AI and HPC as Drivers for Industrial Competitiveness

Alison Kennedy
Director | STFC Hartree Centre
Who are we?

- The UK Science and Technology Facilities Council’s high performance computing, data analytics and cognitive technology centre
- Primary purpose is to apply the latest capabilities in computing technologies to industrial challenges
- Provides businesses (and applied researchers) with access to powerful technologies, facilities and scientific computing expertise
- Based at Sci-Tech Daresbury alongside Daresbury Laboratory in North West England
Our mission

Transforming UK industry by accelerating the adoption of high performance computing, big data and cognitive technologies (AI, ML, DL) through challenge-led research and innovation.
Our purpose

- To lead UK research into the (industrial) application of high performance computing, data & cognitive technologies (with IBM Research, ATOS Bull and other partners)

- To drive the economic development of UK industry through the use of leading edge scientific knowledge and applications
What we do

- **Collaborative R&D**
  We build a team to deliver a solution to a particular challenge

- **Platform as a service**
  Pay-as-you-go access to our compute power

- **Creating digital assets**
  License the new industry-led software applications we create with IBM Research (IROR programme)

- **Training and skills**
  Run specialist training courses and events
Our track record
Faster development process for products like shampoo, reducing physical testing

“The Hartree Centre’s high performance computing capabilities help us achieve better design solutions for our consumers, delivered by more efficient, cost-effective and sustainable processes.”

– Paul Howells, Unilever
Consumable HPC

User

Sent

Received
Getting HPC to “work smart, not hard”

- Typically HPC development is focused on increased speed.
  - The fastest calculation is the one which you don’t run!
- Can we use machine learning to make better decisions on which simulations give the most value?
- Can we use machine learning to improve resolution of information?

‘Cognitive’ workflow uses 1/3 of the calculations to achieve 4 orders of magnitude resolution increase
Case study | Building the cognitive hospital

Transforming the patient experience using cognitive technology and data analytics

“Helping our patients and their families prepare properly for coming into hospital will really reduce their anxiety and could mean they spend more meaningful time with doctors so we are able to get them better faster.”

– Iain Hennessey, Alder Hey Children’s Hospital
HPC capabilities ran simulations four times faster than using in-house systems

“The service enabled us to increase the size of models we use and run projects more quickly and efficiently, without increasing our overheads. That all adds up to improved learn rates and better value for money for clients.”

– Jeremy Smith, QED Naval Ltd
Access to compute capability reduced time to run vehicle simulation models by 50%

“Access to high performance computing with the convenience of operating from our own office is a real game-changer, helping us meet customer cost and timescale requirements and enabling us to thrive in a global market.”

− Tim Williams, Simpact Engineering Ltd
Creating digital assets

Case study | Smart crop protection

Building pest risk prediction models with the potential to:

- Enable the farming industry to more accurately plan preventative measures
- Reduce crop losses
- Drive down insurance rates through lower probability of crop damage
Creating digital assets

Case study | Cognitive waste water treatment

Using artificial intelligence (AI) to dynamically manage the waste water treatment process with the potential to:

• Adapt responsively to environment conditions e.g. weather forecast
• Make water treatment plants more efficient
• Minimise costly regulatory violations
Delivered a series of workshops to SMEs on Industry 4.0 technologies including:

- Building IoT applications with Sigfox
- Blockchain | Changing industries
The Innovation Return On Research (IROR) programme (STFC/IBM Research) gives us the opportunity to bring three key technologies together.

(Current programme uses capital funding to develop reusable digital assets under State Aid rules.)
Industrial Engagement process InCEPT™ ("Accelerated translation of research ideas through improved Industrial Collaborations"), developed in partnership with Cambium LLP
Panther

- IBM POWER8 Firestone
  - 512 POWER8 cores
  - 64 nVidia Kepler K80 GPUs
  - IBM ESS GS4 storage
  - Mellanox Infiniband
- Also has IBM FlashStorage in IB- and CAPI-attached options
- System targets IROR workloads
Paragon

- IBM POWER8 Minsky
  - 656 POWER8 cores
  - 82 nVidia Pascal GPUs
  - IBM ESS GS4 storage
  - Mellanox Infiniband
  - NVMe

- System targets IROR workloads
Purpose: to optimise experiments

Supported by the STFC Hartree Centre’s Innovation Return on Research (IROR) programme, funded by the UK Government’s Department for Business, Energy & Industrial Strategy.
Hyper-parameter tuning for neural networks

Can use BOaaS...

Hyper-parameters:
• Number of neurons in each layer
• Dropout probability for each layer
• Regularisation parameter for each layer
• Etc....

Output:
• Value of the loss function
Drug discovery

Can use BOaaS...

