UCL Research Computing

Dr Owain Kenway, (@owainkenway)
UCL/ISD/RITS/[Acting] Head of Research Computing
UCL Research Computing

“Supporting UCL’s researchers with not enough money, not enough space, not enough people and not enough time”

Dr Owain Kenway, (@owainkenway)
UCL/ISD/RITS/[Acting] Head of Research Computing
Who we are:

- A team within Research IT Services (RITS), which is within ISD (UCL’s central IT division)
- We (two teams of ~4) look after UCL’s central, and UCL-hosted national services
  - RCI → “The systems team” look after the hardware, OS, are root
  - RCAS → “The user support team” look after users, user applications, are ccspapp
  - **Together** design, deliver + support all the services
The Problem:

- UCL is a world leading research institution :)  
  - Research in almost every field :)  
  - *All institutions* see IT as a cost area (i.e. minimise funding) :( 

_{aka_}

We have limited funding but must support wildly diverse needs.

_(This is not just true of Research Computing but the whole of ISD!)_
The Problem:

- Some users need traditional HPC (massively parallel workloads)
- Some users need HTC (thousands of independent jobs)
- Some users need a mix of the two (High Throughput High Performance Computing?)
- Some users need GPUs
- Some users need terabytes of RAM
- Some users need to use massive amounts of (temporary) storage
- Some users are HPC experts, some are novices
- Some users are developing their own code, some are using centrally installed applications
- Some users...
Challenges

- Geography
  - There is no space for anything anywhere in central London
  - There is no power for anything anywhere in central London

- Money
  - IT is incredibly underfunded sector-wide
  - HPC is funded worse than that (a national problem, not a UCL one)
  - Not enough money for kit or staff

- Insatiable demand for compute from researchers
What we do:

- **UCL only services:**
  - **Grace** → High Performance Computing (HPC)
  - **Myriad**, **Legion** → High Throughput Computing (HTC)
  - **Aristotle** → Interactive teaching Linux service
  - **DSH** → secure data storage and compute (not currently under RC control)

- **National services:**
  - **Thomas** (Tier 2 MMM hub)
  - **Michael** (Faraday Institution)

**COMMON software stack across RC controlled services**

- **Parallel**
  - Single job spans multiple nodes
  - Tightly coupled parallelisation usually in MPI
  - Sensitive to network performance
  - Currently primarily chemistry, physics, engineering

- **High throughput**
  - Lots (tens of thousands) of independent jobs on different data
  - High I/O
  - Currently, primarily biosciences, physics, computer science
  - In the future, digital humanities
Many processes on many processors work simultaneously + communicate between each other.
Many processes, operate independently of each other and in any order.
Grace is UCL’s primary HPC service.
- OCF/Lenovo, QDR IB
- ~11K cores
- 16 cores/node
- 1PB of Lustre storage
Myriad is UCL’s High Throughput/Data intensive service:
- OCF/Lenovo, EDR IB (storage only)
- ~1800 cores
- 4 GPUs
- 2 high memory nodes
- 1 PB of Lustre Storage

Free and paid access models
National Services

- Two national services,
  - both for specialised research
  - both “High Throughput High Performance Computing”
    - (i.e. arrays of small parallel jobs)

Slough:
20 blocks with 3:1 OmniPath connectivity, 485TB shared lustre file system

Bloomsbury:
9 blocks with 3:1 OPA, 200TB shared lustre file system

Each block has:
36 nodes with 1:1 OmniPath connectivity

Each node has:
24 cores
128GB RAM
National Services

- Thomas (the Tier 2 MMM Hub)
  - Built off of Grace (OPA)
  - 18,000 cores
  - £4M of EPSRC funding
  - Running costs paid for by partner institutions:
    - Imperial College London, University of Kent, Kings College London, Oxford University, Queen Mary University of London, Queen’s University Belfast, University of Southampton and UCL

- High Throughput High Performance Computing!
- Running for almost two years.
  - >600,000 user jobs to date
  - >2.57x10^8 CPU hours used by user jobs
    - >29,300 years on one core
    - Upper Palaeolithic start!
National Services

- Michael (the Faraday Institution machine)
  - Built off of Grace (OPA) in summer last year
  - 7,000 cores
  - £1M of external funding
Common software stack

