



# Introduction to Gadi

NCI Training

Yue Sun, User Services, NCI

Roger Edberg, Manager, User Services, NCI

11 May 2021

# Overview

- NCI: a collaboration of stakeholders representing Australian research community
- Gadi: NCI peak HPC system
- Get access to Gadi
- Showcase
  - Login
  - Basic application workflow
  - Submit and monitor jobs
  - Monitor project usage
  - Use applications installed on /apps
  - Build your own python/R/Julia packages on Gadi
  - Monitor job efficiency

NCI & Gadi

# National Computational Infrastructure

# What is NCI

- HPC for Australian researchers
  - 55 stakeholders, including universities, research institutes, and Australian Government agencies
  - 2600+ projects
  - 5500+ users
- Services
  - High performance computing
  - Data storage and services: THREDDS and VDI etc
  - Cloud computing



NCI & Gadi

# Gadi: Australia's Peak Research Supercomputer

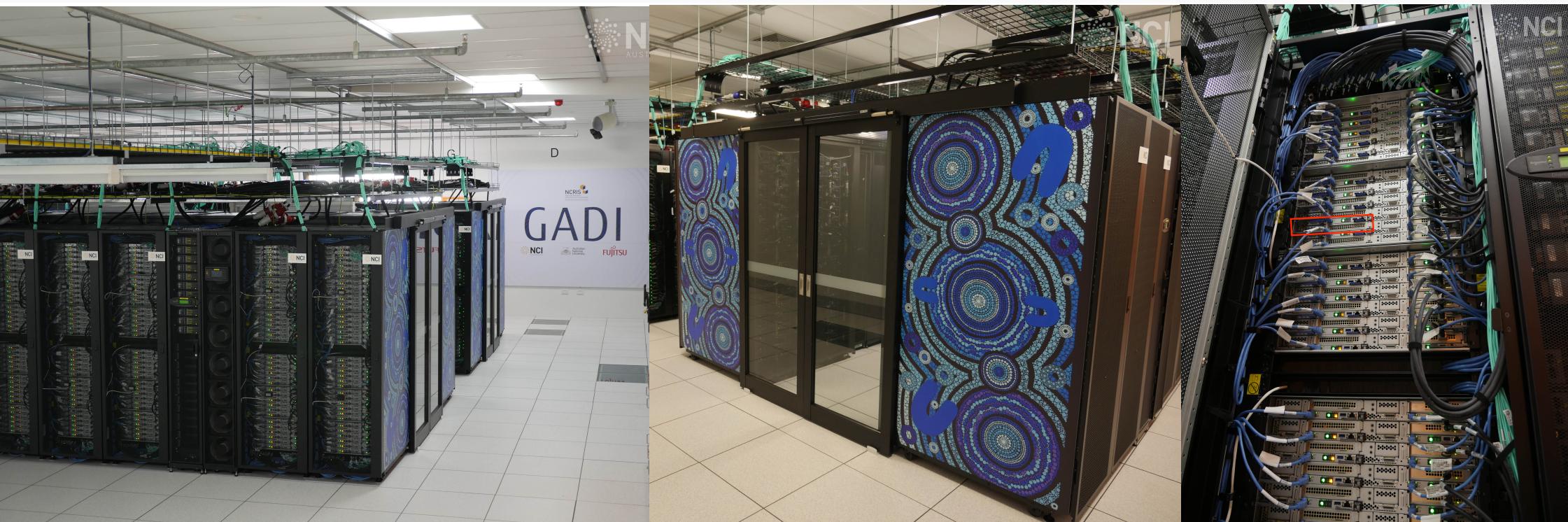
## What is Gadi

- High performance computing system operated by NCI
- Pronounced `gar dee`, means `*to search for*` in the language of the Ngunnawal people
- Located on ANU Acton campus
- Launched in 2019 , replaced Raijin (2013-2020)
- In 2021 Q1, Gadi ran ~4M jobs for 1659 users from 650 projects

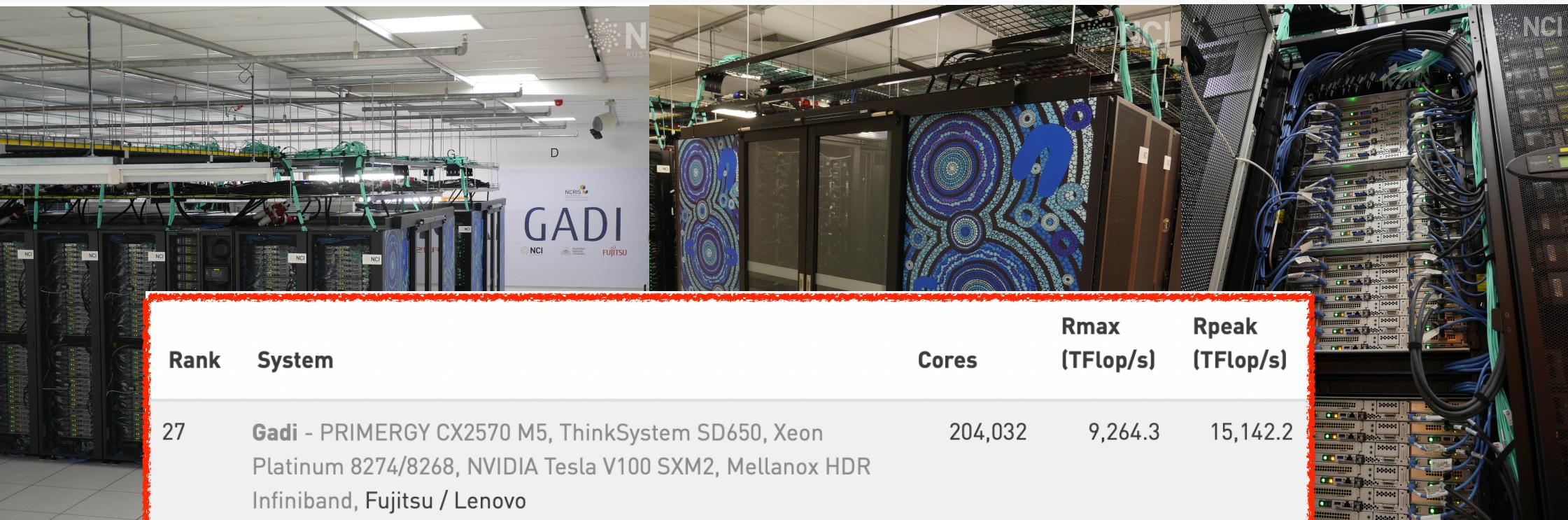
# What is Gadi

- Supercomputer
  - 4000+ compute nodes, including 160 GPU nodes
  - 10 login nodes
  - 6 data mover nodes
  - PiB parallel file systems
  - Mellanox HDR InfiniBand interconnect network in Dragonfly+ topology
  - Application software catalogue
  - PBS Pro server 2021.1
  - CentOS 8.3.2011

