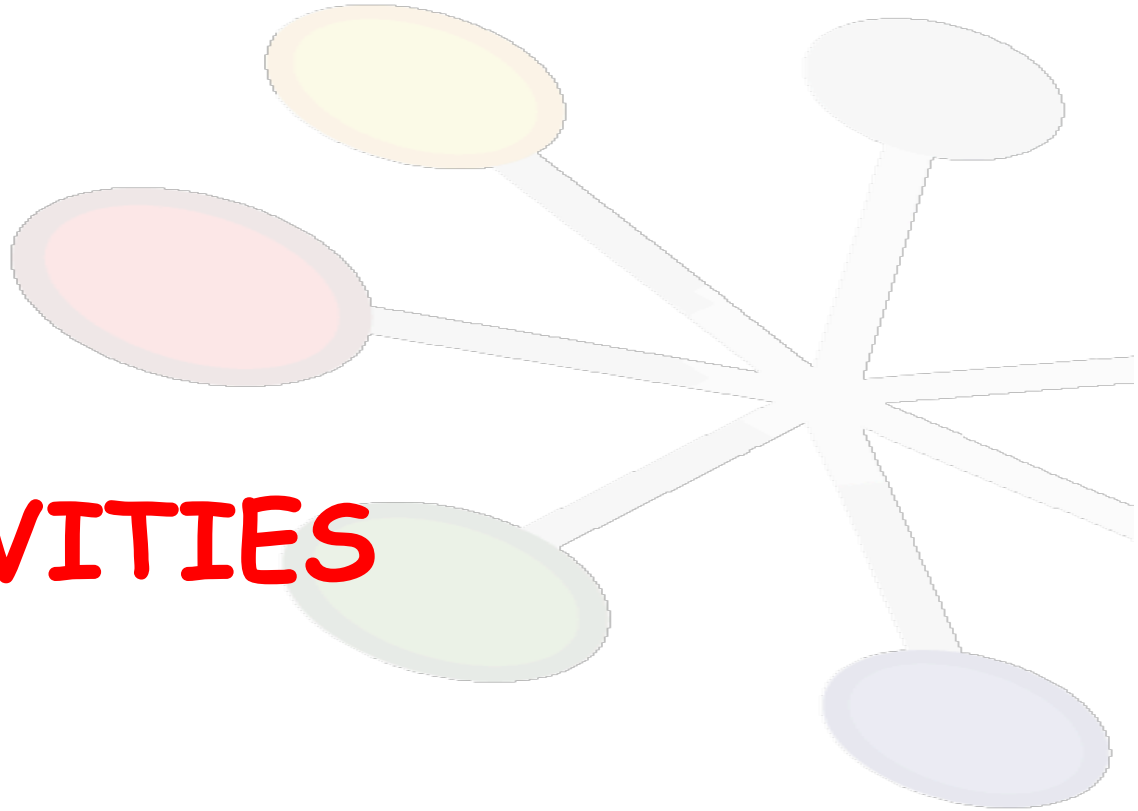
LHCb computing status

Marco Cattaneo

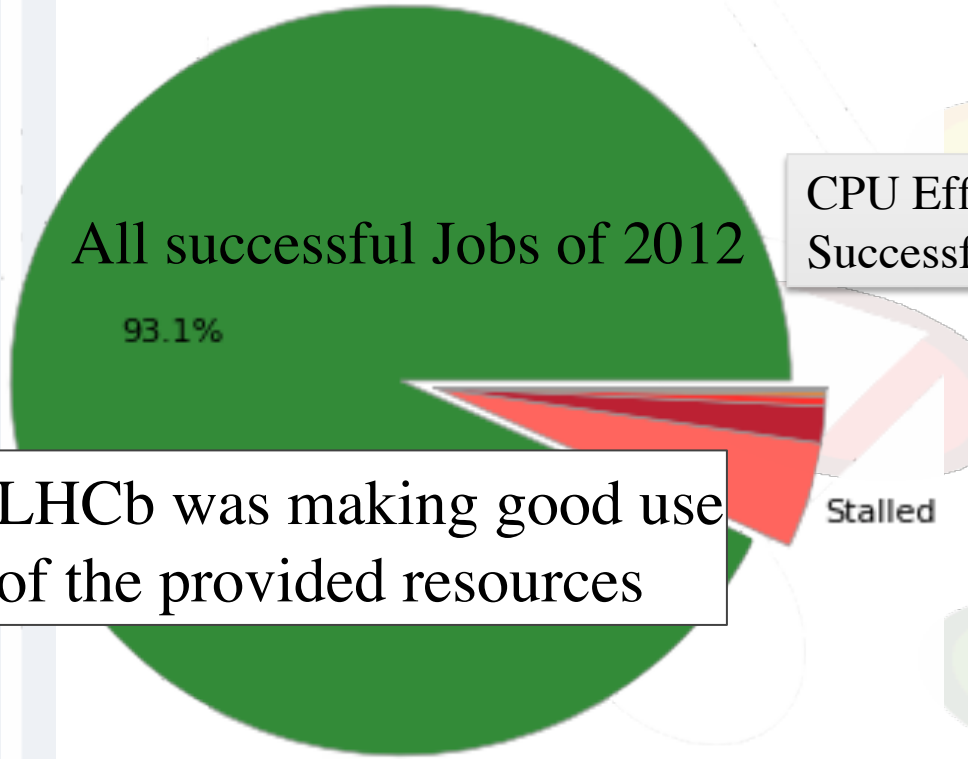
CERN – LHCb

Slides prepared by Stefan Roiser and Philippe Charpentier

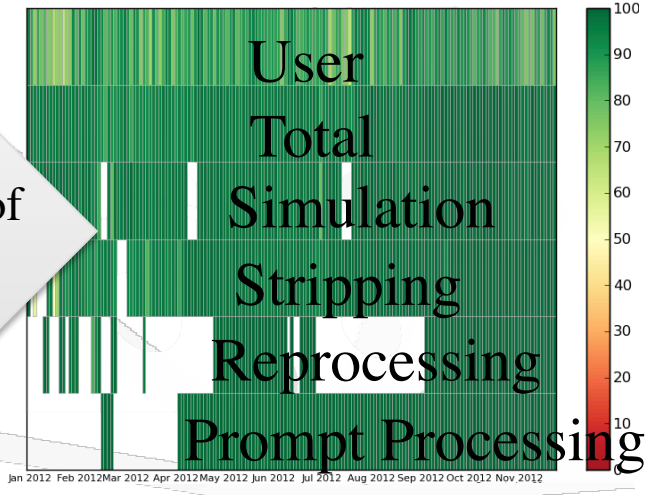


