ATA .raw data format (version: Oct 2021)

- GUPPI files hold complex data products, in 4 or 8 bit formats
- Each file is divided up to 128 blocks, with each block containing an ascii **header** followed by **binary data**.
- An aggregate of multiple files comprise an “observation”
- Each file is ~16 GB in size, depending on the configuration (see below)

File Naming convention:

Each file has a naming convention of:

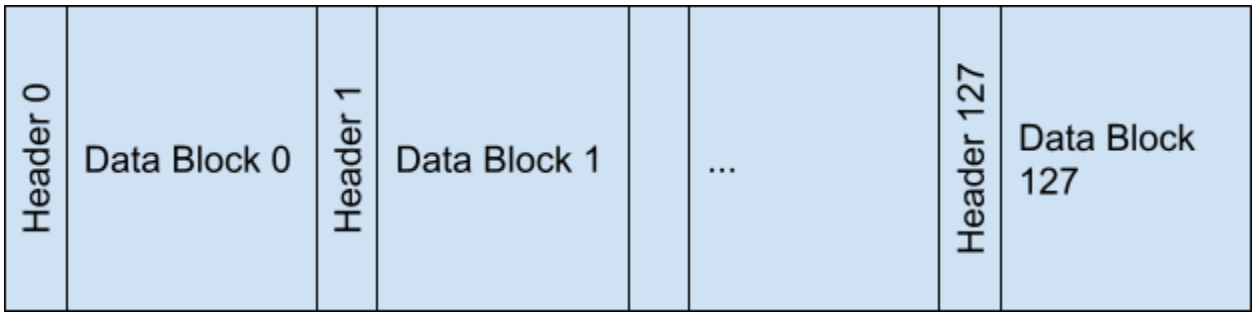
`guppi_MJD_SDAY_US_SOURCE_SCAN.NFILE.raw`

Where:

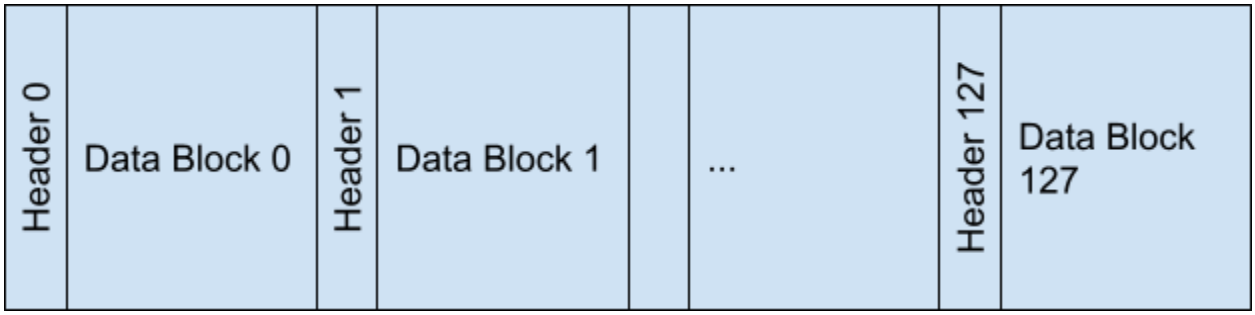
- `MJD`: Is the MJD start of the observation
- `SDAY`: second since the start of the MJD day
- `US`: microsecond after SDAY
- `SOURCE`: source or target
- `SCAN`: 4-digit. Not used for ATA, always 0001
- `NFILE`: 4-digit guppi file number in current observation

Example:

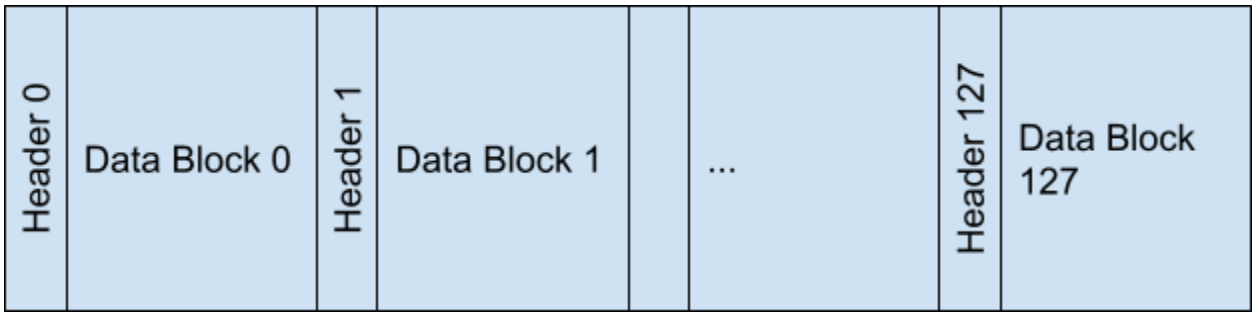
`guppi_59444_23895_918945_J0332+5434_0001.0000.raw`



