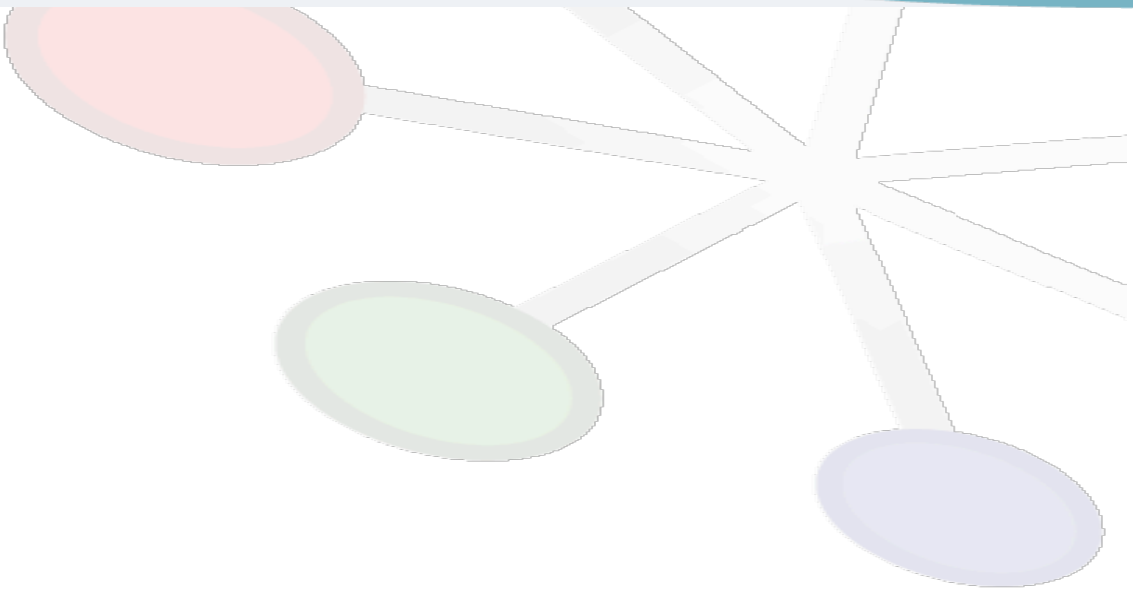
LHCb Grid Operations in 2011



LHCb Data Characteristics

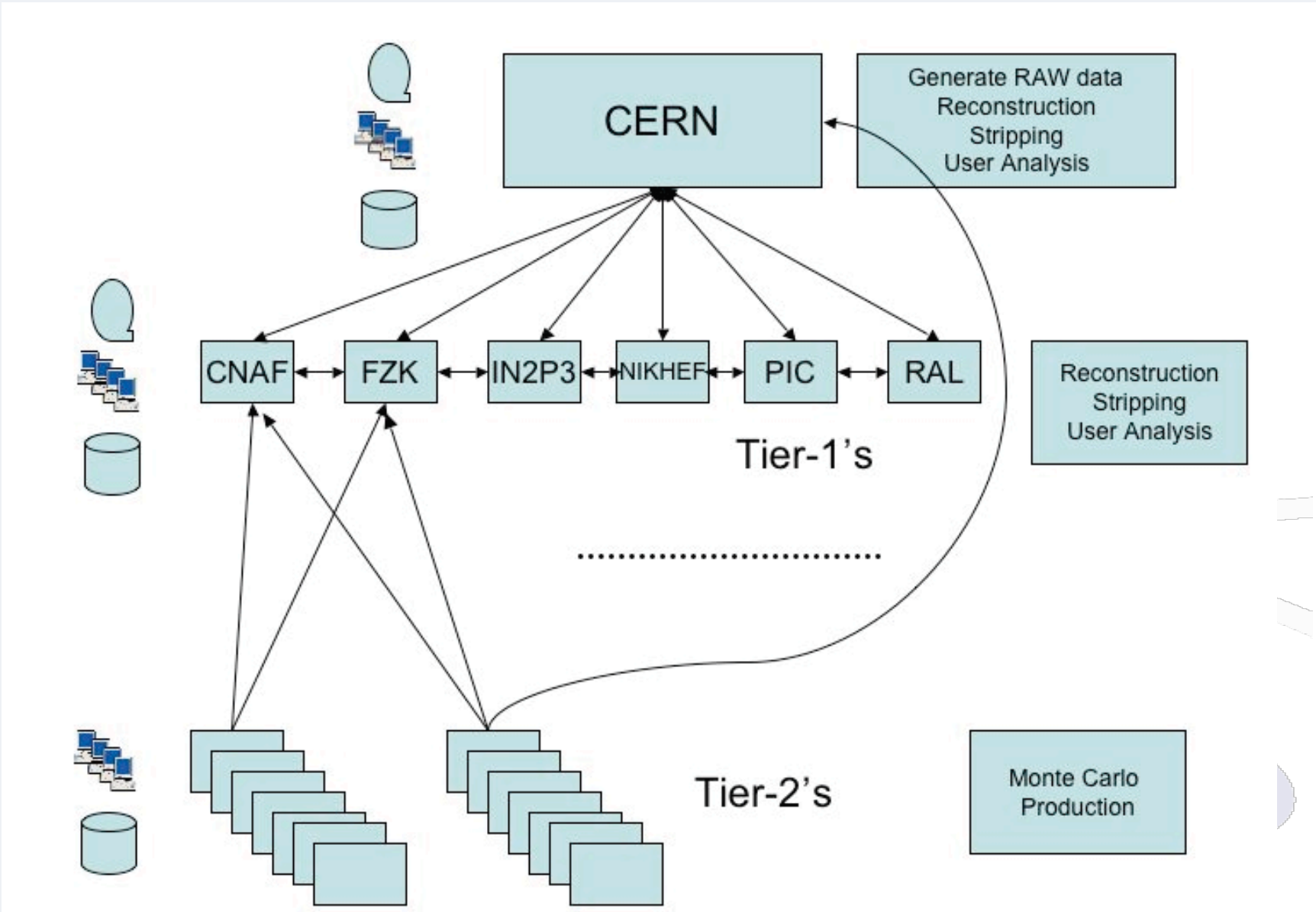
- Expected data size and trigger rates are modest compared to other LHC experiments (TDR)
  - 35 kB RAW event size
  - Trigger rate: 2000 events/s
  - 25 kB RDST (a.k.a. ESD), 85 kB DST (a.k.a. AOD)
  - Typical reconstruction time: 12 HS06.s/event
- Physics research channels are rare
  - b-quark CP violation decay modes (BR ~ 10^{-9} to 10^{-6})
  - Typically a few 10'000s to a million events per year (2 fb^{-1})
    - A needle in a haystack
  - Easier to extract b decay events if only one primary vertex
  - Metrics = average number of visible interactions per beam crossing (μ)
    - For LHC design characteristics μ=0.4 at LHCb
- LHCb is a small experiment
  - Very small Computing Operations Team (< 5FTE)

Guidelines for the Computing Model

- Small processing time, but high trigger rate
  - 24 kHS06 required for reconstruction
    - Typically 2000 CPU slots
  - Tier0 could not provide the necessary CPU power
  - Use Tier1s as well for reconstruction (first pass)
- Most problems for analysis jobs are related to Data Management
  - SE accessibility, scalability, reliability…
  - Restrict the number of sites with data access
  - Use Tier1s for analysis
- High requirements on simulated data
  - Background identification, efficiency estimation for signal
  - Typically 360 HS06.s per event
  - Use all possible non-Tier1 resources for simulation

