Evaluation of the Service Level Agreement (SLA) between STFC's Scientific Computing Department and EPSRC for computational support of Collaborative Computational Projects and High-end Computing Consortia

Summary Report – November 2019

Contents

Executive Summary

The Service Level Agreement ("SLA") is an agreement between STFC and ESPRC for a programme of computational support provided by CoSeC, the Computational Science Centre for Research Communities, part of STFC's Scientific Computing Department. The support is available to a number of scientific communities funded by EPSRC, organised in the form of Collaborative Computational Projects (CCPs) and High-End Computing (HEC) consortia. While the nature of the support provided depends upon the needs of the communities, it can include development of algorithms and code, providing training, code porting and optimisation, maintenance and distribution. The programme that today goes under the name of SLA has been delivering support for 45 years and was first established in 1973 to support the first CCP.

The evaluation

This report summarises the main findings from the **evaluation of their Service Level Agreement (SLA) with the EPSRC,** commissioned by the Science and Technology Facilities Council (STFC) to Technopolis in December 2017. The main objectives of these evaluation were:

- To assess whether the intended aims and objectives of the SLA were being met
- To determine if the current structure of the SLA programme provided the most added value to the communities (providing an understanding of the efficiency and effectiveness of the delivery model of the programme)
- To provide input and evidence to the SLA management team, in anticipation of the EPSRC's Mid-Term Review of the Agreement in October 2018

The approach to the study comprised in-depth research on the SLA and its resulting impacts, an evaluation framework has been designed and a full evaluation accompanies this summary report. This included desk-based research to compile and analyse existing evidence; a programme of interviews with members of the SLA management and delivery team, high-level representatives and CCP and HEC community Chairs; online surveys of supported communities (195 respondents); the development of 9 deep-dives (based on desk research and interviews with community Chairs); and analysis to profile (describe, exemplify, quantify) key benefits and impacts where possible. This summary report draws its findings from the Evaluation Final Report.

The SLA provides unique and critical support to thousands of researchers efficiently and effectively

The evaluation found that the SLA **provides critical support related to a range of different aspects of software engineering**, including a range of code and method development activities to support the further development of research tools. It also largely reflects the needs of the communities that are supported and is making important contributions for the dissemination and training in the use of the tools that it helps develop.

The SLA is beneficial to both the communities it directly supports as well as to the wider UK research landscape. The current context of increased dependence on research software makes the SLA especially relevant and the long-standing history and longevity of the support provided means that the work delivered via the SLA is uniquely differentiated from, and complementary to, other newer forms of software development support. Moreover, through the activities of the SLA staff, the programme is connected to a range of software infrastructure facilities, networks and support mechanisms, including embedded CSE (eCSE) support, the Research Software Engineer Association and Fellows, the Software Sustainability Institute, Distributed Research utilizing Advanced Computing (DiRAC), and Centre Européen de Calcul Atomique et Moléculaire (CECAM) and Psi-K, a worldwide network of researchers working on the advancement of first-principles computational materials science.

The delivery model of the SLA is unique in providing centralised software development support to a wide range of research communities. The SLA programme is an effective system, the combination of its core functions and collaboration with academia supporting outputs with an estimated equivalent commercial value at almost £14m a year.

The SLA **implementation and delivery has improved** over the past years through improved resource management, and it was found to be efficient in light of the additional burden on supporting an increasing number of communities with a decreasing overall budget. The SLA enjoys high customer satisfaction, from both Chairs and wider community members, particularly with regard to the nature of the support available (77% satisfied) and the support delivered (62% satisfied).

Overall, the SLA makes an important contribution in supporting the development and maintenance of software that researchers can rely on in a number of different fields. The value of the programme is underpinned by its longevity, which has supported the consistent and sustained dedication by highly dedicated staff with both domain expertise and code development. This has enabled the accumulation of expertise, skills, long-lasting collaborations and partnerships, software, processes and a research-enabling ethos within the programme.

The SLA supports the research community to deliver more and better science

The SLA has contributed to an increase in the quality of research methods used by the communities it supports. Furthermore, the long-term and widespread use of these codes has been achieved through supporting greater sustainability and reliability of codes.

For a large part of the supported communities, the SLA has played a role in helping them **to secure further UK funding for their research activities**. About a third of those surveyed also reported an increased ability to secure further funding from abroad. Researchers engaging with the SLA often reported that they had achieved or expected to achieve additional scientific outputs such as greater research impact, and the creation of improved tools as part of their involvement with the SLA.

For almost all community Chairs, **loss of access to the SLA would result in severe loss of expertise and accumulated know how and would have a negative impact** in aspects such as the availability of training and the ability to implement and distribute code improvements to the community. The Chairs of communities supported by the SLA stressed that many of their flagship codes were developed with heavy reliance on SLA staff members. Developing and integrating new functionality and redistributing improvements to research codes in a structured process was found to maximise the impact of research software development.

The SLA supports the research community to capitalise on hardware and software infrastructure by providing efficient and sustainable codes and a trained user base

The work of the SLA in maintaining the continued sustainability of code ensures the work and investment over the previous decades is not lost. As such, the SLA is delivering support to further capitalise on the code development funded through a range of means, as well as ensuring this investment of resources is not wasted. Through engagement with community members, SLA staff support a perspective on good software engineering practices but also raise awareness of developments in other parts of the software landscape.

The work of the SLA in improving usability of codes enables greater uptake from users at the edges of these communities, by lowering the barriers to engagement. As such, the SLA supports capability building by both supporting a better understanding of software development and improving accessibility to a wider range of researchers.

Other coordination efforts supporting the internal communication of the communities were also appreciated especially by newer communities. As such, the SLA is also supporting the community to access compute resources available elsewhere. These coordination efforts, particularly with regard to developing the online presence of the communities, were also vital for HECs to help distribute access to ARCHER, improving its uptake.

The SLA supports capability building through training and providing unique career opportunities for software engineers

The training and workshops provided by the SLA were found to be highly relevant and cost effective for teaching computation methods and the effective application of SLA supported codes. The training and presence of the SLA also raised awareness, widening the engagement with the codes and enabling a broader community of researchers to apply these tools effectively and robustly.

The SLA also provides unique career opportunities for software engineers and long-term options for career development within STFC. The programme is able to attract and retain high calibre staff by ensuring that the roles involve cutting edge research activities.

