CLOUD & HPC IN BIOMEDICINE

DATE: 27TH APRIL 2017 | TIME: 09:00 – 17:30 |
LOCATION: CHRISTOPHER INGOLD LECTURE THEATRE
DEPARTMENT OF CHEMISTRY, 20 GORDON ST, LONDON, WC1H 0AJ

TIME SLOT	AGENDA	TITLE OF PRESENTATION
08:45 - 09:15	*Registration / Tea and Coffee in the Nyholm Room*	
09:15 - 09:30	*Introduction by Peter V Coveney, UCL*	
09:30 - 09:50	Ana Minchole, Oxford University	High Performance Computing for the investigation of electrophysiological activity of human heart in control and disease.
09:50 - 10:10	Dieter Kranzmüller, Leibniz-Rechenzentrum	General purpose supercomputing on SuperMUC
10:10 - 10:40	Peter V Coveney, UCL	Biomedical high performance computing within and outside clouds
10:30 - 11:00	Herman van Vlijmen, Janssen	High performance computing needs for free energy calculations at Janssen.
11:00 - 11.15	*Remarks from Chair (Andy Grant) with Q&A*	
11:15 - 11.30	*Break / Tea and Coffee in the Nyholm Room*	
11:30 - 11:50	Gavin Pringle, EPCC, University of Edinburgh	Current Hardware for HPC, HTC and Data Management for CompBioMed
11:50 - 12:10	Neil Morgan, STFC, The Hartree Centre	Accelerating the adoption of Data-centric and Cognitive technologies
12:10 - 12.30	Andy Grant, Bull Atos	New Cloud services - HPC as a Service (HPCaaS) and Deep Learning as a Service (DLaaS)
12:30 - 12.45	*Remarks from Chair (Herman van Vlijmen) with Q&A*	
12:45 - 13.45	*Lunch provided in the Nyholm Room (Poster Session)*	
13:45 - 14:05	Brendan Bouffler, Amazon Web Services	Cloud computing: the scientific method embodied in computational form
14:05 - 14:25	Matt Harvey, Acellera	Pain-free molecular dynamics on AWS with AceCloud
14:25 - 14.45	Kenji Takeda, Microsoft Research	Biomedicine and real HPC in the Cloud – from research to reality
14:40 - 15:20	Alessandro Riccombeni, DNAnexus	DNAnexus: today's cloud network for secure, scalable collaborations across Industry, Academia and Healthcare
15.20 - 15:45	*Break with Tea and Coffee in the Nyholm Room*	
15:45 - 16.00	*Remarks from Chair (Peter V Coveney)*	
16:00- 17:30	Panel Discussion: Herman van Vlijmen, Ana Minchole, Matt Harvey, Kenji Takeda, Brendan Bouffler, Andy Grant	
17:30	*End of meeting – informal networking in the Nyholm Room*	

ABSTRACTS

Ana Minchole, Oxford

High Performance Computing for the investigation of electrophysiological activity of human heart in control and disease.

Cardiovascular diseases are the leading cause of death globally and they represent the 31% of all deaths according to the WHO. One of the most used and non-invasive clinical tool is the 12-lead electrocardiogram (ECG) that records the electrical activity of the heart over time. However ECG is not disease specific and other clinical modalities such as cardiac magnetic resonance (CMR) images are used in combination. Image-based computational models of heart and torso allow simulation of electrophysiology, opening exciting avenues for clinical translation by expanding image studies to link structure and function. The goal is to showcase how HPC simulations enable the investigation of electrophysiological activity of the human heart in control and disease taking into account inter-subject variability.

First, we investigated how anatomical variability of the heart/torso in a cohort of normal subjects affects the ECG. Second, we evaluated the role of variability in size and location of acute ischemia on arrhythmic risk and changes in ECG biomarkers. To do so, we developed a fully working pipeline from MRI images to the simulation of electrical activity in the heart up to body surface and simulation of the ECG. In conclusion, personalization of models using patients imaging data allows risk stratification in pathological conditions such as ischemia and assessment of pharmacological treatments.

Dieter Kranzmüller, LRZ

General purpose supercomputing on SuperMUC - Update from LRZ.

The Leibniz Supercomputing Centre (LRZ) is one of the largest academic and research computing center worldwide, providing a variety of IT infrastructures to a diverse group of users in Munich, Bavaria, Germany, Europe and beyond. The services of LRZ concern not only the computing power of machines such as SuperMUC, but also the respective support structures. However, with a growing demand from communities such as Life Sciences, more sophisticated interaction is needed between users and providers. At LRZ, this is addressed with the partnership computational sciences (piCS), which established a new kind of relation between those providing the access to computing facilities and those requiring these facilities for their everyday work.

Peter V Coveney, UCL

Biomedical high performance computing within and outside clouds.

