CAP Submission to the 2012 STFC Programmatic Review

Prepared by the STFC Computing Advisory Panel (CAP)
October 2012

Final

1. Introduction

Computing underpins all of STFC supported science, and encompasses HPC (High Perfomance Computing), HTC (High Throughput Computing), storage and repositories, high performance networking, software and algorithms, and people.

Experimental Particle Physice depends upon very large distributred HTC resources for LHC operations. Theoretical Particle Physics and Astrophysics depend upon leading edge HPC facilities for modelling and calculations, Observational Astronomy depends upon repositories and high performance networks to link sites such as Lofar and SKA and many future projects have significant compute-intensive needs, Nuclear physics requires grid computing for large experiments, and the STFC facilities depend upon Computing close to the instruments to record and process data and good links and compatible systems/software to users institutions.

Computing is, in many cases, not funded as part of science project approval, but is separately funded either by computing specific projects, consolidated grants, or supplementary bids which are tensioned against all other calls on funding. Therefore we present this general submission to summarise the importance of Computing in addressing the STFC Science Roadmap Challenges. We also include some important (often long outstanding) issues which should be addressed, as well as the need to be able to respond to the changing landscape.

This document draws upon an earlier document [1] entitled “STFC HPC Requirements looking towards future provision of HPC facilities”, April 2011, which concentrated only upon the HPC sector. This is extended to reflect all computing relevant to STFC supported science.

2. Science Computing Challenges

Here we set the context with a very brief summary of science challenges which the Computing infrastructure currently addresses:

- The study of astrophysics (including cosmology) where challenges include: modelling the large-scale structure of the Universe; identification of the nature of dark matter; the formation and evolution of galaxies and clusters of galaxies, and the physics of the intergalactic and intracluster gas; studying fundamental physics in the early universe; the effects of dark energy and modified gravity on structure formation; parameter estimation and model selection using large astronomical datasets generated by cosmic microwave background experiments and large surveys; and theoretical simulations vital for efficient and effective use of STFC supported telescope facilities on ground or in space.

- The study of the ‘local’ astronomical environment – from the upper levels of the Earth’s atmosphere to the stars and gas clouds of the galaxy. Modelling solar and planetary magnetohydrodynamics and galactic dynamos; modelling star and planet formation.

- The study of the theory of the strong force using the numerical techniques of Lattice QCD. The objective is to increase the predictive power of the Standard Model of particle physics and other relevant strongly-interacting field theories by numerical simulation of lattice-regularised quantum field theory. This leads to the calculation of physical quantities which are essential to the interpretation of experimental observations.

- The study of particle interactions at the LHC. The LHC program is searching for new particles and processes at the high energy frontier including the origin of mass, dark matter and supersymmetry, evidence for new physics in quantum loops and the origin of the matter anti-matter asymmetry, and the quark gluon plasma. This requires handling multi-petabytes of event based data which must be reconstructed, stored, reduced and then analysed. In addition a large amount of simulated data is required.

- Specific challenges exist for radio astronomy which utilizes real time signal processing for correlations of signals from distributed instruments. This is needed for eVLBI, LOFAR, SKA which more and more use HPC machines. SKA will have a hugely increased Computing requirement. Other areas such as high energy gamma-ray astronomy using atmospheric Cherenkhov techniques, e.g. CTA, will also have very large Computing requirements.

- The study of the strong force in Nuclear Physics includes topics from hadron structure, hadron spectroscopy and quark-gluon plasma, over the synthesis of nuclei in stars to the study of the structure of nuclei. Nuclear Physics utilizes Grid computing in experiments like ALICE, PANDA and AGATA, but also other solutions, e.g. in the use of cloud computing at Jeffersonlab. Modern, large-scale nuclear physics experiments like CLAS12 will produce raw experimental data as well as MC simulations in the PetaByte range.

- The STFC National facilities including Diamond, ISIS and the Central Laser Facility (CLF). These facilities provide the large-scale experimental instrumentation needed to underpin a wide range of research activities in a number of key areas of science and technology, including physics, chemistry, biology and earth sciences. General themes include studies of the structure and dynamics of condensed matter, molecular and plasma systems using high intensity light and neutron beamlines. Computing close to the instruments is essential.

