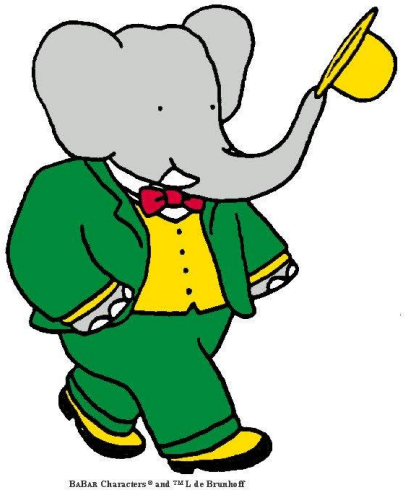
The
BaBar Long Term Data Preservation
and
Computing Infrastructure

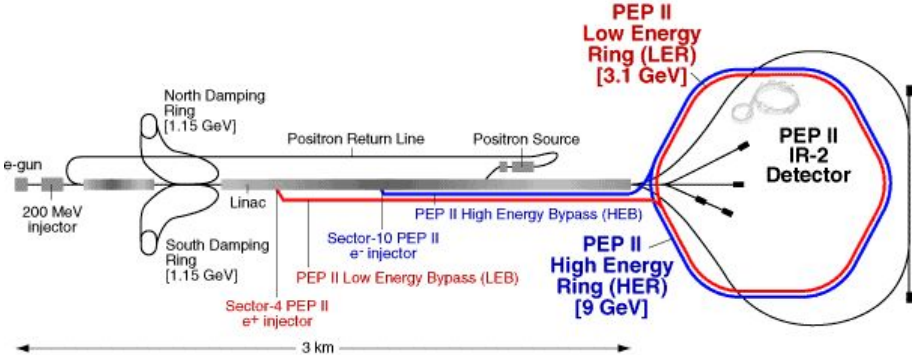


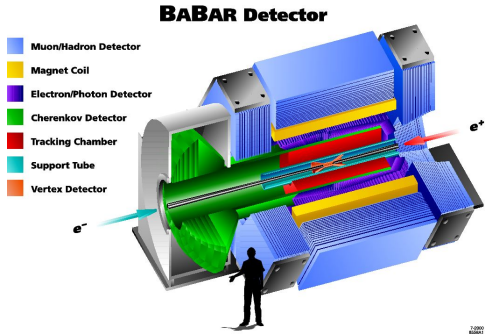
Marcus Ebert
BaBar Computing Coordinator

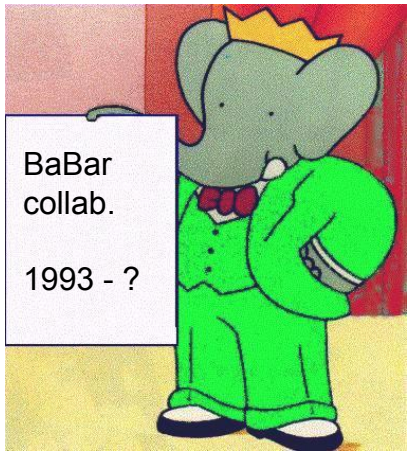
on behalf of
BaBar

The BaBar Experiment

- collider experiment at SLAC NATIONAL ACCELERATOR LABORATORY
- BaBar founded 1993
- data taking 1999-2008







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 analyses published in 2023 (not incl. conference proceedings)

- Beginning of 2021: support for infrastructure at SLAC finally stopped
  - support extended from 2018 to beginning of 2021

- To be able to continue, everything still needed had to be moved away from SLAC
  - very tightly integration of SLAC services and BaBar services, grown over years

- Analysis system and documentation moved to University of Victoria

What is needed for new system?

