UNIVERSITY OF CAMBRIDGE
Yusuf Hamied
Department of Chemistry

How to do virtually everything with UKCA

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

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

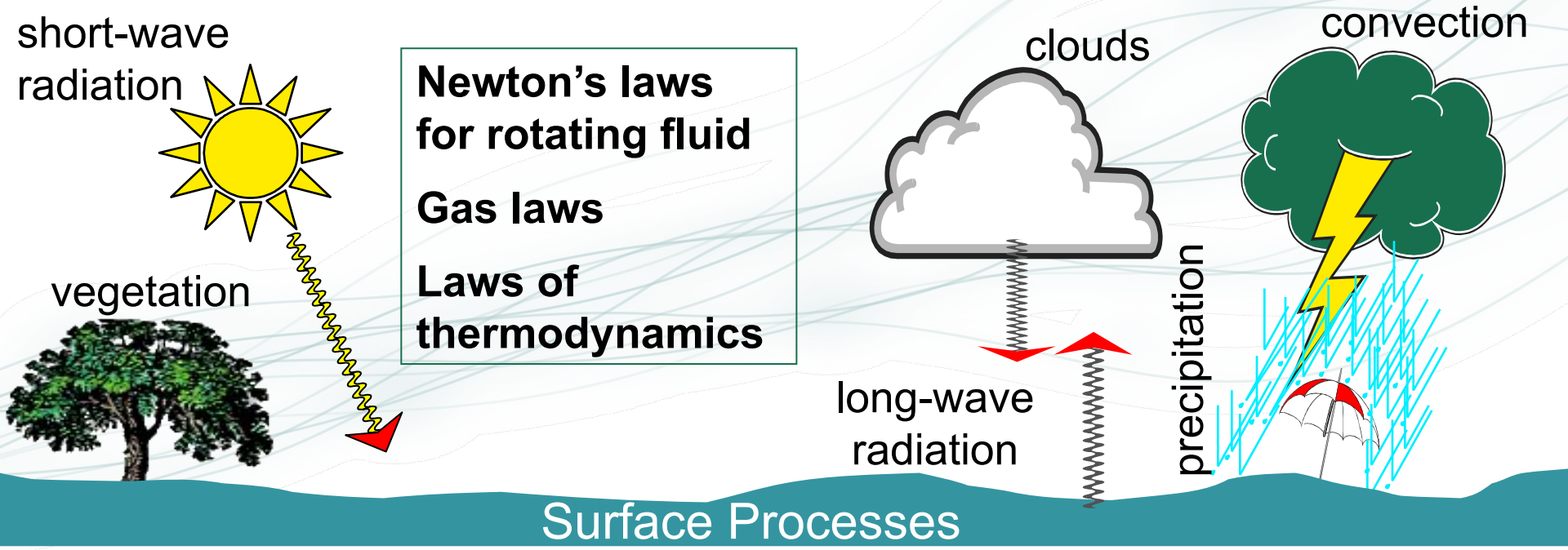
- What is the Unified Model, and how is code developed for it?
- What do I mean by *“virtually”*?
- What do I mean by *“virtually everything”*?
- What do I mean by *“everything”*?

What is the Met Office Unified Model?

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Atmospheric Modelling: integrating our knowledge of atmospheric behaviour forward in time



- The challenge: To reproduce the behaviour of (hazardous) weather systems

Unified Model

Brown et al. (2013)

- Operational forecasts
  - Mesoscale (resolution approx. 4km, 1.5km)
  - Global scale (resolution approx. 17km)
- Seasonal predictions
  - Resolution approx. 60km
- Global and regional climate predictions
  - Resolution around 120km
  - Run for 10-100-… years
- Research mode
  - Resolution 1km - 10m

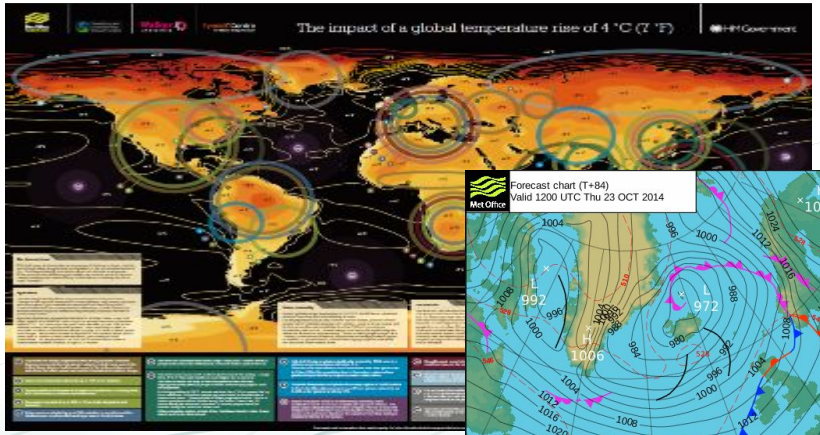
> 25 years old

Met Office

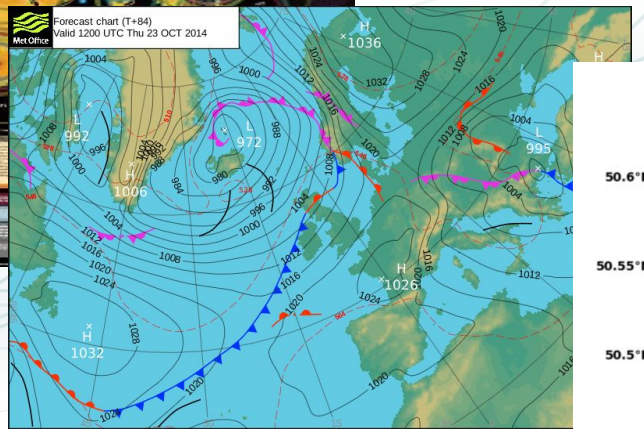
National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

The consequence of unification

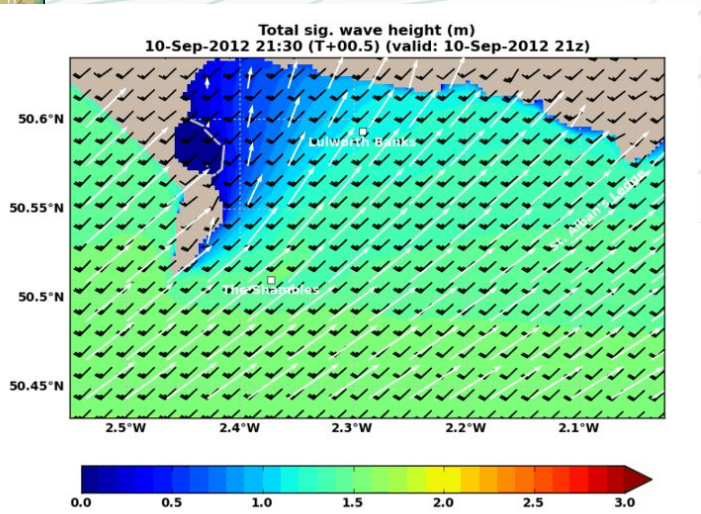
300 km



30 km



300 m



A factor of 1000 between these

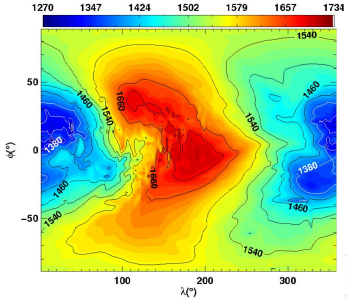
…the same scheme has to continue to work

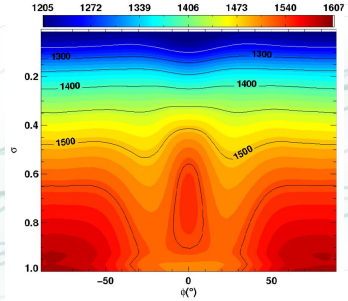
Met Office

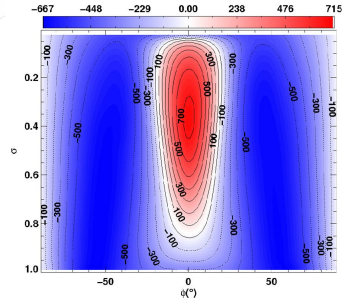
National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

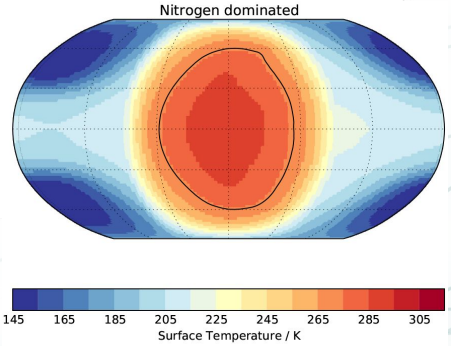
Exoplanets

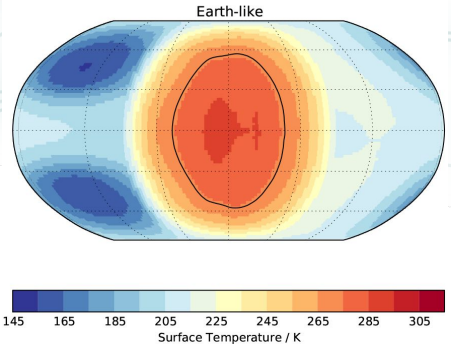
The unified model, a fully-compressible, non-hydrostatic, deep atmosphere global circulation model, applied to hot Jupiters
Mayne *et al.* 2014