3. HPC Facilities – DiRAC

The majority of the HPC infrastructure for STFC science is now managed and provided through the DiRAC consortium which now federates many of the consortia which were previously separate. DiRAC currently provides for several areas of STFC science, including theoretical particle physics, theoretical astrophysics and cosmology. DiRAC has been very successful in creating an efficiently structured and cost effective infrastructure well suited to the needs of its science base.

DiRAC itself had made a project based submission elsewhere in this process and therefore we do not give details here. In summary DiRAC currently (in its second phase) comprises 5 HPC systems deployed at 4 sites. These are:

- (i) the UK's first low-latency Pflop/s system, designed for Lattice QCD calculations;
- (ii) 2 systems to manage high data rates and volumes,
- (iii) 1 system with high levels of communication feedback to model complexity, such as N-body simulations of the early universe
- (iv) 1 system with a 15TB single system memory image for the most complex calculations,

In its recent document [1] CAP noted that in the STFC HPC sector the requirements are most cost-effectively provided through a set of architectures appropriate to different problems. In other words "one size does not fit all". The requirements would not, for example, be cost effectively provided by a single architecture central facility. Also, as is typical of very many compute-intensive science areas, the nature of the task requires access to different scales of resource at different times during the lifecycle of the research.

Existing HPC consortia such as UKQCD or VIRGO have been praised in many reviews as providing a natural focus for science themes, providing the relevant architectures, whilst remaining agile. DiRAC has evolved this model further. Between its first phase starting 2009 and the second phase starting October 2012 it consolidated the UK's STFC-facing HPC resources from 13 hosting sites to 4. It currently has 4 funded technical staff to manage the facility's systems. DiRAC as a whole now provides a coherent federated HPC facility optimally matched to the requirements of its base scientific user community.

DiRAC has launched a competitive open access resource allocation process into which research groups from any area of STFC-supported theoretical science including Nuclear Physics may bid for HPC resources. The process has been designed to accommodate long, complex projects run by existing consortia, shorter pump-priming allocations accessible to users without necessarily any previous extensive HPC experience, and rapid access, time-limited projects predicated on new discoveries. It is also possible for a single project to have exclusive use of an entire facility for an agreed period. CAP regards this as a significant and much-needed reform to the provision of HPC facilities for the STFC community.

To date DiRAC has relied on HEIs and external sources for the bulk of its recurrent costs. The current phase of the project has an operations grant of £1.5M from STFC which will cover only the first year's electricity costs of the three-year project. To ensure DiRAC's sustainability it is vital to establish a stable funding model not just for capital (ie. hardware) but also recurrent support for electricity and support costs.

DiRAC users will also need access to specialist code development skills to assist with reducing run-times and improving code functionality, either via parallelisation or optimisation of existing code, or the design of new more efficient codes. This may entail an uplift in the number of PDRAs and PhD students working in theoretical science in STFC areas, with advanced IT skills beneficial to the wider UK economy.

From 2014, to retain international leadership it is anticipated DiRAC will require upgrades at the level of: (i) 10 Pflop low-latency capability machine; (ii) a 3 Pflop cluster with high communication feedback; (iii) a 1 Pflop 100TB shared-memory system; (iv) a 100PB Data Processing Facility

Recommendations:

- **(i) DiRAC is congratulated for consolidating STFC HPC resources into an powerful and cost effective infrastructure well suited to the needs of its science base.**
- **(ii) DiRAC facilities are essential for all STFC HPC based science and should be supported in the future at the same priority as the science which it enables.**
- **(iii) It is vital that a stable funding model for recurrent costs is established.**

4. HTC Facilities – GridPP

The computing infrastructure for the LHC is provided by the GridPP project.

GridPP itself has made a project based submission elsewhere in this process and therefore we do not give details here. In summary GridPP is centred upon the Tier-1 centre at RAL and federates Tier-2 facilities throughout the UK. The infrastructure comprises approximately 37,000 of HTC processing cores, 21 PBytes of disk storage and 10 TBytes of tape storage. It is connected to CERN by a 20 Gbit/s dedicated optical link, and is interconnected within the UK by SuperJanet-5.

The data is produced in pp collisions is sent to the Tier-0 centre at CERN, from where it is immediately replicated for data security. The data is then reconstructed across the global infrastructure to produced physics format data that is replicated and distributed. These data are then available for further reduction and repeated access for physics analysis. At the same time a continuous campaign for production of Monte Carlo simulated data is carried out at all Tier centres.