# Gadi

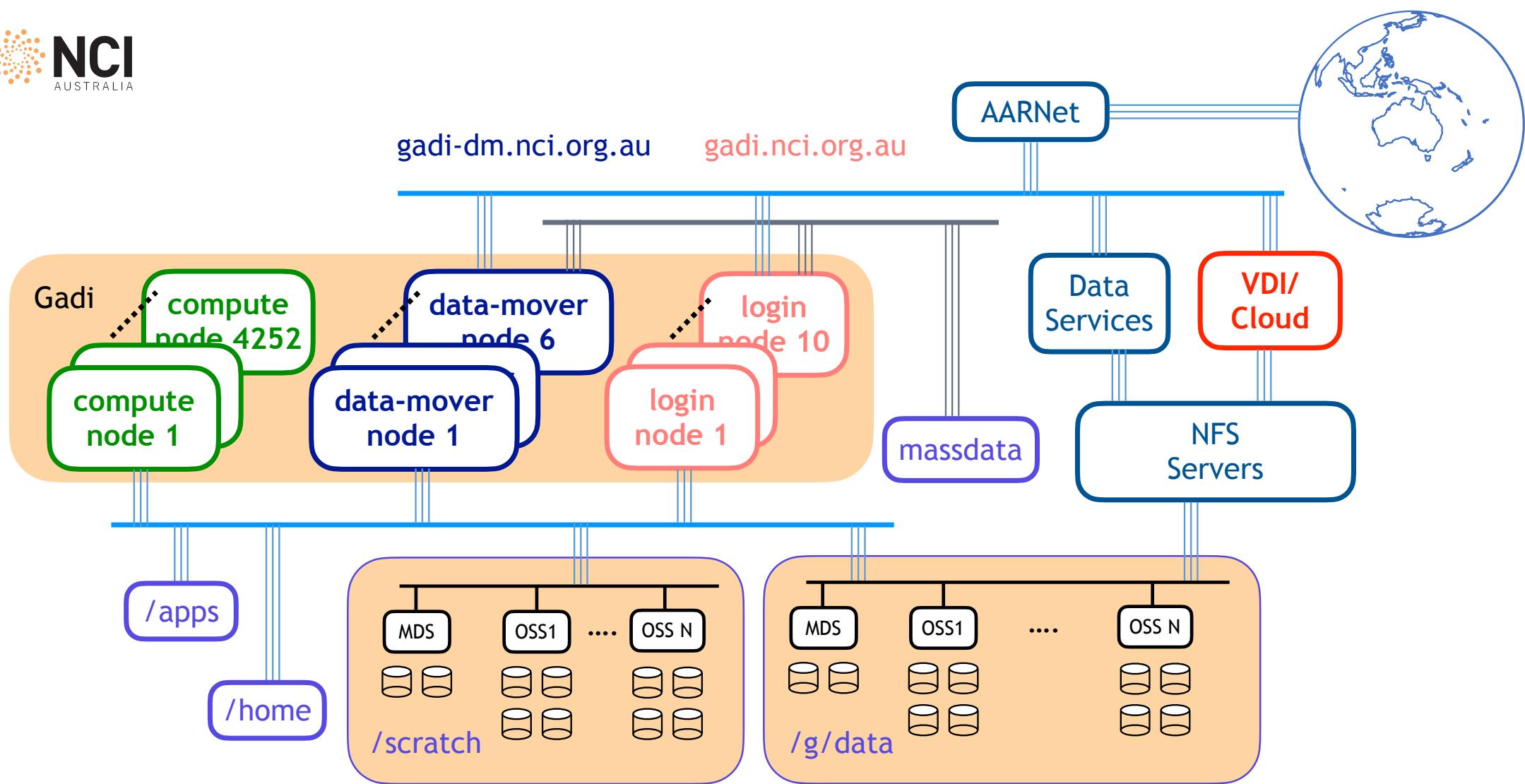


# Gadi



Rank	System	Cores	Rmax (TFlop/s)	Peak (TFlop/s)
27	<b>Gadi</b> - PRIMERGY CX2570 M5, ThinkSystem SD650, Xeon Platinum 8274/8268, NVIDIA Tesla V100 SXM2, Mellanox HDR Infiniband, Fujitsu / Lenovo National Computational Infrastructure (NCI Australia) Australia	204,032	9,264.3	15,142.2

<https://www.top500.org/lists/top500/list/2020/11/>



# Gadi

- Access to more CPU, GPU, memory, storage and faster interconnect between them
  - CPU cores: (Cascade Lake, 155K), (Broadwell, 23K), (Skylake, 6K)
  - GPUs: 160\*4 NVIDIA V100
  - High memory: (4\*clx, 3TiB), (50\*clx, 1.5TiB), (3\*bdw, 3TiB), (10\*bdw, 1TiB)
  - Parallel file systems: (/scratch, 11PiB)\*2, (/g/data, ~11PiB)\*4, (massdata, 25PiB)\*2 extensible to 100 PiB each.
- Application software catalogue
  - MPI compiled for utilising the faster interconnections
  - Python3, R, Julia compiled with libraries such as intel-mkl, FFTW, and possibility to build your own packages
  - Matlab for almost all major research universities, like ANU, UNSW, RMIT, Monash, USyd, Macquarie...
  - Deep Learning Frameworks: tensorflow, pytorch

# Gadi

- Shared resources
  - Read and write operations to and from the shared file systems
  - For jobs use less than a full compute node
- Gadi runs under CentOS 8.3.2011 as of 28 April 2021
- Most applications for scientific computing support CentOS/RHEL
- No root privilege like you have on your local Ubuntu/MacOS/Windows
  - Installation goes to your project/home directories, not the default system directories like /bin
  - Ask for help when changing file/directory ownership and permission
- Not Ideal for GUI, but possible with X forwarding enabled
- 48 hours default limit for normal queue jobs using less than 720 CPU cores\*

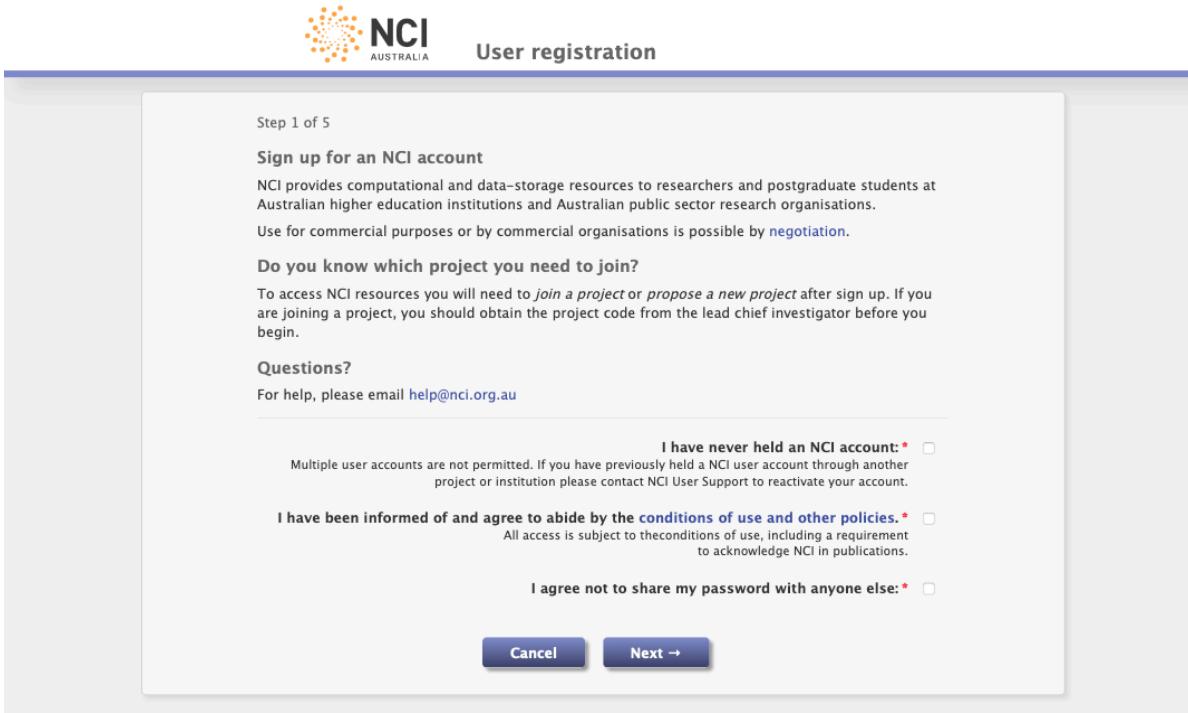
\* walltime limit varies and exception can be granted if necessary, see details here <https://opus.nci.org.au/display/Help/Queue+Limits>



Get Access

# User, Project and Stakeholder

# User Registration at [my.nci.org.au](https://my.nci.org.au)



The screenshot shows the first step of a five-step user registration process. The title bar reads "User registration". The main content area is titled "Step 1 of 5" and "Sign up for an NCI account". It provides information about NCI resources being available to researchers and postgraduate students at Australian higher education institutions and Australian public sector research organisations. It also mentions that use for commercial purposes or by commercial organisations is possible by negotiation. A section titled "Do you know which project you need to join?" explains that access requires joining a project or proposing a new one. Below this, a "Questions?" section provides email support information. Three checkboxes are present: "I have never held an NCI account.", "I have been informed of and agree to abide by the conditions of use and other policies.", and "I agree not to share my password with anyone else.". At the bottom are "Cancel" and "Next →" buttons.

