

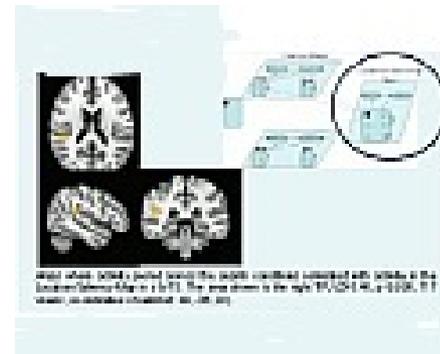
Private Cloud services at UoB

Going back a while

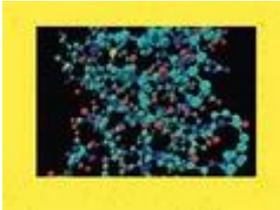
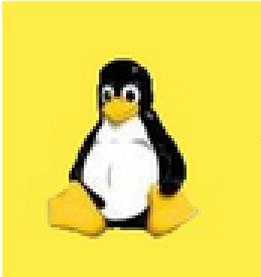
2003 - 5 worker nodes
(HP J6700s, PA-8700
CPU, 8GB RAM, HP-UX),
1 login node and 2 TB
storage



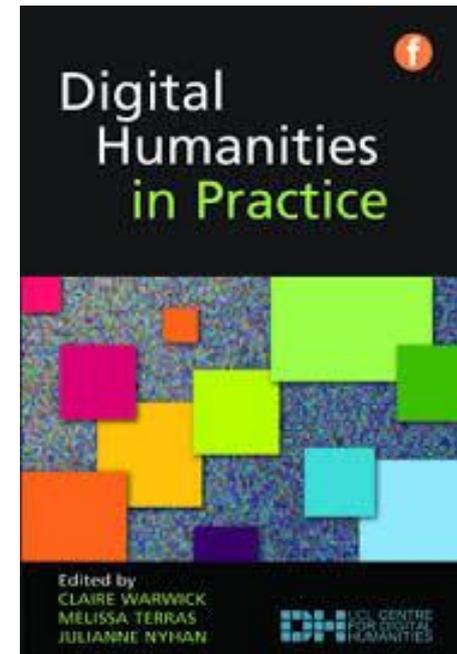
- but Birmingham has never been a leading HPC site in terms of amount of installed kit
- we have focussed on a broad offering (if that's not an oxymoron), attracting many disciplines



our 2003 offering evolved into BEAR



but then some limitations emerged



Centre for Computational Biology



- CCB launched in June 2014 with £2million funding from the University to:
 - promote excellence in Computational Biology, Systems Biology, and Bioinformatics across the range of fundamental and applied sciences, in both the University and allied Health Care arenas
 - federate expertise through world class, independent and collaborative research, as well as teaching to a broad audience ranging from undergraduates to health care professionals
 - provide an environment, both physical and structural, to foster collaborative work and cross-pollination
- recognised the need for compute power
- did not want infrastructure or contract liability; keen to work in conjunction with IT Services and championed central Research IT at high level, including skills requirements such as RSE
- but not all work is suitable for traditional HPC; primarily desktop based
- many workflow applications

