Australia's National
Science Agency

SDHDF: Formal Definition

Formal definition for Single Dish Hierarchical Data Format – a file format for spectral line and continuum data from single dish radio telescopes

Lawrence Toomey

V2.1 (July 31st 2021)

Table of Contents

The SDHDF Definition (v2.1)

SDHDF Definition Overview		
SDHDF Definition Version: 2.1		
Author: Lawrence Toomey		
Copyright: CSIRO 2021		
SDHDF File Overview		
HDF_Object_Name	*HDF_Object_Type*	*Value*
file.hdf	File	SDHDF format file
SDHDF Structure Overview		
HDF_Object_Name	*HDF_Object_Type*	*Value*
/beam_N	Group	-
DESCRIPTION	Attribute	SDHDF group containing data products specific to the antenna beam
NAME	Attribute	beam_N
SDHDF_CLASS	Attribute	sdhdf_beam
/beam_N/band_N	Group	
DESCRIPTION	Attribute	SDHDF group containing data products specific to the frequency band
NAME	Attribute	band_N
SDHDF_CLASS	Attribute	sdhdf_band
/beam_N/band_N/astronomy_data	Group	-
DESCRIPTION	Attribute	SDHDF group containing observation data
NAME	Attribute	astronomy_data
SDHDF_CLASS	Attribute	sdhdf_data
/beam_N/band_N/astronomy_data/data	Dataset	-
DESCRIPTION	Attribute	SDHDF astronomy data
DIMENSION_LABELS	Attribute	time, polarization, frequency, bin
DIMENSION_LIST	Attribute	List of dimension objects
NAME	Attribute	data
NORMALISATION_FACTOR	Attribute	The factor by which the number of samples integrated into each output sample are normalised
SDHDF_CLASS	Attribute	sdhdf_waterfall
UNIT	Attribute	counts
/beam_N/band_N/astronomy_data/frequency	Dataset	-
CLASS	Attribute	DIMENSION_SCALE

DESCRIPTION	Attribute	Weighted mean of channel centre frequencies
FRAME	Attribute	topocentric
NAME	Attribute	frequency
REFERENCE_LIST	Attribute	HDF5 object reference
SDHDF_CLASS	Attribute	sdhdf_frequency
UNIT	Attribute	MHz
/beam_N/band_N/astronomy_data/data_weights	Dataset	-
DESCRIPTION	Attribute	Product of the channel bandwidth (Hz) and integration time (s)
NAME	Attribute	data_weights
SDHDF_CLASS	Attribute	sdhdf_weights
/beam_N/band_N/astronomy_data/data_flags	Dataset	-
DESCRIPTION	Attribute	Flag applied to a frequency channel (boolean)
NAME	Attribute	data_flags
SDHDF_CLASS	Attribute	sdhdf_flags
/beam_N/band_N/calibrator_data/cal32_data	Dataset	-
DESCRIPTION	Attribute	SDHDF calibration dataset (calibrator signal 32 bins)
DIMENSION_LABELS	Attribute	time, polarization, frequency, bin
DIMENSION_LIST	Attribute	List of dimension objects
NAME	Attribute	cal32_data
NORMALISATION_FACTOR	Attribute	The factor by which the number of samples integrated into each output bin are normalised
SDHDF_CLASS	Attribute	sdhdf_waterfall
UNIT	Attribute	counts
/beam_N/band_N/calibrator_data/cal_data_off	Dataset	-
DESCRIPTION	Attribute	SDHDF calibration dataset (calibrator signal off)
DIMENSION_LABELS	Attribute	time, polarization, frequency, bin
DIMENSION_LIST	Attribute	List of dimension objects
NAME	Attribute	cal_data_off
NORMALISATION_FACTOR	Attribute	The factor by which the number of samples integrated into each output bin are normalised
SDHDF_CLASS	Attribute	sdhdf_waterfall
UNIT	Attribute	counts
/beam_N/band_N/calibrator_data/cal_data_on	Dataset	-
DESCRIPTION	Attribute	SDHDF calibration dataset (calibrator signal on)