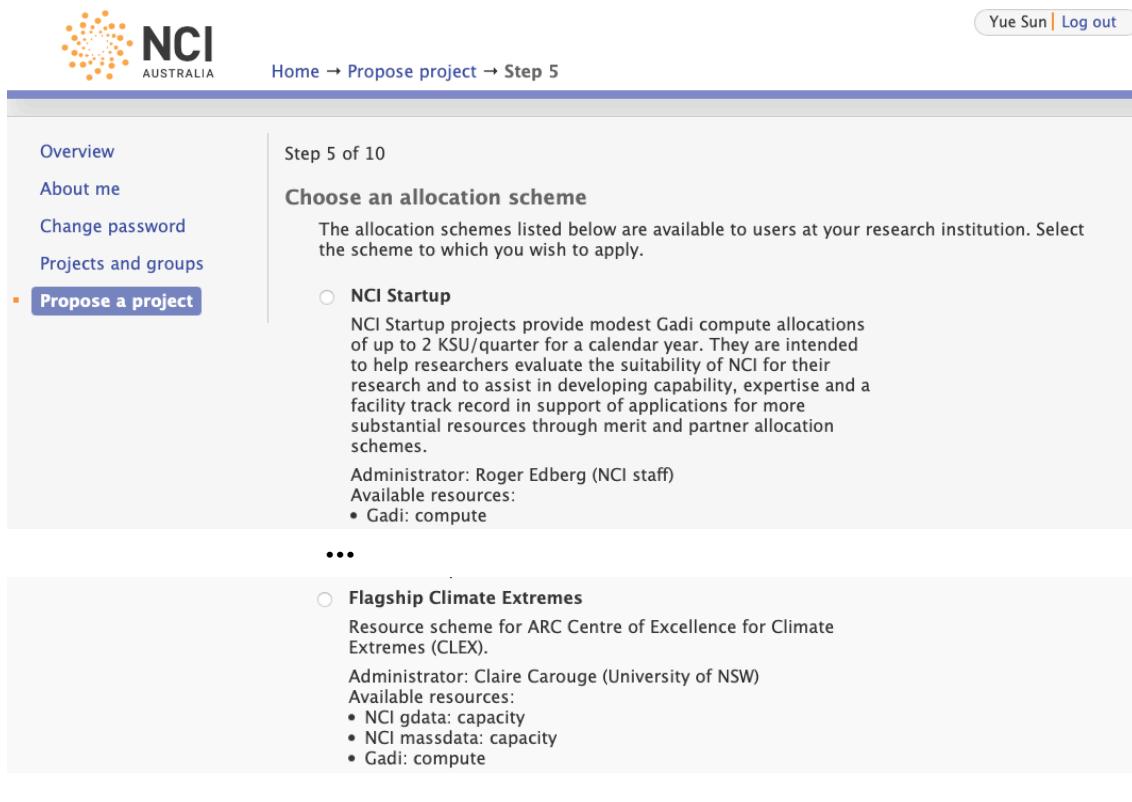
Follow the registration dialogue:

- Confirm no previous account
- Provide your details:
  - Institution email address
  - Mobile number
  - Institution
- Project code

\* see more details here

<https://opus.nci.org.au/display/Help/1.0.0+User+Account+Registration>

# Propose Project after User Registration



Home → Propose project → Step 5

Yue Sun | Log out

Overview

About me

Change password

Projects and groups

Propose a project

Step 5 of 10

**Choose an allocation scheme**

The allocation schemes listed below are available to users at your research institution. Select the scheme to which you wish to apply.

**NCI Startup**  
NCI Startup projects provide modest Gadi compute allocations of up to 2 KSU/quarter for a calendar year. They are intended to help researchers evaluate the suitability of NCI for their research and to assist in developing capability, expertise and a facility track record in support of applications for more substantial resources through merit and partner allocation schemes.  
Administrator: Roger Edberg (NCI staff)  
Available resources:

- Gadi: compute

...

**Flagship Climate Extremes**  
Resource scheme for ARC Centre of Excellence for Climate Extremes (CLEX).  
Administrator: Claire Carouge (University of NSW)  
Available resources:

- NCI gdata: capacity
- NCI massdata: capacity
- Gadi: compute

...

## Steps:

- Acknowledge Terms and Conditions of Use
- Write project proposal
  - Project description
  - Research field classification
  - Resources request
- Choose which scheme to propose to
- Invite project members on board

# Propose Project after User Registration



Home → Propose project → Step 5

Yue Sun | Log out

Overview

About me

Change password

Projects and groups

**Propose a project**

Step 5 of 10

**Choose an allocation scheme**

The allocation schemes listed below are available to users at your research institution. Select the scheme to which you wish to apply.

**NCI Startup**  
NCI Startup projects provide modest Gadi compute allocations of up to 2 KSU/quarter for a calendar year. They are intended to help researchers evaluate the suitability of NCI for their research and to assist in developing capability, expertise and a facility track record in support of applications for more substantial resources through merit and partner allocation schemes.  
Administrator: Roger Edberg (NCI staff)  
Available resources:

- Gadi: compute

...

**Flagship Climate Extremes**  
Resource scheme for ARC Centre of Excellence for Climate Extremes (CLIFX).  
Administrator: Claire Carouge (University of NSW)  
Available resources:

- NCI gdata: capacity
- NCI massdata: capacity
- Gadi: compute

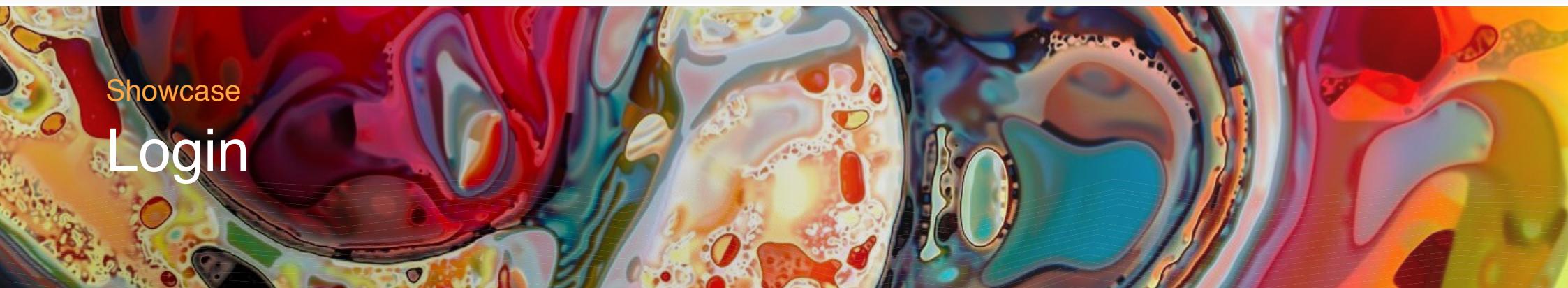
...

## Steps:

- Acknowledge Term and Conditions of Use
- Write the proposal
  - Project description
  - Research field classification
  - Resources request
- Choose which scheme to propose to
- Invite project members on board

## Current Stakeholders

- Universities: ANU, UNSW, Sydney, Melbourne, RMIT, Monash, Deakin, UQ, Adelaide, Macquarie, Wollongong, UTS, UTAS, Victoria University
- Australian Government research agencies, Bureau of Meteorology, CSIRO, and Geoscience Australia
- ARC Centres of Excellence, CLEX, FLEET, Exciton Science
- Consortia: QCIF, Intersect
- Merit schemes: NCMAS, ALCG



Showcase  
Login

## Login to Gadi

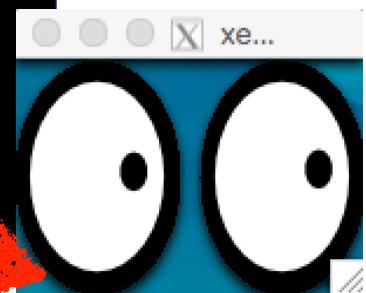
- Open a terminal and login to one of the ten login nodes by doing

```
ssh <username>@gadi.nci.org.au
```

Your username goes here

- X forwarding enabled login and test with xeyes

```
Sue@local:~ $ ssh -Y jjj777@gadi.nci.org.au
jjj777@gadi.nci.org.au's password:
[jjj777@gadi-login-05 ~]$ xeyes
[jjj777@gadi-login-05 ~]$ exit
Sue@local:~ $
```



Showcase

# Basic Application Workflow on Gadi

Showcase

# Submit and Monitor Jobs

## Submit Jobs and Monitor Status

```
qsub gutentag.sh  
qstat -u $USER -sw  
qstat -fx <jobid>  
cat <jobname>.o<jobid>
```

Used: 62.00 GB      **WATERLINE USE**  
d: 100.0MB      **JobFS use**

# Job Submission Script

```
#!/bin/bash
#PBS -P ab01                                Project Ownership
#PBS -q copyq
#PBS -l ncpus=1
#PBS -l mem=2GB
#PBS -l walltime=02:00:00                      Resource Requests → Reserved SU
#PBS -l storage=scratch/ab01+gdata/ab01
#PBS -l wd

hostname
export DSTDIR=/g/data/ab01/xyz123/job_archive
export SOURCEDIR=/scratch/ab01/xyz123/benchmarks/python/tensorflow
cp -vr $SOURCEDIR $DSTDIR > ${PBS_JOBID}.log
```

