DESY/HERA data preservation

*** Preliminary collection of material, 6.3.2017 ***

Achim Geiser, DESY Hamburg, Germany

ZEUS

for DPHEP workshop CERN, Geneva, 14. 3. 2017

HERA



HERA material
 Personal remarks and
 relation to other projects

Remarks to this collection

- A coherent HERA talk will be assembled from the IT/H1/HERMES/ZEUS material presented
- Some parts of the ZEUS material presented here are for internal DESY documentation only
- Beyond the actual HERA material, the talk will (if you allow) also include some personal remarks in a more general DPHEP context. They will be clearly labeled as such.



from PRC open session, 11. 5. 2015

HERA data preservation

- DESY is co-founding member of Collaboration Agreement for the DPHEP project
 Supported by ICFA (May 2014, already reported last PRC)
- HERA workshop at DESY, future physics topics (Nov. 2014)
 World HEP community
- Status of bit preservation



Workshop:

- What do the HERA data still have to say and how are they relevant to other facilities?
- two days with lively discussions and almost 30 presentations <u>https://indico.desy.de/event/futurehera</u>
 - ~ 70 participants, both experimentalists and theorists from across the globe



Future Physics with HERA Data for Current and Planned Experiments

13 November 2014 DESY, Hamburg, Germany

A bright future for HERA physics

A workshop in DESY looked at what the data from HERA can still offer for experiments, now and in the future.

VOLUME 55 NUMBER 2 MARCH 2015

PENSICS AT THE TERA TERA Retractit Allance

HERA papers, past and future

ZEUS

H1





kindly provided by Dirk

HERA Bit-Preservation

- The HERA data archive is finalized
- The online (disk) store is filled and 2 tape copies are written
- Small additions to the heritage data are possible - details about the procedure will be defined in agreement with the experiments
 - First cases now
- The content of the archive and the procedures how to add and restore data had been documented
- Restoring data from the tape archive to the online store had been successfully exercised



For the Statistics Enthusiasts: final storage content

| H1 | HERMES | ZEUS | HERA-B | Туре |
|--------|---------|---------|--------|------------------------|
| 983398 | 6557725 | 1183157 | 846059 | single files |
| 11111 | 9179 | 7318 | 4110 | archive (tar) files |
| 810316 | 774032 | 1182941 | 0 | files online |
| 359 | 57 | 239 | 0 | TiB online |
| 464 | 581 | 368 | 392 | # LTO4 (800G) tapes |
| 134 | 174 | 104 | 110 | #LTO6 (2.4T) tapes |
| 430 | 358 | 239 | 276 | TiB on LTO4/LTO6 tapes |

- In nuce: 1.3 PB and 10 million files
- In addition there are 10 TB data of polarimeter data/simulations included



kindly provided by Stefan

H1 and DPHEP status



February 2017

- H1 publications
- H1 computing and Monte Carlo production
- Reviving old software: GKS for look and event display

Paper production summary

- Paper production 2016: two papers
 - QCD instanton search
 - Jet production at low Q²
- 2017: hope to have five publications
 - Diffractive D* in DIS
 - α_s at NNLO from jets
 - H1/ZEUS F2b,F2c
 - ρ meson diffractive photoproduction
 - H1 electroweak fit

Analyses marked in red: using preserved H1 data



A. Geiser, DESY DPHEP meeting

H1 computing and MC production

- H1 computing
- All SL5 machines switched off
- H1 public web-server: now can only display static web pages

Many H1 collaborative tools are based on cgi-scripts for accessing oracle.

Work-around: for critical tools we have a local web-server running, using port 8080 which is not reachable outside the firewall.

Longer term: have to seek for another solution.

