



Running DUNE analysis workflows on HPC resources

Paul Laycock

12.10.2022



In a nutshell

- In the Software Frameworks section of the DUNE Computing CDR (going public very soon!) we pointed out that DUNE may well use a separate software framework for analysis compared to large scale data production
 - Likely lightweight and written in python based on current experience
- Analysis includes: extraction of oscillation parameters, comparing data and MC, extracting calibrations, measuring cross-sections of various processes...
- (Some) DUNE analysis is well suited to HPCs and existing lightweight python frameworks like PandAna (developed for NoVA) work well, there have been successful runs at NERSC. PandAna is data parallel and happily utilising MPI already
- funcX is pythonic so it likely plays nicely with something like PandAna...

What could we do?

Scaling up from laptops to LCFs was a desirable feature of the analysis framework in the CDR, and python is an expectation
funcX seems like a good fit

funcX may be best for compute-intensive work, we would want to see where it would work best, e.g. jupyter front-ends (laptop to Analysis Facility?) sending work to powerful backends as and when needed:

- Aim: Work out which workflows funcX would benefit, and how easily funcX can be incorporated into existing tools / frameworks

“Sensitivity throws” with lots of systematics may be a good candidate

Could potentially be interesting to look at ML training as well

Et tu, Belle II?

Belle II is very much jupyter-friendly, with an integrated production and analysis framework (basf2) - relatively heavy analysis fits common

- just to throw even more spaghetti at the wall !