



Grid operations in 2015

T1/T2 workshop – 6th edition - Bergen

18 April 2016
Lachezar Betev

T1/T2 workshops

CERN 2009

6th workshop, Bergen 2016



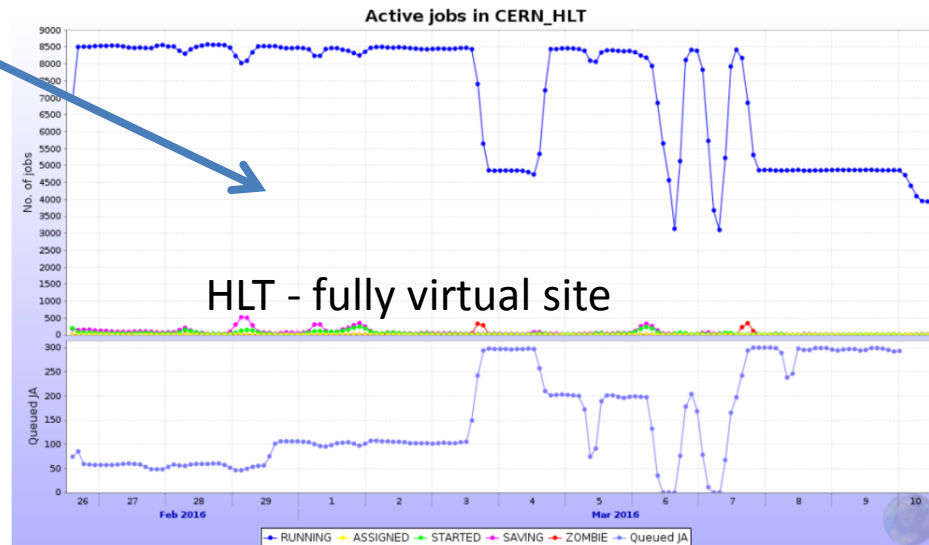
where the weather is always nice!

The ALICE Grid sites today

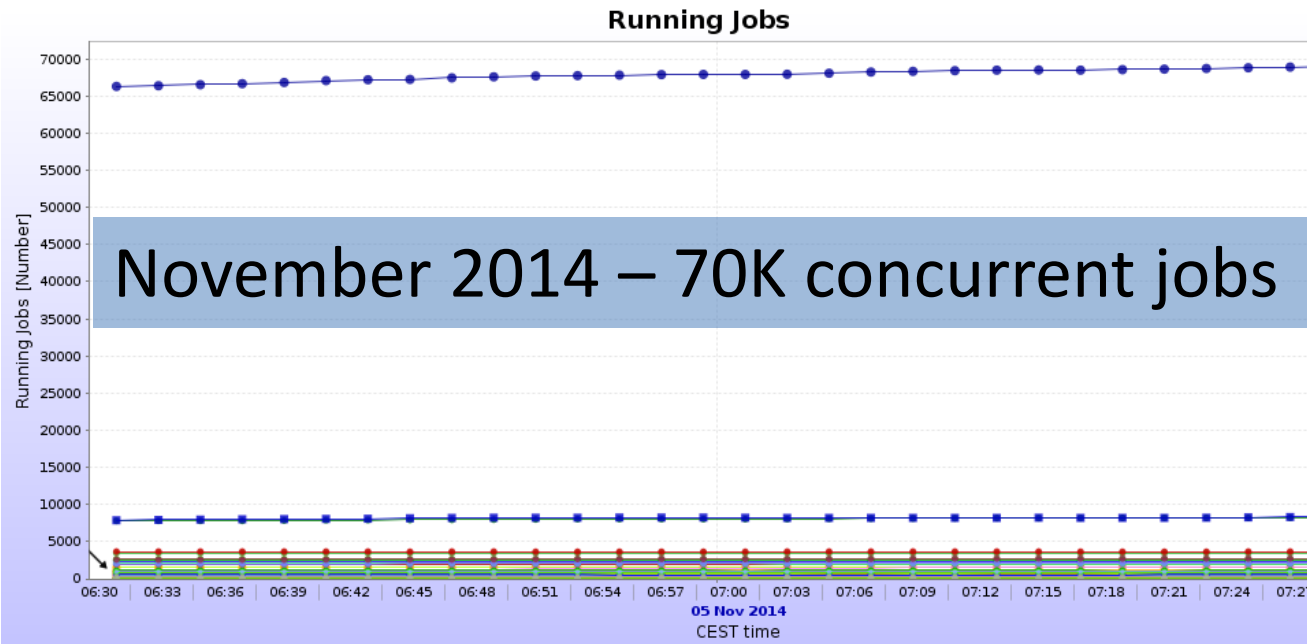


New sites

- ORNL – since August 2015 (replaces LLNL)
- Bandung and Cibinong – since August 2015
- CBPF – since December 2015
- Vienna – since April 2016
- Wuhan (back) – since September 2015
- ALICE HLT cluster
- Altaria (see Maarten's talk)



