Measuring job's CO2 footprint

There are three main ways in which the use of the HPC clusters can be more taxing for the environment than it needs to be:

1. by using more of the cluster RAM (Random Allocated Memory) than needed for your calculations (i.e., the "job" you submit to the cluster),
2. by having your submitted jobs crash
3. by requesting more cores (i.e., computing units) for a job than needed.

These all imply waste of energy. To help minimize them, the GreenAlgorithms4HPC package was installed on the clusters. It can estimate the carbon output and energy consumption of the user, either for a particular job run on the clusters, or over a time period that you specify. In addition, it can also measure how much memory is being used for the jobs, compared to how much is actually required to run the job.

Green Algorithms

The methodolgy is based on Green Algorithms developed by Loïc Lannelongue. He developed the package GreenAlgorithms4HPC which is a plugin to process the accounting information of a cluster HPC in order to provide an estimation of CO2 footprint.

How to use it

You need to load the following module:

```
module load ga4hpc
```

And then you can check your CO2 footprint for a period of time:

```
green_hpc -S 2025-11-24 -E 2025-11-25
```

The following output is generated:

```
################################

        #
#
        #  Carbon footprint on curnagl  #
        #       - user: cruiz1 -        #
        #   (2025-11-24 / 2025-11-25)   #

################################




              --------------
             |   51 gCO2e
|

              --------------



    ...This is equivalent to:
         - 0.055 tree-months
         - driving 0.29 km
         - 0.0 flights between Paris and London
```

You can also get an estimation for a particular job:

```
green_hpc -S 2025-11-24 -E 2025-11-25 --filterJobIDs 41694290
```

> In oder to have information about jobs running in the same day, you should put the day after in the `-E` parameter.

There are several options to filter jobs that you can check with:

```
green_hpc -h
```

How precise is the estimation?

The power usage is based on the TDP (Termal Desing Power) information provided by the manufacturer. This value is a limit of the power comsomption a CPU, GPU could have. The power consumption is estimated as follows:

Power consumption = time * (resources 1 * TDP + resources_2* TPD + ...)

Assumpions and limitations

- Resources are assumed to be used at a 100%. This may lead to slightly overestimated carbon footprints, although the order of magnitude is probably correct.
- Conversely, the wasted energy due to memory overallocation may be largely underestimated, as the information needed is not always logged.
- Only the carbon imprint of cluster use is measured, not the impact of cooling the computers down, or of building the facilities. The estimation does not take into account neither the CO2 produced during manufacturing.

Results of some tests:

config	appli	GA mesured	real
cpu 48 cores	cpu benchmark NAS	0.343	0.3017
2 gpu A100	julia heat equation	0.355	0.350
2 gpu A100	LLM inference	0.376	0.234

Créé 2025-11-25 13:47:04 UTC par Cristian Ruiz
Mis à jour 2025-11-26 12:40:42 UTC par Cristian Ruiz