Pushing HTCondor and glideinWMS to 200K+ Jobs in a Global Pool for CMS before Run 2

J. Balcas[1], S. Belforte[2], B. Bockelman[3], O. Gutsche[4], F. Khan[5], K. Larson[4], J. Letts[6],
M. Mascheroni[7], D. Mason[4], A. McCrea[6], M. Saiz-Santos[6], I. Sfiligoi[6]

[1]Vilnius Univ. (LT), [2]INFN-Trieste (IT), [3]Univ. of Nebraska-Lincoln (US), [4]FNAL (US), [5]NPC (PK), [6]UCSD (US), [7]INFN-Milano (IT)

Abstract

The CMS experiment at the LHC relies on HTCondor and glideinWMS as its primary batch and pilot-based Grid provisioning system. So far we have been running several independent resource pools, but we are working on unifying them all to reduce the operational load and more effectively share resources between various activities in CMS. The major challenge of this unification activity is scale. The combined pool size is expected to reach 200K job slots, which is significantly bigger than any other multi-user HTCondor based system currently in production. To get there we have studied scaling limitations in our existing pools, the biggest of which tops out at about 70K slots, providing valuable feedback to the development communities, who have responded by delivering improvements which have helped us reach higher and higher scales with more stability. We have also worked on improving the organization and support model for this critical service during Run 2 of the LHC. This contribution will present the results of the scale testing and experiences from the first months of running the Global Pool.

The CMS Global Pool

GlideinWMS, which is based on HTCondor, is the main resource provisioning system in CMS. The main components are a central manager, various submission nodes which hold the batch queues, and execute nodes which run on various Grid resources. These execute nodes are submitted by glideinWMS factories upon request by a CMS frontend.

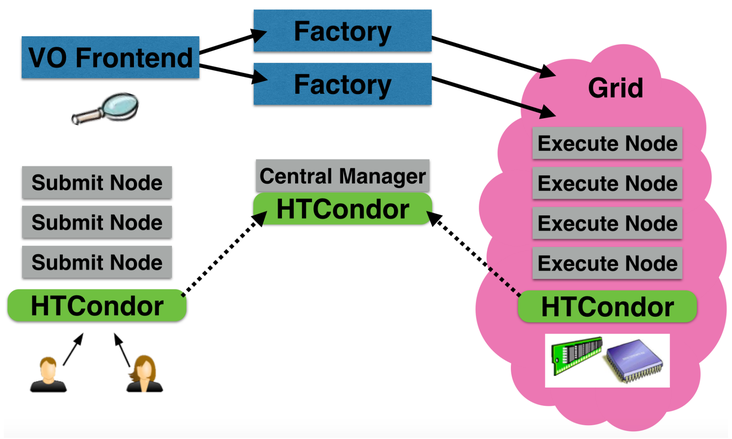

Figure 1. The basic components of glideinWMS for Grid operations.

For the past several years we have operated independent HTCondor pools, one for analysis and another for central data processing. The initial motivation for unifying the varyous pools in CMS into a single "Global Pool" was to be able to rapidly prioritize between different kinds of workflows, e.g. high vs. low priority Monte Carlo production, or to boost reprocessing or a high-stakes analysis[1].

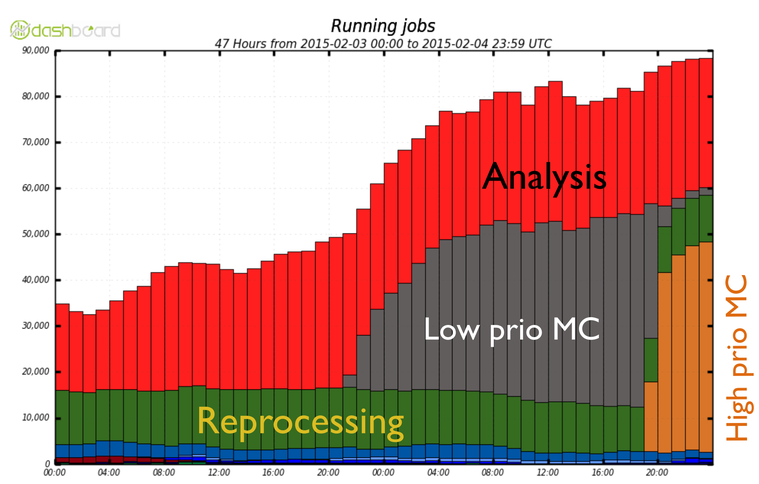

Figure 2. Demonstration of a high priority workflow rapidly taking over the Grid resources in the Global Pool.