 Recent progress in MC production: production on the NAF system is working



2017 production using DESY batch system for new analyses



H1 software

- Customized GKS libraries installed on SL6 \rightarrow it works
- H1 software depending on GKS:
 - Look (histogramming and analysis framework) by
 V.Blobel → working
 - "old" event display: supports some features not available on the new (root-based) tools
 - Special event display for driftchamber analysis (hit-level)

Not critical for "standard" H1 analyses but nice to have



ZEUS + general material

Collected from

- ADMP workshop, CERN, Geneva, 30. 6. 2016
- ZEUS/IsoQuant meeting, 31. 8. 2016
- ZEUS Collaboration meeting, 24. 3. 2015

Summary to Cristi for ICFA meeting

- Data storage (including MC, ~250 TB) with multiple copies in different geographical locations (2 tape + disk copies at DESY (IT), extra copy at MPP) has been finalized in early 2015. Data and MC (flat ROOT ntuples) are mainly being accessed on the National Analysis Facility at DESY, and can also be accessed via the Grid at MPP. The web pages and documentation have been archived at DESY in 2014.
- Analysis of the preserved ZEUS data and MC is alive and well. We (ZEUS) have published 3 papers in 2015, 4 papers in 2016, and will probably have between 2 and 4 in 2017.
- New MC has successfully been generated via MPP, building on earlier preparation work at DESY and is used in current ZEUS publications.
- In the wake of the workshop on Future Physics with HERA data in November 2014 two new groups are joining ZEUS to exploit synergies between HERA and EIC, as well as HERA and heavy ion physics, so we expect further diversified physics output.



What is ZEUS?

- International Particle Physics Experiment which recorded high energy electronproton collisions at the world's (so far) unique lepton-proton collider HERA at DESY in Hamburg, Germany
- Physics data taking: 1992-2007
- one of main physics goals: measure structure of the proton to ~10⁻¹⁸ m, i.e. 1/1000 of proton size ("X ray" of proton with electrons) use e.g. in measurements of Higgs properties at LHC
- also well suited to study general QCD and electroweak physics



Why to preserve HERA data?



Why to preserve HERA data?



Synergy with current experiment:

<u>LHC</u>

- LHC collides protons on protons
- detailed knowledge of proton structure is crucial for many LHC physics topics, e.g. for measurement of Higgs boson properties
- in general, many common physics topics

see also

- HERA-LHC workshops, DESY and CERN
- workshop on Future Analysis of HERA data,

DESY, November 2014, https://indico.desy.de/conferenceDisplay.py?confId=10523





many EIC topics common with HERA



 informal discussions with EIC members on possible common analyses of HERA data started **ZEUS physcis papers**





2017 update

6.03.17

A. Geiser, DESY DPHEP meeting

What do ZEUS data look like?

| Zeus Run 1 | (Simrun 59924) E | vent 208 | date: | 4-06-2006 time: 00:06:30 |
|---------------------------|--------------------------|----------------------------|---------------------------|--------------------------|
| E=55 GeV | E _t =9.44 GeV | E-p _z =2.98 GeV | E _r =52.8 GeV | E _b =2.07 GeV |
| E _r =0.138 GeV | p _t =2.72 GeV | p _x =-2.66 GeV | p _y =0.583 GeV | p_=52.1 GeV |
| phi=2.93 | t _f =3.08 ns | t _b =-0.371 ns | t _r =-100 ns | t_g=2.97 ns |



event display from "Common Ntuple"

ZEUS

complicated data format and content: for useful analysis, need significant expert knowledge + documentation + guidance how to use it 6.03.17 A. Geiser, DESY DPHEP meeting 23

DPHEP data preservation levels

| Preservation Model | Use case |
|---|---|
| 1. Provide additional documentation | Publication-related information search |
| 2. Preserve the data in a simplified format | Outreach, simple training analyses -> education |
| 3. Preserve the analysis level software | Full scientific analysis based on existing |
| and data format | reconstruction |
| 4. Preserve the reconstruction and simulation software and basic level data | Full potential of the experimental data |

Table 3: Various preservation models, listed in order of increasing complexity.

• ZEUS: level 3 (data and existing Monte Carlo (MC) data), level 4 (additional Monte Carlo data)

