

A new CDF model for data movement based on SRM

Thursday, 26 March 2009 08:00 (20 minutes)

Being a large international collaboration established well before the full development of the Grid as the main computing tool for High Energy Physics, CDF has recently changed and improved its computing model, decentralizing some parts of it in order to be able to exploit the rising number of distributed resources available nowadays.

Despite those efforts, while the large majority of CDF Monte Carlo production has moved to the Grid, data processing is still mainly performed in dedicated farms hosted at FNAL, requiring a centralized management of data and Monte Carlo samples needed for physics analysis.

This rises the question on how to manage the transfer of produced Monte Carlo samples from remote Grid sites to FNAL in an efficient way; up to now CDF has relied on a non scalable centralized solution based on dedicated data servers accessed through rcp protocol, which has proven to be unsatisfactory.

A new data transfer model has been designed that uses SRMs as local caches for remote Monte Carlo production sites, interfaces them with SAM, the experiment data catalog, and finally realizes the file movement exploiting the features provided by the data catalog transfer layer.

We describe here the model and its integration within the current CDF computing architecture. We discuss the performance gain and the benefits of the new framework in comparison with the old approach.

Presentation type (oral | poster)

oral

Primary authors: Dr LUCCHESI, Donatella (University and INFN Padova); Dr BENJAMIN, Douglas (Duke University); Dr COMPOSTELLA, Gabriele (CNAF INFN); Dr JHA, Manoj Kumar (INFN Bologna); Dr PAGAN GRISO, Simone (University and INFN Padova)

Presenters: Dr COMPOSTELLA, Gabriele (CNAF INFN); Dr JHA, Manoj Kumar (INFN Bologna)

Session Classification: Poster session

Track Classification: Grid Middleware and Networking Technologies