Common Readout Unit (CRU) – A New Readout Architecture for ALICE Experiment

Jubin Mitra
VECC, Kolkata, India

For the ALICE Collaboration

TWEPP 2015
Topical Workshop on Electronics for Particle Physics
28 September - 02 October 2015
IST - Lisbon - Portugal

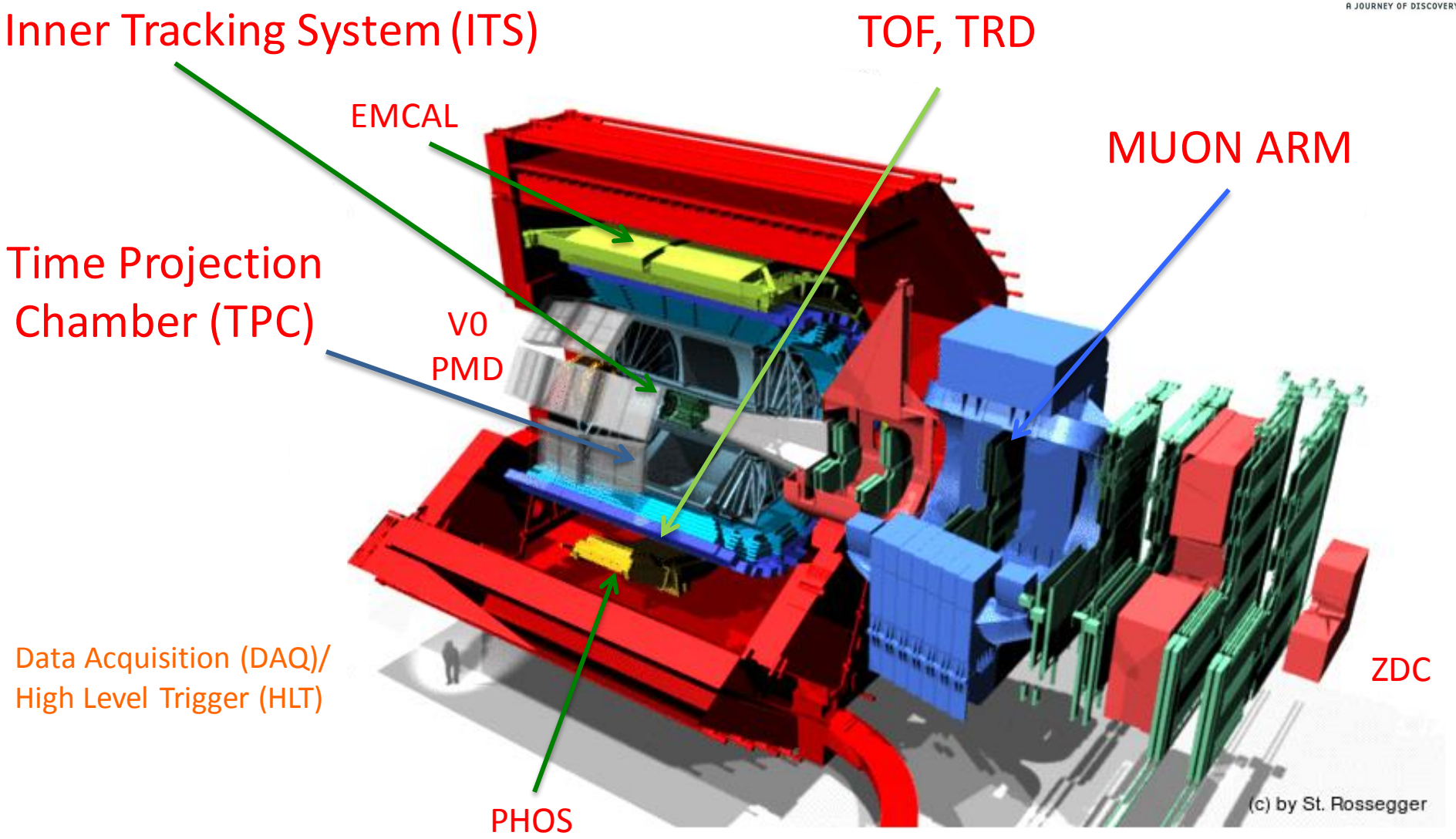
Outlook

1) Motivation of building CRU
2) Understanding CRU connection and functionality
3) How we choose the CRU form factor
4) Firmware Development Status

Part I :

Motivation of Building CRU

ALICE at CERN-LHC (Now)



- **ALICE is a dedicated experiment for study of Quark-Gluon Plasma (QGP) with pp, p-Pb and Pb-Pb collisions at the LHC**

ALICE Upgrade: > 2021

New Inner Tracking System (ITS)

- Improved pointing precision
- Less material -> thinnest tracker at the LHC
- **25x10^9 channels**

Time Projection Chamber (TPC)

- New Micropattern gas detector technology
- **Continuous readout**

New Central Trigger Processor (CTP)

Online Offline Systems (O2)

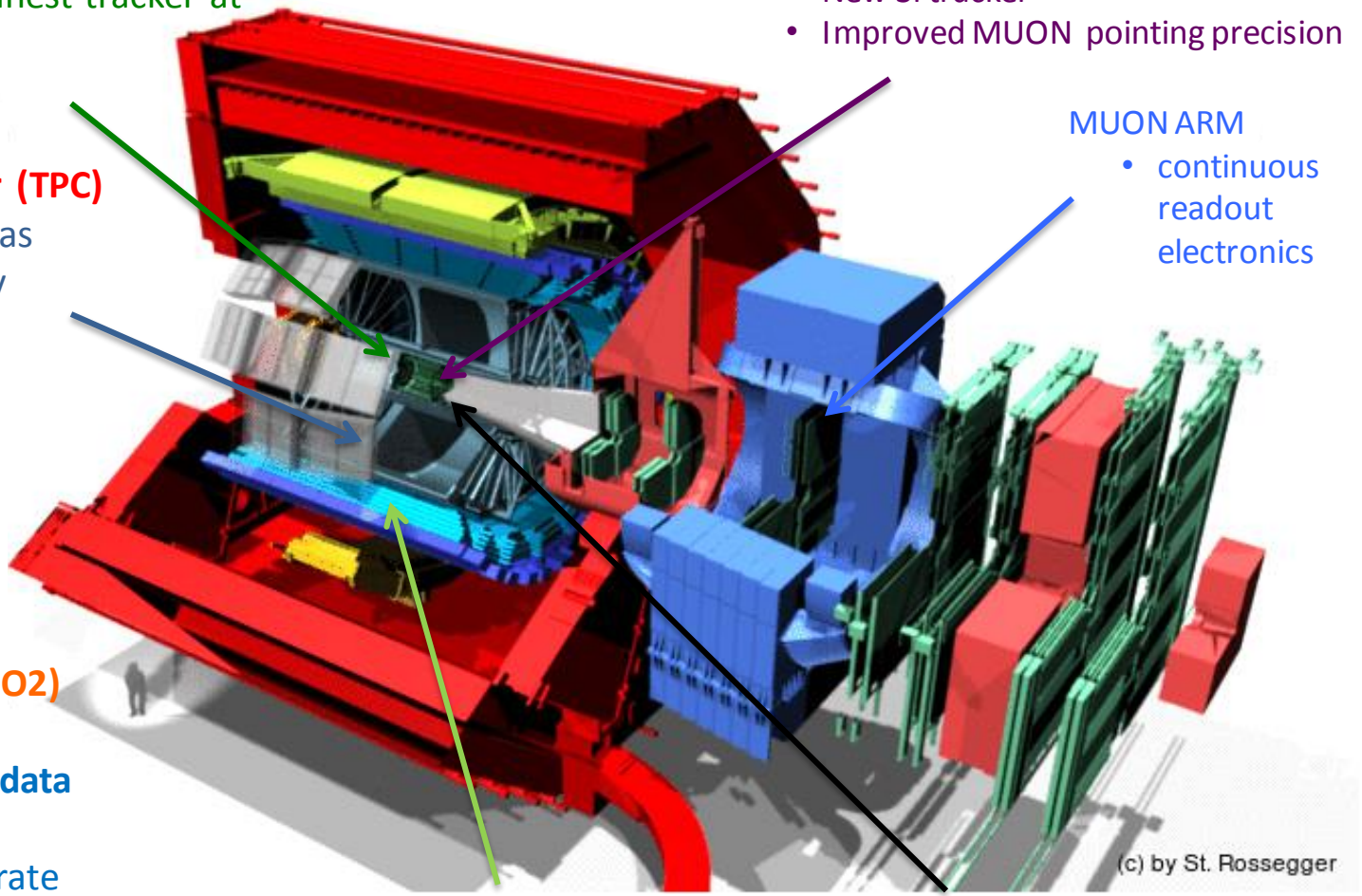
- New architecture
- **On line tracking & data compression**
- 50kHz PbPb event rate

