

Anaconda + Potsdam Institute for Climate Impact Research (PIK)

How Anaconda Helps PIK Teams Steward a Brighter Future

The Potsdam Institute for Climate Impact Research (PIK) is a non-profit research institute addressing scientific questions in the fields of global change, climate impact, and sustainable development. PIK plays an important role in providing scientific advice to policymakers, and PIK scientists frequently rank among the most cited researchers in their fields worldwide. The organization has over 400 employees and received 27 million euros in funding from the German government and external sources in 2020. Of their employees, about 100 are users of Anaconda. These users include data scientists, physicists, oceanographers, land-use scientists, economists, and social scientists.

PIK has been using Anaconda since 2015, when they introduced their current high-performance computing (HPC) cluster. An HPC cluster is a collection of various separate servers, known as nodes, which are connected through a fast interconnect. PIK's HPC cluster is a set of 5,000 CPU cores over 350 compute nodes, all connected via a high-speed network to each other and to a high-performance parallel file system which is designed to store data across multiple networked servers. At the time of installation this machine was ranked about 350 in the top 500 fastest supercomputers in the world, with a performance of about 200 teraflops.



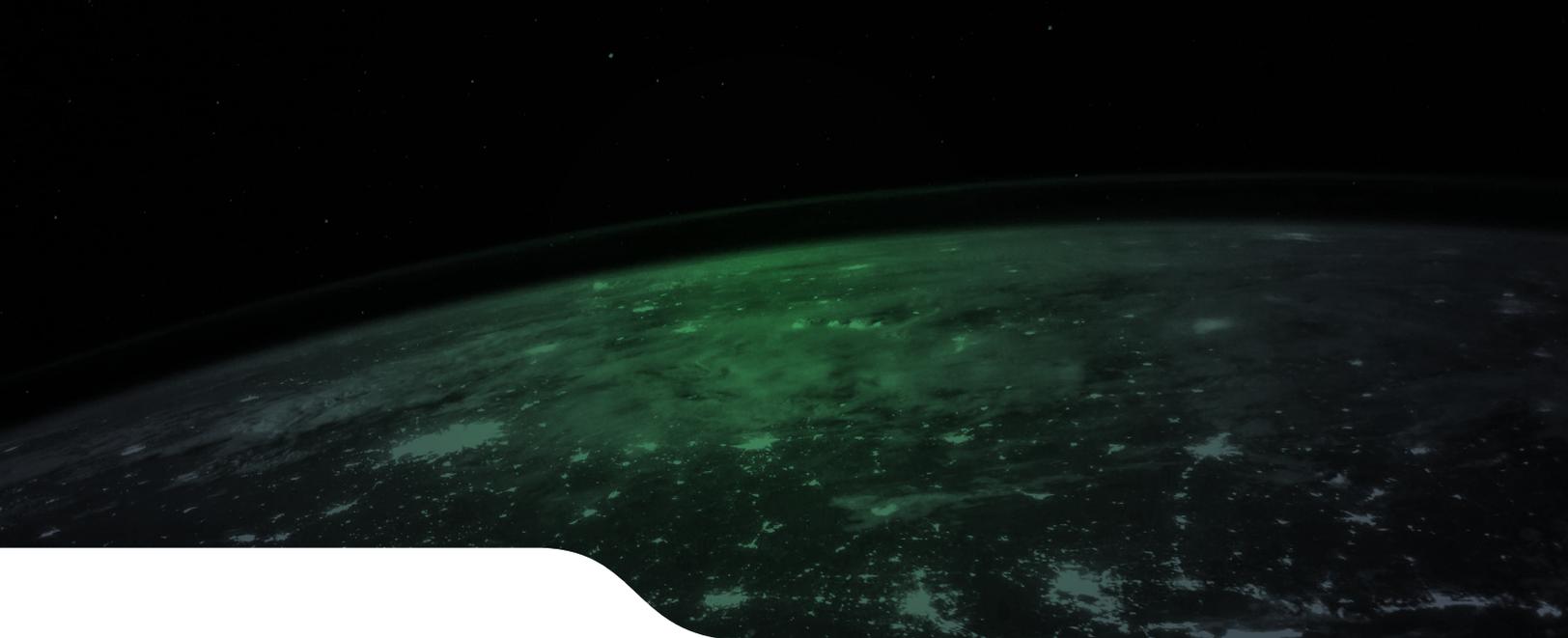
POTSDAM INSTITUTE FOR
CLIMATE IMPACT RESEARCH



Anaconda allows us to use and share the most efficient builds of important numerical [computing] tools for high-quality, reproducible climate science.

Ciaron Linstead

Software Support Engineer, HPC



Anaconda gives PIK scientists more freedom in their choices of software tools, versions, and environments than a traditional HPC system with a centralized software provisioning model can provide. HPC applications typically need to be tuned to make the best possible use of the hardware resources. PIK benefits from Anaconda with its Intel MKL integration, in that it can be sure it has highly efficient builds of the most important numerical and AI Python modules. The integration ensures that PIK scientists have the quickest possible turnaround in their model runs and allows them to concentrate on the science rather than the tooling of environments—which is essential for scientific reproducibility. The quick turnaround makes it much easier for them to try out new techniques and tools.

“My work is part of the HPC section of the IT department. It’s my job to ensure that our scientists have the most up-to-date and efficient tools for developing code, running models and applications, and analyzing results,” said Ciaron Linstead, HPC Software Support Engineer. Anaconda has improved PIK’s data science and machine learning workflows by allowing system administrators to spend less time figuring out how to build important modules and their dependencies from scratch and more time helping scientists get

the best performance out of their code on their HPC cluster. Scientists can create their own Python environments in minutes without waiting for system admin help or approval, which would normally take hours or sometimes days depending on workloads.

Anaconda not only benefits system administrators, but practitioners as well. PIK developers who work in Python, C, C++, Fortran, and R find Anaconda useful for pre- and post-processing of input and output data. Additionally, conda allows practitioners to quickly generate and share software environments, which is essential when it comes to scientific reproducibility—but the practitioner benefits don’t end there.



“We rely on Anaconda to manage the Python packages for our research,” said a team member working on improving climate model simulations with deep generative adversarial networks. “Anaconda is fast and flexible, allowing us, for example, to easily install the latest PyTorch version and its CUDA dependencies. This enables us to train machine learning models efficiently on GPUs.”

Another individual, who is working on predicting dynamic stability of power grids using graph neural networks, had the following to say:

“At PIK we use Anaconda to easily switch between different versions of libraries and can always be up to date with the most recent developments. This is very helpful for conducting research in the quickly evolving field of machine learning.”

Clearly, PIK is an example of an organization that has benefited from adopting Anaconda in a multitude of ways—which, in turn, means that the fields of global change, climate impact, and sustainable development; policymakers; and, ultimately, the people that policy affects (read: all of us) have also benefited. At PIK, Anaconda makes system administrators’ and practitioners’ day-to-day work easier and more efficient and grants scientists more freedom of choice. And through those users and their work, it advances the organization’s mission and betters our world. This is the Anaconda vision—and it can serve your organization, too.

To learn how Anaconda can benefit your organization, reach out to sales@anaconda.com.

With more than 35 million users, Anaconda is the world’s most popular platform to develop and deploy secure Python solutions, faster. We pioneered the use of Python for data science, champion its vibrant community, and steward the open-source projects behind tomorrow’s artificial intelligence (AI) and machine learning (ML) breakthroughs. Our solutions enable practitioners and institutions around the world to securely harness the power of open source for competitive advantage and groundbreaking discoveries. Visit anaconda.com to learn more.