

The BaBar Long Term Data Analysis facility at UVic

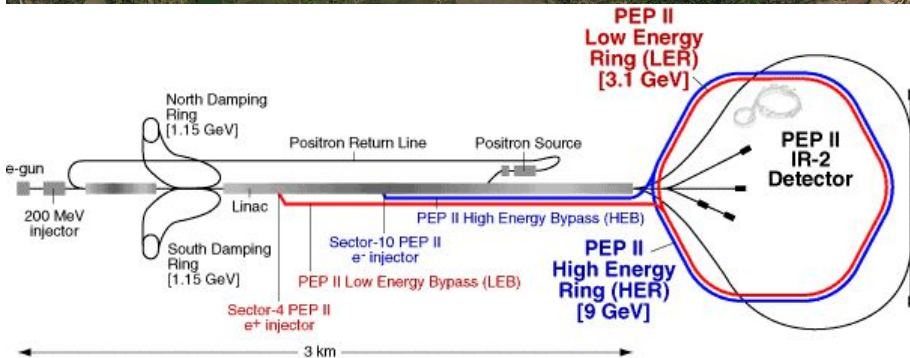
Marcus Ebert

University of Victoria

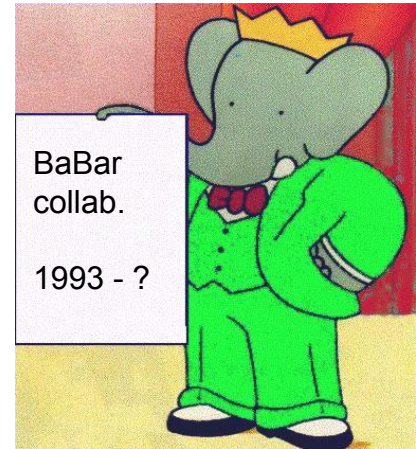
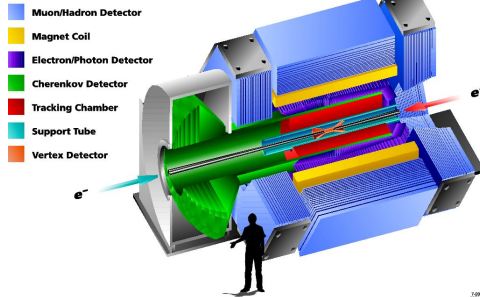
BaBar experiment



- collider experiment at 
- BaBar founded 1993
- data taking 1999-2008



BABAR Detector



BaBar Status

- BaBar stopped data taking in 2008, anticipated to do data analyses until 2018
 - but still actively doing analyses (local, no Grid usage)
 - 223 active authors from 14 countries
 - 27 new analyses publications since 2018 (more than 60 incl. conference proceedings)
 - 5 new analyses submitted in 2023; 3 published in 2023, so far (not incl. conference proceedings)
- Beginning of 2021: support for infrastructure at SLAC finally stopped
 - support extended from 2018 to beginning of 2021
- “Everything” needed to be moved away from SLAC to be able to continue
 - very tightly integration of SLAC services and BaBar services, grown over years
- UVic HEP-RC group offered the setup of new analysis environment

What is “Everything”?

- Data
 - collected collision data and generated MC events (~1.5PB)
 - all in root files
 - metadata stored in mysql database
 - number of events per root file, dataset/

What is “Everything”?

- **Data**

- collected collision data and generated MC events (~1.5PB)
 - all in root files
- metadata stored in mysql database
 - number of events per root file, dataset/

- **Analysis environment**

- software is 32bit, users usually write C++ code and compile their analysis modules
 - does not compile on 64bit-only systems
- depends on older software releases, e.g. perl, xrootd,...
 - latest verified system: SL6.3, gcc 4.4.x, kernel 2.6,...

- **Documentation**

- new users still join, sometimes just for a single analysis
- preserving documentation only way to have someone successfully started

- **Collaboration tools**

- calendar, analysis review, mailing lists, meeting organizer, ...

