

How to Use Your Local Supercomputing Centre



Or, “What to do when your users bite off more than they can chew”

What's the point?

- High Performance Computing (HPC) is stuff which “can't be done on desktops, regular servers etc”
 - Getting a years worth of computation done in a week
 - Increasing the level of detail in a computation (scaling in)
 - Increasing the size of problem you can tackle (scaling out)
- It's a problem solver

HPC in the UK

- **National facilities**

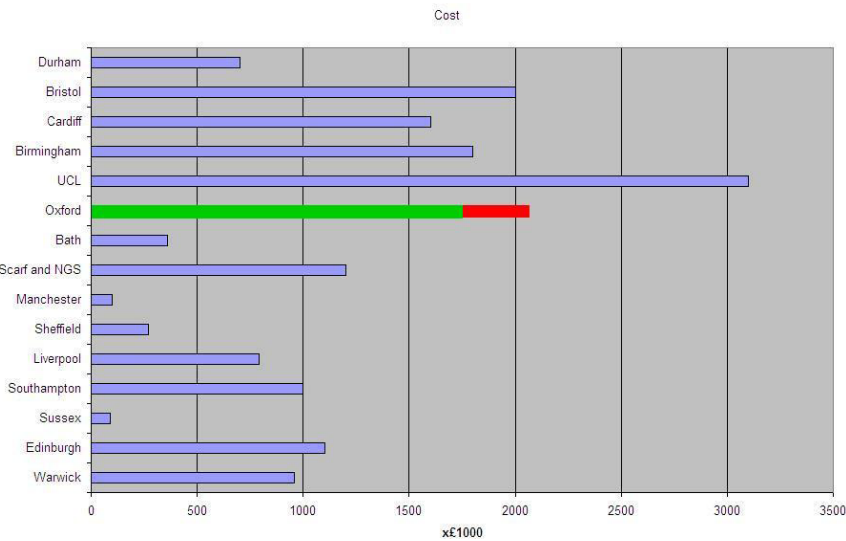
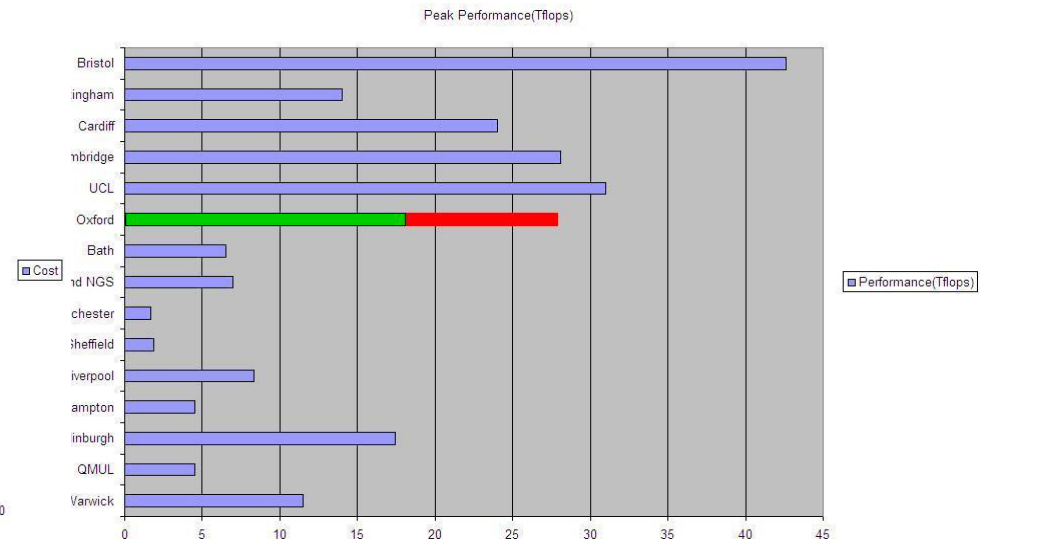
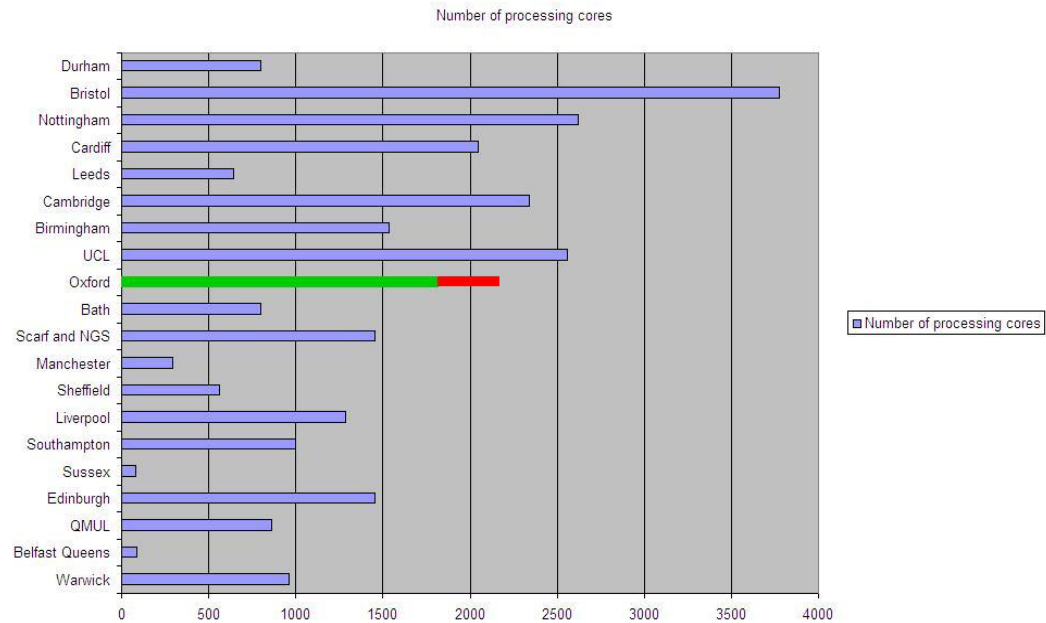
- Hector, HPCx, (Hartree?)

