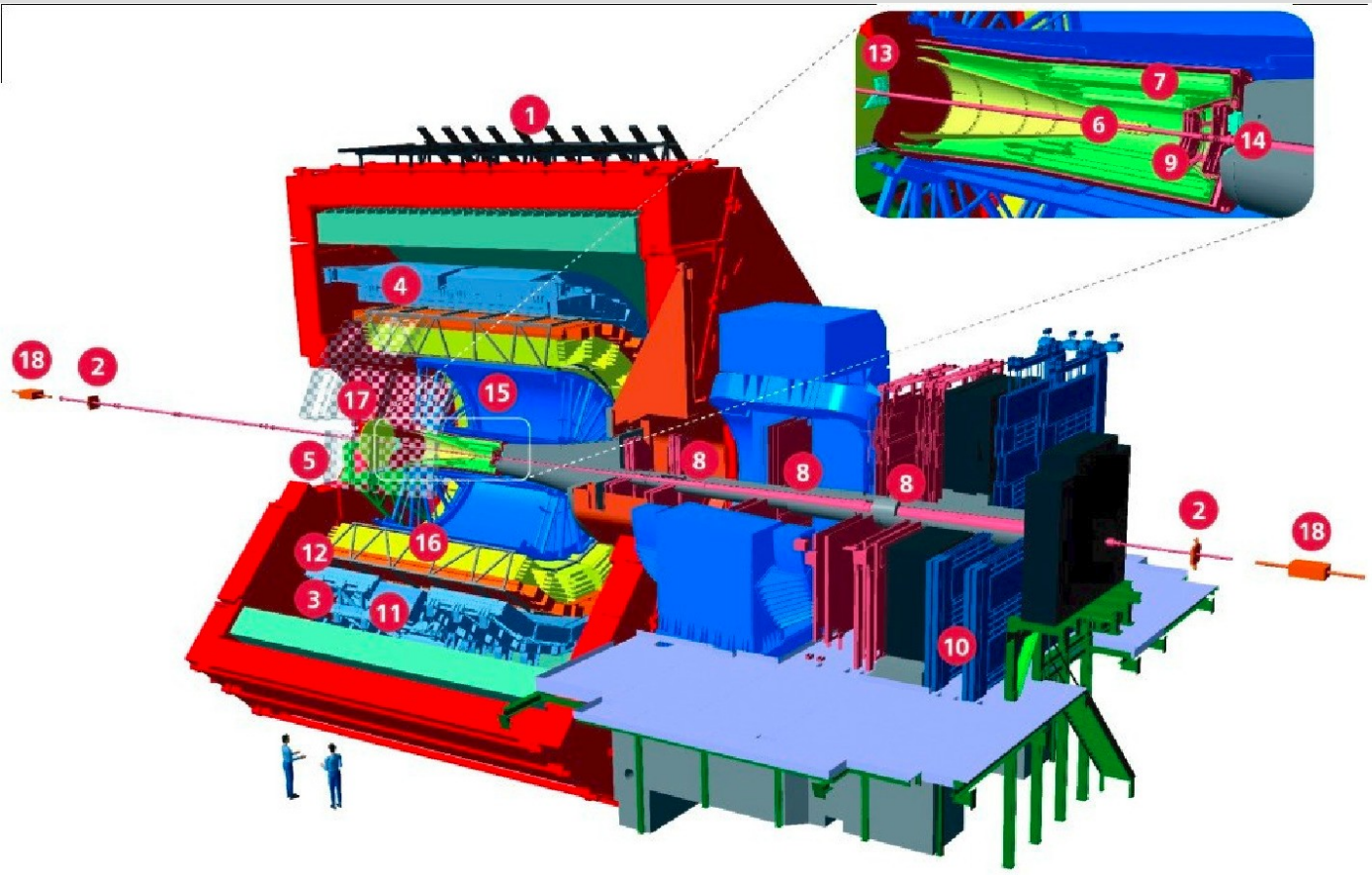
Versatile firmware for the Common Readout Unit (CRU) of the LHC ALICE experiment

Olivier Bourrion
on behalf of the CRU team

Outline

- Introduction
- Requirements
- Firmware description
  - Overview
  - GBT interface
  - TTC interface
  - Data path
  - Readout control
  - PCIe
- Summary

A Common Readout Unit for upgraded detectors



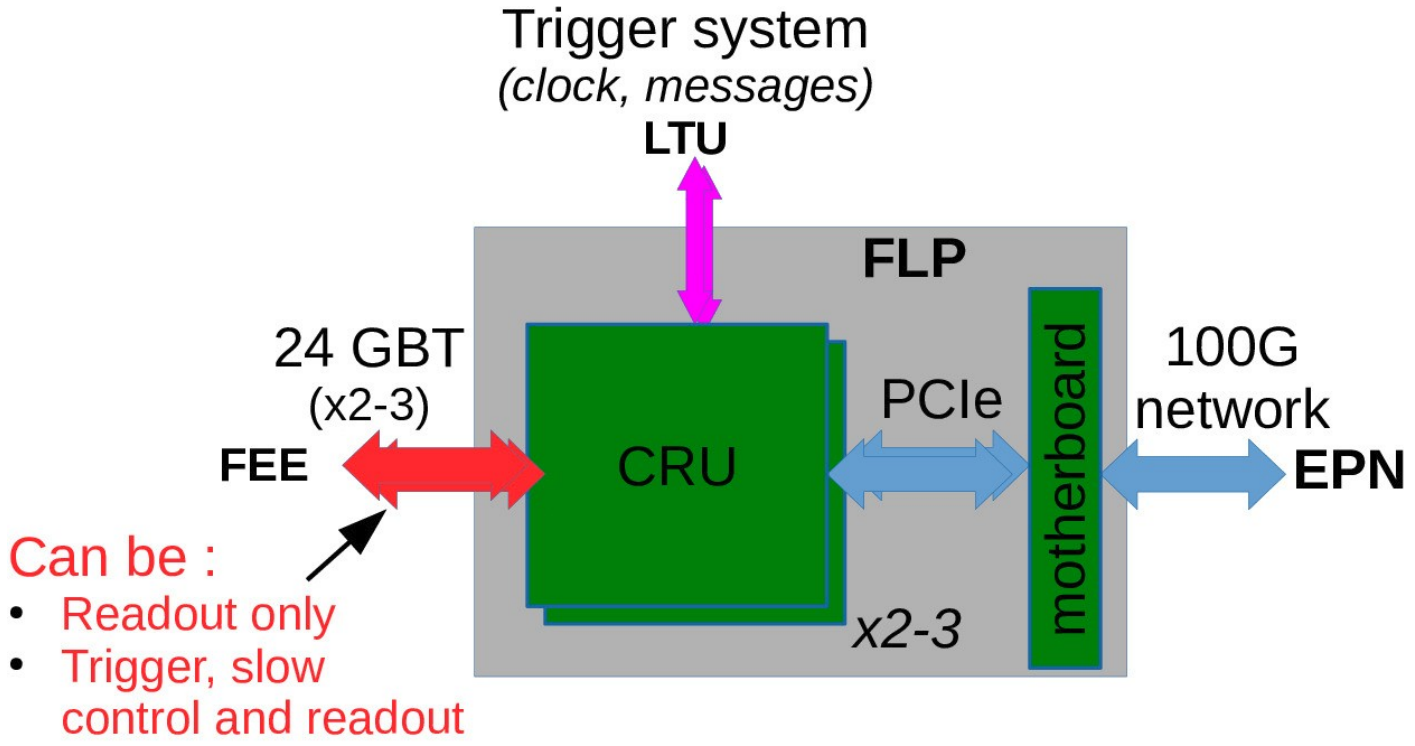
Upgraded detectors

6) ITS-IB | Inner Tracking System- inner barrel
7) ITS-OB | Inner Tracking System- Outer barrel
8) MCH | Muon Tracking chamber
9) MFT | Muon Forward Tracker
10) MID | Muon Identifier
11) CPV | Charged Particule Veto
12) TOF | Time Of Flight
2,13,14,17) FIT | Fast Interaction Trigger
15) TPC | Time Of Flight
16) TRD | Transition Radiation Detector
18) ZDC | Zero Degree Calorimeter

- For run 3 and 4, ALICE moves to a continuous readout
  - ALICE will produce 3.5 TB/s of data
- The upgraded detectors (10) will use the CRU to comply with this scheme
- A common firmware was developed for the CRU to share efforts

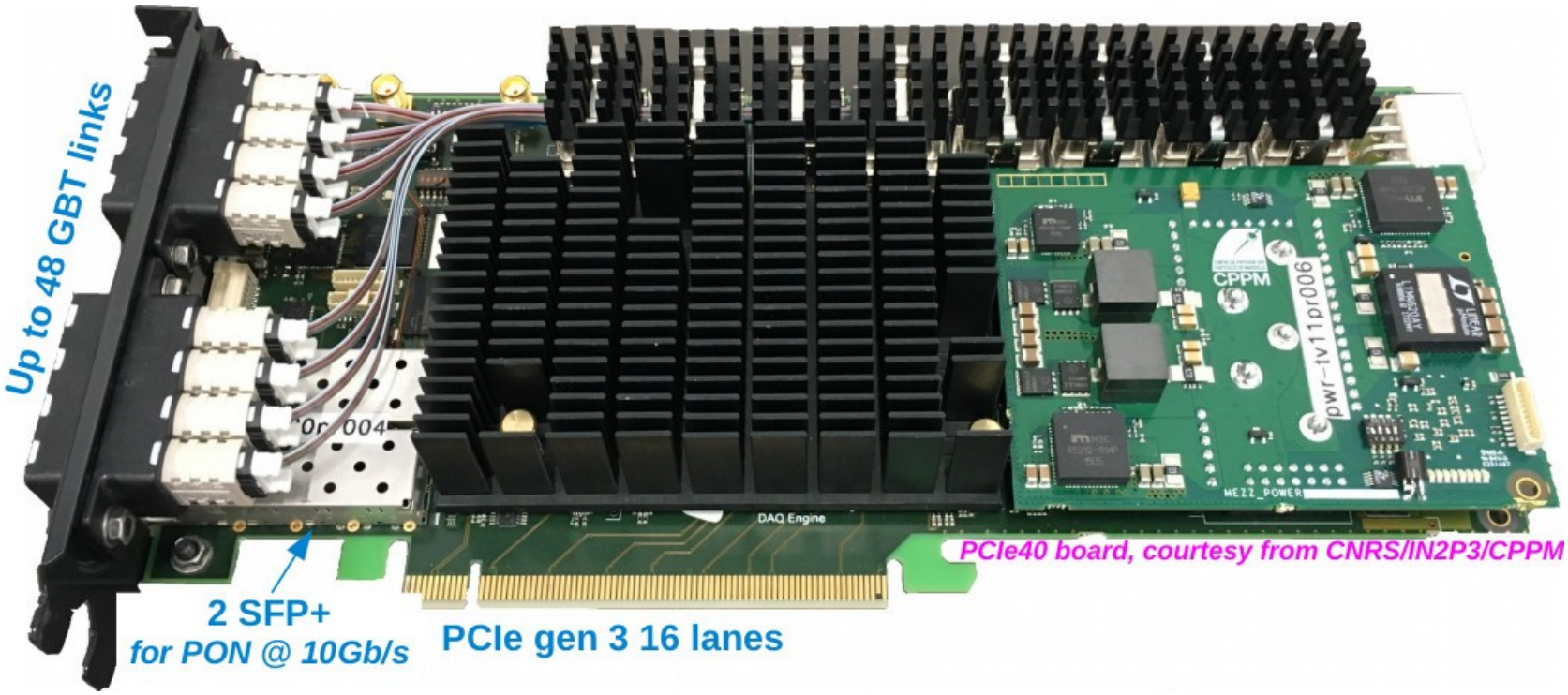
CRU in the system

- The CRU is the interface between :
  - the detector front-end electronics
  - The O² facility
  - The Detector Control System
  - The Trigger System
- O² facility
  - 2-3 CRUs are hosted by First Level Processor (FLP)
  - FLP communicates with Event Processing Nodes (EPN) through the network