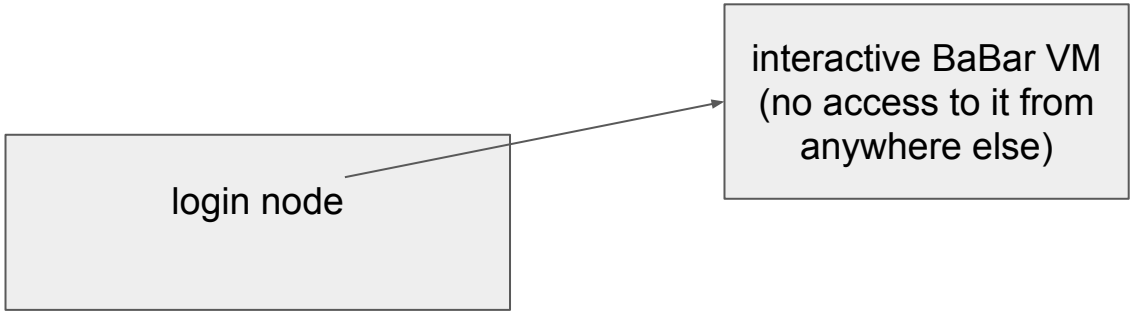
- Data
  - collected collision data and generated MC events (~1.5PB)
    - all in root files
  - metadata stored in mysql database
    - number of events per root file, dataset,...
- **Analysis environment**
  - software is 32bit, users usually write C++ code and compile their analysis modules
    - 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

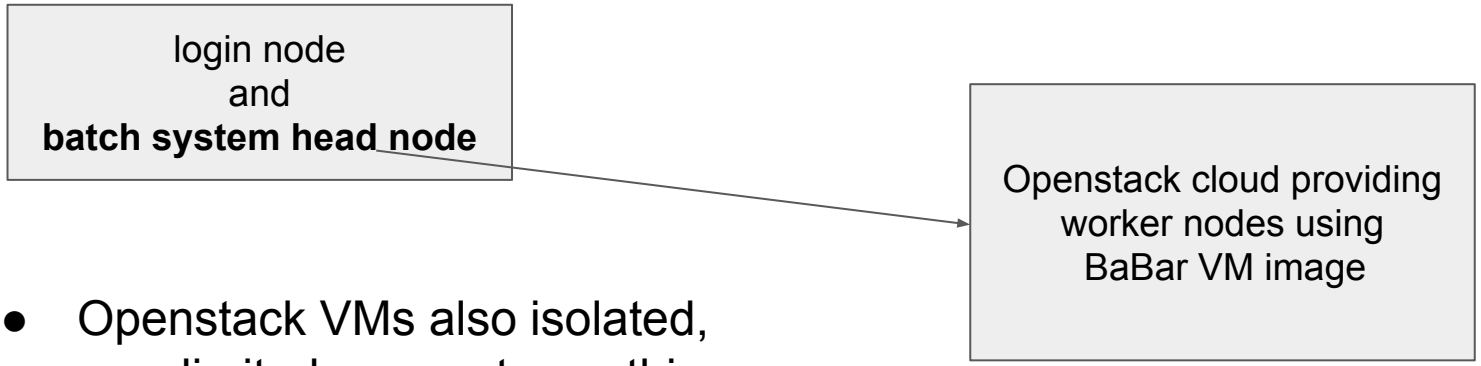


BaBar-To-Go is alternative to use UVic system.