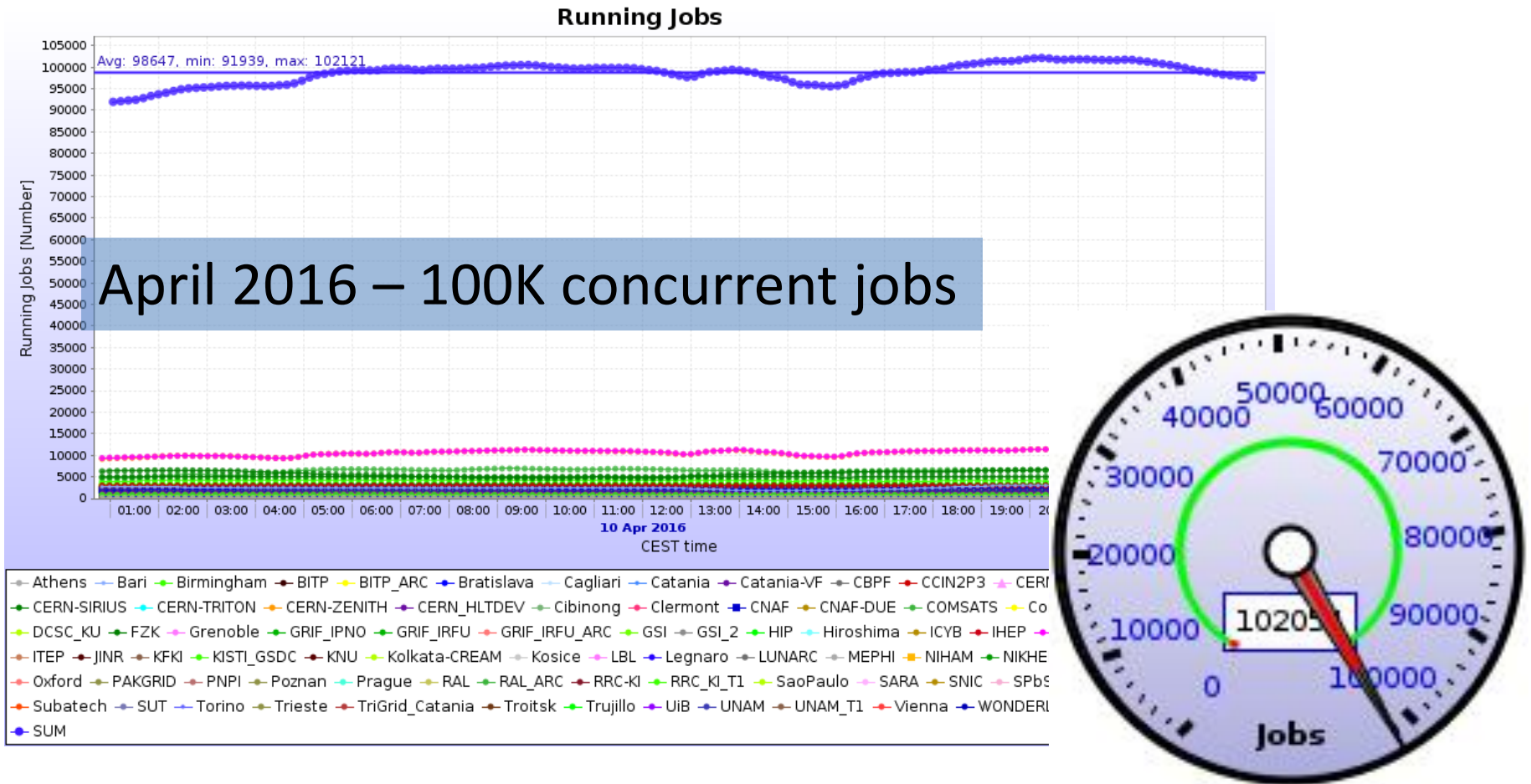
Torino workshop record



- SUM • Athens • Bandung • Bari • Birmingham • BITP • BITP_ARC • Bologna • Bratislava • Catania • CCIN2P3 • CERN • CERN (Wigner)
- CERN (Meyrin) • CERN-TEST • CERN_HLT • Cibirong • Clermont • CNAF • CSC • Cyfronet • DCSC_KU • FZK • Grenoble • GRIF_IPNO • GR
- GSI_2 • Hiroshima • IHEP • ISMA • ISS • ISS_LCG • JINR • KFKI • KISTI_GSDC • Kolkata-CREAM • Kosice • LBL • Legnano • LLNL • LUNAF
- Madrid • MEPHI • NECTEC • NIHAM • NIKHEF • NIPNE • Oxford • PAKGRID • PNPI • Poznan • Prague • RAL • RAL_ARC • RRC-KI • RRC_
- SaoPaulo • SARA • SNIC • SPbSU • Strasbourg_IRES • Subatech • SUT • Torino • Trieste • TriGrid_Catania • Troitsk • Trujillo • UiB • UNAM
- UNAM_T1 • WUT • Yerevan • ZA_CHPC

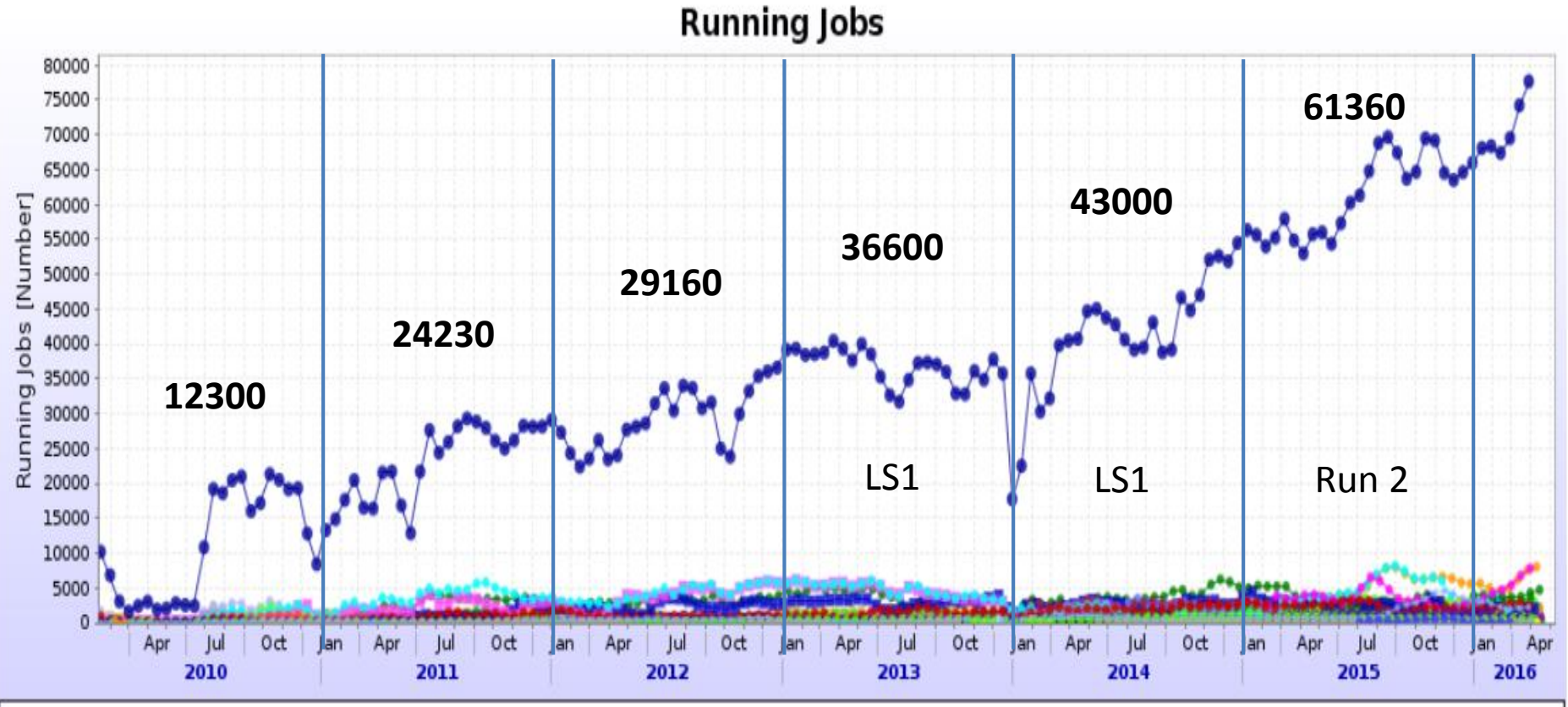


Current record



- “Order of magnitude” record
- Next order (1 Mio) projected in 2034... if we keep the same pace