Recommendations

Our headline recommendations for improvement of the SLA are to:

- Improve certain aspects of SLA monitoring and administration, in order to support improvements to external communication about the SLA and its achievements, and generate additional talent attraction
- Make a case for an increased scope and resources for the programme to address additional demand for support, future-proofing of strategic initiatives, and seizing opportunities for increased collaboration with STFC's unique large-scale facilities
- Use the evaluation framework and the monitoring and plan developed as part of this study as the backbone for further evaluation work under CoSeC

Extended recommendations presented in Appendix A2 of this summary.

Evaluation of STFC's Service Level Agreement with EPSRC

The Service Level Agreement (**SLA**) is an agreement between STFC and ESPRC for a programme of computation support provided by STFC's Scientific Computing Department. The support is available to EPSRC funded scientific communities, the Computational Projects (CCPs) and High-End Computing (HEC) consortia.

~£2.2m /year → ~22 FTE / year of support delivered → 9 CCPs 8 HECs → 3000+ researchers supported

What does the SLA support?

20 flagship scientific codes

The SLA is responsible for the development & maintenance of flagship scientific codes.

3 x more communities than 10 years ago

Over the past 10 years, the number of communities supported has increased almost threefold

Each FTE of SLA supports ~ 144 researchers

Survey responses

87% said computational support extremely important for their research work

Computational support and software engineering underpins a range of coded methods across different R&D areas.

52% said SLA played a role in securing further UK research funding

36% said SLA played a role in securing further international research funding

The SLA provides support, training and tools that underpins funding applications.

82% said SLA helped improve quality of research methods

The SLA provides sustained & consistent expertise and knowledge that underpins the development of high quality research methods

65% respondents said their work would be negatively affected without the SLA

Objectives of the SLA

- **Develop & maintain codes** to maintain relevance for current scientific challenges and hardware capabilities
- **Widen exploitation of methods & codes** through training & scientific collaboration
- **Support collaboration & coordination** of the communities to strengthen UK capabilities in line with EPSRC strategic goals
- **Provide career paths & professional development** for computational scientists & engineers focussed on software
- **Widen engagement with UK & international groups** who develop methods/software for computational research

Value of the SLA

SLA supports outputs with equivalent commercial value of est. ~£13.6m /yr

3,000 - 4,000 Training days → **£850,000/yr** Commercial value of SLA training

est. £3m / yr Estimate of the value of the SLA-related software

~2,500/yr Scientific papers by researchers supported → **est. £3.75m /yr** Value of SLA supported papers

30 - 50 SLA staff papers published per year

£6m/yr Average additional grant income generated for researchers with support from SLA

Context and rationale for the Service Level Agreement (SLA)

R&D depends on software

A strong national capability in computational science is an element of strategic importance for Research and Development (R&D). Indeed, most researchers use research software as one of the key research tools across disciplines, as many phenomena can only be studied through modelling and simulation, while analysis of experimental data is increasingly required to deal with ever larger datasets. The quality of the research outputs is thus dependent upon the expertise of those developing and using these research software tools. **EPSRC's Software as an Infrastructure strategy highlights the importance of software as a cornerstone of the research base within the UK and the SLA plays an active role in contributing to its wider objectives.**

Software as infrastructure

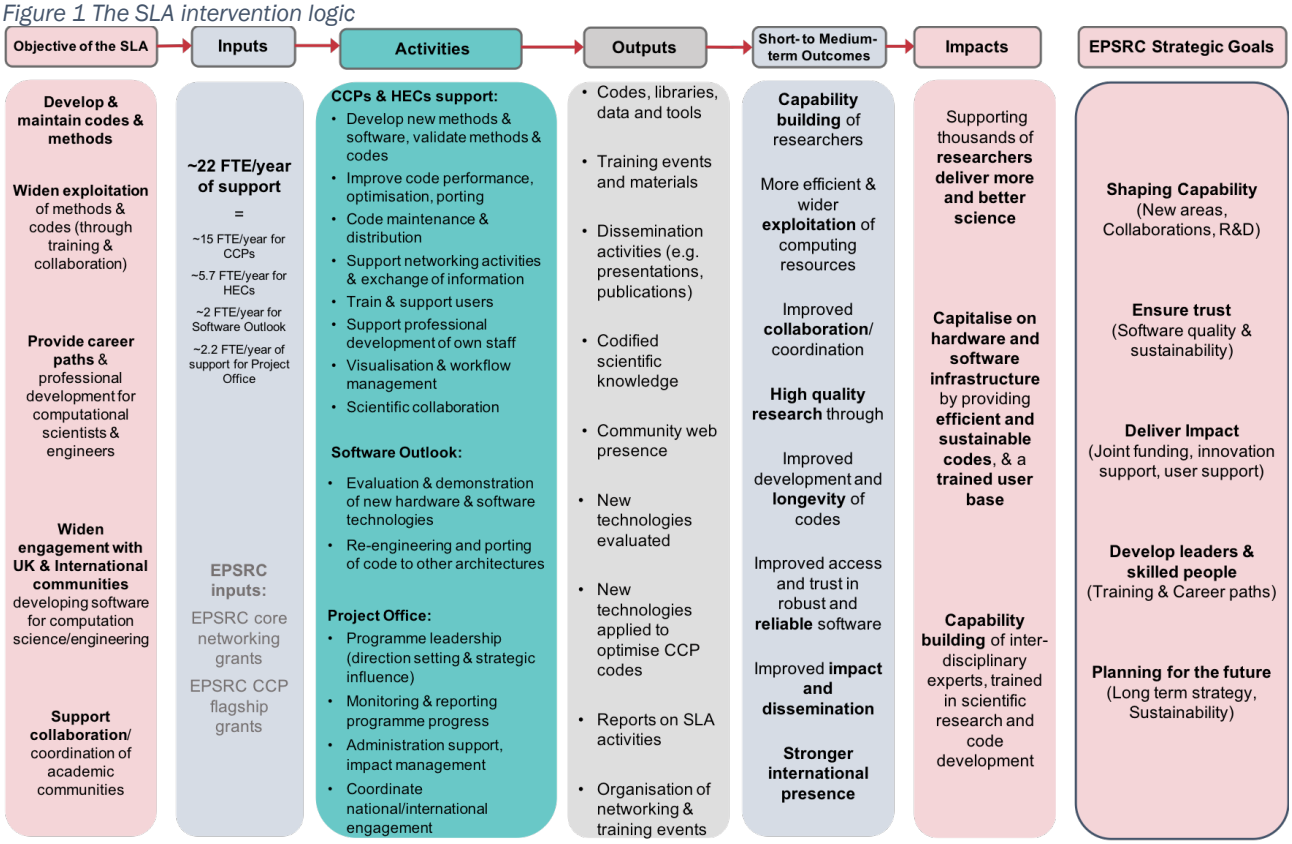
In the same way that hardware is acknowledged to need consistent investment due to maintenance and replacement costs, there is a sense of urgency in the research software development community to showcase the fact that current software for computational science and engineering is an infrastructure in its own right, with longer lifespans and further reach than that of hardware investments.

