Contribution ID: **131**   Type: **Parallel**

End-To-End Solution for Integrated Workload and Data Management using glideinWMS and Globus Online

Tuesday, May 22, 2012 5:25 PM (25 minutes)

Grid computing has enabled scientific communities to effectively share computing resources distributed over many independent sites. Several such communities, or Virtual Organizations (VO), in the Open Science Grid and the European Grid Infrastructure use the glideinWMS system to run complex application work-flows. GlideinWMS is a pilot-based workload management system (WMS) that creates on demand, dynamically-sized overlay Condor batch system on Grid resources. While the WMS addresses the management of compute resources, however, data management in the Grid is still the responsibility of the VO. In general, large VOs have resources to develop complex custom solutions, while small VOs would rather push this responsibility to the infrastructure. The latter requires a tight integration of the WMS and the data management layers, an approach still not common in modern Grids. In this paper we describe a solution developed to address this shortcoming in the context of Center for Enabling Distributed Petascale Science (CEDPS) by integrating glideinWMS with Globus Online (GO). GO is a fast, reliable file transfer service that makes it easy for any user to move data. The solution eliminates the need for the users to provide custom data transfer solutions in the application by making this functionality part of the glideinWMS infrastructure. To achieve this, glideinWMS uses the file transfer plug-in architecture of Condor. The paper describes the system architecture and how this solution can be extended to support data transfer services other than GO when used with Condor or glideinWMS.

Primary author: MHASHILKAR, Parag (Fermi National Accelerator Laboratory)

Co-authors: Dr HOLZMAN, Burt (Fermi National Accelerator Lab. (US)); WEISS, Cathrin (UW Madison); Dr GARZOGLIO, Gabriele (FERMI NATIONAL ACCELERATOR LABORATORY); LACINSKI, Lukasz (University of Chicago); KETTIMUTHU, Raj (Argonne National Laboratory); Mr DUAN, Xi (Illinois Institute of Technology); MILLER, Zach (UW Madison)

Presenter: MHASHILKAR, Parag (Fermi National Accelerator Laboratory)

Session Classification: Distributed Processing and Analysis on Grids and Clouds

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