Preparing experiments' software for long term analysis and data preservation

Yves Kemp, Dmitry Ozerov
DESY, Notkestrasse 85, D-22607 Hamburg
E-mail: yves.kemp@desy.de

Abstract. Preserving data from past experiments and preserving the ability to perform analysis with old data is of growing importance in many domains of science, including High Energy Physics (HEP). A study group on this issue, DPHEP, has been established in this field to provide guidelines and a structure for international collaboration on data preservation projects in HEP. This contribution aims at preparing experiments' software for long term analysis and data preservation. In a first part, we discuss the use of modern techniques like virtualization or Cloud for this purpose. In a second part, we detail the constraints of a supporting IT center for future legacy experiments. In a third part, we present a framework that allows experimentalists to validate their software against a previously defined set of tests in an automated way. We show first usage of the system, and present results gained from the experience with early-bird-users, and future adaptations to the system.

1. Introduction
In August 2009, ICFA, the International Committee for Future Accelerators, endorsed the “Study Group for Data Preservation and Long Term Analysis in High Energy Physics”, DPHEP [1]. This makes clear that High Energy Physics (HEP) embraces a topic of rising importance for all scientific data in the digital era. An overview on the DPHEP initiative is presented in these proceedings and elsewhere [2, 3, 4]. In this contribution, we will concentrate on one particular question: How to best preserve the ability to perform analysis on preserved data.
In a first part we will discuss the use of modern techniques like virtualization or Cloud for the purpose of long term analysis. In a second part, we will show some constraints of an IT center when supported experiments move from an active to a legacy phase, and how long term analysis support is best achieved from an IT center perspective. In a third part, we will present frameworks developed at DESY to help the HERA experiments achieving these goals. We will show the setup of the frameworks, first usage and results gained from early-bird-usage, and how these frameworks could evolve in the future.

2. Usage of modern virtualization based techniques
A naive approach to long term analysis could read like the following: Take the analysis environment one uses today, put it in a virtual machine, and save this virtual machine (VM) until a new analysis should be performed. The code is self-contained in this VM, one can compile new code in this VM. Data is either completely inside the VM - rather improbable for HEP -
or outside of the VM and accessible from the central storage.
While this might seem like a viable solution, several questions arise:

- Is the environment really so isolated and self-contained? Are there really no dependencies to external services like conditions database or license server?
- When the long term analysis event occurs: The old code will probably compile using the old compiler in the old OS - but will it be possible to integrate modern code written for modern compilers and modern systems?
- The VM lives in a network. Will the VM easily integrate with a future network, are there changes needed, or will this just be impossible? Will the network security group allow a potentially outdated OS to run on the public network?
- The storage probably will be outside of the VM because of the usually large amount of HEP data. Currently, often proprietary protocols like dcap or RFIO are used to access data over the network. Special libraries are needed, and a special configuration. Will the old libraries and the old configuration work with a potentially updated system?
- Hardware changes occur, some of these affect the way VMs are run. If for example the old instruction set is no longer supported by a CPU, the old instruction set must be emulated on a new CPU. This process usually is much slower, and most importantly: Using emulation of old hardware can produce different results than the native hardware itself.
- Virtualization technology itself can evolve. Is it clear that a VM of today can be understood by the virtualization layer in some decades?
- Is one VM enough for the analysis to finish in time, or will one need more than one VM? Does one need to maintain the job submission framework in addition to the VM?

The answers to these questions may vary for different cases, but it should be clear that a general usage of such a naive approach for long term analysis has too many pitfalls to take it seriously into consideration. Different solutions are possible [5].

2.1. The BaBar LTDA system
The BaBar community has fixed itself an upper timescale for long term data analysis. A system should work until 2018, allowing all active analyses to be finished and some cross-checks with superseding experiments like Belle 2 to be made [6, 7].

Because of this timescale, some constraints can be easily answered. Adverse hardware and virtualization technology changes can probably be neglected. Other constraints remain valid, like network security, data access, other external dependencies or job submission framework.

BaBar has decided to build a system based on VMs which is totally isolated from the rest of the hosting computing center (security), and which provides all necessary external dependencies and data. User access to the system is done using secure gateways. A sketch of the system can be seen in figure 1.

2.2. The DESY “pizza preservation” approach
In contrast to BaBar, the HERA experiments have no superseding experiment on the horizon. There are no concrete plans for a future $e^\pm - p$ collider. Even if another $e^\pm - p$ collider would be build like LHeC, the HERA energy range would still be unique. In any case, the HERA data and the ability to perform analysis on the HERA data must be preserved for a longer time span [10, 11, 12, 13].

