

NCAS, CEDA and JASMIN: AN OVERVIEW



Climate



Air Quality



Atmospheric
physics



**National Centre for
Atmospheric Science**

NATURAL ENVIRONMENT RESEARCH COUNCIL

Facilities and
Services



Technology



**“National
capability”**



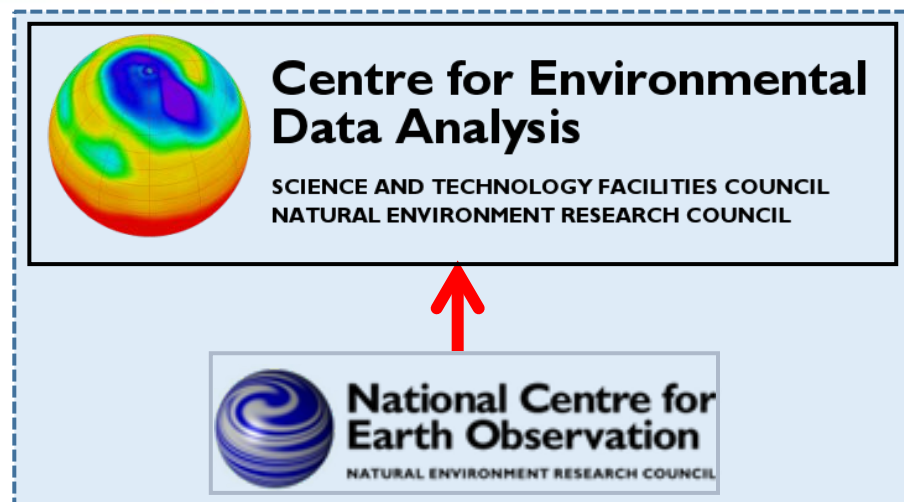
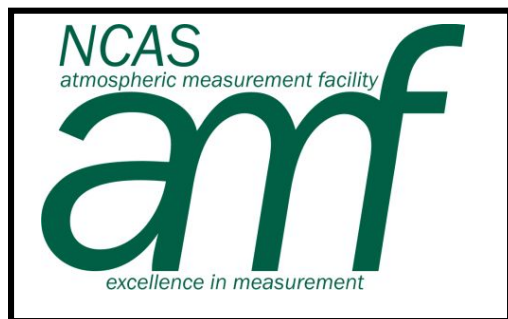
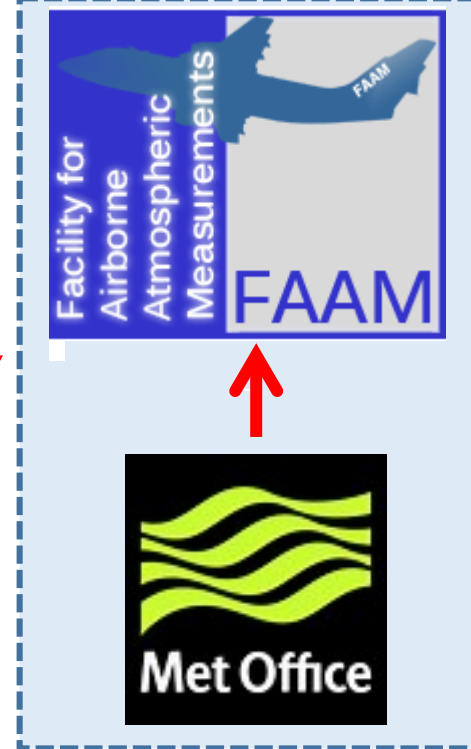
**Centre for Environmental
Data Analysis**
SCIENCE AND TECHNOLOGY FACILITIES COUNCIL
NATURAL ENVIRONMENT RESEARCH COUNCIL