Hyper-parameters:
• Different drug quantities

Output:
• Effect of the drug on the biological target
File System Optimisation

Can use BOaaS...

Hyper-parameters:
• Number of requests
• Number of works
• ‘readahead’ parameter
Output:
• Throughput performance
STFC Hartree Centre / ATOS Bull Collaboration

Hartree Centre
- £300M investment
- Supercomputing and Big Data
- HPC / DS Analytics and Cognitive

Sci-Tech Daresbury
- National Science and Innovation Campus
- Co-location of Science and Business
- 130 companies on site

Daresbury Laboratory
- Accelerator Science
- Nuclear Physics
- Scientific Computing

Innovation Centre
- Lab space - Biomedical, ICT and Engineering

Royal Charter (2007)
Funded by Gov dept for BIS

Science & Technology Facilities Council
UK government body that carries out civil research in science and engineering, and funds UK research in areas including computing technology, particle physics, nuclear physics, space science and astronomy

StFC, Langtree, Halton BC
JV - STFC, Langtree, Halton BC

Hartree Centre / ATOS Bull Collaboration

Images courtesy of STFC

https://www.youtube.com/watch?v=0Kb8iktZRAo
Key scientific areas include:
- Molecular modelling and material science
- Life sciences
- Virtual Engineering and Digital simulation
- Deep Learning

Atos Bull Technologies Solution:
- Bull Sequana X1000 supercomputer
  - ~4 Pflop/s - Intel Xeon and Xeon Phi, GPUs
  - In excess of 82,496 Cores
- Extreme Factory (XCS and XRV)
  - Professional services to tailor the interface
- Business Development resources to co-sell

Scafell Pike - Public HPCaaS offering to UK academics and Industry
Key scientific areas include:

- Joint Academic Data science Endeavour
- National Deep Learning Service
- GPU enabled Computing
- DL enabled HPC application development
- Prototyping and Algorithm Development

Atos Bull Technologies Solution:

- 22x Nvidia DGX1- DL supercomputer
  - 176 P100 GPUS, Deep Learning Frameworks
  - **In excess of** 630,784 CUDA Cores
- Extreme Factory (XCS and XRV)
  - Professional services to tailor the interface
- Business Development resources to co-sell

JADE – Oxford University and the Hartree Centre
Public DLaaS offering to UK academics and Industry

Image courtesy of STFC
## Benefits of HPC and DL as a Service

### Latest technology, with no Capex

<table>
<thead>
<tr>
<th><strong>Increased IT Agility and Flexibility</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Up to date resources. New capabilities added constantly</td>
</tr>
<tr>
<td>- Run <strong>supercomputing simulations</strong> in an easy, flexible, and secure way</td>
</tr>
<tr>
<td>- Focus on the <strong>Use Case</strong> rather than the system</td>
</tr>
<tr>
<td>- Public/Private/Hybrid models supported</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Improved Return on Investment</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Reduced System Admin and Data Centre costs</td>
</tr>
<tr>
<td>- <strong>Mutualize</strong> user environments and workloads</td>
</tr>
<tr>
<td>- Pay only for what you use, no wasted assets</td>
</tr>
<tr>
<td>- <strong>No requirement for Capex</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Integrated and Secure</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Ease of access through <strong>web based portals</strong></td>
</tr>
<tr>
<td>- Integrated ISV applications</td>
</tr>
<tr>
<td>- <strong>Fast remote visualisation</strong>. View results in 3D from any location</td>
</tr>
<tr>
<td>- <strong>Enhanced Security</strong></td>
</tr>
</tbody>
</table>
Key Markets for HPC/DL as a Service

Just a selection, there are many more

Electro-Magnetics  Computational Chemistry Quantum Mechanics  Computational Chemistry Molecular Dynamics  Computational Biology  Structural Mechanics Implicit

Life science  Data Mining  Video Surveillance  Financial and Fraud  Seismic Processing

Computational Fluid Dynamics  Reservoir Simulation  Rendering / Ray Tracing  Climate / Weather Ocean Simulation  Data Analytics

Source: Hyperion 2017
Deep Learning Application

Detecting flaws in airplane wings

Use case is to detect flaws, missing sealant, excess paint, scratches and foreign objects.

Atos - Airbus
Deep Learning Application

Detecting flaws in welding

(a) Tiny cracks with low contrast and different brightness.

(b) Scratches, grind marks, and welds in background.

Once trained the deep learning neural network, can identify welding issues in near real time, from visual images.

Phys.org
Thank you

Slides contributed by IBM Research (Ed Pyzer-Knapp) and ATOS Bull (Shane Rigby)

Find out more:

@ hartree@stfc.ac.uk
🌐 hartree.stfc.ac.uk
LinkedIn /company/stfc-hartree-centre
🐦 @hartreecentre