

# **UCL Computational Resource Allocation Group (CRAG)**

## **MEETING MINUTES**

**14<sup>th</sup> December 2012**

### In Attendance:

1. Prof Nik Kaltsoyannis (Chair) - *Molecular Quantum Dynamics and Electronic Structure*
2. Clare Gryce - *Head of Research Computing and Facilitating Services, ISD*
3. Dr Nicholas Achilleos - *Astrophysics and Remote Sensing*
4. Dr Andrew Martin - *Bioinformatics and Computational Biology*
5. Dr Bruno Silva - *Research Computing Platforms Team Leader (Service Lead), ISD*
6. Dr Ben Waugh - *High Energy Physics*
7. Jo Lampard - *Senior Research IT Services Facilitator, ISD*

### Apologies:

1. Prof Andrew Philips - *Epidemiology*
2. Dr Dugan Witherick - *Astrophysics and Remote Sensing*
3. Prof Dario Alfe - *Thomas Young Centre (Materials Science)*

*Note: Minutes below provide a high level summary of decisions taken and actions assigned by the Group.*

### **1. Approval of Minutes of last meeting on 9<sup>th</sup> November 2012**

The Chair (NK) asked for clarification of the following points:

Item 7 – BCS can easily publish the graph with Legion's utilisation information but wished to consult the group on where this would be best located on the Research Computing website. A decision was made to place it in the Legion Cluster section for the time being (with some text explaining what it is). CG pointed out that the entire Research IT Services website will shortly be redesigned.

Item 10 - Use of the Legion@UCL string. CG has placed this information in the section advising users how to acknowledge the use of Legion in publications. NK asked CG to publicise both this, and the new graph of scheduler information, by emailing the legion-users list (see new Action 98).

The Group then approved the Minutes of the November 9<sup>th</sup> 2012 meeting.

## 2. Update on status of current Actions

The list of current Actions (below) was updated, and new Actions arising were added.

## 3. Any requests for additional resources: (a) scratch quota requests, (b) any other request.

### Scratch Quota requests

The following requests were approved:

*None*

### Other requests:

The following requests were approved:

- Martin Hardcastle
- Andrew Logsdail

Regarding MH's application to reserve 256 cores on Legion for two weeks, BCS stated that (due to the outage of the Torrington Place nodes, to be brought back online on 17/12/2012) it would be easy to reserve the requisite number of cores, but stated that it would be beneficial to know the ID's of any jobs already queuing so that they could be given higher priority ahead of the reservation. NA said he would ask MH to provide them.

## 4. Review of any Centre for Innovation (Cfi) access requests (Chair)

*(Doc: Cfi\_Access\_Application\_v0 6)*

The following applications were submitted and approved:

- Ben Ahmady
- Dr Erlend Davidson
- Dr Andrew Logsdail

AM was pleased to hear that AMBER is already installed on Emerald, as he has a PhD student due to start work on a project which will use it. The GPU community are already aware that there is a dedicated GPU developer (Istvan Reguly of the Oxford e-Research Centre) to help with applications; it was agreed that Legion users should also be advised of this (see new Action 99).

## 5. Review of Legion usage stats using new online reporting framework (Bruno Silva)

BCS presented the Legion usage stats for November 2012 (it was noted that viewing them in IE is not recommended). NK expressed concern that the stats for SysBioMed showed huge wait times in relation to run time; he had also noticed one particular user (uccaeak) whose jobs shared this characteristic, and wondered if there could be a common cause. BCS agreed that there could be a correlation, but that it was not simple to determine the cause without looking at the jobs. He also pointed out that all of the jobs shown on the graph did eventually run. The systems team spends a lot of time looking at the scheduler and highlights any apparent issues; the Research Computing team then contacts the user to check if there is a problem with the job submission script, or if it is simply a question of a very large job requiring significant resources. BCS agreed to contact the user in question and would also check December's statistics to see if the pattern was repeated for the same consortium (see new Action 100).

## 6. Cfl usage stats (Clare Gryce)

CG presented the Q2 utilisation statistics for Emerald and IRIDIS, drawing attention to the differences in UCL's and Oxford's usage. UCL only uses the Cfl facilities for its capability work, whereas Oxford has smaller local facilities so does not impose any barrier to entry, hence they have many more users. The different local facilities at each institution means that capability work takes on a different meaning; this should be discussed, but as there is no competition for resources at the moment there is no immediate concern.

