WLCG TEG: Workload Management

Kaushik De
Univ. of Texas at Arlington

OSG AHM, Lincoln
### March 20, 2012

Overview

- WLCG TEG on Workload Management
  - Goal: evaluate future developments in WMS
  - Chaired by: Davide Salomoni, Torre Wenaus
  - 48 members representing experiments & IT
  - This talk is my personal impression and summary of some of the work done in the WMS TEG so far
  - Though a member of the TEG, I have only been an observer – all the work was done by others
  - Apologies in advance for any inadvertent ATLAS bias in this talk – I know ATLAS flaws too well!

Scope of Work

- The main work areas in TEG-WMS are:

    - Pilots and frameworks

    - Resource allocation and job management

    - Use of information services

    - Security models

    - New computing models

- In this talk I will focus
  on some of these areas

- https://twiki.cern.ch/twiki/bin/view/LCG/WorkloadManagementTechnicalEvolution
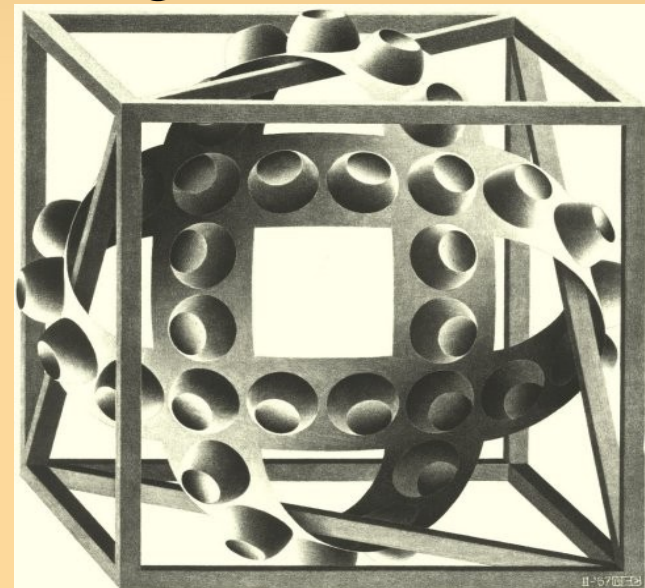
Pilots and Frameworks

- Grid submission layer

  - All four experiments use different products

  - Ongoing work in ATLAS to use glideinWMS

    - Interesting possibility of ATLAS/CMS common solution

- Streamed submission

  - Much discussion on a CE interface to replace external pilot factories – keep N pilots always queued at site

