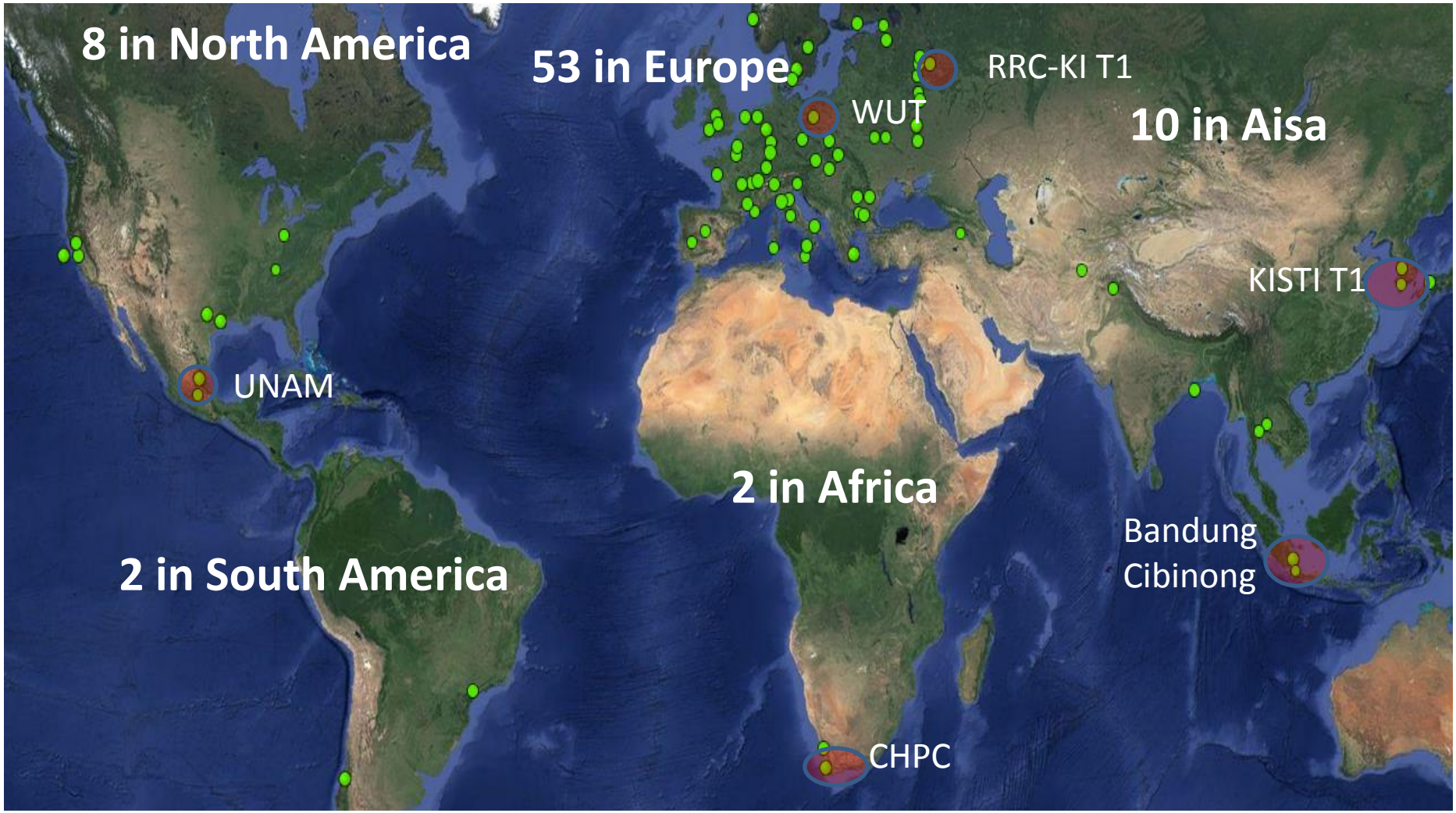
Grid operations in 2014

ALICE Offline week

20 March 2015

Latchezar Betev

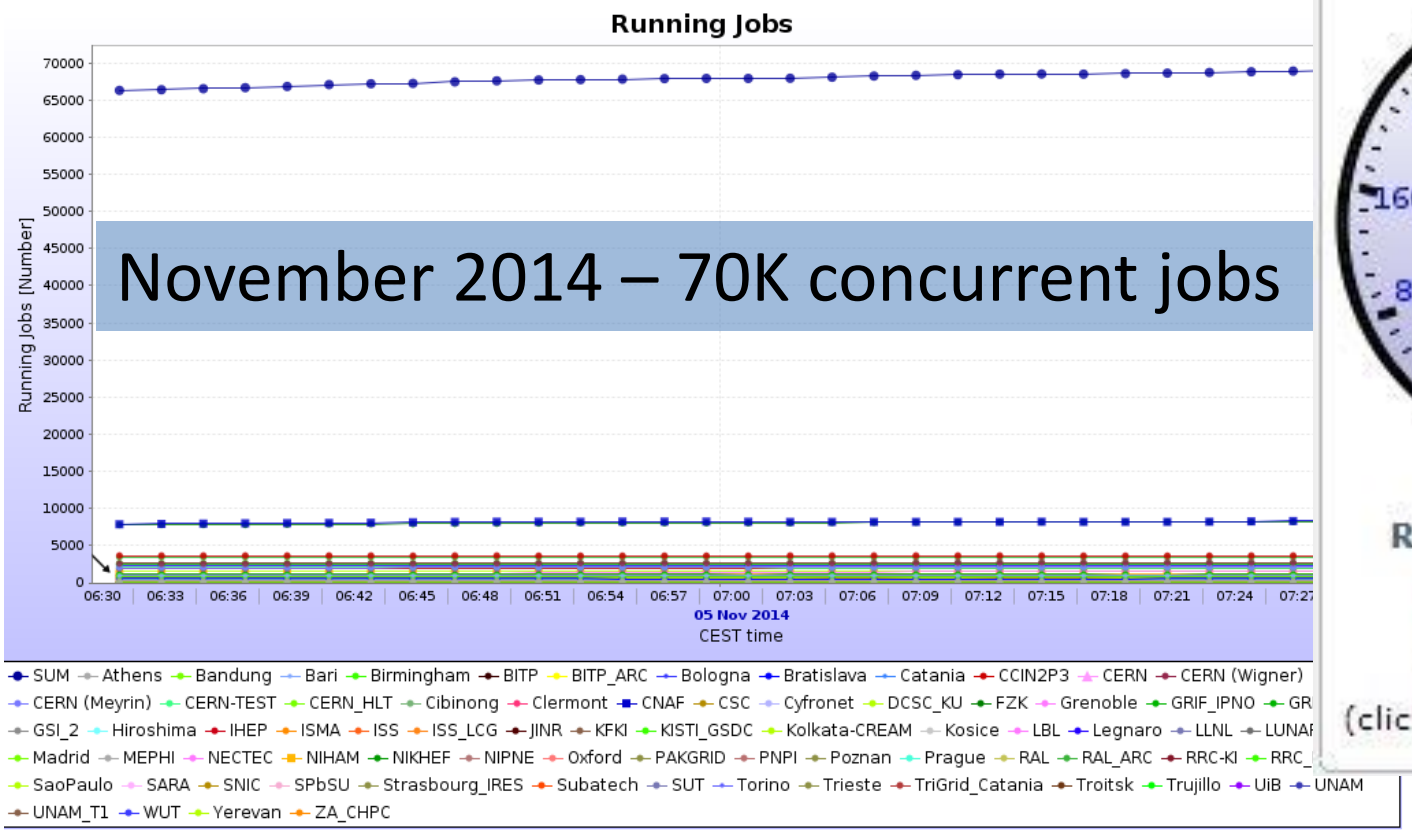
The ALICE Grid sites today



New sites

- KISTI – officially a T1 in WLCG
- UNAM – MoU for T2 in November 2014, towards a T1
- WUT (Poland) in production September 2014
- RRC-KI T1 in production January 2014
- ZA_CHPC x4 capacity in November 2014
- Bandung and Cibinong in production September 2014