CPU resources evolution



Year on year increase

↑
+97%

↑
+20%

↑
+26%

↑
+17%

↑
+43%

Resources evolution – regional comparison

Region	2014	2015	Difference [%]
Africa	380	760	+100
Asia	3500	4580	+31
Europe	36340	52070	+43
North America	2100	2520	+20
South America	680	1433	+110

Resources evolution

- From 2012 to 2015 – CPU increase
 - 2015 - 2.5x the power in 2011
 - @26% average per year – above the WLCG projection
 - Above-flat budget capacity increase

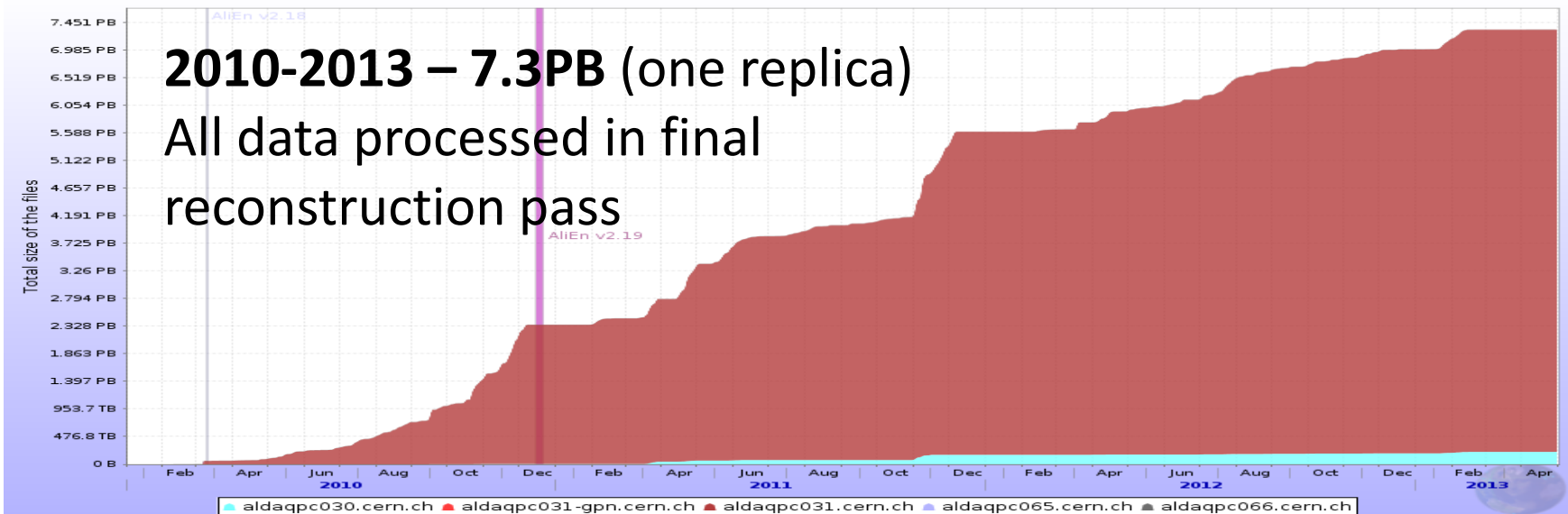
- **This is all excellent news!**



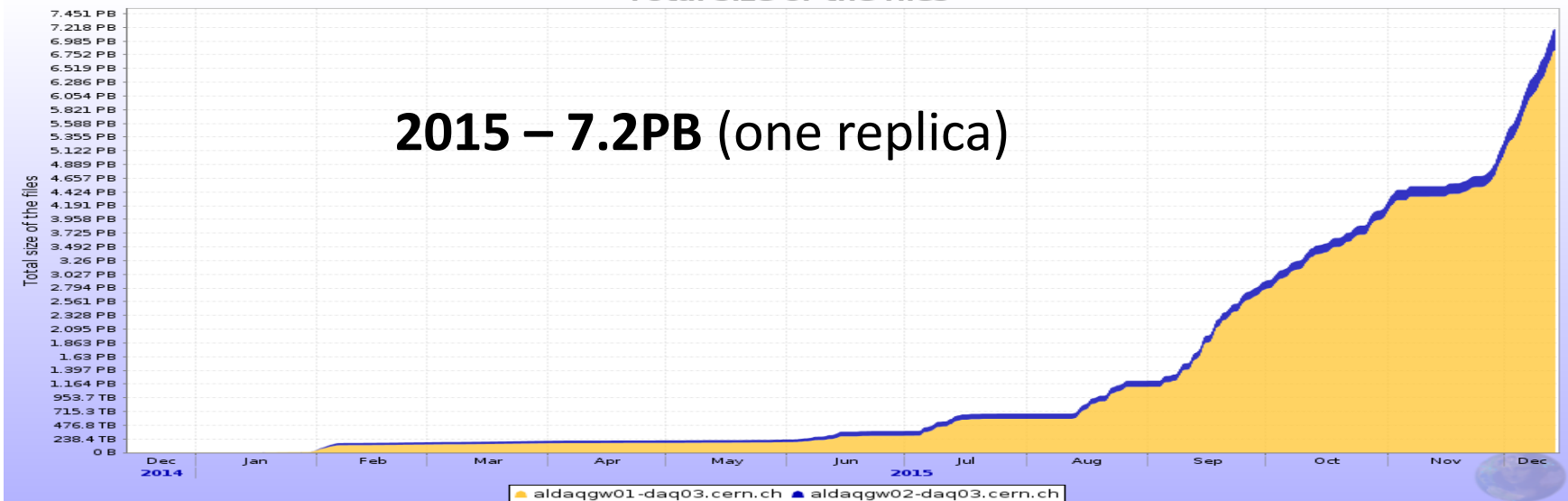
- More potential (opportunistic) CPU resources to come from supercomputers
 - See Pavlo's talk