- **Grids**

- NGS, GridPP etc
- Clouds?

- **Campus HPC centres**

- 20ish large centres
- £40M SRIF3 capital
- Ongoing investment
- Services running cost recovery models

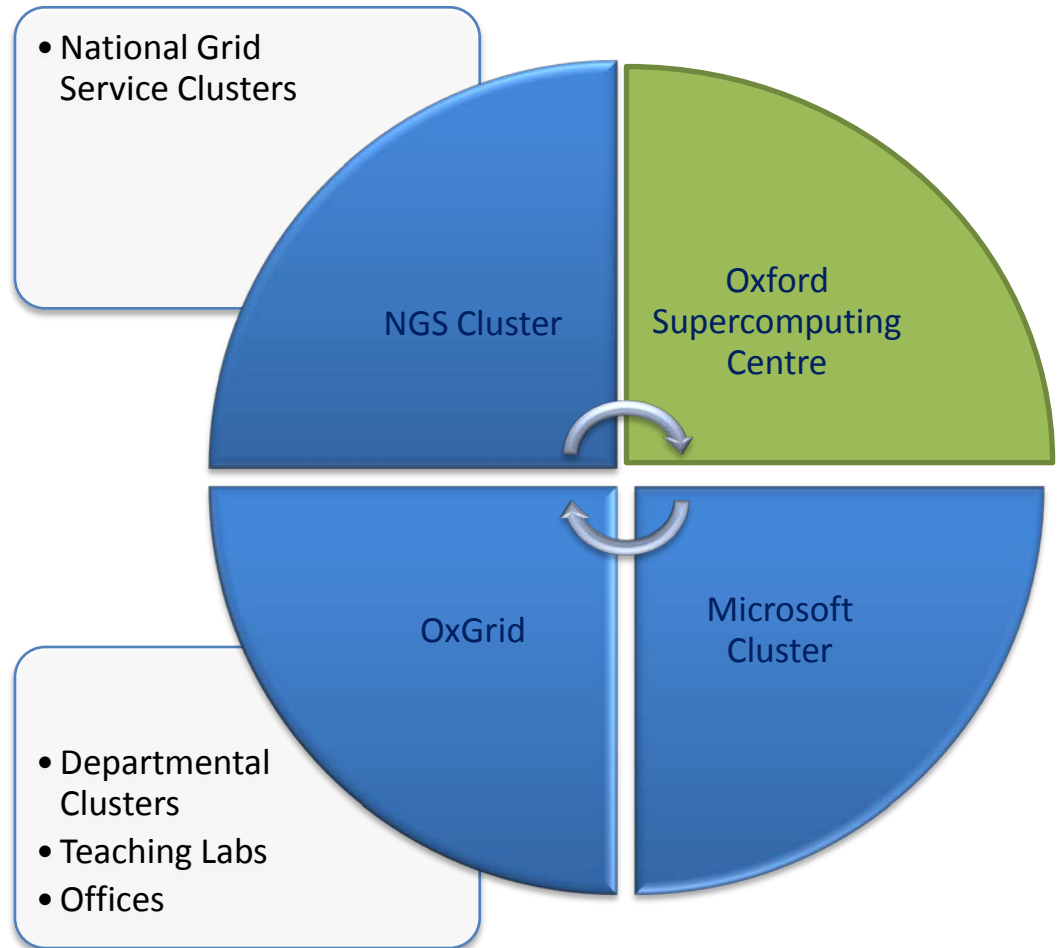


HPC at Oxford e-Research Centre

The Oxford e-Research centre runs a number of HPC services either entirely in-house or in partnership with others:

All of these services are available to Oxford researchers and collaborators.

OSC can be used by undergraduates for teaching.





Confused about what to do?

- There is overlap between our services but as a rule of thumb, OSC's "niche" is one or more of
 - High core count jobs
 - Long duration (weeks)
 - Optimised applications on new hardware
 - Parallel computing
 - Big storage (TBs)
- **Tell us (OeRC) about the computational requirements and we'll try to find the most effective service**





