Big Data & Long-Term Preservation

Challenges, Opportunities and Solutions

Jamie.Shiers@cern.ch

Computing in Atmospheric Sciences 2013

International Collaboration for Data Preservation and Long Term Analysis in High Energy Physics
Overview

• Why so much data?

• How long to preserve it?

• Challenges, Opportunities and Solutions(?)
WHY SO MUCH DATA?
(HOW MUCH DATA?)
\[ \sqrt{s} = 7 \text{ TeV} \int Ldt = 0.02 \text{ fb}^{-1} \quad \text{Apr 18, 2011} \]

**ATLAS** Preliminary

H\(\rightarrow\gamma\gamma\) channel

**ATLAS** Higgs Candidates

- Data
- Background-only

Events / GeV

\[ M_{\gamma\gamma} \text{ [GeV]} \]

Data - Fit

-200
0
200
400
600
800
100
110
120
130
140
150
160
Castor data written 2010-12

**2010-2012 Data written:**
- Total ~19 PB in 2012
- Close to 4 PB/month now

- 15 PB in 2010
- 23 PB in 2011

Data rates in Castor increased
- 3-4 GB/s input
- ~15 GB/s output
CERN has ~100 PB archive
HOW LONG MUST WE KEEP IT?
(HOW LONG SHOULD WE KEEP IT?)
The LHC roadmap to fully exploit the physics potential

- **2009**: LHC startup, $\sqrt{s} = 900$ GeV

- **2010**
  - $\sqrt{s} = 7\sim 8$ TeV, $L = 6 \times 10^{33}$ cm$^{-2}$ s$^{-1}$, bunch spacing 50 ns
  - ~20-25 fb$^{-1}$

- **2013**
  - LS1: Go to design energy, nominal luminosity

- **2014**
  - $\sqrt{s} = 13\sim 14$ TeV, $L = 1 \times 10^{34}$ cm$^{-2}$ s$^{-1}$, bunch spacing 25 ns
  - ~75-100 fb$^{-1}$

- **2018**
  - LS2: Injector and LHC Phase-1 upgrade to ultimate design luminosity

- **2019**
  - $\sqrt{s} = 14$ TeV, $L = 2 \times 10^{34}$ cm$^{-2}$ s$^{-1}$, bunch spacing 25 ns
  - ~350 fb$^{-1}$

- **2022**
  - LS3: HL-LHC Phase-2 upgrade, IR, crab cavities?

- **2023**
  - $\sqrt{s} = 14$ TeV, $L = 5 \times 10^{34}$ cm$^{-2}$ s$^{-1}$, luminosity levelling
  - ~3000 fb$^{-1}$

**2030?**
The LHC roadmap to fully exploit the physics potential

- **2009**: LHC startup, $\sqrt{s} = 900$ GeV
  - $\sqrt{s} = 7\sim8$ TeV, $L = 6 \times 10^{33}$ cm$^{-2}$ s$^{-1}$, bunch spacing 50 ns
  - 15 PB/yr

- **2010**: Go to design energy, nominal luminosity
  - $\sqrt{s} = 13\sim14$ TeV, $L = 1 \times 10^{34}$ cm$^{-2}$ s$^{-1}$, bunch spacing 25 ns
  - 25 PB/yr
  - ∼20-25 fb$^{-1}$

- **2011-2013**: LS1
- **2014-2015**: LS2
- **2016**: Injector and LHC Phase-1 upgrade to ultimate design luminosity
  - $\sqrt{s} = 14$ TeV, $L = 2 \times 10^{34}$ cm$^{-2}$ s$^{-1}$, bunch spacing 25 ns
  - 50 PB/yr
  - ∼75-100 fb$^{-1}$

- **2018**: LS3
- **2019-2021**: HL-LHC Phase-2 upgrade, IR, crab cavities?
  - 100 PB/yr
  - ∼350 fb$^{-1}$

- **2022-2023**: ...
- **2030?**: HL-LHC Phase-2 upgrade, luminosity levelling
  - ∼3000 fb$^{-1}$
Beyond HE-LHC

- First studies on a new “80” km tunnel in the Geneva area
  - 42 TeV with 8.3 T using present LHC dipoles
  - 80 TeV with 16 T based on Nb$_3$Sn dipoles
  - 100 TeV with 20 T based on HTS dipoles
“CERN should undertake design studies for accelerator projects in a global context, with emphasis on proton-proton and electron-positron high-energy frontier machines.”

