

Read and Write Data

Overview of NetCDF

Thanks to all contributors:

Alison Pamment, the Unidata team.

What is NetCDF?

NetCDF is more than just a file format. In the simple view, netCDF is a:

- Data model
- File format
- Application programming interface (API)
- Library implementing the API

Together the data model, file format, and APIs support the creation, access, and sharing of scientific data.

What is netCDF, really?

In the real world, things are more complicated. NetCDF evolves to keep up with needs of its users, and is actually:

- Two data models:
 - **Classic model** (for netCDF-1, netCDF-2, netCDF-3)
 - **Enhanced model** (for netCDF-4)
- Four file formats:
 - **Classic format** and **64-bit offset format** variant
 - **NetCDF-4 format** and **netCDF-4 classic model format** variant
- Two independent flavours of APIs:
 - **C-based** (C, F90, F77, C++, Python, Perl, Ruby, Matlab, ...)
 - **Java-based**

**However, we only use
one of these**

Most scientists, most of the time, will only interact with one of these - so the story isn't so complex after all.

NetCDF Features

Several characteristics make netCDF useful for storing and accessing scientific data.

- **Self-Describing:** A netCDF file may include metadata as well as data: names of variables, data locations in time and space, units of measure, and other useful information.
- **Portable:** Data written on one platform can be read on other platforms.
- **Direct-access:** A small subset of a large dataset may be accessed efficiently, without first reading through all the preceding data.
- **Appendable:** Data may be appended to a netCDF file without copying the dataset or redefining its structure.

NetCDF Features

- **Networkable:** The netCDF library provides client access to structured data on remote servers through OPeNDAP protocols.
- **Extensible:** Adding new dimensions, variables, or attributes to netCDF files does not require changes to existing programs that read the files.
- **Sharable:** One writer and multiple readers may simultaneously access the same netCDF file. With parallel netCDF, multiple writers may efficiently and concurrently write into the same netCDF file.
- **Archivable:** Access to all earlier forms of netCDF data will be supported by current and future versions of the software.

NetCDF's niche

- The netCDF niche is *array-oriented scientific data*.
- It also has a large user community to foster:
 - Support in third-party applications
 - Third-party APIs for many programming and scripting languages
 - Community conventions, such as Climate and Forecast (CF) metadata conventions
 - Standards for interoperability

Installing NetCDF

See documentation:

<https://downloads.unidata.ucar.edu/netcdf/>

(If you want python-netCDF to work you will need the NetCDF libraries written in C).

To install for Python, you can normally just do:

```
pip install netCDF4
```

or:

```
conda install -c conda-forge netcdf4
```


Acknowledgement

The material presented here was primarily taken from the Unidata NetCDF workshop notes at:

<http://www.unidata.ucar.edu/software/netcdf/workshops/2012>