We have some experience in this area

CLIMB

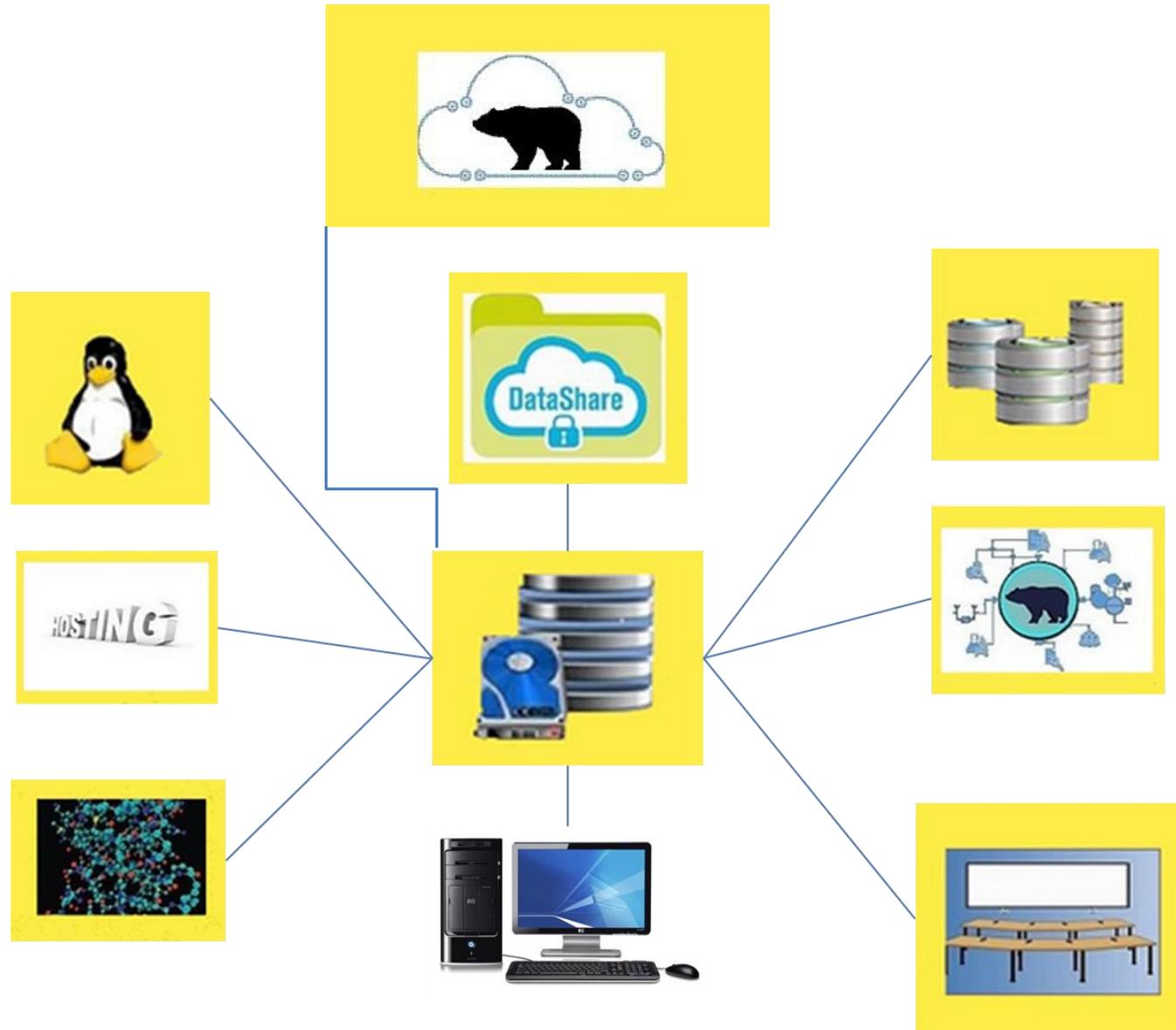


- CLIMB is a world-leading Microbial Bioinformatics Cyber-Infrastructure
- provides free cloud-based compute, storage, and analysis tools for academic microbiologists in the UK
- a collaboration between Warwick, Birmingham, Cardiff and Swansea Universities. Birmingham led the technical implementation.
- OpenStack as the cloud computing platform
- Ceph/GPFS as the storage platform
- 7680 vCPU Cores, 1000 VMs, 500TB local storage

so why not implement something similar locally at Birmingham as part of BEAR?