GridPP also provides resources for non-LHC experiments including T2K, MICE, the ILC and vestigial activity for the Tevatron and Hera experiments.

GridPP is itself a part of the Worldwide LHC Computing Grid (WLCG). GridPP is recognised as providing a very reliable and effective component of the WLGG, and this is due to the close and coherent organisation of the project combined with serious monitoring of delivery against targets.

It is perhaps appropriate to epitomize the success of GridPP by the most recent example, the discovery of (a particle compatible with) the Higgs boson. This was made possible by the WLCG computing infrastructure, and would have been impossible in without it. In fact the computing was cited explicitly several times during the announcment of the discovery at CERN (July 4^{th}) by experimental spokespersons and the CERN Director General.

In the next five years the computing requirements are set to increase substantially due to several factors. In the short term the LHC and the experiments have been very successful in producing and recording integrated luminosity and as a result all experiments are now recording substantially more data in order to maximise LHC exploitation. In addition experiments are planning for the impending long shutdown (2013-2014) and "parking" some data for later processing and or planning major re-processing campaigns. In the longer term, following the long shutdown the LHC will start at double the centre-of-mass energy and all experiments will have worked to ensure that they can run at the maximum data rate, and in particular to maximise the potential to study the Higgs boson characteristics. All of these factors place new burdens on the distributed computing system and the experiment software.

Recommendation:

- **(i) GridPP should be congratulated for successfully handling the LHC data leading to the recent results.**
- **(ii) The GridPP computing infrastructure is essential for LHC exploitation and should be supported in the future at the same priority as the science which it enables.**
- **(iii) Computing planning for the LHC has a long timescale, particularly now the Higgs has been discovered and the LHC approaches high energy running post Long-Shutdown-1. It would be more appropriate if GridPP funding (currently a 4 year cycle) could be extended to > 5 years.**

5. Astronomy

Radio astronomy has traditionally been the area with the largest computing resource requirements. It utilizes real time signal processing for correlations of signals from distributed instruments (eVLBI, LOFAR, SKA). This requires high bandwidth connections to HPC machines. Other branches of observational astronomy are beginning to place more significant demands on computing resources, for example large ground-based and space-based surveys (UKIDSS, Vista, PanSTARRS, Euclid, LSST) are producing, or will soon produce, enormous data volumes every year, with associated challenges to fully exploit these resources. New facilities under development such as the Cherenkhov Telescope Array (CTA) also place new demands. CTA data analysis will be dominated by the need to optimally remove the overwhelming hadron-induced background and to calibrate the atmosphere as a gamma-ray detector for the observation-specific and data selection-specific circumstances. This typically involves Monte-Carlo simulation and thus significant compute resources. CTA data analysis is very likely to adopt a distributed approach, modelled on the computing tiers familiar to the LHC Grid. Similar considerations apply to gravitational wave astronomy.

In the observational astronomy arena, the largest projects are international collaborations. Data processing, distribution and archiving take place in distributed systems, leading to relatively modest computing requirements for the partners who do not have the primary roles in these areas. The flip-side to this is that being able to offer such facilities can be key to securing major project roles. One issue for future UK participation in these projects is thus the availability of advanced storage and processing facilities within the UK. Where such facilities are available, they can be used to leverage leading project roles, an obvious benefit to the UK's scientific profile.

The approach taken for future provision of Computing resources for observational astronomy needs careful consideration. Some aspects, those critical to facility operations in particular, will continue to require dedicated provision. It is possible, however, that in other areas provision for those observational astronomy projects with the largest storage and compute needs could adopt a resource-sharing model similar to that successfully developed for DiRAC. This might be most appropriate for the large scale simulations required to support optimum data analysis strategies and instrument performance characterisation, as well as the more demanding data mining challenges. Pursuing this route would of course require a very careful analysis of the requirements and identification of the appropriate a resource sharing model consistent with project needs.