DIMENSION_LABELS	Attribute	time, polarization, frequency, bin
DIMENSION_LIST	Attribute	List of dimension objects
NAME	Attribute	cal_data_on
NORMALISATION_FACTOR	Attribute	The factor by which the number of samples integrated into each output bin are normalised
SDHDF_CLASS	Attribute	sdhdf_waterfall
UNIT	Attribute	counts
/beam_N/band_N/calibrator_data/cal_frequency	Dataset	-
CLASS	Attribute	DIMENSION_SCALE
DESCRIPTION	Attribute	Weighted mean of channel centre frequencies
FRAME	Attribute	topocentric
NAME	Attribute	cal_frequency
REFERENCE_LIST	Attribute	HDF5 object reference
SDHDF_CLASS	Attribute	sdhdf_frequency
UNIT	Attribute	MHz
/beam_N/band_N/calibrator_data/cal_data_weights	Dataset	-
DESCRIPTION	Attribute	Product of the channel bandwidth (Hz) and integration time (s)
NAME	Attribute	cal_data_weights
SDHDF_CLASS	Attribute	sdhdf_weights
/beam_N/band_N/calibrator_data/cal_data_flags	Dataset	-
DESCRIPTION	Attribute	Flag applied to a frequency channel (boolean)
NAME	Attribute	cal_data_flags
SDHDF_CLASS	Attribute	sdhdf_flags
/beam_N/band_N/metadata	Group	-
DESCRIPTION	Attribute	SDHDF group containing observation metadata
NAME	Attribute	metadata
SDHDF_CLASS	Attribute	sdhdf_meta
/beam_N/band_N/metadata/cal_obs_params	Dataset	-
AEST	Attribute	Australian Eastern Standard Time (HH:MM:SS.s)
AEST_UNIT	Attribute	hours
AZ_OFFSET	Attribute	J2000 Ax (Cross-El) offset
AZ_OFFSET_UNIT	Attribute	arc-sec
AZ	Attribute	Antenna azimuth angle
AZ_DRIVE_RATE	Attribute	Azimuth drive rate
AZ_DRIVE_RATE_UNIT	Attribute	degrees/min

AZ_UNIT	Attribute	degrees
DEC_DEG	Attribute	J2000 declination
DEC_DEG_UNIT	Attribute	degrees
DEC_OFFSET	Attribute	J2000 DEC offset
DEC_OFFSET_UNIT	Attribute	arc-sec
DEC_STR	Attribute	J2000 declination
DEC_STR_UNIT	Attribute	DD:MM:SS.s
DESCRIPTION	Attribute	Metadata specific to the integrations of the observation
EL	Attribute	Antenna elevation angle
ELAPSED_TIME	Attribute	Time since integration start
ELAPSED_TIME_UNIT	Attribute	s
EL_OFFSET	Attribute	J2000 El offset
EL_OFFSET_UNIT	Attribute	arc-sec
EL_UNIT	Attribute	degrees
GB	Attribute	Galactic latitude
GB_UNIT	Attribute	degrees
GL	Attribute	Galactic longitude
GL_UNIT	Attribute	degrees
HOUR_ANGLE	Attribute	Hour angle
HOUR_ANGLE_UNIT	Attribute	degrees
MJD	Attribute	Timestamp at integration midpoint (MJD)
MJD_UNIT	Attribute	days
NAME	Attribute	cal_obs_params
PARA_ANGLE	Attribute	Parallactic angle
PARA_ANGLE_UNIT	Attribute	degrees
RA_DEG	Attribute	J2000 right ascension
RA_DEG_UNIT	Attribute	degrees
RA_OFFSET	Attribute	J2000 RA offset
RA_OFFSET_UNIT	Attribute	s
RA_STR	Attribute	J2000 right ascension
RA_STR_UNIT	Attribute	HH:MM:SS.s
SDHDF_CLASS	Attribute	sdhdf_table
TIME_DB	Attribute	Database Universal Coordinated Time (HH:MM:SS.s)
TIME_DB_UNIT	Attribute	hours
UTC	Attribute	Timestamp at integration midpoint (UTC)
UTC_UNIT	Attribute	hours
UT_DATE	Attribute	Date at integration midpoint (UTC YYYY-MM-DD-hh:mm:ss)
WIND_DIR	Attribute	Wind direction
WIND_DIR_UNIT	Attribute	degrees
WIND_SPD	Attribute	Wind speed
WIND_SPD_UNIT	Attribute	km/hr