Analysis Environment - Overview

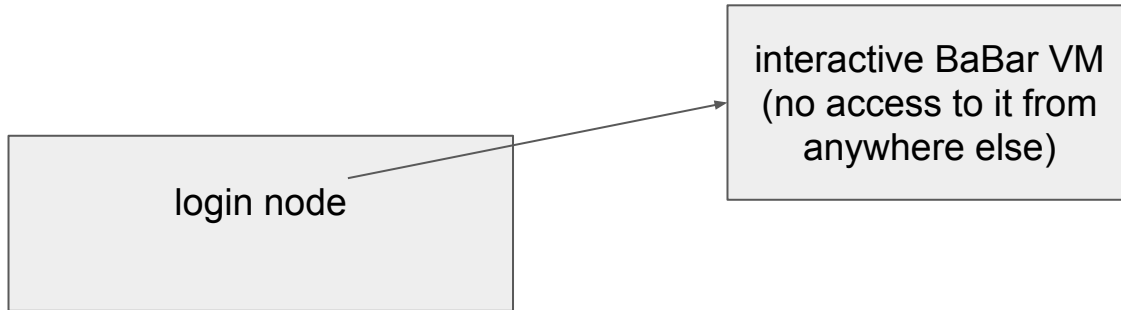
- User need to compile locally -> [user accounts/management](#)
- Users need to run over thousands of data files -> [batch system](#)
- Batch system jobs need to access local user environment -> [shared file system](#)
- All needs to run in an outdated, unsecured environment -> [isolation](#)
- Users want to take their job output home -> [data transfer machine](#)
- jobs need to access data in root files -> [XRootD system](#)
- hardware replacement uncertain -> [redundancy needs to be built in](#)

Isolation

- OS and tools frozen, since long time without security updates
 - BaBar already used a VM based system at SLAC

Isolation

- OS and tools frozen, since long time without security updates
 - BaBar already used a VM based system at SLAC



BaBar-To-Go is alternative to use UVic system.

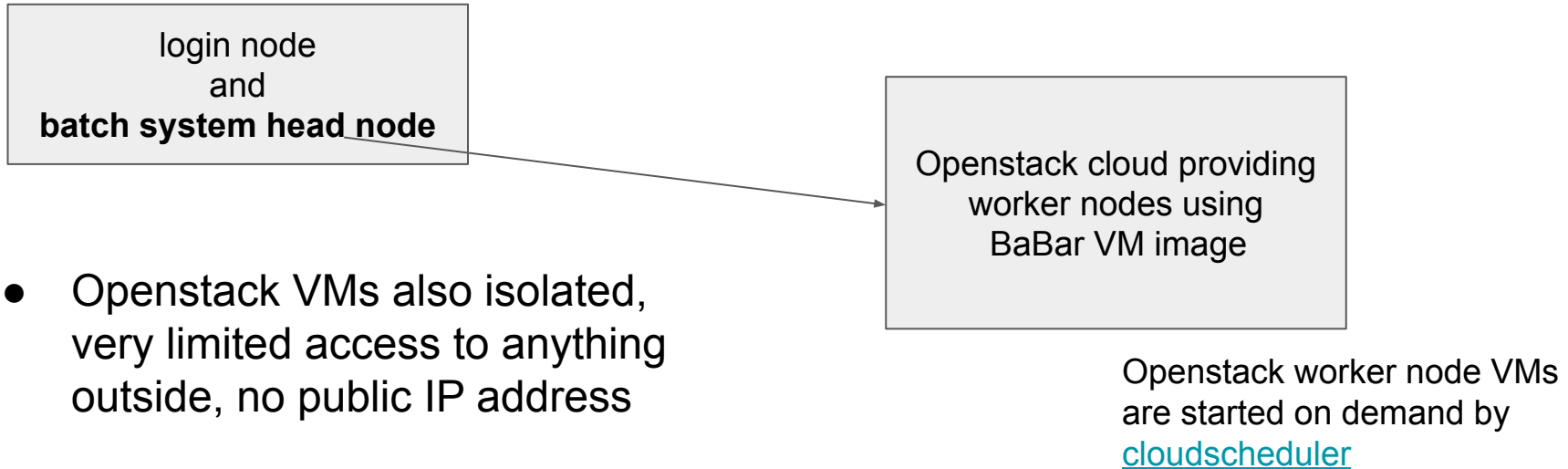
- login node reachable from the outside
 - current OS that gets security updates
- interactive VM can only be accessed from the login node; limited access to outside
 - interactive VM based on BaBar's approved image

