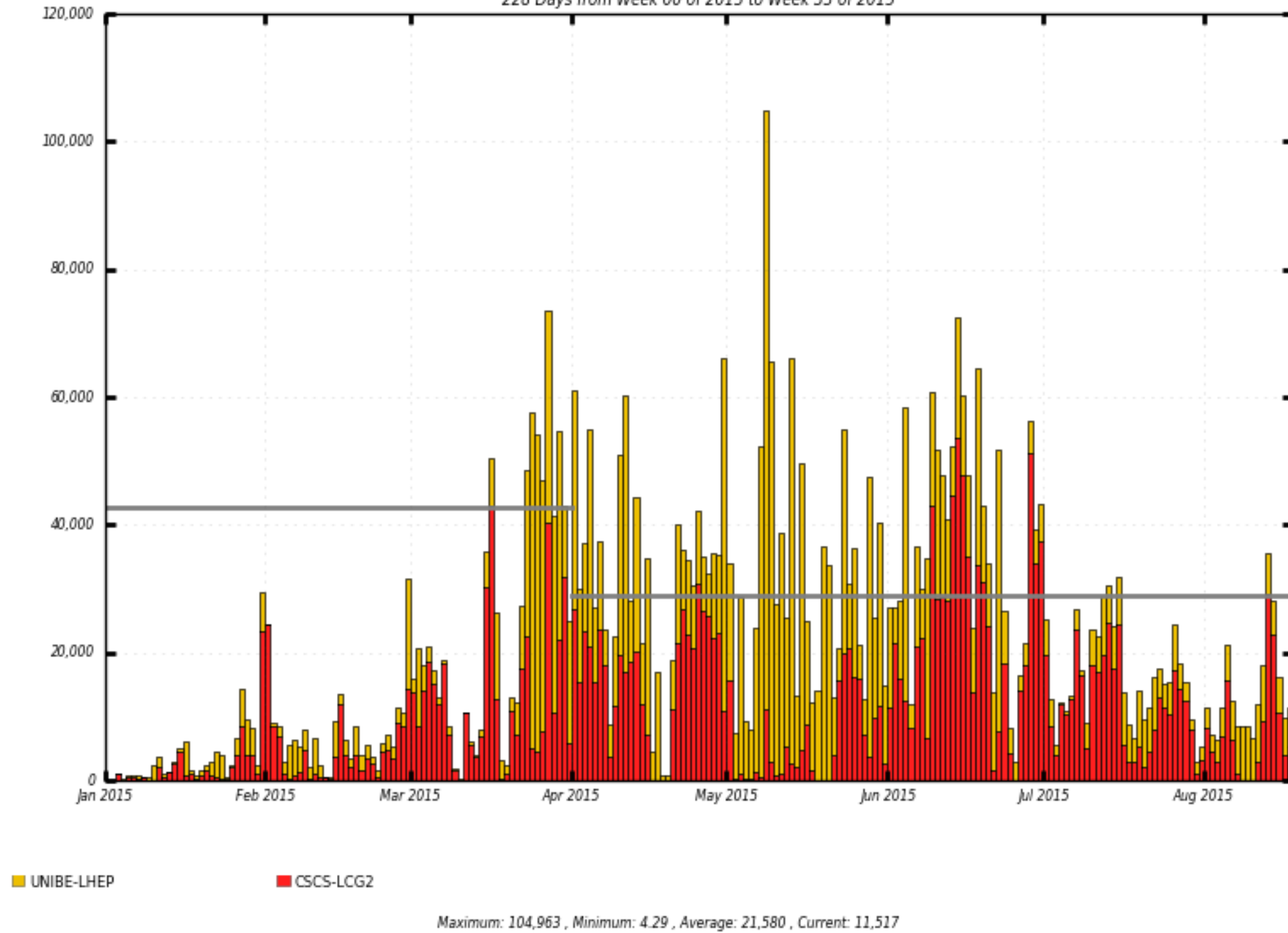
- In the meanwhile we are developing our own temperature monitoring system, with the ambition to automate some actions in case of further incidents