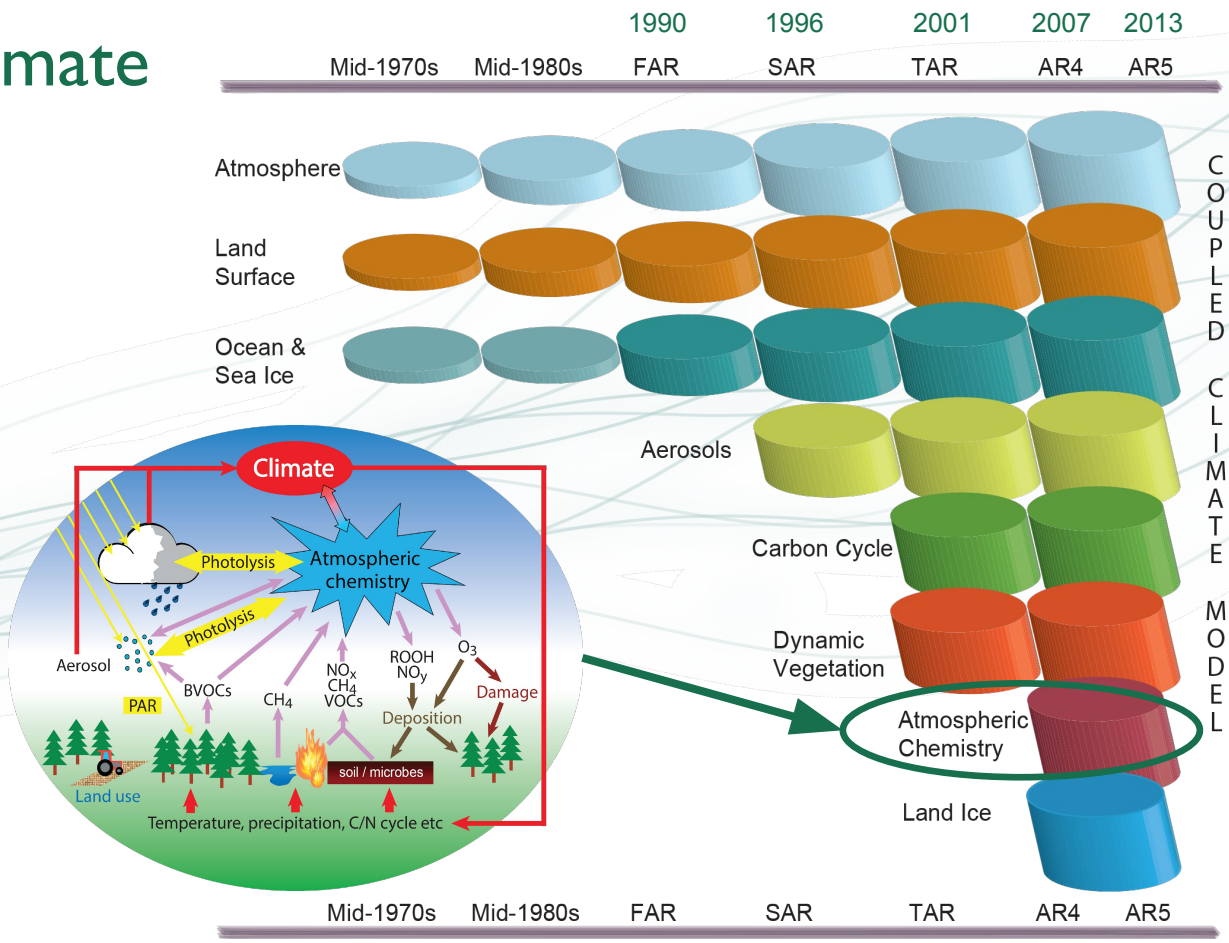
Exploring the climate of Proxima B with the Met Office Unified Model
Boutle *et al.* 2017

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Development of Climate Models

Increasing the **number** and **complexity** of processes represented in climate models.



IPCC, 2013

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

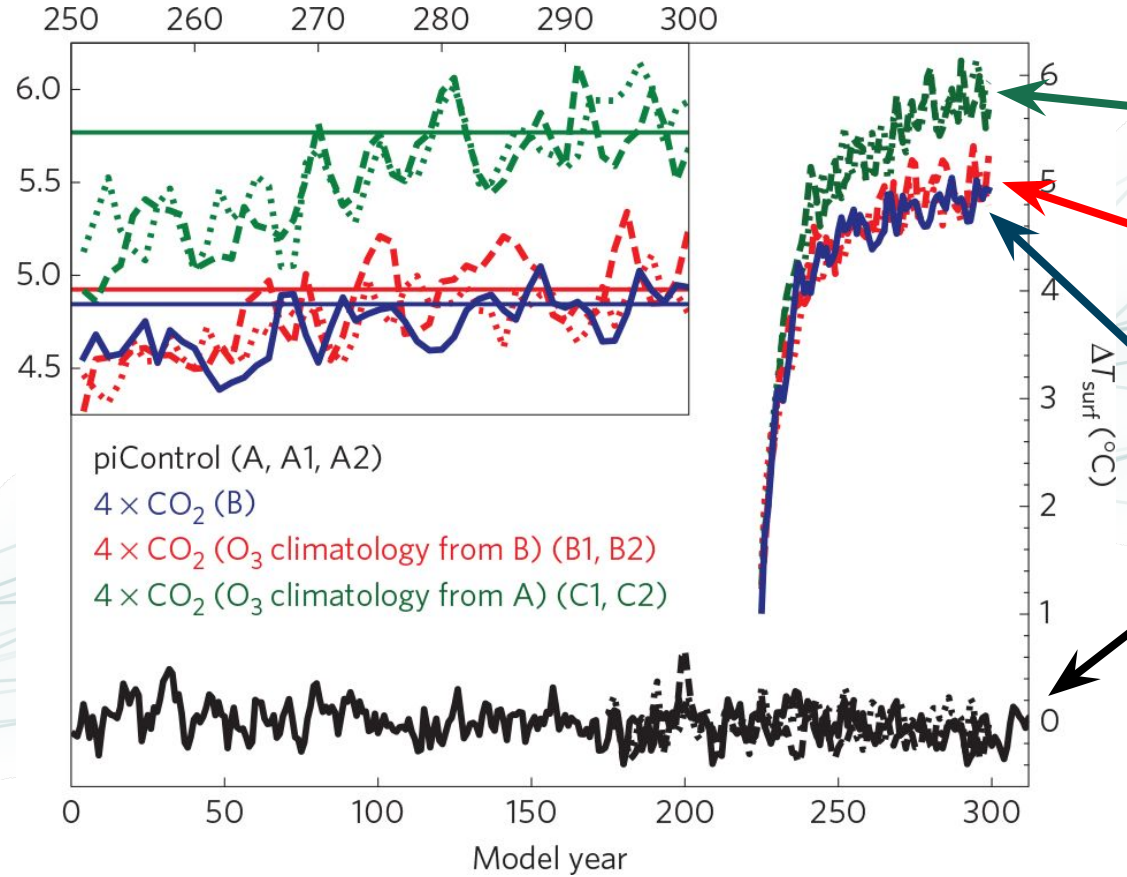
United Kingdom Chemistry and Aerosols

- In Cambridge we work on a part of the UM called the *United Kingdom Chemistry and Aerosols* sub-model, or **UKCA**.
- We develop new chemistry schemes for the model, with a focus on Chemistry-Climate Interactions.
- Here we usually include between 75-240 transported chemical species and hundreds of reactions to allow the model to accurately simulate changes to radiatively important gases such as ozone and methane, which can feed-back on the climate system.

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Chemistry-Climate Interactions



Prescribed pre-industrial ozone from A

Prescribed ozone from B

Interactive chemistry (B)

Pre-industrial control simulation (A)

Nowack *et al.*, *Nature Climate Change*, 2014

Met Office

4xCO2

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Developing code for the Unified Model

Met Office

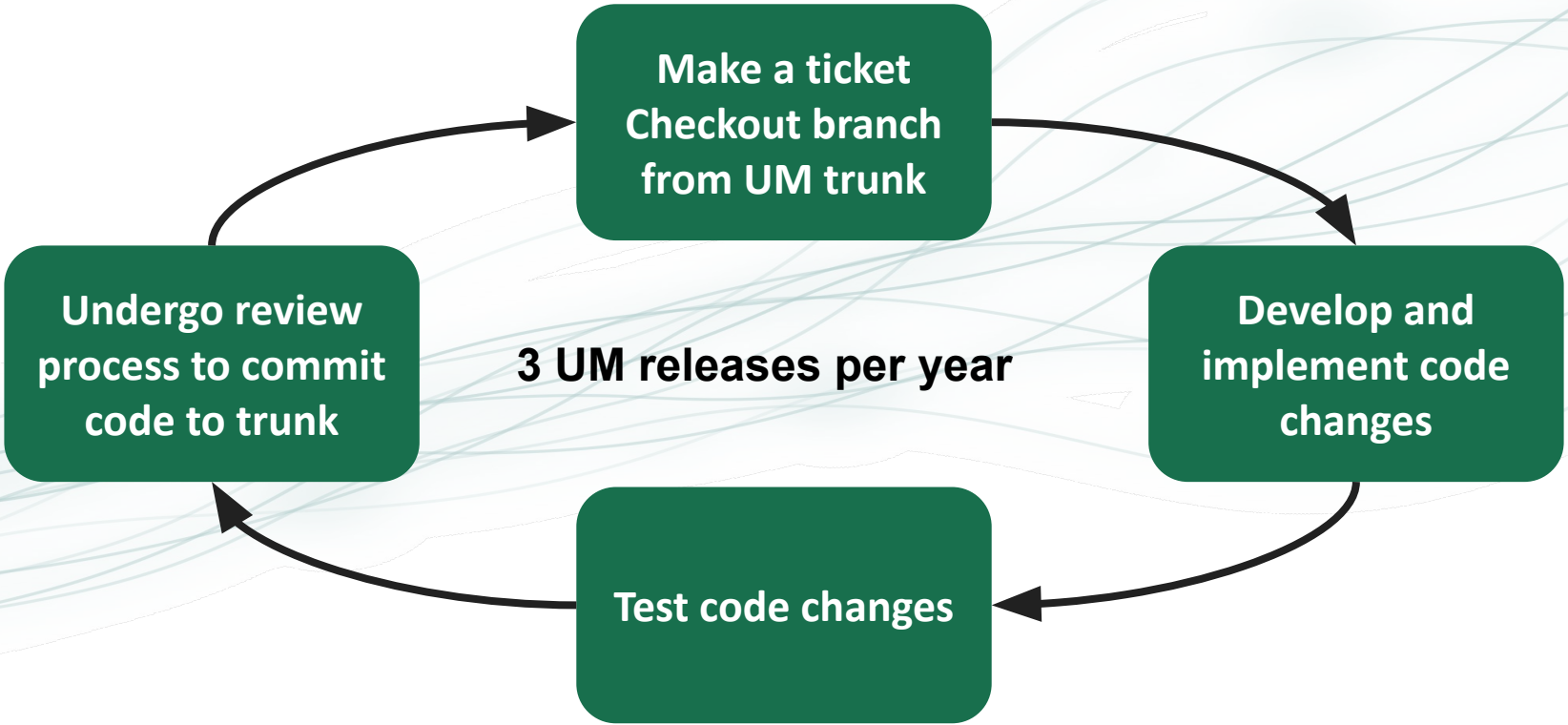
National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

