



Contribution ID: 177

Type: **Poster**

No file left behind - monitoring transfer latencies in PhEDEx

Tuesday, May 22, 2012 1:30 PM (4h 45m)

The CMS experiment has to move Petabytes of data among dozens of computing centres with low latency in order to make efficient use of its resources. Transfer operations are well established to achieve the desired level of throughput, but operators lack a system to identify early on transfers that will need manual intervention to reach completion.

File transfer latencies are sensitive to the underlying problems in the transfer infrastructure, and their measurement can be used as prompt trigger for preventive actions. For this reason, PhEDEx, the CMS transfer management system, has recently implemented a monitoring system to measure the transfer latencies at the level of individual files. For the first time now, the system can predict the completion time for the transfer of a data set. The operators can detect abnormal patterns in transfer latencies early, and correct the issues while the transfer is still in progress. Statistics are aggregated for blocks of files, recording a historical log to monitor the long-term evolution of transfer latencies, which are used as cumulative metrics to evaluate the performance of the transfer infrastructure, and to plan the global data placement strategy.

In this contribution, we present the typical patterns of transfer latencies that have been identified in the operational experience acquired with the latency monitor. We show how we are able to detect the sources of latency arising from the underlying infrastructure (such as stuck files) which need operator intervention, and we identify the areas in PhEDEx where a development effort can reduce the latency. The improvement in transfer completion times achieved since the implementation of the latency monitoring in 2011 is demonstrated.

Authors: RATNIKOVA, Natalia (KIT - Karlsruhe Institute of Technology (DE)); Dr MAGINI, Nicolo (CERN)

Co-authors: Dr SANCHEZ HERNANDEZ, Alberto (Centro Invest. Estudios Avanz. IPN (MX)); SARTIRANA, Andrea (Ecole Polytechnique (FR)); Dr HUANG, Chih-Hao (Fermi National Accelerator Laboratory); MOSCATO, Federica (Fermi National Accelerator Lab. (US)); KLUTE, Markus (Massachusetts Institute of Technology); YANG, Mingming (Massachusetts Inst. of Technology (US)); GUTSCHE, Oliver (Fermi National Accelerator Lab. (US)); ROSSMAN, Paul (Fermi National Accelerator Laboratory (FNAL)); KASELIS, Rapolas (Vilnius University (LT)); XIE, Si (Massachusetts Inst. of Technology (US)); PIPEROV, Stefan (Bulgarian Academy of Sciences (BG)); Dr CHWALEK, Thorsten (KIT - Karlsruhe Institute of Technology); Dr WILDISH, Tony (Princeton University (US))

Presenter: RATNIKOVA, Natalia (KIT - Karlsruhe Institute of Technology (DE))

Session Classification: Poster Session

Track Classification: Distributed Processing and Analysis on Grids and Clouds (track 3)