

# **CERN accounting Issues and status**

CERN IT

# New team

- Left
  - March: Miguel Santos -> CERN Data Protection Office
  - July: Sergio Llana Perez
- Arrived
  - Gavin McCance
  - October: Florentia Protopsalti - though short overlap with Sergio on previous project
  - October: Gorka Bravo

# Two main issues in last six months

- Both “completeness” problems in the data
  - LSF data counted twice at one point
  - Cloud compute CPU data missing

# LSF issue

- Earlier this year, LSF “main instance” data was counted twice (records sent to accounting pipeline from the primary master *and* the backup master)
  - De-duplication might have helped
  - Affected everyone, though noticed by ATLAS
  - Status: corrected and republished

# Cloud issue

- Problem in Openstack Nova schema in identifying accounting group of project, from May this year
  - CPU time query was returned for a subset of projects only
  - Published wall-time unaffected (we use quota)
  - Notably affected CMS (with very large T0 cloud project).  
Tiny cputime / correct walltime —> very low efficiency
  - Status: corrected for RRB, not yet republished
    - Similar issue found and corrected for ATLAS

# Main lessons

- Both issues caused by underlying bugs causing a query to return incomplete (or over-complete) results
  - Need better sanity check at various levels: input data and final product
  - Better checking of final numbers compared to pledge
- Quite tricky, since some usage (notably cloud) is extremely choppy and variable, even during data taking
- Can we (semi-)automatically check against experiment data? Other suggestions?

# Reminder

- For fixed *cloud compute* resources, CERN charges the fixed quota that's available to you even if you don't make VMs in the quota
  - (..and efficiencies are calculated on that)