guppi_59444_23895_918945_J0332+5434_0001.0001.raw

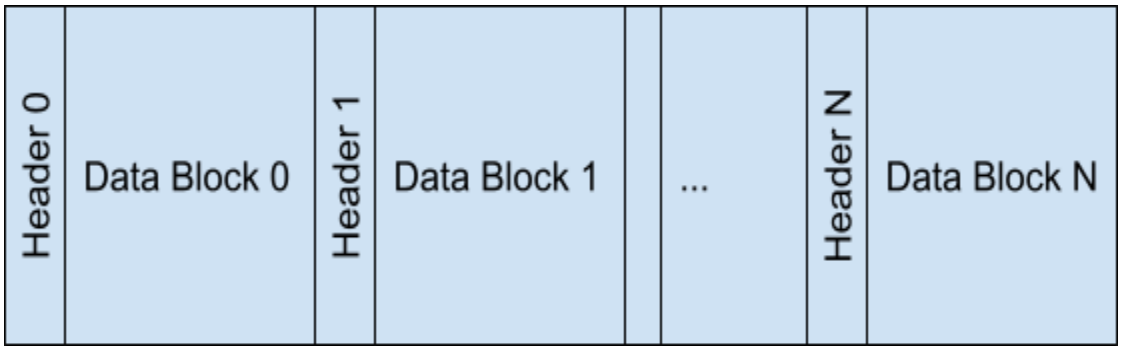


guppi_59444_23895_918945_J0332+5434_0001.0002.raw



.....

guppi_59444_23895_918945_J0332+5434_0001.XXXX.raw



Header block format:

ASCII 80-character fixed-width records, padded with spaces. Each record engulfs a key and value. No newline or NULL characters. Ends with a “END” followed by 77 spaces.
Keys are at most 8-letters long.

Example:

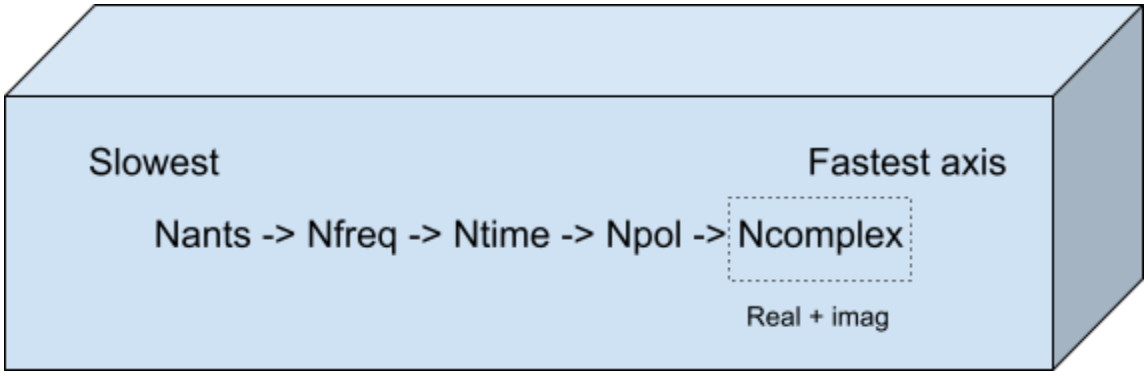
ANTNAMES	1cB,1eB,1gB,1hB,1kB,2aB,2bB,2cB,2eB,2hB,2jB,2kB,2lB,2mB,3cB,3dB,3lB	ANTNMS00	1cB,1eB,1gB,1hB,1kB,2aB,2bB,2cB,2eB,2hB,2jB,2kB,2lB,2mB,3cB,3dB,3lB	ANTNMS01	4jB,5bB,4gB	AZ	197.24592590332031
BACKEND	GUPPI	BINDHOST	enp97s0f1	BINDPORT	10000	BLKIDX	21
BLKSTART	8284973568	BLKSTOP	8284981760	BLOCSIZE	125829120	CHAN_BW	0.5
DAQPULSE	Thu Oct 21 17:19:33 2021	DAQSTATE	idling	DATADIR	/mnt/buf0	DEC	-21.620577496218008
DEC_STR	-21.620577496218008	DIRECTIO	1	EL	25.603609085083008	FENCHAN	2048
IBVBUFST	1/24	IBVGBPS	62.019311	IBVPKTSZ	42,16,8208	IBVPPS	1249986.119
IBVSTAT	running	INSTANCE	0	MAXFLOWS	16	NANTS	20
NBITS	8	NCHAN	192	NDROP	80492	NETBLKMS	5.793670177
NETBLKPS	80.207817078	NETBUFST	1/24	NETSTAT	receiving	NETTHRDS	8
NPKTS	3600224096	NPOL	2	NSTRM	2	OBSBLKMS	0.023590000
OBSBLKPS	62	OBSBW	96	OBSFREQ	1478.75	OBSINFO	VALID
OBSNCHAN	3840	OBSNDROP	0	OBSNPKTS	37519360	OBSSTAT	processing
OBSSTEM	guppi_59444_23895_918945_J0332+5434_0001	OBS_MODE	RAW	OVERLAP	0	PHYSGBPS	7.699763298
PHYSPKPS	1249961.625000000	PIPERBLK	8192	PKTFMT	ATASNAPV	PKTIDX	8284973568
PKTNCHAN	96	PKTNTIME	16	PKTSIZE	6160	PKTSTART	7528000000
PKTSTOP	7543000000	POSTPROC	none	PPCNDARG	-r 1 -s 10 -o auto -n bla	PPCNDINP	turboseti ^turboseti
PPCPARG	/mnt/datax-netStorage-40G/dmpauto-$hnme$.$inst$/$stem$/	PPCPINP	rawspec	PPLOGARG	-H $hnme$ -i $inst$ -s $stem$ -b beg -e end -t $time$ -p $proc$	PPLOGINP	&
PPPULSE	Thu Oct 21 17:19:33 2021	PPRMINP	hpguppi &*.raw,rawspec &*.fil	PPRWSARG	-f 262144 -t 2 -I 1.0 -d /mnt/buf$inst$/rawspec/$stem$/	PPRWSENV	CUDA_VISIBLE_DEVICES:$inst$ --numa -3 $inst$
PPRWSINP	hpguppi	PPSTATUS	WAITING.	PPTBSARG	-M 10 -g y -p 12 -n 1440 -o /mnt/buf$inst$/turboseti/$stem$/	PPTBSENV	CUDA_VISIBLE_DEVICES:$inst$
PPTBSINP	rawspec	PROJID	ENDURANCETEST	RA	9.3227265378553295	RA_STR	9.3227265378553295
SCHAN	736	SOURCE	J0332+5434	SRC_NAME	J0332+5434	STTVALID	0
STT_IMJD	59444	STT_OFFS	0	STT_SMJD	23895	SYNCTIME	1629253639
TBIN	1.9999999999999999e-06	XPCTGBPS	7.500GBps 60.000Gbps				