H1 and HERMES: level 4

Publicly available information on DPHEP and ZEUS data preservation

| | HEP 6 records found | Search took 0.15 seconds. |
|---|--|--|
| Eile Edit View Higtory Bookmarks Tools Help Image: find d × FCC - Future W SLAC Nat G pentaqua W Heisenbe Image: find d Image: find d W Heisenbe Image: find d Image: fi | Status Report of the DPHEP Collaboration: A Global Effort for Sus DPHEP Collaboration (Silvia Amerio (INFN, Padua) <i>et al.</i>). Feb 17, 2015. 60 pp. DPHEP-2015-001 DOI: 10.5281/zenodo.46158 e-Print: arXiv:1512.02019 [hep-ex] PDF References BibTeX LaTeX(US) LaTeX(EU) Harvmac EndNote CERN Document Server ; ADS Abstract Service Detailed record - Cited by 2 records 2. The DPHEP Study Group: Data Preservation in High Energy Physic DPHEP Study Group Collaboration (David M. South for the collaboration). 2013. 6 pp. Published in PoS ICHEP2012 (2013) 536 Conference: C12-07-04 Proceedings References BibTeX LaTeX(US) LaTeX(EU) Harvmac EndNote Proceedings of Science Server; ink to Fulltext | tainable Data Preservation in High Energy Physics cs + DPHEP@DESY documents |
| find data preservation and CN ZEUS Brief f find ("Phys RevLett.105"; :: more Bisplay results: istest first desc. or rank by - Z records found Image: Structure HEP 2 records found 1 The ZEUS data preservation project ZEUS and DESY DPHEP Group Collaborations (J. Malka (DESY) for the collaboration). 20 DO: 10.1109/INSSMIC.2012.6551468 Conference: C12-10-29, p.2022-2023 Proceedings References BibTeX LaTeX(US) LaTeX(EU) Harvmac EndNote Detailed record 2 Conference: C12-10-29, p.2022-2023 References BibTeX LaTeX(US) LaTeX(EU) Harvmac EndNote Detailed record Ciference: C12-05-21.3 Proceedings References BibTeX LaTeX(US) LaTeX(EU) Harvmac EndNote Display/1742-6596/396/2/022033 Conference: C12-05-21.3 Proceedings References BibTeX LaTeX(US) LaTeX(EU) Harvmac EndNote Staled record - Cited by 1 record | Detailed record 3. DPHEP: From Study Group to Collaboration DPHEP Collaboration (David M. South (DESY) for the collaboration). Sep 30, 2013. 6 pr Published in PoS DIS2013 (2013) 267 Conference: C13-07-18 Proceedings e-Print: arXiv:1309.7868 [hep-ex] PDE References BibTeX LaTeX(US) LaTeX(EU) Harvmac EndNote ADS Abstract Service: Proceedings of Science Server: Link to Fulltext Detailed record 4. Status Report of the DPHEP Study Group: Towards a Global Effort DPHEP-2012-001, FERMILAB-PUB-12-878-PPD e-Print: arXiv:1205.4667 [hep-ex] PDE References BibTeX LaTeX(US) LaTeX(EU) Harvmac EndNote CERN Document Server; ADS Abstract Service; OSTI Information Bridge Server; Detailed record - Cited by To Tecords 5. Data Preservation in High Energy Physics DPHEP Study Group Collaboration (David M. South (DESY) for the collaboration). Jan 2 Published in J.Phys.Conf.Ser. 331 (2011) 012005 CHEP-2010 DOI: 10.1088/1742-6596/331/1/012005 Proceedings of plenary talk given at Conference: C10-10-18.4 Proceedings e-Print: arXiv:1101.3186 [hep-ex] PDE References BibTeX LaTeX(US) LaTeX(EU) Harvmac EndNote ADS Abstract Service Detailed record - Cited by 6 records < | p. Tor Sustainable Data Preservation in High-Energy Physics Fermilab Library Server (fulltext availables, tink to Fulltext 2011. 10 pp. INSPIRE itself is a "level 1 data preservation lab Library Server (fulltext availables, tink to Fulltext project |

6.03.17

A. Geiser, DESY DPHEP meeting



DPHEP portal:

- <u>http://hep-project-dphep-portal.web.cern.ch</u>
- ZEUS web page:
- <u>http://www-zeus.desy.de/</u>
- information on ZEUS far from perfect

(manpower ..., in case of availability conflict, content/useability takes preference over (organisation of) documentation)

... but we are proud of what we achieved \bigcirc

see also presentation A. Verbytskyi at DIS2016 conference https://indico.desy.de/contributionDisplay.py?contribId=176&sessionId=7&confId=12482

and ZEUS MPI web page https://www.zeus.mpp.mpg.de/

Data Preservation Challenge:

How to organize the Management?



Challenge: What is the "Data"?