Muon Forward Tracker (MFT)

- New Si tracker
- Improved MUON pointing precision

MUON ARM

- continuous readout electronics



TOF, TRD

- **Faster readout**

New Trigger Detectors (FIT)

Why Do We Need CRU ?

LS1 (2013-14) Present (RUN2)	ALICE UPGRADE	After LS2 * (2019 - 20) Future (RUN3/RUN4 – YEAR 2021)
1 nb^{-1} (PbPb)	**Collisions and Collection**	>10 nb^{-1} (PbPb) 6 pb^{-1} (pp) 50 nb^{-1} (pPb)
$10^{27}cm^{-2}\ s^{-1}$	**At Peak Luminosities**	**6 x $10^{27}cm^{-2}\ s^{-1}$**
8 kHz (PbPb)	**Corresponding to Collision Rate Of**	**50 kHz (PbPb)** 200 kHz (pp and pPb)
500 Hz (PbPb)	**Maximum Readout Rate**	**>200 kHz (PbPb)** **1 MHz (pp)**
Hardware triggers • Event multiplicity • Calorimeter energy • Track p_T	**Trigger Mechanism**	• A minimum bias event (Non- Upgraded detector) • **A self-triggered Continuous fashion (upgraded detector)**

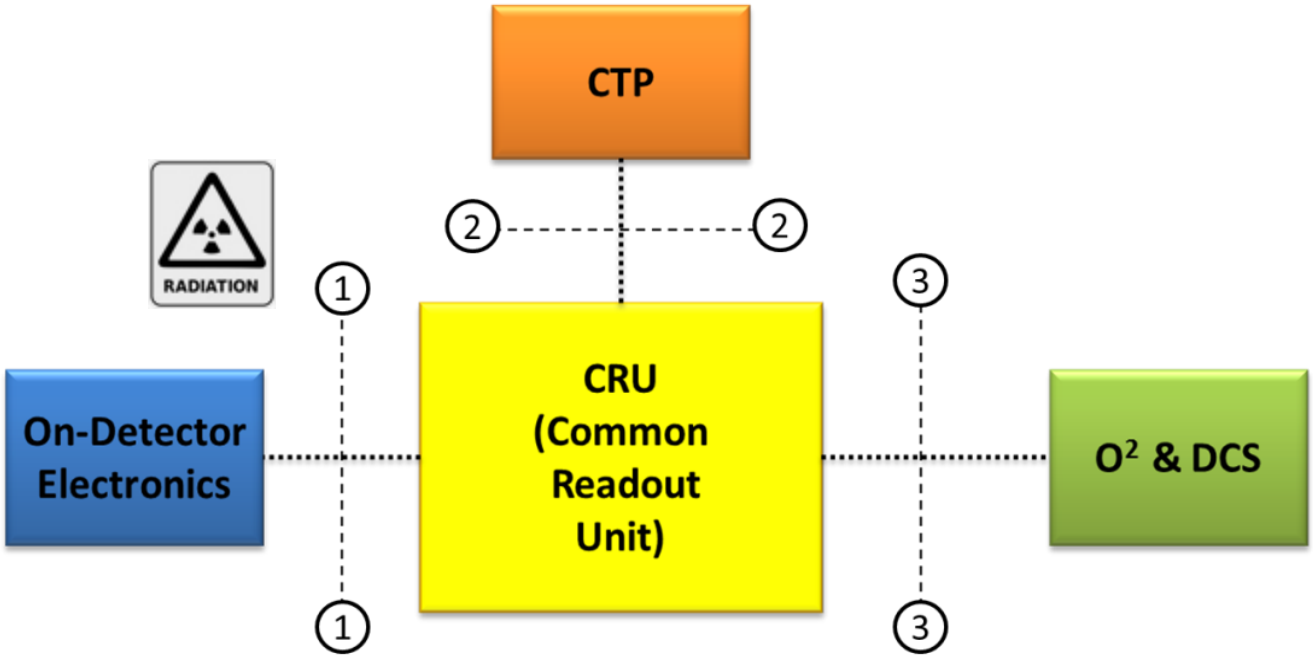
***To summarize*:**

TO HANDLE THE DATA FLOW TRAFFIC OF ABOUT 1 TB/S

Part II :

Understanding CRU Connection And Functionality

Common Readout Unit (CRU)



CRU has three interfaces:

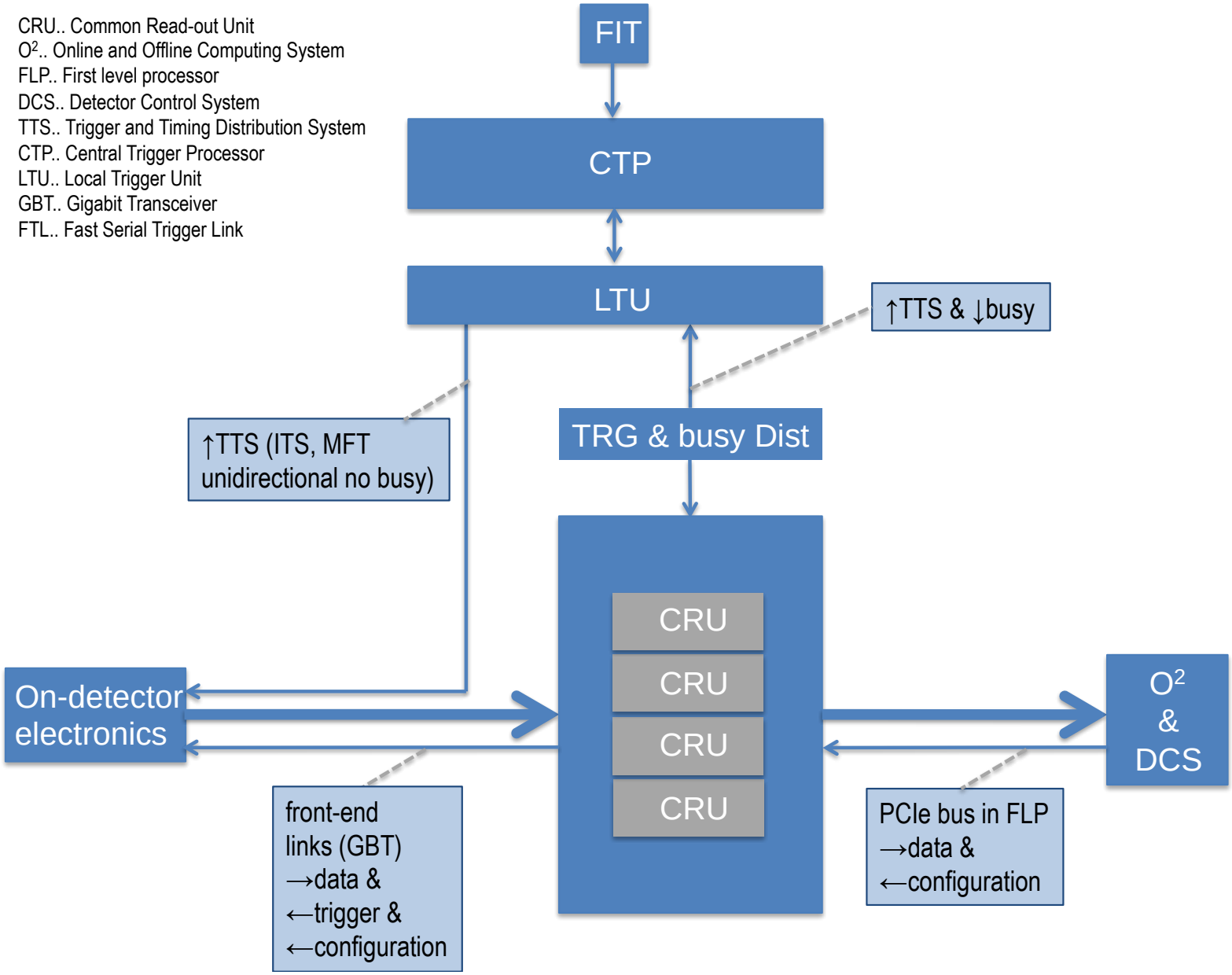
1 – **GBT Link (Radiation Tolerant High Speed Optical Link)**

