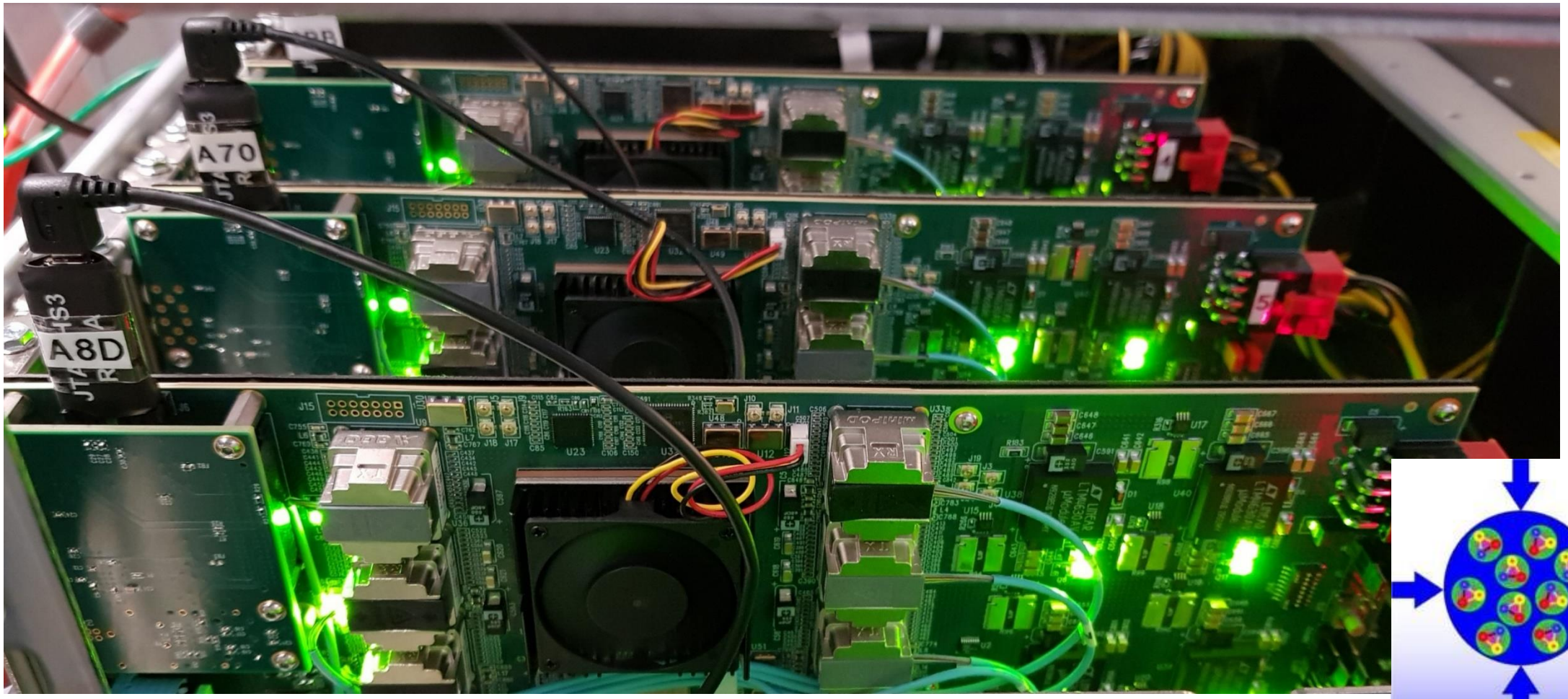
# The CBM DAQ system




David Emschermann
CBM CM44 Students day
CTU, Prague – 15. September 2024

HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES

CBM

# Outline

- CBM DAQ in a nutshell
- mCBM - technology testbench
- Timing and Fast Control System
- Entry Nodes, Processing Nodes, Virgo Cluster
- Core DAQ component: CRI1 and future CRI2

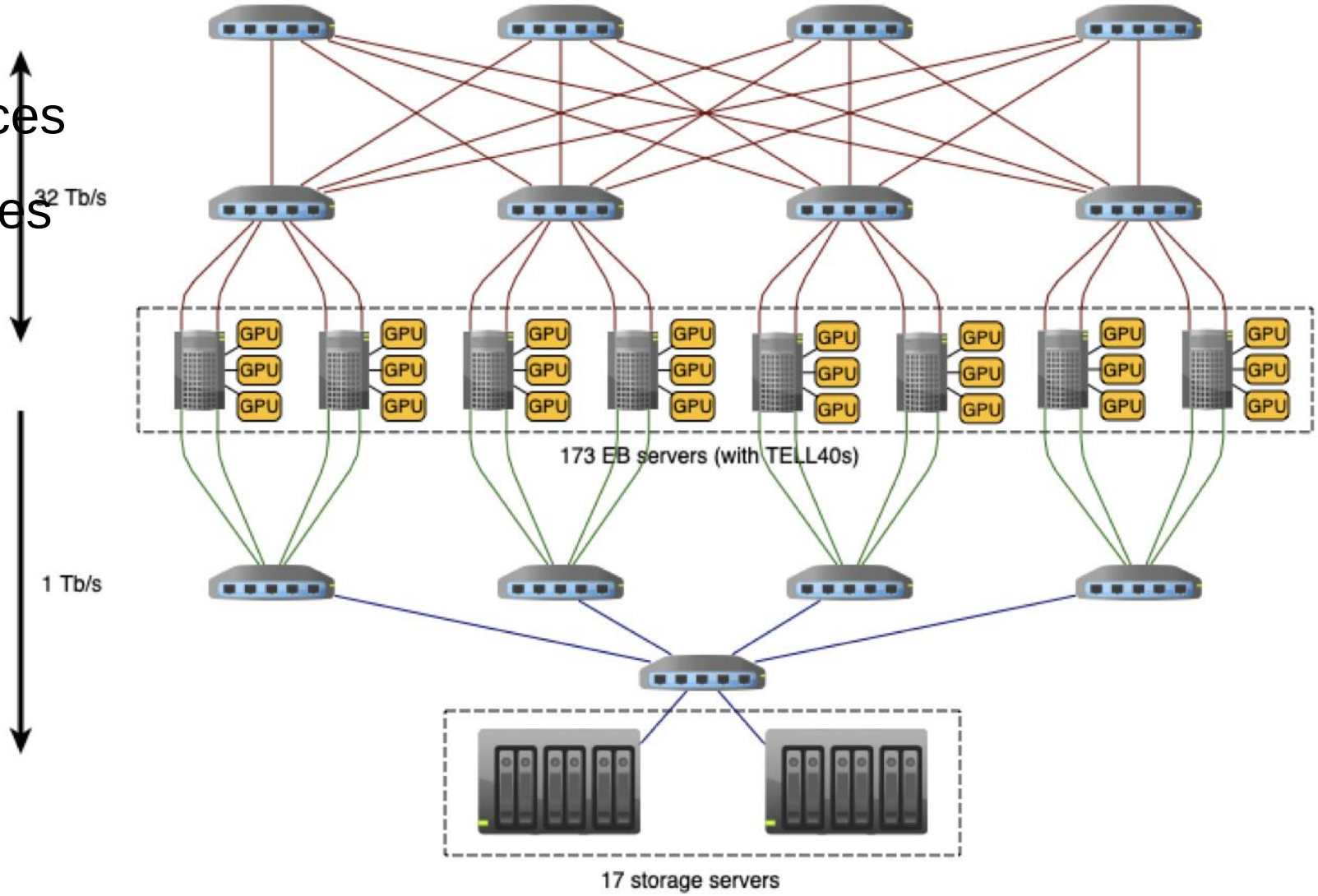
# Generic DAQ scheme (LHCb)

FLIM

Micro slices

Time slices

.tsa



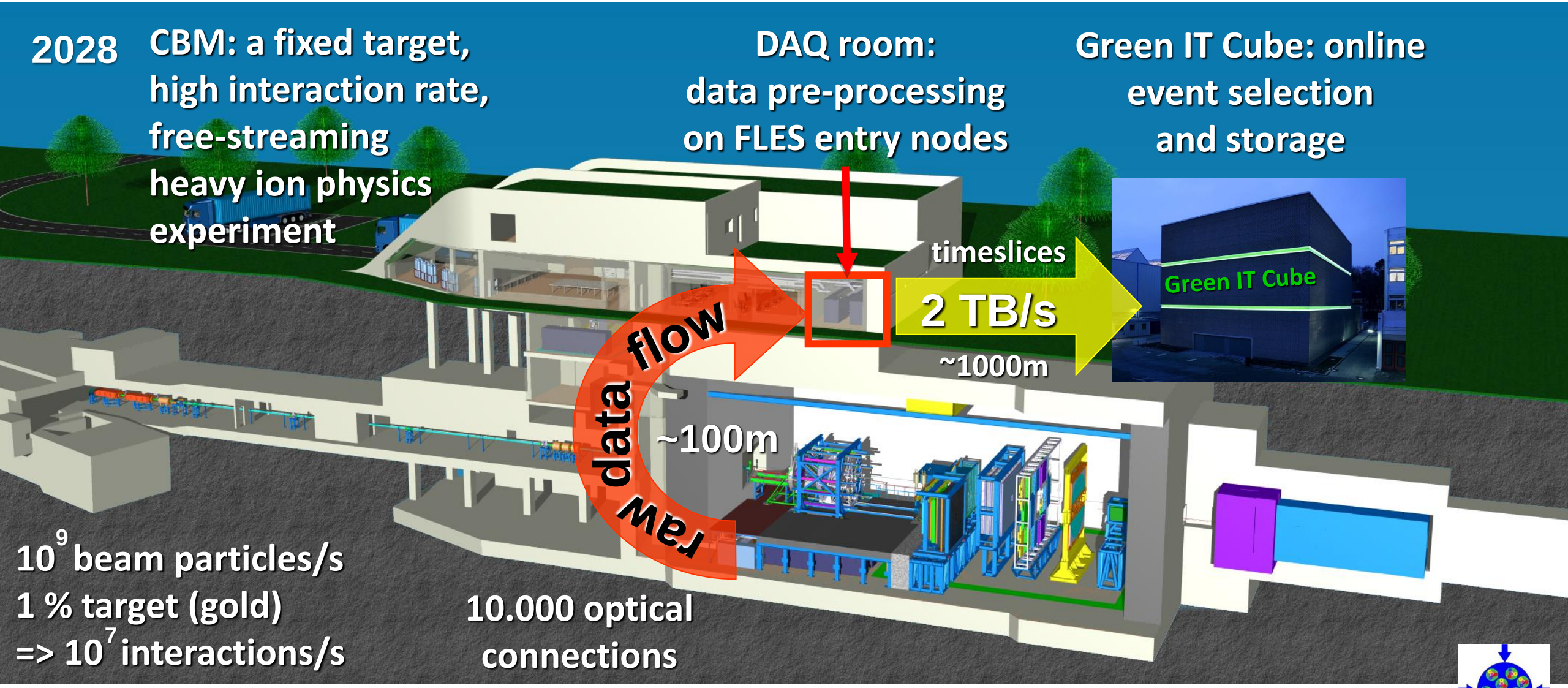

Entry Nodes

Processing Nodes

Online Nodes (Virgo)

Data Storage (Lustre)