WallClock HEPSPC06

228 Days from Week 00 of 2015 to Week 33 of 2015

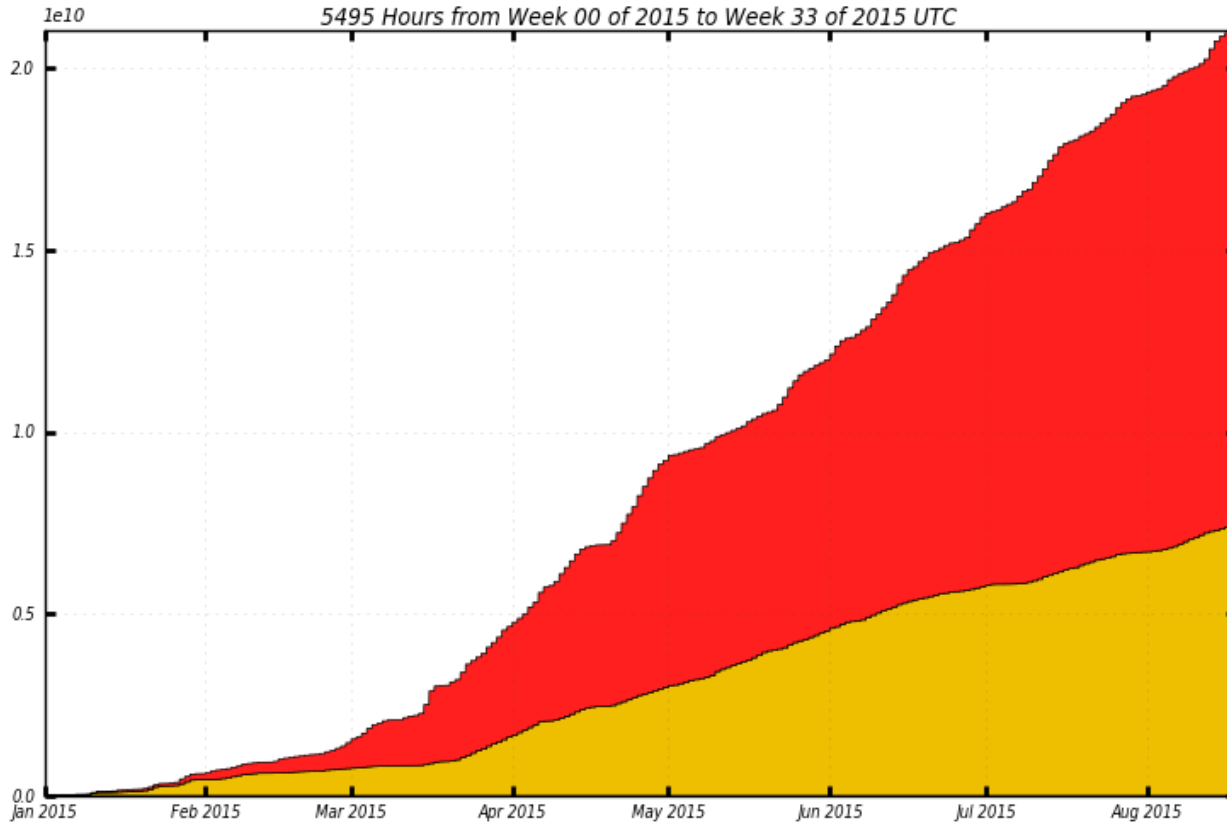




dashboard

CPU consumption Good Jobs in seconds

5495 Hours from Week 00 of 2015 to Week 33 of 2015 UTC



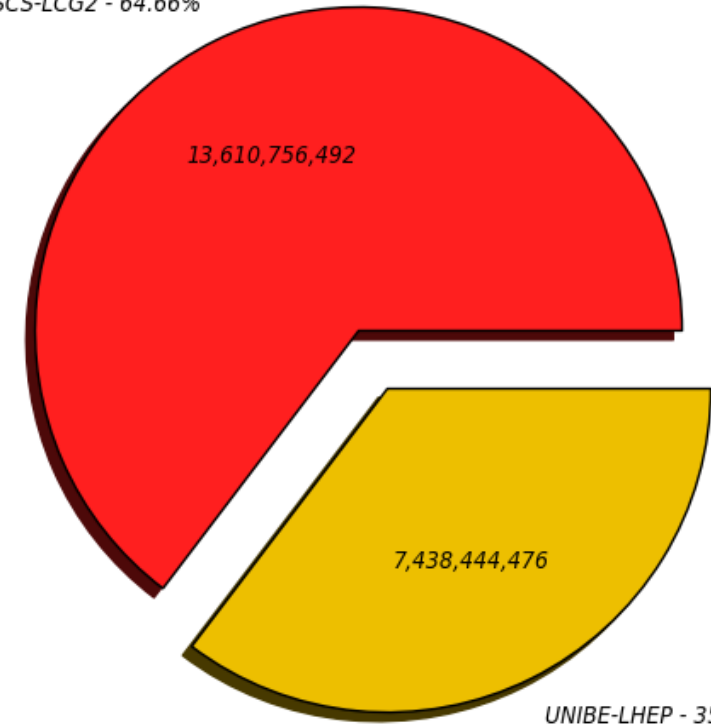
■ CSCS-LCG2 (13,610,756,492) ■ UNIBE-LHEP (7,438,444,476)

Total: 21,049,200,968 , Average Rate: 1,063 /s

dashboard

CPU consumption Good Jobs in seconds (Sum: 21,049,200,968)

