# WLCG Accounting Task Force Update

## Julia Andreeva CERN

## WLCG Management Board 21.06.2016

# Main objectives

- To have by the end of August :
  - EGI accounting portal with a dedicated WLCG view which contains all necessary information in a table and/or graph form available also through APIs. Should be bug-free and user-friendly
  - Data exposed through the portal should be correct. We should have in place an automatic data validation procedure, which would allow us to detect eventual problems with data publishing or/and processing
  - The accounting reports with an agreed content should be generated by the EGI portal. This should be possible to do on demand for any given time range with monthly granularity.

# Task Force organization

- Be-weekly meetings. First meeting took place on the 9$^{th}$ of June

  *https://indico.cern.ch/event/539179/*

- Mailing list : wlcg-accounting@cern.ch

- Twiki page :

  https://twiki.cern.ch/twiki/bin/view/LCG/AccountingTaskForce

# Task Force activities (1)

- Validation of the data provided through the portal via comparison with alternative information sources. Debugging problematic cases. Automatic publishing of the comparison results to the SSB. (Ongoing)

- Testing and validation of the new portal. The latest prototype is available for validation:

  https://accounting-next.egi.eu

(Ongoing)

# Task Force activities (2)

- Set up a sub-group (coordinated by Alessandra Forti and Miguel Coelho dos Santos) which will investigate and document all data transformations happening at all levels (site, APEL, experiment-specific monitoring systems)

raw wall clock (time) ->scaled wall clock (time) -> HS06 normalized wall clock (work)

Conclude whether procedures are correct/accurate enough, derive recommendations regarding normalization and benchmarking.

# First results of data validation (1)

| Experiment | Performed comparison scaled/raw wall clock multiplied by number of cores | Agreement EGI/Experiment | Performed comparison normalized (HS06) wall clock multiplied by number of cores | Agreement EGI/Experiment | Comment |
|---|---|---|---|---|---|
| ATLAS | YES | ~105% | YES | 95% | Using ATLAS Dashboard API |
| ALICE | YES | ~104% | NO | - | Normalized metrics are not available |
| CMS | YES | ~132% | YES | ~132% | Using CMS Dashboard API |
| LHCb | NOT | - | NOT | - | Waiting for API or data extraction from DIRAC |

# First results of data validation (2)

- We do not compare exactly the same thing.
  - Experiment systems measure real payloads, APEL measures pilots, therefore EGI should have slightly higher numbers
  - We do not always have raw wall clock in APEL, since some batch system rather report scaled wall clock. However in this case normalized metrics should still agree (RAL, Prague)

- Agreement we see so far is pretty good and we can conclude that overall accounting data is reliable

- However we detected quite some problems with certain sites which have to be followed up and fixed

# First results of data validation (3)

- Cases to be investigated:
  - EGI usage is much lower (more than 30-40%) both in terms of raw and normalized metrics compared to the experiment systems
  - EGI usage is several times higher both in terms of raw and normalized metrics compared to the experiment systems
  - Big discrepancy for normalized metric, while raw metric does agree

- Most striking discrepancies for ATLAS are already understood

- Enabling automatic publishing of the comparison results to SSB

# First lessons learned

- In order to be able to crosscheck, understand and debug accounting data, we need raw wall clock measurement to be included in the accounting record. This is also important for transparent integration of the accounting of the commercial clouds and HPC.

- Global WLCG topology and configuration system is essential for providing topology description for the EGI accounting