



STFC Cloud Developments

Ian Collier & Alex Dibbo

STFC Rutherford Appleton Laboratory

UK Research & Innovation

HEPiX Spring, Madison, WI

17 May 2018

Contents

- Background & History
- Current design considerations
- New Use cases
- (Some of) What this allows us to do
 - IRIS (UKTO)



Background

- OpenNebula cloud now running for several years on a stable, but pre-production basis.
 - Has served very well.
- Rise of more complex use cases – in particular supporting external user communities → deploying OpenStack
- In final stages of deploying
 - decommissioning OpenNebula
 - migrating remaining users in coming months



Hardware

- OpenNebula: reducing from ~700 to ~300 cores
- OpenStack: ~500 cores, growing to ~3500 end May and ~5000 by end 2018.
- Adding
 - Hypervisors: 108 Dell 6420 sleds (27 x 4-up 2U 6400 chassis) - **Testing complete, about enter production**
 - 16 physical cores each, 6GB RAM/core, 25G NIC
 - Cloud Storage: 12 Dell R730xd (12 bay 2U) - **Testing complete, about enter production**
 - 12 x 4TB each, total 576TB, 25G NICs
 - Replicated CEPH (3x), VM image storage
 - Data Storage: 21 x (Dell R630 + 2 x MD1400) sets - **Just finished testing**
 - Added to ECHO for UKT0 project
 - 4PB raw /~2.9TB configured
 - Network:
 - 7 x Mellanox SN2100 16x100Gb port



OpenStack implementation

- Flexible
 - Within reason want a design that can accomodate whatever comes along – within reason
- Multi tenancy
 - Separating different user communities
- Highly Available
 - Services should be as highly available as possible – OpenStack services behind HAProxy
- Ceph
 - A replicated Ceph cluster called SIRIUS provides block storage for VMs and Volumes
 - 3x Replication
 - Optimised for lower latency



Ceph RBD

- A replicated Ceph cluster called SIRIUS provides block storage for VMs and Volumes
- 3x Replication
- Optimised for lower latency
- Helps speed up launch of VMs
- Are considering use cases that would benefit from local VM storage
- Currently being rebuilt – fresh installation in new machine room location.



Multi Tenancy

- Projects (Tenants) need to be isolated
 - From each other
 - From STFC site network
- Security Groups
- VXLAN private project networks
 - Brings its own problems



Private Networks

- Virtual machines connect to a private network
 - VXLAN is used to tunnel these networks across hypervisors
- Ingress and Egress is via a virtual router with NAT
 - Distributed Virtual Routing is used to minimise this bottleneck – every hypervisor runs a limited version of the virtual router agent.

