

hdf5

1 hdf5

package **hdf5**

current version 1.1

TCL package to read and write **hdf5** files. It permits to deal with a subset of the hdf5 options.

Description

An **hdf5** is a file format to store datasets of numbers in an efficient way for scientific or engineering use. It is basically a collection of **groups** and **datasets** organized in a tree shape. The **group** is similar to a directory of the filesystem. The **dataset** is a vector of integer, floats or doubles. The **dataset** can be optionally compressed. Every group or dataset is referenced in a way similar to a file in a filesystem. For example, `"/group1/groups2/dataset1"` refers to the dataset `dataset1` that is stored inside group `group2` that is stored inside `group1`. `group1` is stored in the root of the file.

Every **group** or **dataset** can contain an arbitrary number of **attributes**, which are collections of name, value pairs.

Capabilities of the package:

- It permits to read and write unidimensional vectors of integer, float or double
- For storing a set of coordinates, for example, one would store a vector of integers and three vectors of double
- The vectors are converted to and from a specialized `tcl_obj` that can hold a vector of integer, float or double. If this vector is used in TCL, it can be converted to/from a tcl string or list. When used directly from c or c++ there are not unnecessary conversions
- It is possible to create groups and list or delete objects in them
- It is possible to apply or read string attributes to every dataset or group as strings. Every string is inherently stored as **utf-8**.

Restrictions

In the TCL package it is only possible to deal with unidimensional vectors of integer, float or double values. Attributes can only be vectors of chars (inherently interpreted as **utf-8**).

Commands:

hdf5

`hdf5 handle filename`

handle becomes a tcl command that represents the file and that can be used to write or read from the file.

If *filename* does not exist, it is created. If it exists and it is already a **hdf5** file, it is opened. If not, an error is raised.

handle set

`handle set ?-vtype int|float|double|string? ?-compress 0-9? ?-ncolumns nc? name value ?name value?`

hdf5

creates or updates a dataset. *name* must be the full path name of the dataset (for example: /group1/group2/datasetname). *value* is a vector. If it is already a intarray or doublearray tcl_obj, the type is inferred from it. Otherwise, it has to be specified with option -vtype.

If -vtype is 'string' then value must be a list of strings.

nc is the number of columns to represent a bidimensional array, by default is a single vector (*nc*==1). The array *value* must contain the values ordered by rows (C-like), not by columns (FORTRAN-like), and of course the amount of data must be a multiple of *nc*.

handle get

handle get ?-dimensions? name1 ?name2 ...?

returns a list of all datasets. Every element of the list is either a intarray, floatarray, doublearray or list of strings. The typical use can be:

lassign [*handle* get name1 name2] value1 value2

if *-dimensions* flag is specified then the dimensions of the array on each direction are returned instead the data.

handle create_group

handle create_group name1 ?name2 ...?

Creates one or more groups. Parent group needs to exist.

handle delete

handle delete name1 ?name2 ...?

Deletes one or more groups or datasets.

handle set_attribute

handle set_attribute ?-type integer|string? obj_name att_name1 value1 ?att_name2 value_2...?

Creates one or more attributes applied to one existing group or dataset.

type is by default set to *integer*, and is a classical vector of chars ended by a NULL char. (inherently interpreted as **utf-8**), the other option is *string* and then a special HDF5 string structure will be used to store the data.

handle get_attribute

handle get_attribute ?-default def_value? obj_name ?att_name?

Returns the values of one attribute. If *att_name* is not given, it returns a dictionary of all name value attributes that the object contains.

if option *-default* is given, when attribute does not exist, the *default_value* is returned. Otherwise, an error is returned

handle glob

handle glob ?-directory dir? ?-types all|group|dataset? pattern

hdf5

Returns a list of groups or datasets. pattern can be a relative or absolute path. It can only contain substitution characters like the "*" in the last component of the path

if option *-default* is given, when the attribute does not exist, the *default_value* is returned. Otherwise, an error is returned.

handle is_group

handle is_group name

Returns true if name is a group.

handle is_dataset

handle is_dataset name

Returns true if name is a dataset.

handle close

handle close

Closes the file and releases the handle.

News

- From version 1.1

-vtype string in command *handle* set, to allow the creation of string sets.

-dimensions option in the command *handle* get, to know the dimensions of multidimensional arrays

- From version 1.0

First version.