

Job Efficiency Working Group meeting 12-12-2014

Present: Costin Grigoras PS-AIP, Christian Nieke IT-DSS, Luca Menichetti IT-DI, Alessandro Di Girolamo IT-DSC, Stefan Stancu IT-CS, Barnd Panzer-Steindel IT-DI, Arne Wiebalck IT-OIS, Jerome Belleman IT-PES, Ulrich Schwickerath IT-PES, Valentina Mancinelli IT-SDC, Domenico Giordano IT-SDC, Eddie Karavakis IT-SDC, Maarten Litmaath IT-SDC, Daniele Spiga IT-SDC and Manuel Guijarro (chair and notes)

1 Agenda

Items were not processed as listed in the agenda. The material provided by each contributor is available at <https://indico.cern.ch/event/356918/>

Jerome Belleman: Using a benchmark program provided by Costin which simply allocates 2GB of RAM, the batch service managers have identified that, in some worker nodes, job slots could not use 2GB of RAM without swapping. This has triggered a reconfiguration of the number of job slots in those WNs which now ensures no need for swapping if the job uses less than 2GB of RAM.

Dirk suggested that it would be good to have a look to the amount of swapping which is happening in Batch.

Bernd showed some plots of job efficiency for the last 6 months. Those plots show around 75% efficiency for all experiments (see material in indico page).

Costin: The benchmark provided by Alice is a reproduction of what was seen in real Alice jobs. Many machines needed 3G, 2G of RAM and very few needed just below 2GB.

Maarten: has the impression that things have improved but yet it is important for Alice to re-check again if it is ok.

Costin: Some machines seem to spend more time in IO wait than others. Helge indicates that could be linked to the fact that new hardware is faster.

Alessandro: Since April there are no so many problems (before problems were real: network hiccups, XROOT not using TTreeCache option, phantom pilot jobs, etc).

Dirk: It is difficult for the experiments to know how many events they process per HEPSPEC.

Costin: No matter what, we should further investigate how to reduce that 10% I/O wait, which Alice sees even in machines that do not swap. The workload at CERN is similar to that of any other T1 where that 10% of IO wait is not seen. **Bernd** disagrees about that fact as he thinks the workload at CERN is different to the workload in other T1s.

Bernd: Memory is not the only culprit for the lose on efficiency. There are other issues, such as phantom pilots jobs occupying batch job slots that had been seen in the past (Atlas). We should make sure we do not have Alice phantom pilots jobs.

Manuel: We should be careful about reserving lots of resources to jobs slots to ensure very good efficiency as that might result on the batch system not optimally using all available slots and a greater waste of resources for I/O intensive jobs.

Costin: Claims this misbehavior (~10 I/O wait) can be seen (live) almost any time. **The cloud team will arrange for a session with Costin to have a look to those jobs and status of the hypervisor.** Bernd suggest to do the same type of exercise for the other experiments. Unfortunately they can not currently find live examples of low efficiency jobs.

Christian: Nice summary of the presentation on how to Measure Job Performance provided in the AWG meeting. Bottom line is that the currently used LSF CPU factors based on HEPSPECs for WNs are OK. In other words, our "meter" is not a problem. Christian also encourages people to use normalized Wall. The normalization of cpu time is $\text{cpu} * \text{cpufactor}$ (factor could be hepspec). But for the normalization of wall time, it should go like this:

Wall = cpu + io , or rather, compute io as $\text{io} = (\text{wall} - \text{cpu})$

Then $\text{normalized_wall} = \text{cpu} * \text{cpufactor} + \text{io}$. In Christian's experiments, normalizing also the io part could give even worse results than doing no normalization. The best way to check the CPU factor is using the machines features, which are meant to be available in all Grid sites. Please do not query LSF to get that information.

Stefan provided an update on the status of the PerfSonar dedicated instance in Wigner and how it is setup using CentOS6 VMs with a special network configuration which makes those VMs similar to having physical machines. They are not using OIS managed Hypervisors. PerfSonar shows asymmetric src-dst dst-src **transfer bandwidth**. This is most likely due to asymmetric packet loss. While investigating a DB replication transfer, it was discovered that even a low rate loss (smaller than 0.01%) can severely impact throughput to/from Wigner, due to the large RTT (25ms). **Congestion algorithm:** It can not really be changed (using the default one in linux) since it might affect many people. **Latency:** src-dst dst-src again not symmetric, but $\text{src-dst_ipv4} \approx \text{dst-src_ipv6}$ and $\text{src-dst_ipv6} \approx \text{dst-src_ipv4}$. Stefan guesses this is due to the hashing of the probe packets on either of the two 100G Wigner links. **The AWG will take care of ensuring all PerfSonar monitoring data makes its way to the central Hadoop.**

Manuel went quickly through the kibana plots on Batch efficiency:

http://pesmon-master00:9200/_plugin/kibana/index.html#/dashboard/elasticsearch/Batch%20Efficiency

The batch team will investigate why those Kibana plots do not show the 75% efficiency showed in Bernd's plots