

Paper for Consideration by S-100WG

Recommended remove of the Attribute chunking

Submitted by:	BSH (Germany)
Executive Summary:	Removing the attribute chunking in Part 10c of S-100 Specification.
Related Documents:	S-100 Specification Edition 4.0.0 HDF5 Users Guide 1.11.x
Related Projects:	S-102 Product Specification and all Product Specifications using HDF5 file format

Introduction / Background

The BSH has undertaken the task to enable the creation of S-102 datasets and has identified an area for improvement in the S-100 specification to provide more clarity for implementation. Findings from work to support the creation of the S-102 datasets indicate the requirement for a change to the definition of the chunking size.

Analysis/Discussion

Part 10c of the S-100 specification describes the storage of data in HDF5 file format. The Data can be stored in multidimensional arrays. S-102 uses a two-dimensional array for storing the bathymetry coverage. Such an array is stored in the HDF5 file format as a Dataset [see S-100 Chapter 10c-5.1.3]. Each Dataset has a property list [see S-100 Chapter 10c-5.1.7] and may have attributes [see S-100 Chapter 10c-5.1.6].

A property list is a collection of *name-value* pairs. This list contains information to the storage strategy. If the storage strategy is “chunked” then the chunking size is also stored in the property list. These are parts of the HDF5 file format specification of the HDF-Group [see HDF5 Users Guide, Chapter 5.4].

In S-100 Table 10c-8, an attribute called *chunking* is defined. Its value describes the chunking dimension size for data values of this feature as string representation.

Conclusions

S-100 defines a redundant attribute for the chunking size. This definition causes data redundancy and inconsistency, which should be avoided. The HDF library already implements a functionality for storing Datasets in chunks.

The property list is part of the HDF5 specification and the use of the information it contains is automatically controlled by the HDF library. Reading a HDF5 file using the storage strategy mechanism of HDF library is transparent for each application.

Recommendations

We recommend removing the attribute *chunking* from S-100 specification.

The attribute is mentioned in the following passages of the S-100:

- Table 10c-2
- Table 10c-8
- Chapter 10c-9.5 (in the text)
- Figure 10c-8
- Table 10c-12

Justification and Impacts

Reduction of redundancies avoids the risk of inconsistency. Furthermore, it optimizes the data size.

Action Required of [HSSC][Relevant HSSC WG]

The [HSSC] [Relevant HSSC WG] is invited to:

- a. endorse
- b. agree
- c. note