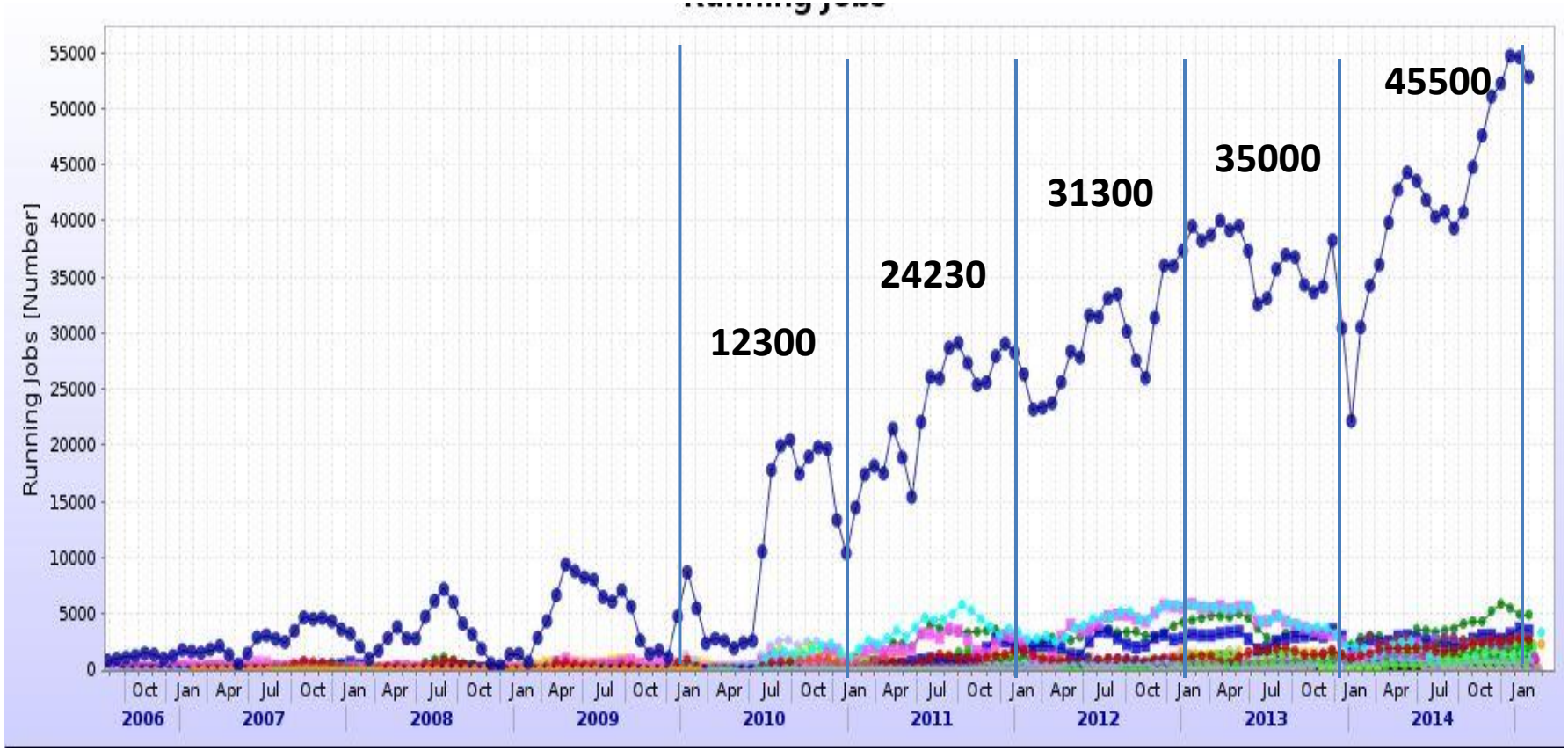
A new job record



November 2014 – 70K concurrent jobs

… and we are breaching 72K jobs every week since February

CPU resources evolution

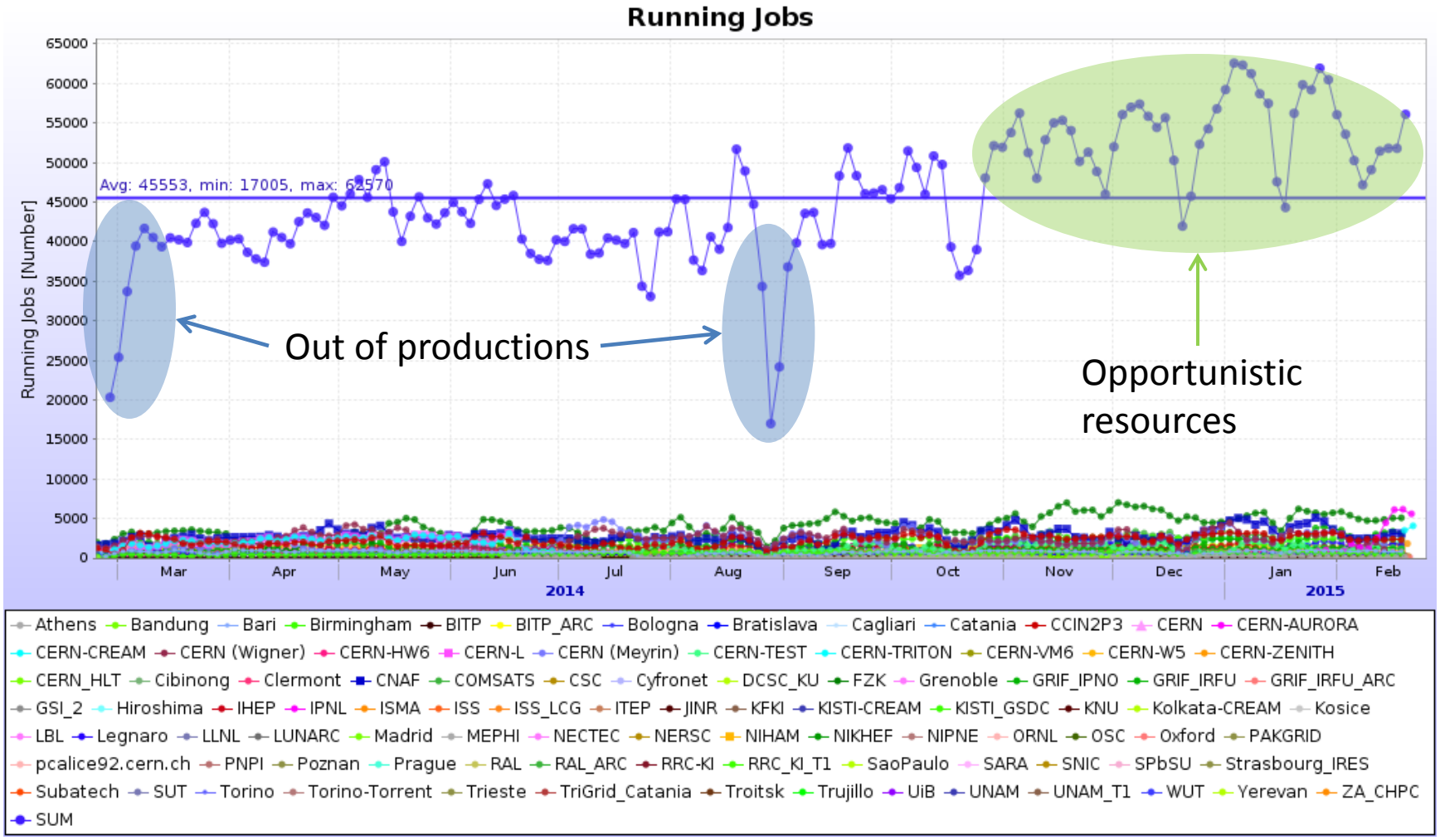


Year on year increase	+97%	+30%	+12%	+30%

Resources evolution

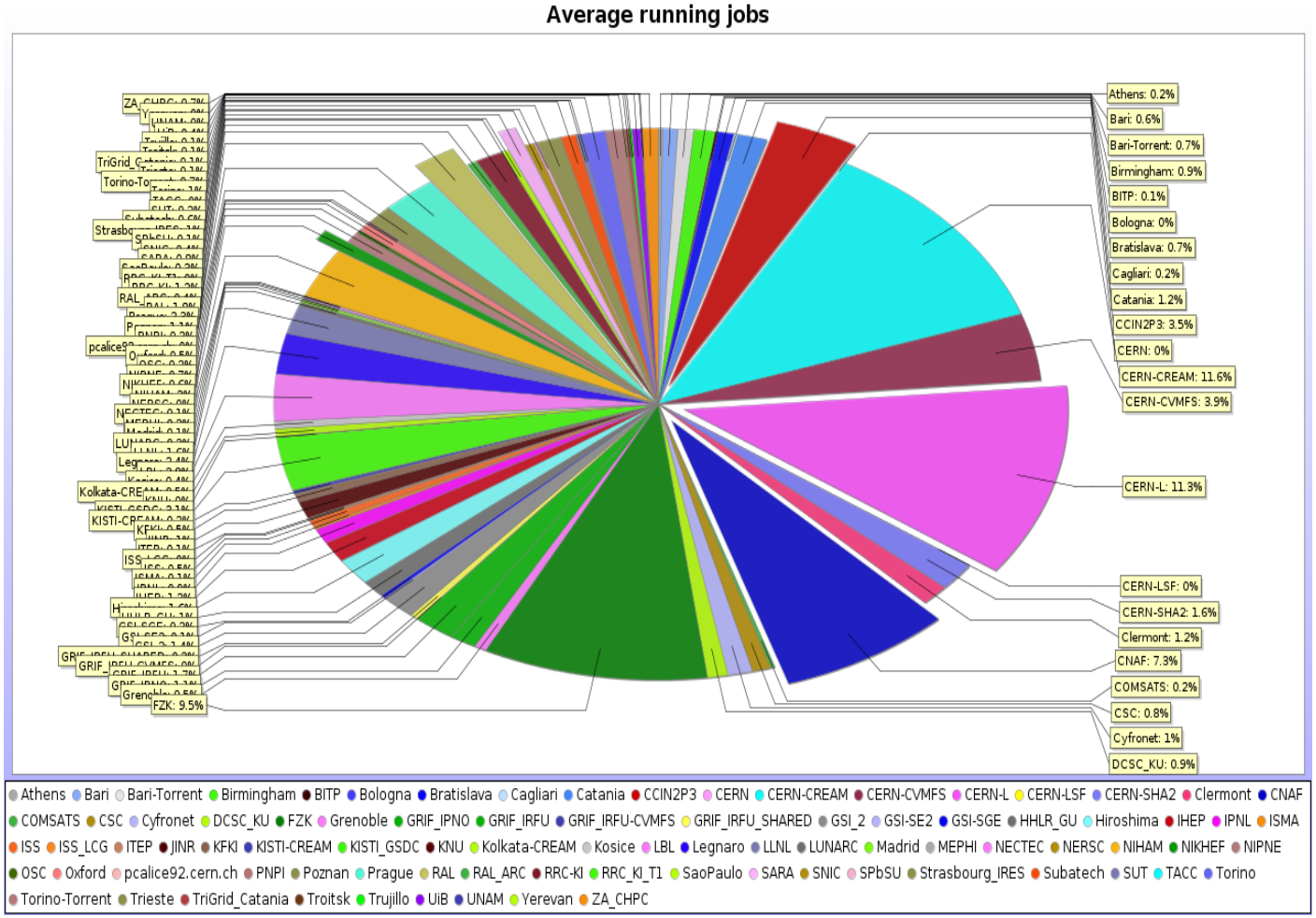
- From 2011 to 2014 – 88% CPU increase
  - @22% average per year – slightly above the WLCG projection
