Prof. Peter Clarke, FRSE
F.IET, F.InstP, C.Eng, C.Phys
School of Physics and Astronomy
University of Edinburgh
James Clerk Maxwell Building
Kings Buildings
Edinburgh EH9 3JZ
Tel: +44 (0)131 650 6798
Fax: +44 (0)131 650 9819

20-June-2013

Submission of evidence to the House of Lords Select Committee on Science and Technology

Enquiry into scientific infrastructure, June 2013

Context of this submission

This submission concerns the e-Infrastructure (i.e. computing, data management, networking and people) that underpins several large collaborative international science endeavours in which the UK participates and which are funded by the STFC Research Council, including:

- particle physics experiments at the Large Hadron Collider and elsewhere;
- theoretical particle physics and astronomy;
- observational astronomy and
- nuclear physics experiment and theory.

This submission is made from the Computing Advisory Panel (CAP) of STFC. The CAP gives advice to STFC in all matters pertaining to computing. The CAP consists of independent members of the scientific community who are external to STFC. This submission has been formed independently of any STFC views and is not made on behalf of STFC.

We have commented only upon those questions where we consider it appropriate to do so, and by and large do not comment in any detail on policy matters, or other areas, which are in the domain of research council interaction with government. In some cases where a question covers two distinct topics we have greyed out parts which we do not cover.

Prof. Peter Clarke (University of Edinburgh)
Prof. Jeremy Frey (University of Southampton)
Dr. Adrian Jenkins (University of Durham)
Prof. Roger Jones (University of Lancaster)
Prof. Ralf Kaiser (University of Glasgow)
Prof. Neal Skipper (University College London)
Prof. Mike Watson (University of Leicester)

Responses

Q: What scientific infrastructure is currently available in the UK, do UK researchers have sufficient access to cutting edge scientific infrastructure and how does this situation compare to that of other countries?

The LHC programme and future nuclear physics experimentation require very large High Throughput Computing (HTC) infrastructures to reconstruct and analyse very large volumes of data (many hundreds of peta-bytes). Within the UK the LHC is provided for by the GridPP project, which comprises a central "Tier-1" facility hosted at the Rutherford Appleton Laboratory, which is federated with 18 regional "Tier-2" facilities hosted across University sites. These deliver 40,000 processors as part of an infrastructure of approximately 350,000 processors around the world. The UK hosts 35 peta-bytes of disk and tape storage for the LHC experiments. GridPP is considered to be very effective by its international LHC peers.

The theoretical physics communities require very high capacity High Performance Computing (HPC) facilities to perform large complex simulations. These facilities require sophisticated architectures with high-speed interconnections between many thousands of processing elements. Within the UK this is provided for primarily by the DiRAC facility, which federates five machines at four UK sites. These machines have different architectures, which are optimised for different scientific tasks. The UK DiRAC facilities are excellent, and are competitive on a global scale – for example DiRAC offers a 1.26 Pflop/s BlueGene/Q system for lattice calculations, which is of the same scale as in the USA and Japan. Dirac host systems ranking at 20,93 and 134 in the top 500. In addition these communities make use of HECToR (the national facility – which will be reported upon in other submissions) and of PRACE (the European facility) as well as some other machines in collaborating countries.

Both GridPP and DiRAC are funded through STFC. Both are very well organised and very cost effective for their science delivery. This is largely due to their strong academic lead (which ensures an optimal match to the science sectors they serve) combined with the strong engagement of STFC in their provision. Both compare well with international peers.

All communities produce large volumes of very valuable data (hundreds of peta-bytes) stored in data archives, which form part of the e-Infrastructure. In astronomy the UK hosts several very important repositories used by scientists worldwide, such as the infrared sky surveys from UKIRT/WFCAM and VISTA, which are part of the growing global Virtual Observatory. Data sets generated by instruments at facilities such as ISIS and Diamond are stored and archived on-site for exploitation by scientists worldwide.

Finally it is essential to understand that the UK e-Infrastructure relies upon a national network with reliable high quality end-to-end bandwidth capabilities between national and international data centres. Within the UK this is provided by Janet. Janet works closely with the scientific communities and as a consequence the UK has an excellent scientific network provision, which is only possible due to this symbiotic relationship between Janet and the communities it serves. Janet is one of the better National Research and Education Networks compared to international equivalents.

