CERN

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

Long-Term Sustainability
A User (Support) View

EOSC Pilot: Science Demonstrator

Jamie.Shiers@cern.ch

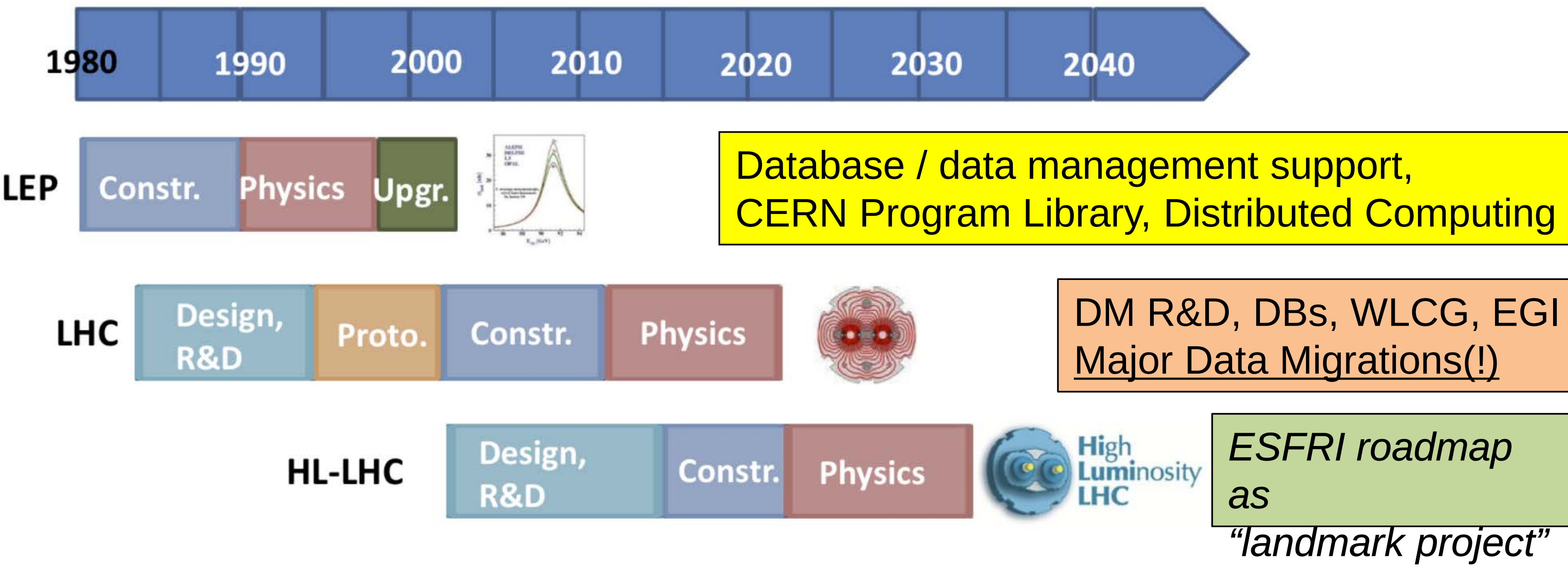
EOSC pilot
The European Open Science Cloud for Research Pilot Project

Slides available at https://indico.cern.ch/event/643419/

CERN

WLCG
Worldwide LHC Computing Grid

LEP / (HL-)LHC Timeline



- Robust, stable services over **several decades**
- Data preservation and re-use over **similar periods**
- "Transparent" and <u>supported</u> **migrations**

"Data Preservation" Demonstrator

- Goal is to demonstrate "best practices" regarding data management and their applicability to LTDP + "open" sharing + re-use
  - ✓ **PIDs for data & meta-data stored in TDRs;**
  - ✓ **DOIs for documentation;**
  - ✓ **S/W + environment.**

- Equivalent to CERN Open Data Portal but using "open" – i.e. non-HEP – solutions
  - **These all exist and are "advertised" in some form**
  - But there are "questions" around:
    Services; Resources; Long-Term Support (& Migration)…
  - **As well as Cost of Entry / "Ownership"**

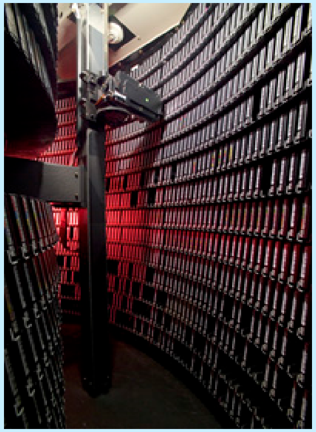
See www.eiroforum.org/downloads/20170425_federated-scientific-data-hub.pdf