2 – To be decided (GBT/**10Gigabit PON**)

3 – **DDL3 link (PCIe Gen 3 x16)**

Read-out and Trigger Distribution System

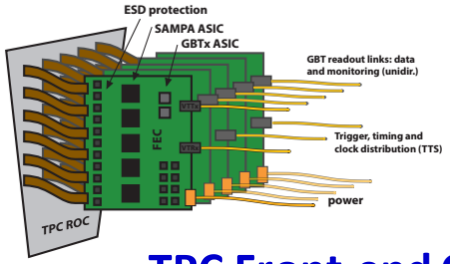
CRU.. Common Read-out Unit
O^2.. Online and Offline Computing System
FLP.. First level processor
DCS.. Detector Control System
TTS.. Trigger and Timing Distribution System
CTP.. Central Trigger Processor
LTU.. Local Trigger Unit
GBT.. Gigabit Transceiver
FTL.. Fast Serial Trigger Link



ALICE Detector readout board types

Detector	Link type	Number of links DDL1 (2 Gbps)	DDL2 (4-5 Gbps)	GBT (4.8 Gbps)	Read-out board type	Number of boards C-RORC	CRU
ACO	DDL1	1			C-RORC	1	
CPV	DDL1	6			C-RORC	1	
CTP	GBT			14	CRU		1
EMC	DDL2		20		C-RORC	4	
FIT	DDL2		2		C-RORC	1	
HMP	DDL1	14			C-RORC	4	
ITS	GBT			495	CRU		23
MCH	GBT			550	CRU		25
MFT	GBT			304	CRU		14
MID	GBT			32	CRU		2
PHS	DDL2		16		C-RORC	4	
TOF	GBT			72	CRU		3
TPC	GBT			5832	CRU		324
TRD	Custom			1044	CRU		54
ZDC	GBT			1	CRU		1
Total		21	38	8344		15	447

Example CRU connection with TPC



TPC Front-end Card

The CRU re-orders the data samples according to their position in the pad row allowing a more efficient cluster search.

Data Storage

Storage Network

DCS

Control and Monitoring

SAMPA

SCA

GBTX

Physics & Monitoring Data

GBT

GBT

Trigger, Control and Configuration

CRU

FPGA

Physics Data

PCIe x16

FLP

Farm Network

EPN

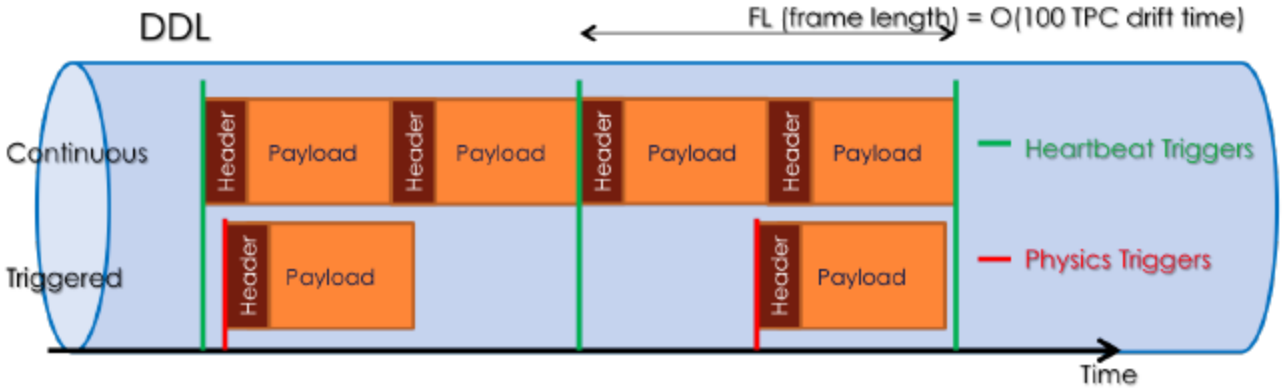
1st stage of data reduction using cluster or tracklet finder

2nd level of data reduction. Reconstruct the tracks and associate them to their primary vertex

Pic courtesy ALICE TPC TDR

Trigger Types

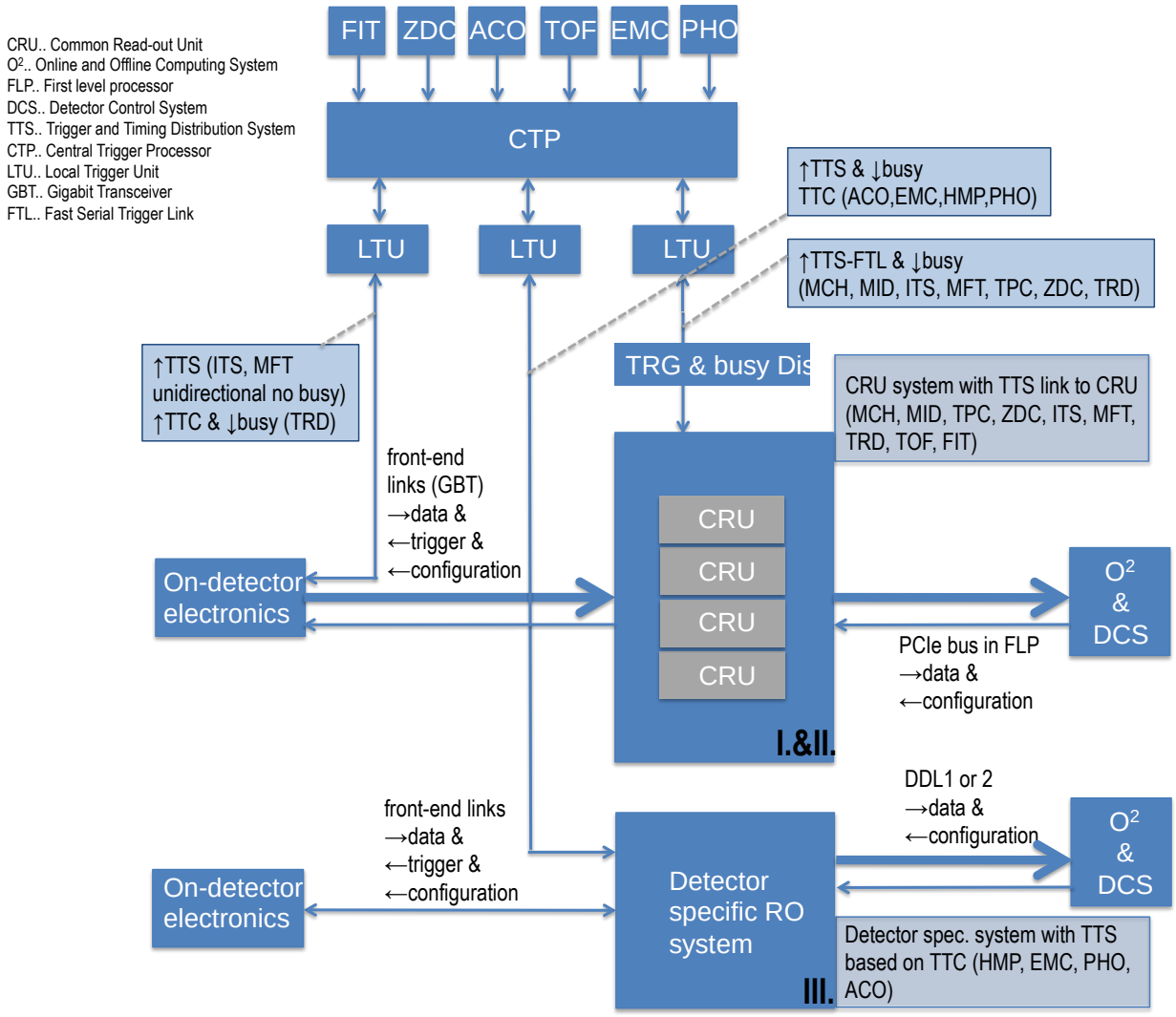
Level	Trigger Input to CTP [ns]	Trigger output at CTP [ns]	Trigger decision at detector * [ns]	contributing detectors
LM	425	525	775	FIT
L0	1200	1300	1500	ACO, EMC, PHO, TOF, ZDC
L1	#6100	#6200	#6400	EMC, ZDC



