



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

Archiving and Preservation "in the Cloud" – Outlook

Joint WLCG – HSF Workshop
March 2018

Jamie.Shiers@cern.ch



Introduction

- We talk in HEP of 3 pillars of data preservation:
 1. The data itself – “bits” hopefully stored in a TDR
 2. The documentation (which can be quite complex) – INSPIREHEP et al
 3. Software & environment – CVMFS / CernVM
- Corresponding solutions / services mature and widely deployed
- Can we run a demo / prototype:
 1. In the European Open Science Cloud?
 2. In a public – private cloud?

Summary

- ESRIN outsources their “archive”: 20PB end 2017, 100PB within 10 years (*get/put i/f*)
- Base services not discipline specific – but some institutes may wish to maintain some control
- Combine e.g. Zenodo with a commercial archive (TDR)? “Solve” large storage requests to CERN
- Retain control of the “tricky bits” whilst out-sourcing the simpler ones?