Science & Technology Facilities Council

JASMIN is changing. What does the future hold?

Matt Pritchard & Ag Stephens

On behalf of the JASMIN team
(STFC:SCD, STFC/NERC:CEDA, NERC:NCAS, NERC:NCEO)

Overview

- Concept of managed, semi-managed and un-managed Virtual Organisations
- Physical and organisational views of the infrastructure
- Interactions between the un-managed and managed components
- Special Virtual Organisations inside JASMIN
- Planned usage of JASMIN services in Phases 2/3

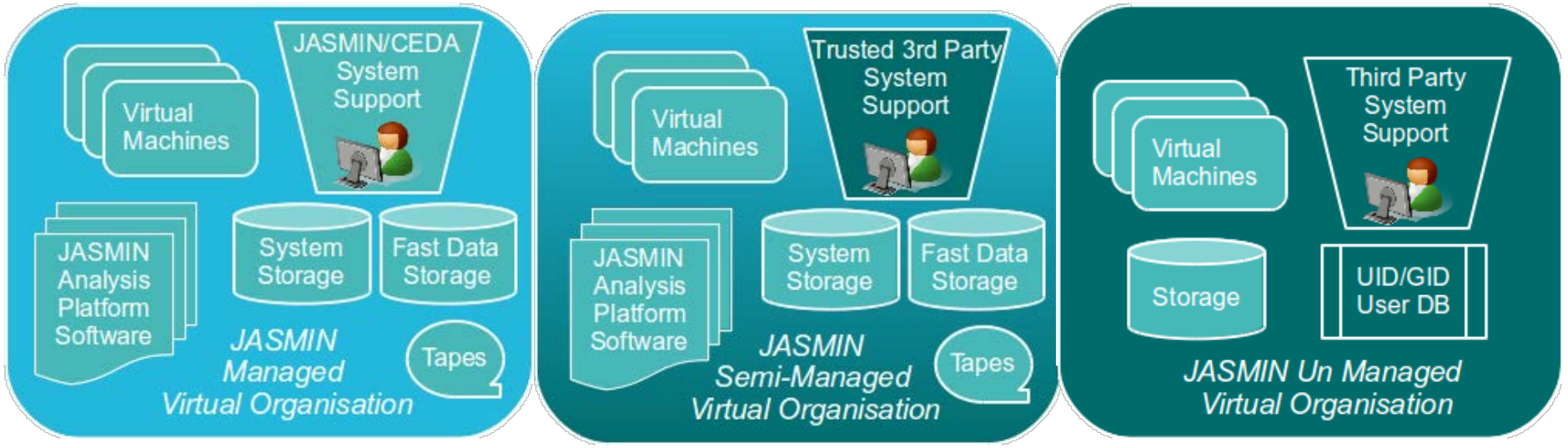
Phase 2/3 expansion

- 2013 NERC Big Data capital investment
  - Wider scope: support projects from new communities, e.g.
    - EOS Cloud
      - Environmental ‘omics. Cloud BioLinux platform
    - Geohazards
      - Batch compute of Sentinel-1a SAR for large-scale, hi-res Earth surface deformation measurement

Sentinel-1a (ESA)

