



LHCb: Operational Report

On behalf of LHCb



Roberto Santinelli- GDB 08th December 2010



- **News & Activities**
 - **GGUS tickets summary**
 - **Site round issues and outstanding issues**
 - **Conclusions**
- 



(Relevant) September GDB's conclusions

- Analysis is not now a so big unknown ($O(100)$ different users) with spikes of the activities when new data become available or conferences approach (ICHEP).
 - ↳ Combined with smallish unmerged files, system accuses the increased load, storages are still the most vulnerable component.
- We still believe that a proper size of servers (disk servers and # slots, SRM, shared area) wrt #batch slots as well as hard division of resources to avoid interferences, has to be addressed.
- Load put in the system in this quarter would not allow for a fair comparison wrt Spring quarter. New services in production new instabilities.
 - ↳ Usage of relatively new services (or not completely exercised) reflected into increased failure rates (CREAM), backlogs forming (LFC and ConditionDB)
- We were/are running in shortage of space because events are larger then expected due to pileup 5 times the design. (LHC delivering few high intensity bunches instead of many low intense).



Hitting everywhere space limits

- Adjusted HLT cuts on the ONLINE.
- revision of the data distribution policy.
- massive cleanup campaigns of previous reco-stripping output.

Introduced in DIRAC a throttling mechanism that limits the load on grid services configurable per type of activity and per site.

CREAM used in production:

- Big improvement since last release as declared by devs.
- Evaluating direct submission at CNAF and NIKHEF

xrootd: encouraging users to use it at CERN for reading (LSF slot not taken for reading)

CERNVMFS for serving the application area:

- PIC, RAL NIKHEF infrastructure in place.
- CNAF CERN given also immediately positive answers.

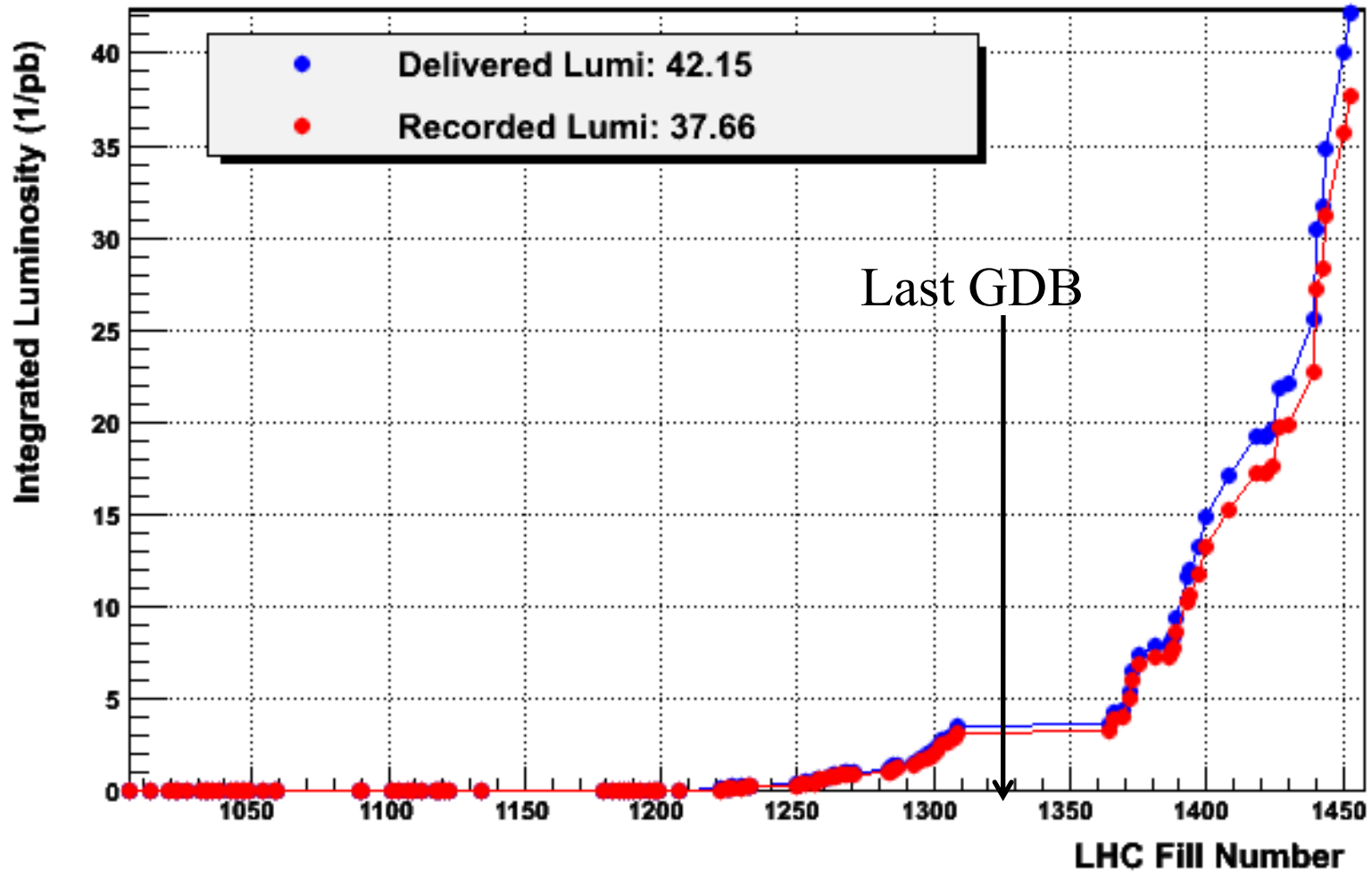
Looking for alternative to Oracle for Condition DB:

- Frontier/Squid, SQLDDDB, Condition Information embedded in the ROOT file.



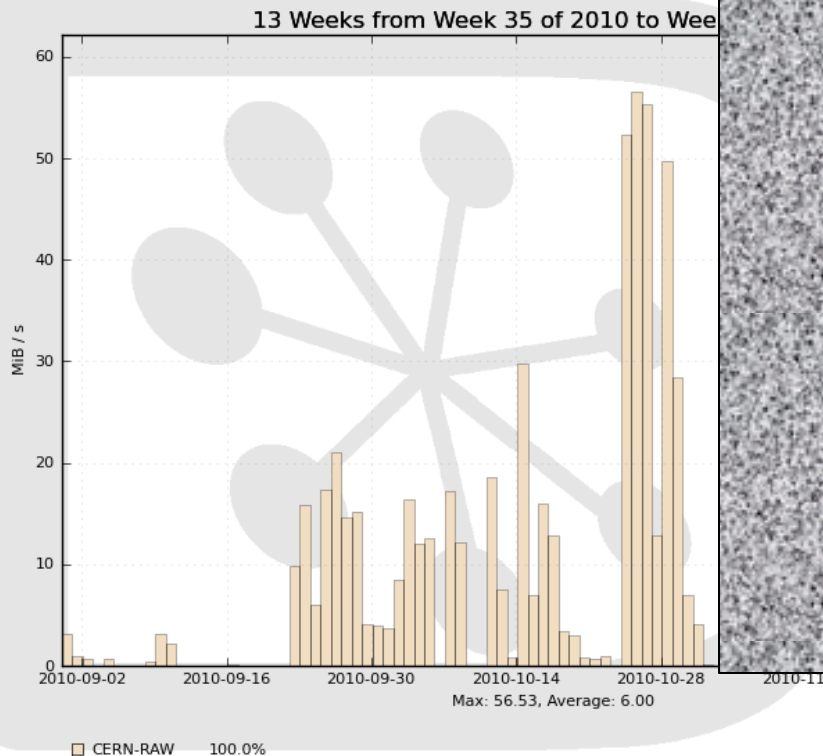
LHCb Integrated Lumi over Fill Number at 3.5 TeV