# Job Submission Script

```
#!/bin/bash

#PBS -P ab01
#PBS -q copyq
#PBS -l ncpus=1
#PBS -l mem=2GB
#PBS -l walltime=02:00:00
#PBS -l storage=scratch/ab01+gdata/ab01
#PBS -l wd

hostname
export DSTDIR=/g/data/ab01/xyz123/job_archive
export SOURCEDIR=/scratch/ab01/xyz123/benchmarks/python/tensorflow
cp -vr $SOURCEDIR $DSTDIR > ${PBS_JOBID}.log
```

Actual Tasks: hostname, cp

Run your own applications

## Submit an Interactive Job

```
qsub -I -lstorage=gdata/c25+scratch/x11,wd job.sh
```

- Add PBS directives so that the job
  - sends you email at start: `-M <abc123>@<gmail.com> -m abe`
  - waits until matlab licenses is available: `-lsoftware=matlab_<unsw>`
  - redirects STDOUT and STDERR into the specific log(s):
    - `-e err.log -o /scratch/c25/abc123/Logs/`
    - `-j eo`
  - waits until 1:55pm to start: `-a 202105111355`

Extended Reading <https://opus.nci.org.au/display/Help/PBS+Directives+Explained>

Showcase

# Monitor Project Usage

```
$ nci_account -P ab01
```

```
Usage Report: Project=ab01 Period=2021.q2
```

```
=====
```

Grant:	471.00	KSU
Used:	406.30	KSU
Reserved:	0.00	SU
Avail:	64.70	KSU

```
=====
```

```
Storage Usage Report: Project=ab01
```

```
=====
```

Filesystem	Used	iUsed	Allocation	iAllocation
gdata1a	513.57 GB	6.00	2.00 TB	312.00 K

```
=====
```

```
$  
$  
$  
$  
$  
$  
$ nci_account -P ab01 -p 2021.q1
```

```
Usage Report: Project=ab01 Period=2021.q1
```

```
=====
```

Grant:	548.00	KSU
Used:	524.56	KSU
Reserved:	0.00	SU
Avail:	23.44	KSU

```
=====
```

```
Storage Usage Report: Project=ab01
```

```
=====
```

Filesystem	Used	iUsed	Allocation	iAllocation
gdata1b	513.57 GB	6.00	2.00 TB	312.00 K

```
=====
```

```
$  
$  
$  
$  
$  
$  
$  
$ nci_account -P ab01 -p 2021.q1 -v
```

```
Usage Report: Project=ab01 Period=2021.q1
```

```
=====
```

Grant:	548.00	KSU
Used:	524.56	KSU
Reserved:	0.00	SU
Avail:	23.44	KSU

```
=====
```

Stakeholder	Grant	Used	Avail
MAS	250.00 KSU	249.65 KSU	354.93 SU
UNSW	298.00 KSU	274.92 KSU	23.08 KSU

User	Used	Reserved
abc321	372.50 KSU	0.00 SU
xyz123	152.06 KSU	0.00 SU

```
Storage Usage Report: Project=a57
```

```
=====
```

Filesystem	Used	iUsed	Allocation	iAllocation
gdata1b	513.57 GB	6.00	2.00 TB	312.00 K

```
=====
```

Stakeholder	Allocation	iAllocation
UNSW	2.00 TB	312.00 K

# Monitoring Grant and Usage

- Project can receive compute/storage grant from multiple schemes
- One SU
  - supports a single CPU core job submitted to normal queue with up to 4GiB memory request to run for 30 minutes
  - has the nominal value of 4 cents for grant application purpose
- Storage
  - 72 GB /scratch/\$PROJECT by default, more available to meet job demand, auto-purge policy applied
  - /g/data/\$PROJECT needs approval from scheme managers

Extended Reading <https://opus.nci.org.au/display/Help/2.2+Job+Cost+Examples>

```
$ nci-files-report -f scratch --group xy12
```

FILESYSTEM	SCAN DATE	PROJECT	GROUP	USER	SPACE USED	TOTAL SIZE	COUNT
scratch	2020-12-27	am1	xy12	xyz123	2.0G	2.0G	3968
scratch	2020-12-27	ab01	xy12	xyz123	48.0K	11.8K	21
scratch	2020-12-27	ab01	xy12	abc321	32.0K	4.8K	12
scratch	2020-12-27	xy12	xy12	xyz123	396.0K	211.8K	97
scratch	2020-12-27	xy12	xy12	abc321	29.7G	29.7G	7258

```
$ nci-files-report -f scratch --project xy12
```

FILESYSTEM	SCAN DATE	PROJECT	GROUP	USER	SPACE USED	TOTAL SIZE	COUNT
scratch	2020-12-27	xy12	c25	xyz123	2.7G	2.7G	1652
scratch	2020-12-27	xy12	xy12	xyz123	396.0K	211.8K	97
scratch	2020-12-27	xy12	xy12	abc321	29.7G	29.7G	7258

```
$ nci-files-report -f scratch --user xyz123
```

FILESYSTEM	SCAN DATE	PROJECT	GROUP	USER	SPACE USED	TOTAL SIZE	COUNT
scratch	2020-12-27	am1	xy12	xyz123	2.0G	2.0G	3968
scratch	2020-12-27	ab01	xy12	xyz123	48.0K	11.8K	21
scratch	2020-12-27	xy12	c25	xyz123	2.7G	2.7G	1652
scratch	2020-12-27	xy12	xy12	xyz123	396.0K	211.8K	97

Showcase

# Use Applications installed on /apps

# Module Environment

- Multiple versions and conflicting packages
- module commands
  - module avail python3
  - module list
  - module show python3/3.9.2
  - module load python3/3.9.2
  - module unload python3

```
$ module avail python3
-----/apps/Modules/modulefiles -----
python3-as-python  python3/3.7.4  python3/3.8.5  python3/3.9.2
```

```
$ module list
Currently Loaded Modulefiles:
 1) pbs
```

```
$ module show python3/3.9.2
-----
/apps/Modules/modulefiles/python3/3.9.2:
```

```
prepend-path      PATH /apps/python3/3.9.2/bin
prepend-path      C_INCLUDE_PATH /apps/python3/3.9.2/include/python3.9
prepend-path      CPLUS_INCLUDE_PATH /apps/python3/3.9.2/include/python3.9
prepend-path      CPATH /apps/python3/3.9.2/include/python3.9
prepend-path      FPATH /apps/python3/3.9.2/include/python3.9
prepend-path      LIBRARY_PATH /apps/python3/3.9.2/lib
prepend-path      LD_LIBRARY_PATH /apps/python3/3.9.2/lib
prepend-path      LD_RUN_PATH /apps/python3/3.9.2/lib
prepend-path      MANPATH /apps/python3/3.9.2/share/man
prepend-path      PKG_CONFIG_PATH /apps/python3/3.9.2/lib/pkgconfig
module          load intel-mkl/2020.3.304
conflict        python3
setenv          PYTHON3_BASE /apps/python3/3.9.2
setenv          PYTHON3_ROOT /apps/python3/3.9.2
setenv          PYTHON3_VERSION 3.9.2
module-whatis   {python3, version 3.9.2}
-----
```

```
$ module load python3/3.9.2
Loading python3/3.9.2
  Loading requirement: intel-mkl/2020.3.304
```

```
$ module list
Currently Loaded Modulefiles:
 1) pbs    2) intel-mkl/2020.3.304    3) python3/3.9.2
```

```
$ python3
Python 3.9.2 (default, Mar 29 2021, 10:41:26)
[GCC 8.3.1 20191121 (Red Hat 8.3.1-5)] on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> import numpy as np
>>> np.__version__
'1.20.0'
>>> exit()
```

```
$ module load python3/3.8.5
MODULE ERROR DETECTED: GLOBALERR python3/3.8.5 cannot be loaded due to a
conflict.
(Detailed error information and backtrace has been suppressed, set
$MODULES_ERROR_BACKTRACE to unsuppress.)
```

```
Loading python3/3.8.5
  ERROR: python3/3.8.5 cannot be loaded due to a conflict.
  HINT: Might try "module unload python3" first.
```

# Software Catalogue

- Modules on Gadi are named as <software>/<version>
  - always load a specific version of the software application under interest
  - `module load <software>` loads the default version which changes over time
- `module load` does the following
  - modifies search/exec path
  - loads prerequisite modules [ intel-mkl/2020.3.304 for python3/3.9.2]
  - handles software/module conflicts
  - configures environment to define how the application runs