**Emerald:** UCL has few active users, in contrast to Bristol – although only 1 or 2 additional UCL projects would change this situation. UCL's usage was around 5% for September - November 2012. The allocation is 25% per month, though this is only a soft limit. A maximum of 8% of both Emerald and IRIDIS are reserved for use by external users and this usage is kept under review. Academic researchers have been advised that they should not rely on the availability of the Cfl resources for long term work as they are intended to be purely for development purposes; eventually, it is envisaged that there will be a charging model for paid work. There is still a degree of under-utilisation but this is shrinking; there is no prospect of getting 90% utilisation with capability work as there is no backfilling of jobs, but 70% should be achievable – in fact, this has already been exceeded. STFC own a 10% share of Emerald, but it was not clear whether this is included in the pie charts; CG to clarify for the next report (see new Action 101). BCS thought that it would be very interesting to know which application domains are using it, and whether they are standard or newly developed applications; this question had already been asked but obtaining that type of information would probably have to be done manually rather than via the job scheduler (PDS). Reporting definitely needs to be improved as the project board will require more information for the next quarter; at the moment it is not possible to analyse jobs even by wait time.

There is a need to raise awareness of GPU; BCS and Istvan Reguly (who is doing hands-on CUDA work) would shortly be putting on a 1-day seminar, hopefully with someone from NVIDIA attending. BCS mentioned that GPU work can be done on a desktop computer provided that it has a decent graphics card; even a £20 graphics card would allow a prospective developer to make a start. CG also emphasised that raising awareness of the available support is essential; GPU programming and the Cfl facilities are briefly covered in the Introduction to Legion course. BCS will circulate a list of the applications which are available on Emerald to the legion-users mailing list and other research computing mailing lists. (see new Action 102).

**IRIDIS:** usage is much higher, though still behind that of Oxford. Inadequate reporting again means that even the job classes cannot be seen, nor is there any data on applications. NK said he would like to see what the job size distribution was, as there needs to be some way of checking to see if it is heavily dominated by jobs using only a few cores. The operations group have agreed to prioritise jobs requiring >128 cores, with smaller jobs starting to backfill. Prioritisation is likely to increase to 512 cores when UCL obtains a larger share of the resource in April (from buying out part of Bristol's share for 2013-14). Each institution has nominated 2 users to carry out capability testing. Richard Catlow has nominated Ben Slater, who will start his work in the New Year. One of Bristol's users has already found that their work does not scale to more than 64 cores! CG felt that enough knowledge of how to support jobs requiring >2,000 cores has now been gained that job scheduling should become routine within the next couple of months; as a consequence, the user's experience should improve considerably.

The same reporting structure will be presented for the next quarter, with CG to obtain additional data on job class distribution (see new Action 103).

## **7. Verbal update on Legion IV project** (*Bruno Silva*)

BCS gave a brief overview of the current status of Legion IV phase 1; the period of testing after the filestore upgrade had now ended, with the new scratch area up and running and showing good performance. No problems with job throughput had been reported and there should be no requirement for further outages to update the nodes. BW enquired whether this would be the case for CPU replacements in phase 2. CG replied that although further on-site expansion had initially been planned before moving the hardware off-site, because of the opportunity to purchase a larger share of IRIDIS for 2013-14, the RIISG had taken the decision to focus on off-site expansion. BW added that hosting of further hardware in the KLB had posed an additional problem. The consequences of this decision were a dip in Legion capability in the interim due to contraction of on-site facilities, balanced by the increased capacity on IRIDIS; IRIDIS users would need to be supported to the same level as those currently on Legion. Legion IV would be terminated at the end of phase 1, after the software (Lustre) stage of the filestore upgrade. Some of the funds

released by abandoning further Legion upgrades would be used to fund the increased IRIDIS share for 2013-14, with the balance to eventually be used for the off-site project. The Legion project manager will start on the design of the off-site hardware project in early 2013; CG emphasised that the off-site servers would still be managed by UCL.

