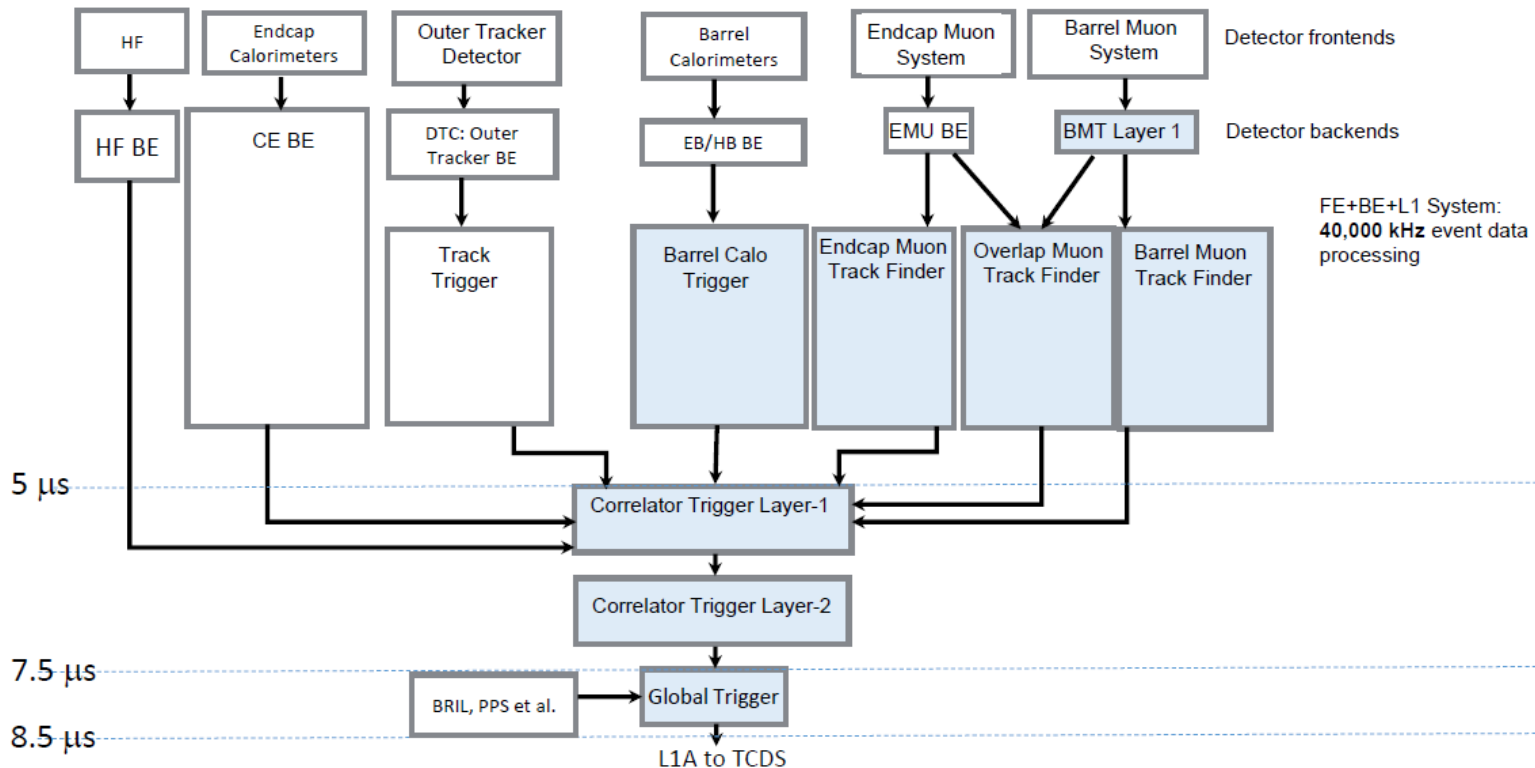


# Phase-2 Level-1 Trigger Architecture Options

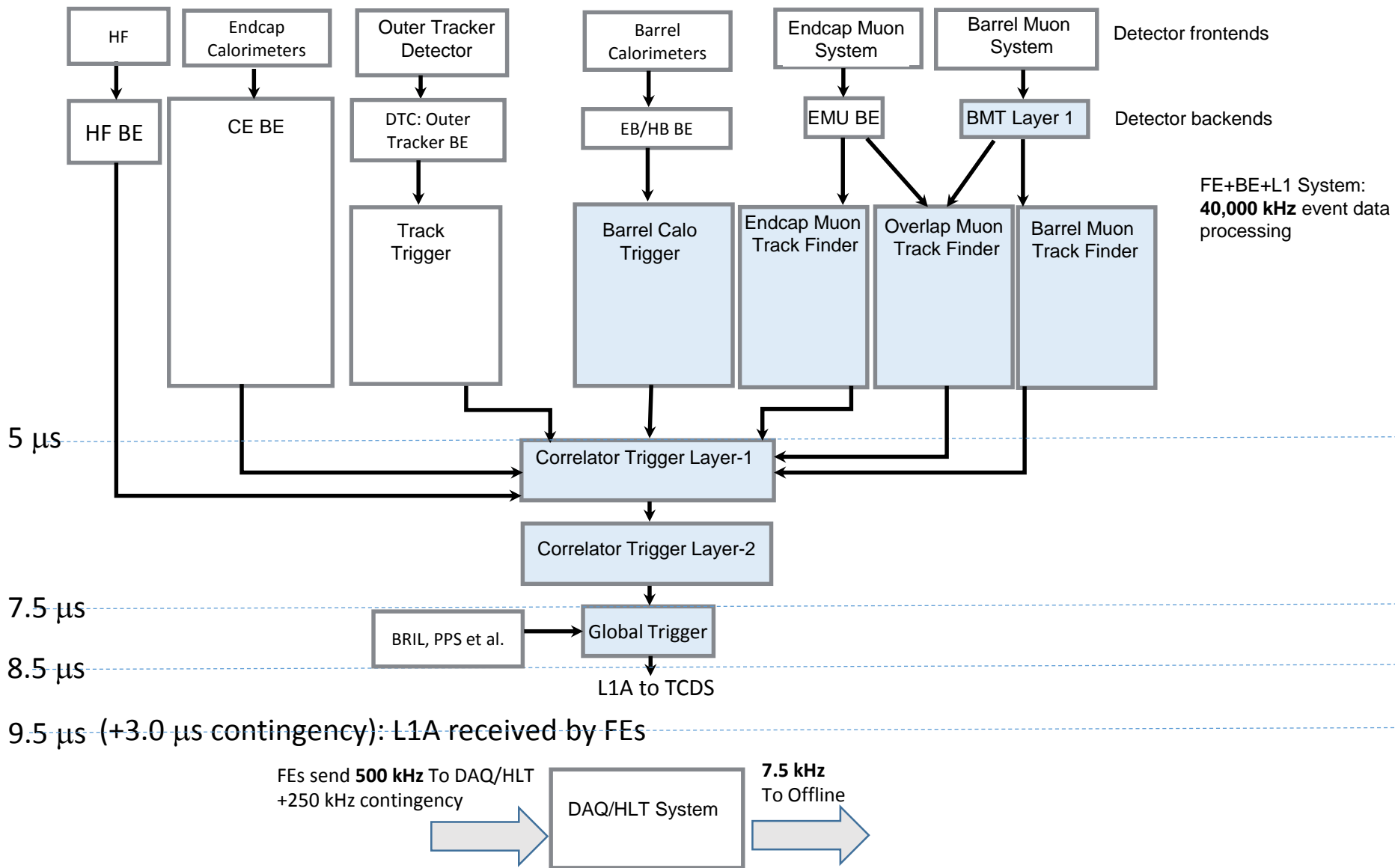
Jeffrey Berryhill, Fermilab

## Phase-2 Muon/Trigger Workshop



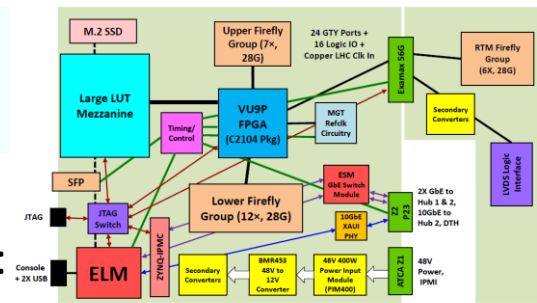
# CMS HL-LHC Readout and Trigger Electronics