Usage of **Interaction Trigger** → **Triggered Readout**

Usage of **Heartbeat Trigger** → **Continuous Readout**

Trigger, Timing and Clock Distribution System (TTS)

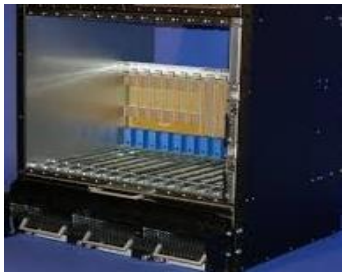


- ❑Configuration I: Transmission of non-critical trigger and data via CRU
- ❑Configuration II: Transmission of critical trigger directly to the detector and data via CRU
- ❑Configuration III: Back-end compatibility for Legacy detectors

Part III :

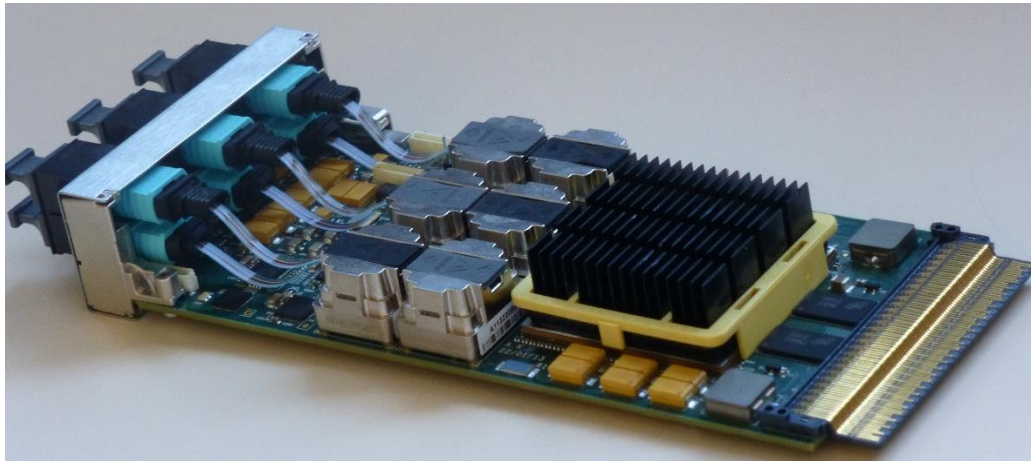
How we choose the CRU form factor ?

CRU Form Factor Evaluation

Features	Prototype version 1	Prototype version 2
DDL3	***10 Gigabit Ethernet*** 	***PCIe Gen 3***  
Trigger and Busy line Distribution	ATCA 	Processor 
Advantage	Modularity	Directly connected to the O^2
Disadvantage	Not Enough memory for data clustering not possible	Compatibility of the board depends on future PCs PCIe form factor

CRU CANDIDATE BOARDS EVALUATED

AMC 40



Developed By:
CPPM Marseille

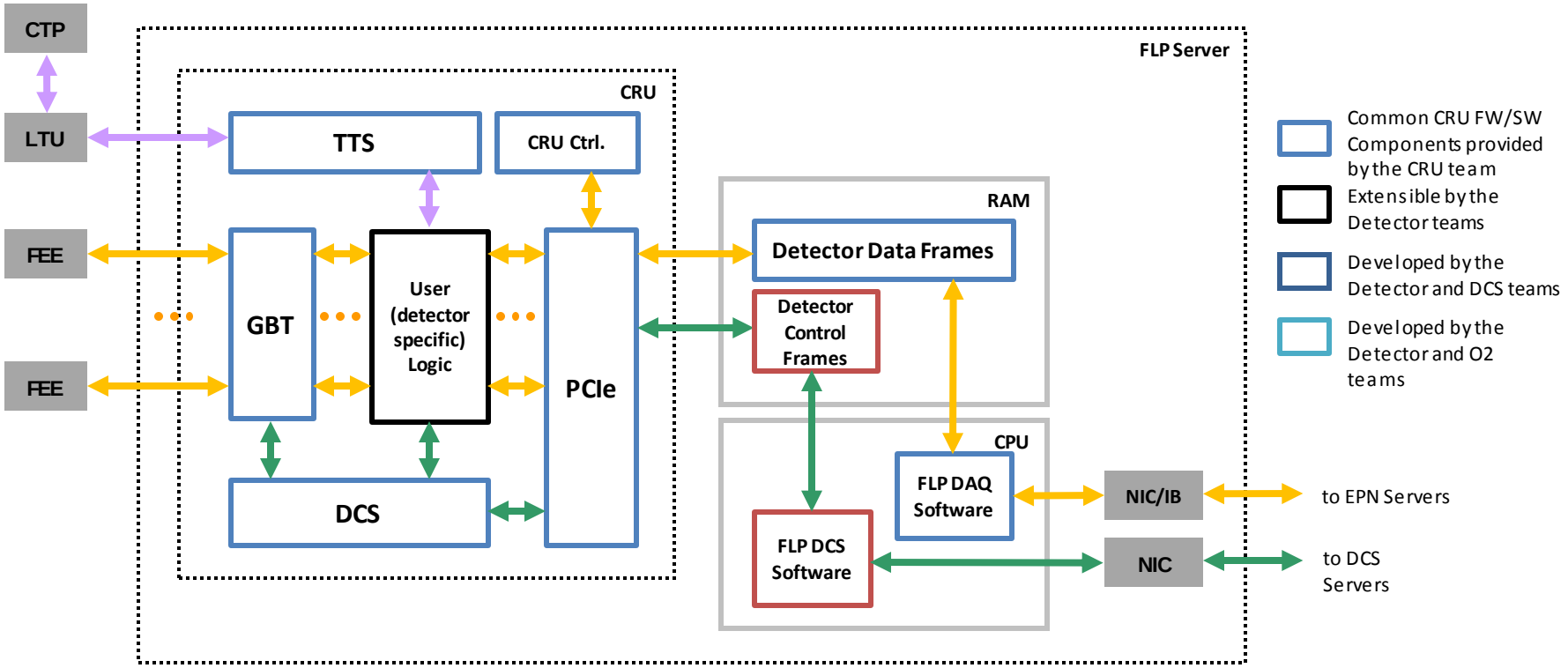
Stratix V vs Arria 10 FPGA

	Stratix V (High End FPGA)	Arria 10 (Mid End Latest FPGA)
	5SGXEA7N2F45C3	10AX115S4F45I3SGES
Chip Technology	28 nm	20 nm
Core voltage	0.85V	0.95V (For ES) else 0.9V
ALMs	234720	427200
Total I/Os	1064	960
GXB Channel PMA and PCS/ HSSI channels	48	72
PCIe Hard IP Blocks	4	4
Memory Bits	52428800	55562240
DSP Blocks	256	1518
27 x 27 Multiplier	256	1518
Fractional PLL	28	32
DLLs	4	-
I/O PLLs	-	16
Global Clocks	16	32
HPS CPU Core	-	0