  - Approved by TEG – hopefully common solution

- Common job wrapper for pilots

  - Difficult – may need factorization of common parts

GlideinWMS in ATLAS

- Constant (but low) level of effort in US for many years to integrate glideinWMS with PanDA
  - Rod Walker (LMU Munich) is leading current push
  - Need to get OSG involved – meeting at BNL soon
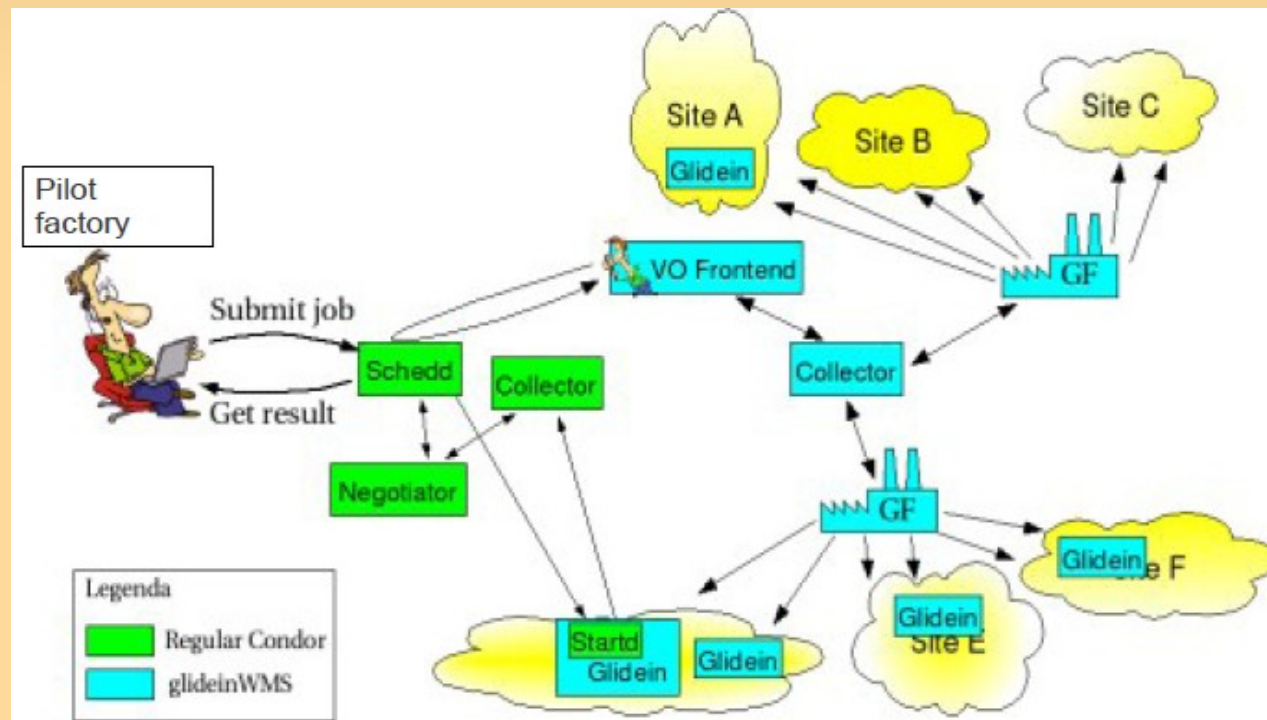
Resource Allocation and Job Management

- Some overlap with pilots & framework

- Whole node & multi-core jobs

- CPU affinity & CPU sets

- I/O vs CPU intensive job tagging

- Definition/requirements for CE

Multi-core Jobs

- LHC experiments typically run 1 job per core

- Development of multi-core exp. app's ongoing

- Sites are already highly multi-core

  - Sites are being asked to partition resources into whole node queues, which is inefficient

- Useful to have extensible system, to be ready when frameworks and sites will need them

  - Some middleware support already available

  - Need to match experimental requirements

- OSG can help here with common solution

New Computing Models

- Very active work area
- Also has overlap with pilots and frameworks
- Two main themes:
  - Virtualization
  - Cloud computing
- Many talks already in this meeting
  - Excellent introduction and status this morning
  - These results provided feedback for TEG

Virtualization

- Draft Recommendations:

  - virtualization is a decision of the site

  - sites should ensure that the virtualization penalty is lower than the gain in TCO

Cloud Computing

- Draft Recommendations:

  - Sites are free to use virtualization. If they do, it must be transparent to users

  - Sites which decide to go beyond virtualization, and setup an internal cloud, are encouraged to open part of the resources via an EC2 access to VOs which are interested in this.

  - In order to profit most from clouds at different sites, some level of synchronization between sites is needed. This includes the sharing of images and the definition of common virtual machine templates.

Cloud Computing

- In the described IaaS models, experiments need to integrate their frameworks with clouds, and sites need to adapt their applications to use internal clouds. We recommend to found a body which involves all parties to synchronize these efforts, with the aim to avoid independent developments and multiple solutions for the same problems.

- Accounting: Cloud access does not come for free. The TEG recomments that

  - VOs making use of cloud resources are fully accounted for their cloud usage.

  - fixed quotas per VO should be setup to ensure that each VO get their pledges.

- Possibilities to enable fair share at the cloud level need to be investigated

Summary

- Final report is being assembled

- Good exercise before LHC shutdown

- Experimental involvement is very important

- OSG has big role to play

- More info:
  https://twiki.cern.ch/twiki/bin/view/LCG/WorkloadManagementTechnicalEvolution