Parallel Computing

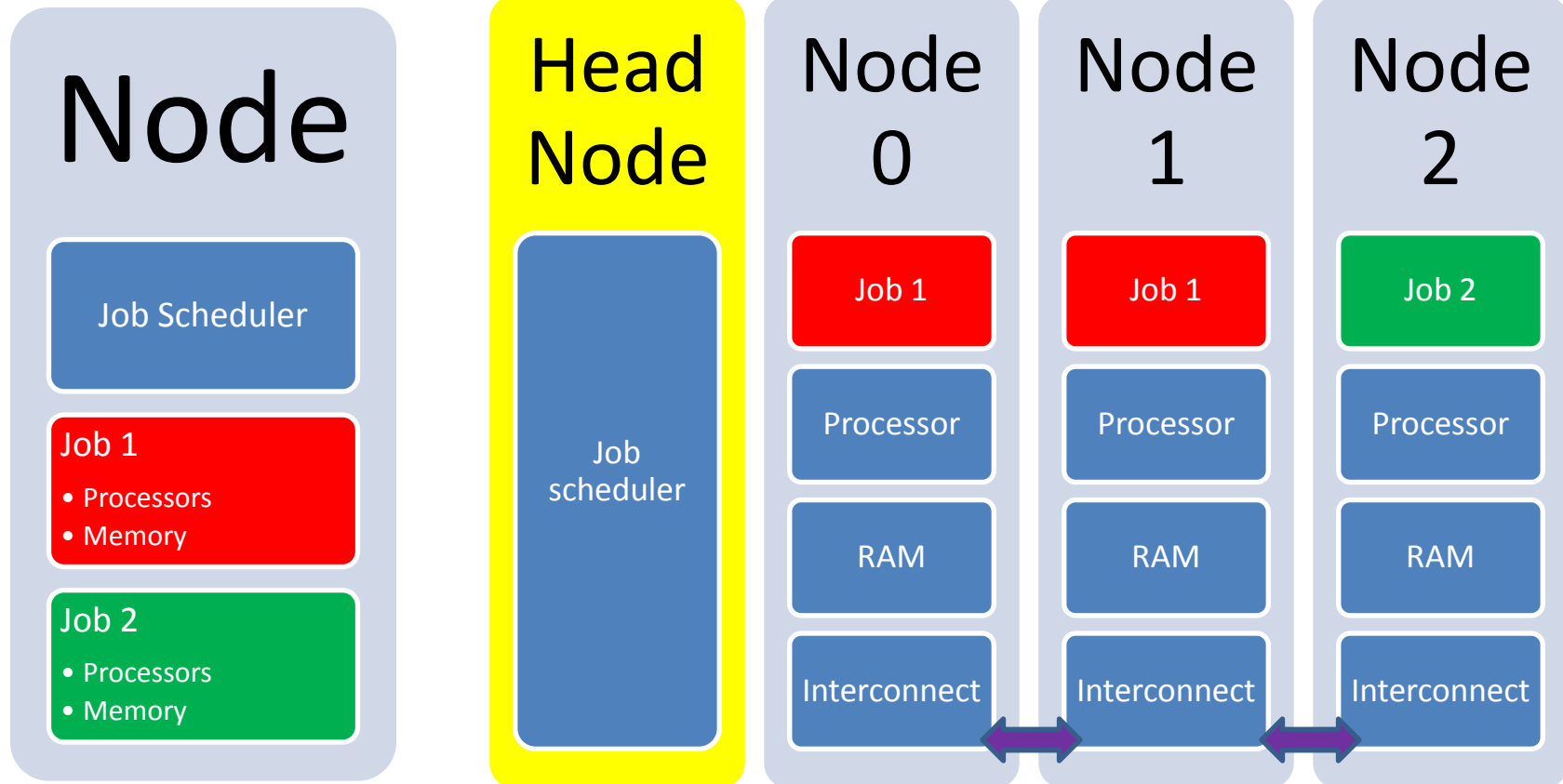
- Most of the work we carry out uses parallel computing
- Split large problems into smaller ones
- Coordination brings the smaller solutions together into one big one
- Ant colonies work rather like a message passing cluster





Parallel Computers

- Shared Memory
- Distributed Memory (Clusters)





OSC Hardware

Production Systems

- 1800 cores in total
- 26TB scratch disks
- 30TB storage disks
- 19 Tflop peak
- **Annual refresh of 1/3 of the centre each year**

R&D Systems

- Nvidia Tesla cluster (“Skynet”)
- Cloud (possibly)
- Apple (given to Biochem)

Future Systems

- MS Cluster may merge into OSC
- NGS service may run on OSC
- Looking at summer procurement of Nehalem cluster and/or Cloud and/or GPU cluster



Orac (Shared memory)

- 256 Itanium2
- 1TB RAM
- 6 TB NAS Storage



RedQueen (Cluster)

- 512 Intel (Dual) Cores
- 0.5TB RAM
- 5 TB Luster Parallel file system
- Infinipath interconnect



Queeg (Cluster)

- 512 AMD (Dual) Cores
- 1TB RAM
- 5 TB Luster
- Infinipath



Hal (Cluster)

- 528 Intel (Dual) Cores
- 1TB RAM
- 10 TB NAS storage
- NumaFlex Interconnect

OSC Statistics

- Groups = 109
- Users ~ 220
- Utilisation=20-50%
- Uptime=99.3%
- 1MW data centre
- PUE ~ 1.5

- Why the low utilisation?
- Users don't like the MRF fEC model of charging



Getting An Account

1. Set up a project with a group leader
 2. Users attach to the new project
 3. Claim 25,000 hours of free time
- OSC has a credit system. 1 credit = 1 core second

Department/Group

Project 1 (Dave)

Project 2 (Alice)

Alice

Bob

Dave

Alice



OSC Services



Jon Lockley
(Manager)



Mihai Duta
(Applications)



Mark Hylton
(Research Facilitator)



Jo Shields
(SysAdmin)