2015 RAW data collection

Total size of the files



Total size of the files



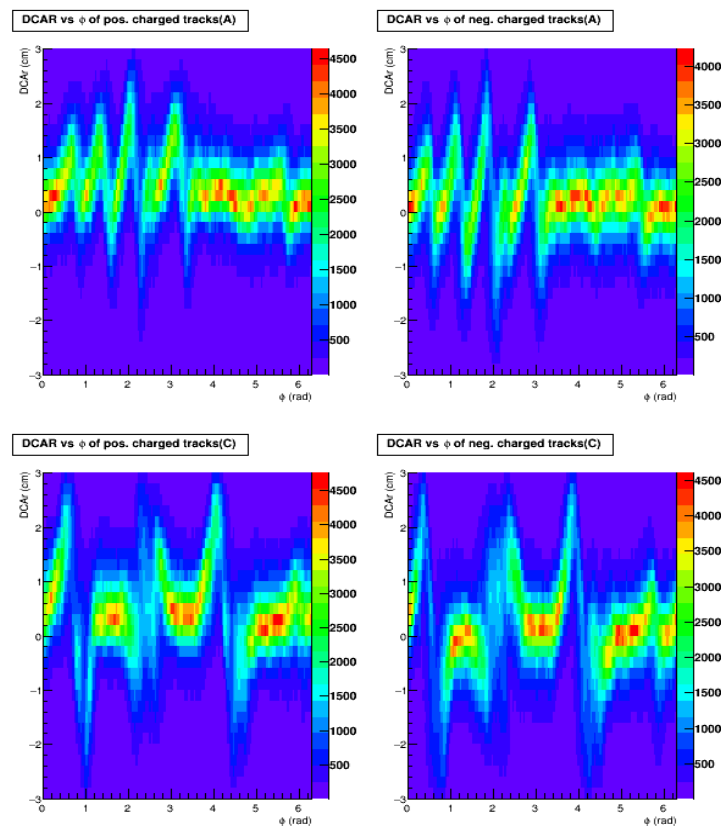
Status of 2015 data processing

- Muon+Calo cycles (reduced detector set)
 - Good for muon analysis and calorimeter calibration
 - All periods completed (including Pass2)

Production	Description	Status	Run Range	Runs	Chunks	Size	Chunks	Size	Events		
LHC15m_muon_calor_pass2	LHC period LHC15m - Muon+Calorimeters reconstruction pass 2	Completed	243374 - 243984	23	5,175	1.329 TB	5,154	99%	150.8 GB	11%	11,934,877
LHC15k_muon_calor_pass2	LHC period LHC15k - Muon+Calorimeters reconstruction pass 2	Completed	238682 - 239144	48	77,916	121.6 TB	77,163	99%	1.35 TB	1%	85,223,329
LHC15n_muon_calor_pass2	LHC period LHC15n - Muon+Calorimeters reconstruction pass 2	Completed	244340 - 244628	27	75,647	121.2 TB	75,541	99%	3.466 TB	2%	186,242,616
LHC15g_muon_calor_pass2	LHC period LHC15g - Muon+Calorimeters reconstruction pass 2	Completed	227750 - 231568	159	151,666	222.7 TB	147,054	96%	2.254 TB	1%	131,882,789
LHC15h_muon_calor_pass2	LHC period LHC15h - Muon+Calorimeters reconstruction pass 2	Completed	232465 - 234057	172	404,756	523.8 TB	402,662	99%	11.34 TB	2%	956,547,719
LHC15l_muon_calor_pass2	LHC period LHC15l - Muon+Calorimeters reconstruction pass 2	Completed	239319 - 241544	170	497,522	807.1 TB	486,593	97%	8.944 TB	1%	427,740,933
LHC15j_muon_calor_pass2	LHC period LHC15j - Muon+Calorimeters reconstruction pass 2	Completed	236892 - 238622	224	525,353	764.9 TB	518,265	98%	12.96 TB	1%	798,344,183
LHC15i_muon_calor_pass2	LHC period LHC15i - Muon+Calorimeters reconstruction pass 2	Completed	235196 - 236866	201	857,570	1.324 PB	851,534	99%	9.117 TB	0%	587,758,874
LHC15o_muon_calor_pass1	LHC period LHC15o - Muon+Calorimeters reconstruction pass 1	Completed	244824 - 246994	190	2,011,443	2.438 PB	2,009,745	99%	38.16 TB	1%	581,595,428
LHC15n_muon_calor_pass1	LHC period LHC15n - Muon+Calorimeters reconstruction pass 1	Completed	244340 - 244628	27	75,647	121.2 TB	75,436	99%	4.38 TB	3%	185,867,613
LHC15m_muon_calor_pass1	LHC period LHC15m - Muon+Calorimeters reconstruction pass 1	Completed	243374 - 244284	24	5,445	1.397 TB	5,174	95%	208.7 GB	15%	11,984,275
LHC15l_muon_calor_pass1	LHC period LHC15l - Muon+Calorimeters reconstruction pass 1	Completed	239319 - 241544	170	497,522	807.1 TB	460,504	92%	17.91 TB	2%	405,211,740
LHC15k_muon_calor_pass1	LHC period LHC15k - Muon+Calorimeters reconstruction pass 1	Completed	238682 - 239144	48	77,916	121.6 TB	77,906	99%	1.7 TB	1%	86,010,725
LHC15j_muon_calor_pass1	LHC period LHC15j - Muon+Calorimeters reconstruction pass 1	Completed	236892 - 238622	224	525,353	764.9 TB	524,572	99%	16.83 TB	2%	811,327,388
LHC15i_muon_calor_pass1	LHC period LHC15i - Muon+Calorimeters reconstruction pass 1	Completed	235196 - 236866	202	857,870	1.324 PB	847,062	98%	8.916 TB	0%	629,905,933
LHC15h_muon_calor_pass1	LHC period LHC15h - Muon+Calorimeters reconstruction pass 1	Completed	232465 - 234057	172	404,756	523.8 TB	394,648	97%	9.467 TB	1%	935,304,148
LHC15g_muon_calor_pass1	LHC period LHC15g - Muon+Calorimeters reconstruction pass 1	Completed	227750 - 231568	159	151,666	222.7 TB	148,503	97%	2.239 TB	1%	153,982,393
					7,203,223	10.09 PB	7,107,516		149.4 TB		6,986,864,963

Status of 2015 data processing (2)

- Substantial IR-induced distortions in the TPC
- Affect both p-p and Pb-Pb data
- Sophisticated correction algorithms development in the past 6 months
- Data reconstructed partially (first physics, Lower IR runs)
- Bulk of reconstruction still pending



MC generation

- 150 individual MC cycles
 - Most of these for physics papers in preparation
 - “First physics” analysis of 2015 data
- Total of 2,790,430,404 events
- As for RAW data processing, the bulk of 2015 MC is still to be run
 - Both p-p and Pb-Pb

... and back to resources - disk

- Overall increase of disk in 2015 – 15%
 - This is a lot less than the 43% CPU increase, but OK
- “golden ratio” is **3/1**, i.e. 3cores/1TB disk,

Region	Cores	Installed disk [TB]	Ratio (lower is better)
Africa	760	100	7.6
Asia	4580	1480	3.1
Europe	52070	22000	2.3 (CERN EOS++)
North America	2520	2000	1.26 (ORNL EOS)
South America	1433	200	7.2

