

GridPP

UK Computing for Particle Physics

Identifying Problematic Worker Nodes using pilots

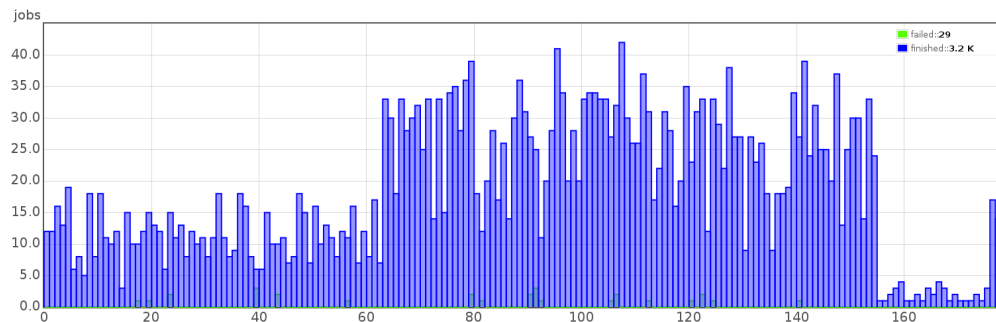
Christopher J. Walker, Alastair Dewhurst



Queen Mary
University of London

- Summary of Discussion at Hepsysman on proposal to pass back more information from pilot monitoring back to sites.
 - <https://indico.cern.ch/event/318354/session/5/contribution/5/material>
- Significant support for the idea
 - We have some of this information already.
- Two concepts
 - Send results back and amalgamate across experiments
 - Communicate results to local monitoring systems

- Central monitoring
 - Simple for sites
- Panda monitoring already useful for some problems
 - Extend (or provide similar) to monitor pilots
 - Per worker node
 - Per CE
- Ideally across VOs
 - Including non-LHC



- Put files on worker nodes
 - Wouldn't work if the WN disk is broken or full (one cause of failure).
 - Would be cleaned up if in /tmp etc by RAL
 - Would require checks to be written by sites.
- Send to syslog (Favoured option)
 - Option favoured by sites and receiving much of the discussion
 - Sites probably already monitoring syslog
 - Appears in /var/log/messages (where sysadmins look for problems)
 - Some monitoring systems pick up errors here already
 - Central logging a WLCG requirement

- Must ensure appropriate levels are used
 - INFO, WARN, ERROR
 - Successful jobs: Could report with level INFO - Further discussion needed to determine if this is desirable
- HTML like error codes + free form text liked (507 Insufficient storage in ATLASDATADISK or whatever)
 - http://en.wikipedia.org/wiki/List_of_HTTP_status_codes
 - This needs to be standardised across experiments
 - Sysadmins need to know what test was and how it failed
- Interface to sysadmin is syslog
 - VOs can send directly to syslog, or can use logger
 - If required (and not clear it is), machine features could be used to substitute a command other than logger - but with same syntax

- Avoid this being used as a stick to beat sites
- Sites need to know what the test was attempting, and how it went wrong.

- Lots of support from sites
- Central monitoring useful
- Local reporting via syslog considered better than dropping files on worker nodes
 - Easier to monitor
 - More likely to work