

T. Hartmann, A. Forti, G. Roy, J. Belleman, D. Traynor, Carles, A. McCrea, A. Perez Calero Yzquierdo A. Filipcic, A. Lahiff, D. Crooks, A. Sedov, Andrea, Stefan Roiser, J. Hernandez Calama, S. Skipsey J. Templon, A. McNab, C. Walker, A. DiGirolamo, C. Wissing, Rod Walker, Manfred Alef

Comments after the CMS presentation

Simone question: this mechanism works better if you have the length of the pilot much larger than the length of the executable. You waste time at the end so the longest is the pilot the least you loose. Based on CMS jobs what is the optimal or the minimal life time for the pilot job. Antonio: We need to tune it. If the last job has to be killed the relative loss is small. Taking into account what the sites want.

Chris: slide 29 once you got the 8 slots... it means sites have already done a significant chunk of the scheduling. multi-VO support is not compatible with long pilots aside of the draining issue to apply changes. Need exchanges with sites. Long pilots from CMS would interfere with atlas workload. In an atlas site giving spare cycle to CMS like QMUL this can result in Atlas not getting the resources when needed. Rod: that's easy you can confine CMS in a shorter queue. But that means the CMS model still needs work on the batch system to add extra queues.

Jeff: One of the things that make things easier in scheduling is entrophy, reducing entrophy makes scheduling more difficult. To do that requiring longer jobs or more resources makes entrophy more difficult. Confusing predictability with ability to schedule. You need also predictability. Predictability doesn't matter at all for single core. Predictability might help with multicore but reduced entrophy hits at all level. If you have alrge entrophy you can fill the gaps but to fill the gaps you need to know how long the job will be.

If there's peaks and valleys that's site going to waste and that's thousands of euros wasted.

Your efficiency of the CMS model depends on filling the pilots and the ability to guarantuee pilots are full. It doesn't help with other VOs, high predictability would force system administator to allocate resources to avoid interference with other users different patterns. Even if the experiments are free to fill the pilot as they want having pilots all of the same length doesn't help the batch system scheduler. High entrophy helps the batch system scheduler. If the pilots cannot be guarantued to be full the pilot could kill itself if after sometime it doesn't receive any workload. However this would result in a degraded predictability. Short MC jobs can be used to mop up the waste space inside the pilot.

Time is guarantued by the batch system and queries the machine-job features are not necessary, a job knows the time it has left.

Antonio: Number of cores should be tested by application people. They should tell what is the best for the application. Simone: 8 is a magic number in atlas it wasn't chosen randomly but it was a compromise between reducing the memory consuption (too few cores) and avoid the serial component taking over (too many).

Jeff can you mix different streams multicore and single core? It would be useful if experiments could turn the know and increase one or the other according to necessity.

What is ATLAS pilot lifetime in general? do they have predictable length? No, analysis jobs are typically short mostly below 4h, production is a mixture depending on the application.

Discussion on advantages of predictability vs high entrophy. High entrophy helps the scheduler filling the gaps in any case,high predictability helps with multicore but doesn't work well with a mixture of single core. In general it is recognised that high entrophy is preferable as it works in any case.

-- AntonioPerezCalero - 03 Feb 2014

This topic: LCG > Minutes20140204

Topic revision: r2 - 2014-02-20 - AlessandraForti



Copyright &© 2008-2021 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.
or Ideas, requests, problems regarding TWiki? use [Discourse](#) or [Send feedback](#)