LHCb Grid Operations

The LHCb Computing Model



LHCb GRID OPERATIONS

LHCb Computing Model

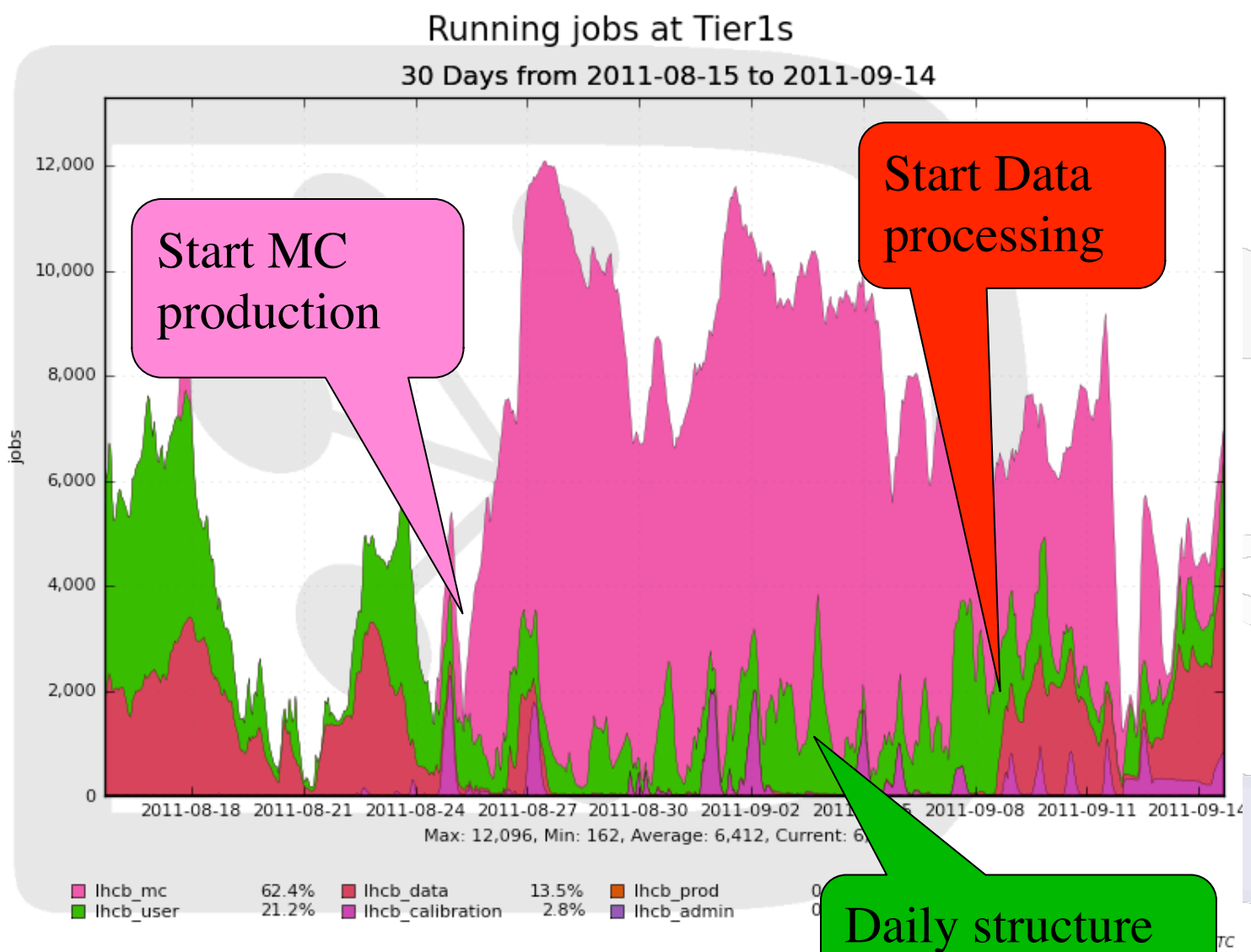
- First pass reconstruction takes place at Tier0 and all Tier1s
- RAW distribution to Tier1s and reconstruction jobs proportional to the CPU pledges 2011
  - 15% reconstruction at CERN
  - Some adjustments made by hand on Tier1 shares
  - Output SDST on T1D0 at local SE
- Stripping (a.k.a. skimming) done at the reconstruction site
  - Output streamed DSTs (10 to 12) uploaded to local T0D1
  - Small files automatically merged into up to 5 GB files (locally)
- DST placement
  - One replica local (from the job)
  - One replica at CERN
  - Two more replicas at Tier1s (only one for MC)
  - Two replicas as Archive on T1D0 (CERN + one Tier1)

LHCb Grid Operations

The LHCb Grid solution

- Based on DIRAC (developed within LHCb)
  - Now used also by many other communities
- LHCb specific extensions (LHCbDirac)
  - E.g. production system, bookkeeping system etc…
- Based on central services run at CERN
  - Redundancy services running at Tier1s:
    - VOBox (collecting requests from jobs)
    - LFC read-only mirrors
    - Oracle Conditions DB mirrors (3D streaming)
- Data Management System
  - File staging, bulk transfers, bulk removal
  - Using LFC, SRM-enables Ses, FTS services at Tier1s
- Workload Management System
  - Using pilot jobs (pioneers!) and central task queues
  - Allows a lot of flexibility on implementing policies