Part IV :

Firmware Development Status

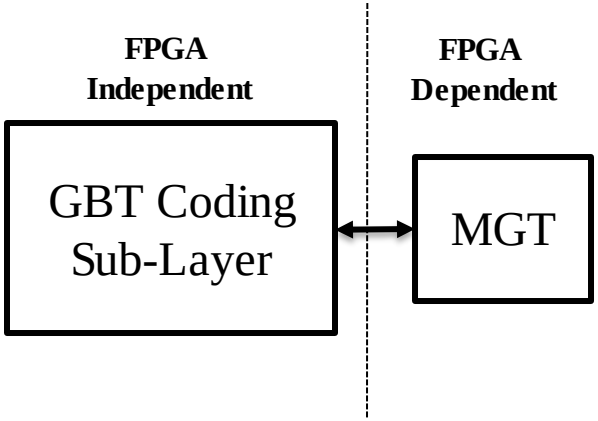
CRU internal block connections



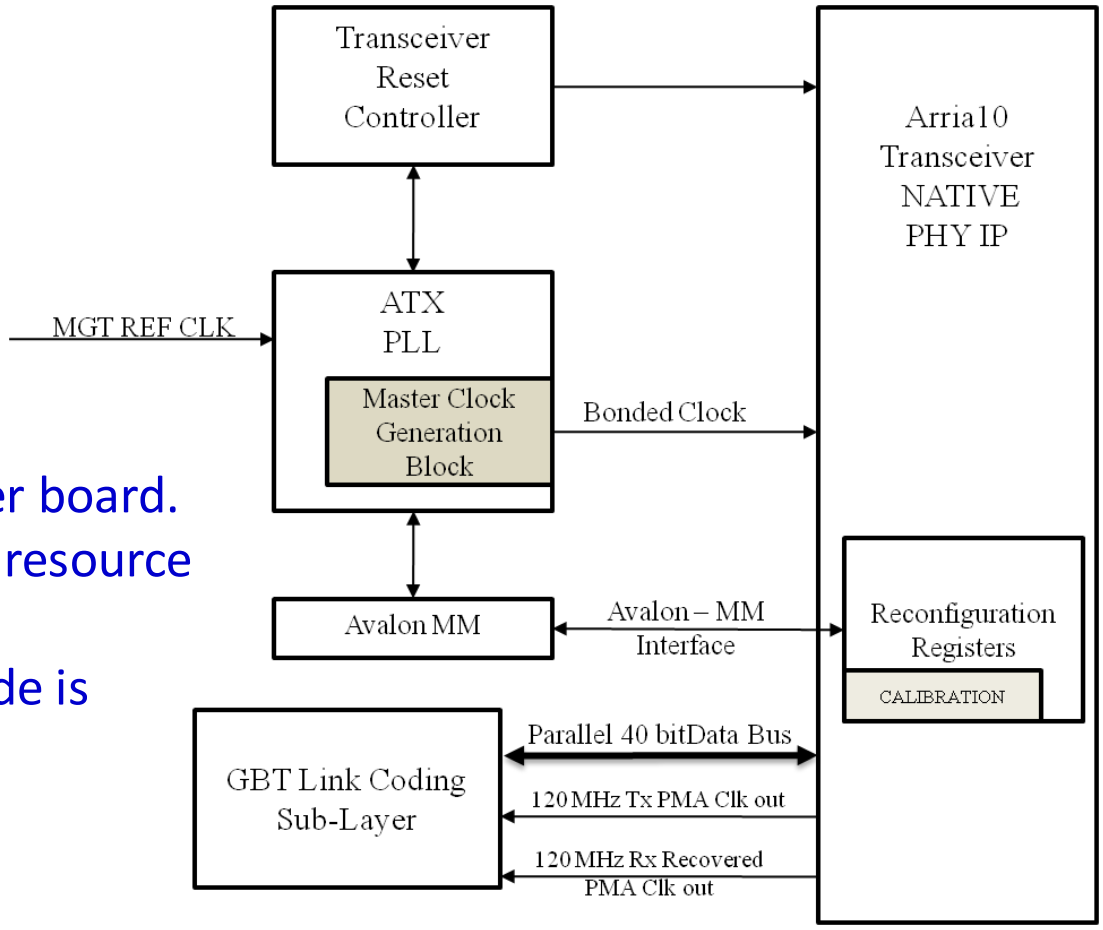
CTP – Central Trigger Processor
DCS – Detector Control System
EPN – Event Processing Node
FLP – First Level Processor
GBT – Giga-Bit Transceiver
LTU – Local Trigger Unit

GBT : Design Implementation on Arria 10

GBT Interface links are error resilient data communication protocol developed by CERN for high energy physics experiment.



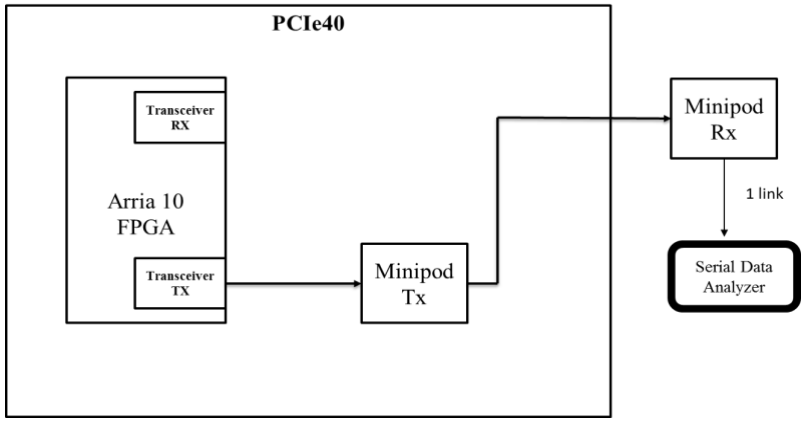
MGT connection for Arria 10



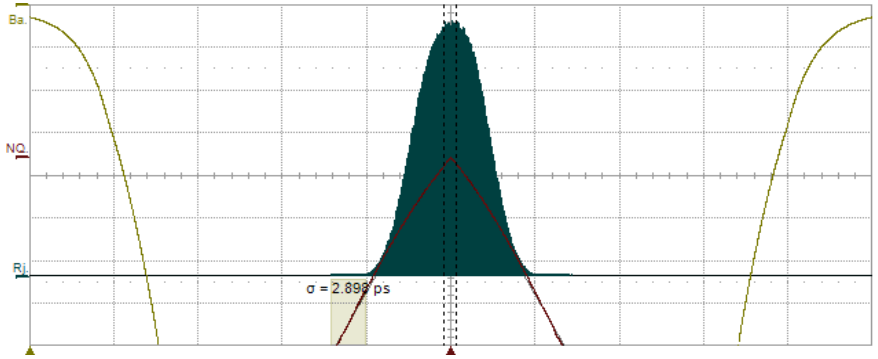
- CRU requires **12-32 GBT links** per board.
- **Channel bonding** -minimize the resource utilization
- 6 channel PMA (x6) bonding mode is supported in Arria 10

GBT : Operating at 4.8 Gbps Using 120 MHz External Jitter Cleaner

Test Setup



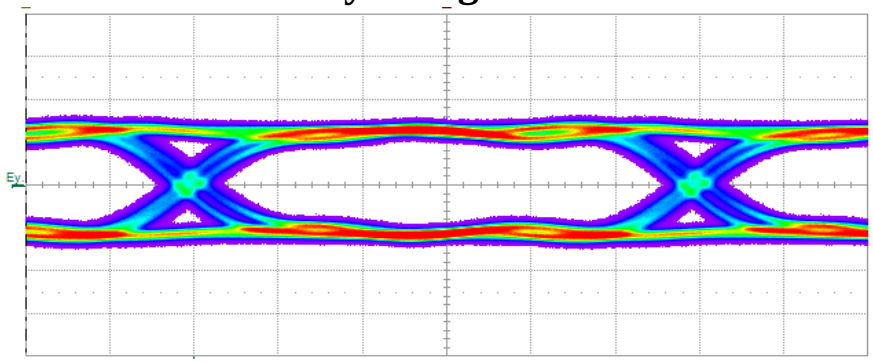
Jitter Measurement



The random component of the Jitter is specified using statistical terms.