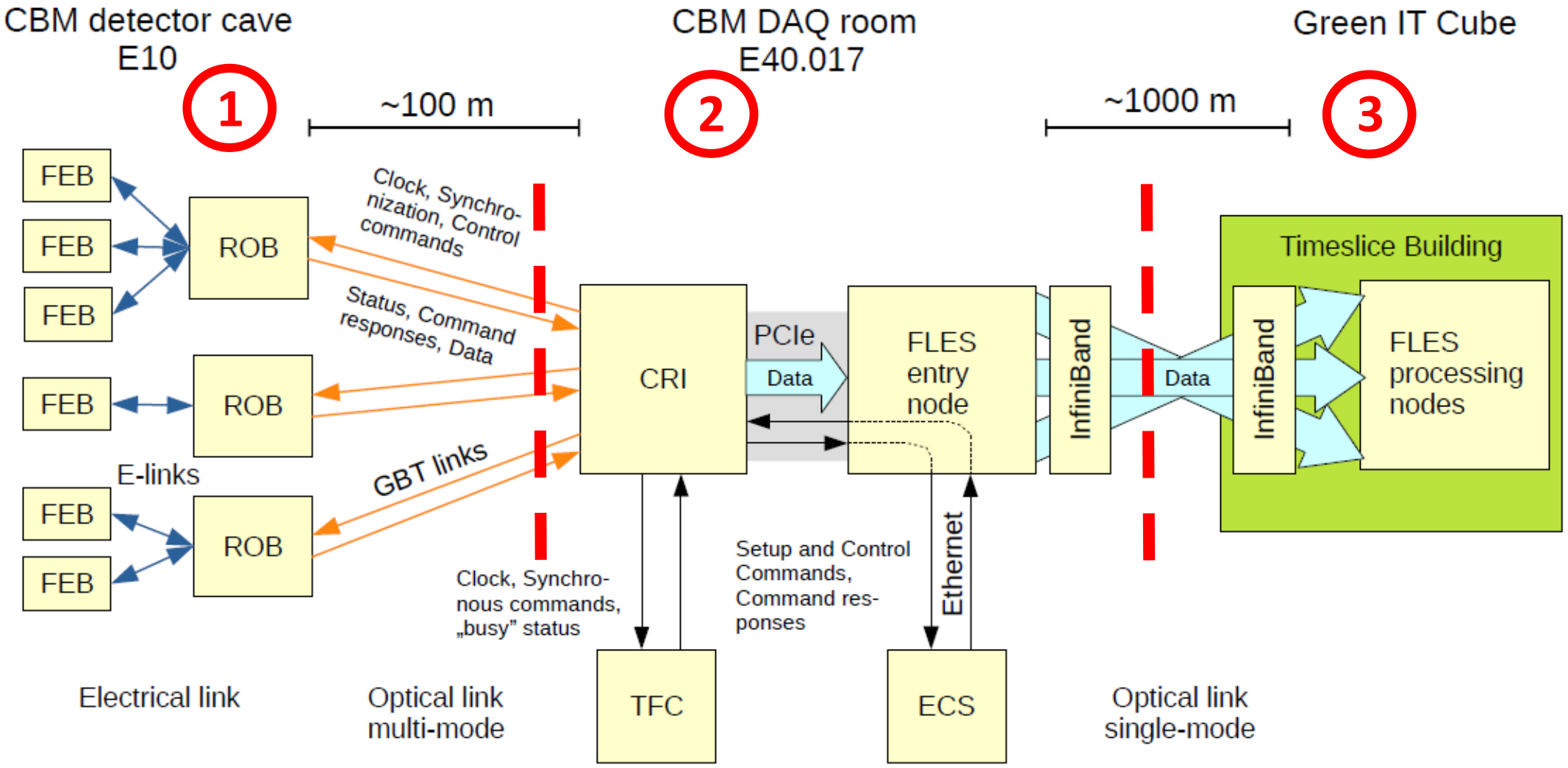
# The CBM data flow at SIS100



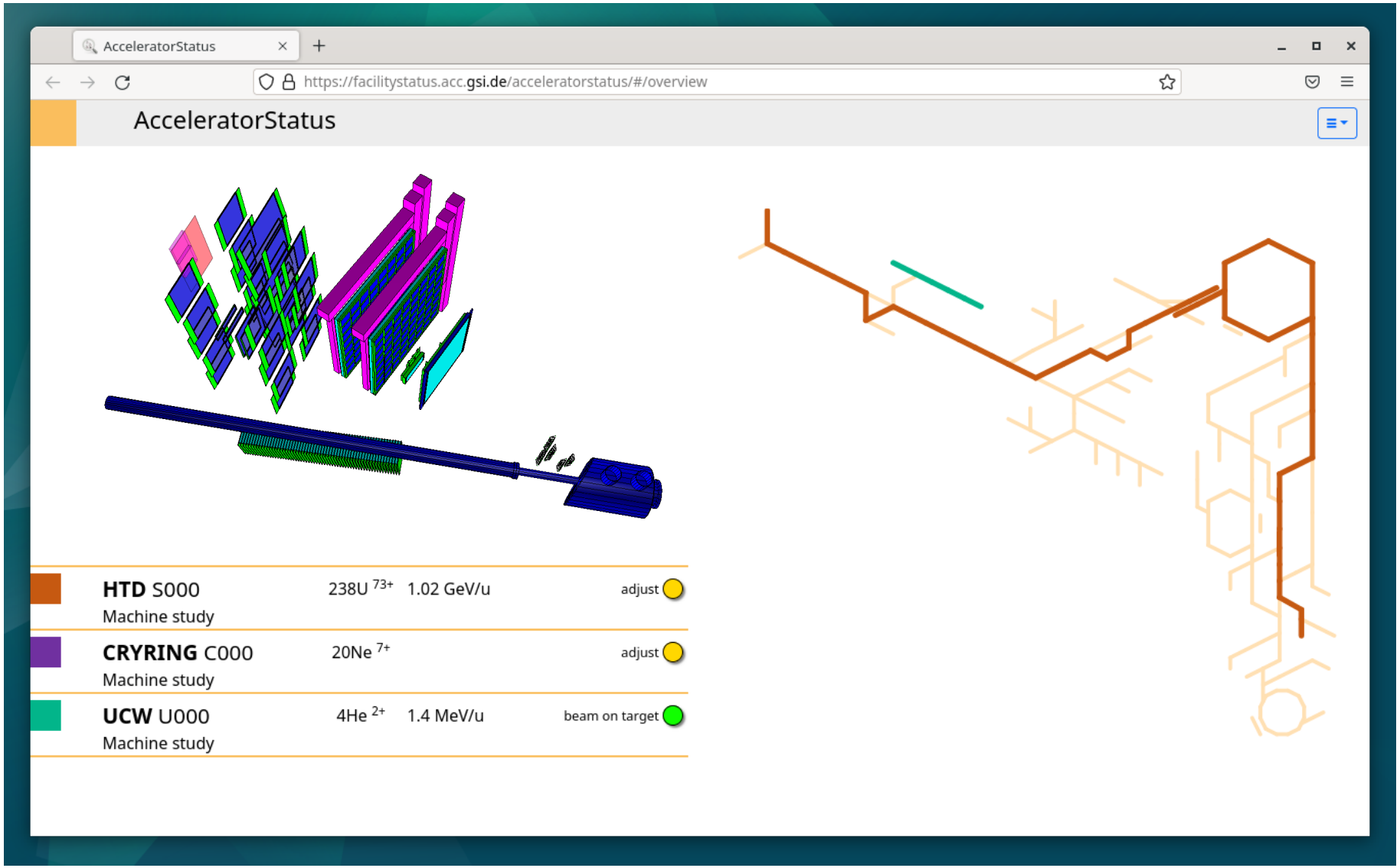
# CBM to Green-IT-Cube connection


FAIR construction site – October 2021

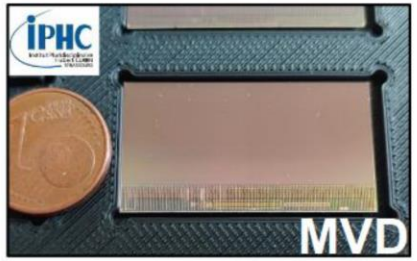
# The CBM readout and control architecture (CRI based)