ZE	Attribute	Antenna zenith angle
ZE_DRIVE_RATE	Attribute	Zenith drive rate
ZE_DRIVE_RATE_UNIT	Attribute	degrees/min
ZE_UNIT	Attribute	degrees
/beam_N/band_N/metadata/obs_params	Dataset	-
AEST	Attribute	Australian Eastern Standard Time (HH:MM:SS.s)
AEST_UNIT	Attribute	hours
AZ_OFFSET	Attribute	J2000 Ax (Cross-El) offset
AZ_OFFSET_UNIT	Attribute	arc-sec
AZ	Attribute	Antenna azimuth angle
AZ_DRIVE_RATE	Attribute	Azimuth drive rate
AZ_DRIVE_RATE_UNIT	Attribute	degrees/min
AZ_UNIT	Attribute	degrees
DEC_DEG	Attribute	J2000 declination
DEC_DEG_UNIT	Attribute	degrees
DEC_OFFSET	Attribute	J2000 DEC offset
DEC_OFFSET_UNIT	Attribute	arc-sec
DEC_STR	Attribute	J2000 declination
DEC_STR_UNIT	Attribute	DD:MM:SS.s
DESCRIPTION	Attribute	Metadata specific to the integrations of the observation
EL	Attribute	Antenna elevation angle
ELAPSED_TIME	Attribute	Time since integration start
ELAPSED_TIME_UNIT	Attribute	s
EL_OFFSET	Attribute	J2000 El offset
EL_OFFSET_UNIT	Attribute	arc-sec
EL_UNIT	Attribute	degrees
GB	Attribute	Galactic latitude
GB_UNIT	Attribute	degrees
GL	Attribute	Galactic longitude
GL_UNIT	Attribute	degrees
HOUR_ANGLE	Attribute	Hour angle
HOUR_ANGLE_UNIT	Attribute	degrees
MJD	Attribute	Timestamp at integration midpoint (MJD)
MJD_UNIT	Attribute	days
NAME	Attribute	obs_params
PARA_ANGLE	Attribute	Parallactic angle
PARA_ANGLE_UNIT	Attribute	degrees
RA_DEG	Attribute	J2000 right ascension
RA_DEG_UNIT	Attribute	degrees
RA_OFFSET	Attribute	J2000 RA offset
RA_OFFSET_UNIT	Attribute	s

RA_STR	Attribute	J2000 right ascension
RA_STR_UNIT	Attribute	HH:MM:SS.s
SDHDF_CLASS	Attribute	sdhdf_table
TIME_DB	Attribute	Database Universal Coordinated Time (HH:MM:SS.s)
TIME_DB_UNIT	Attribute	hours
UTC	Attribute	Timestamp at integration midpoint (UTC)
UTC_UNIT	Attribute	hours
UT_DATE	Attribute	Date at integration midpoint (UTC YYYY-MM-DD-hh:mm:ss)
WIND_DIR	Attribute	Wind direction
WIND_DIR_UNIT	Attribute	degrees
WIND_SPD	Attribute	Wind speed
WIND_SPD_UNIT	Attribute	km/hr
ZE	Attribute	Antenna zenith angle
ZE_DRIVE_RATE	Attribute	Zenith drive rate
ZE_DRIVE_RATE_UNIT	Attribute	degrees/min
ZE_UNIT	Attribute	degrees
/beam_N/metadata	Group	-
DESCRIPTION	Attribute	SDHDF group containing observation metadata
NAME	Attribute	metadata
SDHDF_CLASS	Attribute	sdhdf_meta
/beam_N/metadata/band_params	Dataset	-
CENTRE_FREQ	Attribute	Band centre frequency
CENTRE_FREQ_UNIT	Attribute	MHz
DESCRIPTION	Attribute	Metadata specific to the frequency bands of the antenna beam
DUMP_TIME	Attribute	Band dump time
DUMP_TIME_UNIT	Attribute	s
HIGH_FREQ	Attribute	Band range high frequency
HIGH_FREQ_UNIT	Attribute	MHz
LABEL	Attribute	Band label
LOW_FREQ	Attribute	Band range low frequency
LOW_FREQ_UNIT	Attribute	MHz
NAME	Attribute	band_params
N_CHANS	Attribute	Number of channels in band
N_DUMPS	Attribute	Number of spectral data dumps in band
N_POLS	Attribute	Number of polarisations (1, 2, 4)
POL_TYPE	Attribute	Polarisation type (AABBCRCI for 4 pol coherence data where AA and BB are the direct products of the two input channels A and B and CR and CI are the real and imaginary