The Met Office Unified Model

um | Unified Model

- Approximately 900,000 lines of code (mainly Fortran 2003).
- Over 200 active developers.
- Uses the *Rose* graphical namelist editor and the *Cylc* workflow engine, with the code held in subversion repositories, managed using *FCM* ("Flexible Configuration Management" - mostly a wrapper around subversion).
- Majority of the development work is done by the Met Office.

Met Office Unified Model code development process

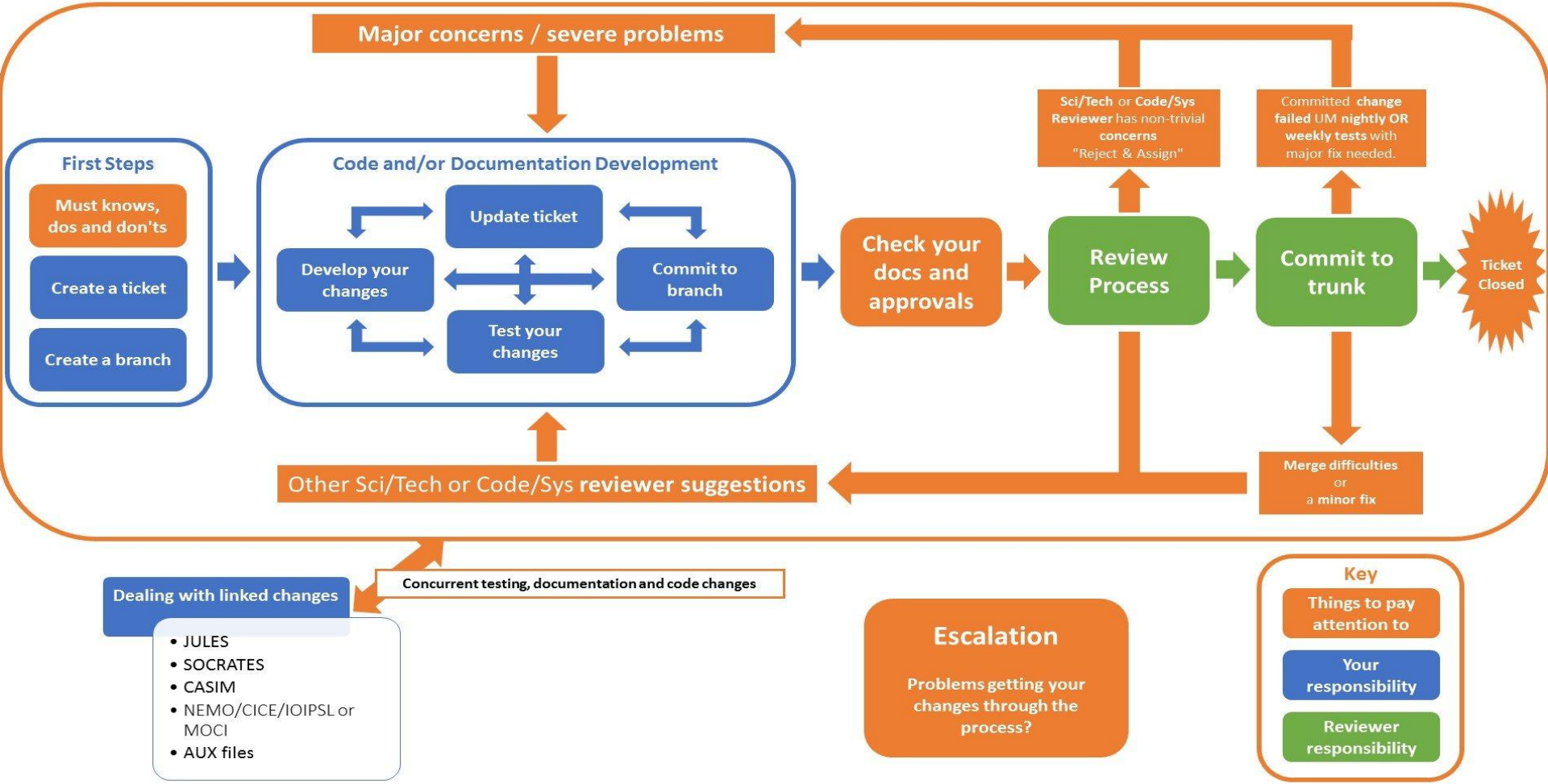


Met Office

Similar process for science improvements to be included in operational configurations.

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

UM Development Working Practices



Code development process

- Science changes often require testing with long simulations that will take several weeks or more to run.
  - Diagnostics are then run through assessment and validation tools to produce many plots of standard metrics.

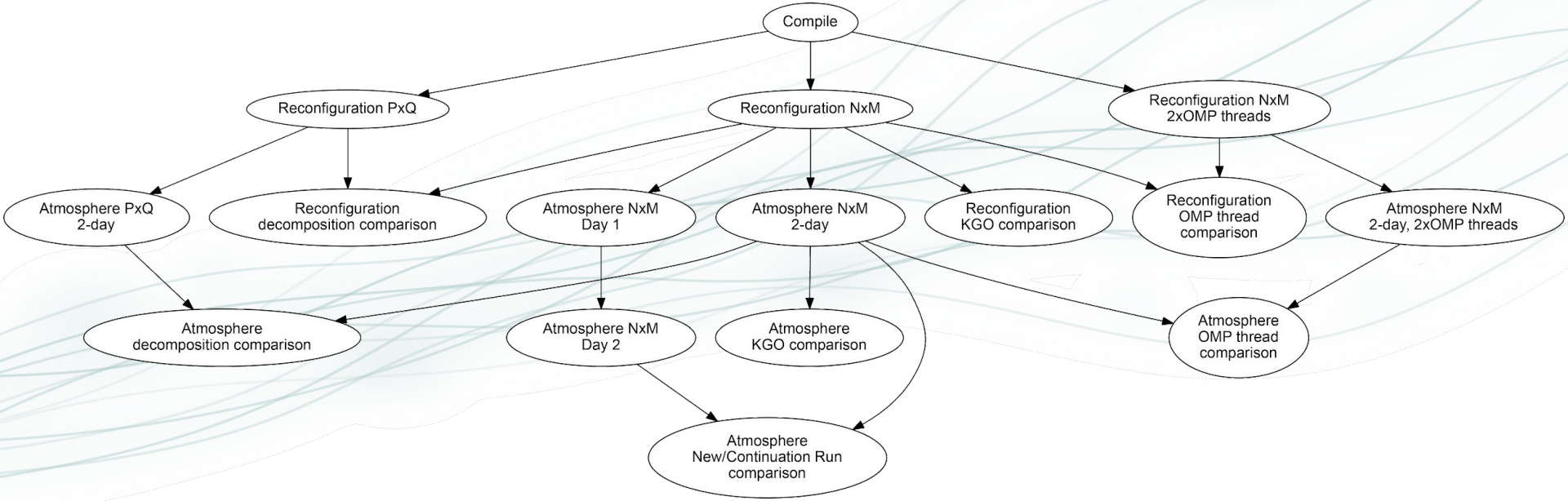
- All code changes are tested using the *rose-stem* utility, using a set of standard tests that compare against "***Known Good Output***" *(KGO).*
  - e.g. a climate resolution UKCA testing job takes 7 minutes to run 2 model days on 144 cores, requiring 112GB of memory.