Project

Kick-off meeting: 11th Nov. 2013 (Daresbury)

Study: VHE-LHC with TLEP

Kick-off meeting: February 2014 (CERN)
Data Summary

• Total data volume for LHC: $O(1\text{EB})$

• Scientific programme: a few decades

• This is much shorter than required / desired in other disciplines

• Also much shorter than funding agencies push for
2020 Vision for LT DP in HEP

• **Long-term – e.g. LC timescales**: disruptive change

  – By 2020, all archived data – e.g. that described in Blueprint, including LHC data – easily findable, fully usable by designated communities with clear (Open) access policies and possibilities to annotate further
    ➢ Via DPHEP Portal; to be setup as from now...

  – Best practices, tools and services well run-in, fully documented and sustainable; built in common with other disciplines, based on standards

  ➢ Vision achievable, but we are far from this today
## DPHEP models of HEP data preservation

<table>
<thead>
<tr>
<th>Preservation Model</th>
<th>Use Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Provide additional documentation</td>
</tr>
<tr>
<td>2</td>
<td>Preserve the data in a simplified format</td>
</tr>
<tr>
<td>3</td>
<td>Preserve the analysis level software and data format</td>
</tr>
<tr>
<td>4</td>
<td>Preserve the reconstruction and simulation software as well as the basic level data</td>
</tr>
</tbody>
</table>

### Use Case Details

- **Documentation**: Publication related info search
- **Outreach**: Outreach, simple training analyses
- **Technical Preservation Projects**: Full scientific analysis, based on the existing reconstruction
- **Technical Preservation Projects**: Retain the full potential of the experimental data

> These are the original definitions of DPHEP preservation levels from the 2009 publication

- Still valid now, although interaction between the levels now better understood

> Originally idea was a progression, an inclusive level structure, but now seen as complementary initiatives

> Three levels representing three areas:

- **Documentation**, **Outreach** and **Technical Preservation Projects**
**LHC Open Access Policies**

- Simplified example from LHCb. *(CMS similar)*
- *Can we harmonize policies:*
  - Across experiments? Across labs?

<table>
<thead>
<tr>
<th>Level 1 data: Published results</th>
<th>All scientific results are public. ...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 2 data: Outreach and education</td>
<td>[Samples made public.] The data are for educational purpose only, not suitable for publication</td>
</tr>
<tr>
<td>Level 3 data: Reconstructed data</td>
<td>LHCb will make reconstructed data (DST) available to open public; 50% 5 years after data is taken, 100% after 10 years. Associated software will be available as open source, together with existing documentation. Publications must include disclaimer.</td>
</tr>
<tr>
<td>Level 4 data: Raw data</td>
<td>[Not directly accessible to collaboration] But must still be preserved!</td>
</tr>
</tbody>
</table>
LTDP: Component Breakdown

• Can break this down into three distinct areas
  – (OAIS reference model is somewhat more complex: this is a zeroth iteration)

• “Archive issues” (BE)

• Digital Libraries & “Adding Value” to data (FE)

• “Knowledge retention” – the Crux of the Matter
Where to Invest – Summary

Tools and Services, e.g. Invenio: could be solved. (2-3 years?)

Archival Storage Functionality: should be solved. (i.e. “now”)

Support to the Experiments for DPHEP Levels 3-4: must be solved – but how?
Archive Issues (back-end)

✓ We (HEP) have significant experience of 100PB+ distributed data stores

✓ Plan is to coordinate long-term “bit preservation” issues via HEPiX

✓ And with other disciplines e.g. via IEEE MSST, Research Data Alliance, 4C, APA etc.

✗ Sustainable models for long-term multi-disciplinary data archives still to be solved

➢ H2020 funding targeted for this
Digital Libraries

✓ Significant investment in this space, including multiple EU (and other) funded projects

✓ No reason to believe that the issues will not be solved, nor that funding models will not exist, e.g. adapted from “traditional” libraries

✓ Related topics: “linked data”, “adding value to data” – again with projects / communities

✔ Should work closely with these projects / communities, not start new initiatives
Documentation projects with INSPIRE

- Internal notes from all HERA experiments now available on INSPIRE
  - Experiments no longer need to provide dedicated hardware for such things
  - Password protected now, simple to make publicly available in the future

- The ingestion of other documents is under discussion, including theses, preliminary results, conference talks and proceedings, paper drafts, ...
  - More experiments working with INSPIRE, including CDF, D0 as well as BaBar
## Summary of information from the (pre-LHC) experiments

