TOP Feature Extraction

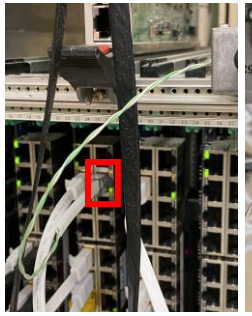
Progress Updates & Plans

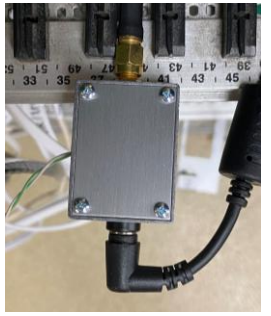
Harsh Purwar

TOP Feature Extraction Meeting

April 21st, 2025

Project status





- **The development of 1st version of TOP Feature Extraction is now complete.**
- We began testing it at UH last week (manually, looking at the data words, event sizes, etc.)
- UH test bench has also been slightly updated to ensure hits in every event, makes manual testing easier (+higher occupancy) – thanks a lot Matt!

- Next steps:
  - Saving data to file and test with basf2 unpacker module [**…ongoing**]

- Excluded from version 1 (To do for version 2):
  - Loading of pedestal values from disk and subtracting them from hits
  - Loading thresholds from disk and using them in feature extraction

TOP Feature Extraction

(Source code tagged with **v1** on Gitlab)

Version 1 (v1) is now on Gitlab.

- Repository: pcie40_software
- Tag: v1, branch: top-feature-extraction-by-daq-temp (`git checkout tags/v1 -b ……`)

```
num_words_in: 1099, num_words_out: 135,  |  TOP Feature Extracted Data:
# of words : 135
 data   0 : 64a00204 00000000 a4269900 03010000 80b92d81 0000f1d8 05000000 0000577f
 data 007 : 372d0f0f 80b92f81 0000f1d8 05000000 0000577f 1a220f0f 80b92d83 0000f1d8
 data 00f : 05000000 0000577f 04290f0f 80b92f83 0000f1d8 05000000 0000577f 1e1c0f0f
 data 017 : 80b92d91 0000f1d8 05000000 0000577f 6c790f0f 80b92f91 0000f1d8 05000000
 data 01f : 0000577f 6c690f0f 80b92d93 0000f1d8 05000000 0000577f 8c750f0f 80b92f93
 data 027 : 0000f1d8 05000000 0000577f 98690f0f 80b92da1 0000f1d8 05000000 0000577f
 data 02f : 90590f0f 80b92da3 0000f1d8 05000000 0000577f 9e4b0f0f 80b92db1 0000f1d8
 data 037 : 05000000 0000577f ac470f0f 80b92fb1 0000f1d8 05000000 0000577f aa390f0f
 data 03f : 80b92db3 0000f1d8 05000000 0000577f be370f0f 80b92fb3 0000f1d8 05000000
 data 047 : 0000577f ca3b0f0f 80b92dc1 0000f1d8 05000000 0000577f ce270f0f 80b92dc3
 data 04f : 0000f1d8 05000000 0000577f dc290f0f 80b92dd1 0000f1d8 05000000 0000577f
 data 057 : 73ad0f0f 80b92fd1 0000f1d8 05000000 0000577f 71ad0f0f 80b92dd3 0000f1d8
 data 05f : 05000000 0000577f 819b0f0f 80b92fd3 0000f1d8 05000000 0000577f 7d9d0f0f
 data 067 : 80b92de1 0000f1d8 05000000 0000577f 838b0f0f 80b92de3 0000f1d8 05000000
 data 06f : 0000577f 818b0f0f 80b92df1 0000f1d8 05000000 0000577f 916d0f0f 80b92ff1
 data 077 : 0000f1d8 05000000 0000577f 916b0f0f 80b92df3 0000f1d8 05000000 0000577f
 data 07f : a15b0f0f 80b92ff3 0000f1d8 05000000 0000577f 9f5b0f0f 3a0a0000
Words after TOP FE: 144
```

Thorough tests with TOP Unpacker

Planned for this week

- Idea:
  - Write TOP feature extracted data to a root file with basf2
  - Run basf2 unpacker to unpack the data file and log all/any errors
  - Then may be plot waveforms (might look weird without pedestal subtraction though)
  - This will also be done at KEK with a local run when the system is up and running (soon)

- Status:
  - Compiling basf2 software on PCIe40 host server at UH after suggested modifications from the DAQ group.