**National Centre for
Atmospheric Science**
NATURAL ENVIRONMENT RESEARCH COUNCIL



**National Centre for
Earth Observation**
NATURAL ENVIRONMENT RESEARCH COUNCIL





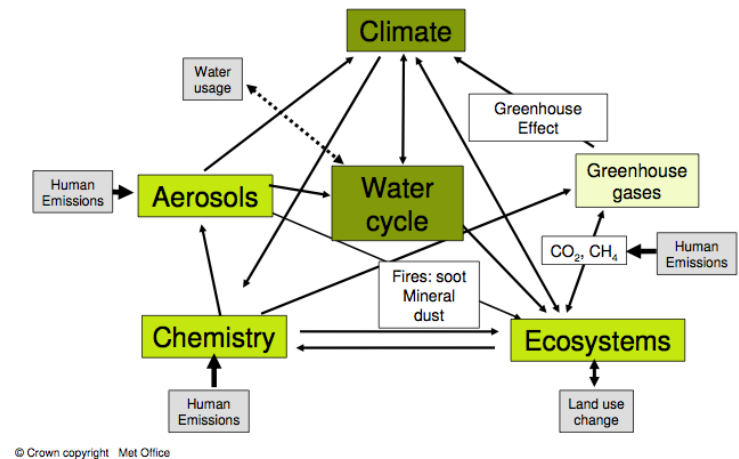
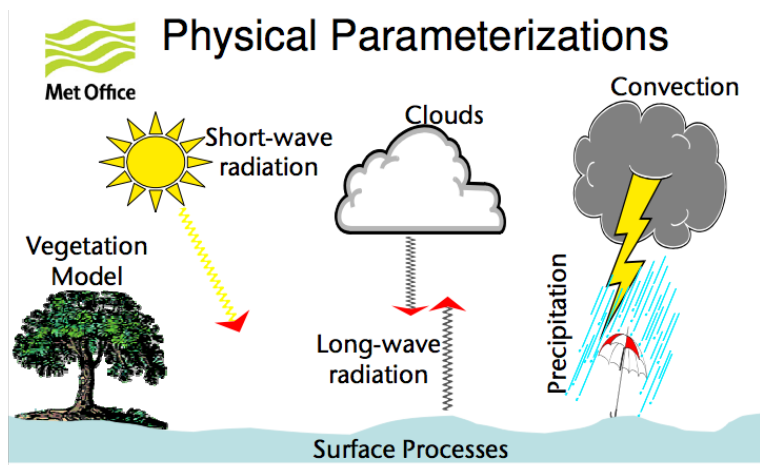
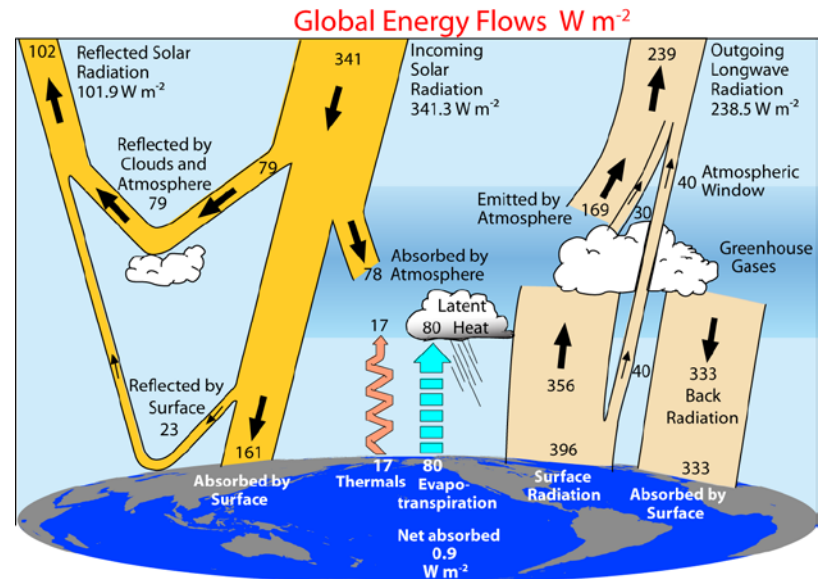
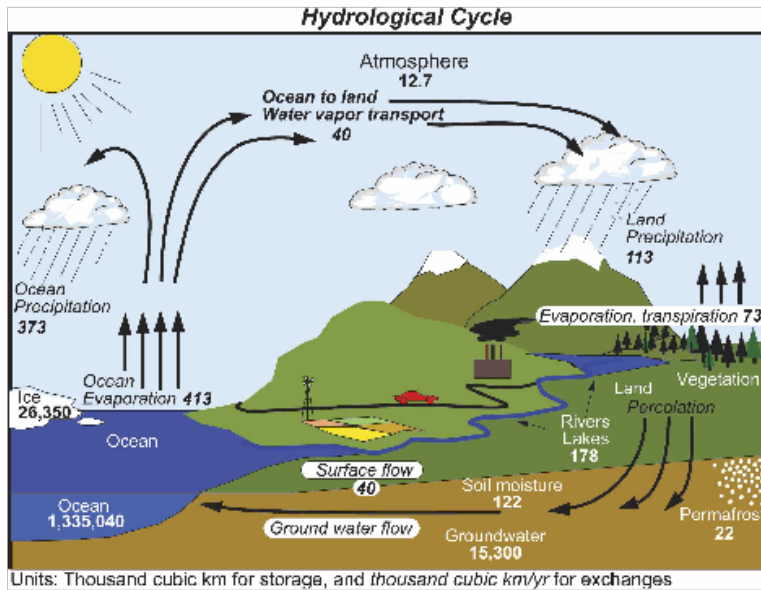
NCAS Computational Modelling Services (CMS)



NCAS Computational Modelling Services (CMS)

- CMS staff undertake NCAS activities in support of computational science (particularly High Performance Computing (HPC) and numerical modelling)
- Provide underpinning infrastructure for the UK academic atmospheric and polar science community to support climate, weather, and earth-system research.
- Provide training for scientists:
 - NCAS Introduction to Scientific Computing course
 - NCAS Introduction to Unified Model course
 - NCAS CF Python tools course

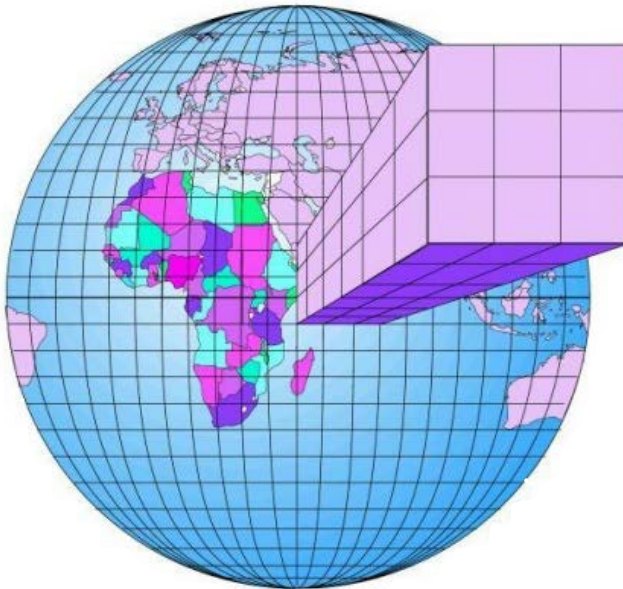
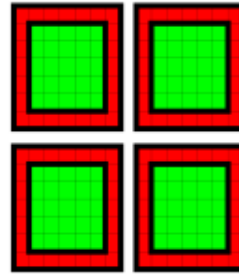
Climate Modelling



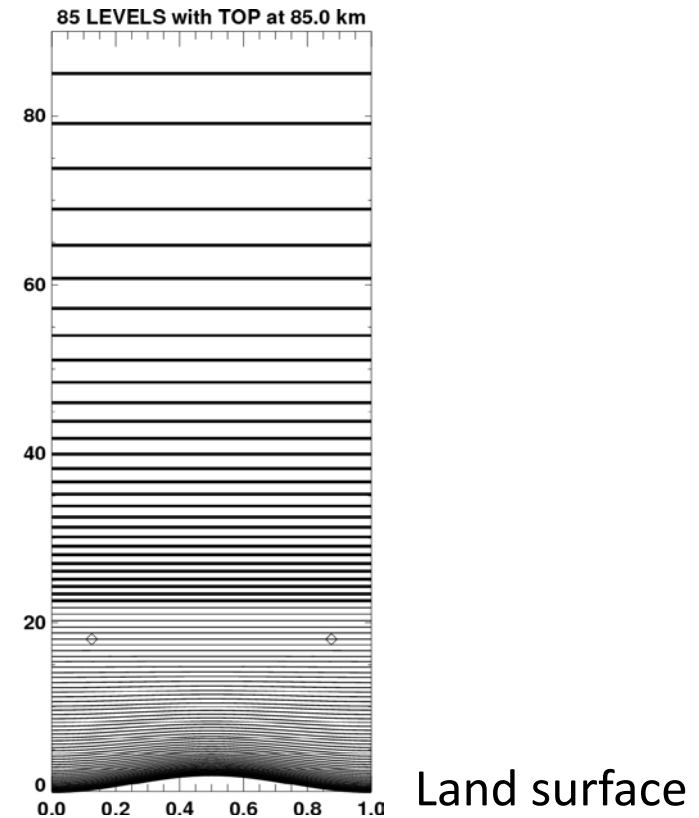
(Trenberth et al, 2007,2009)