MC filling the gaps



2011 Running Conditions

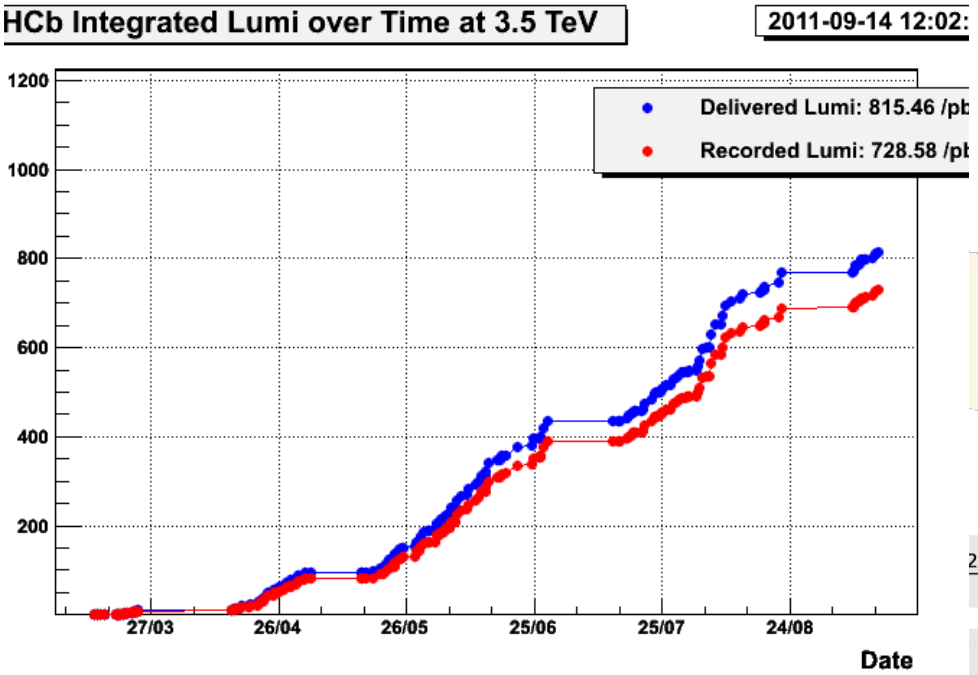
- LHC decided to optimise luminosity by increasing the bunch density
  - Good for ATLAS and CMS, less good for LHCb (large pileup)
- Required rapid adaptability (trigger and CM)
  - Be able to use large pileup events for B physics
  - Introduction of charm physics in the LHCb program (high luminosity)
- Consequences
  - Larger events (50 kB vs 35 kB), higher multiplicity
    - Longer reconstruction time
  - High trigger rate (3 kHz vs 2 kHz)
  - Larger DST (a.k.a. AOD) (130 kB vs 100 kB)
  - Higher stripping retention (15% vs 10%)
- However… we and more importantly you managed to adapt to these new conditions!
  - Thanks!