CSCS-LCG2 - 64.66%



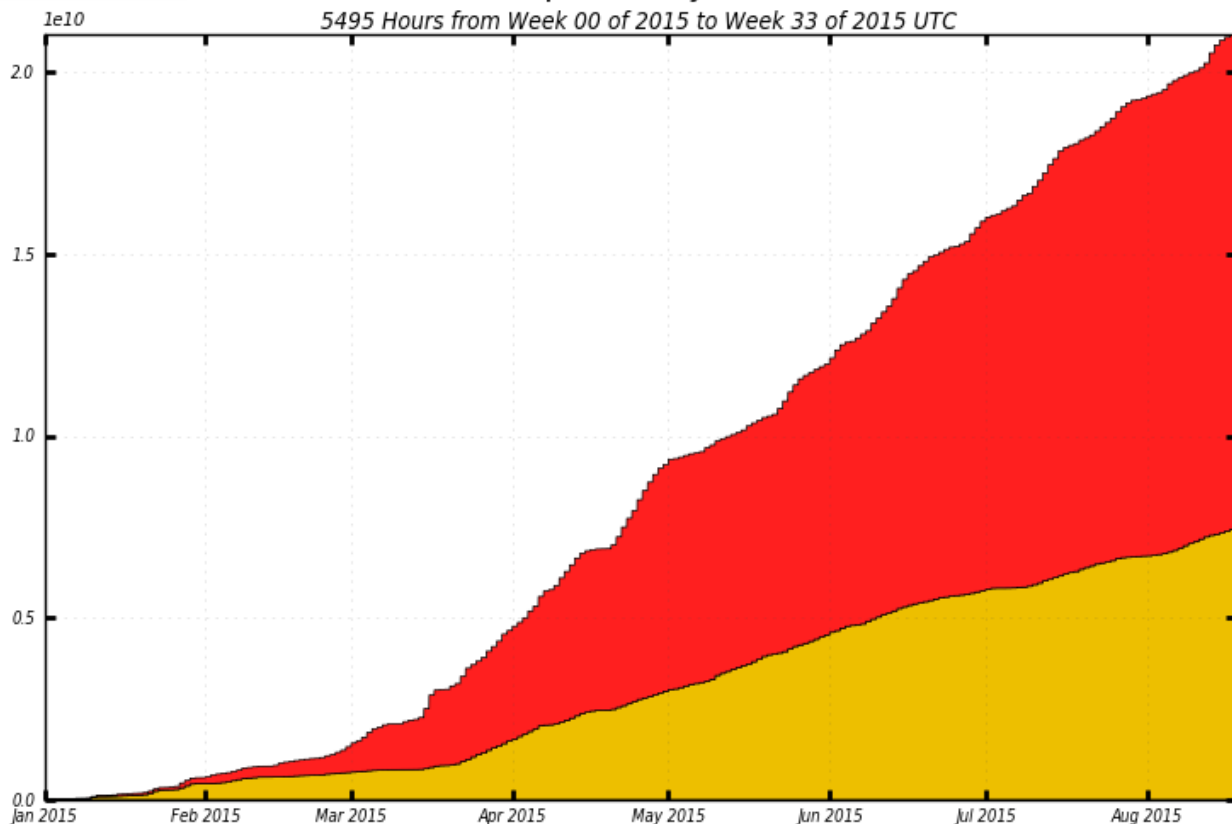
■ CSCS-LCG2 - 64.66% (13,610,756,492)

■ UNIBE-LHEP - 35.34% (7,438,444,476)



CPU consumption Good Jobs in seconds

5495 Hours from Week 00 of 2015 to Week 33 of 2015 UTC



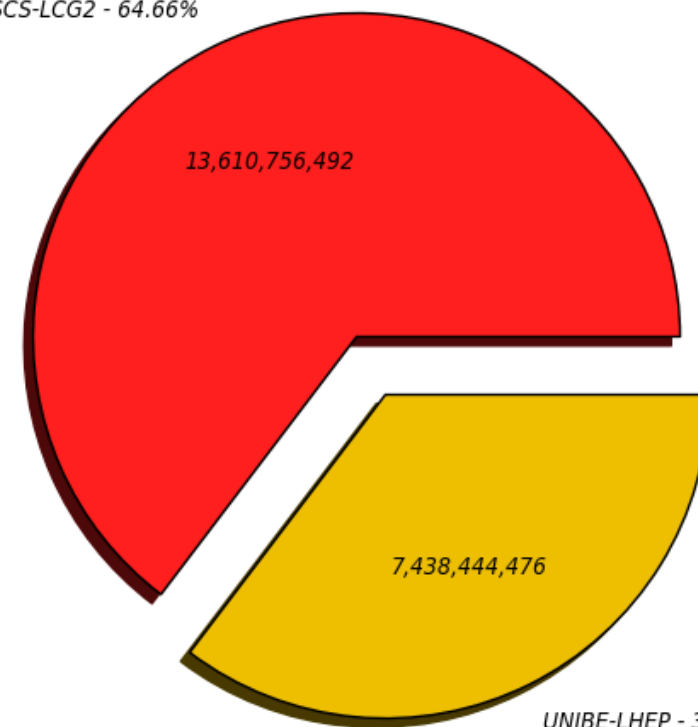
■ CSCS-LCG2 (13,610,756,492) ■ UNIBE-LHEP (7,438,444,476)