PARALLEL Implementation

- Regular, Static, Lat-Long Decomposition
- Mixed mode MPI/OpenMP
- Asynchronous I/O servers
- Communications on demand for advection
- Multiple halo sizes



Horizontal resolution



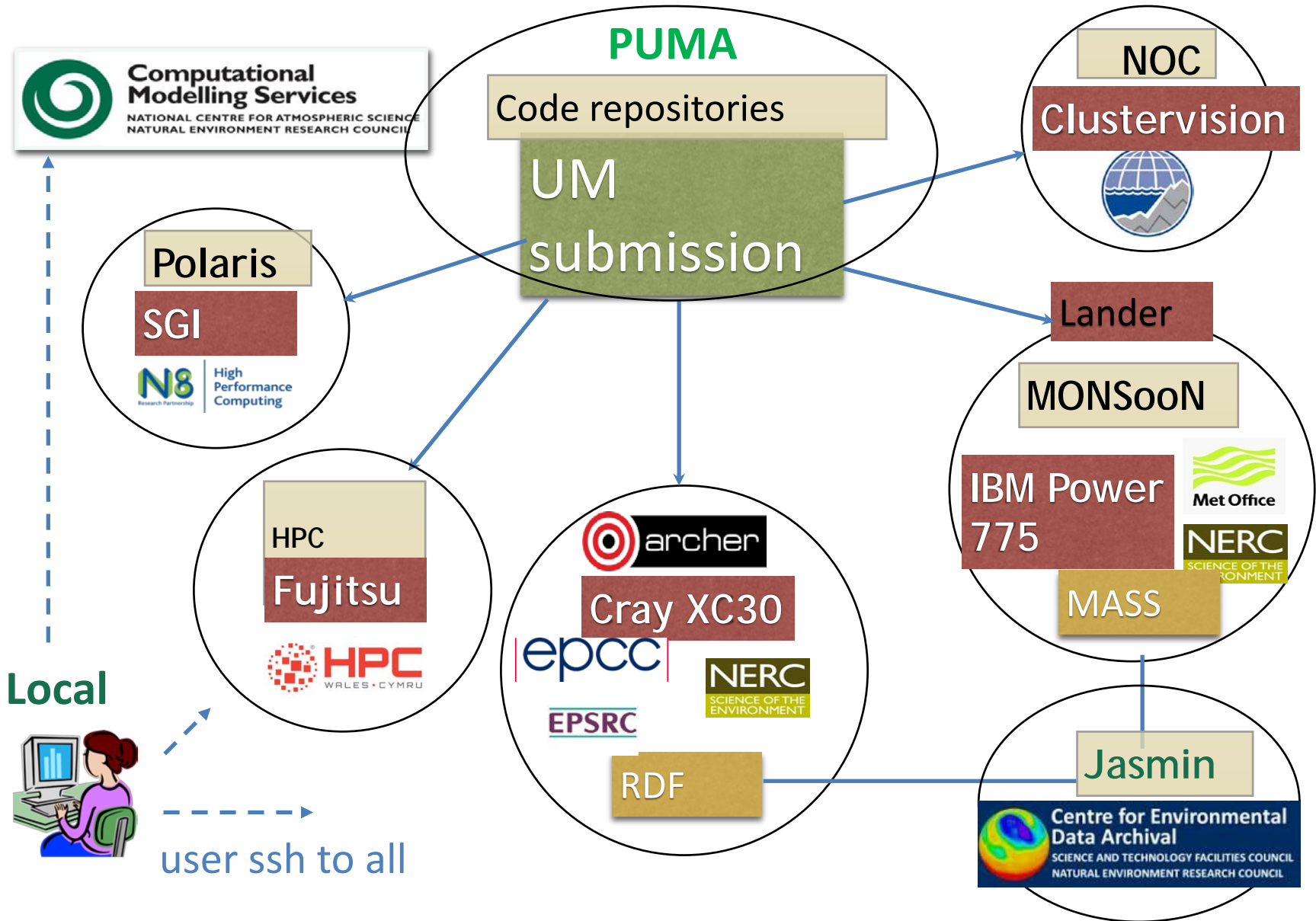
Vertical resolution

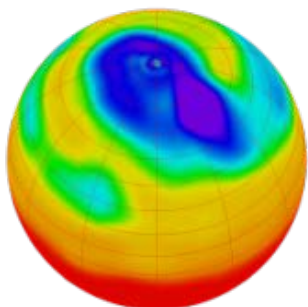
Global Models

N96	N144	N216	N320	N512	N768	N1024	N2048
(192 x 145)	(288 x 217)	(432 x 325)	(640 x 481)	(1024 x 769)	(1536 x 1152)	(2048 x 1536)	(4096 x 3073)
~135 km	~90 km	~60 km	~40 km	~25 km	~17 km	~12 km	~6 km

	NWP	Climate
Run length	<i>10 day operational forecast, 15 day ensemble forecast</i>	<i>Months (seasonal) Years, decades, centuries+</i>
Global resolution	<i>Testing: N320 (40 km) with 15 min ts</i> <i>Operational: N768 (17 km) with 7.5 min ts</i>	<i>Low resolution: N96 (135 km) with 20 min ts</i> <i>High resolution: N512 (25 km) with 15 min ts</i>
Dynamics	<i>Non-bit reproducible</i>	<i>Bit-reproducible</i>

NCAS supported MACHINES





Centre for Environmental Data Analysis

SCIENCE AND TECHNOLOGY FACILITIES COUNCIL
NATURAL ENVIRONMENT RESEARCH COUNCIL



Centre for Environmental
Data Analysis
SCIENCE AND TECHNOLOGY FACILITIES COUNCIL
NATURAL ENVIRONMENT RESEARCH COUNCIL



National Centre for
Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL



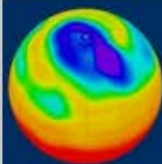
National Centre for
Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

NERC Data Centres

NERC supports five data centres covering a range of discipline areas:

- British Oceanographic Data Centre (Marine)
- Centre for Environmental Data Analysis Archive (Atmospheric and Earth Observation)
- Environmental Information Data Centre (Terrestrial and freshwater)
- National Geoscience Data Centre (Geoscience)
- Polar Data Centre (Polar and cryosphere)





Centre for Environmental Data Analysis

SCIENCE AND TECHNOLOGY FACILITIES COUNCIL
NATURAL ENVIRONMENT RESEARCH COUNCIL

[Contact us](#) | [CEDA Help](#) | [News](#)



[Login](#) | [Register](#)

► CEDA ► Data Centres

[Data search](#) Search this site

Go

[About CEDA](#)

[Data Centres](#)

[JASMIN](#)

[Projects](#)

[For Academics](#)

[For Business](#)

[Training](#)

[Contact Us](#)

[Help](#)

Data Centres

The Centre for Environmental Data Analysis is responsible for the running of the following data centres:

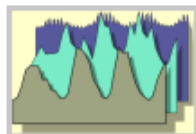
CEDA Archive

The CEDA Archive operated the atmospheric and earth observation data centre functions on behalf of NERC for the UK atmospheric science and earth observation communities. It covers climate, composition, observations and NWP data as well as various earth observation datasets, including airborne and satellite data and imagery. Prior to November 2016 these functions were operated by CEDA under the titles of the British Atmospheric Data Centre (BADC) and the NERC Earth Observation Data Centre (NEODC).