LHCb GRID OPERATIONS

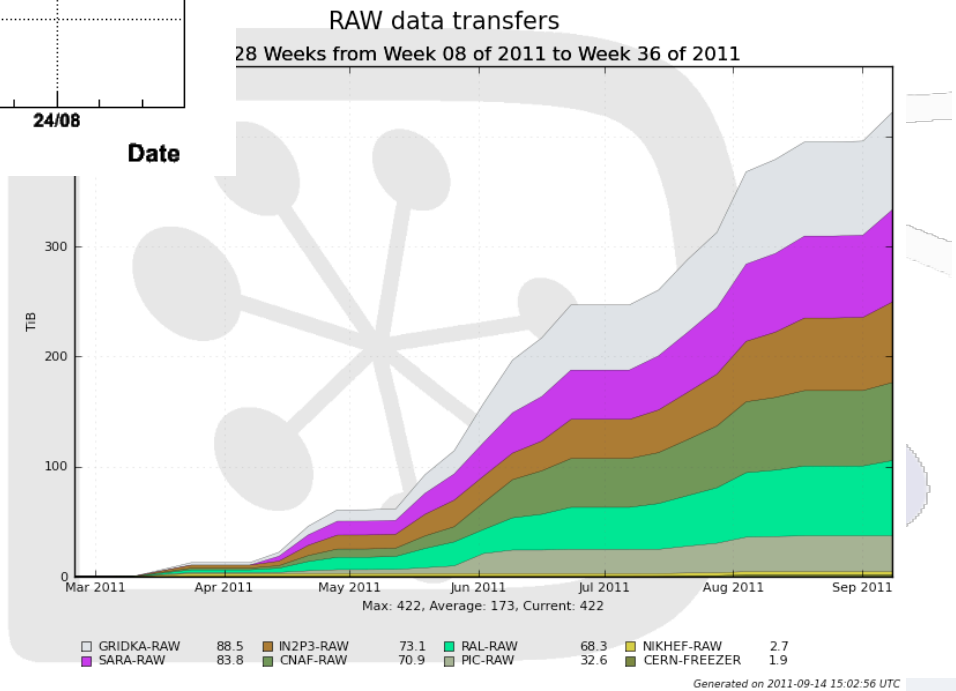
LHCb new requirements for 2011

- We had to change our requirements for 2011 in order to be able to cope with these changes
  - More CPU, more disk and more tape requested
  - Presented to and approved by C-RSG in March 2011
    - … but no changes in pledges from sites
    - Fortunately bilateral agreements were made with most sites
- CPU limitations
  - Not expected during data taking
    - We can even run MC at Tier1s and CERN
  - … but during reprocessing (need 4 times as much CPU power)
    - This is the bottleneck
- Disk limitations
  - Also comes at reprocessing time
    - We must keep previous processing on disk + 2010
  - 0.4 PB missing at CERN, 1.5 PB missing at Tier1s (250 TB per Tier1)