- **You must be able to show that your change works when turned on and doesn't break anything when turned off.**

rose-stem

- At UM vn12.2 there are 328 UM testing jobs (inc. 33 UKCA jobs):
  - HPC jobs with Cray, GCC/Intel & GNU compilers
  - Linux jobs GCC/Intel, GCC/PGI, Clang/Intel, & GNU compilers
- Additional restart file creation tests and further tests for code standards, metadata, utilities, creation of boundary conditions, etc.
- Tests include KGO, restartability, OpenMP, & processor decomposition tests, with a range of optimisation levels:
  - “high”, “fast”, “safe”, “debug”, “rigorous”.
- If a KGO test “breaks” you need to explain why and get approval to change the saved output.
- **Tests protect your code and scientific configurations from being broken by another change.**

rose-stem - Met Office testing framework

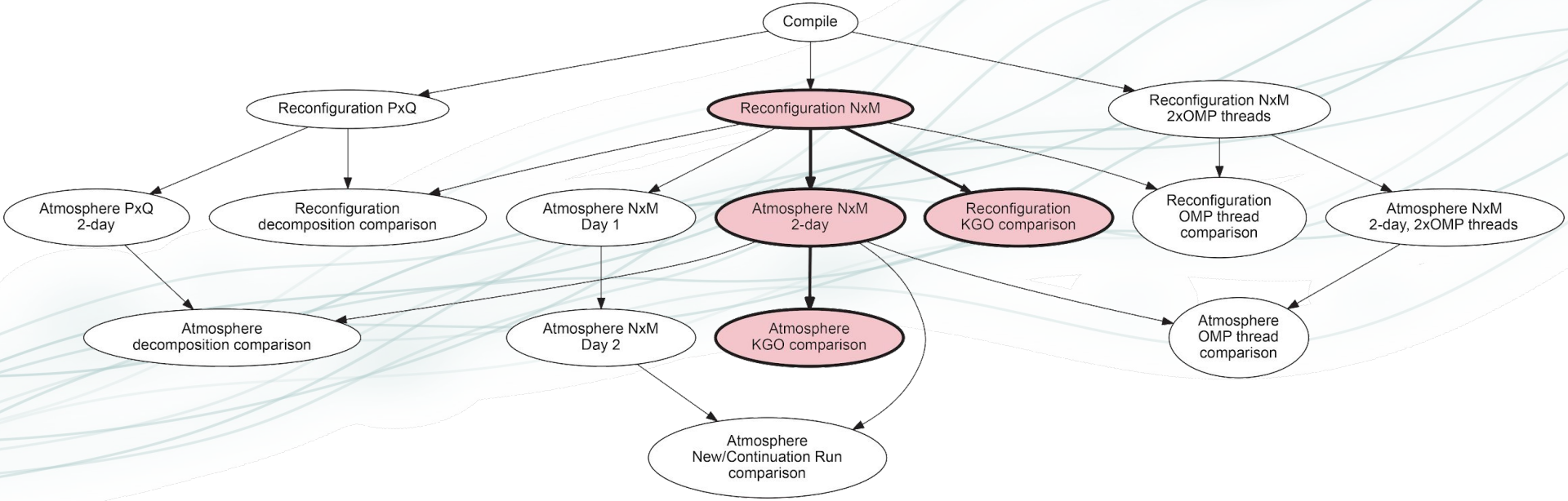


Done with different levels of compiler optimisation.

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

rose-stem - “Known Good Output” or KGO tests

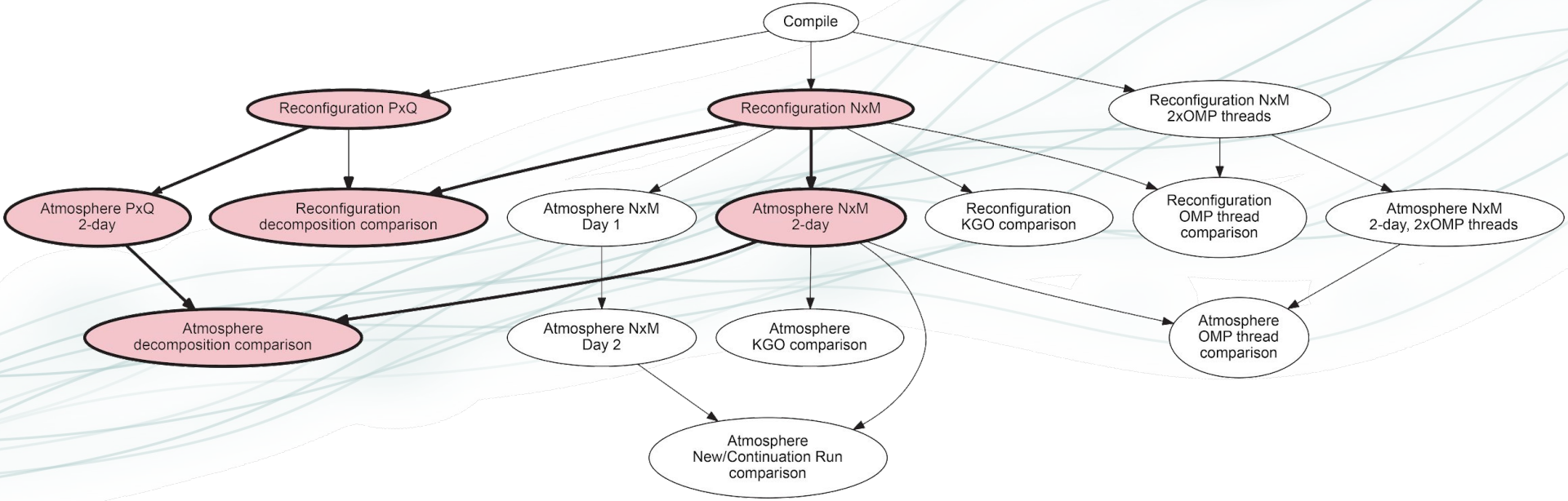


KGO will likely be different with different levels of compiler optimisation.

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

rose-stem - processor decomposition tests

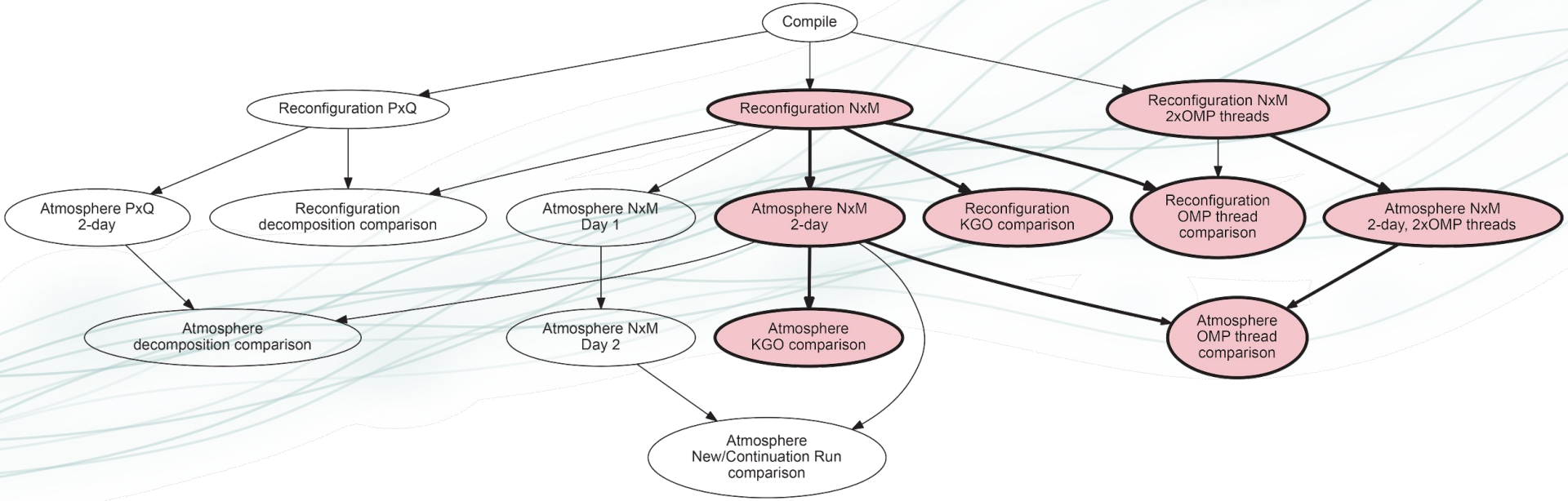


Some changes may break KGO but should still pass the processor decomposition tests.

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

rose-stem - OpenMP tests

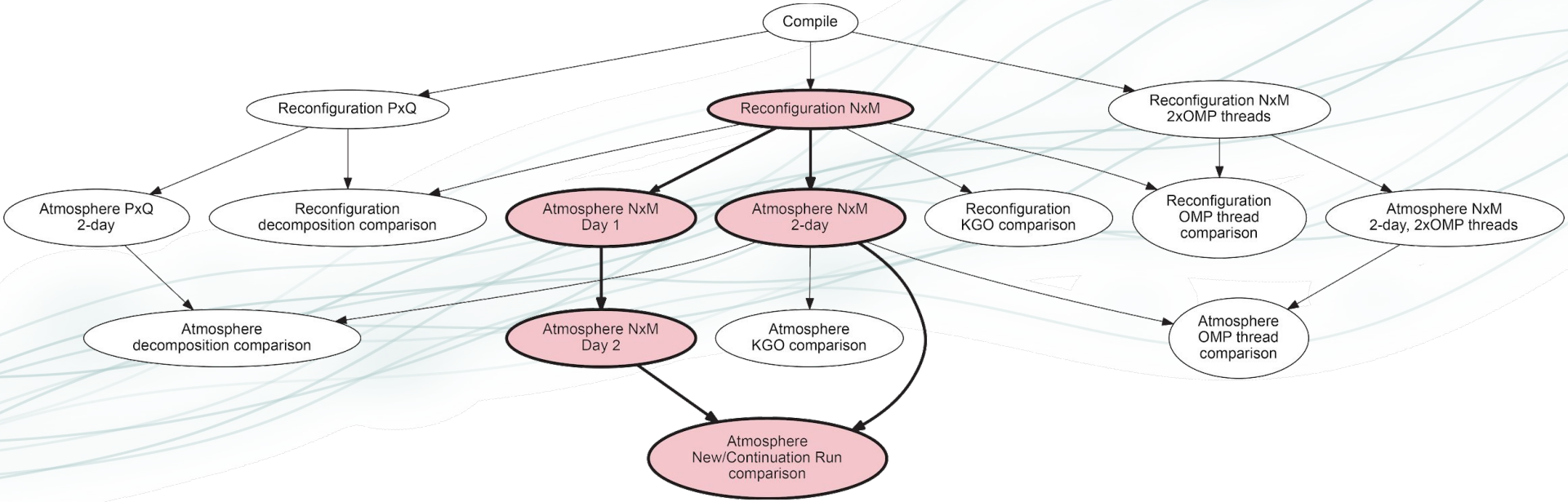


KGO should be identical with or without OpenMP, although the OMP test should still pass even if the KGO one does not.

