

Summary of workload management

Torre Wenaus (BNL), Frank Wuerthwein (UCSD)

Joint WLCG & HSF Workshop 2018

Naples

March 29 2018

Presentations

- ATLAS, CMS presentations
 - Asked to address how HL-LHC challenges impact WM and particularly how the challenges inform potential R&D projects
 - Approach to summarizing: the commonalities and distinctions in what they reported
- ARC, PanDA, DIRAC, OSG presentations
 - Asked to bring perspective as a WM tool/service provider to future commonality in WM, uniform interfaces to resources, roadmap
 - Approach to summarizing: tried to skip the ‘this is X’ material in favor of ‘this is the future roadmap for X’, highlighting prospects for R&D and common effort
 - Bottom line impression:
 - Don’t see prospects for fewer WM offerings in the future coming from merging or fading away
 - But many prospects for common effort and R&D, among these efforts and with wider HL-LHC directed R&D and common project activities
 - As expressed by Brian: common approaches, common capabilities, maybe not standardization

ATLAS & CMS - commonality & distinctions

- Agree: computing can't be allowed to blow the budget. A constrained resource. Technical and sociological challenge to expose to analysts the costs and tradeoffs of their processing.
- Agree: expectation management is vital. Understanding of what we can do, by when.
- Agree: addressing CPU shortfall demands much work, but we see paths.
- Agree: the bigger problem and challenge is disk storage.
- Agree: there will be managed production stages in the processing chain when we can expect the inputs for a campaign will not all fit on disk.
- Differ: CMS has one agreed compact analysis data product, ATLAS has ~100; if this persists, ATLAS likely will be hit by this earlier in the chain than CMS.
- Agree: The best bet for reducing storage budget is to rely more on tape and less on disk.
- Agree: tape usage will need to be orchestrated by WM & DM (eg train processing in 'carousel' mode) to reach the needed efficiency and turnaround time to meet analyst expectations. Generally, a full (re)architecting of the workflow + storage stack with WM and DM closely interacting.
- Agree: Interested to make greater tape usage part of near term R&D.
- Agree: we need to calculate, model, analyse just what such workflows imply.
- Differ: ATLAS couples this interest with granular processing/streaming R&D in the data lake context.
- Agree: common interest in data/analysis facility to explore 'declarative programming' for analysis.

ARC

- Roadmap points to
 - Streamlining: simpler config, REST, git
 - **resource discovery, info system**
 - **caching extensions**
 - Native containers
 - **alternatives to X509**
 - **extending aCT to other user communities**
 - **advanced, dynamic workflows**
 - **input, output streaming**
 - **dynamic optimization of task mixture**
 - **Harvester as common (with PanDA) approach**
 - **supporting resource diversity**
 - **metrics and info gathering**
 - **(sub)event level workflows**

PanDA

- Roadmap points to
 - **close integration with key (distinct) systems: info system, DDM system(s), metadata system(s)**
 - **reliance on key middleware (HTCondor)**
 - **Harvester: common, coherent, optimized provisioning for diverse resources**
 - **integration with networks**
 - new physics workflows, MC organization
 - **address new computing model**
 - **ML, analytics**
 - **granularity, data streaming**
 - **WM support for co-processor capabilities and workflows that utilize them**
 - **tighter WMS-DDM coupling: data lakes, fine grained processing (datasets => events)**
 - **data lakes R&D (Google 'data ocean')**

Bold = potential for R&D, common effort

DIRAC

- Roadmap points to
 - **Ongoing broad multi-experiment support through Consortium**
 - **Support other types of grids/fabrics as they appear and are interesting to DIRAC users**
 - **Dirac pilots as the federator of “any” computing resource**
 - Support for CE free resources via ssh
 - Support for CE & BS free resources via “Make a machine a pilot machine”, e.g. pilot launch in VM contextualization
 - **Clean, agnostic, extensible API**
 - **“Vision, and roadmap, won’t deviate while approaching Run3 and HL-LHC”**

OSG

- Roadmap points to
 - **our community should approach new sites with one voice, particularly in terms of providing an experiment-independent baseline**
 - **we should tackle greenfield projects together especially as we look at new site functionality**
 - **Shouldn’t a common compute service be a baseline starting point for discussing new resources?**
 - but large HPCs break such a baseline
 - **realistically look for common approaches, common capabilities; maybe not standardization**
 - **let community pick best-of-breed**