- KEK TOP setup:
  - Martin got the interlock issue fixed.
  - During weekend, he mentioned he had some DAQ issues while taking a local run with one BS.
    I'll try to find some time to investigate it with Martin this week.

TOP FE data words > raw input data words

- This could happen when we append raw waveforms to the data.
- Yamada-san has been informed, and he is looking into it.

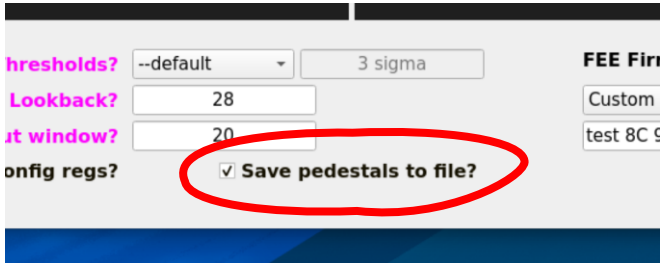
- **Anyway, not a problem for Physics/Cosmic runs.**

```
// Copy TOP feature-extracted-buffer to chunks
for(unsigned int id = 0; id < chunks.size(); id++){
  unsigned int sizeChunk = chunks[id].get()->getSize() / sizeof(unsigned int);
  int offset_words = 0;
  if( id == 0 ){
    if( sizeChunk <= this->B2L_FEE_HSLB_header_words + this->B2L_FEE_HSLB_trailer_words ){
  char errmsg[500];
  char hostnamebuf[50];
  gethostname(hostnamebuf, sizeof(hostnamebuf));
  sprintf( errmsg, "[FATAL] %s ch=%d : Event size is too small. %d words. Exiting..: %s %s %d\n",
       hostnamebuf, linkId, sizeChunk, __FILE__, __PRETTY_FUNCTION__, __LINE__ );
  printfMessage(errmsg);
  exit(1);
    }
    offset_words = this->B2L_FEE_HSLB_header_words;
  }else{
    offset_words = 0;
  }
  if( new_detbuf_words - fe_pos > sizeChunk - offset_words ){
    memcpy( chunks[id].get()->getRawEvtData() + offset_words, fe_buf + fe_pos,
        (sizeChunk - offset_words)*sizeof(unsigned int));
    fe_pos += ( sizeChunk - offset_words );
  }else{
    memcpy( chunks[id].get()->getRawEvtData() + offset_words, fe_buf + fe_pos,
        (new_detbuf_words - fe_pos)*sizeof(unsigned int));

    // Copy B2L_HSLB_FEE trailer
    if( ( sizeChunk - offset_words ) - ( new_detbuf_words - fe_pos ) >= this->B2L_FEE_HSLB_trailer_words ){
  memcpy( chunks[id].get()->getRawEvtData() + offset_words + ( new_detbuf_words - fe_pos ),
      orig_data.get() + eve_words - this->B2L_FEE_HSLB_trailer_words,
      this->B2L_FEE_HSLB_trailer_words*sizeof(unsigned int) );
    }else{
  int pos_trailer = ( sizeChunk - offset_words ) - ( new_detbuf_words - fe_pos );
  int tmp_err_flag = 0;
  if( id + 1 < chunks.size() ){
    if( chunks[id + 1].get()->getSize() / sizeof(unsigned int) >= this->B2L_FEE_HSLB_trailer_words - pos_trailer ){
    memcpy( chunks[id].get()->getRawEvtData() + offset_words + ( new_detbuf_words - fe_pos ),
        orig_data.get() + eve_words - this->B2L_FEE_HSLB_trailer_words,
        pos_trailer*sizeof(unsigned int) );

    offset_words = 0;
    memcpy( chunks[id + 1].get()->getRawEvtData() + offset_words,
        orig_data.get() + eve_words - this->B2L_FEE_HSLB_trailer_words + pos_trailer,
        ( this->B2L_FEE_HSLB_trailer_words - pos_trailer )*sizeof(unsigned int) );
    }else{
      tmp_err_flag = 2;
    }
  }else{
    tmp_err_flag = 1;
  }
```

Raw pedestal data files



- These raw pedestal data files can be created by reading out pedestal values for each channel (uses PS) and dumping the data to text/binary files.
- Usually, it can be done right after a BS is power-cycled (not during a run though).
- An option to do this easily was included in the TOP Power-cycle and Config GUI.
- **However, this seems to be broken now, likely due to changes made in the PCIe40 readout software over the year or so.**

- I need to investigate this and see how it can be fixed.
- For development and local testing, for now we can use the previously saved files with pedestal values.

