VLBA Sensitivity Upgrade Memo 42 rev. 1

Description of the VDIF format used by the VLBA, VLA and GBT

including the

VDIF Extended Data Version 3 definition

Walter Brisken, Matt Luce, & Paula Metzner

16 Apr 2014 *Updated: corrected Word 1 of diagram*
15 Nov 2013 *Original*

1 Introduction

This memo describes the flavor of VDIF produced by the VLBA, VLA and GBT. The data from all instruments conform to the VDIF standard (see specification at http://www.vlbi.org/vdif/). VDIF is a very flexible format allowing many different ways to represent the same data. The sections below describe the conventions adopted by the VLBA, VLA and GBT, a new VDIF header Extended Data Version that is used by the VLBA and GBT, and some errata from early deployments. The VLBA and GBT use the same devices with the same firmware to produce VDIF data and hence have identical data formats. The VLA VDIF data comes from the WIDAR correlator and is independently implemented.

2 Adopted conventions

The following choices have been made that limit the variety of VDIF data that are produced by the VLBA, VLA and GBT:

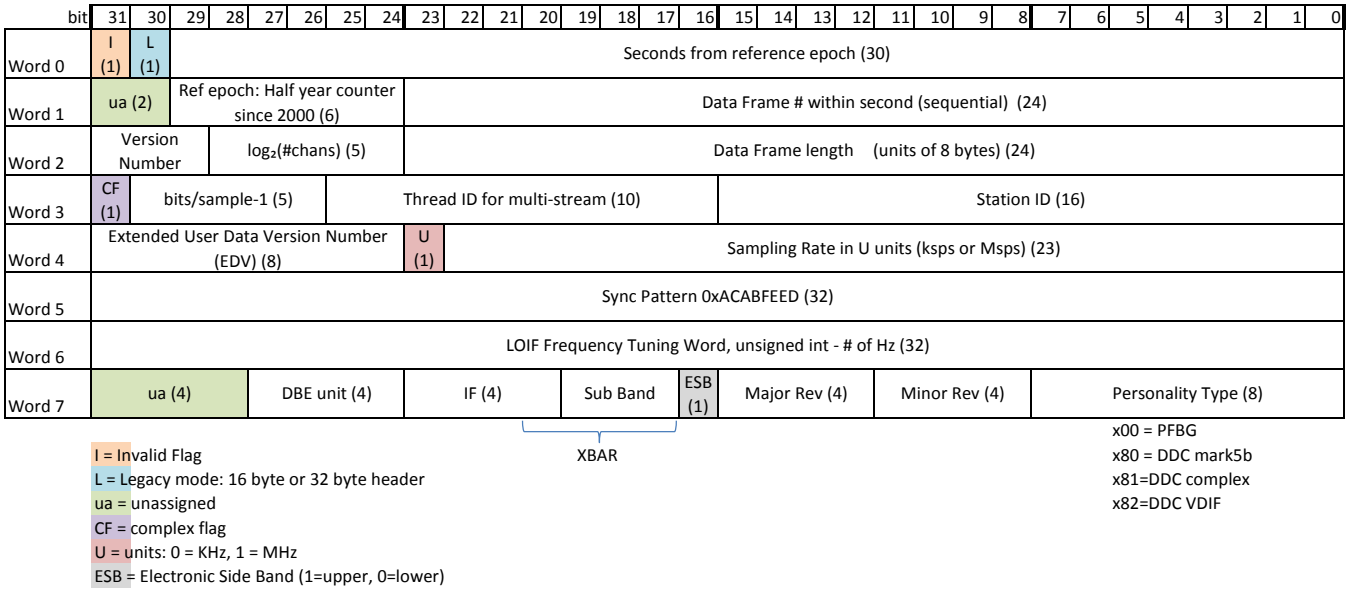
1. Each VDIF packet contains 5000 bytes of payload.
2. Full length (32-byte) VDIF headers are always used.
3. All VDIF packets thus have a total length of 5032 bytes.
4. As of now, all data is in the form of 2-bit real samples. It is possible other options will be made available in the future.
5. Each VDIF thread produced contains a single base-band channel.
6. As of now, between 1 and 8 threads are produced. It is possible that more channels could be made available in the future.
7. As of now, all threads produced are of identical data rates. It is possible that this restriction will be lifted sometime in the future.
8. When ordered alphabetically by channel name, as specified in the `.vex` file, the thread Ids start at zero for the first channel and increment by one for each subsequent channel. The vex2 format will have a proper method to designate VDIF frames so this approach will not be required once vex2 is adopted by the VLBA and VLA.

