



Contribution ID: 394

Type: poster

Using HEP Systems to Provide Storage for Biologists

Tuesday 28 September 2004 10:00 (1 minute)

Protein analysis, imaging, and DNA sequencing are some of the branches of biology where growth has been enabled by the availability of computational resources. With this growth, biologists face an associated need for reliable, flexible storage systems. For decades the HEP community has been driving the development of such storage systems to meet their own needs. Two of these systems - the dCache disk caching system and the Enstore hierarchical storage manager - are viable candidates for addressing the storage needs of biologists. Both incorporate considerable experience from the HEP community.

While biologists have much to gain from the HEP community's experience with storage systems, they face several issues that are unique to the biological sciences. There is a wider diversity in experiments, in number and size of datafiles, and in client operating systems in biology than there is in HEP. Patient information must be kept confidential. Disparate IT departments set up firewalls that separate client systems and the storage system.

Vanderbilt University is developing a storage system with the goal of meeting biologists' needs. This system will use Enstore for its robustness and reliability, and will use the flexible door-based architecture of dCache to provide storage services to biologists via web-portal, the dCache copy command, and custom applications. This system will be deployed using an automated tape library, several secure central servers, and nodes placed near biologists' existing compute infrastructure to ensure locality of caches and secure data channels between researchers and the central servers.

Author: CALEF, M. (VANDERBILT UNIVERSITY)

Presenter: TACKETT, Alan

Session Classification: Poster Session 1

Track Classification: Track 6 - Computer Fabrics