2010-12-04 12:00:03



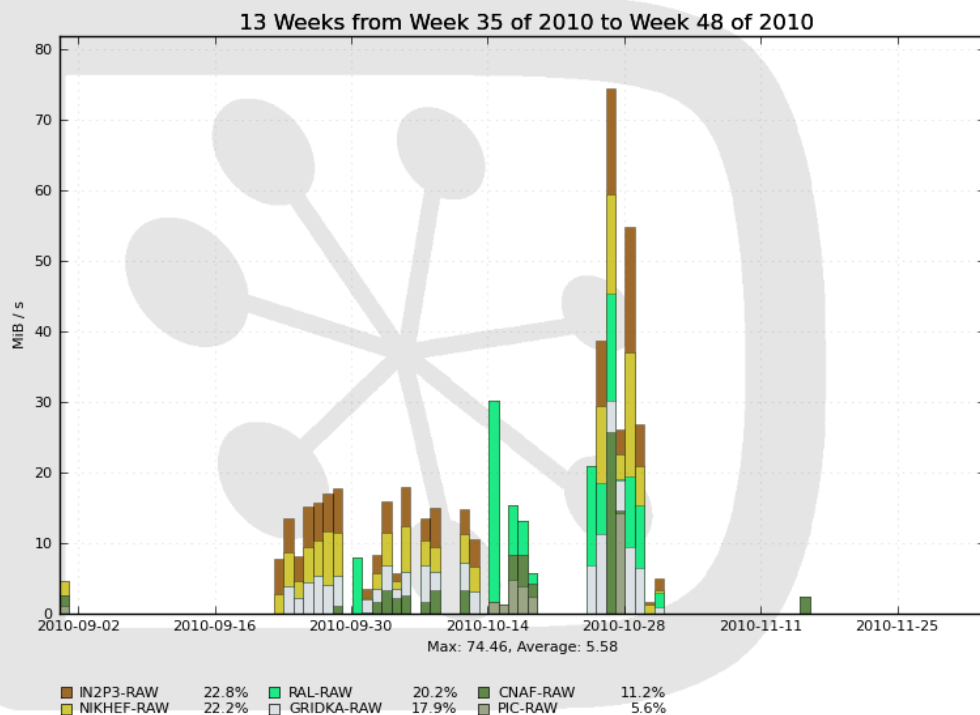
Throughput by Destination

Data Transfer



LHCb: not heavy ion physics

Throughput by Destination



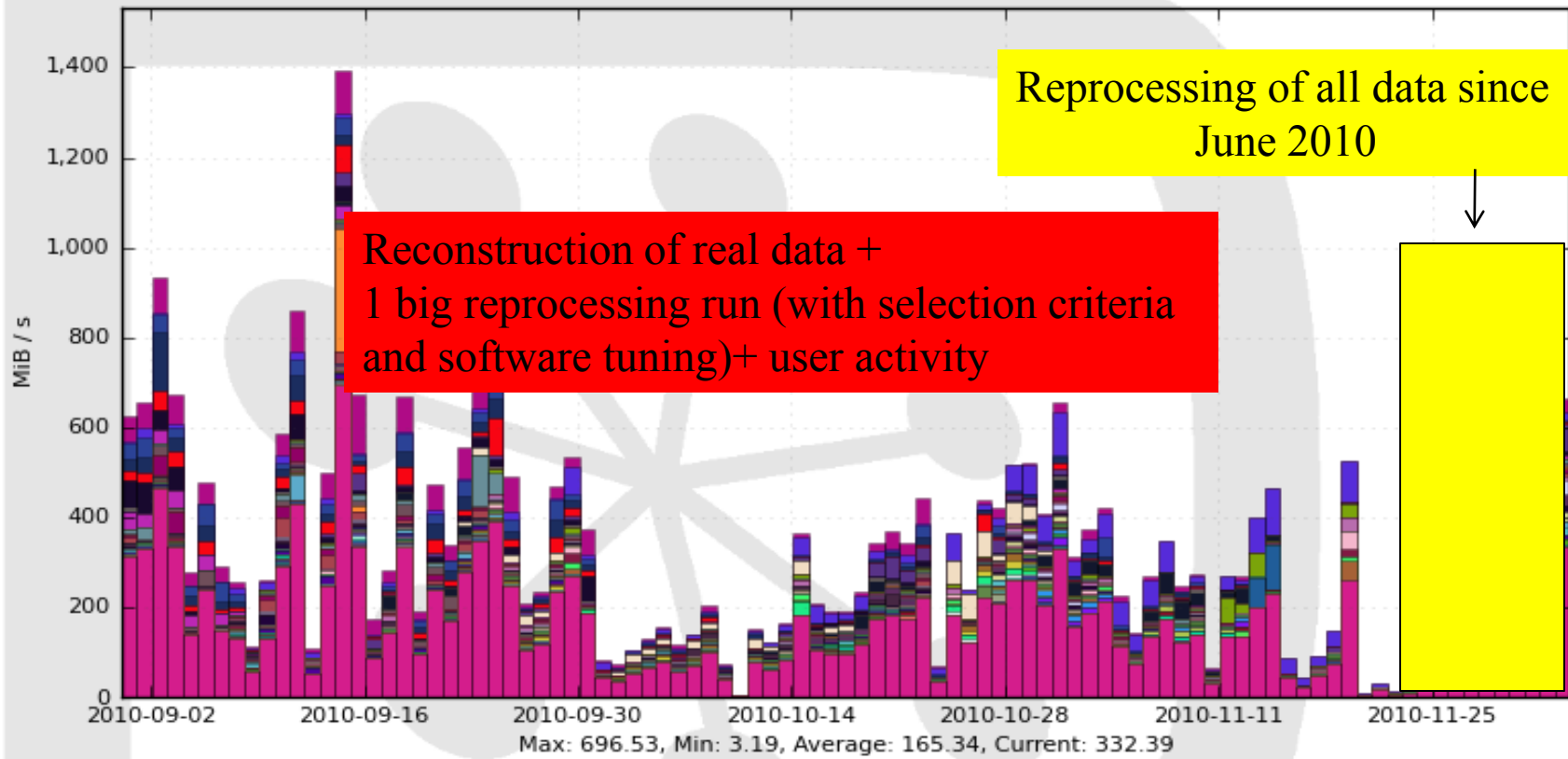
The ratio to each sites closely **reflects** the share defined (in turn based on top of pledges from each site).



