HTCondor-CE accounting with APEL

Ste Jones
(Liverpool University, GridPP)
EGI, Amsterdam, May 2019

Background

- Liverpool T2 uses HTCondor-CE (since January.)
- It worked well, but it had no APEL parser to provide accounting.
- We tried HTCondor-CE PIC changes to APEL client parser (single input f le). Functionally perfect. But not completely compatible.
- PIC changes conf lcted with CREAM-CE + HTCondor, since it overrides existing HTCondor parser with a new version, and a new f le format.

Goals

- We extended APEL parsers to support HTCondor-CE in a way that is compatible with everything else, i.e. no conf ġ change required for non-HTCondor-CE sites who update APEL client.
- We directly support HTCondor-CE + HTCondor batch system, but architecture “supports” other backends, i.e. option to extend further for PBS, LSF, SGE, SLURM (and perhaps others in due course.)
- To do this, we preserved (to the largest extent) the existing data f ƀw/f le format conventions, giving minimal code changes (BTW: APEL parser is already well designed in this respect.)
- It is to be released with UMD as a standard way to link APEL with HTCondor-CE.

Design/implementation

- We retained standard use of two input f les; one from CE (blah log), another from batch system (event log).
- We developed data extraction scripts that use HTCondor's powerful (printf-like) formatting language to produce these data f les. Hence alternative batch systems are adopted by writing one new data extraction script for the new event log format (just reuse the blah script.)
- A small code change needed (~ 4 lines) for a new, optional scaling factor f ėld in existing HTCondor APEL parser. The change is transparent to existing CREAM-CE/HTCondor sites. Version 1.8.0-1 of the APEL client parser software contains this change.
- To provide support for heterogeneous clusters, a scheme is used to obtain node scaling factor via some ClassAd. An example scheme is given in the documentation.

Remaining (vaguely related) issue

- This is a new general requirement for all sites using APEL client.
- Until now, APEL client obtains CE benchmark reference via BDII (Glue 1) and puts it in the accounting records to be sent.
- Problems: HTCondor-CE only gives Glue 2; and in any case the BDII is “soon” going away, it is said. Although the dates for this are not set.
- Solution: ~ 20 line code change to allow admin to hard conf ġure the scaling benchmark for CE in the APEL client. No query to BDII. Change still awaiting acceptance test/release.
- Note: To “get the show on the road” a workaround is used for the time being - a one-off “static data” SQL insert done by sysadmin/build system. See documentation.

Documentation

End user documentation

https://twiki.cern.ch/twiki/bin/view/LCG/HtCondorCeAccounting#Technical_setup

Scaling factor scheme

https://twiki.cern.ch/twiki/bin/view/LCG/HtCondorCeAccounting#Implement_scaling_factor

Test

https://twiki.cern.ch/twiki/bin/view/LCG/HtCondorCeAccounting#Tests_on_a_HTCondor_CE

Design notes

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

Performance in latest test (all Feb)

On the CE:

 # cd /var/lib/condor/accounting

 # ls -rt batch-201902* > /tmp/t

 # for f in `cat /tmp/t`; do ~/scripts/workDone.pl < $f; done | perl -ape'$s+=$F[0]}{$_=$s' ; echo

3953875.81 # WHAT WE CLAIM WE SENT FOR FEB.

(workDone.pl audit script listed below)

Performance in latest test (all Feb)

On the EGI Accounting Portal:

Research Infrastructure/T2/UK/NORTHGRID/Feb 2019 Feb 2019/Sum Wallclock Work/Row Var Submit Host

hepgrid6.ph.liv.ac.uk:9619…

- 3,953,212 # WHAT APEL CLAIMS IT GOT FOR FEB.

Accuracy comparison: ~ 0.017%

Ongoing work

- **Additional packaging work is ongoing with HTCondor-CE team to implement a couple of "nice-to-haves".**
  - Make a RPM of the necessary interface scripts, within the HTCondor-CE release.
  - The existing HTCondor-CE puppet module may also be enhanced to install / configure the RPMs.
- **But system is already functionally complete and ready to use without this work. See links above to relevant set up documents.**