UNIVERSITY OF CAMBRIDGE
Yusuf Hamied
Department of Chemistry

Using the JASMIN Unmanaged Cloud for UKCA Training

Luke Abraham
NCAS Cambridge
n.luke.abraham@ncas.ac.uk

With thanks to Richard Smith, Matt Pryor, Steven Sharpe, Frances Dee, Louise Whitehouse, Scott Archer-Nicholls, Thomas Aubry, James Weber, Mohit Dalvi, Lauren Marshall, Zosia Staniaszek, Alex Archibald, & Bryan Lawrence

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Overview

- What is UKCA?
- UKCA Training
- Virtual Machine
- Setting-up on JASMIN
- Training this year

What is UKCA?

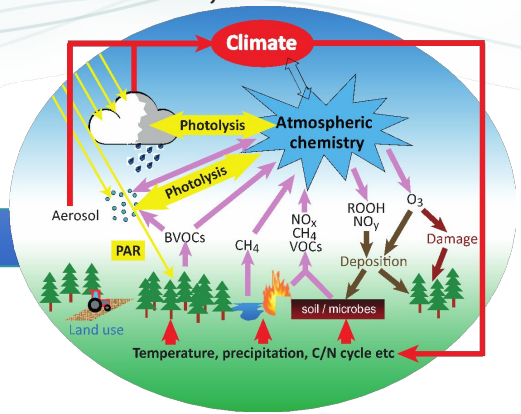
UKCA is an atmospheric composition model, currently built as part of the Met Office's Unified Model (UM).

UKCA is not a *particular* collection of chemistry and aerosol schemes, but is a **framework** for putting chemistry and aerosol schemes into the UM.

It was originally designed for long climate simulations, but is now also used for air quality forecasts.

It is part of the UKESM1 Earth system model, used internationally.

Unified Model



UKCA

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

UKCA Training

Training courses covering the use of UKCA have been run since 2013, and have been supported by NERC and NCAS.

Training is available at UM versions 8.2, 8.4, 10.4, 10.9, & 11.8 on the UKCA website, **`www.ukca.ac.uk`**

Students are asked to solve one big task, broken down into a series of separate tutorials:

Create two new species, ALICE and BOB, then add in emissions of ALICE and the reaction

ALICE + OH ⟶ BOB + Secondary Organic Compound (Sec_Org)

before adding in the dry deposition of ALICE and the wet deposition of BOB. You should also output the fluxes through the reaction and deposition processes.

Later you will look at several aerosol diagnostics.

