High Performance Computing for Industry: The Case of the Centre for Innovation
The SES consortium is the most powerful cluster of research intensive universities in the UK, harnessing the insights and resources of the Universities of Oxford, Cambridge and Southampton, Imperial College London and University College London.

UCL Advances is affiliated with UCL Enterprise, which provides UCL's structures for engaging with business for commercial and societal benefit. UCL Advances is responsible for promoting a culture of entrepreneurship on campus and engagement with entrepreneurs and small businesses beyond UCL's boundaries.

EPSRC is the main UK government agency for funding research and training in engineering and the physical sciences.
This report was conducted by University College London (UCL Advances and UCL Information Services Division) as part of the “Centre for Innovation - SME Engagement Springboard” project, funded by the EPSRC, which focussed on increasing the impact of high performance computing in industry.

February 2015

Contributors:
Riam Kanso, Timothy Metcalf, Daniel Hulme, Alastair Moore, Timothy Barnes and Clare Gryce.

© 2015 – All rights reserved
Executive Summary

Background

High Performance Computing in Industry: Barriers to Utilisation

Case Studies

Case Study 1: An SME’s Perspective on Access to GPU facilities
Case study 2: Using a GPU Platform for Materials Modelling
Case study 3: Drug Discovery using HPC in the Pharmaceutical Sector
Case 4: Software Licensing Models from a Software Company’s Perspective

Skills and Awareness

Cloud Computing: An Alternative?

Cost and Flexibility

Expertise

Scale

Security

Case Study Findings: Lessons Learned

Conclusions

Appendix: Cfl Utilisation Metrics June 2013 - December 2014

EMERALD

IRIDIS
EXECUTIVE SUMMARY
High Performance Computing simulates, models, predicts and solves highly complicated scenarios and problems. It dramatically reduces the time and expense involved in developing new products and services, and the process of translating research into innovation. It is already being used extensively in a wide range of research areas, from climate change prediction, drug testing and development to the design of new materials, such as environmentally friendly catalysts. Sectors that are benefiting are just as diverse: from automotive and aerospace to biotechnology and finance.

Established in 2012, the Centre for Innovation (Cfl) is a shared High Performance Computing facility. It is jointly owned and operated by the Universities of Bristol, Oxford, Southampton and University College London. It was set up with £3.8m of funding from the Engineering and Physical Sciences Research Council, and is part of the Science and Engineering South (SES) Consortium.

Making the most of the potential that HPC has to offer is highly dependent on the level of industry access to research and development facilities, such as those offered by the Cfl. It also relies on other factors, such as increasing awareness of what HPC has to offer and ensuring that there are enough people with the necessary skills and knowledge of HPC and how it can be used by industry.

This report looks at industry access to the Cfl’s e-Infrastructure services. There is a particular emphasis on small and medium enterprises (SMEs). The experiences of four SMEs that have used the Cfl facilities provided case study material. We learned some important lessons about the Cfl’s engagement with industry and more generally about industry access to HPC.

Much of this evidence lends further weight to, and often mirrors, the barriers identified in the report commissioned by the Department for Business, Innovation and Skills in 2012, Developing e-infrastructure in the UK’s engineering and manufacturing industries.

The barriers to industry access identified and issues highlighted are often interlinked and inevitably overlap:

- **People, processes and tools**: The Cfl would benefit from an increased investment in further developing internal expertise, in particular the training and recruiting of more staff with specialist knowledge of business processes and tools.

- **Technical incompatibility**: The software and middle-ware required by some industrial users can often be incompatible with the HPC system.

- **Different policies and infrastructures**: From firewalls to inflexible policies, some industrial clients have systems in place that can make access to HPC resources difficult. Such clients need to be encouraged to configure their systems for flexible access to HPC.

- **Meeting different service needs**: The Cfl needs to be able to accommodate different types of industrial clients with flexible, tiered levels of support, service and access.

- **Meeting different financial needs**: On-demand cloud resources and software pay-as-you-go models could reduce the risk of some SME’s being unable to take promising projects forward due to limited finances. Some SME’s find software licences prohibitively expensive. Many software companies that offer HPC packages are reluctant to change their software licensing models in order that the needs and often limited finances of SMEs can be
accommodated.

- **Heightening awareness of HPC:** More SME’s would invest significantly in HPC if they were aware of the benefits. More actions could be undertaken to promote this message to industry in general. One approach could be to include HPC sessions in ICT business-orientated courses.

- **Security:** Concerns about security are present on both sides of the HPC business relationship. Restricted access models clearly offer strong security, but can also hinder fast-paced work. There is a need for well-defined policies and technology standards between the client and provider.

Conclusion:

From innovative drug development to motorsport engineering, the case studies outlined in this report show a promising landscape of innovative projects that can be carried out with HPC.