IPCC Data Distribution Centre

The [Intergovernmental Panel on Climate Change \(IPCC\)](#) DDC provides climate, socio-economic and environmental data, both from the past and also in scenarios projected into the future. Technical guidelines on the selection and use of different types of data and scenarios in research and assessment are also provided.



UKSSDC

The UK Solar System Data Centre (UKSSDC), co-funded by STFC and NERC, curates and provides access to archives of data from the upper atmosphere, ionosphere and Earth's solar environment.

www.ceda.ac.uk



Centre for Environmental Data Analysis
SCIENCE AND TECHNOLOGY FACILITIES COUNCIL
NATURAL ENVIRONMENT RESEARCH COUNCIL

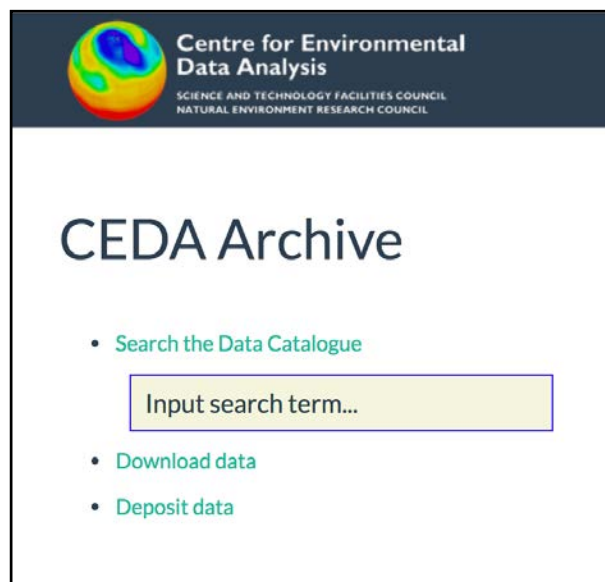


National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL



National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

CEDA Archive (July 2018)



Centre for Environmental Data Analysis
SCIENCE AND TECHNOLOGY FACILITIES COUNCIL
NATURAL ENVIRONMENT RESEARCH COUNCIL

CEDA Archive

- Search the Data Catalogue

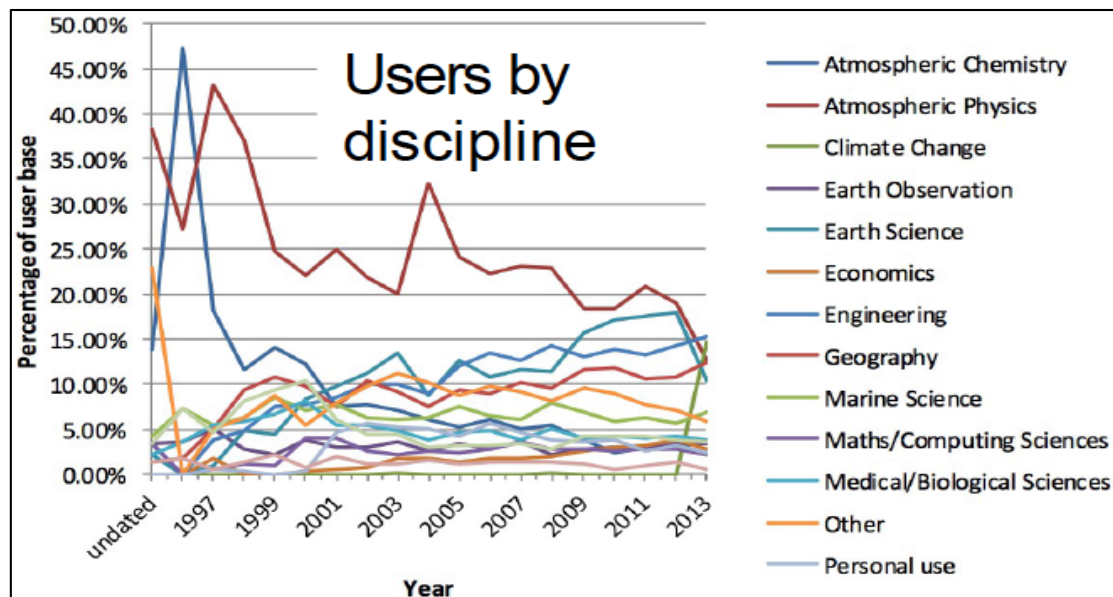
- Download data
- Deposit data

~ 556 datasets collections,
~5272 datasets

~ 154 million files

> 44,000 registered users

Data Type	Data Volume (Petabytes)
Earth Observation	4
Atmospheric Science	2
Total	6PB



http://catalogue.ceda.ac.uk

precipitation

Search

2028 Results

Sort by Relevance

Filter by record type

☐ Datasets (1898)

☐ Dataset Collections (33)

☐ Projects (31)

☐ Instruments (26)

☐ Platforms (20)

☐ Computations (20)

Met Office

UKCP09: Met Office gridded land surface climate observations - precipitation and temperature indices at 5km resolution

View parent collections

NASA

Global Precipitation Measurements (GPM) Integrated Multi-satellite Retrievals (IMERG) L3 Half Hourly 0.1 degree x 0.1 degree

View parent collections

NASA

CMORPH 0.25 degree daily precipitation estimates

View parent collections

Register/Login for access

Explore

More Info

Register/Login for access

Explore

More Info

Register/Login for access

Explore

More Info



What is the role of CEDA?

- **Preserve the science record** (for data produced by NERC funded research).
- **Facilitate data use** (for any data that compliments NERC research in Atmospheric sci or EO).
- **Support data standards** (for international science community), e.g. ESA Climate Change Initiative; CMIP data request; Climate-Forecast (CF) metadata convention.
- **Engage with global community.** CEDA staff participate in:
 - World Climate Research Programme (WCRP);
 - Committee on Earth Observation Satellites (CEOS);
 - Earth System Grid Federation (ESGF);
 - Intergovernmental Panel on Climate Change (IPCC) Data Distribution Centre.

