

RESEARCH COMPUTING PLATFORMS WORKING GROUP

NOTES

Date: Monday 12th December 2011

Time: 12:00 – 14:00

Venue: 14 Taviton St, Room 128

Members Present:

1. John Brodholt [JB] – Chair
2. Richard Catlow [RC] – Executive Dean, MAPS
3. Eric Fraga [EF] – Engineering
4. Nora H De Leeuw [NL] – (Thomas Young Centre)
5. Jacky Pallas [JP] – Platform Technologies
6. Christine Orenge [CO] – Structural and Molecular Biology
7. Clare Gryce [CG] – ISD Research Computing
8. Bruno Silva [BCS] – ISD Research Computing
9. Thomas Jones [TJ] – ISD Research Computing

Representatives:

10. Stefan Zasada [SZ] (for Peter Coveney – Computational Life and Medical Sciences)
11. Dugan Witherick [DW] (for Nick Achilleos – Miracle Consortium)
12. Vincent Plagnol [VP] (for David Jones – Bioinformatics)

Matters Discussed

1. Recap on brief for Working Group

CG briefly recounted the brief for the Working Group (*Slides presented at meeting are appended as APPENDIX A - see slide 3*). It was noted that the document describing the new IT governance structure, which includes the RIISG was still to be distributed to the Group [**ACTION CG**]

2. Minutes and Actions of last meeting (18th Nov)

CG reported that the Minutes of the last meeting were still in preparation and would be distributed as soon as possible.

3. Research Computing Platforms Strategic Roadmap

CG gave a presentation outlining two possible directions for the strategic development of UCL's central HPC Facilities (see *APPENDIX A slides 4 to 14*) and asked the meeting for guidance as to which option should be pursued. An outline project programme for years 2011/12 and 2012/13 was presented in both cases. Options presented:

1. To consolidate all (central) HPC to single off-site facility,
2. To retain on-site facility, and establish a second offsite HPC facility.

The short term replacement of the current high performance file-system platform was recommended in both cases

The meeting gave unanimous support for the (second) option to maintain an on-site HPC facility in the short to medium term, noting the following concerns:

- That there were 'too many' risks involved in pursuing a consolidation strategy at the current time.
- The need to maintain an HPC facility close to UCL Desktops, for data transfer reasons.
- The possible 'psychological' impacts of geographical distance.
- Current uncertainties regarding operational models and processes for a geographically remote facility.

4. AOB

A presentation on the ISD SmartIT re-organisation, and the new Research IT Services Department, was requested for the next meeting [**ACTION CG**]

5. Date of next meeting

It was agreed that the next meeting should be held in February.

List of Current Actions

Shaded Actions will be deleted after next meeting.

No	Date	Minute	Owner	Action	Status	Notes
1	18/11/11			Review Membership	OPEN	<i>Minutes of 18/11 meeting still to be confirmed</i>
2	18/11/11	1.2.1	R.Ch	Circulate RIISG minutes	OPEN	<i>Minutes of 18/11 meeting still to be confirmed</i>
3	18/11/11	1.4.1	CG	Circulate current HPC Purchasing "Straw man" document.	CLOSED	<i>Distributed by email on 23/12/11</i>
4	18/11/11	1.5.3	CG	To draft a document outlining future resource allocation policy, based on discussions.	OPEN	<i>Minutes of 18/11 meeting still to be confirmed</i>
5	18/11/11	1.6.1	LRAW	Establish policy for cluster/consortia leaders.	OPEN	<i>Minutes of 18/11 meeting still to be confirmed – clarification of Action required.</i>
6	18/11/11	1.7.1	CG	Progress discussions within ISD regarding research computing use of Desktops.	OPEN	<i>Minutes of 18/11 meeting still to be confirmed</i>
7	18/11/11	3.1	CG	Poll for date of next meeting.	CLOSED	Completed.
8	12/12/11	1	CG	To distribute document describing new IT Governance structure.	CLOSED	<i>Distributed by email on 09/01/12</i>
9	12/12/11	4	CG	To give presentation on SmartIT and RITS to next meeting.	OPEN	

**Research Computing Platforms RIISG
Working Group**



12th December 2011

Agenda

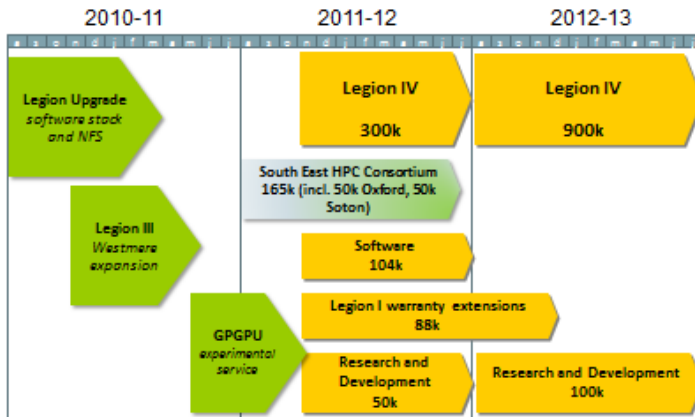
1. Recap on brief for Working Group
2. Minutes and Actions of last meeting (18th Nov)
3. Research Computing Platforms Strategic Roadmap
4. Date of next meeting