However, the case study evidence also shows that the CfI, as it stands, is not able to provide the required level of service for industry to be able to make full and effective use of the HPC services on offer.

We feel that this is having a detrimental effect on realising the full potential of HPC, which is already proving to be such a vital part of research and innovation in the UK. For this potential to be realised there needs to be a significant level of ongoing and further investment in the CfI’s resources and personnel.

In addition, because CfI was set up as a computing resource for academics, more could be done to make industry aware of what HPC has to offer. One way of bridging this divide would be to invest more in the provision of HPC academic advice and guidance services for industry.
BACKGROUND
The Centre for Innovation (CfI) is a shared High Performance Computing (HPC) facility jointly owned and operated by the Universities of Bristol, Oxford, Southampton and UCL. The CfI was established in 2012 with £3.8m of funding from the Engineering and Physical Sciences Research Council (EPSRC). CfI is a part of the Science and Engineering South (SES) consortium, the most powerful cluster of research-intensive universities in the UK, harnessing the insights and resources of the Universities of Oxford, Cambridge and Southampton, Imperial College London and UCL.

The CfI was set up as the largest of several regional or ‘Tier 2’ HPC facilities intended to support and promote increased levels of industrial engagement and collaboration based around the use of e-Infrastructure and short-term proof of concept activities by providing access to its HPC services.

“e-Infrastructure refers to a combination and interworking of digitally-based technology (hardware and software), resources (data, services, digital libraries), communications (protocols, access rights and networks), and the people and organisational structures needed to support modern, internationally leading collaborative research be it in the arts and humanities or the sciences.”

Following a major review of the UK’s e-Infrastructure in 2011/12, it is anticipated that future funding for this area across academia will be accompanied by a drive to dramatically increase industrial uptake, primarily via regional e-Infrastructure ‘hubs’ such as the CfI.

The CfI comprises two HPC systems: EMERALD (www.cfi.ses.ac.uk/emerald/), the largest GPGPU cluster in the UK, and IRIDIS (www.cfi.ses.ac.uk/iridis/), a CPU cluster.

The CfI engages with external parties in a number of ways:

- Performing collaborative research

Figure 1 | Emerald is being used to model cell mutations in cancer.

The idea behind this project is to drive an increase in industry use of the e-Infrastructure services provided by the CfI, particularly aimed at small and medium enterprises (SMEs). This would involve improvements in facilitating engagement, collaboration and knowledge exchange between academic researchers, industrial and other commercial organisations based around their use of the CfI’s e-Infrastructure services.

By referring to the CfI website (www.cfi.ses.ac.uk/show-cases/), one comes across exciting projects in the partner universities related to climate modelling, drug development and design, the modelling of new materials such as environmentally friendly catalysts and many others.
Other projects in their take off phase have enormous potential for innovation and impact. One of these projects, which is led by UCL and one of its spin-out companies NPComplete Ltd, aims to optimise the design of computer chips, resulting in faster, cheaper, lighter and more efficient products. Optimisation of circuit complexity commands a very substantial silicon gate count and/or CPU load, a problem that lends itself to HPC.

It is worth noting that, from the outset, Cfl was set up for academic researchers and was configured to suit such a need. The facility has been running at high capacity, as can be seen in the utilisation metrics in Appendix B. Industry outreach, as a result, has not achieved its full potential as encouraged by Government.

**High Performance Computing in Industry: Barriers to Utilisation**

This report explores what the most significant barriers are for optimising industry engagement, and what can be done to overcome them. After a brief review of relevant literature, we describe the findings from a series of case studies in which we engaged with SMEs to use the Cfl facilities. The process has allowed us to draw some conclusions about the suitability of the Cfl for industry engagement, and provide some recommendations for the future.

High Performance Computing in Industry: Barriers to Utilisation

The commercial use of HPC has been growing rapidly for the past two decades within traditional sectors such as aerospace, automotive and new sectors such as biotechnology and finance. Reaping the benefits of the new or enhanced products and services enabled by HPC requires investment in industrially relevant HPC software applications and the development of appropriate delivery models.