  - Due to new sites (!) and above-flat budget capacity increase
- Storage capacity is growing at ~15% per year
  - Also slightly above flat-budget scenario
  - Remains critical in terms of what we can store – timely cleanup and reviews must continue

Yearly job profile

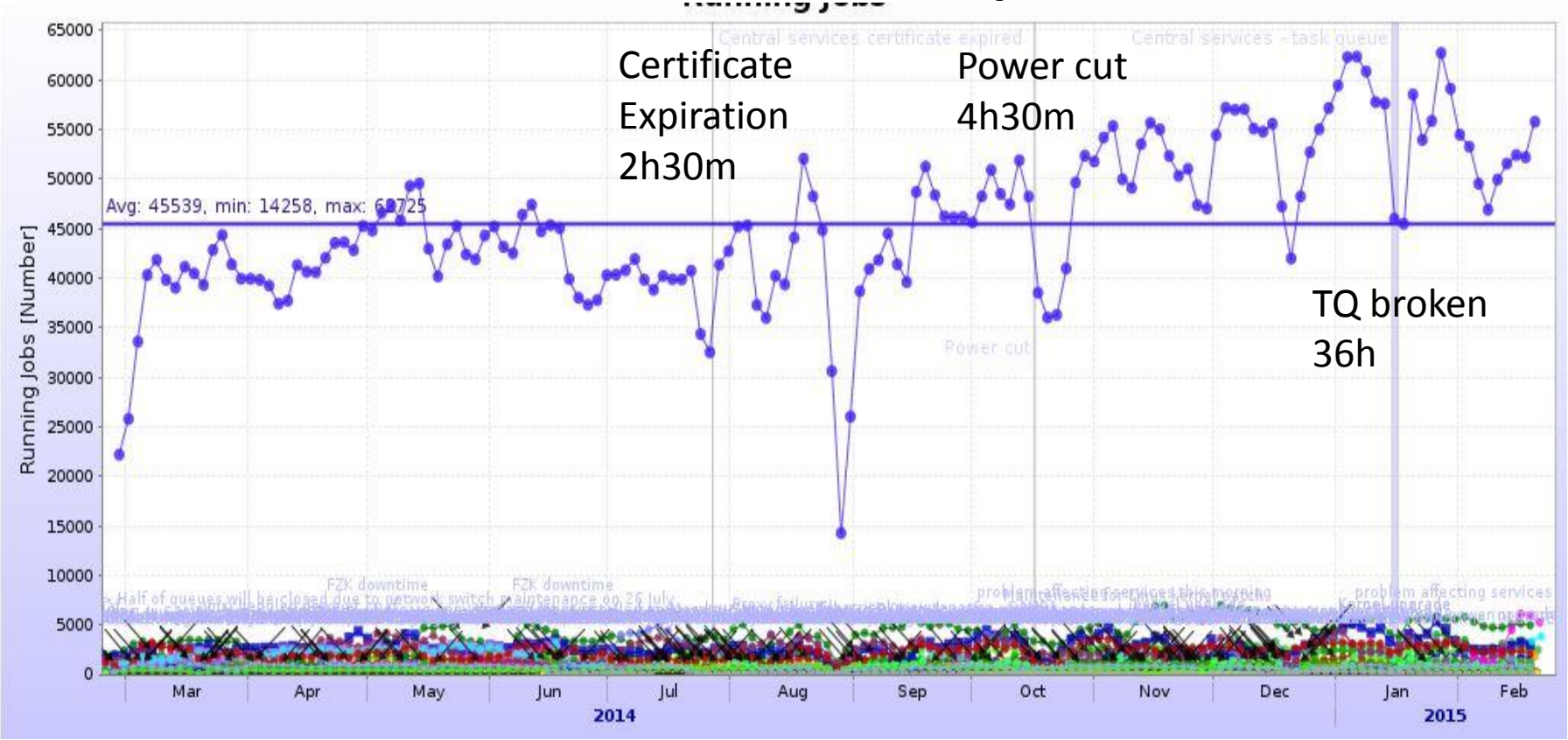


Resources distribution

Continuous and remarkable 50/50 share between large (T0/T1) and smaller computing centres