Recommendation:

- **(i) Future planning should be cognisant of the increasing Computing needs of Astronomy in the next five years.**
- **(ii) Provision of future resources to support observational astronomy should consider whether a resource-sharing model, similar to that adopted for HPC and HTC, could provide a workable and cost-effective solution.**

6. Nuclear Physics

In many ways, Nuclear Physics in the UK is following the same trends as Particle Physics. For example, contemporary collaborations in Nuclear Physics continue to increase in size, and while they have not reached the size of ATLAS or CMS, NuSTAR (760 collaborators), PANDA (450) and AGATA (400) are very large collaborations indeed. ALICE (1200) is a special case, as it is itself an LHC experiment. The amount of data generated by Nuclear Physics experiments now ranges from 10 TByte for a single average AGATA experiment to several PetaBytes for one year of data taking with CLAS12. This leads to much the same kind of requirements on data processing, storage and networking for Nuclear Physics as for Particle Physics.

One way to address these computing requirements has been the adoption of Grid Computing. The ALICE experiment at LHC is an obvious example which is already served by GridPP. AGATA (since 2007) and PANDA have also followed this route. This does not constitute a separate requirement for Nuclear Physics, but Nuclear Physics experiments should be taken into account when the future of Grid computing is considered.

Cloud computing is a different way to address the computing needs of a large experiment, and it is the way that the CLAS12 collaboration at Jeffersonlab has chosen. It combines economies of scale in the management, commodity hardware and virtualisation for service provision and encapsulating the application environment, while details of the physical resources remain hidden. The CLAS12 Reconstruction and Analysis Framework (CLARA) is an implementation of a service oriented architecture (SOA) which as such has a large overlap with cloud computing. For practical purposes it is important as this means that the computing resources will be concentrated at Jeffersonlab.

Theoretical Nuclear Physics is a particularly interesting field where future computing requirements are concerned. To state it simply: one reason why High Performance Computing has not played an important role in Theoretical Nuclear Physics yet, is that up to now it has not been 'high performance' enough. Lattice QCD is still working on the description and construction of individual hadrons. However, in the near future LQCD calculations of three- and four-baryon systems, such as the triton and alpha particle, will allow for the extraction of various three-body interaction parameters that are currently poorly constrained empirically. Of particular importance is the three-nucleon interaction, which has implications for nuclear structure and nuclear reactions. Internationally, 'nuclear' lattice-related methods are gaining in importance, like lattice-based effective field theories, using nucleon degrees of freedom as opposed to those of quarks that have made progress in calculating neutron matter. First ab initio calculations of ^{12}C states, most notably the Hoyle state, have been carried out. In the UK, in the near future, access to DiRAC will be important to Theoretical Nuclear Physics. We welcome the fact that DiRAC has already made itself available to this sector, and that a nuclear theorist serves on the resource allocation committee, Generally it can be expected that as HPC resources increase, Nuclear Physics will take up an increasing share of these resources in the future.

Recommendation:

(i) **It should be noted that the requirements of both Nuclear Theory and Experiment in the future will make a increasing call upon both HPC and HTC facilities**

(ii) **Consideration should be given as to whether further extending the access of GridPP and DiRAC to explicitly include experimental and theoretical Nuclear Physics would provide benefit to the science in an efficient manner.**

7. STFC Facilities.

We do not address STFC facilities in detail in this document as (i) much of the science falls under EPSRC as does the offline computing (remote from the data taking at the facilities) and (ii) this is the subject of a separate sub-panel.

For completeness we give a short summary of the computing requirement close to the facilities. The STFC facilities (ISIS, CLF, Diamond) schedule experimental time for short periods a long time in advance. During these periods computational facilities are required which are dedicated to, and close to the instruments. These are needed both to process and store data, as well as in some cases real time modelling. HPC resources are also used by the CLF (SCARF LEXION II). There facilities need to integrate with the users home institutional resources, some of which are now significant regional HPC systems.

We have noted in out previous document that the requirement is pseudo-real time, i.e. resources must be available at the time of using the facilities. This means that it is in general not possible to consider using, for example, shared resources on a remote large facility with a disjoint scheduling process.

8. Networks

Whilst the STFC network provision is not itself part of this programmatic review, networks are integral to the Computing infrastructure. For completeness we summarise the situation.

Networks are a first class and essential piece of the computing infrastructure needed to do the science. This includes everything from simple communications, collaborative work, data transfer, linkage of large computing resources, and in extreme cases the use of dedicated light path services. All wide area networking is provided through the national network provider, ja.net.

