Journeys in Research Computing, Genomics and Atmospheric Physics – Challenges, Successes and Lessons Learnt

Dr Man-Suen Chan – Linux IT Officer, Department of Physics, University of Oxford

UNIVERSITY OF OXFORD

Outline

- Case study of Research Computing in two units in a university department environment (Oxford University). Both approximately 100 people.
- MRC Functional Genomics Unit (Department of Physiology, Anatomy and Genetics).
- Atmospheric, Oceanic and Planetary Physics (Department of Physics)
- Organisation
- Science
- People
- Machines
- Storage and Data management Common to both
- General Observations

Research Computing Delivers

How to Deliver Research Computing

Some Observations about

How to Deliver Research Computing

Some caveats

- Both Units are in the same university so things may be very different elsewhere.
- I have a Life Sciences background and worked for a much longer time at the FGU so observations are not symmetrical.
- I do not pretend this is large scale computing in any way.
- This is my point of view

MRC Functional Genomics Unit

- I worked there 2005 – 2015
- Unit of the Medical Research Council (Intramural until 2013 then "University Unit")
- Embedded in the Department of Physiology, Anatomy and Genetics
- Managed by "Committee of Programme Leaders"
- When I was there increased from 40-100 staff and 4-8 groups.
- Approximately half of unit were engaged in computational work
- IT setup of Unit largely independent of host department
- Since then Unit has closed (loss of core funding)

Genomics Projects

- Comparative genomics
- Non-coding transcripts – lncRNAs
- Disease genomics – neuro-degenerative diseases
- Next-generation sequencing – became more important and so data volumes increased significantly!

CGAT

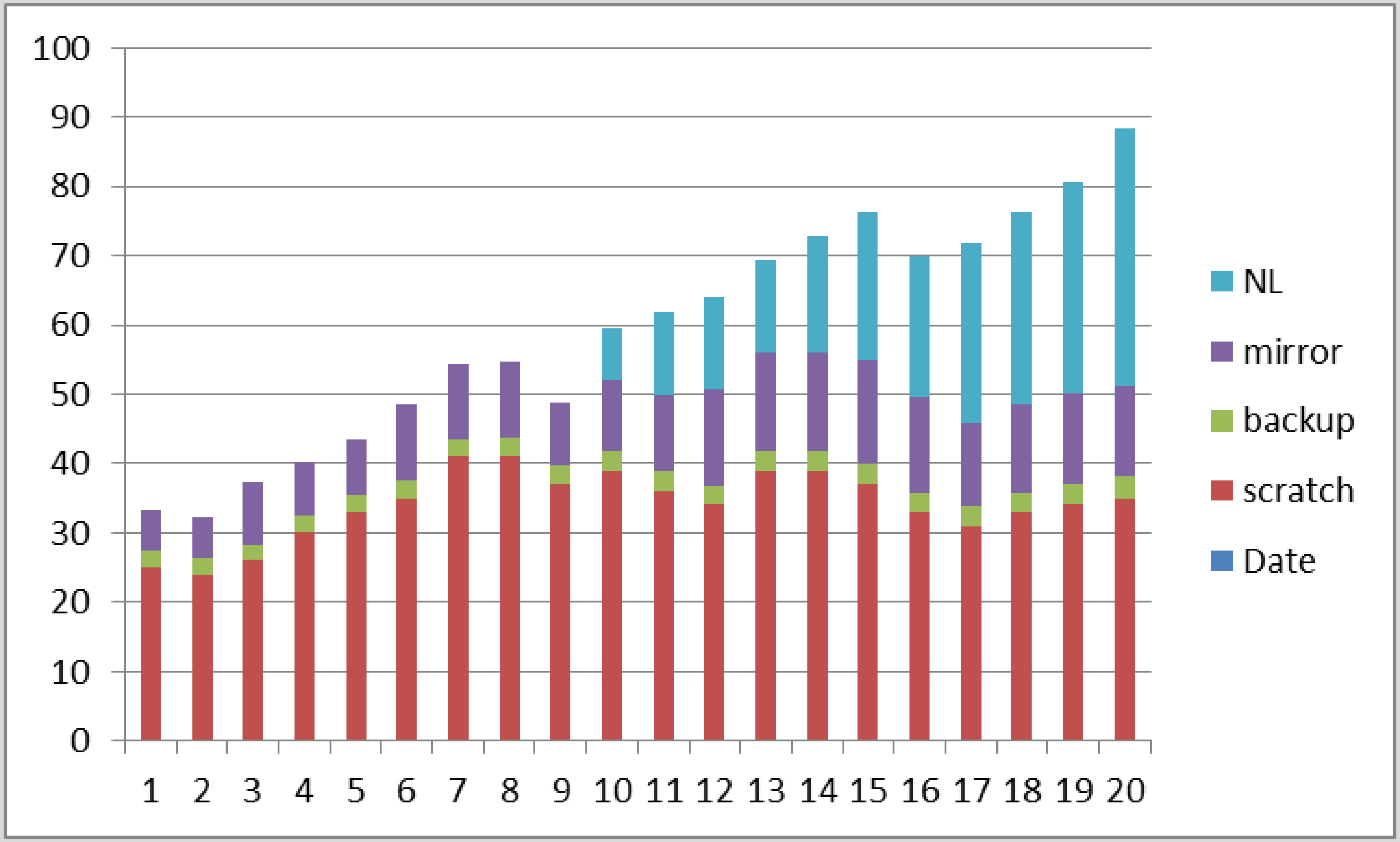
- Computational Genomics Analysis and Training
- Training programme for post-doctoral biologists going into computational genomics
- Taught computational methods and programming
- Collaborative projects with lab based scientists
- Developed a tightly specified project-based approach to data management
- Free pipeline toolkits for repeatable data analysis
- Made possible by the environment of the training programme and may be difficult to replicate in other academic environments. (NB herding cats ..)

