

# Installing Python modules on supercomputers

It is often recommended by desktop Python developers to use tools such as pip or anaconda to install additional Python modules together with their prerequisites. There are important reasons not to do this on supercomputers.

Firstly, these tools often install the modules into your /home directory, which is often very low performance. When Python dynamically loads modules there is excessive load placed on the home filesystem, resulting in both low performance for yourself and everyone else interacting with the supercomputer. This can be somewhat mitigated by installing software in the higher performance /group, but the problem remains. We have seen running python at scale take many minutes before any processing is done, due to Python's startup IO.

The best solution is to install Python modules into a **container**, and use the containers on supercomputers. Pawsey provides several Python containers with common modules pre-installed:

- [Pawsey HPC Python Docker image](#)

Using a container will keep the Python startup IO to within the compute node's memory, which is very fast. Running pip and anaconda to create containers for research is fine if the containers are kept under version control, since you keep a record of the versions of all Python modules used in the research.

If you cannot use containers, we recommend installing specific versions of Python modules into your /group area, along with specific versions of all prerequisite modules. This is important for reproducibility of your research, and ensures the home filesystem is not touched while loading Python. This installation can be tedious, but should be a once-off task. Pawsey's maali tool can be used for this purpose, and if it is not aware of a Python module then you can request help from the Pawsey Service Desk.

## Related articles

- [Exploring files in Lustre](#)
- [Use of tmp by Running Jobs](#)
- [Deleting large numbers of files on scratch and group](#)
- [Python 2 End of Life](#)
- [Installing Python modules on supercomputers](#)