STFC has excellent relations with, ja.net. The network is commonly known as SuperJanet-5, or SJ5. These relations, built and maintained over many years, ensure that we can work in close cooperation with ja.net such that they are fully appraised of STFC science, and are therefore able to anticipate requirements. Ja.net view provision of world class networks for science as a high priority goal, and value the expertise which exists within the STFC domain. The relation is therefore mutually beneficial. This situation should not be taken for granted and it is very much in the interest of STFC science to proactively maintain these relations.

Recommendation:

- **(i) STFC should retain its excellent working relationship with ja.net (JANET) which ensures that network provision continues to enable the science.**
- **(ii) The provision for connectivity to the main science centres (including the STFC campuses, and principal science sites) should remain high priority.**

9. Issues

9.1. Recurrent Costs

The longest outstanding (unresolved) issue is that of recurrent costs for Computing.

At present the approach to recurrent costs is diverse and sporadic. In general STFC does not pay for infrastructure as SRF/MRF facilities. This means that recurrent costs for power in particular often come from host institutions. Often such arrangements are one-off leading to uncertainty for the future. Similarly operational staff often remain a hidden cost. Both aspects lead to planning uncertainty.

This situation has arisen in some measure as it has led to a very cost effective provision of computing as seen by STFC because Universities have paid for the power. But the situation is changing. Firstly as institutions develop further towards fEC principles, and as they install the capability to see the power bills form smaller units, then these costs become very visible. Secondly there is a prevailing presumption from on high that RCs (in general) should pay for use of facilities via SRF/MRF rather than fund equipment per se – University managements will more and more expect this leading tension.

The issue is particularly pertinent to DiRAC as mentioned earlier, but can be expected to pervade all Computing provision in the future.

Recommendation: It is essential that STFC embraces the need to develop a consistent and holistic funding model which deals with both the capital and recurrent costs of computing.

9.2. Software: sustainability, development and future opportunities.

Some of the STFC sectors depend upon mature code bases which present an investment in effort, and a corresponding user base, which is akin to that of other facilities or even an experimental apparatus.

These codes need maintenance and development as well as major “step” changes to embrace the evolving processor architectures. It is often forgotten that a relatively small investment in code optimization may lead to a substantial improvement in efficiency, which reduces the overall Computing resource requirement and even “reduces time to insight” - in both cases saving real funds. Nevertheless it remains relatively difficult to obtain resources for these activities.

There are also opportunities to be at the leading edge of software development. For example in the future exascale computing brings the need for radical changes to algorithms and software to address the need for fault tolerance, strong scaling and much steeper memory hierarchies. The steep increase in core count on all types of machine is already demanding innovative approaches to make efficient use of processors.

In addition the competition for access to international facilities is very intense and, in addition to excellent and ambitious science, will increasingly favour projects that have well organised software and support (at local and/or national levels).

It has been a continuing thrust of computing reviews over many years that the importance of these activities should be better recognised, and that support should be available for these activities, but in practice the science project peer review bodies are not practically open to such bids. STFC should consider mechanisms to address this important area, possibly including trying to secure e-Infrastructure funds.

Recommendation: Efforts should be made to identify mechanisms to support software sustainability and development.

9.3. The Value of Data

A more recent issue is that of Data Management Planning. There has been a significant change in the perception of the value of publically funded data as being a significant resource in its own right, which in many cases should be properly preserved for reuse. Policies pertaining to this exist at Government level, Research council level and Institute level.

STFC has a specific data management policy. It is expected by STFC that data from the science it supports is preserved, in some cases indefinitely. It is also expected that, in general, data collected with STFC funding is made openly available to the public, entailing professional curation expertise. Both of these aspects have significant cost and time implications, in particular any proactive approach to open data sharing. Such costs should be recognised openly and cannot be assumed to be absorbed as a minor incremental burden on the research funding itself. Training is needed for the new generation of researchers to use tools such as electronic research notebooks to maximise the advantages of sharing and collaboration and minimise the time and effort needed to achieve this.

CAP has provided a comprehensive commentary on the STFC Policy [2] and has had a very fruitful dialog with STFC in respect of guidelines.

There should be a significant effort placed on the generation of research tools to aid the management of data for better quality, provenance, access, re-use and this should be treated like the development of any other tool that facilitates and enables scientific research in STFC areas. The Computation Science Facilities in STFC (e-Science, Hartree, and facilities in general) are ideal places to develop and disseminate best practice and give the UK a world lead.