iTDR design

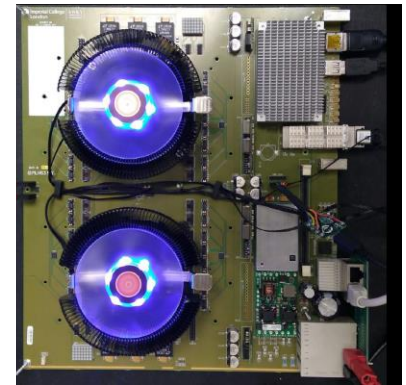


# Building blocks

- Processing units come in two different FPGA form factors:
  - 96-link** data in/out packages (C2104) supporting **16-28 Gbps**:  
e.g. Xilinx VU9P (VUxP supports up to 96 32-Gbps)
  - 64-link** data in/out packages (B2104) supporting **16 Gbps**:  
e.g. Xilinx KU115 (KUxP can support up to 32 32-Gbps)
- Different speed grades are available in each case (~30% speed boost available)
- Different logic resources are available in each case (~2X more DSPs)
- ATCA Boards in our R&D program can accommodate **one (Apx) or up to two (Serenity) such chips per board**. Two-chip boards can daisy-chain resources with low latency, if required.
- Each **ATCA crate houses up to 12 ATCA boards** + 2 slots for DAQ, each rack houses 2 crates
- For a given throughput, data can be typically organized into geometric regions, or different time slices (TMUX), or both.



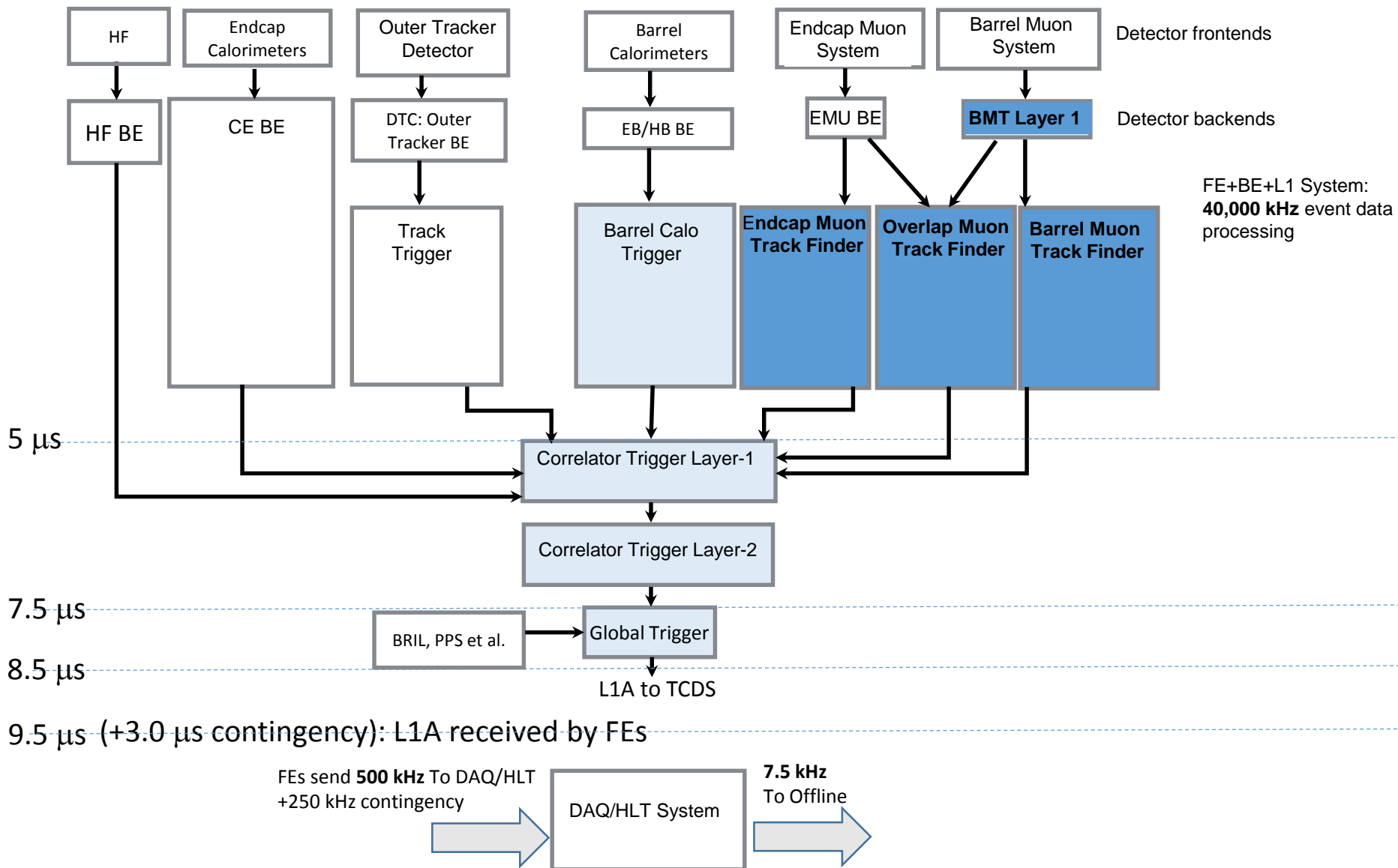
Note: Some connections omitted to enhance clarity



TMUX\*Regions is  
~invariant amount of  
computing power

# Muon Trigger System

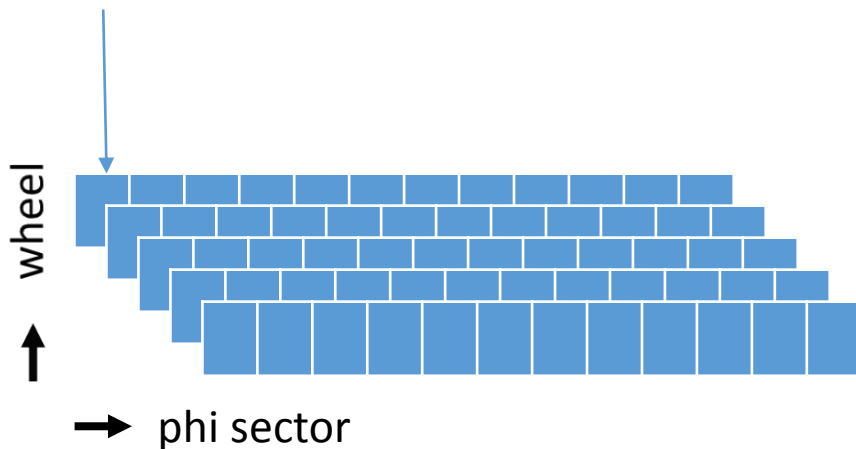
iTDR design



# Barrel Muon Trigger layer 1

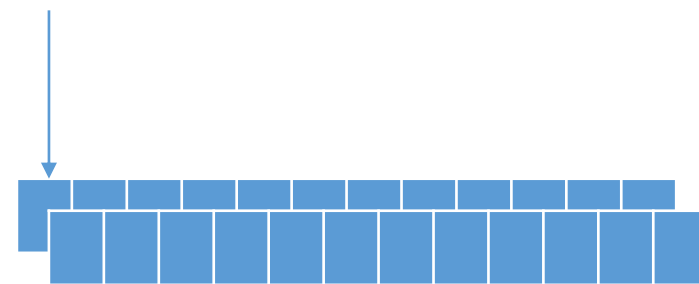
- DT chambers organized in 12 phi sectors \* 5 wheels
- DT FE cards deliver **48 links of phi view** at 10 Gbps per sector per wheel
- **RPC/HO deliver 6 links** at 10 Gbps per sector per wheel
- DT chamber phi view + RPC/HO can be sent to **60 64-link cards (1/sector/wheel)**
- DT chamber theta view can be sent to **24 64-link cards (1/sector/2.5 wheel)**

