Uni.lu HPC School 2020
PS08: Advanced Distributed Computing with Python

Uni.lu High Performance Computing (HPC) Team

E. Kieffer

University of Luxembourg (UL), Luxembourg

http://hpc.uni.lu

High Performance
Computing &
Big Data Services

hpc.uni.lu

hpc@uni.lu

@ULHPC

Latest versions available on Github:

UL HPC tutorials: https://github.com/ULHPC/tutorials

UL HPC School: http://hpc.uni.lu/hpc-school/

PS08 tutorial sources: ulhpc-tutorials.rtfd.io/en/latest/python/advanced/

Summary

Main Objectives

1. How to parallelise your python code ?

2. Hereafter, we are going to see two alternatives :
   - ↪ High-level approach with ipyparallel for scikit-learn
   - ↪ Low-level approach with scoop

Summary

Scikit-learn + ipyparallel

- **Scikit-learn** is a well-known python scientific package:
  - ↪ Machine learning algorithms (e.g. SVM)
  - ↪ Data analysis approaches (e.g. PCA)
  - ↪ Data mining techniques (e.g. Clustering)

- Scikit-learn algorithms can be parallelised

- Especially useful for Hyper-parameters search

- Scikit-learn relies on **ipyparallel** and **joblib** libraries to parallelism algortihms

Ipyparallel

- Originally designed under **Ipython**

- IPython's architecture for parallel and distributed computing

- Support many different styles of parallelism:
  - ↪ Single program, multiple data (SPMD) parallelism
  - ↪ Multiple program, multiple data (MPMD) parallelism
  - ↪ Message passing using MPI
  - ↪ Task farming
  - ↪ Hybrid approaches combined the above ones

- Ipyparallel can detect a job scheduler (e.g. Slurm) when started on a HPC platform

Practical session

Please go to https://ulhpc-tutorials.readthedocs.io/en/latest/python/advanced/scikit-learn/

Summary

Scoop + deap

- **Deap**

- Python evolutionary computing library:
    - $\hookrightarrow$ Genetic algorithms
    - $\hookrightarrow$ Particle swarm algorithms
    - $\hookrightarrow$ Evolutionary strategies
    - $\hookrightarrow$ Estimation of Distribution algorithms

- Deap relies on **scoop**

Scoop

- SCOOP => Scalable COncurrent Operations in Python

- Applications of SCOOP:
    - ↪ Evolutionary algorithms
    - ↪ Monte Carlo simulations
    - ↪ Data mining
    - ↪ Data processing
    - ↪ Graph traversam

- Very simple to use
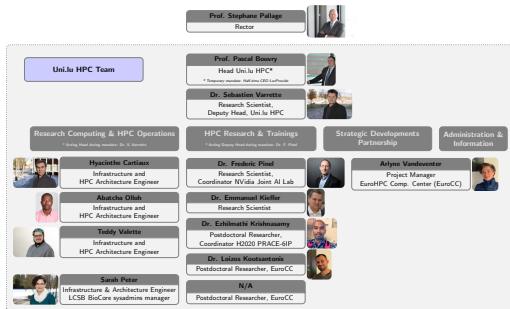
- Override default map (reduce) function

Practical session

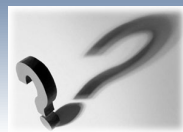Please go to https://ulhpc-tutorials.readthedocs.io/en/latest/python/advanced/scoop-deap/

Thank you for your attention...

Questions?

High Performance Computing @ Uni.lu

- ① **Introduction**
- ② **Parallel machine learning with ipyparallel**
- ③ **Parallel evolutionary computing with scoop**

University of Luxembourg, Belval Campus
Maison du Nombre, 4th floor
2, avenue de l'Université
L-4365 Esch-sur-Alzette
mail: hpc@uni.lu

https://hpc.uni.lu