#### **8. Management and use of 7-day queue for Gaussian jobs - discussion (Chair)**

NK asked how many users had access to the 7-day queue; BCS did not have that information but would find out. He said that because usage of the Gaussian 7-day queue had been restricted to jobs using the Research Computing binary, it was much harder for anyone to cheat. BLCR would soon allow users to checkpoint their jobs, which would benefit everyone whilst removing one potential reason for needing to use the 7-day queue. The queue also caused headaches when planning system outages so it would be better to get rid of it if there was no longer any need. NK agreed that the justification for the queue should be re-examined. Analytical frequencies calculations cannot be checkpointed, but could run in the Linda version; using Linda would probably remove the need for the 7-day queue as those jobs would be likely to run in 48 hours. It was agreed that users should be asked to reapply for access to the queue before the February meeting, with the warning that if they had not done their access would then be revoked. BCS to email the legion-users list (see new Action 104). There had recently been a discussion about purchasing MOLPRO and Gaussian licensing on IRIDIS with funding probably being available. Although Gaussian jobs do not scale very well to very large numbers of cores, users needing to run a lot of 128 core jobs would find it difficult to do this on Legion, probably justifying their use of IRIDIS. NK should be involved in any discussion about this so that the right decision can be made.

#### **9. Role of Consortia leaders - discussion (Chair)**

To be discussed at next meeting.

#### **10. Usage of Legion by taught students and master students – brief report (Clare Gryce)**

A list of UG and taught PG users with accounts on Legion (excluding those who also had a staff role) was circulated at the meeting. BW said that role data may not be quite accurate, and also pointed out that taught PGs are still carrying out research projects, this just being a smaller proportion of their work. BCS felt that taught students were being directed to use old, inappropriate machines (Socrates and Plato, which are running AIX) for gaining practical programming experience, which is of limited value. BW felt that it would be appropriate to have a reserved area on Legion for James Hetherington's software carpentry courses. Whilst BCS did not want to jeopardise research nodes with this type of usage, he felt that either using some of the old nodes or purchasing some low-powered hardware would be a good short-term solution. This might also be used to run applications such as

Mathematica, MATLAB and SAS. The group agreed that a blanket ban on Legion usage would be too harsh, and also inadvisable as the value of educating the next generation of research programmers should not be underestimated. Some departments (e.g. Chemistry) have dedicated clusters for this purpose, though others clearly do not as there is an obvious demand for a central resource. CG said she felt that ISD's Teaching & Learning group were not in the best position to champion the cause and any application for additional hardware for teaching would be likely to fall foul of the bid process. NK felt that the discussion was now taking a different tack to the agenda item, and suggested that an item on teaching and learning use should be added to the agenda of the next meeting. BCS to bring a proposal for a T&L policy to be discussed at the next meeting (see new Action 105).

## **11. Next Meeting Date and Agenda**

- 11<sup>th</sup> January 2012 from 1pm – 3pm, Cruciform Foyer 102 Seminar Room 2, Cruciform Building, Gower Street, London, WC1E 6BT

### **Agenda (Items) for the next meeting:**

#### **Standing items:**

1. Approval of Minutes of last meeting
2. Update on status of current Actions
3. Any requests for additional resources: (a) scratch quota requests, (b) any other request.
4. Review of any Centre for Innovation (Cfi) access requests (Chair)
5. Reporting data/stats to be made available to users

#### **Additional items:**

6. Management and use of 7-day queue for Gaussian job – discussion (chair)
7. Role of Consortia leaders – discussion (chair)
8. Verbal update on Legion IV project (Bruno Silva)
9. Proposal for a policy allowing Teaching and Learning usage of Legion by UG students and taught PG students – discussion (Bruno Silva)

## LIST OF CURRENT ACTIONS

Shaded (closed/completed) items will be deleted in the next version.

	Actions	Status	Owner
73	Devise proactive strategy to inform users of the availability of Condor.	<p>(25/11/11): Email consortium leaders and speak to members of the Genomics community (Jacky Pallas perhaps) regarding Condor to move this forward.</p> <p>(27/01/2012): Pending completion of on-going works to improve usability of Condor service.</p> <p>(24/02/2012): IN PROGRESS. Usability works still in progress; OK to liaise with ISD Datacentre Services to create shared storage for users to install applications (e.g. R &amp; Python) via Research Computing.</p> <p>(30/03/2012): IN PROGRESS. Technical issues regarding mounting of shared storage currently undergoing testing.</p> <p>(24/05/2012): IN PROGRESS. Technical issues regarding mounting of shared storage currently undergoing testing</p> <p>(26/06/2012): IN PROGRESS. Shared storage areas have been mounted on Condor machines, applications require placement.</p> <p>(25/07/12): Can now mount and start moving users. Going to make Condor useable from Legion itself, and should be available by next year. Will be a hybrid system.</p> <p>(10/09/2012): Modification to Condor successful. Users will be contacted to move their web jobs for Condor.</p> <p>(12/10/2012): We have found unexpected problems mounting the K drive – we are investigating what the problem is. Nevertheless, Condor is usable.</p> <p>(9/11/2012): Further testing of K drive to be undertaken and results to be reported at next meeting.</p> <p><b>(14/12/2012): After further testing, the problem appears to be with either DNS or firewall; some nodes in the Bartlett were removed as job submission was not possible. A temporary solution has been implemented (not using institutional filestore).</b></p>	BCS
76	Peter Harrison request for additional Unity resource	<p>(27/01/2012): To follow up with requestors regarding additional information about check-pointing and problem decomposition.</p> <p>(24/02/2012): Update from DG, Peter Harrison has been working closely with RC team and adapting code where</p>	BCS