48+6 links in for 1 sector \* 1 wheel



BMT Layer-1 phi view

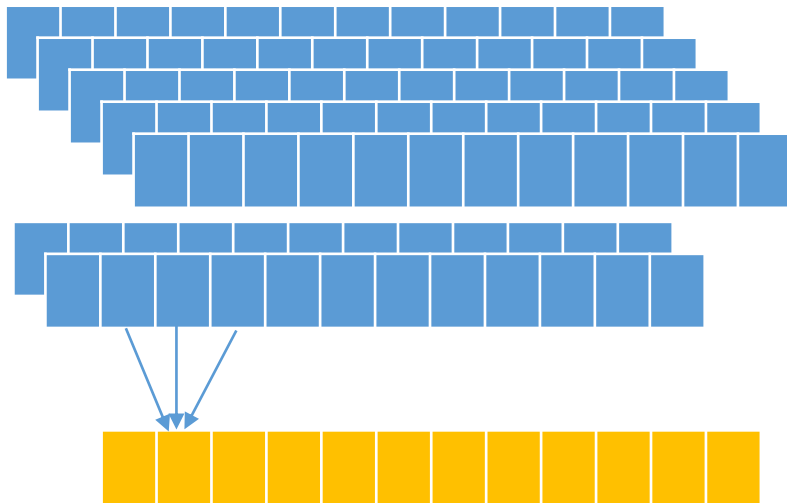
30 links in for 1 sector \* 2.5 wheels



BMT Layer-1 theta view

# BMT layer 1 output to layer 2 (track finders)

- Assume that each stub is 2x the size of current one= 64 bits and we produce 4 instead of 2 stubs per chamber
- Each sector could make 4 stubs x 4 chambers = 16 stubs
  - Therefore 16 stubs x 64 bits = 1024 bits /BX/sector = 4 16G links/sector
- **iTDR scenario:** Layer 2 performs SA barrel muon track-finding, TMUX = 1
  - A geometric partitioning of 5 wheels\*3 phi spanning **12 64-link boards** meets requirements



BMT Layer 1 phi view: 60 boards

BMT Layer 1 theta view: 24 boards

BMT Layer 2: 12 boards

BMT Layer-2: 12 64-link boards

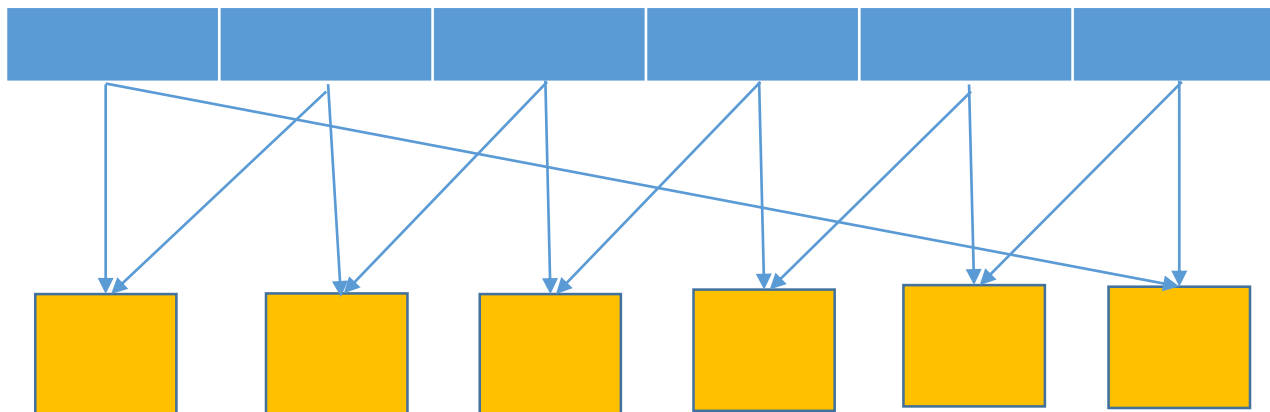
1 board sees all 5 wheels, phi+theta view, 3 phi sectors (=1 sector + nearest neighbors)

# EMTF Input

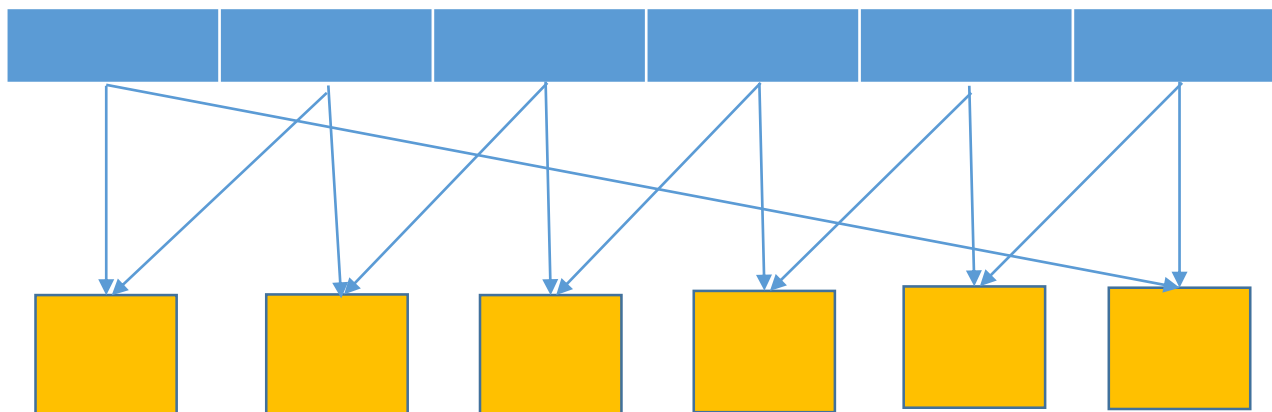
- Prior assumptions, per 60° sector, with neighbor sharing (+N):
  - **CSC:**
    - 3.2 Gbps links from legacy MPCs (40+9) ← Assumes no CSC data concentration
  - **RPC endcap via CPPF/New RPC electronics:**
    - 12.5 Gbps links from LBs + RE1/3+RE2/3 (8+2) ← Assumes link concentration for RPC as current CPPF
  - **iRPC forward:**
    - RE3/1: 10 Gbps links (3+1) ← iRPC electronics will be distinct from RPC Link board replacement
    - RE4/1: 10 Gbps links (3+1)
  - **GEM:**
    - ME0: 10 Gbps links (3+1)
    - GE1/1: 10 Gbps links (8+2)
    - GE2/1: 10 Gbps links (9+2)
  - **DT:**
    - 10 Gbps links, 1/30° for MB2/1+MB2/2 (2+1) ← (Provision to possibly cover overlap with barrel)
- **Summary:**
  - 95 links per 60° sector [Includes sharing]
  - 12 cards total, ~1100 total input links for ~7 Tbps bandwidth
  - In a TMUX =1 arrangement, 12 96-link cards can cover (2 endcaps \* 6 sixty-deg sectors)

# EMTF

Endcap system Backend, organized in 2 endcaps and six 60deg sectors, 1100 links



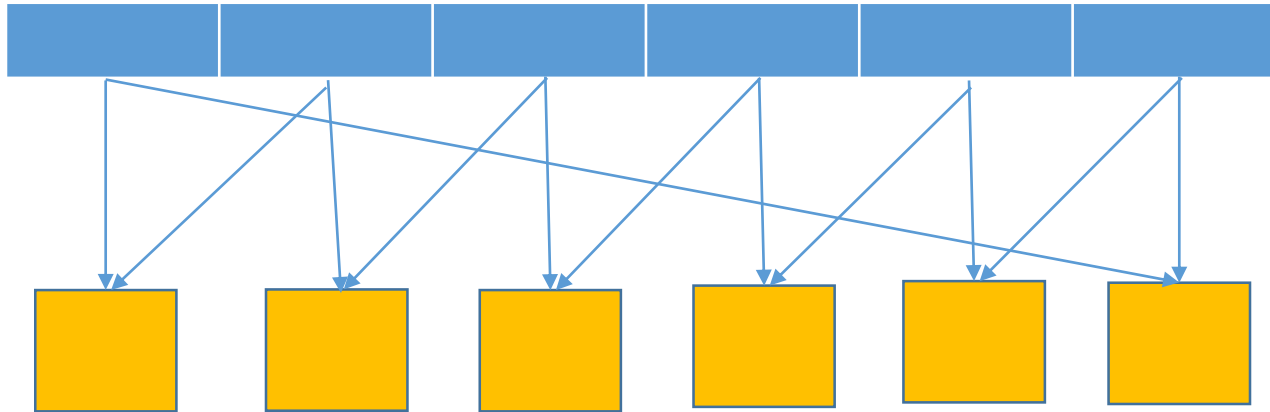
EMTF: for each endcap, six 96-link boards consuming one 60 deg sector + nearest neighbor ( $2 \times 6 = 12$  boards total).