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

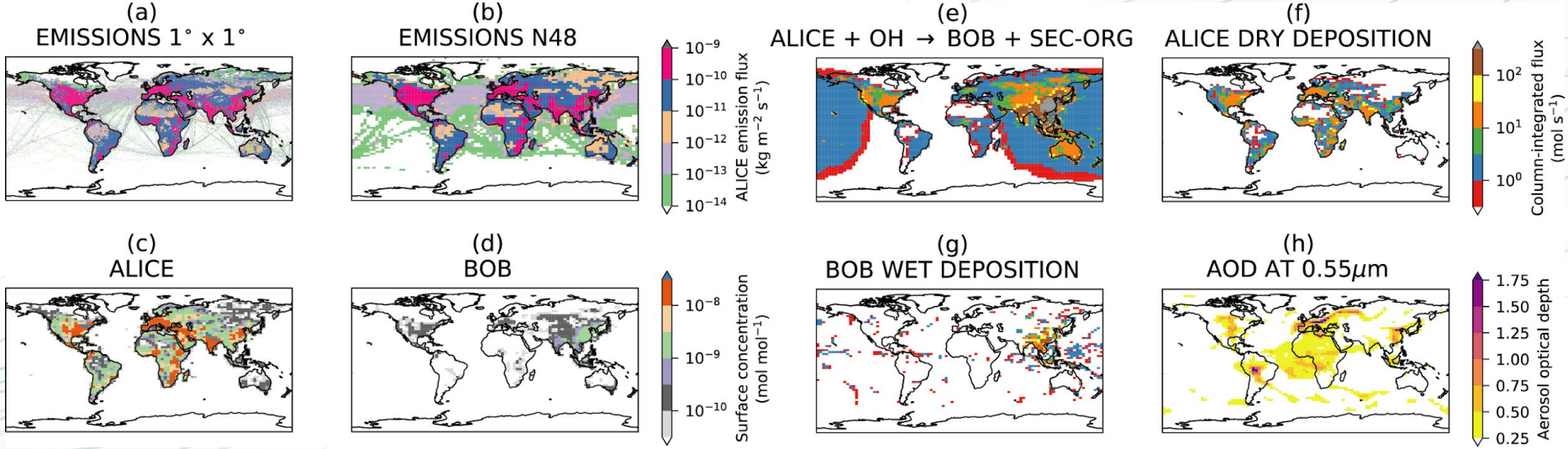
UKCA Training

These tasks cover most of what new users might want to do with UKCA:

- Adding new chemical species
- Adding new emissions
- Adding new chemical reactions & deposition
- Defining new model output
- Regridding input data
- Creating NetCDF input files
- Processing & viewing model output

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

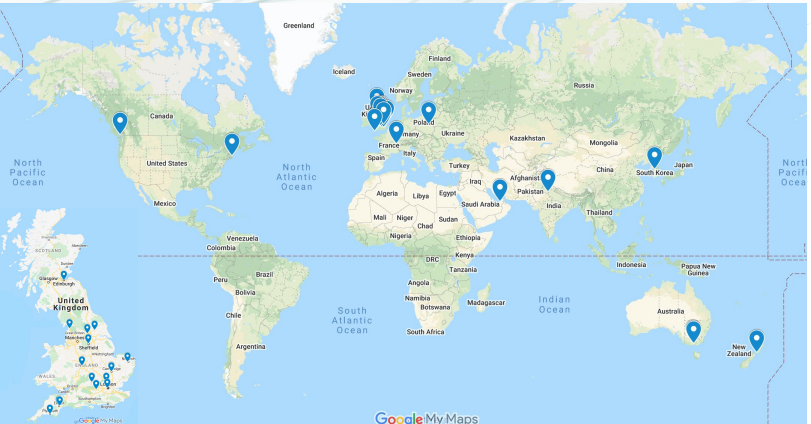
UKCA Training



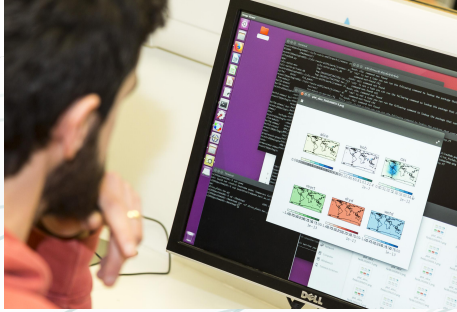
National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

UKCA Training

- Face-to-face training has been held in Cambridge every January, supported by NCAS
- These courses made use of the PUMA service and ARCHER HPC to run the model
- Over 120 people have attended in-person training events, from 27 institutions from around the world



The vn10.9 & vn11.8 tutorials have also been developed for use on the Met Office Virtual Machine