A report published by the UK’s e-Infrastructure Leadership Council, outlines the barriers preventing industry from taking up modern ICT technologies. These are a combination of technological, social and economic factors:

1. Awareness
2. Skills
3. Software usability and licensing models.
4. IP ownership
5. Security
6. Connectivity

**Awareness**

An obvious yet often unappreciated barrier to the uptake of HPC by industry and SMEs in particular, is the lack of awareness, of both:

- The relevance and benefit of HPC to their business
- The availability of resources

**Skills**

There is a severe shortage of skills required to make the best use of the technology available. Excellent hardware and software fall short of purpose without the skills and educational resources necessary to exploit them. The ELC report emphasises two ways in which this skills shortage can be addressed:

- Direct and intensive training for business employees
- A more general shift in educational policy that starts at the undergraduate level to encourage students to utilise available e-infrastructure facilities

**Software Usability and Licensing Models**

Given that the software is the point of contact between an HPC user and the e-infrastructure, there are currently a number of issues that need to be resolved regarding its usability. These are cost, complexity, applicability and licensing models.
Specialist software can be prohibitively expensive, even for larger companies. In addition specialist software can only be operated by people with the requisite skills. Considerable investment in training is required to address this issue. The ELC report has the following suggestions to improve software usability. These offer some shorter term options rather than waiting for longer-term educational/training changes to take place:

- Simplification of the user interface
- Provision of software specialists to operate the application on an SME’s behalf when the software is required for short periods
- Provision of online community environments that combine education, collaboration and simulation resources

Inflexible licensing models are thought to prevent SMEs from using HPC for business opportunities. There is a need for more innovative and flexible licensing, more in line with try-before-you-buy and pay-as-you-go models. Perpetual ownership of licenses is not necessary or realistic in a dynamic production environment where companies need to experiment with software for short periods of time or update and change their tools and processes. A cost-effective model needs to be put in place and a license for use from a third party HPC provider is often the most economical option.

A lot of software developed is now open-source (offering universal access and a free license), although it may not always suffice for an SME’s business needs in terms of functionality and capability.

**Intellectual Property Ownership**

It is not uncommon for IP issues to arise in the context of collaborations between academia and industry. It is increasingly obvious that current IP management models are not configured to support the requirements of innovation, where a fast-paced and flexible approach is required to ensure that ideas and commercial development rights are protected. It is inconceivable that a single business can retain all the IP, where a number of partners are involved. As inter-operability increases, so will the flow of information and the need for new models will increase.

**Security**

While interconnectivity is providing significant opportunities for more effective access to resources, security is a major concern. This is more of an issue for Cloud-based pay-per-use providers, and is elaborated in a subsequent section on cloud computing.

The issue of security is wide ranging and has an impact on research along multiple pinch-points, depending on the subject area. Taking an approach that is either rigid or not robust enough will hinder the benefits of access to distributed HPC capability.

**Connectivity**

Successful HPC access is dependent on always being able to supply the latest in network speed and bandwidth capability. Although the UK prides itself on the JANET network, any increase in demand will require an up-grade of our connectivity capability. Many SMEs will not have access to JANET and will be constrained by standard commercial offerings. If work is high data volume, cost could be very high and time scales slow. Connectivity is also relevant for SMEs wanting to access external HPC services over the Cloud. They currently either have to pay for a dedicated, high-cost network connection or have to make do with slow performance.
CASE STUDIES
Case Study 1:  
An SME’s Perspective on Access to GPU facilities

Company: Zenotech

Zenotech is a UK-based SME that produces computational dynamics software, and have a demand for Graphics Processing Units. They stopped using EMERALD when the charging model came into effect in May 2014. This case study highlights the barriers encountered by Zenotech and potential implications for similar industrial players.

Powering innovation with HPC

Access to academic facilities like the EMERALD supercomputer is vital to the success of new start-ups in the High Performance Computing field, says Jamil Appa, co-founder of Simulation on Demand start-up Zenotech.

Appa left a successful career at BAE Systems two years ago, in order to set up Zenotech with co-founder David Standingford. Their partnership brought together more than thirty years of computational engineering experience.

After years creating software for his internal customers, Appa had valuable insight into what would be of use to engineers in other businesses. He could also see the barriers to entry that these businesses faced with the products that were then available to them.

“The tools that are available in the commercial space have high licensing costs, you have to buy a ‘seat’, even if you only use a product for a couple of weeks a year, they are also perceived as very complicated to use. There are lots of engineers who would benefit from using aerospace grade tools with a flexible licensing policy, so that was the aim of the business we set up,” Appa says.

Zenotech has developed an easy-to-use GPU accelerated high performance computational fluid dynamics solver. ZCFD offers unlimited resources through an on-demand payment model. To develop and test this sort of software requires major resources and needs both access to an array of GPUs and true scalability.

“It’s been invaluable. When we started, there was no other resource in the UK that could provide GPUs at that sort of scale,” says Appa. “Plus, if I had to pay full commercial rates to test at scale, I’d be investing a huge amount in development. That’s one of the major barriers to the whole area of HPC software development, the infrastructure needed. And not just for testing but just for proving things. It might be the case that a particular algorithm just doesn’t work at a given scale, or that a problem just can’t be solved. And the cost of finding that out, for an SME like ours, we were just two people until we recently hired our first employee, well, it would just be an impossibility,” he says.

The testing on EMERALD wasn’t always straightforward, and sometimes involved working directly with hardware manufacturers, including NVIDIA, to iron out the glitches.

