AccessDev Transition to Gadi @ NCI

https://opus.nci.org.au/display/DAE/UK+Met+Office+Environment+on+NCI

Run Climate and Weather Applications Nonstop For 3 Months

10 November, 2023

Some accessdev history

- 2010 BOM and CSIRO move to NCI
  - accesscollab machine at NCI to host UMUI
  - access-svn for UM code
- 2013 accessdev cloud machine set up (CentOS 6)
- 2014 UM vn10.0, MOSRS, fcm, rose, cylc
  - MOSRS mirror available from accessdev
- 2021-22 NCI MOSRS mirror in ki32, NCI maintained rose/cylc environment in hr22

Agenda

9:40-9:55	Overview and Highlights	Supported Modes, Timeline, Suite Changes, Ongoing Assistance
10:00-11:00	Work with persistent sessions	Launch a Cylc Job in Localhost Mode from Gadi Login Node
Morning Tea Break		
11:15-12:15	Best Practise Special Notes	Supported Modes, Configuration Changes, Suite Run Directories
Lunch Break		
13:00-14:00	One-On-One: Problem Solving: Rui, Yue, Martin, Dale	
14:00-14:30	Known Issue Addressed	SSH configuration, 2-step fcm_make, general porting notes
14:30-15:00	Sharing Suite Progress	Look up Job Progress using Cylc Review in ARE VDI
15:00-15:30	Discussion	How to better support the community • Containerised workflow • Migration to Cylc8 • Other workflows on accessdev • …

UK Met Office Environment on NCI

Created by Yue Sun, last modified by Ben Evans about 3 hours ago

Many weather and climate models are based on or coupled with the UK Met Office model code, including many of the ACCESS climate and earth systems models, Australian weather model configurations, and the UM suites with coupled model configurations. To support these models, we have created the UK Met Office Environment at NCI.

The environment has been integrated with the system including Gadi, ARE and supporting services. It consists of: the Cylc workflow system; rose suites; the Met Office MOSRS repository and our local replica repository; and the model software and tools.

We note that we are progressively transitioning from the aging **accessdev** environment toward this new Met Office Environment described here.

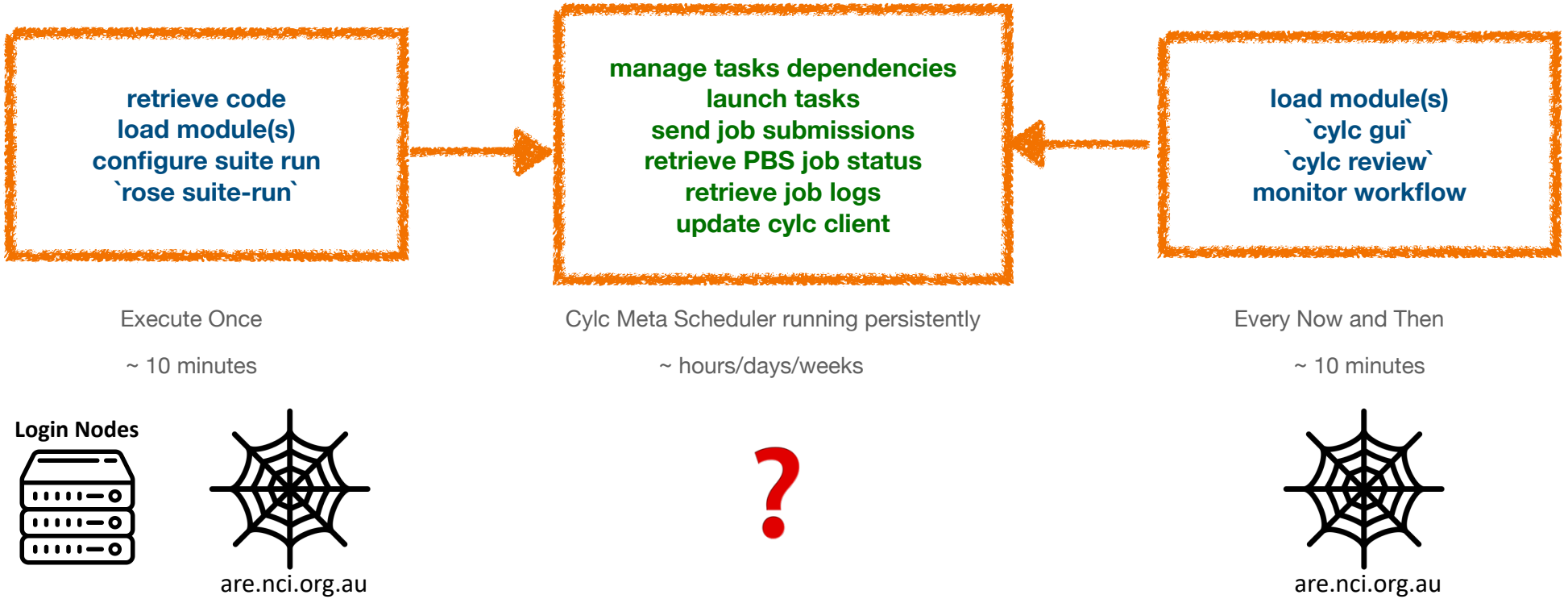
https://opus.nci.org.au/display/DAE/UK+Met+Office+Environment+on+NCI