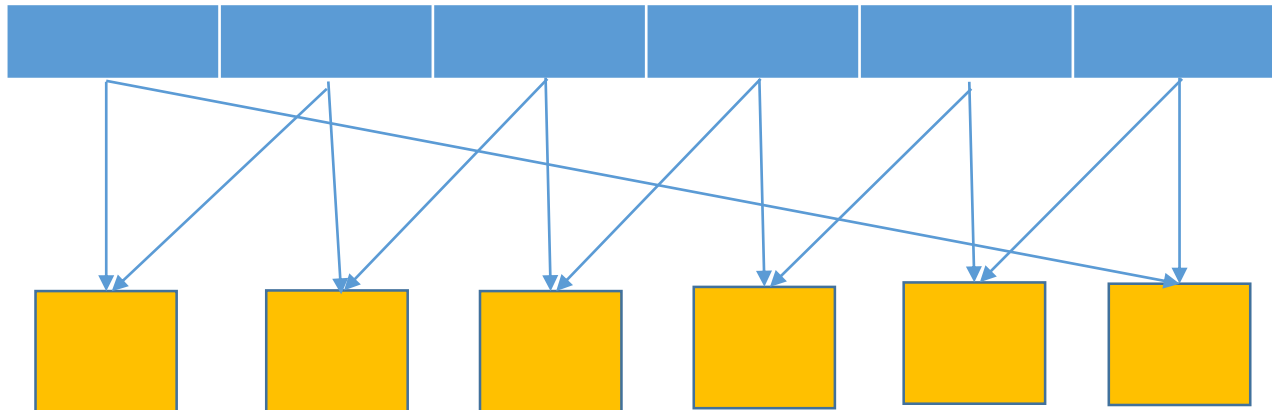


# EMTF/OMTF

Endcap system Backend, organized in 2 endcaps and six 60deg sectors, 1100 links

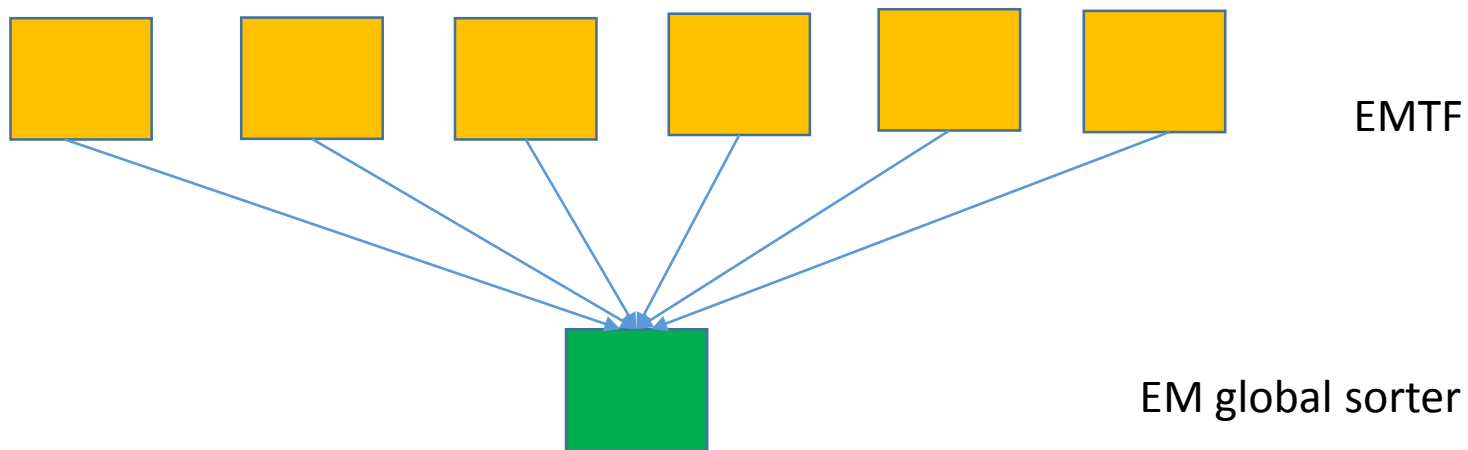


In this scenario, OMTF is realized as part of each EMTF board FW. OMTF functions could also be realized as a few physically separate boards.



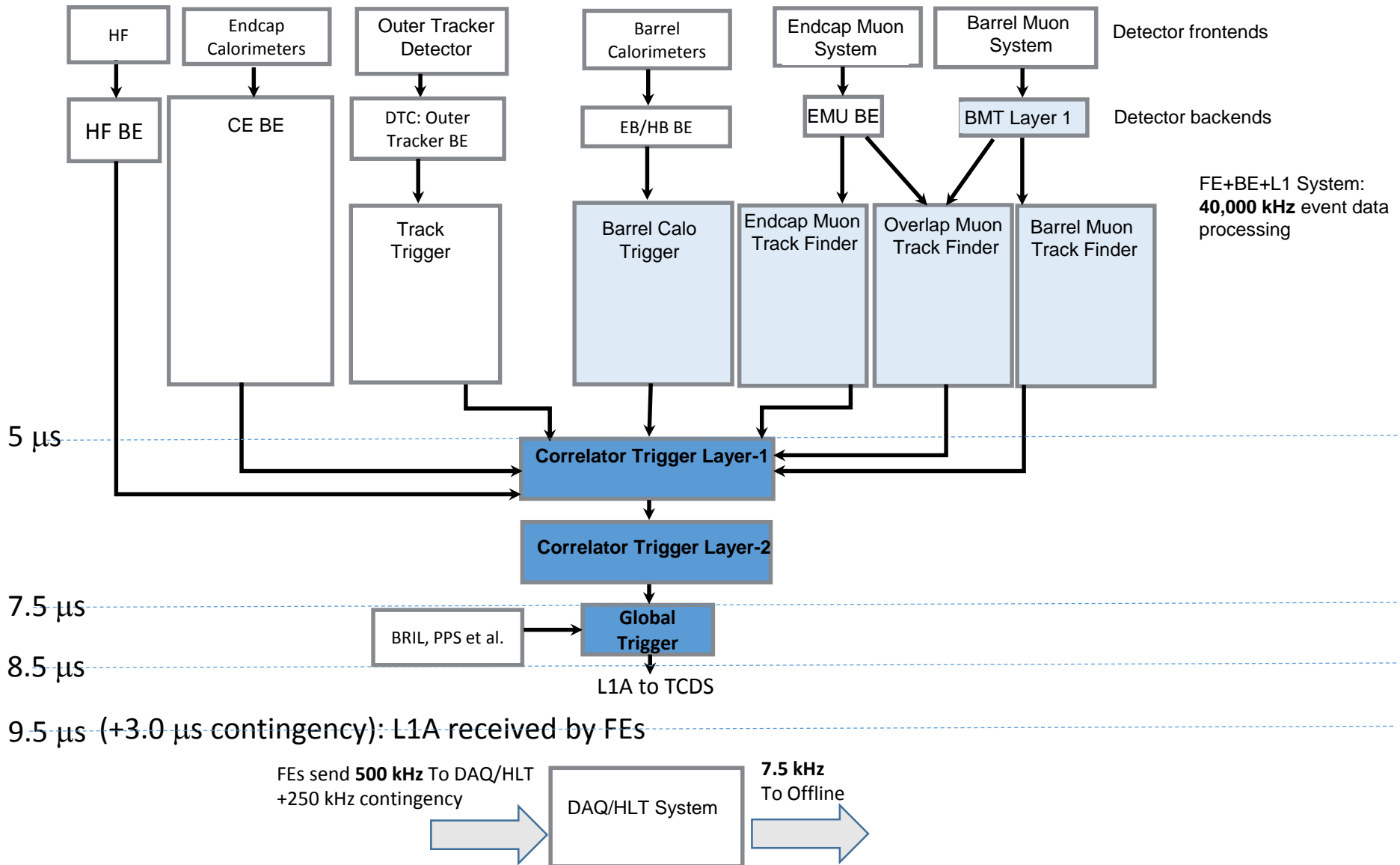
# Muon sorting/duplicate removal

- In a regionalized track finding design, there will be duplicate tracks in neighboring regions
- To limit output payload downstream, there can also be an advantage to PT or quality sorting
- In Phase 1 this was accomplished with the GMT and the endcap global sorter
- 1-2 ATCA boards receiving BMTF/OMTF/EMTF output provide the same function
- Alternatively these functions could be deferred to the correlator.