CEDA Projects

[About CEDA](#)[Data Centres](#)[Services](#)[Projects](#)[For Academics](#)[For Business](#)[Training](#)[Contact Us](#)[Help](#)

Projects



Characterisation of metadata to enable high-quality climate applications and

services - CHARMe

CHARMe is a 2 year FP7 funded project aiming to link commentary metadata (e.g. annotations, supporting information about the data) and datasets. The project will deliver repositories of commentary metadata with interfaces for users to populate and interrogate the information. This will enable users to assess if the of climate data are fit for purpose.

CEDA is working with 8 other UK and European partners, and has key roles on the data model, software development, implementation in archives, and application to climate services.



InfraStructure for the European Network for Earth System Modelling - Phase 2 (IS-ENES II)

IS-ENES II is a FP7-Project, funded by the European Commission under

Climate Information Portal for Copernicus (CLIPC)

The CLIPC platform will complement existing GMES/Copernicus pre-operational components by providing access on decadal to centennial climate variability data to a wide variety of users. The data will include satellite and in-situ observations, climate models and re-analyses, transformed data products to enable impacts assessments and climate change impact indicators. Supporting data quality and related information will also be made available.

CEDA is leading the project, coordinating a consortium of 22 partners, and leads the access to climate data work package. This work package will provide the software infrastructure to create a single point of access for climate model data from various sources: climate model data, in situ and satellite observations, and re-analyses.



ESPAS - Near-Earth Space Data Infrastructure for e-Science

The ESPAS project aims to provide e-infrastructure necessary to support the access to observations, modelling and prediction of the Near-Earth

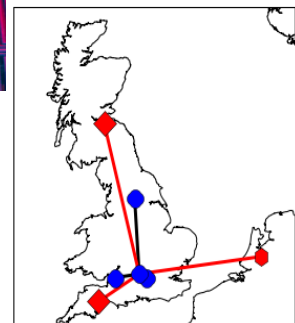
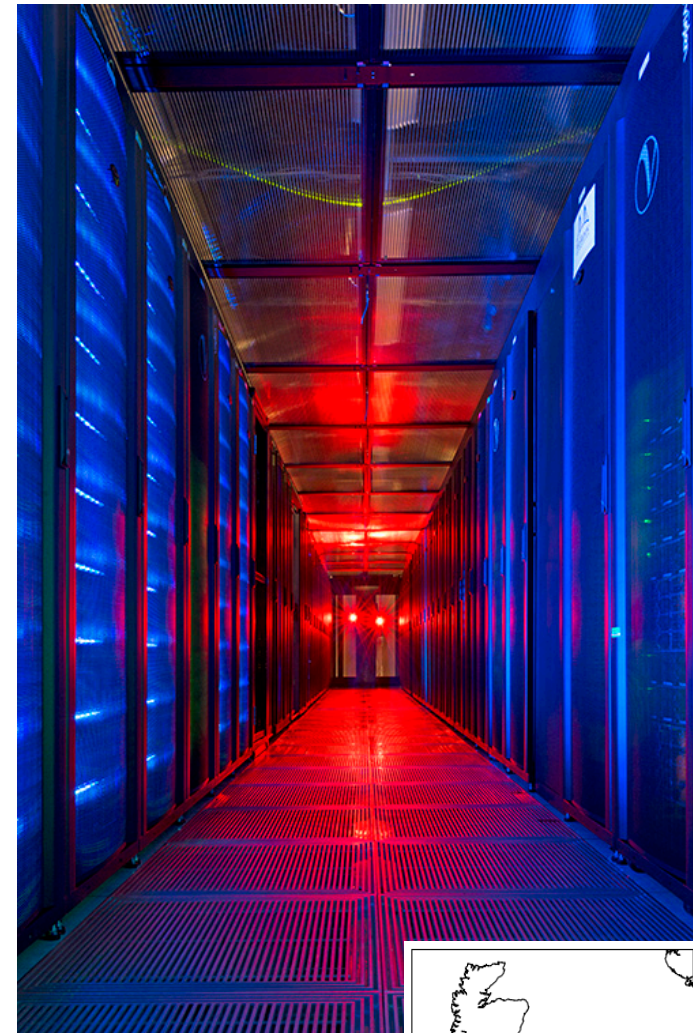
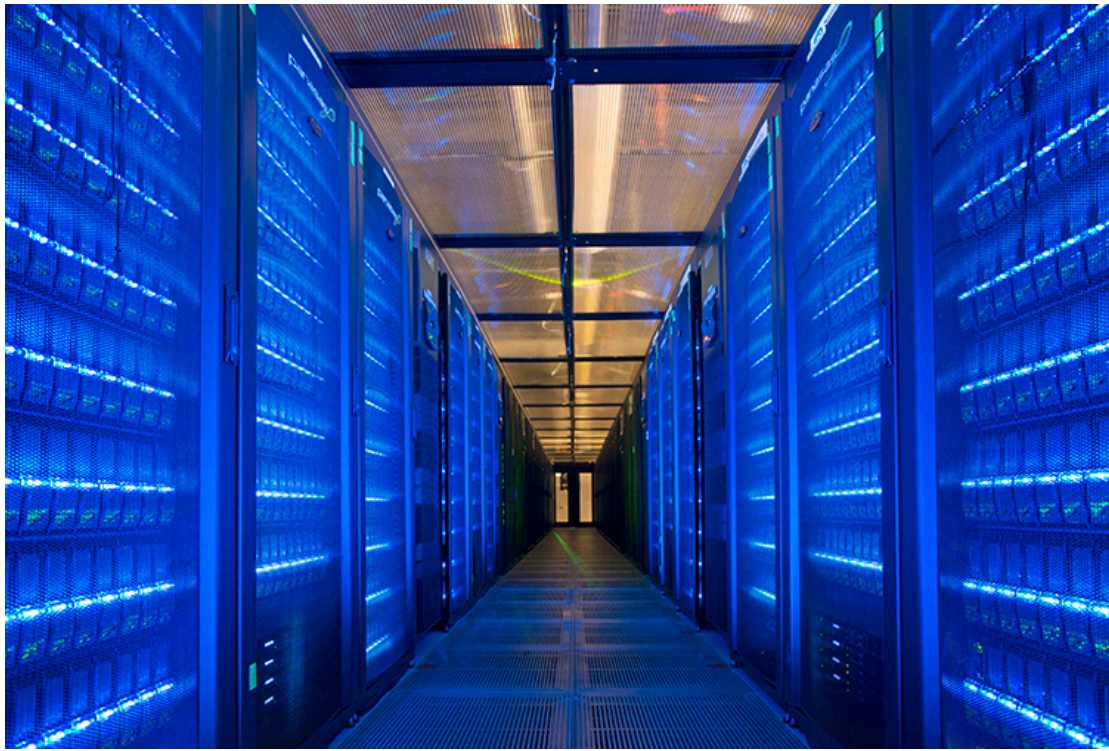
Lots
more

JASMIN Overview



The JASMIN facility is a "super-data-cluster" which delivers infrastructure for data analysis.