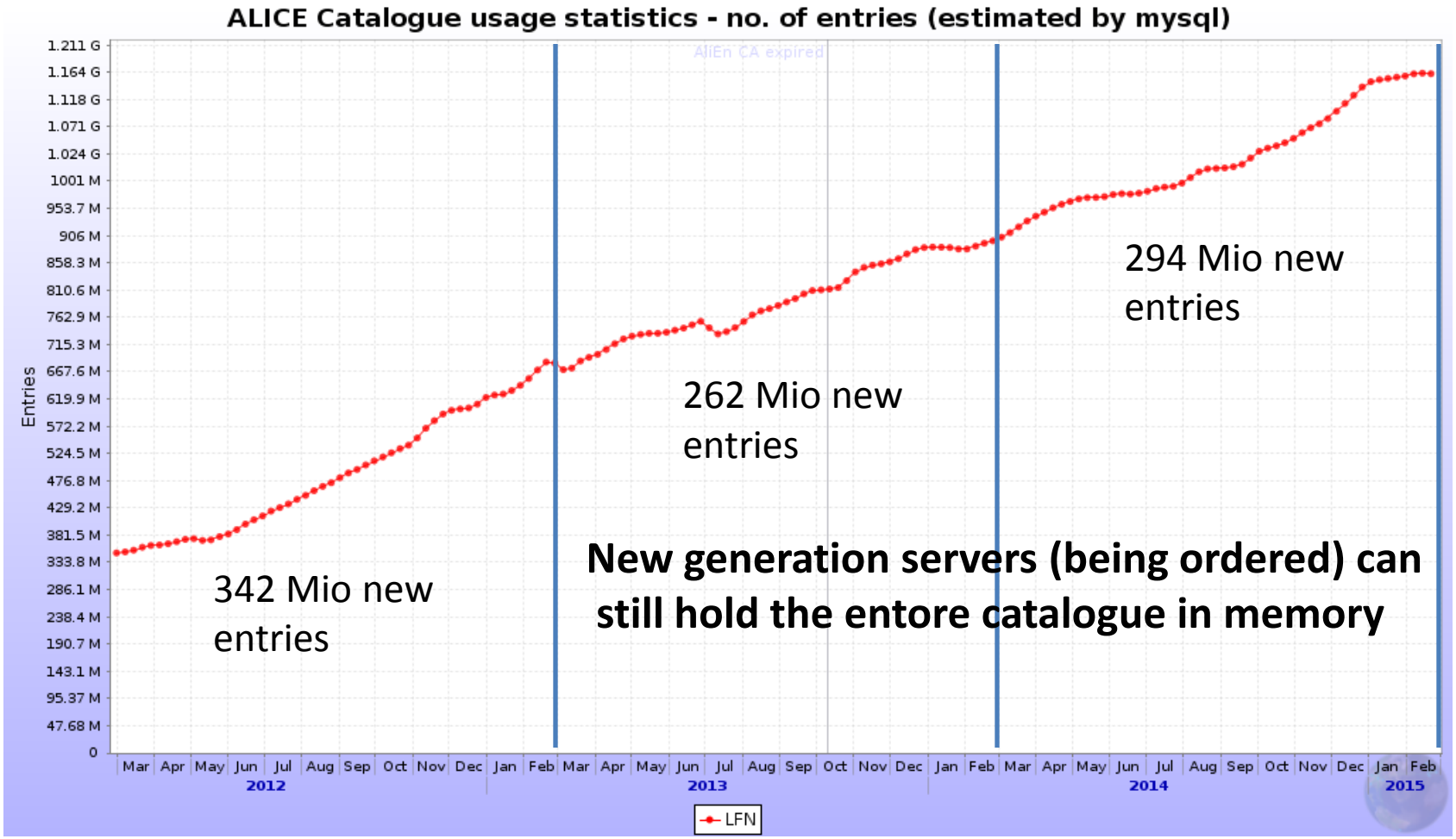


Central services operation

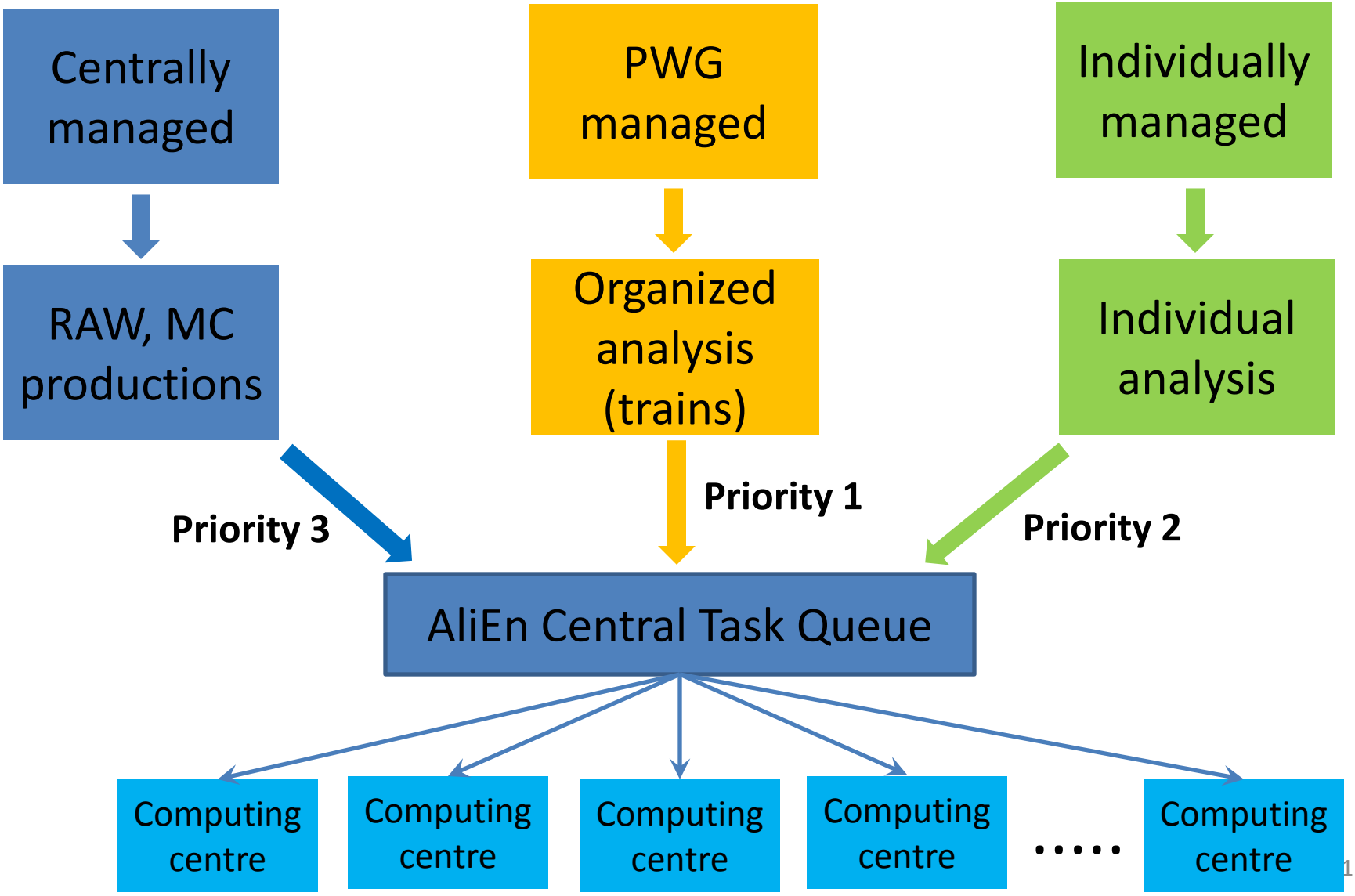


- Total downtime 43 hours => **99.5% availability**
- The blue grass above sites profile – site updates announcements, see individual sites for details