## Software Groups

- Restricted Modules are available to specific groups of users
- Software groups control access to license modules
  - Try: module avail matlab
- License Modules tell the application where to checkout license

```
$ module avail matlab
-----/apps/Modules/restricted-modulefiles/matlab_anu -----
matlab_licence/anu

-----/apps/Modules/restricted-modulefiles/matlab_usyd -----
matlab_licence/usyd

-----/apps/Modules/restricted-modulefiles/matlab_vu -----
matlab_licence/vu

-----/apps/Modules/restricted-modulefiles/matlab_utas -----
matlab_licence/utas

-----/apps/Modules/modulefiles -----
matlab/R2019b  matlab/R2020b
```

```
$ getfacl /apps/Modules/restricted-modulefiles/matlab_anu
getfacl: Removing leading '/' from absolute path names
# file: apps/Modules/restricted-modulefiles/matlab_anu
# owner: apps
# group: z30
user::rwx
group::rwx
group:matlab_anu:r-x
mask::rwx
other::---
```

```
$ module show matlab_licence/anu
-----
/applications/Modules/restricted-modulefiles/matlab_anu/matlab_licence/anu:

setenv      MLM_LICENSE_FILE 12345678@xxxxxxxxxx.anu.edu.au
conflict    matlab_licence
module-whatis {matlab_licence, version anu}
-----
```

```
$ module load matlab/R2020b
$ module load matlab_licence/anu
$ matlab -nodesktop -nodisplay
```

```
          < M A T L A B (R) >
Copyright 1984-2020 The MathWorks, Inc.
R2020b Update 5 (9.9.0.1592791) 64-bit (glnxa64)
February 4, 2021
```

```
To get started, type doc.
For product information, visit www.mathworks.com.
```

```
>> eig(magic(3))

ans =

15.0000
 4.8990
-4.8990

>> exit()

$ ls
test.m
$ matlab -nodesktop -nodisplay -r "var1='$PBS_JOBFS',var2=$PBS_NCPUS, test, exit"
```

Showcase

# Build Your Own Python/R/Julia Packages on Gadi

## Example I: python3 pandas

```
module purge
module load python3/3.9.2
InstallDir=/scratch/$PROJECT/$USER/.local/python
mkdir -p $InstallDir
pip3 install -v --no-binary :all: -prefix=$InstallDir pandas

export PYTHONPATH=$InstallDir/lib/python3.9/site-packages:$PYTHONPATH
python3
>>> import pandas as pd
```



## Example II: python3 neural-structured-learning

```
module purge
module load tensorflow/2.4.1
pip3 install -v --no-binary :all: --prefix=$InstallDir neural-structured-learning==1.2.0

python3
>>> import neural_structured_learning as nsl
```

```
$ module purge
$ module load tensorflow/2.4.1
Loading tensorflow/2.4.1
  Loading requirement: intel-mkl/2020.3.304 python3/3.9.2 cuda/11.0.3 cudnn/8.1.1-cuda11 nccl/2.8.4-
cuda11.0 openmpi/4.0.3
$ module list
Currently Loaded Modulefiles:
 1) intel-mkl/2020.3.304   3) cuda/11.0.3           5) nccl/2.8.4-cuda11.0      7) tensorflow/2.4.1
 2) python3/3.9.2          4) cudnn/8.1.1-cuda11    6) openmpi/4.0.3
$ 
$ 
$ 
$ 
$ 
$ echo $InstallDir
/scratch/ab01/xyz123/.local/python
$ pip3 install -v --no-binary :all: --prefix=$InstallDir neural-structured-learning==1.2.0
.
.
.

Successfully installed neural-structured-learning-1.2.0
WARNING: You are using pip version 21.0.1; however, version 21.1.1 is available.
You should consider upgrading via the '/apps/python3/3.9.2/bin/python3.9 -m pip install --upgrade
pip' command.
Removed build tracker: '/scratch/ab01/xyz123/tmp/pip-req-tracker-4a2cicit'
$ 
$ 
$ 
$ 
$ echo $PYTHONPATH
/apps/tensorflow/2.4.1/lib/python3.9/site-packages:/scratch/ab01/xyz123/.local/python/lib/python3.9/
site-packages:
$ python3
Python 3.9.2 (default, Mar 29 2021, 10:41:26)
[GCC 8.3.1 20191121 (Red Hat 8.3.1-5)] on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> import neural_structured_learning as nsl
2021-05-06 14:08:33.613622: I tensorflow/stream_executor/platform/default/dso_loader.cc:49]
Successfully opened dynamic library libcudart.so.11.0
>>> nsl.__file__
'/scratch/ab01/xyz123/.local/python/lib/python3.9/site-packages/neural_structured_learning/
__init__.py'
>>> nsl.__version__
'1.2.0'
>>>
```

## Example III: R randomForest

```
module purge
module load intel-compiler/2019.5.281
module load R/4.0.0
R
```

```
>install.packages("randomForest",repos="https://mirror.aarnet.edu.au/pub/CRAN/")
...
lib = "/apps/R/4.0.0/lib64/R/library" is not writable
Would you like to use a personal library instead? (yes/No/cancel) yes
Would you like to create a personal library '~/.R/x86_64-pc-linux-gnu-library/4.0' to install packages into? (yes/No/cancel) yes
```

```

$ module purge
$ module load intel-compiler/2019.5.281
$ module load R/4.0.0
$ R

R version 4.0.0 (2020-04-24) -- "Arbor Day"
Copyright (C) 2020 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> install.packages("randomForest", repos="https://mirror.aarnet.edu.au/pub/CRAN/")
Warning in install.packages("randomForest", repos = "https://mirror.aarnet.edu.au/pub/CRAN/") :
  'lib = "/apps/R/4.0.0/lib64/R/library"' is not writable
Would you like to use a personal library instead? (yes/No/cancel) yes
Would you like to create a personal library
'~/R/x86_64-pc-linux-gnu-library/4.0'
to install packages into? (yes/No/cancel) yes
trying URL 'https://mirror.aarnet.edu.au/pub/CRAN/src/contrib/randomForest_4.6-14.tar.gz'
Content type 'application/x-gzip' length 80074 bytes (78 KB)
=====
downloaded 78 KB

* installing *source* package ‘randomForest’ ...
** package ‘randomForest’ successfully unpacked and MD5 sums checked
** using staged installation
** libs
icc -I"/apps/R/4.0.0/lib64/R/include" -DNDEBUG -I/usr/local/include -fpic -g -O2
-c classTree.c -o classTree.o
icc -I"/apps/R/4.0.0/lib64/R/include" -DNDEBUG -I/usr/local/include -fpic -g -O2
-c init.c -o init.o
icc -I"/apps/R/4.0.0/lib64/R/include" -DNDEBUG -I/usr/local/include -fpic -g -O2
-c regTree.c -o regTree.o
icc -I"/apps/R/4.0.0/lib64/R/include" -DNDEBUG -I/usr/local/include -fpic -g -O2
-c regrf.c -o regrf.o
icc -I"/apps/R/4.0.0/lib64/R/include" -DNDEBUG -I/usr/local/include -fpic -g -O2
-c rf.c -o rf.o
ifort -fpic -g -c rfsub.f -o rfsub.o
icc -I"/apps/R/4.0.0/lib64/R/include" -DNDEBUG -I/usr/local/include -fpic -g -O2
-c rfutils.c -o rfutils.o
icc -shared -L/apps/R/4.0.0/lib64/R/lib -L/usr/local/lib64 -o randomForest.so
classTree.o init.o regTree.o regrf.o rf.o rfsub.o rfutils.o -lifport -lifcoremt -limf
-lsvml -lm -lipgo -lirc -lpthread -lirc_s -ldl -L/apps/R/4.0.0/lib64/R/lib -lR
installing to /home/123/xyz123/R/x86_64-pc-linux-gnu-library/4.0/00LOCK-randomForest/
00new/randomForest/libs
** R
** data
** inst
** byte-compile and prepare package for lazy loading
** help
*** installing help indices
** building package indices
** testing if installed package can be loaded from temporary location
** checking absolute paths in shared objects and dynamic libraries
** testing if installed package can be loaded from final location
** testing if installed package keeps a record of temporary installation path
* DONE (randomForest)

The downloaded source packages are in
  '/scratch/ab01/xyz123/tmp/RtmpIqJUxr/downloaded_packages'
> library(randomForest)
randomForest 4.6-14
Type rfNews() to see new features/changes/bug fixes.
> library()

Packages in library ‘/home/123/xyz123/R/x86_64-pc-linux-gnu-library/4.0’:
randomForest          Breiman and Cutler's Random Forests for
                        Classification and Regression

Packages in library ‘/apps/R/4.0.0/lib64/R/library’:
base                  The R Base Package
boot                  Bootstrap Functions (Originally by Angelo Canty
                        for S)
class                 Functions for Classification
cluster                "Finding Groups in Data": Cluster Analysis
                        Extended Rousseeuw et al.
.
.
.
parallel               Support for Parallel computation in R