Machines

- Over nearly 10 years clearly there was change and evolution was on the lines of development in other sites
- Deployment was relatively space constrained but less so in economical or power availability. Cluster was made up of IBM Blade Center although towards the latter time period 4 in 1 server chassis was used.
- Storage increased from a few TB to 100s TB and evolved from separate RAID boxes, then consolidated into a fibre channel SAN and after to scale-out NAS (Isilon) at each time choosing the most appropriate available technology

UNIVERSITY OF OXFORD

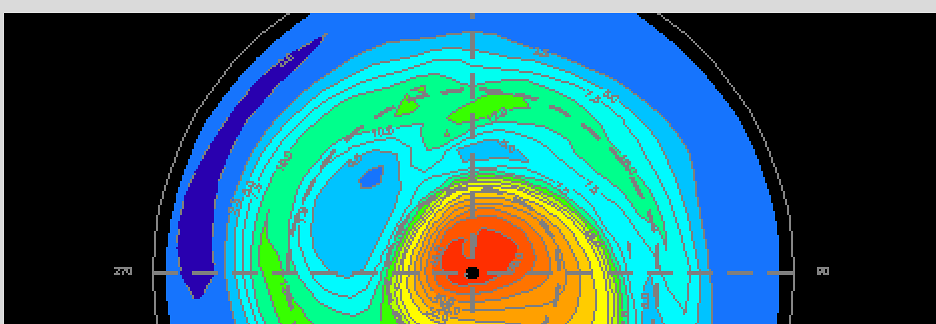
Slide of data growth



Data usage in terabytes on FGU Computational partitions for 20 months to May 2014

