

Parallel-NetCDF: A High Performance API for NetCDF File Access

Overview

Parallel-NetCDF is a library providing high-performance I/O while still maintaining file-format compatibility with Unidata's NetCDF.

NetCDF gives scientific programmers a space-efficient and portable means for storing data. However, it does so in a serial manner, making it difficult to achieve high I/O performance. By making some small changes to the API specified by NetCDF, we can use MPI-IO and its collective operations.

- [Download](#) has the latest release and development links as well as information about svn access.
- [Documentation](#): papers, presentations, articles, and other resources
- [Benchmarking](#): tools and suggestions for evaluating pnetcdf performance

A note about Large File Support

As of parallel-netcdf-0.9.2, we ship with support for "CDF-2" formatted data. With this format, even 32 bit platforms can create netcdf datasets greater than 2GB in size. See the file README.large_files in the source tree for more information.

The maintainers of the serial NetCDF library added support for the CDF-2 format in netcdf-3.6.0. The support was based largely on work from Greg Sjaardema.

File and Variable Limits

Both Parallel-netCDF and NetCDF share limitations on file and variable sizes. More information can be found on the [FileLimits](#) page.

Required Software

Parallel-NetCDF requires an MPI implementation with MPI-IO support. Most MPI libraries have this nowadays. A parallel file system would also go a long way towards achieving highest performance.

Related Projects

Parallel-NetCDF makes use of several other technologies.

- [ROMIO](#), an implementation of MPI-IO, provides optimized collective and noncontiguous operations. It also provides an abstract interface for a large number of parallel file systems.
- One of those file systems ROMIO supports is [PVFS](#), a high performance parallel filesystem for linux clusters.

Mailing List

We discuss the design and use of the Parallel-NetCDF library on the `parallel-netcdf@mcs.anl.gov` mailing list. Anyone interested in developing or using parallel-netcdf is encouraged to join. Visit [the list information page](#) for details.

The URL for the list archive is <http://lists.mcs.anl.gov/pipermail/parallel-netcdf/>. You can browse even older mailing list messages at the older [mailing list archives](#)

Project Members

- Rob Latham, Rob Ross, Rajeev Thakur (Argonne National Lab)
- Kui Gao, Alok Choudhary, Wei-keng Liao (Northwestern University)
- Jianwei Li (NWU, since graduated)
- Bill Gropp (formerly ANL, now UIUC)

Citations

When referring to the Parallel-NetCDF project, please use our "permanent" URL: `www.mcs.anl.gov/parallel-netcdf`. The 'trac' or 'www-unix' URLs could change.

If you are looking for a reference to use in a published paper, please cite our SC2003 paper