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

rose-stem - restart tests

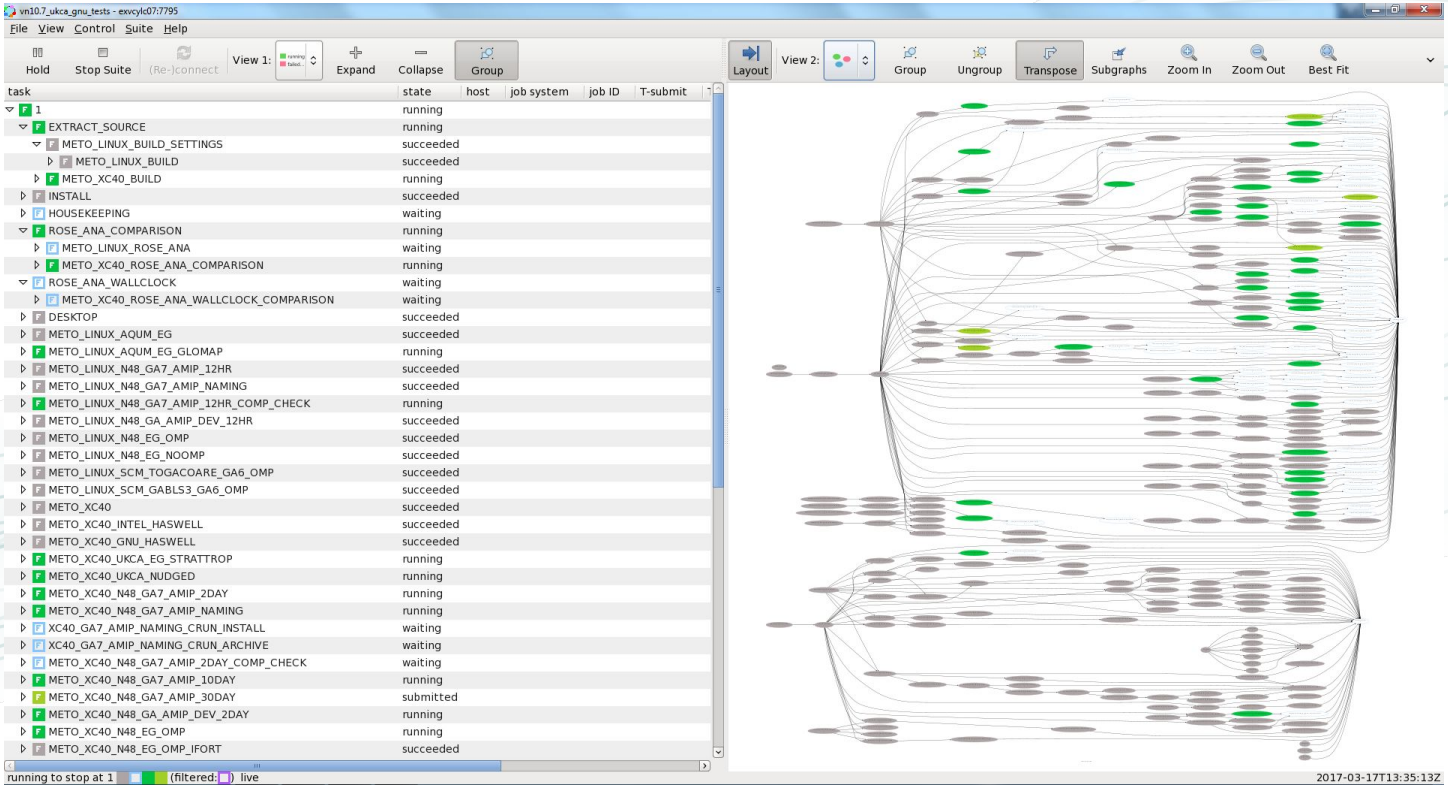


Some changes may break KGO but should still pass the restartability tests.

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

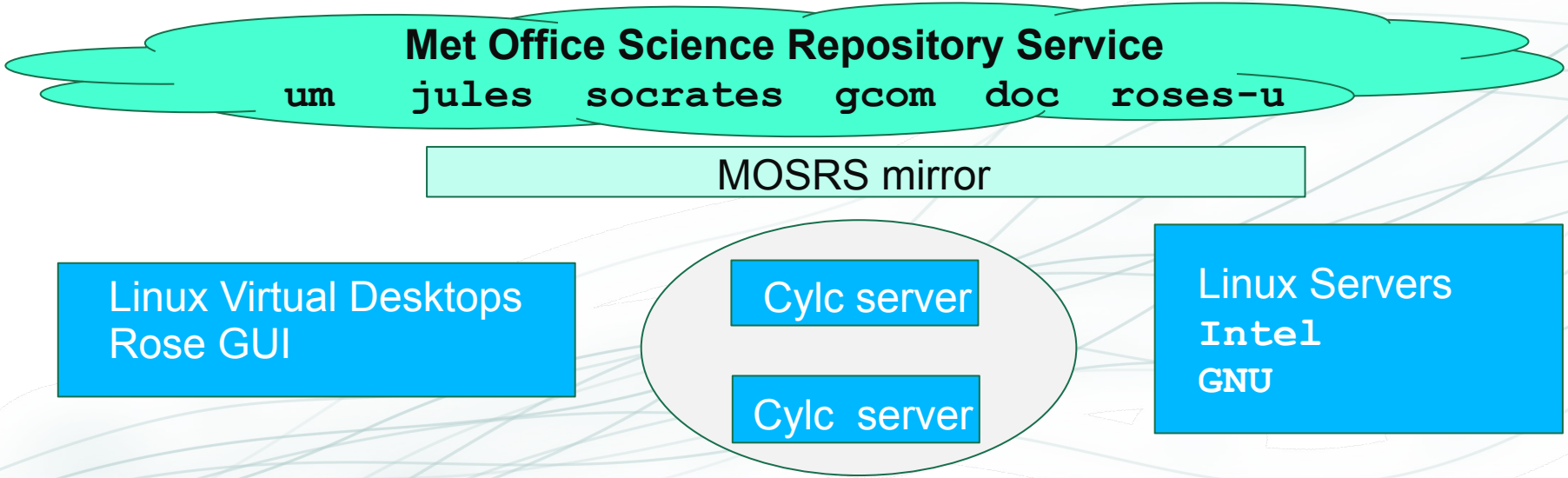
rose stem --group=developer,ukca



Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Met Office Infrastructure (2021)



Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Outside the Met Office

But what if you don't have access to the Met Office infrastructure?

What can UK university students and researchers access?

- You should be able to get access to the UM source code and the tools to run the UM.
- You *may* be able to get access to HPC resources, e.g. ARCHER2 or Monsoon2 etc.
- You *may* be able to get access to an analysis platform, e.g. JASMIN.
- Access to rose-stem and the internal Met Office tools could be limited or not available at all.

What do I mean by "virtually"?

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

What do I mean by *"virtually"*?

- A Virtual Machine configuration has been developed to allow people to easily use FCM, Rose, & Cylc.
- Uses Vagrant and VirtualBox (for a PC or server) or AWS.
- UM Systems Team have set-up running the UM in an Ubuntu guest image.

JASMIN

Met Office Virtual Machine on GitHub:

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

aws

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Uses for the VM

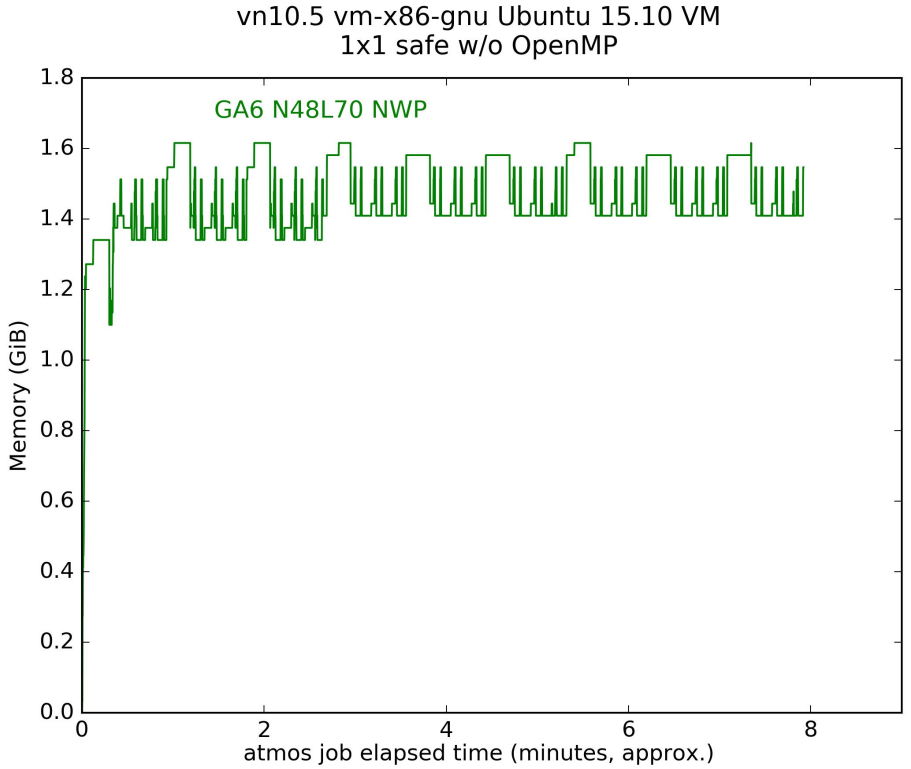
1. Developing & testing code changes
    - Faster turnaround time
        - Approximately 10 minutes to compile from scratch, with only a few seconds to recompile.
        - Small versions of standard jobs run in a few minutes on 2 cores.
2. Training

What do I mean by *"virtually"*?

