World-Class Research

Drawing on technologies and expertise from Atos and Dell Technologies, Supercomputing Wales delivers leading-edge compute facilities to researchers across Wales.

Customer profile

SUPERCOMPUTING WALES
UWCHGYFRIFIADURA CYMRU

Academic and Collaborative Research
United Kingdom

“Without these pipelines, the identification of tuberculosis would take months. This takes it down to a few weeks.”

Anna Price
Research Software Engineer Anna Price, Supercomputing Wales

Needs

Supercomputing Wales needed a high-performance computing resource to empower the work of researchers across the nation of Wales.

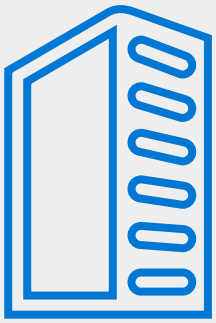
Results

- Fuelling research breakthroughs
- Creating highly skilled research jobs
- Helping Wales capture more research funding

Solutions at a glance

Two supercomputer hubs with:

- Dell EMC PowerEdge servers with Intel® Xeon® processors
- Dell EMC PowerSwitch S3048
- Dell EMC PowerVault MD3420 storage



The HPC facilities comprise two hubs with **13,080 cores** delivering **1 petaflop**

A national mission

Using AI and data in manufacturing to maximize assets. Modelling the power of the sea as a low-carbon energy source. Analyzing 20 years of data on the atmosphere of the sun. Improving the speed and accuracy of tuberculosis testing. This is just a small sample of the diverse scientific investigations powered by the high-performance computing clusters of Supercomputing Wales.

Supercomputing Wales is a strategic programme of investment in the university sector in Wales that is changing the way supercomputing facilities are used to support research activities. The programme is led by Cardiff University, in a consortium with Aberystwyth, Bangor and Swansea universities.

Supercomputing Wales provides researchers across Wales with access to powerful computing facilities for science and innovation projects. The programme's facilities are used by research groups at Cardiff, Swansea, Bangor and Aberystwyth along with companies and other partners working on collaborative projects.

At a higher level, Supercomputing Wales is helping the nation capture more research funding, increase scientific partnerships, create highly-skilled research jobs and support collaborations among research partners. Funded in part by the European Regional Development Fund through the Welsh Government, the program supports Wales in its efforts to compete globally for research and innovation that requires state-of-the-art computing facilities to simulate and solve complex scientific problems.

With its high-powered HPC resources, Supercomputing Wales is facilitating a step change in supercomputing activity across strategically important sectors of the Welsh economy — including nano-scale materials and advanced engineering; energy and the environment; and life sciences and health.

A centre of excellence

At the heart of the Supercomputing Wales initiative is a supercomputing centre of excellence spearheaded by two global leaders in digital transformation, Atos and Dell EMC.

The joint Atos Dell Technologies Supercomputing Centre of Excellence provides Welsh researchers with a full suite of leading-edge HPC equipment, software and services. The centre provides a supercomputer, two Atos BullSequana S Datalake appliances, and on-demand cloud-based supercomputing simulation and services powered by Extreme Factory, the HPC on-demand solution from Atos.

The centre also offers community benefit provisions from Atos and its partners, including in-depth technology workshops for staff and guest lecturers by Atos scientific community experts.

In addition, Supercomputing Wales employs Research Software Engineers (RSEs), embedded with research teams in specific domains, to develop algorithms and customised software that harnesses the power of the supercomputing facilities to perform multiple computational tasks simultaneously at very high speeds.

Two supercomputer hubs

The supercomputer hubs under the Supercomputing Wales umbrella are based at Cardiff and Swansea universities, with research teams across the consortium universities accessing the hubs through high-speed network connections. In total, the facilities consist of over 13,000 cores connected to high-speed memory and storage.

Connected to high-speed **memory and storage**

Table 1 shows the specifications for the supercomputing hubs, named Hawk (based at Cardiff University) and Sunbird (located at Swansea University).

Table 1. Supercomputer specifications.