Batch System

- BaBar used Torque/Maui/LSF before leaving SLAC
- HEP-RC group uses HTCondor to start Openstack VMs on demand as worker nodes
 - needed to write wrapper scripts
 - framework/users -> torque/maui/LSF commands -> wrapper script -> HTCondor commands
 - HTCondor command output->wrapper script->torque/maui/LSF style output->framework/users

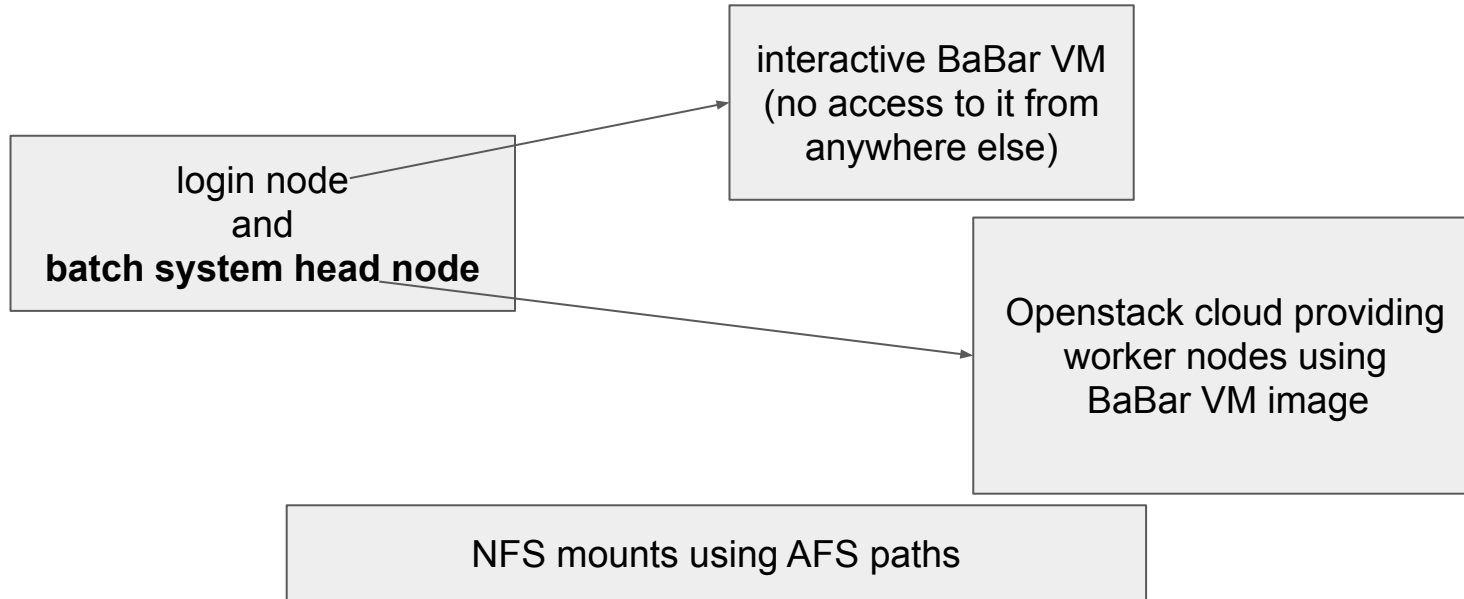
Batch System

- BaBar used Torque/Maui/LSF before leaving SLAC
- HEP-RC group uses HTCondor to start Openstack VMs on demand as worker nodes
 - needed to write wrapper scripts
 - framework/users -> torque/maui/LSF commands -> wrapper script -> HTCondor commands
 - HTCondor command output->wrapper script->torque/maui/LSF style output->framework/users



