

ATLAS Grid Workflow Performance Optimization

Tuesday, July 10, 2018 2:15 PM (15 minutes)

The CERN ATLAS experiment grid workflow system manages routinely 250 to 500 thousand concurrently running production and analysis jobs to process simulation and detector data. In total more than 300 PB of data is distributed over more than 150 sites in the WLCG.

At this scale small improvements in the software and computing performance and workflows can lead to significant resource usage gains.

ATLAS is reviewing together with CERN IT experts several typical simulation and data processing workloads for potential performance improvements in terms of memory and CPU usage, disk and network I/O.

All ATLAS production and analysis grid jobs are instrumented to collect many performance metrics for detailed statistical studies using modern data analytics tools like Elasticsearch and Kibana.

This presentation will review and explain the performance gains of several ATLAS simulation and data processing workflows and present analytics studies of the ATLAS grid workflows.

Primary authors: ELMSHEUSER, Johannes (Brookhaven National Laboratory (US)); DI GIROLAMO, Alessandro (CERN); FILIPCIC, Andrej (Jozef Stefan Institute (SI)); LIMOSANI, Antonio (University of Sydney (AU)); SCHULZ, Markus (CERN); SMITH, David (CERN); SCIABA, Andrea (CERN); VALASSI, Andrea (CERN)

Presenter: ELMSHEUSER, Johannes (Brookhaven National Laboratory (US))

Session Classification: T3 - Distributed computing

Track Classification: Track 3 – Distributed computing