“I have a good relationship with them and could work directly with the technical team,

Figure 2 | Motorsport simulation undertaken on 24 GPUs with a 20x acceleration in time to solution.
which made things easier. But I always made sure the EMERALD sys admins were aware of what was happening, so that they could implement changes and learn from each hiccup that we solved,” Appa says.

In the long term, Appa hopes to create a commercial contract with EMERALD, to run his own ZFD software for clients, and also to offer on-demand access to HPC Cloud resources for clients who need a fast, scalable system.

“I want to offer a choice. If clients want a cheap as chips option from AWS (Amazon Web Services), we can offer that. But others will want access to GPUs at scale, or simply need more hand holding, in which case that might be EMERALD. We can offer full support and the whole spectrum in between. Some people will want more reliability, and others will just want a quick look at something at scale.”

“For commercial use to be successful, there will need to be a change to the access processes for EMERALD, he says. As with other academic resources, sign-up can be cumbersome.”

“Rapidity and ease of use is key. People want to log on, and do their work. If it takes a week for someone to come along and get an account set up, then you’ve lost the business.”

Appa expects that an upcoming upgrade to EMERALD will be the right time to sign up commercially, and he has already been working with their team on what the details might be. He is confident that the system can provide what he and his customers need.

Appa hopes to work with a range of vendors, each of them 'carving out their own special offerings' in that space, and EMERALD is very much part of that. For clients wanting to use GPUs, we need that resource. You can get access to GPUs in Amazon, yes, but those systems aren’t architectured to use this sort of large software at scale.

“We wouldn’t be where we are today without having access to facilities like EMERALD, both for product development, and for the service we want to offer our clients,” he says.

Case study 2:
Using a GPU Platform for Materials Modelling

Company: Johnson Matthey

Johnson Matthey (JM) is a leading enterprise in sustainable technologies. It develops and manufactures a wide range of high technology products, many of which provide environmental and quality of life benefits from clean energy to the recycling of precious metals. JM has strong ties with UCL and the Doctoral Centre in Materials Modelling.

Background

Recent developments in software have allowed codes used for materials modelling to be ported to GPU platforms (such as EMERALD), with the aim of achieving significant speed-up of the modelling process compared to CPU platforms. In this case study, JM worked with CFI to utilise GPU processors for quantum chemical calculations, which are widely used in the materials modelling space. This project utilised the open-source software GPAW,

![Figure 3](image-url) A visualisation of a molecular modelling conducted by Johnson Matthey.
with the intention of using a test set of molecules for evaluation purposes.

JM worked with the EMERALD facility on attempting to compile and install the necessary GPAW code requirements for porting to a GPU platform, and run some simple evaluations to test, speed and accuracy of the code. In this process, several barriers surfaced.

Progress

Following an optimistic beginning, the first real barrier to commencing the project surfaced. Interestingly, it did not originate on the EMERALD side but rather with the IT infrastructure of the industrial partner. Many corporate entities have a significant firewall and security around their internal network that does not lend itself easily to allowing remote access to HPC facilities (or other external computing facilities). There was a certain amount of reticence about opening up the network and several technical issues were also encountered. It took two months for the industrial IT group to open a satisfactorily secure port that would allow access to the EMERALD system. There remains an open question of how a commercial code, which would need to check-out licences through the firewall, would function in this environment.

After a seemingly successful compilation of the code, initial tests indicated a runtime error.

Further tests are now underway to determine whether this is due to a bug or due to an incompatibility with the modules used for compilation. The developers were contacted and it was suggested to try an older version of several modules.

This unforeseen technical issue has put an unanticipated demand on time and effort for the client. It is also worth noting that EMERALD cannot currently provide flexible access for mobile clients, which is a source of delay for innovative research projects.

For instance in this case the client works between sites in London and South Africa. Having created their account remotely from South Africa, EMERALD can only be accessed via that gateway. On subsequent trips to London access to EMERALD has not been possible due to the added security requirement. A more flexible access mechanism could have made the operation more efficient.

Conclusion

This case study was not completed to its full potential due to the problems mentioned above, but important lessons were learned:

- Corporate firewalls are a barrier to allow remote access to HPC facilities
- Software incompatibilities with an existing HPC system are a source of hindrance to progress, especially when both the client and the provider have a limited amount of technical support available
- Restricted access models may be a good security measure but are a hindrance to fast-paced flexible work
- Again, lack of recognised priority or business driver on the side of the client and provider impacts progress when challenges are encountered.

Case study 3: Drug Discovery using HPC in the Pharmaceutical Sector

Company: CBK Sci Con Ltd

This case study is designed to investigate whether Cfl meets the needs of the pharmaceutical sector in their innovative drug discovery processes. CBK Sci Con Ltd is a UK-based SME who will front the interaction with pharma clients in its position as an HPC consultancy with strong research ties to UCL.