Aggregate data throughput

Throughput by Destination

13 Weeks from Week 35 of 2010 to Week 48 of 2010



| | | | | | |
|-----------------|-------|----------------|------|------------------|------|
| ■ CERN_M-DST | 11.9% | ■ IN2P3-DST | 2.3% | ■ RAL-RDST | 1.8% |
| ■ CERN-USER | 11.8% | ■ IN2P3-USER | 2.2% | ■ CERN-DEBUG | 1.7% |
| ■ GRIDKA-DST | 5.9% | ■ IN2P3_M-DST | 2.2% | ■ NIKHEF-RDST | 1.7% |
| ■ NIKHEF-DST | 5.9% | ■ RAL_M-DST | 2.1% | ■ GRIDKA-RDST | 1.6% |
| ■ CNAF-DST | 5.4% | ■ PIC-DST | 2.0% | ■ GRIDKA-USER | 1.5% |
| ■ CERN-RDST | 4.6% | ■ GRIDKA_M-DST | 1.9% | ■ NIKHEF-USER | 1.3% |
| ■ RAL-DST | 4.2% | ■ NIKHEF_M-DST | 1.8% | ■ CNAF-USER | 1.2% |
| ■ CERN-RAW | 3.6% | ■ RAL-USER | 1.8% | ■ PIC-USER | 1.2% |
| ■ CERN_MC_M-DST | 3.5% | ■ IN2P3-RDST | 1.8% | ... plus 48 more | |

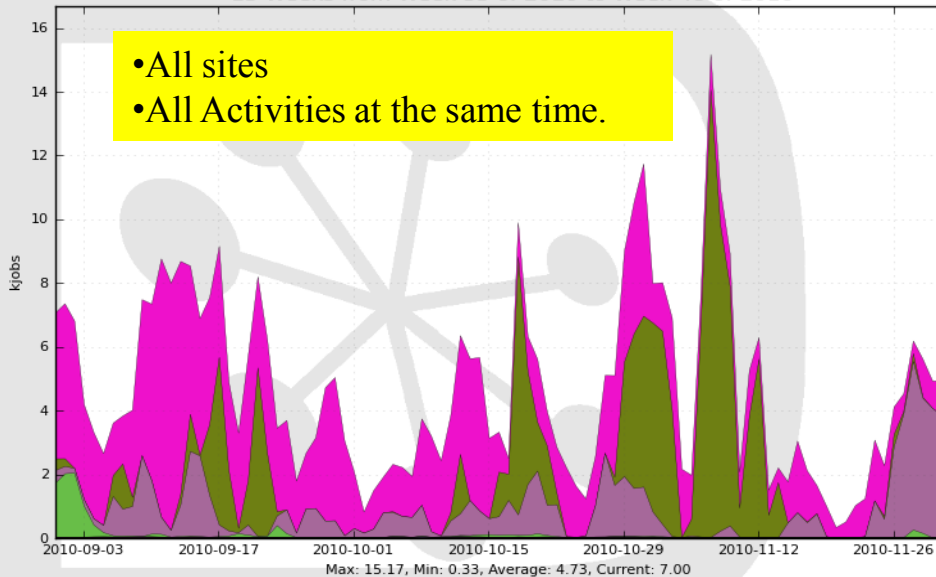


Activities run over the last 3 months

- Real Data Reconstruction/Stripping/Merging. One major reprocessing end of November over all data since June.
- Important user activities with spikes corresponding to new merged data made available.
- MC production and calibrations