Example Services – LTDP

Service	HEP	Non-HEP	Issues
Trustworthy DR	CERN CASTOR+EOS (ISO 16363)	EUDAT (?) (DSA / WDS)	How to get access to even modest resources?
PID / DOI systems			"Long-term" support; availability of services
Digital Library	CERN Document Server, INSPIRE-HEP **(Invenio-based)**	B2SHARE, Zenodo **(Invenio-based)**	CERNLIB documentation example (20 years)
Software + Environment (+build system)	CVMFS, CernVM	Ditto	***"Software without environment is just bad documentation"***

- **For a user (community) to go "shopping around" to find the right services, resources and support is a (major?) challenge / impediment**
- **More (and more complex) services needed to support data processing, distribution and analysis (full data lifecycle=WLCG4LHC)**

What is (HEP) data? (And its not just “the bits”)



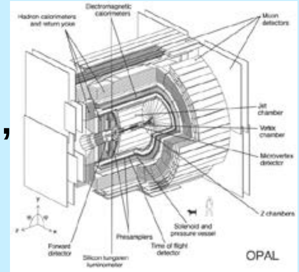
Digital information

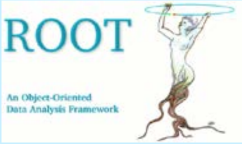
The data themselves, volume estimates for preservation data of the order of a few to 10 EB

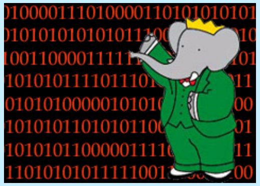
Other digital sources such as databases to also be considered

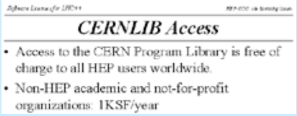
Software

Simulation, reconstruction, analysis, user, in addition to any external dependencies



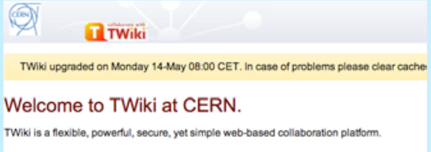






Meta information

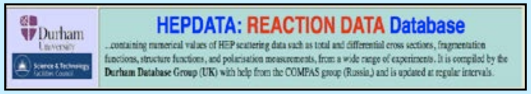
Hyper-news, messages, wikis, user forums..