Standard Deviation = 2.898 ps

Eye Diagram

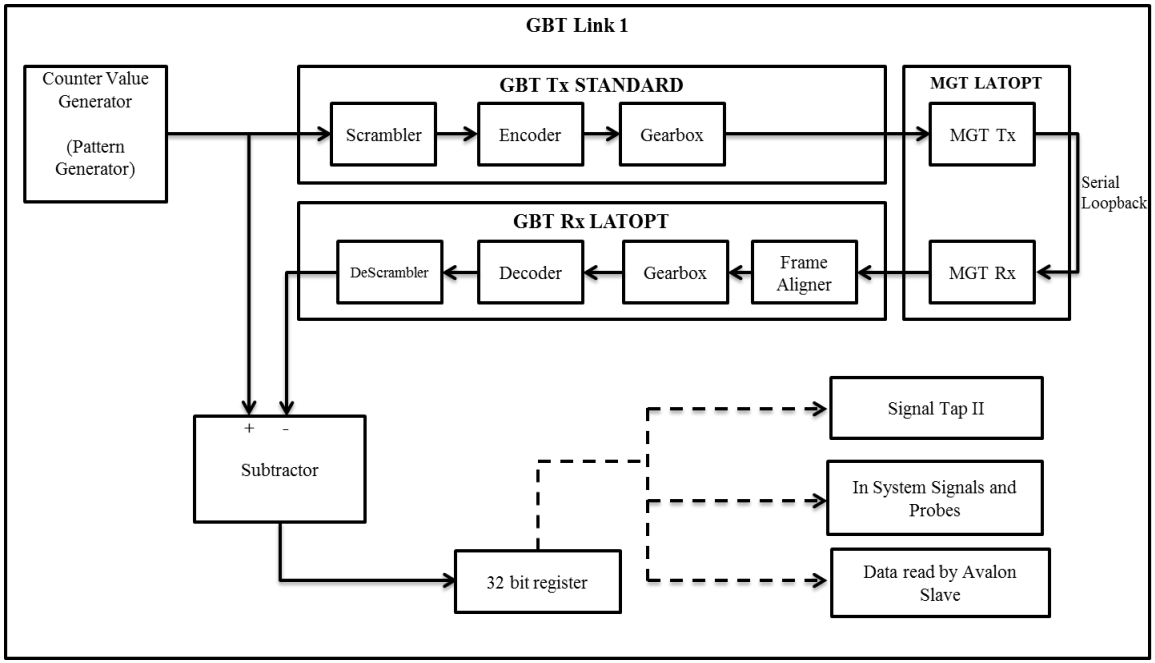


Eye Width = 176.8 ps

Eye Height = 373 mV

Bit/Rate = 4.7996

GBT : Latency Measurement



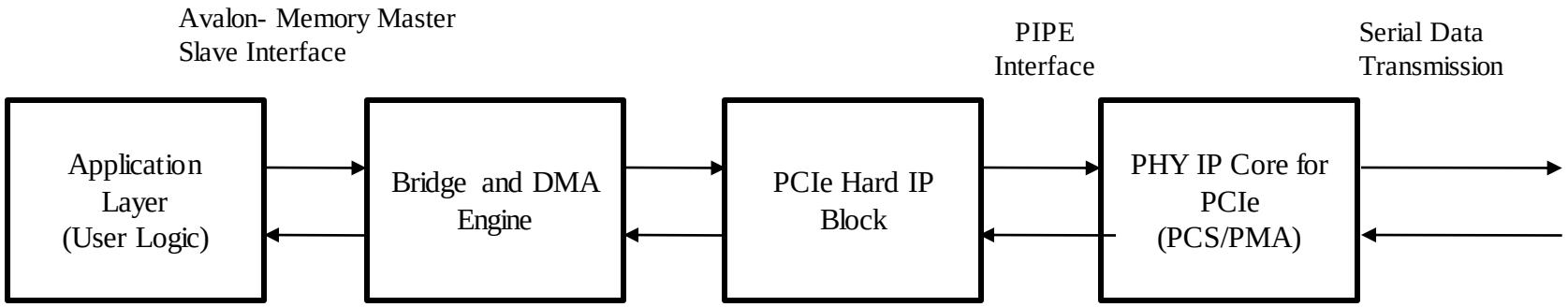
Latency measured between

Tx Data Frame – GBT Tx – MGT Tx (Serialization) – Optical Loopback --MGT Rx (De-serialization) – GBT Rx – Rx Data Frame

	GBT			
	Tx Standard	**Tx Standard**	**Tx Latency Optimized**	**Tx Latency Optimized**
	Rx Standard	**Rx Latency Optimized**	**Rx Standard**	**Rx Latency Optimized**
Latency Measured	***450 ns***	***350 ns***	***200 ns***	***150 ns***

 # Transmission Side Phase latency is very stringent as it is used for timing information transmission

DDL3: PCIe Link Testing and DMA Performance Measurement



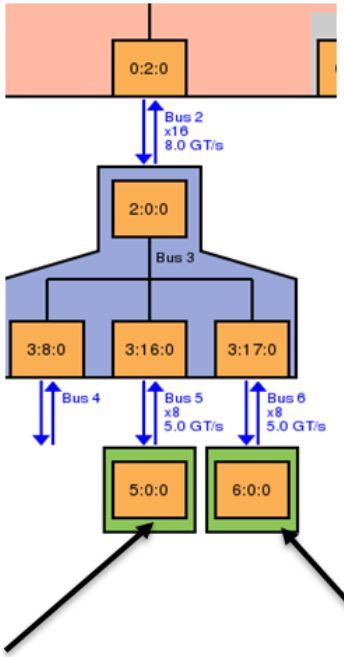
PCIe Gen2 x8 Performance Measurement:

Signalling Rate = 5 Gbps per lane x 8 = 40 Gbps

Useful Data Throughput = 32 Gbps

As, Gen2 use 8B/10B encoding which introduces a 20% overhead

PLX 8747 Switch →



PCIe gen2 x8 from Arria 10 FPGA

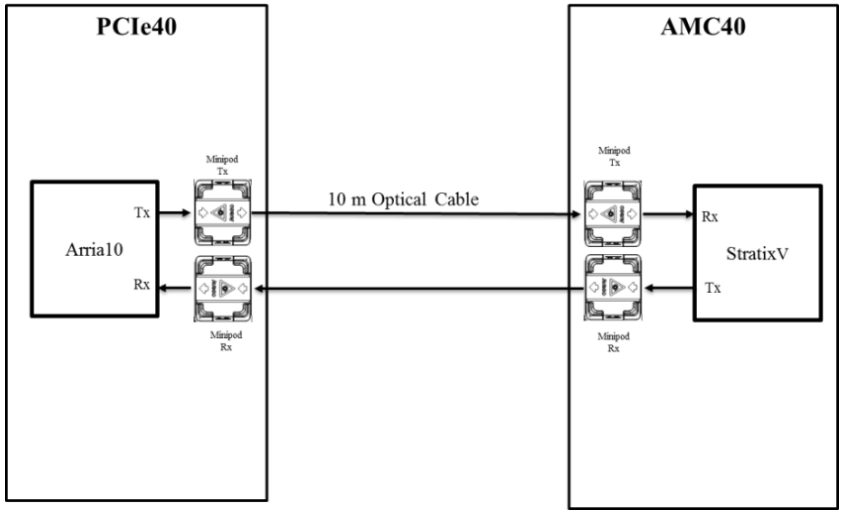
PCIe gen2 x8 from Arria 10 FPGA

 # In Arria 10 Engineering Sample1 we have PCIe Gen2 , but Engineering Sample2 and production chips run in GEN3

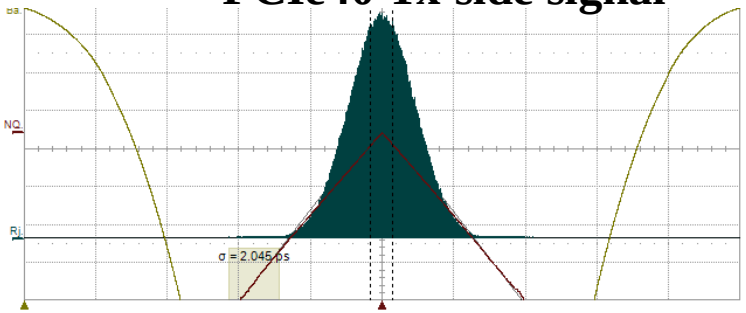
Using PLX8747 PCIe Switch for multiplexing 2 banks of x8 lanes

Avago MiniPOD™ performance study

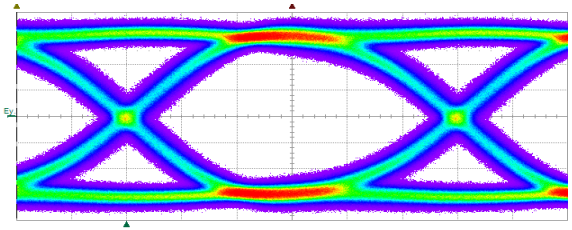
12 channel Transceiver tool kit design for board to board communication between Stratix V and Arria 10 at 10.312Gbps



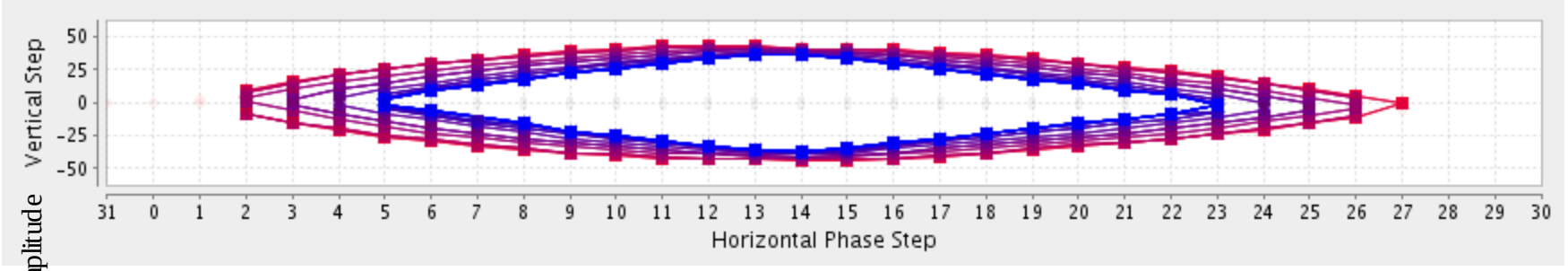
PCIe40 Tx side signal



Random Jitter = 2.048 ps



Eye Width = 79.4 ps
Eye Height = 373 mV



PCIe40 Tx signal as received in AMC40 using Transceiver Tool kit (TTK)

A Test Configuration for Firmware Resource Estimation

For power and resource estimation of CRU a low level interface is made :

- 48 bidirectional GBT link + x16 PCIe Gen 3 + SFP+ (Transceiver Toolkit design)

- It is composed of total 65 High speed links each attached with its dedicated ***pattern generator*** and ***pattern checker***

- The entire design is integrated in Altera QSYS Integration tool

- In this design it is assumed to operate PCIe40 in extended configuration

- No user or glue logic is taken into account in this ***preliminary resource estimation***

Aggregated Links	Logic Utilization	HSSI SERDES Utilization
48 GBT link + x16 PCIe Gen 3 + SFP+ (Transceiver Toolkit design)	34,614 / 427,200 (8%)	65/72 (90%)
36 GBT link + x16 PCIe Gen 3 + SFP+ (Transceiver Toolkit design)	32,247 / 427,200 (7.5%)	53/72 (74%)
24 GBT link + x16 PCIe Gen 3 + SFP+ (Transceiver Toolkit design)	29,771 / 427,200 (7%)	41/72 (57%)

Summary and Outlook

- The CRU in ALICE is designed to cope up with increased beam energy and luminosity of LHC for RUN3 and beyond

- CRU acts as the interface between:
  - the on-detector electronics
  - Online and Offline computing system (O2)
  - the Trigger Timing System (TTS)

- Already done:
  - Survey of prototype boards
  - Implementation of GBT for Arria 10
  - PCIe functional testing and how to use the IP cores
  - A resource estimation before full firmware implementation

- Plans for near future:
  - Finishing the Interface specification (External and Internal)
  - Integration of CRU firmware/hardware building blocks
  - Built and test pilot system
  - Detector specific firmware development (Done by the sub-detector groups)

Technical Team Members

Variable Energy Cyclotron Centre (VECC)

Jubin MITRA
Shuaib Ahmad KHAN
Tapan Kumar NAYAK

University of Jammu

Anik GUPTA

Bose Institute

Sanjoy MUKHERJEE

University Of Calcutta

Rourab PAUL
Amlan CHAKRABARTI

Wigner Research Centre for Physics

Erno DAVID
Tivadar KISS

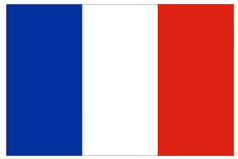
CERN

Filippo COSTA

Leaders of Interfacing Groups:

- Peter CHOCHULA (DCS)
- Marian KRIVDA (TTS)
- Pierre Vande VYVRE (O2)
- Alex KLUGE (Electronics Coordinator)

With Active help and Support from LHCb group:



CPPM, Marseille

Jean-Pierre CACHEMICHE
and others

Questions

Backup

ALICE Upgrade Strategy

❖ Goal:

- High precision measurements of rare probes at low transverse momentum, which cannot be selected with a trigger
- Target to record Pb-Pb collisions at higher luminosity to gain a factor 100 in statistics over present Run1+Run2
- Readout all Pb-Pb interactions at a maximum rate of 50kHz (i.e. $L = 6 \times 10^{27}\ cm^{-2} s^{-1}$) which is roughly 6 times of present rate
- Perform online data reduction based on reconstruction of clusters and tracks
- **A separate data processing unit is needed for detector data** multiplexing, processing and formatting before online data reduction.