	Cardiff Hub – Hawk	Swansea Hub – Sunbird
Key specifications	• 201 nodes • 8,040 cores • 46 TB of total memory	• 126 nodes • 5,040 cores • 48.4 TB of total memory
Compute nodes	• CPU: 2x Intel® Xeon® Gold 6148 CPU @ 2.40GHz with 20 cores each • RAM: 192 GB - A subset for large memory applications has 384 GB - 13 nodes have dual NVIDIA P100 GPUs	• CPU: 2x Intel® Xeon® Gold 6148 CPU @ 2.40GHz with 20 cores each • RAM: 384 GB per nodes - 4 nodes have dual 8x NVIDIA V100 GPUs
Storage	• 692 TB (usable) scratch space on a Lustre filesystem • 420 TB of home directory space over NFS	• 808 TB (usable) scratch space on a Lustre filesystem • 231 TB of home directory space on a Lustre filesystem
Interconnect	• Mellanox EDR 100Gb/s InfiniBand compute and storage network	• Mellanox EDR 100Gb/s InfiniBand compute and storage network

Fuelling ground-breaking research

Supercomputing Wales supports research across a wide range of academic disciplines, from traditional science, technology, engineering and mathematics fields to environmental, e-science, health, medical and social sciences.

Here are a few examples of the ground-breaking research initiatives that are capitalising on the computational resources of Supercomputing Wales.

Swansea University: Using AI and data in manufacturing to maximize asset

The EPSRC-funded Transfer Learning for Robust, Reliable and Transferable Cyber Manufacturing Systems project, led by Cinzia Giannetti of Swansea University College of Engineering, works with industry partners to increase the use of artificial intelligence in manufacturing and help them use data to make better decisions.

“The main body of my research is looking at developing machine learning technologies, and algorithms that can use the vast amounts of data collected during manufacturing processes and use them to enable companies to make better, faster and optimal decisions,” says Giannetti.

Machine learning can, for example, help a company to run predictive maintenance to maximize the use of assets and machines, she notes.

“The idea there is to use data to develop and train very complex predictive models using deep learning and then use those models to predict the remaining useful life of equipment, and determine the right time to do maintenance.

“The research itself involves collecting and synthesizing large volumes of data collected through different machines and sensors. We then want to visualize this data and use it to develop a predictive model,” she says.

“Given the amount of data, the precision we want to achieve and the complex architecture of Deep Neural Networks, we need an adequate computing facility,” Giannetti says.

The Research Software Engineers at Supercomputing Wales have been vital in getting this work done, she says.

“They provided training and, in the initial stages, helped us to perform some optimization to run the models more efficiently,” Giannetti says. “The biggest advantage of Supercomputing Wales is not only the availability of high-performance hardware but actually having people that can support us in running the programs — the training, the support, the help in setting up the system. We’ve become quite independent now, but at the beginning it was crucial to us.”

Bangor University: Modelling the power of the sea as a low-carbon energy source

Using the power of the sea is nothing new: people have been using tidal range — the rise and fall of sea levels — since medieval times to drive mills and grind grain. New technologies now allow us convert tidal and wave power into electricity, and research from the School of Ocean Sciences at Bangor University aims to understand the available resources and how they can best be harnessed.

“The good thing about tides is they’re predictable, so you can see what contribution they can make for a hundred years into the future. There’s the tidal range, and there’s tidal stream,

where you harness the kinetic energy of the tide using a horizontal axis turbine, like an underwater windmill," says Simon Neill, Professor in Physical Oceanography.

"And then we have wave energy. It's less advanced than tidal energy because waves by their very nature are quite de¬structive, and of course you want to put your devices somewhere where the waves are energetic — so at present the devices have to be really overengineered to withstand those forces," says Neill.

Neill's research focuses on understanding the resource itself, and identifying the optimal locations for energy devices.

"Waves, for example vary over a long timescale, and you can't really say anything meaningful about the resource until you have at least a 10-year record," he says. "So you could go and put a wave measuring device in place and wait 10 years, or you can use modelling. We use measurements to build a model and see how things will vary over 10, 20 or even 50 years. You can look at things like climate change. No one really knows what the wave climate will look like in 2050, so the only thing you can do is run a model."

The team also looks at the interaction between energy extraction and the resource.

"If I put a 100-megawatt array in place, I wouldn't really know what the impact on the environment would be until post construction," Neill says. "By modelling you can see what the impacts would be, and change the shape or the spacing between devices to minimize impacts on the environment."

Tidal, and especially wave models, are extremely computationally expensive, and can take up to three months to complete, even on a facility as extensive as Supercomputing Wales, but the facility and team make it easy for Neill.

"Before this we had to spend a lot of our time setting up and optimizing our models — and that was just taking away from the science," he says. "Having this dedicated team who deal with all the technical issues is great. After some discussions about a new model application, they'll take care of all the technical set up and optimization, and provide me with some scripts to run the model, so I can focus getting outputs."

