

# DPHEP @ DESY

Current efforts in data preservation

*Karsten Schwank*, Dirk Krücker, Patrick Fuhrmann, Birgit Lewendel,  
David South

April 13, 2015



# Content

## 1 HERA

Overview

## 2 DPHEP

Areas of Data Preservation

DPHEP Activities

Strategy

Small Files Service

## 3 Summary



# HERA



- HERA was the largest particle accelerator at DESY
- Active from 1991 to 2007
- Served experiments H1, ZEUS, HERMES, HERA-B
- The world's only high-energy electron-proton collider
- Investigation of the Proton
  - ⇒ Unique Dataset



# Data Preservation

## 1 HERA

Overview

## 2 DPHEP

Areas of Data Preservation

DPHEP Activities

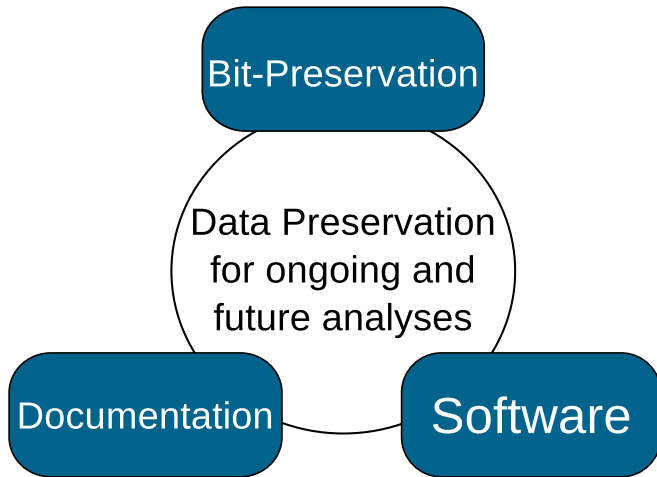
Strategy

Small Files Service

## 3 Summary



# Aspects of Data Preservation



# Activities

- Many activities over the past years
  - Documentation, accumulated by the experiments over the years, has been collected, stored and catalogued; non-digital documentation has been partly digitized.
  - Selected software is ensured to run on SL6, partly preserved in VMs
  - An ICFA study group has analyzed the various aspects of data preservation and explored various solutions.
    - Detailed status report in 2012 (see <http://arxiv.org/abs/1205.4667>)
  - DPHEP agreement signed in 2014 by many institutes including DESY
- Importance of data preservation is clearly acknowledged by all involved parties.



# Strategy

What steps are taken to address bit-preservation of HERA data?

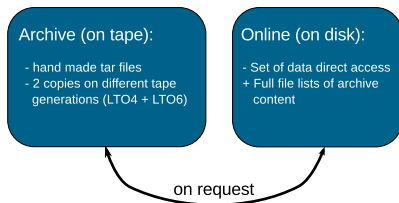
- To ensure data is preserved beyond the end of the experiments, solutions have to be provided by DESY as institute.
  - Smooth transition from experiment specific solutions to institutional solution
  - IT department and Library
- Short- and long-term availability - **2-fold strategy**
  - Assure the long-term availability of the data → **tape archive**
  - Support on-going analyses work and keep still used data easily accessible → **disk pools**
- Our main effort during the last months was related to bit-preservation



# Storage Structure DPHEP on DESY dCache

How is this 2-fold strategy set up?

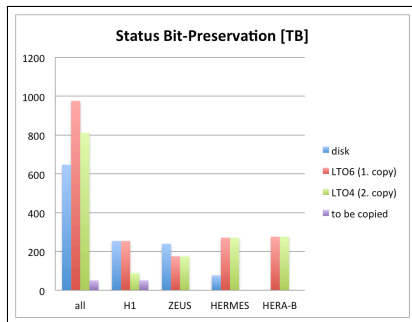
- Archive part ( $\approx 1.2PB$ ):
  - All data identified for preservation will be available in 2 redundant tape copies
  - The archive is not generally accessible, i.e. requires manual intervention
- Online part ( $\approx 650TB$ ):
  - *dCache* disk pools where experiment data is available
  - Can be mounted r/o where needed (e.g., batch systems, Workgroup Server)





# Status of Archive part

What is the current status of the Archive part?

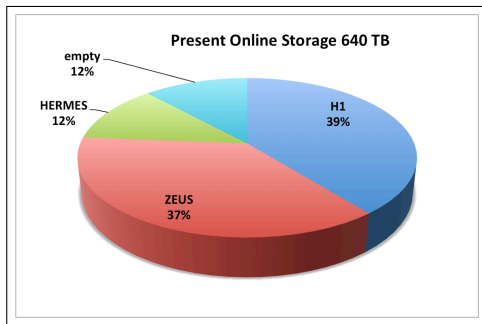


- About 75% of all data already on tape
- 50TB still on H1 *dCache*



# Status of Online part

What is the current status of the Online part?



- Already almost filled → experiments have to restrict themselves



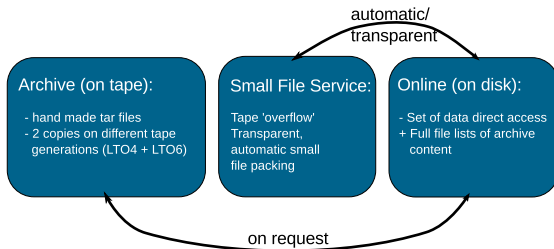
# Motivation for Extended Online Storage

## Data Management View:

- Wasted (expensive) disk space for seldomly used data
- Additional work load related to handling of millions of small files

## User View:

- As much data online as possible
- Avoid (slow) interaction with storage admin



# Extended Online Storage

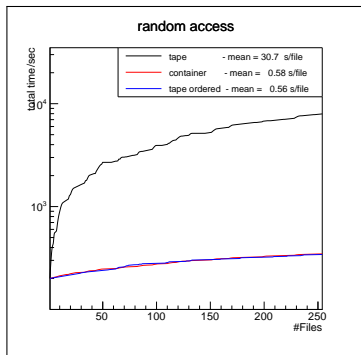
Increase online space by automatic data swaps to and from tape with external *dCache* service:

- “Swapping” is automatically handled by *dCache*
- Files are transparently packed and unpacked from container files using *dCache*'s tape interface
- Can be attached to any recent *dCache* version
- Allows easy direct access to potentially all data w/o manual intervention
- All file metadata is stored in *dCache* → service failure does not prevent data from being accessed



# Performance - Tape

Large container files allow optimized restoring of files from tape



**Figure :** The service improves random access to files on tape by a factor of  $\approx 50$  and provides data rates close to reading files from tape sequentially



# Performance - Unpacking

Overhead introduced by unpacking files from containers

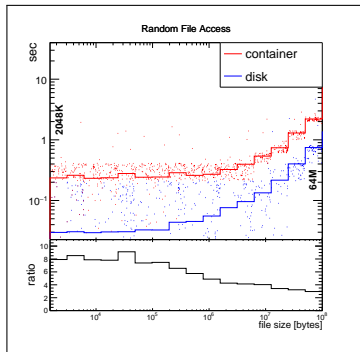


Figure : The overhead per file introduced by unpacking is  $\approx .2s$

⇒ Details on poster “Transparent handling of small files with dCache”  
(Poster Session B #284)



# Summary

- HERA dataset is unique → preservation for ongoing and future analyses is important
- Preservation is done by *DESY* on an institutional level
- Majority of files has already been archived
- We take care of the data in a 2-fold strategy (Online and Archive)
- Currently evaluating *Small File Service* to be used for Extended Online Storage