UKMO Environment Supported

Rose & Cylc workflow software stack (hr22)	Software repository mirrors (ki32)	Spack package manager (ki32)	Computing platforms
• Cylc 7 • Cylc 8 • fcm • fab • mosrs-setup	• MOSRS • Nemo • …	• Rose/Cylc suite dependencies • Containerised model packages • Gitlab CI/CD pipeline*	• ARE VDI • ARE JupyterLab • Persistent Sessions • Interim special project with the extended PBS time limit

* coming soon

New Home to Cylc Meta-Scheduler



New Home to Cylc Meta-Scheduler

Execute Once

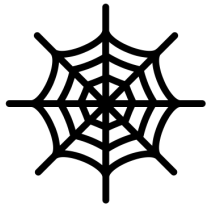
~ 10 minutes

Cylc Meta Scheduler running persistently

~ hours/days/weeks

Every Now and Then

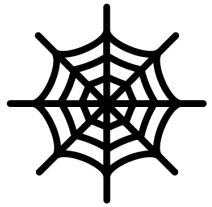
~ 10 minutes



are.nci.org.au

persistent sessions

demo.jpg777.ab01.ps.gadi.nci.org.au

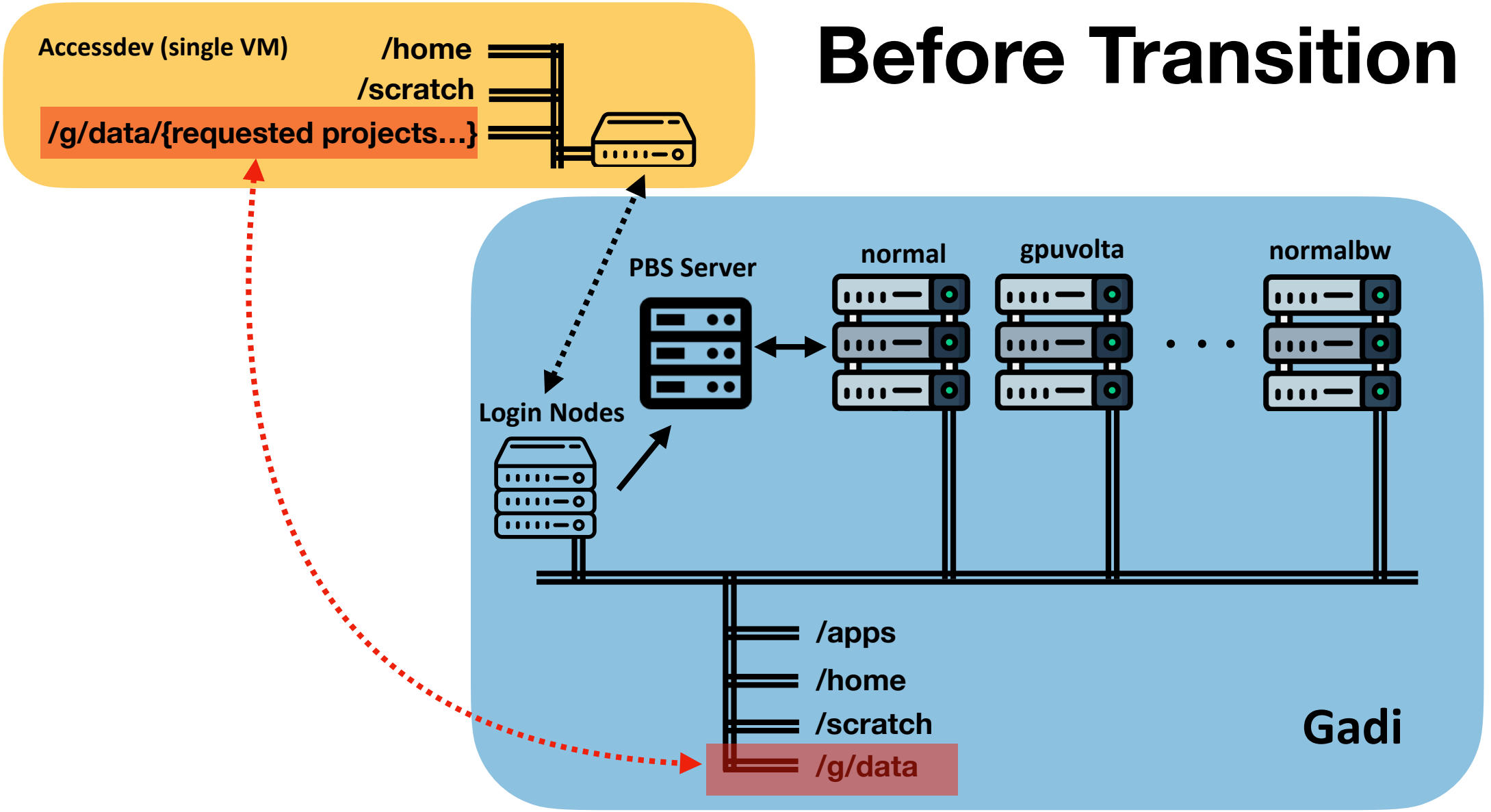


are.nci.org.au

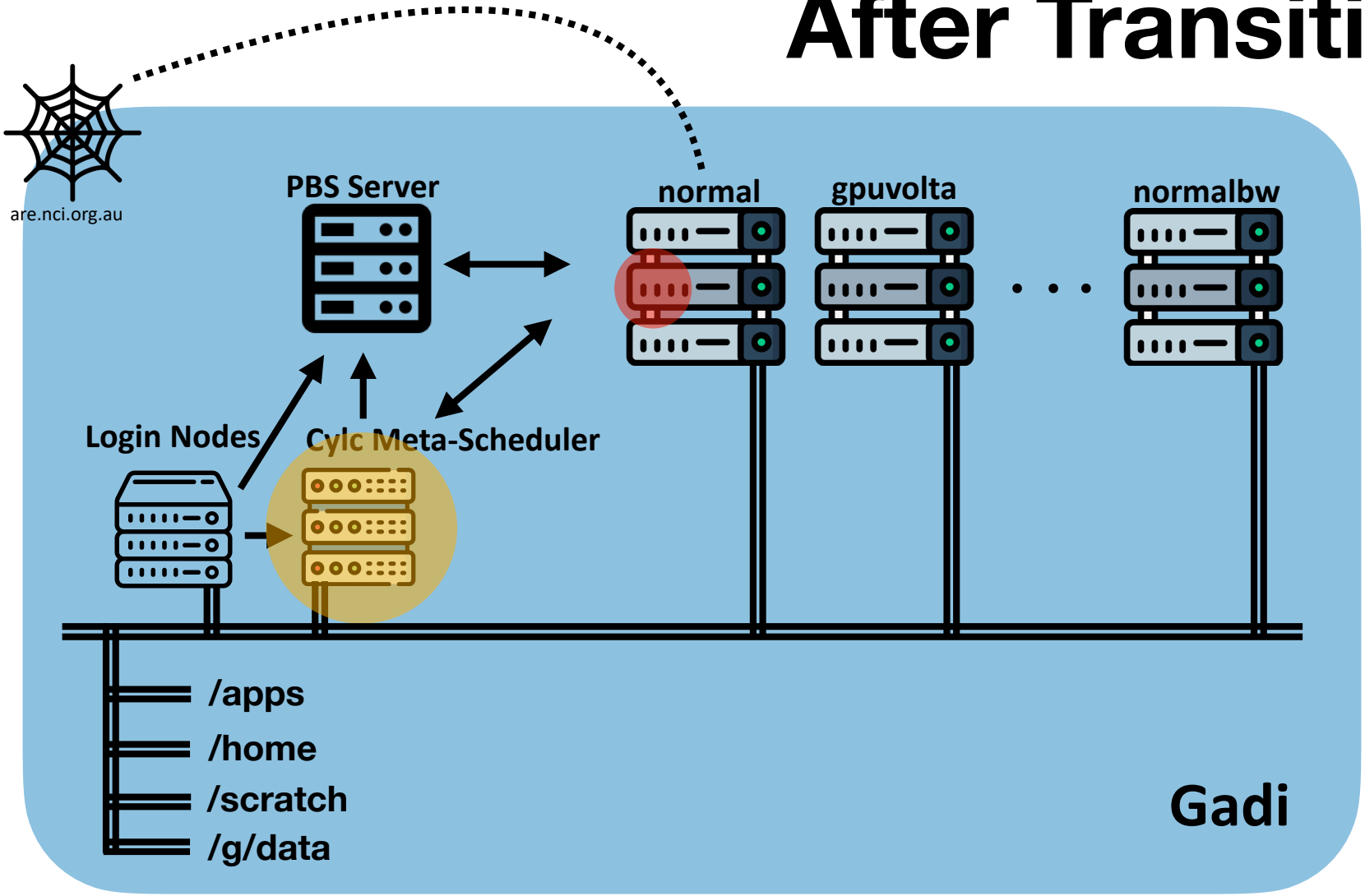
Persistent Sessions @ NCI

- User Specific
- Access to/from Gadi
  - PBS server
  - ARE VDI on compute nodes
  - login nodes
  - /home, /scratch, and /g/data
- Expected Life Span
  - ~ 3 months
  - From maintenance to maintenance