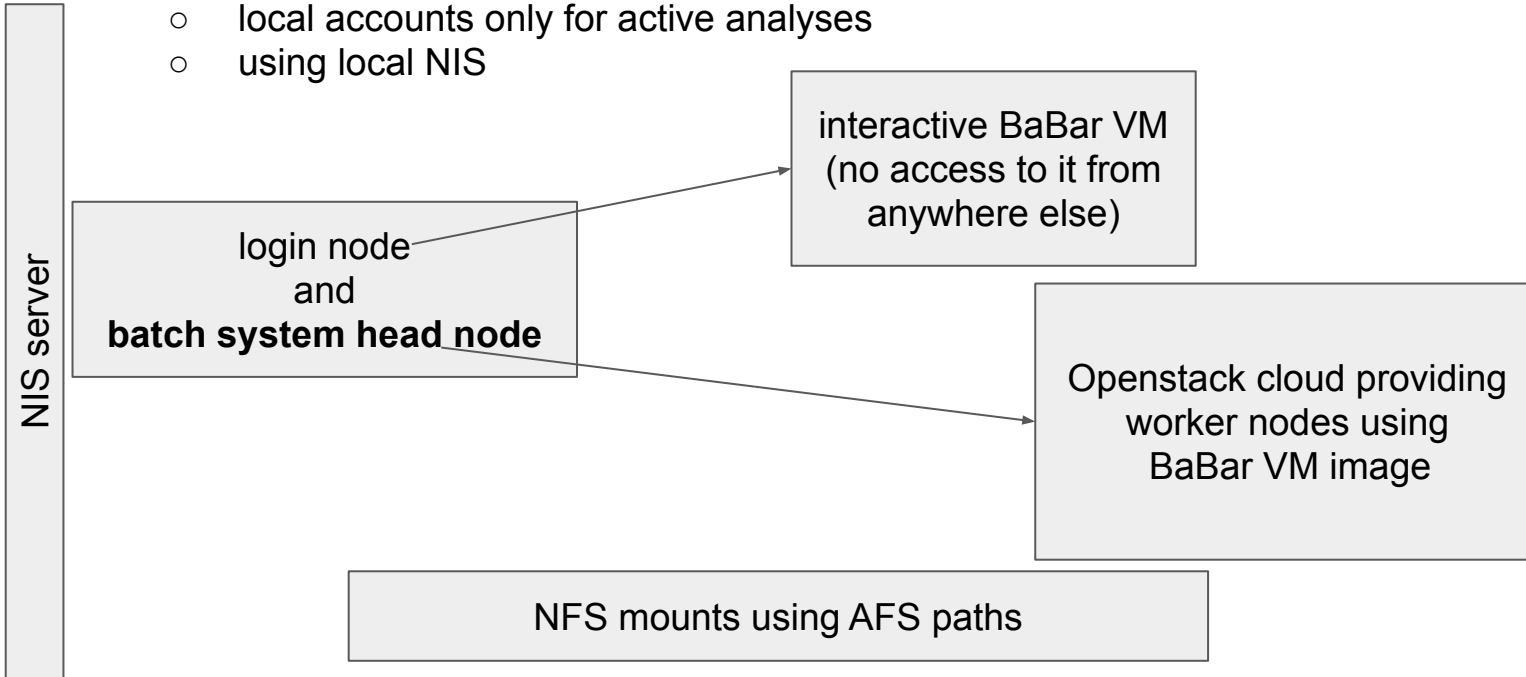
Shared File System

- AFS at SLAC
 - all of BaBar's software in a well defined directory structure
 - use NFS on new system