CaStLeS

Compute and Storage for Life Sciences



CaStLeS background

- Wider remit than just CCB, for example our £8million Phenome Centre, which is studying metabolism in human health and disease to benefit the UK and global population
- resource requests evaluated by a board of 3 academics in Life Sciences
- emphatically **not** a general-purpose virtual desktop; dedicated for research needs
- a range of [standard VMs](#) are offered
- there can be a requirement to run a specific flavour of Linux, especially Ubuntu, which is not available on the BlueBEAR HPC service
- VMs can be supplied with pre-defined software stacks, for example Galaxy
- potentially installing the GVL Launcher as a gateway to a series of rich genomics resources
- potentially making the demonstrator Galaxy instance available to requesting groups as a self-contained/private facility for that group

CaStLeS Technology

- same hardware (2x10 core E5-2640 v4 Broadwell CPUs, 128 GB RAM) as HPC nodes, so can be reprovisioned (not dynamically) in the light of requirements
- OpenStack based, as in CLIMB
- Primarily Linux VMs; Ubuntu is often requested
- Some Windows VMs for research applications that require Windows, such as some of the Illumina assembly toolkits
- VMs have access to the same applications base as the HPC nodes
- VMs have the same access to user storage as the other BEAR services
- will be extended to a wider BEAR Cloud service

Non-technical considerations for CaStLeS

- evaluation of relative merit of requests for VM resources
 - delegated to an academic board
- efficient use of available resources for VMs
 - initially we provisioned persistent VMs, which led to wasted resources on the underlying real hardware
 - this is recognised by the CaStLeS management board as a limitation on the current service
 - we have a summer placement student working on an on-demand provisioning portal
- there has been use of VMs for what is basically a batch compute requirement
 - some batch nodes are being incorporated into the CaStLeS environment
- applications licensing
 - the availability of VMs has resulted in greater demand for applications such as STATA which are not heavily used by traditional HPC users. IT Services only fund a small number of getting started licences - who should pay for the increased demand?

Non-technical considerations for CaStLeS

- single node (shared memory) parallelism only - but that is appropriate for many Life Sciences needs, and distributed-memory parallelism is better suited to an HPC service (usually)
- increased requirement for training, especially on registering for the service, logging on to the VM (for Linux VMs) and use of the storage infrastructure
- persuading researchers that they do not need root access
 - root access is very, very bad for data security in an OpenStack environment
 - when asked why they really need it the usual reply is ‘I always work as root on my own desktop’ without understanding that it’s rarely necessary
 - sudo much safer

Why an on-premises, centrally-funded Cloud service?

- no financial burden or complications on researchers (or us)
- widen existing service to non-traditional Research Computing users
- local support from Research Computing team who have developed and control the cloud environment
- development of existing relationship with researchers - team is well respected on campus
- tightly integrated with existing services, especially central data store
- immediate access to current applications, and/or provision bespoke VMs
- short-term and immediate access by researchers to familiar environment for developing grant proposals
- straightforward continuation of research if end-user workstation fails
- develop support for Research Computing from non-traditional areas

Our University private cloud service

- currently available as a pilot service
- tightly integrated into other BEAR services, with fast access to secure on-premises storage at its heart
- VMs for interactive applications, such as image analysis, which are computationally and data demanding
- VMs for areas such as CFD where convergence of a simulation can be monitored, to avoid wasted compute cycles - limited to shared-memory parallelism for the solver
- delivering a Jupyter Notebook service underpinned by serious computational power
- possible other on-demand services being considered:
 - running a Matlab parallel processing cluster which can be directly accessed from client desktops or a VM
 - enabling users to take snapshots of running systems and share these with other users, which may be of particular interest to researchers publishing papers as they will be able to take a snapshot of the system with all software/configuration to enable reproducible science
 - integration with our metadata management system which is currently under development (iRODS) which will allow users to tag and categorise data.

Lessons Learnt from the CaStLeS service

- the worth of a project has to be decided by an academic committee, not IT Services
- training of, and support for, new, non-traditional HPC users must not be under-estimated
- need clarity on applications licensing - what is supplied centrally and limitations on use
- need a portal for automated provisioning of VMs (virtual machines) 'on demand', with finite resources and lifetime
- users tend to over-state requirements; discuss real needs at point of application
- be clear that this is a research-based offering, not a generic desktop

Summary

- working in conjunction with influential Centres within the University from their establishment has been vital to developing the services
- integrating the centrally-funded private cloud offering into an existing service with storage at its heart has raised the profile of the Research Computing services with disciplines who would not otherwise use our services