Total: 21,049,200,968 , Average Rate: 1,063 /s



CPU consumption Good Jobs in seconds (Sum: 21,049,200,968)

CSCS-LCG2 - 64.66%



■ CSCS-LCG2 - 64.66% (13,610,756,492)

■ UNIBE-LHEP - 35.34% (7,438,444,476)

Switzerland Normalised CPU time (kSI2K) by SITE and DATE.

CUSTOM VOs. January 2015 - August 2015.

The following table shows the distribution of Normalised CPU time (kSI2K) grouped by SITE and DATE (only information about the selected VOs is returned).

Normalised CPU time [units 1K.SI2K.Hours] by SITE and DATE

SITE	Jan 15	Feb 15	Mar 15	Apr 15	May 15	Jun 15	Jul 15	Aug 15	Total	%
CSCS-LCG2	836,523	1,161,810	2,138,947	3,146,993	1,388,701	2,647,161	2,873,324	1,448,033	15,641,492	70.74%
UNIBE-LHEP	584,036	293,640	802,492	940,100	746,960	938,340	557,467	1,605,851	6,468,886	29.26%
Total	1,420,559	1,455,450	2,941,439	4,087,093	2,135,661	3,585,501	3,430,791	3,053,884	22,110,378	
Percentage	6.42%	6.58%	13.30%	18.48%	9.66%	16.22%	15.52%	13.81%		

[Click here for a CSV dump of this table](#)

[Click here for a Extended CSV dump of this table](#)

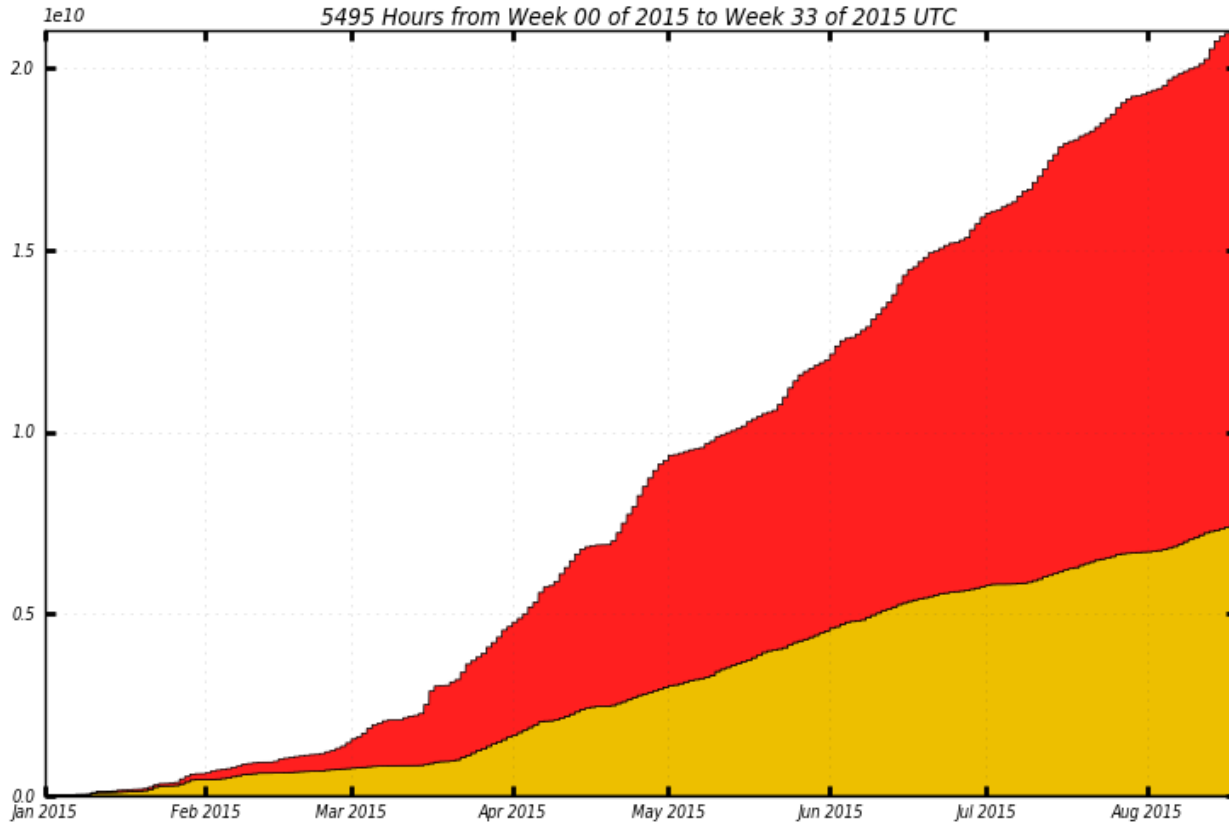
[Click here for XML encoded data](#)