		parts of the cross product A* B; AA+BB for 1 pol data with summed orthogonal products; AABB for 2 pol data)
SDHDF_CLASS	Attribute	sdhdf_table
/beam_N/metadata/cal_band_params	Dataset	-
CENTRE_FREQ	Attribute	Band centre frequency
CENTRE_FREQ_UNIT	Attribute	MHz
DESCRIPTION	Attribute	Metadata specific to the frequency bands of the antenna beam
DUMP_TIME	Attribute	Band dump time
DUMP_TIME_UNIT	Attribute	s
HIGH_FREQ	Attribute	Band range high frequency
HIGH_FREQ_UNIT	Attribute	MHz
LABEL	Attribute	Band label
LOW_FREQ	Attribute	Band range low frequency
LOW_FREQ_UNIT	Attribute	MHz
NAME	Attribute	band_params
N_CHANS	Attribute	Number of channels in band
N_DUMPS	Attribute	Number of spectral data dumps in band
N_POLS	Attribute	Number of polarisations (1, 2, 4)
POL_TYPE	Attribute	Polarisation type (AABBCRCI for 4 pol coherence data where AA and BB are the direct products of the two input channels A and B and CR and CI are the real and imaginary parts of the cross product A* B; AA+BB for 1 pol data with summed orthogonal products; AABB for 2 pol data)
SDHDF_CLASS	Attribute	sdhdf_table
/config	Group	-
DESCRIPTION	Attribute	SDHDF group containing configuration parameters as defined at the time of the observation
NAME	Attribute	config
SDHDF_CLASS	Attribute	sdhdf_config
/config/backend_config	Dataset	-
DESCRIPTION	Attribute	Astronomy backend configuration
NAME	Attribute	backend_config
SDHDF_CLASS	Attribute	sdhdf_table
/config/cal_backend_config	Dataset	-
DESCRIPTION	Attribute	Astronomy backend configuration (calibration)

NAME	Attribute	backend_config
SDHDF_CLASS	Attribute	sdhdf_table
/metadata	Group	-
DESCRIPTION	Attribute	SDHDF group containing observation metadata
NAME	Attribute	metadata
SDHDF_CLASS	Attribute	sdhdf_meta
/metadata/beam_params	Dataset	-
DESCRIPTION	Attribute	Metadata specific to the antenna beam
LABEL	Attribute	Beam label
NAME	Attribute	beam_params
N_BANDS	Attribute	Number of frequency bands
SDHDF_CLASS	Attribute	sdhdf_table
SOURCE	Attribute	Source name
/metadata/history	Dataset	-
DATE	Attribute	Date (UTC YYYY-MM-DD-hh:mm:ss)
DESCRIPTION	Attribute	Metadata specific to the processing history of the file
NAME	Attribute	history
PROC	Attribute	Process name
PROC_ARGS	Attribute	Process command arguments
PROC_DESCR	Attribute	Process description
PROC_HOST	Attribute	Host machine running the process
SDHDF_CLASS	Attribute	sdhdf_table
/metadata/primary_header	Dataset	-
CAL_MODE	Attribute	Calibration mode (ON | OFF)
DATE	Attribute	File creation date (UTC YYYY-MM-DD-hh:mm:ss)
DESCRIPTION	Attribute	General observation metadata
FILE_FORMAT	Attribute	File format
FILE_FORMAT_VERSION	Attribute	File format version
HDR_DEFN	Attribute	File format definition
HDR_DEFN_VERSION	Attribute	File format definition version
INSTRUMENT	Attribute	Backend instrument name
NAME	Attribute	primary_header
N_BEAMS	Attribute	Number of beams
OBSERVER	Attribute	Observer name
OBS_TYPE	Attribute	Observation type (TRACK|SCAN)
PID	Attribute	Project ID
RECEIVER	Attribute	Receiver name
SCHED_BLOCK_ID	Attribute	Schedule block ID

SDHDF_CLASS	Attribute	sdhdf_table
TELESCOPE	Attribute	Telescope name
UTC_START	Attribute	Observation start (UTC YYYY-MM-DD-hh:mm:ss)
/metadata/schedule	Dataset	-
DATE	Attribute	Date (UTC YYYY-MM-DD-hh:mm:ss)
DESCRIPTION	Attribute	Metadata specific to the scheduling of the observation
NAME	Attribute	schedule
SCHED_BLOCK	Attribute	Schedule block contents
SCHED_HOST	Attribute	Schedule host machine
SCHED_VERSION	Attribute	Schedule template version
SDHDF_CLASS	Attribute	sdhdf_table
/metadata/software_versions	Dataset	-
DESCRIPTION	Attribute	Metadata specific to software packages used for creating or processing the file
NAME	Attribute	software_versions
PROC	Attribute	Process name
SDHDF_CLASS	Attribute	sdhdf_table
SOFTWARE	Attribute	Software package
SOFTWARE_DESCR	Attribute	Software package description
SOFTWARE_VERSION	Attribute	Software package version

Appendix B: SDHDF Version History

SDHDF Versions

SDHDF v2.1 (Latest: July 31st 2021)

