Miss Sadie Bartholomew

March 6, 2020

Regarding the Summer School on Effective HPC for Climate and Weather, 2020

Dear Dr. Julian Kunkel, Dr. Sophie Valcke, Dr. Luciana Pedro & Prof. Bryan Lawrence,

Though I have experience in two effective research software engineer roles in climate & weather area for groups concerned largely with aspects from high-performance computing, I have not partaken in any formal education or training on topics in HPC. There are a number of topics I feel (with pity) ignorant about & though I am trying to read up on many of these, it is hard to find a good entry point assuming minimal initial knowledge & set in a helpful context. For that reason, above all, I think the summer school advertised would be a valuable learning (& networking) opportunity for me, coming at a perfect stage in my career, as I can learn about these exciting topics in a structured & meaningful way from those with experience, & including dedicated practical sessions to solidify the learning objectives.

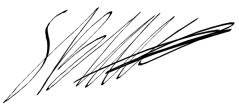
My idea for a project to conduct in the academic group sessions is **a comparison of "high-level" approaches for parallel I/O with netCDF conducted in Python via the Python interfaces *netcdf4-python* & *mpi4py*** to netCDF & MPI respectively. In particular, the aim would be for the group to code up a small number of approaches for writing, & separately reading back, an array to a netCDF file in parallel. I envision there could be (up to) four such approaches implemented (depending on time): to/from netCDF4 using *HDF5* & to/from classic netCDF with *parallel-netCDF*, where for both such cases there would be two variants, one via the MPI-IO supported by *mpi4py* & the other via POSIX. The group would then configure & use *Darshan* to gain insight into the I/O behaviour of each case finally use the knowledge obtained inside (& outside of) the summer school to interpret any patterns in those results. Keywords for this project are: **storage**, **I/O**, **parallel (programming)**, **netCDF**, & **profiling**.

It think such a project would benefit all in the group because it ties together many of the topics discussed in a way that does not require skills beyond the prerequisites listed for the target audience, & it is flexible such that the number of approaches coded up could be adapted based on the progress made mid-week so that the group would move onto Darshan summary stage regardless & almost certainly have interesting results to discuss & present despite any difficulties encountered hindering progress.

To share the gathered information, firstly I would take detailed notes & later revise them carefully with added links to suitable external material to form a self-contained learning resource, probably as a set of Jupyter Notebooks, which I would publish to GitHub. With permission I could share this via the NCAS-CMS site & via mailing lists to members of NCAS & relevant departments at the University of Reading. Moreover, as a member of NCAS-CMS, a group responsible in part for providing support & advice on topics covered in the school & advice to the climate & weather community, I could volunteer to act as a point of contact for researchers in the domain who have questions relating to one of the topics, where I could source & either put them in touch with, or collect reply with relevant information from, a relevant expert who might be able to provide guidance. I have also been planning to start writing technical blog posts, probably for the 'DEV community' site, & the topics taught at the school seem very apt to explore for, & within, published blog posts.

I appreciate that places on the school are limited, especially so with subsidy, but not only I am confident that I would benefit greatly from the school personally, to help me succeed in my current role & environment, & to develop my wider knowledge & skills in HPC, data & computing, if selected I promise to put effort into disseminating the teachings from the school. Thank you for considering my application.

Yours sincerely,



Sadie Bartholomew