Before Transition



After Transition



Overview & Highlights

- Supported Modes
- Suite Changes
  - SSH
  - Directories
- Transition Timeline
- Ongoing Assistance
  - Data migration
  - Weekly Open Mic Virtual Lunch
  - NCI Help Desk

Revisit Localhost Mode

- Localhost: Jobs submitted directly from persistent sessions
- Access-Dev Compatible: SSH to login nodes to submit jobs

$ROSE_ORIG_HOST	Localhost Mode	Access-Dev Compatible Mode
ARE VDI Gadi Login Node	Cylc client on ARE VDI/Login node & server on the persistent session	Special ~/.ssh/config Required Cylc client on ARE VDI/Login node & server on the persistent session
Persistent Session	Both Cylc client & server on the persistent session	Both Cylc client & server on the persistent session

Suite Changes

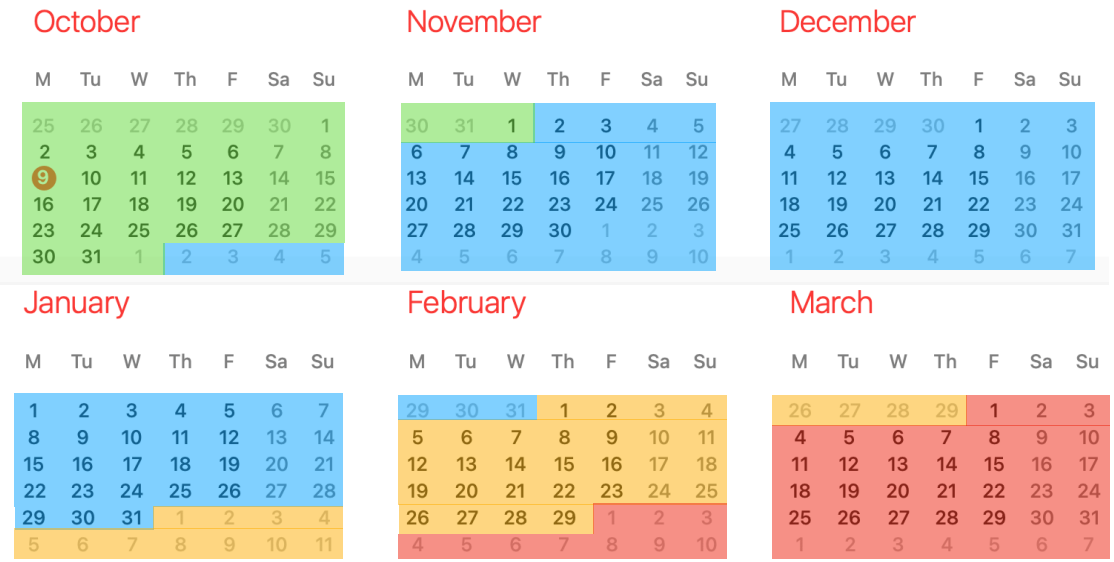
- Suite Run Directories
  - default to /scratch/$PROJECT/cylc-run/<suite-id>
  - define user level configuration in
    - `~/.metomi/rose.conf ` for Cylc7
    - `~/.cylc/flow/global.cylc` for Cylc8
- Remote Host
  - redefine “remote = gadi” to “remote = localhost” in `suite.rc`

```
[runtime]
        [[root]]
             …
        [[GADI]]
                …
             [[[remote]]]
                    host = localhost
```

  - Change PBS storage options:
    - Add “gdata/hr22+gdata/ki32”

Transition Timeline

- Stages
  - Test runs in PS
  - Production runs in both PS & AD
  - Restricted access to accessdev
  - Decommission of accessdev



- Key Dates
  - Production runs in persistent sessions supported since 2 Nov
  - Limited access since 1 Feb 2024: only for data access
  - Decommission @ 1 Mar 2024

Ongoing assistance

- Data migration to Gadi
- Open mic starts from 30 Oct @ every Monday 1pm
- NCI Help Desk
  - ticket tag with “accessdev transition”
  - email subject contains “accessdev transition” and cc
    - yue.sun@nci.org.au
    - cws_help@nci.org.au