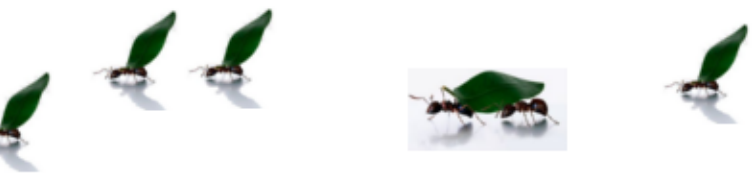
Tolis Sisiaridis
(SysAdmin)

- Systems management
- Project management
- Applications support
 - Abaqus, Abinit, ADF, BLACS, CASTEP, CFD-ACE, CGNS, CP2K, DDT, FFTW, GAMESS-US, Gaussian, Gromacs, Gromos, HDF5, Hypre, PARADIS, ParMetis, PETSc, R, Scalapack, Siesta, Stata...
- Software licensing

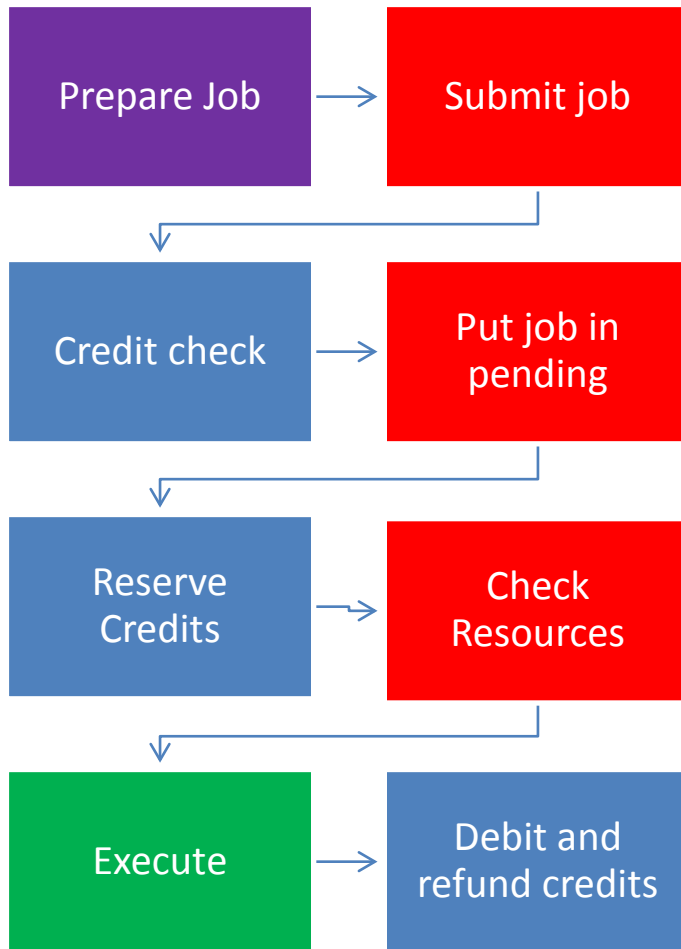


Behind the Scenes

- Jobs are managed with a job scheduler – PBS Pro. Jobs can be run in both batch and interactive mode.
- User accounts etc managed with Open LDAP
- Project credits are managed with GOLD
- HPC on Suse Linux
- Infrastructure on Debian