- 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

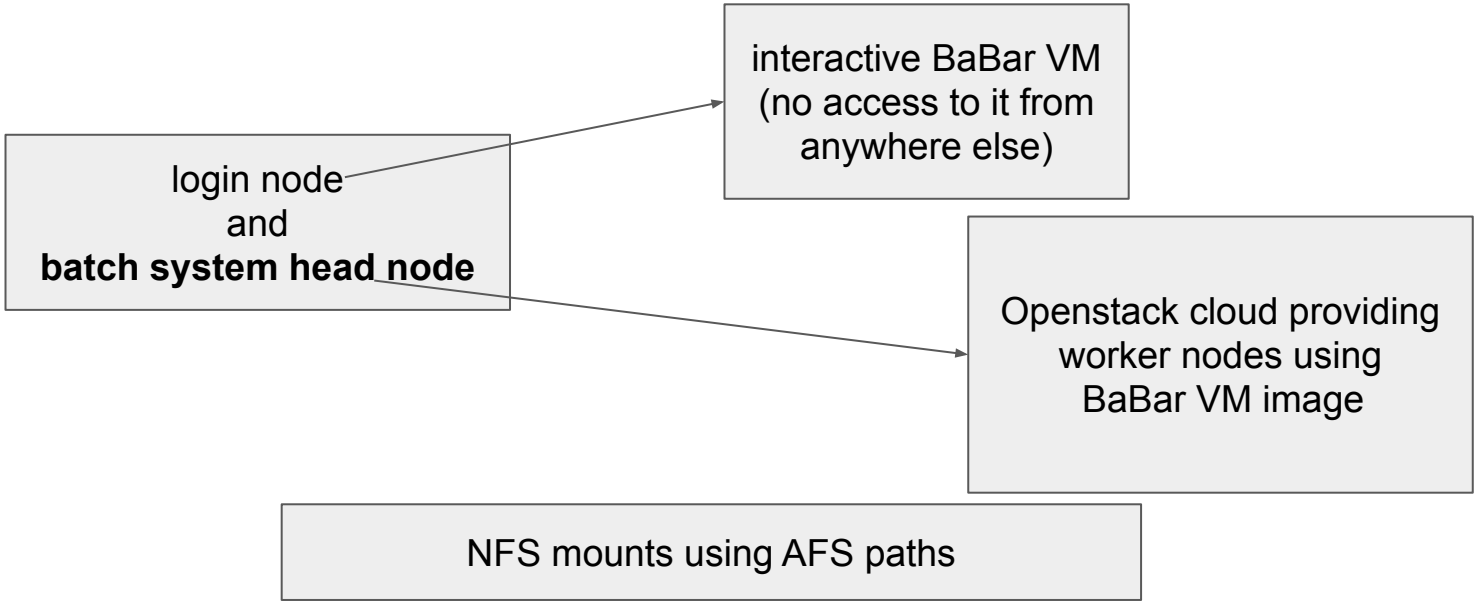


- Openstack VMs also isolated, very limited access to anything outside, no public IP address

Openstack worker node VMs are started on demand by [cloudscheduler](https://csv2.heprc.uvic.ca/public/) (https://csv2.heprc.uvic.ca/public/)