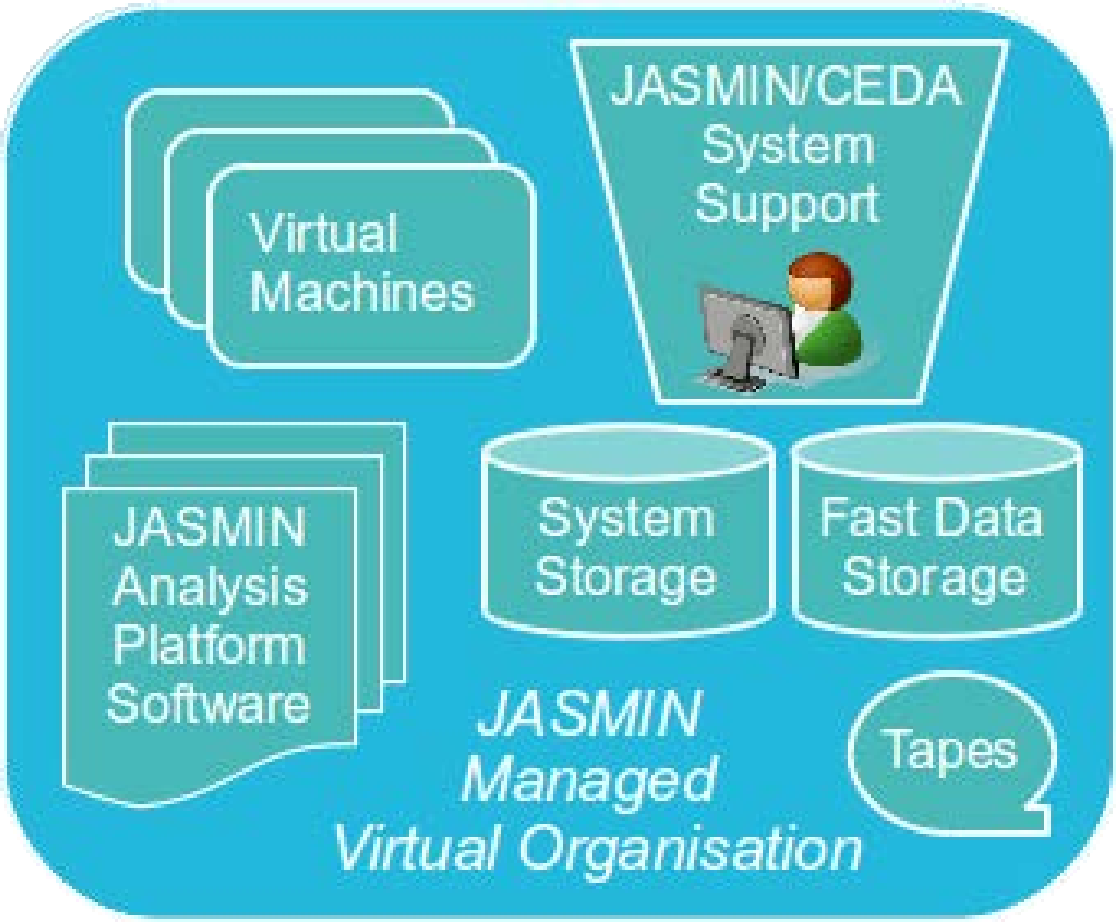
	Phase 2 by March 2014	Phase 3 by March 2015
JASMIN hard upgrade	+7 Petabytes disk +6 Petabytes tape +3000 compute cores network enhancement	+o(2) Petabytes disk +o(800) compute cores network enhancement
JASMIN soft upgrade	Virtualisation software Scientific analysis software Cloud management software Dataset construction Documentation	

Virtual Organisations (VOs)



Platform as a Service (Paas) → Infrastructure as a Service (IaaS)

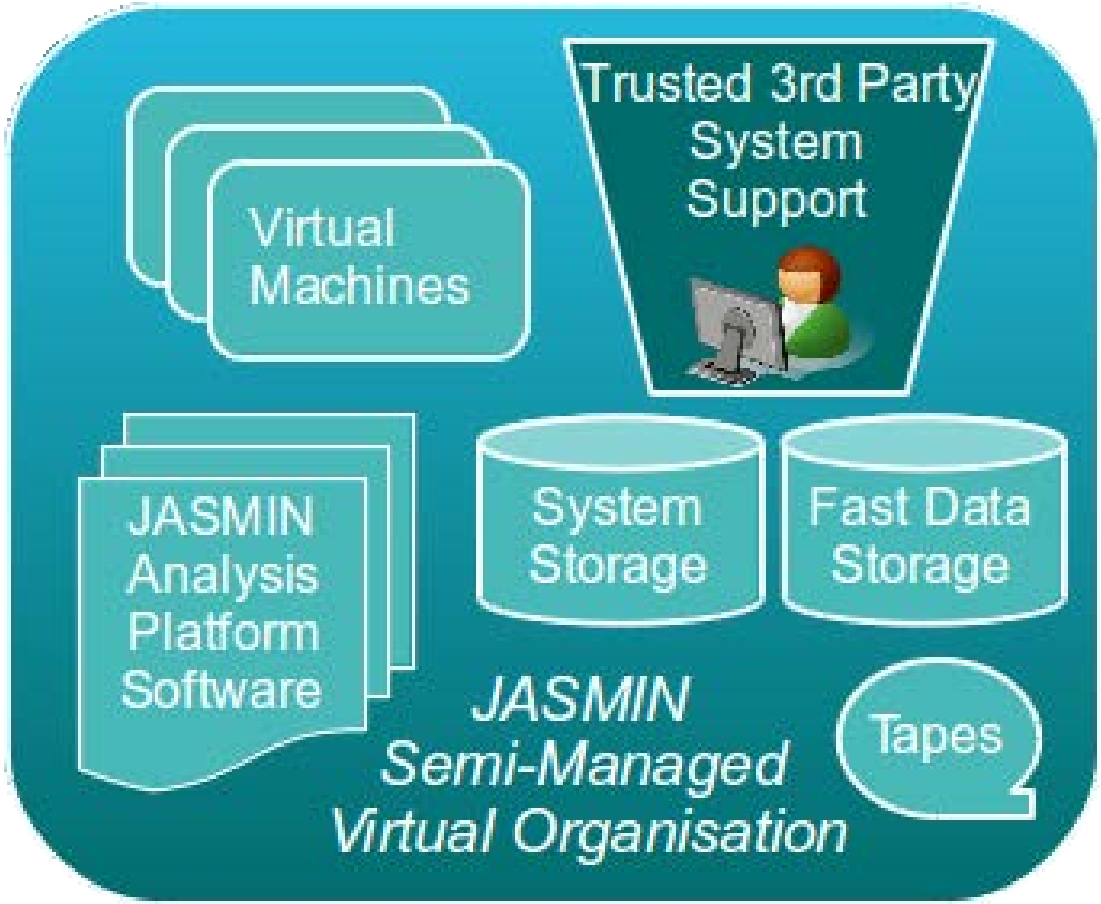
Managed Virtual Organisations (e.g. CEDA)



