ISTITUTO NAZIONALE ASTROFISICA

ISTITUTO DI RADIOASTRONOMIA

Rapporto Interno n. 457/12

LOFAR data formatting

F. Schilliro - IRA Noto

G.Naldi - IRA Medicina

A. Mattana - IRA Medicina

M. Bartolini - IRA Medicina

Data formatting for Best3-LO experiment has to be done according LOFAR Format encoding procedure [1] between a RSP (Remote Station Processing) and a CEP (Central Processor) of a LOFAR system. This format is mainly based on 'beamlet' data, which contains the real and imaginary parts of both polarizations for a number of sub-bands of the total bandwidth of the receiving system.

In the case of the experiment Best3-LO were used only a fraction of the resources made available by the LOFAR system, given the relative simplicity of the architecture used. The purpose of this document is to describe the operations of data packaging out of the Medicina system , which must then be processed by the correlator, which requires certain specifications in terms of bandwidth of the signal, data format of spectrum and time interval over which carry the same spectrum.

1.Input Data

The main problem is to synchronize Medicina data with Exloo core station 1 receiving system, for correlation process. This last one has a 160 MHz of sampling frequency, a passband of 0-80 MHz with 512 frequency bins (channels) out to a polyphase filter bank. Medicina back-end works with a sampling frequency of 10 MHz (5 Mhz of a real bandwith) but a 2.5 MHz two-sided complex band out to a downconverter system. In order to ensure the same frequency granularity for correlating the two spectra, Medicina bandwith is divided in 16 frequency bins (channels), so that is possible to have a common spectral resolution of 0.156250 MHz, the same used at Exloo system.
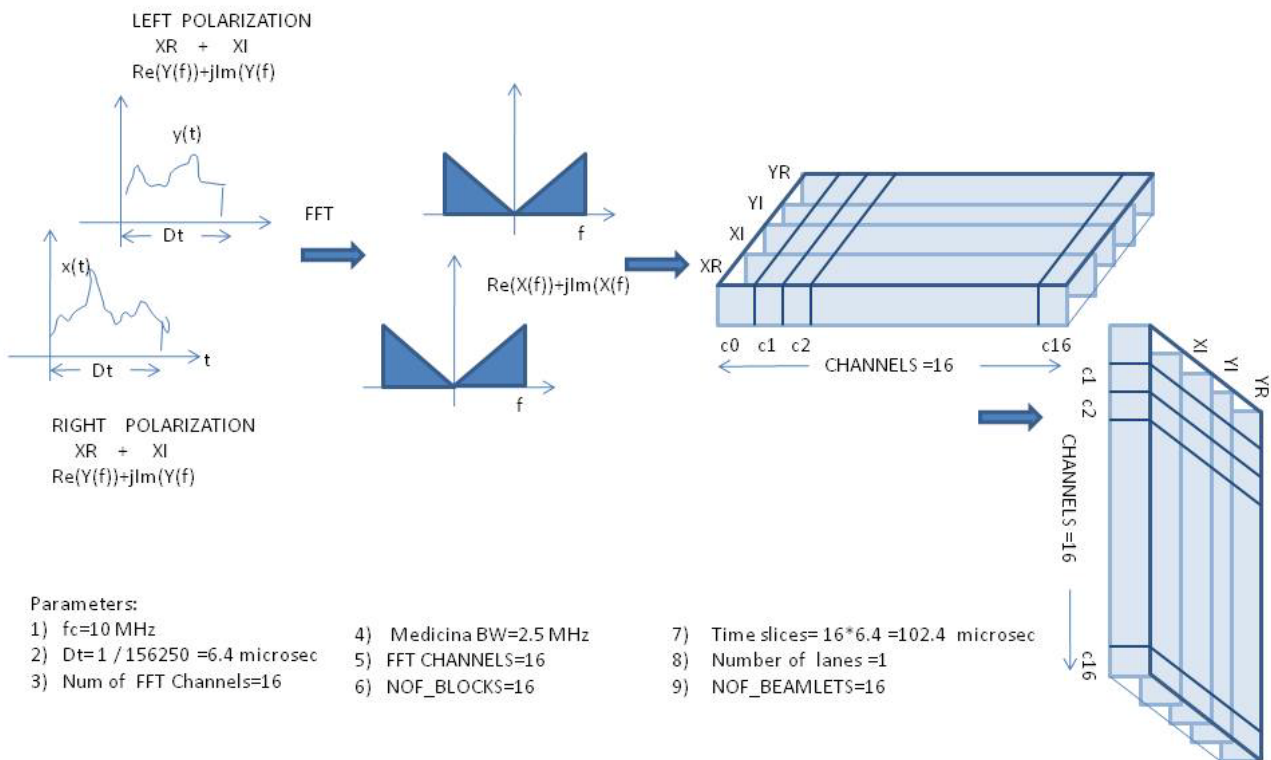