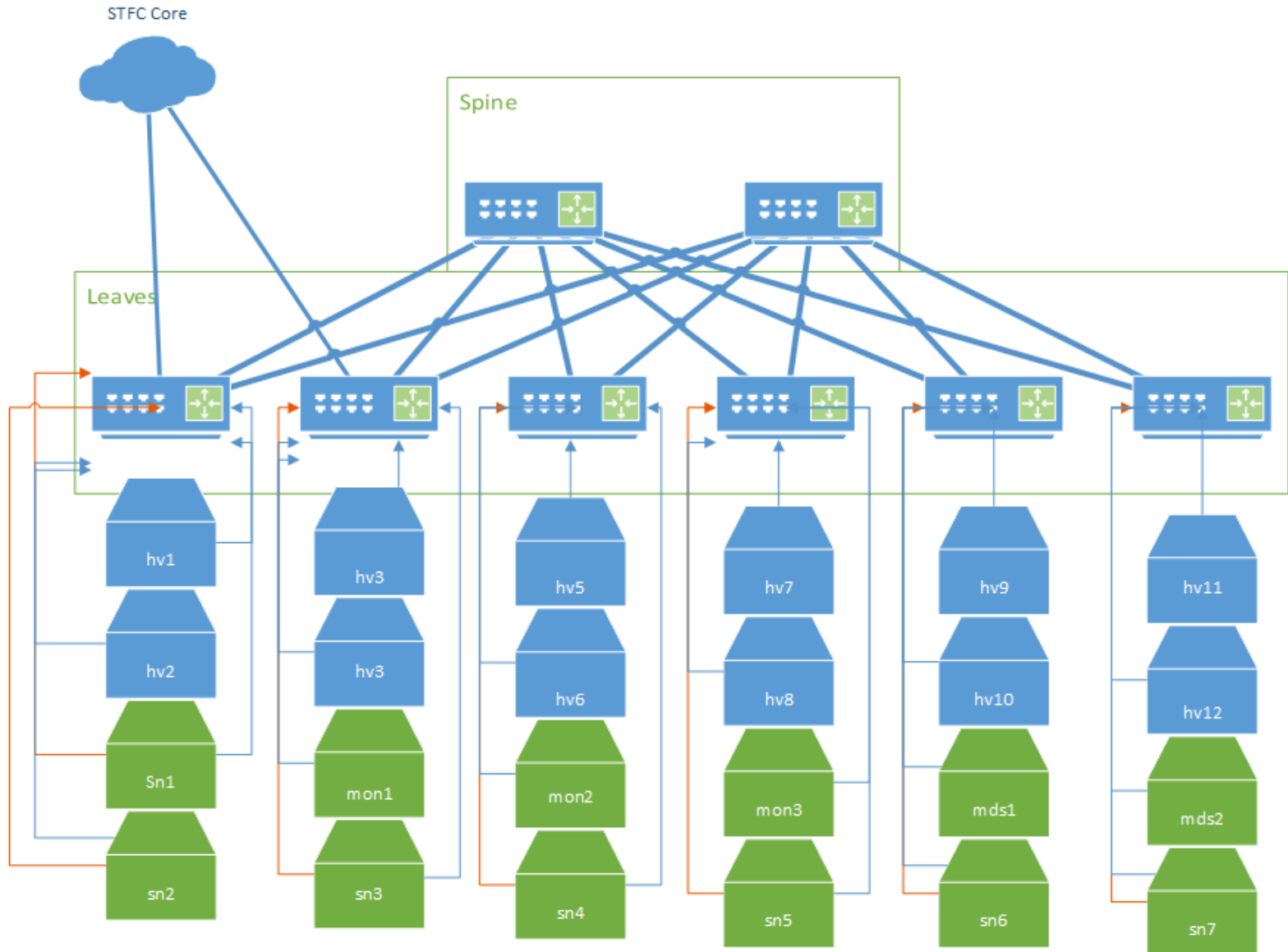


VXLAN

- VXLAN by default showed significant overheads
 - VXLAN performance was ~10% of line rate
- Tuning memory pages, CPU allocation, mainline kernel
 - Performance is ~40% of line rate
- Hardware offload
 - VXLAN offload to NIC gives ~80% of line rate
- High Performance Routed network + EVPN
 - 99+% of line rate



Cloud Network



Flexible

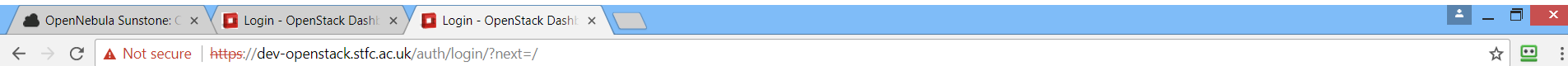
- Availability Zones across site
 - 1st will be in ISIS soon
 - In discussion with other departments
- GPU support
- AAI
- APIs
 - Nova, EC2, OCCl
- Design decisions should, so far as possible, preclude anything
- ‘Pet’ VMs as well as ‘cattle’



Current User Communities

- STFC Scientific Computing Department “internal”
 - Self service VMs
 - Some programmatic use
- Tier1
 - Bursting the batch farm
- ISIS Neutron Source
 - Data-Analysis-as-a-Service
- Jenkins Build Service for other groups in SCD
- Central Laser Facility – OCTOPUS project
- Diamond Lightsource - Xchem
 - Cloud bursting Diamond GridEngine
 - Xchem data processing using OpenShift
- WLCG Datalake project
- Quattor Nightlies
- Range of H2020 & similar projects

AAI – EGI CheckIn - Horizon



Welcome to the SCD Cloud

This service provides a private IaaS cloud resource for SCD users. To start, select the login link at the top of the page and enter your federal username and password.