Given the ever-increasing suite of codes and the evolving nature of hardware, it follows that greater attention needs to be given to ensure software maintenance, optimisation and application support.

The Service Level Agreement (SLA) between STFC's Scientific Computing Department and EPSRC for computational support of Collaborative Computational Projects (CCPs) and High-end Computing Consortia (HECs) aims to address some of the challenges associated with developing software infrastructure in line with EPSRC's strategic goals.

The SLA aims to support the development of methods and software tools to enable the UK research community to do more and better science. It provides support to ensure the continued development and long-term maintenance of software codes and to nurture the communities by building synergies, human capital and coordination activities (Figure 1).

Figure 1 The SLA intervention logic



The SLA delivery model

The delivery model

The CCP concept dates back from the 1970s, making this one of the longest standing activities in the field in the UK. **The longevity of this mechanism of computational support is a key part of its strength and one of the main reasons the CCP and HEC model is well respected internationally and is envied by other countries seeking to develop computation support structures.**

EPSRC and STFC agree on a Service Level Agreement for a 5-year period, defining the key features of the support that STFC will provide to different research communities across EPSRC-relevant research fields (Figure 2). The communities eligible for support are those that are successful in CCP/HEC calls issued by EPSRC. In addition to SLA support, these communities also receive a core allocation of funds for networking and core activities (directly from EPSRC) and the CCP or HEC 'badge'. This allocation process accounts for the attribution of most of the SLA resources.

Given this relationship, the objectives of the SLA clearly reflect and align with the strategy goals of EPSRC's Software as infrastructure strategy, as highlighted against the SLA's intervention logic (Figure 1).

Overall, the current SLA has around £2.2m/year for around 22 FTEs to support a UK academic community of over 3,000 people.

SLA work strands

Development of theory, algorithms, and software resulting in the long-term, continued expansion and updating of the software programs, also including the consolidation of existing codes into a more sustainable community software package.

Maintenance, distribution, license management, dissemination and demonstration of software.

Management of scientific data through the development of visualisation and workflow management tools, database and archiving, and verification and validation activities etc.

User support and training: this ranges from organising and teaching at large events to individual support and training, help to organise training events and contributing teaching and training materials.

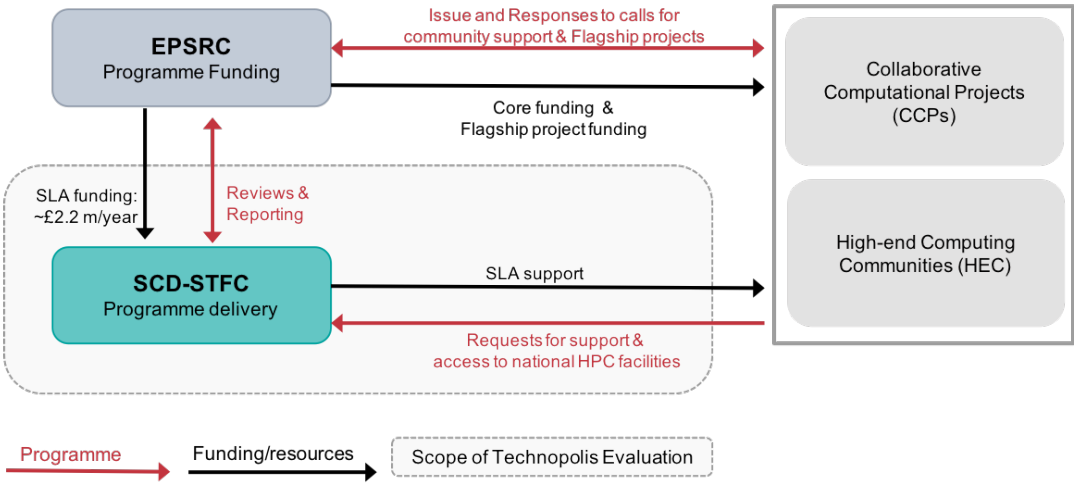
Collaboration on scientific projects is often offered to the community members who are not experts in the specific computational methods or software.

Porting, optimisation, and benchmarking for HPC and new architectures.

The SLA Project Office works to providing underpinning **coordination and networking support** for the communities and administrative support for activities e.g. such as scientific and executive meetings, and communities websites.

The Software Outlook activity focuses on the evaluation of new software technologies essential for the timely and cost-effective exploitation of current and near-future systems, while demonstrating how these can be applied to existing applications.

Figure 2 The SLA delivery model



Communities supported by the SLA

There were 9 CCPs and 5 HEC Consortia supported by the SLA and in scope of this evaluation (Table 1).[1] These numbers are set to rise as the SLA was brought into CoSeC during 2017, the Computational Science Centre for Research Communities, which encompasses the SLA programme as well as other activities.

Table 1 Communities currently supported by the SLA

CCP (Main code)	Title	Chair	Start Date	Community Size (members)	FTE/year
CCP5 (DL_POLY etc.)	The Computer Simulation of Condensed Phases – from atomistic to mesoscopic	Prof Neil Allan	1980	600	3.2
CCP9 (Questaal)	Computational Electronic Structure of Condensed Matter	Prof Stewart Clark	1981	450	2.4
CCP-Plasma (GS2 / BOUT++)	The Plasma-CCP Network	Prof Tony Arber	2007	150 (with HEC Plasma)	0.75
CCP-NC (MagResView)	NMR Crystallography	Dr Jonathan Yates	2011	60	1.3
CCPQ (R-Matrix, TNT, Quantics)	Quantum dynamics in Atomic Molecular and Optical Physics	Prof Graham Worth	2011	150	1.86
CCP-BioSim (FESetup)	Bio-molecular simulation at the life sciences interface	Prof Adrian Mulholland	2011	345 (with HEC-BioSim)	1.2
CCPi (CCPi CIL)	Tomographic Imaging	Prof Phillip Withers	2012	380	1.2
CCP-Mag (KKR)	Computational Magnetism	Prof Julie Staunton	2015	44	0.74
CCP PET-MR (SIRF)	Synergistic PET-MR Reconstruction	Prof Kris Thielemans	2015	80	1.2