Atmospheric Oceanic and Planetary Physics

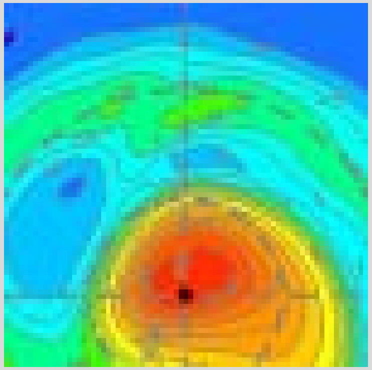
- One of the subdepartments of the University of Oxford Department of Physics, there are six research subdepartments and one admistrative one.
- Physics is a large department >1000 people
- AOPP has approximately 100 people in 17 research groups

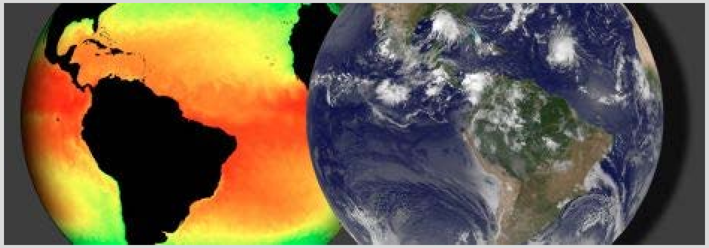


Organisation of Computing in Physics

- Integrated IT team of 17 people. Organised by Subdepartment and OS. I would be AOPP/Central Linux
- Central Physics provides core functionality: network provision (including wireless), Active Directory, printing, desktop and laptop systems (with choice of OS)
- Subdepartments provide Research Computing hardware and support. Research Computing specialists co-located with subdepartment
- Total department resources include 8600 CPU Cores 4.6 Petabytes of storage.
- Linux team provides a standard desktop (ubuntu) which authenticate to AD and configuration managed using Puppet

Atmospheres, Oceans and Planets





Atmospheric Dynamics
Tim Woollings
Climate and Ocean Physics
Laure Zanna
Climate Dynamics
Myles Allen, David Andrews, William Ingram
Climate Processes
Philip Stier
Earth Observation Data Group
Roy Grainger, Anu Dudhia
Ice and Fluid Dynamics
Andrew Wells
Physical Oceanography
David Marshall
Predictability of Weather and Climate
Tim Palmer, Antje Weisheimer
Stratosphere and Climate
Lesley Gray, Scott Osprey

Exoplanets
Raymond Pierrehumbert, Suzanne Aigrain, Patrick Irwin, Caroline Terquem, Niranjan Thatte, Peter Read, Vivien Parmentier
Geophysical Fluid Dynamics
Peter Read
Planetary Science

AOPP Computing Challenge 2015

- Historically AOPP was a separate department and some systems still not integrated, accounts were separate from the Active Directory
- The other subdepartments already had large well organised Research Computing infrastructures in particular Particle Physics
- There were several small infrastructures (9) for group(s) which typically consisted of one nfs file server and four cluster nodes, each with different batch scheduler or none.
- Storage devices were of variable age and brand, some were very small, RAID boxes were not monitored. 645TB available on 21 arrays.
- Software available was generally different for each infrastructure
- No configuration management system
- The division into separate group infrastructures is entirely logical to the academics because of the funding arrangements (but it is not operationally efficient).

Diverse storage systems 2015

Array name (redacted)	Group	Hardware	Capacity	RAID card	monitoring possible
1	CP	Supermicro	82TB	LSI	yes
2	CP	Supermicro	82TB	LSI	yes
3	EODG		82TB	LSI	yes
4	EODG	Supermicro		Areca	no
5	EODG	Supermicro	38TB	Adaptec	yes
6	EODG	HP+nexsan	11+13TB	N/A	yes
7	EODG	Supermicro	11+24TB	Adaptec	yes
8	GFD	Supermicro	11TB	Adaptec	yes
9	hydro	HP+nexsan	24TB	N/A	yes
10	MAD	DNUK	9TB	3ware inc	no
11	ocean	transtec	6TB	LSI Ultra320 SCSI	no
12	Predictability	Supermicro	38TB	Adaptec	yes
13	AOPP	transtec	14TB	Atto tech Ultra scsi320	no
14	MAD	Supermicro	37TB	Adaptec	yes
15	MAD	Supermicro	37TB	Adaptec	yes
16	MAD	Supermicro	40TB	Adaptec	yes
17	plan	HP+nexsan?	35TB	N/A	yes
18	plan	dell	3.5TB	Dell	yes
19	plan	Supermicro	5.TB	3ware inc	no
20	predictability	Dell	37TB	LSI	yes
21	plan		5.5TB	old	no