You should be aware that the SCD Cloud is still in active development. Please read the [Terms of Service](#) before proceeding.

For any questions, problems or ideas email us at cloud-support@helpdesk.gridpp.rl.ac.uk

Log in

Authenticate using

Keystone Credentials

If you are not sure which authentication method to use, contact your administrator.

Domain *

User Name *

Password *

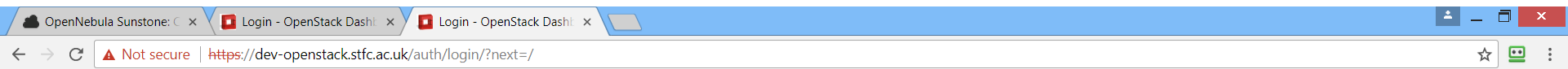
CONNECT

Infrastructure



Science & Technology
Facilities Council

AAI – EGI CheckIn – Horizon 2



SCD Cloud [Terms](#) [SLA](#) [FAQs](#)

Welcome to the SCD Cloud

This service provides a private IaaS cloud resource for SCD users. To start, select the login link at the top of the page and enter your federal username and password.

You should be aware that the SCD Cloud is still in active development. Please read the [Terms of Service](#) before proceeding.

For any questions, problems or ideas email us at cloud-support@helpdesk.gridpp.rl.ac.uk

Log in

Authenticate using

EGI CheckIn

If you are not sure which authentication method to use, contact your administrator.

CONNECT

Infrastructure

The cloud is based on [OpenStack](#) for the virtualisation and [CEPH](#) for the storage. For the hardware, we use

28x Dell R420 for the hypervisors

30x Dell R520 for the storage nodes

This gives us a total of **3.5TB** of memory, **896** processing cores and a storage capacity of approximately **750TB**.



© Copyright STFC 2016 | [Terms of Service](#) | [Service-level Agreement](#) | [Frequently Asked Questions](#) | [GitHub](#) | [Issue Tracker](#)



Science & Technology
Facilities Council

AAI – EGI CheckIn

OpenNebula Sunstone: x Login - OpenStack Dashb x EGI | Select your identity x

Secure | <https://aai-dev.egi.eu/proxy/module.php/discopower/disco.php?entityID=https%3A%2F%2Faai-dev.egi.eu%2Fproxy%2Fmodule.php%2Fsaml%2Fsp%2Fmetadata.php%2Fsso&return=https%3A%2F%2Fa...>

Log In

LOG IN WITH GOOGLE

OR

CHOOSE ANOTHER ACCOUNT

Search...

Google

A. T. Still University

AAF Virtual Home

AAI@EduHr Single Sign-On Service

Aalborg University

Aalto University

AARC DIY Identity Provider

Aarhus School of Marine and Technical Engineering

Aarhus Universitv

OR

Google DARIAH EGI LOG-IN Facebook IGTF LinkedIn

ORCID

English

grnet
Networking Research and Education

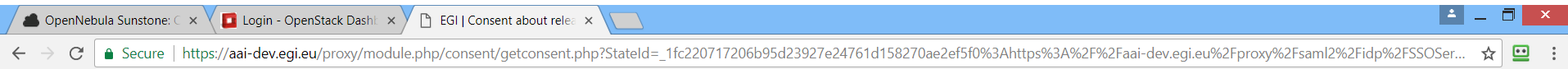
Copyright © 2016-2018 | Check-in is an EGI service provided by GRNET,
co-funded by the EGI Foundation (EGI.eu) and EOSC-hub | Powered by RCIAM

Terms Privacy



Science & Technology
Facilities Council

AAI – Google Login



Check-in

EGI AAI OpenID Connect Provider Proxy (DEV) requires that the information below is transferred.