Disk storage

- From a “cloud” view – the ratios are more or less OK
- There is a site by site disparity
 - Compensated in the “cloud”
 - Absence (or inadequate capacity) on individual sites hurts efficiency – **sites should make an effort to keep to the ratio 3/1 as close as possible**
 - Remote sites (with respect to the bulk) SEs are receiving less data
- Network is a factor, despite the continuous readjustment of ‘storage discovery’ algorithms
 - This is addressed in the case of Asian sites by common network initiatives – “Asia-Tier workshop @KISTI”

Disk storage (2)

- The global disk space needed for the 2015 data processing will be about 5-6PB (RAW + MC)
 - This is taking into account the new 'single ESD replica' scenario
 - We will be able to run everything within the currently available disk + the foreseen increase in 2016

Popularity and cleanup

ALICE number of accesses in time X

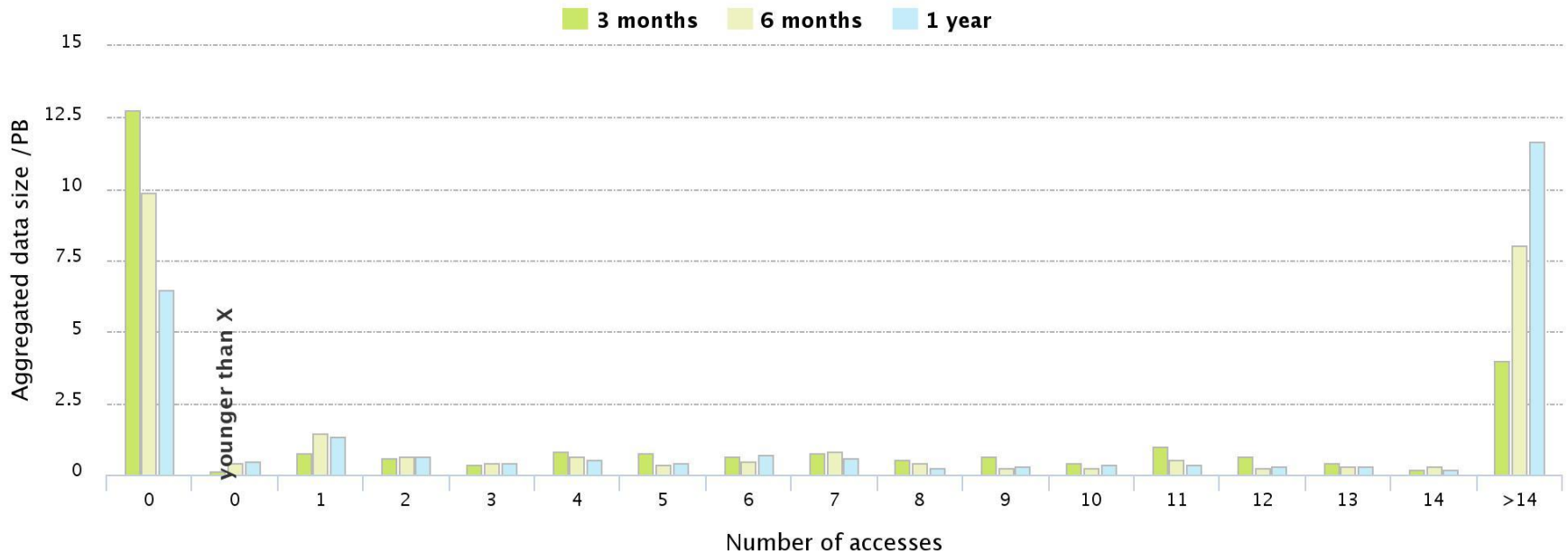
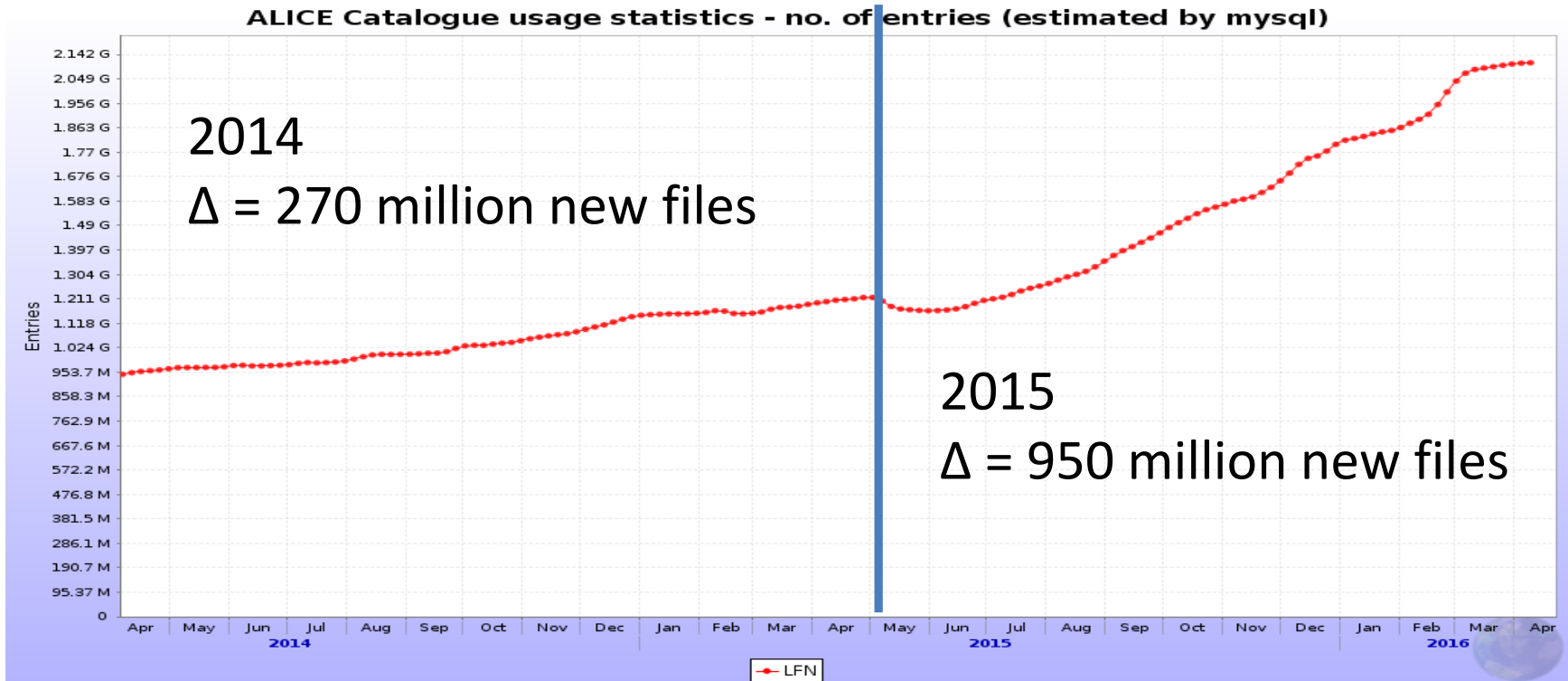