National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

UKCA Training using PUMA/ARCHER

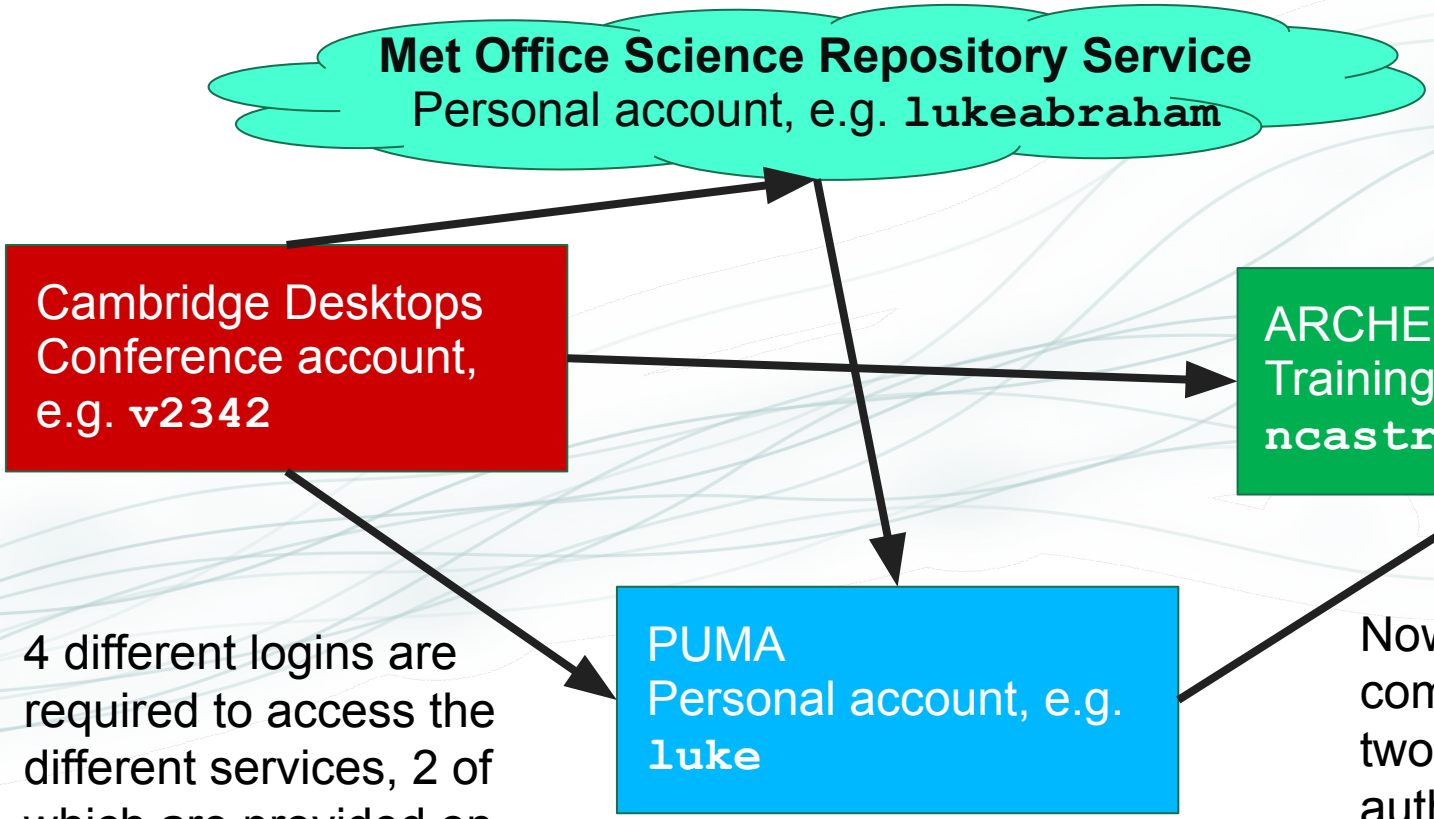
As the UM is installed on ARCHER (& now ARCHER2), superficially this seems like a good system to use for UKCA Training.

- No technical set-up work is required for the UM
- Students can familiarise themselves with using the HPC and in using UM suites similar to what they would use in their research

However…

- Compile times on ARCHER were long due to the use of shared nodes, meaning that a “manual compile” option had to be implemented
- Several accounts and servers are used, leading to numerous setup issues and confusion
- Often the difficulties encountered by students were to do with connecting to or using ARCHER, not UKCA

Workflow



ARCHER to ARCHER2 transition also a concern for 2021 course

4 different logins are required to access the different services, 2 of which are provided on the day of the course

Now made even more complicated by two-factor authentication

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Met Office Virtual Machine

- A Met Office Virtual Machine has been developed for running FCM, Rose, & Cylc
- Uses VirtualBox and Vagrant
- UM Systems Team have set-up running the UM in an Ubuntu guest image.

Geosci. Model Dev., 11, 3647–3657, 2018
https://doi.org/10.5194/gmd-11-3647-2018

Geoscientific Model Development
Open Access
EGU

Using a virtual machine environment for developing, testing, and training for the UM-UKCA composition-climate model, using Unified Model version 10.9 and above

Nathan Luke Abraham[1,2], Alexander T. Archibald[1,2], Paul Cresswell[3], Sam Cusworth[3], Mohit Dalvi[3], David Matthews[3], Steven Wardle[3], and Stuart Whitehouse[3]

[1]Department of Chemistry, University of Cambridge, Cambridge, CB2 1EW, UK
[2]National Centre for Atmospheric Science, UK
[3] Met Office, FitzRoy Road, Exeter, EX1 3PB, UK

Correspondence: Nathan Luke Abraham (luke.abraham@atm.ch.cam.ac.uk)

Received: 8 May 2018 – Discussion started: 18 May 2018
Revised: 10 August 2018 – Accepted: 23 August 2018 – Published: 6 September 2018

Virtual Machine on GitHub:

`https://github.com/metomi/metomi-vms`

See UM Documentation Paper X10 for more information on using this system with the UM, and Abraham *et al.* (2018) for its use with UKCA training.

