

FEATURE 1

Course: Large-scale AI engineering FS25

Author: Maja Gwóźdź

Date: May 15, 2025

Graceful exit mechanism

We present a simple functionality that allows the user to save intermediate checkpoints before timeout is reached (precisely, the checkpoints are created 10 seconds before the timeout, so submitting a job shorter than 10 seconds is incompatible with the feature). The assumption we make is that there are no other errors during the run, so we handle only the timeout and the maximum number of iterations.

The job automatically re-submits if there are still iterations left and it terminates before timeout has been reached if the specified number of iterations has been achieved. The checkpoints are stored in the following path: /iopsstor/scratch/cscs/\$USER/checkpoints.

The method is included in one sbatch file (please see the attached n2.sbatch file) and in the minimal example that runs multiple iterations and receives the signal (please see the attached compute2.py file). The two files can be easily adapted.

Below is a sample terminal output that shows the functionality (the case when the run is 10 seconds before the timeout):

[compute] iter=988200

[compute] iter=988300

[compute] iter=988400

[compute] SIGUSR1 received

[wrapper] Timeout checkpoint → resubmitting...

```
[compute] resumed at 1472381
[compute] iter=1472400
[compute] iter=1472500
[compute] iter=1472600

[compute] iter=999800
[compute] iter=999900
[wrapper] Reached MAX_ITER (1000000) → chain ends.
```

The code has been tested on multiple iterations (either low or high, with the maximum reaching 5 000 000 iterations) and on various timeout values (ranging from 20 seconds to 10 minutes).