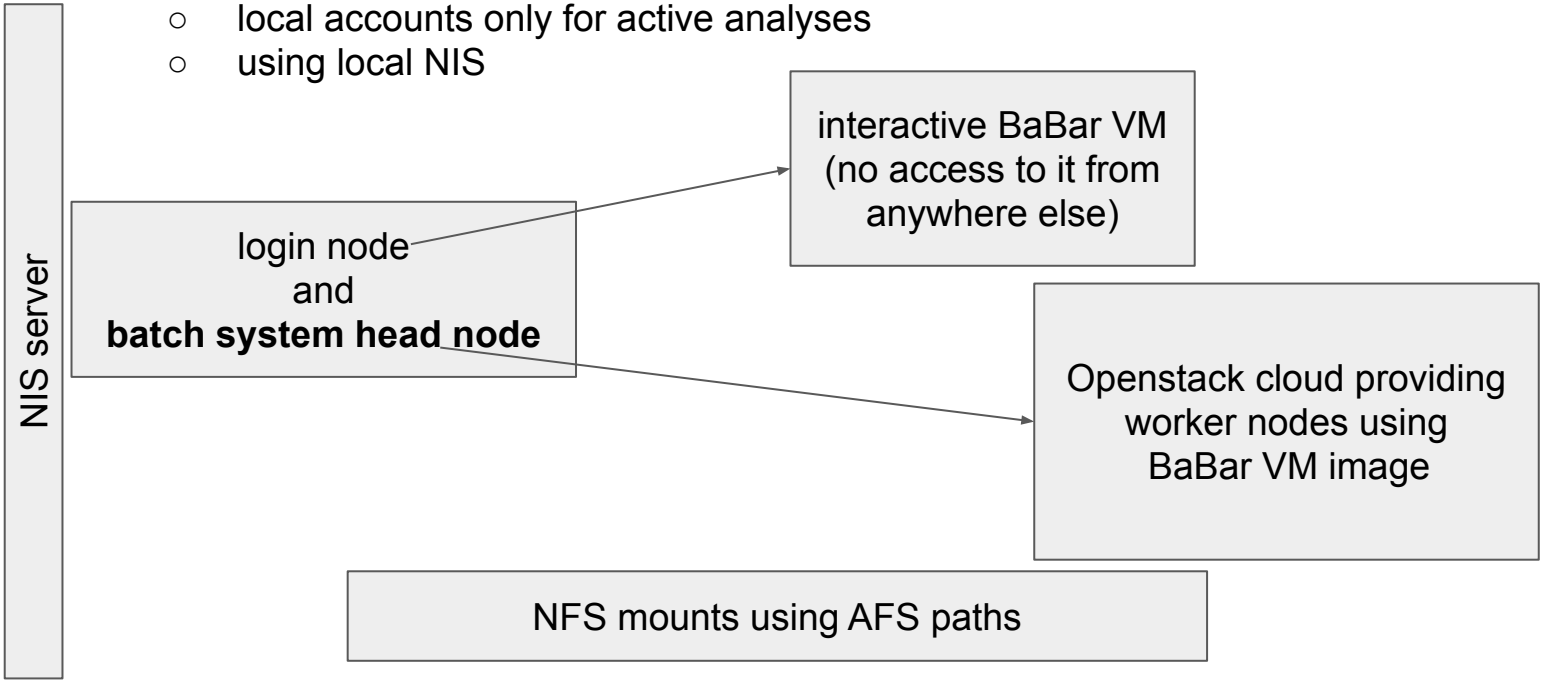
Shared File System

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



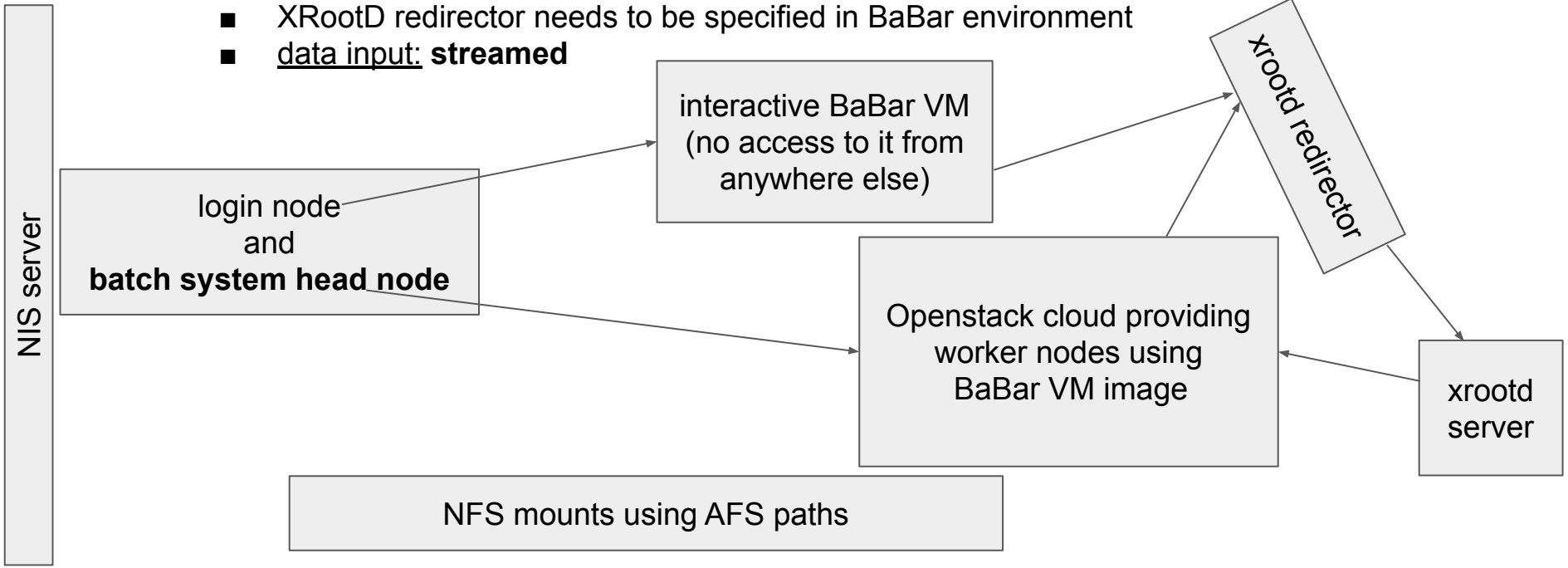
User Accounts and Management

- 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

- GridKa offered to store data and MC files from the latest processing run (AllEvents, skims, conditions db,...) for active usage
- GridKa also continues to host the metadata db (mariadb)