2012 ACTIVITIES

2012 Processing Overview
(all plots of this talk since 1 Jan)

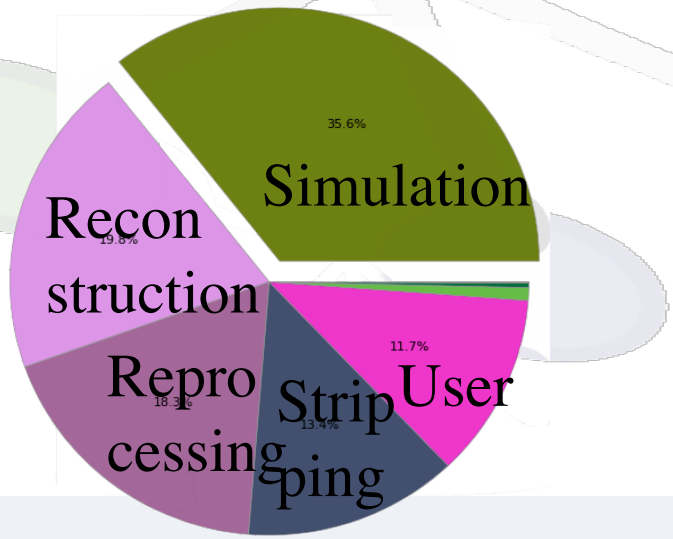


CPU Efficiency of Successful jobs

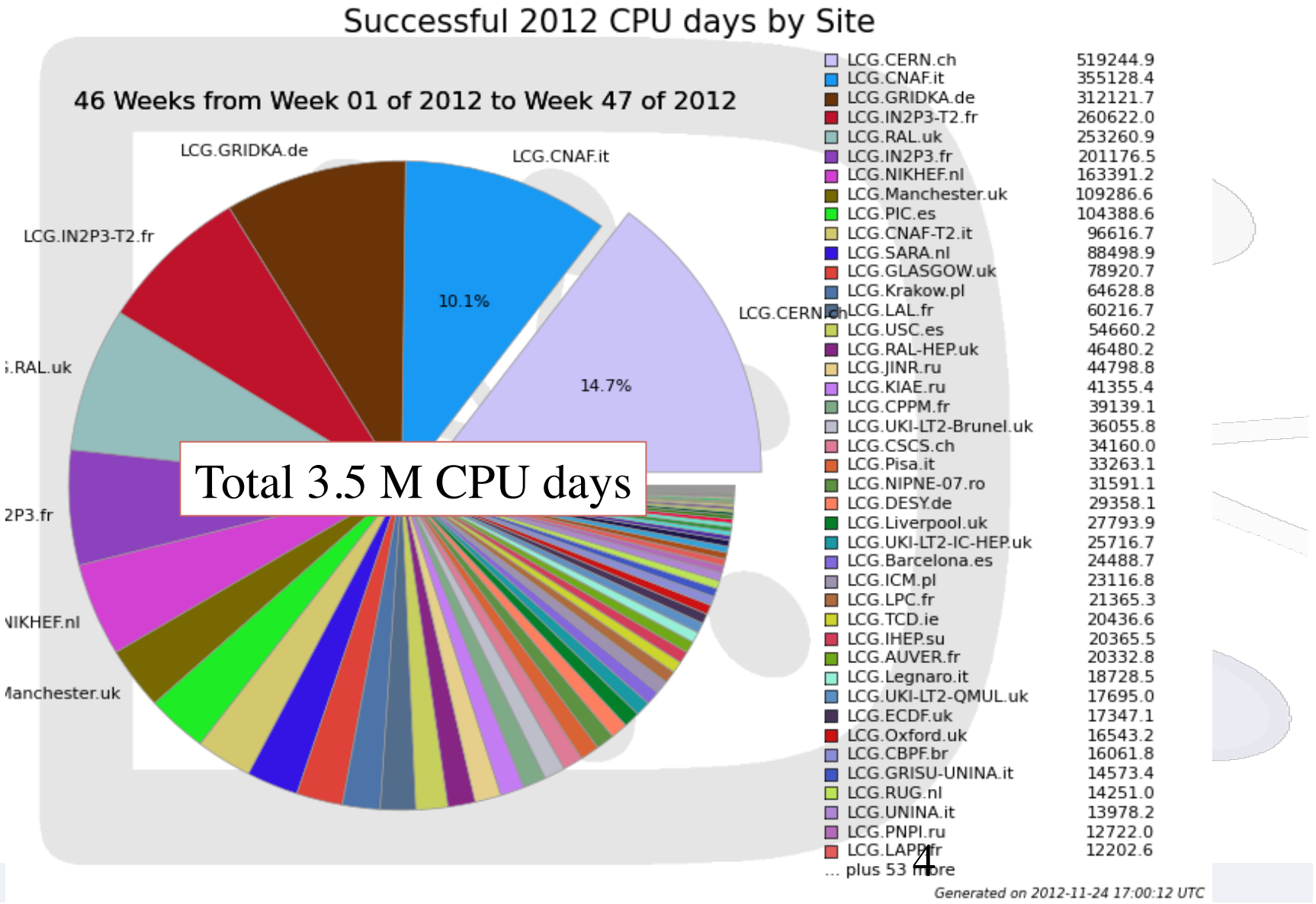


LHCb was making good use of the provided resources