Dato in ingresso, spettro, beamlet

Parameters:
1) fc=10 MHz
2) Dt=1 / 156250 =6.4 microsec
3) Num of FFT Channels=16
4) Medicina BW=2.5 MHz
5) FFT CHANNELS=16
6) NOF_BLOCKS=16
7) Time slices= 16*6.4 =102.4 microsec
8) Number of lanes =1
9) NOF_BEAMLETS=16

Fig. 1

So it is possible to have in both cases a data time slice of 1/156250=6.5 microsec, that are 64 samples of Medicina data flow on which spectrum is calculated and composed of 16 frequency bins, for both Left and Right polarization, (even if only X is used, Y is set to zero). Beamlet out to Medicina system is composed of 4 real values (real and imaginary part of both polarizations) for 16 subbands, as described in fig.1.

2. Beamlet

Beamlet is defined as a data that includes values for a specific sub-band spectrum of the input signal, containing both real and imaginary values for each polarization, thus containing a total of four real fields. Figure 2 shows a field of 16 beamlets corresponding to channels used in Medicina and Exloo back-ends (Table 1), selected from a maximum of 248 possible.
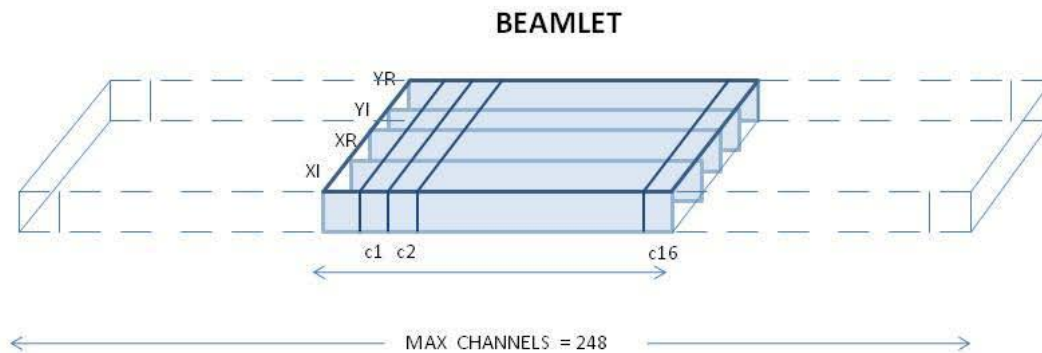

Fig. 2

The name derives from the fact that each analog signal in time domain is a weighted sum of all the signals coming from selected antennas for beamforming. Each field contains data from 16 beamlets signal for 6.4 microseconds, so it is necessary to package the various fields of beamlets in order to have a data packet containing larger chunks of data and then transmit multiple signals over time. This packaging takes into account the 'time slices', ie a number of beamlet that are joined together. This number in our case (but in most cases LOFAR) is 16 and is configured in the data in the value NOF_BLOCKS. In this way, 16 blocks of beamlets are packed in a data packet, for a total of 16 * 4.6 = 102.4 microsec data, already converted in frequency domain. Below it is describes how the routing of the beamlets and slices of time.