- IN2P3 hosts since a long time a second copy of all BaBar data, incl. raw data, as backup (not for active usage) and agreed to continue to do that
- CERN offered to also host a copy of all data

Data Access

- Data available to analyses: ~1.5PB

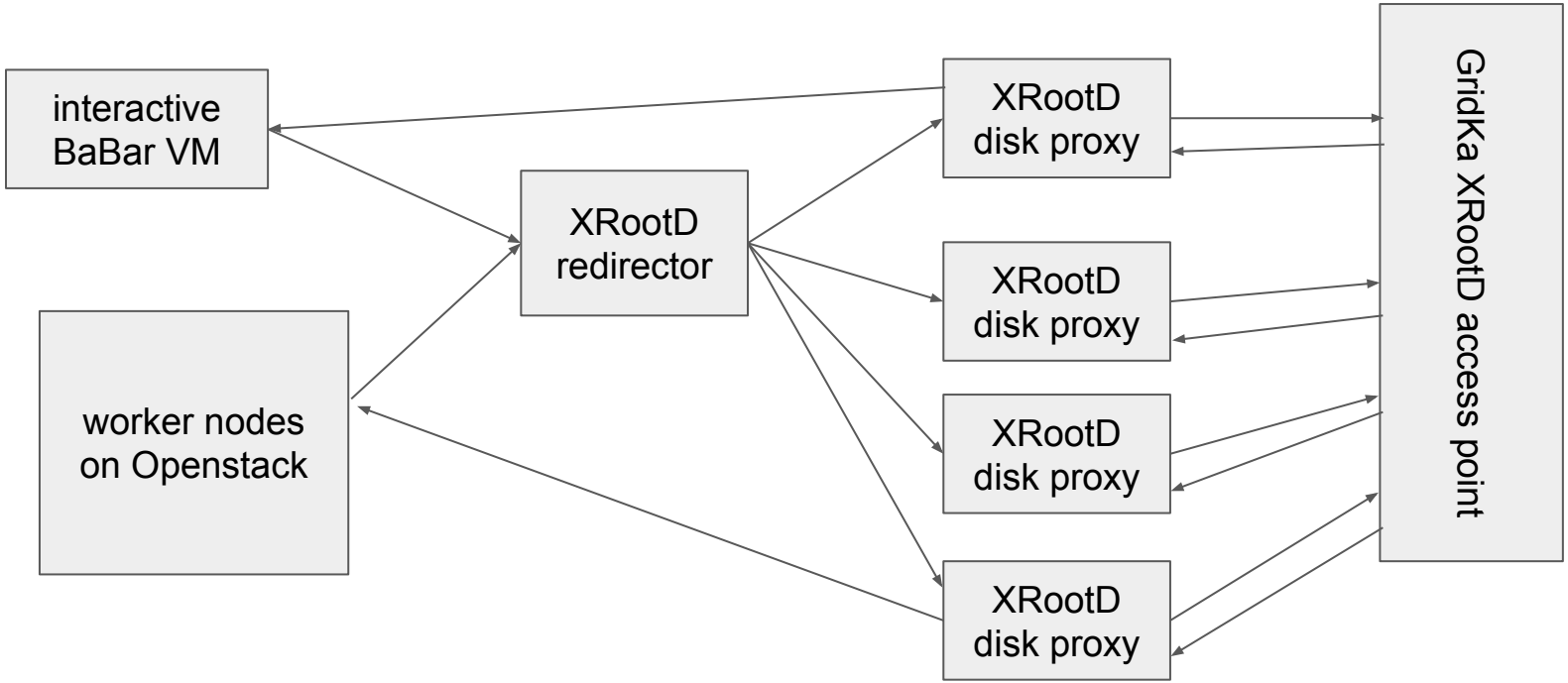
- Framework at UVic needs to access data at GridKa via streaming...
  - works surprisingly well for normal event data
    - workflow: read event, process, read event, process,...
  - but conditions data is also read via streaming
    - large amount of data each job needs to read...