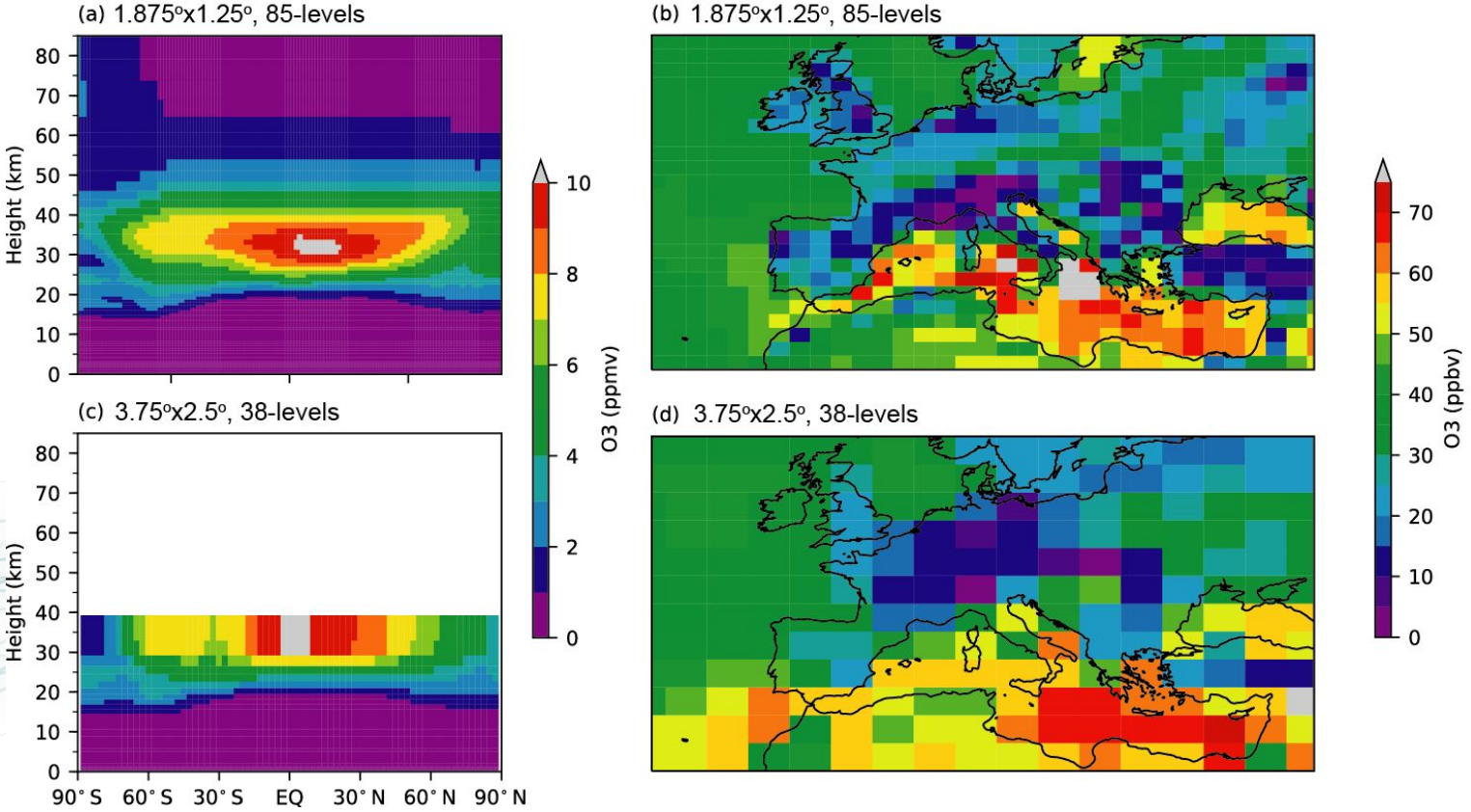
Has also been run on Azure cloud.