In technical terms it is half super-computer and half data-centre and it provides a globally unique computational environment.



- JASMIN is a world leading, unique hybrid of:
- 16PB high performance storage ($\sim 250\text{GByte/s}$)
- High-performance computing ($\sim 4,000$ cores)
- Non-blocking Networking ($> 3\text{Tbit/sec}$), and Optical Private Network WAN's
- Coupled with cloud hosting capabilities

To address “one of NERC’s most strategically important challenges: the improvement of predictive environmental science.” Prof. Duncan Wingham, NERC Chief Exec.

Hosted by STFC Scientific Computing Department

***“Computing Expertise across
length scales from processes within
atoms to environmental modelling”***

- Applications development and support,
- Compute and data facilities and services
- Research and Training
- Numerical Analysis

Data Services

- STFC: Facility Archives (ISIS, Diamond)
- LHC: UK Hub (Tier 1 archive)
- BBSRC: Institutes data archive
- MRC: Data Support Service
- NERC: CEDA backup and JASMIN elastic tape



High Performance Computing

- Emerald GPU cluster for Oxford, UCL, Southampton, Bristol.
- SCARF HPC for RAL
- Hartree: Blue Joule bluegene HPC
- Hartree: Blue Wonder idataplex HPC
- JASMIN: NERC super data cluster

Close working partnership with industry

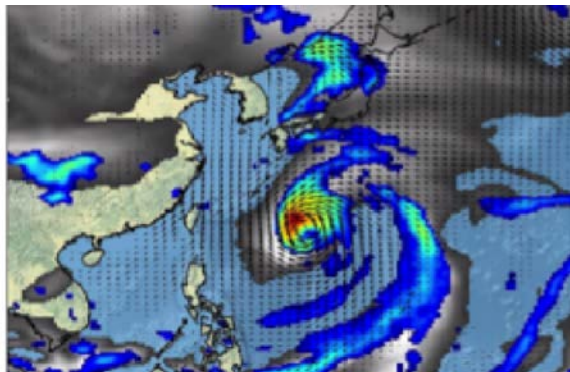




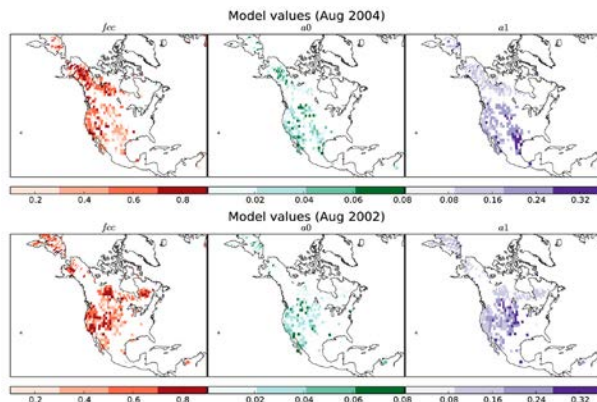
Processing big data: the issues

- Parallel processing in the Environmental Sciences has historically focussed on running highly-parallelised models.
- Data analysis was typically run sequentially because:
 - It was a smaller problem
 - It didn't have parallel resources available
 - The software/scientists were not equipped to work in parallel
- The generation of enormous datasets (e.g. UPSCALE – around 300Tb) means that:
 - Processing big data **requires** a parallel approach
 - Fortunately, platforms, tools, and programmers are becoming better equipped

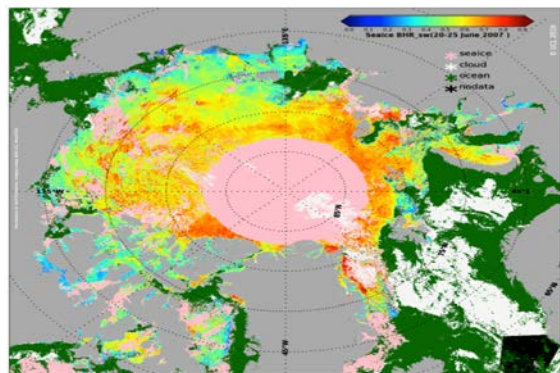
~150 Science projects on JASMIN to date



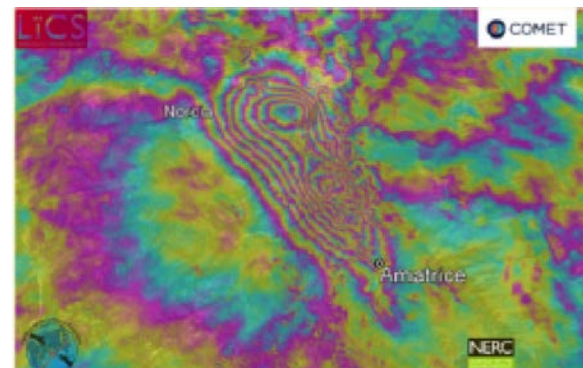
High Res Climate Model analysis



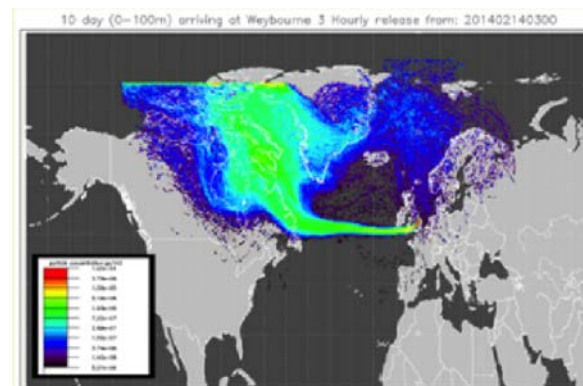
Deriving the impact of fire on vegetation from earth observation data



Climate variables from European and US instruments/satellites



Fault analysis



Atmospheric dispersion

JASMIN in pictures

JASMIN

jasmin-login1

SSH login gateway

jasmin-xfer1

Data transfers

Key:



General-purpose resources



Project-specific resources



Data centre resources

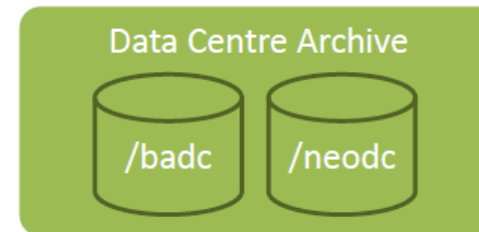
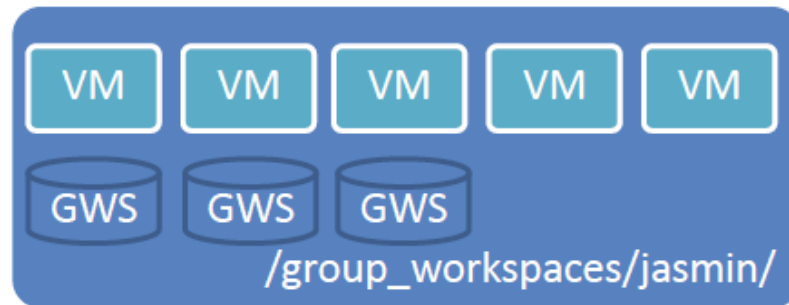
firewall

