

Wide Area Network Access to CMS Data Using the Lustre Cluster Filesystem

Monday, 23 March 2009 08:00 (20 minutes)

The CMS experiment will generate tens of petabytes of data per year, data that will be processed, moved and stored in large computing facilities at locations all over the globe. Each of these facilities deploys complex and sophisticated hardware and software components which require dedicated expertise lacking at many of the university and institutions wanting access to the data as soon as it becomes available. Also, the standard methods for accessing data remotely rely on grid interfaces and batch jobs that while powerful, significantly increase the amount of procedural overhead and can impede a remote user's ability to analyze data interactively, develop and debug code and examine detailed information.

We believe that enabling direct but remote access to CMS data will greatly enhance the analysis experience for remote users not situated at a CMS Tier1 or Tier2.

The Lustre cluster filesystem allows remote servers the ability to mount filesystems over the wide-area-network as well as over the local-area network as it is more commonly used. It is also has an easy-to-deploy client, is reliable and performs exceptionally well. In this paper we report our experience using the Lustre filesystem to access CMS data from servers located a few hundred kilometers away from the physical filesystem. We describe the procedure used to connect two of the Florida Tier3 sites located in Miami and Daytona Beach to a storage element located in the University of Florida's, located in Gainesville, Tier2 center and its High Performance Computing Center. We include details on the hardware used, kernel modifications and tunings, report on network bandwidth, system I/O performance and compare these benchmarks with actual CMS application runs. We also propose a possible scenario for implementing this new method of accessing CMS data in the context of the CMS data management system. Finally we explore some of the issues concerning remote user access with Lustre, and touch upon security concerns.

Primary author: Prof. JORGE LUIS, Rodriguez (Florida Int'l University)

Presenter: Prof. JORGE LUIS, Rodriguez (Florida Int'l University)

Session Classification: Poster session

Track Classification: Distributed Processing and Analysis