Azure

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Climate resolution

VM-suitable resolution



By lowering the resolution and model top, turning off some diagnostic sections, and reducing output, it is possible to run a UKCA configuration on a VM with 2x CPUs with 6GB of RAM

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Setting-up Training VMs on JASMIN

While the VM can be run on a PC or in the Cloud, there are a number of technical limitations that need to be overcome.

The current VM uses Vagrant & VirtualBox to configure the VM, with a lot of manual set-up to install the UM and required input files once the VM has been provisioned.

Richard Smith at CEDA ported this set-up for a single VM into Ansible so that it could be run on the UMC. I then extended this set-up further to

1. Pre-install the UM, provide the required input files, and configure the Iris Python package as required for some of the tasks
2. Adapt the Ansible scripts to enable multiple VMs to be provisioned at once, and also to create login and NFS servers

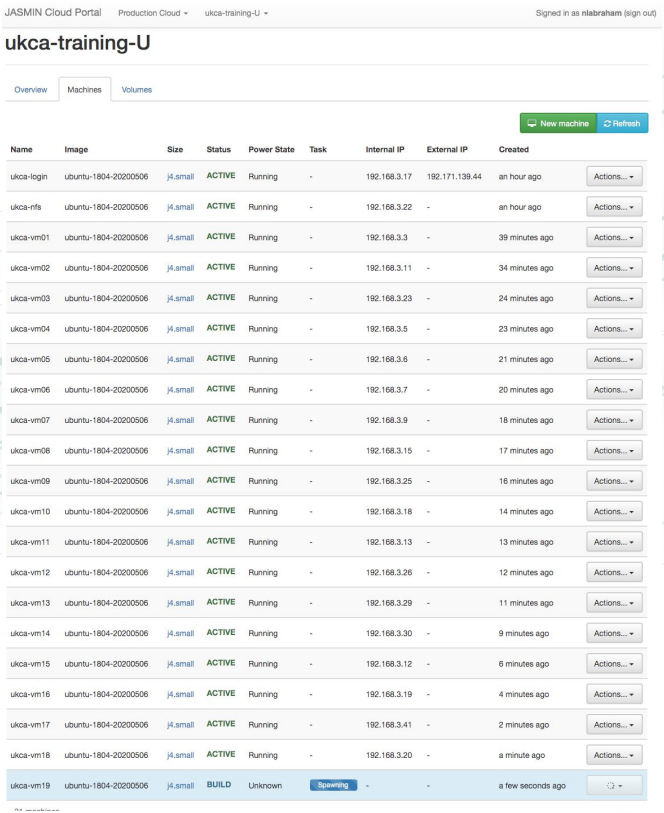
Setting-up Training VMs on JASMIN

Required 21 j4.small (2 CPUs, 8 GB RAM) VMs and 1 data volume:

25GB volume for UM install
1x NFS server
1x login server
19x VMs (14 for students, 5 for demonstrators)