Any Questions?

https://opus.nci.org.au/display/DAE/UK+Met+Office+Environment+on+NCI

Run Cylc7/8 suites using Persistent Sessions

- Create a Persistent Session
- Load Modules
- Checkout Suites
- Run Suite

Create a Persistent Session

```
$ persistent-sessions list
                                UUID  PROJECT       ADDRESS      CPUTIME  MEMORY
cac76067-5ad0-663c-9791-cfd1c53828f6     ab11      10.9.4.4 00:30:00.091  144.9M
18acc940-e61c-7907-c853-9c9b9013b1ab     ab11      10.9.4.3 00:02:26.939  118.8M
$
$ persistent-sessions start workshopdemo
session d26db958-3238-00db-027b-03ef79cd4acb running - connect using
  ssh workshopdemo.abc111.xy99.ps.gadi.nci.org.au
$
$ persistent-sessions list
                                UUID  PROJECT       ADDRESS      CPUTIME  MEMORY
d26db958-3238-00db-027b-03ef79cd4acb     xy99     10.9.4.12 00:00:00.033    3.0M
cac76067-5ad0-663c-9791-cfd1c53828f6     ab11      10.9.4.4 00:30:00.091  144.9M
18acc940-e61c-7907-c853-9c9b9013b1ab     ab11      10.9.4.3 00:02:26.939  118.8M

$ persistent-sessions -h
$ persistent-sessions list -h
```

Load Modules

Replace cylc7/23.09 with cylc/8.2.1 if working with Cylc8 suites

```
$ export CYLC_SESSION=workshopdemo.abc111.xy99.ps.gadi.nci.org.au
$ module use /g/data/hr22/modulefiles
$ module load cylc7/23.09
Using the cylc session workshopdemo.abc111.xy99.ps.gadi.nci.org.au
Loading cylc7/23.09
  Loading requirement: mosrs-setup/1.0.1
$ module list
Currently Loaded Modulefiles:
 1) pbs   2) mosrs-setup/1.0.1(default)   3) cylc7/23.09
```

What does the module do?

- **env | grep CYLC**
- **Cylc7: rose config | grep -v ^#**
- **Cylc8: cylc config -v**

Checkout Suites

```
$ mosrs-auth
INFO: You need to enter your MOSRS credentials here so that GPG can
cache your password.
Please enter the MOSRS password for testuser:
INFO: Checking your credentials using Subversion. Please wait.
INFO: Successfully accessed Subversion with your credentials.
INFO: Checking your credentials using rosie. Please wait.
INFO: Successfully accessed rosie with your credentials.
$ rosie co u-cs809
[INFO] u-cs809: local copy created at /home/111/abc111/roses/u-cs809
```

How to redirect the suite run directory to /g/data?

- **cat ~/.metomi/rose.conf**
- **cat ~/.cylc/flow/global.cylc**

Run Suites

Replace rose suite-run with cylc install & cylc play <u-cz535> if working with Cylc8 suites

```
$ cd ~/roses/u-cs809
$ rose suite-run
$ readlink ~/cylc-run/u-cs809
/scratch/xy11/abc111/cylc-run/u-cs809
$ readlink ~/cylc-run/u-cz535/runN/log
/scratch/xy11/abc111/cylc-run/u-cz535/run4/log
$ readlink ~/cylc-run/u-cz535/runN/share
/scratch/xy11/abc111/cylc-run/u-cz535/run4/share
$ readlink ~/cylc-run/u-cz535/runN/work
/scratch/xy11/abc111/cylc-run/u-cz535/run4/work
```

What happens next?

- **ps auwxff | grep $USERNAME**
- **ssh $CYLC_SESSION “ps auwxff” | grep $USERNAME**

Back @ 11:15

https://opus.nci.org.au/display/DAE/UK+Met+Office+Environment+on+NCI

Best Practise Special Notes

- ROSE_ORIG_HOST and meta-scheduler
- Supported Cylc 7 Suite Modes
- Hands-on excises
- Advanced excises
- FAQ

ROSE_ORIG_HOST and meta-scheduler

- ROSE_ORIG_HOST: The name of the host where the ‘rose suite-run’ or ‘cylc install’ command was invoked.
  - Gadi login node or ARE VDI
    - `ssh -X gadi.nci.org.au`
  - Persistent session (from Gadi login node or ARE VDI)
    - `ssh -Y <name>.$USER.$PROJECT.ps.gadi.org.au`
- Metascheduler: The host executing the cylc main process. It always utilises persistent session specified by
  - `$CYLC_SESSION` or
  - `~/.persistent-sessions/cylc-session`.