<table>
<thead>
<tr>
<th></th>
<th>BaBar</th>
<th>H1</th>
<th>ZEUS</th>
<th>HERMES</th>
<th>Belle</th>
<th>BESIII</th>
<th>CDF</th>
<th>DØ</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>End of data taking</strong></td>
<td>07.04.08</td>
<td>30.06.07</td>
<td>30.06.07</td>
<td>30.06.07</td>
<td>30.06.10</td>
<td>2017</td>
<td>30.09.11</td>
<td>30.09.11</td>
</tr>
<tr>
<td><strong>Type of data to be preserved</strong></td>
<td>RAW data Sim/rec level Data skims in ROOT</td>
<td>RAW data Sim/rec level Analysis level ROOT data</td>
<td>Raw data Sim/rec level Analysis level ROOT data</td>
<td>RAW data Sim/rec level</td>
<td>RAW data Sim/rec level</td>
<td>Raw data Rec. level ROOT files (data+MC)</td>
<td>Raw data Rec. level ROOT files (data+MC)</td>
<td></td>
</tr>
<tr>
<td><strong>Data Volume</strong></td>
<td>2 PB</td>
<td>0.5 PB</td>
<td>0.2 PB</td>
<td>0.5 PB</td>
<td>4 PB</td>
<td>6 PB</td>
<td>9 PB</td>
<td>8.5 PB</td>
</tr>
<tr>
<td><strong>Desired longevity of long term analysis</strong></td>
<td>Unlimited</td>
<td>At least 10 years</td>
<td>At least 20 years</td>
<td>5-10 years</td>
<td>5 years</td>
<td>15 years</td>
<td>Unlimited</td>
<td>10 years</td>
</tr>
<tr>
<td><strong>Current operating system</strong></td>
<td>SL/RHEL3 SL/RHEL 5</td>
<td>SL5</td>
<td>SL5</td>
<td>SL3 SL5</td>
<td>SL5/RHEL5</td>
<td>SL5</td>
<td>SL5 SL6</td>
<td>SL5</td>
</tr>
<tr>
<td><strong>Languages</strong></td>
<td>C++ Java Python</td>
<td>C++ Fortran Python</td>
<td>C++ Fortran</td>
<td>C++ C++ Fortran</td>
<td>C++</td>
<td>C++ Python</td>
<td>C++</td>
<td></td>
</tr>
<tr>
<td><strong>Simulation</strong></td>
<td>GEANT 4</td>
<td>GEANT 3</td>
<td>GEANT 3</td>
<td>GEANT 3</td>
<td>GEANT 3</td>
<td>GEANT 4</td>
<td>GEANT 3</td>
<td>GEANT 3</td>
</tr>
<tr>
<td><strong>External dependencies</strong></td>
<td>ACE CERNLIB CLHEP CMLOG Flex GNU Bison MySQL Oracle ROOT TCL XRootD</td>
<td>CERNLIB FastJet NeuroBayes Oracle ROOT</td>
<td>ROOT</td>
<td>ADAMO CERNLIB ROOT</td>
<td>Boost CERNLIB NeuroBayes PostgreSQL ROOT</td>
<td>CASTPR CERNLIB CLHEP HepMC ROOT</td>
<td>CERNLIB NeuroBayes Oracle ROOT</td>
<td>Oracle ROOT</td>
</tr>
</tbody>
</table>
Software Strategies

• A 3-pronged approach is being considered:
  
  – Validation frameworks to (semi-)automate continuous migrations
  – Virtualisation tools to preserve complete environments during LHC lifetimes (decades)
  – Software techniques to help design and implement sustainable software

• Given the (very) long lifetime of the LHC, we will have time + opportunity to evaluate pros & cons
  
  – e.g. during LS2, LS3 etc.
Where are we now?

1. **Initial** (chaotic, ad hoc, individual heroics) – the starting point for use of a new or undocumented repeat process.

2. **Repeatable** – the process is at least documented sufficiently such that repeating the same steps may be attempted.

3. **Defined** – the process is defined/confirmed as a standard business process, and decomposed to levels 0, 1 and 2 (the last being Work Instructions).

4. ** Managed** – the process is quantitatively managed in accordance with agreed-upon metrics.

5. **Optimizing** – process management includes deliberate process optimization/improvement.
## Data Preservation Maturity Model

<table>
<thead>
<tr>
<th>Level</th>
<th>Metric</th>
<th>Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Reproducible results by “citizen scientists”</td>
<td>Desired(?) by funding agencies: people able to reproduce an analysis should be awarded “a degree” – beyond what can realistically be afforded?</td>
</tr>
<tr>
<td>3</td>
<td>Reproducible results where consumer ≠ producer and outside immediate community</td>
<td>Stronger demonstration of long-term preservation. Knowledge stored is sufficient for physicist outside immediate community to reproduce results</td>
</tr>
<tr>
<td>2</td>
<td>Reproducible results where consumer ≠ producer but within same “larger community”, e.g. LHC (ATLAS / CMS; CDF / D0, ...)</td>
<td>Highly desirable for “minimal” long-term preservation. “Knowledge” stored is sufficient for a physicist from a different collaboration (but within same overall programme) to reproduce results</td>
</tr>
<tr>
<td>1</td>
<td>Reproducible results where consumer = producer</td>
<td>Required during lifetime of collaboration</td>
</tr>
<tr>
<td>0</td>
<td>N/A</td>
<td>Data is lost: logically or physically. This is probably the reality for the bulk of pre-DPHEP experiments (and even some of those??)</td>
</tr>
</tbody>
</table>