Publications

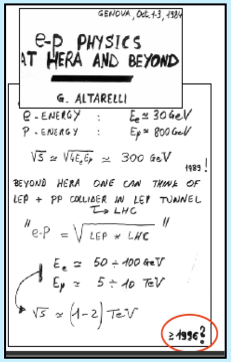
arXiv.org

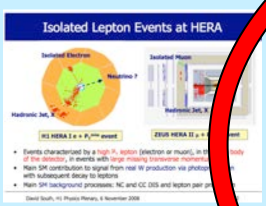
INSPIRE

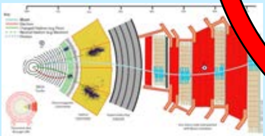


Documentation

Internal publications, notes, manuals, slides







Expertise and people

DESY

User requirements / expectations

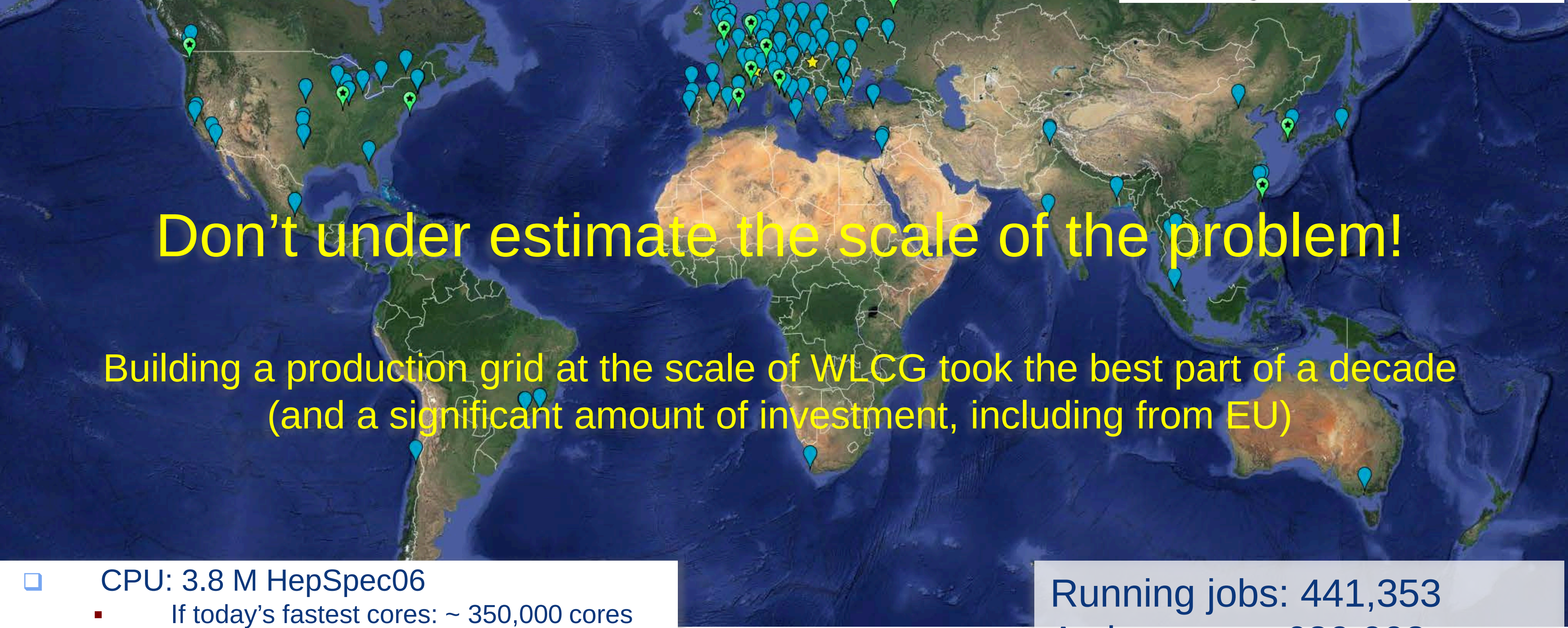
- (Large) user requirements often exceed available resources / budgets (and existing resources typically **fully utilised**)
- ➢ **Negotiation phase to converge**

- Service expectations (e.g. max 10' downtime) quasi-impossible to achieve
- ➢ **Focus on response targets, critical services and reporting metrics**
- Regular operations meetings de-fuse situations before they arise

- **How to scale these "solutions" to large numbers of communities in an EOSC?**
- **Community-based support, e.g. for ESFRIs, probably needed**

- ➢ **WLCG could be a successful model to look at**

The Worldwide LHC Computing Grid



October 2016:
-63 MoU's
-167 sites; 42 countries
-Tier0, Tier1s & Tier2s
-O(1) , O(10), O(100)

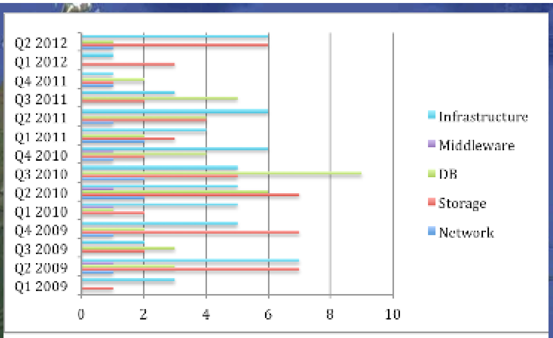


Figure 1 - Service Incidents by Area

Don't under estimate the scale of the problem!

Building a production grid at the scale of WLCG took the best part of a decade
(and a significant amount of investment, including from EU)

- CPU: 3.8 M HepSpec06
  - If today's fastest cores: ~ 350,000 cores
- Disk 310 PB
- Tape 390 PB

Running jobs: 441,353
Active cores: 630,003
Transfer rate: 35.32 GiB/sec

WLCG Service Challenges

- **As much about people and collaboration as about technology**
- **Getting people to provide a 24 x 7 service for a machine on the other side of the planet for no clear reason was going to be hard!**
- Regional workshops – both motivational as well as technical – plus daily Operations Calls
- In a grid, ***something*** is broken all of the time!
- Clear KPIs, "critical services" & response targets: **measurable improvement in service quality despite ever increasing demands**

Targets for response, intervention and resolution based on severity. Monitored regularly – *not guarantees!*

DMPs, the EOSC and ESFRIs

- An EOSC must support **multiple disciplines**

- Therefore, we need a ***lingua franca*** i.e. someway of getting them to talk together
  - And / or to the service providers!

