I/O PROFILING TO IMPROVE STORAGE PERFORMANCE AT DIAMOND LIGHT SOURCE ON AN ALTAIR GRID ENGINE CLUSTER

Liam McClean, Senior Software Engineer, Altair / January 12, 2021

Diamond Light Source, the UK's National Synchrotron Science Facility

No two high-performance computing (HPC) teams and architectures will ever be the same, but the team at Diamond Light Source handles a wider variety of workloads than many. Performance of both in-house and third-party applications is therefore vital.

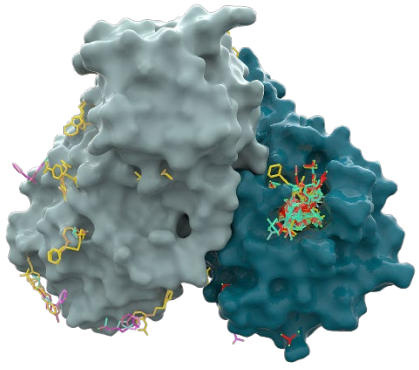
The Diamond Light Source team, which employs Altair® Grid Engine® for workload management, used Altair Mistral™ to identify straightforward improvements that could be made to improve performance and cut down runtime.

> *"By using Mistral, our team has already made a marked improvement to various applications that we maintain in-house. We have been able to reduce the impact of noisy neighbors, reduce runtime and identify applications with bad I/O. We intend to keep using Mistral to profile more applications and improve the overall architecture of our systems for in-house and third-party tools."*
>
> —Fredrik Ferner, senior computer systems administrator, Diamond Light Source

Diamond Light Source

Diamond Light Source is the UK's national synchrotron or particle accelerator. Unlike particle accelerators such as those used in the labs at CERN to look at the result of colliding particles, Diamond Light Source uses the radiation from speeding electrons to beam bright light into a range of experiments arranged around the edges of the accelerator. It works like a giant microscope, harnessing the power of electrons to produce bright light that scientists can use to study anything from fossils to jet engines to viruses and vaccines, including those for COVID-19.

The machine accelerates electrons to near light speeds so that they give off light 10 billion times brighter than the sun. These bright beams are then directed into laboratories known as "beamlines." Here, scientists use the light to study a vast range of subject matter, from new medicines and treatments for disease to innovative engineering and cutting-edge technology.



Representation of the Chemicals Binding to the Main Protease of the SARS-CoV-2 Virus

Each beamline is used in an experiment that generates vast amounts of data, which must be stored in real time and processed quickly. Experiments cover a range of fields from industrial engineering to microbiology and cystalography, which means the HPC team needs to be able to handle a wide range of data rates and compute applications.

I/O Profiling Tools

The Altair Mistral and Altair Breeze™ I/O profiling tools provide visibility into how applications are accessing data. While Mistral offers live system telemetry and scalable I/O profiling, Breeze provides detailed dependency analysis and data hygiene capabilities. The tools provide always-on, system- and storage-agnostic APN monitoring and profiling that allows the user to scale rapidly or migrate to new platforms seamlessly.

Diamond Light Source has been using these tools to profile the wide range of applications run by the team and to check them for data access efficiency. In particular, they have been using Mistral.

Not all applications are written or maintained in-house at Diamond Light Source. The team wanted to see if there were any obvious changes that could be made to the maintained tools to improve performance, or changes to the way third-party applications are used.

The Results

Unnecessary metadata – Following a profile of various workloads using Mistral, the Diamond Light Source team noticed one particular application doing a lot more stat() calls than expected. This eventually turned out to be a number of programs where static files were tested for in a loop even though they either existed at the start of the program or were created by it at the end. The same file was stat'ed many times over, even though there was no change to it throughout the execution. By amending this program, run time was reduced and the metadata load on the file system was improved.

Python can take a long time to load – The profile using Mistral also revealed a number of inefficient program startups. Unfortunate setups with long pythonpath chains resulted in lots of attempts to load python modules from different locations until they were found in the last configured path. The programs searched for the python modules in every location in the path, trawling the file system each time when the module was in the last location in the path. By cutting down the pythonpath settings, the team managed to decrease startup times and increase the performance of the application significantly.

Catching noisy neighbors with live I/O telemetry – Several other findings revealed inefficiencies that slowed the system down, impeding the work of different users. For example, the team discovered an instance where one particular program in a specific invocation completely saturated the network of a number of compute nodes. The program used only a single core for execution and shared node with other jobs on the node, so just by looking at the node it wasn't obvious which program it was. The problem was quickly resolved by looking at the live telemetry from Mistral because Mistral logs not only the hostname, but also the job ID of any noisy neighbors. It's easy to see how the application is overloading the network and preventing other compute nodes on that network from operating.

Conclusion

The Diamond Light Source team has implemented many of the improvements uncovered by Mistral. By reducing runtime and improving the overall architecture, productivity has increased.