Brief for Working Group

- Research Information and IT Services Group convened, first meeting held 16th November
- 4 Working Groups agreed:
 - Research Data
 - Research Management Applications
 - Researcher Applications
 - Research Computing Platforms
 - To look at strategy, services and policies in these areas in more detail.
- Next meeting of RIISG on 26th January

To circulate – Document on IT Governance Structures

'Strawman' Programme Proposal (ISC budgets agreed)



- Approved by RIISG
- Opportunity to move funding between fiscal years

Research Computing Platforms – Drivers and Context

Internal Drivers

- Demand for compute capacity continues to increase
 - Both parallel and serial workloads
 - Continued growth in demand from Life and Medical Sciences anticipated
- Increase in data-intensive workloads
 - Need for fully integrated networked solutions
- Constraints on permanent staffing levels
- South East HPC Consortium evolving

External Drivers

- Drive towards Shared Services (regional model?)
- Uncertainty regarding national tier

Current Legion Service Issues

1. Wolfson House Data Centre issues (power capacity and heat distribution)
2. Job Wait Times
 - 'capability' workloads particularly hard hit
3. Data transfer rates (getting data in and out of Legion)
 - Campus Network bottlenecks
 - 'Gateway' nodes (current login nodes serve as bottleneck)
4. I/O throughput constraints (internal to Legion)
 - Compute nodes limited to 1Gb/s
 - Lustre HW platform not optimised for high I/O workload (number of file operations requested)
 - % of I/O intensive work has increased (esp. from Biosciences – work to parallelise apps is progressing)
 - Overall volume of workload has increased (Legion III capacity)
5. Data bottlenecks
 - Fragmented and complex data handling environment owing to highly hierarchical storage (\$TMPDR, /scratch, /home); (arising from points 3 and 4)
 - Insufficient capacity for secure storage of large data volumes pre and post processing

Current Service Issues – Impacts

1. **Wolfson House Data Centre issues (power capacity and heat distribution)**
 - Service vulnerability (periodic capacity reductions)
 - Impediment to HPC infrastructure and service consolidation
2. **Job Wait Times**
 - Reduced research throughput
 - Impediment to HPC infrastructure and service consolidation
3. **Data transfer rates (getting data in and out of Legion)**
 - Reduced research throughput
 - Challenges in supporting some workloads (esp. 'Big Data', active databases)
4. **I/O throughput constraints (internal to Legion)**
 - Inefficient use of staff time (supporting workarounds)
 - Reduced efficiency of user's work (dealing with complexity of environment)
 - Longer overall job completion times owing to multiple internal data transfers
 - Challenges in supporting some workloads (esp. High Throughput Sequencing)
5. **Data bottlenecks**
 - Barrier to entry for less experienced users
 - Increased support load

Current Service Issues – Solutions

1. **Wolfson House Data Centre issues**
 - Project started to deploy water cooled doors to 2 Legion CU's (funded from other budgets)
2. **Job Wait Times**
 - Increase compute capacity
3. **Data transfer rates (getting data in and out of Legion)**
 - Increase spec of login nodes
 - Strategic enhancements to Campus Network
4. **I/O throughput constraints (internal to Legion)**
 - Upgrade Lustre HW platform
 - Eliminate bandwidth limits on compute hardware
 - Rationalise internal data handling architecture
5. **Data bottlenecks**
 - Establish data 'third tier' ('Data Staging') area
 - Use same platform as for Lustre, flexible provisioning

Strategic Research Computing Platforms Roadmap – options

1. **Consolidate all (central) HPC to single off-site facility**
 2. **Retain on-site facility, establish second offsite HPC facility**
 - Off-site optimised for 'capability' workloads
 - On-site optimised for 'capacity' high-throughput workloads
- *Recommend replacement of current high performance file-system platform Q1 2012 in both cases*

Consolidate to single off-site HPC facility

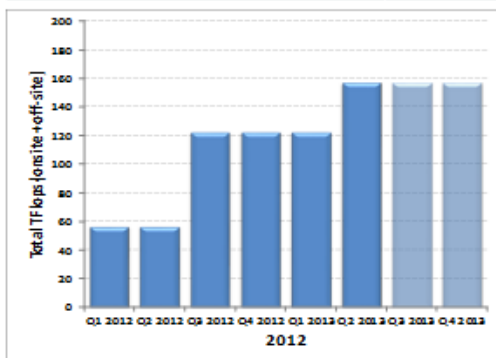
Key Benefits *	Key Risks/Issues *
Efficiency of systems administration	Possible network bottleneck
Alignment with emerging political and funding agenda	Implementation risk – dependency on JANET Data Centre project
Alignment with SE Consortium	Non-alignment with SE Consortium
Greater flexibility to rapidly deploy experimental/niche systems on-site	



