Research computing: consultation report

First draft of 27 September 2007
Cuna Ekmekcioglu

1. Introduction

Research efforts rely heavily on IT infrastructure, people, and a broad range of IT services to collect, analyse, and distribute information across an expanding range of academic disciplines and geographic locations. Ever-larger data sets are being collected and shared, simulations and visualization are becoming routine tools, and the co-evolution of science and computing increasingly requires scientists to have solid grounding in information management. Bearing this in mind, the University aims to attract, nurture and support its researchers, and offer a world-class support environment that enhances the productivity and capability of their research effort. To achieve this goal, there will be a need for wider consultation to identify and analyse the requirements of computation and data intensive applications of the research community.

2. Aim and objectives

The purpose of this consultation is to identify and analyse, from the perspective of the staff, the research computing support service requirements across the University, and to inform the development of IS services in the research computing area.

The objectives of the consultation phase of the research computing are:

- To ensure that Information Services has a good understanding of the research computing support service requirements across the University.

- To identify the level of support required for existing services and the critical services which do not exist but should be established.

- To carry out a full requirements gathering exercise across the University which will include:

  - computational infrastructure (including high performance throughout computing, storage and networking),
  - support for operating environments, middleware, virtual eResearch environments and portals being developed nationally,
  - data storage and management services, including assured backup,
• support for collaborative research environment,
• support for visualisation and information discovery,
• support for research software,
• other services which may arise as a result of consultations.

- To inform the development of IS services in the research computing area.

3. Methodology

Research computing consultation process was a qualitative study, involving two phases: issue capture and analysis of these issues.

Capturing issues

The issue capture process involved an online survey which was circulated within the University research community to determine the current state and gather requirements concerning the computing needs of the researchers. This survey was then followed up by individual meetings.

Analysing issues

The issues arising from both the survey and the interviews are analysed in Section 4 with the intention of informing the development of IS support services in the research computing area.

4. Research computing at the University of Edinburgh: analysis of the consultation results

4.1 General points

The online survey returned 503 responses over a period of 2 months. Of the total, 292 responses were from CSE, 136 from CHSS, and 74 were from CMVM (Appendix: Chart 1). 72% of the total responses were from academic/research staff and the rest (28.8%) from PhD students (Appendix: Chart 2).

4.2 Current state in research computing

4.2.1 Research tools in use

The distribution of tools being used by research staff is given below. Access Grid received the lowest hits whereas graphical tools and libraries received the highest. There was an interesting comment on Access Grid:
“Good web conferencing would be helpful - maybe Adobe Breeze is the right thing here. Access grid is a bit of a joke - I'm not sure why we still try to get it to work.”

It is also interesting to see that wikis, blogs and other social software are used by 64.4% and visualization tools gets a high rate by 61%. Desktop video conferencing (26.7%) is slightly more popular than Video Conferencing Suite (23.4%). One of the participants commented:

“please note that I don't use videoconferencing because of the wholly inadequate support in the school …”

which might raise a question of support of this tool within the University.

<table>
<thead>
<tr>
<th>Tools</th>
<th>% of use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access Grid</td>
<td>12.5</td>
</tr>
<tr>
<td>Podcasts</td>
<td>16.6</td>
</tr>
<tr>
<td>XML for publishing research papers</td>
<td>20.6</td>
</tr>
<tr>
<td>Videoconferencing suite</td>
<td>23.4</td>
</tr>
<tr>
<td>Desktop Videoconferencing</td>
<td>26.7</td>
</tr>
<tr>
<td>Cluster computing</td>
<td>32.0</td>
</tr>
<tr>
<td>Instant messaging</td>
<td>37.0</td>
</tr>
<tr>
<td>Simulations</td>
<td>47.4</td>
</tr>
<tr>
<td>Maths packages, e.g. Matlab, Mathematica</td>
<td>53.7</td>
</tr>
<tr>
<td>Libraries, e.g. NAG, Numerical Recipes, GNU</td>
<td>57.6</td>
</tr>
<tr>
<td>Programming tools, e.g. C, C++, Fortran, Java</td>
<td>58.8</td>
</tr>
<tr>
<td>Statistical packages, e.g. SPSS, SAS, Minitab</td>
<td>61.4</td>
</tr>
<tr>
<td>Visualization tools</td>
<td>61.0</td>
</tr>
<tr>
<td>Wikis, blogs and other social software</td>
<td>64.4</td>
</tr>
<tr>
<td>Graphical tools and libraries</td>
<td>73.4</td>
</tr>
<tr>
<td>Other tools</td>
<td>43.3</td>
</tr>
</tbody>
</table>

The frequency of the use of each of these tools is given in Appendix (Chart 3). Among the other tools used by research staff, the most popular ones are Latex, databases, reference management tools, Microsoft Office and Adobe software.