Catalogue stats



Computing tasks and workflow

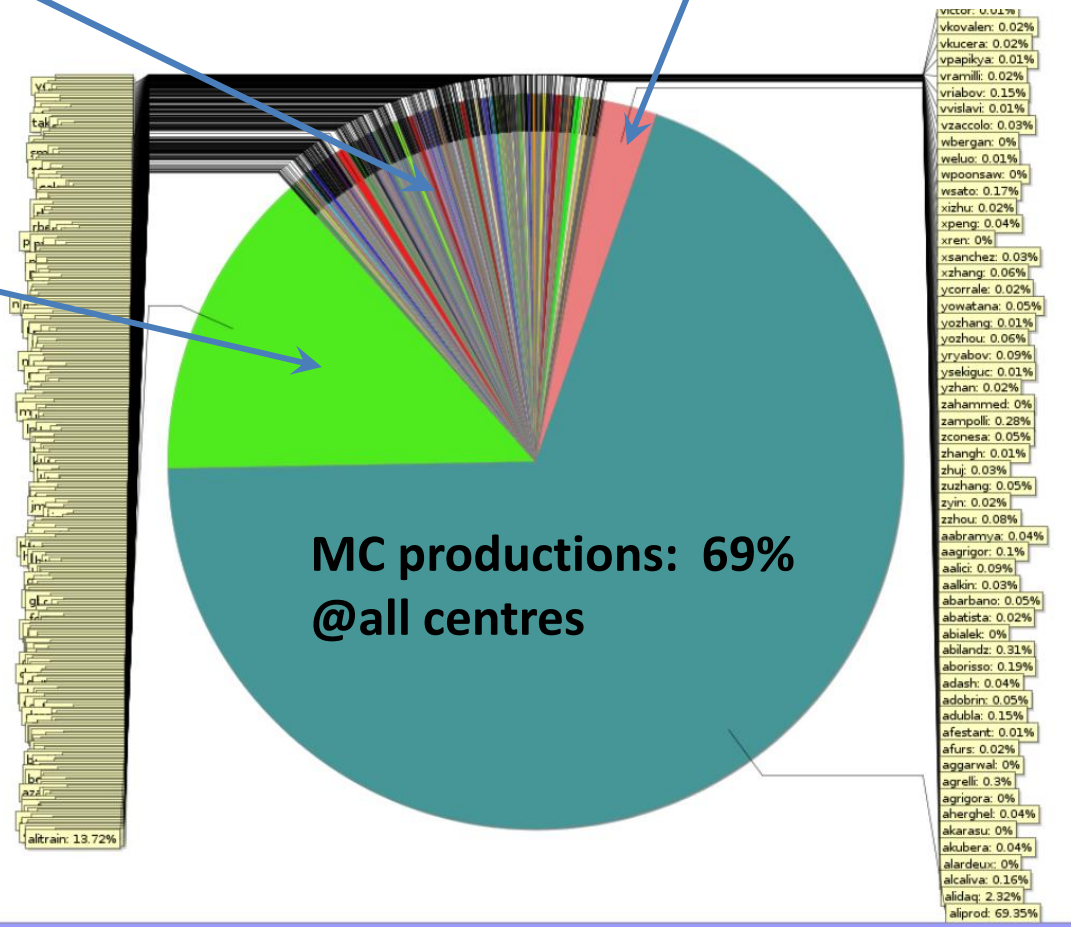


Wall time resources share 2014

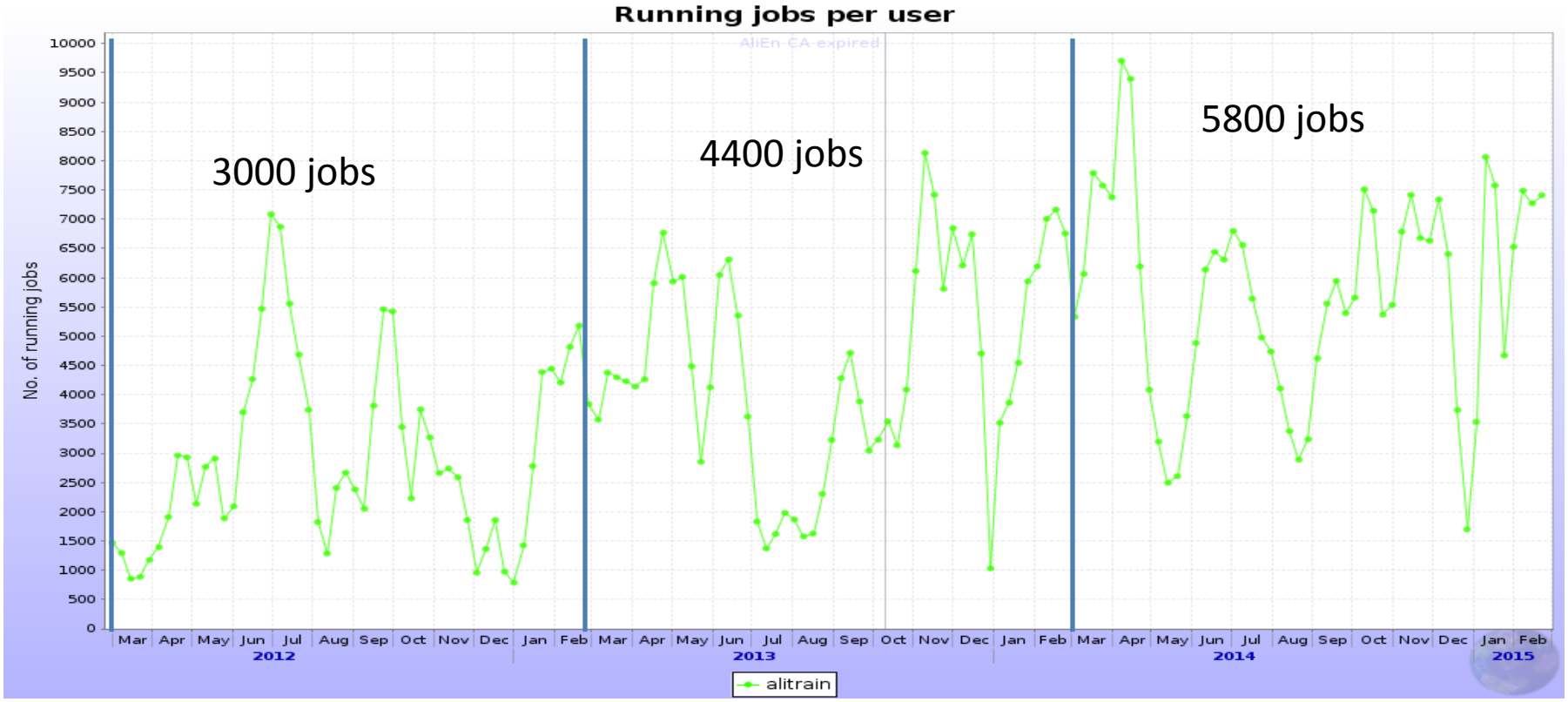
Individual analysis: 12%
@all centres
432 users