Thank you

-- Harsh Purwar

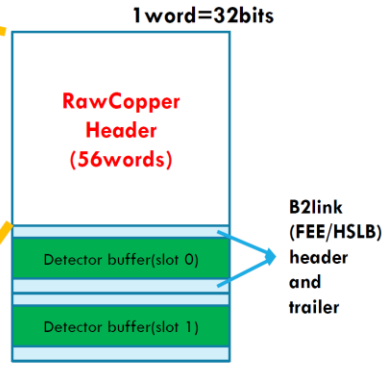
Backup slides

B2L, HSLB, COPPER/PCIe40 data formats

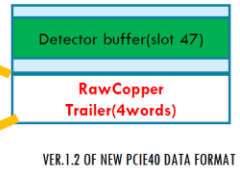
New data format for RawCOPPER hdr. and trl. (ver. alpha-200809)

- Remove COPPER header/trailer (which is currently removed on a readout PC.)
- Position in header will be removed to reduce header size.
- Error information in trailer (which link, what kind of error)
- XOR checksum (less CPU power to re-calculate on host servers or HLT)

1 Number of total words
2 0x7f7f | Format ver.(8bit) | Number of words in this block (8bit)
3 exp no. (10bit=1024), run no.(22bit=4194304 including subrun)
4 event number(32bit)
5 From B2link FEE header 2 (TT-ctime | Trig-type)
6 From B2link FEE header 2 (TT-utime)
7 Node ID
8 b2link CRC error bit (4) | |truncation mask (truncated or not) / type of data (compressed,calibration, …)
9 Position of ch0 data
10 Posiont of ch1 data
…
56 Position of ch48 data