Submitting a Job



User



Job Scheduler



GOLD credit manager

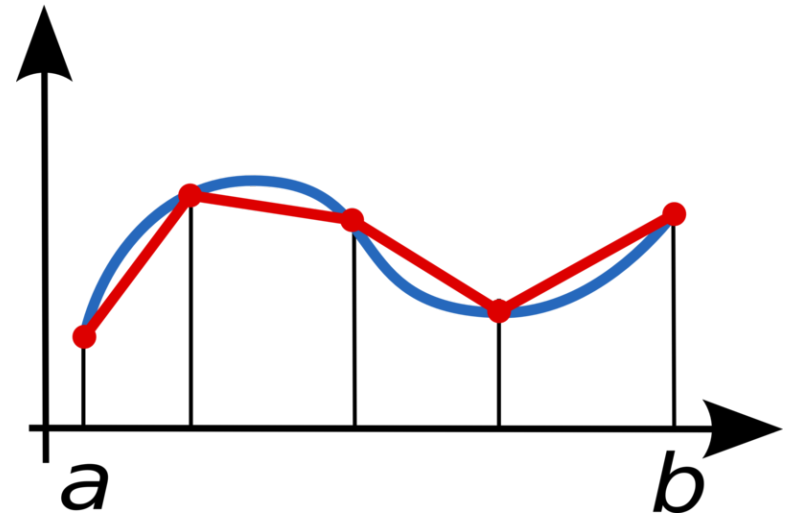


Compute Node

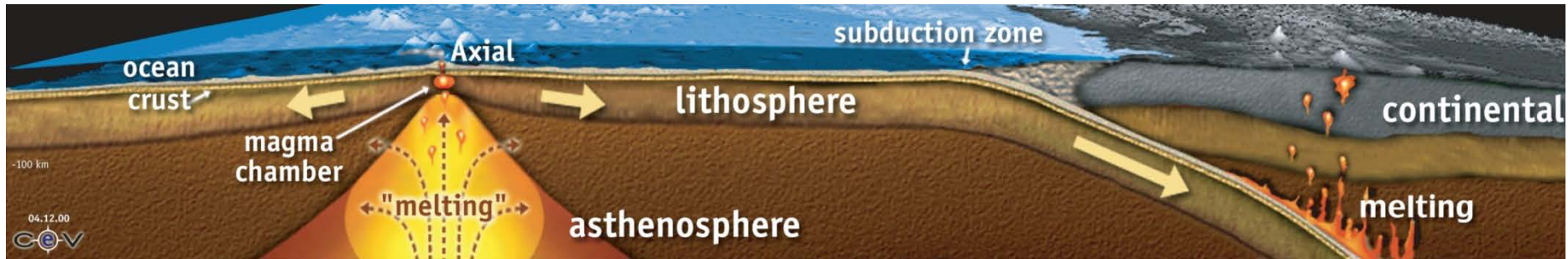


Example

- Let's create and submit a Hello World job on REDQUEEN (a cluster)
- We're going to integrate a simple function (find the area under a curve) using the trapezoidal rule
- We'll split the trapezoids across eight processors and collect the results at the end from each machine



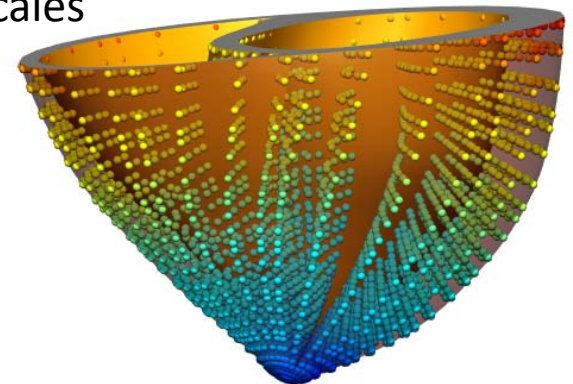
Magma Dynamics (Richard Katz, Earth Sciences)



- Simulation is the only probe available
- Finite volume in 2D or 3D
- Coupled multi-physics
- Non-linear algebraic description
- Uses “HAL” and “QUEEG”

Heart Simulation (Steven Niederer, ComLab)

- Simulation of electrical activity on clinically relevant time scales



Airports

(Georgina Santos, Transport Studies)

- A study of hubbing, slot coordination and the cause of delays
 - 38 million flights
 - 1,300 airlines
 - 444 airports
- Used “Stata” on ORAC



Search & Rescue

(Helen Flynn, ComLab)

- Robots sent in after disaster (eg earthquake) to identify objects
- Victims are identified (as people), locations mapped
- Surrounding objects such as cars mapped to ease location of victim