RHEL6 based

Users known/managed by SCD/CEDA

Semi-managed Virtual Organisations

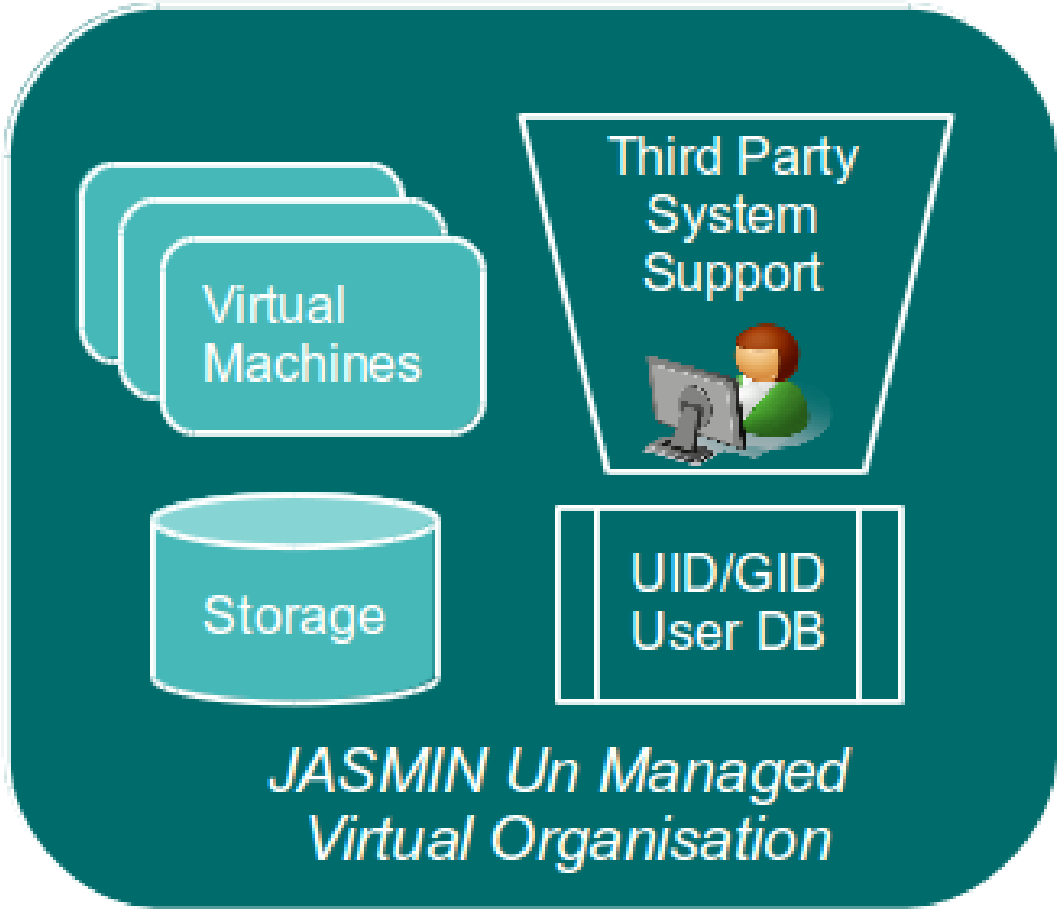


RHEL6 based

Users known/managed by SCD/CEDA

Trusted users able to have Root access

Un-managed Virtual Organisations (e.g. EOS Cloud)