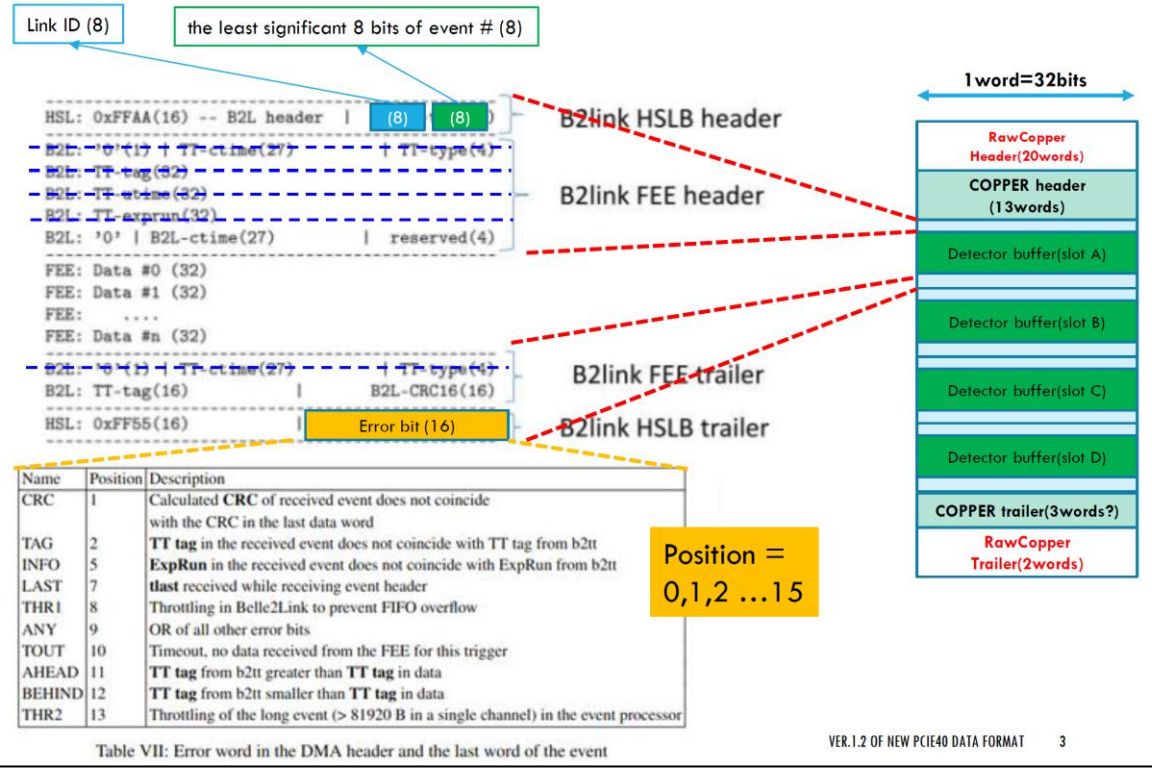


1 Error bit
2 Error slots
3 XOR checksum
4 termination word of this block = 0x7fff0006



New data format for B2L/HSLB hdr. and trl. (ver. alpha-200809)

- Basically unchanged
- Added some info in ffaa header and ff55 trailer



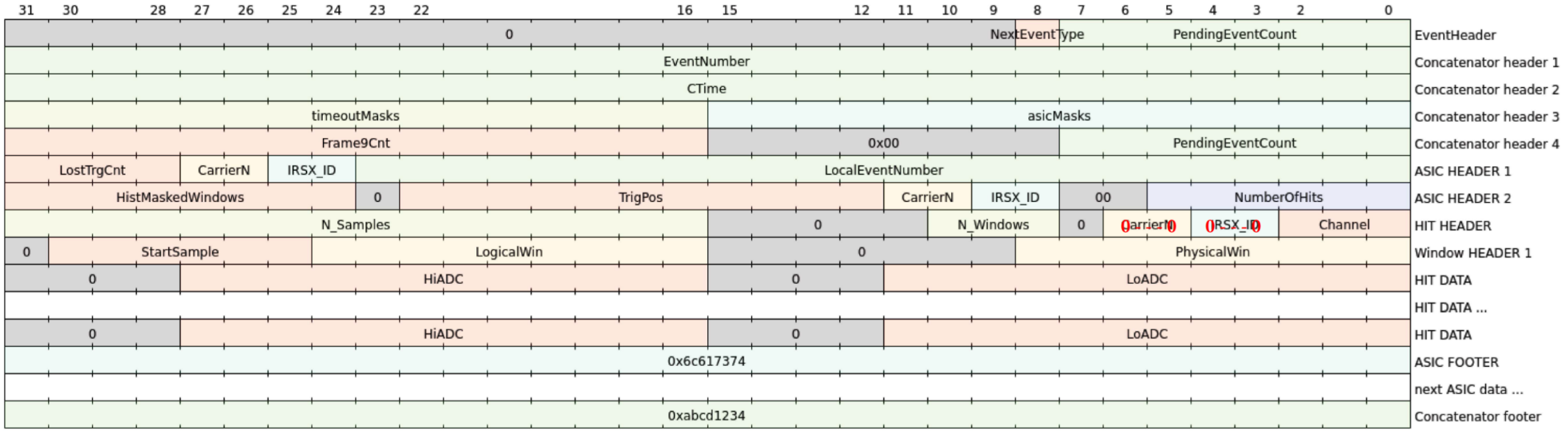
Name	Position	Description
CRC	1	Calculated **CRC** of received event does not coincide with the CRC in the last data word
TAG	2	**TT tag** in the received event does not coincide with TT tag from b2tt
INFO	5	**ExpRun** in the received event does not coincide with ExpRun from b2tt
LAST	7	**tlast** received while receiving event header
THR1	8	Throttling in Belle2Link to prevent FIFO overflow
ANY	9	OR of all other error bits
TOUT	10	Timeout, no data received from the FEE for this trigger
AHEAD	11	**TT tag** from b2tt greater than **TT tag** in data
BEHIND	12	**TT tag** from b2tt smaller than **TT tag** in data
THR2	13	Throttling of the long event (> 81920 B in a single channel) in the event processor

Position = 0,1,2 …15

Table VII: Error word in the DMA header and the last word of the event

Raw data format

Note: Hit header bits 3, 4, 5 and 6 are always 0.



TOP Production Data Format

Production Debugging 4.1

2.2 **Belle 2 TOP Data Format (Production Data)**

Note that the data listed below does NOT include protocol headers; trigger type, ctime, utime, and trgtag are included in Belle2Link headers.

