ALICE
Internal Note

C.E.R.N.
Wigner R.C.

ALICE COMMON READ-OUT UNIT
ALICE-CRU

CRU Specification

Issue:	Draft
Revision:	1
Reference:	CRU
Created:	08/31/2015
Last modified:	22/01/2016
Prepared by:	E. David, T. Kiss

Abstract

This document specifies the ALICE CRU (Common Read-out Unit): the actual PCIe40 CRU hardware, the different communication forms and usage scenarios between the CRU and the detector FEs (Front-Ends) over the GBT link and the internal firmware User Logic Interface in details.

Document Status Sheet

1. Document Title: CRU Specification			
2. Document Reference Number: CRU			
3. Issue	**4. Revision**	**5. Date**	**6. Reason for change**
Draft	1	22 January 2016	First draft

Table of Contents

1 Introduction

To be done.

2 CRU Hardware Specification

2.1 Main components

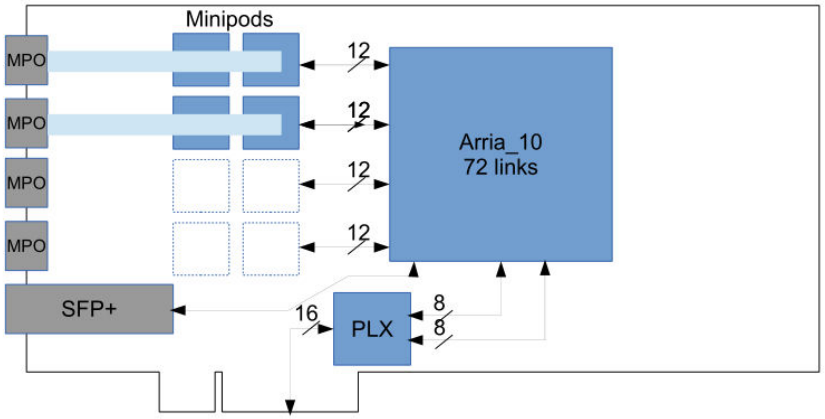


Figure 1: Block diagram of the PCIe40 board

2.1.1 FPGA

The FPGA is an Arria 10 10AX115S4F45I3SGES from Altera. This is the largest one in the Arria 10 family with 1150 kLE. It is pin compatible with a smaller one with 900 kLE (see Figure 2) which will allow a cost optimization if the application can fit in this matrix. This FPGA is also pin compatible with the future Stratix 10 with the same amount of cells which paves the way for a future speed upgrade if required.

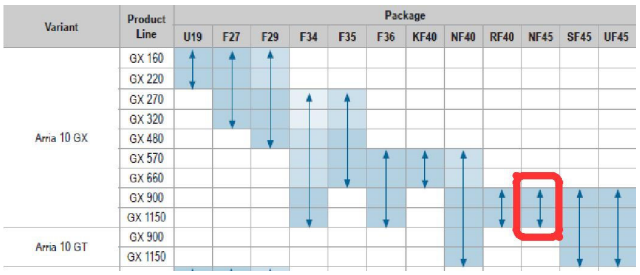


Figure 2: Pin compatible migration possibilities for the selected FPGA

Compared with the FPGA implemented on the AMC40 pins we have 1.8 times more cells than before. This should make the implementation more comfortable.

The main features are the following:

- Number of Logical Elements: 1 150 000
- Internal memory: 8.4 MB
- Variable precision DSP blocks: 1518

- 18 x 19 multipliers: 3036
- Fractional PLLs: 32
- I/O PLLs: 16
- High speed serial links at 12.5 Gbits/s: 72
- Maximum speed on serial links : 12.5 Gbps
- Maximum internal speed : 500 MHz

2.1.2 Configuration devices

An EPQS512 quad serial memory device download the configuration of the FPGA at power-up. This device can be accessed by the JTAG interface or the PCIe interface using the CvP protocol though the FPGA.

2.1.3 Optical transceivers

The board is able to interface Front-Ends with up to 48 bidirectional optical links at 10 Gbits/s each for a total bandwidth of 0.48 Tbits in each direction, even if the connection with Front-Ends is currently limited to 4.8 Gbits/s only per link. The board will be able to manage the new LP_GBT protocol when available.

This interface is implemented with 4 optical transmitters AFBR_81uVxyZ and 4 receivers AFBR_82uVxyZ from Avago also called MiniPODs as shown in Figure 3. Each of them handle 12 optical links. Each transmitter or receiver is linked to a MTP/MTP connector on the front plate.

Figure 3: MiniPOD optical interfaces from Avago

The board also embeds one SFP+ bidirectional device devoted to TFC interface. It is also possible to insert a PON device to broadcast the clock to several Front-Ends.

2.1.4 PCIe bridge

The FPGA can only manage PCIe Gen3 interfaces with 8 lanes whereas the CPU needs a PCIe Gen3 with 16 interface. The adaptation is made with a PCIe bridge PEX8747 from PLX.

The PEX8747 has 48 PCIe Lanes, implemented as 16 lanes per station across three stations. The stations are connected to one another by the internal non blocking fabric.

The adaptation 8 lanes/16 lanes is done by connecting two 8 lanes ports of the bridge to the FPGA and a 16 lanes port to the PCIe connector of the card as shown in Figure 4.

Two 8 lanes port are not used in this scheme.

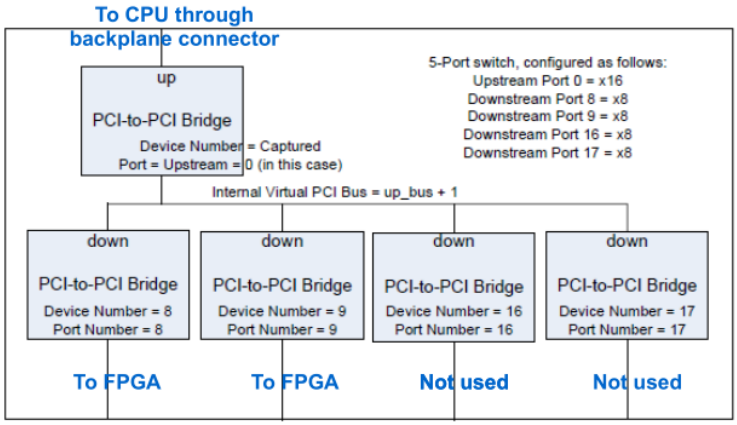


Figure 4: PEX8747 configuration

2.1.5 JTAG management

Three sources of JTAG allow to program the Arria10 GX FPGA:

- an external 10 pins connector located on the front board
- a JTAG link piloted by the PCIe
- a JTAG link piloted by an embedded USBblaster

JTAG source selection is assured by a hub implemented in a MAX V FPGA. The choice of JTAG source is made either by external straps. The overall JTAG connection is shown in Figure 5.