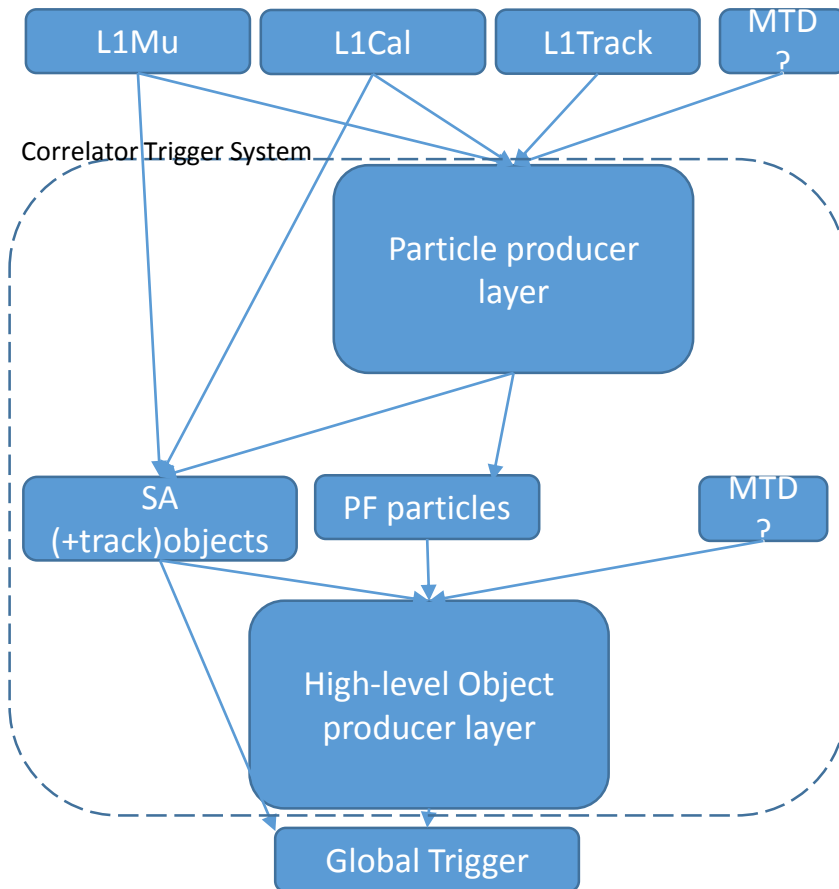


# Correlator and Global Trigger System

iTDR design



# Correlator Functional Diagram



- Particle flow candidate creation
  - Track-muon match
  - Track-cluster match
  - PF assignment
- SA track objects
- Vertex reconstruction
- SA calo-track matching
- **SA muon-track matching**
- PUPPI weighting/vertex assignment for all of the above
- Early MTD timestamping of all charged particles?

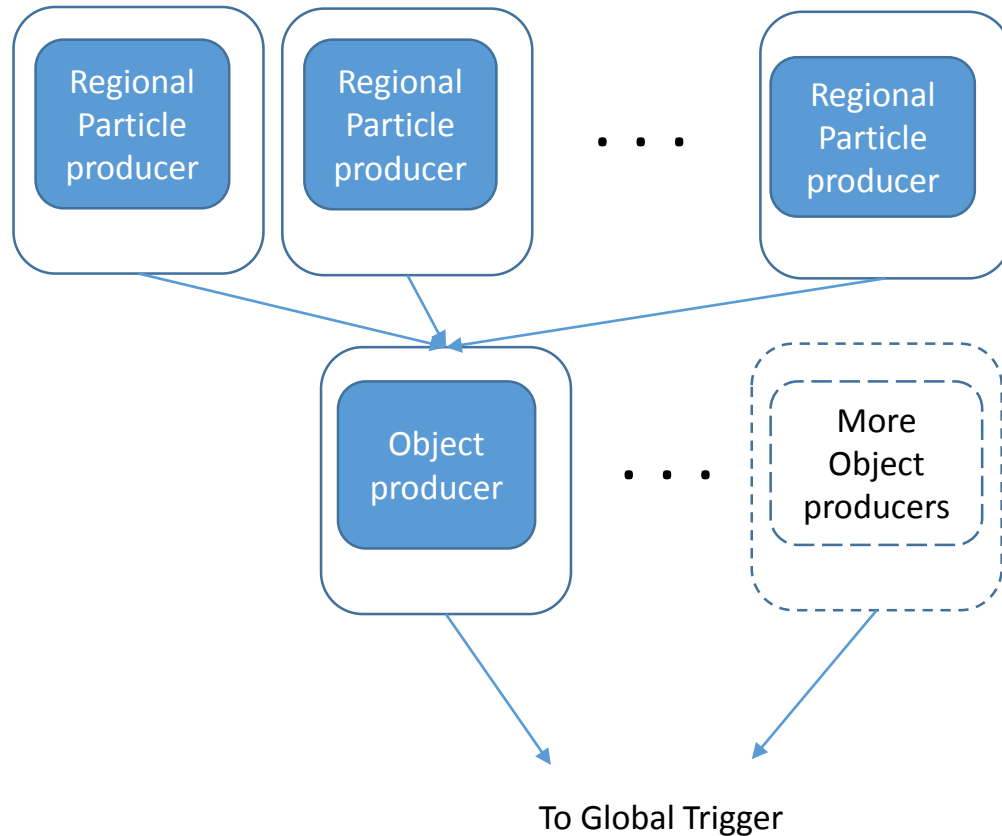
- Data volume is predominantly tracks and clusters for PF
- Potentially large data reduction factor from discarding PF candidates of “no interest”

- PF jets
- PF taus
- PF sums
- **PF isolated photons/leptons**
- **Track-isolated photons/leptons**
- **Displaced/OOT objects**
- **Non-trivial composites of above (jet-tagging)**
- **Late MTD cleanup, late pixel matching w/ROIs?**

- Functional division of labor possible in HW
- Need a lot of flexibility/extensibility here

# Correlator Trigger Physical Partitioning

TMUX1 time slices



2 physical layers

TMUX=3,6,12,18 with regionalization possible

Variable number of object producers/time slice  
Object producers could see entire PF and SA content  
Object producers upgradeable  $\leq$ TMUX2 boards at a time.

# Correlator Layer 1 design

- Some possible solutions for TMUX 6 and TMUX 18 shown below. TMUX 9 can also work in this scenario.

	V1	V2
<b>TMUX</b>	<b>6</b>	<b>18</b>
<b>Total FPGAs</b>	<b>24</b>	<b>36</b>
<b><math>\Phi \times \eta</math></b>	<b>2 x 2</b>	<b>1 x 2</b>
<b>Payload</b>	<b>20000 (PF+P), 40000 (PF Calo)</b>	<b>20000 (PF+P), 40000 (PF Calo)</b>
<b>link (Rx) / FPGA</b>	<b>49</b>	<b>35</b>
<b>link (Tx) / FPGA</b>	<b>2, 3</b>	<b>2, 3</b>

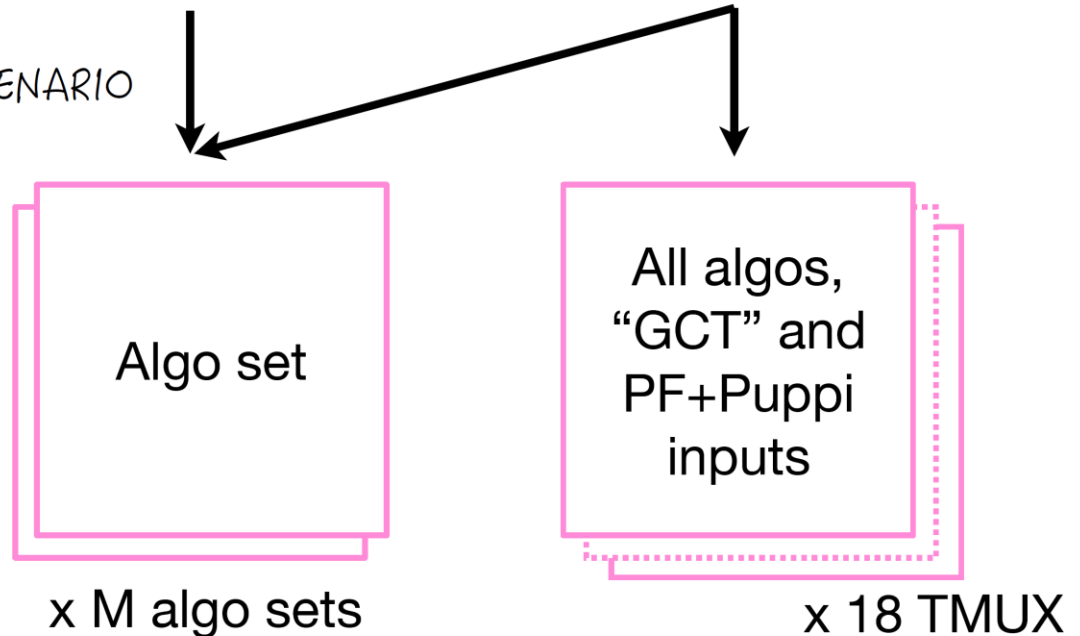
# Correlator Layer1 to Layer 2

	V1	V2
<b>TMUX</b>	6	18
<b>Total FPGAs</b>	24	36
<b><math>\Phi \times \eta</math></b>	2 x 2	1 x 2
<b>Payload</b>	20000 (PF+P) , 40000 (GCT)	20000 (PF+P) , 40000 (GCT)
<b>link (Rx) / FPGA</b>	49	35
<b>link (Tx) / FPGA</b>	2, 3	2, 3