Important header parameters:

NBITS	Number of bits per component of complex sample. 4 or 8 for the ATA
NCHAN	Number of channels from each FEngine: equates to OBSNCHAN/NANTS. Isn’t officially implemented, but will replace the NSTRM parameter.
SCHAN	The index of the first channel received from the antennae.
PIPERBLK	Name suggests something else, but it is used as number of spectra per block
NANTS	Number of antenna-streams (F-Engine outputs).
NPOL	Number of polarisations
OBSNCHAN	Total number of frequency channels, aggregated over all the ants

	= NCHAN*NANTS
OBSBW	Observation bandwidth (MHz)
OBSNDROP	Number of dropped packets so far, i.e. since the start of the observation
NDROP	Number of dropped packets since the start of the hashpipe acquisition program (aggregated over multiple previous observations)
ANTNAMES	(depreciated because we can't fit all ant names in single value, see below ANTNMS)
ANTNMS00	Comma separated antenna names. ATA eg: "1eC,1eB,1gB,..." (The "B" refers to the tuning "B". ATA has 4 tunings, "A", "B", "C", "D")
ANTNMS*xx*	Continuation of ANTNMS*xx-1*.
BLOCSIZE	Number of bytes in datablock
TBIN	Time span of every sample in the datablock in seconds.
OBSFREQ	Centre frequency
NSTRM	The number of packets per time-index, indicating how many channels are received from each antenna as PKTNCHAN*NSTRM.
CHAN_BW	The frequency width of each channel, in MHz.
RA_STR	
SYNCTIME	UNIX second at which the synchronisation of the RFSoCs took place.
AZ	Azimuthal angle (of the first antenna listed in ANTNMS)
EL	Elevation angle (of the first antenna listed in ANTNMS)

Data block format:



- Depending on the observation configuration, 2 data formats can be obtained:
  - “Widefield” format (120 MB block size).
    Nants = 5
    Nfreq = 1280
    Ntime = 8192 (for 4bit, would become 4096 for 8bit)
    Npol = 2
  - “Beamforming” format, all antennas aggregated on a single nodes (each of the 2 instances recording in 120 MB blocks):
    Nants = 20
    Nfreq = 192
    Ntime = 8192
    Npol = 2

- Ncomplex can either be 4 or 8 bits for each component. 4bit mode will be depreciated once the 8bit mode is fully tested.

NOTE: The same file (data) format will exist on RAM, i.e. header+data blocks