jasmin-sci1

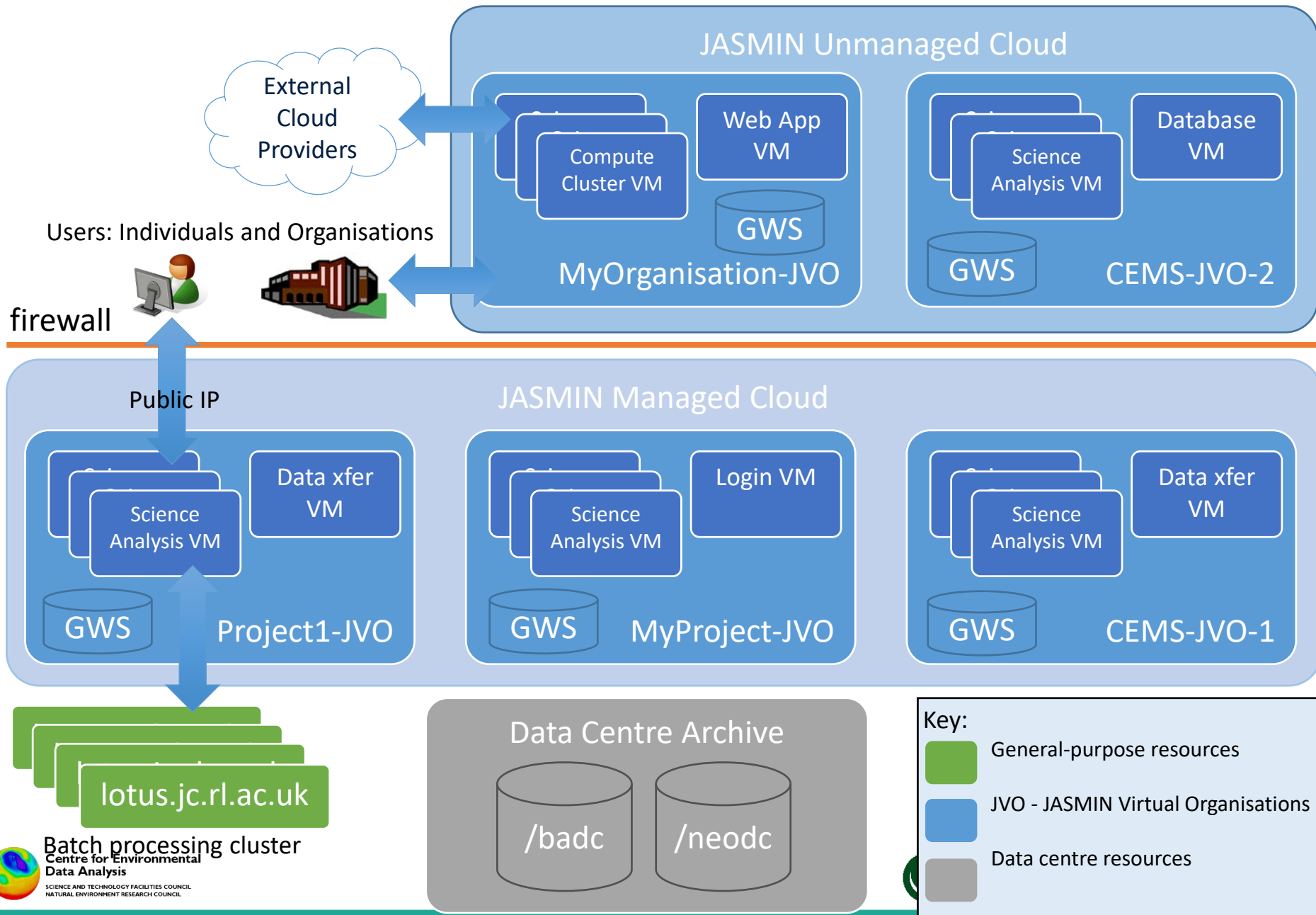
Science/analysis

lotus.jc.rl.ac.uk

Batch processing cluster



JASMIN Cloud Infrastructure



Internal network: vital to JASMIN / CEMS performance

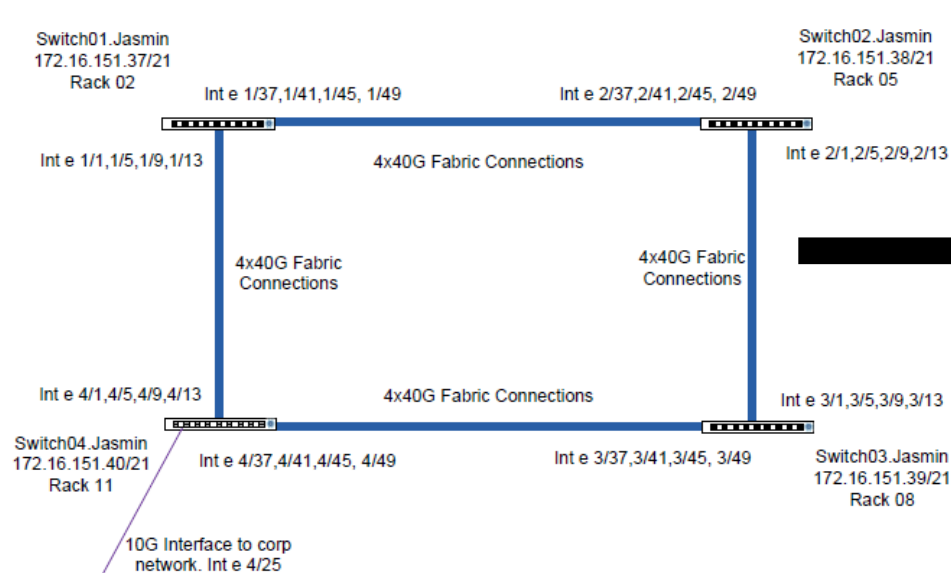
172.16.144.0/21 = 2,000 IPs

130.246.136.0/21

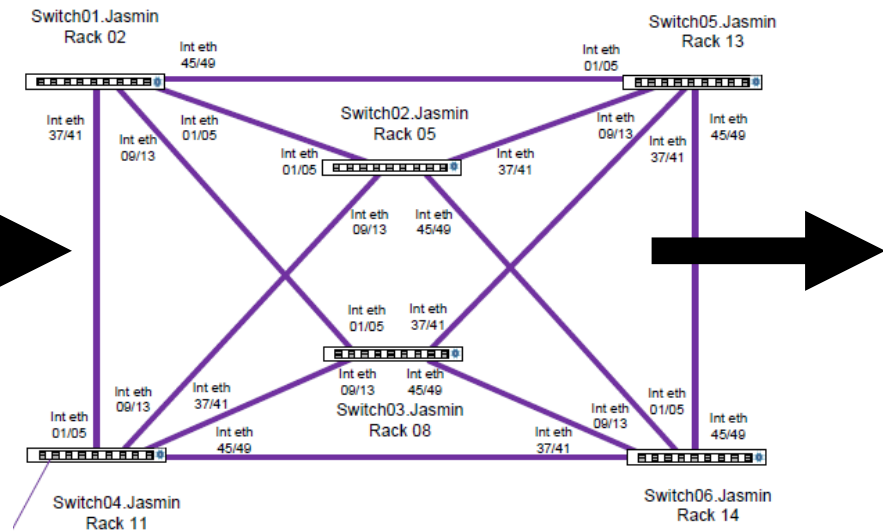
Flat Overlaid L2

160->240 Ports @ 10Gb

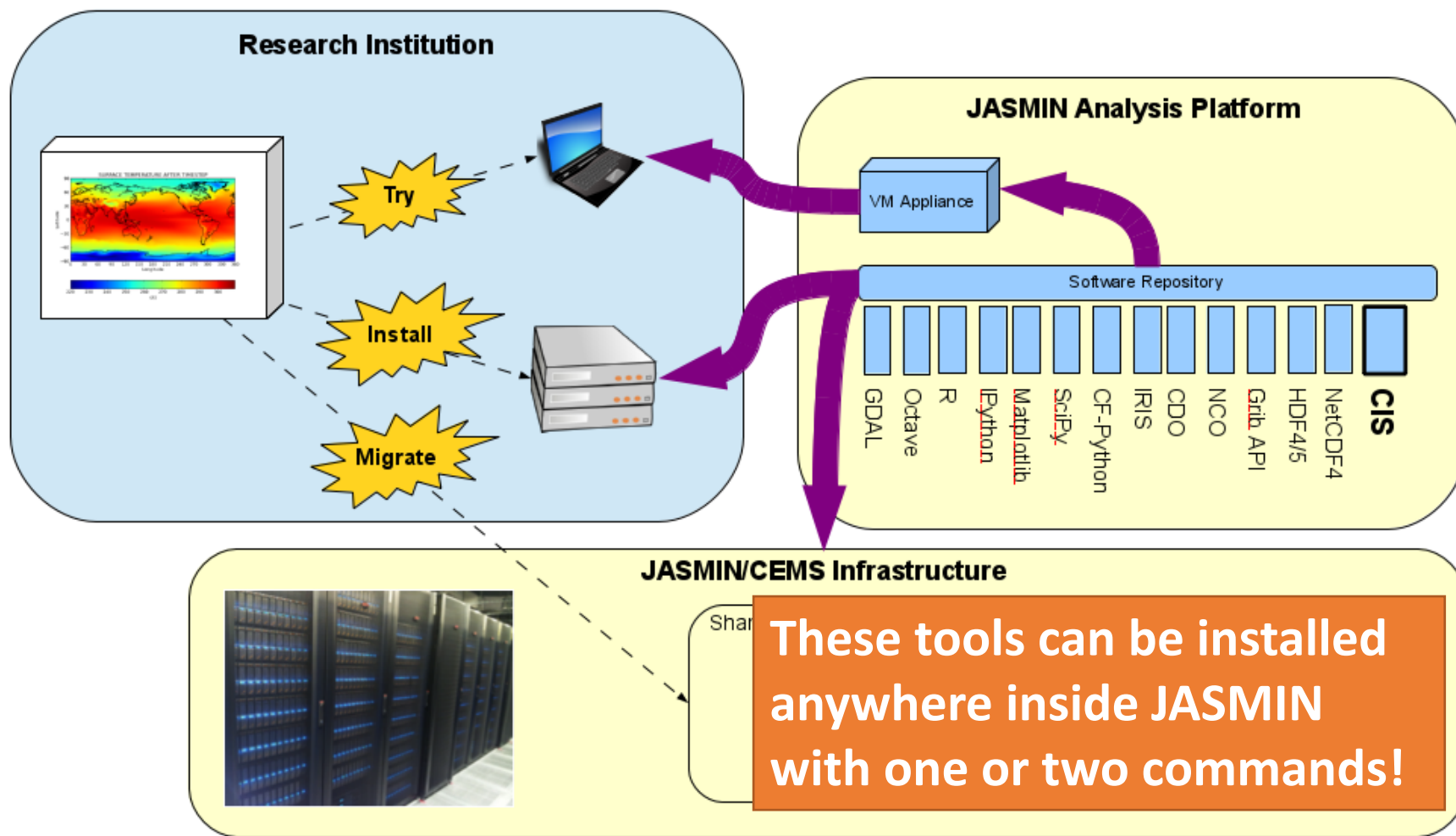
RAL Jasmin/CEMS Gnodal Physical
Topology



RAL Jasmin/CEMS Gnodal Physical
Topology Installed July 2012



The “JASMIN Analysis Platform”: a re-usable, re-deployable bundle of common tools



Further Information

- NCAS website: ncas.ac.uk
- CMS website: cms.ncas.ac.uk
- CEDA website: ceda.ac.uk
- JASMIN website: jasmin.ac.uk