All successful Work grouped by Activity

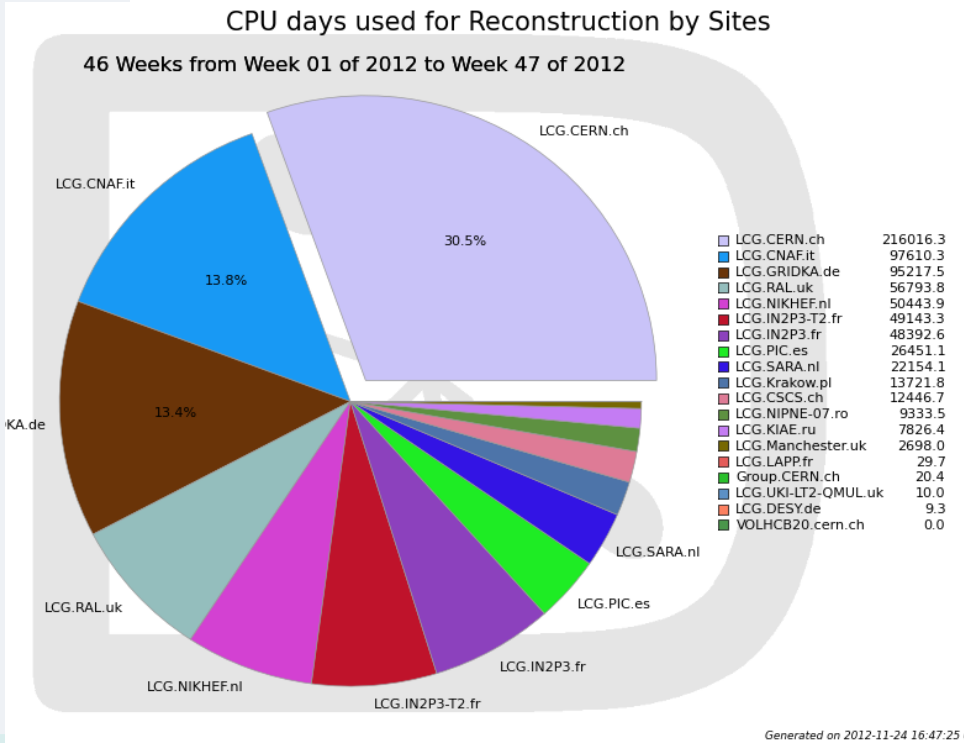


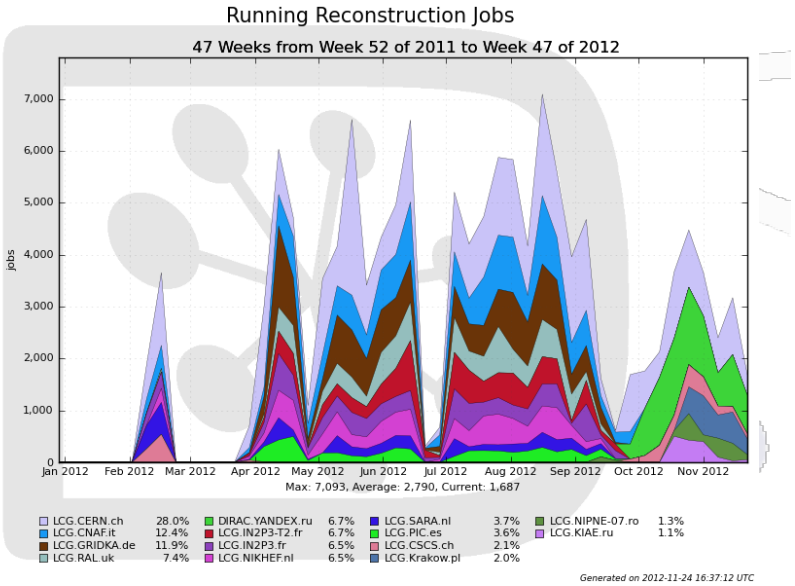
Successful Work in 2012 by Site



Prompt Reconstruction

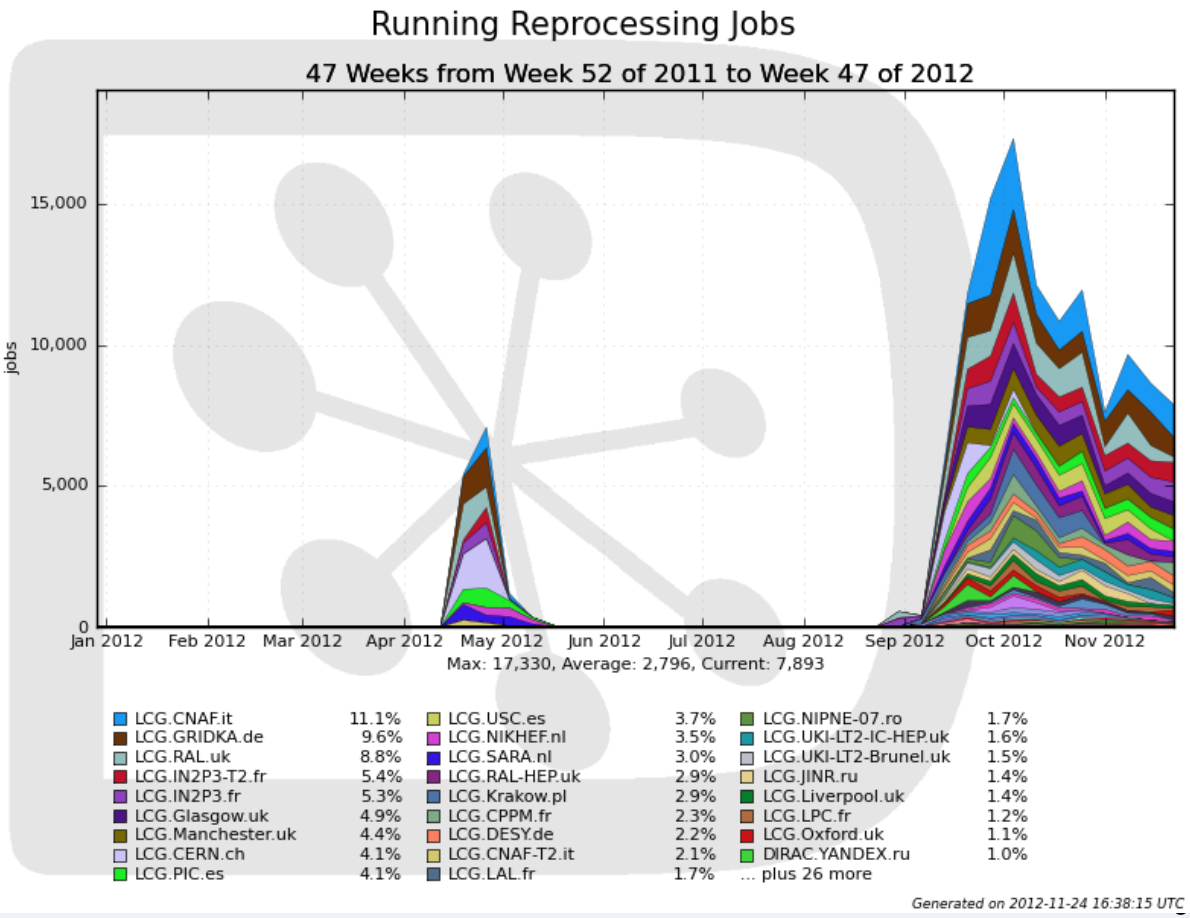
- First pass reconstruction of detector data
  - Usually 100 % of RAW files are processed.
  - Since reprocessing started (September), only partial (~30%) reconstruction at CERN + "attached T2s"
    - Only used for Data Quality and detector calibrations