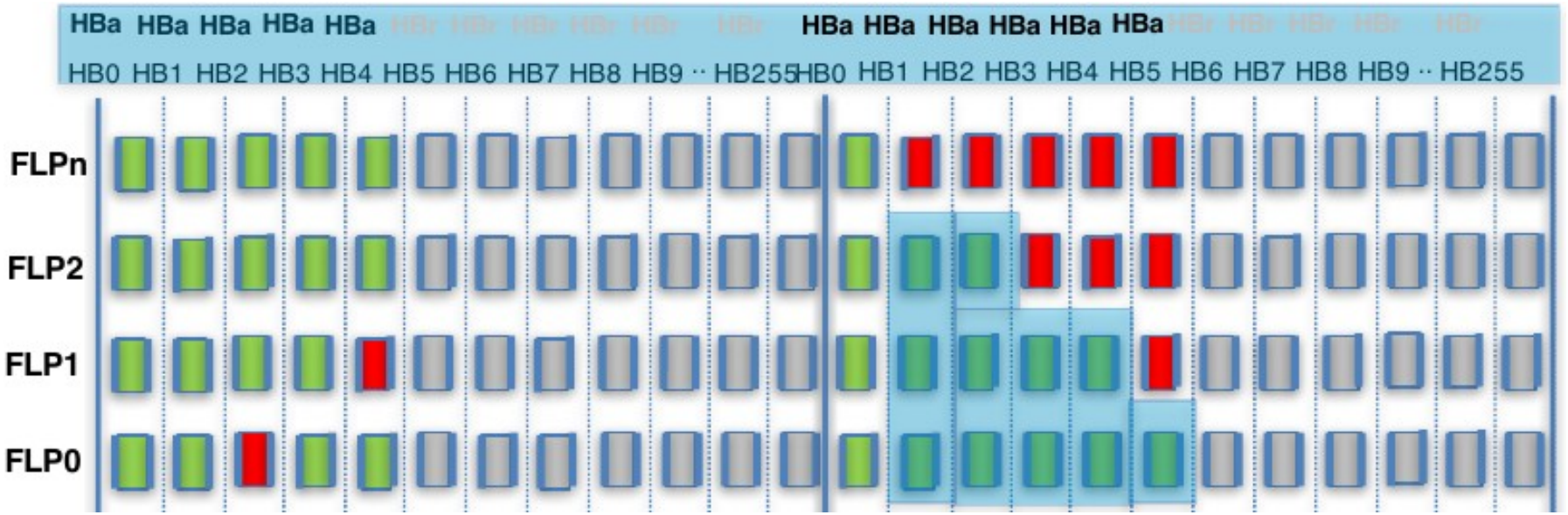
CRU Hardware

- Hardware was designed for and by LHCb (named PCIe40)



- One Small Factor Pluggable (SFP) cage used as interface with the trigger system
- Up to 24 GBT links (or 36 for TRD) used
- PCIe interface with FLP, and thus DAQ and DCS

Continuous readout in ALICE



- Continuous stream is sliced in Time Frames of 22 ms (size of data sample requested by reconstruction)
- Time Frames are divided in 256 Heart Beat Frame (HBF) of 1 orbit duration (89.4 µs).
- CRU collect data continuously **and check successful** HBF reception in each FLP
- For each HBF an (not) ACKnowledge is delivered to the Central Trigger Processor
- Upon missing ACK
  - => CTP requests all CRU to drop the HBF until the end of Time Frame to allow recovery.

Detector requirements

- GBT link type ("**gbt-mode**" or "**wide**")
  - Downstream, either:
    - Trigger messages (TTS) or only clocking
    - Slow control
  - Upstream (FEE → CRU)
    - Readout (streaming mode or packet mode)
    - Slow control acknowledge
  - A maximum of 24 GBT links
- Interface with trigger system
  - Various trigger bit
  - Busy through CRU (TRD only)
- PCIe readout (gen3 x16)

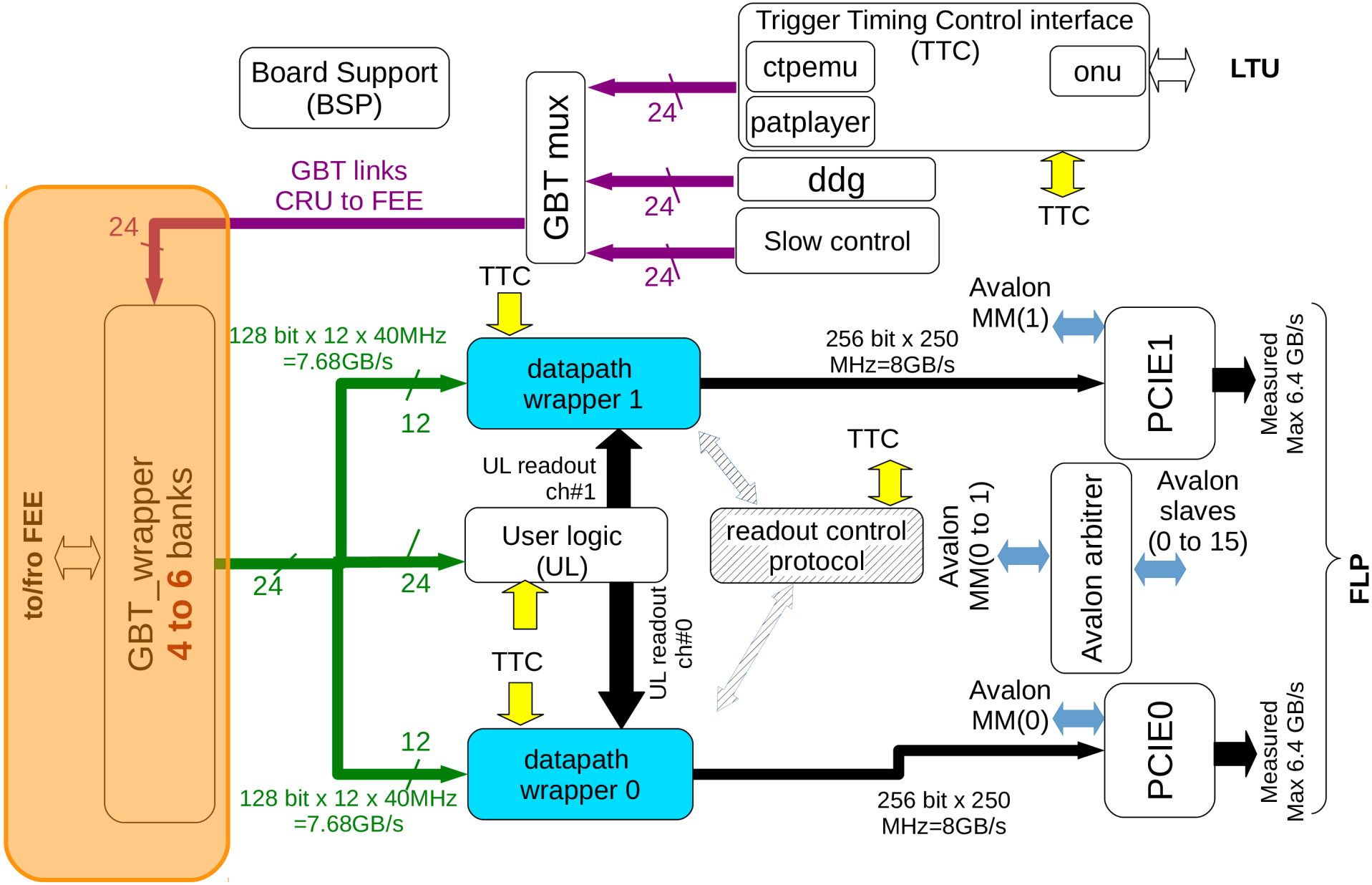
Detector	Number of CRU	TX/RX per CRU	User Logic	GBT with FEC (''gbt-mode'')	TTS through CRU	DCS through CRU
CPV	1	24/24	N	Y	Y	Y (SCA)
FIT	2	24/24	N	Y	Y	
ITS	26	12/24	**TBD**	Y	clock only	Y (single word transfert)
MCH	32	24/24	**Y**	Y	Y	Y (SCA)
MFT	11	12/24	N	Y	clock only	Y (single word transfert)
MID	2	24/24	**TBD**	Y	Y	Y (SCA)
TOF	4	24/24	**TBD**	Y	Y	Y (SCA)
TPC	360	24/24 (50%) 0/24 (50%)	**Y**	**N**	Y	Y (SCA)
TRD	37	**0/36**	**Y**	**not GBT**		
ZDC	1	24/24	N	Y	Y	

=> Most of the detectors need similar features (**in blue**)

Common Firmware overview

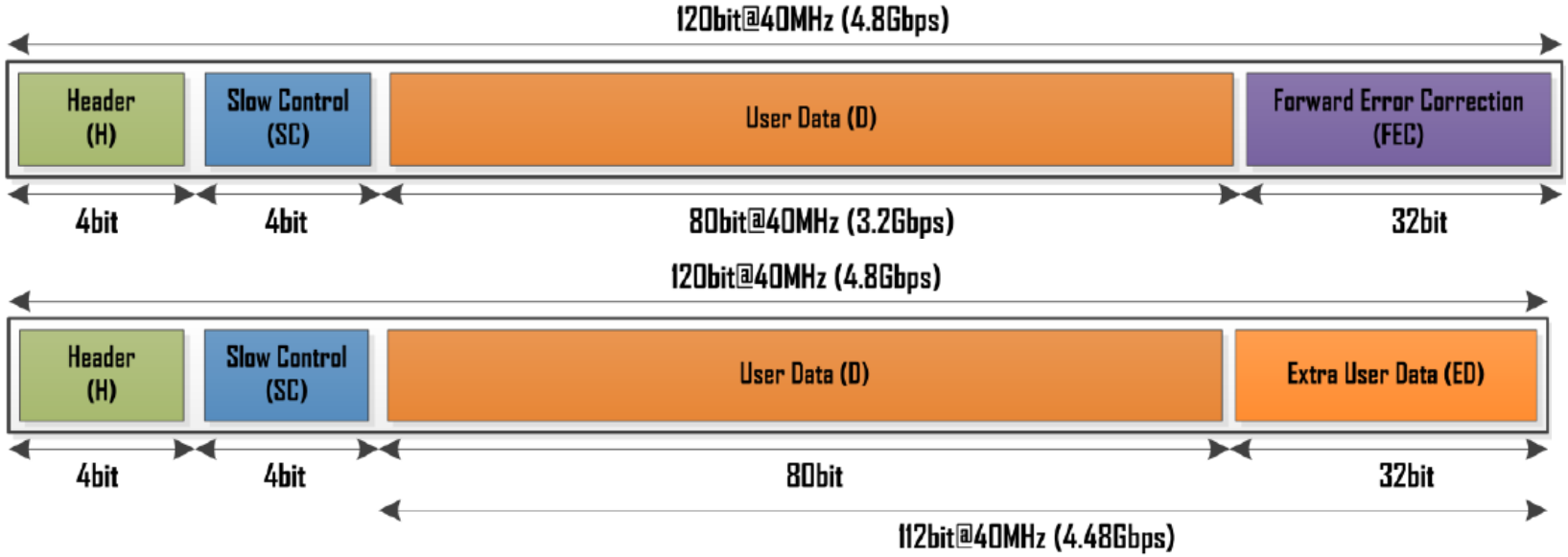
- The firmware was developed to fit Alice needs :
  - Share common features : PCIe, trigger and timing, GBT
  - Readout in **raw mode** (no processing in CRU) for all detectors
  - Clock and trigger distribution
  - FEE configuration
- Detector specific features allow for specific **User Logic** for some detectors
  - Online data processing (baseline correction, zero suppression)
    - User logic is inserted in common firmware => specific compilation
  - At any moment, can switch from **raw mode** to **User Logic**
- Self-testing capabilities
  - Data loop-back and automatic checks

Firmware description: GBT



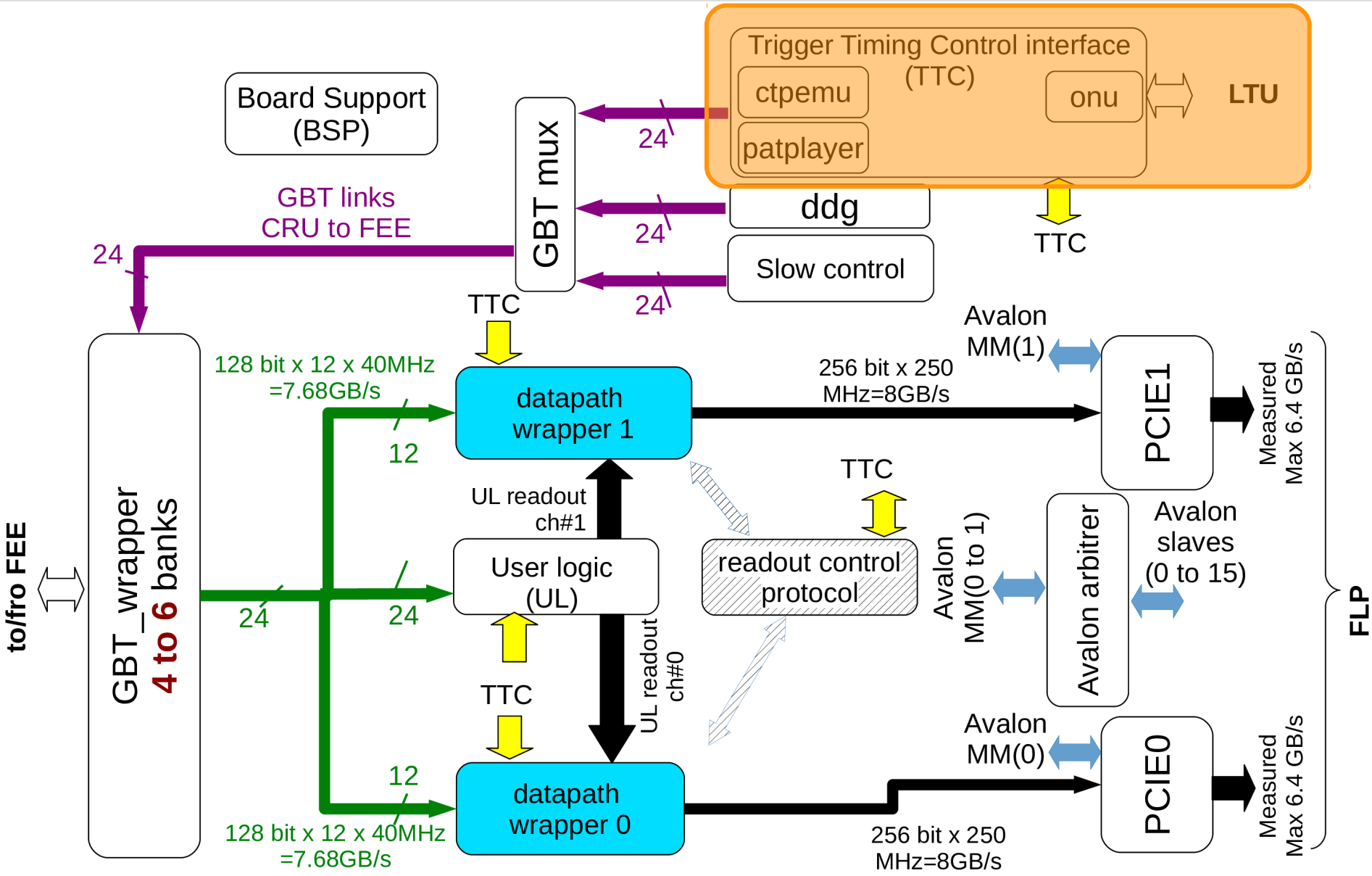
Giga Bit Transceiver interface (GBT)