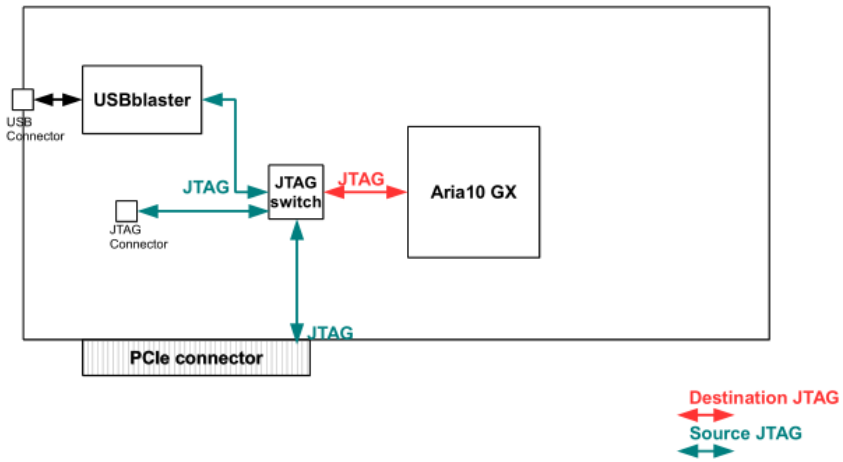


Figure 5: JTAG hub

2.1.6 Clocks

LHC clocks

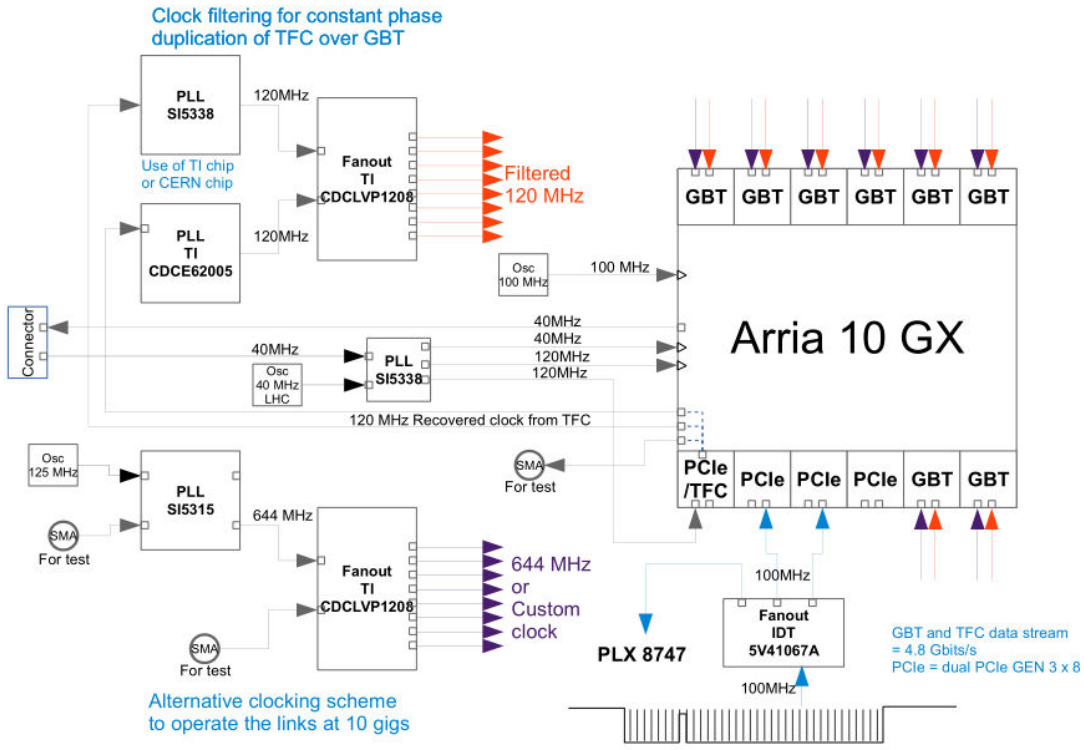


Figure 6: Clock paths

PCIe Clocks

A 100 MHz clock signal coming from the PCIe connector is fanned out to the two PCIe interfaces of the FPGA.

2.1.7 PCIe connector

The PCIe connector is a 82 pins vertical edge card connectors supporting x1, x4, x8, and x16 Link widths to suit different bandwidth requirements.

Pinout and form factor are detailed in xxx (to be done).

The PCIe40 board is compliant with PCI Express CEM r3.0 specification

2.1.8 JTAG connectors

Two 10 pin JTAG headers are mounted inside the board for programming the Arria10 GX through an USB blaster or to program the JTAG hub implemented in the Max V FPGA. It is shown in Figure 7.

Figure 7: Front plate JTAG interface 10 pin header (M40-4011046 from Harwin)

2.2 Power supply

2.2.1 Power supply estimation

The overall power consumption of the PCIe40 board is estimated to be 120 W.

2.2.2 Power source

Because it overpass the maximum 75W power consumption that can be drawn from the PCIe connector, an PCI Express 2 x 4 auxiliary power connector and cable assembly will be used as specified in the PCIe specification.

2.2.3 Power supply tree

The power tree for supplying the FPGA is illustrated in Figure 8. Other devices like PLLs, temperature sensors are connected on the same tree through passive filters.

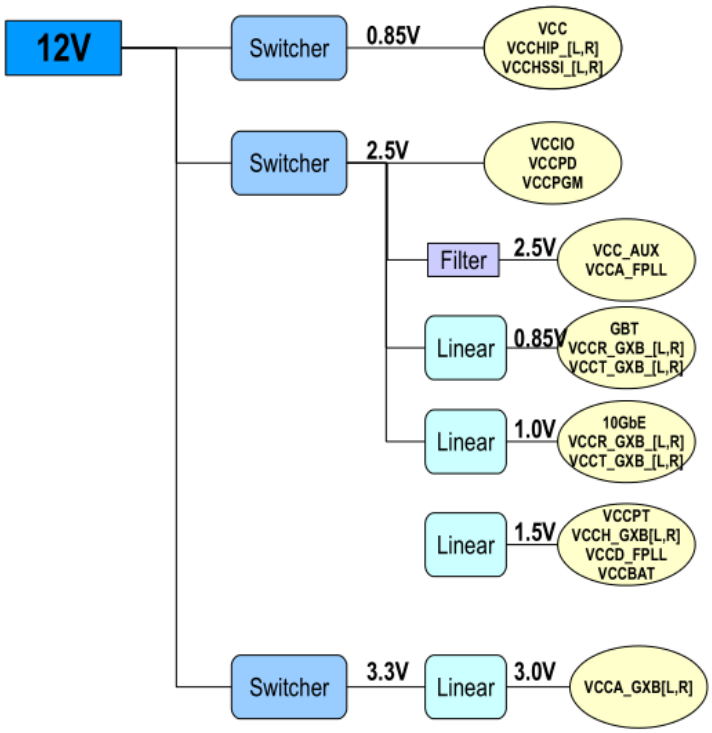


Figure 8: Power supply tree for supplying the Arria10 GX FPGA

2.3 Cooling

2.3.1 Requirements

An air flow of 2 m/s is required to maintain the FPGA temperature under the maximum bearable limit.

2.3.2 Implementation

To be defined.

2.4 Mechanics

2.4.1 Dimensions

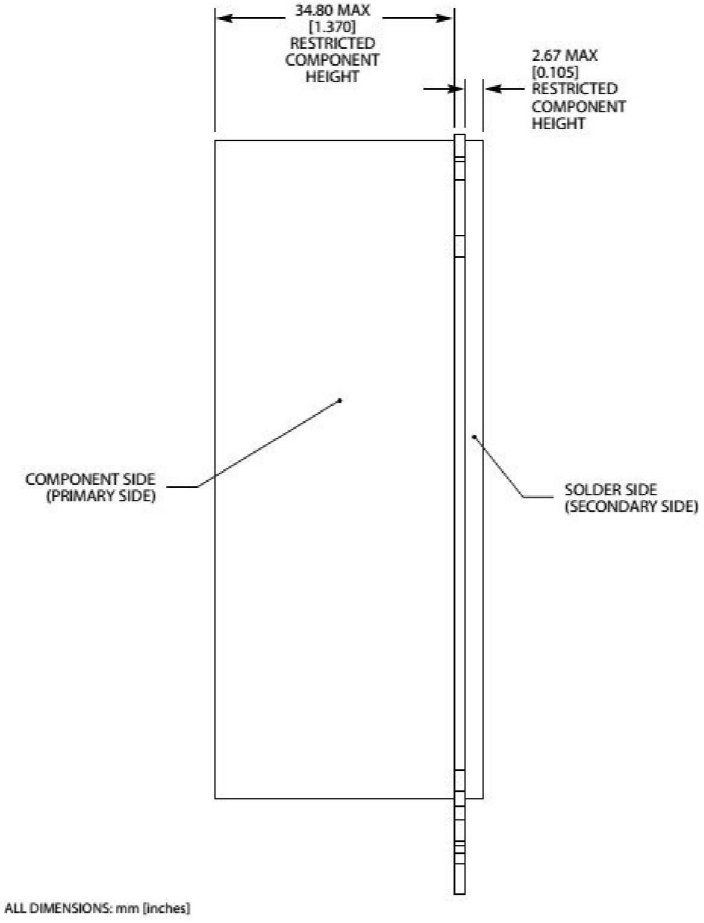


Figure 9: Board dimensions

The form factor is Standard Height (111.28 mm) Half Length (167.65 mm).