HEC (Main Code)	Title	Chair	Start	Community Size	FTE/year
UKCP (CASTEP)	United Kingdom Car-Parrinello Consortium	Prof Matt Probert	1990	150	1
MCC (CRYSTAL, ChemShell, DL_POLY)	UK Materials Chemistry Consortium	Prof Scott Woodley	1994	464	2
HEC-BioSim (Longbow)	High-End Computing Consortium in Biomolecular Simulation	Prof Syma Khalid	2013	345 (with CCP BioSim)	0.8
UK-COMES (DL_MESO)	UK Consortium on Mesoscale Engineering Sciences	Prof Kai Luo	2013	150	0.6
HEC-Plasma (GS2, BOUT++)	Plasma High-End Computing Consortium	Prof Tony Arber	2013	150 (with CCP Plasma)	0.2
UK-AMOR † (R-Matrix)	UK Atomic, Molecular and Optical Physics R-matrix Consortium	Prof Jonathan Tennyson	2018	40	0.2
UKTC * † (Code_Saturne)	UK Turbulence Consortium	Dr Sylvain Laizet	2018	47	0.4
UKCTRF* † (SENGA+)	UK Consortium on Turbulent Reacting Flows	Prof Nilanjan Chakraborty	2019	47	0.5

* Working as a consortium under the CCPEngSci umbrella along with UKAAC and the UK Fluids Network. † Outside the scope of this evaluation

[1] 3 further HEC communities were out of scope of the evaluation (UKTC, UKCTRF and UK-Amor) as their SLA support began after the evaluation

The appropriateness and relevance of the SLA

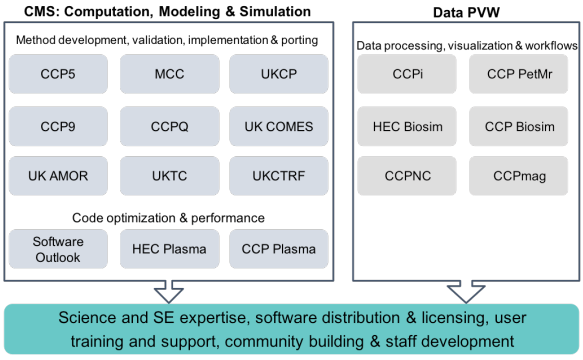
Delivery of SLA support is tailored and heterogenous depending on community needs

The support provided by the SLA is bespoke and highly tailored to the needs of each research community.

The SLA communities are very heterogeneous, while some are large and have been active for decades, others are much newer and comprise only a handful of principal researchers and small research groups.

While some CCPs/HECs share parts of the same research communities, completely new research topics such as tomographic imaging or plasma physics are now in scope of the support provided by SLA staff (see Figure 4). This comes to underline the increasing value of software as an infrastructure for new research areas over the past decade.

Figure 3 Different modes of SLA software support



The current delivery model is working well for the supported communities and SLA staff

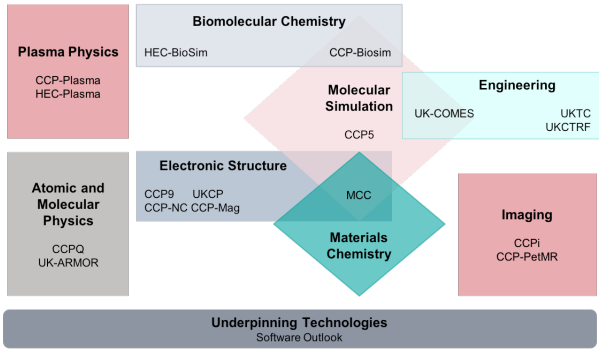
The support provided by the SLA is valuable for the communities it supports because it provides a unique form of expertise and particular support functions that would be challenging to source from elsewhere. In particular, the SLA served to fill critical gaps in software maintenance and optimisation that are not well addressed by other members of the software development landscape due to external pressures such as the pressure to publish. Over 80% of respondents were satisfied with the nature of support provided.

The centralised function of the SLA under STFC remit was found to be important and valuable as it allows staff to get involved in different projects and with different communities, allowing them to cross-fertilise ideas, generate their own expertise and share it with others. Indeed, project leads highlighted the importance of being in touch with the communities. The current delivery model is particularly conducive to this and contrasts with other modes of computational support (for example, via those research software engineers whose remit is to support researchers in a single university or institution). This is also a feature that was said to keep SLA staff engaged and interested in the work that they do, improving staff retention and creating a critical mass of wide-ranging expertise.

The SLA is well placed in the software support landscape, providing complementary support and strong links with other elements of the software infrastructure landscape

Moreover, a number of examples were provided to highlight how the SLA had formed beneficial connections with other aspects of software infrastructure landscape, whether that be through participating in funded projects (e.g. eCSE projects) or positions on management and advisory boards of different infrastructures (e.g. DiRAC and the E-infrastructure Experts Group). The SLA also provides links to other networks both nationally and internationally (e.g. Psi-K and CECAM).

Figure 4 Research fields covered by SLA-supported communities



The SLA is achieving its objectives across a large community of users. With high levels of customer satisfaction underpinned by responsive support and a unique delivery model, the programme is being managed and delivered well despite increased resource pressures.

The management and delivery of the SLA

The management and delivery of the SLA is working well, and has kept improving over time

The SLA management has implemented improvements in areas such as accountability, tracking of the work delivered, relevance of the work plans, improving project planning and shortening resourcing cycles. A 5-year planning cycle with yearly operational work plans, plus the structured feedback mechanisms through the project leads and the SLA steering committee have been highlighted as key features to this improvement.

In the past inefficiencies were mainly due to staff resourcing issues (vacancies, difficulty in filling jobs, long-term sickness, etc.) and there is now a wide acknowledgement that these problems seem to have been addressed for the most part through greater flexibility in resource deployment and management. Finding qualified personnel is still a challenge, but both the SLA management and the communities themselves are now tackling this much more proactively.

