

README

May 2004

This directory contains changes to the ROMS to enable the use of parallel netCDF. The changes were written by the HDF Group at NCSA, the National Center for Supercomputing Applications.

HDF Group

National Center for Supercomputing Applications

University of Illinois at Urbana-Champaign

605 E. Springfield, Champaign IL 61820

<http://hdf.ncsa.uiuc.edu/>

This code requires Parallel NetCDF 0.93 is from Argonne National Laboratory. See:

<http://www-unix.mcs.anl.gov/parallel-netcdf>

This software is an unsupported prototype. Use at your own risk.

This work was funded by the MEAD expedition at the National Center for Supercomputing Applications, NCSA. For more information see:

<http://www.ncsa.uiuc.edu/expeditions/>

Details

This code is an implementation of the ROMS NetCDF history file writer, modified to use parallel netCDF. It was only tested with one case and only on IBM SP platforms. *This code is not even a prototype.* NCSA won't be responsible for or support any uses of this code.

To use it, the following two packages are needed

1. Parallel NetCDF 0.9.3 from Argonne National Lab and Northwestern University (<http://www-unix.mcs.anl.gov/parallel-netcdf/>)
2. ROMS 2.0 from <http://marine.rutgers.edu/po/index.php?model=roms&page=RomsPackages>.

Changes

The following changes were made. The *.diff files list the differences can be found under the source tar file.

1) Source files that need to be modified:

- cppdefs.h -- added another C-preprocessor option, HISTORYP. If this macro is defined, then parallel NetCDF is invoked to generate history file.
- globaldefs.h -- Parallel NetCDF function macros need to be added in this file.
- output.F -- Added ifdef HISTORYP block to generate history file with parallel NetCDF. Added another module USE mod_pnetcdf

- `inp_par.F` -- Around line 140,

```
!-----
!   Check C-preprocessing options and definitions.
!-----
```

A string to store CPP definitions has been input only through Master Node. For parallel IO, that has to be the same for all processors.

This can be fixed in the future by ROMS developers; however, here it is what I did:

```
!-----
! For parallel output, these two lines can be commented out
!       IF (Master) CALL checkdefs
!       IF (Master) CALL my_flush (out)
! For parallel output, all processors should get these values.
CALL checkdefs
CALL my_flush(out)
```

2) Files added to the distribution:

1. Parallel NetCDF IO source files need to be added to ROMS:

```
def_hisp.F
def_infop.F
def_varp.F
mod_pnetcdf.F
nf_fwritep.F
opencdfp.F
wrt_hisp.F
wrt_infop.F
```

These files are modified copies from the ROMS source, the corresponding sequential NetCDF IO source files are:

```
def_his.F
def_info.F
def_var.F
mod_netcdf.F
nf_fwrite.F
opencdf.F
wrt_his.F
wrt_info.F
```

In these file, netCDF functions are renamed, the prefix `nf` is replaced by ROMS `nfmpi` in the parallel NetCDF source files.

Also, the netCDF include file must be changed:

```
#include "pnetcdf.inc"
```

instead of

```
#include "netcdf.inc"
```

3) Makefile and MakeDepend

The Makefile and MakeDepend must be changed.

- Link to parallel NetCDF library needs to be added in Makefile.
- New IO source files need to be added in MakeDepend.

Some implementation tips

We have to assure that each process writes correctly, two special cases for parallel NetCDF implementation are:

1. There are about 20 1-element NetCDF variables in ROMS that have to be written using MPI independent IO through the parallel NetCDF library. The functions *nfmpi_begin_indep_data(int ncid)* and *nfmpi_end_indep_data(int ncid)* need to be included for every independent Parallel NetCDF write function calls. These two functions are used extensively in *wrt_infop.F*.
2. Inside *nf_fwritetp.F*, the array index for parallel IO has to be adjusted since
 - a. the model currently used halo-points for its upper and lower bounds
 - b. raw data at each processor should be written to the disk in parallel.

This case is handled by:

Set *Nghost* = 0; Use subroutine *get_bounds* to obtain the starting and ending points of physical subdomain. Adjust the stagger-C points offset by 1.

Linking with parallel NetCDF

Makefile and MakeDepend should be changed to compile the new modules and to link with parallel netCDF. The history file writer was only tested on IBM SP platforms. On these platforms, *Mpxlf90_r* should be used to compile with MPI-IO library.

Suggestions about ROMS parallel NetCDF work

1. NetCDF Version 4 will support parallel IO. This library will be available from Unidata in 2005. See:
<http://www.unidata.ucar.edu/packages/netcdf-4/>
We recommend users wait for Unidata's NetCDF4 rather than parallel netCDF from Argonne.
2. This code is missing many features. Additional work is needed, such as:
 - Parallel reader
 - Testing on other platforms, right now only IBM SP
 - Optimize one-element array writing and reading
 - Parallel version for other types output files
 - More hybrid IOs, part of parallel and part of sequential