2.4.2 Face plate

The layout of MTP/MTP , SFP+ , JTAG and USB connectors is shown in Figure 10.

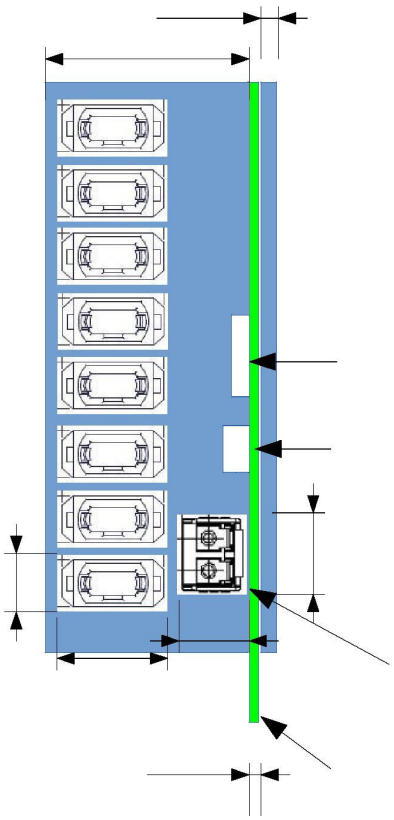
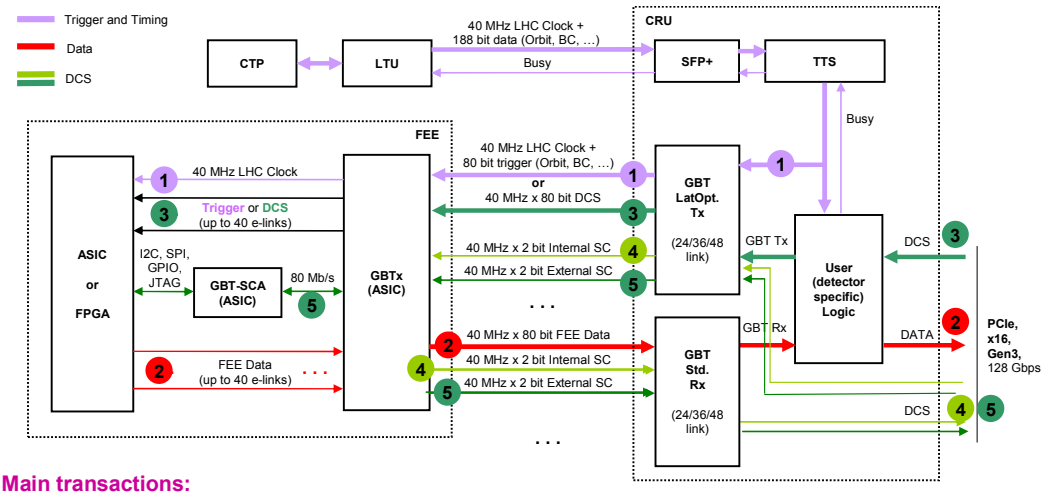


Figure 10: Face plate

3 CRU Front-End Interface Specification

3.1 Introduction

The aim of this chapter is to specify in details the different communication forms between the Common Read-out Unit (CRU) and the detector front-end electronics (FE) over the bidirectional GBT link.



- **Main transactions:**
  1. Delivering the TTS information or Readout Control from CRU to FE
  2. Receiving the detector data packets in serial and parallel form from the FE
  3. Delivering custom packets from CRU to FE FPGA in parallel form
  4. Sending and receiving packets to and from GBTx internal register block
  5. Sending and receiving packets to and from GBT-SCA ASIC

Figure 11: Different communication forms between the CRU and the FEs

Figure 11 shows a generic CRU and FE setup with different communication forms. The communication forms organized into three groups based on the functionality:

- Trigger related communications (clock, trigger, readout)
- Detector data transfer related communications
- Slow control related communications

After evaluating the different detector requirements the following required GBT downlink and uplink feature usages has been identified:

GBT Downlink features:

- LHC clock distribution – The CRU delivers the 40 MHz LHC clock with deterministic phase to the FE over the GBT downlink.
- Trigger information distribution – The CRU receives the full TTS information from the LTU (up to 192 bit) and is able to deliver some subset of this trigger information (up to 80 bit) over the GBT downlink with deterministic latency.

- Readout control distribution – Some detectors will implement the readout control based on the trigger information inside the CRU instead sending the raw trigger information to the FE. This mode allows sending any custom trigger information with deterministic latency.
- Parallel packet based communication – Unidirectional packet based communication from CRU to FE.
- Parallel single word communication - Unidirectional single word based communication from CRU to FE.
- GBTx ASIC communication (GBT IC field) – GBTx serial packet sending from CRU to FE.
- GBT-SCA communication (GBT EC field) – GBT-SCA serial packet sending from CRU to FE.

GBT Uplink features:

- Raw GBT data stream
- Parallel packet based communication – Unidirectional packet based communication from FE to CRU.
- Parallel single word communication - Unidirectional single word based communication from FE to CRU.
- GBTx ASIC communication (GBT IC field) – GBTx serial packet sending from FE to CRU.
- GBT-SCA communication (GBT EC field) – GBT-SCA serial packet sending from FE to CRU.

GBT Downlink	TPC	MCH	MID	TOF	FIT	ZDC	ITS	MFT	TRD	CTP
1. LHC clock distribution	X	X	X	X		?				
2. Trigger information distribution				X	?	?				
3. Readout control distribution	X	X	X							
4. Parallel packet based communication							X	X		
5. Parallel single word communication							X	X		X
6. GBTx ASIC communication (IC field)	X	X	X		?	?	?	?		
7. GBT-SCA communication (EC field)	X	X	X		?	?	X	X		
GBT Uplink										
1. Independent serial raw data streams	X	X	X							
2. Parallel packet based communication				X	?	?	X	X	?	X
3. Parallel single word communication							X	X		X
4. GBTx ASIC communication (IC field)	X	X	X		?	?	?	?		
5. GBT-SCA communication (EC field)	X	X	X		?	?	X	X		

Table 1: GBT Link Feature Chart

Table 1 summarizes the different GBT downlink and uplink feature utilizations by different detectors.

3.2 Terminology

CRU – Common Read-out Unit

Core/Common CRU firmware – CRU firmware provided by CRU team

Detector specific logic – CRU firmware modules developed by detector teams

GBT – GigaBit Transceiver

CTP – Central Trigger Processor

LTU – Local Trigger Unit

TTS downlink – LTU -> CRU direction

TTS uplink – CRU -> LTU direction

GBT downlink – CRU -> FE direction

GBT uplink – FE -> CRU direction

3.3 Trigger distribution related features

The following chapters describe the trigger distribution related features in details. Each feature is unidirectional and handled by the GBT downlink.

3.3.1 LHC clock distribution

The CRU firmware is able to receive the 40 MHz LHC clock information from CTP/LTU and reconstruct the extract LHC clock with known phase relation to the origin. The firmware is able to synchronize the GBT frame header to the LHC clock rising edge and propagate the LHC clock information to the FE over the GBT downlink.

The LHC clock distribution feature has no payload overhead, the latency optimized GBT-FPGA implementation provides this feature.

3.3.2 Trigger information distribution

In this mode the CRU will propagate a subset of the whole TTS information received from CTP/LTU over the GBT downlink (80 bit out of 192 bit). This feature will be implemented in the common CRU firmware.

This mode can be used when there is a capable onboard logic on the FE side and able to work based on the raw TTS information.

3.3.3 Readout control distribution

In this mode the TTS information will be not sent directly to the FE, but instead to a detector specific firmware module inside the CRU firmware and that detector specific logic will control the actual GBT downlink depending on the specific FE needs.

This mode can be used when there is no capable logic on the FE to but only a dedicated ASIC like logic and the more complex readout control can be moved to the CRU side.

3.4 Data transport related features

The following chapters describe the data transport related features in details. Two main modes have been identified: raw data stream mode and the packet based mode.