```

## Example IV: R Rtsne

```
module purge
module load intel-compiler/2019.5.281
module load R/4.0.0
vi ~/.R/Makevars
R
```

```
>install.packages("randomForest",repos="https://mirror.aarnet.edu.au/pub/
CRAN/")
```

```

$ module purge
$ module load intel-compiler/2019.5.281
$ module load R/4.0.0
$ 
$ 
$ cat ~/.R/Makevars
CC=gcc
CXX=g++
$ 
$ 
$ R

R version 4.0.0 (2020-04-24) -- "Arbor Day"
Copyright (C) 2020 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)
.
.
.
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> install.packages("Rtsne", repos="https://mirror.aarnet.edu.au/pub/CRAN/")
Installing package into '/home/123/xyz123/R/x86_64-pc-linux-gnu-library/4.0'
(as 'lib' is unspecified)
also installing the dependency 'Rcpp'

trying URL 'https://mirror.aarnet.edu.au/pub/CRAN/src/contrib/Rcpp_1.0.6.tar.gz'
Content type 'application/x-gzip' length 2952876 bytes (2.8 MB)
=====
downloaded 2.8 MB

trying URL 'https://mirror.aarnet.edu.au/pub/CRAN/src/contrib/Rtsne_0.15.tar.gz'
Content type 'application/x-gzip' length 67595 bytes (66 KB)
=====
downloaded 66 KB

* installing *source* package 'Rcpp' ...
** package 'Rcpp' successfully unpacked and MD5 sums checked
** using staged installation
** libs
g++ -I"/apps/R/4.0.0/lib64/R/include" -DNDEBUG -I../inst/include/ -I/usr/local/include -fPIC -g
-O2 -c api.cpp -o api.o
g++ -I"/apps/R/4.0.0/lib64/R/include" -DNDEBUG -I../inst/include/ -I/usr/local/include -fPIC -g
-O2 -c attributes.cpp -o attributes.o
g++ -I"/apps/R/4.0.0/lib64/R/include" -DNDEBUG -I../inst/include/ -I/usr/local/include -fPIC -g
-O2 -c barrier.cpp -o barrier.o
g++ -I"/apps/R/4.0.0/lib64/R/include" -DNDEBUG -I../inst/include/ -I/usr/local/include -fPIC -g
-O2 -c date.cpp -o date.o
g++ -I"/apps/R/4.0.0/lib64/R/include" -DNDEBUG -I../inst/include/ -I/usr/local/include -fPIC -g
-O2 -c module.cpp -o module.o
g++ -I"/apps/R/4.0.0/lib64/R/include" -DNDEBUG -I../inst/include/ -I/usr/local/include -fPIC -g
-O2 -c rcpp_init.cpp -o rcpp_init.o
g++ -shared -L/apps/R/4.0.0/lib64/R/lib -L/usr/local/lib64 -o Rcpp.so api.o attributes.o barrier.o
date.o module.o rcpp_init.o -L/apps/R/4.0.0/lib64/R/lib -lR
installing to /home/123/xyz123/R/x86_64-pc-linux-gnu-library/4.0/00LOCK-Rcpp/00new/Rcpp/libs
** R
** inst
** byte-compile and prepare package for lazy loading
** help
*** installing help indices
** building package indices
** installing vignettes
** testing if installed package can be loaded from temporary location
** checking absolute paths in shared objects and dynamic libraries
** testing if installed package can be loaded from final location
** testing if installed package keeps a record of temporary installation path
* DONE (Rcpp)

* installing *source* package 'Rtsne' ...
** package 'Rtsne' successfully unpacked and MD5 sums checked
** using staged installation
** libs
g++ -I"/apps/R/4.0.0/lib64/R/include" -DNDEBUG -I'/home/123/xyz123/R/x86_64-pc-linux-gnu-library/
4.0/Rcpp/include' -I/usr/local/include -fopenmp -fPIC -g -O2 -c RcppExports.cpp -o RcppExports.o
g++ -I"/apps/R/4.0.0/lib64/R/include" -DNDEBUG -I'/home/123/xyz123/R/x86_64-pc-linux-gnu-library/
4.0/Rcpp/include' -I/usr/local/include -fopenmp -fPIC -g -O2 -c Rtsne.cpp -o Rtsne.o
g++ -I"/apps/R/4.0.0/lib64/R/include" -DNDEBUG -I'/home/123/xyz123/R/x86_64-pc-linux-gnu-library/
4.0/Rcpp/include' -I/usr/local/include -fopenmp -fPIC -g -O2 -c normalize_input.cpp -o
normalize_input.o
g++ -I"/apps/R/4.0.0/lib64/R/include" -DNDEBUG -I'/home/123/xyz123/R/x86_64-pc-linux-gnu-library/
4.0/Rcpp/include' -I/usr/local/include -fopenmp -fPIC -g -O2 -c sptree.cpp -o sptree.o
g++ -I"/apps/R/4.0.0/lib64/R/include" -DNDEBUG -I'/home/123/xyz123/R/x86_64-pc-linux-gnu-library/
4.0/Rcpp/include' -I/usr/local/include -fopenmp -fPIC -g -O2 -c tsne.cpp -o tsne.o
g++ -shared -L/apps/R/4.0.0/lib64/R/lib -L/usr/local/lib64 -o Rtsne.so RcppExports.o Rtsne.o
normalize_input.o sptree.o tsne.o -L/apps/intel-ct/2019.5.281/mkl/lib/intel64 -lmkl_intel_lp64
-lmkl_intel_thread -lmkl_core -liomp5 -lpthread -lifport -lifcoremt -limf -lsvml -lm -lipgo -lirc
-lpthread -lirc_s -ldl -fopenmp -L/apps/R/4.0.0/lib64/R/lib -lR
installing to /home/123/xyz123/R/x86_64-pc-linux-gnu-library/4.0/00LOCK-Rtsne/00new/Rtsne/libs
** R
** inst
** byte-compile and prepare package for lazy loading
** help
*** installing help indices
** building package indices
** testing if installed package can be loaded from temporary location
** checking absolute paths in shared objects and dynamic libraries
** testing if installed package can be loaded from final location
** testing if installed package keeps a record of temporary installation path
* DONE (Rtsne)

The downloaded source packages are in
  '/scratch/ab01/xyz123/tmp/RtmpfWBjN7/downloaded_packages'

> library()

Packages in library '/home/123/xyz123/R/x86_64-pc-linux-gnu-library/4.0':

randomForest      Breiman and Cutler's Random Forests for
                  Classification and Regression
Rcpp              Seamless R and C++ Integration
Rtsne             T-Distributed Stochastic Neighbor Embedding
                  using a Barnes-Hut Implementation

Packages in library '/apps/R/4.0.0/lib64/R/library':

base           The R Base Package
.
.
.
> quit()
Save workspace image? [y/n/c]: n
$ 
$ 
$ vi ~/.R/Makevars
$ cat !$
cat ~/.R/Makevars
#CC=gcc
#CXX=g++

```

## Example V: Julia NLopt

```
module purge
module load julia/1.5.3
mkdir -p /g/data/$PROJECT/.julia
export JULIA_DEPOT_PATH=/g/data/$PROJECT/.julia
julia

> ]
(@v1.5) pkg> add NLopt
```



Showcase

# Monitoring Job Efficiency

## Resource Utilisation Rate

```
nqstat_anu <jobID1> <jobID2> ...
```

				%CPU	WallTime	Time Lim	RSS	mem	memlim	cpus
12345678	R	abc123	x11 myTest	33	10:53:56	20:00:00	58.7GB	58.7GB	200GB	96
19145286	R	abc123	x11 atmos_ma	96	01:32:41	03:30:00	369GB	369GB	2625GB	768
19149497	R	abc123	x11 coupled.	84	00:34:25	04:30:00	320GB	320GB	1440GB	720
19149708	R	abc123	x11 netcdf_c	71	00:36:30	02:00:00	12.0GB	12.0GB	12.0GB	1
19150248	R	abc123	x11 atmos_ma	86	00:22:27	03:30:00	345GB	345GB	2625GB	768

## If under use, look into the job

```
qcat -e <jobID>
```

```
qps -Lopid,nlwp,lwp,stat,sgi_p,pcpu,cputime,comm <jobID>
```