Chart generation startdate: April 2016

- Cleanup consists of removing really old MC productions (done) and second ESD replica for productions with low access
- List of productions to be sanitized is available and data will be removed as necessary

Catalogue stats



- Substantial increase of files in a data taking year
 - Partially due to new procedures and tasks associated with RAW data processing
 - So far no issues with catalogue (see Miguel's talk)

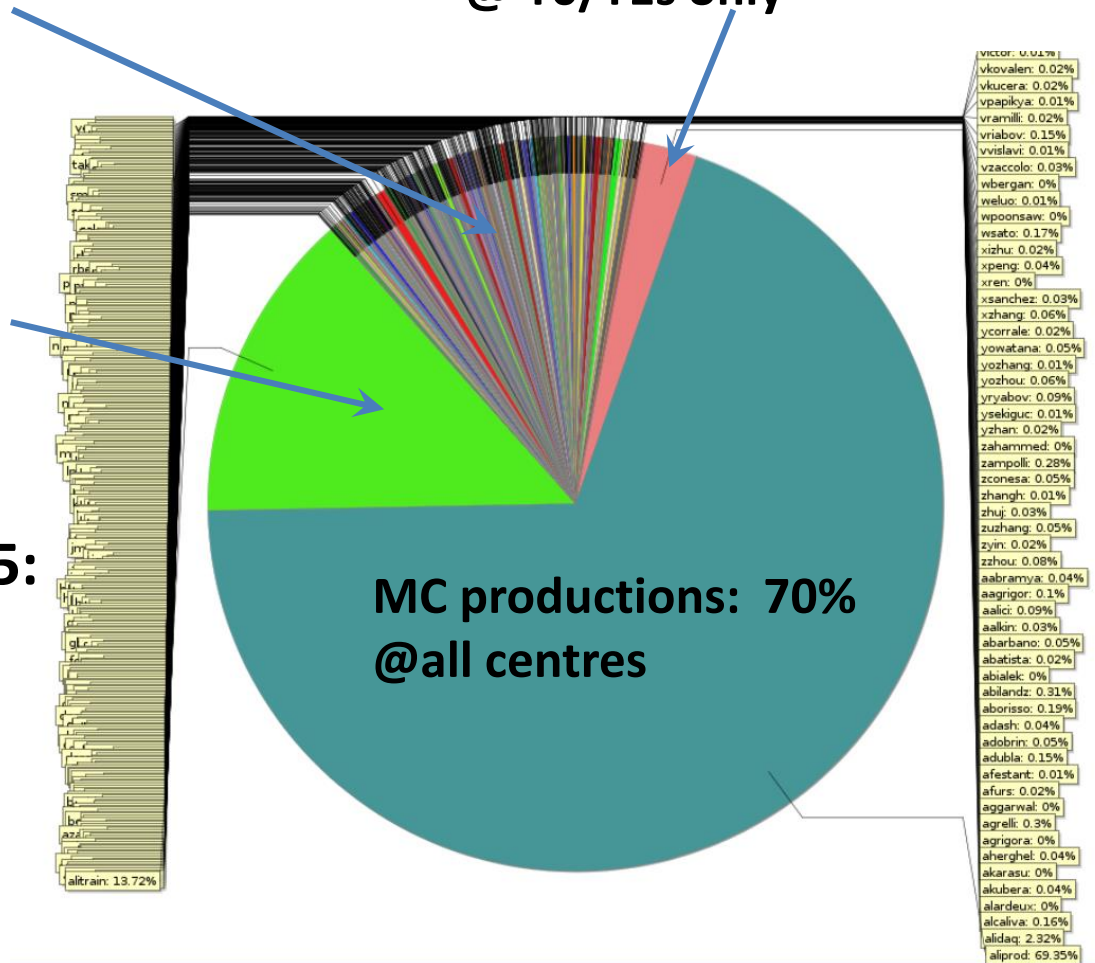
Wall time resources share 2016

Individual analysis: 6%
@all centres
460 users

RAW data processing: 9%
@ T0/T1s only

Organized analysis: 15%
@all centres

Share ~same as in 2015:
RAW 3->9%