- **IMHO, DMPs could provide just that!**
  - Even though guidelines would need to be broadened to cover data acquisition, processing, distribution and analysis in more detail!

- DMP w/s for ESFRI(-like) projects proposed: to be re-scheduled now that EOSC goals / plans more clear

Benefits of collaboration: LTDP

1. The elaboration of a clear "**business case**" for long-term data preservation
2. The development of an associated "**cost model**"
3. A common view of the **Use Cases** driving the need for data preservation
4. Understanding how to address Funding Agencies requirements for **Data Management Plans**
5. Preparing for **Certification** of HEP digital repositories and their long-term future.

Director Generals' Viewpoints

- *Software/Computing should not* ***limit*** *the detector performance and LHC physics reach*
- *the Software must be* ***easy-to-use and stable***
- ***not to hinder*** *the fast delivery of physics results (and a possible early discovery …)*

CHEP 2004, Interlaken

- To find the Higgs you need the Accelerator, the Detectors and **the Grid!**

"Higgs discovery day", CERN, 2012

Services are (just) services

- No matter how fantastic our { TDRs, PID services, Digital Library, Software repository } etc is, they are there to support **the users**

- **Who have to do the <u>really</u> hard work!**
  - **E.g. write the software, documentation, acquire and analyse the data, write the scientific papers**

- Getting the degree of public recognition as at the Higgs discovery day was a **target KPI**!

~30 years of LEP – what does it tell us?

- Major migrations are **unavoidable** but hard to **foresee**!
- **Data** is not just "**bits**", but also **documentation, software + environment + "knowledge"**
  - "**Collective knowledge**" particularly hard to capture
    - Documentation "refreshed" after 20 years (1995) – now in Digital Library in PDF & PDF/A formats (was Postscript)
- Today's "**Big Data**" may become tomorrow's "**peanuts**"
  - 100TB per LEP experiment: **immensely challenging** at the time; now "trivial" for both CPU and storage
  - With time, **hardware costs** tend to zero
    - O(CHF 1000) per experiment per year for archive storage
  - **Personnel costs** tend to O(1FTE) **>> CHF 1000!**
    - Perhaps as little now as 0.1 – 0.2 FTE per LEP experiment to keep data + s/w alive – no new analyses included

See DPHEP Workshop on "Full Costs of Curation", January 2014: https://indico.cern.ch/event/276820/

ODBMS migration – overview (300TB)

- **A triple migration!**
  - Data <u>format</u> and <u>software</u> conversion from Objectivity/DB to Oracle
  - Physical <u>media</u> migration from StorageTek 9940A to 9940B tapes
- **Took ~1 year to prepare; ~1 year to execute**
- **Could never have been achieved without extensive system, database and application support!**
- Two experiments – many software packages and data sets
  - **COMPASS** <u>raw event data</u> (300 TB)
    - Data taking continued after the migration, using the new Oracle software
  - **HARP** <u>raw event data</u> (30 TB), <u>event collections</u> and <u>conditions data</u>
    - Data taking stopped in 2002, no need to port event writing infrastructure
  - In both cases, the migration was during the "lifetime" of the experiment
  - System integration tests validating read-back from the new storage

Open Science: A 5-Star Scale?

- **We have a 5-star scale for Open Data**
  - **Sir Timothy Berners-Lee**

- **We have a proposed 5-star scale for FAIR data management (+TDRs)**
  - **Peter Doorn and Ingrid Dillo**

- **How about a 5-star scale for "Open Science: Open to the World"?**
  - **The EOSC**

What are the right metrics?

- **As easy to use as Amazon?**

- **Cheaper (and better) than doing it in-house?**

- **A majority of ESFRIs use it as their baseline?**

➢ ***"To find dark matter, you need the EOSC"***?