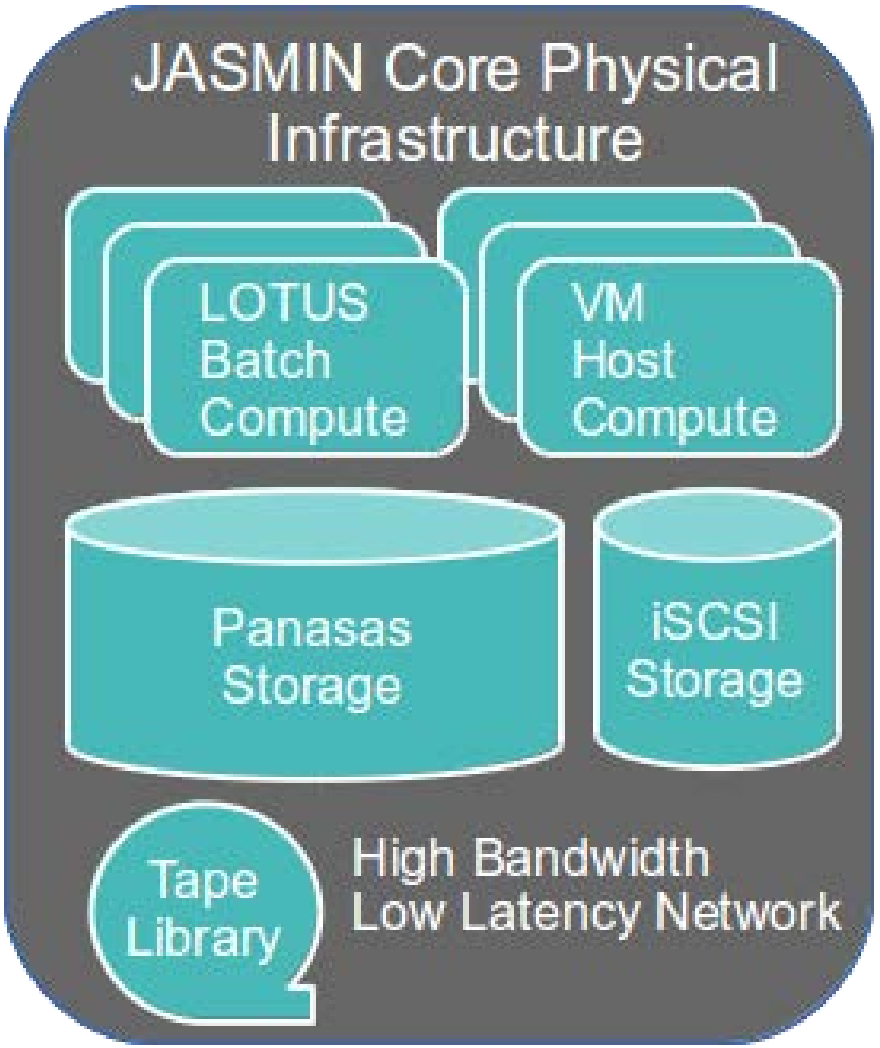


Operating system decided by VO manager

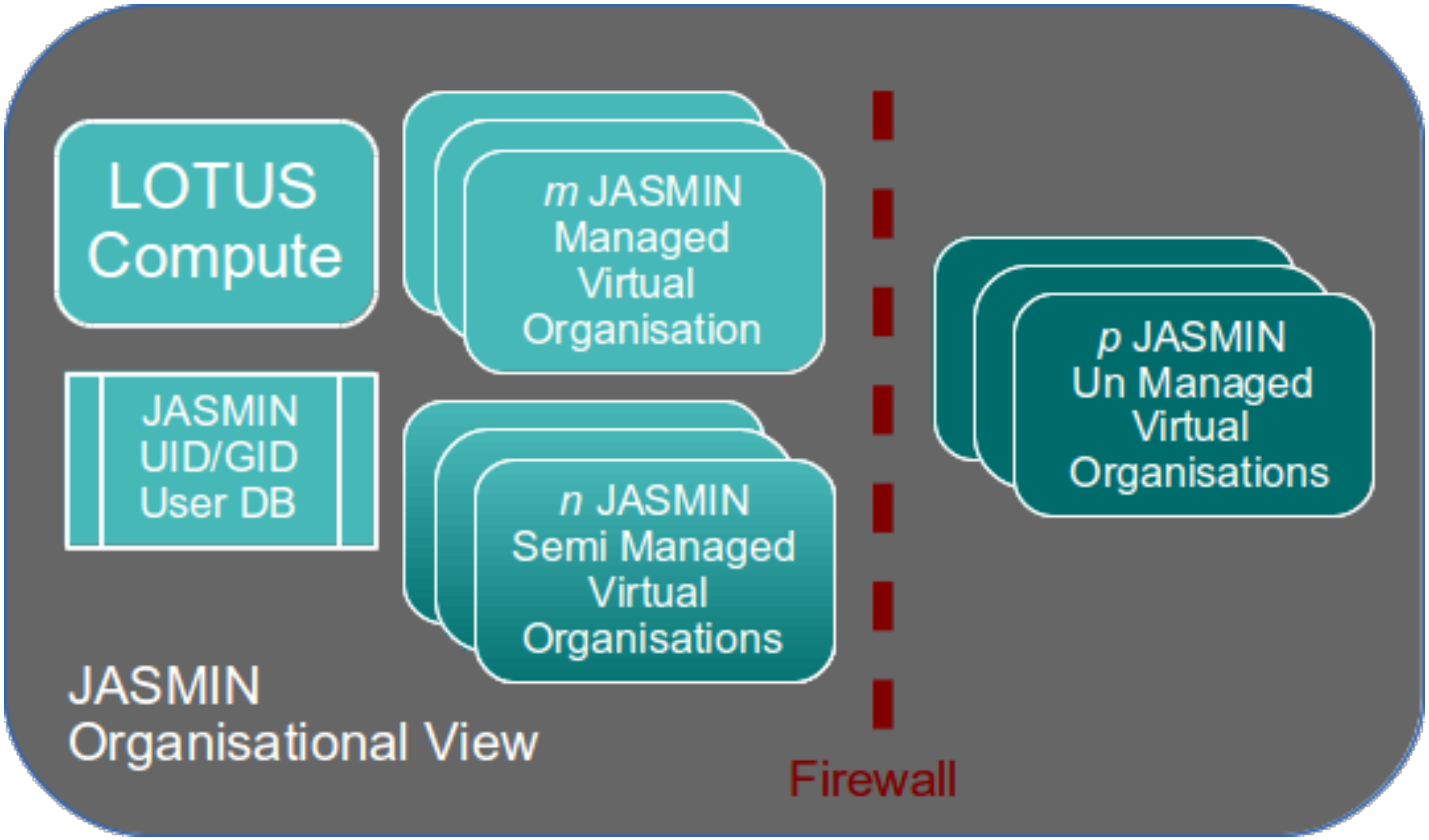
Users known/managed by VO manager

Storage: local iSCSI

Physical View of JASMIN Infrastructure



Organisational View of Infrastructure

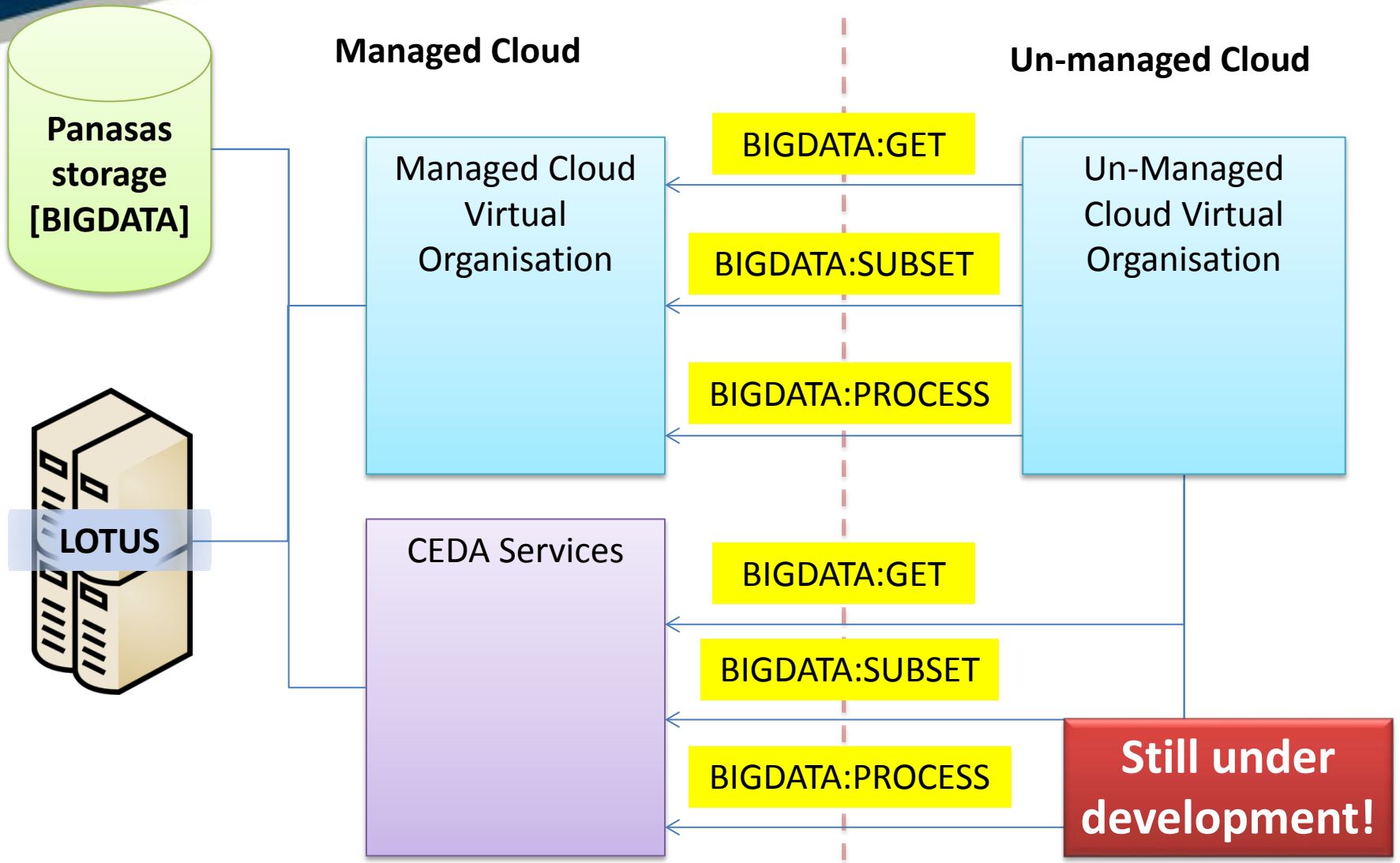


Use of Cloud Technologies

The main technologies employed for this are:

- **VMWare** – the virtualisation platform
- **vCloud**
  - Providing a web interface (vCloud Director) to allow VO managers to create/allocate/deploy resources within their own part of the network.
  - Includes REST API
- **JASMIN Cloud Portal**
  - CEDA-developed web interface giving access to **specific features of vCloud (those we're happy for users to have!)**
  - Abstraction layer on top of the REST API

Interactions between managed and un-managed



Some Special Virtual Organisations

