

# Groups

Creating and manipulating groups of objects inside an HDF5 file (H5G)

- `H5G_CLOSE`
- `H5G_CREATE`
  - `H5G_CREATE1` \*
  - `H5G_CREATE2`
- `H5G_CREATE_ANON`
- `H5G_FLUSH`
- `H5G_GET_COMMENT` \*
- `H5G_GET_CREATE_PLIST`
- `H5G_GET_INFO`
- `H5G_GET_INFO_BY_IDX`
- `H5G_GET_INFO_BY_NAME`
- `H5G_GET_LINKVAL` \*
- `H5G_GET_NUM_OBJS` \*
- `H5G_GET_OBJINFO` \*
- `H5G_GET_OBJNAME_BY_IDX` \*
- `H5G_GET_OBJTYPE_BY_IDX` \*
- `H5G_ITERATE` \*
- `H5G_LINK` \*
- `H5G_LINK2` \*
- `H5G_MOVE` \*
- `H5G_MOVE2` \*
- `H5G_OPEN`
  - `H5G_OPEN1` \*
  - `H5G_OPEN2`
- `H5G_REFRESH`
- `H5G_SET_COMMENT` \*
- `H5G_UNLINK` \*

\* Use of these functions is deprecated in Release 1.8.0.

---

## Groups in HDF5:

A group associates names with objects and provides a mechanism for mapping a name to an object. Since all objects appear in at least one group (with the possible exception of the root object) and since objects can have names in more than one group, the set of all objects in an HDF5 file is a directed graph. The internal nodes (nodes with out-degree greater than zero) must be groups while the leaf nodes (nodes with out-degree zero) are either empty groups or objects of some other type. Exactly one object in every non-empty file is the root object. The root object always has a positive in-degree because it is pointed to by the file super block.

## Locating objects in the HDF5 file hierarchy:

An object name consists of one or more components separated from one another by slashes. An absolute name begins with a slash and the object is located by looking for the first component in the root object, then looking for the second component in the first object, etc., until the entire name is traversed. A relative name does not begin with a slash and the traversal begins at the location specified by the create or access function.

## Group implementations in HDF5:

The original HDF5 group implementation provided a single indexed structure for link storage. A new group implementation, in HDF5 Release 1.8.0, enables more efficient compact storage for very small groups, improved link indexing for large groups, and other advanced features.

- The *original indexed* format remains the default. Links are stored in a B-tree in the group's local heap.
- Groups created in the new *compact-or-indexed* format, the implementation introduced with Release 1.8.0, can be tuned for performance, switching between the compact and indexed formats at thresholds set in the user application.
  - The *compact* format will conserve file space and processing overhead when working with small groups and is particularly valuable when a group contains no links. Links are stored as a list of messages in the group's header.
  - The *indexed* format will yield improved performance when working with large groups, e.g., groups containing thousands to millions of members. Links are stored in a fractal heap and indexed with an improved B-tree.
- The new implementation also enables the use of link names consisting of non-ASCII character sets (see `H5P_SET_CHAR_ENCODING`) and is required for all link types other than hard or soft links, e.g., external and user-defined links (see the `H5L APIs`).

The original group structure and the newer structures are not directly interoperable. By default, a group will be created in the original indexed format. An existing group can be changed to a compact-or-indexed format if the need arises; there is no capability to change back. As stated above, once in the compact-or-indexed format, a group can switch between compact and indexed as needed.

Groups will be initially created in the compact-or-indexed format only when one or more of the following conditions is met:

- The *low version bound* value of the *library version bounds* property has been set to Release 1.8.0 or later in the file access property list (see `H5P_SET_LIBVER_BOUNDS`). Currently, that would require an `H5P_SET_LIBVER_BOUNDS` call with the *low* parameter set to `H5F_LIBVER_LATEST`.

When this property is set for an HDF5 file, all objects in the file will be created using the latest available format; no effort will be made to create a file that can be read by older libraries.

- The creation order tracking property, `H5P_CRT_ORDER_TRACKED`, has been set in the group creation property list (see `H5P_SET_LINK_CREATION_ORDER`).

An existing group, currently in the original indexed format, will be converted to the compact-or-indexed format upon the occurrence of any of the following events:

- An external or user-defined link is inserted into the group.
- A link named with a string composed of non-ASCII characters is inserted into the group.

The compact-or-indexed format offers performance improvements that will be most notable at the extremes, i.e., in groups with zero members and in groups with tens of thousands of members. But measurable differences may sometimes appear at a threshold as low as eight group members. Since these performance thresholds and criteria differ from application to application, tunable settings are provided to govern the switch between the compact and indexed formats (see [H5P\\_SET\\_LINK\\_PHASE\\_CHANGE](#)). Optimal thresholds will depend on the application and the operating environment.

Future versions of HDF5 will retain the ability to create, read, write, and manipulate all groups stored in either the original indexed format or the compact-or-indexed format.