- GBT is part of the “radiation hard optical link project” → bi-directional optical link
- Allows “*GBT-frame*” or “*wide-bus*”



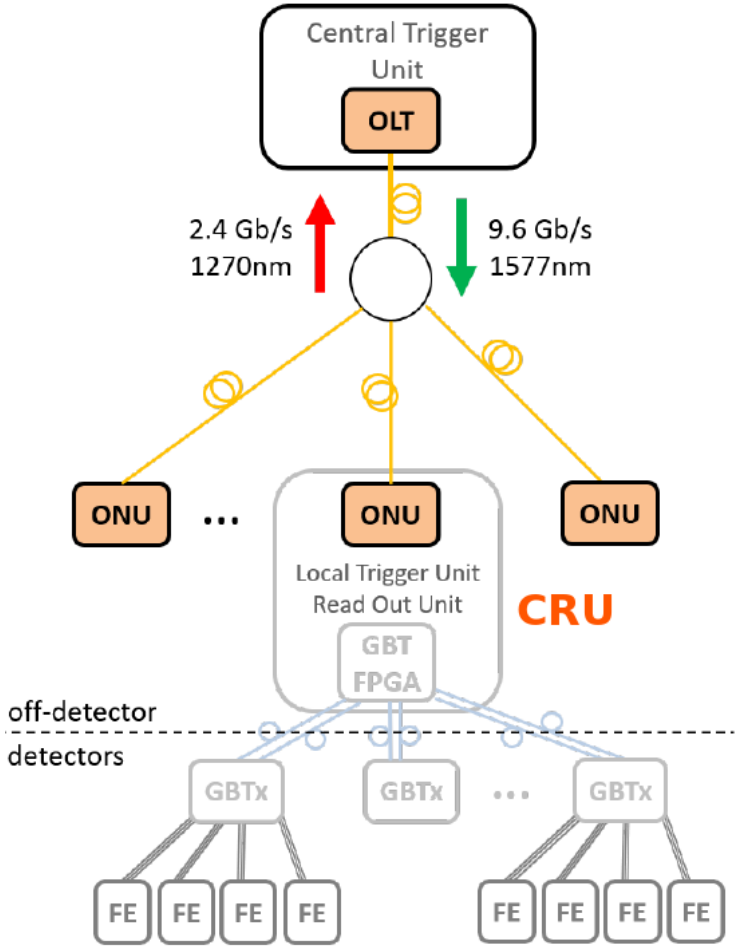
- CRU uses up to 24 GBT links, down-link is latency optimized
- GBT_wrapper contains a modified version of the GBT-FPGA provided by CERN electronics group
  - **Dynamic switching** between “GBT-frame” and “wide-bus” => covers more use cases
- Test modes
  - External loop-back (with fibers, readout electronics) => allow validation of front-end ↔ CRU communication
  - The internal loop-back (via transceiver) => Allows checking during operation

Firmware description: TTC



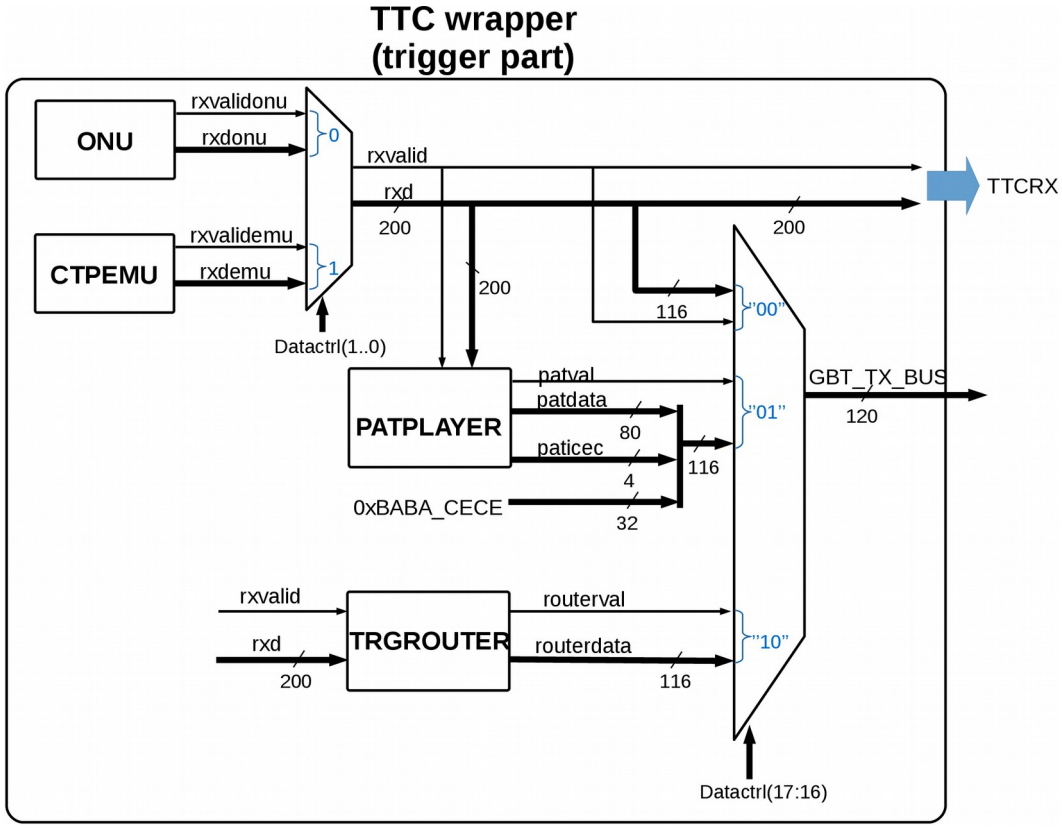
Trigger Timing Control interface (1/2)

- Uses the ONU provided by CERN electronics group
  - CRU system clock is recovered from Passive Optical Network
  - Recovered clock is used and forwarded to FEE by GBT
  - Reception of trigger messages (200 bit @ 40MHz)
    - Trigger bit, bunch crossing number, orbit number, ...
  - Emission of Heart Beat ACKnowledge (56 bit)
    - Time domain multiplexing up-link, slot available every number of ONU x 125 ns
    - 4 ONU → slot every 600 ns !

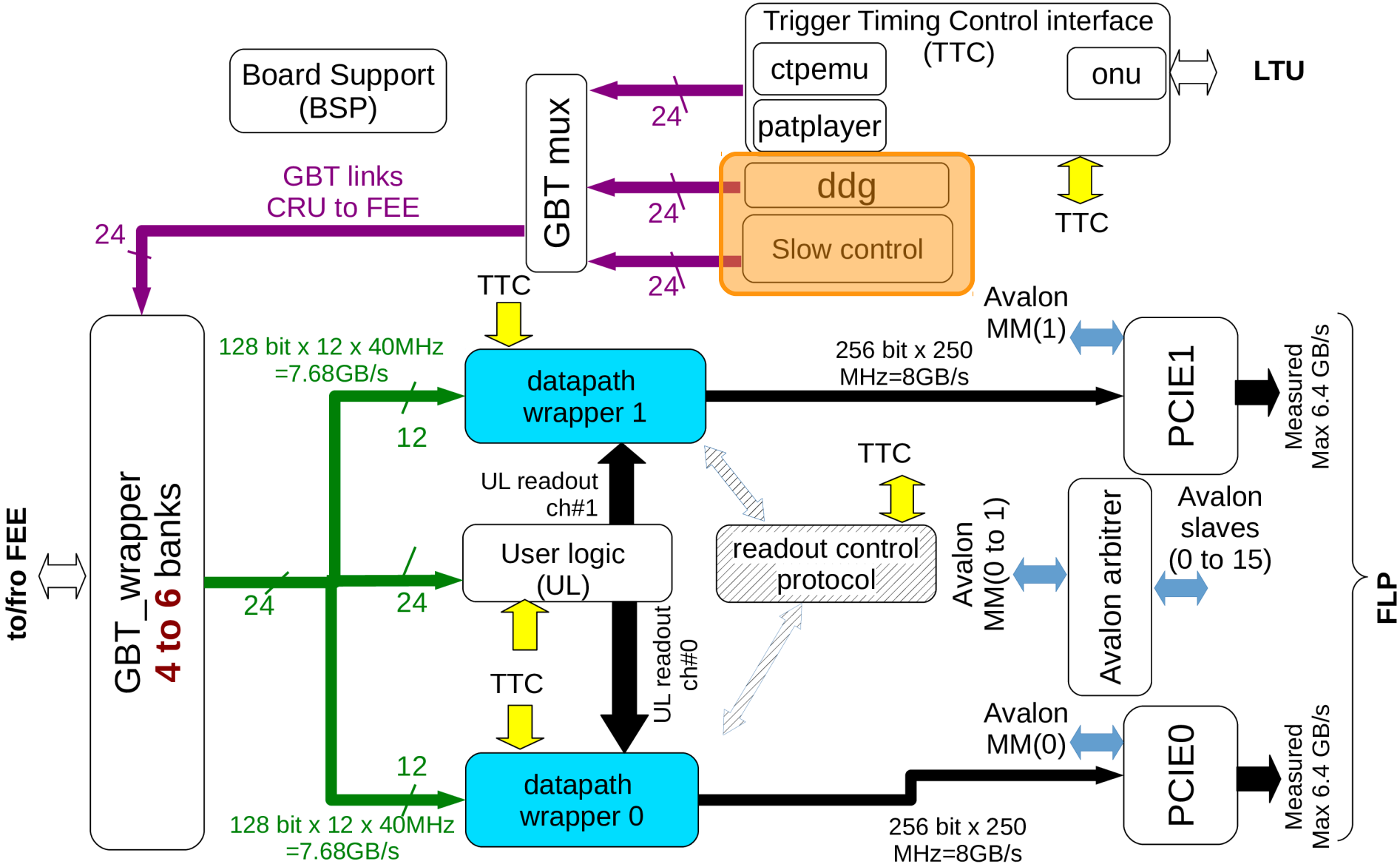


Trigger Timing Control interface (2/2)

- A trigger emulator is implemented :
  - To mimic trigger messages
  - To simulate readout flow control
  - Can be used by detectors for functional testing with local oscillator
- Pattern player
  - Generates a programmable sequence to transmit to FEE, fired by a trigger bit
  - An example: SYNC – RST sequence for SAMPA chip
    - For Time Projection Chamber (TPC), Muon Forward Tracker (MCH), ...
- Trigger router
  - Used to replicate, or reroute a trigger bit to several positions (replicate bit on various e-links)
  - For Muon IDentifier (MID)