		<p>possible. CRAG agreed to extend 72hr wall-clock to 10 days on Unity. CLOSED to be REVISTED at a later date</p> <p>(24/05/2012): DG to contact the user regarding progress of his work.</p> <p>(26/06/2012): Re-assigned to BCS</p> <p>(25/07/2012): BCS to speak to contact Peter Harrison.</p> <p>(10/09/2012): PH requests continued use of Unity and will provide review for next CRAG meeting.</p> <p>(12/10/2012): CRAG approved Peter Harrison's indefinite request with proviso that Legion is cited in research papers. BCS to provide update to CRAG in 3 months.</p> <p><b>(14/12/2012): ONGOING – update next month</b></p>	
88	Consolidated metrics and information for reporting and users	<p>(24/05/2012): Merge items 70,79,83,84,85 and report at next meeting</p> <p><b>70 – (26/06/2012): ONGOING</b></p> <ol style="list-style-type: none"> <li>1) Total node utilisation – The amount of time a node is being utilised by any job vs. total up-time (simple binary measure per node) summed over all nodes.</li> <li>2) Used/requested memory – The ratio of effectively utilised memory to requested memory.</li> <li>3) Requested memory vs. Requested cores plotted over time – Simultaneous (side by side) plot of CPU and memory utilisation over time, possibly normalised for each component.</li> </ol> <p>Points 1 and 3 have now been implemented as database entries and the RC team is now working on presenting these as a report. Implementation of Point 2 requires further investigation.</p> <p><b>25/07/12 – Database entries are now ready and will be presented at next meeting.</b></p> <p><b>14/12/2012 – ONGOING; Item 2) cannot be implemented as SGE has a bug in the current version that reports actual memory utilised by a job inaccurately.</b></p> <p><b>83 - (26/06/2012)</b></p> <p>To provide report and analysis on the Serial vs. Parallel parts of the system for top 5 users.</p> <p>25/07/2012 - DONE</p> <p><b>It was agreed that a discussion needs to be had at a later date, in the form of a presentation at a future meeting, about what information is made available to users and what information is made available to the group – add</b></p>	BCS



		<p><b>item to agenda.</b></p> <p><b>85 - (26/06/2012): ONGOING</b></p> <p>Develop simple utilities of transparency for infiniband jobs, a simple map of what is running and where.</p> <p>25/07/2012 – ONGOING</p> <p><b>14/12/2012 - ONGOING</b></p> <p>(12/10/2012): Note that in future discussion in meetings re usage should be 'by exception', with all members to review reports using online tool (link has already been circulated) prior to each meeting. Next step will be to extract data from scheduler for live up to minute reporting.</p> <p>(9/11/2012): Academic members to review reports using online tool forward comments as to which information should NOT be available to users</p> <p><b>14/12/2012 - DONE</b></p> <p>(9/11/2012): BCS to anonymise data for viewing service metrics</p> <p><b>14/12/2012 – DONE</b></p> <p><b>(14/12/2012): Action reduced to Infiniband utilities.</b></p>	
90	Provide information to CRAG about backfill policy and job characteristics of top user.	<p>(25/07/2012) – Action requested</p> <p>Requested to clarify whether the job distribution observed in July was due to backfill settings and user job profile.</p> <p>(10/09/2012): BCS to provide data/stats for backfill jobs.</p> <p>(9/11/2012) – ONGOING. Investigating ways to provide useful information.</p> <p><b>(14/12/2012): BCS has spoken to the systems team; hard evidence is difficult to produce as the data is complex but the scheduler appears to be backfilling jobs correctly; by definition, any job that runs by finding space in the cluster that is reserved for a future job is NOT preventing other jobs from running; this future job will also backfill if it can, hence leading to the complexity. Serial jobs will not run on Infiniband nodes unless they are backfilling. BCS displayed a graph showing serial jobs on X and W nodes – there were not a huge quantity (there were a fair quantity of single node jobs but these were all multicore). Utilisation is typically ~ 90% indicating no spare capacity. Chair was happy to accept that backfilling is not a problem and will speak to DA.</b></p>	BCS