3.4.1 GBT raw data stream mode

In this mode the raw GBT payload (80 bits @ 40 MHz) is not processed by the core CRU firmware but passed directly to the detector specific user logic module inside the CRU firmware. This mode allows any custom communication form over the GBT uplink between the FE and the CRU.

3.4.2 Parallel packet and single word based protocol

In this mode the core CRU firmware implements an ALICE specific protocol over the raw GBT data stream. This protocol is similar to the existing DDL protocol where the FE is able to send a well formed packetized detector data toward the CRU. The CRU receives the well formed packetized data and transfers it into the FLP servers memory.

The main properties of the packet based protocol:

- Unidirectional - Due to the very diverse GBT downlink/uplink usage scenarios requiring a bidirectional connection between the CRU and the FE would be a limiting factor.
- Operates with the full 80 bit GBT payload words where the GBT words classified as data or control words based on the GBT data valid feature.
- Requires a capable FPGA presented on the FE card which is able to build a well formed SDH packages.

Figure 12 and Figure 13 shows the GBT DATA and CONTROL word structure.

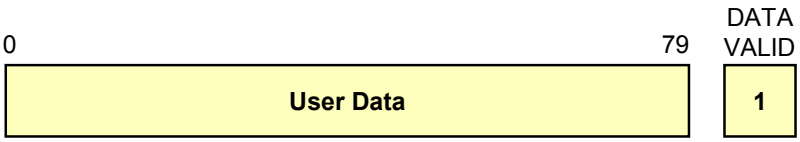


Figure 12: The 80 bit GBT Word Marked as DATA

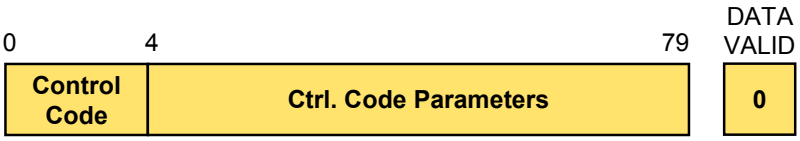


Figure 13: The 80 bit GBT Word Marked as CONTROL

Each transmitted 80 bit GBT payload falls into two categories: data or control. If the data valid bit is 1 then the 80 bit GBT word must be treated as part of the data transfer. If the data valid is 0 then the 80 bit GBT word must be treated as control word where the exact control word type is encoded in the “Control Code” field.

Currently four control codes are supported: IDLE, SOP, EOP and SWT (see Figure 14, Figure 15, Figure 16, Figure 17).

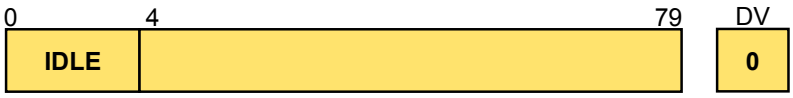


Figure 14: IDLE Control Word Format

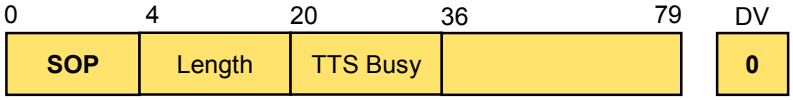


Figure 15: Start of Packet Control Word Format

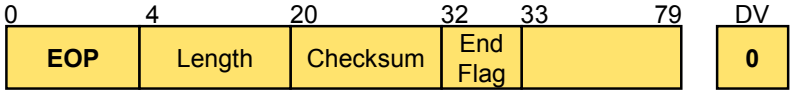


Figure 16: End of Packet Control Word Format

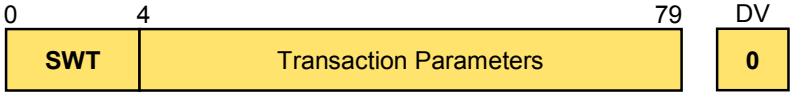


Figure 17: Single Word Transaction Control Word Format

Packet based transfer

The most typical usage of the packet based transfer is the FE sending the detector data to the CRU. Figure 18 shows an example CRU-FE message/packet transmission. The message must be start on GBT word aligned (80 bit) position. The message starts with one or more header control words, and then it follows by the payload words, and ends with a tail control word. The purpose of CRU-FE message is to encapsulate the upper layer messages (SDH or DCS packet).

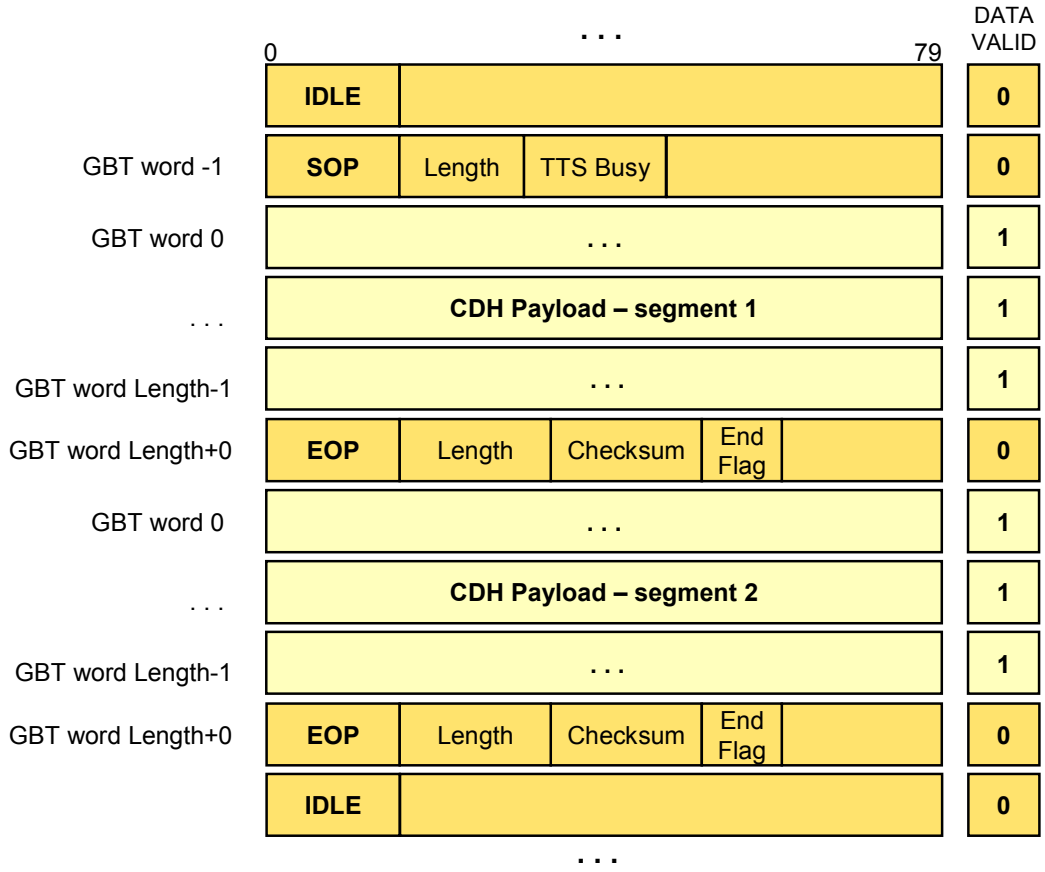


Figure 18: Packet Transmission Example

Figure 19 shows the packet sending sequence from the FE to the CRU in detailed steps.

