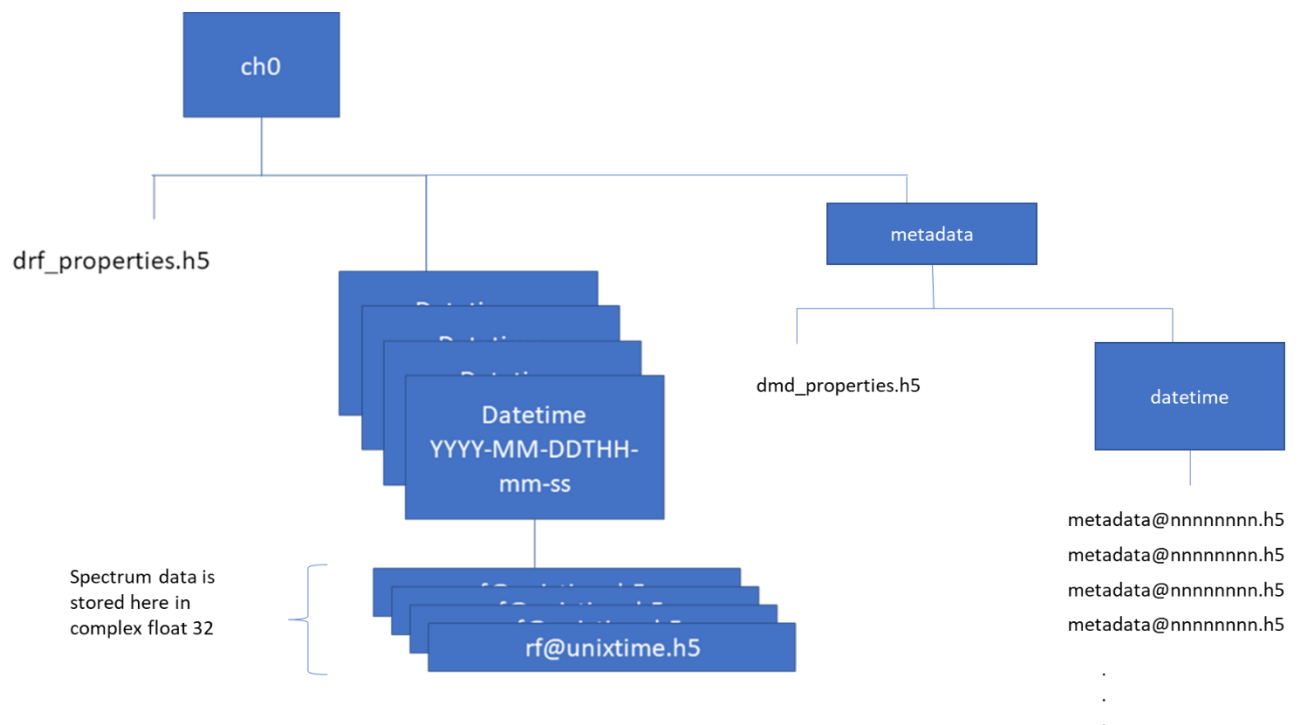


An Introduction to Digital RF

Digital RF (DRF) is a data storage design that is optimized for saving radio spectrum data. Developed by MIT Haystack Observatory, it allows you to include metadata both about the entire package of data being saved, and also about individual measurements (if you need that level of detail) (for details, see https://github.com/MITHaystack/digital_rf). Behind the scenes, it uses the HDF5 (“Hierarchical Data Format”) protocol (refer to <https://hdfgroup.org/>), a standard for saving scientific/mathematical data.

HDF (and DRF in particular) store data in a “dataset” format that is more like a database than a conventional file. This means that you can’t simply view/edit it using a text editor or Excel. Rather, you must use the necessary utilities that understand how to navigate through the binary file structure of the datasets. This minor inconvenience is compensated for the flexibility of the format, which allows you to store a variety of different data types, including lists, arrays, and images and associate metadata (descriptions of the data) with the stored information at every level. Software to navigate inside DRF files is available in several languages; for this project, we use Python. The following illustrates the logical architecture of a typical DRF dataset; the blue boxes are directories.



From upper left top, we have:

1. ch0 – by convention, the top level directory of a DRF dataset.
2. drf_properties.h5 – metadata pertaining to the entire dataset
3. Datetime directories – titled with datetime, these directories contain the detailed spectrum data files

4. Under the datetime directories, individual files contain the spectrum data. If this DRF dataset contains more than one channel (i.e., monitoring of multiple frequencies such as 5 MHz, 10 MHz, 15 MHz), these are stored as columns in these files. The data are complex (real and imaginary, each in a column), 32-bit floating point.
5. Metadata: under this directory is a properties file and one or more datetime directories, which contain a series of metadata files; these are used in cases where individual spectrum data packages or measurements have metadata associated with them at the detailed level. (Note that the gnuradio DRF code stores metadata down at this level).

Needless to say, this sophisticated dataset format requires some special handling at every step. This is the reason for the relatively complex code for uploading your collected data to the PSWS Network Central Control System; Observations are stored in one-day packages (about 35 MB each for 10 Hz bandwidth monitoring).

If you wish to work with the data, sample code is available from sources such as:

- The Digital RF repository in github.com contains DRF sample code and documentation
- Sample code for analyzing and plotting spectrum data is available from the author
- For getting into the low level details, you can find documentation on the HDF internal format at the hdfgroup.org website