LAYER 1 TO LAYER 2 SCENARIO

At high Layer 2 TMUX, just one layer 2 board per time slice is possible

If Layer 2 TMUX is small, or downshifted from Layer 1, can have some functional division of labor per time slice

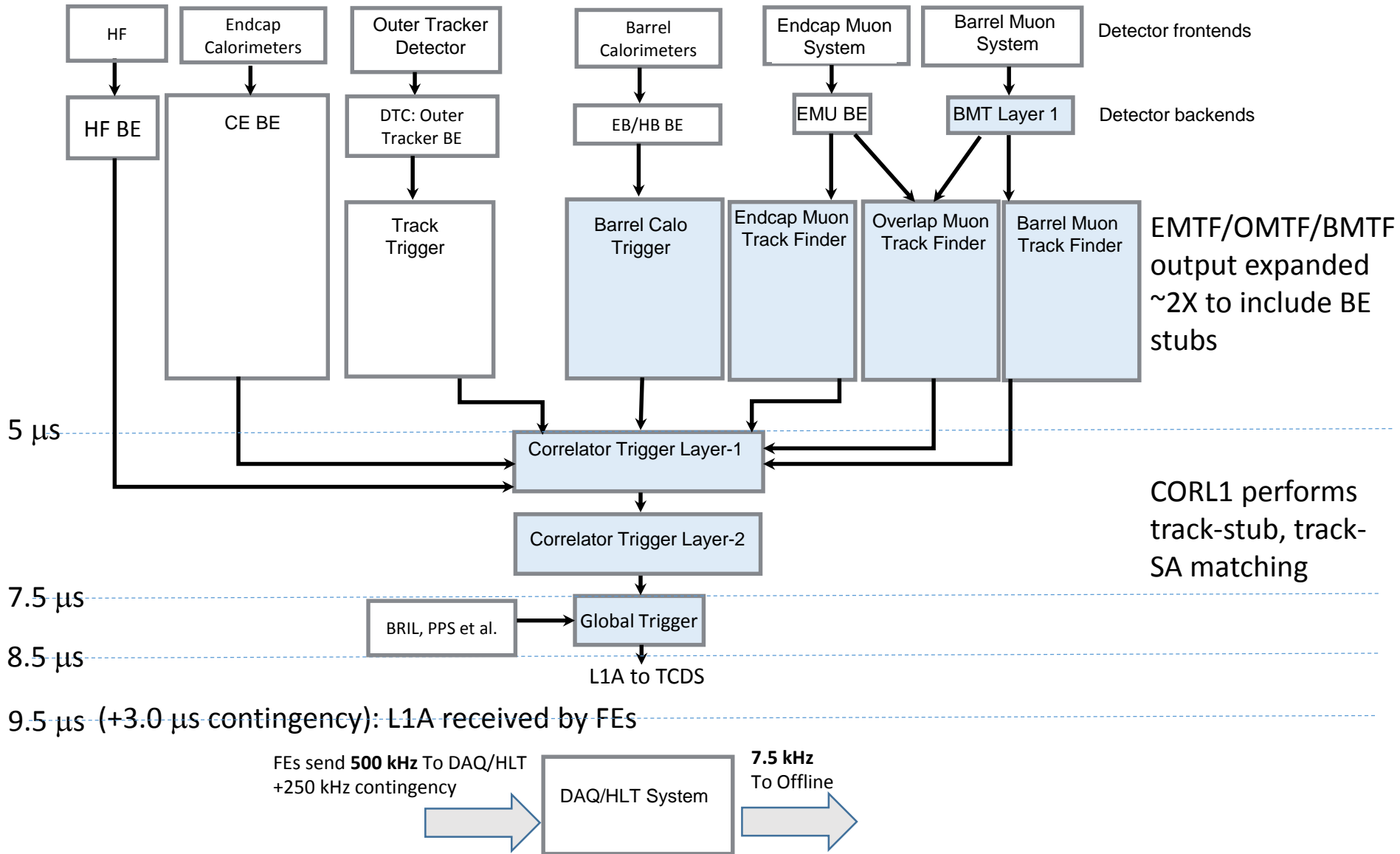


# Beyond iTDR muon scenarios: track-muon matching

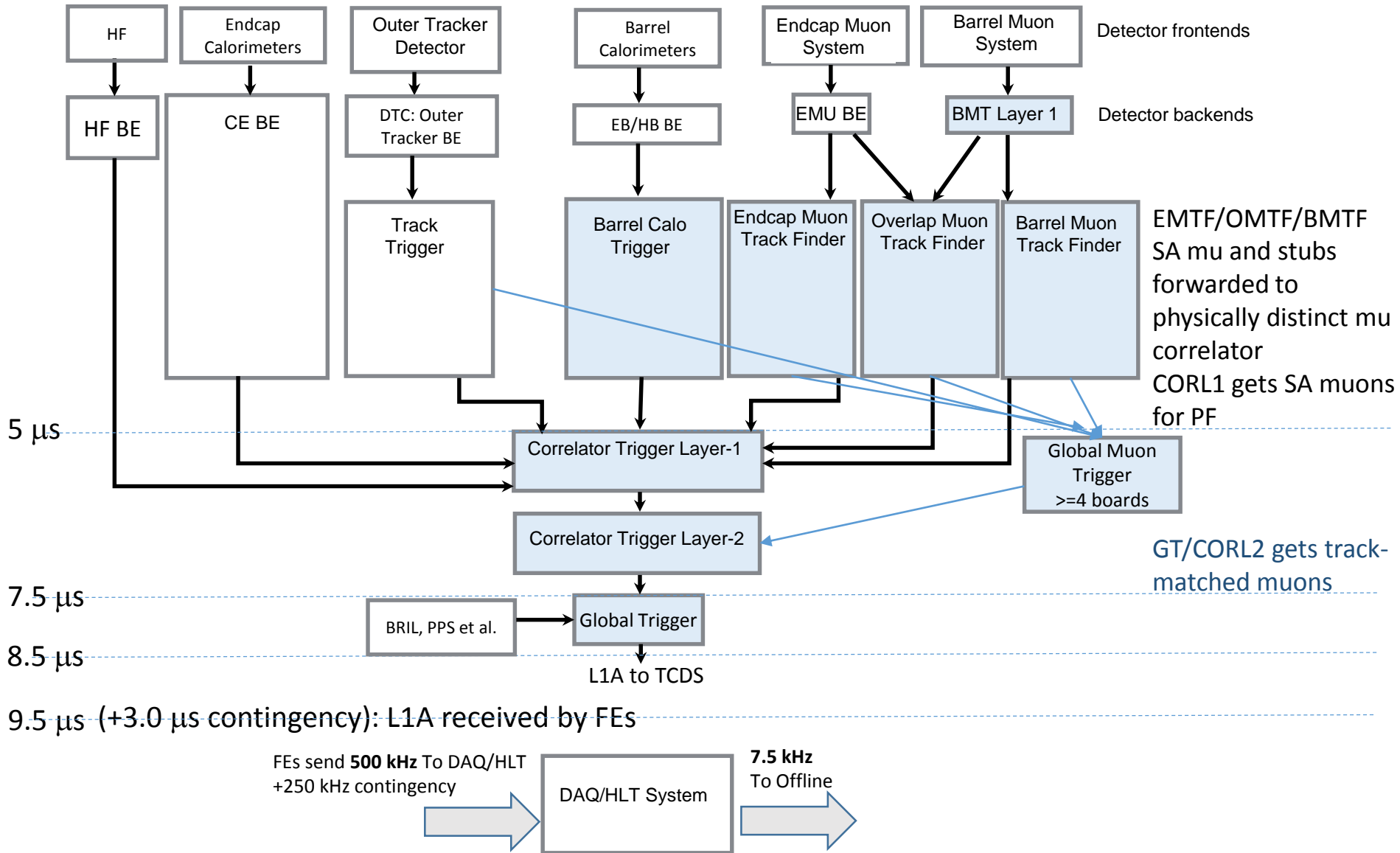
- There are scientific advantages to **matching muon stubs with tracker tracks**.
  - Higher efficiency than tracks+SA muons, with acceptable rate
  - More robust L1 muon reconstruction (ensures against muon chamber aging or downtime)
- There are potential operational benefits to **physically partition track-muon matching** from the correlator system
  - Muon triggers ultimately may have no/minimal calorimeter dependence
  - Track-muon matching is an additional firmware burden on the correlator system
  - A separate physical path to GT which does not go through the correlator
  - Similar arguments can be made for partitioning other non-PF functions of the correlator system (SA track objects, SA calorimeter objects)
- We are exploring scenarios that introduce one or both of these features
- Request from CE BE to process endcap muon candidates for MIPs
  - Requires EMTF SA or track-matched muons interfaced to CE BE to help determine DAQ readout upon L1A