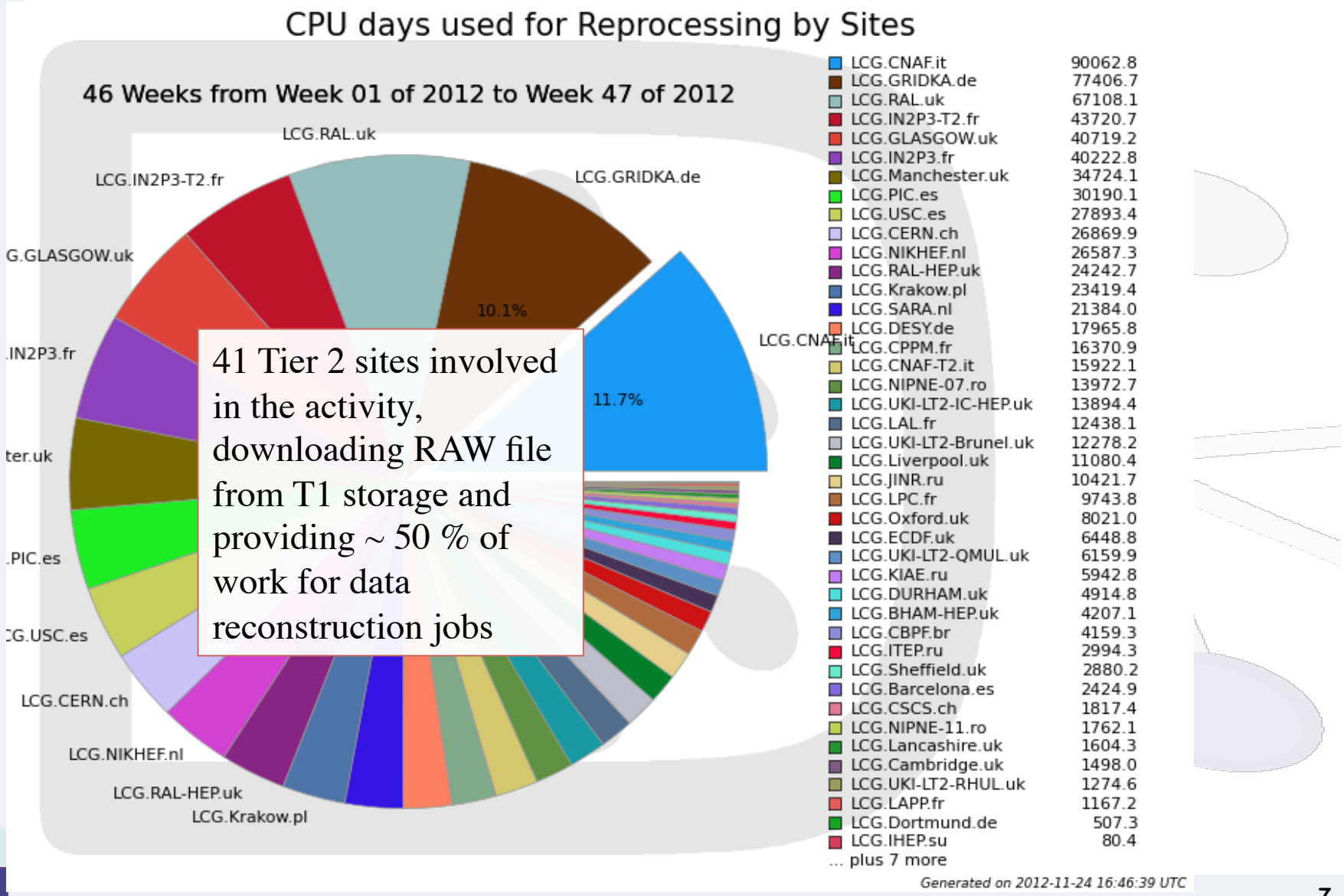


Data Reprocessing

- Reprocessing of 2012 data started mid Sept.
- Pushing system to its limits
  - Running up to 15k reconstruction jobs was a very good stress test for post LS1 prompt processing
  - Data processing smooth
  - Hitting limit with data movement, e.g. staging