- `qcat` : print the job's standard streams
  - -e/-o for standard error/out stream
  - -s for submission script
- `qps` : take a snapshot of the current processes in the job
  - launches a `ps` query on each node hosting the job
  - accepts most flags `ps` would take

```
$ nqstat_anu 12345678
                                         %CPU  WallTime  Time Lim      RSS      mem  memlim  cpus
12345678 R abc123 x11 myTest          0 10:59:39 48:00:00  228MB  228MB 180GB    96
```

```
$ qps -Lopid,nlwp,lwp,stat,sgi_p,pcpu,cputime,comm 12345678
```

Node 0 (gadi-cpu-clx-2962):

PID	NLWP	LWP	STAT	P	%CPU	TIME	COMMAND
232633	1	232633	Ss	*	0.0	00:00:00	bash
232641	1	232641	S	*	0.0	00:00:00	pbs_demux
232682	1	232682	S	*	0.0	00:00:00	12345678.gadi-p
232697	4	232697	S1	*	0.0	00:00:00	mpirun
232697	4	232702	S1	*	0.0	00:00:00	mpirun
232697	4	232703	S1	*	0.0	00:00:00	mpirun
232697	4	232704	S1	*	0.0	00:00:00	mpirun

Node 1 (gadi-cpu-clx-2971):

PID	NLWP	LWP	STAT	P	%CPU	TIME	COMMAND
884675	3	884675	Ssl	*	0.0	00:00:00	orted
884675	3	884687	Ssl	*	0.0	00:00:00	orted
884675	3	884688	Ssl	*	0.0	00:00:00	orted

```
$ qcat -e 12345678
```

```
./myTest: error while loading shared libraries: libmkl_intel_lp64.so: cannot open
shared object file: No such file or directory
./myTest: error while loading shared libraries: libmkl_intel_lp64.so: cannot open
shared object file: No such file or directory
./myTest: error while loading shared libraries: libmkl_intel_lp64.so: cannot open
shared object file: No such file or directory
....
```

```
$ qcat -s 12345678
```

```
#!/bin/bash
#PBS -N myTest
#PBS -P c25
#PBS -q normal
#PBS -l walltime=48:00:00
#PBS -l mem=180GB
#PBS -l ncpus=96
#PBS -l storage=gdata/c25
#PBS -l wd

module load openmpi/4.0.1
mpirun -np $PBS_NCPUS ./myTest
```

```
$ nqstat_anu 12345679
```

				%CPU	WallTime	Time Lim	RSS	mem	memlim	cpus	
12345679	R	abc123	x11	myTest	33	10:53:56	48:00:00	58.7GB	58.7GB	180GB	96

```
$ myqps 12345679
```

```
qps
```

```
Node 0 (gadi-cpu-clx-1957):
```

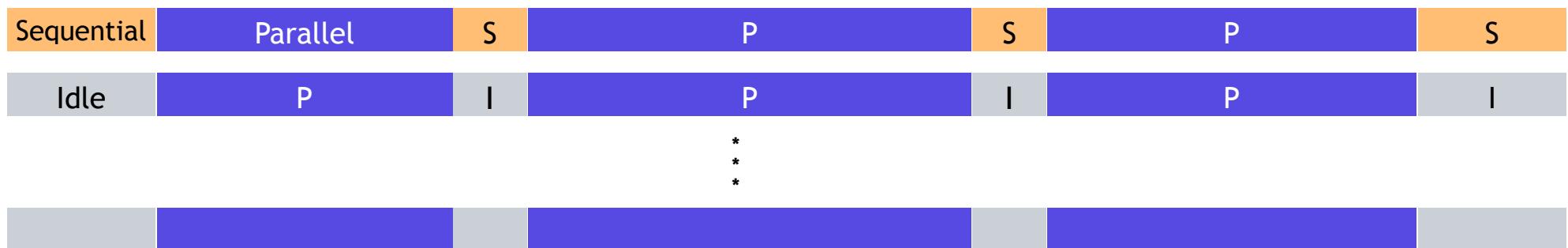
PID	NLWP	LWP	STAT	P	%CPU	TIME	COMMAND
186877	1	186877	Ss	*	0.0	00:00:00	bash
186885	1	186885	S	*	0.0	00:00:00	pbs_demux
186926	1	186926	S	*	0.0	00:00:00	12345679.gadi-p
186961	4	186961	S1	*	0.0	00:00:00	mpirun
186961	4	186966	S1	*	0.0	00:00:00	mpirun
186961	4	186967	S1	*	0.0	00:00:00	mpirun
186961	4	186968	S1	*	0.0	00:00:00	mpirun
186969	51	186969	R1	47	90.1	09:56:32	myTest
186969	51	186972	S1	*	0.0	00:00:00	myTest
186969	51	186975	S1	*	0.0	00:00:00	myTest
186969	51	186978	S1	*	0.1	00:00:50	myTest
186969	51	186980	S1	*	32.6	03:36:20	myTest
186969	51	186981	S1	*	32.1	03:32:56	myTest
186969	51	186982	S1	*	32.9	03:38:02	myTest
186969	51	186983	S1	*	33.1	03:39:17	myTest
186969	51	186984	S1	*	33.3	03:40:34	myTest
186969	51	186985	S1	*	32.7	03:37:01	myTest
186969	51	186986	S1	*	32.9	03:38:07	myTest
186969	51	186987	S1	*	33.2	03:39:55	myTest
186969	51	186988	S1	*	32.2	03:33:36	myTest
186969	51	186989	S1	*	32.1	03:33:06	myTest
186969	51	186990	S1	*	32.8	03:37:37	myTest
186969	51	186991	S1	*	33.3	03:40:34	myTest
186969	51	186992	S1	*	32.8	03:37:33	myTest
186969	51	186993	S1	*	33.0	03:38:57	myTest
186969	51	186994	S1	*	32.8	03:37:06	myTest
186969	51	186995	S1	*	32.7	03:36:59	myTest
186969	51	186996	S1	*	32.5	03:35:26	myTest
186969	51	186997	S1	*	32.8	03:37:36	myTest
186969	51	186998	S1	*	32.7	03:36:36	myTest
186969	51	186999	S1	*	32.4	03:34:42	myTest
186969	51	187000	S1	*	31.6	03:29:19	myTest
186969	51	187001	S1	*	32.6	03:36:03	myTest
186969	51	187002	S1	*	32.3	03:33:55	myTest
186969	51	187003	S1	*	33.1	03:39:09	myTest
186969	51	187004	S1	*	32.7	03:36:36	myTest
186969	51	187005	S1	*	33.0	03:38:30	myTest
186969	51	187006	S1	*	33.0	03:38:28	myTest
186969	51	187007	S1	*	33.2	03:39:45	myTest
186969	51	187008	S1	*	33.2	03:39:59	myTest
186969	51	187009	S1	*	32.3	03:34:23	myTest
186969	51	187010	S1	*	33.8	03:44:10	myTest
186969	51	187011	S1	*	32.8	03:37:39	myTest
186969	51	187012	S1	*	33.3	03:40:27	myTest
186969	51	187013	S1	*	32.5	03:35:44	myTest
186969	51	187014	S1	*	32.8	03:37:40	myTest
186969	51	187015	S1	*	32.9	03:38:15	myTest
186969	51	187016	S1	*	33.3	03:40:45	myTest
186969	51	187017	S1	*	33.1	03:39:29	myTest
186969	51	187018	S1	*	32.5	03:35:23	myTest
186969	51	187019	S1	*	33.0	03:38:32	myTest
186969	51	187020	S1	*	32.3	03:34:22	myTest
186969	51	187021	S1	*	32.4	03:34:33	myTest
186969	51	187022	S1	*	32.7	03:36:36	myTest
186969	51	187023	S1	*	33.4	03:41:14	myTest
186969	51	187024	S1	*	33.2	03:40:24	myTest
186969	51	187025	S1	*	32.1	03:32:52	myTest
186969	51	187026	S1	*	33.5	03:42:04	myTest

```
Node 1 (gadi-cpu-clx-1975):
```