Background

Increasing numbers of pharmaceutical companies realise that they could be using
HPC simulation in their product design and development, but lack their own large scale computational facilities to perform such simulations. They are also reluctant to invest in compute and data storage facilities until the return on investment has been roven in this sector. The Binding Affinity Calculator (BAC) is a decision support tool for simulating targeted drug delivery that can improve the efficiency and effectiveness of the treatment of people with Cancer, HIV, and an increasing number of diseases.

![Molecular model of drug binding in a case of HIV.](image)

Because of the amount of data and the complexity of the calculations involved BAC needs access to an HPC facility like EMERALD and some middleware and software installation requirements as outlined below.

In order to perform a large part of the simulation chain on the EMERALD system, commercially licensed versions of the NAMD and AMBER MD simulation codes are required. Additionally, to incorporate the system in CBK’s automated workflow platform, a suitable job submission interface needs to be installed on EMERALD, such as QCG-Computing or Globus GRAM.

Progress

The simulation code installation was circumvented due to the existence of these codes on EMERALD already with academic licenses. As this project was a joint CBK-UCL venture, the academic licences were used in addition to an access profile which belonged to the UCL researcher who agreed to participate in this case study. It is worth noting that the purchasing of industrial licences would have caused a significant additional expense and delay due to installation.

The installation of the middleware (QCG-Computing) posed significant problems due to incompatibilities with EMERALD and caused a delay of two months and took up significant amount of CFI support staff time. The middleware developers were contacted to solve this problem and they requested an account on the system to investigate the issue. There was reluctance on behalf of the CFI staff to grant them this as there was a view that such access to the live EMERALD environment was inadvisable with progress again slowed by resource constraints.

Eventually, CFI staff created a test environment and deployed a job, but the case study time-frame expired and was terminated, as there was no guarantee of success even if a dedicated machine was used for this purpose.

Conclusion

This case study faced serious obstacles from the outset and was not completed successfully. As with the previous case study, important lessons were learned:

- The CFI is under-resourced to fulfil time-consuming middleware installation requirement that are necessary for certain projects to take off.
- Some specialised software tools have not been rigorously tested across a wide range of platforms and installation may pose problems that challenge both vendors and systems administrators. HPC Test environments are costly and may not be readily available, this has a further impact on the ability to resolve this issue.
- The level of support from the staff (communication, progress update) is inadequate for an industry client who would expect a return on a commercial service.
A service worth taking to industry must meet the demand of a prospective client. It must also ensure that there are alternatives available for compute such as Cloud providers.

**Case 4: Software Licensing Models from a Software Company's Perspective**

**Company: BIOVIA (Dassault Systemes)**

Dassault Systèmes BIOVIA brand ([www.3ds.com/products-services/biovia](http://www.3ds.com/products-services/biovia)) provides 3D modelling software for chemical, materials and bioscience research for the pharmaceutical, biotechnology, consumer packaged goods, aerospace, energy and chemical sectors, among others. BIOVIA was previously known as Accelrys, whose solutions are widely used in the materials modelling sector across academia and industry.

In an interview with BIOVIA at its Cambridge, UK headquarters, the issue of licensing models for academic and industry users was discussed. This was done to shed some light on whether existing models were preventing some SMEs from accessing the software, and consequently, exploiting the high performance computing infrastructure.

BIOVIA provides modelling and simulation tools like BIOVIA Materials Studio and BIOVIA Discovery Studio to numerous working groups and consortia that are working on improving the effectiveness and accuracy of predictions as well as the speed of modelling processes. The company also enhances and completes the models upon which it bases its predictive science through collaborations with these groups and their thought leadership.

BIOVIA places value on its relationship with the academic sector and the company has historically provided access to BIOVIA Materials Studio to academics for a number of reasons. For example, academics contribute to the development of specific

---

**Figure 5** BIOVIA Materials Studio is a complete modeling and simulation environment that enables researchers in materials science and chemistry to predict and understand the relationships of a material's molecular structure with its properties and behaviour.
forcefields that the company has shared with its customers. It would be very expensive for BIOVIA to develop those independently. Additionally, academics that are thought-leaders in their fields promote Materials Studio in their research work, assist with seminars on their work and are available for technical and future-looking discussions. In several cases, they have been actively involved in providing feedback on new product functionality. Thus, the relationship with academics is perceived to have significant return on investment.

In the same vein, BIOVIA has used academic centres of excellence to provide Materials Studio to other academic institutions in a particular country, enabling them to host, support and promote the use of Materials Studio to other academic customers within the region.