Medicina	Exloo	Frequency Band
0	376	138.750000-138.828125 & 141.171875-141.250000 **unusable**
1	377	138.828125 - 138.984375

2	378	138.984375 - 139.140625
3	379	139.140625-139.296875
4	380	139.296875 – 139.453125
5	381	139.453125 – 139.609375
6	382	139.609375 – 139.765625
7	383	139.765625 – 139.921875
8	384	139.921875 – 140.078125
9	385	140.078125 – 140.234375
10	386	140.234375 – 140.390625
11	387	140.390625 – 140.546875
12	388	140.546875 – 140.703125
13	389	140.703125 – 140.859375
14	390	140.859375 – 141.015625
15	391	141,015,625 - 141,171,875

Tab.1

3. Data packaging and routing

Payload of data packet is formed by means of certain functions applied directly to the beamlet to be transmitted:

1) Buffering of beamlets in a NOF_BLOCKS numbers (typically 16 time blocks, as in our case);

2) Routing: transmission typically takes place on specified lanes in number from one to a maximum of four. Each lane is managed by a single RSP (Remote Station Processor) so that, for example, when a package is composed of 248 beamlets, it is routed to four lanes from 62 (NOF_BEAMLETS parameter) beamlets each. The routing is done by selecting the beamlets according to their position set by B matrices , each related to a different lane and each formed according to the law:

$$B_{lane}(row,col)=(lane+row*nof\_lane)_{th} \qquad (1)$$

where *lane* is the i-th lane run by a particular RSP, *row* is the row index that varies between beamlets (0: NOF_BEAMLETS-1), *col* is the column index that varies between the time slices (0: NOF_BLOCKS-1). As an example, consider the case where you have to route 16 time blocks of 248 beamlets in lane 0; inside the beamlets according to formula (1) you must select the beamlets indexes:

```
  0     4     8    12    16    20    24    28    32    36    40    44    48
 52    56    60    64    68    72    76    80    84    88    92    96   100
104   108   112   116   120   124   128   132   136   140   144   148   152
156   160   164   168   172   176   180   184   188   192   196   200   204
208   212   216   220   224   228   232   236   240   244
```

for each time slice. Then you have to line up and link them to a new array based on time slices and read the contents of this matrix by rows, and then packed in the payload encoding each of the four values of the beamlets formed in 'little endian'. This is what is called **'interleaved mode'**, as in the case of LOFAR CEP (Central Processor). Figure 3 depicts the method used in the case of Medicina, using a single lane, a number of 16 beamlets and 16 time slices. In this case the routing is very simple, Blane array contains only values from 1 to 15 on lane 0.
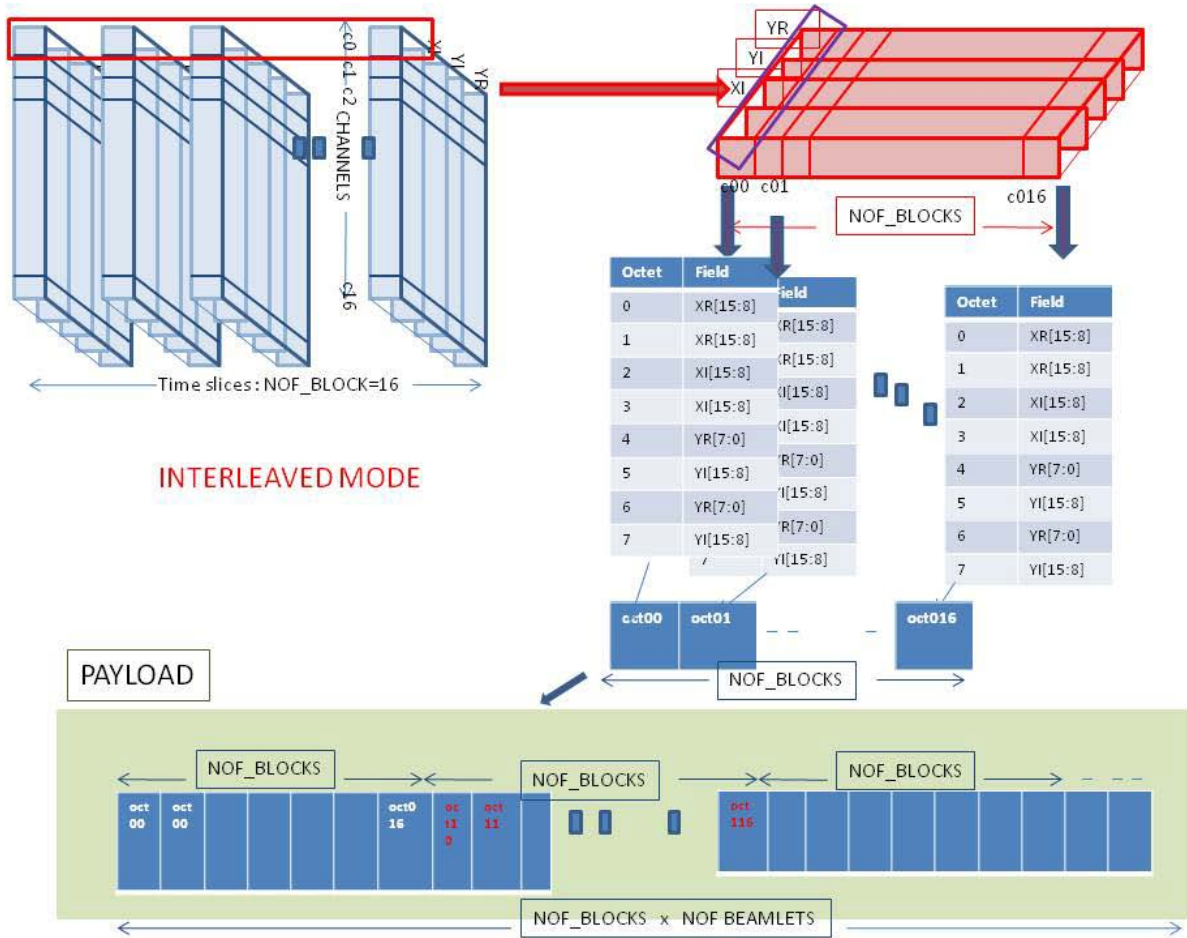

