Title

A gLite FTS based solution for managing user output in CMS

Abstract

The CMS distributed data analysis workflow assumes that jobs run in a different location to where their results are finally stored. Typically the user output must be transferred across the network from one site to another, possibly on a different continent or over links not necessarily validated for high bandwidth/high reliability transfer. This step is named stage-out and in CMS was originally implemented as a synchronous step of the job execution. However, our experience showed the weakness of this approach both in terms of low total job execution efficiency and failure rates, wasting precious CPU resources.

The nature of analysis data makes it inappropriate to use PhEDEx, CMS' core data placement system. As part of the new generation of CMS Workload Management tools, the Asynchronous Stage-Out system (AsyncStageOut) has been developed to enable third party copy of the user output. The AsyncStageOut component manages glite FTS transfers of data from a temporary store at the site where the job ran to the final location of the data on behalf of that data owner. The tool uses python daemons, built using the WMCore framework, talking to CouchDB, to manage the queue of work and FTS transfers. CouchDB also provides the platform for a dedicated operations monitoring system.

In this paper, we present the motivations of the asynchronous stage-out system. We give an insight into the design and the implementation of key features, describing how it is coupled with the CMS workload management system. Finally, we show the results and the commissioning experience.

Author

Daniele Spiga (CERN)

Co-authors

Mattia Cinquilli (CERN) ; CMS DMWM Group

Presentation type

parallel

-- SpigaDaniele - 20-Sep-2011

This topic: LCG > CHEP12AsyncStageOut
Topic revision: r1 - 2011-09-20 - SpigaDaniele