RAW data processing: 3%
@ T0/T1s only

Organized analysis: 16%
@all centres

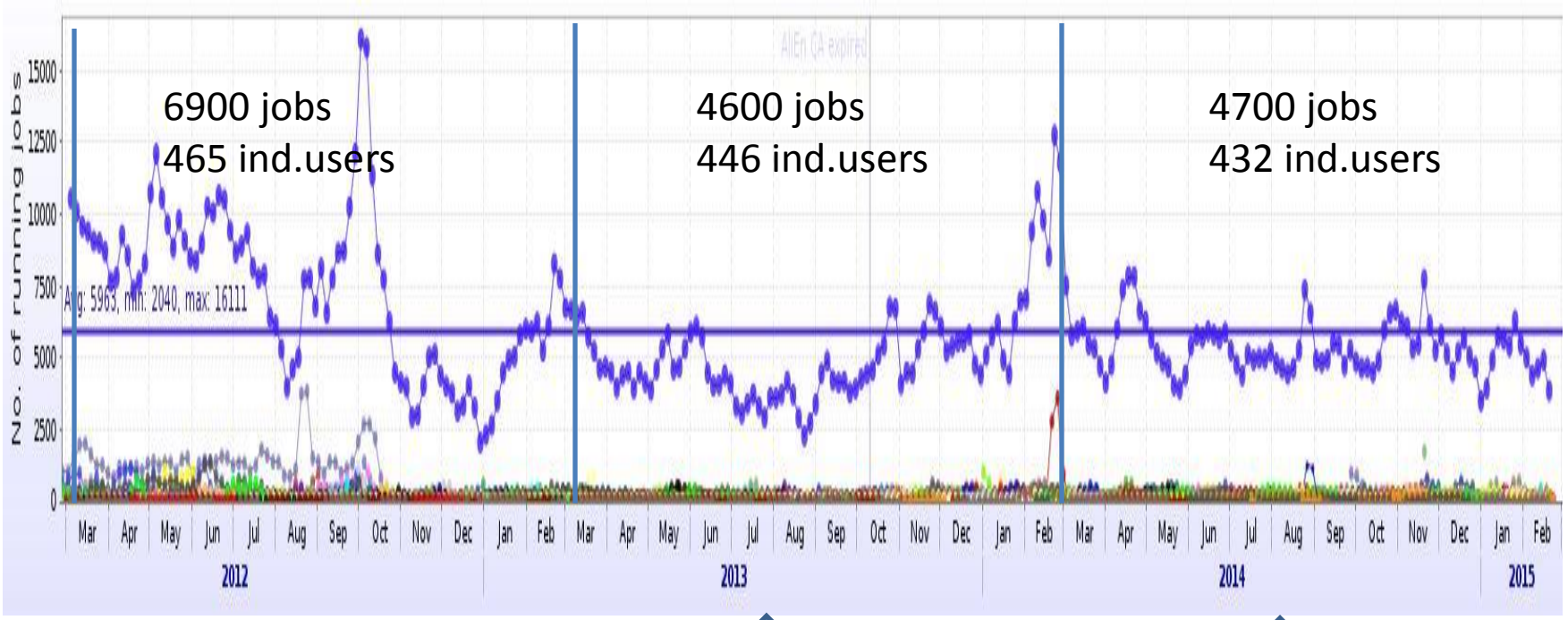


Organized analysis



Year on year increase +47% +32%

Individual analysis



Year on year increase Individual analysis

-50%

+3%

Year on year increase organized analysis

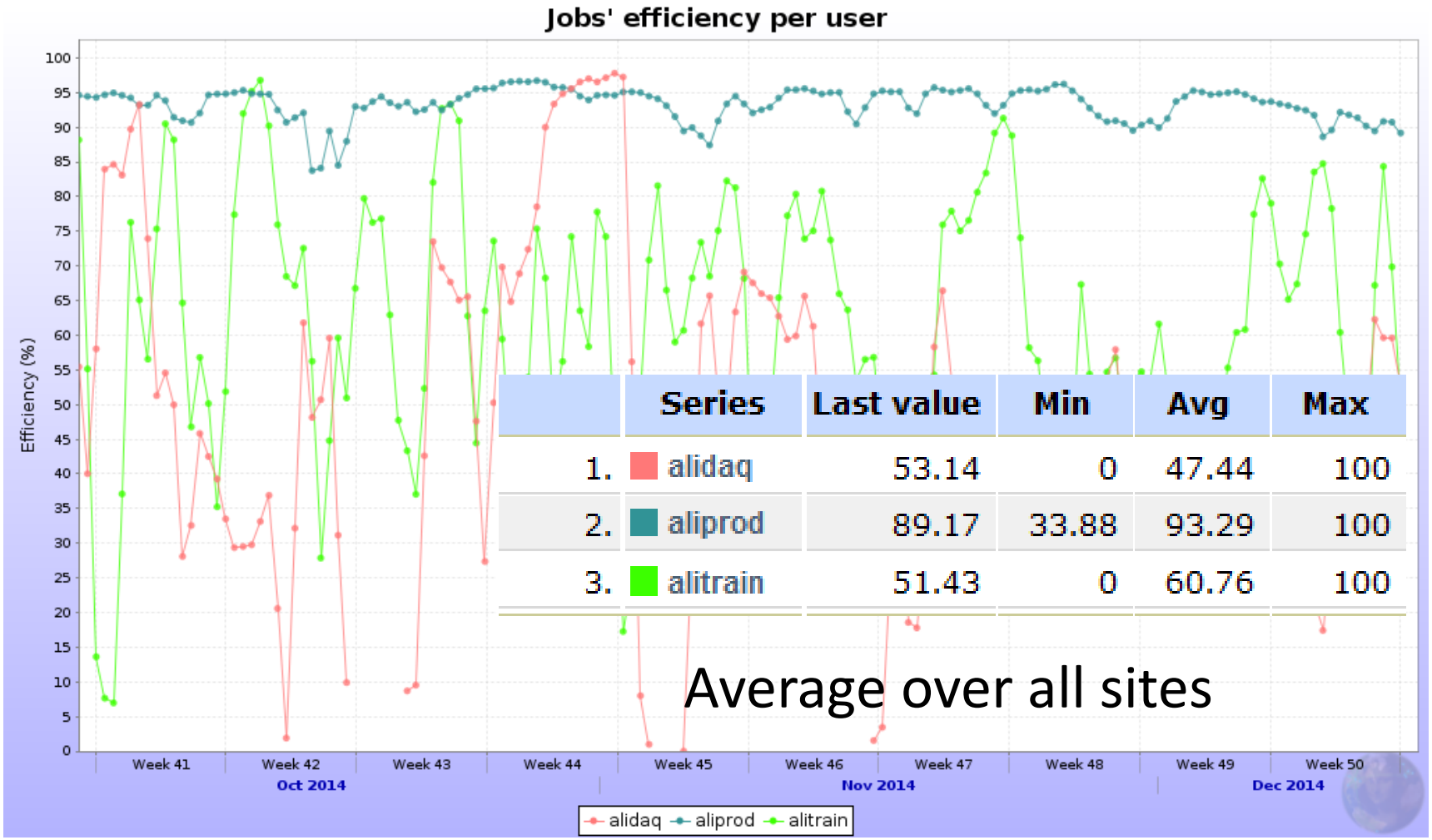
+47%

+32%

Analysis evolution

- From 2012 to 2014 the individual user analysis has decreased by 50%
  - It has remained at the same level of resources utilization between 2013 and 2014
- The organized analysis fully compensated the 'loss' of individual already in 2013
- Since 2013, the amount of resources used by analysis has grown by **35%**, all of it organized
- The number of individual users has remained steady at ~445
- There is still ample room to increase the share of the organized analysis