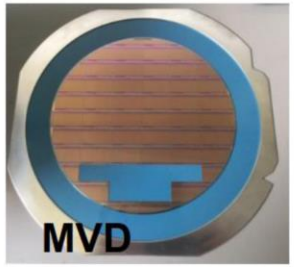


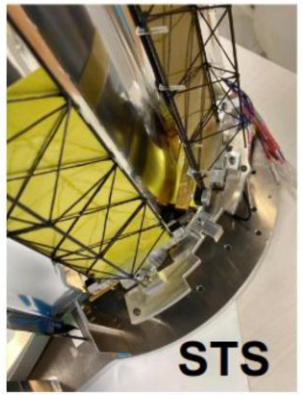
# mCBM – technology testbench

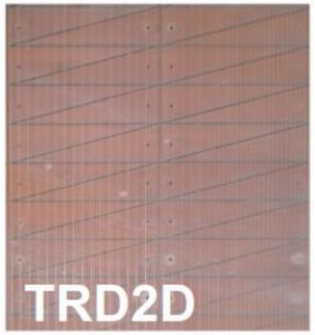


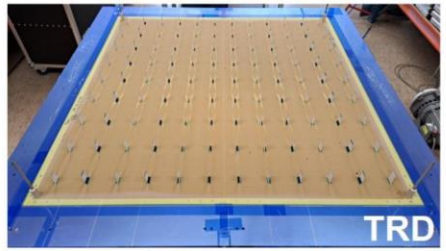
# CBM – pre-series detector coponents



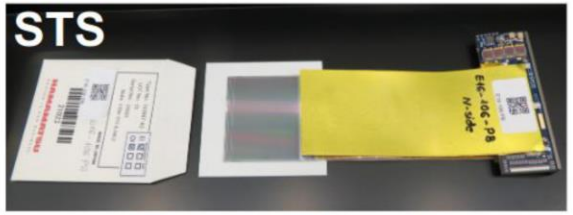



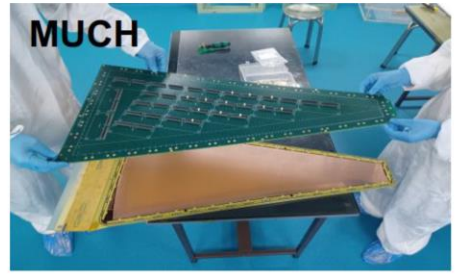



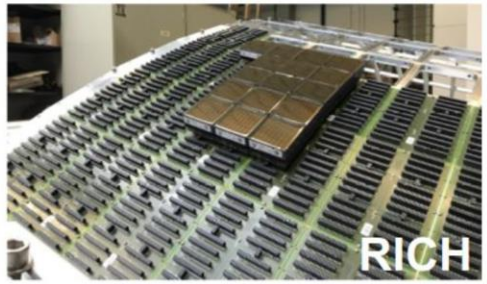



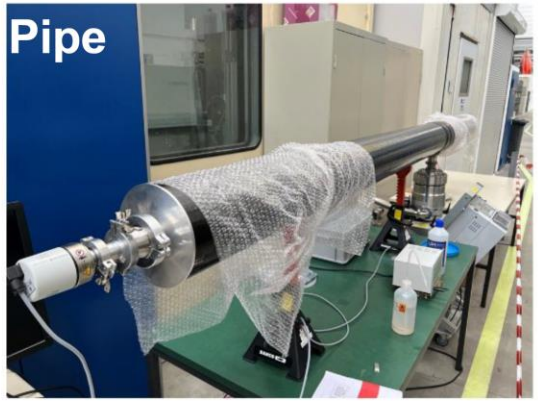


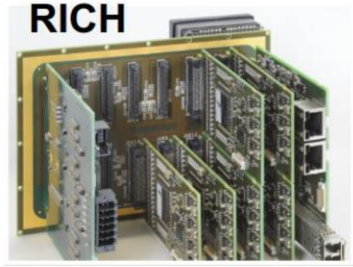


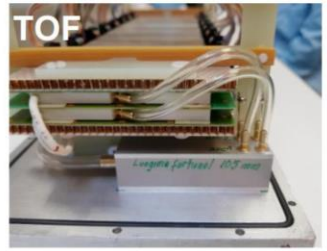



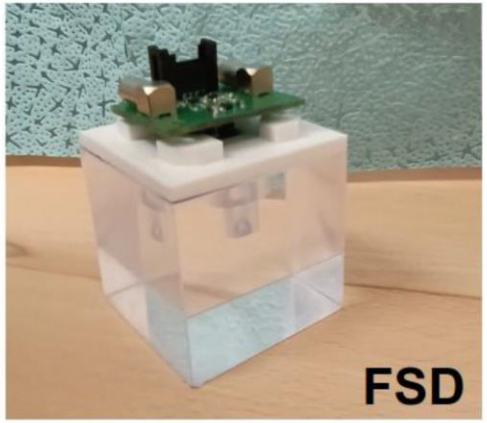





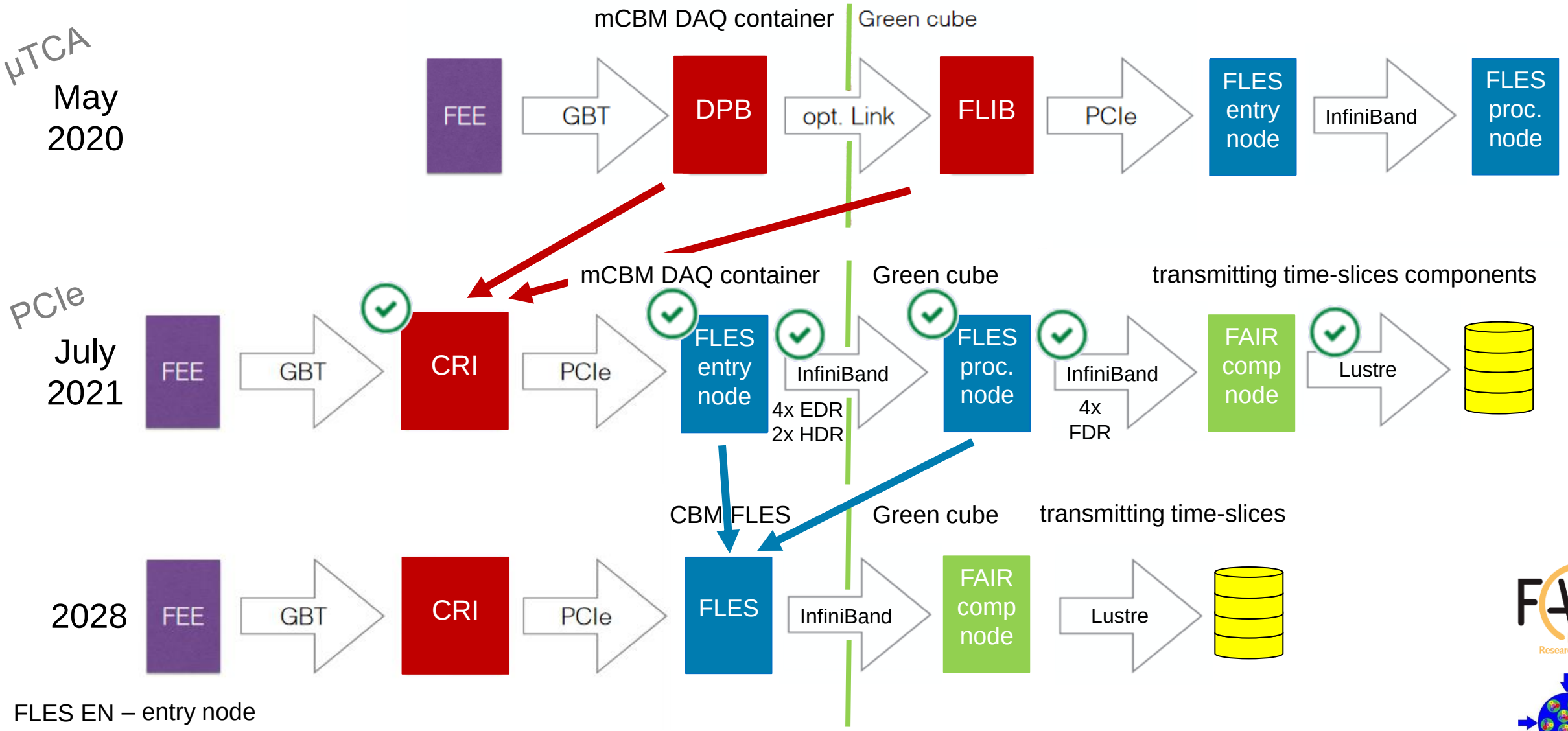


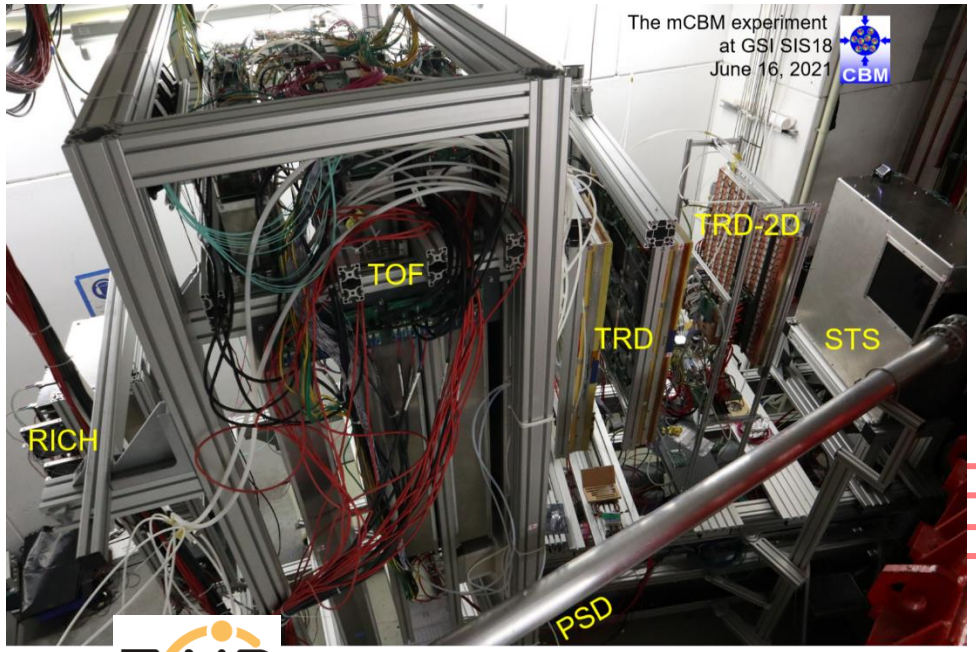




# Evolution of the readout chain at CBM in the past 5 years




FLES EN – entry node
FLES PN – processing node
FAIR CN – compute node

# mCBM @ SIS18 - data transport with CRIs




optical fibers 50m

DAQ container

CRI (PCIe) @ FLES entry nodes

optical fibers

300 m


FLES processing nodes

FAIR Phase 0 Research Program

TFC (CRI based)

triggerless-streaming FEE **assigning time stamps to hits**

1 m Copper

GBTx

50 m optical

CRI FPGA **µSlice** building

FLES entry nodes

300 m optical InfiniBand

Green IT Cube

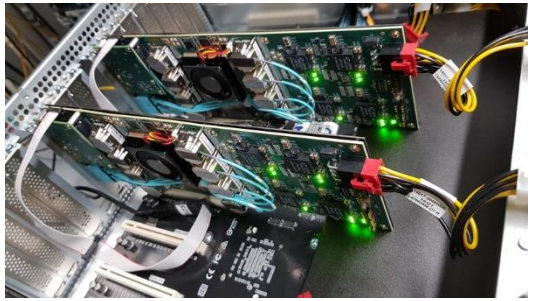
Timeslice Building

Event Reconstruction

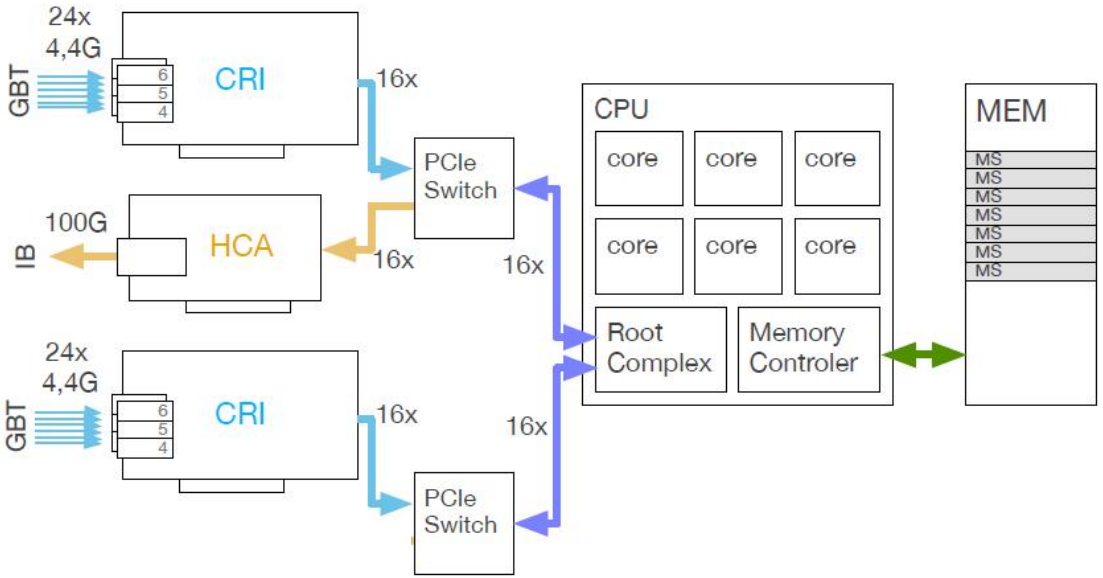
Event Selection

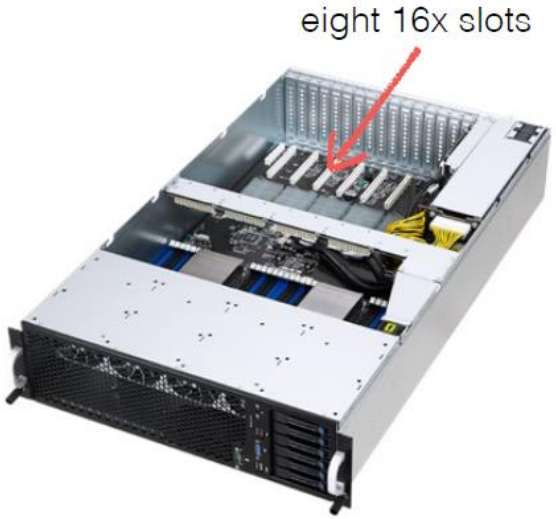
Archiving

# Entry Node configuration - CRI data path



**2 CRI cards installed in one entry node**



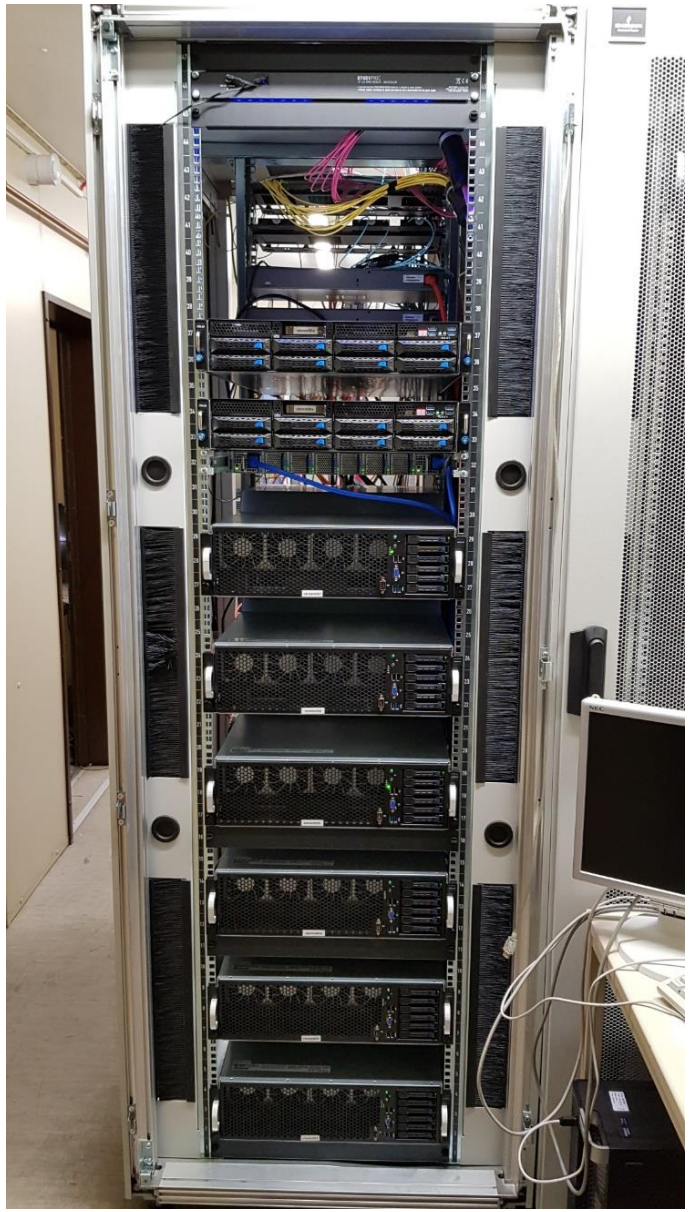

ASUS ESC8000
-> prototype for mCBM

basic configuration of CRI entry nodes for 2021:

- 2x CRI - max 235 Gbps (in)
- 1x HDR HCA - max 120 Gbps (out)

- 13x CRI boards are installed in the CRI rack

# The CRI DAQ rack prototype



- This is the CBM DAQ prototype rack (est. September 2020)
- all data from mCBM subsystems are transiting here (scale-up 24x for SIS100)

**mcbmcri** - JTAG server
**devel09** - 1x TFC-Master

**devel08** - 1x CRI

**devel07** - 2x CRI

**devel06** - 3x CRI

**devel05** - 2x CRI

**devel04** - 2x CRI

**devel03** - 1x CRI

**devel02** - 1x CRI

- this rack hosts the TFC system and FLES Entry Nodes fitted with CRI1

Up to 3 CRI cards and 1 HDR HCA are installed in each Entry Node

Information and usage details are in the Redmine Wiki:

https://lxcbmredmine01.gsi.de/projects/mcbm/wiki/CRI_operation

# White Rabbit and PTP interface to GSI

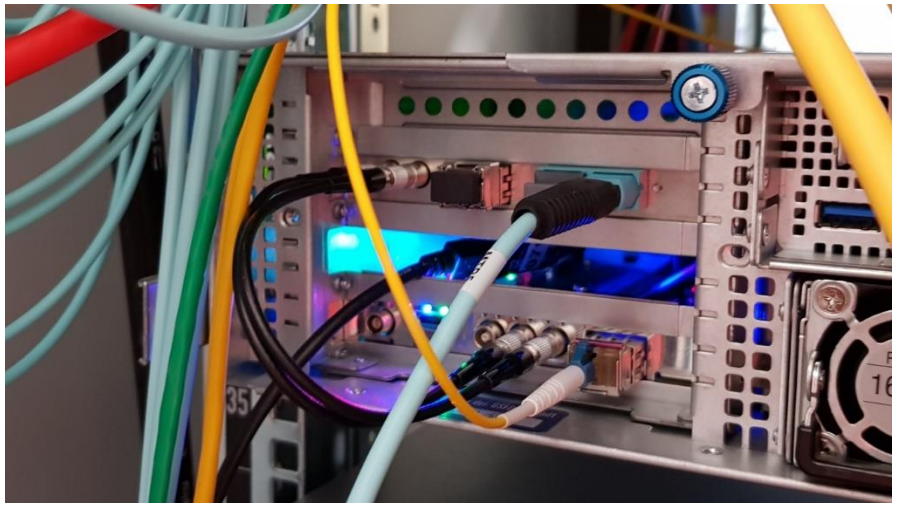
At mCBM we are commissioning prototype DAQ components to be used for the day-1 readout of CBM.


- mCBM is linked to the White Rabbit network of GSI
- serves as time source to the TFC and PTP master to FLES
- allows to receive spill on/off information from the accelerator

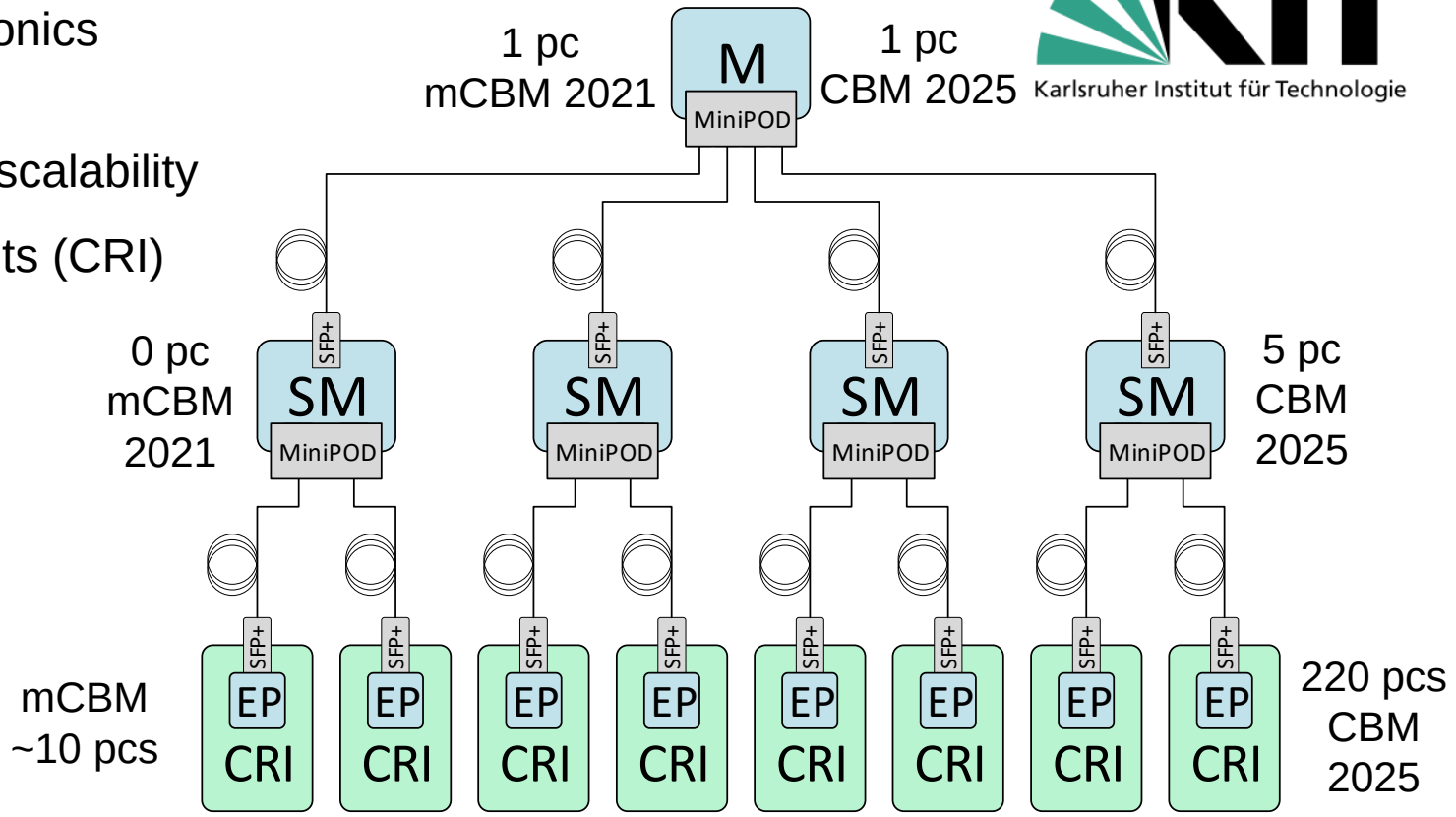
WR interface added 09/2021

# TFC System – Synchronous operation of CRI and attached FEE

- The Timing and Fast Control system (TFC) synchronises the data processing electronics experiment-wide over optical fibres
- Organised as a hierarchical network for scalability
- Distributes timing information to endpoints (CRI)
- Based on CRI cards



TFC-Master CRI




1x TFC-Master commissioned at mCBM in July 2021.
To be scaled up with TFC-Submasters for operation at CBM.

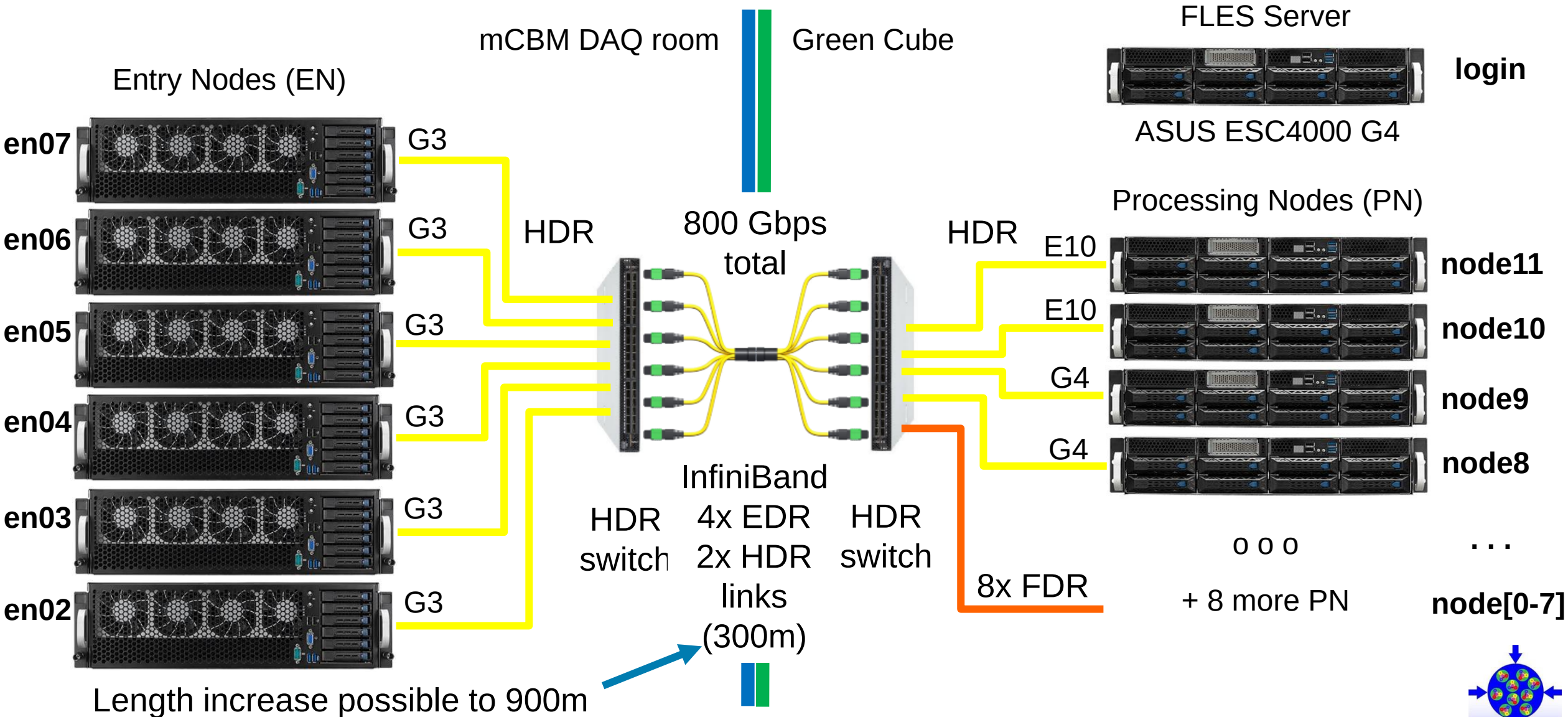
# TFC with 1:47 or 1:24 hardware

The TFC system needs to synchronize 213 CRI (Online TDR Part I, Table 5.1) to one common time source (WhiteRabbit accelerator interface).

