

Handling of WMS server Errors and Timeouts in DIRAC *TaskQueue* Director

In order to successfully fill up the LCG computing resources with pilots for LHCb, with current peaks of over 25 K jobs per day, DIRAC *TaskQueue* Director Agent needs to sustain a submission rate of at least twice this number to allow for site and wms inefficiencies as well as to have some margin for increasing rates. This means a pilots submission rate over 1000 pilots per hour.

For an efficient Operation, this must be achieved and maintained without the need of human intervention to manually include and exclude servers.

To do so, the Director submits (and previously list-match if required, ie not for SAM jobs, with a configurable [currently=15 minutes] caching time) in parallel (python-) threads to a randomly selected WMS server out of a configurable list [currently=11 servers at all our Tier1 sites].

All configurable parameters are read on every iteration of the Director (from 1 to few minutes depending on the execution time of the threads) so they can be updated on real time.

Quite often we observed that due to bugs in the server and client gLite-wms implementation, these commands (glite-wms-job-submit/glite-wms-job-list-match) fail or, even worst, take unreasonable large time to complete with cases of complete blocking of the execution thread. According to our experience, in a limited number of cases these errors are due to configuration issues on the server side, while the vast majority of cases are due to different overloads on the server side that result in unpredictable behaviours, return codes, return messages,...

The director sets a maximum execution time for any of this commands of 120 seconds and kills the execution if this time is reached, retrieving all the information provided to *stdout* and *stderr* up to this moment.

Additionally two caches with configurable different expiration times [currently=1 hour] are created (*failingWMSCache*, *ticketsWMSCache*), their usage is detailed bellow.

When an error code is returned or a timeout is detected, the Director takes several configurable actions:

- tries to exclude the affected server from the list considered in the current iteration, since several threads are running some other might have found the problem shortly before the error, or some might already have been directed to this server
- if the server is included in the *failingWMSCache*, the error is ignored
- the server is added to the *failingWMSCache*
- an error message starts to be prepared with the command failing, the failure (Timeout or command Error), and the full *stdout* and *stderr* collected, and a *errorAddress* is set as destination.
- if the server is in the *ticketsWMSCache* (means that failed once was already added to both the *failingWMSCache*, and the *ticketsWMSCache*; the *failingWMSCache* was cleared but the server failed again before the *ticketsWMSCache* expired) then the error message is converted into an alarm message by:
 - ◆ adding extra lines in the top of the message ("Submit GGUS Ticket for this error if not already opened It has been failing at least for %s hours", properly filled depending on the current configuration), and
 - ◆ the destination is changed to a configurable *alarmAddress* destination.
- if the server is not in the *ticketsWMSCache*, it is added with an extended expiration with respect to the one used for *failingWMSCache*.
- both error and alarm destinations are currently set to dirac.alarms at gmail.com, the account can be accessed via the usual lhcb password, but these are configurable values and can be disabled by setting them to a null string in the configuration.
- if the destination is valid, the Notification system is used to send the message to the requested

destination.

This configuration has allowed, during the last week, to achieve the necessary submission rate, in excess of 1k pilots per hour when necessary, without any need of manual intervention and with numerous errors from the WMS servers that in the past would have meant almost a full blockage of the pilot submission mechanism.

-- RicardoGraciani - 10 May 2009

This topic: LHCb > WMSErrorHandling

Topic revision: r1 - 2009-05-10 - RicardoGraciani



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](#)