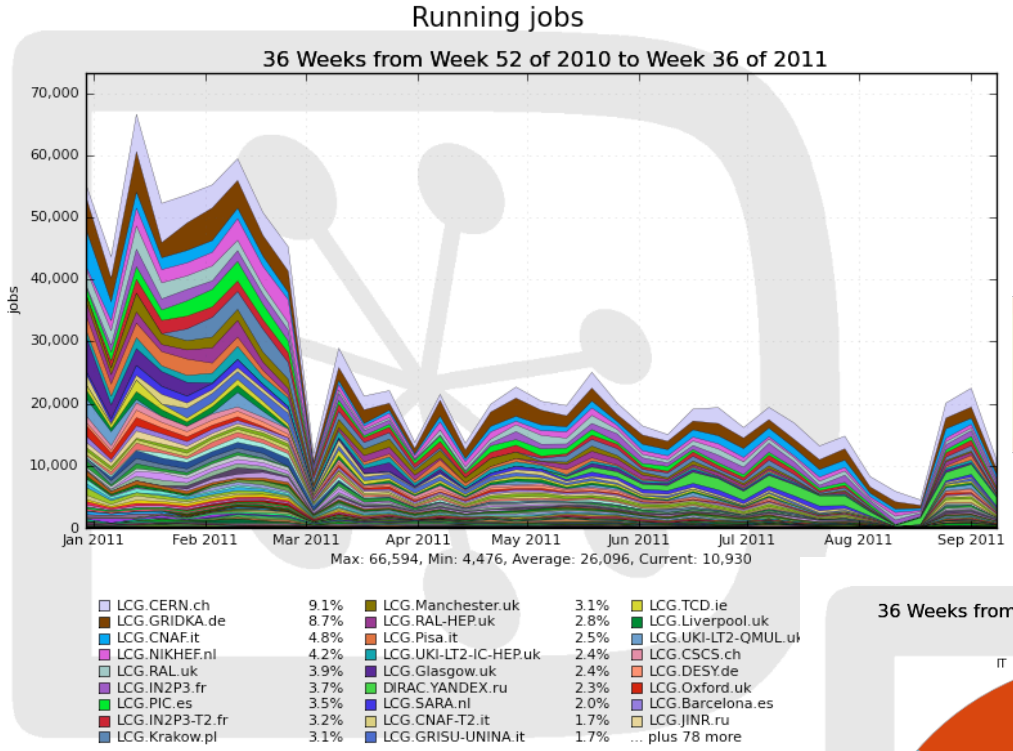
2011 data taking

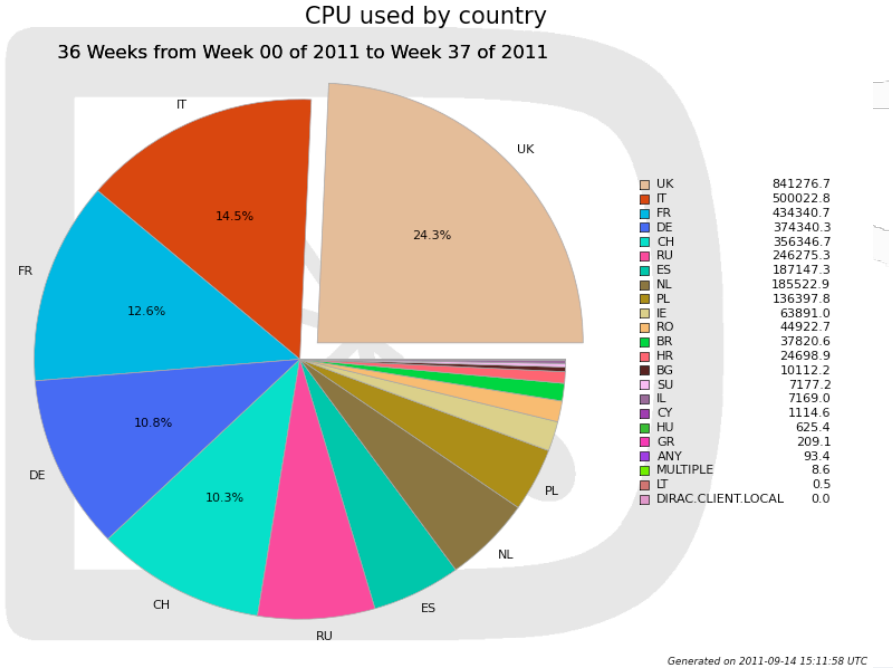


- 90% DAQ efficiency
- 1 fb-1 target in reach
- 420 TB of RAW data



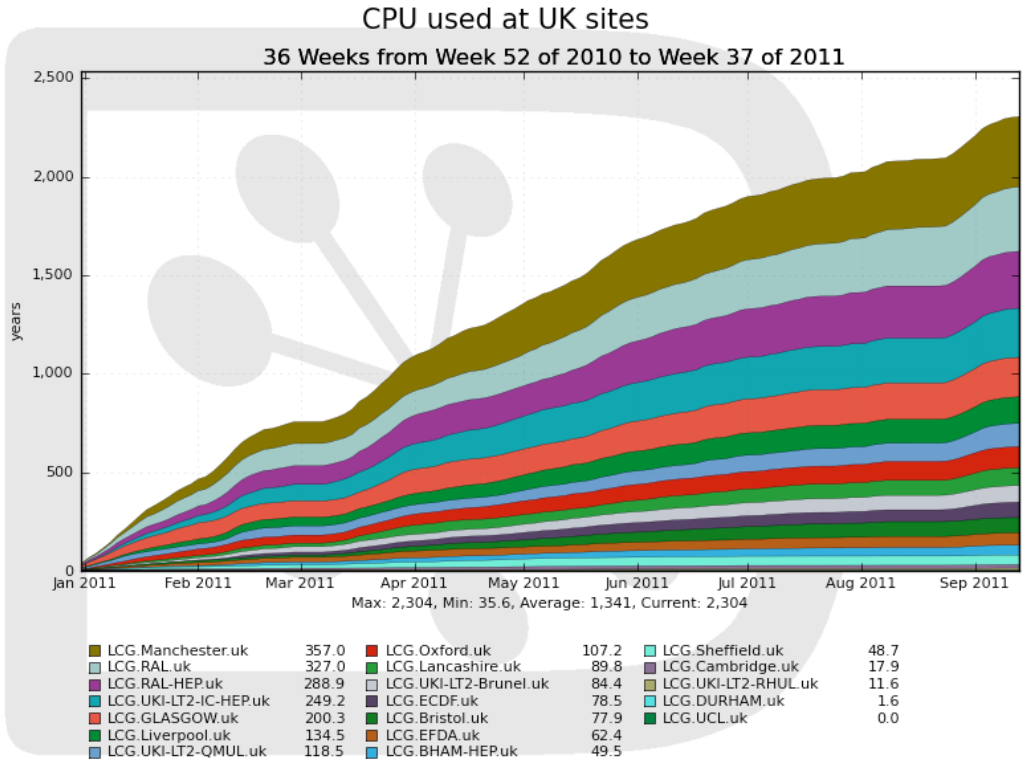
Jobs in 2011



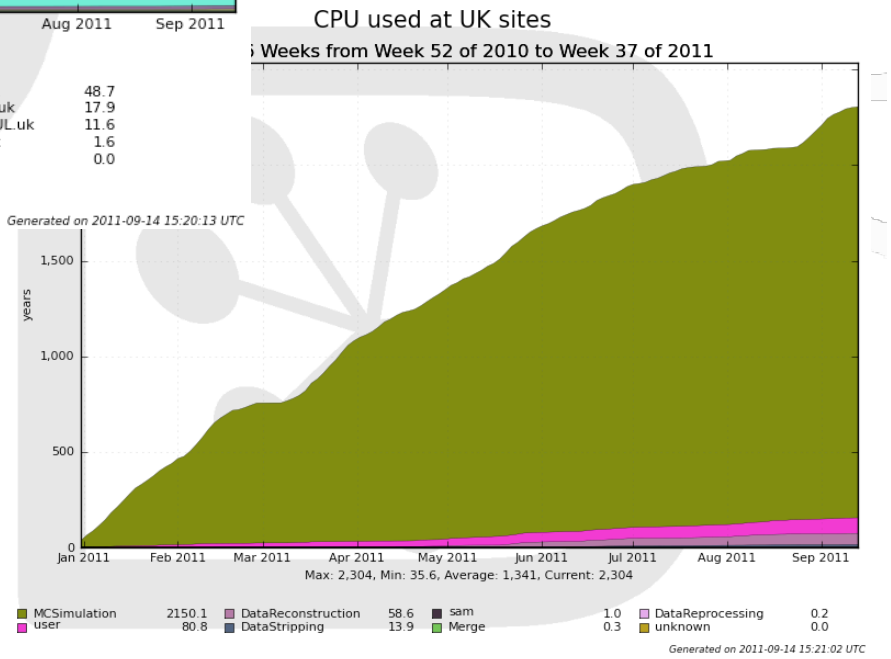


- Over 60,000 jobs simulataneously
- UK provides 24% of CPU

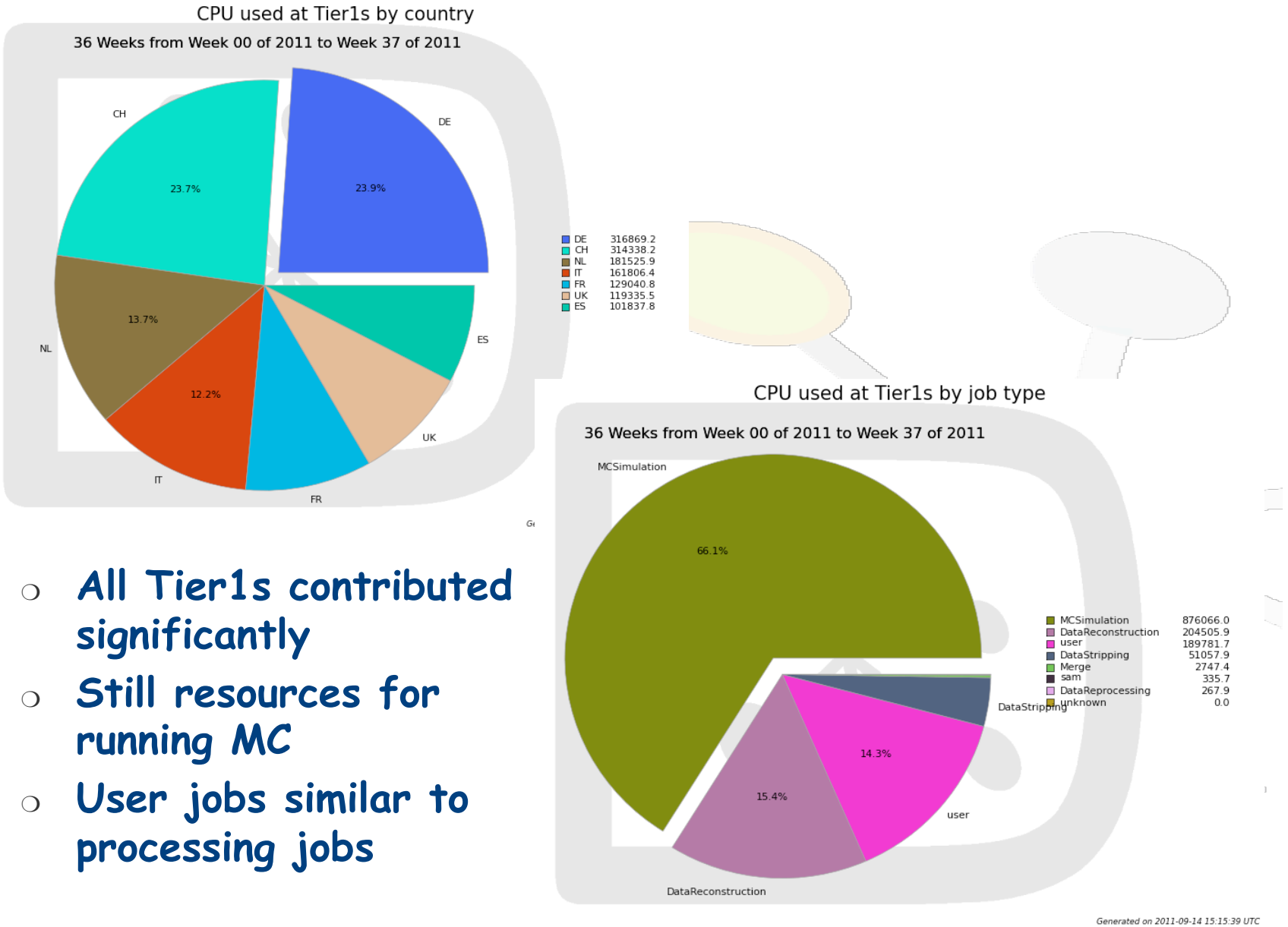
UK jobs



- On average 3000 running jobs
- Dominated by MC jobs
- Manchester > RAL !



Tier1 usage



- All Tier1s contributed significantly
- Still resources for running MC
- User jobs similar to processing jobs

LHCb Grid Operations

2011 re-processing

- Due to higher luminosity, not enough CPU power at Tier1s
- LHCb Reconstruction Centers
  - In addition to CERN and Tier1s, use major Tier2s also for reconstruction
  - RAW data downloaded from a Tier1
    - Not necessarily in the same country
    - Add Tier2s to Tier1s in order to balance the storage usage and RAW availability
