Ideas for Testing Parallel NetCDF-4

April 12, 2005
Revised: April 28, 2005

This document defined a test plan for parallel netcdf4. Part 2 describes revisions based on a technical analysis.

Introduction

As part of the HDF-Unidata collaboration, HDF has been asked to work on a test suite for parallel netcdf-4. This note proposes goals and roadmap for this activity.

In work to date, Unidata has written one test for parallel features, along with a simple script to run the test. This was released in the alpha 2 tarball.

The HDF group has added several additional tests that cover the basic features used by WRF: reading and writing two, three, and four dimensional arrays of integer and floats, in independent and collective mode. There is also a test for writing and reading attributes (int and text) in parallel. The test script was hand coded for several platforms.

Extendable datasets were not working in alpha2, so no tests could written for this feature. Also, it is not clear that user defined chunking has been implemented for netcdf-4, so we cannot test this feature yet.

Adding tests has required considerable effort to make the configure and make work correctly. The configure problems urgently need to be addressed, but the test suite should be developed without worrying about the configure. For purposes of testing, the goal is just to get the tests to work, not to debug the overall configure.

The testing so far has already uncovered one bug, and some questions. The overall approach will be to concentrate on developing tests, not on fixing bugs. Problems will be documented and reported to Unidata, and the tests will be completed as well as possible. We will neither try to fix the bugs nor wait for them to be fixed.

Test Goals

The overall goal is to create a reasonable test suite for parallel netcdf-4, and to make it run on at least copper (AIX) and heping (Linux). If all goes well, we will try other platforms.

The design of the tests remains an open question. The initial goal is to use the WRF model as a guide to design the tests. We have done the initial phase of this. The current tests need to be cleaned up, and completed. Remaining work on the current tests includes

- Check the results and give appropriate output in all cases.
- Cover more data types
- Use larger arrays

- Use chunking and filters
- Hyperslab I/O

When these targets have been met, if resources permit, the test suite should be expanded to cover other cases TBD. These might include:

- Cases inspired by the PHDF5 test suite (e.g., fill value, open/close, etc. behavior)
- Cases based on netcdf-4 semantics, e.g., selecting independent vs collective IO.

Test Plan

Here is a proposed test plan.

April 12-22 (using alpha3)

- Get alpha3 working on 2 platforms
- Clean up current tests
- Use larger arrays
- Add more data types

Milestone: new tests work on 2 platforms, copper and heping

- Write note to document missing/broken features that we can't test:
    - Extendable
    - Chunking
    - Filters?
    - Others?

Milestone: report to HDF and Unidata on status of tests

April 25-29

- Add tests for hyperslab I/O

Milestone: new tests working on 2 platforms

May 2-13

- Add additional tests, TBD
    - Based on experience and understanding
    - Based on PHDF5 tests
    - If alpha4 is available test new, fixed features

Milestone: new tests working on 2 platforms

May 16-27

- Port tests to additional platforms
- Document and send tests to Unidata

Milestone: test suite checked in to Unidata CVS

Part 2: Revisions (April 29)

This section documents suggested changes to the Test Goals and Plan above.

Test Goals

Configuration and build of alpha3 has proved to be challenging.

We learned that parallel netCDF4 does not compile on heping (Linux) with the Portland Group compilers (gcc works). This is the default compiler on that platform for our tests, but it isn't critical to debug this problem for this phase of the testing.

The test goals included "Cover more data types". We think it is not important to cover the whole range of data types in parallel, these are covered in the serial tests. We will not expend effort on data type tests.

The tests goals included "Use chunking and filters". Chunking and filters do not seem to implemented in alpha3. These tests are deferred. Similarly, extendible datasets are not working with nc_open.

Test Plan

The original schedule should be discarded. Here is a revised schedule.

April 18-29 (using alpha3)

- Get alpha3 working on 1 platform
- Write note to document missing/broken features that we can't test:
    - Extendable
    - Chunking
    - Filters?
    - Others?

Milestone: report to HDF and Unidata on status of tests

May 2-13

- Clean up current tests
- New tests
    - MPI-IO and MPI-POSIXIO
        - Write_chunk
        - Write_contiguous
        - Read_chunk
        - Read_contiguous
    - Collective IO
        - Write_chunk
        - Write_contiguous
        - Read_chunk
        - Read_contiguous

    - Attributes
    - Test with big file
    - tests for hyperslab I/O

Milestone: new tests working on 2 platforms

May 16-31

- Port tests to additional platforms
- Document and send tests to Unidata

Milestone: test suite checked in to Unidata CVS

June 1-??

- Add additional tests, TBD
    - Based on experience and understanding
    - Based on PHDF5 tests
    - If alpha4 is available test new, fixed features

Milestone: new tests working on 2 platforms

Status

Kent is working on the tests as described in the revised test plan.