- To run a basic UM test suite, the VM needs around 3GB of memory.
- Initial thoughts were that the main issue with running UKCA would (*probably!*) be the memory requirement.
  - A standard UKCA job requires an additional 137 3D fields to be added to the restart file, along with further temporary arrays allocated at run-time and any extra required for diagnostic output.
  - Usually require at least 112GB of memory at climate resolution (192x144x85 grid points).

Memory usage for um-atmos job

Basic test suite uses ~1.6GiB



GA6 job runs for 9 model hours
- 27 timesteps

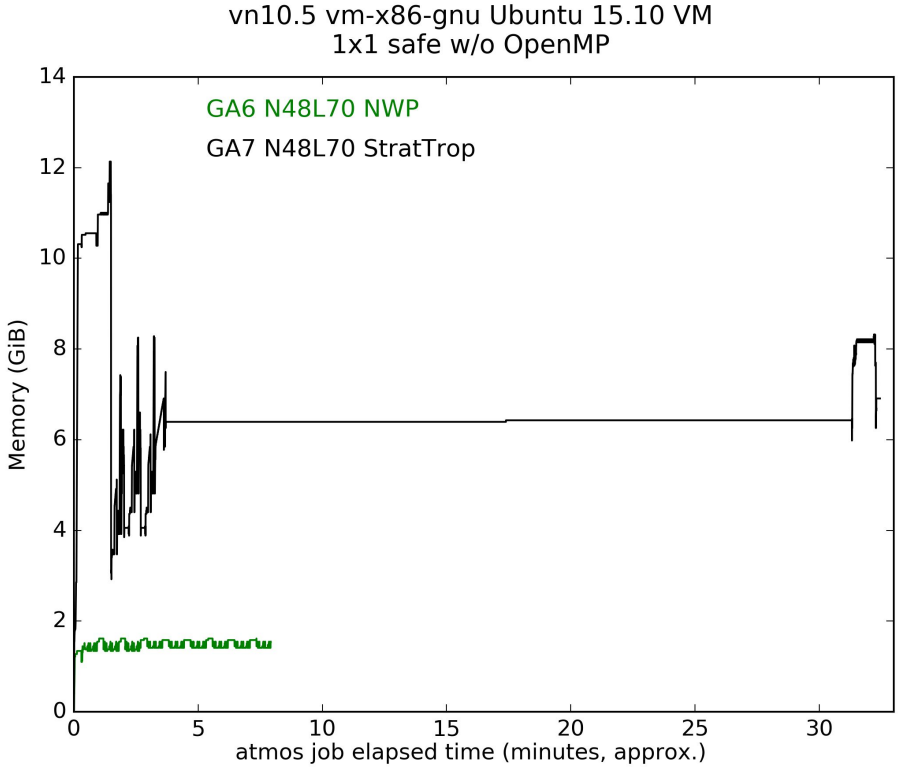
Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Memory usage for um-atmos job

Basic test suite uses ~1.6GiB

Low resolution UKCA suite uses ~12.2GiB



GA6 job runs for 9 model hours
- 27 timesteps
GA7 jobs run for 1 model hour
- 3 timesteps for L70

Met Office

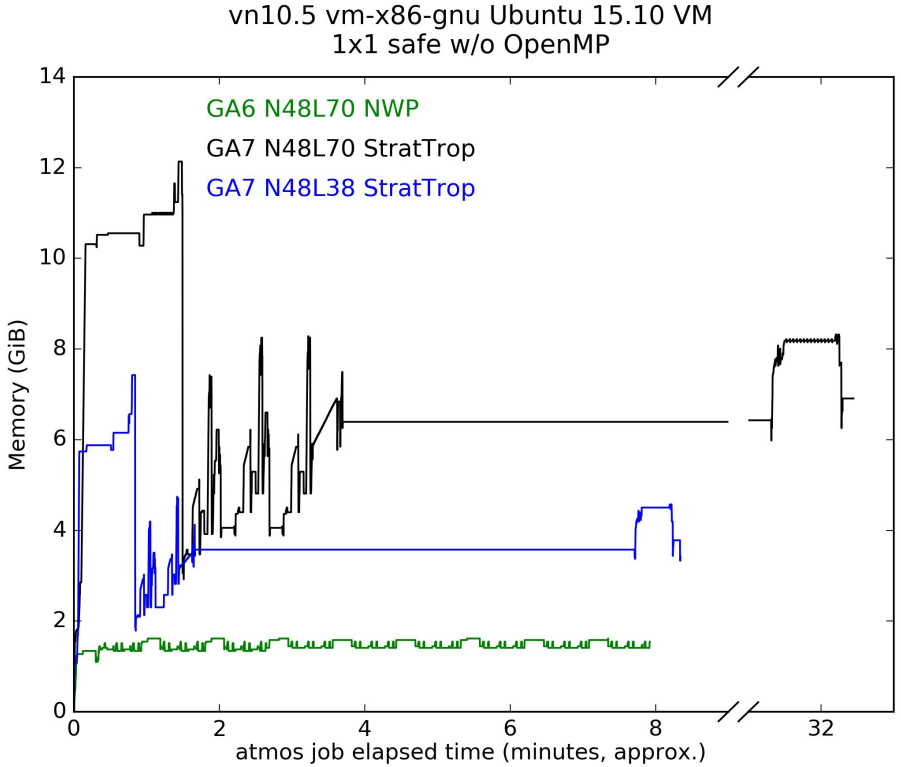
National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Memory usage for um-atmos job

Basic test suite uses ~1.6GiB

Low resolution UKCA suite uses ~12.2GiB

Low resolution UKCA suite (low top) uses ~7.4GiB



GA6 job runs for 9 model hours
- 27 timesteps
GA7 jobs run for 1 model hour
- 3 timesteps for L70
- 2 timesteps for L38

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

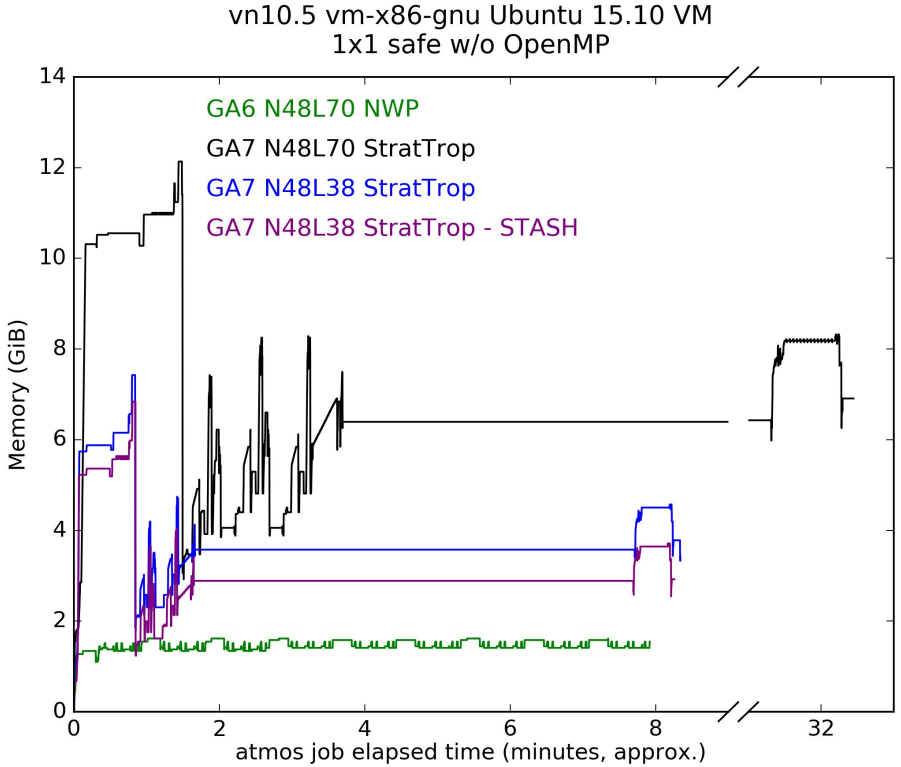
Memory usage for um-atmos job

Basic test suite uses ~1.6GiB

Low resolution UKCA suite uses ~12.2GiB

Low resolution UKCA suite (low top) uses ~7.4GiB

Low resolution UKCA suite (low top) uses ~6.9GiB when all diagnostic output requests are removed



GA6 job runs for 9 model hours
- 27 timesteps
GA7 jobs run for 1 model hour
- 3 timesteps for L70
- 2 timesteps for L38

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Memory usage for um-atmos job

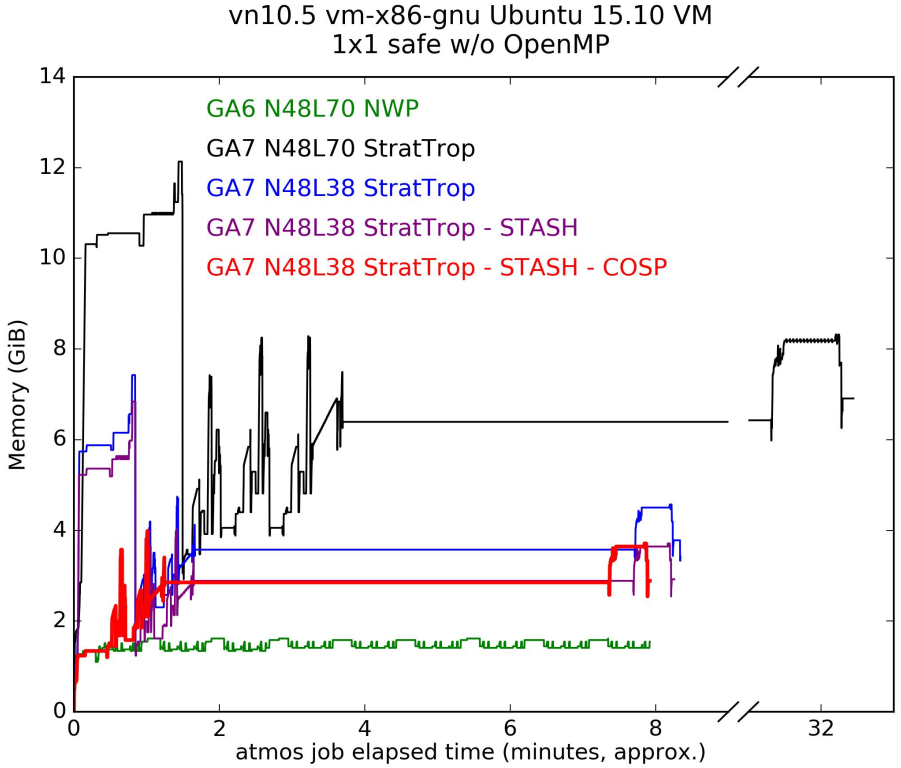
Basic test suite uses ~1.6GiB

Low resolution UKCA suite uses ~12.2GiB

Low resolution UKCA suite (low top) uses ~7.4GiB

Low resolution UKCA suite (low top) uses ~6.9GiB when all diagnostic output requests are removed