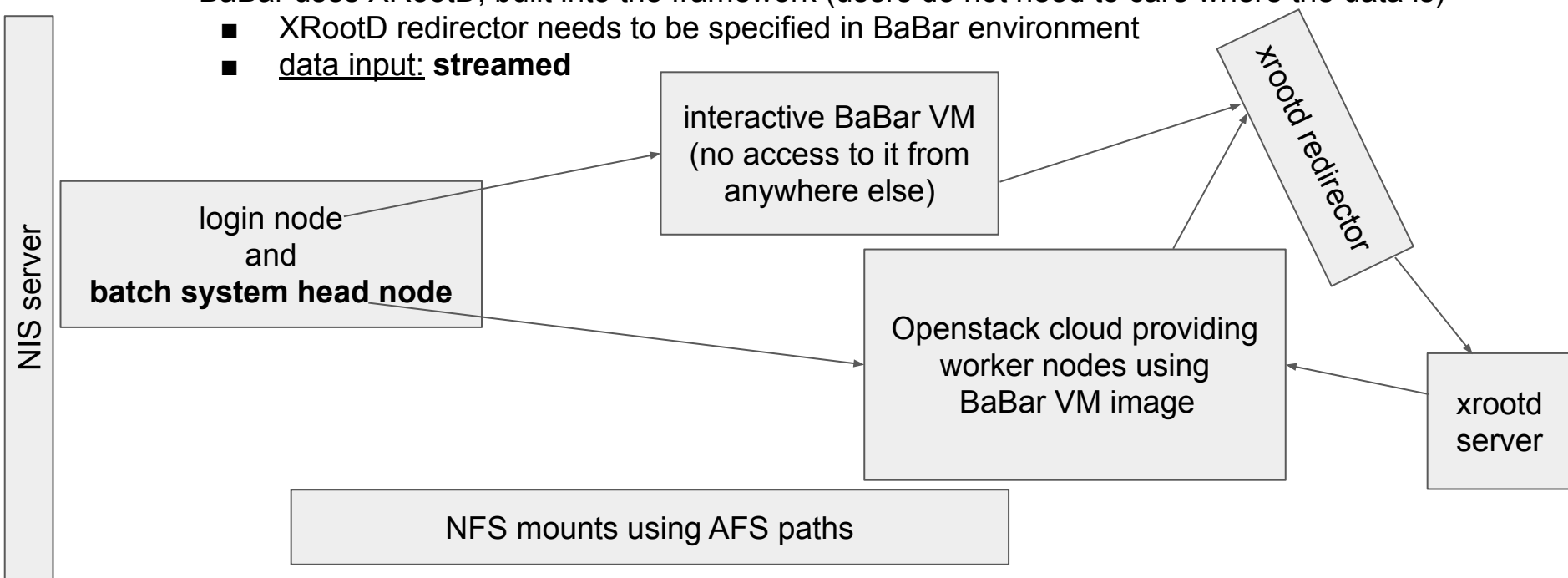
User Accounts and Management

- everyone in BaBar had an account at SLAC
 - can't do that here, most people do not need it anymore
 - local accounts only for active analyses
 - using local NIS



Data Access

- access needed from interactive machine and from worker nodes
 - BaBar uses XRootD, built into the framework (users do not need to care where the data is)
 - XRootD redirector needs to be specified in BaBar environment
 - data input: streamed