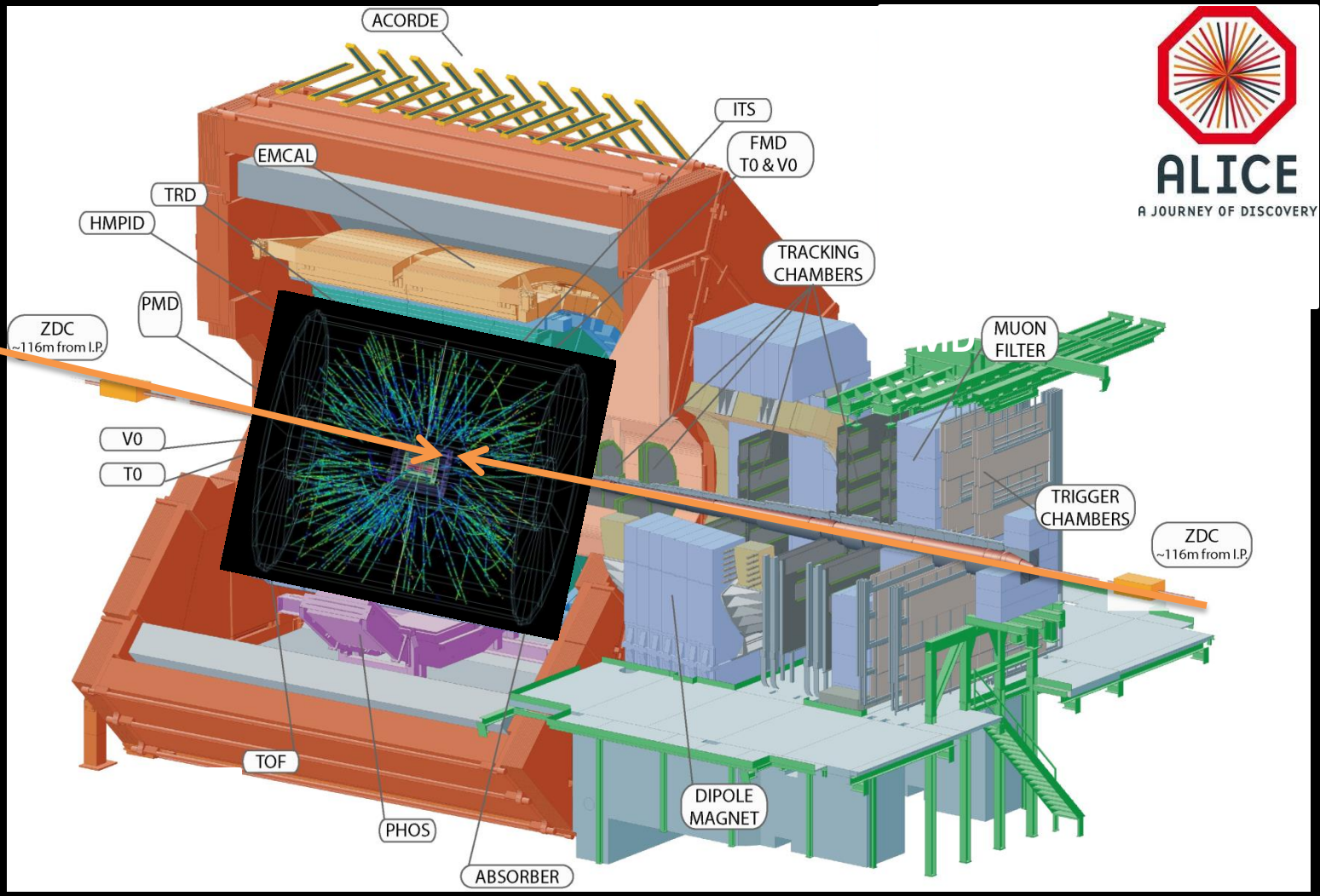
Implementation of a novel readout architecture
Common Read-out Unit (CRU).

UPGRADE EFFORT

Detector Data Rate And Channels

Det	# channels	Run1&2 RO rate [kHz]	upgrade RO rate [kHz]	FE ASIC
TPC	5×10^5	3.5	50	17000 SAMPA
MCH	10^6	1	100	33000 SAMPA
ITS	25×10^9	0.5	100	25000 ASICs
MFT	500×10^6		100	896 ASICs
MID	21×10^3	1	100	4000 FEERIC
ZDC	22	8	100	
TOF	1.6×10^5	40	100	
FIT	$160 + 64$	80	100	
ACO	120	100	100	
TRD	1.2×10^6	1	39	
EMC	18×10^3	3.7	42	
PHS	17×10^3	3.7	42	
HMP	1.6×10^5	2.5	2.5	

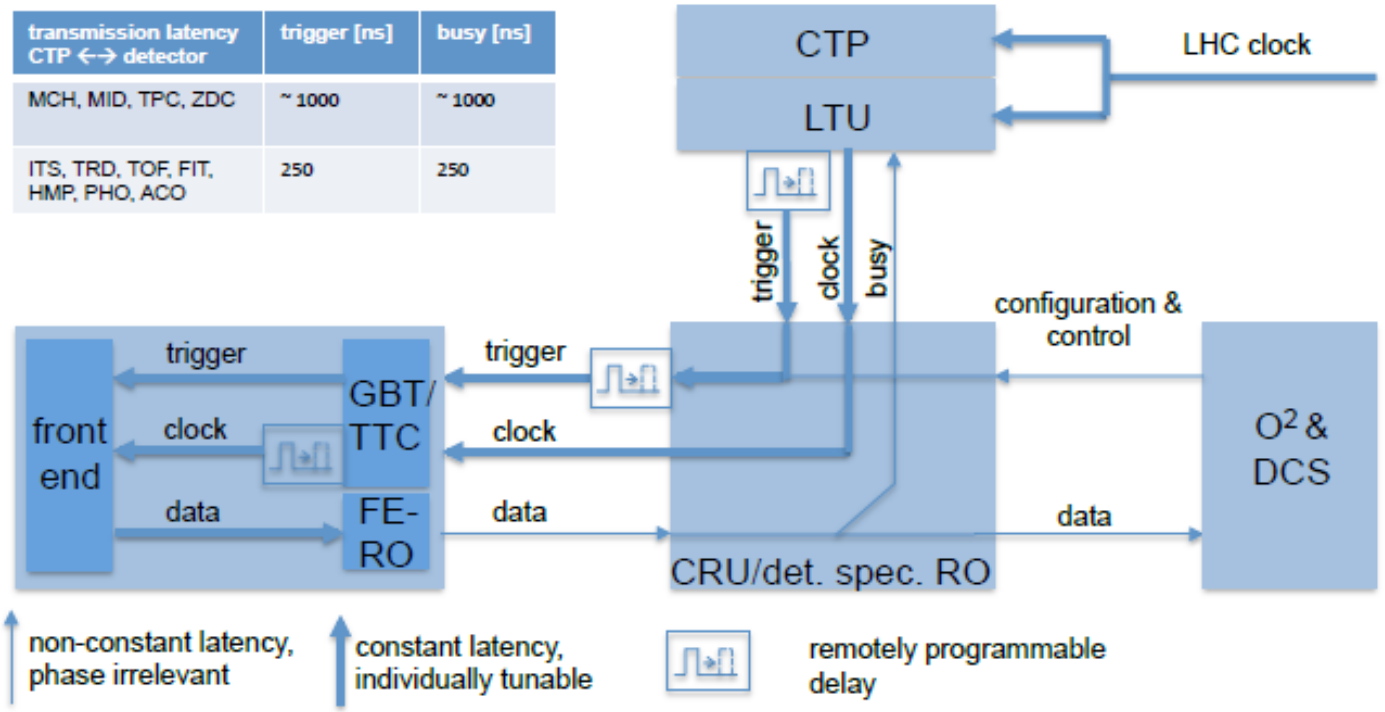
ALICE is giving answer to some of our basic questions using the extraordinary tools provided by the LHC



Collisions at the CERN Large Hadron Collider (LHC): proton-proton, proton-lead, lead-lead

Read-out control signal flow through CRU

transmission latency CTP ↔ detector	trigger [ns]	busy [ns]
MCH, MID, TPC, ZDC	~1000	~1000
ITS, TRD, TOF, FIT, HMP, PHO, ACO	250	250



- **CLOCK:** Common Reference Timing Signal + the LHC clock
- **TRIGGER:** PHYSICS TRIGGER + HEARTBEAT TRIGGER + BUNCH CROSSING ID + ORBIT COUNTER VALUE
- **BUSY:** When **Trigger rate > Detector Readout Capabilities.** One signal covers for entire sub-detector.
- **DATA:** Detector specific data payload with header and timestamp