Additional motivations were to reduce operational load by having a unified infrastructure and the ability to bring new and different types of resources into a Global Pool. The main challenge, however, is that a glideinWMS or HTCondor pool on the scale of the resources available to CMS had never been attempted before.

The Scale Challenge

Currently the WLCG resources pledged to CMS by the sites accessible to the Global Pool is about 108,000 batch slots. Using the Global Pool infrastructure, however, we can discover over time the totality of the resources available to the pool, which we currently estimate to be about 200,000 batch cores.

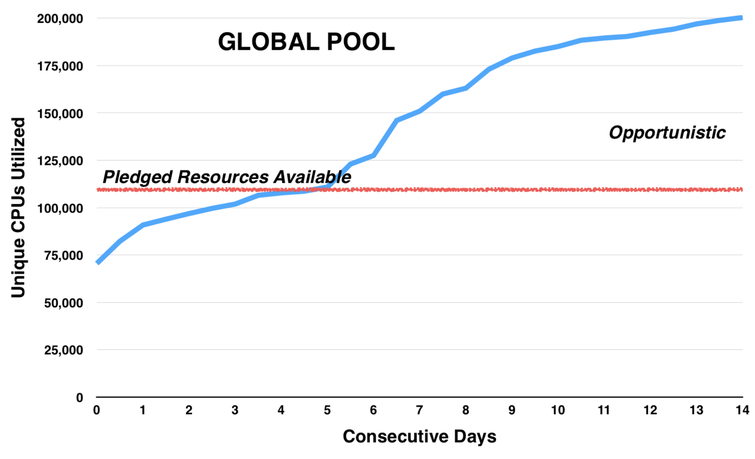

Figure 3. Batch cores discovered by the Global Pool over two weeks.

On some of these resources CMS must compete for access, but this estimate gives an idea of the scale necessary to reach during Run 2, also taking into account that the resources requested (and CMS needs) will grow from year to year.

Tier-0, HLT, etc.

CMS has further unified the resource provisioning system by including new types and combinations of facilities and workflows that we did not have during Run 1, such as using the HLT (High Level Trigger) farm during LHC inter-fills, and running the Tier-0 as part of the glideinWMS system.

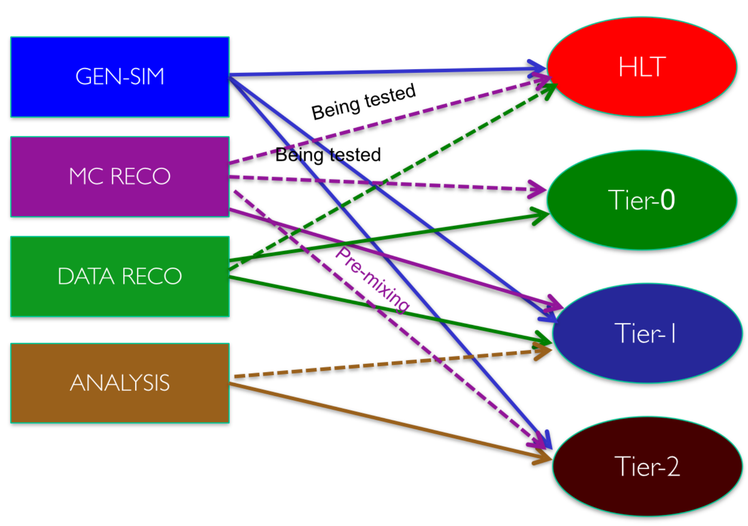

Figure 4. Various CMS workflow types and the resources where they run.

This expansion, however, further increases the scale at which the HTCondor pool must operate. In order to mitigate the risk during Run 2 that any scaling limitations we might encounter do not impact data taking, we opted to run the Tier-0 as an independent yet highly similar pool which can "flock" extra jobs to the Global Pool when needed[2]. CMS also has won some initial allocations on HPC clusters such as SDSC, to which we can submit workflows[3] with glideinWMS and which we would like to include in the Global Pool as well.