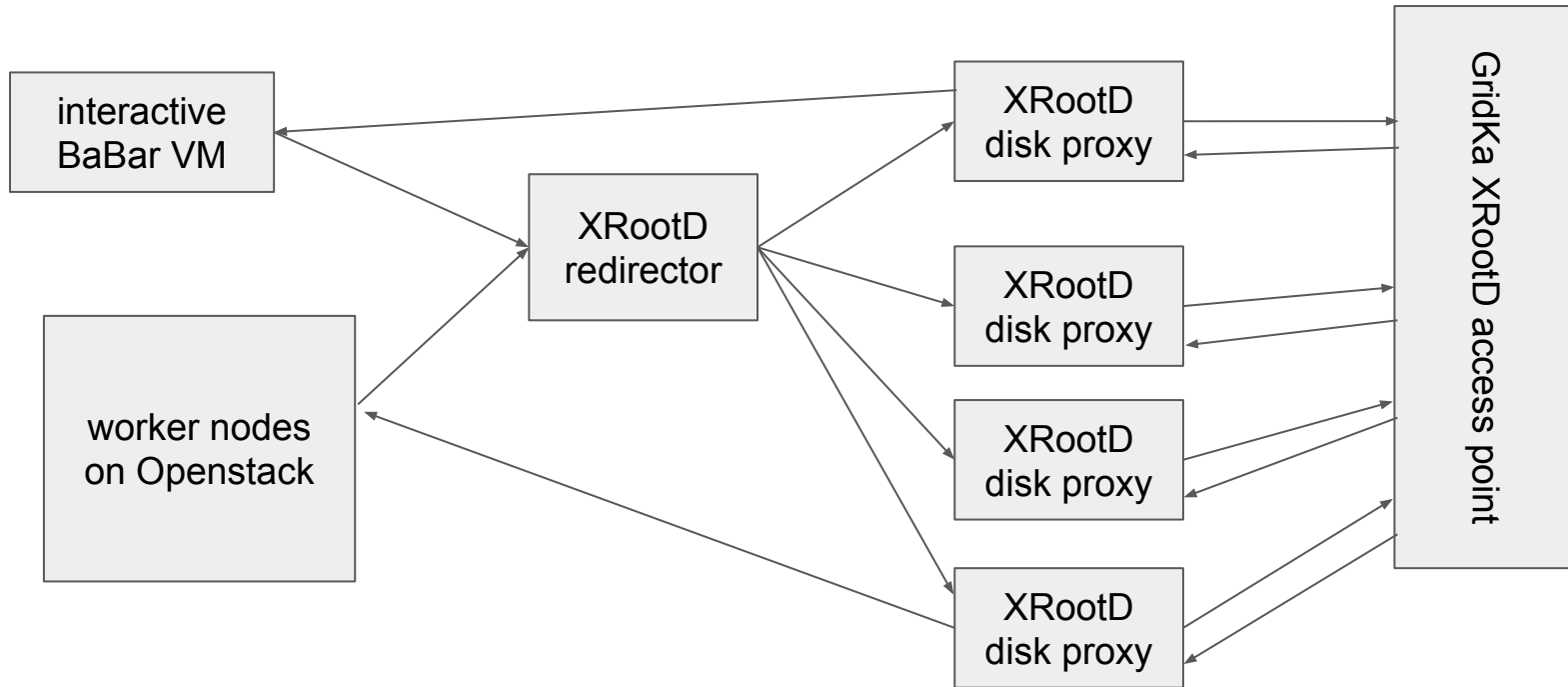
Data Access

- Data available to analyses: ~1.5PB
 - no storage available at UVic for it
- GridKa agreed to host the BaBar data to be used by analyses
 - BaBar site since a long time
 - had already some data on site; anything missing was copied to GridKa
 - also the metadata db to find the data files needed in an analysis was already there
 - only update of content needed
 - BaBar environment configuration specifies XRootD and db access point for the data and db
- Framework at UVic needs to access data at GridKa via streaming...
 - works surprisingly well for normal event data
 - workflow: read event, process, read event, process,...
 - but conditions data is also read via streaming
 - large amount of data each job needs to read



Data Access

- direct access to GridKa ---> access via local cache system



Data Transfer Machines

- for copying user's job output out of the system
- job output is on NFS
- do not want to block the login machine with data transfers
 - usual transfer tools disabled on the login machine
- new VM on Openstack system created (current and updated OS)
 - uses same NIS and NFS as the other BaBar VMs
 - only be used for data transfers out of UVic

Documentation

- HTML content: Historic documentation in html pages; need local account to edit
 - content all under a common BaBar directory
 - copy to UVic and use with a new Apache server
- Wiki: main documentation system since ~2013; need browser access to edit
 - content in mysql db
 - use mysql dump with new server at UVic
- HyperNews: forum for discussions
 - largely integrated into the SLAC infrastructure (permissions, email system, ...)
 - retired, but keep as archive in r/o mode
 - content all under a common BaBar directory
 - copy over to UVic and setup as r/o archive

Issues that came up during migration

- binaries linked against compiled libraries
 - sometimes tests done in user's home directory
 - binaries compiled and tested there, binaries and libs copied over to central software directory
 - binary still looks for lib in home directory...
- sometimes absolute path used instead of relative ones
 - replicate the AFS structure in NFS
- full URLs used in documentation
 - in html pages easy to fix - keep same structure on new server and change hostname via *sed*
 - in wiki pages more complicated; all data is in db
 - have users fix it "on the go"
 - SLAC implemented redirect to new server at UVic ("sed" equivalent)

