

UCL Computational Resource Allocation Group (CRAG)

MEETING MINUTES

10th September 2012

In Attendance:

1. Prof Nik Kaltsoyannis (Chair) - *Molecular Quantum Dynamics and Electronic Structure*
2. Dr Bruno Silva - *Research Computing Platforms Team Leader (Service Lead), ISD*
3. Prof Dario Alfe - *Thomas Young Centre (Materials Science)*
4. Dr Ben Waugh – *High Energy Physics*
5. Clare Gryce - *Head of Research Computing and Facilitating Services, ISD*
6. Thomas Jones - *Research Platforms Team Leader (Infrastructure Lead), ISD*
7. Dr Dugan Witherick - *Astrophysics and Remote Sensing*

Apologies:

1. Prof Andrew Philips - *Epidemiology*
2. Dr Vincent Plagnol - *Next Generation Sequencing*
3. Dr Nicholas Achilleos - *Astrophysics and Remote Sensing*
4. Dr Andrew Martin – *Bioinformatics and Computational Biology*

Note: Minutes below provide a high level summary of decisions taken and actions assigned by the Group. CRAG members are referred to the slide deck for the meeting for further information on matters presented.

1. Approval of Minutes of last meeting on 25th July 2013

The Group approved the Minutes of the July 25th 2012 meeting. The following updates were noted:

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

No requests have been received.

Any other requests:

Serge Guillas, requested that his current 'priority queue' arrangement be extended to provide a further 5,000 CPU hours.

CG explained the background to this previous request to the new members of the group. It was agreed that the extension request should be declined until this facility has been rolled out and made available as a service to all Legion users.

4. Miracle Consortium - proposal for use of remaining paid-for resources

DW explained that the Miracle Consortium was transitioning to use the new DiRAC facilities, but would continue to also use Legion according to the terms of the previous agreement.. CG explained that Miracle Consortia users had not used all CPU hours provided under the terms of the previously agreed MoU, resulting in excess unused hours. It was agreed that a priority 'queue' would be set up for use by the Consortium; this would be provide for 12 months for UCL users and 6 months for external Miracle users (from the date of set up) and that any CPU hours remaining after this time would be 'lost'. This would be subject to monthly monitoring by CRAG.

5. Legion Service Stats for July and August

The reports were presented and it was noted that currently there were 127 users, service usability was at 99% and total CPU available was 98%. CG explained that the latter was due to pre-emptive and automated drain of queues to specific Computational Units (CUs) related potential data centre issues.

6. & (8A) Legion usage stats for July and August (using new reporting framework). Review what reports does the CRAG require, in what format?

BCS gave an update on CPU and wait times for each Consortium. BCS explained why jobs would be rejected (and hence show very extended wait times) although this would be very rare. BCS presented the new web reporting tool. It was deemed that the following information was missing:

- a) Breakdown of jobs for top 5 users.
- b) Metrics for wait times of job class over time.

CG explained that the web reporting tool information was currently available for CRAG members only. However once the tool and contents are stable, it would require reviewing as to how it may be presented to users for transparency and accountability.

7. Review of next version of KPIs' 'strawman' document

CG explained the current draft KPIs document. BW suggested a re-edit explaining that some points were KPIs while others were management information reports.

8. 8b Unity

8c Condor:

Review of monthly reporting (what reports does the CRAG require, in what format?)

CG enquired as to what the type of information the CRAG required for these machines. BCS clarified Information is to be provided as part of the reporting tool,

9. Next Meeting Date and Agenda

- 12th October 2012 from 1pm – 3pm, Cruciform Foyer 102 Seminar Room 2, Cruciform Building, Gower Street London, WC1E 6BT.

