

Uni.lu HPC School 2019

PS10b: Python II (Advanced) Parallel Machine learning and Evolutionary Computation



Uni.lu High Performance Computing (HPC) Team

E. Kieffer

University of Luxembourg (UL), Luxembourg

<http://hpc.uni.lu>



Latest versions available on Github:



UL HPC tutorials:

<https://github.com/ULHPC/tutorials>

UL HPC School:

<http://hpc.uni.lu/hpc-school/>

PS10b tutorial sources:

ulhpc-tutorials.rtfid.io/en/latest/python/advanced/





Summary

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

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

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

Scikit-learn + ipyparallel

- **Scikit-learn** offers numerous:
 - ↪ 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 algorithms



Ipyparallel

- Originally designed under **lpython**
- 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

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

Scoop + deap

- Deap
- Python evolutionary computing library:
 - ↪ Genetic algorithms
 - ↪ Particle swarm algorithms
 - ↪ Evolutionary strategies
 - ↪ 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/>

Questions?

<http://hpc.uni.lu>

High Performance Computing @ uni.lu

Prof. Pascal Bouvry
Dr. Sebastien Varrette
Valentin Plugaru
Sarah Peter
Hyacinthe Cartiaux
Clement Pariset
Dr. Frédéric Pinel
Dr. Emmanuel Kieffer

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



1 Introduction

- 2 Parallel machine learning with ipyparallel
- 3 Parallel evolutionary computing with scoop