Especially, the constraints neglected on purpose by the BaBar system need attention. Guaranteeing the correctness of an already known result is a bare minimum needed when planning for long term analysis. The focus here must therefore more be on preserving the recipe for a result producing system than on preserving the result producing system. This can
be compared to “pizza preservation”: If one wants to preserve a pizza for a short amount of time (some days-some month), putting it in the freezer is a valid procedure. If the task is to preserve the pizza for an extended period, like a couple of years, preserving the recipe is a valid procedure. It is not enough though: The correctness of the result must be checked at regular intervals to detect variations. Continuous checking and adaption to changing environments is an integral part of the procedure.

Virtualization is an important state-of-the-art tool for automated test procedures. DESY IT has developed a framework for enabling long term analysis of the HERA experiments based on workflows using virtualization. This framework will be presented later in these proceedings.

2.3. Using “The Cloud” for long term analysis

If virtualization techniques are used, cloud middleware can be used to steer the virtual machines and to build workflows. At first sight, it does not matter whether the cloud infrastructure resides in the computing center of the institute, or at a large (commercial) cloud provider. When large datasets are used, the location of the cloud infrastructure does matter, as data transfers into and out of the cloud present non-negligible costs and management effort.

Current cloud middleware APIs are subject to change, and HEP is not part of the API standardization bodies. Change is not in the hands of HEP collaborations. If these limitations are known clouds can be a helpful tool - but not a silver bullet per se.

2.4. Relation to other virtualization efforts in HEP

Currently, CernVM has formed an ecosystem which enables experiments to build VM images according to the particular needs of users, containing a definable OS version, a definable VO specific software stack, and an image format matching the VM technology needed by the user. Originally targeted at users developing on their laptop or desktop, the CernVM project also claims for enabling long-term data preservation [8, 9].
By its nature, CernVM is strong in recreating environments, and recreating them for different and evolving virtualization platforms. CernVM can be an important tool in a “pizza preservation” approach - but it lacks the probably most important part: continuous testing. Users of CernVM for long term analysis should therefore organize for a separate testing framework.

3. Supporting legacy experiments: An IT center view

HEP experiments are hosted by a laboratory, and usually, the IT division of this laboratory plays an active role in each experiment’s IT organization. DESY for example was and still is Tier-0 center for the HERA data, the DESY IT division ran the necessary systems for storing and analyzing the data in conjunction with the experiments. If experiments move from an active phase to a legacy phase, the relation between the IT division and the experiments might change due to several factors:

- In the active phase the experiment had a relatively large user community, if not the majority of users of a laboratory. This changes in the legacy phase: Only a small community of sporadic users is left, the majority of IT users are working on different projects.
- In the active phase of the experiment, development is ongoing, and changes can be adapted. This is problematic in the legacy phase, little or no development is done, so integrating necessary changes can be difficult.
- This also reflects on the available manpower: There are many responsible persons both in the experiment and in IT during the active phase, in the legacy phase, only a small number of people are responsible on both sides.
- In the active phase, the financial support is clear, whereas in the legacy phase the financial support is potentially unclear. This can affect long term analysis capability: Even if technically an analysis could be done, if there are no financial resources to provide the necessary hardware infrastructure no analysis can be done.

The hosting laboratory should be aware of these constraints and follow the transition from an active state to a legacy state of the experiment. This is done, for example, at DESY for the HERA experiments. It has been shown at DESY that the cooperation to such a project by especially two departments was very helpful: The IT department and the DESY library. Both departments have in common that they are organized in a “persisting” manner, in contrast to experiments which only have a limited active phase.

These lab departments should take over those parts of the experiments necessary for the whole project of long term data preservation and data analysis. This includes hosting documentation (paper, web pages, ...), digitizing and indexing information, archiving experimental data, and more. A natural prerequisite is that there is a laboratory behind the experiment, and that this laboratory shows some interest in the legacy, long-term-analysis phase of the experiment.