Surname	Dibbo
Given name	Alexander
Display name	Alexander Dibbo
Mail	apdibbo@googlemail.com
Affiliation at home organization	staff@STFCTesting
Identity assurance profile	https://aai.egi.eu/LoA#Low
Unique, non-reassignable, persistent, pseudonymous user ID	d98aa2fd27edc0e621352511190d7e1eb911014f7fb15df6dce93b4c39e255c9@egi.eu

☐ Remember

YES, CONTINUE

NO, CANCEL

English ▲



Copyright © 2016-2018 | Check-in is an EGI service provided by GRNET, co-funded by the EGI Foundation (EGI.eu) and EOSC-hub | Powered by RCIAM

[Terms](#) [Privacy](#)

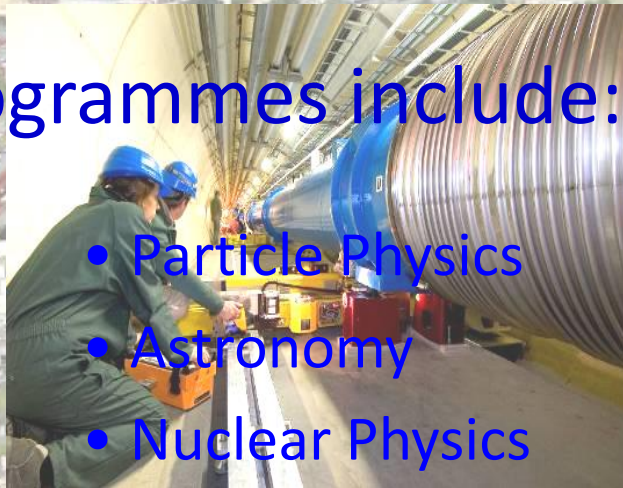


Science & Technology
Facilities Council

What do we do?

Programmes include:

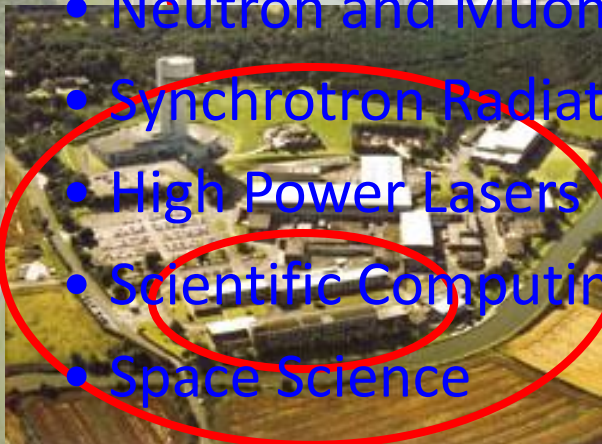
- Particle Physics
- Astronomy
- Nuclear Physics



Large Hadron Collider

National Laboratories include:

- Neutron and Muon Source
- Synchrotron Radiation Source
- High Power Lasers
- Scientific Computing/Hartree Centre
- Space Science
- Electronics, Cryogenics



Daresbury Laboratory

International include:

- CERN
- ESQ
- SKA
- ESS
- ESRF
- ILL



Square Kilometre Array



ESRF & ILL, Grenoble

GridPP: HTC for the LHC



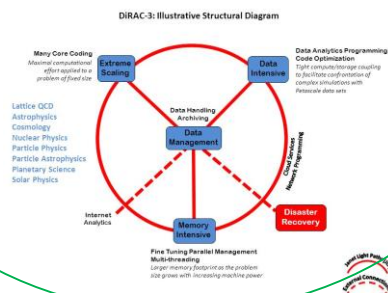
Hartree Centre: Industry



Scientific Computing Department



DiRAC: HPC for Theory & Cosmology



HTC and data storage for Astronomy, Nuclear, Astroparticle



**Science & Technology
Facilities Council**

Wider UK landscape

- Effort running for several years now under ‘UKT0’ banner
- Bringing together different communities
 - HTC and HPC for ‘STFC’ science
- Linking together diverse resources into coherent and complementary support for existing & emerging communities
- “UKT0” (now IRIS) is an umbrella project which
 - represents users
 - ‘contracts’ to resource providers
 - eventually should provide effort for cross-cutting activities