4.2.2 Support services for research

Most of the support services for research are rated well. The most popular services are built-in system help files (73.8%) followed by online training.
materials (72.4%), face-to-face training courses (69.4%), online communities and mailing lists (68.3%), and central services support person/help desk (67.4%).

**4.2.3 Main obstacles to carry out research**

"Lack of familiarity with the software" received the greatest hit among the listed main obstacles to carrying out research. This was followed by "difficulty in finding information on relevant computing facilities and services" and "lack of available software". It is worth mentioning that the "lack of support available to help installing and using appropriate software" was high in the ranking.

<table>
<thead>
<tr>
<th>Main obstacles</th>
<th>No. of hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of familiarity with the software</td>
<td>204</td>
</tr>
<tr>
<td>Difficulty in finding information on relevant computing facilities and services</td>
<td>167</td>
</tr>
<tr>
<td>Lack of available software</td>
<td>147</td>
</tr>
<tr>
<td>Lack of support available to help installing and using appropriate software</td>
<td>135</td>
</tr>
<tr>
<td>Lack of IT support for research administration</td>
<td>117</td>
</tr>
<tr>
<td>Lack of computing hardware facilities</td>
<td>93</td>
</tr>
<tr>
<td>Other</td>
<td>87</td>
</tr>
<tr>
<td>Lack of familiarity with research methods</td>
<td>86</td>
</tr>
<tr>
<td>Lack of advice on how to publish data sets for journals</td>
<td>34</td>
</tr>
</tbody>
</table>

Among the "other" obstacles listed by research staff, the most prominent ones are: lack of sufficient disk space, back-up facilities, and appropriate training. Also, quite a high number of staff expressed dissatisfaction with the existing “managed desktop” finding it an obstacle to carrying out their research efficiently and effectively.

**4.2.4 Data storage and back-up**

16.8% of the staff participated in the survey state that they work with data ≤100 MB whereas the majority (42.5%) has to store data ≤100 GB (Appendix 3). Their data is mainly backed-up on the memory sticks, CD/DVD, and local hard drive. Some data is backed-up on school or college servers (Appendix: Chart 4 & 5).

26.4% of the data collected by research staff would be impossible to re-create in the event of a loss. According to the 44.3% of the responses, it would take months to years to recreate the lost data (Appendix: Chart 6).
56.9% of the consulted staff wants to retain data whereas 30% state that they are obliged to retain it. A 16.1% would like to retain data for over 20 years (Appendix: Chart 7 & 8).

4.2.5 Tools in use for publishing research information

According to the survey results, quite a wide range of tools are being used for publishing research information the most prominent being the Web (341) followed by institutional repository, archive server (arXiv) and wikis.

4.2.6 What other support research community needs?

The wish list given by consulted staff is too long to include in this report. However, it will be made available on request. The main themes arising from this list are given below in no particular order with some quotes and highlights from the survey results. It is worth mentioning that there is a high demand for better data storage and back-up facilities, and for more training in how to use the available software, and statistical and maths packages.

Software

There is a definite need for:

- More software.
  "Free software repository, like Napier!"
- The ability to upgrade to more recent versions of software.
- Greater availability of statistical packages via university licenses.
- More essential software available through free licenses by the University.
- Overview of available software.
- More advice on software.
  "A directory of software (not research) skills of the University personnel? (Example: somebody very good with web services or Java could help me, even if she is using them for Geology while I'm going to use them for Informatics)."
  "A regular update of what research software is available; How to use it"
  "I feel that the support on Qualitative software packages is quite lean"

Hardware

In most cases existing hardware is not fast and powerful enough to cater for the needs of research community. The community needs:
• Better central support for Macs and Linux.
• Faster and more powerful machines.
• Enhanced laptop support.
• Faster network connections.

**Data storage and back-up**

This comes up as a very weak point in the IT facilities provided by the University. The University should provide research community with:

• Simple, secure and preferably an automatic data back-up service,
• Larger storage space on servers

"Additional network storage space as currently a lot of my data is not backed up - it is too large to keep on my hard drive, and is stored only on my external hard drive."

"More storage available on the University system; I can barely do any of my work (which often involves very large sound files) on University computers, which makes my work difficult since I do not have a copy of SPSS of my own, so I have to use the University computers."

"Easy to access high volume storage, linked efficiently to the machines I need to use, along with a non-snooty support that responds within less than 2 months (preferably within 12 hours would be nice)"

"An end to data being walked around the campus on memory sticks, network everything!"

• More robust archiving services

"Serious data archiving (TB level) would be very useful."