CEDA: Centre for Environmental Data Archival

- Will provide archival services for the community.
- Data held in the archive will be managed, and made available to all the managed and semi-managed VOs directly (and indirectly to the un-managed VOs).
- Will provide “generic” access platforms for virtual organisations that do not wish to manage their own platforms and users who do not belong to specific VOs.

EOS Cloud

- Cloud services for the environmental 'omics community
- Delivered by JASMIN on behalf of the Centre for Ecology and Hydrology

CEMS: The facility for Climate, Environment and Monitoring from Space

CEMS
Climate, Environment & Monitoring from Space

- Will acquire and archive (via CEDA) key third party datasets needed by the NERC science community.
- Will provide services for the Earth Observation Community, in particular, in partnership with Satellite Applications catapult (SAC), the UK and European space industry.
- The academic component will run on JASMIN, the bulk of the industrial component, in the SAC, with access to CEDA data.

What else is planned in JASMIN Phase 2/3?

ECMWF on JASMIN

- Running data services on large ECMWF data.

EOS Cloud

- Cloud services for the environmental bioinformatics community
- Delivered by JASMIN on behalf of the Centre for Ecology and Hydrology.

Met Office Climate Cloud?

- Prototype developed within CEMS (on JASMIN platform).
- Discussions underway with Satellite Applications Catapult about possible deployment approaches.

Running OSTIA on LOTUS

- OSTIA to be run on LOTUS through cylc/rose ported to JASMIN.

JULES on JASMIN

- Web-interface to the JULES land-surface model.

NERC Environmental Workbench

- CEH-led activity to develop a set of cloud-based tools allowing users to carry out scientific workflows.