IRIS – eInfrastructure for STFC Science

- Initiative to bring STFC computing interests together
- Formed bottom up by the science communities and compute providers
- Association of peer interests (PPAN + National Facilities + Friends)
 - Particle Physics: LHC + other PP experiments (GridPP)
 - DiRAC (UK HPC for theoretical physics)
 - National Facilities: Diamond Light Source, ISIS
 - Astro: LOFAR, LSST, EUCLID, SKA,
 - Astro-particle: LZ, Advanced-LIGO
 - STFC Scientific Computing Dept (SCD)
 - CCFE (Culham Fusion)
- Not a project seeking users - users wanting and ready to work together

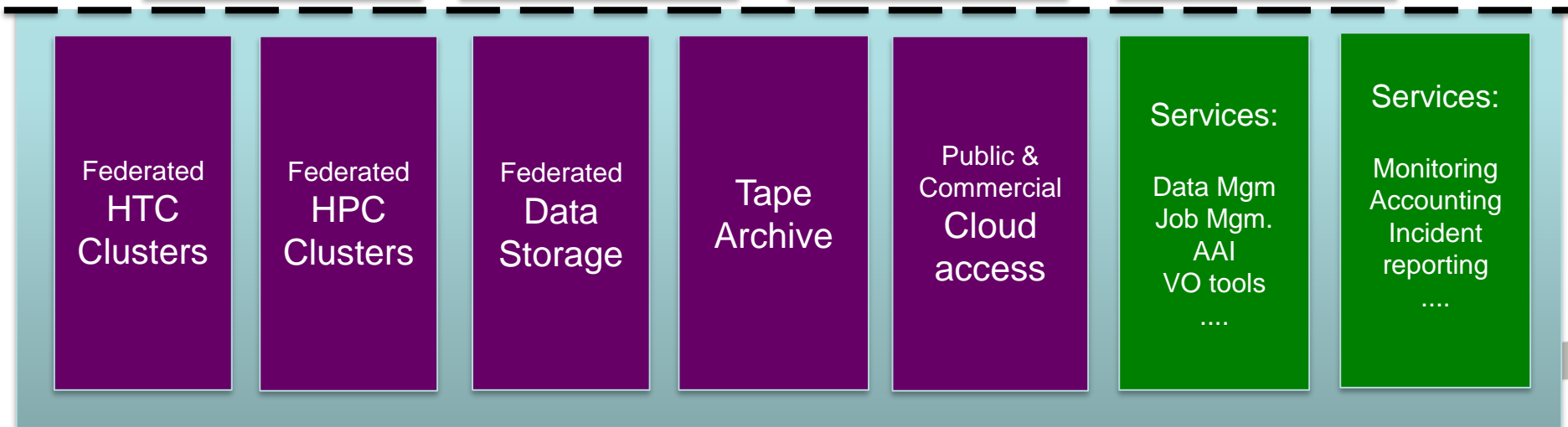


Evolution – UKT0 -> IRIS

- 2015-2017 Loose association
 - → Community Meeting Autumn 2015
- 2017/2018 first funding £1.5M to be spent by 31 March'18
 - interim-PMB (to manage short term completion)
 - Grant delivered resources into machine rooms
 - Cambridge, Edinburgh, Manchester, RAL,
 - First (proto) Collaboration Meeting – March 2018
 - Project was successful and on time (spent money/wrote software)
- 2018 Approved BEIS “keep lights on funding” of £4M p.a. for 4 years for UKT0/ALC.
 - Now finalising Organisational structures and delivery planning



UKTO 'idea'