BIOVIA has also made provisions for working groups, consortia and NGOs that consume large amounts of structured and unstructured data. For example, a recent project with British Telecom and the UK National Health Service was only possible because BIOVIA’s Pipeline Pilot was able to comb through the anonymity records of 57 million patients to produce data-sets which could then be further analysed by researchers to look for patterns in cancer occurrence in certain population groups. As another example, in 2010 (when BIOVIA was Accelrys), BIOVIA supported the Drugs for Neglected Diseases Initiative (DNDi), providing free licenses to Pipeline Pilot and cheminformatics (computer science and chemistry) software.

The DNDi project provided hosted data for several not-for-profit, public-private partnerships (PPPs) that are leading the charge against neglected Diseases such as malaria, chagas, schistosomiasis and human African trypanosomiasis (sleeping sickness) that affect millions of people in the developing world.

When it comes to small and medium enterprises, however, the landscape is still immature. Dr. Marc Meunier, part of the Modelling and Simulations group in BIOVIA, was asked whether the company had any special provisions for SMEs, or whether they were placed in the same category as larger companies. He replied that although BIOVIA has historically been able to meet some particular requirements on a case-by-case basis they do not offer special provisions for small companies.

There are a few important factors to consider. As suggested by Dr. Meunier, the company has not faced considerable demand from the SME sector to warrant a revision of policies. This may, in part, result from the fact that the software is rather specialised. In addition these companies might not have the skill capability in-house to make best use of BIOVIA’s software.

Moving away from BIOVIA as a case example, software accessibility and licensing models have been identified by the UK e-Infrastructure Leadership Council as a significant barrier to SMEs wanting to make use of the high performance computing e-infrastructure. A five-figure sum for software can be quite a challenge for small companies who have even less incentive to make such an investment if they are not interested in the perpetual use of a license or where the real costs outweigh the possible returns. Companies sometimes need to experiment with software for short periods of time before deciding whether it meets their needs. It has been suggested that the licensing models of software companies should be revised, to include pay-as-you go or try-before-you buy models to lower the barriers to access.

On this topic, BIOVIA states that such innovative models are under discussion, where the company might look to implement new policies in line with the current demands in industry. However, according to Dr Meunier, the biggest barrier to entry for an SME is not the cost of the software but the
lack of internal expertise. This might oblige a company to hire an expert, which would add a lot of expense and most likely, end up more costly than the software license.

Disclaimer: Any views or opinions presented in this article are solely those of the employee and do not necessarily represent those of the company. Employees of BIOVIA are expressly required not to make defamatory statements and not to infringe or authorize any infringement of copyright or any other legal right by written communications. Any such communication is contrary to company policy and outside the scope of the employment of the individual concerned. The company will not accept any liability in respect of such communication, and the employee responsible.
SKILLS AND AWARENESS
As discussed previously, companies may not be aware of the benefits that HPC can bring to their business. Where awareness exists, it is important to re-iterate that there is a shortage of skills necessary to fully exploit HPC to deliver the expected benefit.

While requisite skills may be present in the academic network, they may not necessarily be readily available. Researchers tend to be engaged in their funded research and may not be willing or able to stop and work on commercial consultancy work for a number of weeks or possibly months. However, the establishment of a coordinated platform of expertise across academic institutions that provide guidance and advisory services, with the possible inclusion of doctoral students in interesting projects, could achieve substantial mutual benefit.

It is equally important for publicly accessible e-infrastructures like the CfI to work closely with local companies to provide hands-on training and experience with HPC systems and the tools and support for existing users to make the transition from academia to industry. There is a need to understand business priorities such as how to achieve competitive advantage, the importance of time to market, cost/value and articulating a clear value proposition of how HPC can support and drive these areas.

There are a growing number of entities that understand the need of building bridges between academia and industry through training. One example is the ASI Data Science and Business Analytics Company (www.theasi.co). It selects outstanding PhD’s and Post-Doc’s from science, technology, engineering and mathematics from the world’s top universities to join a fellowship program; an eight week bootcamp in the fundamentals of data science, helping them to make the transition from academia to industry. The fellowship content is also customised to address each partner company’s business needs.

Taking the example of the ASI, there has been a surge of recent projects in the programme that require elevated amounts of computational power. These are mostly imaging related machine learning problems in, for example, the fashion and entertainment industries. Certain machine learning techniques that are used in customer segmentation across a number of industries also tend to require more processing power. Thus, there is a growing interest to incorporate more HPC in training programmes.

For such an initiative, linking up with an academic institution could provide valuable opportunities. The Research IT Services Department (RITS) at UCL is specifically tasked with providing IT services to enable and support UCL research and includes the group that cover core areas of e-Infrastructure: Research Computing Platforms (such as High Performance Computing, HPC), Research Software Development and Research Data Services. RITS have a rich training programme with experienced lecturers covering a range of the aforementioned topics.