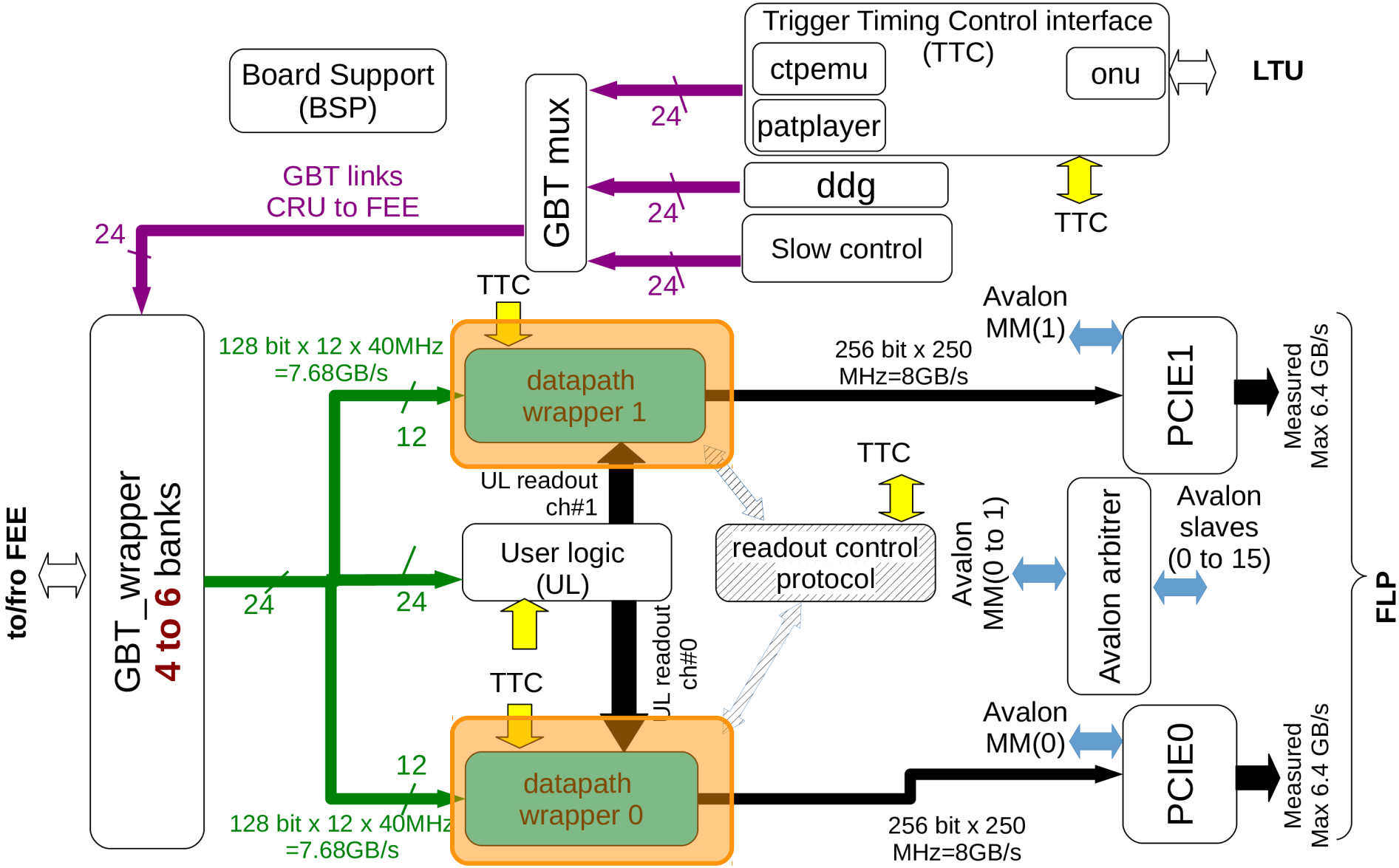
Firmware description: other down-link features



Other down-link features

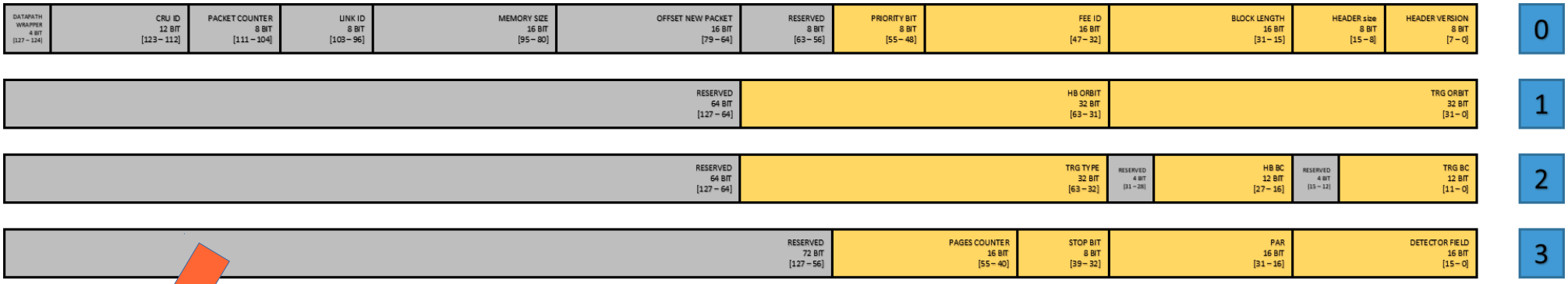
- DDG (Dedicated Data Generator)
  - Produce streaming or packet data
  - Packet data with fixed/random length, fixed/random idle
- Slow control
  - SWT (Single Word Transaction) is a slow control protocol intermingled with data flow (for ITS)
  - GBT-SCA (Slow Control Adapter) produces slow control transactions for GBTx
  - Both are controlled by PCIe

Firmware description: DATAPATH WRAPPER



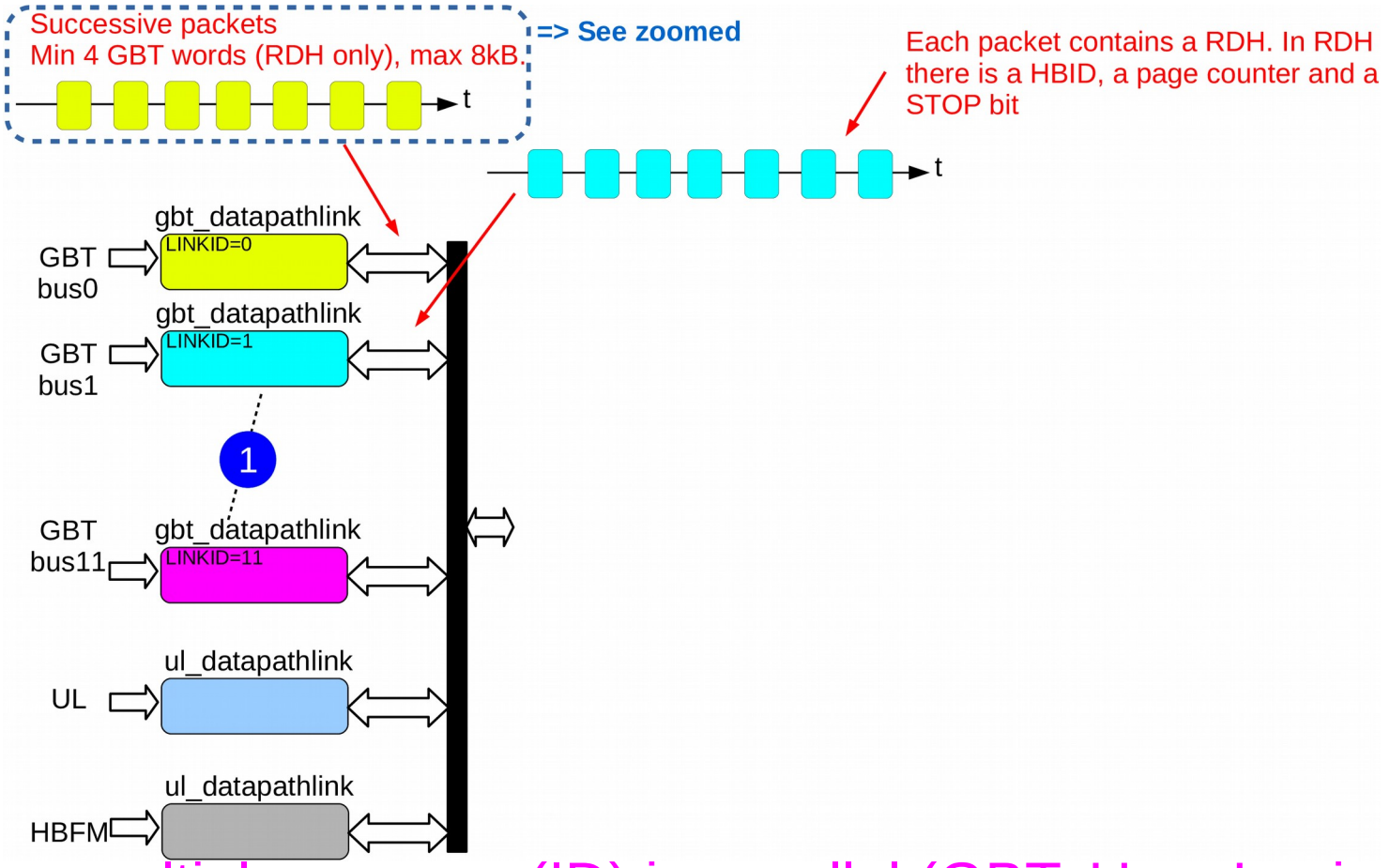
DATAPATH WRAPPER overview

- Collect data from multiple sources
  - Produces up to 7.68 GB/s per datapath wrapper
- Construct data packets for streaming detectors → chop the stream
  - Insert Reduced Data Header (RDH) at regular interval



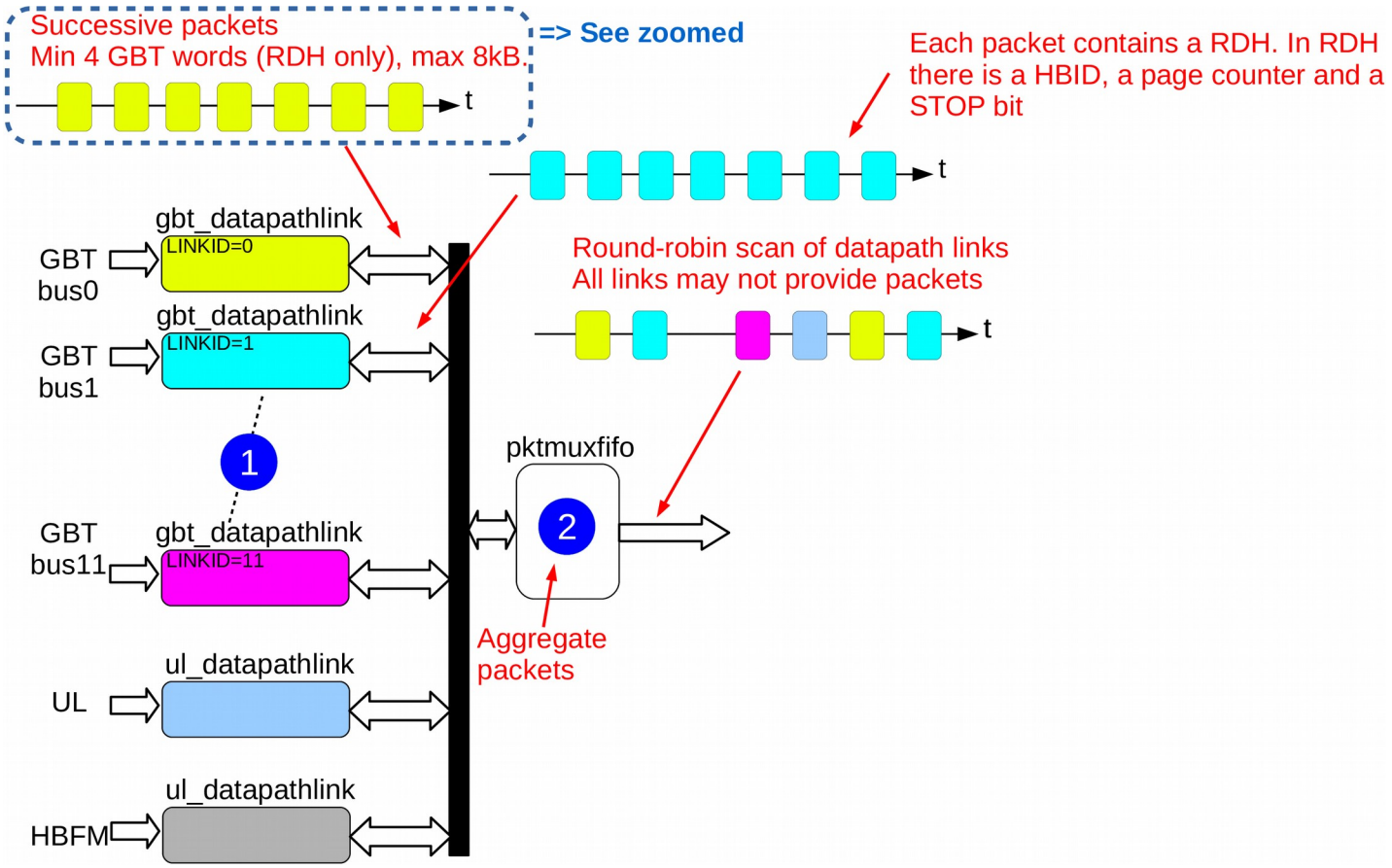
  - **RDH** : Heart Beat ID, Link ID, STOP bit, page counter, length, …
- Removes all data packets from a HBF if requested by the trigger system
- Finally, presents the data to the PCIe endpoints

DATAPATH WRAPPER
Operation



1) Data from multiple sources (ID) in parallel (GBT, User Logic, ...)

