

Archival, anonymization and presentation of HTCondor logs with GlideinMonitor

Thursday, 20 May 2021 15:26 (13 minutes)

GlideinWMS is a pilot framework to provide uniform and reliable HTCondor clusters using heterogeneous and unreliable resources. The Glideins are pilot jobs that are sent to the selected nodes, test them, set them up as desired by the user jobs, and ultimately start an HTCondor schedd to join an elastic pool. These Glideins collect information that is very useful to evaluate the health and efficiency of the worker nodes and invaluable to troubleshoot when something goes wrong. This data, including local stats, the results of all the tests, and the HTCondor log files, is packed and sent to the GlideinWMS Factory. To access this information, developers and troubleshooters must exchange emails with Factory operators and dig manually into files. Furthermore, these files contain also information like email and IP addresses, and user IDs, that we want to protect and limit access to. GlideinMonitor is a Web application to make these logs more accessible and useful: it organizes the logs in an efficient compressed archive; it allows to search, unpack, and inspect them, all in a convenient and secure Web interface; via plugins like the log anonymizer, it can redact protected information preserving the parts useful for troubleshooting.

Primary authors: MAMBELLI, Marco (Fermilab (US)); HEIN, Thomas (University of Illinois Chicago); YANCEY, Mirica (Valparaiso University)

Presenter: MAMBELLI, Marco (Fermilab (US))

Session Classification: Monitoring

Track Classification: Distributed Computing, Data Management and Facilities