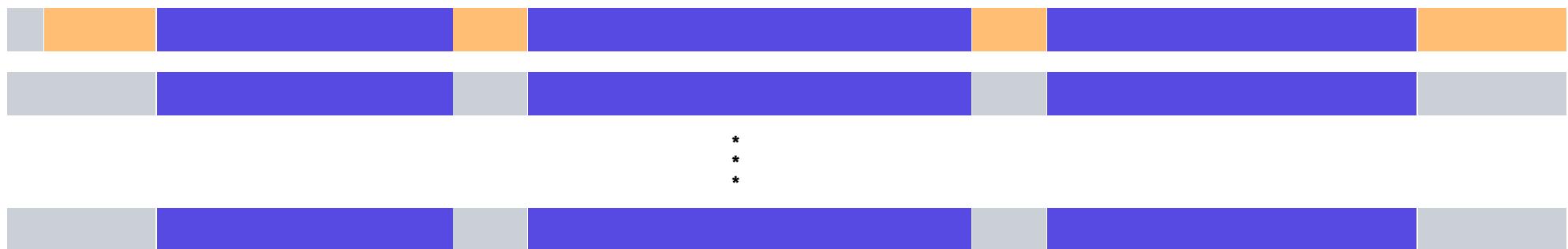
PID	NLWP	LWP	STAT	P	%CPU	TIME	COMMAND
174219	3	174219	Ssl	*	0.0	00:00:00	orted
174219	3	174231	Ssl	*	0.0	00:00:00	orted
174219	3	174232	Ssl	*	0.0	00:00:00	orted
174233	51	174233	R1	3	88.7	09:47:10	myTest
174233	51	174236	S1	*	0.0	00:00:00	myTest
174233	51	174239	S1	*	0.0	00:00:00	myTest
174233	51	174242	S1	*	0.1	00:00:51	myTest
174233	51	174244	S1	*	31.5	03:28:34	myTest
174233	51	174245	S1	*	31.2	03:26:32	myTest
174233	51	174246	S1	*	31.0	03:25:46	myTest
174233	51	174247	S1	*	31.5	03:28:41	myTest
174233	51	174248	S1	*	31.7	03:30:12	myTest
174233	51	174249	S1	*	31.3	03:27:16	myTest
174233	51	174250	S1	*	30.9	03:24:32	myTest
174233	51	174251	S1	*	31.1	03:26:24	myTest
174233	51	174252	S1	*	31.4	03:27:57	myTest
174233	51	174253	S1	*	31.1	03:25:51	myTest
174233	51	174254	S1	*	30.7	03:23:32	myTest
174233	51	174255	S1	*	31.6	03:29:31	myTest
174233	51	174256	S1	*	31.6	03:29:31	myTest
174233	51	174257	S1	*	31.3	03:27:34	myTest
174233	51	174258	S1	*	31.6	03:29:41	myTest
174233	51	174259	S1	*	31.5	03:28:43	myTest
174233	51	174260	S1	*	31.3	03:27:15	myTest
174233	51	174261	S1	*	31.5	03:28:31	myTest
174233	51	174262	S1	*	31.4	03:27:50	myTest
174233	51	174263	S1	*	31.7	03:30:09	myTest
174233	51	174264	S1	*	31.7	03:29:51	myTest
174233	51	174265	S1	*	31.4	03:28:21	myTest
174233	51	174266	S1	*	31.4	03:28:06	myTest
174233	51	174267	S1	*	31.4	03:28:16	myTest
174233	51	174268	S1	*	31.3	03:27:45	myTest
174233	51	174269	S1	*	31.4	03:28:10	myTest
174233	51	174270	S1	*	31.6	03:29:09	myTest
174233	51	174271	S1	*	31.6	03:29:38	myTest
174233	51	174272	S1	*	31.2	03:26:37	myTest
174233	51	174273	S1	*	31.5	03:28:45	myTest
174233	51	174274	S1	*	31.6	03:29:41	myTest
174233	51	174275	S1	*	31.0	03:25:45	myTest
174233	51	174276	S1	*	31.0	03:25:25	myTest
174233	51	174277	S1	*	31.0	03:25:37	myTest
174233	51	174278	S1</				

# Join-Fork Paradigm

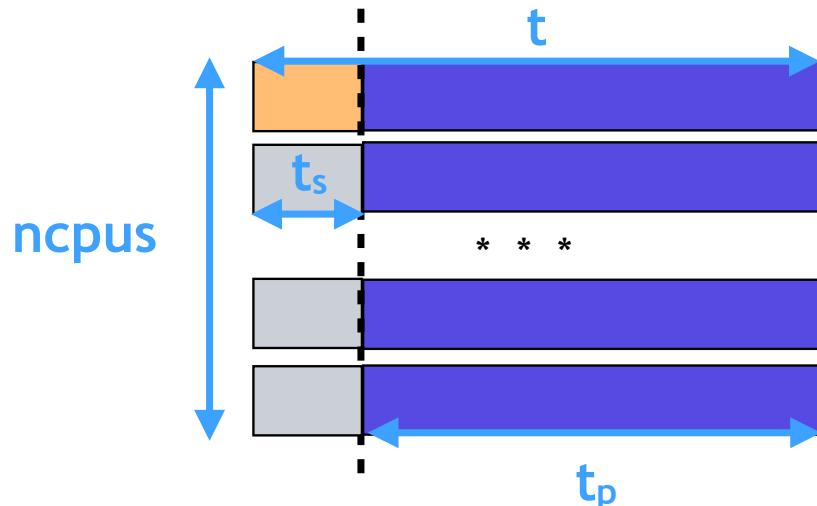
Node 0



Node 1



# Theoretical CPU Utilisation Rate in a Ideal Case

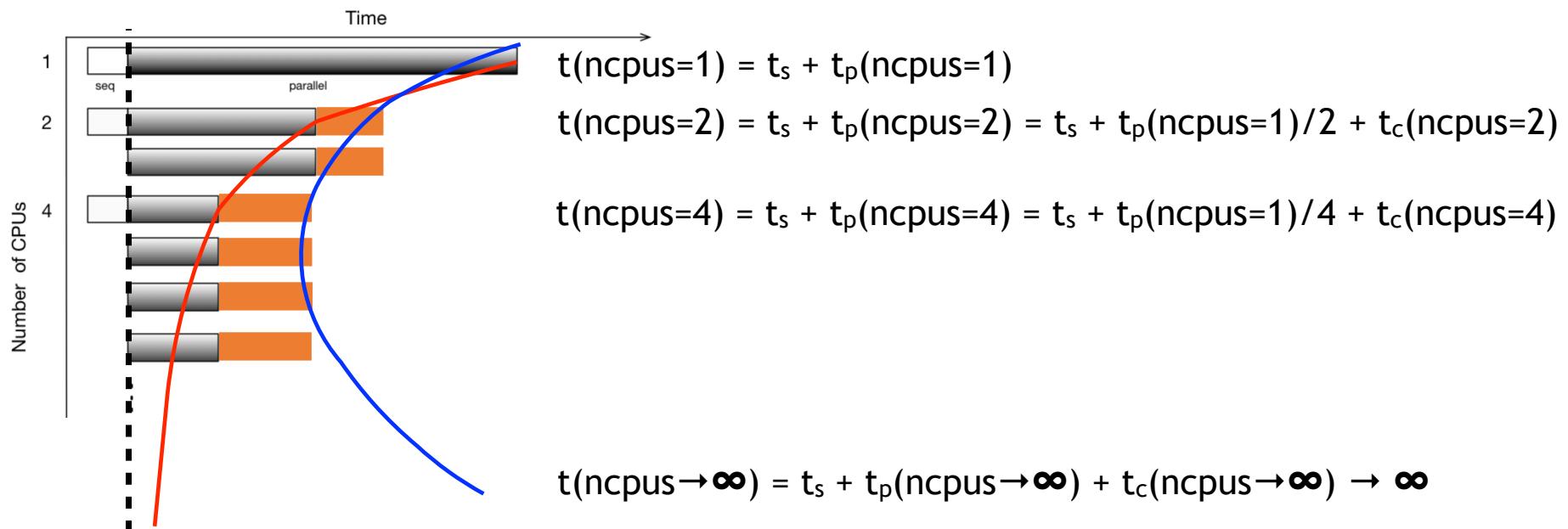


$$\eta(ncpus) = \frac{t_s + ncpus*t_p(ncpus)}{ncpus*t(ncpus)}$$

$\eta(ncpus \rightarrow \infty) \rightarrow 0$   
 given a constant  $ncpus*t_p + t_s$

ncpus	tp	t	$\eta$
1	99	100	1
12	8.25	9.25	0.90
28	3.54	4.54	0.79
48	2.06	3.06	0.68
192	0.52	1.52	0.34

# Execution Time Increases Beyond Sweet Spot



Help Desk

# Asking Questions

## Any More Questions

- Gadi user guide: <https://opus.nci.org.au/display/Help/Gadi+User+Guide>
- User support tickets: [help@nci.org.au](mailto:help@nci.org.au)
- Online support sessions: 1-2pm AEST, first Thursday of each month