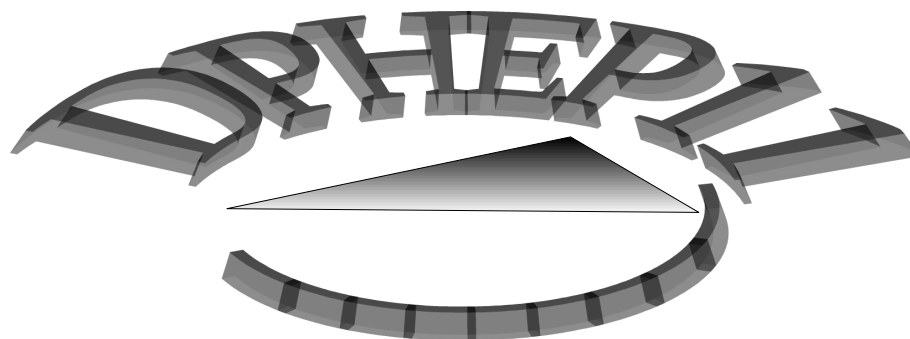


Common Solutions and Common Tools for Data Preservation



A comparison of the models used so far
for data preservation at DESY and BaBar
&
Identifying ways of sharing tools.

By H. Neal on behalf of all the discussion
participants

Data Access Models

	BaBar	Desy
MODEL	<p>Migrate to RHEL6 or SL6 and then freeze the OS using virtualization and make secure using firewall and default deny access.</p> <p>OS on servers continues to be kept up-to-date by IT.</p>	<p>Virtualization used for the validation but analysis and production done on maintained platform as long as possible</p>
THEORY	<p>Freeze while expertise for performing any fixes needed for migration is still present and provide stable platform. Almost all dependencies are preserved in the frozen image.</p>	<p>Continuing migrating to new platforms until an uncurable problem is encountered. Provide validation that will identify as quickly and as precisely as possible any problems.</p>

Problem resolution

	BaBar	Desy
ACTION	Freeze while expertise for performing any fixes needed for migration is still present and provide stable platform. Almost all dependencies are preserved in the frozen image.	Continuing migrating to new platforms until an uncurable problem is encountered. Provide validation that will identify as quickly and as precisely as possible any problems.

Archival System Solutions

- Solution depends on:
 - Required resources
 - Frequency of use
 - Amount of support available

- Private cloud
- Cloud shared between projects at site
- commercial/research clouds

For projects with higher demands

Lower demands less often and little development

Modularity of Archival System

- Concern expressed about data access:
 - If system moved will host be able to provide the needed file access
 - Who can predict long term future of DCACHE/XROOTD?
 - Must have ability to have standardized simple access to files
 - Ex: NFS

Validation

- Warning: cores will be diversifying and differences can occur in precision calculations but hopefully at levels insignificant for the physics
- Current validation algorithms need to be made more human; a physics based judgement not just a strict reproduction.
- BaBar will also need periodic validations and is interested in using the tools/facility produced by DESY
.... but how?

Validation Client Issue

- Offer is to run BaBar validation suite at DESY and have the system validate the results
 - Issues:
 - Reconstruction validation requires raw data input; data not available to non-associates
 - Running BaBar applications requires that BaBar software release be present; code not available to non-associates
 - Possible solution:
 - An agreement to allow restricted access to code and small data sample for the sole purpose of validation

Tools/Services for All

- Validation tools
 - All projects involving functional code preservation will need validation, can a common tool be made
- Virtualization experience
 - Valuable tools for preservation and several DPHEP participating groups now have much experience to share
- Documentation of mature projects
 - Details are important for development of new initiatives