- "Data" = recorded events, simulated events,
 - + related software, knowledge, and documentation
 - original ZEUS data format and core software from 1990's
 - maintenance of software, simulation and analysis framework
 needed ~4 FTE/year (experiment) + IT
 - e.g. porting from SL4 to SL5 took about 2 years
 - -> not sustainable long term
- -> go for simplified ZEUS data format:
 - "Common Ntuples" = flat ROOT ntuples
 - almost no dedicated software maintenance needed
- -> for new simulation: freeze software and run compiled executables in virtualized environment
 - see also https://wwwzeus.mpp.mpg.de

managed at MPI

^{...} not achieved

Analog and digital archive

- analog archive in DESY library
- ZEUS technical notes digitized on INSPIRE (via DESY library)
- frozen plain html documentation web pages (DESY web office)
 Introduction to ZEUS analysis

Introduction

General

Are you the new member of the ZEUS community and want to start a new analysis??? Or have you already started analyzing data but still have a lot of questions? We hope that on these web pages you will find an answer to most of your questions.

Analysis

Tools

 knowledge preservation in "human neural networks" (ZEUS collaboration)



Other

Common Ntuple analysis model

 ZEUS Common Ntuple: Motto: keep it simple! flat (simple) ROOT-based ntuple (same format as PAW ntuple converted with h2root) containing high level objects (electrons, muons, jets, energy flow objects, ...) as well as low level objects (tracks, CAL cells, ...)

• Well tested !

almost all recent ZEUS papers based on Common Ntuples

• "Easy" to use

6.03.17

several recent ZEUS papers based on results produced by Master students from remote institutes, using resources at DESY: analysis on DESY NAF/BIRD computing farm

PhD students can produce a ZEUS paper within only a fraction of their PhD time (e.g. ~6 months - 1 year)

| date: | 4-06-2006 time: 00:06:30 |
|---------------------------|--------------------------|
| E _f =52.8 GeV | E _b =2.07 GeV |
| p _y =0.583 GeV | p_=52.1 GeV |
| t _r =-100 ns | t _g =2.97 ns |







"When will the project be finally done?"

• my answer:

(usually hard to digest for host labs, funding agencies, committees ...)

if taken serious, a data preservation project will never be "done", unless and until one gives up on useability of the data

Challenge: Bit preservation

HERA Bit-Preservation

 at DESY: common approach for all three HERA experiments

status 06/2015 (now complet<mark>e</mark>)



| Size | of a | lata | set | S compiled | by D. Z | otkin/A.G. | |
|------------|------------|-----------------|------|-----------------------|-----------|-------------------------|---------|
| Root files | (officiall | <u>y preser</u> | ved) | units: Tb | (status 4 | .9.13) | ZEUS |
| HERA II | v02 | v06 | v08 | HERA I VO8 | total | | |
| Data | 1.9 | 5.2 | 7.0 | 1.7+1. | 17. | | |
| MC | 10.5 | 64.0 | 70. | 4.8 <mark>+4</mark> . | 153. | +30 for fu [.] | ture MC |

~ 100 million inclusive DIS events (Q²>5 GeV², triggered almost bias-free)

~ 100 million semi-inclusive photoproduction events (mainly via p_T >4 GeV dijet trigger) smaller sets of more specialised triggers/samples (e.g. heavy flavours, vector mesons, ...) ~ equal sample sizes for e+, e-, righthanded/lefthanded polarisation

~ 4 billion MC events, for almost any analysis

generation of additional MC samples might be possible (see talk A. Verbytskyi)

can technically read/analyze full ZEUS data set on NAF/BIRD at DESY within ~1 day (for even faster access, many analyzers produce their own mini-ntuples for analysis) 6.03.17 A. Geiser, DESY DPHEP meeting 34

Challenge: Virtualisation

- generation of new MC requires detector simulation
 - can not avoid use of full fledged "legacy" ZEUS software from the 1990's
- porting to new operating systems not realistic for ZEUS with realistic manpower
- -> run existing software in "Virtual box" environment with "old" operating system

(implemented and maintained by MPI based on content development work at DESY)

interface to new generators via HEPMC data format

Challenge: sustainability

- Bit preservation: long term agreement with DESY IT, ensured ⁽³⁾
- ZEUS data and existing MC analysability: simplified format minimizes maintenance. Maintenance ensured through semi-official long term manpower (0.1 FTE), ensured [©]
- ZEUS event display relies on unofficial temporary manpower -> will break down as soon as person leaves
- Situation of additional ZEUS MC simulation (no support at DESY due to lack of manpower) unclear

Conclusions and Outlook

HERA data are scientifically unique and worth preserving !