The SLA is responsive and relevant to community needs

For the majority of the supported communities, the SLA was addressing the research needs of the community. Notably, where these needs were not being met this was largely attributed to resource limitations.

The feedback mechanisms through the project leads and the SLA steering committee are key features that ensured the specific work plans with the SLA remained relevant to the community needs. As such, the SLA itself is not static but instead an evolving collection of expertise reflecting the needs and function of the academic community.

The thematic fit between the research interests of the communities supported and the focus on computational support were at the top of the reasons for engaging with the SLA.

Overall, the SLA enjoys high customer satisfaction, underpinned by good working relationships with SLA staff

Chairs generally have good working relationships with the SLA staff members and close cooperation in specific areas means that there is certain familiarity with STFC-SCD members, although increased face time with SLA staff and more community engagement were still requested.

On the other hand, awareness of the SLA by the wider research community seems to largely be through word of mouth. While the contributions of the SLA are strongly positive, their somewhat behind-the-scenes approach means there is limited awareness of the team's existence and its manifold support functions beyond the Chairs and more active contributors to CCPs and HECs themselves. This has been bolstered in the last few years by the increasing presence of SLA team representatives at steering and outreach meetings and activities with each of the communities. However, further strengthening this aspect will help with the attraction of new talent to the SLA. Inclusion of the support provided by the SLA under the name of CoSeC should improve community awareness of this support.

The current distribution of resources across the different work strands was found to be appropriate, through stakeholders did highlight challenges with regard to the availability of the most appropriate staff member. It was suggested that the supported communities could provide useful networks to support staff recruitment.

Current funding is only appropriate for the day-to-day maintenance of codes and current level of training activities, due to increased resource pressures

While the level of funding was found to be appropriate for the SLA's current functions, over the past 10-year period, the number of communities supported has increased almost threefold, and overall resources and funding for the SLA have decreased in real terms during the same time period.

Current funding levels are no longer sufficient to continue to support the more strategic, long-term view of software development

The reduction in available funds overall means that the SLA may not have enough resources to continue to take a more strategic and long-term view, especially in light of the larger communities that need support and the increased complexity and size of codes.

From this perspective, the scale and funding for the SLA has not increased commensurately in relation to the size of the need.

The outcomes and impacts from the SLA

The SLA has enabled impacts in a variety of different ways, from supporting more and better research, enabling the communities to be more efficient, expanding networks of researchers and on capability building and the development of skills.

Much of the value of the investment in the SLA stems from its longevity, which has enabled the accumulation of expertise, skills, long-lasting collaborations, software, ethos and reputation, therefore the impact analysis needs to be considered within this context.

The SLA supports a large community of users for a cost-effective budget

The SLA plays a central role in supporting **17 CCP and HEC communities** and their engagement with over 3,000 computational scientists across their respective communities; development and maintaining more than **20 flagship scientific codes**; delivery of 3,000-4,000 training days per year; is associated with the publication of around 2,500 annual papers.

With a budget of c. **£2m a year**, SCD delivers ~ 22 FTEs of highly qualified scientists and engineers specialised in scientific computing and with wide-ranging experience of the use and development of scientific software. Each FTE of SLA supports roughly 144 researchers.

Securing the equivalent volume of support through local external contractors would cost significantly more (estimated ~ **£4m**) but more importantly wouldn't offer the flexible and quality-assured support underpinned by decades of cumulative specific domain knowledge and experience provided by the dedicated SCD team. Such a contract would also lack the strong, collaborative relationships between the SLA and the scientific community it supports.

The SLA programme supports outputs with an estimated equivalent commercial value close to £14m a year

Much of the value of the investment in the SLA has come from its longevity, which has enabled the SLA programme to build up substantial expertise, skills, long-lasting collaborations and partnerships, software, processes, training experience, reputation, and a research-enabling ethos.

A summary of the monetised value of some aspects of the SLA are presented in the table below. Taken together, the equivalent commercial value of these combined benefit streams is close to £14m a year.

Overview of the analysis	Value added per year
Based on a conservative estimate of code commercial value at £1,500 annual subscription price.	An equivalent income from SLA related software of around **£3m** a year
The 'social value' of the scientific publications from the SLA is given a proxy of £1,500 within an open access model	This estimates the SLA's 2,500 'additional' papers as worth **£3.75m** a year
The volume of training delivered by the SLA has a much higher equivalent commercial value	The commercial value of the SLA training is **£850,000** a year
The SLA supports the research community in securing additional grant income	Additional grant income generated at an average of **£6m** a year

The SLA underpins a wide range of scientific research topics across the physical sciences and engineering.

Through its support of the CCPs and HECs, the SLA serves a number of fields with important sectoral applications. Examples of research aided by the programme include catalysis and energy storage; brain scanning technologies that may make it possible to identify and track the signs of dementia; engineering of matter into new and useful materials, and nuclear fusion which promises to generate large amounts of carbon-free energy in the future.

The SLA supports improvements to the quality and sustainability of codes

The SLA has contributed to an increase in the quality of research methods used by the community it supports, with over 80% of the surveyed community noting this type of impact, ensuring they are state of the art. Indeed, this type of impact was highlighted as being of the greatest value to the research community surveyed. The SLA staff provide a very particular perspective with regards to code development. They approach the problem, not only with a solid foundation of understanding of the research domain, but with the perspective of software engineering.

"Software development in academia is often a little 'hacky'. The expertise provided by the SLA guides the community to a more structured and more scientific approach to software." -Community Chair

For a large part of the community, SLA supported codes have helped them to secure further funding (from the UK and abroad) for their research activities. 52% stated that their involvement with SLA played a role in securing further research funding in the UK, while for 36% stated it supported successful applications to international research programmes. This additional grant income is estimated at an average of £6m annually.

Integrating new functionality and redistributing improvements to research codes in a structured process was found to maximise the impact of research software development.

The SLA provides a hub of expertise

Chairs of supported communities have also stressed that loss of access to the SLA would result in severe loss of expertise and accumulated know how and have a negative impact in aspects such as the availability of training and the ability to distribute code improvements to the community.