dashboard

CPU consumption Good Jobs in seconds

5495 Hours from Week 00 of 2015 to Week 33 of 2015 UTC



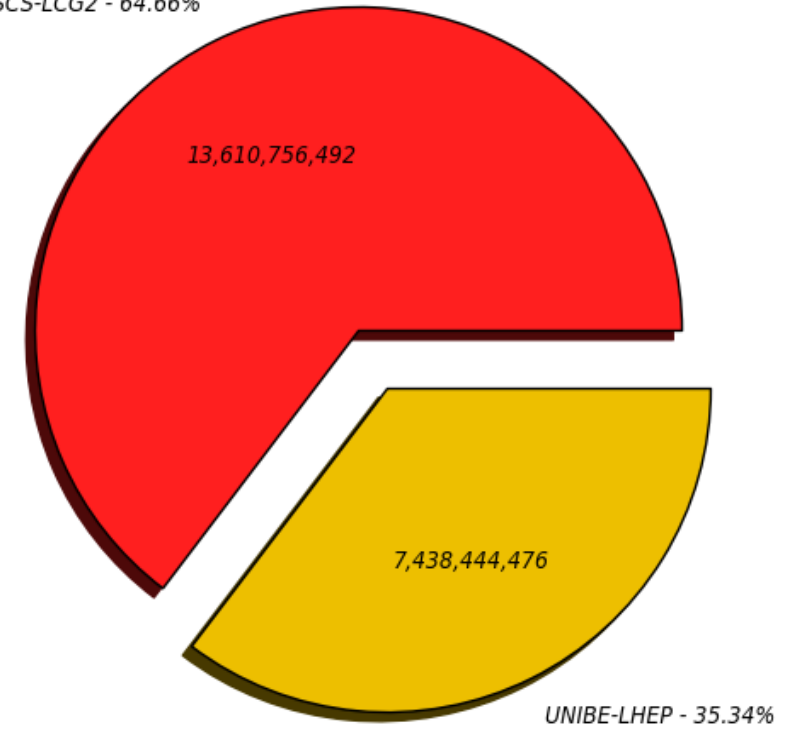
■ CSCS-LCG2 (13,610,756,492) ■ UNIBE-LHEP (7,438,444,476)

Total: 21,049,200,968 , Average Rate: 1,063 /s

dashboard

CPU consumption Good Jobs in seconds (Sum: 21,049,200,968)

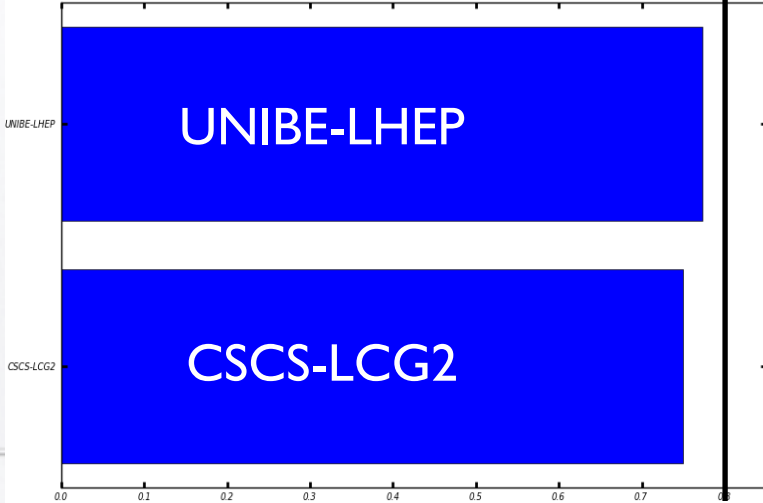
CSCS-LCG2 - 64.66%



■ CSCS-LCG2 - 64.66% (13,610,756,492) ■ UNIBE-LHEP - 35.34% (7,438,444,476)