- 9 years after end of data taking in 2007, thanks to data preservation, ZEUS scientific output continues at a significant rate, for very little cost (expect ~10% of total scientific output, if long term sustainability is achieved). Made possible through substantial support by collaboration, host lab (DESY), and external institutes!
- ZEUS has successfully implemented its long term "Active Data Management Plan" for data preservation, worked out 2006-2012, and in full operation since Jan 2015. Integrated into DPHEP strategy.
- Bottleneck: Long term data preservation needs long term manpower: don't need "much" (~O(‰) of original project investment, spread over 20 years), but 0 will not do ...



managment

Data management in particle physics

this part is about plans for long term Active Data Management !

actual ZEUS experience (also other experiments):

 passive data management (just storing the data somewhere) will not work long term,

Active data management is crucial

- Data must include metadata, and preservation of software, knowledge, and useability
- Management must include manpower needed for long term management, both at IT and user level

ZEUS

"Active Data Management Plan"

wasn't called like that at the time, but

 a three page "bottom-up ZEUS ADMP"
 can be found in the
 2012 DPHEP study group document (see previous slide)

... and we conceptually implemented more or less exactly what we planned ③ with some practical variations (of course at that time it was already half way done)

of ZEUS papers vs. time









2017 update

6.03.17

A. Geiser, DESY DPHEP meeting



How to measure the success?

personal measure used on previous slide:

expected # of additional scientific papers total # of scientific papers

compared to

estimated integrated cost of data preservation estimated integrated total cost of project

arguable - but is there a better one?

HERA vs. LEP vs. Tevatron



Some ingredients for success of actual project

- Make sure you start the 'user mode' well (>~ 2 years) before the temporary manpower ends (-> need to be able to fix "hickups" !) ③ ZEUS: user data preservation mode gradually started 2011-2013
- Ensure strong support of host lab or other funding body during the 'long term benefit' phase 🙂 ZEUS: scientific support OK, long term manpower/minimal funding support more difficult than expected/hoped for
- Make sure to get the necessary dedicated long term manpower ZEUS need: ~2/3 short term (and funding!) going along with this support $\overline{\mathbf{S}}$ ~1/3 long term (~20 year integral) people understand the need to maintain storage, networks and tape vaults, and to provide some minimal CPU power, but rarely understand the (size of the) manpower need for knowledge preservation, software preservation, and user support ...

-> this is the main point upon which many of the past projects personal have failed and many of the current projects risk to fail

view

"EU data principles"

- Discoverable: ZEUS and DPHEP web pages, conferences, workshops, ...
- Accessible: ZEUS data are not (yet) open data

(would need more manpower/funding)

but "Free Access to ZEUS Data" programme for PhD students and physicists (e.g. EIC), data accessible at DESY, + on data grid via MPI

- Intelligible: bottleneck! currently OK, but would strongly profit from more manpower (keep experts involved!)
- Assessable: quality/reproduceability is ensured by the ZEUS collaboration
- Useable: Yes! (papers based on these data continously being published) all recent ZEUS papers are open access (DESY rule)

add:

 Sustainable: bottleneck! Can't do without some funding, in particular for long term manpower! "librarian" attitude to preservation could be useful!

<u>Comparison ZEUS data</u> preservation/CMS open data

- example for synergy!
- both use ROOT-based higher level data format ("level 3") with very similar content (CMS data format somewhat more complicated)
- both use virtual machine environments for software environment / preservation (well-developed for CMS open data analysis, under development by MPI for ZEUS for additional MC generation and analysis, not needed for baseline data and MC analysis at DESY)
- CMS more advanced in "transparent" access (open to general public)
- ZEUS more advanced in formal preservation of secondary information
- current practical synergy is that I work on both, can easily switch between the two (very similar conceptual approaches), and one profits from the other

Internal slides for this meeting

How to analyze ZEUS data at DESY?

