Python Refresher

As a general-purpose programming language with a growing user-base amongst data scientists, Python is increasingly used for data analysis and machine learning applications at the Leibniz Supercomputing Centre (LRZ). In this course we will introduce the different high performance computing resources provided by LRZ and by demonstrating several techniques and best practice examples, we will empower participants to use Python effectively on these systems.

Contents (tentative)

9:15 - 10:30 Python syntax

10:30 - 10:45 Break Q&A

10:45 - 12:00 Python data structures and libraries, conda, virtualenv, pip, pylab

12:00 - 13:00 Lunch Break

13:00 - 14:15 Computing libraries: Numpy / Pandas / Dask.array / theano / tensorflow

Plotting and Graphs: scipy, pylab, matplotlib

14:15 - 14:30 Break Q&A

14:30 - 15:45 Parallel programming: Dask.distributed / MPI4py / ipython jupyter, batch jobs

15:45 - 16:00 Break Q&A

16:00 - 17:15 Native Programming: Cython/f2py/pyCUDA, numba, numexpr

Interactive Programming: ipython, jupyter

Package management: conda

Numerical Arrays: Numpy, numexpr, numba

Compiling to C: Cython

parallel programming: MPI4py

Using Fortran: f2py

Using CUDA: pyCUDA

ML: tensorflow, theano, SymPy

Plots and Graphics: pylab and scipy

Profiling and Debugging

Big Data: pandas, dask