= 42 CPUs & 172GB of RAM

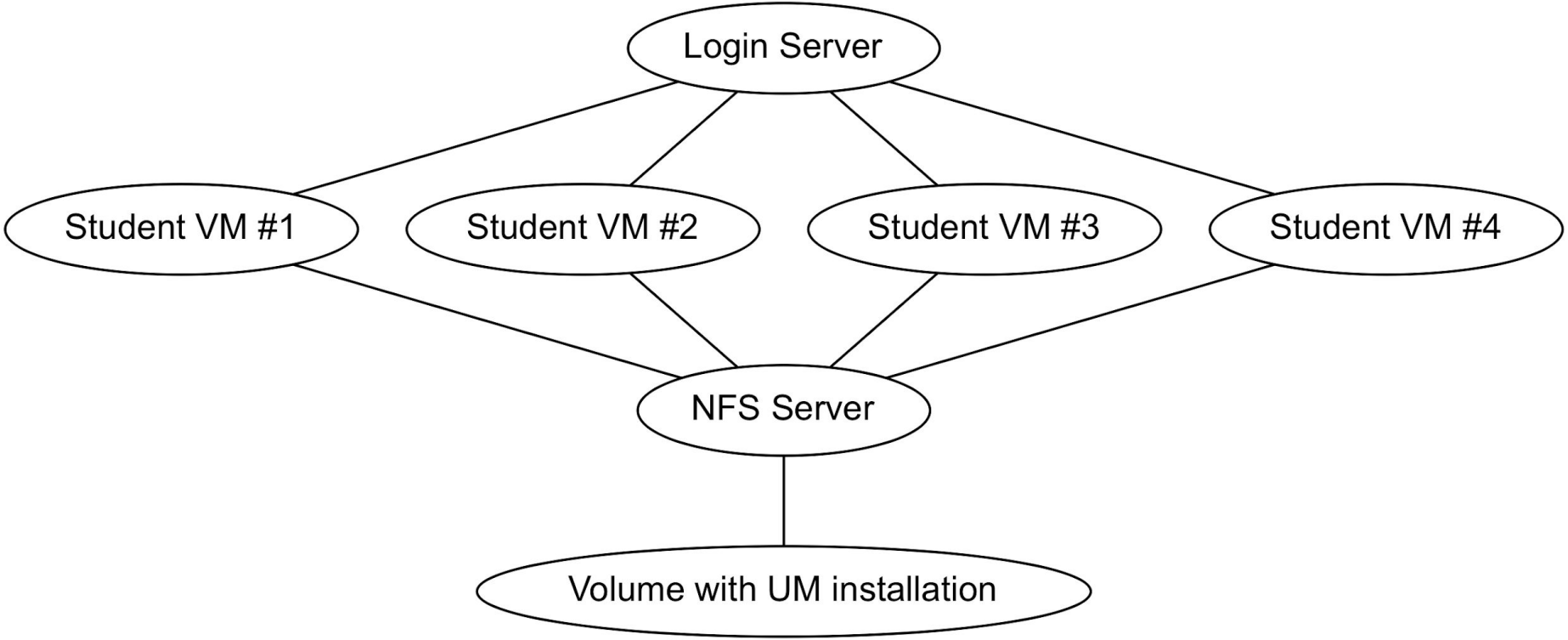
Once UM has been installed on the data volume, the VMs can be spun-up and spun-down as required in an hour or so.



Once the course finished, the VMs were shut down and the resources released back to the UMC.

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Setting-up Training VMs on JASMIN



Ansible playbooks for this system are available via GitHub:
`https://github.com/theabro/ukca-playbook`