- Fix frequency dataset bandwidth bug (where nchan-1 should be nchan), and define as channel centre not channel start
- Update data weights definition
- Add data flags definition for RFI flagging
- Add parsing of normalisation flags from the telescope backend
- Remove redundant beam dimension from astronomy and calibrator datasets
- Minor small fixes and attribute updates

SDHDF v2.0

- Implement configuration file for use by any telescope
- Add weights group to definition
- Enable metadata copy if existing, to prevent unnecessary database query duplication
- Metadata for wind direction and wind speed parameters are now the right way around
- Enable splitting observation by frequency band
- Fix metadata query string not making it into history table
- Replace deprecated H5py create_scale with make_scale

SDHDF v1.9.3

- Enable file splitting mode with max file size threshold of ~10GB, effective from and including UTC 2020-09-24-12-27-43
- Add new 'schedule' dataset to hold DHAGU metadata
- Update DHAGU template query for PostGres (replacing ElasticSearch)
- Minor additions/fixes to modules and metadata

SDHDF v1.9.2

- Incorporate 32-bin calibrator dataset
- Fix missing PROC_ARGS in history dataset
- Add metadata query string to history dataset
- Fix missing reference to InfluxDB in software_versions dataset
- Minor additions/fixes to modules and metadata

SDHDF v1.9.1

- Fix OBS_TYPE
- Minor additions/fixes to modules and metadata

SDHDF v1.9

- Added OBS_TYPE to primary_header dataset
- Hierarchical structure updated to allow for multiple beams
- Add position offsets to obs_params datasets
- Minor additions/fixes to modules and metadata

SDHDF v1.8

- Incorporated an interpolation routine for InfluxDB output
- Added telescope config class
- InfluxDB observation metadata (including pointing) are accurate to within +/-1 second of time
- Minor additions/fixes to modules and metadata

SDHDF v1.7

- Re-processed observations from and including UTC_START: 2018-12-19-07:40:30
- InfluxDB observation metadata (including pointing) are accurate to within +/-5 seconds of time
- Metadata comprises the SDHDF definition by default
- Added 'history', 'rfi_excision', 'software_versions', 'weights' and 'cal_weights' datasets

SDHDF v1.6

- 'time' and 'cal_time' datasets changed to 'obs_params' and 'cal_obs_params', now incorporating time and observation metadata from InfluxDB
- 'obs_params' and 'cal_obs_params' dataset values now all refer to the timestamp at the start of the integration
- Enabled data from partial integrations to be incorporated
- Minor fixes to unit datasets
- Implemented header metadata comparison with input data file contents (header does not necessarily reflect contents)
- Added 'band_header' dataset for sub-/zoom-band metadata

- Added extra parameters to 'primary_header' dataset

SDHDF v1.5

- Updated time axis method to fix intervals, add integration index and time since start columns to time datasets, and set MJD to integration mid-point
- Minor fixes to array shapes
- Updated SDHDF definition with backend header descriptions
- Fixed missing dimension scales
- Fixed incorrect cal-on cal-off bin ordering
- No InfluxDB metadata
- Processed observations from UTC_START: 2019-04-09*

SDHDF v1.4

- Implemented fix to correctly order integrations in time
- Implemented separate datasets for cal-on and cal-off
- No InfluxDB metadata
- Processed observations from and including UTC_START: 2019-04-02-01:25:28

SDHDF v1.3

- First official complete run with real data
- Medusa calibration (preproc) data included
- Implemented independent frequency datasets to allow for sub-bands with different frequency resolutions
- Primary header intact but not finalised
- No InfluxDB metadata
- Processed observations from and including UTC_START: 2018-12-19-07:40:30

SDHDF v1.2

- Data from second commissioning run (P737)
- Incorrect and/or missing header information
- Medusa calibration (preproc) data included
- No InfluxDB metadata

SDHDF v1.1

- Data from first commissioning run (BL, P737)
- Incorrect and/or missing header information
- No calibration (preproc) data
- No InfluxDB metadata

SDHDF v1.0

- Data from first commissioning run (P737)
- Incorrect and/or missing header information
- Preliminary SDHDF structure
- Draft naming scheme implemented
- No Medusa calibration (preproc) data
- No InfluxDB metadata

As Australia's national science agency and innovation catalyst, CSIRO is solving the greatest challenges through innovative science and technology.

CSIRO. Unlocking a better future for everyone.

Contact us

1300 363 400
+61 3 9545 2176
csiroenquiries@csiro.au
www.csiro.au

For further information

CSIRO Space & Astronomy
Lawrence Toomey
+61 0 0000 0000
lawrence.toomey@csiro.au
csiro.au/cass