Supported Cylc 7 Suite Modes

Accessdev	Gadi: Login nodes, ARE VDI, persistent session
Network connection	Network connection
Seperate /home Selected /g/data	Common /home, /scratch and /g/data
No PBS job submission	PBS job submission
Need remote connection to push jobs to Gadi	All jobs can run locally or being pushed to PBS
Suite: Accessdev mode	Suite: Localhost mode Accessdev compatible mode

```
$ cat ~/roses/u-da543/suite.rc
[runtime]
      [[root]]
      …
      [[local]]
         [[[remote]]]
            host = localhost
         [[[job]]]
            batch system = background
      …
      [[HPC]]
         [[[remote]]]
            host = localhost
         [[[job]]]
            batch system = pbs
      …
      [[GADI]]
         [[[remote]]]
            host = gadi
         [[[job]]]
            batch system = pbs
      …
```

Localhost mode (host=localhost)

Accesdev compatible mode (host=gadi,gadi.nci.org.au,…)

Connection between ROSE_ORIG_HOST and metascheduler

- Suite localhost mode: No remote connections.
- Suite accessdev mode: Connection between Gadi (gadi.nci.org.au) from $ROSE_ORIG_HOST
  - ISSUE: Gadi login node, VDI and persistent session can not connect to Gadi (gadi.nci.org.au)
  - SOLUTION: We produced specific ssh config to enable connections within rose/cylc/fcm.
    - Persistent session ==> Gadi (gadi.nci.org.au) utilises the existing ssh key.
    - Gadi login node & VDI ==> Gadi (gadi.nci.org.au) needs to create the ssh key pair (via an initialisation step).

All settings have no impact on user's global ssh environment.

$ROSE_ORIG_HOST	Localhost Mode	Accessdev Compatible Mode
ARE VDI Gadi Login Node	No remote host detection	Connection from VDI/Gadi login nodes to Gadi (gadi.nci.org.au) requires a specific SSH key pair produced by an initialization script
Persistent Session	No remote host detection	Connection from persistent session to Gadi (gadi.nci.org.au) utilises the existing SSH key pair

Localhost mode is highly recommended in Gadi.

Cylc 8 Platforms

```
$ cat ~/roses/u-cz535/flow.cylc
```

```
[runtime]
    [[root]]
    …
    [[local]]
        inherit = root
        platform = localhost
    [[HPC]]
        inherit = root
        platform = pbs
```

* Some platforms such as gadi, gadi_pbs, gadi_localhost and gadi_background are still supposed in the current cylc 8 module, in order to transition existing Cylc7 suites to Cylc 8. They will deprecated in the future cylc8 modules.

General workflow

- **ps auwxff | grep $USERNAME**
- **ssh $CYLC_SESSION “ps auwxff” | grep $USERNAME**

```
login $ ssh -Y workshopdemo.abc111.xy99.ps.gadi.nci.org.au
 demo $ module use /g/data/hr22/modulefiles
 demo $ module load cylc7/23.09
 demo $ moss-auth
 demo $ rosie co u-da543
 demo $ cd ~/roses/u-da543
 demo $ rose suite-run
```

Replace cylc7/23.09 with cylc/8.2.1 if working with Cylc8 suites

You can monitor the workflow progress later from Gadi login nodes, ARE VDI or persistent session.

Hands-on Excise: Make Your Environment Ready for running Cylc in Gadi

Cylc7 Environment: https://github.com/NCI-HPCDI/cylc-workflow/blob/main/cylc7-guide.md
- ROSE_ORIG_HOST=persistent session
- ROSE_ORIG_HOST=gadi login nodes

Cylc8 Environment: https://github.com/NCI-HPCDI/cylc-workflow/blob/main/cylc8-guide.md
ROSE_ORIG_HOST= ARE VDI (for both executing and monitoring)

After the excises, your environment will be ready to run both cylc7 and cylc 8 workflows in Gadi.

Advanced Excise: Cylc 7 Working Directories

Default: the whole ~/cylc-run directory is linked to /scratch/$PROJECT/$USER/cylc-run

You can change these links in ~/.metomi/rose.conf

Default	Customised

No changes needed.	Add the following configurations into ~/.metomi/rose.conf
	
	`[rose-suite-run]`
	`root-dir=*=/g/data/$PROJECT/$USER`
	`root-dir{share}=*=/scratch/$PROJECT/$USER`
	`root-dir{work}=*=/scratch/$PROJECT/$USER`

`$ readlink ~/cylc-run/u-da543`
`/scratch/fp0/jpf777/cylc-run/u-da543`
	`$ readlink ~/cylc-run/u-da543`
	`/g/data/fp0/jpf777/cylc-run/u-da543`
	`$ readlink ~/cylc-run/u-da543/work`
	`/scratch/fp0/jpf777/cylc-run/u-da543/work`
	`$ readlink ~/cylc-run/u-da543/share`
	`/scratch/fp0/jpf777/cylc-run/u-da543/share`