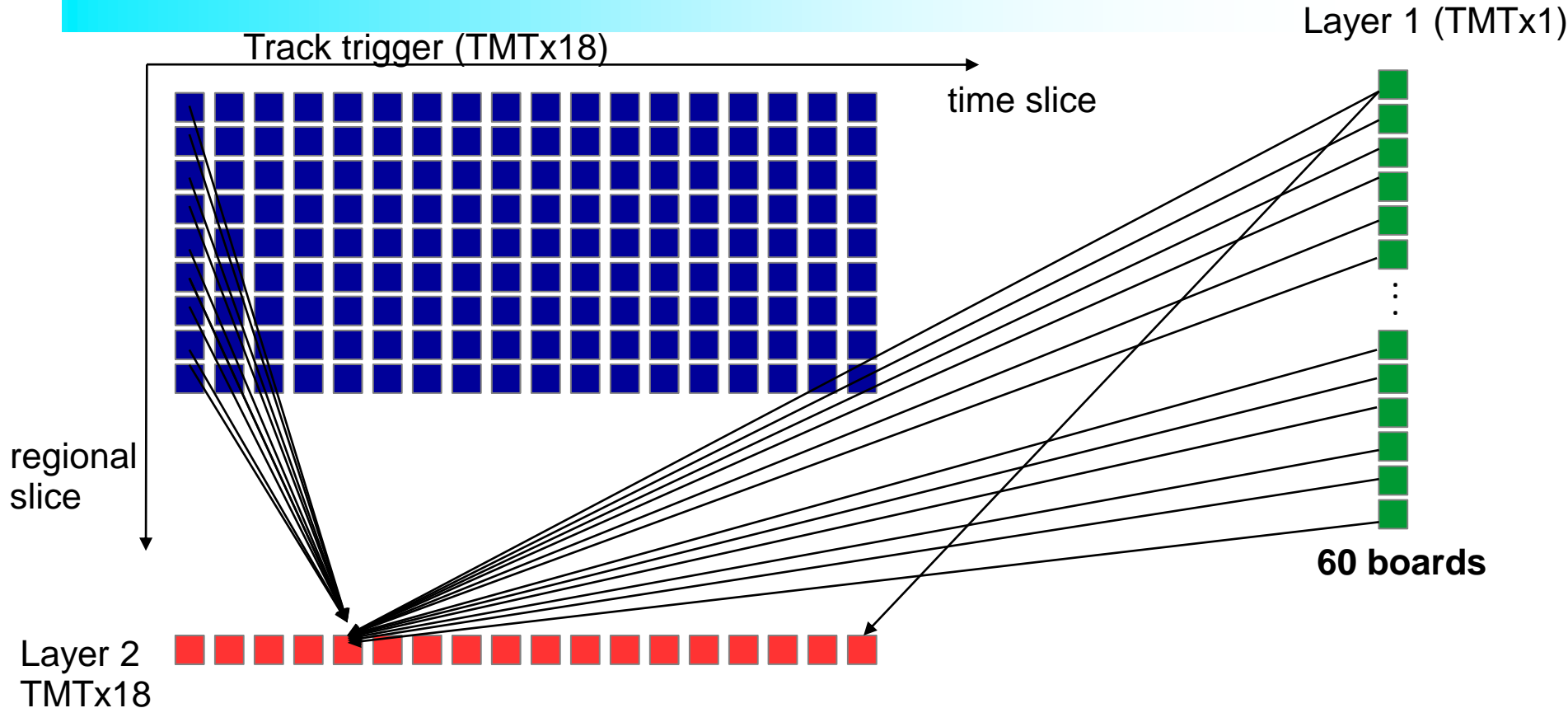
# Muon-Track Matching Scenario A: defer to CORL1



# Muon-Track Matching Scenario B: GMT



# Possible BMT Layer-2/Track Trigger interface



- Each layer 1 board receives 57 links @ 10 G + RPC and outputs 18 links @16G
- Each Layer 2 TMT 96-link board receives
  - 2 fibres \*9 regions =18 fibres from track trigger @ 25G
  - +60 fibres from Layer 1 @16G

# EMTF + Muon Correlator Architecture

Barrel Muon System



60 MB sectors

Track Trigger



Track Finding Processor

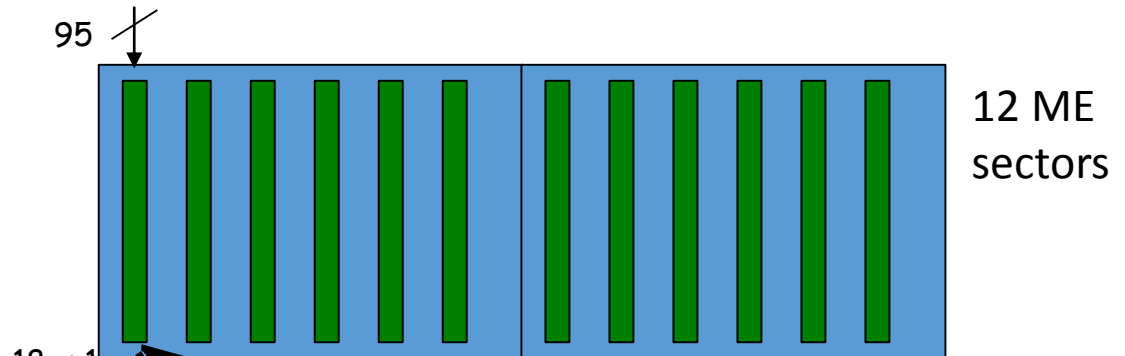
9 regions by ~18 time slices, 2 links each



x TM



(Standalone ME Muons) + Stub Data Concentrator



18 x 1

...