3 Extended Data Version 3

The VLBA has created its own VDIF Extended Data Version (EDV) to add some additional information to the VDIF headers. This has been assigned EDV 3 by the VDIF committee. It only makes sense to use EDV 3 on data where a single channel is recorded per thread.

The intention of this EDV is to allow specification of sample rate, tuning, and origin of the data (which data source number and which physical analog input was used). It preserves the location and value of a sync word introduced by the NICT EDV 1.

The details of this EDV are shown in the diagram below:



Explanation of some of the new fields:

- **Sample rate** The sample rate of this stream is determined by two subfields. Word 4, bit 23 is set to 0 for units of kHz and 1 for units of MHz. Word 4, bits 0 to 22 contain an unsigned integer specifying the sample rate in the units set by bit 23.
- **Sync Pattern** The entirety of word 5 contains a fixed bit pattern `0xACABFEED`.
- **Tuning** Word 6 contains a 32-bit unsigned integer specifying in Hz the digital back end tuning frequency (not the sky frequency). For the RDBE as implemented by NRAO (and in use by VLBA and GBT), the legal values for this would range between 512000000 and 1024000000.
- **DBE unit** Word 7 bits 24-27, interpreted as a 4-bit unsigned integer, represents the DBE number at the site producing the data. For the VLBA and GBT this number will be 0 or 1.
- **IF** Word 7 bits 20-23, interpreted as a 4-bit unsigned integer, represents the intermediate frequency input number to the digital back end. For the RDBE at the VLBA and GBT there are only 2 inputs which could take values 0 and 1.

- **Sub Band** Word 7 bits 17-19, interpreted as a 3-bit unsigned integer, represents the DBE internal subband that was used to generate the channel. In the DDC personality of the RDBE, as currently implemented, each 512 MHz IF input is divided into three sub-bands. The central sub-band is complex and can be logically addressed from either band edge. In total there are 4 logical sub-bands, labeled 0 through 3, that this value can assume.

- **ESB** Word 7 bit 16 contains the electronic sideband of the baseband channel. This is not the net sideband but the sideband relative to the input IF. A value of 1 indicates upper sideband and 0 indicates lower sideband.

- **Version** Three fields spanning word 7 bits 0 through 15 are used to indicate the firmware type and version. Additional personality types can be added as needed.

At this time only the VLBA and GBT have adopted this EDV and as of this writing no production VDIF data has been generated by GBT.

4 Errata and special notes

Some early VDIF data from the VLBA or VLA did not fully conform to the above descriptions. Below are some known examples of that. At the time of this writing the data being produced by the VLBA is thought to be fully conformant with the above description.

1. Early VDIF data from the VLBA, GBT and VLA had its validity bit set incorrectly (all data was marked as invalid but in fact was valid).

2. Early VDIF data from the VLA did not conform to the thread numbering convention (#7 above).

3. There was a short period where VLBA and GBT data indicated EDV 2, not EDV 3. During a transition period, some data contained a mixture with some threads specifying EDV 2 and some specifying EDV 3.

4. All three instruments record data on a Mark5C unit which can experience a small amount (typically about 1 part in 10^5) of packet loss. Any software or hardware device consuming such data should handle this.

5. All VLBA VDIF data produced before Nov 20, 2013 suffered from an error in quantization. The *-low* and *+low* states had their bit sequences swapped (i.e., used the Mark5B representation). Only the VLBA and GBT are affected by this problem.