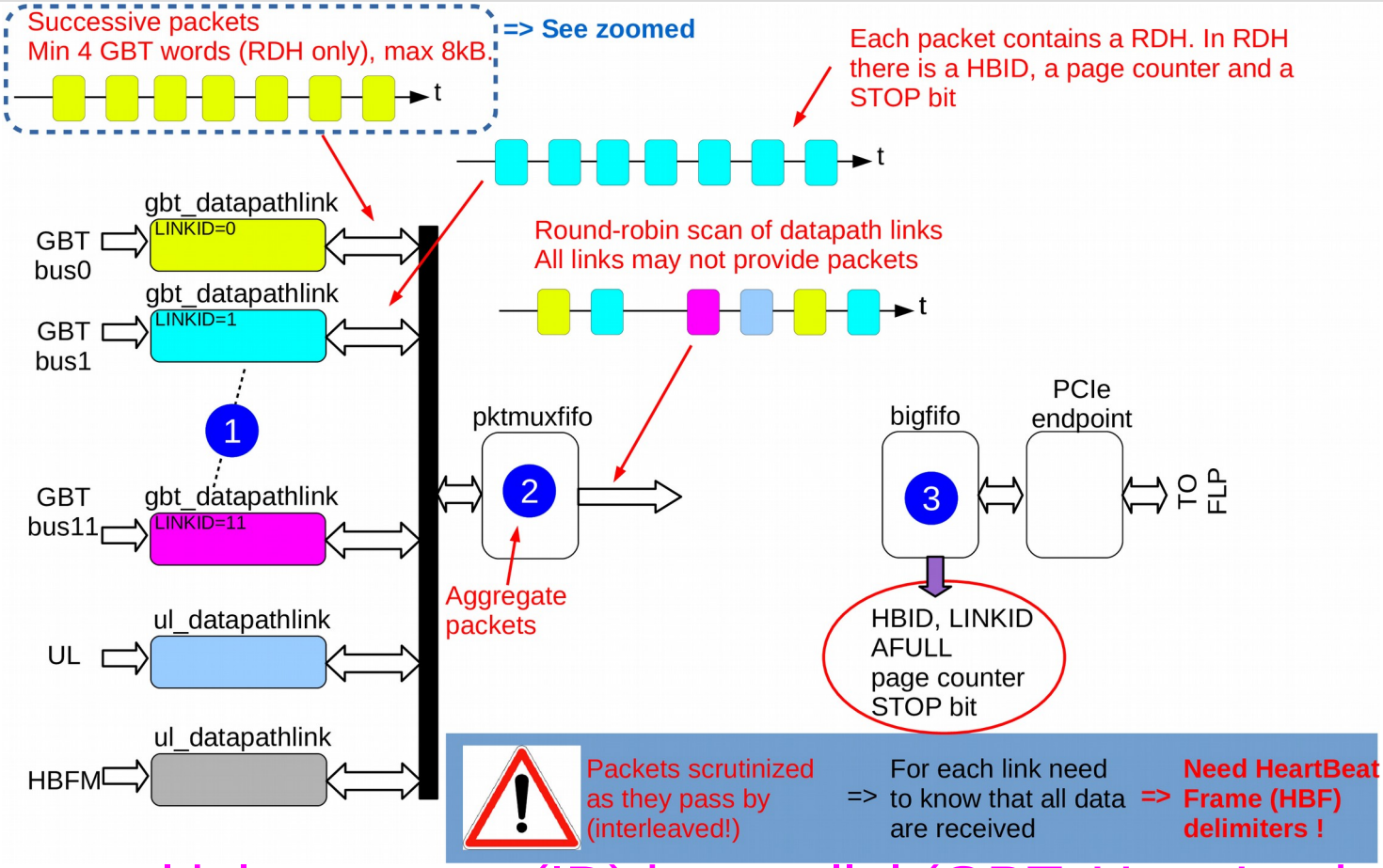
DATAPATH WRAPPER
Operation



1) Data from multiple sources (ID) in parallel (GBT, User Logic, ...)

2) Aggregation in pktmuxfifo → packets are interleaved

DATAPATH WRAPPER Operation

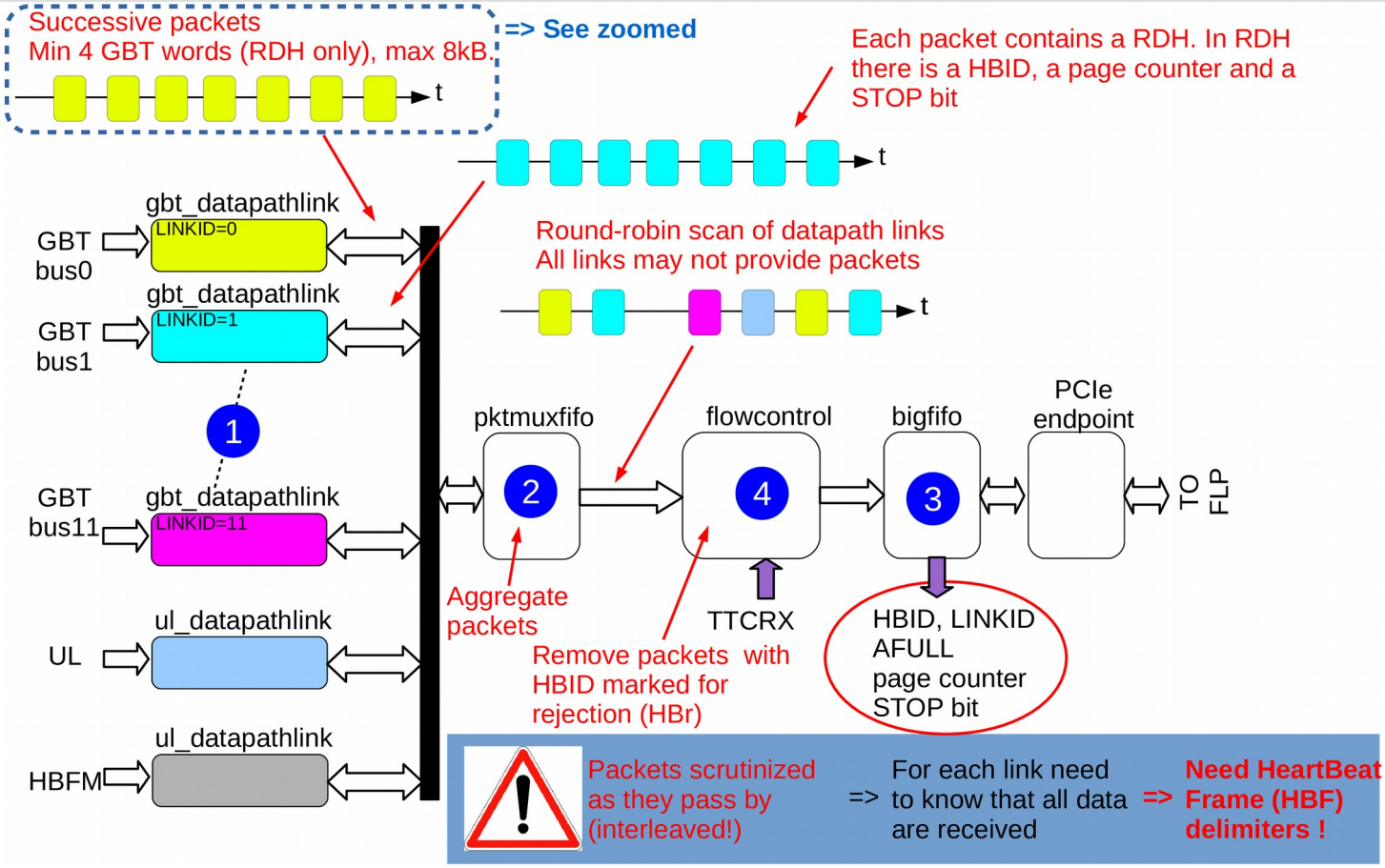


1) Data from multiple sources (ID) in parallel (GBT, User Logic, ...)

2) Aggregation in pktmuxfifo → packets are interleaved

3) Identify packets as they pass-by ⇒ **SCRUTINIZER**

DATAPATH WRAPPER Operation



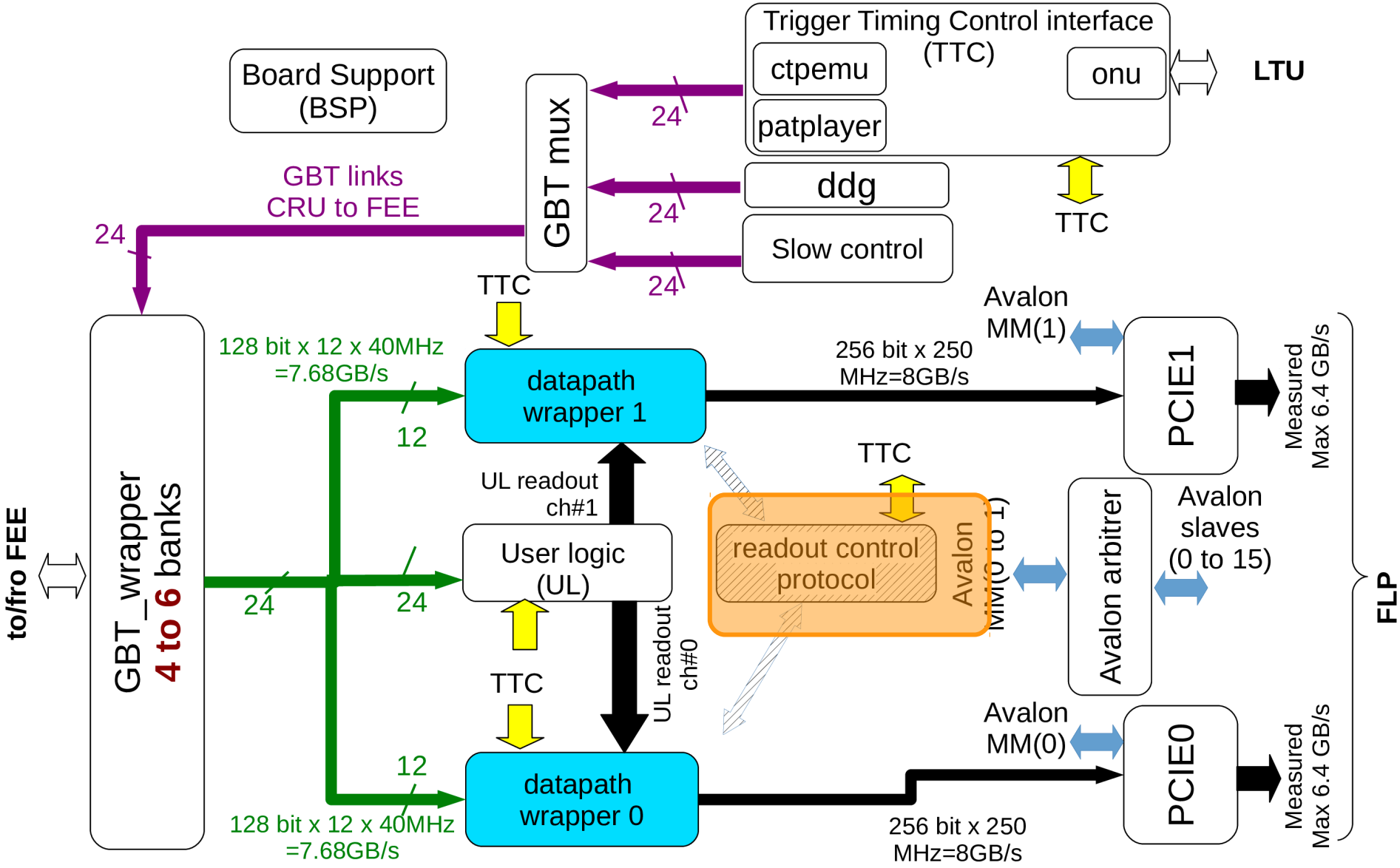
1) Data from multiple sources (ID) in parallel (GBT, User Logic, ...)

2) Aggregation in pktmuxfifo → packets are interleaved

3) Identify packets as they pass-by ⇒ **SCRUTINIZER**

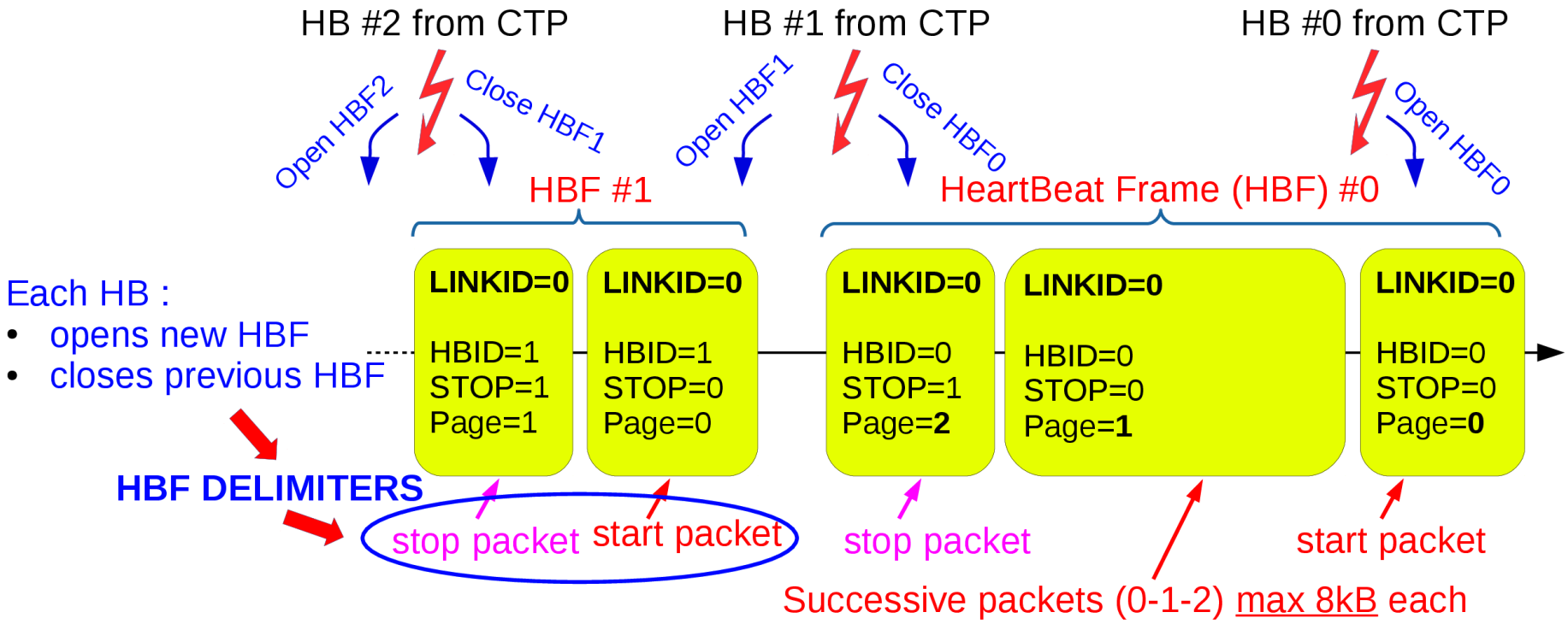
4) Remove packets marked for deletion if necessary