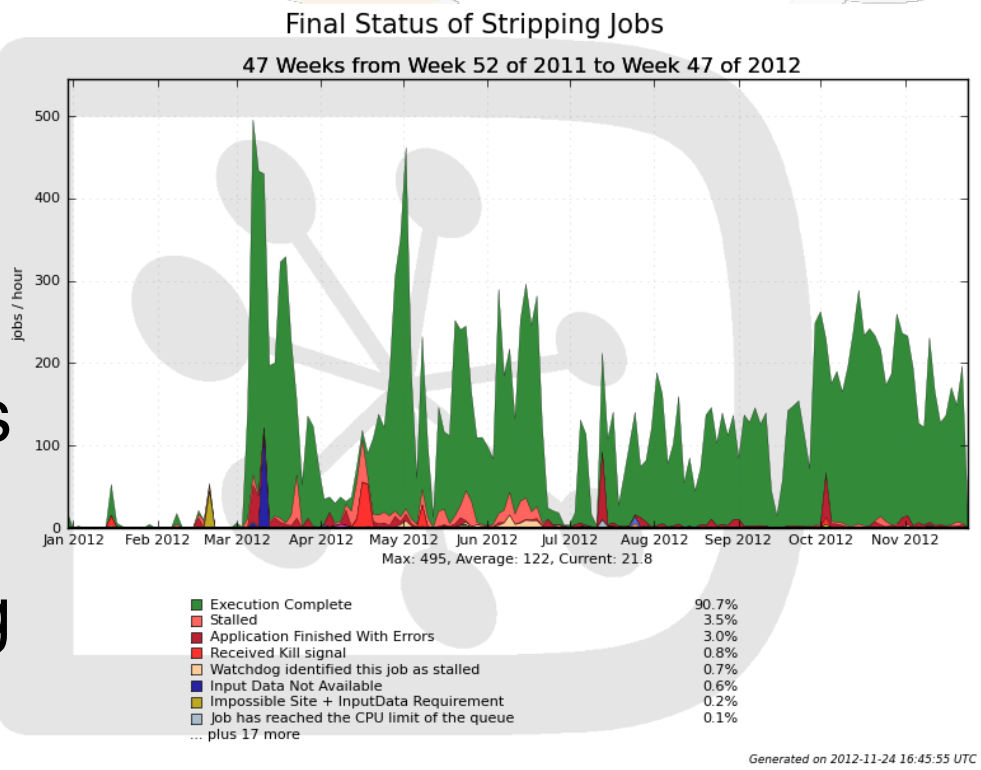


2012 Data Reprocessing

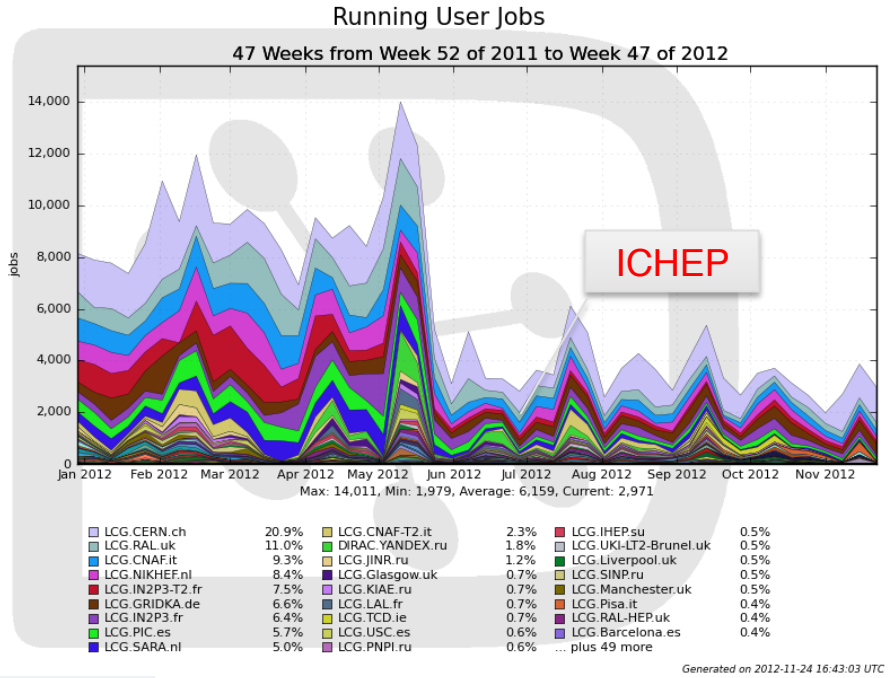


41 Tier 2 sites involved in the activity, downloading RAW file from T1 storage and providing ~ 50 % of work for data reconstruction jobs

