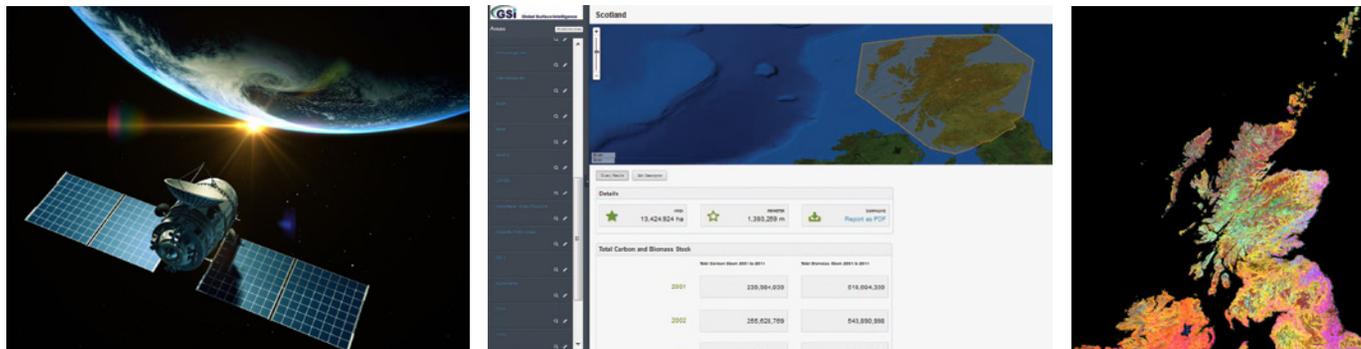


Helping your customers generate value from big data



Images: GSi

Global Surface Intelligence (GSi) is a data services company that provides machine learning and predictive analytics on large and complex data sets ('big data'). One particular area of GSi specialisation is analysing satellite information and remote-sensing Earth Observations (EO) to solve its customers' business problems in new and novel ways.

One example of GSi's value is in helping its customers determine commercially valuable information on forest growth and crop yield predictions. GSi has developed bespoke machine-learning software that learns to recognise what satellites are "seeing" when covering important global assets such as forestry and crops. By understanding what is on the ground, often across vast areas, GSi provides stakeholders with invaluable ongoing insight into the value of those forests or crops.

The challenge

GSi's service proposition is founded on the GSi-Platform that collates and analyses vast amounts of satellite image data. The first phase of the platform was based on in-house desktop-based compute. This was extremely inefficient in terms of time-to-results, and a better platform architecture was required that provided both efficient data management and faster data analytics.

How we helped

EPCC analysed GSi's business requirements and examined the GSi-Platform workflow and software. We

then moved GSi's code from its desktop and laptop computers to our supercomputers at EPCC, which required some re-engineering of the code and refining of the mathematics to enable it to run better and faster on our HPC systems. We also profiled the code to identify areas that could be fine-tuned to further improve software performance.

By analysing GSi's workflow, we found those areas where we could use parallel computing methods to run jobs to different processors (or cores) on our supercomputer, allowing many analysis tasks to be computed simultaneously.

The benefits

The resultant software is estimated to be 100,000 times quicker than other similar software when run on conventional computing systems; the exponential benefits of running GSi software on HPC machines are huge. As a result of this work, a full set of satellite imagery for the entire Earth's surface can now be analysed in a few hours rather than days.

Further, GSi now has a commercial relationship with EPCC and utilises our HPC-platform-as-a-service to provide the basis of its customer facing large-scale data management and analytics services.

The close working relationship with EPCC and GSi typifies a relevant and vital collaboration between industry and academia that helps a local SME tackle truly global challenges.

"The vast data ingested by the GSi-Platform is stored efficiently at EPCC. High bandwidth, low latency interconnect reduces the need for copy-managing the data through other means. This presents a huge commercial advantage to GSi in reducing time and effort to provide EO analysis of land assets. At all times, EPCC provides expertise and advice to GSi in maximising efficiencies of using HPC in EO and big data management." **Ronnie Galloway, GSi**

EPCC: the UK's leading supercomputing centre

Supercomputing capability straight to your desktop



Introducing Cirrus: our latest computational service for industry.

Cirrus is an SGI ICE XA supercomputer comprising more than 5,000 cores – the equivalent of hundreds of desktop computers.

This facilitates calculations that would be impossible, or much slower, when carried out on conventional desktop computing systems, delivering results in hours to days instead of weeks or months.



Our Accelerator service: supercomputing on demand



Cirrus is just one part of our Accelerator service, which delivers high-performance computing capability at a fraction of the cost of buying and operating in-house HPC services.

Accelerator can be used as a:

- Transformative HPC resource accelerating development and discovery lifecycles
- Flexible HPC resource smoothing out demand peaks
- Contingency over internal HPC infrastructure failure

Accelerator provides access to:

- ARCHER and Blue Gene: our high-end compute systems for large-scale simulation and modelling challenges
- Cirrus: a midrange, industry-standard Linux cluster. An ideal platform for applying commercial software tools to solve a range of CFD and FEA simulation and modelling problems
- RDF: our large-scale data facility giving access to petabyte-scale data storage and archive facilities

With the security of our exceptional expertise

With over 80 highly-qualified permanent staff, we ensure you get the most from our systems. We can help you with:

- General HPC support
- Data management and analytics
- Computational modelling & simulation
- Training & consultancy
- Software development

To discuss our services for business, contact George Graham at EPCC:

g.graham@epcc.ed.ac.uk
+44 (0) 131 651 3460
+44 (0) 777 370 8191