Running jobs by JobType

13 Weeks from Week 35 of 2010 to Week 48 of 2010

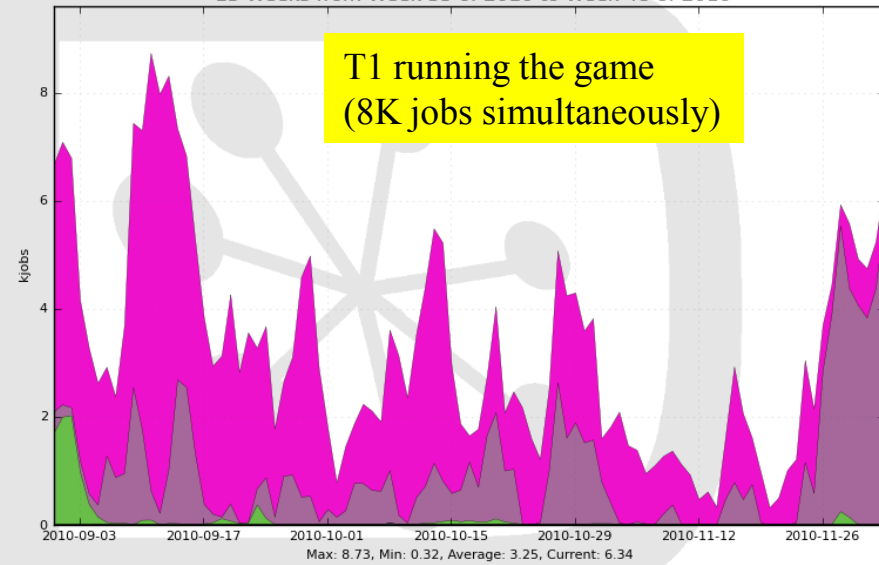


| | | | | | |
|--------------------|-------|----------|------|---------------|------|
| user | 49.7% | Merge | 2.1% | DataStripping | 0.0% |
| MCSimulation | 28.0% | sam | 0.5% | unknown | 0.0% |
| DataReconstruction | 19.6% | Hospital | 0.0% | | |

Generated on 2010-12-04 15:32:07 UTC

Running jobs by JobType

13 Weeks from Week 35 of 2010 to Week 48 of 2010



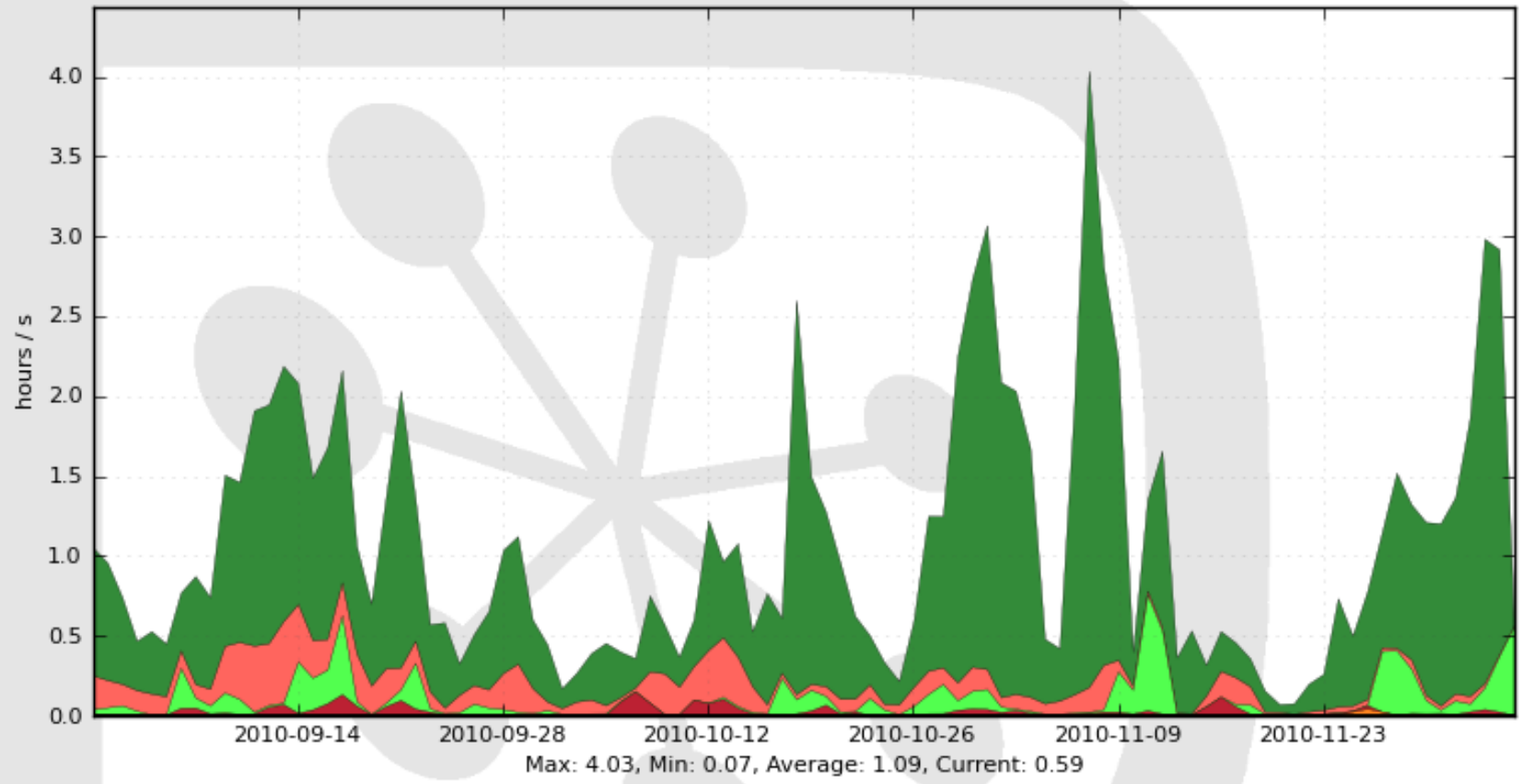
| | | | | | |
|--------------------|-------|-------|------|---------------|------|
| user | 68.3% | Merge | 3.0% | DataStripping | 0.0% |
| DataReconstruction | 28.5% | sam | 0.1% | MCSimulation | 0.0% |

