

HPHI Oversight Committee Meeting

Easter Term Meeting – 22nd May 2018, 12.00pm – 1.00pm

1. Present

Professor Ed Bullmore (EB), Dr Guy Williams (GW), Professor Bertie Gottgens (BG), Professor Evis Sala (ES), Professor John Suckling (JS), Dr Paul Calleja (PC), Richard Bartlett (RB), Andrew Leedham (AL)

2. Apologies

Ferdia Gallagher (FD), James Brenton (JB), Lydia Drumright (LD) and Ari Ercole (AE)

3. Minutes of the previous meeting

The minutes of the previous meeting on 23rd February were approved

4. Matters arising

None

5. Review of action points

Action points from the previous meeting were reviewed.

Recruitment for the embedded post at WBIC is proving problematic. The HPC team at UIS has recently lost several team members. PC mentioned that HPC are just about to offer a post and there may be a chance that some of this post's time could be spared for a 0.5 FTE embedded post, but this was not clear at the moment as it was primarily needed to fill a vacancy in the RCS team. The funds are with HPC, and it was suggested that it may be worth looking at alternative ways to use them if the post is not filled.

AL and Paul Sumption (RCS) met with Caldas' group, which will be discussed later in the meeting.

ACTION: PC and GW to meet to discuss how best to proceed with recruitment and if any support could be offered with current staffing levels, and, if post is not filled, whether funds could be used for an alternative purpose **to be completed by 5th June 2018.**

6. Review of usage

Clinical Cloud HPC Cluster usage

Clinical Medicine usage was averaging at 10 -15% for core hours, which is low, but usage is likely to increase in the future, and there may be a need to leave some headroom. BG mentioned that a training course, given by Paul Sumption and Simon Flood was taking place today (22nd), it was hoped that this would increase usage by Bertie's group. Serena Nik-Zainal's group's usage is also likely to increase in the future.

ES asked if her group could use this resource. PC offered to give a seminar to give a better understanding of what the system is. It was suggested that, if utilisation of the system remains low, the Committee would consider other groups that could be invited to use it from the Clinical School. It's still not clear where the level of demand is going to be. BG usage is low, but it was felt this would increase.

It was noted that statistics outlining Lustre storage was confusing. The figures for WBIC were included under the 1.4PB assigned to the Genomics side of HPHI whereas they should have been listed separately.

AL informed the Committee that he and Paul Sumption had met with Caldas' group to discuss their use of storage. They were asked if they could produce a pipeline for their research which reduced the load on lustre storage, possibly utilising some tape/cold store. The group offered that they could reduce their requirement for lustre storage to about 250TBs. It was suggested that, to allow for a fairer distribution of storage, Caldas' group could buy additional storage and move some of their data over.



The Committee decided that Caldas' group should not be using such a large portion of the lustre storage and that it would be fairer if they were to reduce this to 250 TB. PC volunteered to talk to the group and encourage them to reduce their usage of free storage possibly by purchasing additional storage.

WBIC

The Committee discussed WBIC and Storage. WBIC usage was around 40% for the period Feb-May 2108. There was a discussion about what constitutes good uptake/usage and felt that 50% usage was deemed good and that higher percentages would impact queue times.

It was noted that the stats showed the number of jobs run rather than the number of users on the system. WBIC currently has 64 users which seems a reasonable amount. These would all be accessing data but not necessarily running jobs.

Data usage was relatively high 421TB out of 693TB. GW suggested his team would enforce policies with regard to data retention to ensure data hygiene. There was also a question whether there was an outreach issue and that if the WBIC role had been recruited, this person would be able to provide advice on HPC and data hygiene.

ACTION: PC to ask Stuart Rankin to look at providing statistics showing activity ie submitting jobs vs accessing data, and to clarify statistics for Lustre Storage use separating out WBIC from Genomics.

ACTION: PC to make arrangements for seminar for Professor Salis' group

ACTION: PC to meet with Professor Caldas' group to discuss reducing his group's use of storage down to 250 TBs, as suggested by the HPHI O/S Committee

7. Nik-Zainal group's HPHI requirements for long term research needs and representation

Professor Nik-Zainal is currently discussing use of HPHI resource with PC and will potentially be a big user in the long term that is expected to increase usage of cores. The group is in transition from Sanger. They will require longer run time queues than other groups. (NB currently given SL3 - 3 days, 7 days was requested)

8. Feedback on ISO process

GW reported good progress had been made so far on the project. ISO board has met several times, but is still scoping at the moment and hopes to have a plan in place by the end of this the summer.

Clarification was requested on whether the ISO certified platform would be hosted on HPHI resource. GW/PC clarified that the HPHI platform would not be certified as a whole and explained that current thinking would involve separate tenancies in OpenStack.

JS requested guidance on the current provision for storage of personal data. RB offered to discuss this with him after the meeting.

9. AOB

None

10. Date of next meeting

TBC