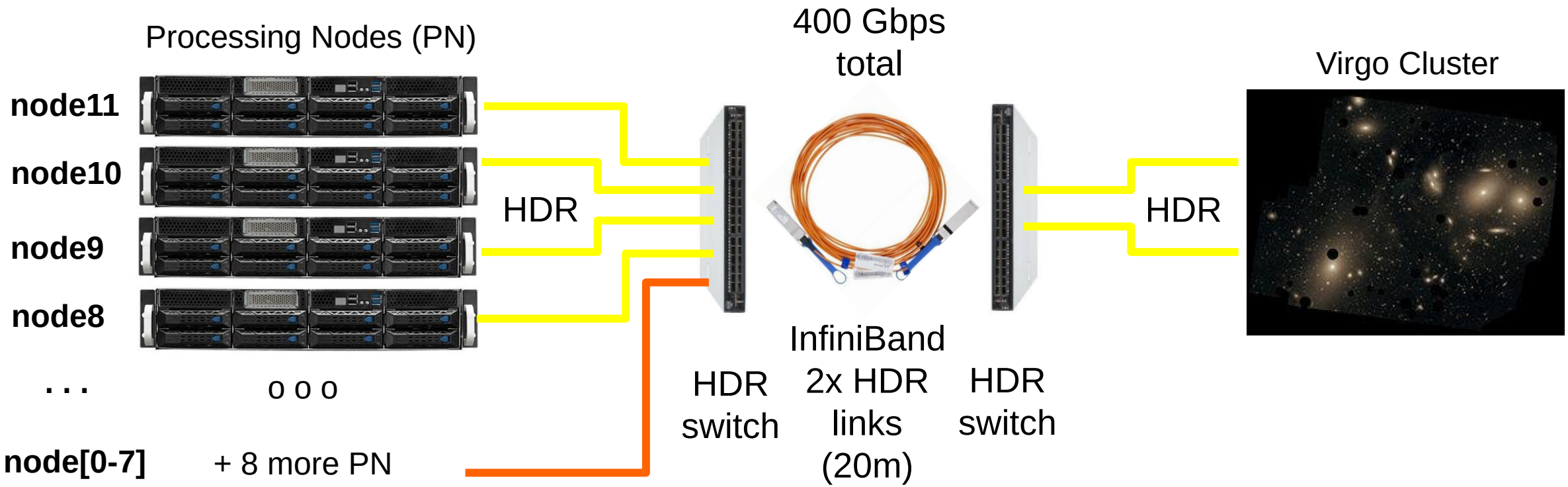
- Using CRI1:47 with 1 input and 47 outputs
- 213 / 47 => 5 TFC-Submaster (max 235)
- 5 / 47 => 1 TFC-Master
- TFC System consisting of 5+1 = 6 CRI

- Using CRI1:24 with 1 input and 24 outputs
- 213 / 24 => 9 TFC-Submaster (max 216)
- 9 / 24 => 1 TFC-Master
- TFC System consisting of 9+1 = 10 CRI

# The link from the mCBM Entry Nodes to the Processing Nodes in the GC

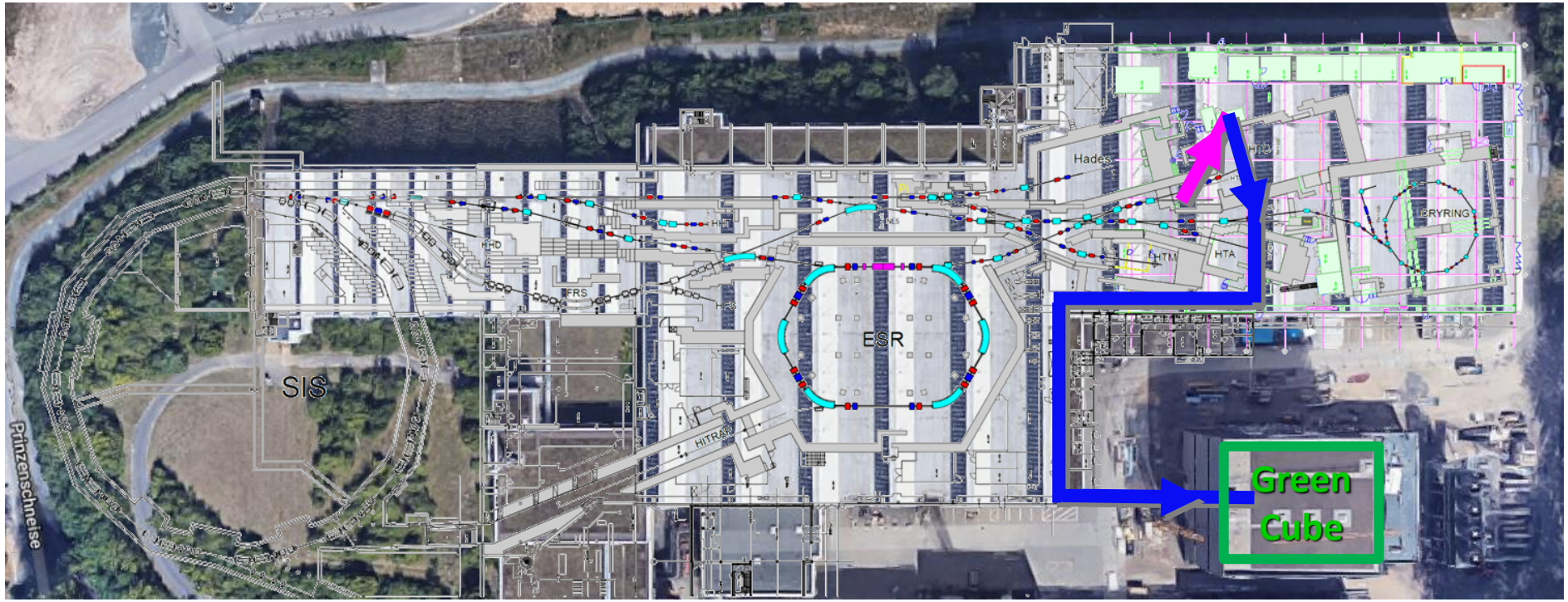


# The link between FLES and Virgo (= GSI IT cluster) inside the GC



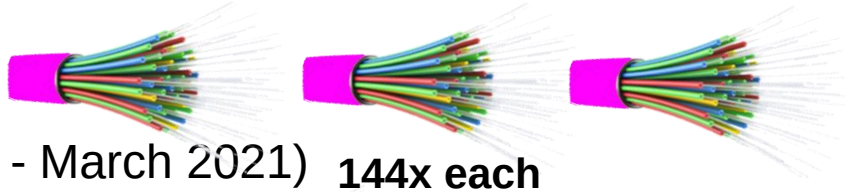

This is the connection of the CBM FLES to the Virgo cluster of the GSI IT.
The links were upgraded to HDR in 2023.

# mCBM optical links to DAQ container and GreenCube




**432x** multi-mode OM4 fibers, **50 m** long:

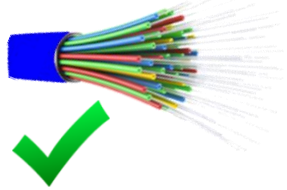
mCBM cave – DAQ container (installation April 2018 - March 2021)



**144x each**

**144x** single-mode OS2 fibers, **300 m** long:

DAQ container – Green Cube (installation in March 2018)



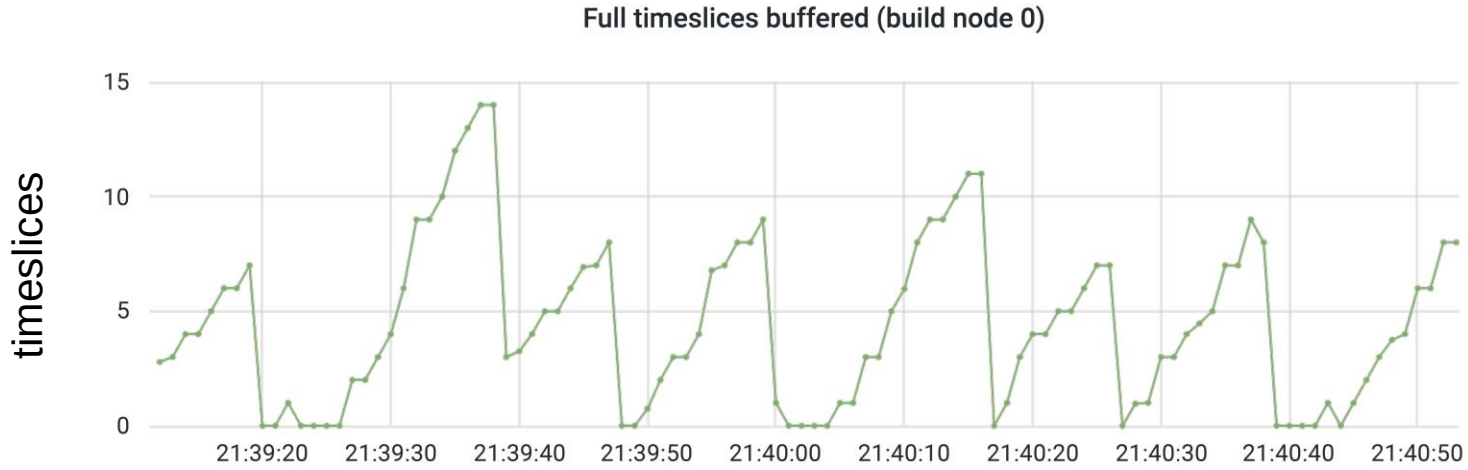
# Matching of fiber lenghts between CBM / mCBM and the GreenCube


We will send our mCBM data forward, backward and forward to bridge a similar distance as later with CBM @ SIS100.

# Data path performance – FLES input and output data rates

FLES input

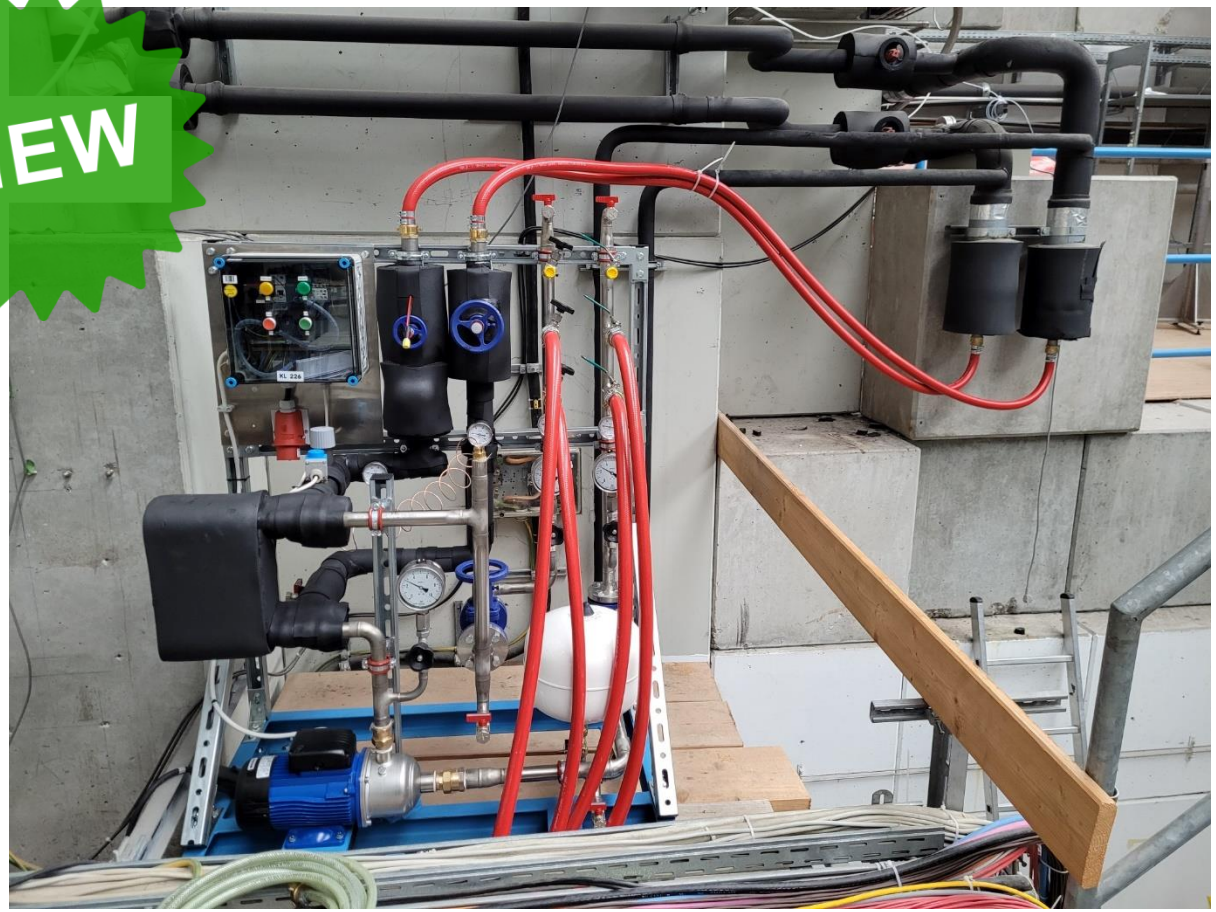

FLES output




run 2448
June 16, 2022
Au + Au, T = 1.23 AGeV
av. collision rate: 300 - 400kHz
av. data rate 2.4 GB/s to disc