> *"The SLA gives us access to high quality research software engineering. SCD staff have useful skill-sets within STFC, work well, and also bring their shared knowledge of the codes and equipment. Because the SLA staff are embedded within a wider community, they help us connect with the wider community of software infrastructure" - Community Chair*

The SLA works to coordinate the research community and prevent them from "re-inventing the wheel"

In terms of cost-effectiveness the SLA is important for preventing the research community from "re-inventing the wheel," building a foundation upon which further development can occur and preventing development activities within the communities from being lost. This is a key differentiator between the SLA and other forms of software engineering currently undertaken at research institutions. From this perspective the SLA provides good value for money, although the benefits are challenging to quantify, due to the position of SLA staff as deeply embedded within the communities they support.

The SLA supports exploitation of long-term investments in software

Furthermore, the work of the SLA in maintaining and ensuring the continued sustainability of code ensures the work and investment over the previous decades is not lost. As such, the SLA is delivering support to further capitalise on the code development funded through a range of means, as well as ensuring this investment of resources is not wasted.

Key pieces of software developed by some of the long-standing communities supported by the SLA are now being taken for granted by entire fields of research. Substantial capital investments in software built over long periods of time could be lost if the associated development efforts and maintenance stopped.

The SLA helps the research community to capitalise on existing hardware, software and research infrastructure

The SLA is an important element of the UK's software infrastructure landscape, complementing other software development support but also serving to provide a bridge to other research groups and facilities. The SLA provide valuable contributions to code development and efficiency, porting codes for other hardware platforms and supporting users to effectively use HPC through training and software tools (e.g. LONGBOW). In doing so, the SLA supports researchers to use national and local hardware structure. Based on a conservative estimate of code commercial value at £1,500 annual subscription price, the SLA supports and provides software with an equivalent income of around £3m a year.

The SLA also supports researchers to access and capitalise on the existing software infrastructure. This includes collaborations with RSE fellows and other RSEs, as well as accessing a wide network of collaborators world-wide (e.g. through Psi-K and CECAM).

Collaboration with and application of SLA codes within the Hartree Centre and the large experimental facilities such as Diamond Light Source, ISIS and the Central Laser Facility, provides further connections to industrial and experimental research communities, capitalising on STFC and government investments.

> *"The consequence of losing the SLA would be quite severe. The alternative of hiring a post-doc as a short-term appointment would lose the continuity aspect that is key to the SLA. Software maintenance is a very important job, and few post-docs would have the experience to write code and user guides properly to such a high professional standard." -Community Chair*

The structure of the CCP programme is internationally respected and SLA supported tools have facilitated cooperation between the UK and international research community

The development of such widely available research tools also facilitates wider use outside the UK and other countries with strong computational science research acknowledge the value of the structure of the CCP programme. The value of this wider programme of computational support is in part attributed to the longevity of support provided and the sustained activity of the communities.

The international reach of the tools also facilitates greater interoperability and cooperation between UK based researchers and the international scientific community.

The SLA provides high quality training in a cost-effective way

The SLA supports cost-effective capability building through providing reported high-impact support for training events and workshops. The SLA training programme is mainly focussed on teaching how to use methods and codes that are developed by the programme. The available data shows that this is a cost-effective way of delivering training to large numbers of researchers. The volume of training delivered by the SLA has an equivalent commercial value around an estimated £850,000 per year.

Through engagement with community members, SLA staff support a perspective on software development more in line with software engineering but also raise awareness of developments in other parts of the software landscape. Indeed, over 70% of survey respondents noted that this training had an impact of their expertise and knowledge.

> *"For younger researchers and PhD students, the training is an important step. It helps them to understand what's behind the software and optimise its use. Because major software these days is very sophisticated, more training is necessary for them to be able to use it effectively." -Community Chair*

The work of the SLA also lowers barriers to engagement

By improving the understanding of software engineering practices, the SLA not only improves software development standards but also lowers the barrier to engagement. Improving the usability enables greater uptake from users at the edges of these communities which also lowers the barrier to engagement. As such, the SLA supports capability building by both supporting a better understanding of software development and improving accessibility to a wider range of researchers.

The SLA provides a unique career path for software engineers

The SLA provides unique career opportunities for software engineers and long-term options for career development within STFC, although talent attraction and onboarding have been challenging historically, given the bespoke nature of support needed and competition for talent from the private sector. Despite this, the SLA does have low levels of staff turnover due to the unique interdisciplinarity of the role and the opportunities to deliver academic research in parallel.

Consequences of the loss of the SLA

The loss of SLA support would have a significant negative impact, the quality and the quantity of research undertaken in the field would suffer (both nationally and internationally), and training would not be readily available for Academia, industry, and the next generation of researchers in the field.

For 65% of survey respondents, the loss of the SLA would have a negative impact on their research. In particular, 20% of respondents noted they could not deliver their research at all without the contributions of the SLA, almost half of whom were research group leaders. It is therefore likely that negative impact of the absence of the SLA would also extend to their research groups too. 19% of respondents noted that their research would have been delivered at a slower rate and a further 26% would have experienced reduced scale or scope on their research.

Substantial capital investments in software built throughout long periods of time could be lost if the associated development efforts and maintenance stopped. Therefore, the community is also eager to stress the negative impacts of loss of SLA support. If that were to happen the quality and the quantity of research undertaken in the fields supported would suffer (both nationally and internationally), and training would not be readily available for industry and the next generation of researchers in the field.

Appendix

A1 Cost-effectiveness analysis

The evaluation questions relating to value for money are the following:

- From the evaluation questions of the present study:
  - To what extent does the SLA provide value for money?
- From the questions of the EPSRC mid-term review:
  - Do the outputs from the programme represent good value for money?
  - Community coordination: does the Programme enable efficiencies/savings through its community co-ordination activities?
  - Suitability of the centralised model of funding and comparison to alternatives. What are the advantages and disadvantages?

Assessing value for money is in the first instance a question about efficiency and the relationship between the quantum of outputs produced as a result of the resources used to produce them. It is also a question of effectiveness, and the relationship between the resources consumed and the outcomes realised as a result of that public expenditure. It should also assess whether the realised costs and benefits are in line with expectations, as long as there is sufficient data available.

