

A case for application-aware grid services.

Monday, 13 February 2006 11:00 (20 minutes)

In 2005, the DZero Data Reconstruction project processed 250 tera-bytes of data on the Grid, using 1,600 CPU-years of computing cycles in 6 months. The large computational task required a high-level of refinement of the SAM-Grid system, the integrated data, job, and information management infrastructure of the RunII experiments at Fermilab. The success of the project was in part due to the ability of the SAM-Grid to adapt to the local configuration of the resources and services at the participating sites. A key aspect of such adaptation was coordinating the resource usage in order to optimize the typical access patterns of the DZero reprocessing application. Examples of such optimizations include database access, data storage access, and worker nodes allocation and utilization.

A popular approach to implement resource coordination on the grid is developing services that understand application requirements and preferences in terms of abstract quantities e.g. required CPU cycles or data access pattern characteristics. On the other hand, as of today, it is still difficult to implement real-life resource optimizations using such level of abstraction. First, this approach assumes maximum knowledge of the resource/service interfaces from the users and the applications. Second, it requires a high level of maturity for the grid interfaces. To overcome these difficulties, the SAM-Grid provides resource optimization implementing application-aware grid services. For a known application, such services can act in concert maximizing the efficiency of the resource usage. This paper describes what optimizations the SAM-Grid framework had to provide to serve the DZero reconstruction and montecarlo production. It also shows how application-aware grid services fulfill the task.

Primary authors: BARANOVSKI, Andrii (Fermilab); GABRIELE, Garzoglio (FERMI NATIONAL ACCELERATOR LABORATORY); MHASHILKAR, Parag (Fermilab)

Co-author: WICKE, Daniel (University of Wuppertal)

Presenter: GABRIELE, Garzoglio (FERMI NATIONAL ACCELERATOR LABORATORY)

Session Classification: Poster

Track Classification: Grid middleware and e-Infrastructure operation