- (for additional possibilities at MPI see contributions A. Verbytskyi)
 need:
 - interest in some physics topic 😳



- agreement with ZEUS management and DESY to obtain
- ZEUS user account at DESY
 - -> access to NAF/BIRD analysis farm via
 - ZEUS NAF server (can log on from remote)
- basic knowledge of ROOT
 - (no special ZEUS software to learn!)
- basic knowledge of particle physics

Possible HERA collider physics topics

as discussed at Future Analysis with HERA data workshop

- BSM:
 - Provide standard candles against which new physics searches can be calibrated
- Proton structure:
 - FL combination, integration of high x results into PDF fit, finalize heavy flavour combinations and fit, improved transverse momentum dependent PDFs, investigation of low x phenomenology, ...
 - -> understand the proton, understand QCD, provide detailed descriptions for other colliders
 - Are we starting to hit the nonperturbative limit?
 - Can we make further decisive measurements from existing data?
 - Can we achieve improved theoretical interpretations from existing results?
 - Can statements about new physics at high scales be made from the low energy data?

Diffraction and DVCS

- Finalize inclusive diffractive measurements, make them more differential
- Finalize measurements of elastic vector meson production and compare to improved theory models and to other experiments
- Measure elastic scalar model production, test odderon hypothesis
- Finalize measurements of DVCS 6. 03. 17 A. Geiser, DESY DPHEP meeting

Possible HERA collider physics topics

as discussed at Future Analysis with HERA data workshop

- Jets:
 - Finalize (ZEUS) measurements, combine,
 - make more differential measurements, event shape measurements,
 - apply NNLO theory, remeasure alphas
- Hadronic final states:
 - Study multiparton interactions and other nonperturbative effects
 - (re)measure photon structure
 - (re)measure QCD instanton production
 - Search for exotic resonances
 - Complete total gamma-p cross section
- Heavy Flavours:
 - Intrinsic charm
 - NNLO measurements of c- and b-masses
 - Multi-differential heavy flavour cross sections
 - More cross section combinations
 - Improved measurements of charm fragmentation functions

Offline and



D. Kruecker/IT

data preservation at DESY

Offline report J. Malka (summary here)

General ZEUS Offline coordination (Janusz Malka) ended in December 2014 as announced. Big thanks again to Janusz!
-> this report (developments since last collaboration meeting in September 2014) might be the last of this kind.

DESY data preservation report -> this contribution

MPI data preservation report -> see contribution A. Verbytskyi/MPI



- General ZEUS Offline coordination (J. Malka): ended
- Data preservation: hardware and storage at DESY:
 D. Krücker (DESY/IT), ended, now via "tickets"@DESY IT

Account management: M. Wing (UCL), official Remaining desktops: I. Martens (DESY/CMS) inofficial (good will)

 Data preservation: DESY/NAF analysis access to existing root data, support for interpretation of CN variables

A. Geiser (DESY/CMS), semi-official (open ended) + O. Zenaiev (DESY/CMS) inofficial, temporary, for event display

- Software preservation system (SP system) (DESY/IT + J. Malka/K, Wichmann): abandoned
- Web archive/documentation at DESY:
 K. Wichmann (DESY/FLA), official until may 2016(?) (+ student support, ended)
- MC generation + bookkeeping at DESY ended December 2014 as forseen (no manpower) attempt to add new MC generated at MPI might be failing due to lack of dedicated manpower
- Data preservation at MPI: A. Verbytskyi (official, temporary)

DATA - format

slide J. Malka Sept. 2014

- the Data and MC samples available in Root Common Ntuples format are documented on the common ntuple web page: (zeusdp) http://www-zeus.desy.de/ZEUS_ONLY/analysis/comntp/
 - * the DPHEP data are the CN from versions

| * v02d | data | 11 |
|-------------|----------|-------|
| * V02e,v02f | mc | П |
| * v06b | mc | П |
| * v06d | data | П |
| * v07a | data, mc | I |
| * v08b | data, mc | I, II |

483631 root files and 175T capacity

 The log-files from the production needs to be copied – ongoing done. Some files lost.