		<b>CLOSED</b>	
91	Establish policy for requesting Priority CP hours	<p>(10/09/2012): CG to circulate draft policy paper to CRAG members and inform Serge Guillas that his request is under review. All CRAG members to report back on implementation of Priority Queue.</p> <p>(12/10/2012): TJ to investigate implementation of Priority Access using 'Projects' method as discussed and agreed by group. Gold Accounting Software – RC to investigate by further testing. NA to provide local Miracle users for testing.</p> <p>(9/11/2012): Still pending. Meanwhile, TJ to set up priority access for Miracle jobs as previously agreed using same set up on Miracle as for Harvest project (Serge Guillas).</p> <p><b>(14/12/2012): Done for Miracle jobs; TJ to present Gold accounting software information at the next CRAG.</b></p>	TJ
92	Review of next version of KPIs' 'strawman'	<p>Re-edit KPIs document and present to WG1 Research Computing working group.</p> <p>(9/11/2012): document re-edited, to be submitted to next RIISG</p> <p><b>(14/12/2012): CG to take to the next working group.</b></p> <p><b>CLOSED</b></p>	CG
94	User viewable 'live' service info	<p>(9/11/2012): To add core service statistics and view of present user windows of opportunities and number of nodes available to website.</p> <p><b>(14/12/2012): Service statistics added to webpage. View of present window of opportunities is not necessary owing to the fact that backfill actually takes care of automatically filling scheduling gaps. - ONGOING</b></p>	BCS
95	Centre for Innovation (Cfi) access requests	<p>(9/11/2012): BCS to review Cfi request from Dr Andrew Logsdail and report back regarding the Global Arrays toolkit over QLogic on Legion.</p> <p><b>(14/12/2012): BCS has done this.</b></p> <p><b>CLOSED</b></p>	BCS
96	Record of Cfi applications	<p>(9/11/2012): It was agreed that a spreadsheet record of all Cfi requests, including reasons for rejection where appropriate, should be maintained.</p> <p><b>14/12/2012 - ONGOING</b></p>	CG

97	Cfl usage statistics	(9/11/2012): CG to provide reports on usage data of 25% share of Cfl facilities at future meetings. <b>(14/12/2012): On the agenda for the current meeting. CLOSED</b>	CG
98	Publicising use of Legion@UCL and new scheduler information graph	<b>(14/12/2012): CG to publicise the location on the website which advises how to acknowledge use of Legion (Legion@UCL) and the new graph of scheduler information, by emailing the research-computing list.</b>	CG
99	Cfl GPU applications developer	<b>(14/12/2012): There is a dedicated GPU developer to help with applications on Emerald; Legion users to be advised of this.</b>	CG
100	Legion jobs showing unusually long wait times	<b>(14/12/2012): BCS to contact the Legion user who had a job with a very large wait time to run-time ratio (uccaeak) to see if there was a problem; also to check December's statistics to see if the pattern was repeated for the SysBioMed consortium.</b>	BCS
101	Cfl usage statistics	<b>(14/12/2012): CG to clarify whether STFC's share of Emerald is included in the under-utilisation figures.</b>	CG
102	Cfl Emerald applications	<b>(14/12/2012): BCS to circulate a list of the applications which are available on Emerald to the research-computing mailing list</b>	BCS
103	Cfl IRIDIS job classes	<b>(14/12/2012): CG to obtain data on IRIDIS job class distribution.</b>	CG
104	Legion 7-day Gaussian queue	<b>(14/12/2012): BCS to email the research-computing list advising Legion users that they need to reapply for access to the 7-day Gaussian queue before the end of February, if still required, or access will be revoked.</b>	BCS
105	Legion Teaching and Learning use	<b>(14/12/2012): BCS to bring a proposal for a Legion Teaching &amp; Learning policy for discussion at the next meeting.</b>	BCS