Doable but very slow processing

---> use XRootD proxy system

Data Access

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



Documentation

- different systems used:
  - html web pages: in AFS within well defined directory structure, r/w rights via ACL, every BaBar user had a SLAC account; edit html files directly in AFS
  - Wiki: added ~2012 to have self contained system editable by anyone in the collaboration via web browser

- html web pages: visible to public or specific groups via .htaccess files, difficult to maintain content, for historic purpose
- Wiki: visible only to BaBar members, easy to maintain content, main BaBar documentation
- two new web servers at UVic and a single public web page (rest got access restricted to BaBar members)

Collaboration Tools

- SLAC based mailing lists ---> Caltech mailing lists
  - only created what is still needed

- old meeting agendas were HTML pages, registration based on SLAC systems

---> switch to use CERN Indico

- **Hypernews** was deeply integrated into SLAC
  - sending emails for posts to SLAC emails, notify SLAC systems in case of issues, people joining need SLAC UNIX account,... - but all content of posts in text files

---> moved Hypernews to UVic, made read-only, and removed any mailing feature -> still readable and archive of any communication happened in the past

---> replacement: CERN egoups
  - also nicely integrated with CERN Indico for accessing BaBar meetings

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

- 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

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

Summary

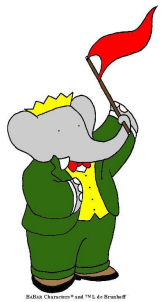
<u>To run an old and outdated analysis environment on current infrastructure:</u>

- Keep analysis and documentation framework in a well defined directory structure
- Outdated analysis environments can be preserved in VM image
- Running on clouds can make use of such VM image
- Running on clouds to not depend on specific hardware/worker node machines
- Data access via xrootd gives good choices for data server setups
- Keeping a tape backup of framework and documentation
- Keeping data backup at independent sites
- Using mirrored server infrastructure to account for old hardware

Conclusion

Running analyses in an old and outdated environment is possible and can be done safely and very well using current infrastructure solutions like clouds.

Big Thanks
to the
GridKa, CERN, IN2P3, INSPIRE, Caltech, and UVic HEP-RC groups!



Other Collaboration Tools

- Analysis documents, notes, and Analysis metadata
  - old content archived to INSPIRE
  - new documents will be added too for long term preservation

<u>new system for active analyses and management:</u>

- Google drive folder for each analysis
  - for documents and other informations
- Google sheets for metadata of each analysis
- review done using CERN egroups (each analysis has its own)
- specific folders for SpeakersBureau, PublicationBoard,...

Open Data

- making data openly available is possible, but not useful by itself
- to make use of the data one also needs
  - Analysis framework
  - Documentation
  - Communication with collaboration members

‘BaBar Associates’ open-access:

- anyone can join (== data access for anyone)
  - full access to communications and documentation tools and archives
  - analyses for publication to be done within BaBar publication framework
    - e.g. going through the full review process
  - https://babar.heprc.uvic.ca/www/join_BaBar.html

Access to BaBar framework: analysis system at UVic, BaBar-To-Go (VM) at home