Q:Is sufficient provision made for operational costs and upgrades to enable best use to be made of the UK's existing scientific infrastructure?

We comment on electricity operating costs. It has been a long standing problem that although injections of capital for upgrades have by and large been available (although not always in the most timely way), unfortunately recurrent electricity operating costs are not consistently planned for at the same time. This is due to the disjoint between "capital" and "recurrent" costs which pervades UK infrastructure provisioning. This leads to significant uncertainties at times, and to ad hoc short-term solutions which can disrupt the competitive environment for UK scientists. This continues to be one of the biggest areas of concern for many UK scientists who depend upon the e-Infrastructure.

We feel it is also appropriate to comment on a second long-standing problem (we believe this is pervasive within most other domains within the UK). This is the area of software engineering support, which must be considered as a holistic part of the infrastructure. The software which is run upon the infrastructures must be properly engineered, maintained and, crucially, developed for the future to adapt to changing architectures and new algorithms. Talented software engineers are needed for this, but such posts tend to fall between two stools – being not considered as front line science posts, but neither are they operational or support posts. It is also evident that the career path of such people is in general not well provided for in Universities. A solution to this problem would enable far more efficient use to be made of the hardware infrastructure and improve research competitiveness. Proposals to capitalise such costs should be considered.

Q: Is it used to full capacity; and, if not, what steps could be taken to address this?

The infrastructures reported upon here are used to full capacity.

In the bigger picture better use of the UK's existing scientific infrastructure could be made if more progress was made in joining up the components in areas such as authentication and authorization spanning diverse administrative domains, accounting, monitoring, and data management, so that it becomes easier for collaborations to make use of different resources that may be available or better suited to a particular task.

Q:What substantial increases in scale would allow new areas or domains of science to be explored (analogous to Large Hadron Collider and Higgs boson)?

Q:What are the long-term needs for scientific infrastructure and how are decisions on priorities for funding usually made?

The importance of the HTC computing infrastructure to the LHC (and hence the Higgs boson discovery) is manifest. It successfully handled data at a scale, and at a speed, that was previously unimaginable. Some of the data on which the announcement of the Higgs boson on July 4th 2012 was based, had been taken less than two weeks earlier and within a week it had been included in the final analyses. The LHC will turn on again in 2015 bringing access to a new high-energy domain, which brings the possibilities of new discoveries such as candidates for dark matter. A commensurate large increase in e-Infrastructure capacity is needed to match this. The nuclear physics community working at the large international laboratories is also on the verge of requiring large HTC facilities.

In the HPC sector advances in theoretical physics will come from being able to carry out simulations at the next scale, be this for the discretisation of space-time, or setting the smallest unit of some astronomical or cosmological process to be simulated. Since there are at least four dimensions of time and space then each factor of two of improvement in scale requires a factor of 10-20 in computational power. Calculations are also needed to interpret the experimental measurements from the next phase of the LHC, and each new order of precision demands the inclusion of at least an order of magnitude more quantum mechanical processes. Thus DiRAC would benefit from access to 10-20 times the computing power on the timescale of 2015.

We can foresee an explosion in the amount of data being produced in the so-called "Big Data" era. The e-Infrastructure required to handle this is not only a question of scale (how many hundreds of peta-bytes) but must also include the facilities to curate the data and make it openly available for future exploitation. Open access is required by policies from Government levels down (e.g. the most recent G8 science Minsters statement released on12-June-2013). Making data openly available for further exploitation carries real resource implications. A substantial investment in hardware, services and people for the whole data lifecycle is needed if this vision is to be realised.

Looking to the longer term the UK involvement in the Large Synoptic Survey Telescope (LSST) would require the ability to mine multi-Petabyte databases, and the future Square Kilometer Array (SKA) international radio telescope project will make even greater demands upon the e-Infrastructure than ever before.

Q: Is it more important to invest in large, national infrastructure or medium infrastructure?

We wish to make a very important point here, which is that it is not a question of one or the other - both are needed for different phases of the scientific process. It is a scale and agility issue. Medium scale facilities (e.g. topical or regional centres) tend to provide agile access and fast turn around for smaller and/or many exploratory computations, which are crucial for developing insight in a given domain. In fact, for some activities only medium scale resources are required.