Recommendation: Resources need to be identified to implement the STFC policy on data preservation and sharing.

10. Science Roadmap

This table shows on the left the STFC Science Roadmap Challenges taken directly from the STFC site. On the right are shown the **major Computing infrastructures** which underpin the activities addressing these challenges. The **SJ5 network**, which is not shown explicitly, is essential to all of these activities. This nicely illustrates that the Computing infrastructure is essential to all Science Challenges.

Science Roadmap Challenge	DiRAC	GridPP	Astronomy Computing	Nuclear-Computing
A: How did the universe begin and how is it evolving?				
A:1. What is the physics of the early universe?	✔	✔	✔	
A:2. How did structure first form?	✔		✔	
A:3. What are the roles of dark matter and dark energy?	✔		✔	
A:4. When were the first stars, black holes and galaxies born?	✔		✔	
A:5. How do galaxies evolve?	✔		✔	
A:6. How are stars born and how do they evolve?	✔		✔	✔
B: How do stars and planetary systems develop and is life unique to our planet?				
B:1. How common are planetary systems and is ours typical?	✔		✔	
B:2. How does the Sun influence the environment of the Earth and the rest of the Solar System?	✔			
B:3. Is there life elsewhere in the universe?				
C: What are the fundamental constituents and fabric of the universe and how do they interact?				
C:1. What are the fundamental particles?	✔	✔		
C:2. What is the nature of space - time?		✔		
C:3. Is there a unified framework?	✔	✔		
C:4. What is the nature of dark matter?		✔	✔	
C:5. What is the nature of dark energy?		✔	✔	
C:6. What is the nature of nuclear and hadronic matter?	✔	✔		✔
C:7. What is the origin of the matter - antimatter asymmetry?	✔	✔		
D: How can we explore and understand the extremes of the universe?				

D:1. How do the laws of physics work when driven to the extremes?	✔	✔	✔	✔
D:2. How can high energy particles and gravitational waves tell us about the extreme universe?		✔	✔	
D:3. How do ultra-compact objects form, what is their nature and how does extreme gravity impact on their surroundings?	✔		✔	

11. Summary of Recommendations

(i) DiRAC is congratulated for consolidating STFC HPC resources into an powerful and cost effective infrastructure well suited to the needs of its science base.

(ii) DiRAC facilities are essential for all STFC HPC based science and should be supported in the future at the same priority as the science which it enables.

(iii) It is vital that a stable funding model for recurrent costs is established.

(iii) GridPP should be congratulated for successfully handling the LHC data leading to the recent results.

(iv) The GridPP computing infrastructure is essential for LHC exploitation and should be supported in the future at the same priority as the science which it enables.

(v) Computing planning for the LHC has a long timescale, particularly now the Higgs has been discovered and the LHC approaches high energy running post Long-Shutdown-1. It would be more appropriate if GridPP funding (currently a 4 year cycle) could be extended to > 5 years.

(vi) Future planning should be cognisant of the increasing Computing needs of Astronomy in the next five years.

(vii) Provision of future resources to support observational astronomy should consider whether a resource-sharing model, similar to that adopted for HPC and HTC, could provide a workable and cost-effective solution.

(iii) It should be noted that the requirements of both Nuclear Theory and Experiment in the future will make a increasing call upon both HPC and HTC facilities

(iv) Consideration should be given as to whether further extending the access of GridPP and DiRAC to explicitly include experimental and theoretical Nuclear Physics would provide benefit to the science in an efficient manner.

(viii) STFC should retain its excellent working relationship with ja.net (JANET) which ensures that network provision continues to enable the science.

(ix) The provision for connectivity to the main science centres (including the STFC campuses, and principal science sites) should remain high priority.

(x) It is essential that STFC embraces the need to develop a consistent and holistic funding model which deals with both the capital and recurrent costs of computing.

(xi) Efforts should be made to identify mechanisms to support software sustainability and development.

(xii) Resources need to be identified to implement the STFC policy on data preservation and sharing.

[1] STFC HPC Requirements looking towards future provision of HPC facilities, April 2011, http://www.stfc.ac.uk/resources/PDF/110418-CAP-HPC-Requirements.pdf

[2] Comments on the STFC Scientific Data Policy, Jan 2012, http://www.stfc.ac.uk/Resources/pdf/120100-CAP-SDP-Comments.pdf