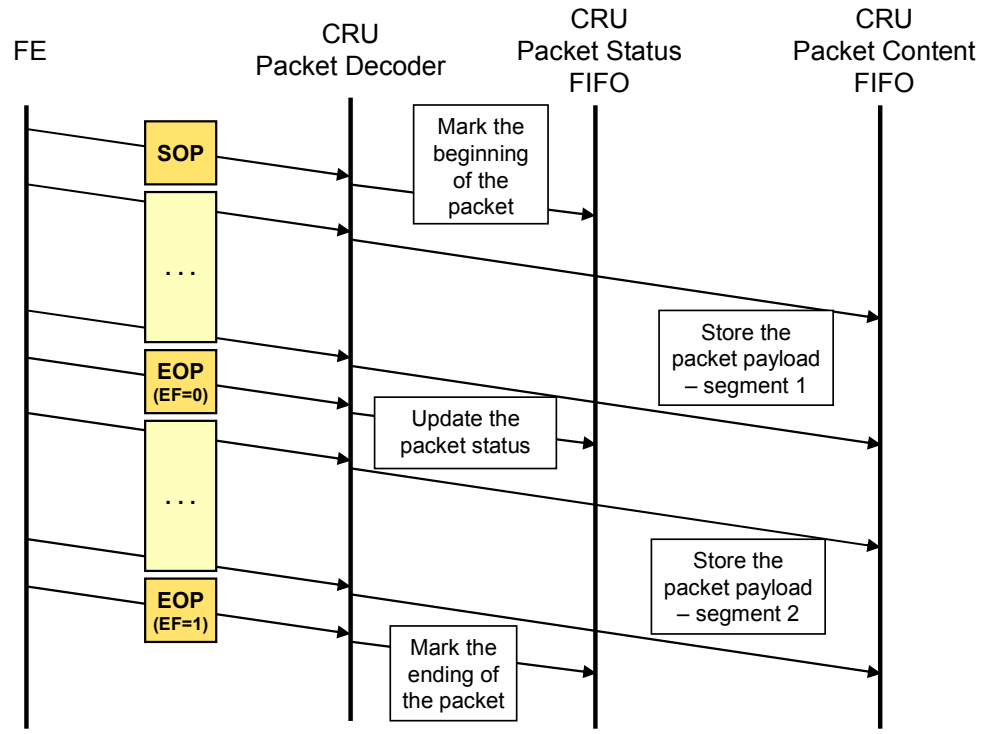


Figure 19: FE to CRU Detector Data Packet Transfer

Figure 20 shows the packet sending sequence from the CRU to the FE in detailed steps.

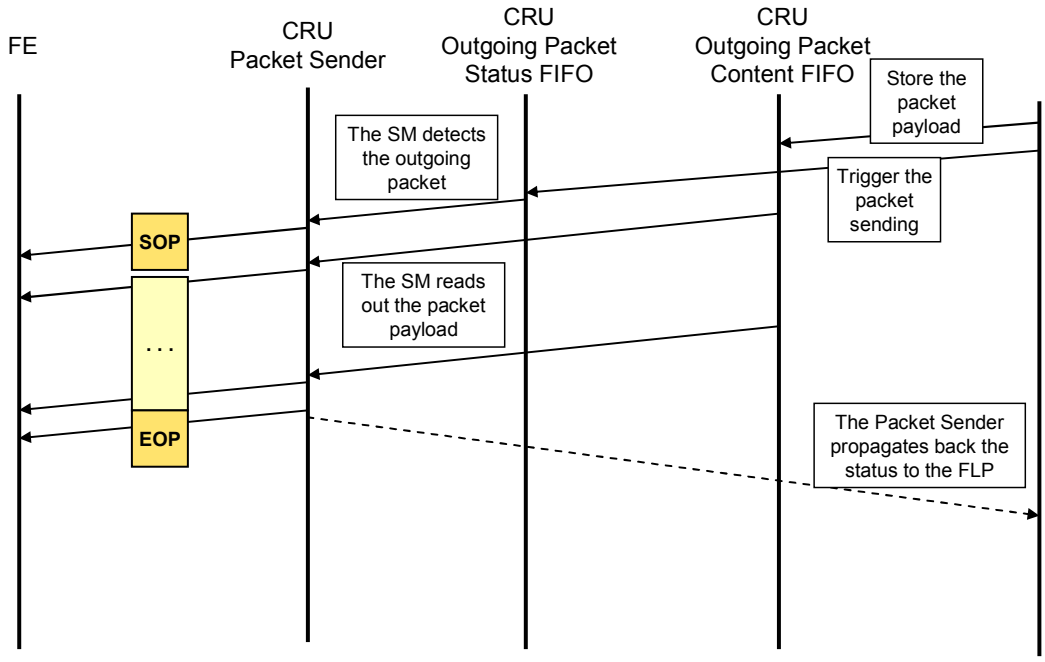


Figure 20: CRU to FE Packet Transfer

Single word transaction

While the packed based data transfer can be optimal for a large detector data packet transmissions there are some cases when a simpler protocol would be more suitable for the communication between the CR and the FE (e.g. controlling and monitoring some registers from the CRU). For this reason the single word transaction is also included in the specification.

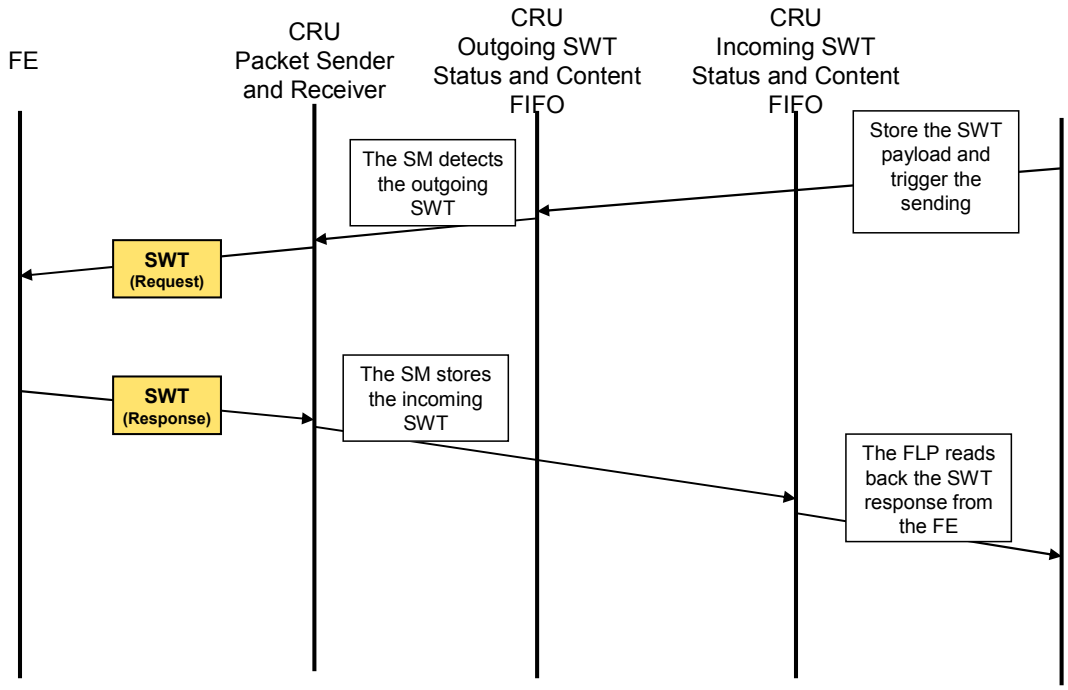


Figure 21:CRU to FE Single Word Transaction

Figure 21 shows the single word transaction steps between the CRU and the FE.

3.5 Slow control related features

The following chapters describe the supported slow control related features. The GBT frame contains two dedicated bit field for the slow control purposes: the 2 bit IC filed for GBTx ASIC communication and the 2 bit EC field for GBT-SCA communication. The common CRU firmware supports both channel and also enables to the FE to utilize the packet based protocol for slow control.

3.5.1 GBTx ASIC communication

In the case when the GBTx ASIC pin configSelect (K7) tied to 0 there is a possibility to access the GBTx ASIC internal registers from CRU firmware through the 2 bit IC field. The CRU firmware and tools fully support this feature.

3.5.2 GBT-SCA communication

The GBT frame contains 2 dedicated bits for external slow control applications (the EC field). This 2 bit @ 40 MHz field is translated to a bidirectional 80 Mbps E-Link:

- 80 MHz ePort clock output: SCCLKN (pin R6), SCCLKP (pin P6);
- 80 Mbps ePort data input: SCINN (P13), SCINP (P12);
- 80 Mbps ePort data output: SCOUTN (P3), SCOUTP (P4)

The CRU firmware and tools fully support the communication with GBT-SCA over this dedicated E-Link.

3.6 Supported FE Architectures

The CRU firmware supports the most common detector front-end architectures built from CERN developed building blocks (like Versatile Link VTRx and VTTx, GBTx ASIC, GBT-SCA, custom front-end ASICs, etc).