dashboard

Average Efficiency Good Jobs



0.8

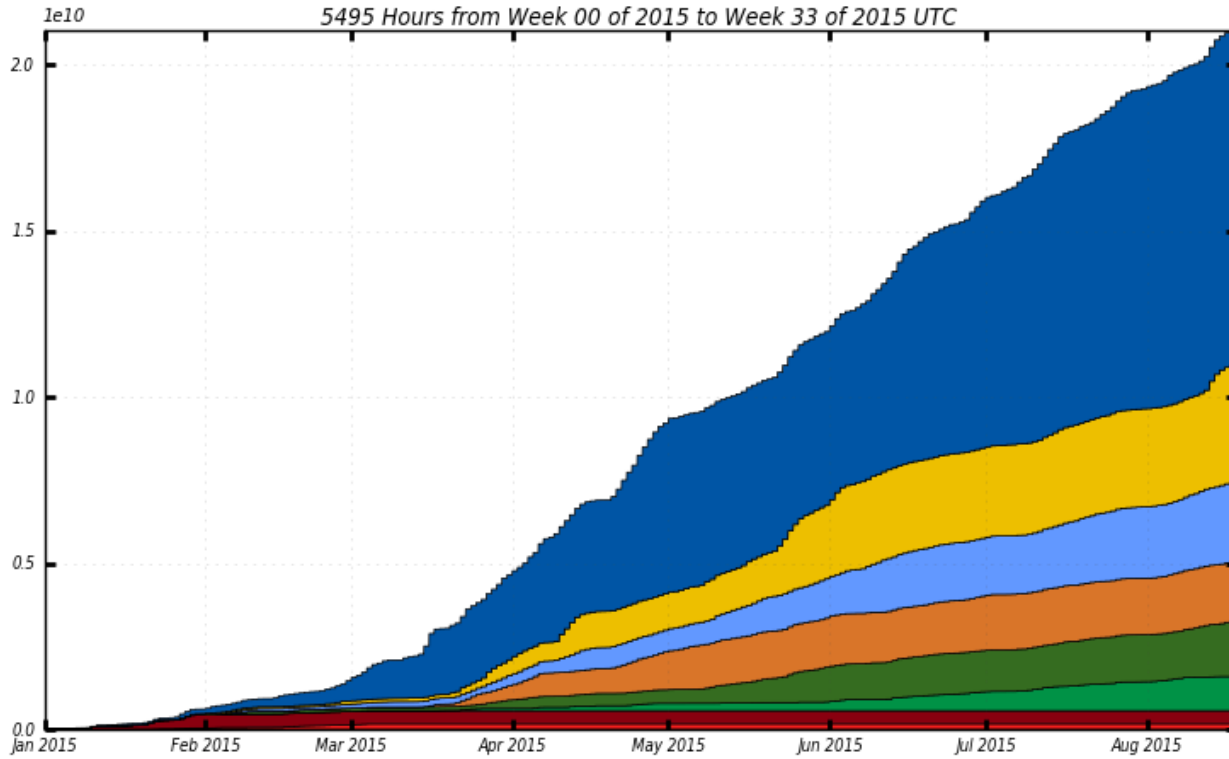
WallClock consumption in seconds





CPU consumption Good Jobs in seconds

5495 Hours from Week 00 of 2015 to Week 33 of 2015 UTC

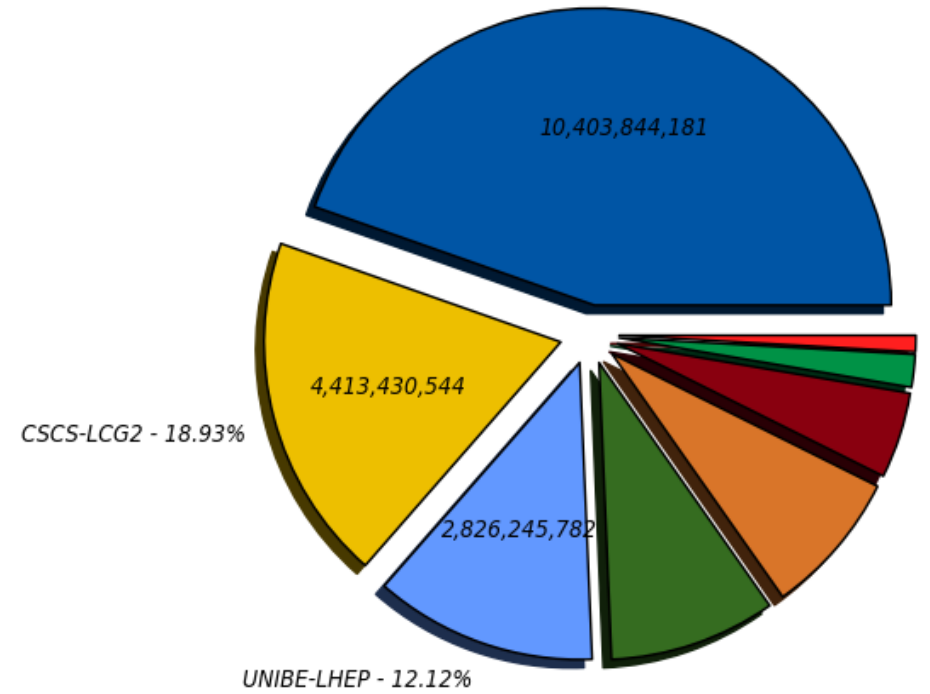


- CSCS-LCG2_MCORE (10,073,300,929)
- UNIBE-LHEP (2,401,542,986)
- UNIBE-LHEP_UBELIX (1,659,725,330)
- CSCS-TODI (401,527,498)
- CSCS-LCG2 (3,537,455,563)
- UNIBE-LHEP_MCORE (1,789,387,203)
- UNIBE-LHEP_UBELIX_MCORE (1,026,923,217)
- ARC_MCORE (159,338,242)