DATA - storage

slide J. Malka Sept. 2014

- * The DPHEP disk online storage plus two archive tapes copy
 - The data were copied from the HERA dCache to DESY-SE (DPHEP now read-only no more stageing
 - * Test check the number of events per file and checkvar = (sum of total CAL energy)/ (no of track) and sum over one file
 - * Cross check HERA dCache vs on-line DPHEP storage
 - Tests were run over all CN which should be keep for dpHEP (vo2,vo6,vo7,vo8)
 - All CN are available and readable on on-line dpHEP storage completed
 - The list of archived files agree with contented of online part completed

Documentation

slide J. Malka Sept. 2014

- * Our archive webserver will be http://zeusdp.desy.de/
 - the first archiving done year ago needs an updated

is

the active webpages will be attached to afs

K. Wichmann, V. Myronenko, O. Turkot, I. Martens

- * Internal notes, thesis achieved in INSPIRE
- completed The documentation of the CN samples is ongoing
- The documentation of the ZEUS software is ongoing has
- * INDICO will replace ZEMS

Hardware

slide J. Malka Sept. 2014

The NAF2/BIRD cluster will be available for users gateway servers nafhh-zeuso1 and nafhh-zeuso2 are installed

is

* The online DPHEP storage is accessible from NAF2/BIRD

have been

* The workgroup servers will be switched off on 31Deceber2014 hopefully, you did (actually, slightly later)

please secure you data, source codes, etc

Analyses beyond 2014

slide J. Malka Oct. 2014

- Data in Common Ntuple format: <u>http://zeusdp.desy.de/ZEUS_ONLY/analysis/comntp/comntp_main.php</u>
 — Stored in dp-hep online storage
- Computing on NAF2 batch facility

 have been
 All workgroups servers will be shut down
- ROOT as analyses software
- Documentation on archive webserver:
 - <u>zeusdp.desy.de</u> -> now again <u>www-zeus.desy.de</u>
- ZEUS active webpages on AFS
- Repository for historical analysis/MC code on web server
- Set of historical generator steering files on web server

ZEMS

- Zeus Electronic Meeting Management System was replaced by Indico
- The final update of zems archive has been done: searchable!
 <u>http://zeusdp.desy.de/ZEUS_ONLY/meetings/zems/index.html</u>
- The submit/updated access to online zems has been blocked

| > C 🗋 zeusdp.desy.de/ZEU | JS_ONLY/meetings/zems/index.html | | |
|---|----------------------------------|--|--|
| ZERS | Select a group you want to view: | | |
| | ANALYSIS FORUM | | |
| | COLLAB. MEETING | | |
| Zeos Electronic Meeting Management System | ZEUS-PHYSICS | | |
| | DIFFRACTION AND LOW X | | |
| ZEMS is an integrated system used for collaborative | EXOTICS & SFEW | | |
| meetings. | FL | | |
| | HEAVY FLAVOUR | | |
| WSC ENTRY WSC | QCD/HFS MEETINGS | | |
| | PHYSICS COORDINATOR | | |
| | MONDAY WEEKLY MEETINGS | | |
| | OFFLINE | | |
| | TRACKING | | |
| | 6M TAGGER | | |
| | LUMI | | |
| | CALORIMETER & HES | | |
| | GFLT | | |

Many thanks to Ingo



Overview Bit-preservation

Dirk Krücker DPHEP meeting 26.2.15

6.03.17

A. Geiser, DESY DPHEP meeting

Data preservation for the HERA Experiments

PRC slide D. Kruecker Oct. 2014

- DPHEP Collaboration is now officially installed
 - MoU signed by : CERN, DESY, IN2P3, MPI Munich, HEP, IHEP, KEK
- Legacy dataset
 - Defined by now see next slide
 - but can still be extended
 - Transfer into long term format ~80%
- Documentation
 - Web servers
 - Archive server available for all experiments
 - Transition ongoing
- Software preservation
 - Migration to SL6 successful
 - New software preservation system "old" one was lost ...
 - Integration into IT infrastructure for long term "survival"

Zeus

slide D. Kruecker Feb. 2014

- All PAW ntuples (data&MC) are copied to the new store
 - 4.3TB data / 56TB old MC /2.5TB new MC
 - not yet written to tape but you can read them check!
 - 4 problematic files out of 208201 withstand the reading attempts
 - Only 4 MC files see appendix for file names
 - lost files are documented in: /pnfs/desy.de/dphep/online/"exp"/DPHEP/lists/00README.lost (does not exist yet for H1 and ZEUS)
- 1/2 million log files
 - 2 tapes seems to be unreadable with 298 + 1263 log files
 - cannot be mounted by yet unknown reason
 - (3 tapes with this kind of failure compared to ~9000 tapes in the robot
 - bad coincidence that 2 of these 3 belongs to Zeus)