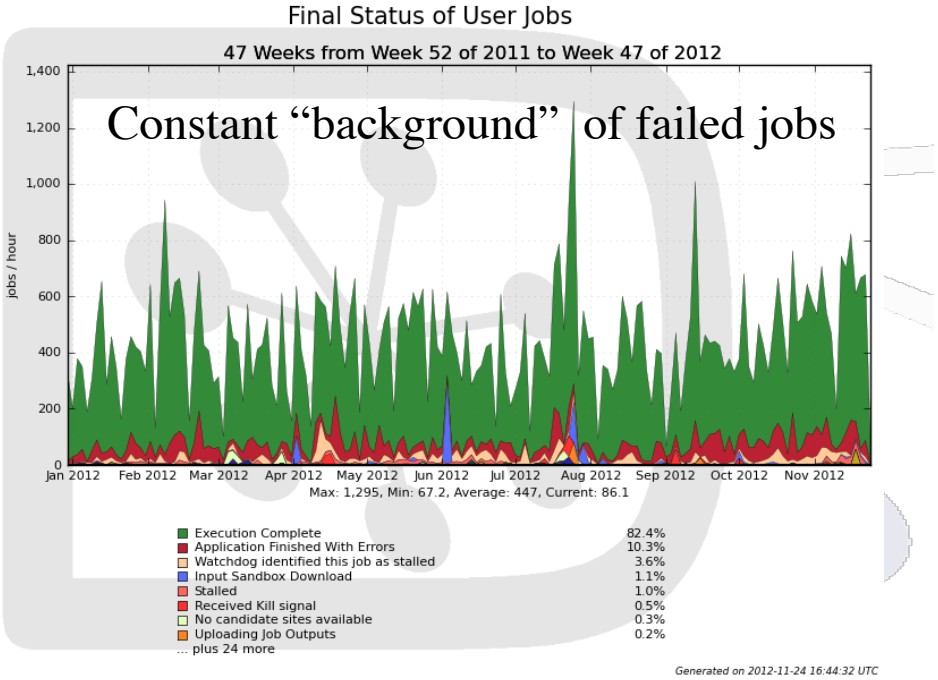
- Mostly running at Tier2 sites
  - But also at Tier0/1 when resources are available
  - Usually running with lowest priority
  - Low error rate as also true for other production activities
  - Backlog building up due to reprocessing



User Activities



Higher activities until summer, since then fewer running jobs



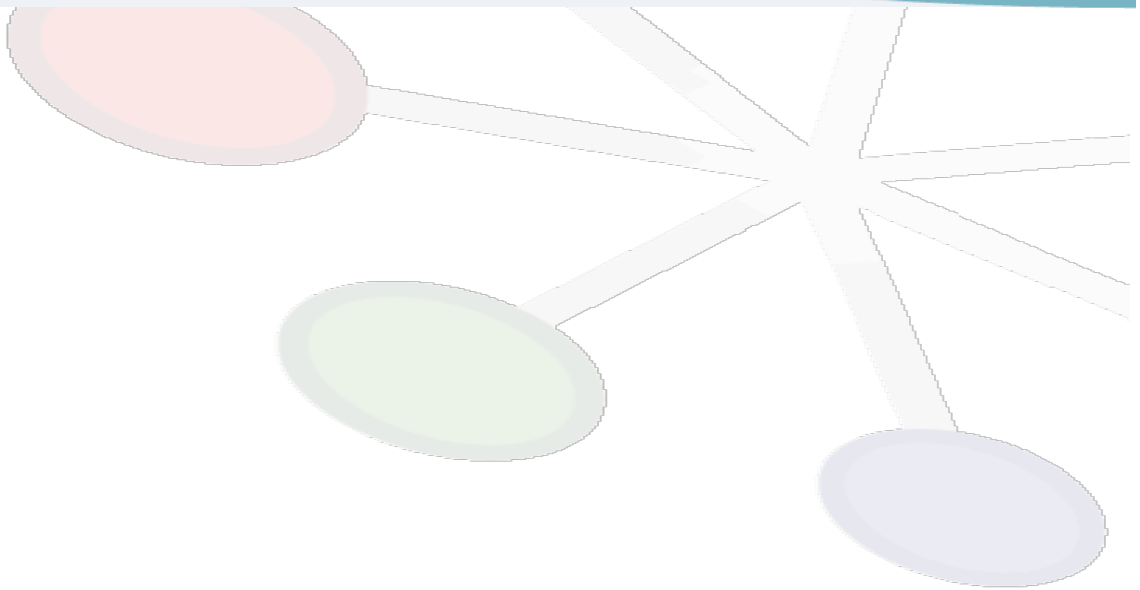
Next Processing Activities

Activity	Approx Time + Duration
2012 data reprocessing	Sep '12 – Jan '13
2013 pA prompt processing	Jan-Feb '13
2011 data reprocessing	Beginning '13 (~ 1 ½ months)
Incremental stripping	~ 2 x / year in 2013 (~ 2 months)
2011/12 data reprocessing	During 2014 (~ 5 months)