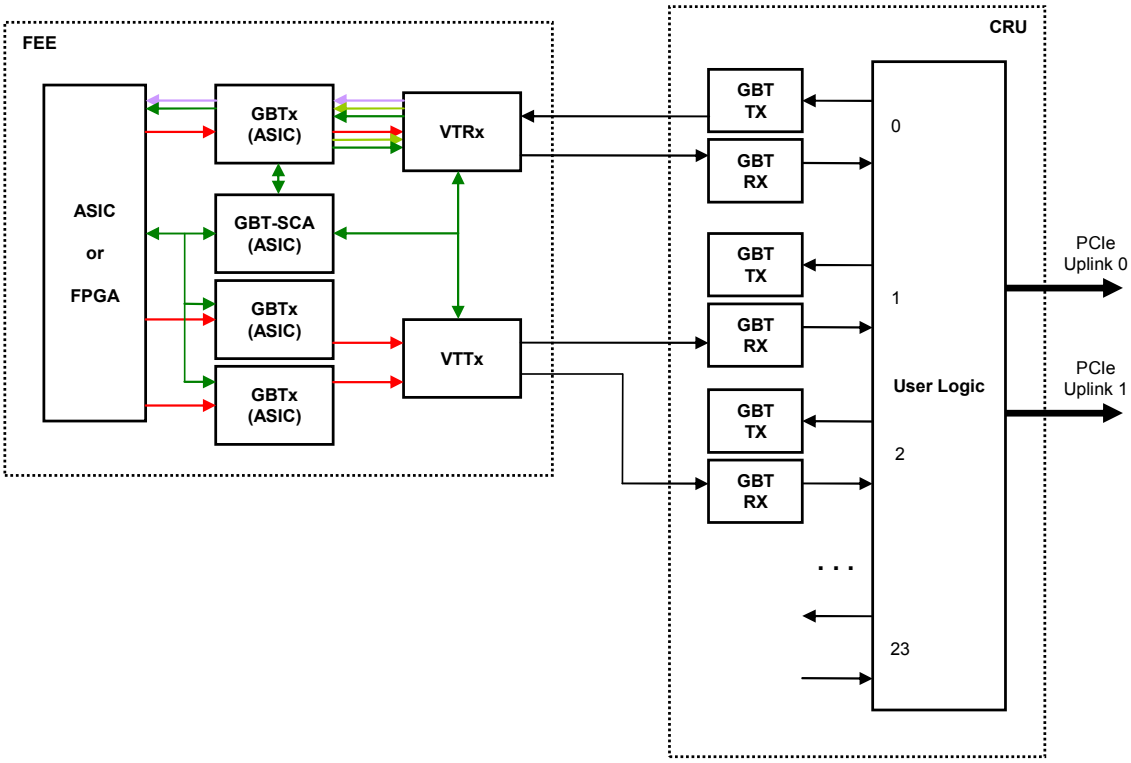


Figure 22: Supported Front-End Architectures

Figure 22 summarize the supported scenarios. The CRU provide 24 fully bidirectional GBT links but it is not mandatory for the FE card to utilize both direction. A typical FE card will be based around at least one bidirectional VTRx module and several additional VTTx transmit only module.

The primary GBTx ASICs configuration/reconfiguration is supported through the GBTs IC (internal control) channel. The additional GBTx ASICs should be connected to the GBT-SCA I2C master ports to enable the configuration/reconfiguration through the CRU and in this case the configuration/reconfiguration goes through the GBTs EC (external control) channel.

3.7 Detector specific usage scenarios

The following chapters will present the detector specific scenarios in more details.

3.7.1 TPC

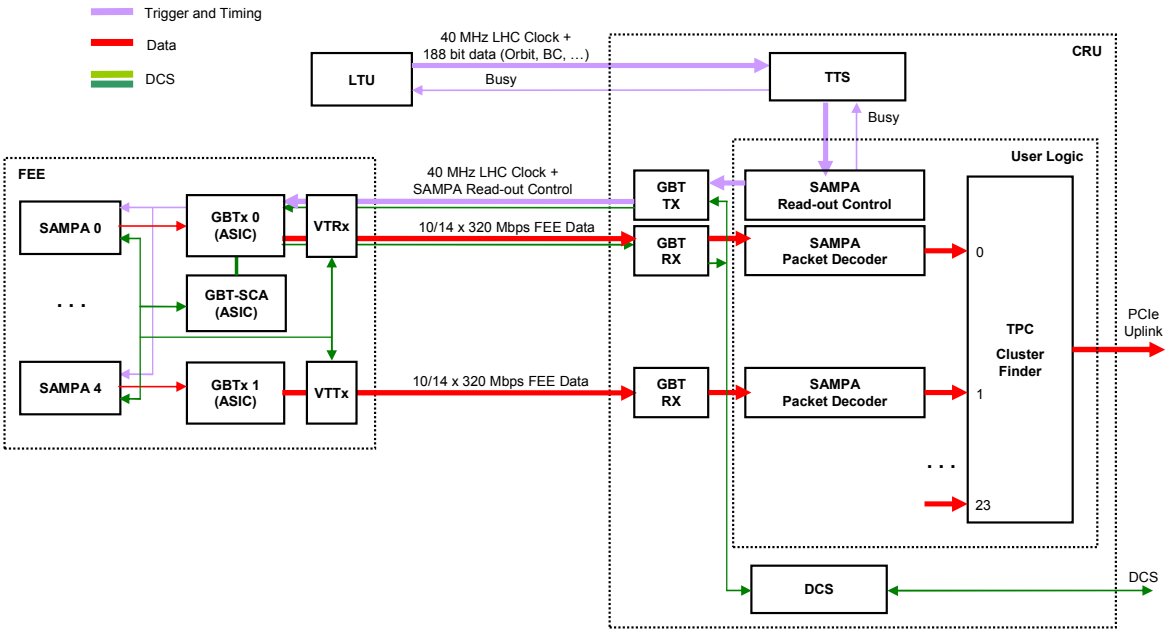


Figure 23: TPC Usage Scenario Block Diagram

To be synchronized with latest TPC readout scheme.

The TPC FE contains five SAMPA ASICs, two GBTx ASIC and one GBT-SCA ASIC. Depending on the detector traffic requirements it will use at least one bidirectional GBT link (VTRx) and one additional transmit only GBT link (VTTx).

The SAMPA ASICs initial configuration is supported through the GBT-SCA I2C master ports, and also if second GBTx ASIC configuration needed it can be achieved through the GBT-SCA I2C master port.

The SAMPA ASIC requires a continuous read-out control through several e-links connected directly to the SAMPA pins (reset, heartbeat, bunch crossing and sync). This read-out control logic must be implemented inside the CRU firmware's User Logic module by the TPC team where the TTS information is fully available.

Each SAMPA ASIC is able to produce outgoing packages on up to four/six e-links. Those e-links are aggregated by the GBTx ASIC chips. On the CRU side the "SAMPA Packed Decoder" module implemented inside the User Logic block by TPC team.

If the TPC chose to use the GBT Uplink in Wide Bus mode (112 bits @ 40 MHz) the CRU GBT block and the User Logic interface must be rebuilt by Wide Bus support by the TPC team.

The TPC Cluster Finder is implemented inside the CRU User Logic module by the TPC team.