Out with the old, in with the new

- Leverage the existing available authentication and configuration management systems in Physics, all systems now authenticate with the AD and are managed using Puppet. They are essentially using the same system as the centrally managed desktops. This is also important as the staff allocation for AOPP was reduced from 1.4 to 1.0 FTE.
- One batch system, slurm. New shared queue and all old hardware added to a legacy queue. Priority queues available to those groups who buy extra hardware
- New shared storage based on gluster plus an archive server shared between all of AOPP.
- The shared storage system is a scale out clustered storage but is managed by quota dependent on investments made.
- A “project system” developed to manage the shared storage space accountably and fairly
- Importantly this is underpinned by management agreement, via the sub-departmental computing committee that all future purchases of research computing equipment must be in a shared system and also that all computers are set up in a central physics configuration.

Reusable technologies – shared storage and project system (FGU)

- Shared storage in FGU was based on Isilon which is a scale-out NAS system "ready made" with hardware, software and support.
- Storage was divided into projects. The project system was first used by CGAT but then adopted widely in FGU
- Recall definition of project "a piece of planned work or an activity that is finished over a period of time and intended to achieve a particular purpose" Cambridge English Dictionary.
- Projects had subdirectories backup, data, doc,src and web
- Project had a common naming scheme eg JB001_SMITH_MOUSEGEN
- The contents of the project are defined using an XML file in the Dublin Core format (commonly used for electronic information resources) to allow accountability, parsing and general usability

</metadata>dc="**http://purl.org/dc/elements/1.1/**" xmlns:xsi="**http://www.w3.org/2001/XMLSchema-instance**">
<dc:title>Population transcriptomics of intergenic lncRNAs in the human prefrontal cortex</dc:title>
<dc:creator>Wilfried Haerty</dc:creator>
<dc:subject>transcriptomics</dc:subject>
<dc:subject>lncRNA</dc:subject>
<dc:subject>human</dc:subject>
<dc:subject>prefrontal cortex</dc:subject>
<dc:description>Thousands of long non-coding RNAs (lncRNAs) have previously been identified in the human genome. … </dc:description>
<dc:publisher>unpublished</dc:publisher>
<dc:contributor>MRC FGU</dc:contributor>
<dc:contributor>Lieber Institute for Brain Development</dc:contributor>
<dc:contributor>Yang Li</dc:contributor>
<dc:contributor>Chris Ponting</dc:contributor>
<dc:contributor>Daniel Weinberger</dc:contributor>
<dc:contributor>Thomas Hyde</dc:contributor>
<dc:date>2013-02-20</dc:date>
<dc:type>RNA-Seq</dc:type>
<dc:format>BAM</dc:format>
<dc:format>GTF</dc:format>
<dc:identifier>CP001_HAERTY_LILNCRNA</dc:identifier>
<dc:source>LIBD</dc:source>
<dc:language>English</dc:language>
<dc:relation>CP000_TEST_TEST</dc:relation>
<dc:coverage>human</dc:coverage>
<dc:coverage>schizophrenia</dc:coverage>
<dc:rights>private data, property of the LIBD</dc:rights></metadata>

Reusable technologies – shared storage and project system (AOPP)

- Much less money available to set up! **Why?** Budget 2015 £1000/TB (FGU) → £140/TB (AOPP)
- Therefore must use DIY method, standard servers (Dell Poweredge R730XD), (open source) gluster. However I built it using very similar architectural principles. Although it is somewhat less user friendly to manage it works well and meets our needs.
- We are using the same project system and after promoting it in AOPP (wiki and talk at retreat) it is now well received and accepted. Project quotas are managed using gluster directory quotas.
- We chose not to use the internal directory structure which had been used in FGU. This was felt to be too prescriptive and other service models already exist in central physics, gitlab and group web spaces for example.
- The project naming scheme and XML project definition has been adopted in its entirety
- Good example of using previous solutions and experience to new department in very different subject area.