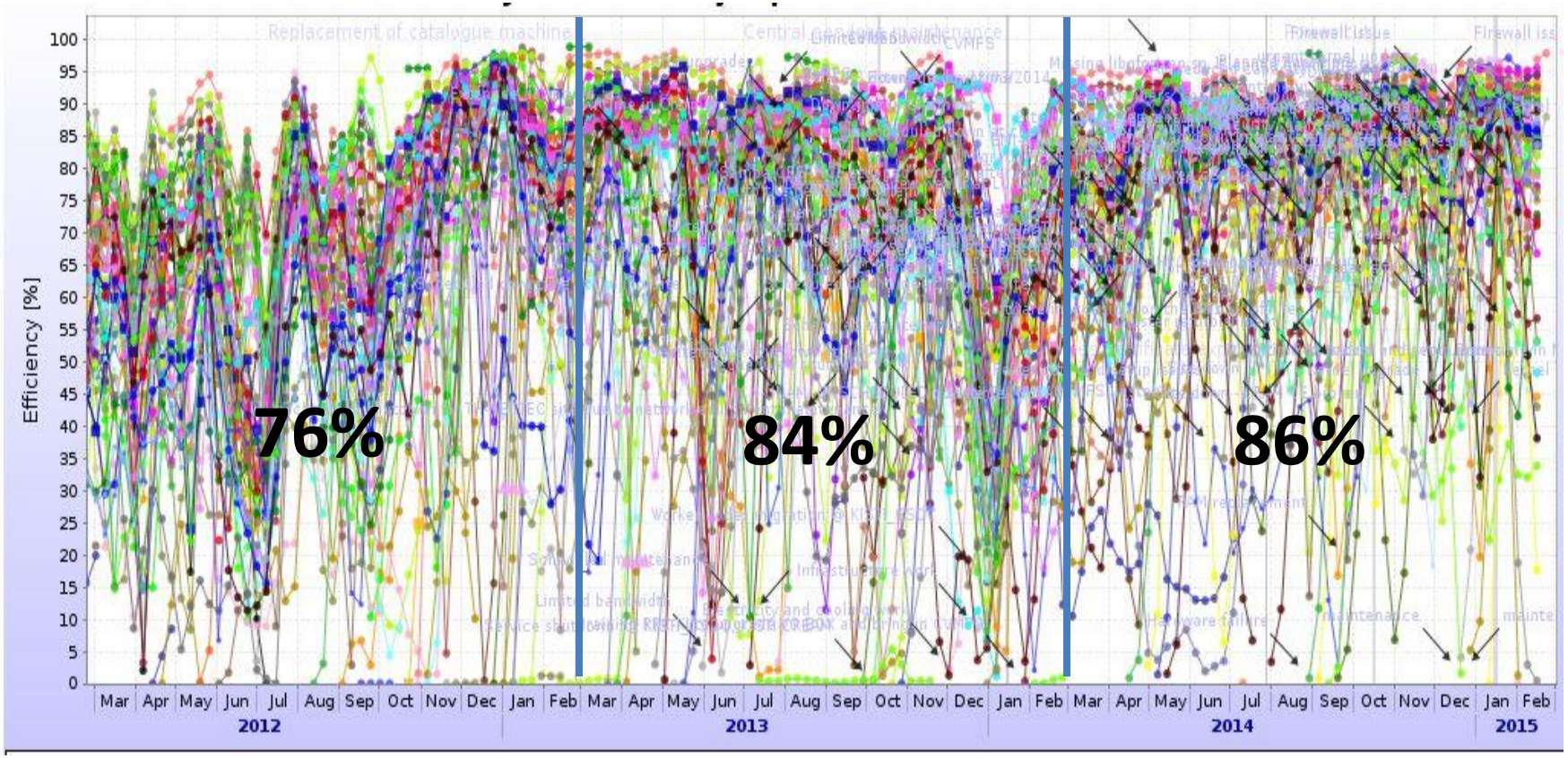
Efficiency per workflow



	Series	Last value	Min	Avg	Max
1.	alidaq	53.14	0	47.44	100
2.	aliprod	89.17	33.88	93.29	100
3.	alitrain	51.43	0	60.76	100

Average over all sites

Grid efficiency



Year on year change

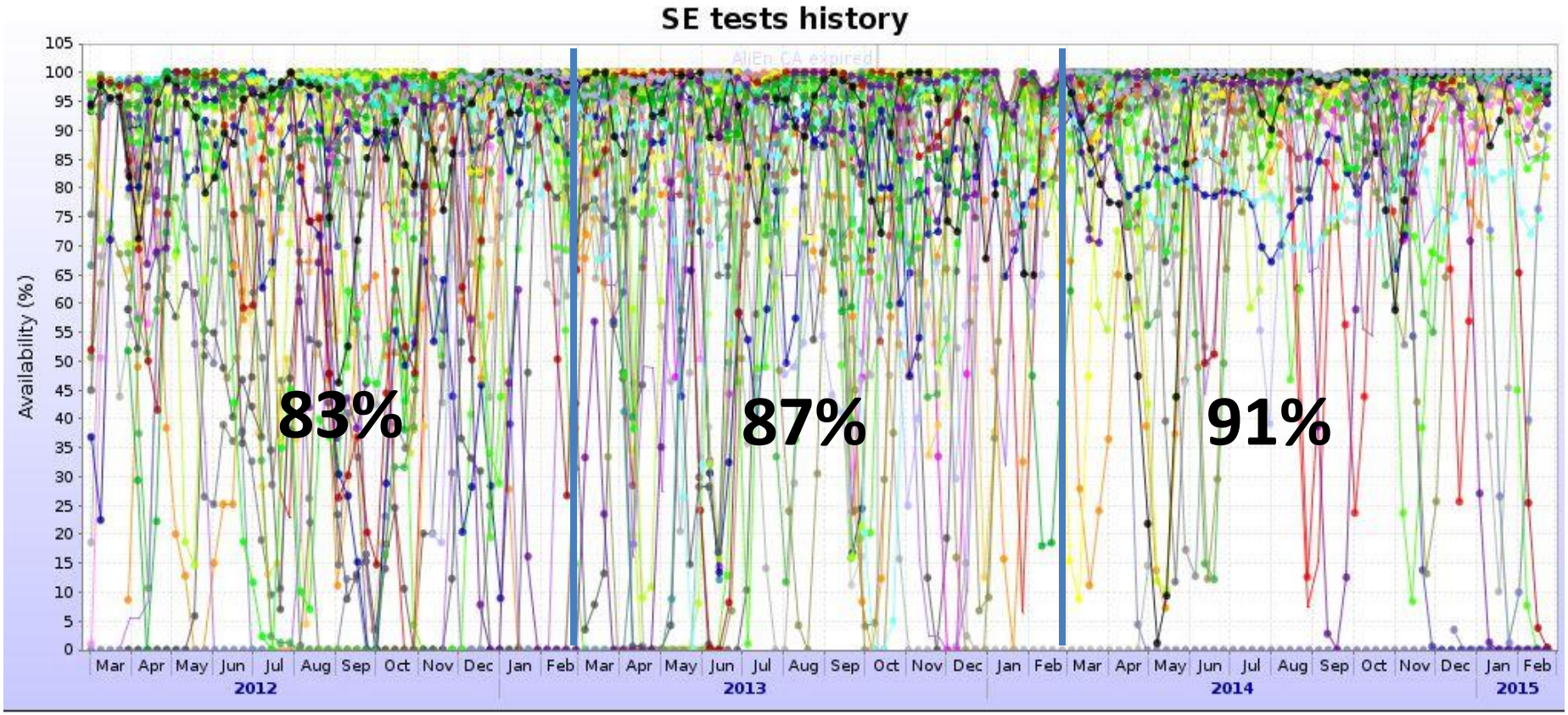
+8%

+2%

Grid efficiency evolution

- Since the re-introduction of TTree Cache, the efficiency has stabilized at ~85%
  - The dramatic decrease of individual analysis also helped the efficiency increase
- In the past year, there is a slight upward trend, could be attributed to the better availability of storage (see next)
- We could expect a slight (2-5%) increase
  - If the individual analysis is decreased by factor 2
  - If the current efficiency level of the other activities remains the same