There could be an opportunity to implement days of training related to HPC in existing programmes. The training could be conducted by RITS staff, with the use of pre-existing training material owned by ISD. This initial step could lead to more incorporation of HPC training and awareness-raising in industry-oriented training programmes.
CLOUD COMPUTING: AN ALTERNATIVE
An issue worth addressing here is whether industry, especially SMEs, could have more opportunities to make use of HPC facilities by using Cloud computing.

The word Cloud is very widely used, but it might not mean the same thing to everyone. According to the US National Institute of Standards and Technology, Special Publication 800-145:

“Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.”

The essential criteria for cloud computing are a) on demand self service; b) broad network access; c) resource pooling; d) rapid elasticity; and e) measured service.

Despite this definition, the term cloud is still very broad. It is used to refer to a variety of services and different layers (i.e., infrastructure, software, platform). Clouds can be implemented in different ways: public, private, and hybrid. Clouds can be both on- and off-premises.

Ignoring these variations is risky, as there is not a single coherent value proposition that a Cloud can bring to industry. Amazon Web Services and Google are two examples of cloud providers that are capturing the interest of companies that want to enter the compute space without significant resource commitments. They typically involve no minimal requirements and no start-up costs, which is perceived as attractive to the growing SME market.

Cost and Flexibility

Indeed, one of the most frequently cited examples for opting for cloud use is the cheaper cost, especially when economies of scale are involved. Cloud service providers usually have no minimum charge, or set up costs, relying mostly on an affordable pay-as-you-go model. Considering CfI’s pricing scheme, there is a set-up charge for use of either EMERALD or IRIDIS, followed by a price per core per hour depending on the machine used. Thus, at first sight, one can say that cloud use offers a cheaper and more flexible option.

However, this probably applies to situations where highly-standardised commodity services are involved. Standardisation of existing applications may turn out, in the end, to be prohibitively expensive. Research and innovation by its very nature does not lend itself to standardisation.

It is undeniable, however, that cloud provision offers the opportunity for experimentation and short services. While this may be true, it does not necessarily provide a long-term robust and secure service.

Expertise

e-Infrastructure providers such as the CfI have a strong network of experts with access to a range of academic research centres that are world leading in the computational space. These facilities are used across a variety of sectors. Ongoing in the CfI pipeline are exciting projects in the partner universities related to climate modelling, drug design, the modelling of new materials such as environmentally friendly catalysts and many others.

This academic research also has considerable overlap with the type of applications that are relevant to the industrial sector.

Thus, a public infrastructure with ties in higher education, like the CfI, has the potential to offer domain-specific expert advice to industries that are interested in making use of computing infrastructure. Indeed, not all companies are similar in their experience with computational methods. For those who are computationally savvy and have no doubt as to the code they are
trying to run and how they wish their data to be handled, cloud resources are an attractive proposition. However, for those who need guidance on the process, providers such as the CFI could offer a more attractive support and advise service than a cloud provider.

**Scale**

While cloud providers are increasing in popularity, the scale at which they are operating at the moment is insufficient to meet the demands of a computationally intensive project. A full-blown version of Case Study 3 would require something in the region of 10 Million Core hours. The issue of scale is echoed by Zenotech in Case Study 1, who could not find the scale of GPU that they are looking for in the cloud.

**Security**

Security concerns are often raised in conjunction with the cloud, and media hype adds to these perceptions. Recent examples of major disruption to apparently secure web services have raised serious concerns. Access to cloud-based services can be denied, data can be copied or data can be altered resulting in serious negative impact.

Cloud providers have explicit security policies that they agree on with the users. When a user moves computer systems and data to the cloud, security responsibilities become shared between the user and the cloud service provider. The provider is responsible for securing the underlying infrastructure that supports the cloud whilst obeying the law of the land (which is often unclear when the company is domiciled in a different country to the hosting infrastructure). The user is responsible for all content they put on or connect to the cloud. It must be noted that policy/process controls are inherently weaker than technical controls.

Security concerns related to the cloud may change over time as services become more mature and technologies become more robust. However, at the moment, some industries (i.e., the pharmaceutical sector) have qualms about sending data on proprietary compounds to the cloud. The risks of unauthorized access to research data include:

1. Loss of Intellectual property to a competitor.
2. Negative impact on reputation and unethically achieved gain for competitors.
5. Access route used as a wormhole to perform further breaches.
Security concerns related to the cloud may change over time as services become more mature and technologies become more robust.
CASE STUDIES FINDINGS: LESSON LEARNED
These case studies have shed some light on the prospect of CfI providing HPC services to industry. There are some important lessons to learn from this exercise:

- An increased investment in specialist staffing, business processes and tools are required for successful industry engagement:

The CfI has been set up as an academically oriented service, with a level of staffing appropriate for such a role. The level of staffing available and embedded business processes falls short of business expectations. Traditionally academia operates on different (usually longer) time scales and priorities to industry. As an example, in case study 3, a client would not have been happy to wait for two months for an answer about whether a specific middleware item could be installed.