Firmware description: READOUT CONTROL



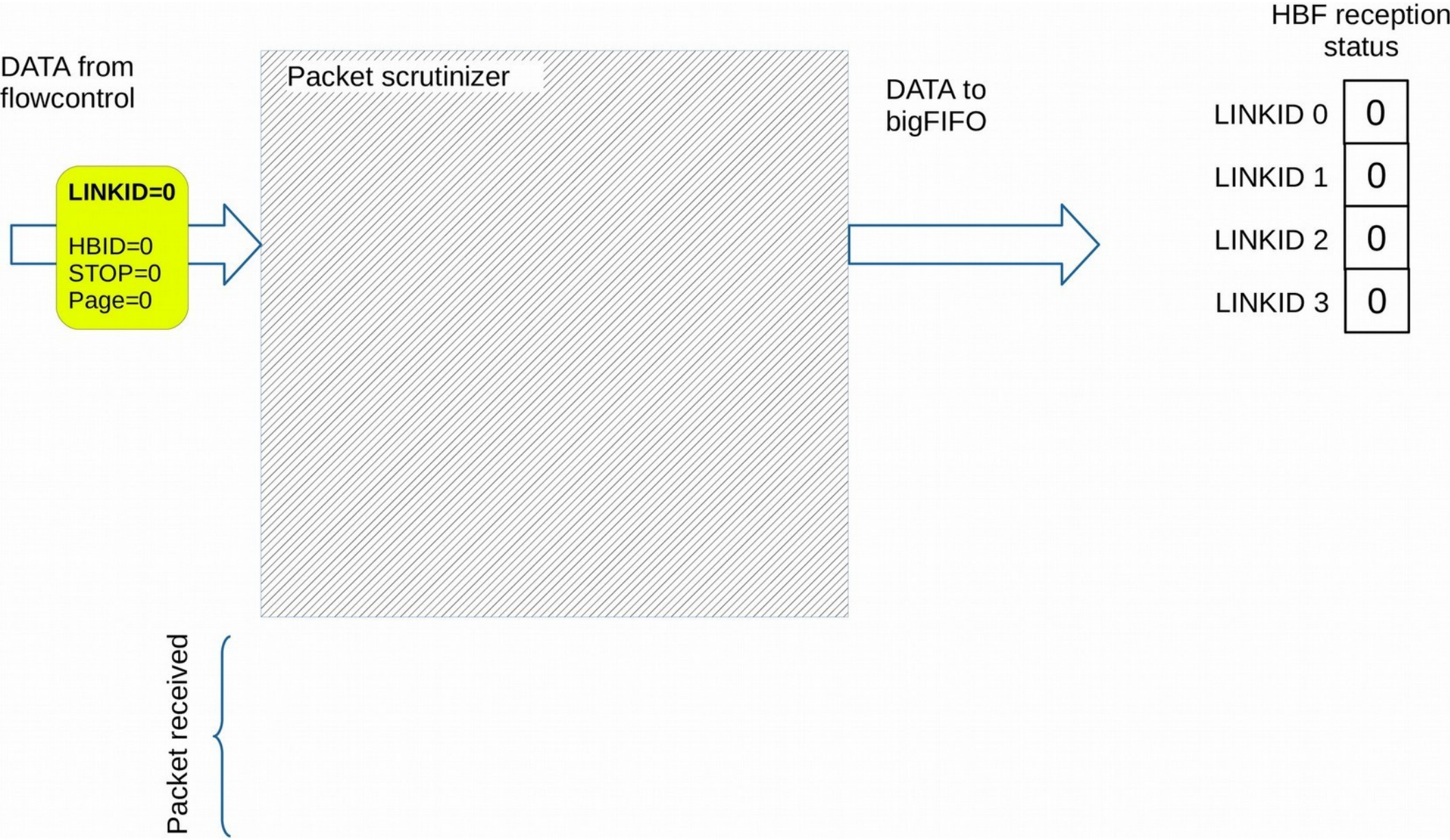
READOUT CONTROL
Overview

- Checks interleaved packets as they are flying-by
- **Successful HBF reception:** if for each LINKID, START and STOP packets were received and all packets of links were consecutive

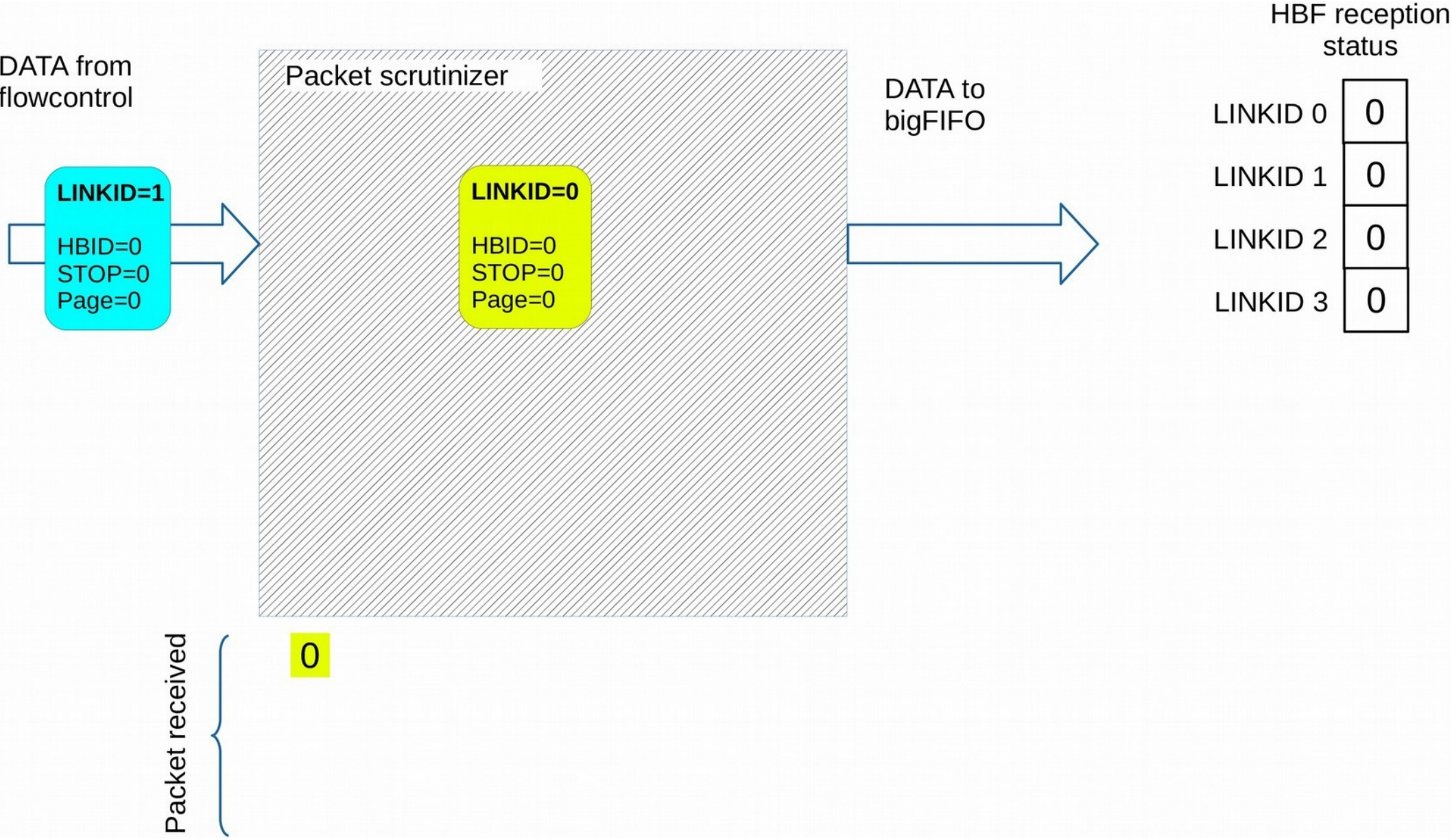


- HBF reception status is transmitted upstream to the Trigger system : trigger message