Ind.users-> 12->6%



Organized analysis

Running jobs per user



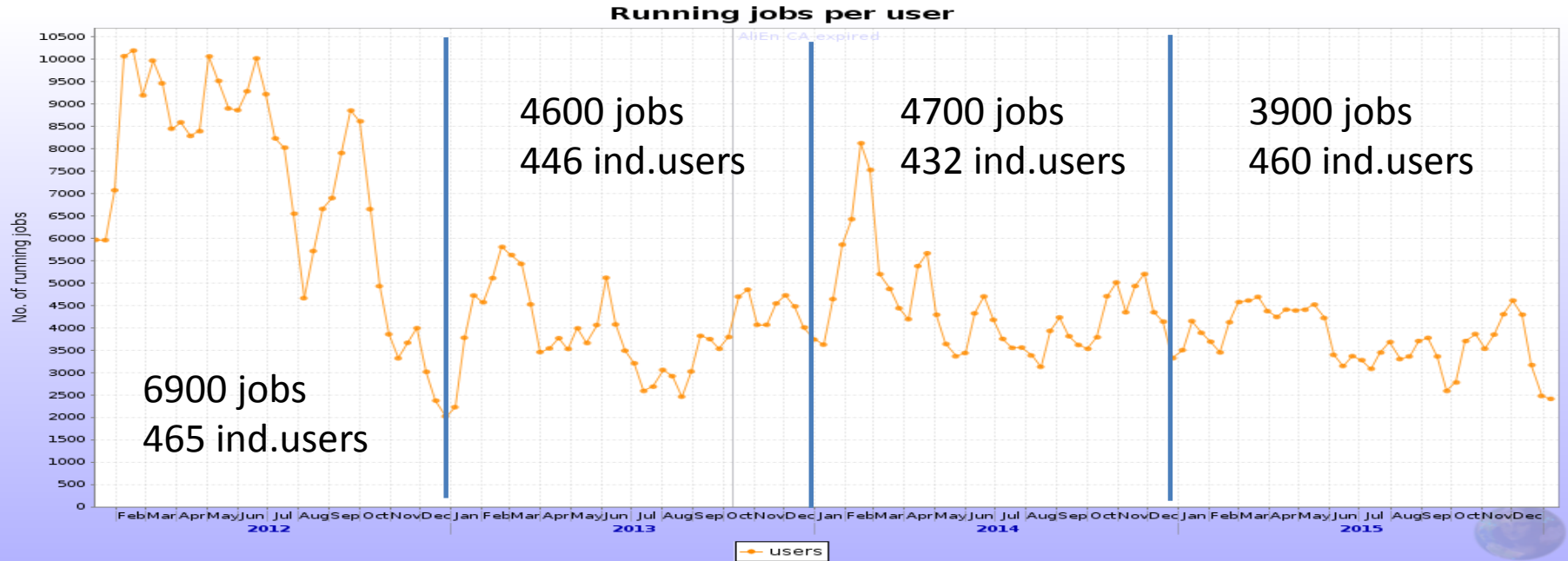
Year on year increase

+47%

+32%

+57%

Individual analysis



**Year on year increase
Individual analysis**



-50%



+3%



-17%

**Year on year increase
organized analysis**



+47%



+32%



+57%

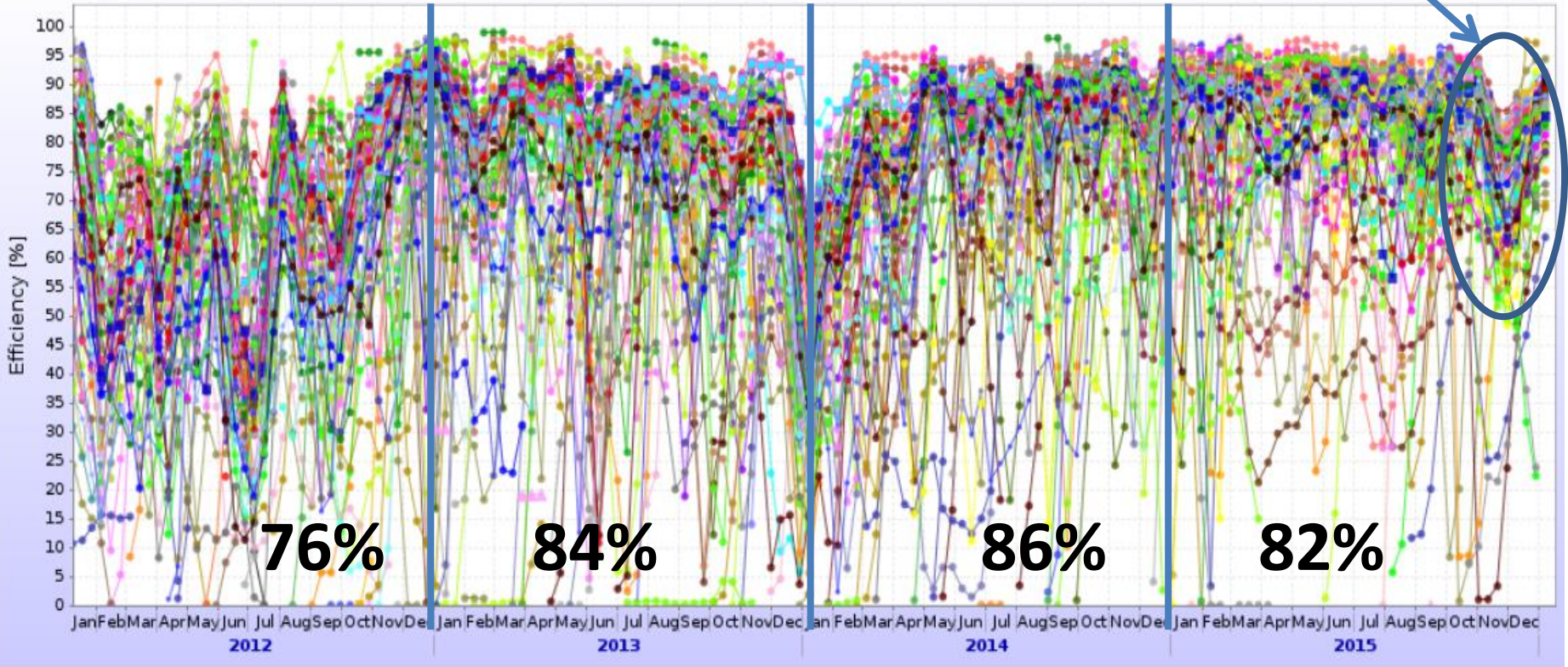
Analysis evolution

- From 2012 to 2015 the individual user analysis has decreased by **57%**
- The organized analysis share of total resources remains the same (**@15%**), increases in absolute terms (**57% in 2015**)
- The number of individual users has remained steady at ~450

Grid efficiency

Jobs efficiency (cpu time / wall time)