- Industry requirements regarding Software and Middleware can often by incompatible with the HPC system:

It is sometimes the case that the software or middleware required for a certain project is incompatible with the HPC system in question. For example, case Study 3 required a type of middleware called QCG Computing, which did not support the submission system on EMERALD. This type of barrier is rather difficult to overcome, as alternative software needs to be recommended. This also relates to the issue of staffing and the expertise of support staff in question. Without a sufficient and appropriate level of resource allocated to ensure the smoothness of the CfI Industry operation, it would be difficult to ensure that alternative software can be proposed.

- Client companies also have the responsibility to configure their systems for flexible access to HPC:

Case study 2 showed us that getting around the company’s firewall took an excessive amount of time. It is therefore equally important that the policies and infrastructures are well structured on industry’s end to allow a flexible and on-demand access to compute resources.

Now, the barriers mentioned at the beginning of the report will be taken and scrutinised from the specific perspective of the CfI.

**Awareness**

We recognise that there might be an untapped market for SMEs who are not aware of the benefit that HPC can bring to their business. Business would be prepared to invest significant amounts of money if they believed they would receive a positive return on such an investment.

There is a lack of recognised business drivers – for both suppliers and customers to invest in high performance computing. This could be alleviated, in part, by including HPC sessions in ICT business-oriented courses.

**Skills**

BIOVIA, the software company that featured in Case Study 4, would agree with the ELC report that skills present an important barrier for SMEs to make best use of their computational software, thus exploiting HPC facilities. However, Case Study 1 shows an opposite picture where the respective SME involved did not have a problem with the required skill set required. They faced other problems, of which price and accessibility were the highest. Each SME is different and would thus expect a different level of service from an infrastructure provider. Computationally savvy SMEs with a standardised problem may only be interested in purchasing cycles or opt for a cloud provider due to the flexibility that offers.

For other SMEs, who need a varying level of hand-holding, infrastructure providers like the CfI must be able to offer an appropriate level of support and expertise. This would
range from help to set up the code properly and configure it on the HPC systems in question, but also a higher level of guidance and technical advice.

*A successful compute infrastructure must be able to provide a tiered level of service based on the client’s level of comfort with compute.*

**Licensing models**

Case Studies 1 and 4 addressed the issue of licensing models from the perspectives of both an SME and a software company respectively. While some SMEs might find software licenses prohibitively expensive, it is still the perspective of a company like BIOVIA that the level of demand from the SME sector does not warrant a change in policies yet. While there is talk of innovative models like pay-as-you-go, it might actually take a while before this issue is resolved. SMEs may need to look for open-source software as an alternative.

Not all software companies find an argument for changing their software licensing models to accommodate the financial restrictions of SMEs.

**Security**

It is worth mentioning that security concerns are multi-faceted, coming from both the perspective of the client and the provider. A client might be unwilling to send confidential data to a resource provider. Also, delays may come from the client’s perspective due to the need to get past the company’s firewall, as we saw in Case Study 2. A service provider may also have security qualms about the specific software that an SME may require.

Security concerns are present on both sides of the compute business relationship. There need to be well-defined policies and a clear understanding between the client and provider.

**The Cloud alternative**

Cloud resources are an acceptable alternative for industries that are interested in on-demand compute with a flexible model, provided that they have the necessary skills to exploit the resources in-house and do not require very large amounts of compute.

**Conclusions**

The case studies showed a promising landscape of innovative projects that could be conducted with HPC. The CfI, however, as it stands, is not able to provide the required level of service for industry to be able to make full and effective use of the HPC services on offer. The CfI is already providing a consistent high-quality service for researchers from the CfI consortium. For the needs of industry to be accommodated there needs to be a significant level of investment, including, at least:

- Increase in the level of support staff and personnel with the relevant specialist knowledge and skills across both industry and academia.
- Revision of the current costing model
- Strengthening of the network of academic advisory services
- Provision of flexible, tiered levels of support, service and access.
- Revision of the security policy for better definition of and provision of standards between the client and provider.

The return on investment in financial terms for such a serious industry outreach agenda has not been established for the CfI. None of the UK HPC regional centres have achieved a significant financial return yet. There are clearly different requirements, as pertains to compute, between researchers and industry. A better investment might be in strengthening the skills/advisory network around the CfI partner institutions, which would then be able to offer an advanced level of skills and service.
A successful compute infrastructure must be able to provide a tiered level of service based on the client's level of comfort with compute.
APPENDIX: CfI UTILISATION METRICS
June 2013 - December 2014