If the hosting experiment agrees to some kind of support, it is up to the involved partners to organize the legacy phase in an efficient way. From the view of an IT division, this might be expressed in three wishes:

- **Be as mainstream as possible.** On the level of the operating system, this means that legacy experiments should be somewhat reactive and move with other user communities as long as possible. Best would of course be to plan their software such that this move is not too difficult. Also dependencies should be as mainstream as possible, avoiding personalized versions of otherwise common tools, and not using deprecated versions.
- **Be as flexible as possible.** Things will change in the infrastructure. Server names, IP addresses, firewall rules, protocols, .... The software and the framework should be such that these changes can be adapted.
4. A framework for the DESY long term data analysis

At DESY, the HERA experiments have opted for a preservation strategy based on continuous testing. To help experiments, IT has implemented a framework which should be able to:

- Use virtualization techniques to repeatedly run well defined tests
- Perform checks against different and evolving environments
- Automatically check these results against predefined values
- Only notify when results differ from these values
- Separate the duties of IT from those of the experiment

At the basis of this framework is a workflow based on a generic, atomic test life-cycle, shown in figure 2. A request gathering, design and implementation were performed by Marco Strutz during a master thesis. More details are available in [14, 15].

The first test results show that this framework is well suited for repeated similar and isolated actions. The framework was not intended to serve as a development phase. Unfortunately, in the project, it was seen that many tests needed development, and experiments were limited by the long turnaround times. It also turned out that experiments needed dedicated storage in three categories:

- Common input to virtual machines like database files

**Figure 2.** Sketch of the generic, atomic test life-cycle.
• Transient storage for inter-VM exchange in complex workflows. An example would be that a first job in a VM compiles the software and puts it into a tarball. This tarball should then be input to further jobs in different VMs.

• Easy-access store for logs and result files.

Following the long turn-around times and the missing storage options of the first framework, we decided to use a different framework during the test development phase of the HERA experiments. This framework does not have automation and offers a simple method for file storage: Storage local to the VM.

The new framework for HERA experiments test development is based on VMs. The fast turn-around times are achieved by always-on VMs. Job submission is done via a batch-like interaction, no direct “ssh” to the VM is possible. All steps done to the VM are recorded like in a recipe. The VMs do not present any HERA specifics. It is very important that the VMs do not have access to AFS and HERA information stored there: No hidden dependencies can be used.

The current version of the framework is not final and will evolve with the complexity and needs identified when developing the tests of the experiments.

5. Summary

We have shown that virtualization technologies are important tools for long term analysis of HEP data, but that a naive utilization will not ensure an analysis in the future. Two models, one from BaBar and one from the DESY HERA experiments are shown.

We have presented some aspects of long term analysis that make the support of an IT division at a hosting laboratory easier and enhance the chance for a successful later analysis. In addition, we have presented frameworks developed at DESY to enable continuous testing and described first experiences with these frameworks.

References

[1] ICFA Study Group on Data Preservation and Long Term Analysis in High Energy Physics http://www.dphep.org

[2] DPHEP Study Group 2009 Data preservation in high energy physics Preprint arXiv:0912.0255 [hep-ex]

[3] South D 2010 Data preservation in high energy physics Preprint arXiv:1101.3186v1 [hep-ex]

[4] South D 2012 Data preservation in high energy physics CHEP 2012 conference To appear in these proceedings.

[5] Kemp Y 2009 Use of virtualization techniques for data preservation and long term analysis DPHEP workshop DESY 2009 http://www.dphep.org/sites/site_dphep/content/e20/e36625/e43750/dlpta.pdf

[6] Neal H and Cartaro T 2011 Data preservation at BaBar/SLAC 5th DPHEP Workshop FNAL 2011 http://indico.cern.ch/conferenceDisplay.py?confId=116485

[7] Cartaro C et al 2012 Implementation and use of BaBar Long Term Data Access. CHEP 2012 conference To appear in these proceedings.

[8] Buncic P et al A practical approach to virtualization in HEP The European Physical Journal Plus 126(1) 1-8 10.1140/epjp/i2011-11013-1

[9] Larsen D et al 2012 Long-term preservation of analysis software environment CHEP 2012 conference To appear in these proceedings

[10] Steder M and South D 2012 The H1 data preservation project CHEP 2012 conference To appear in these proceedings

[11] Avetysian E 2012 The HERMES data preservation project CHEP 2012 conference To appear in these proceedings

[12] Wichmann K and Malka J 2012 The ZEUS data preservation project CHEP 2012 conference To appear in these proceedings

[13] Ozerov D, Rostovtseva I and Goloubkov G Preservation of the HERA-B Collaboration heritage J. Phys. Conf. Ser. 331 (2011) 042018.

[14] Strutz M 2011 Virtualisation technologies and cloud computing in data preservation for high energy physics Master’s thesis Berlin University of Applied Sciences (HTW) Berlin (available in university library).

[15] Kemp Y Strutz M and Hessling H 2011 A validation system for data preservation in HEP ACAT 2011 Uxbridge, London to appear in Journal of Physics: Conference Series