Advanced Excise Guide: https://github.com/NCI-HPCDI/cylc-workflow/blob/main/advanced_guide.md

Advanced Excise: Cylc 8 Working Directories

Default: log, share and work are linked to /scratch/$PROJECT/$USER
Change these settings via ~/.cylc/flow/global.cylc

```
$ cat ~/.cylc/flow/global.cylc
#!Jinja2
[install]
    source dirs = ~/cylc-src, ~/roses
    [[symlink dirs]]
        [[[gadi]]]
            log = /g/data/{{environ['PROJECT']}}/{{environ['USER']}}
            share = /scratch/{{environ['PROJECT']}}/{{environ['USER']}}
            work = /scratch/{{environ['PROJECT']}}/{{environ['USER']}}
        [[[localhost]]]
            log = /g/data/{{environ['PROJECT']}}/{{environ['USER']}}
            share = /scratch/{{environ['PROJECT']}}/{{environ['USER']}}
            work = /scratch/{{environ['PROJECT']}}/{{environ['USER']}}
```

```
$ readlink  ~/cylc-run/u-cz535/runN/log
/g/data/fp0/jpf777/cylc-run/u-cz535/run1/log
$ readlink  ~/cylc-run/u-cz535/runN/share
/scratch/fp0/jpf777/cylc-run/u-cz535/run1/share
$ readlink  ~/cylc-run/u-cz535/runN/work
/scratch/fp0/jpf777/cylc-run/u-cz535/run1/work
```

Advanced Excise Guide: https://github.com/NCI-HPCDI/cylc-workflow/blob/main/advanced_guide.md

Advanced Excise: User Script with Remote Connections

Most remote connections may be unnecessary; consider removing them.

If you still require these remote connections, please add the following configurations to ~/.ssh/config

```
Host gadi.nci.org.au
   Match exec "echo '%l' | grep -q 'ps.gadi.nci.org.au'" host gadi,gadi.nci.org.au,localhost
        HostName localhost
  IdentityFile ~/.persistent-sessions/%l/user.key
  StrictHostKeyChecking no
  Port 2222
```

Excise: remove “#” from the following line in ~/roses/u-da543/suite.rc and run the suite again after adding the above configurations to ~/.ssh/config

```
#     scp -r STREAM gadi.nci.org.au:~/
```

Possible risks of encountering conflicts in other work.

Advanced Excise Guide: https://github.com/NCI-HPCDI/cylc-workflow/blob/main/advanced_guide.md

Back @ 13:00/14:00

https://opus.nci.org.au/display/DAE/UK+Met+Office+Environment+on+NCI

One-On-One: Problem Solving

- Breakout Rooms: Rui/Yue/Martin/Dale
- Opus Documentation

  https://opus.nci.org.au/display/DAE/UK+Met+Office+Environment+on+NCI

Known Issue Addressed

- 2-step fcm_make
- Suite Restart
- General Porting Notes

Suite migration

- Recommend using localhost and skipping the ssh configuration

- Many suites (e.g. ACCESS-CM2) use a 2 step fcm make to build UM
  - Extract code on accessdev, mirror to gadi
  - Build on gadi

- Using localhost breaks the rose mirror logic
  - Easy to switch to a single step fcm make in recent UM versions (>= 12.1)
  - With older versions (e.g. CM2) can change the fcm make configuration to force the mirror to work or to do it in a single step

Removing mirror step in portable suites

- GC5 AMIP test suite (u-ct505)

```
{# Test to determine when to use double fcm_make_um task #}
{% set DBLE_FCMUM = ['archer', 'nci_gadi'] %}
```

- RA3 suite

```
# Build task for single nesting suite executable.
{% if ONE_STEP_BUILD %}
  [[fcm_make_um]]
      inherit = BUILD_HPC
{% else %}
  [[fcm_make_um]]
      inherit = BUILD_LOCAL
  [[fcm_make2_um]]
      inherit = BUILD_HPC
{% endif %}
```