“Data Preservation” Demonstrator

- Goal is to demonstrate “best practices” regarding data management and their applicability to LTDP + “open” sharing + re-use
  - ✓ **PIDs for data & meta-data in TDRs;**
  - ✓ **DOIs for documentation;**
  - ✓ **S/W + environment.**

- Equivalent to CERN Open Data Portal but using “open” – i.e. non-HEP – solutions
  - These all exist and are “advertised” in some form
  - But there are “questions” around: Services; Resources; Long-Term Support (& Migration)…
  - As well as Cost of Entry / “Ownership”

CERN

See www.eiroforum.org/downloads/20170425_federated-scientific-data-hub.pdf

“Data” Preservation in HEP

- The data from the world’s particle accelerators and colliders (HEP data) is both **costly** and **time consuming** to produce

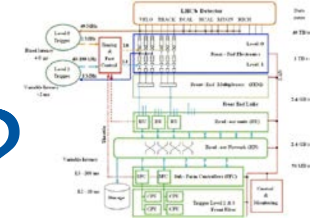
- HEP data contains a wealth of **scientific potential**, plus high value for **educational outreach**.

- Many data samples **are unique**, it is essential to preserve not only the data but also the full capability to reproduce past analyses and perform new ones.

➢ **This means preserving data, documentation, software and "knowledge".**

Requires different (additional) services (resources) to those for analysis

What Makes HEP Different?



- We **throw away** most of our data before it is even recorded – "triggers"
- Our detectors are **relatively stable** over long periods of time (years) – not "doubling every 6 or 18 months"
- We make "**measurements**" – not "**observations**"
- Our projects typically last for **decades** – we **need** to keep data usable during at least this length of time
- We have **shared** "data behind publications" for more than 30 years… (HEPData)

Measurements are repeatable: observations are not

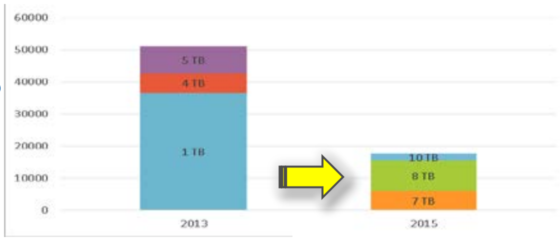
CERN Services for LTDP

1. State-of-the art "**bit preservation**", implementing practices that conform to the ISO 16363 standard
2. "**Software preservation**" - a key challenge in HEP where the software stacks are both large and complex (and dynamic)
3. Analysis **capture and preservation**, corresponding to a set of agreed Use Cases
4. Access to **data behind physics publications** - the HEPData portal
5. An **Open Data portal** for released subsets of the (currently) LHC data
6. A **DPHEP portal** that links also to data preservation efforts at other HEP institutes worldwide.

- **These run in production at CERN and elsewhere and are being prototyped (in generic equivalents) in the EOSC Pilot**

Bit Preservation: Steps Include

- Controlled media **lifecycle**
  - **Media kept for 2 max. 2 drive generations**



- Regular media **verification**
  - When tape written, filled, every 2 years…
- **Reducing** tape mounts
  - Reduces media wear-out & increases efficiency
- Data **Redundancy**
  - For “smaller” communities, a 2nd copy can be created: separate library in a different building (e.g. LEP – **3 copies at CERN**!)
- **Protecting** the physical link
  - Between disk caches and tape servers
- Protecting the **environment**
  - Dust sensors! (Don’t let users touch tapes)

Constant improvement: reduction in bit-loss rate: 5 x 10^{-16}

See “The Lost Picture Show: Hollywood Archivists Can’t Outpace Obsolescence”
IEEE Spectrum

LTDP Conclusions

- **As is well known, Data Preservation is a Journey and not a destination.**
- Can we capture sufficient "**knowledge**" to keep the data usable **beyond** the lifetime of the original collaboration?
- Can we prepare for **major migrations**, similar to those that happened in the past? (Or will x86 and Linux last "forever")
- For the HL-LHC, we may have **neither** the storage resources to keep all (intermediate) data, **nor** the computational resources to re-compute them!

➢ **You can't share or re-use data, nor reproduce results, if you haven't first preserved it (data, software, documentation, knowledge)**

Big Data: From LEP to the LHC to the FCC

From LEP (1989 – 2000) to the LHC (2009 – 2035) to the "FCC"

- **"Big data" from hundreds of TB to hundreds of PB to (perhaps) hundreds of EB**
- **FCC-ee option: "repeat" LEP in just 1 day!**
- **FCC-hh: 7 times LHC energy, 10^{10} Higgs bosons**