Total: 21,049,200,968 , Average Rate: 1,063 /s



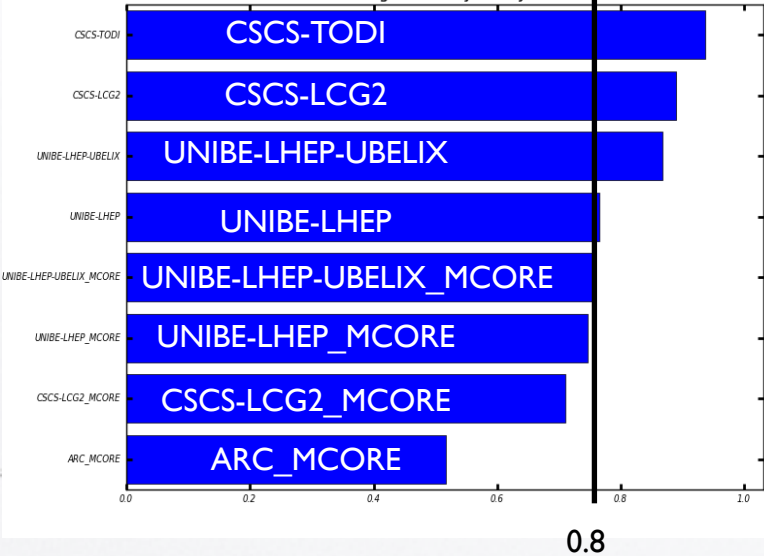
CPU consumption All Jobs in seconds (Sum: 23,315,520,288)
CSCS-LCG2_MCORE - 44.62%



- CSCS-LCG2_MCORE - 44.62% (10,403,844,181)
- UNIBE-LHEP - 12.12% (2,826,245,782)
- UNIBE-LHEP_UBELIX - 7.89% (1,838,788,042)
- CSCS-TODI - 1.78% (414,727,132)
- CSCS-LCG2 - 18.93% (4,413,430,544)
- UNIBE-LHEP_MCORE - 9.09% (2,118,380,337)
- UNIBE-LHEP_UBELIX_MCORE - 4.68% (1,092,146,868)
- ARC_MCORE - 0.89% (207,957,402)