# The DAQ rack water cooling system

NEW



Day-1 setup:

- 16 deg C inlet to rack
- 18 deg C return from rack

- Connected to the cooling backbone of the TH hall (03/2024)
- Heat exchanger and pump for secondary circuit cooling two DAQ racks

# The CBM DAQ cooling

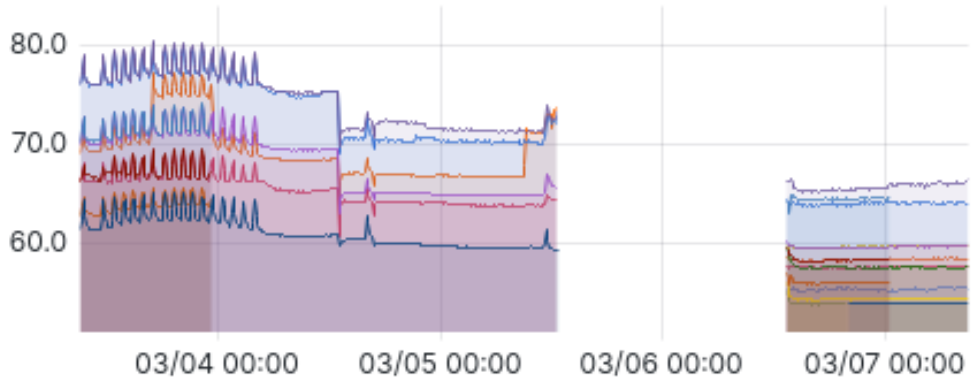
This cooling setup is a **GAMECHANGER**!!!

DAQ operation during any season

- Racks were installed in September 2020, but only air cooled with for the first 3 years
- Cooling configuration in the CBM DAQ identical to racks in the Green IT Cube
- Significantly lower temperature in the container (FPGA <70°C)
- Reduction of noise level due to closed doors in the backside

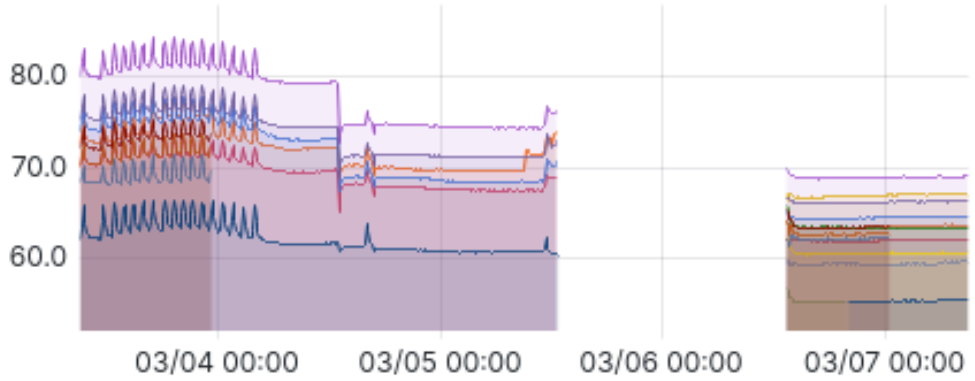
# Temperature difference with water cooled racks

FPGA temperature




| Name | Min | Max | Mean |
|---|---|---|---|
| devel08 0/0 | 53.7 | 55.1 | 53.8 |
| en02 0/0 | 59.3 | 59.8 | 59.6 |
| en02 1/0 | 55.0 | 55.5 | 55.2 |
| en03 0/0 | 58.0 | 77.3 | 66.3 |
| en04 0/0 | 57.4 | 69.1 | 63.2 |
| en05 0/0 | 63.7 | 79.9 | 71.1 |

PEX temperature

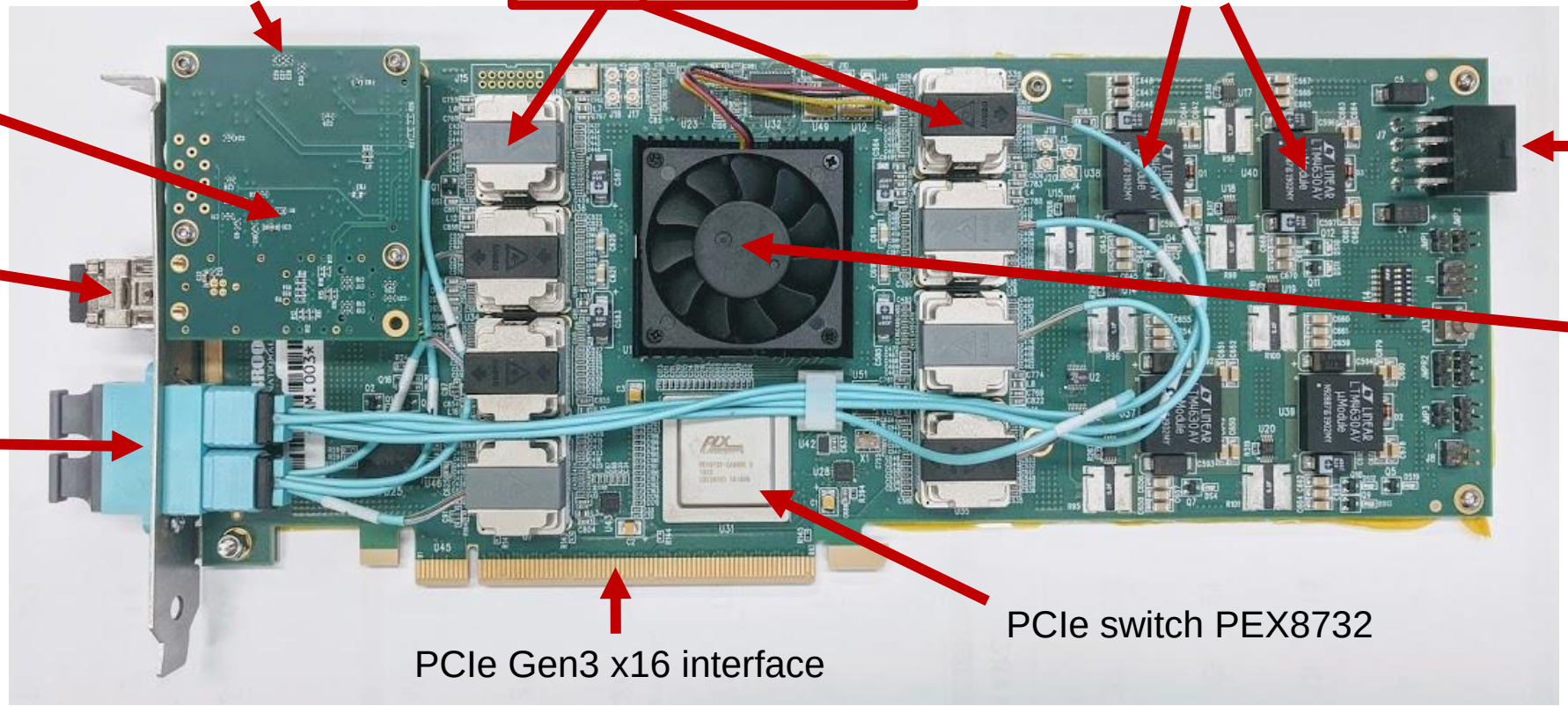



| Name | Min | Max | Mean |
|---|---|---|---|
| devel08 0/0 | 55.0 | 56.7 | 55.1 |
| en02 0/0 | 66.4 | 67.1 | 66.9 |
| en02 1/0 | 59.1 | 59.7 | 59.3 |
| en03 0/0 | 63.2 | 77.7 | 69.7 |
| en04 0/0 | 61.7 | 73.4 | 67.4 |
| en05 0/0 | 64.2 | 77.9 | 69.9 |

> High power load to be tested next week with dry runs

> FPGA temperature reduced to below 70 °C

# The Common Readout Interface card (CRI1) aka BNL-712 v2




- idea: 2018
- production: 2019 -> 2023

- All CBM subsystems are using the CRI1 from 2021 to transfer data into the FLES
- Development of BNL for ATLAS (FELIX)
- CBM owns ~ 32 CRI1 by now (2024)

- Common production with sPHENIX (BNL)
- Some components are EOL since spring 2021
- CBM@FAIR will need **200 pcs** of a successor, the CRI2

# The FELIX family

CRI1

46x FLX-182 boards are in production for ATLAS

FLX-155 passed ATLAS design review on 1st Feb 2024

FLX-711 @2017

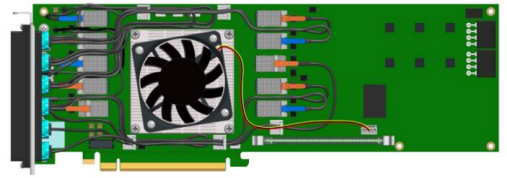
FLX-712 @2018

FLX Demo @2019

FLX-181 @2021

FLX-182 @2022

FLX-155 @2024 (being developed)



PCIe Gen3 x16
48x 10 Gb/s
DDR3

PCIe Gen3 x16
48x 10 Gb/s

PCIe Gen3 x16
25 Gb/s with
varous optical
modules

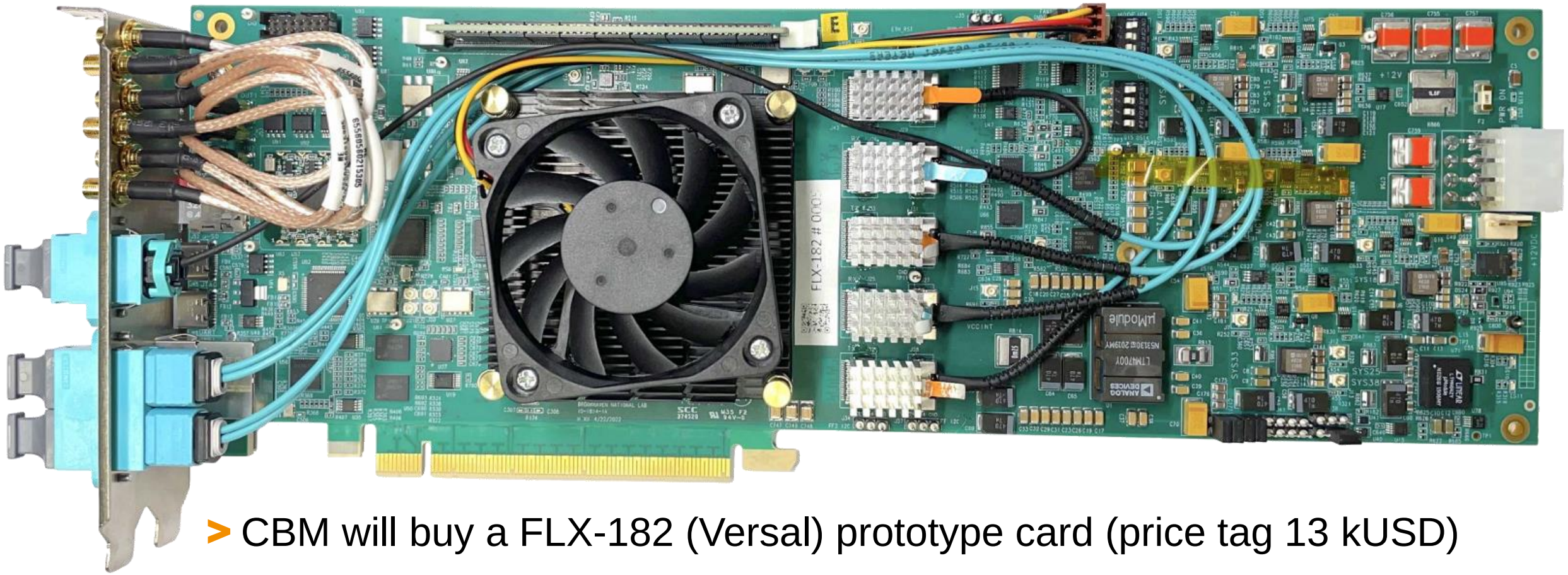
PCIe Gen4 x16
FMC+ 16x 25 Gb/s
12x 16 Gb/s
DDR4
PetaLinux

PCIe Gen4 x16
28x 10/25 Gb/s
100 GbE
DDR4
PetaLinux

PCIe Gen5 x16
48x 10/25 Gb/s
100/400 GbE
DDR4
PetaLinux

> FLX-182 (2023) and FLX-155 (2024) are both an option for a CRI2!