• Deployed across all our resources (inc Thomas + Michael)
  – ~750 user applications + development tools, presented through environment modules
  – Scripts + data from one machine can be run “seamlessly” on another
  – Same interface presented to users
  – AUTOMATED
Common software stack

- This is **not** a stack “just for traditional HPC users” (Fortran/C/C++)
- Supports Python (Cpython, Anaconda, PyPy), R, Julia, Perl (+ Bioperl), Java, Clojure, Common Lisp, Scheme, Mono (.Net), Lua, Go, Racket, Ruby, JavaScript, Matlab…
- ML tools like Tensorflow (GPU, MKL variants), Caffe, OpenCV…
- Allow departmental sysadmins access to install specialist applications centrally!
Common software stack

- Most of the code that builds/runs stuff on our clusters is in Github e.g.
  - Build scripts: [https://github.com/UCL-RITS/rcps-buildscripts](https://github.com/UCL-RITS/rcps-buildscripts)
  - GERun: [https://github.com/UCL/GERun](https://github.com/UCL/GERun)
  - ... and others

METHOD NOT APPLICATION!
- Works on our clusters but maybe no-where else!
- (relies on /shared/ucl/apps existing, not well documented etc.)

Open an issue to ask for a package to be installed.
We install software...

- From the easy:
  
  $ pip3 install numpy

- To the hard:
  - Multiple incompatible dependencies
  - Bazel
  - MPI/Cuda builds...
  - For more examples, see Kenneth Hoste's excellent FOSDEM talk “How to Make Package Managers Cry” on Youtube
    https://www.youtube.com/watch?v=NSemlYagjIU
We answer user tickets...

- E-mail rc-support@ucl.ac.uk for help and advice (not just for our services!)
- Manned by the RCAS team on a rota
Future developments: HPC

- Grace is now more than three years old!
- Time to design and procure a replacement!
- Design project in 2017/18
- Procurement in 2018/19
Almost all time consumed is <2048 cores
Future developments: HPC

Parallel file-systems

scratch
/home

Login nodes

Ubiquitous high speed interconnect

1:1 ~2000
core island
(cpu only)

1:1 ~2000
core island
(cpu only)

1:1 ~2000
core island
(cpu only)

1:1 ???
island

GPU/alt arch

1:1 ???
island

GPU/alt arch

RDS

high speed (40 gbs) link

"Grace 18" service straw man
Dr Owain Kenway, 1st Aug 2018
Future developments: HPC

Biggest application usage is VASP, GROMACS, CP2K
Future developments: HPC

- Based on application usage we have started a Github repository with standard benchmarks for this and future procurements.
- Will grow as we design benchmarks for other systems.
- Based on Archer+TeraGrid benchmarks:

- Vendors use benchmarks to prepare their bids, we use them to do acceptance tests.

DEFINITELY A LEARNING EXPERIENCE!
Future developments: HPC

- Project to procure Grace’s as yet unnamed replacement under way
  - Pilot machine mid-late 2019
  - Looks more like Thomas than Grace
  - Have many bids from different vendors currently being evaluated
    - (they were doing datacentre surveys today!)
  - Free and paid access models a la Myriad (~2000 core chunks!)
Future developments: User Interface

- **File system access:**
  - Presenting home directories via CIFS → people **not** using Linux will **also** be able to mount their home directories on their desktop!
    (Linux users can already do this with FUSE/SSHFS)
  - Mounting Research Data Storage on login nodes
    (all sorts of exciting authentication challenges!)

- **Usability:**
  - Work with CS to present VDI front-end
  - Other ways of accessing resource?
  - Booking system for training courses on Aristotle
Future developments: Collaborations

- From mid 2019 we will be collaborating with CS in running a centrally funded R&D cluster.
- Collaboration is tentatively called RCNIC (Research Computing & Networking Innovation Centre).
- Early access to technologies for researchers e.g. FPGAs, Arm etc outside of a defined service.
  - Successful technologies will be adopted in future ISD service offerings
- Software collaborations (e.g. Linux VDI, our software stack, service reporting…)
- **Not just CS** – we’re keen to bring in other departments that do HPC (Physics, Engineering etc.).
Thanks!

IT’S TIME FOR PIZZA