Contribution ID: **430**   Type: **poster**

Efficient Multi-site data movement using Constraint programing for data hungry science

Thursday 26 March 2009 08:00 (20 minutes)

For the past decade, HENP experiments have been heading towards a distributed computing model in an effort to concurrently process tasks over enormous data sets that have been increasing in size as a function of time. In order to optimize all available resources (geographically spread) and minimize the processing time, it is necessary to face also the question of efficient data transfers and placements. A key question is whether the time penalty for moving the data to the computational resources is worth the presumed gain.

Onward to the truly distributed task scheduling we present the technique using Constraint Programming (CP) approach. The CP technique schedules data transfers from multiple resources considering all available paths of diverse characteristic (capacity, sharing and storage) having minimum user's waiting time as an objective. We introduce a model for planning data transfers to a single destination (data transfer) as well as its extension for an optimal data set spreading strategy (data placement). Several enhancements for solver of CP model will be shown, leading to a faster schedule computation time using symmetry breaking, branch cutting, well studied principles from job-shop scheduling field and several heuristics. Finally, we will present the design and implementation of a corner-stone application aimed at moving datasets according to the schedule. Results will include comparison of performance and trade-off between CP techniques and Peer-2-Peer model from simulation framework as well as the real case scenario taken from a practical usage of CP scheduler.

Presentation type (oral | poster)

oral

Authors: Dr LAURET, Jerome (BROOKHAVEN NATIONAL LABORATORY); Mr ZEROLA, Michal (Nuclear Physics Inst., Academy of Sciences)

Co-authors: Dr SUMBERA, Michal (Nuclear Physics Inst., Academy of Sciences, Praha); Prof. BARTAK, Roman (Faculty of Mathematics and Physics, Charles University, Praha)

Presenter: Mr ZEROLA, Michal (Nuclear Physics Inst., Academy of Sciences)

Session Classification: Poster session

Track Classification: Grid Middleware and Networking Technologies