# FLX-182



1x ✓ to be ordered

- CBM will buy a FLX-182 (Versal) prototype card (price tag 13 kUSD)
- The FLX-155 will provide 48x links (= 2x FLX-182) on same footprint
- Both are similar to our CRI2 specs, with faster PCIe and Versal FPGA

# Three CRI2 hardware options


**Common Readout Interface Board (CRI) 2.0 Hardware Specifications**

Version 1.4

DAQ Working Group
October 23, 2023

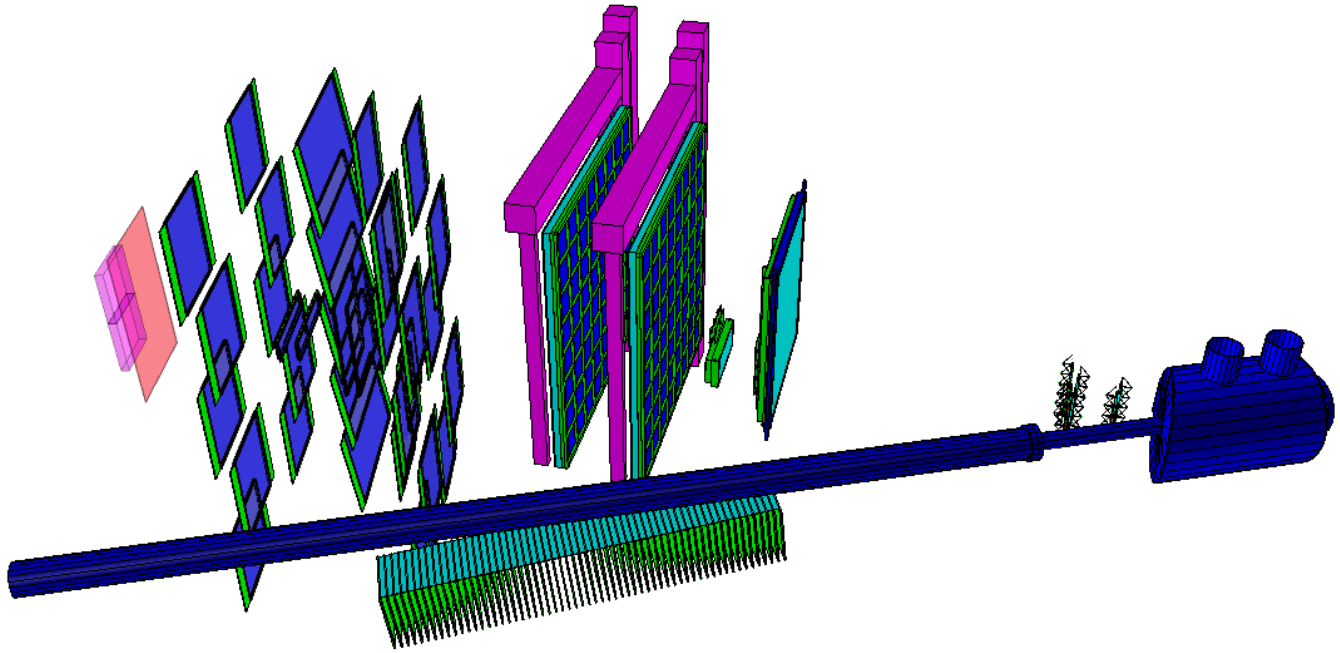
There are 3 options to choose from for a CRI2:

- 1)  – development of a readout board (36 GBT links) according to our own specification
- 2) FLX-182 – development of BNL for ATLAS HL-LHC phase (24 GBT links), CBM will buy 1 board
- 3) FLX-155 – development of BNL for ATLAS HL-LHC phase (48 GBT links), available from autumn 2024
- This is a technology choice, which will define our readout system until ~2035
- CBM needs to pick one of these options

# Summary

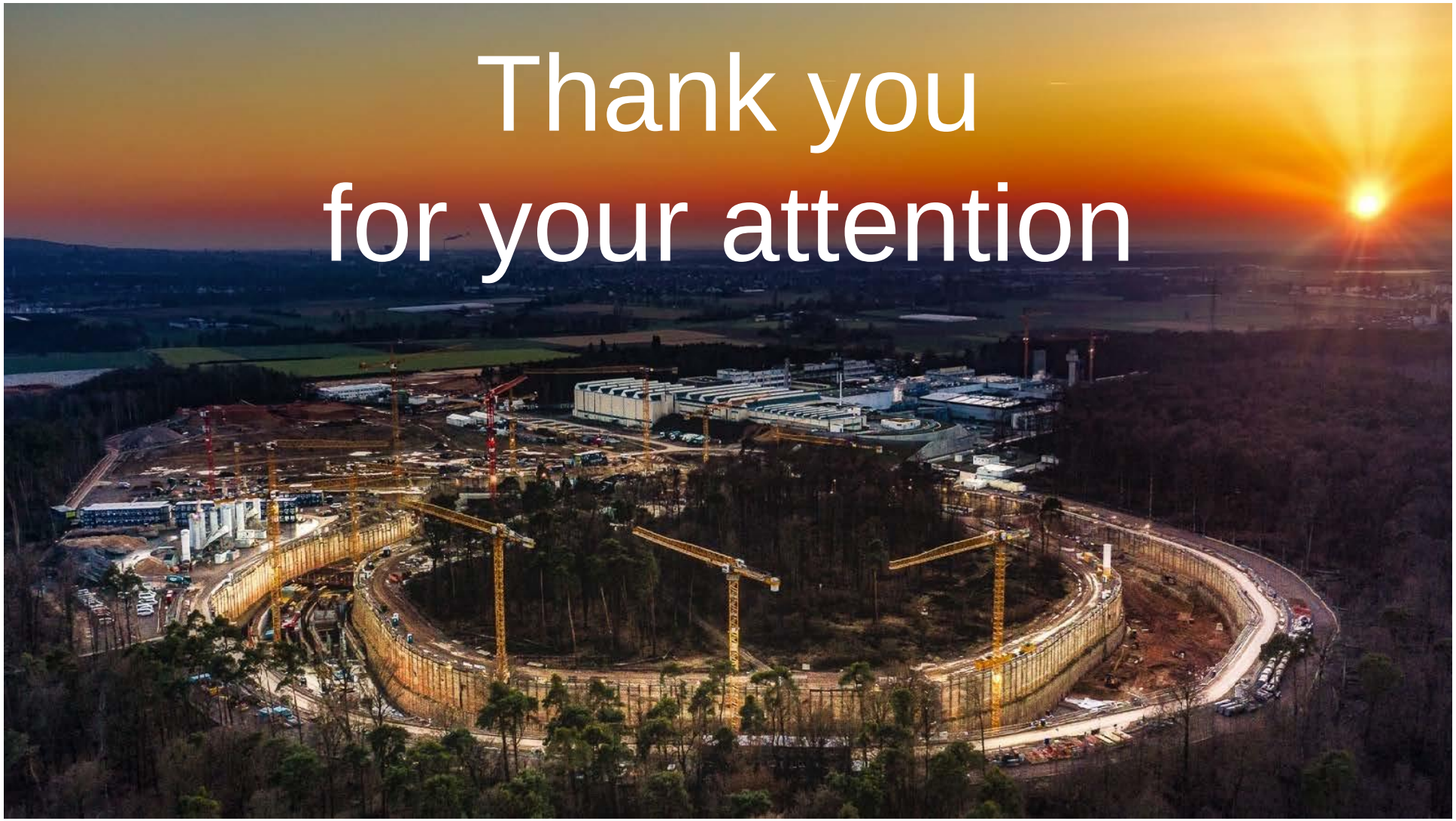
- CBM has setup a prototype DAQ chain over the past ~5 years
- The full data readout and processing chain is being commissioned
- The mCBM setup needs to be scaled up (x20) for CBM @ FAIR
- CBM will require 200x PCIe based FPGA cards (CRI2) for day-1
- A hardware platform for the CRI2 card will be selected till 2025

# The end



Thank you
for your attention

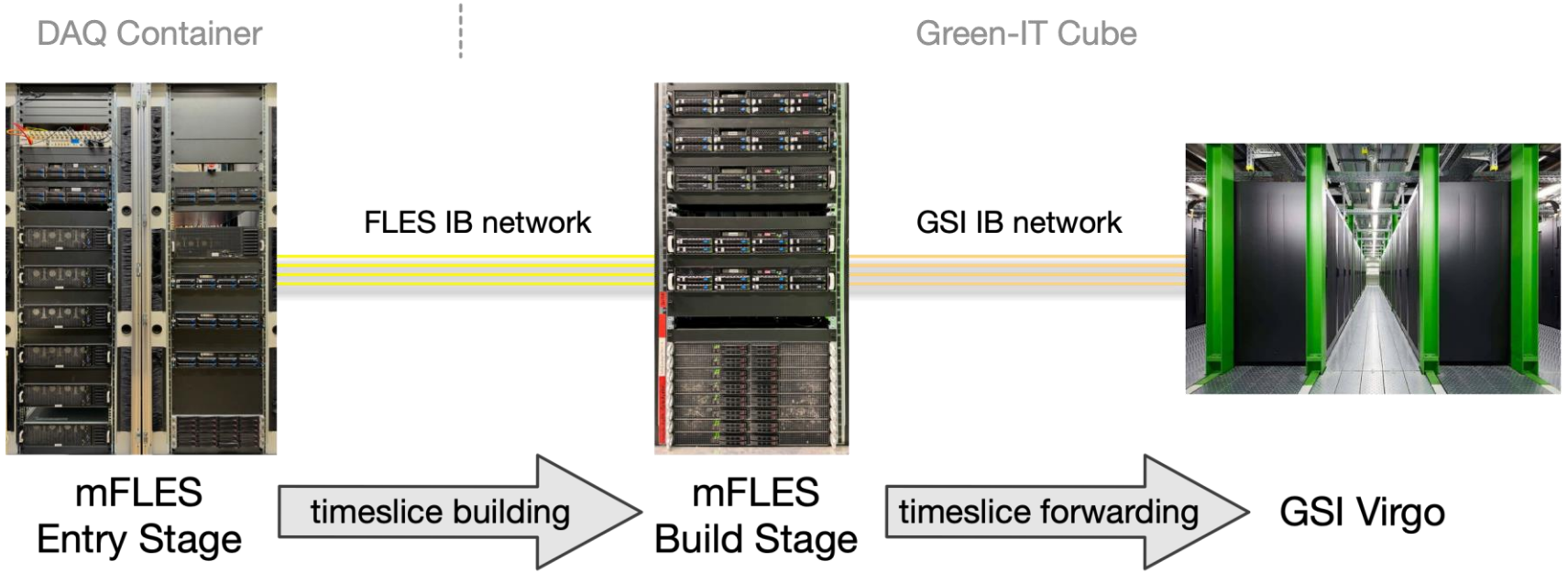
# The end



# Bonus slides

# mFLES Status

- The mFLES setup is our workhorse for FLES development and mCBM
  - Sole readout and control system for mCBM
  - Demonstrator and development platform for FLES software
- Constantly evolving and growing setup
  - First installation in Green-IT-cube in 2012
- Setup includes all key components needed for CBM@SIS100



# mFLES Entry Stage

- Located in the Target Hall mCBM DAQ container next to the mCBM cave
- White Rabbit uplink to GSI machine timing system
- Multimode fiber connection to detector systems
- Two TFC master nodes
- 6 entry nodes with a total of 12 CRIs
- 300m long-range InfiniBand connection to GC
  - 800 GBit/s bandwidth
- Nodes can work in two modes:
  - Stand-alone development (develXX)
  - FLES clusters node (enXX)