The Centre for Computational Science at UCL is concerned with many aspects of theoretical and computational science, from chemistry and physics to materials, life and medical science. Our different computational techniques span time and length scales from the macro, through the meso- to the nano- and microscales. In pursuing these research interests, we routinely make

use of combinations of globally distributed petascale computing resources together with high performance visualisation, and high bandwidth, low latency networks.
To fully exploit the vast range and scale of resources available, we have had to develop mechanisms to automate and manage the execution of complex chains of high performance simulations through workflow tools. In addition, our methods both rely on and produce vast amounts of data, which need to be moved carefully and rapidly.
Our work in the biomedical domain—including drug discovery and selection as well as blood flow, in both of which we work with industry and healthcare providers—requires careful attention to verification and validation, as well as management of confidential data. Speed to solution and quality of service are of paramount importance. As a consequence, we have become increasingly interested in what cloud infrastructures operating at scale can offer to support our research.

Herman van Vlijmen, Janssen

High performance computing needs for free energy calculations at Janssen.

The prediction of free energies of binding for small molecules that bind to protein targets is of very high importance in drug discovery. The current availability of large scale CPU and GPU systems is enabling us to use free energy predictions prospectively in a timely manner in discovery projects. For our calculations we have been accessing a combination of resources, including internal GPU servers, high performance computing centers, and a virtual private cloud at AWS. It is very likely that the scale of free energy calculations will increase significantly and access to supercomputing centers is a very interesting option for Janssen, provided that appropriate security is in place that allows handling of confidential data.

Gavin Pringle, EPCC

Current Hardware for HPC, HTC and Data Management for CompBioMed.

This talk will present the current computing resources available to the CompBioMed consortium and users and plans for the near future for High Performance Computing, High Throughput Computing and Data Management, through direct access and the cloud.

Neil Morgan, STFC, The Hartree Centre

Accelerating the adoption of Data-centric and Cognitive technologies.

STFC's Hartree Centre is in its third phase of funding from the UK government. Its mission, to transform UK competiveness by accelerating the adoption of HPC, big data and cognitive technologies. The current phase of the centre is focused on a wide-ranging R&D collaboration with IBM Research and we have just entered into a new partnership with Atos-Bull with the arrival two new Bull clusters that will form the basis of our HPC and Deep Learning as a Service cloud offerings. The talk will expand on the importance of R&D in the centre, the role of future data centric technologies and how new partners can get involved.

Andy Grant, Bull Atos

Atos Bull have recently launched two new cloud services in partnership with the Hartree Centre.

The HPC as a Service (HPCaaS) and Deep Learning as a Service (DLaaS) are specialist cloud services using a combination of optimised HPC components and GPU deep learning systems. In this talk we will cover the key aspects of these new services and discuss why general purpose cloud environments typically do not work for HPC workloads. We will also discuss the Atos Cognitive Decision Execution (Codex) environment, specifically the big data workflow component which allows high performance analytics tasks to be combined into a template solution which can be deployed either on premise or in the cloud. We will demonstrate how specific biomedical components, such as an optimised genome sequencing pipeline, can be embedded within the workflow tool as part of a hybrid cloud solution, allowing the most appropriate underlying hardware solution to be used for the job in hand."

Brendan Bouffler, AWS

Cloud computing: the scientific method embodied in computational form

The scientific method is fundamentally about exploring novel ideas to find where they fail, so we can adjust our theories & models, test them again, and in doing so, improve our understanding of the world. In a cloud computing world where servers become software that are created with a script, computation itself becomes something that be iteratively improved. A new exploration of an idea can be launched in a few minutes, scaled up or down based on needs (not limitations) and entirely transplanted to different physical environments if it turns out that the computation has different evolving needs to those we anticipated at the beginning. Success in science almost never involves guessing the right answer the first time, so why treat computation any differently

Matt Harvey, Acellera

Pain-free molecular dynamics on AWS with AceCloud

Kenji Takeda, Microsoft

Biomedicine and real HPC in the Cloud – from research to reality

Advances in biomedical tools and techniques are creating a revolution in healthcare. A major challenge is translating research into reality, in terms of scale, on-demand availability, and reliability. Cloud computing provides an ideal unified platform for research, development, and deployment of biomedical applications - from computational fluid dynamics and structural analysis, to genomics and drug discovery. In this talk we will show how the unique capabilities of the Microsoft Azure's trusted cloud platform, including UK data centres with N3 connectivity, are being used to accelerate biomedical projects, enabling rapid deployment of solutions in the real-world. We will describe how real high-performance computing with GPUs is now possible in the cloud, scaling to thousands of cores with InfiniBand networking and bare-metal performance; genomics using GATK/BWA can be accelerated by 700% using Microsoft Genomics; and drug discovery is being enabled using signal pathway modelling in the cloud. Finally, we will explain how the Azure for Research program can provide awards, training, and guidance on how you can use cloud computing for your own research (www.azure4research.com).

Alessandro Riccombeni, DNAnexus

DNAnexus: today's cloud network for secure, scalable collaborations across Industry, Academia and Healthcare

The DNAnexus platform is a global network for genomics and scientific discovery. Our cloud platform was designed to answer three core challenges: large-scale, on-demand data analysis; secure data sharing for global, distributed teams; integration of multidisciplinary datasets and software. These challenges were solved by building DNAnexus as a multi-cloud, multi-region platform allowing scientists to collaborate globally at scale, integrating diverse data types and software solutions, all within DNAnexus' secure and clinically compliant environment.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 675451 (CompBioMed project)