**Training**

Although there is already good training facilities for staff, this could be further improved by providing a better and more training in available specific packages. There is a high demand for information and training on a wide range of specific software particularly for qualitative data analysis packages such as SPSS, SAS, Graphpad Prism, N-Vivo, and other packages such as Matlab, Maple, Latex, Endnote, Cadence, and for advanced programming. There is also need for more on-line training and self-help tutorials.

"Better software training for research staff - substantial turnover every year, predominantly self-taught and using software to only fraction of potential"

**Awareness**
This is most probably one of the most frequently raised issues that IS need to deal with urgently. Research community needs to have:

- Knowledge of what resources are available at the University so that they can make use of the existing resources in their research.
- Well designed and easy to understand web pages. At present support information is all over the place on many different sites.
- Support to keep up to date with new tools.

"Help with keeping up to date with the new computing tools that are continually appearing, pointing out how they might be useful to a researcher (or teacher). I generally only hear about these by chance. There needs to be a link between those closely involved with computing who know about these tools, and the research community who are concentrating on other areas e.g. biology but who can usefully use the new computer techniques and tools."

**Access to IT facilities**

Again, research community needs easy and better access to:

- University facilities from non Windows computers.
- Software and some hardware.
- Workstations for number crunching.
- Online databases.
- Local image storage server.
- HPC time and clearer guidelines on how to get it.
- University support for local HPC facilities.
- Servers from home.
- Support on-line, by mail or by phone.
- Computing labs with additional software (such as ENVI, Erdas Imagine, STATA, S+, SPSS etc) and an easy to find list of such facilities.

**Website building and maintenance**

There is a definite need for:

- More help in building and maintaining websites

  "Support with designing and maintaining websites. These are now an integral way of promoting research, and the first thing someone looks if interested in your work."

  "Very bad support for web pages. This needs urgent attention"
• Easy access and construction of personal pages.

**Video conferencing and desktop conferencing**

According to the survey results there is a need for better support for video conferencing and desktop conferencing. Staff would like to use these facilities in their research however they either don’t have enough information or adequate support to start using them.

**Communication between IT staff and research community**

The results of the survey suggest that there is a lack of communication between IT staff and the research community as well as the local IT staff and central IT services.

"It would be great if there could be more communication between IT and the research community about how IT can actually help preserve valuable data. It is a headache navigating IT at UoE (it was easier at Harvard and the last company I worked for)."

"Support from IT person who knows how to talk in a language that research staff can understand, and is not condescension... if this was general this could lead to much better research."

"I would just like the IT staff to be more friendly and accessible - I'm not trying to do anything particularly difficult in computing terms but I have no experience and need help."

"In house IT support could be better. There is a weak link between in house IT support and Central IT services"

**eLearning**

Some of the participating staff pointed out that they need:

• Advise from MALTS/eLearning department when the research area is actually e-learning.

• Help with design of e-learning courses.

A number of staff also expressed a need for more support on cluster computing, parallel computing, grid computing facilities and visualization tools.

Finally, it is worth mentioning that there were quite a good number of comments from staff stating that they are happy with the current support.

**5. Research computing in other universities**

This section summarizes how a number of other universities are providing IT support to their researchers so that this information could be used by Information Services in designing a realigned support strategy. The examples given in this section include highlights from only a few UK Universities and universities outside
UK. We would like the Advisory Group members to suggest other examples to be included in this section.

5.1 UK universities

The selected UK universities are all major research intensive universities belonging to the Russel Group. These universities have dedicated sections on their websites with comprehensive information on research computing services they provide.

University of Bristol

The University provides computing support for its research community through the Advanced Computing Research Centre. This Centre provides key services such as High Performance Computing (HPC), Access Grid, Condor. The University has a dedicated website to Research computing: http://www.acrc.bris.ac.uk/ which can also be accessed from Information Services website.

University of Bristol is also a member of Worldwide Universities Network (WUN) which is a partnership of 17 research-led universities from Europe, North America, South East Asia and Australia. Within this network WUNgrid is an initiative which brings together the international research expertise at the WUN member institutions to work on a series of innovative collaborative projects which are made possible by the emerging Grid technology.

The partner institutions of the Worldwide Universities Network are international leaders in Grid and E-Science projects. They are also holders and guardians of substantial institutional archives and research collections of significance to the wider research community. The other British members of WUNgrid are: University of Leeds, University of Manchester, University of Sheffield, University of Southampton, and University of York.

Cardiff University

Presently, Cardiff University is working on an ambitious program called “The Modern IT Working Environment” (MWE) to deliver a new and comprehensive IT environment for everyone in the University. The MWE will provide a personalised, single point of access to the online resources, which will support Cardiff University users in all aspects of their learning, teaching, research and administration. The MWE infrastructure is being developed to help the University to achieve the following goals:

1. Collaborative research of international standing
2. Learning and Teaching excellence
3. A modern and distinctive Student Experience
4. Lean administration with high quality management information

Also, 2007 sees the creation of a new Division that will be known as “ARCCA” – Advanced Research Computing @ Cardiff to support the University’s research. ARCCA will provide, co-ordinate, support and develop High End Computing (HEC) and eResearch services for researchers at Cardiff University. It is planned to be fully operational by the end of 2007.