  - Once the RAW file is staged at Tier1, all LRCs attached to that storage are eligible to run the job
    - Availability decides (pilots jobs!)
  - SDST uploaded to the same Tier1
  - Stripping/merging running at Tier1
  - LRSs are "coprocessors" to Tier1s
  - Only requires good network connectivity (using general network for downloading directly to WN)
  - Tests successful with more than 20 Tier2s

LHCb Grid Operations

2011 reprocessing planning

- Applications fully ready 23 September
- LHCbDirac adaptations also ready and commissioned by 23 September
- Last validation the week after
- Start of large reprocessing: last days of September

- New feature: file staging throttling
  - Not stage more than what can be processed
  - Avoid jobs starting when the file has gone!
  - Avoid multiple recall from tape
  - Tested and working well
  - Requires to know the size of the cache buffer(s)
- Memory footprint problem fixed!
  - … however still old releases of applications running
  - Special memory extension can be removed after the end of data taking (November)

LHCb

LHCb Grid Operations

Conclusions

- Very fruitful data taking year for LHCb
  - 1 fb^{-1} in reach (unexpected!)
  - Very rich physics program in sight: even access to new physics
  - Already as precise as b-factories and Tevatron!
- Smooth distributed computing over the year, but…
  - … very manpower intensive!
  - Operations team, production managers, site contacts
    - LHCb team is small (less than 20 people in total, many being developers!)
  - LHCb will be short in disk at Tier1s, so… please…
  - UK was the biggest provider of resources! Thanks!
  - RAL was running very reliably as Tier1
  - Looking forward to using partly UK Tier2s for the end-of-year reprocessing (Manchester, Glasgow, Liverpool… ?)