Support for EU Horizon 2020 Projects

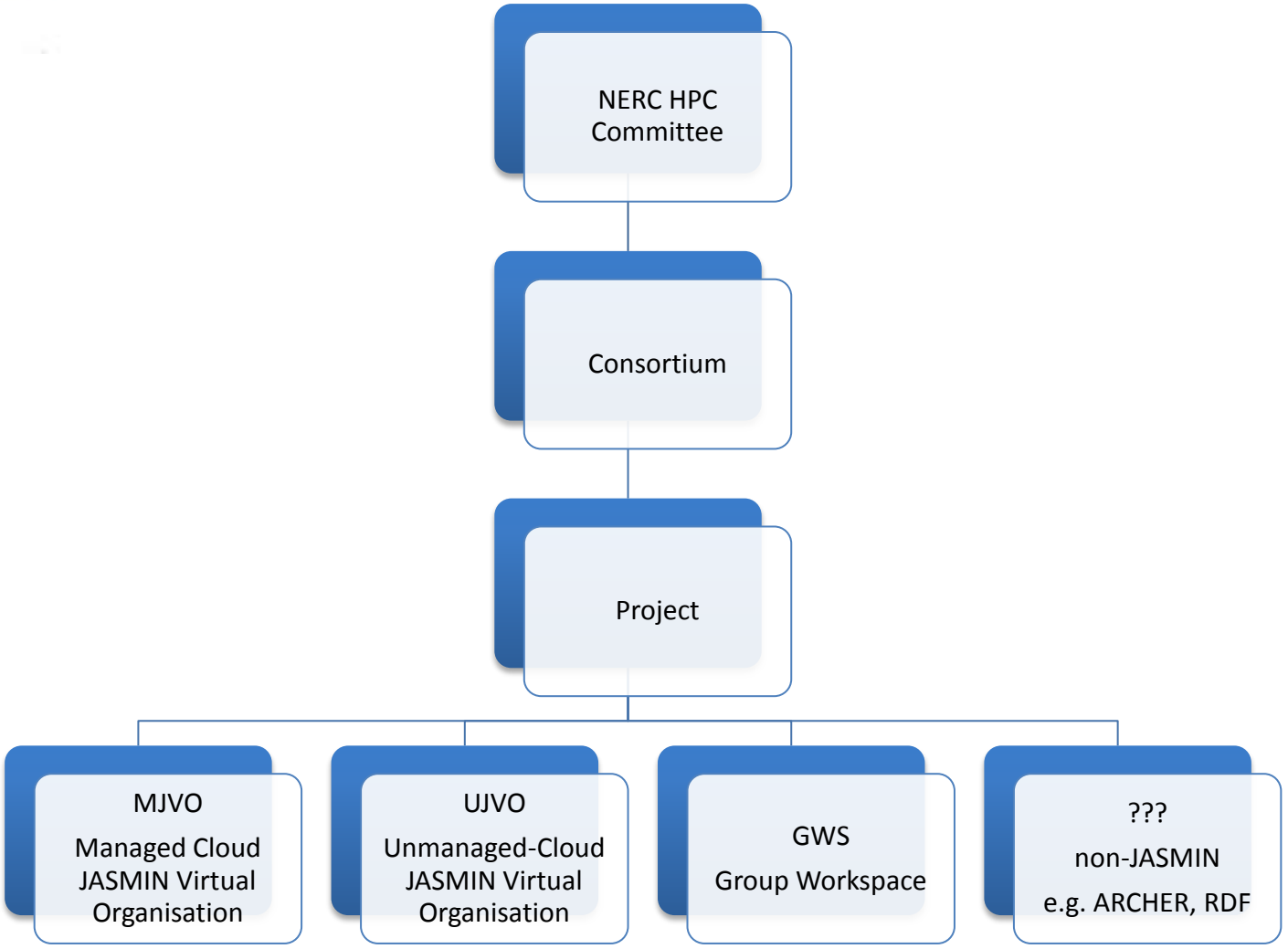
- EUSTACE (Met Office led; funded)
- PRIMAVERA (Met Office/NCAS led; submitted)
- Many more…