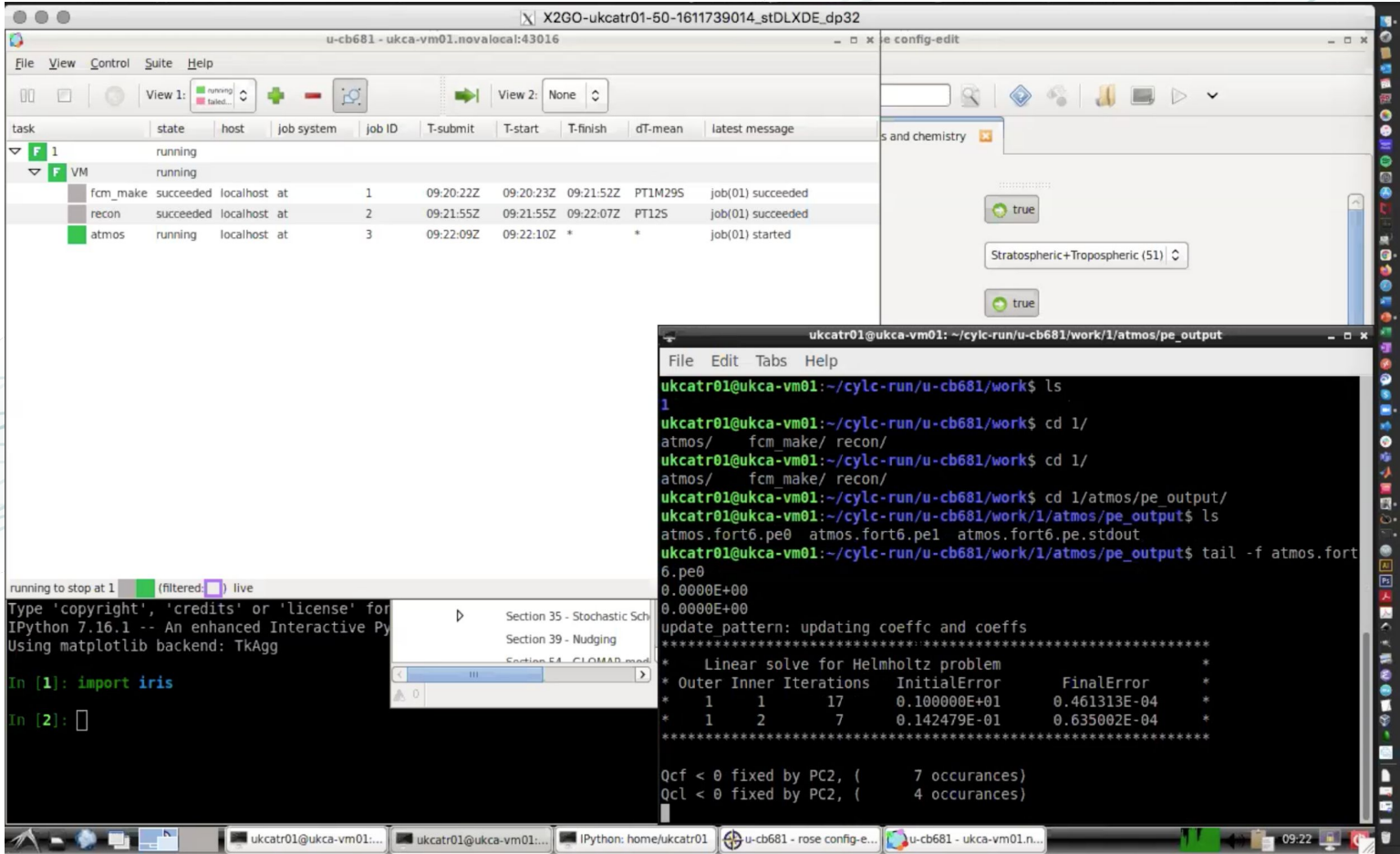
National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Setting-up Training VMs on JASMIN

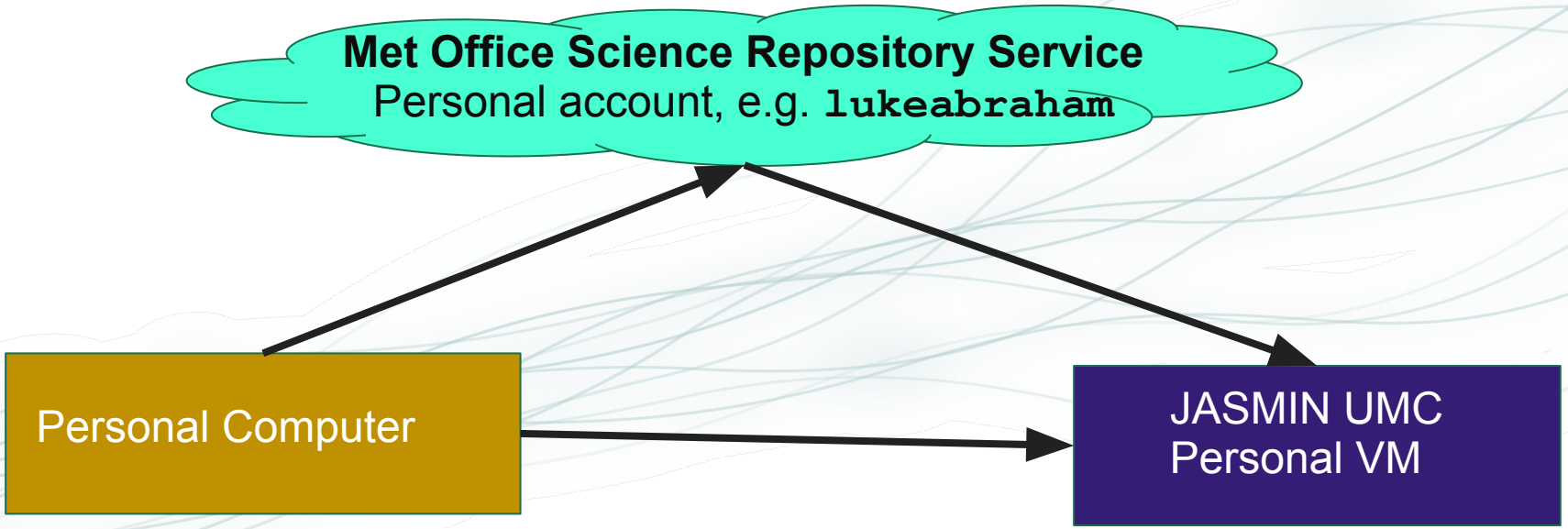
https://www.youtube.com/watch?v=5V3RBCYTQvg