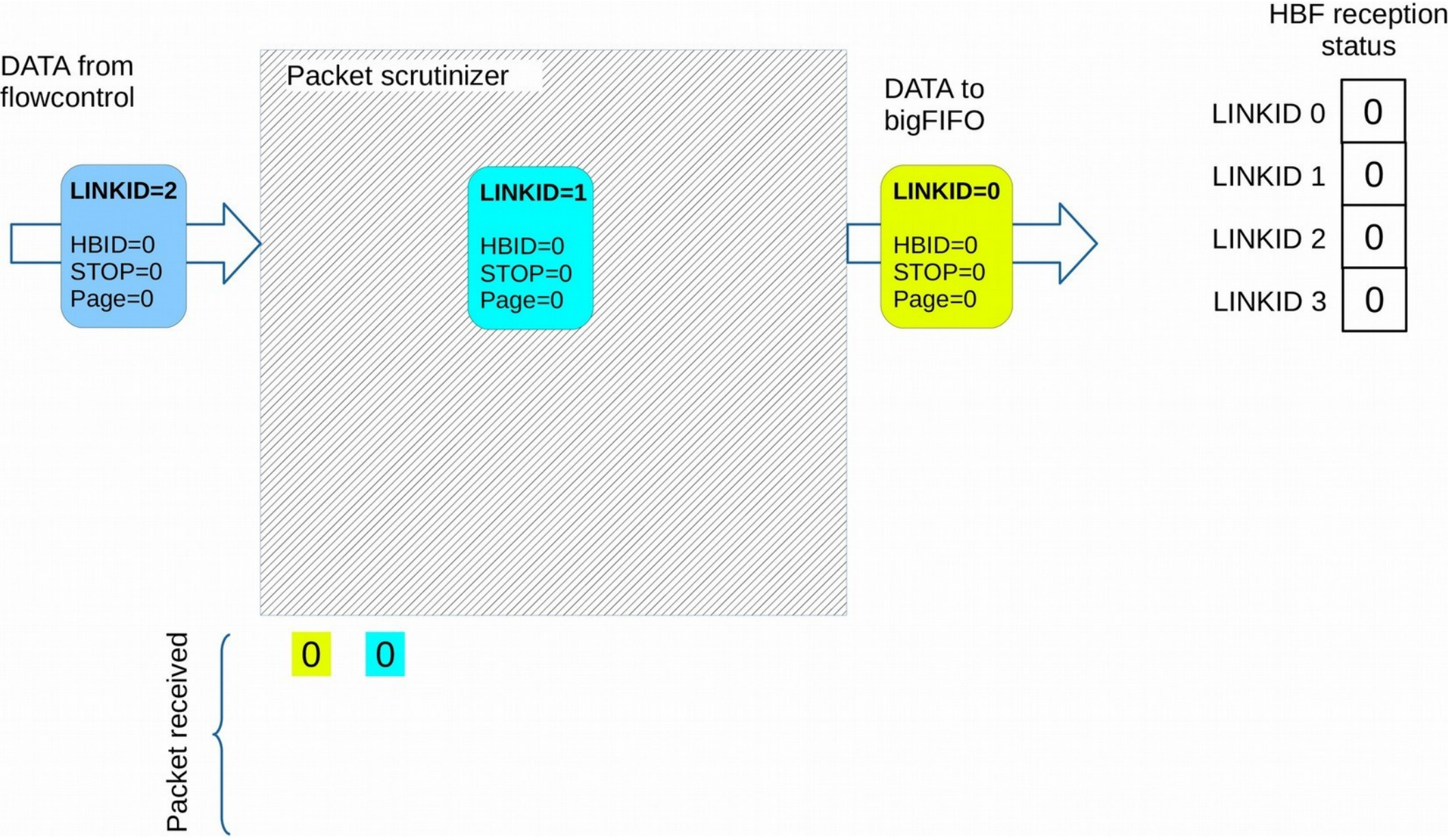
READOUT CONTROL
How the HBF are checked



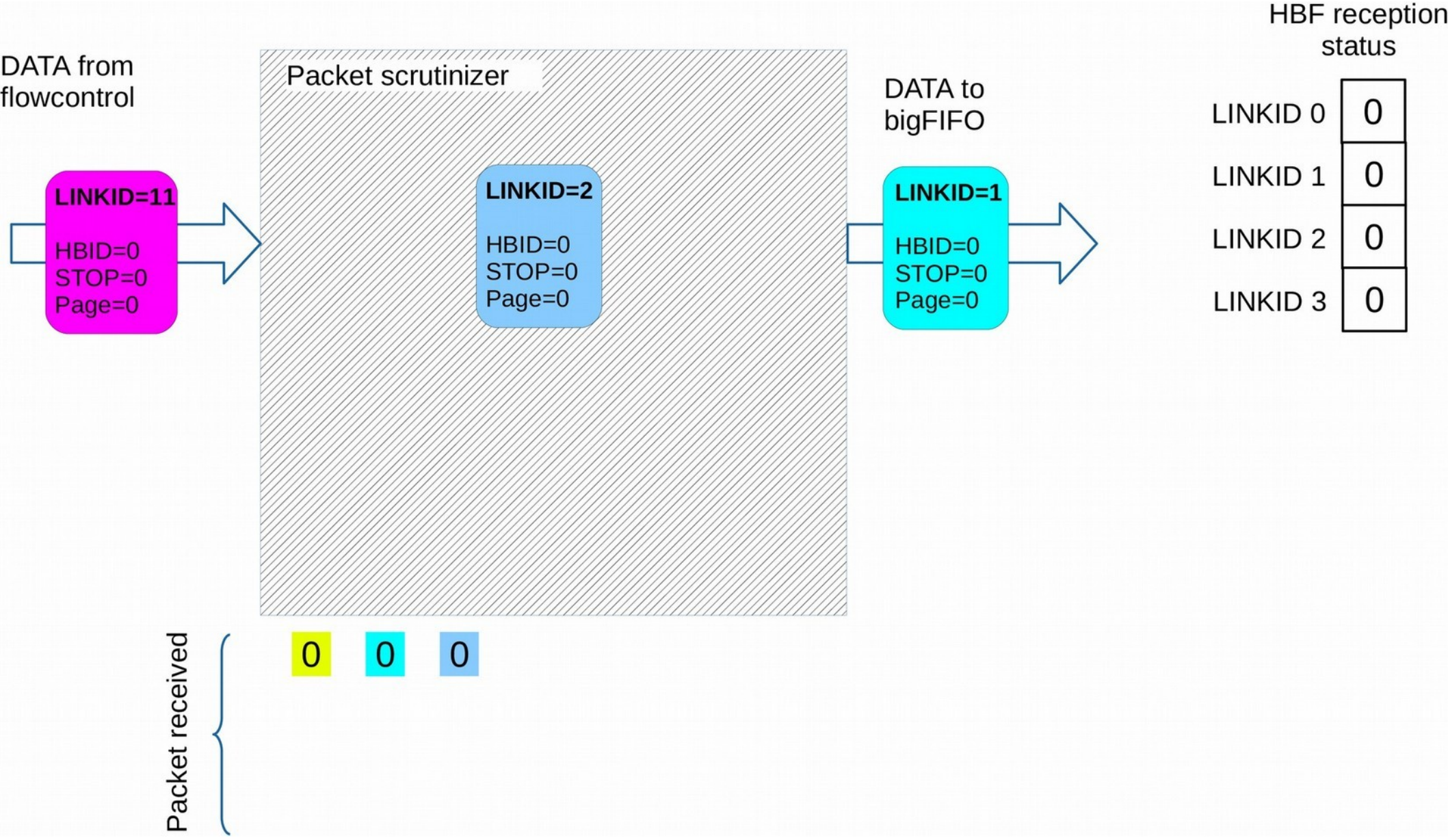
READOUT CONTROL
How the HBF are checked



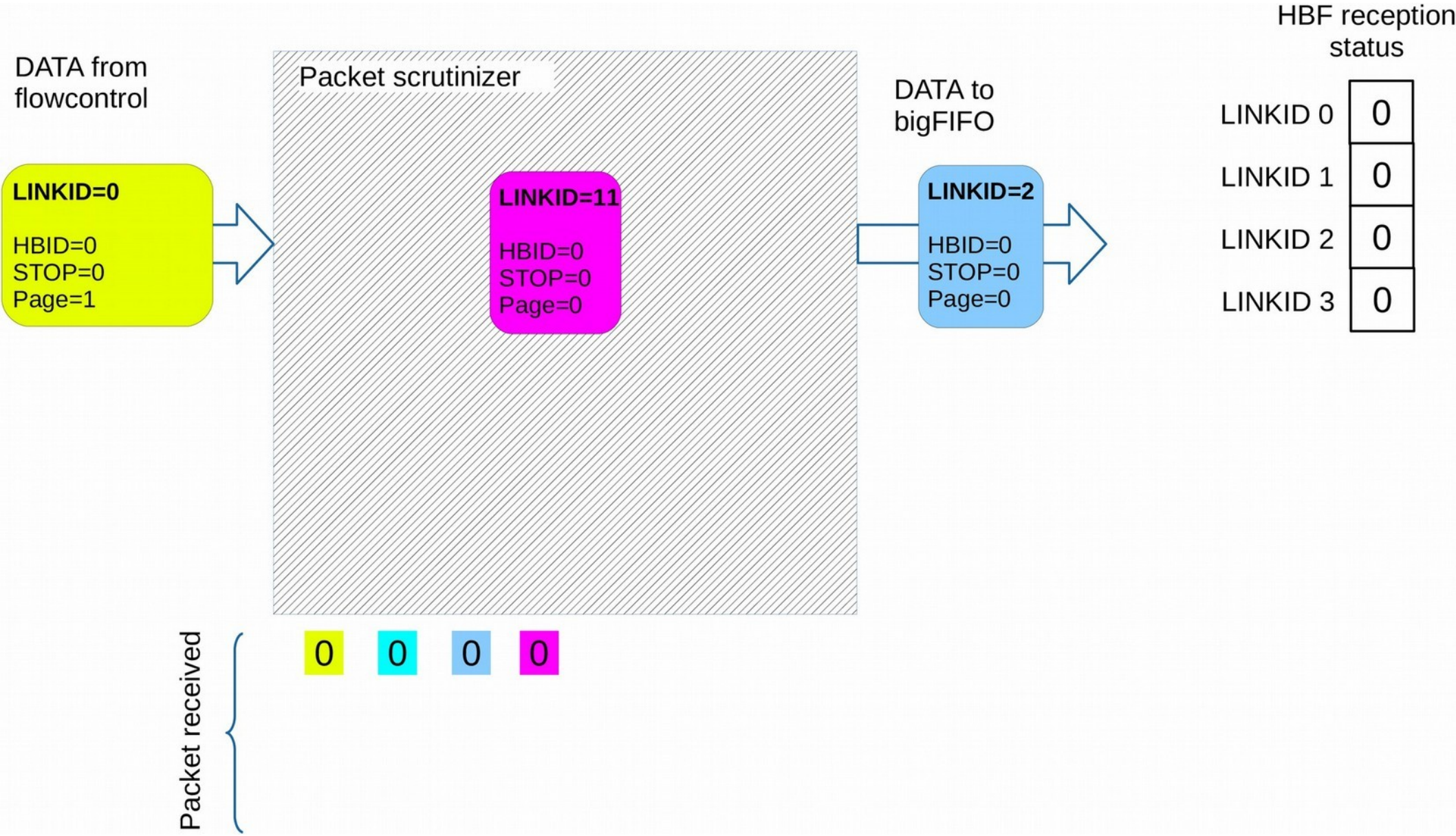
READOUT CONTROL
How the HBF are checked



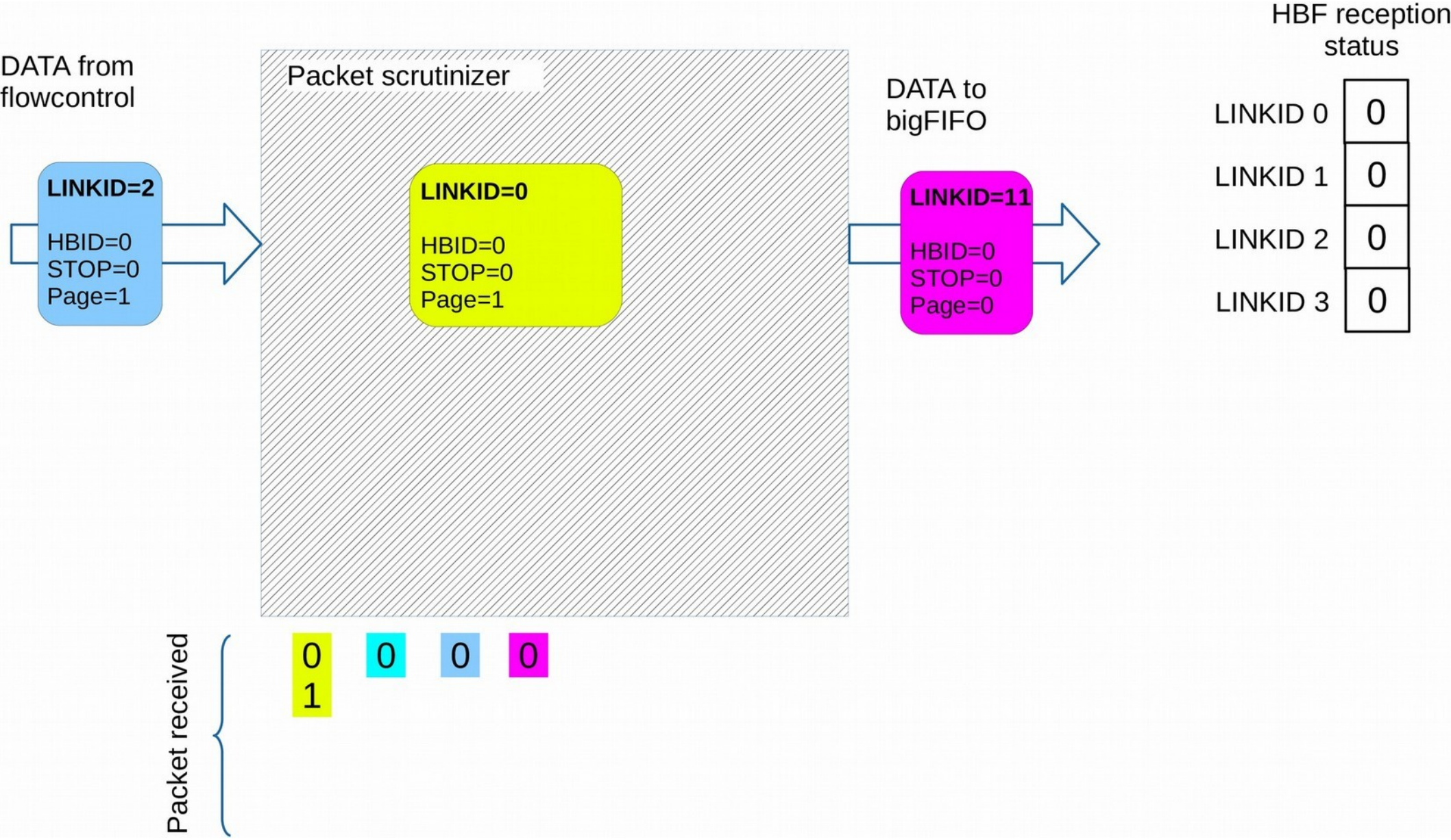
READOUT CONTROL
How the HBF are checked



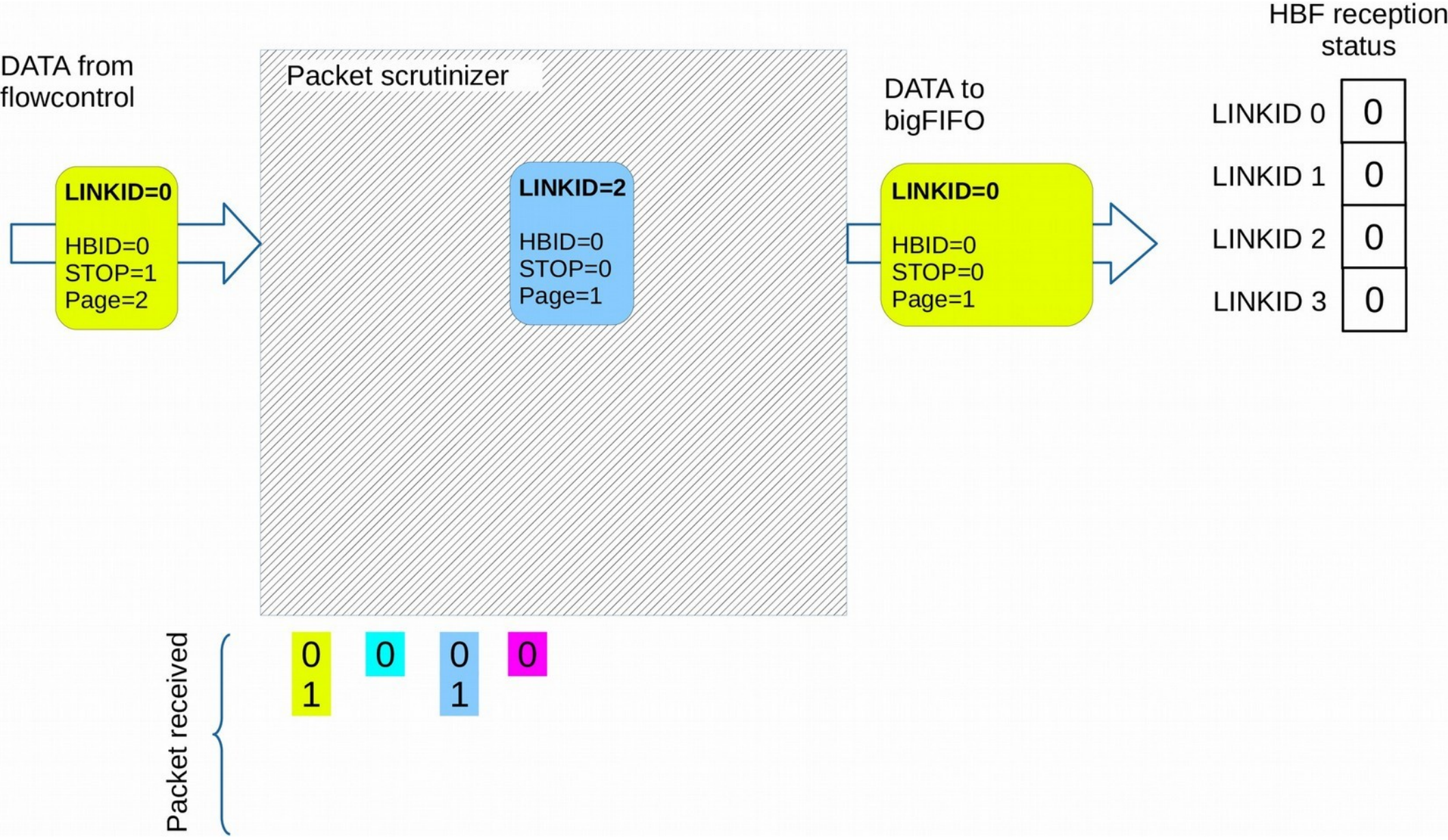
READOUT CONTROL
How the HBF are checked



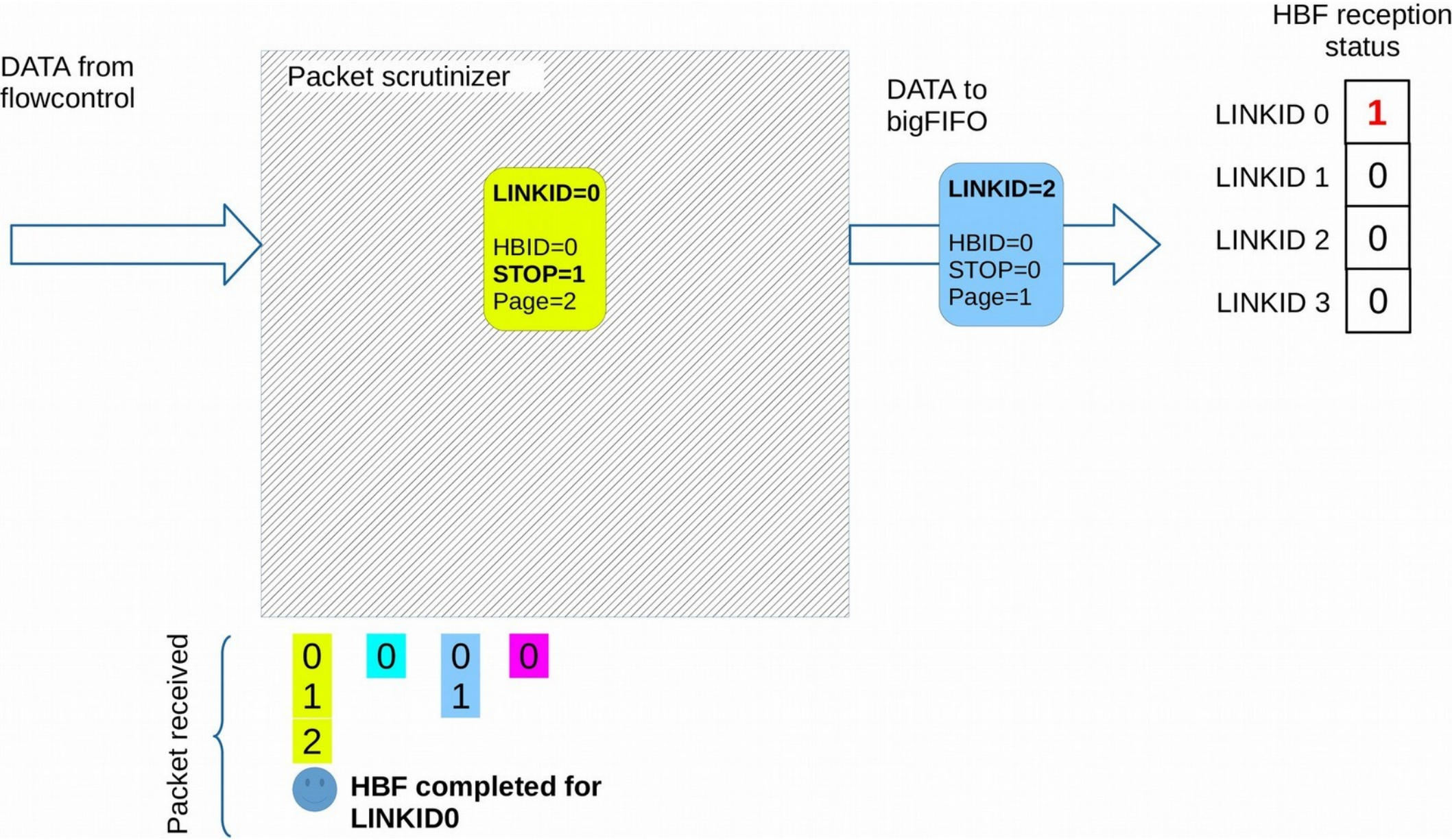
READOUT CONTROL
How the HBF are checked



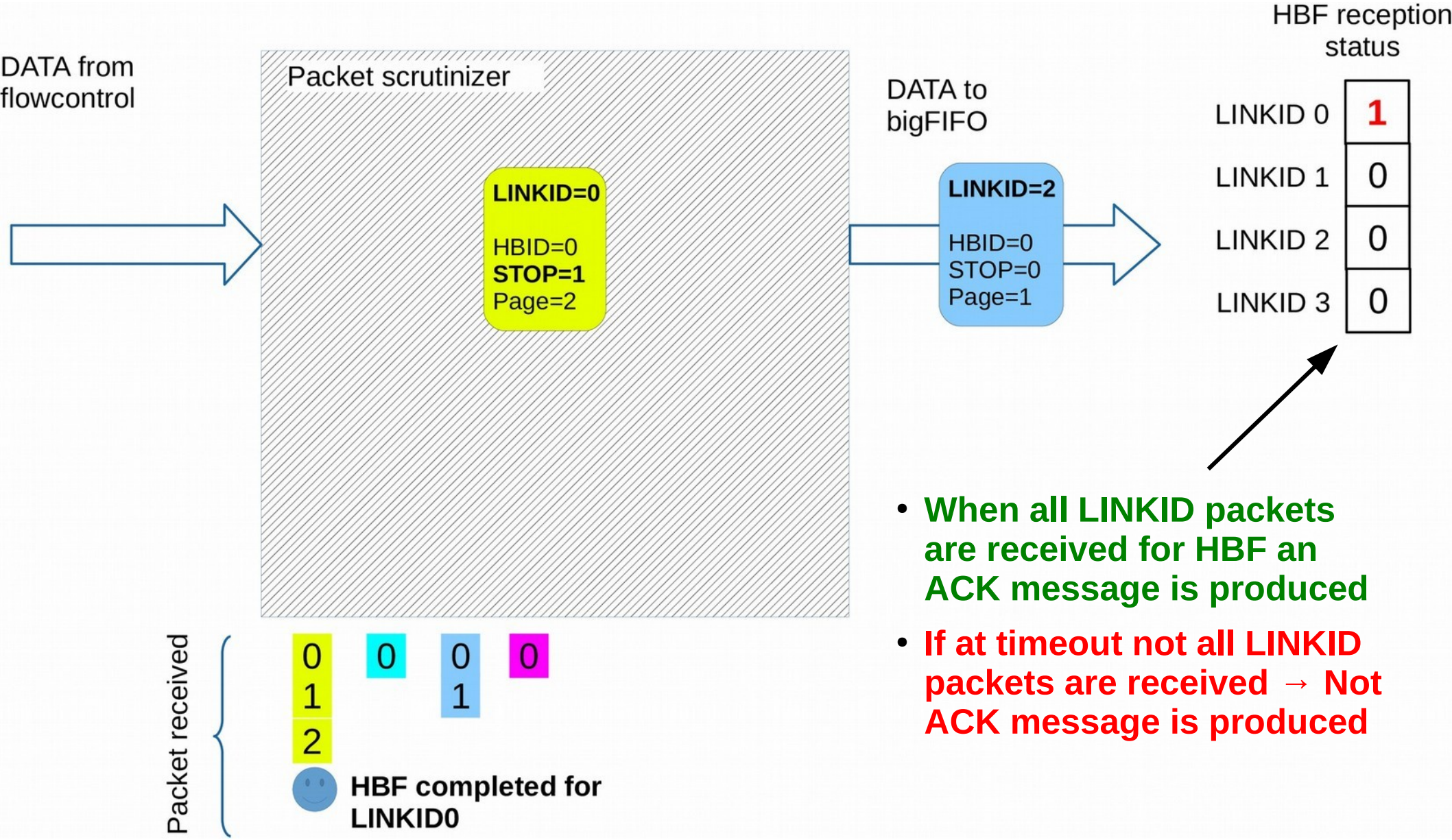
READOUT CONTROL
How the HBF are checked



READOUT CONTROL
How the HBF are checked

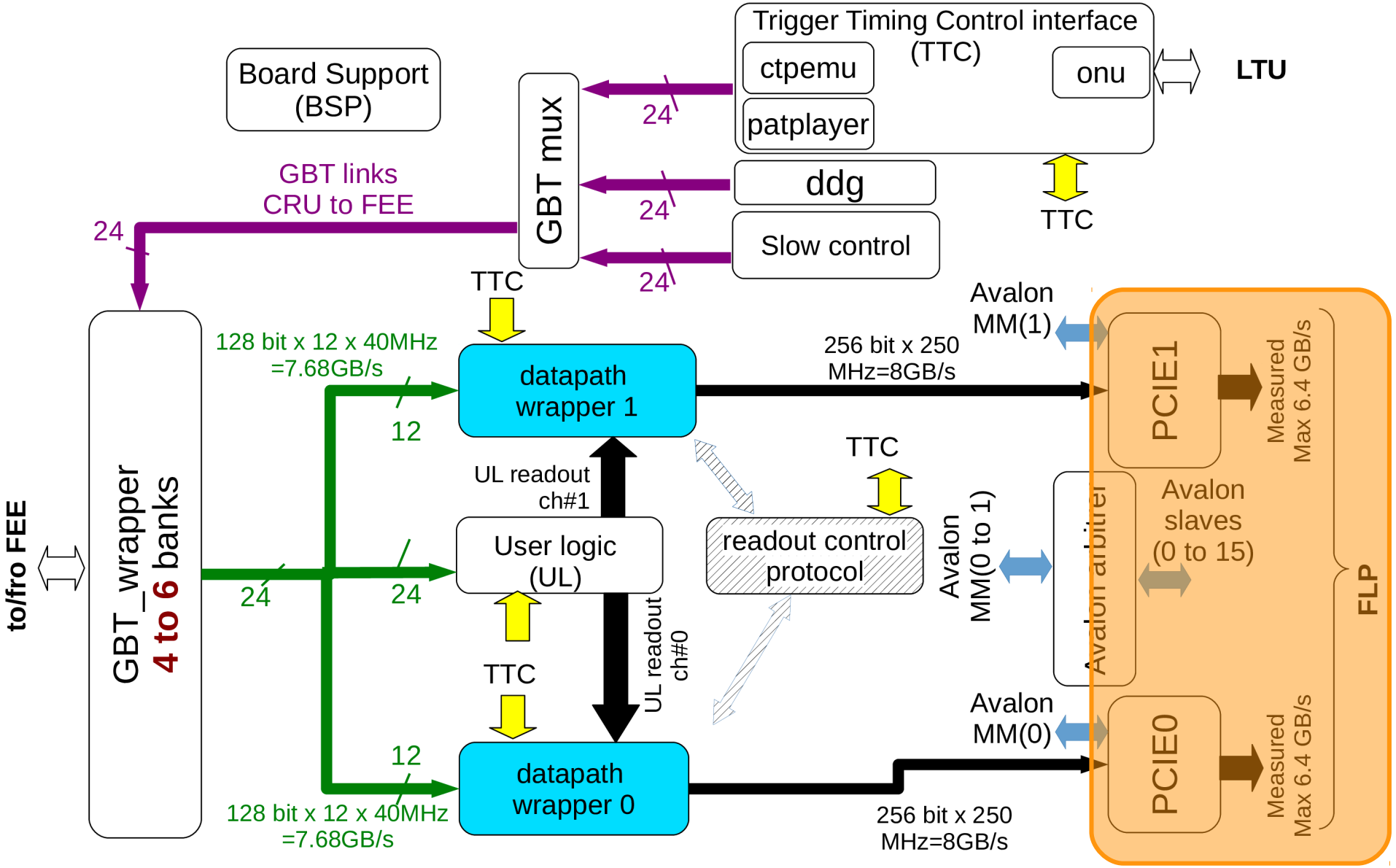


READOUT CONTROL
How the HBF are checked



- **When all LINKID packets are received for HBF an ACK message is produced**
- **If at timeout not all LINKID packets are received → Not ACK message is produced**

Firmware description: PCie



CONCLUSION

- An adaptable and common firmware was developed to cover the needs of the upgraded detectors
  - Development and validation efforts were shared
  - Lab software framework is delivered along with the firmware (python library)
  - CRU and its common firmware is already used extensively and successfully by several detectors
- Continuous readout mode was already validated
- Resources usage of the common firmware (with GBT dynamic switching)
  - For 24 GBT links : 121,297 / 427,200 ALM (28 %)