

Data and metadata management - distributed data access and distributed databases

Wednesday, March 25, 2009 12:00 PM (30 minutes)

Data and meta data management at Petabyte scale remains at the key challenges for the High Energy Physics community. Efficient distribution and reliable access to Petabytes of distributed data in files and relational database

will be required to exploit the physics potential of LHC data and the resources available to the experiments in the world wide LHC computing grid. In this presentation we will summarise the software and deployment infrastructure for distributed data management at CERN and the WLCG partner sites and review the upcoming challenges for sustainable production deployment. We will focus on common technical challenges for the storage and distribution systems as the experiments are ramping up their distributed production and analysis work at CERN and the tier sites and outline the impact of new technologies such as data access protocols, virtualised storage, clustered file systems and changing storage media roles in the medium and long term.

2

Dirk Duellmann is deputy leader of the data management group in CERN's IT department and has been involved in data base and application development projects including the LCG Persistency Framework project and the LCG Distributed Database project over the last 15 years. He studied physics and computer science at the Universities of Münster and Hamburg and holds a PhD in experimental particle physics.

Primary author: Dr DUELLMANN, Dirk (CERN)

Presenter: Dr DUELLMANN, Dirk (CERN)

Session Classification: Plenary

Track Classification: Plenary