* Subject to further analysis

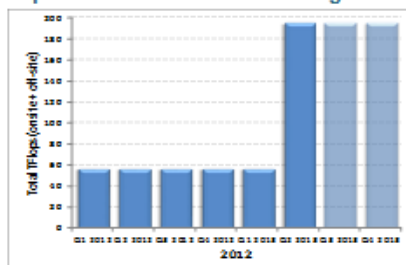
Two HPC facilities (on-site and off-site)

Key Benefits *	Key Risks/Issues *
Capacity compute close to UCL data sources and on UCL Research Network	Possible higher systems administration cost
Alignment with SE Consortium	Non-alignment with SE Consortium
Some cross-site resilience possible	Conflict with other major projects and services for on-site Data Centre space



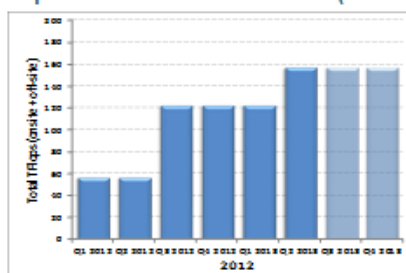
* Subject to further analysis

Option 1 - Consolidate to single off-site HPC facility



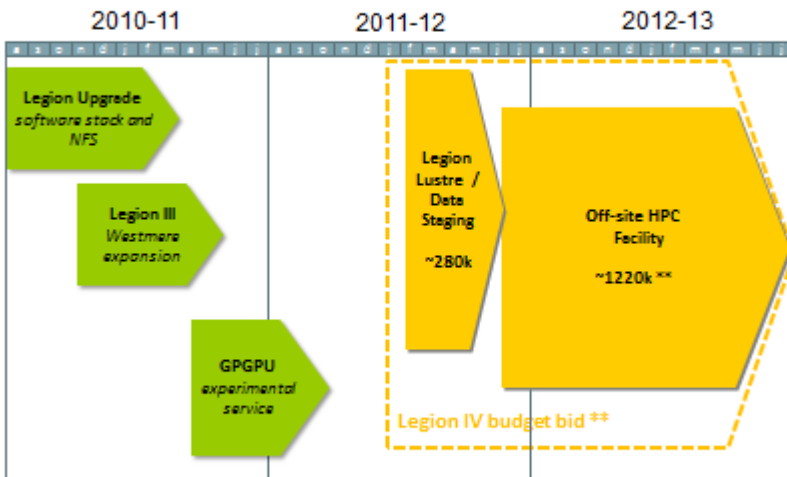
Illustrative programme:
 Q1 2012 – Q1 2013: Current Legion service
 Q2 2013: Legion service reduced (Legion III only) + off-site facility (17 high density racks in new external DC)
 Q3 – Q4 2013: as per Q2. Further extension of off-site facility possible

Option 2 - Two HPC facilities (on-site and off-site)



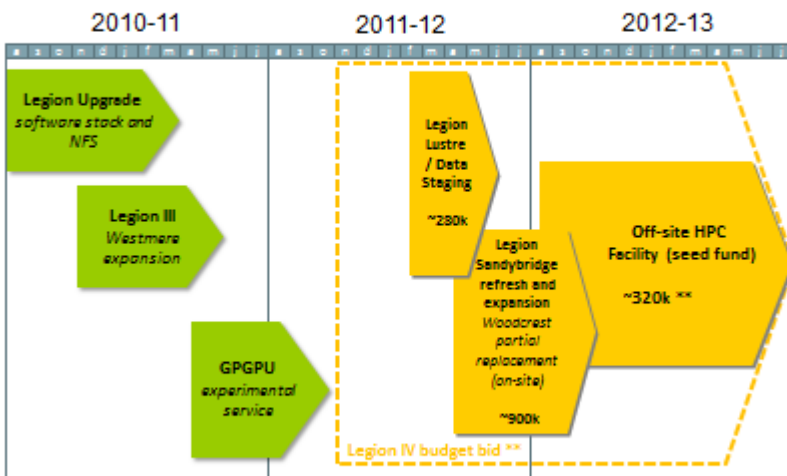
Illustrative programme:
 Q1 – Q2 2012: Current Legion service
 Q3 2012 – Q1 2013: Current Legion service - 5 CU's + Sandybridge expansion
 Q2 2013: Legion service = Legion III + Sandybridge expansion, + off-site facility (10 high density racks in new external DC)
 Q3 – Q4 2013: as per Q2. Further extension of off-site facility possible

Option 1 (Consolidate to single off-site HPC facility) – Programme Proposal



** Figures based on current 'strawman' budget for 2012-13.
Full programme likely to require additional funding

Option 2 (Two HPC facilities (on-site and off-site) – Programme Proposal



** Figures based on current 'strawman' budget for 2012-13.
Full programme likely to require additional funding