



Contribution ID: 277

Type: **Oral presentation to parallel session**

OASIS: a data and software distribution service for Open Science Grid

Tuesday, October 15, 2013 5:25 PM (22 minutes)

The Open Science Grid (OSG) encourages the concept of software portability: a user's scientific application should be able to run in as many operating system environments as possible. This is typically accomplished by compiling the software into a single static binary, or distributing any dependencies in an archive downloaded by each job. However, the concept of portability runs against the software distribution philosophy of many Linux packages, and becomes increasingly difficult to achieve as the size of a scientist's software stack increases. Despite being a core philosophy, portability has become a deterrent to the adoption of the OSG.

It is necessary to provide a mechanism for OSG Virtual Organizations (VO) to install software at sites. Since its initial release, the OSG Compute Element has provided an application software installation directory to VOs, into which VOs assume they can create their own sub-directory, install software into that sub-directory, and have the directory shared on the worker nodes for their sites (typically via NFS). The OSG provides guidelines for the size and UNIX permissions of such directories.

The current model lacks the ability to manage the software archive; there are shortcomings with regard to permissions, policies, versioning, and the lack of a unified, collective procedure or toolset for deploying software across all sites. Therefore, a new mechanism for data and software distributing is desirable. The proposed architecture for the OSG Application Software Installation Service (OASIS) is a server-client model: the software and data are installed only once in a single place (or a reduced number of places), and are automatically distributed to all client sites simultaneously.

Central file distribution offers other advantages, including server-side authentication and authorization, activity records, quota management, data validation and inspection, and well-defined versioning and deletion policies.

Currently the file transfer mechanism in OASIS is implemented using the CERN Virtual Machine Filesystem (CVMFS) as underlying technology.

The proposed architecture, as well as a complete analysis of the current implementation, will be described in this paper.

Authors: BOCKELMAN, Brian Paul (University of Nebraska (US)); HOVER, John (Brookhaven National Laboratory (BNL)-Unknown-Unknown); DE STEFANO JR, John Steven (Brookhaven National Laboratory (US)); Dr CABALLERO BEJAR, Jose (Brookhaven National Laboratory (US)); QUICK, Rob (OSG - Indiana University); TEIGE, Scott Werner (Indiana University (US))

Presenter: Dr CABALLERO BEJAR, Jose (Brookhaven National Laboratory (US))

Session Classification: Distributed Processing and Data Handling A: Infrastructure, Sites, and Virtualization

Track Classification: Distributed Processing and Data Handling A: Infrastructure, Sites, and Virtualization