RAW calibration
Muon+calo reco



Year on year change



+8%



+2%



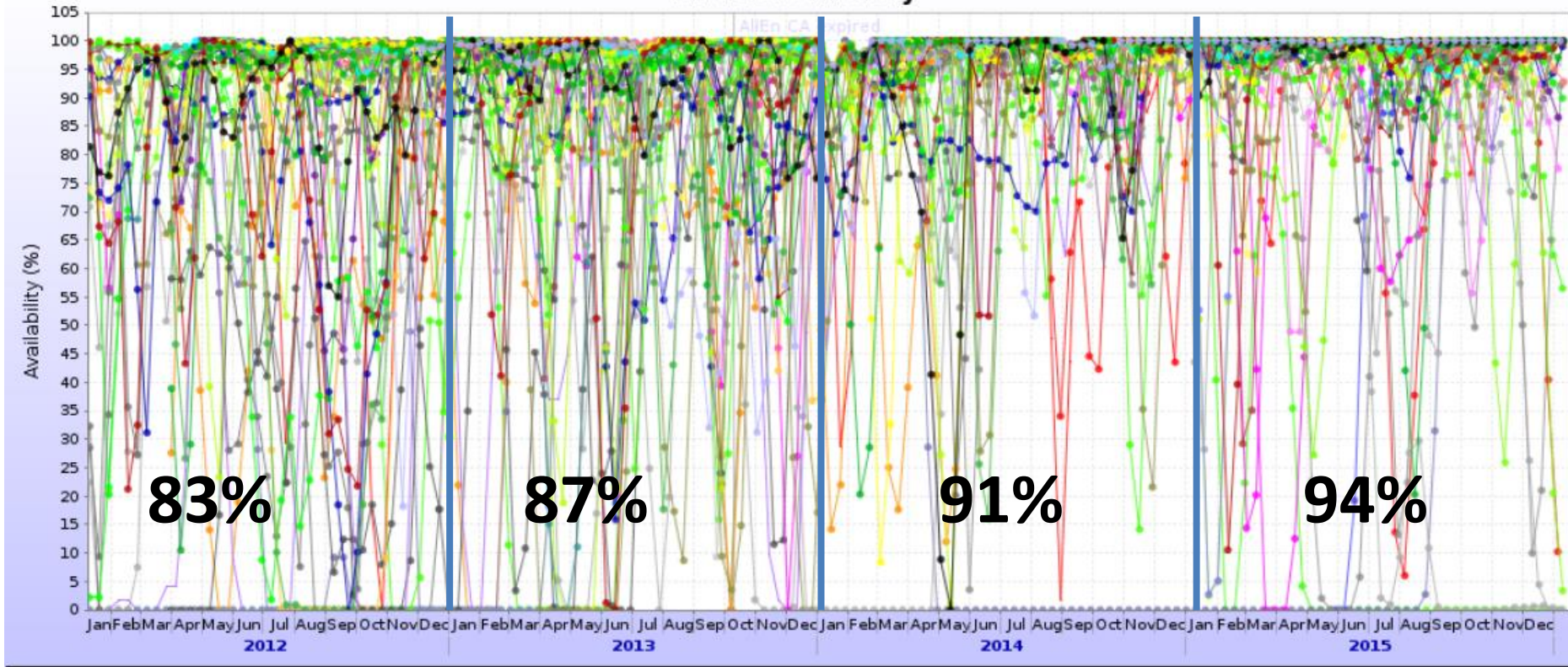
-4%

Grid efficiency evolution

- Efficiency is ~flat
- Decrease in 2015 largely attributable to specific RAW data reconstruction cycles
 - Decrease in individual analysis helps to compensate the lower RAW reco efficiency
- Specific effort to increase the efficiency of calibration tasks
 - Part of the new calibration suite development
- In general, ~85% efficiency is perhaps the maximum we can expect

Storage availability

SE tests history



Year on year change

↑
+4%

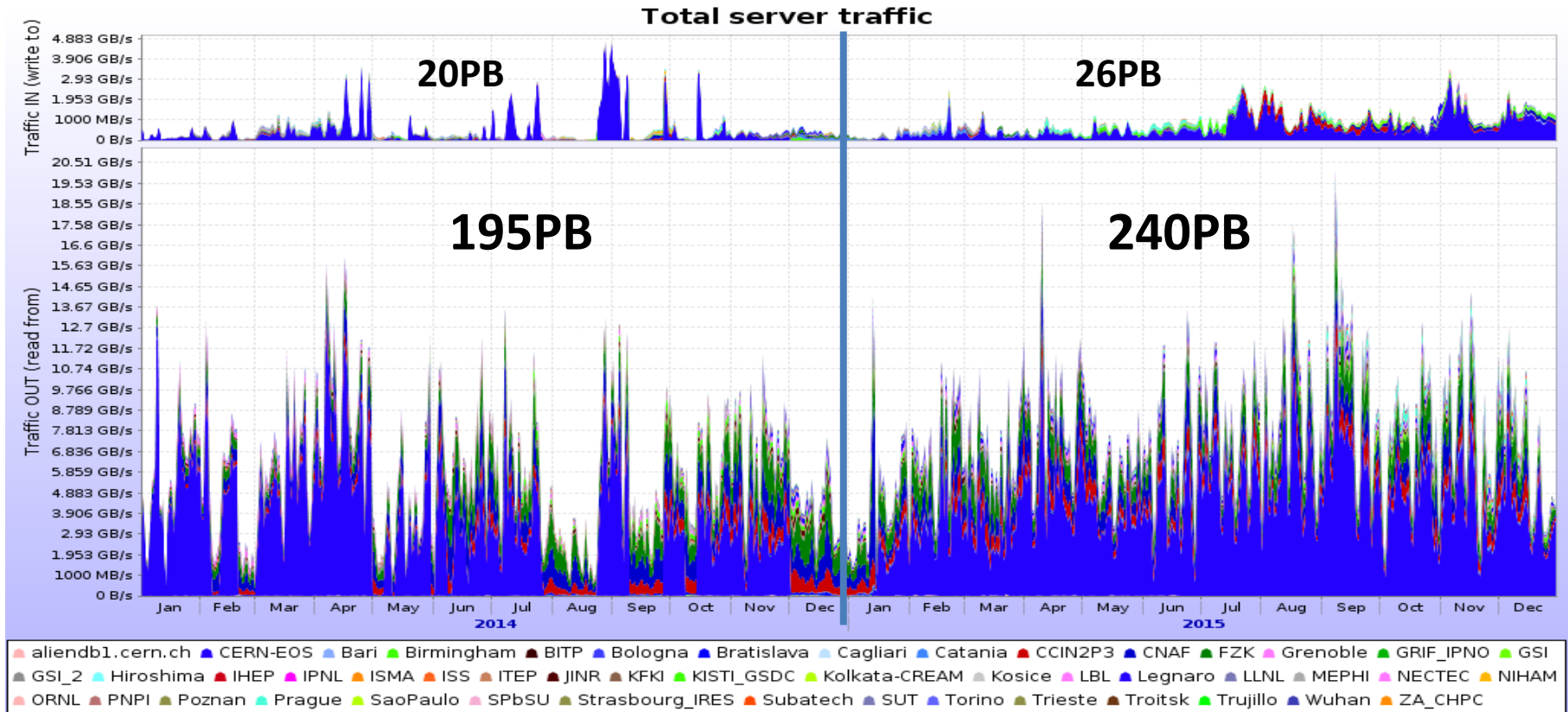
↑
+4%

↑
+3%

Storage availability evolution

- SE stability is still improving
 - SEs are independent, no correlation in downtime
- 9 of 62 SEs are below 90% availability
 - <5% of the total capacity
- 15 of 62 (25%) below the target 95% availability
 - <15% of the total capacity
- The average availability is pulled down by small SE
 - Hence the effect on analysis is minimal
- Replica model is changing to 1 ESD copy only
 - The analysis is AOD-mostly (last groups are moving to AODs now)
 - Saving of about 20% of disk

Storage use



Year on year change

Write +30%

Read +23%

Ratio 9.2

Storage use evolution

- New, more refined, storage monitoring
 - Server side, client side (both local and remote)
 - Check it out!
- Read volume increase – proportional (but not equal) to the increase of analysis activities
 - Less individual, more organized, more AODs
- Cleanup remains a constant activity
 - ‘Orphan files’ – not in the catalogue - < 1%
 - Deletion of files and replica reduction – as needed

Summary

- 2015 was (yet another) very successful year for Grid operations
- Spectacular data taking (all RAW data safe)
- Impressive resources increase and delivery
 - No (or negligible) effect of site upgrades
- Efficiency remains high, site stability remains high
- New projects started, in line with the ALICE upgrade plans (some covered in subsequent presentations)
- The operation is smooth – all ‘pilots’ know their ‘equipment’
- ***On track for the second year of LHC Run2***

Thank you to all who contributed to 2015
being another great Grid year!

Thanks to all speakers

For those who still did not do it – please upload your presentations

Tusen takk to our Bergen hosts

Let's continue...