Low resolution UKCA suite (low top) uses ~4.0GiB when all diagnostic output requests are removed and the COSP (*CFMIP Observation Simulator*) diagnostics package is not turned on



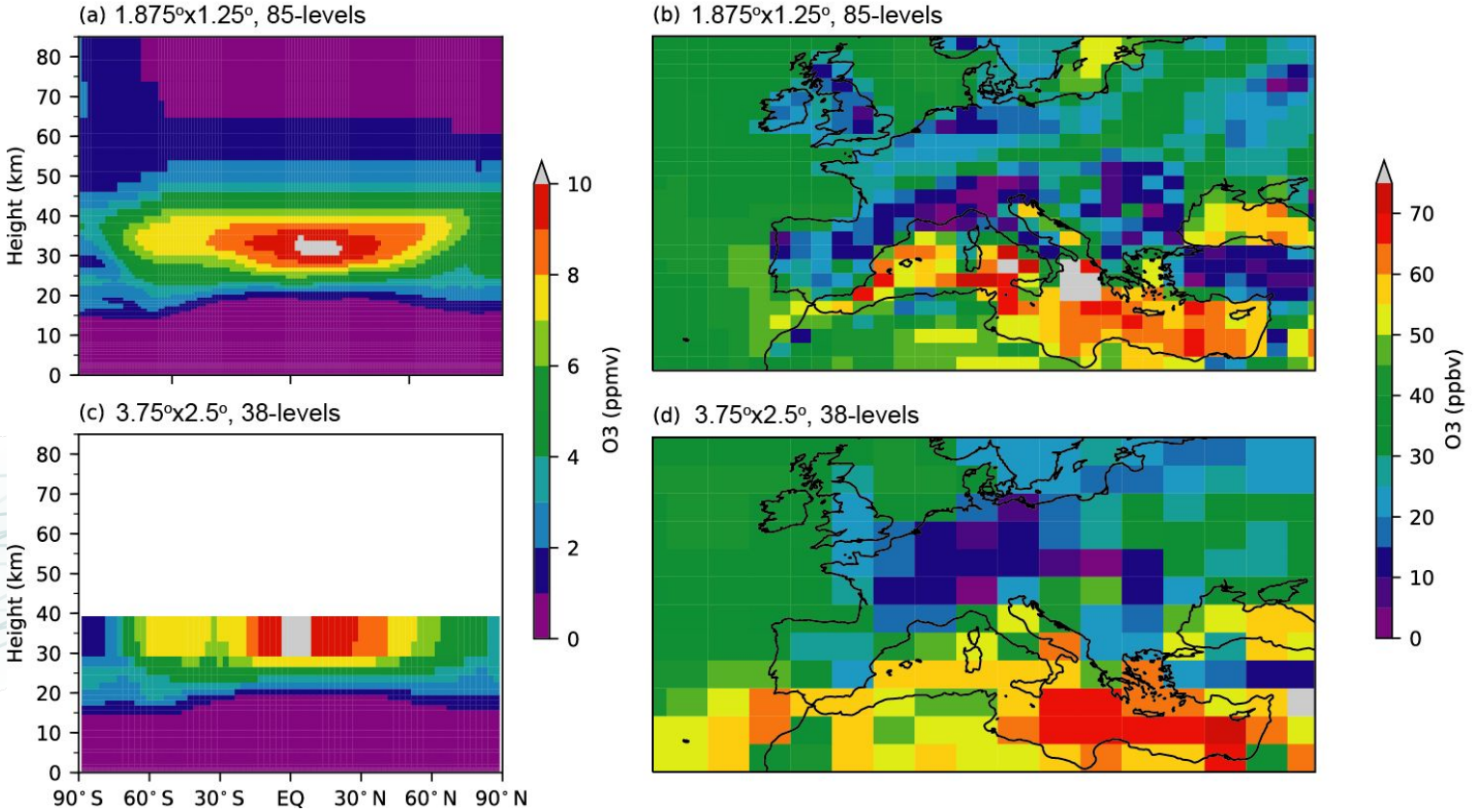
GA6 job runs for 9 model hours
- 27 timesteps
GA7 jobs run for 1 model hour
- 3 timesteps for L70
- 2 timesteps for L38

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Climate resolution
192x144x85
grid points

VM-suitable resolution
96x72x38
grid points



Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

What do I mean by "virtually everything"?

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Helpful(?!) Errors

- However, there seemed to be an issue with using the GNU compiler, as the code failed with the following error on timestep 3:

```
?  Error from routine: EG_BICGSTAB_MIXED_PREC
?  Error message: NaNs in error term in BiCGstab
?        This is a common point for the model to fail if it
?        has ingested or developed NaNs or infinities
?        elsewhere in the code.
?        See the following URL for more information:
? https://code.metoffice.gov.uk/trac/um/wiki/KnownUMFailurePoints
```

- After some debugging, the issue *appeared* to be with the UKCA routine **`asad_hetero.F90`**, which couples the chemistry and aerosol schemes.

Precision

The UM uses compiler flags to make **`REAL`**s double-precision, so rather than using

```
REAL(KIND=8) :: a
```

or

```
INTEGER, PARAMETER :: dp=SELECTED_REAL_KIND(15,300)

REAL(KIND=dp) :: a
```

the UM will use

`-r8` (Intel), **`-s default64`** (Cray), **`-fdefault-real-8`** (GNU)

AND there is still the occasional **`1.0d0`** statement.

But what about `TINY(1.0d0)`?

- Intel (`-r8`):

```
tiny(1.0)    =   2.225073858507201E-308
tiny(1.0d0)  =   2.225073858507201E-308
```

- Cray (`-s default64`)

```
tiny(1.0)    =  2.22507385850720138E-308
tiny(1.0d0)  =  2.22507385850720138E-308
```

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

But what about `TINY(1.0d0)`?

- Intel (`-r8`):

```
tiny(1.0)    =    2.225073858507201E-308
tiny(1.0d0)  =    2.225073858507201E-308
```

- Cray (`-s default64`)

```
tiny(1.0)    =   2.22507385850720138E-308
tiny(1.0d0)  =   2.22507385850720138E-308
```

- GNU (`-fdefault-real-8`)

```
tiny(1.0)    =    2.2250738585072014E-308
tiny(1.0d0)  =    3.36210314311209350626267781732175260E-4932
```

Effectively `a = 1.0d0 = 1.0`

but `a = tiny(1.0d0) = 0.0`

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

But what about `TINY(1.0d0)`?

- Intel (`-r8`):

```
tiny(1.0)    =    2.225073858507201E-308
tiny(1.0d0)  =    2.225073858507201E-308
```

- Cray (`-s default64`)

```
tiny(1.0)    =   2.22507385850720138E-308
tiny(1.0d0)  =   2.22507385850720138E-308
```

- GNU (`-fdefault-real-8`)

```
tiny(1.0)    =    2.2250738585072014E-308
tiny(1.0d0)  =    3.36210314311209350626267781732175260E-4932
```

- GNU (`-fdefault-real-8 -fdefault-double-8`)

```
tiny(1.0)    =    2.2250738585072014E-308
tiny(1.0d0)  =    2.2250738585072014E-308
```

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

What do I mean by "everything"?

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Rose stem (again!)

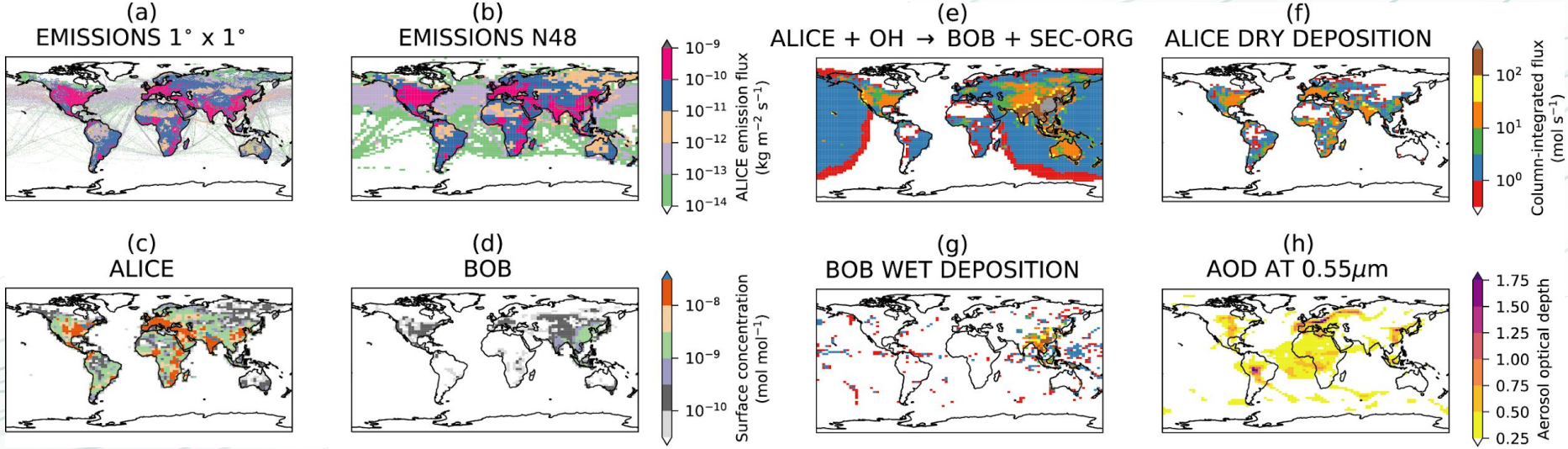
- These compiler flag changes went into the UM trunk at vn10.7.
- There are *rose-stem* testing jobs for several UM configurations (now also including UKCA) on the VM.
- Equivalent low-resolution UKCA jobs have also been added to the Met Office HPC tests.
    - One is included as part of the standard *developer* group that must be used when making any change.