<?xml version="1.0" encoding="UTF-8"?>
<metadata xmlns:dc="**http://purl.org/dc/elements/1.1/**" xmlns:xsi="**http://www.w3.org/2001/XMLSchema-instance**">
<dc:title>Modelling Jupiter's atmospheric spin-up using the MITgcm</dc:title>
<dc:creator>Roland Young</dc:creator>
<dc:subject>Jupiter</dc:subject>
<dc:subject>GCM</dc:subject
><dc:subject>Moist convection</dc:subject>
<dc:subject>MITgcm</dc:subject>
<dc:description>Simulations using the MITgcm studying jet formation in Jupiter's atmosphere under passive and active cloud conditions (moist convection). Also includes test runs of the Jupiter MITgcm, and analysis of these simulations.</dc:description>
<dc:publisher>Data not yet published</dc:publisher>
<dc:contributor>Roland Young</dc:contributor>
<dc:contributor>Peter Read</dc:contributor>
<dc:date>11 May 2016</dc:date>
<dc:type>GCM output</dc:type>
<dc:format>NetCDF</dc:format>
<dc:format>MITgcm custom binaries</dc:format>
<dc:format>IDL .sav</dc:format>
<dc:identifier>PLR001_YOUNG_JUPSPIN</dc:identifier>
<dc:source>MITgcm</dc:source>
<dc:language>English</dc:language>
<dc:relation>N/A</dc:relation>
<dc:coverage>Jupiter</dc:coverage>
<dc:rights>For internal AOPP use only.</dc:rights>
</metadata>

Reflections. Science and machines

In spite of these units being from very different scientific areas and with different computational approaches this actually had a relatively minor effect on the operations of IT.

- The only significant difference in the design of the infrastructure is the use of Infiniband in AOPP because of the heavy use of MPI.
- Standard “HPC” software components were used such as batch schedulers and environment modules.
- The storage design principles used were the same for both. As with all storage designs it is important to take into account performance and to ensure the system has no bottlenecks

Reflections – organisation and funding

- University departments are collections of research groups doing different projects that are not always working together
- They are not “production facilities” doing large experiments together as a team. Enterprise computing models do not always work.
- However they are also not groups of small businesses – academics are employees of their university. Internet service provider models may not work either.
- The way that the management structure provisions computing is very influential for example through computing committees
- Each group may get its own funding and may be reluctant to share equipment, however for efficiency this may be desirable and it is up to management to encourage this (and for funding bodies to allow it)

Reflections - People

- Scientific Computing needs professional IT support. DIY by researchers is not viable these days,
- Specialist computing support is best provided at a human scale so that the computing specialist is known to the users. This is difficult at scales more than approx 200 people.
- One size does not fit all. Large scale HPC centres are of course necessary but so are local resources (for pre- and post- analysis, prototyping and smaller projects) and personnel. To be effective local IT support need system access to install software and configure systems.
- Use of private clouds may increase efficiency by separating physical and virtual machines.
- Thought needs to be given to formulate fulfilling and stable roles for the IT support personnel. Since the effort required for one research group is usually less than one person it is useful where departments underwrite core permanent posts into which grants can contribute
- Where formulated correctly Research Computing can be a very interesting career and working with scientists can be very rewarding.

Thank you

- Staff and students at the former MRC Functional Genomics Unit. Funding was provided by the MRC.
- Staff and students in AOPP
- Photographs from https://www.physics.ox.ac.uk
- Physics IT team