Removing mirror step in newer suites (u-aa124)

```
# Extract, Mirror
  [[fcm_make]]
    inherit = FCM, LINUX
    [[[environment]]]
      ROSE_APP_OPT_CONF_KEYS = mirror

# Pre-process, Build
  [[fcm_make2]]
```

https://code.metoffice.gov.uk/trac/roses-u/changeset/271476/a/a/1/2/4

```
============================================================================
--- suite.rc    (revision 271475)
+++ suite.rc    (revision 271476)
@@ -8,3 +8,3 @@
 {%- if BUILD %}
-        fcm_make => fcm_make2 => \
+        fcm_make => \
 {%- endif %}
@@ -36,3 +35,3 @@
         [[[remote]]]
-            host = gadi.nci.org.au
+            host = localhost
         [[[directives]]]
@@ -57,10 +50,3 @@

-    # Extract, Mirror
     [[fcm_make]]
-        inherit = FCM, LINUX
-        [[[environment]]]
-            ROSE_APP_OPT_CONF_KEYS
= mirror
-
-    # Pre-process, Build
-    [[fcm_make2]]
         pre-script = """
```

ACCESS CM2 suite migration

- Driver scripts
  - Used to come from https://trac.nci.org.au/svn/access_tools/access-cm2-drivers
  - Replaced by `https://github.com/ACCESS-NRI/access-cm2-drivers.git`
  - Clone task needs to run on copyq to access network
- UM build
  - 2 step with forced mirror or 1 step
- MOM and CICE builds
  - Still 2 step builds but the rsync step isn't necessary

CM2 migration example: u-db130 from u-br565 PI control

- Two step UM fcm make

  https://code.metoffice.gov.uk/trac/roses-u/changeset/271474/d/b/1/3/0

  In app/fcm_make/file/fcm-make.cfg

```
# Need to set explicitly for mirroring to work when using localhost
mirror.target = $CYLC_SUITE_SHARE_DIR/$ROSE_TASK_NAME
mirror.prop{config-file.name} = 2
```

- One step fcm make

  https://code.metoffice.gov.uk/trac/roses-u/changeset/271479/d/b/1/3/0

  In app/fcm_make/rose-app.conf

```
mirror=
steplist=extract preprocess-atmos preprocess-recon build-atmos build-recon
```

Moving long running suites

- Stop suite on accessdev at completion of a cycle
  - E.g. for a CM2 suite with 6 monthly runs, `cylc stop SUITE 20000107`
- Make sure all suite changes are committed and then check out on gadi (or just rsync the roses/SUITE directory)
- Add `gdata/hr22` to storage directives
- Change all suite instances of `host = gadi` to `host = localhost`
  - `COMPUTE_HOST` in `rose-suite.conf` for CM2
- *No need to change build related tasks because not needed in restart*
- `rose suite-run -- --warm --hold START_POINT`

  `START_POINT` is date stamp of the step following where suite was stopped, e.g. 20010101

Moving long running suites

```
~/roses/u-br565% svn diff
Index: rose-suite.conf
===================================================================
--- rose-suite.conf     (revision 257668)
+++ rose-suite.conf     (working copy)
@@ -17 +17 @@
-COMPUTE_HOST='gadi.nci.org.au'
+COMPUTE_HOST='localhost'
Index: suite.rc
===================================================================
--- suite.rc    (revision 257668)
+++ suite.rc    (working copy)
@@ -144 +144 @@
-            -l storage = gdata/{{PROJECT}}+gdata/access+gdata/hh5
+            -l storage = gdata/{{PROJECT}}+gdata/access+gdata/hh5+gdata/hr22
@@ -510 +510 @@
-            -l storage = gdata/{{PROJECT}}+gdata/access+gdata/hh5
+            -l storage = gdata/{{PROJECT}}+gdata/access+gdata/hh5+gdata/hr22
```

Example migrated suites

- u-aa124 Simple vn13.3 n48
- u-cz789 CM2 with 0.25 ocean
- u-da593 vn13.1 CM3 development with NUOPC coupling
- u-da793 vn13.1 AMIP CABLE development
- u-db130 migrated version of CM2 PI control

Traps

```
[[[remote]]]
                host = {{ ROSE_ORIG_HOST }}
```

will fail to submit

- Missing `gdata/hr22` storage directive
  - Task will submit & fail but can't connect with cylc so GUI stuck showing it as submitted. job.err has

```
/local/spool/pbs/mom_priv/jobs/100468201.gadi-
pbs.SC: line 91: /g/data/hr22/apps/cylc7/cylc_7.9.7/
lib/cylc/job.sh: No such file or directory
```

Extras

- Rose stem testing
  - UM and JULES roses stem NCI configurations are being updated to run with persistent sessions
- Cylc 8
  - Many suites run ok in backwards compatibility mode
  - CM2 suites don’t yet (problems with directory paths)

Share Suite Run Progress

https://opus.nci.org.au/display/DAE/Sharing+Job+Progress+with+Collaborators

- Start Cylc Review Server
- Monitor Job Progress for Collaborators
- Notes on Data Access
- Share Outside Gadi
  - Nirin instance to host the web server
  - /g/data/ accommodates the static HTML directories