Loads on sites storage systems

- Reprocessing: Reconstruction + Stripping + Merging
  - Reconstruction run on "attached T2 sites"
  - Staging all RAW data from tape
  - Reco output (FULL.DST) migrated to tape (via disk BUFFER)
  - Replication of Merging output (DST) on multiple sites
- Incremental Stripping: Stripping + Merging
  - Staging of all FULL.DST files
  - Producing up to ~ 20% additional DST files
  - Replication of DST on multiple sites

Storage status & forecast




Tape crisis

- As announced, tape is a real issue:
  - More data than expected: longer LHC run, higher HLT rate
  - Usage of FULL.DST instead of SDST
    - 5 GB instead of 2 GB for a 3 GB RAW file
    - Mitigated by removing previous processing's SDST
    - FULL.DST on disk until stripping is complete
      - Less stress on the tape cache disks
- Mitigated by removing non-CERN archive of many datasets (now only one archive for all datasets)
  - Stopped archiving at GRIDKA, IN2P3, PIC and SARA
- Sites have kindly agreed to let us use more than pledged
  - RAL provides an additional 1 PB of tape
  - Other sites on best effort
    - SARA "very close" to the end
  - Agreement between LHCb and ALICE for loan of tapes
    - GRIDKA, CNAF if needed

Situation w.r.t. pledges

Site	Storage	Files	Space (TB)	Installed	Pledge 2012	Pledge 2013
CERN	T1D*	1'774'655	4'605.1		6'400.0	6'500.0
	T*D1	6'553'141	2'182.3	3'150.0	3'500.0	4'000.0
CNAF	T1D*	445'567	1'548.6		900.0	1'600.0
	T*D1	1'009'805	825.6	1'382	1'400.0	1'300.0
GRIDKA	T1D*	404'033	1'343.8		1'050.0	1'050.0
	T*D1	1'002'565	848.9	1'470.0	1'610.0	1'450.0
IN2P3	T1D*	409'660	1'332.5		1'000.0	1'400.0
	T*D1	1'066'455	856.8	1'019.0	1'090.0	1'200.0
PIC	T1D*	128'113	428.3		306.0	551.0
	T*D1	689'790	454.0	672.0	485.0	439.0
RAL	T1D*	383'720	1'259.1		1'116.0	2'010.0
	T*D1	1'215'712	1'130.5	1'994.0	1'767.0	1'600.0
SARA	T1D*	383'509	1'247.6		952.0	2'100.0
	T*D1	958'996	681.6	789.0	810.0	1'008.0
Tier1s	T1D*	2'154'602	7'159.9		5'324.0	8'711.0
	T*D1	5'943'323	4'797.4	7'326	7'162.0	6'997.0

- CERN: some disk ready to be installed
- Part of tape cache not included for dCache sites
- WARNING: decrease of requirements/pledges in 2013!!

Disk pledges 2013 vs 2012

Site	Fraction of req. 2012	Fraction of pledges 2012	Pledge 2012	Pledges 2013 from 2012 frac.	Pledge 2013	Pledges vs req. 2013
CNAF	14.7%	19.5%	1'400	1'681	1'300	15.1%
GRIDKA	16.9%	22.5%	1'610	1'933	1'450	16.9%
IN2P3	11.5%	15.2%	1'090	1'309	1'200	14.0%
PIC	5.1%	6.8%	485	582	439	5.1%
RAL	18.6%	24.7%	1'767	2'122	1'600	18.6%
SARA	8.5%	11.3%	810	973	1'008	11.7%
Tier1s	75.4%	100%	7'162	8'600	6'997	81.4%

- 2012: 75.4% of requirements pledged
  - Requirements were mitigated and reviewed
- 2013: new fraction at some Tier1s
  - When applied to a "reduced" requirement, decrease of pledges
  - Proposal: use 2012 fraction of pledges for 2013
    - More inline w.r.t. 2012 pledges

Conclusions

- Very successful year for both LHC and LHCb
  - Computing model adjusted to mitigate effect of increased volume of physics data
- Tape usage way over 2012 pledges
  - Additional capacity on best effort from sites
    - Reaching limits at some sites
      - Offer of extra loan from Alice, but not at tightest sites
  - 2013 tape pledges mostly OK
- Clarification needed on calculation of Tier1 fractions of pledges