3.7.2 MCH

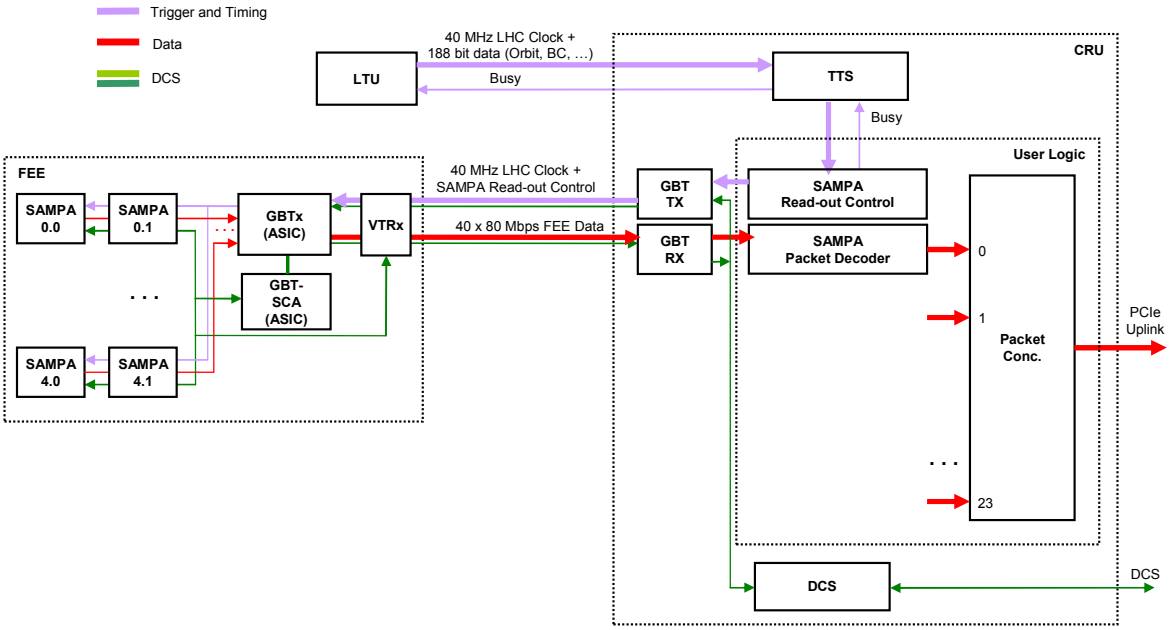


Figure 24: MCH Usage Scenario Block Diagram

To be done.

3.7.3 MID

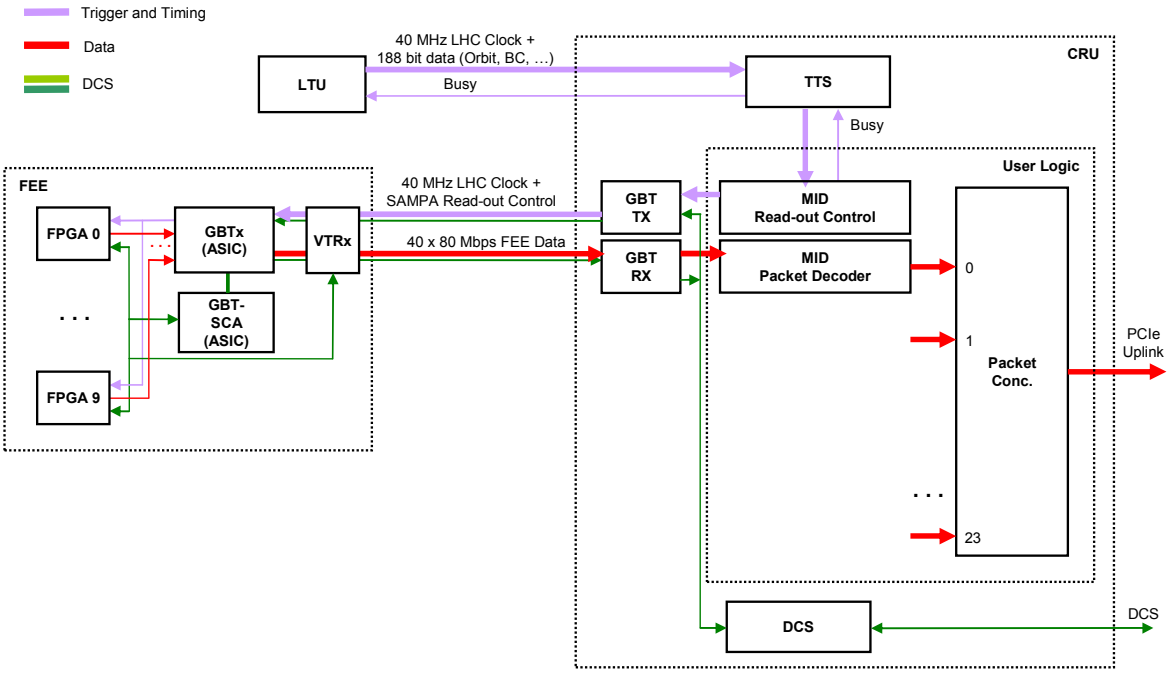


Figure 25: MID Usage Scenario Block Diagram

To be done.

3.7.4 ITS

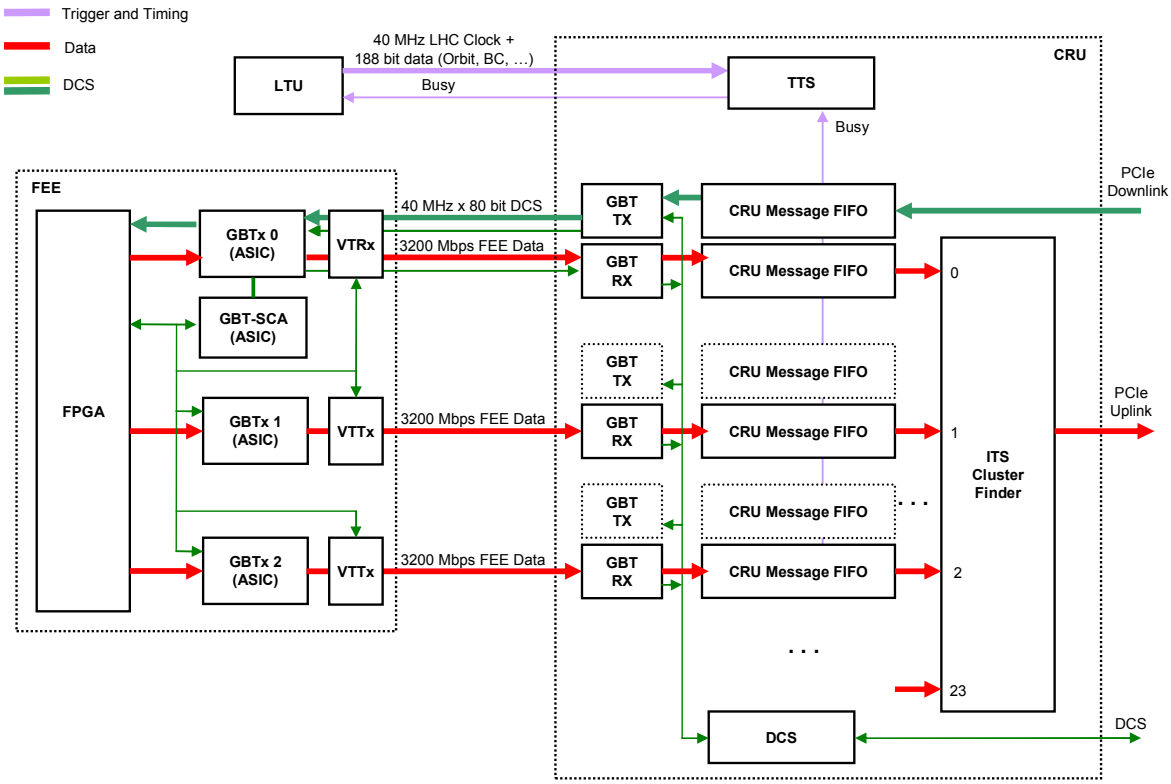


Figure 26: ITS Usage Scenario Block Diagram

The ITS FE contains an onboard FPGA (Xilinx Kintex 7), one GBTx ASIC and one GBT-SCA connected to one bidirectional VTRx module. The two additional GBT unidirectional uplinks (VTTx module) are driven directly from the FPGA (see Figure 26).

In the case of ITS the GBT downlink is available for the slow control purposes because the ITS FE is receiving the TTS information directly from the LTU and not through the CRU.

The ITS FE is able to build well formed packages with SDH and implement the packed based protocol to propagate the detector data to the CRU.

If there is no need for detector data processing in the CRU but just transferring the aggregated detector data to FLP server's memory then the common CRU firmware can be used. If the ITS team choose to do Cluster Finding inside the CRU firmware then it has to implemented inside the user logic module.