- Scale (complexity) is probably “exponential”
<table>
<thead>
<tr>
<th>Level</th>
<th>Metric</th>
<th>Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Reproducible results by “citizen scientists”</td>
<td>Desired(?) by funding agencies: people able to reproduce an analysis should be awarded “a degree” – beyond what can realistically be afforded?</td>
</tr>
<tr>
<td>3</td>
<td>Reproducible results where consumer ≠ producer and outside immediate community</td>
<td>Stronger demonstration of long-term preservation. Knowledge stored is sufficient for physicist outside immediate community to reproduce results</td>
</tr>
<tr>
<td>2</td>
<td>Reproducible results where consumer ≠ producer but within same “larger community”, e.g. LHC (ATLAS / CMS; CDF / D0, …)</td>
<td>Highly desirable for “minimal” long-term preservation. “Knowledge” stored is sufficient for a physicist from a different collaboration (but within the overall programme) to reproduce results</td>
</tr>
<tr>
<td>1</td>
<td>Reproducible results where consumer = producer</td>
<td>Required during lifetime of collaboration</td>
</tr>
<tr>
<td>0</td>
<td>N/A</td>
<td>Data is lost: logically or physically. This is probably the reality for the bulk of pre-DPHEP experiments (and even some of those??)</td>
</tr>
</tbody>
</table>
Conclusions

• Working through international coordination and standardisation bodies, we (HEP) are offering experience, tools and services...

• And expecting to benefit significantly from others in areas where we are less advanced

• Open Data and Data Preservation have arrived – we need to work together on sustainable solutions

• Join us at the RDA plenary 2 in Washington (Preservation Infrastructure Interest Group)
BACKUP
Revitalisation of JADE software

- started in 1995 at RWTH Aachen (P. Movilla-Fernandez; Diploma- and PhD thesis), proceeded until ~2003 at MPP.

- conversion, translation, partly rewriting of Fortran-IV, Mortran, Sheltran, assembler routines

- complete installation on IBM RS6000 AIX platform using xlf compiler

- successfully revitalised and validated entire JADE core software:
  - reconstruction software
  - simulation software
  - event display and JADE graphics package (now in colour!)

→ generation of full-simulation MC events, using modern MC generators plus the experience from LEP
Lessons to be learned....

- keeping data and software from old experiments for future (improved and new) analyses is absolutely neccessary and beneficial
- especially vital for the time long after initial analyses terminated!
- especially vital for processes and energies not being repeated at future projects
- in order to minimise efforts and to separate physics from archeology, several requirements should be met:
  - data and software formats must comply to sound, internationally accepted and widely supported standards
  - data format and software must be concisely and consitently documented
  - same formats and archiving procedures for different projects and experiments
  --> archiving & re-use of data & software must be planned while experiment is still in running mode!
  - proper and sustainable archiving must be assured (at accel. labs plus 1 loc. outside)
  - dedicated manpower to assure transfer of data and software to new platforms and hardware will be vital
THE STANDARD MODEL

### Fermions

<table>
<thead>
<tr>
<th>Quarks</th>
<th>Leptons</th>
</tr>
</thead>
<tbody>
<tr>
<td>u up</td>
<td>ν_ε electron neutrino</td>
</tr>
<tr>
<td>c charm</td>
<td>ν_μ muon neutrino</td>
</tr>
<tr>
<td>t top</td>
<td>ν_τ tau neutrino</td>
</tr>
<tr>
<td>d down</td>
<td>e electron</td>
</tr>
<tr>
<td>s strange</td>
<td>μ muon</td>
</tr>
<tr>
<td>b bottom</td>
<td>τ tau</td>
</tr>
</tbody>
</table>

### Bosons

<table>
<thead>
<tr>
<th>Force carriers</th>
</tr>
</thead>
<tbody>
<tr>
<td>γ photon</td>
</tr>
<tr>
<td>Z Z boson</td>
</tr>
<tr>
<td>W W boson</td>
</tr>
<tr>
<td>g gluon</td>
</tr>
</tbody>
</table>

Higgs* *Yet to be confirmed

Source: AAAS

BEFORE!