On the first point (efficiency), the SLA represents good value for money delivering a large volume of wide-ranging support and playing a central role in:

- Support to **11 standing committees** and their engagement with several thousands of computational scientists across their respective communities
- Development and maintenance of **23** scientific codes, and management of access to those codes
- Delivery of 3,300 plus training days in 2017/18
- Associated with the publication of around 2,500 articles and papers each year
- Associated with the generation of around £6m in annual research income (Fs)

With a budget of c. **£2m a year**, SCD delivers around 22 FTEs of highly qualified scientists and engineers specialised in scientific computing and with wide-ranging experience of the use and development of scientific software. Securing the equivalent volume of support through local external contractors would cost at least as much and would fall short of the flexible and quality-assured support provided by the dedicated SCD team. Moreover, local provision would not be able to benefit from the critical mass and weight of experience available across SCD and would also be at greater risk of being cut or redeployed as part of wider efficiency savings. A central, coordinated service is the right solution. The EPSRC would need to appoint a very good IT consultancy to match the kind of service offer from STFC, and we estimate the annual price would be closer to £4m a year for 20 FTEs with a much more stringent management of requests for services and resourcing.

On the second point (effectiveness), we found no single robust means by which to estimate the global monetary value of the service's outcomes for the research community or society more generally and have therefore used a variety of techniques to estimate the equivalent commercial value of its individual core functions (from code development to training).

SLA codes are used commercially in some areas, so we attempted to arrive at an estimate of the total income that might be generated each year, were all of the licences for the many SLA-related software codes valued at the kinds of prices one finds in the commercial marketplace. We struggled to find a sufficient number of pricing data to allow us to run this 'revealed preference' analysis with confidence, however, we did find some market intelligence, with figures ranging from £1,500 a year for one SLA-derived code up to $22,000 for another. Being conservative, using the £1,500 figure as a proxy for an annual subscription price for the 2,000 active users of the CCPForge – access to multiple codes – would produce an equivalent annual income of around £3m, which is substantially greater than the annual cost to EPSRC. We were not able to estimate a value separately for the efficiency gains realised by the scientific community working together to define common standards, however, our interviews and surveys suggest this does avoid multiple and potentially inconsistent development trajectories.

We also sought to compute a value for the additional annual output of scientific publications, using typical Article Publication Charges (APCs) as a means by which to arrive at a proxy for the average social value of any scientific article. Notwithstanding the limitations, using the £1,500 average price proposed by Professor Dame Janet Finch in her review of open access publishing would mean the SLA's 2,500 'additional' papers are worth £3.75m a year in APC-equivalent payments.[2]

SLA secures further funding for community researchers and their groups, estimated at £6m annually. The resources dedicated to scientific collaboration also result in securing further funding for community researchers and their groups. Data from the SLA Management on further projects from the research communities supported over the past two years comprises 57 projects totalling more than £35m, with SLA staff involved in 80% of them as named investigators. The SLA management team estimates that this means the SLA supports the research community in securing around £6m in annual grant income (£35m in total income for two years and assuming an average term of 3 years: 35/2/3=5.89). This income is unlikely to constitute new money coming into the UK academic systems inasmuch as most of those grants were linked to established UKRI programmes and would have been won by other UK-based researchers. The one exception is for an unknown minority of new projects that relate to international awards (e.g. Horizon 2020 contracts) or new commercially-financed studies.

Lastly, we used the volume of training delivered – just one of the five SLA core functions – as a basis for estimating an equivalent commercial value (at an average of £250 per person per day of training delivered) of at least £850k, or 40% of the annual SLA budget.

Taken together, the equivalent commercial value of these combined benefit streams is close to £14 million a year.

Table 2 Summary of monetary value of various SLA services

Type of 'value for money' analysis conducted	Results - Value added per year
Estimate of the value of the SLA-related software – estimated at an annual income based on a conservative estimate of code commercial value at £1,500 annual subscription price for the 2,000 active users of the CCPForge.	An equivalent income from SLA related software of around £3m a year
The 'social value' of the scientific publications from the SLA is given a proxy of £1,500 within an open access model	This estimates the SLA's 2,500 'additional' papers as worth £3.75m a year
The volume of training delivered by the SLA has a much higher equivalent commercial value	The commercial value of the SLA training is £850,000 a year
The SLA supports the research community in securing additional grant income	Additional grant income generated at an average of £6m a year

A2 Recommendations for the SLA

This section presents the recommendations from the SLA evaluation, focusing on the process evaluation, monitoring arrangements and overall suggestions for strengthening and future-proofing the programme.

Recommendations from the process evaluation

While there were no major issues with the general delivery structure of the SLA, the following main ideas were flagged repeatedly in our interview programme as potential areas for improvement in the programme delivery:

- **Improved internal communications** between the different SLA activities (the project office, the software outlook team, the project leads within STFC, and the communities' Chairs). Although this has been flagged as an area where there is steady improvement, interviewees pointed out that higher levels of awareness of what is on offer under the umbrella of the SLA could improve its effectiveness. Some also expressed an interest in further news or information about the activities being undertaken to support other communities within the programme. This was also echoed by survey respondents who stressed the need to improve internal communication within their own research communities and with STFC.

[2] Report of the Working Group on Expanding Access to Published Research Findings (2012) Accessibility, sustainability, excellence: how to expand access to research publications.

- **Improved external 'branding of the programme**, to improve poor visibility of the SLA brand' in general. Significant efforts were made in the past to ensure the SLA is tightly connected with the CCP programme, so much so to the extent that no difference can be defined by most participants between the two. This poor branding and visibility however placed limitations on the degree to which the SLA has been able to effectively engage with different research communities and with other elements within the research software ecosystem. Some interviewees noted that they expect that the emergence and continued development under CoSeC will serve to address some of these issues, by giving CoSeC a more recognisable name in the wider research community.
- **The emergence of the RSEs will require greater collaboration and coordination** in the future, and it is important to understand how the two models will complement each other.
- There is a **need to define career pathways and to support progression of individuals** specialising in research software: While the SLA provides a unique opportunity in which software engineers are able to specialise and hone their craft, there is also a need to ensure strong career path options to help retain talent within the SLA team. Chairs noted that finding the right people is a vitally important aspect of the service that the SLA provides and past experiences with staff leaving and hard to fill vacancies are still an element of concern to the continuity of services to the CCPs and HECs. Notably, a couple of Chairs highlighted the successful recruitment of SLA staff through their involvement in the process. As such, greater engagement with the relevant communities may help address some of these resourcing challenges.

Further to this, it was also highlighted that the higher levels of demand of SLA staff time left less time for staff development.

- The fact that the SLA is delivered centrally by STFC opens the door for **more collaboration with STFC's large-scale facilities,** and some feel this opportunity is still under exploited. Interviewees have pointed to the existence of large communities of researchers around these big facilities that are not eligible for SLA support but would benefit greatly from it. The current delivery model for the SLA may make it difficult to address this part of the research community and we note that this would have resource implications to an already stretched team.