Generated on 2010-12-04 15:30:47 UTC



CPU usage by FinalMinorStatus

13 Weeks from Week 35 of 2010 to Week 49 of 2010

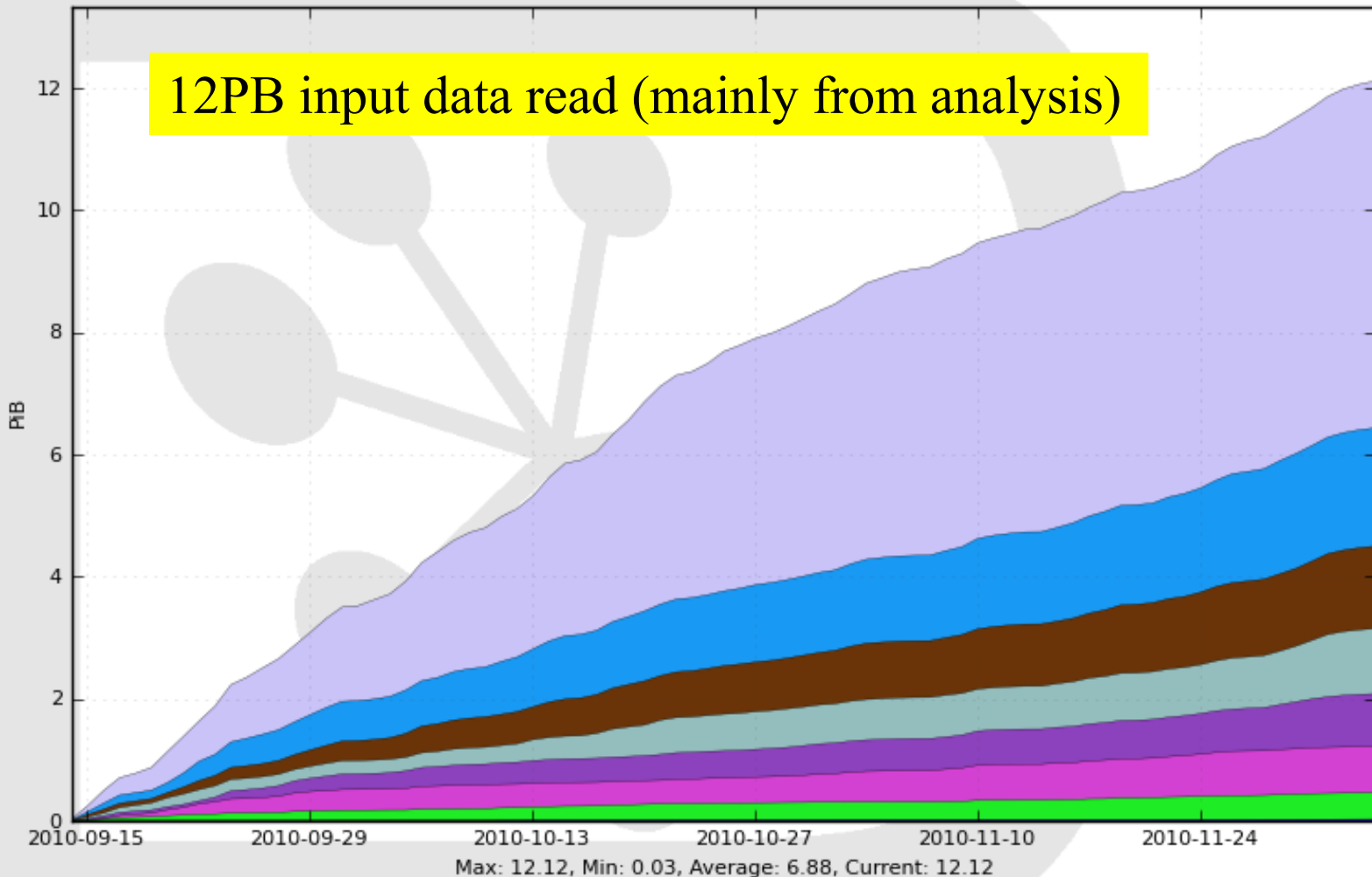


| | |
|--|-------|
| Execution Complete | 78.7% |
| Stalled | 11.2% |
| Pending Requests | 7.7% |
| Application Finished With Errors | 2.3% |
| Uploading Job Outputs | 0.2% |
| Input Data Resolution | 0.0% |
| Received Kill signal | 0.0% |
| Job has exceeded maximum wall clock time | 0.0% |
| ... plus 20 more | |



Cumulative Input data by Site

11 Weeks from Week 37 of 2010 to Week 49 of 2010



| | | | | | | | |
|-------------|-----|---------------|-----|---------------|-----|-------------|-----|
| LCG.CERN.ch | 5.7 | LCG.GRIDKA.de | 1.4 | LCG.IN2P3.fr | 0.9 | LCG.PIC.es | 0.4 |
| LCG.CNAF.it | 1.9 | LCG.RAL.uk | 1.1 | LCG.NIKHEF.nl | 0.8 | LCG.SARA.nl | 0.0 |

GGUS tickets (were 203 in the previous quarter)

- 152 GGUS tickets in total:
 - 21 normal tickets
 - 4 ALARM tickets (1 *real ALARM*)
 - 127 TEAM ticket
- 23 GGUS tickets with shared area problems in total
- 73 GGUS tickets (test ALARM included) open against T0/T1:
 - CERN-PROD: 11 (1 *real ALARM*)
 - FZK-LCG2: 10
 - INFN-T1: 6
 - IN2P3-CC: 14
 - NIKHEF-ELPROD: 3
 - PIC: 5
 - RAL-LCG2: 14
 - SARA-MATRIX: 10

Please note that GGUS is also a communication tool: not necessarily the number of tickets reflects the effective shortage of the site!





Disk server instabilities observed at different times, hitting a bug in CASTOR, shortage of space in some service classes.

- 1st -7th Sept: Disk servers being killed by the full steam activity. Throttled the activity at the LRMS level. Merging also killing the disk servers in a second time 17th -20th Sept: self-limiting jobs DIRAC side (merging backlog formed)
- 27-29thSept: CASTOR DT
- 1st -5th Oct: Issue with ConditionDB (mis-configuration DIRAC side)
- 30th Sept -8th Oct : SRM problem: 0 file size & wrong locality reported by CASTOR: due to internal timeout in CASTOR bug. Reconfigured DB.
- 11th -13th Oct: Issue with checksum reported 0 by SRM: moving to 64 bit disk server the checksum will be available for both tape and disk back end
- 27-31st Oct: Issue retrieving staging files, malformed tURL. Restarting SRM fixed but it remains a mystery (UNSOLVED)
- 11th-15th Nov: Problem accessing data using rootd (diskservers problem)



No Important issues. A lot of user jobs wrt production jobs (higher I/O network activity than production activities and then some degraded CPU/wall).