Storage availability



Year on year change

+4%

+4%

Storage availability evolution

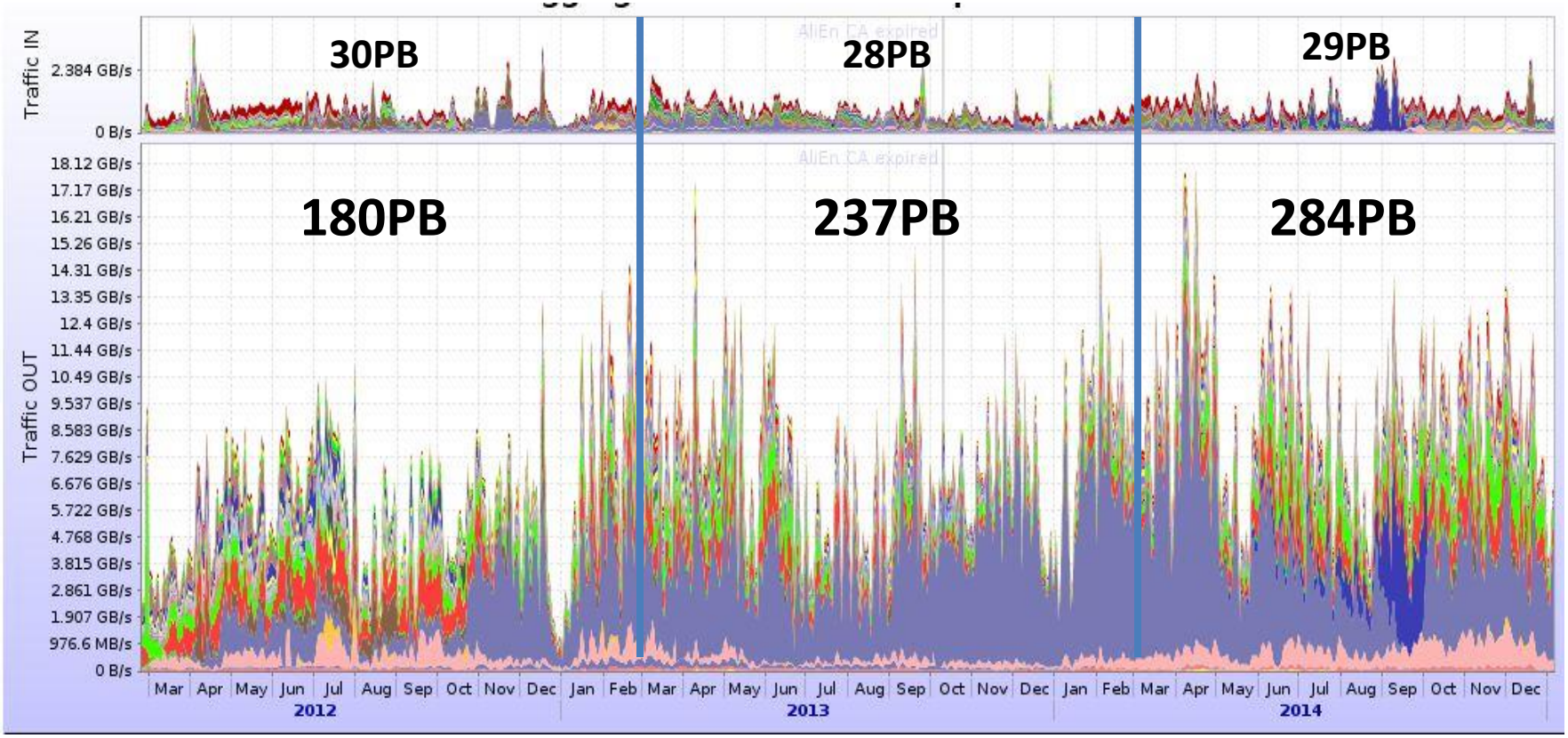
- Constant improvement in availability
  - SEs are independent, no correlation in downtime
- Directly affecting the workload efficiency
- Room for further increase!
  - Allowed downtime for availability >99% = 88 hours

Top 15 SEs, one year average

	Series	Last value	Min	Avg
52.	Strasbourg_IRES - SE	100	0	99.72
17.	Grenoble - SE	100	0	99.62
33.	Legnaro - SE	100	0	99.57
48.	SNIC - DCACHE	98.39	0	99.46
23.	IPNL - SE	98.36	0	99.43
12.	CNAF - SE	97.61	0	99.42
37.	NDGF - DCACHE	96.25	0	99.36
21.	Hiroshima - SE	98.36	0	99.33
9.	CERN - EOS	98.34	0	99.24
27.	JINR - SE	100	0	99.19
10.	CERN - OCDB	98.34	0	99.17
36.	MEPHI - EOS	100	0	99.15
13.	CNAF - TAPE	100	0	99.12
56.	Torino - SE	93.76	0	99.1
34.	LLNL - SE	91.97	0	99.06

- Current replica model (2 copies) => probability for both replicas to be inaccessible @91% SE availability = 0.8%
- @95% availability = 0.25%

Storage use



Year on year change	Write -7%	+4%
	Read +32%	+20%
	Ratio r/w 8.5	9.8

Storage use evolution

- Increase in read volume – directly correlated with the increase in analysis activity
  - Improved ratio read/write
- In 1 year ALICE overwrites the entire disk storage completely
  - Timely cleanup is critical to keep the SEs in good health
  - … and to have free space for the new data
  - The disk cleanup is a continuous activity
  - Minimal amount of ‘dark data’ and files with low popularity

Resources usage 2014

2013	CPU (KHS06)	Disk (PB)	Tape(PB)
Tier 0	90	8.3	12.0
Tiers 1	110	10.1	6.0
Tiers 2	190	12.8	-

Requirements

2013	CPU (KHS06)	Disk (PB)	Tape(PB)
Tier 0	90	8.1	14.0
Tiers 1	117	11.0	13.7
Tiers 2	186	14.1	-

Pledges

	CPU (KHS06)	Disk (PB)	Tape(PB)
Tier 0	43	6,6	10,8
Tiers 1	119	7,6	5,5
Tiers 2	189	7,1	-
All Tiers	372 (351)	23 (21,3)	16,3

Usage

Summary

- 2014 was (another) successful year for Grid operations
- Despite the absence of data taking, the Grid resources use was uninterrupted
  - In fact it has increased, as was the available capacity
- New centers have entered production – the Grid is expanding above the ‘flat budget’ scenario
- Substantial increase of analysis, most of it organized
- Efficiency remains high, and can be increased further
- The computing centres operation continues to be smooth
  - Software and hardware updates have negligible effect on general Grid availability