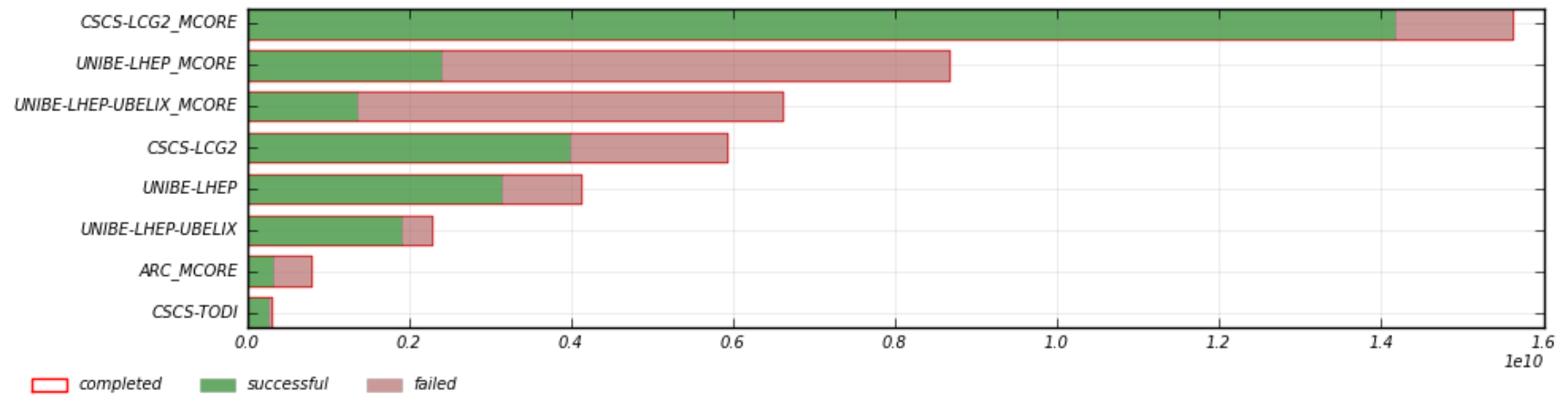


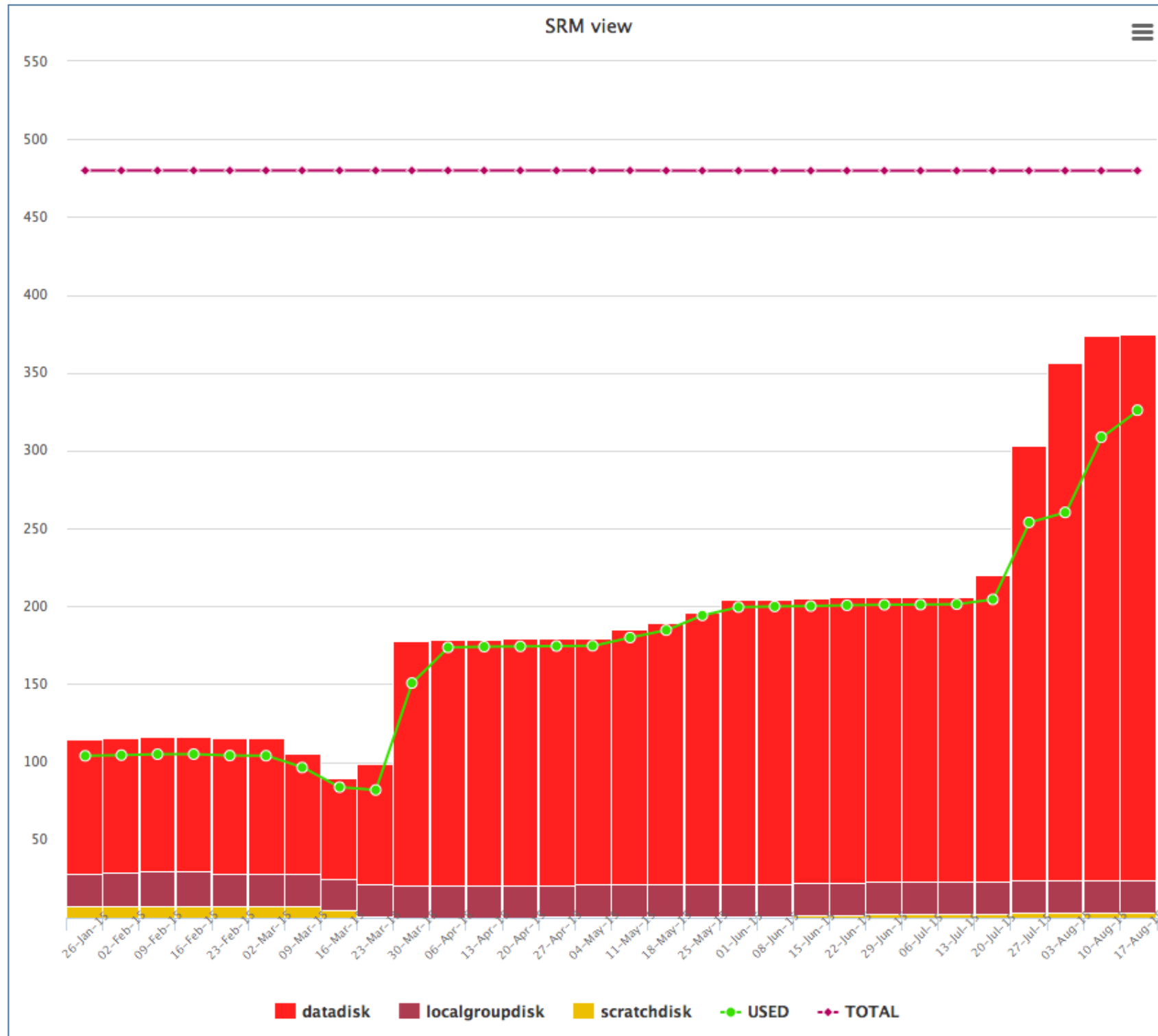
Average Efficiency Good Jobs



0.8

WallClock consumption in seconds





Plans for the future

➔ Re-gain/maintain stability and restore capacity

• Re-installation of both clusters with ROCKS 6.2

- Re-build from scratch of both CEs
- Move to SLURM (better control over resources e.g. mem)
- Support for newest hardware:
 - => 6 new disk servers (204 spindles) for Lustre to replace Thumpers
 - => ~1300 HEPSPEC06 (purchased) + 1300/2600 (to be procured) + 2600/1300 (to be procured in 2016)
- Additional hardware:
 - => 1300 HEPSPEC06 (6 IBM decommissioned from CSCS)
- Restore to working order >~3000 HEPSPEC06 (SunBlades)

• Temperature monitoring and remote management network

- Monitoring of the inlet water temperature and several ambient sensors at top of racks
- Automate orderly shutdown or power-off of WN's when above some threshold
- System will need some training
- Will need a remote management network