- Some Issues with Condition DB until mid of September (connectivity, number of max sessions)
- 2nd Sept. Suffered some shortage of space on MC-(M)-DST tokens (pledged provided)
- 7th Sept: SRM failure (human error)
- 18th Oct. Users jobs timing up (couple of WN mis-configured)
- 28th Oct. Number of files corrupted (due to space full)
- Issue with the network core switch (19 till 22 on November)



Mainly instabilities with SRM and dcap movers hanging.
Shortage of space in some tokens

- **6th sept.** Jobs stalled problem with dcap movers to be cleaned
- **13th Sept.** jobs stalling with dcap movers hanging problem
- **5th October:** prob with CREAMCE
- **20th Oct.:** Instabilities with SRM
- **22nd Oct.:** DST space getting full
- **12th Nov:** transfers to MC-M-DST failing timing out
- **28th Nov.** timeout in turl resolution and transfer (disk servers pb)



NOT ONLY SHARED AREA! Disk servers issues, WN with buggy kernel

- 8th Sept. failures transferring to dst space token
- 23th Sept. back from DT problem with the shared area. Backlog forming
- 1-8th October: Segfault problem due to newer WN with buggy kernel
- 12-22nd Oct.: Faulty disk servers with many jobs timing up accessing
- 26th Oct.: files with turl not retrievable
- 12th Nov: many jobs timing up with shared area



Shared area issue.

- **Opened 4 tickets (the first #59880 dated 08 July)**
 - Affecting SW installation job (installing apart)
 - Production and User activities setting up the environment
- **Exhaustive reports received (after initial lack of communication)**
 - It is rather about a tricky problem to be fixed.
 - LHCb provided collaboration (testing, info, increasing timeouts→1hr)
 - Joint effort with other sites, many directions of investigation found!



Major issues with Condition DB and shared area. Fair share issue at SARA in December.

- Sept: suffering disk space shortage
- 14th Sept: CREAMCE issue replicated on October
- 11-18th Oct: Major issue with the ConditionDB (inconsistency tags)
- 17th Nov: data unavailable Pool not restarted
- 25th Nov: problem with CREAMCE
- 26th Nov: problem with authz in SRM
- Shared area issues spotted
- 6th Dec: FS issue

General difficulty to deal with two sites in one (see the SARA share problem). We assume a share for a site not a share for two sites!



Shortage of space (pledged provided) shared area problems mainly.

- Sept: suffering disk space shortage
- 1-3th Sept: Shared area issue
- 13th-14th Sept: Files UNAVAILABLE (instability dcache pool, network server)
- 17th Sept: :20% of transfers failing due to a variety of problems with SE
- 22th sept: problems with FTS
- 1st - 5th Nov and again 8-9 : problem with shared area.



- Sept: LFC replication to T1 slow: one flag was not enabled
- Some slowness with AFS shared area (few occasions in October and beginning of November)
- Xrootd problem accessing data on all service classes:
 - ↳ Affecting users with FQAN /lhcb/Role=NULL
 - ↳ Problem spotted the 8th Of Nov.
 - ↳ Decision taken on Friday 12th to patch a grid-mapfile like file
 - ↳ Monday 15th Problem understood on the **xrootd redirector** and its grid mapfile.

LHCb_DST incident: all transfers and data accesses were failing for all week-end (27-30 Nov, problem still there yesterday)

- ↳ Problem too high number of concurrent accesses.
- ↳ System degrading and starting random I/O → throughput dropped.
- ↳ Further incoming requests piling up → “snow ball effect”

Solution:

- ↳ More h/w added
- ↳ Bug found (internal communication in castor of failing uploads): timeout of LSF too high and slot left hanging for 30 minutes occupying and preventing further requests honored.
- ↳ Quoting Ponce: “The key problem is in the mismatch between resource scheduling and polling protocols from clients (GFAL).”

Spindles Vs Space

- ↳ Effective need is 2 TB → 40TB disk servers (#2) are not enough!
- ↳ Use different technologies: fast small disks?
- ↳ Change the workflows in order to avoid small files





- Taken $\sim 40 \text{ pb}^{-1}$ and reprocessed all data since June 2010 w/o major problems with users constantly creating load.
- CREAMCE improved very much.
 - The release 1.6.0 fixed many of the problems observed
 - Developers added in the LHCb e-logbook notification for CREAMCE
- ConditionDB access is still a problem but no longer a show stopper.
 - Round robin of servers, remote CondDB access through the WAN : CPU/Wall degradation
 - Fixed internal (DIRAC) source of inconsistencies and abandoned LFC to store connection string information.
 - Looking towards new solutions alternative to Oracle (Frontier/Squid)
- Shared area is still a serious problem (IN2p3 BUT not only)
 - We increased the timeout as done by ATLAS customizable per site: CPU/Wall degradation
 - SAM tests are felt as a reliable source of debug
 - Looking towards new solutions alternative to NFS/AFS (CERNVMFS)
- Smooth data access is still a must. Bugs, mis-configurations, service instabilities, inadequate dimensioning of servers
 - Disentangling spindles from storage space, new workflows, new technologies
 - Xrootd might help on CASTOR reading (not triggering LSF jobs)