Scale Tests and Feedback

During 2014 we worked closely with both the HTCondor and glideinWMS development teams and the OSG[4] to find and fix problems that might limit the scalability of the system. In particular the scale testing performed by the OSG using CMS resources and CMS's own scale testing were both invaluable to identifying improvements that could be made in the communication between the various HTCondor components, in the Negotiator cycle, scheduler stability, etc. In particular, the OSG scale tests have demonstrated that stable operation of a HTCondor pool is possible at a scale of 200,000 parallel running jobs.

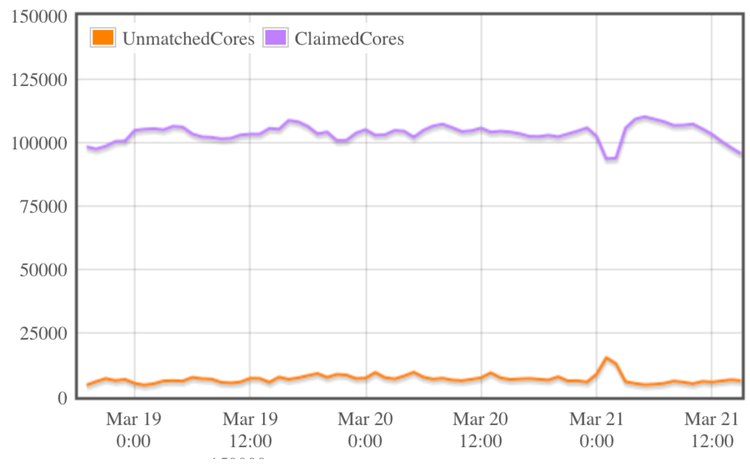

Figure 5. Stable operation of the Global Pool at scales of the pledged resources available has been achieved.

Support Model

The consolidation of glideinWMS operations in CMS into a single Global Pool has achieved significant economies of effort. Furthermore, we take full advantage of the "High Avaialbility" mode of glideinWMS to locate critical services in multiple availability zones. When one critical service (such as the central manager) goes down, another machine can take over the functionality in a seamless way.

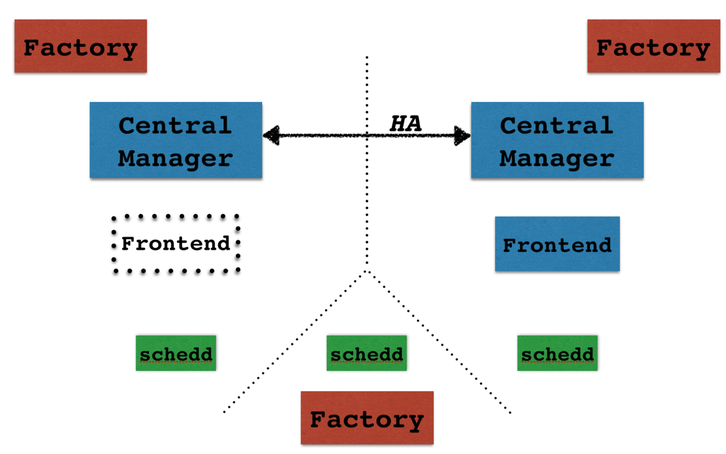

Figure 6. High-Availability (HA) mode is essential for smooth operations in all pools including the Tier-0.

Conclusions

We are currently running a Global Pool for glideinWMS in CMS which serves physics analysis, central data production and reconstruction, overflow from the Tier-0, and opportunistic and special allocations at HPC centers. We are confident that it can scale to our needs during Run 2, at least to 200,000 parallel running jobs and beyond, based on the testing we have made during the past year and the improvements made to HTCondor and glideinWMS partially as a result of feedback from those tests.

Related Work:

1. O. Gutsche et al., "Using the glideinWMS System as a Common Resource Provisioning Layer in CMS", CHEP15 Oral Presentation. #289.
2. D. Hufnagel et al., "The CMS Tier-0 Goes Cloud and Grid for LHC Run 2", CHEP15 Oral Presentation, #119.
3. D. Hufnagel et al., "Enabling Opportunistic Resources for CMS Computing Operations", CHEP15 Oral Presentation, #123.
4. E. Fajadro et al., "How Much Higher Can HTCondor Fly?", CHEP15 Poster, #6.

Contact Information: jletts@ucsd.edu

This work was partially supported by the U.S. Department of Energy and the National Science Foundation.