Fig. 3

The figure also describes the composition of octets in the payload to be packaged,simply by which we can derive the amount of bytes, which is 16 time blocks multiplied for 16 beamlets , multiplied for 8 bytes of data for a total of 2048 bytes.

3) Application header formation: it consists of data fields from which is possible to recognize the transmitting station, and other parameters useful for decoding the payload. In the case of the experiment in question was only required to include the time reference to the packet header, or TIMESTAMP fields and BLOCK_SEQUENCE_NUMBER, calculated based on the following criteria:

a) TIMESTAMP: The number of seconds since January 1, 1970 (0:00:00 hours), encoded in little endian format in four bytes.

b) BLOCK_SEQUENCE_NUMBER: is the fraction of a second divided by 156250, which indicates the instant of sampling and packaging and it is related to the granularity of sampling, which we recall is of 6.4 microseconds (1/ 156250) . This fraction ranging from 0 to 156249 is encoded in 4 bytes in little endian format.

3. Tests

An experiment to test the LOFAR data formatting is performed by sending a data packed with a Matlab synthesized signal at a time, in order to verify the right behavior of routines.

Signal frequency: **1.234 MHz tone + random noise.**

- Channel of tone : **7** (Medicina) to read in Tab.1 corresponding to band 139.765625 – 139.921875 .

- Signal duration: 0.064 seconds.

- Time: April 15, at 19:19:00 (1.302895140000000e+009seconds elapsed from 1/1/1970)

- Initial block sequence number: 0

- Data format for a block of 16 beamlet (payload):

- `packet=[ts1 ts2 ts3 ts4  bn1 bn2 bn3 bn4  octet_i];`

    <-   Time   -> <- block sequence number-> <- **payload**->

Data file is a text file composed by 625 row of packet that contains 16 beamlet for 16 time slice for a total of 0.064 seconds of signal.

4. Conclusions

Have been written Matlab routines to format beamforming data from time domain to beamlet and so on, in LOFAR data packets to send to correlator . Routines are usable for all the cases, even if in this case the use of beamlets, lanes and B matrix is limited to minimum (16 beamlets).

References

[1] W. Lubberhuizen, E. Kooistra – **RSP-CEP Beamlet Data Interface –** Astron, 2004