And many more NERC "Big Data" and Earth Observation projects…

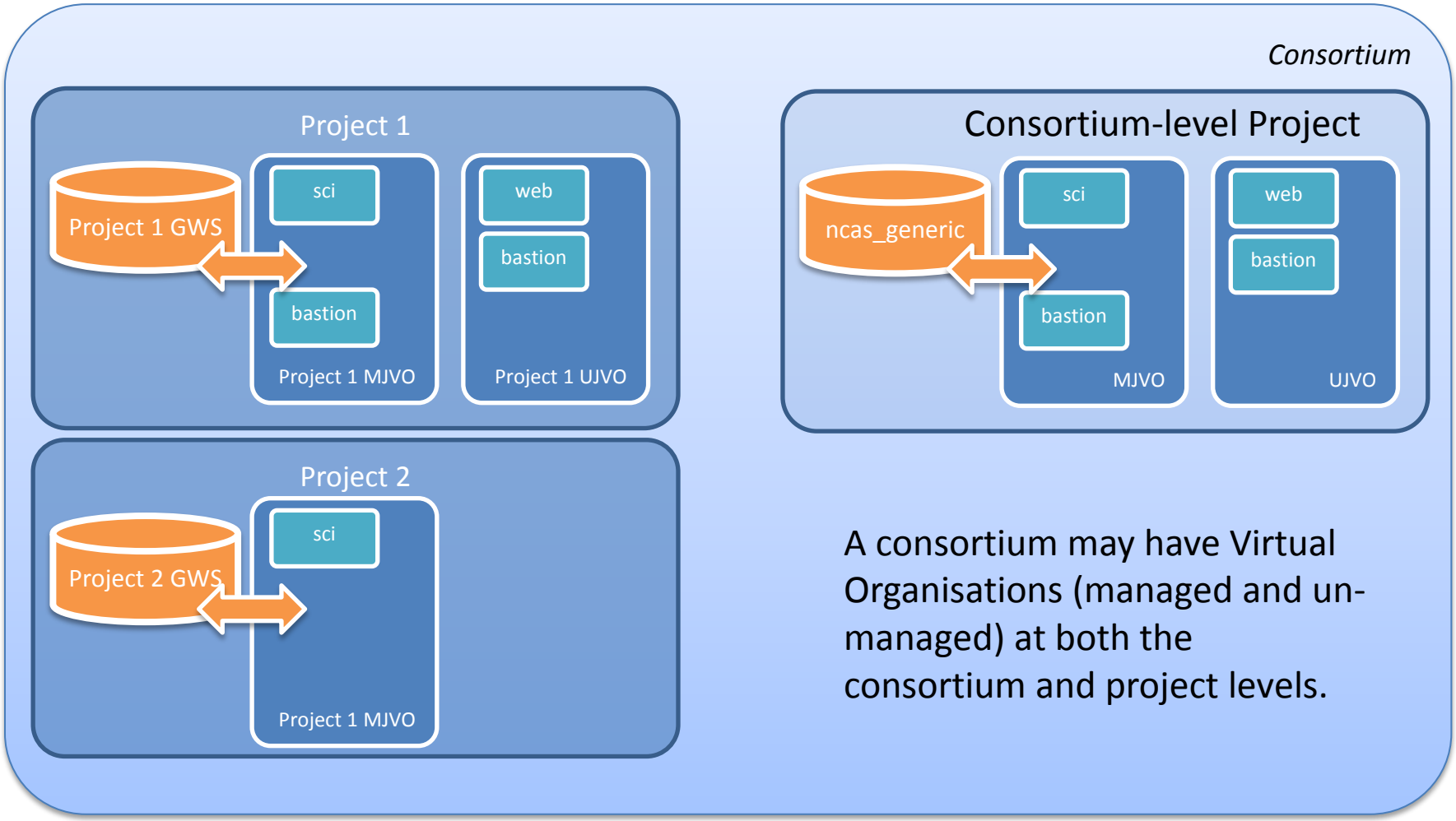
Management via Consortia

Name	Manager
Atmospheric & Polar Science	Grenville Lister
Oceanography & Shelf Seas	
Solid Earth & Mineral Physics	
Genomics	
Ecology & Hydrology	
Earth Observation & Climate Services	Victoria
Geology	
Archive	Sam
Director's cut	Bryan

Management via Consortia



Consortium and project level resources



A consortium may have Virtual Organisations (managed and un-managed) at both the consortium and project levels.

Finding out more…

JASMIN documentation pages

http://www.jasmin.ac.uk

JASMIN Phase 2 Launch page (with links to slides)

http://jasmin.ac.uk/what-is-jasmin/jasmin-launch-event/

JASMIN Expansion in Phases 2 and 3

http://jasmin.ac.uk/what-is-jasmin/jasmin-expansion-phases-2-and-3/

JASMIN paper

Lawrence, B.N. , V.L. Bennett, J. Churchill, M. Juckes, P. Kershaw, S. Pascoe, S. Pepler, M. Pritchard, and A. Stephens. **Storing and manipulating environmental big data with JASMIN.** *Proceedings of IEEE Big Data 2013, p68-75,* doi:10.1109/BigData.2013.6691556