Larger national and international infrastructures tend to provide the capability that is essential for the largest tasks (e.g. reconstruction of raw data, the largest simulations and tasks requiring massive parallelization). In this case it would be wrong to assume this means a single machine or architecture, for this would most likely be sub-optimal for all applications.

The correct mix is therefore a nationally available facility, which federates a range of architectures and scales.

Q: Since the last Comprehensive Spending Review, a series of additional announcements have been made on investment in scientific infrastructure. How were the decisions on investment reached and what have been the impacts of this approach to funding scientific infrastructure?

The DiRAC facility for theoretical physics benefitted significantly from the capital investment in HPC facilities (£12.32 million in 2009, from the Government's Large Facilities Capital Fund and additional £15 million from BIS in 2011). These investments were very timely and were used very effectively. As mentioned above, DiRAC is an excellent facility and the UK now has access to facilities that allow calculations, which are competitive in size and scale with the rest of the world. Without this injection of capital the UK would have continued to fall behind international peers.

Some resources (£3M) went to ensure that the very last stretches of the network provision (that which is outside the jurisdiction of Janet) suitably match the capability of the national Janet. Approximately 20 universities have benefitted in a way which would not have been possible otherwise, and this is an excellent example of an area which fell between the funding for Janet and funding for scientific projects.

Q: If the current funding level is maintained or reduced, what would be the longer term impacts on scientific infrastructure in the UK? Governance structures

The UK e-Infrastructure relies upon maintenance of the funding level in real terms in order to obtain the increased resources need to match the increase in data which is far more than linear - in some cases exponential. If funding is at least maintained in real terms, then through “Moores law” we get much more resource per £ and can by and large “keep up”. However it may be noted that the adequacy of the current infrastructure has relied upon the (fortuitous) capital injections from infrastructure initiatives, indicating the baseline funding level is perhaps too low in some areas.

If funding is reduced in real terms, then we will slowly lose the ability to create, retain and process scientific data in a timely manner. We will fall behind the other countries and this will impact the ability of the scientific competitiveness of the UK.

Q: Are effective structures in place for funding of medium-sized scientific infrastructure and enabling sharing among Higher Education Institutes and Research Institutes?

Once a funding pulse is made available, then the funding structures operated within the STFC research council are excellent. This is due to the close and positive engagement that STFC maintains with the scientific communities in order to understand the (evolving) needs of the scientific areas. This leads to the knowledge and flexibility needed to make the best and most cost effective provision for the science.

However, at a higher level, the UK funding structures do not in general operate with the predictable regularity and timescale needed for forward planning of e-Infrastructures with lifetimes of decades rather than years. In the case of data archives, the commitment required is at least decade. Funding "initiatives" tend to have a 2-3 year outlook. In contrast one may note that this problem does not exist for Janet, which has a much more appropriate funding structure, and it would indeed be impossible to run Janet without this being so.

As an example, there was a very long delay introduced into the release of funds for HPC provision some 5 years ago caused by restrictions imposed upon STFC from structures above. This led to the UK falling well behind the rest of the world until the recent infrastructure initiatives. This was a very clear example of where structure hindered UK competitiveness.

Q: To what extent do funding structures in the UK help or hinder involvement in EU and international projects, and should the level of UK involvement be improved?

Q: To what extent are EU and international programmes effective in promoting collaborative investment in scientific infrastructure projects?

Several EU framework programmes have been effective in fostering international collaboration to build new infrastructures. The most notable in this context are the LHC computing Grid (e.g. DataGrid, EGEE, EGI,) and the PRACE facility. These were a success due to the relatively sharp focus upon the scientific goals in each case. We have some worry that as EU programmes seek to widen their focus (typically to try to include many more diverse interests in each project) they may lose focus and effectiveness.

The only issue which arises, and this is perhaps inevitable, is that national funding cycles are not always aligned. This can lead to problems in being able to make international commitments at the appropriate time.

It is very important to recognise the crucial role of the EU in promoting the international Geant network infrastructure (run by DANTE) which interconnects the national research and education networks (Janet in our case). It is only through close international cooperation that Europe has a very effective and pervasive end-to-end network for research.