In ~2.5 us

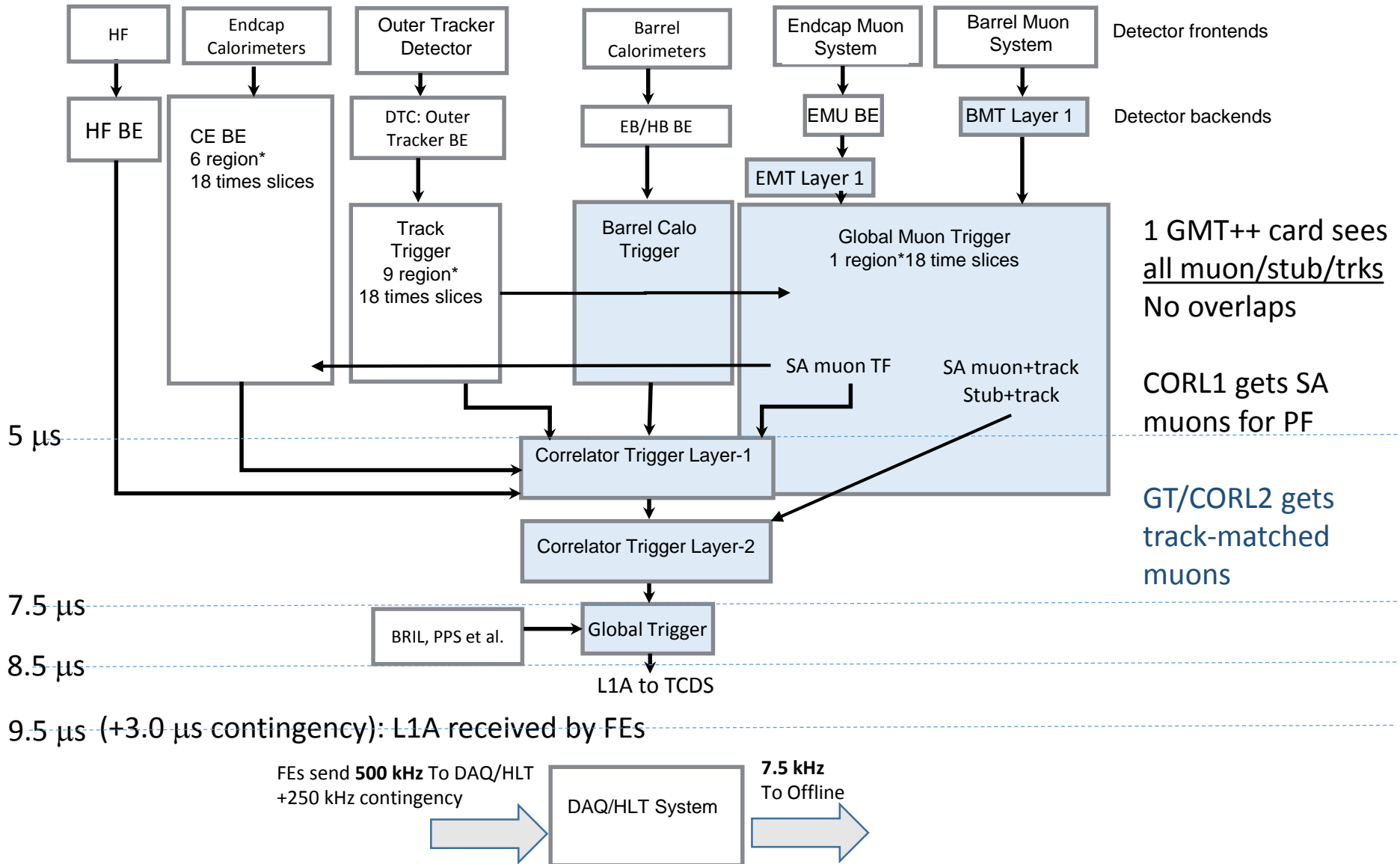
Correlator layer

18 processors  
(1 per time slice) take  
all muon and all  
Tracker data

12

9 x 2

# Muon-Track Matching Scenario C: GMT++

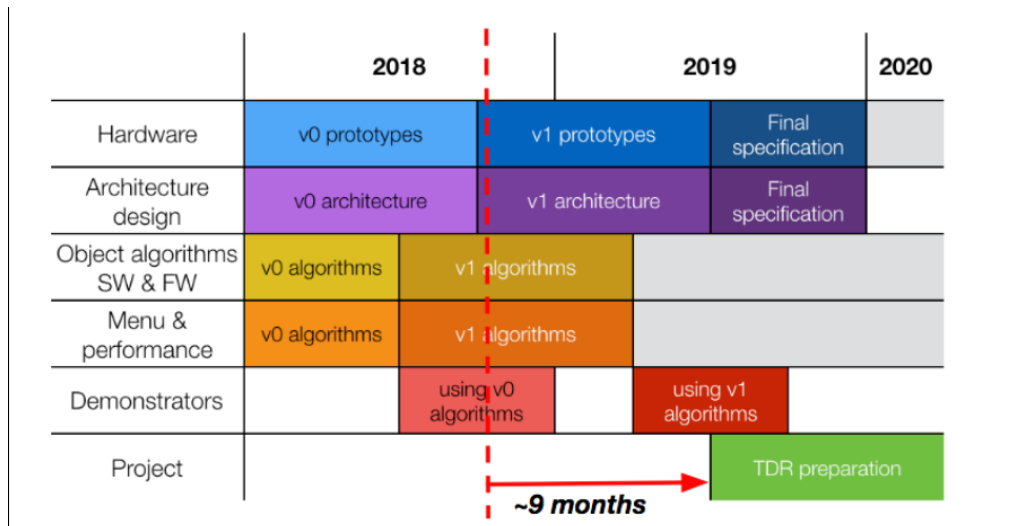


# Muon Trigger: To learn for TDR baseline

- In scenario A (defer to correlator):
  - full estimate of stub data bandwidth
  - FW resources needed on top of PF reco
- In scenario B (iTDR + GMT):
  - Full estimate of stub data bandwidth
  - How many boards?  $\geq 4$  for I/O
- In scenario C (GMT++):
  - Cost/resources for EMT or BMT concentration
  - FW resources needed for 1 board/time slice

# Timeline for TDR

- CWR draft needed by Sept. 2019
- For complete set of architecture choices, require additional study of resource usage per FPGA and data concentration options
- TIMELINE
  - By late winter workshop, reduce to a ~few options that are technically feasible. Evaluate latency and data organization merits of each.
  - By early summer workshop, decide on a baseline option for TDR and define TDR demonstrator architecture specs.
  - TDR demonstrator results delivered for TDR draft Fall 2019.
  - Specification continues through Q4 to prepare for preproduction phase in 2020



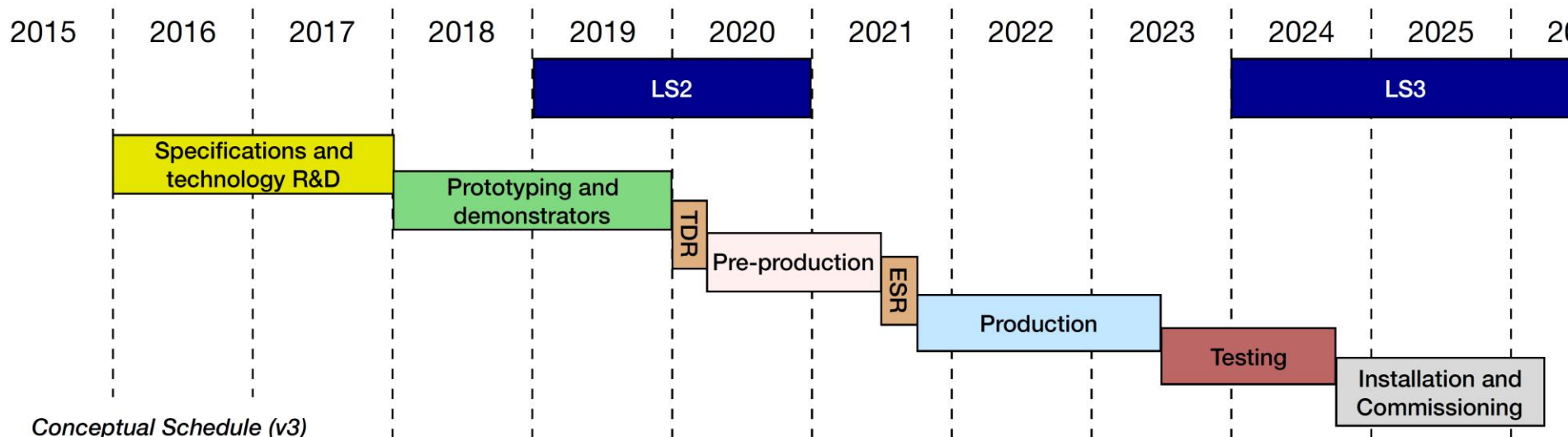
# Demonstrator Specification

- Next three months: prototypes evaluated for the different production lines
- Spring 2019: algorithm firmware demonstration on viable prototypes
- Summer 2019: slice configuration and demonstration on a best-effort basis
  - Input transmitter and/or output receiver can be an actual prototype, if available, or emulated by an available board(s) with the necessary links
  - With DTH/TCDS if available, otherwise can emulate
- Essential triplets:
  - BCP → RCT → GCT
  - DT inputs → BMT Layer 1 → BMT Layer 2
  - CSC inputs → EMTF → EMGS
  - GCT+CE+TT+Muon → CORL1 → CORL2
  - CORL2 → GT → L1A



# Timeline for Construction/Commissioning

- Pre-TDR: establish baseline and change control for interfaces and TPGs
- Pre-ESR (2021Q3): finalize interfaces. Slice tests of all for final design.
- Full production batches delivered to CERN 2023Q3
- 2.75 years available for testing and commissioning as interfacing electronics are installed (currently reserving ~6 months float).
- Pre-beam commissioning:
  - Internal relative timing/TMUXing of L1 (with ECAL pulses, e.g.) and available interfaces
  - Muon cosmics in LS3, run 3 muon data possible for some ingredients (GEM)
  - With Tracker inserted starting 2026



Conceptual Schedule (v3)

# Summary

- We have evaluated board and link counts for each subsystem, using primarily 16 Gbps links connecting 96-link or 64-link chips
- Key algorithms have been evaluated on candidate chips and are expected to meet latency budgets and resource limits
  - CORL2 and GT chief exceptions. These are also the most flexibly defined systems.
- Counts in the scenarios considered are consistent with range specified in iTDR
- Still considering options for different architecture choices (TMUX, regionalization, physical partitioning of functions), to converge on a baseline in six months.
- Demonstration of essential chains scheduled for next summer as part of the TDR