Redundancy/Reliability

Hardware overview:

- XRootD proxy server: **old machines**
- XRootD redirector: VM on an **old machine**
- login machine: VM on an **old machine**
- BaBar interactive VM: VM on an **old machine**
- NIS server: VM on an **old machine**
- web server: on VM on an **old machine**
- babar wiki: VM on an **old machine**
- babar Hypernews: VM on an **old machine**
- NFS server: **one new server**, multiple **old machines**

old machine==out of warranty

Redundancy/Reliability:

- protect against disk failure
- protect against server failure

Redundancy/Reliability

login machine VM
NIS Server VM
interactive VM
XRootD redirector VM

hardware raid1 OS
ZFS mirror data disks

- spare server setup
the same way
- ZFS send/receive

Redundancy/Reliability

login machine VM
NIS Server VM
interactive VM
XRootD redirector VM

hardware raid1 OS
ZFS mirror data disks

XRootD proxy server

hardware raid1 for OS
ZFS raidz3 data disks

- spare server setup the same way
- ZFS send/receive

- multiple servers available
- just cache, loose no data

Redundancy/Reliability

login machine VM
NIS Server VM
interactive VM
XRootD redirector VM

hardware raid1 OS
ZFS mirror data disks

- spare server setup the same way
- ZFS send/receive

XRootD proxy server

hardware raid1 for OS
ZFS raidz3 data disks

- multiple servers available
- just cache, loose no data

Web documentation VM
Wiki VM
Hypernews (HN) VM

hardware raid1 for OS
ZFS raidz3 for data disks

- web content on NFS
- HN content on NFS
- images backed up
- daily mysql dump to NFS

Redundancy/Reliability

login machine VM
NIS Server VM
interactive VM
XRootD redirector VM

hardware raid1 OS
ZFS mirror data disks

- spare server setup the same way
- ZFS send/receive

XRootD proxy server

hardware raid1 for OS
ZFS raidz3 data disks

- multiple servers available
- just cache, loose no data

Web documentation VM
Wiki VM
Hypernews (HN) VM

hardware raid1 for OS
ZFS raidz3 for data disks

- web content on NFS
- HN content on NFS
- images backed up
- daily mysql dump to NFS

4 NFS server:
NFS \$HOME
NFS job output
NFS framework
NFS documentation

all use:
ZFS raidz2/3
hardware raid1 for OS

- spare server setup in the same way
- ZFS send/receive
- extra backup of framework and documentation

Collaboration tools

- Google Drive:
 - analysis documentation (one folder per analysis)
 - for new analyses; old analysis documentation archived at INSPIRE
 - Management folders
 - Publication Board
 - SpeakersBureau
 - general: calendar, author list, list of orchids,...
 - use of google sheets and docs
- CERN tools
 - egroups: for discussions and for accessing INDICO
 - INDICO: meetings (Collaboration meetings, practice talks, analysis reviews,...)
- CALTECH mailing list
 - to reach different groups of collaboration members for general announcements
 - others can reach specific groups inside of BaBar (management, speakers bureau,...)

Summary

It is possible to move an experiment's infrastructure to a new home.

A bit of planning on the experiment's side - while manpower is available - helps a lot.

Issues are in details: different queue systems, users patching/compiling/linking software outside of the production area, using absolute vs relative file names in software, absolute vs relative links in web/wiki pages

System at UVic running successfully since end of 2021.
(BaBar-To-Go is alternative)

Summary

It is possible to move an experiment's infrastructure to a new home.

A bit of planning on the experiment's side - while manpower is available - helps a lot.

Issues are in details: different queue systems, users patching/compiling/linking software outside of the production area, using absolute vs relative file names in software, absolute vs relative links in web/wiki pages

System at UVic running successfully since end of 2021.
(BaBar-To-Go is alternative)

Thank you!