Recommendations from the monitoring plan

This M&E plan does suggest that the SLA will need to launch some additional data collection and monitoring activities, in order to make a better use of the developed evaluation framework in future evaluation exercises. These include:

- **Feedback on training activities:** Given the role of the SLA project office and SLA staff in providing support for the delivering of training activities, they are well placed to coordinate participate feedback on this training. We would recommend a short questionnaire (maximum 7 questions) to be sent out to training attendees following the event to gather their view of the value, the relevance and the delivery of this training. This will naturally also support the communities further in their own awareness of the communities' interests and needs going forward.
- **An impact or 'business' case for each of the communities supported:** Though the annual reports collated by the SLA provide detail of the work delivered by the SLA staff members in support of each of the communities, these reports are highly technical and rarely clearly outline the impact or the added value of this contribution. Interviews with Chairs were often able to speak to the value of the overarching work rather than individual work packages. As such, there is an opportunity to utilise and build upon the existing reporting processes to build a more thorough and continuous processes of collecting information regarding the impact/improvements attributed to the work of the SLA staff. This too may also feed into the generation of impact case studies with the SLA project team. The feedback provided by the Chairs of each community during the EPSRC mid-term review self-assessment is a good starting point for providing a big picture narrative of the aims of each community and how the SLA supports them in that.

In carrying out the research, we have also identified several other developments that may provide a platform for strengthening future monitoring and evaluation activity:

- **Expand the proposed monitoring arrangements to all the CoSeC communities:** While this framework has been developed to evaluate the support provided to the CCP and HECs supported under the SLA, it is likely that future evaluations may also include other supported communities that lie outside the remit of the SLA, though still within that of CoSeC. As such, we recommended that these data collection activities are ex-

panded to also include these other communities as well. This will help in the future should CoSeC be interested in evaluating the entire programme of activities but will also provide a coherent and standardised format for all data collection activities across the group.

- **Conduct future monitoring and data collection under the brand of CoSeC:** The more recent efforts to support a more clearly defined role and brand of CoSeC has raised awareness of the CoSeC brand. As such, future data collection and evaluations could be better placed under this more widely recognisable name.
- **Standardise as possible the mentions of 'Supported by SLA/CoSeC':** Though it was acknowledged that efforts are already in place to encourage greater referencing to the SLA support in publications, it does not seem that there is a standardised process or means of doing this. We recommend that a standard referencing processes be defined by CoSeC and actively encouraged across the communities at all opportunities. As such, this would allow for improved view of the academic impact of the work conducted by the SLA.

Final recommendations

The SLA is beneficial both to the communities being supported and to the wider UK research community. Nevertheless, there are a number of improvements that could be made to improve its delivery and impact.

While the contributions of the SLA are strongly positive, their somewhat behind-the-scenes approach means there is limited awareness of the team's existence and its manifold support functions beyond the Chairs and more active contributors to CCPs and HECs themselves. This may not be problematic; however, it does suggest that the evident benefits from software harmonisation or specialist training in research software developed with SLA support could be enhanced further.

There are also opportunities for **greater coordination and coherence of the support provided** to the different communities, this has been acknowledged by both community Chairs and SLA staff, and could be supported by stronger internal communication activities. Project leads are well placed to deal with supporting these internal communication efforts.

The current level of funding overall means that the SLA may not have enough resources to take a more strategic and consolidated long-term view of the support it is offering and end up focusing on day to day activities. Increasing the visibility of the SLA across the different communities would no doubt increase demand for support of one kind or another, which could bring its own difficulties in terms of service quality and priority setting (in the absence of further investment and the expansion of the SLA function). We recommend that the SLA leadership take the current situation as an opportunity to clarify their ambitions to EPSRC and to **make the case for an increased scope and resources for the programme** should they see this as an appropriate direction. There is currently a demand to support communities not currently covered by the SLA (e.g. astronomy) and to provide more resources to existing communities.

The project office would benefit from the use of more **standardised 'marketing' material** on what the SLA is and what its achievements are, in order to support activities to generate greater awareness of the SLA across the communities. For example, we have seen this from the RSE programme being a much more recent initiative. The reorganisation of activities under the CoSeC brand and umbrella will no doubt contribute to wider brand recognition of the SLA, and there has already been an effort from SLA staff to attend community conferences and training events to deliver the SLA pitch. The materials produced during this study and in the self-assessment report presented to the EPSRC's mid-term review of the SLA can be used as a repository of information to disseminate. The production of public year-end reports can make the programme and its achievements more widely known.

There is a need to **standardise and strengthen the collection of monitoring information** about the communities being supported by the SLA, both from the communities' perspective as well as about the wider research community that is being targeted. Considering current GDPR regulations, we would recommend that distribution lists are updated and populated with information that allows categorising the communities and beneficiaries along the main classification dimensions that can facilitate further monitoring and evaluation work. The lack of information about the constituencies that the SLA supports can hinder strategic activities to increase capacity building and impact of the work carried out. For example, understanding the proportion of early career researchers targeted by the different communities could be used to advertise future employment opportunities for SLA work and wider engagement with RSEs and other experienced actors.

For this same reason, **the evaluation framework and the monitoring plan developed as part of this study can serve as the backbone for further evaluation work**. The evaluation framework details the intervention logic model of the SLA, the main evaluation questions, the methodologies, and the data collection tools used in this evaluation, while

the M&E plan details the desired KPIs from an evaluation standpoint, methodological challenges, and possible planning of future evaluation exercises. While we expect that these will be adapted by the SLA team to best suit their purpose, **we recommend that this evaluation mindset is maintained, and that information compiled in this study and the EPSRC mid-term review recently conducted is kept as baseline information for future evaluation work**.

Finally, there is a need to **work together with EPSRC to align the SLA objectives and phrasing across the different materials and SLA legal documents**. Ideally, the SLA legal agreement, call documents, SLA/CoSeC websites and programme outreach documentation should be aligned with regards to the description of objectives, activities, ultimate goals of the SLA, and division of labour (in broad terms) between STFC and the communities. This would allow to clarify some of the confusion that still exists around the objectives and role of the SLA, the type of support that can be delivered, and the differences between CCPs and HECs. The distinction between the CCP and HEC concepts could be better outlined or dispensed with, depending on which strategic direction the SLA leadership wants to pursue.