Hey George, see you tomorrow at 1pm? Here something you may find useful: https://examples.dask.org/machine-learning/hyperparam-opt.html

One interesting thing to think about when reading this would be: how this differ when applied to EnTK and its task-based model (instead of process/method-based model as with Dask)?

16 batches }

16 pipelines (batches) :
→ 1 stage → 8 Tasks : concurrency using all cores
→ 8 stages → 1 Task : guarantee of sequentiality using 1 core only
→
→

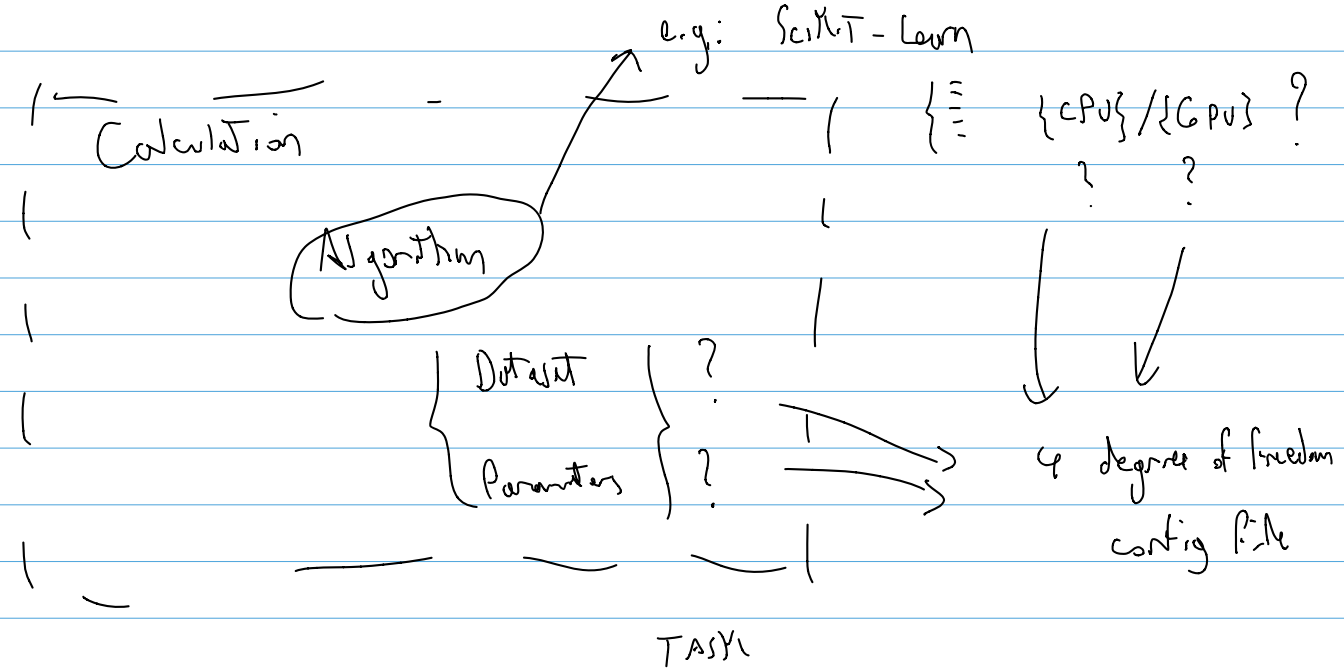
Tuna. project ? or VM

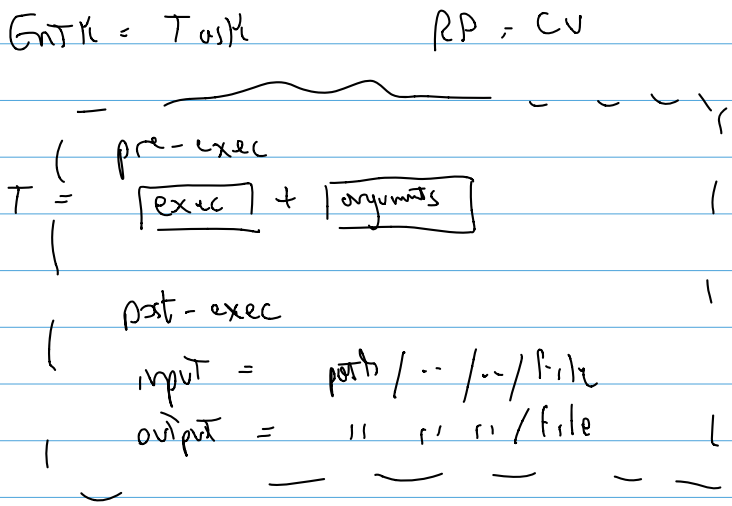
run on VM : max chances of ssh connection not fail (Tunnel)

Tuna. project → RADICAL pays for Two physical machines already near Germany

- → They don't cost us.
- → access the same through ssh on my machine
- → everything preinstalled (rabbitMQ, MongoDB, Docker, GSISSH)

HPO





CPU vs GPU

general — specialized → more efficient, faster

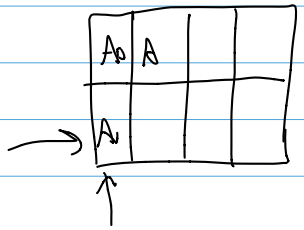
everything in CPU ⟶ can on GPU

but not the way around

distributed vs parallel ?

image :

pixels : (processes)



threads

main diff : independence

distributed is independent

parallel is dependent on IPC communication

Useful resources for distributed / parallel :

http://www.cs.cmu.edu/afs/cs/academic/class/15210-f15/www/tapp.html