Agenda Items:

- 1 Longevity of Legion user accounts added or renewed.
- 2 What to do regarding:
 - 7-day jobs (Gaussian)
 - Privileged Users benefiting from these 7-day

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 & 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. ?? is going to start speaking to 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>	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 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>	DG/BCS

87	Reports on key publication reports to be used for promoting Legion	<p>(24/4/2012) – Action opened: NK to email Consortium Leaders asking them to solicit lists of key publications that have been enabled via use of Legion, for use on RC website and to encourage more requests for ‘additional resources.</p> <p>(26/6/2012): Link publications on RC website and provide list of Consortia publications for standard Legion use.</p> <p>(25/07/2012): Now have a complete list from main users of Legion. To be discussed with CG when she returns from leave.</p> <p>(10/09/2012): CG to provide update on next meeting re linking and tagging enabling research content to be dynamically generated to allow readily refreshed content.</p>	NK
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>70 – (26/06/2012): ONGOING</p> <ol style="list-style-type: none"> 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. 2) Used/requested memory – The ratio of effectively utilised memory to requested memory. 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. <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>25/07/12 – Database entries are now ready and will be presented at next meeting.</p> <p>83 - (26/06/2012)</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>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 item to agenda.</p>	BCS

		<p>85 - (26/06/2012):</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>	
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>	BCS
91	Establish policy for requesting Priority CP hours	<p>(10/09/2012): CG to circulate draft policy paper to CRAG members and inform Sergey Yurchenko that his request is under review. All CRAG members to report back on implementation of Priority Queue.</p>	CG
92	Review of next version of KPIs' 'strawman'	<p>Re-edit KPIs document and present to WG1 Research Computing working group.</p>	CG

CURRENT ACTION ITEM 76

Peter Harrison request for additional Unity resource

From: <Harrison>, Peter <p.w.harrison@ucl.ac.uk>

Date: Thursday, 2 August 2012 14:42

To: Bruno Silva <b.silva@ucl.ac.uk>

Subject: Re: CRAG request update

Dear Bruno,

I will amend that to the current manuscript when it returns from review and ensure to include it on all future publications. Best wishes,

Peter

On 2 Aug 2012, at 14:39, "Silva, Bruno" <b.silva@ucl.ac.uk> wrote:

Dear Peter,

Many thanks for your update.

To acknowledge work done on Research IT Services facilities, we ask researchers to include in the Acknowledgements section their papers the text in the pages below.

For Legion:

http://www.ucl.ac.uk/isd/staff/research_services/research-computing/services/legion-upgrade/userguide/acknowledginglegion

For Unity:

http://www.ucl.ac.uk/isd/staff/research_services/research-computing/services/unity/guide/AcknowledgingUnity

Best Regards,

Bruno

From: <Harrison>, Peter <p.w.harrison@ucl.ac.uk>

Date: Monday, 30 July 2012 11:19

To: Bruno Silva <b.silva@ucl.ac.uk>

Subject: Re: CRAG request update

Dear Bruno,

The work is progressing well, but we have suffered a number of setbacks with problems with our data, preferred methodology of the research community, and from glitches on unity restarting and terminating jobs (hopefully all of which have now been resolved).

We do still require the wall clock time for our analyses if it can be left in place. As usage on Unity has been very light over the past couple of months I have actually been running a greater number of longer jobs, as this is more efficient in terms of loading and checkpointing datasets, however I can avoid using the long queue for all but essential jobs (i.e. those with checkpoints longer than 72 hours) if you and CRAG prefer. Instead running a series of 72 hour jobs which comes under normal CRAG queue conditions. If the wall clock limit could be left in place for when it is required that would be very useful to my group. As CRAG initially stipulated I have been running my analyses over a longer period in order to keep the memory footprint as low as possible.

I have one paper in review at the moment in Genome Biology Evolution, and a further three papers are currently being written up, one by me and two by colleagues that utilise the data I have generated on Unity. The plan is for this data to generate about 10-20 papers over the next 4 years, depending on how much we discover, all of which will have required the special request on Unity. Is there anything you would like placed within the papers to acknowledge usage of Research computing facilities? I will pass on links to publications as and when they are available. The majority of major papers could be a year or more away yet as a lot of further downstream analyses and verifications need to be conducted.

We have also started to obtain new datasets, which may require the same analyses to be conducted. Obviously these will generate further publications of their own. If the analyses can be conducted under the same CRAG request that I have already made that is great, otherwise I can submit additional requests as and when new data is available.

I am hoping that David Gregory will be able to install newer versions of the software that I use soon, which should speed everything up a little.

I hope this is all satisfactory and that we are generating sufficient impact to warrant the increased resource requirement. Please feel free to contact me if you require further information on any of the above. Best wishes,

Peter