

(I (robl) am still working on this document, but wanted to save my work...)

## Big Picture

David A. Wheeler's 'SLOCCount' says we have about 40k lines of code, but a lot of that belongs to the test cases and the generated fortran bindings. There are really just one big directory of code, one big directory of generated fortran wrappers for the code in that first directory, and a directory of tests. When we started parallel-netcdf, we started with netcdf-3.5.0. We changed the programming interface a bit to accommodate parallel I/O constructs like MPI communicators and MPI Info objects. Then we ripped out all the I/O routines and put MPI-IO in its place. If you are familiar with the serial netcdf code, then parallel-netcdf will look familiar to you (though this was much more the case when comparing parallel-netcdf-1.1.0 and netcdf-3.6.3: the latest versions of these two projects are starting to diverge quite a bit).

## src/lib

The biggest contribution parallel-netcdf makes, as mentioned above, is the fact that it uses MPI-IO for I/O, so we'll start there for this walk-through. 1.2.0 introduced a big re-organizing of the source code, resulting in a lot less duplication. All the routines essentially perform subarray I/O via MPI-IO calls. Set the file view with a subarray or hvector and then perform the `MPI_File_read_all` or `MPI_File_write_all` call.