Word	Bits 31 30 29 28 27 26 25 24	23 22 21 20 19 18 17 16	15 14 13 12	11 10 09 08 07 06 05 04	03 02 01 00	Hit
0	Type (=0x04)	Version (=0x01)	0xA	SCROD _ID		N/A
1	EXTRA	NumWordsBonus	Phase(0-8)	NumWordsCore		
2	SKI RSVD(0000)	ctime (11 LSBs)	Revo9 Counter			
3	ASIC Masks (TimeoutMasks | Register Masks)		eventQueueDepth	eventNumberByte		
4	Carr IRSX Channel	Window	0xB	tFine	WF H/S **Heap Window**	1
5		vPeak	Integral			1
6		vRise0		vRise1		1
7		vFall0		vFall1		1
8	SampleRise	dSampPeak dSampFall	**HeaderChecksum**			1
	. . .					
N*(5+EXTRA)+4	SD_type	Slow data	1 0 1	Nhits		N
	RAW HITS APPENDED HERE.					

Waveform Flag Heap/Stack Flag

"1 0 1 x" = 0xC or 0xD

= status bits

= reserved (0 for now)

= unsigned

= signed

Sum of all 16-bit values in "hit header" = 0x0000

Event size = (N*5+2) * 4 bytes e.g., for an event with 20 hits --> 408 bytes

at 30 kHz trigger rate, this gives 11.67 Mb/s

8000 is max words

NumWordsCore from hits alone, max is (13*MAX_HITS) = 3328

Max remainder then is 4672, so we should hae 13 bits reserved for it? Per raw hit, we have 18 words, so we can do a max of... 259 words

*Check with Luca on maximum number of hits per channel.

**Waveforms at the very end. Start with some kind. https://www.phys.hawaii.edu/repos/belle2/itop

***Waveform header, waveforms, waveform footer.

Slow data types

5	FPGA temperatures
9	board temperatures
1	Humidity sensor
24	FPGA power values
10	FW/SW versions
128	Trigger scalers
1	pedestal measurement (10-bin average)
178	subtotal

TOP Production Data Format

Production Debugging 4.2

2.2 **Belle 2 TOP Data Format (Production Data)**

Note that the data listed below does NOT include protocol headers; trigger type, ctime, utime, and trgtag are included in Belle2Link headers.

Word	Bits 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16	15 14 13 12	11 10 09 08	07 06 05 04 03 02 01 00	Hit
0	Type (=0x04)	Version (=0x02)	0xA	SCROD _ID	N/A
1	EXTRA	NumWordsBonus	Phase(0-8)	NumWordsCore	
2	SKI RSVD(0000)	ctime (11 LSBs)	Revo9 Counter		
3	ASIC Masks (TimeoutMasks | Register Masks)		eventQueueDepth	eventNumberByte	
4	Carr IRSX Channel Window	0xB	tFine	WF SB **Integral**	1
5	vPeak	**H/S Heap Window 1**		**H/S Heap Window 0**	1
6	vRise0	vRise1			1
7	vFall0	vFall1			1
8	SampleRise dSampPeak dSampFall	**HeaderChecksum**			1
...					
N*(5+EXTRA)+4	SD_type Slow data	1 0 1	Nhits		N
	RAW HITS APPENDED HERE.				

= status bits

= reserved (0 for now)

= unsigned

= signed

SB: Straddle Bit

WF: Waveform Flag

Heap/Stack Flags, per Window

Sum of all 16-bit values in "hit header" = 0x0000

"1 0 1 x" = 0xC or 0xD

Event size = (N*5+2) * 4 bytes e.g., for an event with 20 hits --> 408 bytes

at 30 kHz trigger rate, this gives 11.67 Mb/s

8000 is max words

NumWordsCore from hits alone, max is (13*MAX_HITS) = 3328

Max remainder then is 4672, so we should hae 13 bits reserved for it? Per raw hit, we have 18 words, so we can do a max of... 259 words

*Check with Luca on maximum number of hits per channel.

**Waveforms at the very end. Start with some kind. https://www.phys.hawaii.edu/repos/belle2/itop

***Waveform header, waveforms, waveform footer.

Slow data types

5	FPGA temperatures
9	board temperatures
1	Humidity sensor
24	FPGA power values
10	FW/SW versions
128	Trigger scalers
1	pedestal measurement (10-bin average)
178	subtotal