# mFLES Build and Processing Stage

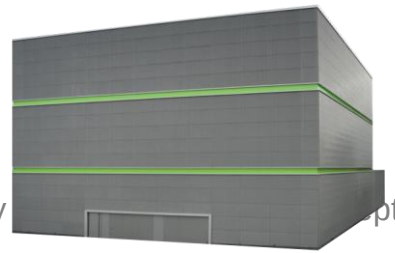
- Build/Processing stage in Green-IT cube
- Heterogeneous setup
  - 8+4 local processing nodes
  - up to 32 cores/64 threads, 256 GB RAM
  - Head node for infrastructure services and login (cbmfles01)
- Local buffer storage for data recording
  - 56 TB fast NVMe SSD buffer
  - 320,8 TB HDD buffer
- 200 GB/s InfiniBand HDR to mCBM and Virgo
  - Application level routing between separate IB fabrics
- Local online processing or timeslice forwarding to Virgo cluster

2nd gen Processing Nodes with local storage

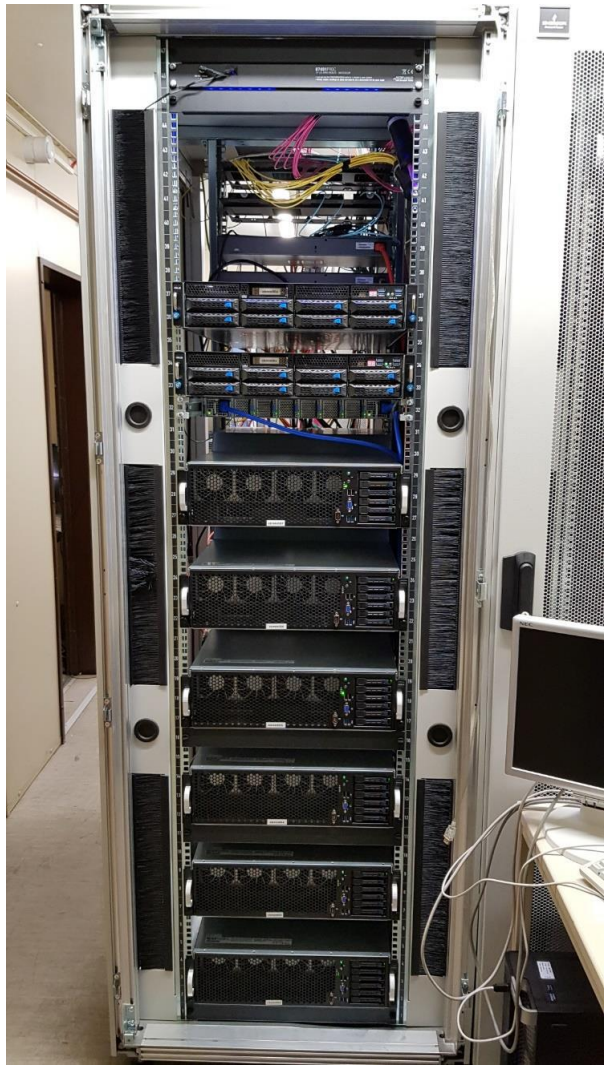
Head Node

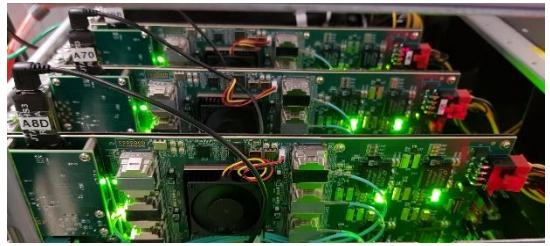
2nd gen Processing Nodes with local storage

1st gen Processing Nodes

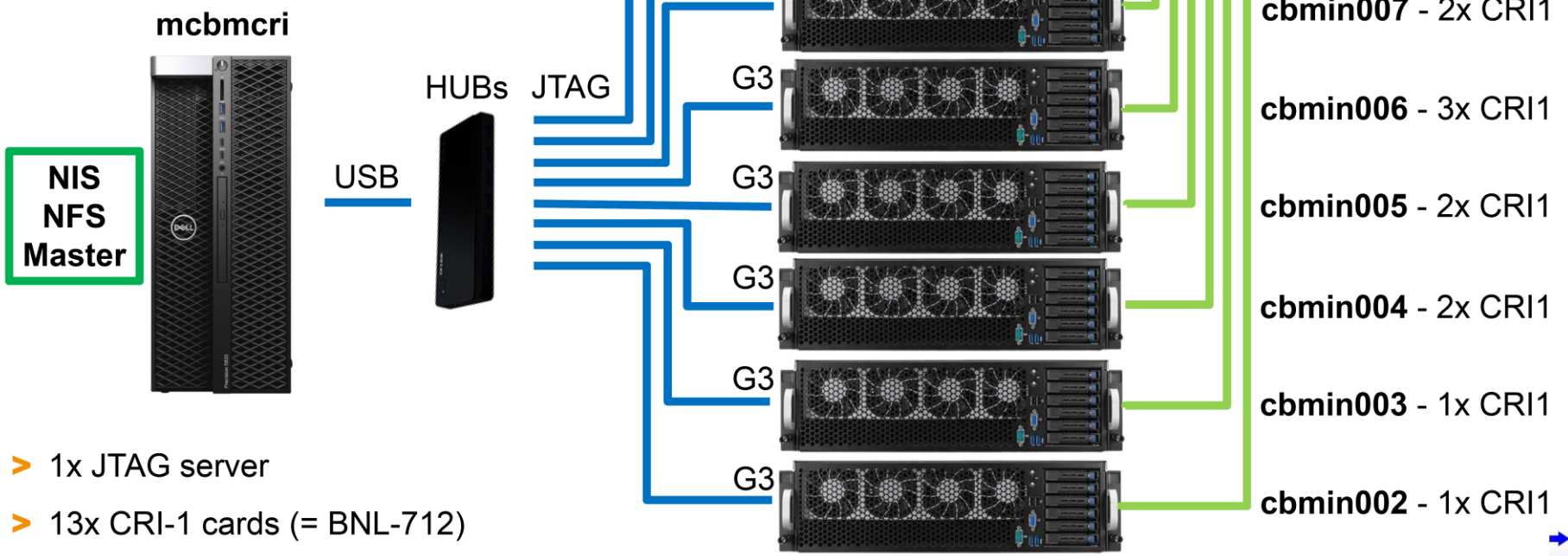


# DAQ / Data Transport using CRIs





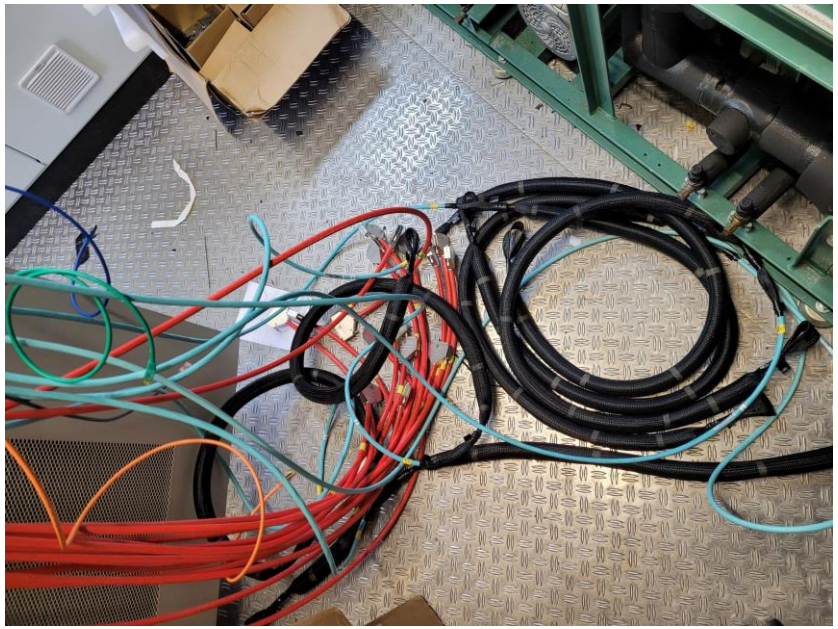
- This is the CBM DAQ prototype rack
- 1x Timing and Fast Control (TFC) server
- 6x Entry Nodes (EN)




- 1x JTAG server
- 13x CRI-1 cards (= BNL-712)

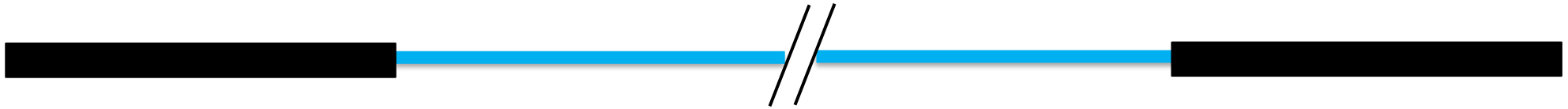
# Single mode fiber (as used between mCBM and Green-IT-Cube)

**144x core trunk fibers for mCBM**



**six 144x core trunk fibers in C17**

300 m long OS2 cable: mCBM to Green-IT-Cube
about 1/3 of the length required at SIS 100 / FAIR



endpiece: ~ 1m long
up to 8 cm in diameter

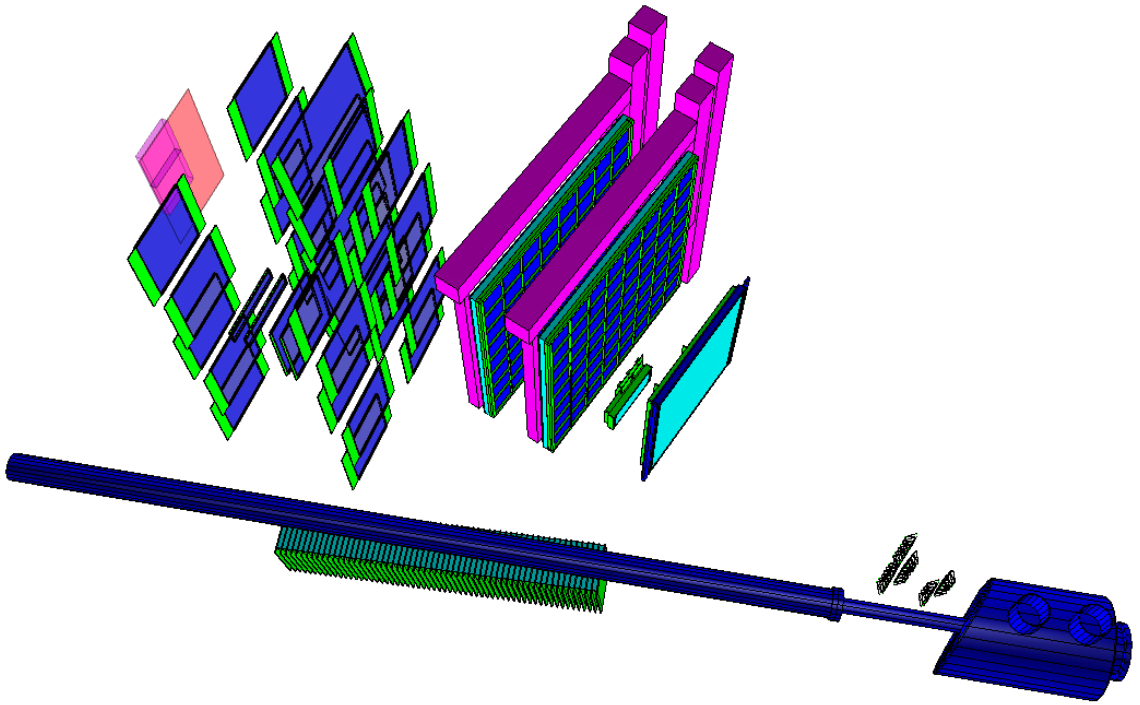
fiber: about 1000m long
2 cm in diameter

endpiece: ~ 1m long
up to 8 cm in diameter

# FAIR construction site – October 2021


FAIR construction site – October 2021

# The end



Thank you
for your attention