3.7.5 MFT

The MFT's usage scenario is very similar to the ITS usage scenario.

3.7.6 TOF

To be done.

3.7.7 CTP

To be done.

4 CRU Firmware User Logic Interface Specification

4.1 Introduction

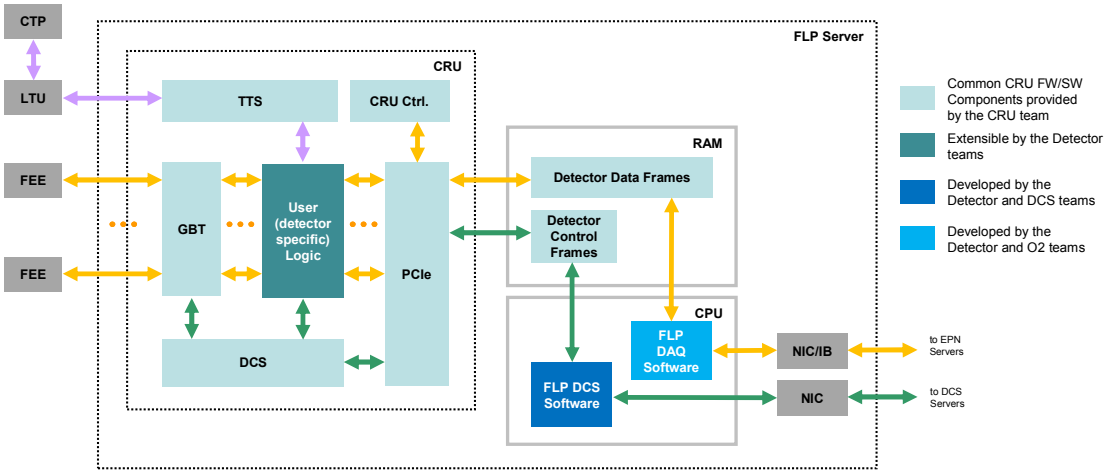


Figure 27: CRU Firmware Block Diagram

To be done.

4.2 Port Specifications

TTS Downlink (LTU -> CRU) Ports

Port Name	Direction	Clock Domain	Description
tts_rx_clk_i	Input	N/A	
tts_rx_rdy_o	Output	tts_rx_clk_i	
tts_rx_valid_i	Input	tts_rx_clk_i	
tts_rx_data_i[N-1:0]	Input	tts_rx_clk_i	

TTS Uplink (CRU -> LTU) Ports

Port Name	Direction	Clock Domain	Description
tts_tx_clk_i	Input	N/A	
tts_tx_rdy_i	Input	tts_tx_clk_i	
tts_tx_data_o[15:0]	Output	tts_tx_clk_i	

GBT Downlink (CRU -> FE) Ports

Port Name	Direction	Clock Domain	Description
Gbt_tx_clk_i[N-1:0]	Input		
Gbt_tx_rdy_i[N-1]	Input	gbt_tx_clk_i	
Gbt_tx_isdata_o[N-1]	Output	gbt_tx_clk_i	
Gbt_tx_valid_o[N-1:0]	Output	gbt_tx_clk_i	
Gbt_tx_data_o[N*84-1:0]	Output	gbt_tx_clk_i	

GBT Uplink (FE -> CRU) Ports

Port Name	Direction	Clock Domain	Description
Gbt_rx_clk_i	Input		
Gbt_rx_rdy_o	Output	gbt_rx_clk_i	
Gbt_rx_isdata_i	Input	gbt_rx_clk_i	
Gbt_rx_valid_i	Input	gbt_rx_clk_i	
Gbt_rx_data_i[N*84-1:0]	Input	gbt_rx_clk_i	

PCIe Uplink (CRU -> FLP) Ports

Port Name	Direction	Clock Domain	Description
pcie_clk250mhz_i[1:0]	Input	N/A	
pcie_st_sink_clk_o[1:0]	Output	pcie_clk250mhz_i	
pcie_st_sink_data_o[511:0]	Output	pcie_clk250mhz_i	
pcie_st_sink_valid_o[1:0]	Output	pcie_clk250mhz_i	
pcie_st_sink_ready_i[1:0]	Input	pcie_clk250mhz_i	

Avalon-MM Slave Ports - PCIe BAR Read/Write access

Port Name	Direction	Clock Domain	Description
mms_clk	Input	N/A	
mms_reset	Input	mms_clk	
mms_address[15:0]	Input	mms_clk	
mms_write	Input	mms_clk	
mms_writedata[31:0]	Input	mms_clk	

mms_read	Input	mms_clk	
mms_readdata[31:0]	Output	mms_clk	
mms_rvalid	Output	mms_clk	

PCIe Downlink (CRU -> FLP) Ports

To be done.

GBT-SCA Downlink (CRU -> FE) Ports

To be done.

GBT-SCA Uplink (FE -> CRU) Ports

To be done.

5 CRU Software Specification

5.1 Introduction

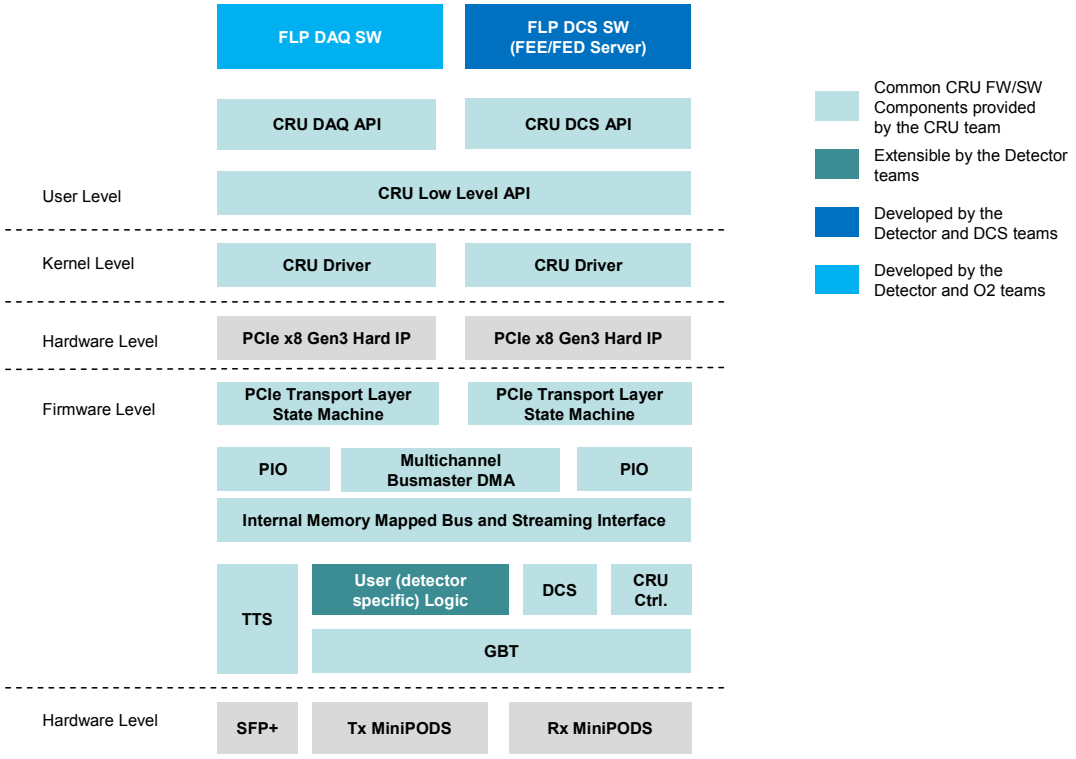


Figure 28: CRU Software Stackup Overview

To be done.

5.2 Specifications

The CRU Driver layer provides:

- Access and enumeration of the PCIe resources (BARs, interrupts)
- Interrupt service routine
- NUMA aware memory allocation for DMA
- Synchronization support for multi-threading

The CRU API layer provides:

- Controlled access to the driver features

- DMA data transfer management functions
- Low level C functions for accessing the register based interfaces provided by the CRU firmware (for example functions accessing the GBT-SCA I2C masters and JTAG port)

To be done.