Aberystwyth University: Analysing 20 years of data on the atmosphere of the sun

The solar system physics team at Aberystwyth University uses Supercomputing Wales to study the sun's atmosphere, analyzing data collected over the past two decades by a range of satellites and ground-based telescopes.

The sun's atmosphere is still "quite mysterious, with a lot we don't understand," says Dr Huw Morgan. "What Supercomputing Wales allows us to do is to analyze long periods of data, up to 20 years' worth of data, in quite a short period. We have around 15 terabytes of data stored on Supercomputing Wales and that data can be analyzed in several ways, so there are different processing streams depending on what we are trying to study."

One current field of study is the shape of the solar atmosphere.

"We have two-dimensional images, but we don't know the three-dimensional shape," Morgan says. "So one of the most advanced methods we've developed is to use tomography, similar to medical tomography in that you take images from many different angles and work out the three-dimensional shape of what you're looking at. That involves taking thousands of images over the course of a month, as the sun rotates."

The work Morgan's team runs on Supercomputing Wales differs from the heavily parallelized code run by many other physics departments, he says.

"Someone doing quantum physics really needs the parallel processing power of Supercomputing Wales," Morgan says. "We do run jobs in parallel, but running one year of data in one job, another in a second and so on — so we can process 20 years of data in one 20th of the time. It's totally invaluable to our work — we wouldn't be doing the research we're doing without Supercomputing Wales."

The solar system physics team works closely with Supercomputing Wales support staff. For example, Colin Sauze, a Research Software Engineer for the facility, is based in the office upstairs from Morgan and regularly helps the team out with problems, or runs workshops on specific issues.

Access to the facility has never been an issue, with computing time usually available within "seconds, or minutes," Morgan says.

Cardiff University: Improving the speed and accuracy of tuberculosis testing in Wales

Around two million people die from tuberculosis each year. These deaths are mostly in developing countries, but the disease remains an issue in countries like Wales, where around 2,000 people are tested for tuberculosis each year.

Supercomputing Wales is involved in a collaboration with Public Health Wales, CLIMB-BIG-DATA, Cardiff University, Public Health England, the University of Oxford, Health Protection Scotland and the European Bioinformatics Institute (EBI) to implement scalable bioinformatics pipelines for the pathogen genomics of tuberculosis.

"Essentially it's about better testing for patients in Wales," says Supercomputing Wales Research Software Engineer Anna Price. "There are current tools that exist, academic software for detecting tuberculosis, but patients need software that's fast, reproduceable and automated."

Price has developed two software pipelines for the project.

"The first pipeline runs quality control on the genomic data of a patient sample when it arrives from the sequencing machine, then identifies the species in the sample and whether or not it is tuberculosis. The second pipeline, which feeds into this one, builds the reference database that is used to detect the species," Price says.

"High performance computing is crucial to all of this. There's a lot of data to crunch," she says. "When a sequencing machine reads the DNA of a sputum sample from a patient, you don't get a nice long strand of DNA, but millions of fragments, of short reads of DNA. So a big part of the work in the main pipeline is quality control, and trying to reassemble these fragments into the original genome, like a big jigsaw puzzle!

"And then you try to answer the questions 'what species to I have here?', and that's what the reference database is used for," she says. "It contains DNA fragments from lots of species of mycobacterium and we run the sample past to see if we can get a match to get a species ID and confirm whether this is mycobacterium tuberculosis."

While there are many reference databases already in use, they are not specific enough for this case, Price says.

"This bespoke database is just mycobacterium, and the process of building it has been automated because it's quite an involved process, feeding it labelled data, labelled species, and making sure the data is high quality," she says. "It's self-correcting as well, in that the more data you feed it the better it gets over time. Without these pipelines, the identification of tuberculosis would take months. This takes it down to a few weeks."

Key takeaways

With the HPC systems and expertise of Supercomputing Wales and the Atos Dell Technologies Supercomputing Centre of Excellence, researchers across Wales have access to the resources they need to simulate and solve complex scientific problems — and keep the nation of Wales at the forefront of theoretical science and technical innovation.

As Professor Roger Whitaker, academic director for Supercomputing Wales, notes, "Supercomputing Wales, with its supercomputing hub facilities and new group of Research Software Engineers, is a unique national technological asset that is crucial to underpin our scientific excellence in Wales."[1]

1 Atos news release, "Atos wins deal to establish supercomputing Centre of Excellence in Wales," June 21 2018.