



# SCD in Horizon 2020

Ian Collier

RAL Tier 1

GridPP 33, Ambleside, August 22<sup>nd</sup> 2014

# What is Horizon 2020?

Horizon 2020 is the biggest EU Research and Innovation programme ever with nearly €80 billion of funding available over 7 years (2014 to 2020) – in addition to the private investment that this money will attract.

It promises more breakthroughs, discoveries and world-firsts by taking great ideas from the lab to the market.



# What is Horizon 2020?

Horizon 2020 is the biggest EU Research and Innovation programme ever with nearly €80 billion of funding available over 7 years (2014 to 2020) – in addition to the private investment that this money will attract.

It promises more breakthroughs, discoveries and world-firsts by taking great ideas from the lab to the market.

Sounds great!



# STFC SCD in many bids

- 26 bids in total at last count
- Covering wide range of activities
  - Shan't try to describe them all
- Some key phrases
  - Crisis management for extreme weather events
  - VREs for nanotechnology, social science, structural biology...
  - Improved Nuclear Fuel Operational Modelling tools for Safety
  - Numerical Linear Algebra for Future and Emerging Technology



# STFC SCD in many bids

- A couple others here are involved in
  - EGI-Engage
  - AARC
  - RAPIDS
- I know about:
  - PanDaaS
  - INDIGO-DataCloud
  - ZEPHYR



# INDIGO-DataCloud

## OBJECTIVES (from bid doc)

1. Development of a Platform based on open source software, without restrictions on the e-Infrastructure to be accessed (public or commercial, GRID/Cloud/HPC) or its underlying software.
2. Providing support for PaaS at the IaaS level
3. Provide high level access to the platform services in the form of science gateways and access libraries.
4. Development of exploitation mechanisms for heterogeneous infrastructures in a wide context: hybrid (public-private) infrastructures deployed in Grid, Cloud and HPC mode.
5. Provide software lifecycle services and related support to project developers and infrastructure providers in a way that can be sustainable for user communities



# ZEPHYR

Zetabyte-Exascale Prototypes for Heterogeneous  
Year-2020+ scientific data factory Requirements  
(ZEPHYR)

Objective:

ZEPHYR will prototype the architectural choices and investigate technological solutions to address the challenge of managing science data at the Zettabyte scale with trillions of objects



# ZEPHYR Work Packages

- WP1 Project Management and transversal tasks
- WP2 Efficient Application Interfaces for data mining and access at the Zetabyte-Exascale level
- WP3 Clean slate approaches for multi-centric storage virtualization
- WP4 Sustainable integration of Local Data Centres







# Questions?

# INDIGO-DataCloud

Addresses topics 4&5 from EINFRA-1 Call

(<http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/2137-einfra-1-2014.html>)

(4) Large scale virtualisation of data/compute centre resources to achieve on-demand compute capacities, improve flexibility for data analysis and avoid unnecessary costly large data transfers.

(5) Development and adoption of a standards-based computing platform (with open software stack) that can be deployed on different hardware and e-infrastructures (such as clouds providing infrastructure-as-a-service (IaaS), HPC, grid infrastructures...) to abstract application development and execution from available (possibly remote) computing systems. This platform should be capable of federating multiple commercial and/or public cloud resources or services and deliver Platform-as-a-Service (PaaS) adapted to the scientific community with a short learning curve. Adequate coordination and interoperability with existing e-infrastructures (including GÉANT, EGI, PRACE and others) is recommended



# ZEPHYR

Addresses topic 7 from EINFRA-1 Call

(<http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/2137-einfra-1-2014.html>)

(7) Proof of concept and prototypes of data infrastructure-enabling software (e.g. for databases and data mining) for extremely large or highly heterogeneous data sets scaling to zetabytes and trillion of objects. Clean slate approaches to data management targeting 2020+ 'data factory' requirements of research communities and large scale facilities (e.g. ESFRI projects) are encouraged.