University of Manchester

At the University of Manchester, Research Computing Services provide advanced computing facilities, services and training for research and teaching. They have a section on the university web site dedicated to research computing services: http://www.rcs.manchester.ac.uk/rcs. Some of the key services they provide are: high performance and grid computing, visualization, Access Grid, HPC systems management, research computing consultancy and procurement support. Research Computing Services also has an active research and development program to identify, develop and evaluate new research computing technologies and services. The University is a member of the Worldwide University Network (WUN) grid.

University of Sheffield

The University of Sheffield has a section dedicated to research computing services on their Corporate Information and Computing Services website: http://www.shef.ac.uk/cics/researchcomputing/service.html. The University is a member of the Worldwide University Network (WUN) and WUNgrid.

University of Southampton

Research-based computational facilities are offered on a number of different hardware or software 'platforms', based mainly on the Unix operating system. A variety of multi-access computer servers are available, including Beowulf computational clusters and a Condor cycle-stealing Windows computational resource. The University hosts one of the UK's eScience centres, as part of the eGrid. University of Southampton is also a member of Worldwide Universities Network (WUN) and WUNgrid.

University's Information Systems Services System website has a section on Research support: http://www.soton.ac.uk/iss/index.shtml.

5.2 Universities outside UK

In addition to some UK universities given above some other universities outside UK are listed below.

University of Arizona
Center for Computing and Information Technology's Research Computing Support Group provides technical and scientific support to faculty, staff, and graduate students of the University of Arizona who are involved in computational research. Among the key services they provide are: high performance computing support, scientific visualization support, statistical consulting, Unix/Linux workstation support, research software licensing. They have a website dedicated to research computing support: [http://ccit.web.arizona.edu/index.php?id=research](http://ccit.web.arizona.edu/index.php?id=research)

**University of Alberta**

The Academic ICT (AICT) Research Support Group provides high-performance computing and support services to researchers at the University of Alberta and partner institutions across Canada. They have a website dedicated to research support at [http://www.ualberta.ca/CNS/RESEARCH/](http://www.ualberta.ca/CNS/RESEARCH/)

**UNC Charlotte**

University Research Computing (URC) is one of several support groups within Information & Technology Services (ITS) at the University of North Carolina (UNC) Charlotte. Their mission is to support the unique computing needs of UNC Charlotte's diverse community of research faculty by developing shared computing facilities and offering specialized services that would be difficult for individual research groups or departments to provide internally. They have a mission statement, a strategic plan and advisory committees in place. Their website is available at: [http://www.its.uncc.edu/URC/](http://www.its.uncc.edu/URC/)

**University of Wisconsin – Madison**

Among U.S. universities, UW-Madison is a pre-eminent research institution, ranking as a leader in: Federally and non Federally funded research, number of doctorates granted, and total research expenditures. To support that rich environment, the campus and DoIT are working to build a solid computing infrastructure that includes:

- A high-speed campus network with robust links to other research institutions
- Secure, reliable hosting for running applications and storing critical data
- Collaboration tools for sharing information across the campus and around the world
- Specialized software for analysis of data
- Systems for grants administration

The university is a member of the Worldwide Universities Network (WUN) and WUNgrid.
6. Closing comments

As a result of the consultations carried out, it has become apparent that research computing support needs to be a collaborative activity between Information Services and the research community.

Communications between the Information Services and the research community needs to be improved and Information Services should become more aligned to the culture of research.

A centrally provided data storage system, and a simple, secure, and preferably automatic, data back-up facility will be welcomed by research staff. Providing further training in areas where there are gaps will help staff to carry out their research more effectively and efficiently.

There is also an urgent need to develop greater awareness and understanding of what resources are available for staff to use in their research. This will tie in well with the idea of dedicating a section of Information Services website to research computing with relevant links to resources and facilities available for the research community at the University of Edinburgh.
Appendix

Chart 1: Distribution of responses by colleges

![Chart 1](image1)

Chart 2: Distribution of responses by staff roles

![Chart 2](image2)
Chart 3: Research tools: frequency of use
Chart 4: Amount of data stored by staff

Chart 5: Who manages the server if/when data is stored on servers
Chart 6: Work needed to recreate the data in the event of a loss

Chart 7: Who retains data?
Chart 8: How long data needs to be retained?

- < 5 years: 166
- < 10 years: 138
- < 20 years: 45
- > 20 years: 67