Contribution ID: **72** Type: **Oral Presentation**

Parallel NFS : EMI data in competition with high-end industry storage solutions

Monday, 11 April 2011 16:00 (30 minutes)

Overview

Providing storage solutions for large international e-science infrastructures, like EGI, is an interesting challenge, as such infrastructures are supposed to support a large variety of diverse science communities. Most of those communities have well established but unfortunately different patterns with respect to data management and data access. The only way to tackle this problem, in a generic way, is to provide industry standard mechanisms in these areas. Besides other solutions, EMI data is approaching the idea of standards-in-storage by offering access to its storage elements through the NFS 4.1 (pNFS) protocol currently being adapted by leading industry storage providers. (e.g. IBM, Panasas, NetAPP etc). This approach makes EMI data competitive to expensive industry solutions and as a side effect, EMI data and the corresponding infrastructures are no longer in charge of maintaining the data access software on their data client entities and with such, resources are freed.

Impact

In current deployments of EMI data storage elements, specific proprietary software has to be installed and maintained on data client nodes. Moreover for RFIO, dCap and others, libraries have to be linked with the applications which makes this approach useless for communities where the source code of the application is not available. For other products, data clients might even have to be purchased. (e.g. GPFS).
With the common availability of pNFS in EMI storage elements, client software will be provided by the OS vendors, which will free resources in EMI and very likely in the infrastructure support (EGI) as well. Furthermore, with pNFS, the repository of EMI storage elements can be mounted into the file system as any other local storage resource. Any unmodified application will have access to data using plain POSIX I/O, an essential requirement for non-HEP communities.
Last but not least, pNFS enables EMI data and with that the EGI infrastructure, to provide storage solutions which can easily compete with expensive industry products and can be used as drop-in replacement.

Description of the work

For the last years, the Centre for Information Technology Integration (CITI), as part of the University of Michigan has been working on a specification of a successor of the overwhelmingly successful Network File Protocol (NFS), focusing on today's needs in terms of speed, reliability, security and modern distribution pattern of data sources.
The effort was supported by the major storage solution vendors, including Sun Microsystems, IBM, NetAPP, Pansas. Mircosoft and dCache.org, as a non funding partner. This work resulted in a specification which is now being implemented by all participants of the effort.
Parallel NFS (pNFS) is a part of the NFS v4.1 standard that allows clients to access storage devices directly

and in parallel. The pNFS architecture eliminates the scalability and performance issues associated with NFS servers in deployments today. This is achieved by the separation of data and metadata, and by moving the metadata server out of the data path.
The design of pNFS perfectly matches the design of the storage elements provided by EMI. dCache, as being part of the CITI group for some some years already offers a production version of pNFS with its current releases. DPM provides a beta version of pNFS available for early testers. A production release is expected before EMI-2. StoRM will provide pNFS as soon as IBM makes pNFS available for GPFS which is expected in 2011.
Besides the intense development work for DPM and dCache in this field, dCache.org has been putting significant efforts into testing of pNFS with available Linux clients. A small but realistic Tier II has been build including CPU and storage resource at DESY. pNFS client server interactions have been tested with plain I/O, ATLAS and CMS analysis software and with specific setups provided by the ROOT team. DPM will catch-up in terms of testing by making the DPM/pNFS implementation available to DPM-sites in the UK for community feedback.

URL

http://www.dcache.org/manuals/2010/20101102-hepix-patrick-nfs41.pdf

Conclusions

The just in time availability of Parallel NFS in industry (IBM, NetApp, Oracle ..) and Open Source products (Linux kernel) allows EMI data to provide professional storage elements, which can easily compete with expensive high end storage solutions. The discomfort of currently deployed systems, in terms of product dependent protocols and client software which needs to be installed and maintained, will be overcome. Significant testing of pNFS in EMI data products as well as in industry products is in progress and more results will be presented at the time of the Vilnius meeting.

Primary author: Dr FUHRMANN, Patrick (DESY/dCache.org)

Co-authors: Mr BAUD, Jean-Philippe (CERN); Mr ROCHA, Ricardo (CERN); Dr ZAPPI, Riccardo (CNAF/INFN); Mr MKRTCHYAN, Tigran (DESY/dCache.org)

Presenter: Dr FUHRMANN, Patrick (DESY/dCache.org)

Session Classification: EMI: Software for Distributed Computing Infrastructures

Track Classification: Data Management - Technology