

Tools



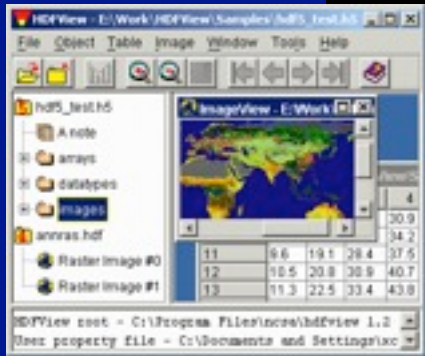
Various tools provide access to HDF5 files, including the data, metadata, and hierarchical structure, without having to write new software.

Commercial off-the-shelf tools supporting HDF5 include IDL, Matlab, Mathematica, Intel® Array Visualizer and Array Viewer, EnSight, and Tecplot.

HDFView, illustrated here, uses multiple panels to display file structure, raw data, metadata, and images. The file represented below is derived from NASA earth observation data; to the right, from black hole gravity wave data. In each case, the full image appears in the background.

For more tools, see <http://hdfgroup.org/tools/>.

Visualization courtesy of John Shelton, NERSC/Lawrence Berkeley Laboratory, using data computed on the NERSC SP2 by Dennis Polney and the Cactus Team, Albert Einstein Institute



The HDF5 technology suite also includes several command line tools. These include utilities for listing, viewing, and managing files and content (h5dump, h5ls, h5diff, h5repack); for data conversion (h5import, gif2h5, h52gif); for examining library performance and managing library and application configuration (h5perf, h5redeploy, h5cc, h5pcc, h5fc, h5pfc, h5c++).

Conversion utilities (h4toh5, h5toh4) and the free-standing H4toH5 Conversion Library facilitate the conversion of files between HDF4 and HDF5 formats.

The HDF Group

The **HDF Group** is a not-for-profit corporation whose mission is to ensure the long-term accessibility to HDF data through the sustainable development and support of HDF technologies.

The HDF Group is dedicated to evolving HDF technologies to serve the needs of users in ever-changing computational environments, while at the same time maintaining its commitment to ensure the accessibility of data stored in HDF for the coming decades, even centuries.

The HDF Group funds its activities through a variety of services focused on helping organizations achieve maximum value from HDF5 technologies.

- Consultation on data models and optimal use of HDF5 features
- Porting the HDF5 library and tools to new platforms
- Installation, maintenance, and priority support
- Custom software development
- Application performance tuning
- Training classes

With over 20 years of experience dealing with large and complex data, and developing HDF software, The HDF Group is uniquely qualified to deliver these specialized services.

The HDF project started at NCSA and the University of Illinois in 1987. The HDF Group completed its transition to an independent corporation in mid-2006.

Contact Information

Website: <http://www.hdfgroup.org/>

Help Desk: help@hdfgroup.org

General inquiries:

<http://www.hdfgroup.org/about/contact.html>

The HDF Group
1800 South Oak Street, Suite 203
Champaign, Illinois 61820



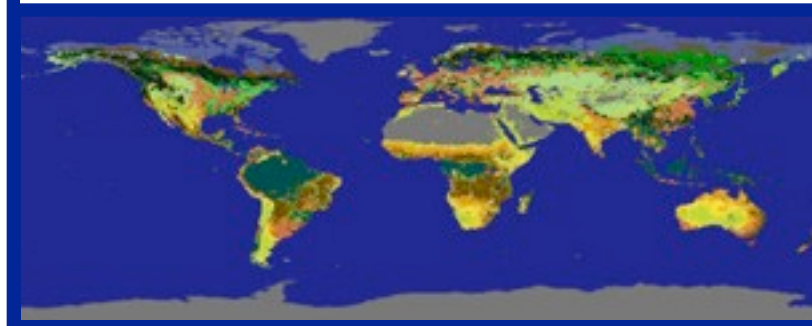
HDF5



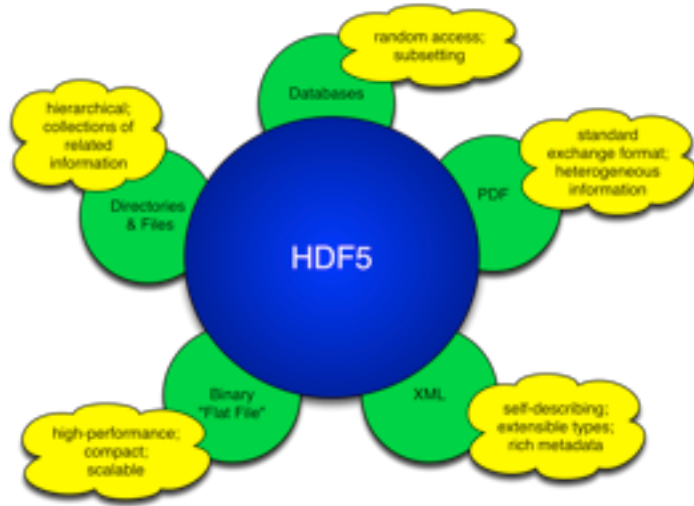
A technology platform to organize, store, share, and access large and complex data

Hierarchical Data Format (HDF) Version 5

- Data model, file format, library, tools
- Portable, self-describing, compact
- A free and open standard



HDF5 is like...



Why HDF5?

- Datasets are big and getting bigger.
- Data comes in many types and structures.
- Metadata takes a variety of forms.
- Data must move from place to place.
- Scientists and policy makers must share data and tools.
- Applications and tools need to create and access data easily.
- Software must work across platforms.
- Applications need parallel I/O on parallel systems.
- Fast I/O enables access to very large datasets.

Building on HDF5

Higher-level libraries, such as NASA's HDF-EOS5 and



Unidata's netCDF-4 use HDF5 as the data management layer.



What is HDF5?

HDF5 is a general-purpose platform for storing, managing, archiving, and exchanging scientific data.

HDF5 consists of an open source data model, file format, library, and tools.

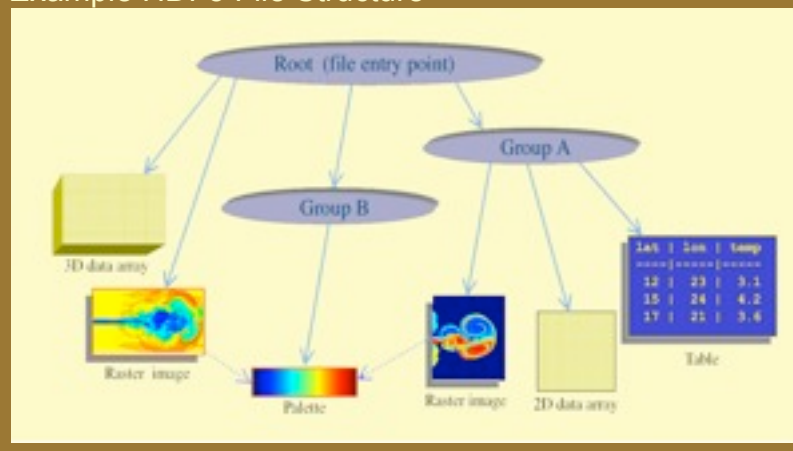
HDF5 can efficiently store almost every kind of scientific data structure, including images, arrays of vectors, structured and unstructured grids, and text.

HDF5 addresses many data management needs of scientists and engineers working in high-performance, data intensive environments.

HDF5 is used worldwide by government, industry, and academia in a wide range of science, engineering, and business disciplines.

HDF5 is developed, owned, licensed, distributed, and maintained by The HDF Group.

Example HDF5 File Structure



HDF5 Features

Flexible and extensible data model and binary file format

- ✓ Unlimited variety of *datatypes*, including integer, floating point, variable-length, compound, and user-defined
- ✓ *Datasets* (multi-dimensional arrays of any datatype) are user-configurable, unlimited in size and number, and extensible
- ✓ User-defined *groups* organize datasets and create hierarchies
- ✓ *Metadata* can be associated with any *object* (dataset or group)
- ✓ *HDF5 files* are unlimited in size and number of objects

Library and API advanced features

- ✓ Compression, encryption, error detection, and user-defined operations can be automatically applied as data is written and read
- ✓ Multiple storage layouts within an HDF5 file support complex data subsetting and reduce access time
- ✓ *Virtual file layer* offers HDF5 file I/O for a variety of storage options including MPI-IO in parallel environments, POSIX-IO, direct-IO, memory, multiple files, and networks

Portable

- ✓ Self-describing data model
- ✓ HDF5 files portable across platforms
- ✓ Library tested with multiple architectures (including 32- and 64-bit), operating systems, and compilers
- ✓ C, Fortran90, C++, and Java APIs