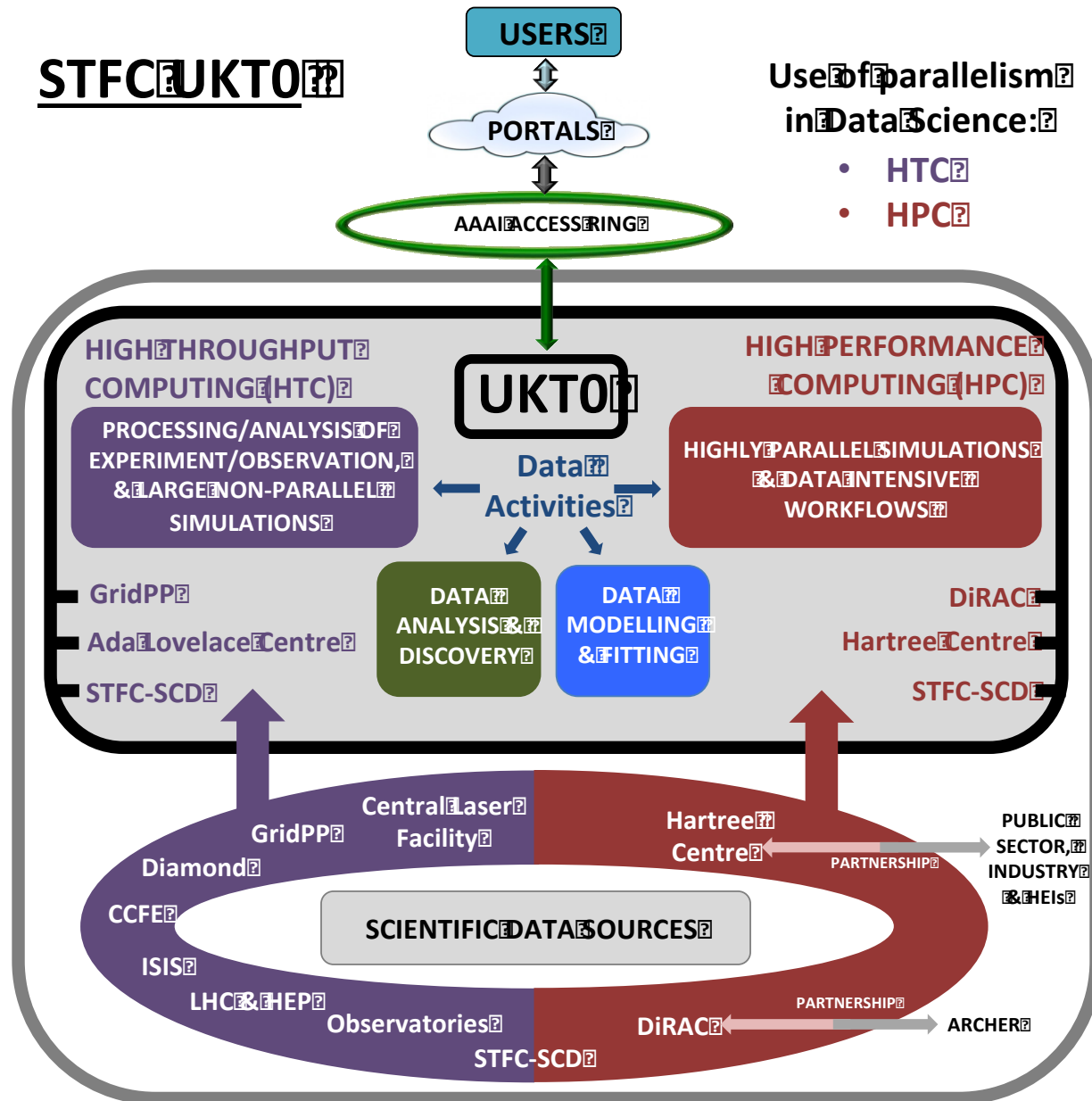
Share where it makes sense to do so



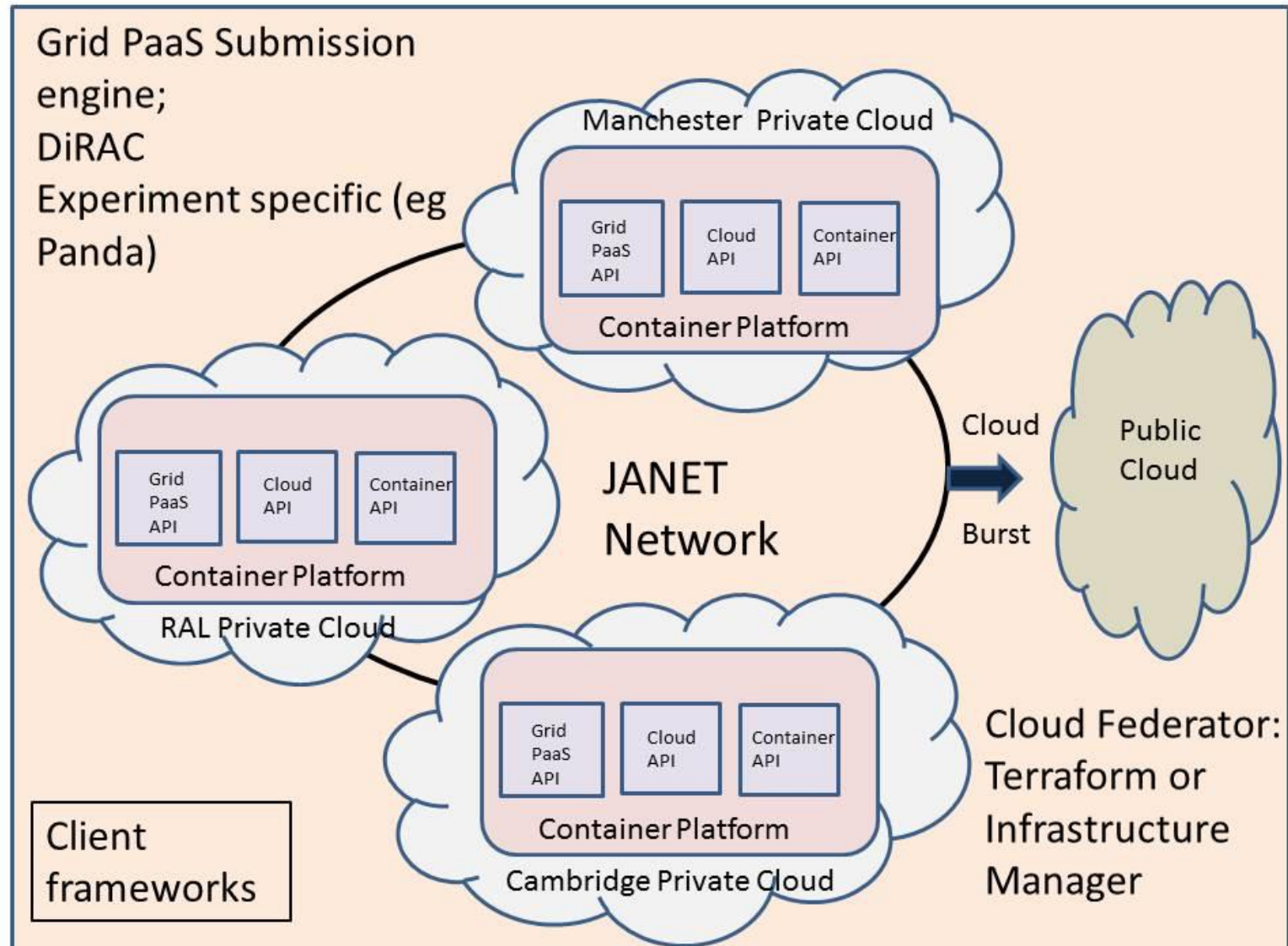
**Science & Technology
Facilities Council**

Long Term Aspiration

STFC UKTO



Initial Architecture



Investment – next 4 years

- Bulk is from Department of Business, Energy & Industrial Strategy (BEIS)
- £10.5M eInfrastructure compute and storage assets
- £5.5M for eInfrastructure Digital assets
- Minimal resource ☹
 - £250K from BEIS in years 3+4
 - £250K resource from STFC in years 1+2





Questions?

ian.collier@stfc.ac.uk

alexander.dibbo@stfc.ac.uk