Training

- What if we could use the VM for training?
- This would mean:

1. Everyone runs on their own dedicated computer.
    - Simpler set-up than for current production runs.
2. Researchers doing the tutorials can easily do the training tasks without needing a supercomputer account.
    - “Try before you buy”

UKCA Training



Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Conclusions

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Conclusions

- Testing environments, as well as production environments, are required to be able to develop code changes in a timely manner.
- A Virtual Machine environment has been developed by the Met Office that easily allows users to install their own copy of the UM.
- Standard tests have also been implemented on this system that can be used when developing new code changes.
- This creates a standard system that all non-Met Office developers can use to quickly implement and test changes, prior to running on a HPC.

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

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Step	ARCHER (XC30)	XCS-C (XC40)	Virtual Machine
Cray `cce` initial compile	34 minutes	15 minutes	-
Cray `cce` incremental compile	5-7 minutes	3 minutes	-
Intel `ifort` initial compile	19 minutes	9 minutes	-
Intel `ifort` incremental compile	6-7 minutes	1 minute	-
GNU `gfortran` initial compile	-	4 minutes	8-10 minutes
GNU `gfortran` incremental compile	-	2-3 minutes	45 seconds
Reconfiguration task, used to produce the initial conditions file.	3-4 minutes (Intel, 6×4)	15-30 seconds (GNU, 4×9)	25-30 seconds (GNU, 1×2)
Atmosphere task	40 seconds (Intel, 6×4)	45 seconds (GNU, 4×9)	12 minutes (GNU, 1×2)

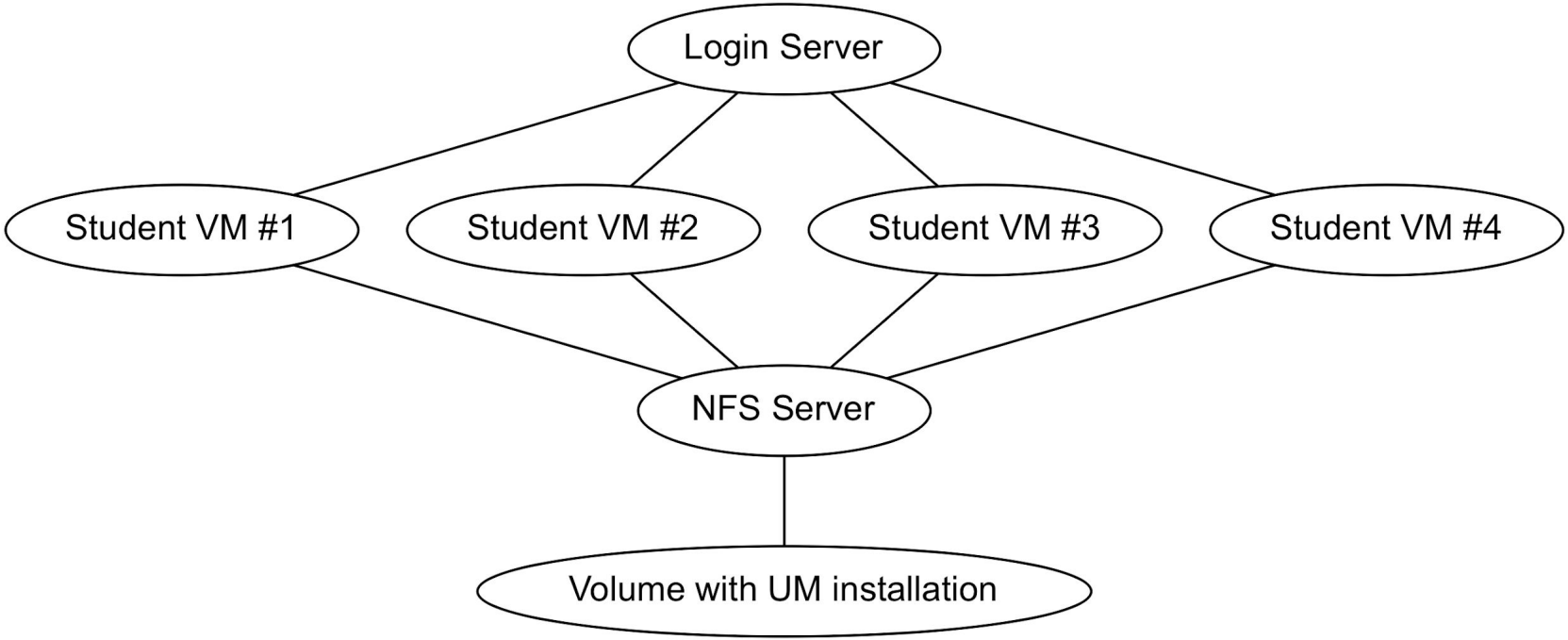
Compiler settings	**Safe**	**Safe**	**Rigorous**
Number of UM OpenMP threads	**0**	**2**	**2**
Approximate run-time on 2-core VM (1×2) (minutes)	8	11	29
Approximate run-time on 4-core VM (1×2) (minutes)	8	4	17
Approximate run-time on 4-core VM (1×4) (minutes)	5	8	22
Approximate run-time on 8-core VM (1×4) (minutes)	5	3	14
Approximate run-time on 8-core VM (1×8) (minutes)	3	6	26
Approximate run-time on 16-core VM (1×8) (minutes)	3	3	22

Compile type	GNU `gfortran` compiler flags	Number of UM OpenMP threads	Total VM memory required (GB)
safe	`-O2 -Werror`	0	6
safe	`-O2 -Werror` `-fopenmp`	2	6

rigorous	`-O0 -Wall`
	`-ffpe-trap=invalid,zero`
	`-fbounds-check`
	`-Warray-bounds`
	`-fcheck-array-temporaries`
	`-finit-real=nan`
	`-fimplicit-none`
	`-fopenmp`
		2	8

Training VMs on JASMIN

JASMIN



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

Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

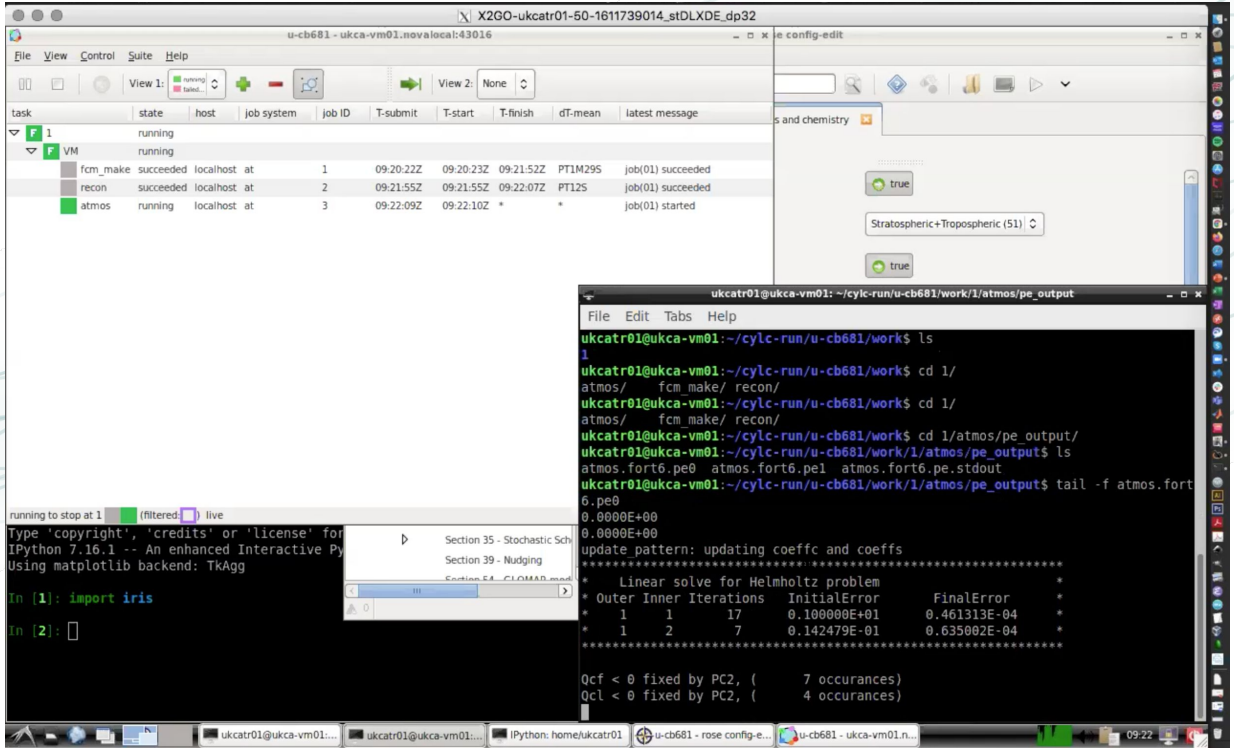
Training VMs on JASMIN

JASMIN

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

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

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



Met Office

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL