Accounting Review Summary from the pre-GDB related to CPU (wallclock) accounting

Julia Andreeva CERN-IT

GDB

13th April 2016

The goals of the accounting review

- We have quite some issues and open questions related to the accounting

- Effort is dispersed across various task forces and working groups (IS task force, benchmarking WG, cloud procurement work, WLCG resource reporting task force, etc…)

- Need to summarize issues we have, define an action plan and propose how we can better organize this work

Topics which were discussed

- CPU (wallclock) accounting
  - Multicore accounting
  - Accounting of the resources provided through cloud interfaces
  - Accounting of the opportunistic resources
  - Review of the EGI accounting portal and accounting reports generated by REBUS
- Storage space accounting. In particular what are our plans for no-srm storage

Various kinds of CPU accounting we deal with

- User-level accounting. Is covered by the experiment specific systems. The accounting instance in this case is a payload job, not a pilot, therefore the pilot overhead is not taken into account. Covered by Experiment-specific accounting systems. Used for operations.

- Batch system accounting. Is covered by APEL. The accounting instance in this case is a pilot job. Used by experiments for CRCB, by sites, funding agencies and WLCG management.

- Though having different information sources and operating with different accounting instances both systems should roughly agree in terms of CPU and wallclock time unless there is a problem with pilots

- **Experiments deal on their own with user level accounting and are doing well. Accounting review did not cover this area.**

Multicore accounting (1)

- Multicore accounting needs N_cores to be provided as part of the information reported per job to classify jobs, scale walltime use and calculate efficiencies

    (efficiency: CPU/(Wall*N_cores)

- LHC VOs currently running multicore are CMS and ATLAS, LHCb in development, ALICE not using them (yet).

- For them, sites reporting to the accounting portals is basically solved:

    - T1s are OK
    - T2s are OK, except a few CEs

Multicore accounting (2)

- Multicore accounting is enabled in the current accounting portal in a dedicated WLCG view. The new portal should also provide multicore functionality

- Validation of information presented in the portal is needed. It can be crosschecked with data in the CMS and ATLAS Dashboards which provide user level accounting with multicore functionality enabled

Accounting of the resources provided through cloud interfaces

- When we deal with resources provided with cloud interfaces and in particular with commercial clouds, there is one more type of accounting which is introduced in addition to user-level accounting and batch system accounting – resource provider accounting. In this case we normally deal with a VM as an accounting instance

- User-level accounting not necessary agrees with resource provider accounting. But when they do not agree you might need to investigate why, since it might point to an efficient use of the resource you pay for. So the possibility to crosscheck between user-level accounting and resource provider accounting is a must

- When pledged resources are provided through cloud interfaces, we still need to account them in APEL to demonstrate how/whether MoU commitments are fulfilled

- On the other hand LHC experiments want to account also opportunistic resources.

Accounting of the resources provided through cloud interfaces (private clouds)

- Work has been started by the WLCG Resource Reporting Task Force. The work was inline with APEL cloud accounting work. By the end of the last year the publishers for OpenStack and OpenNebula as well as normalization method became available and cloud accounting data showed up in the EGI accounting portal. The task force was terminated.
- This effort is continued by CERN IT-CM group for CERN cloud:
  - Adopting cASO for producing Cloud reports of CERN Openstack
  - Data collected from Nova
  - Patch to support Keystone v3 (for upstream)

Accounting of the resources provided through cloud interfaces (commercial clouds)

- For the last year CERN cloud procurement activity monitoring and accounting system based on ganglia was put in place
- New project for using HTCondor as the Batch System on fixed resources in commercial Cloud(s)
- The resource provider cloud usage reports used for billing provide accounting for various cloud providers (AWS, Google, Azure, Rackspace). They are being evaluated by the CERN IT-CM group
- All this is work in progress

Review of the accounting portal and accounting reports and REBUS (1)

- WLCG Operations team asked for input from the experiments, sites and WLCG project office

- This input was presented at the WLCG operations coordination meeting last Thursday

- Short summary will be presented further in this talk, for more details see:

  http://indico.cern.ch/event/514077/

Some highlights regarding accounting portal

- Experiments, sites, WLCG project office and members of the scrutiny group expressed their concern about the current state of the portal in terms of simplicity of navigation and finding necessary information as well as in terms of data trustworthiness.

- Implementation of new features and bug fixes takes too long, terms of many months

- In particular it concerns multi-core accounting which looks still buggy and not yet deployed in production

- Better coordination with the developers of the accounting portal is required

Some highlights regarding REBUS

- REBUS is used for pledges, installed (available) capacities and generation of the accounting reports
- Accounting reports are constructed taking data from REBUS itself and accounting data from the accounting portal
- A historical view within one year would be very useful for accounting purposes
- A notification when changes are performed would be a useful functionality
- Would be good to have pledge value per site not per federation (not a technical issue)
- EGI Accounting portal is used for REBUS T2 reports, but historic data is not available for decommissioned sites
- The functionality of REBUS might be reviewed if the new WLCG configuration system is accepted. REBUS can be turned into a service for report generation taking data via APIs from various external sources

Conclusions and action list

Multicore accounting

- Seems to be in good shape

- Proper reporting of the number of cores is enabled at all T1 and most of T2. Corner cases to be followed up

- Main effort required for validation of the multicore information in the accounting portal

- Information can be crosschecked with CMS and ATLAS Dashboard where multicore accounting is available

Accounting of the resources provided through cloud interfaces (private clouds)

- The good start was the work done by the WLCG Resource Reporting Task Force. However, it looks like there is still a lot of effort required before we have production solutions for instrumenting our cloud resources for proper reporting

- With OpenStack we progress well thanks to work of the CERN IT-CM group. What about other platforms for cloud computing?

- Need to understand whether other WLCG sites apart from CERN work on it. If yes, need to streamline this effort. And collaborate with EGI.

Accounting of the cloud resources provided by commercial clouds

- Work on accounting of the commercial clouds is in progress and is driven by the CERN IT CM group and in the scope of CERN cloud procurement activity.

- Currently no actions are required.

Accounting of the opportunistic resources

- Some of the experiments believe that it would be nice to have a capability to expose usage of the opportunistic resources in the accounting portal

- These resources are already accounted in the experiment-specific systems which can expose accounting information though APIs

- The proposal is to evaluate a possibility to inject information of usage of the opportunistic resources in APEL, the same way it is done for OSG accounting info which is injected from GRATIA. Is there an issue with the topology? Cloud resource might not be described in GOCDB

Review of the accounting portal and accounting reports

- **Agreed** that though available capacity is required for operations, it does not need to be a part of the accounting reports
- **Agreed** that manual fixes of the accounting numbers for the accounting reports should be avoided and if accounting data is not properly reported to APEL the problem should be fixed there instead
- **Agreed** that WLCG needs a dedicated simplified instance of the accounting UI, which provides customized menus and contains ONLY metrics required by WLCG
- **Agreed** that we use GGUS for reporting bugs and submitting requests to APEL and accounting portal developers
- **Agreed** that better coordination between APEL/accounting portal developers and WLCG represented by the WLCG operations team is required, in particular in what concerns validation of the accounting portal (in terms of functionality) and checking of the trustworthiness of the provided information.

Some more proposed and agreed changes on the accounting reports and accounting portal

- Use a consistent metric with the one in which pledges are expressed wallclock HS06 instead of CPU HS06.

- Use the same unit HEPSPEC06 days both for T1 and T2 reports.

- Add a simple metric 'Elapsed time multiplied by number of cores' in order to easily crosscheck information with the experiment specific accounting systems

Proposal for setting up the CPU accounting Task Force

The goals are :

- **to agree with all stakeholders on the content of the accounting portal and accounting reports (metrics, units, aggregation in terms of time and set of attributes like site, VO, country, etc…);**

- **coordinate with APEL and accounting portal development team and follow up on the progress**

- **validate the new version of the portal, ideally a dedicated customized UI for WLCG, and make sure that the requirements of the stakeholders are addressed**

- **validate accounting data, ensure that it is correct in the APEL repository so that manual fixes could be avoided**

- **follow up on whether/ how we enable injection of the opportunistic resources into APEL and accounting portal**

- **In case CRIC is accepted as a WLCG central configuration system, re-evaluate role and functionality of REBUS which can be turned into a simpler service responsible for pledges and report generation**

- **assess the situation with accounting of the pledged resources provided through private clouds , understand whether we have all components in place to properly account such resources , make sure that they are correctly publishing to APEL. In case this requires substantial development, testing and evaluation effort , this can be a question of launching its own dedicated task force**

Thank you for all participants of yesterday and today meetings for fruitful discussion and valuable contribution