Students could connect to the VMs using a number of methods - e.g. X2Go, MobaXTerm, or Terminal/X11

Set-up video with demonstration available on the UKCA YouTube channel.



New! Workflow on JASMIN



Now, the only username/password that needs to be remembered is for MOSRS.
The VMs use SSH keys that are provided to the students.

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

UKCA Training in 2021

Due to the pandemic, UKCA training was delivered online in February 2021.

There were 14 participants, with 5 demonstrators (3/4 at any one time).

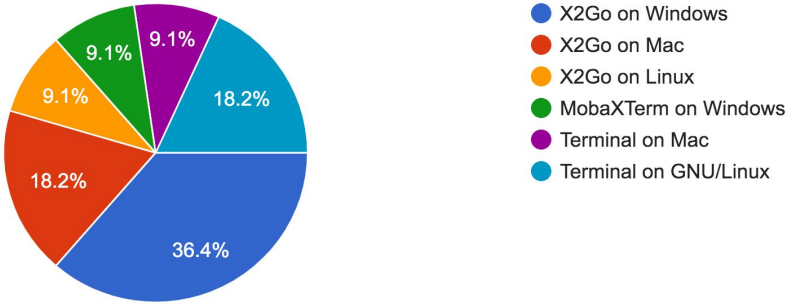
Course was delivered in a mixture of self-guided time, Zoom sessions, & support via Slack

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Feedback

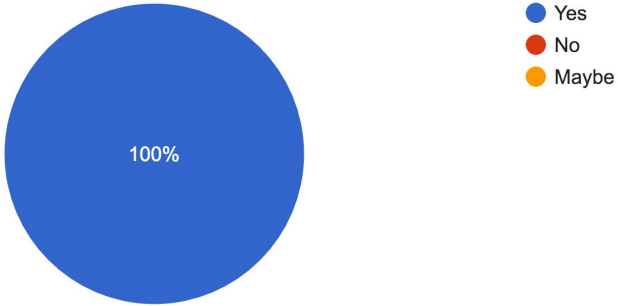
How did you connect to your Virtual Machine?

11 responses



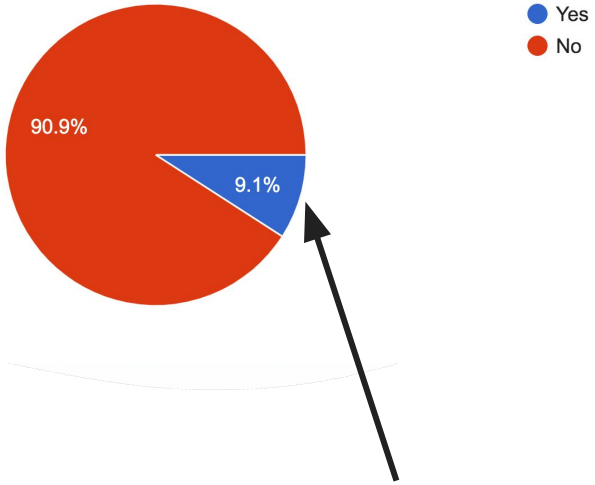
Did you find the Virtual Machine easy to use?

11 responses



Did you have any problems using your Virtual Machine?

11 responses



Problems with cutting/pasting text between VM and host

National Centre for Atmospheric Science

NATURAL ENVIRONMENT RESEARCH COUNCIL

Conclusions

- The JASMIN Unmanaged Cloud is an excellent platform to run online training events
- Effort is required to configure the VMs as required for the needs of the course
- Once done, the experience for the students is far better than when using a service such as ARCHER
- The use of Ansible playbooks means that the environment used can be easily saved and repeated as needed for future courses