

One click dataset transfer: toward efficient coupling of distributed storage resources and CPUs.

Monday, 5 September 2011 15:15 (25 minutes)

The massive data processing in a multi-collaboration environment with geographically spread diverse facilities will be hardly “fair” to users and hardly using network bandwidth efficiently unless we address and deal with planning and reasoning related to data movement and placement. The needs for coordinated data resource sharing and efficient plans solving the data transfer paradigm in a dynamic way are being more required. We will present the work which purpose is to design and develop an automated planning system acting as a centralized decision making component with emphasis on optimization, coordination and load-balancing.

We will describe the most important optimization characteristic and modeling approach based on “constraints”. Constraint-based approach allows for a natural declarative formulation of what must be satisfied, without expressing how. The architecture of the system, communication between components and execution of the plan by underlying data transfer tools will be shown. We will emphasize the separation of the planner from the “executors” and explain how to keep the proper balance between being deliberative and reactive. The extension of the model covering full coupling and reasoning about computing resources will be shown.

The system has been deployed within STAR experiment over several Tier sites and has been used for data movement in the favour of user analyses or production processing. We will present several real use-case scenario and performance of the system with a comparison to the “traditional” - solved by hands methods. The benefits in terms of indispensable shorter data delivery time due to leveraging available network paths and intermediate caches will be revealed.

Finally, we will outline several possible enhancements and avenues for future work.

Primary authors: Dr LAURET, Jérôme (Brookhaven National Laboratory); Mr ZEROLA, Michal (Academy of Sciences, Czech Republic); Dr BARTÁK, Roman (Charles University)

Co-author: Dr ŠUMBERA, Michal (Academy of Sciences, Czech Republic)

Presenter: Mr ZEROLA, Michal (Academy of Sciences, Czech Republic)

Session Classification: Monday 05th - Computing Technology for Physics Research

Track Classification: Track 1: Computing Technology for Physics Research