Textual Analysis

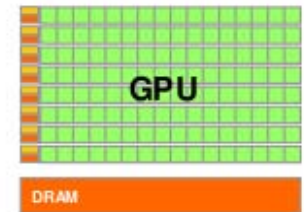
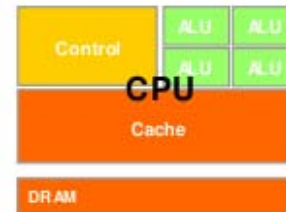
(Scott Moser, Politics)

- Analysis of 200 years of parliamentary speeches as record in Hansard
- Manually reading the text at one page per minute takes 15 years
- Using parallel implementation of “R” on clusters



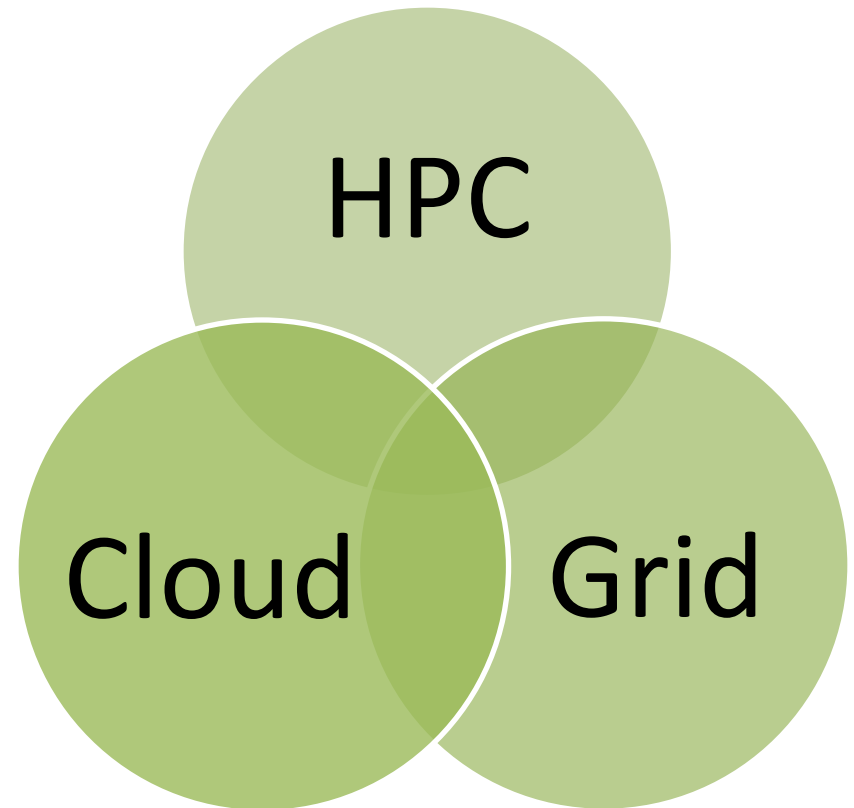
What next?

- OSC has just commissioned a new GPU based cluster – SKYNET
- Access is via regular OSC accounts but there is no charging
- Will users be able to port code to justify a bigger system?



Cloud

- Lots of problems, not many answers
- Probably not an HPC* technology except in some very specific cases
- Cloud has worth in other e-Science areas – talk to OeRC about requirements and ideas for clouds
- OeRC has a prototype Eucalyptus service running – what do we do with it?



*Subject to your definition of HPC



Charging (Booooo!)

- All computing really costs money (sadly)
- OSC has to operate under fEC – we get real electricity bills to pay!
- Current model is as an MRF (looking at SRF)
- We try to support projects without funding as much as possible
- Teaching usage is always free
- 2008/9 rate was 10p per core hour
- 2009/10 rate should be significantly lower



A decorative border of small, dark ant-like icons surrounds the slide. The icons are arranged in a roughly rectangular shape, with a higher density on the right side and top edge.

OSC's next 12 months

- What should we procure next?
 - Storage, probably
 - Cluster, probably
 - Expand our GPU cluster, 50:50
 - Cloud, not so likely
- Firstly we need to sort out this MRF issue...
- Thanks for listening