Design and Development of the DAQ and Timing Hub for CMS Phase-2

September 20, 2018

Jeroen Hegeman on behalf of the CMS DAQ group

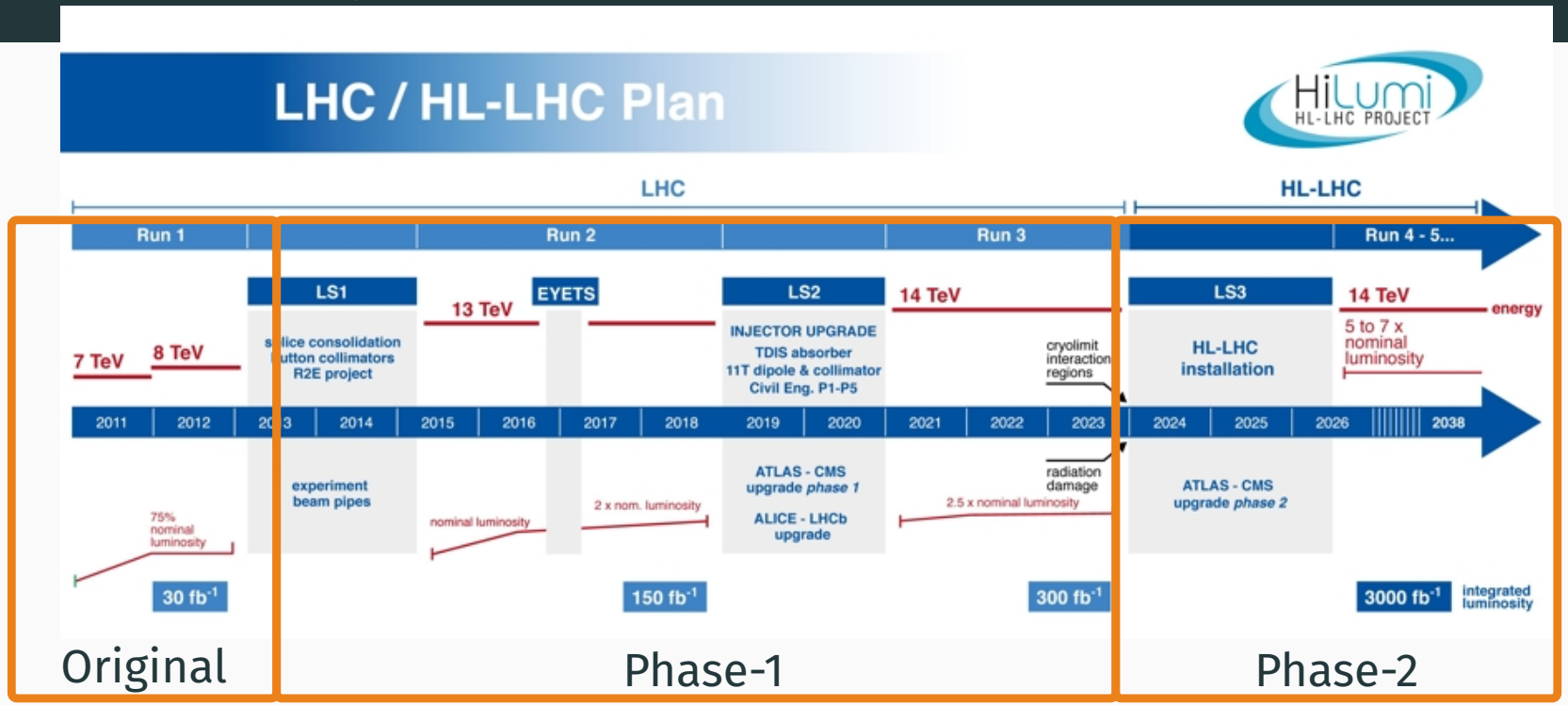
TWEPP 2018
Topical Workshop on Electronics for Particle Physics

Outline

- Introduction to ‘Phase-2 CMS’
- DAQ Design for CMS Phase-2
- CMS Phase-2 DAQ and TCDS Hub (DTH)
- Next steps

Introduction to ‘Phase-2 CMS’

High-Luminosity LHC



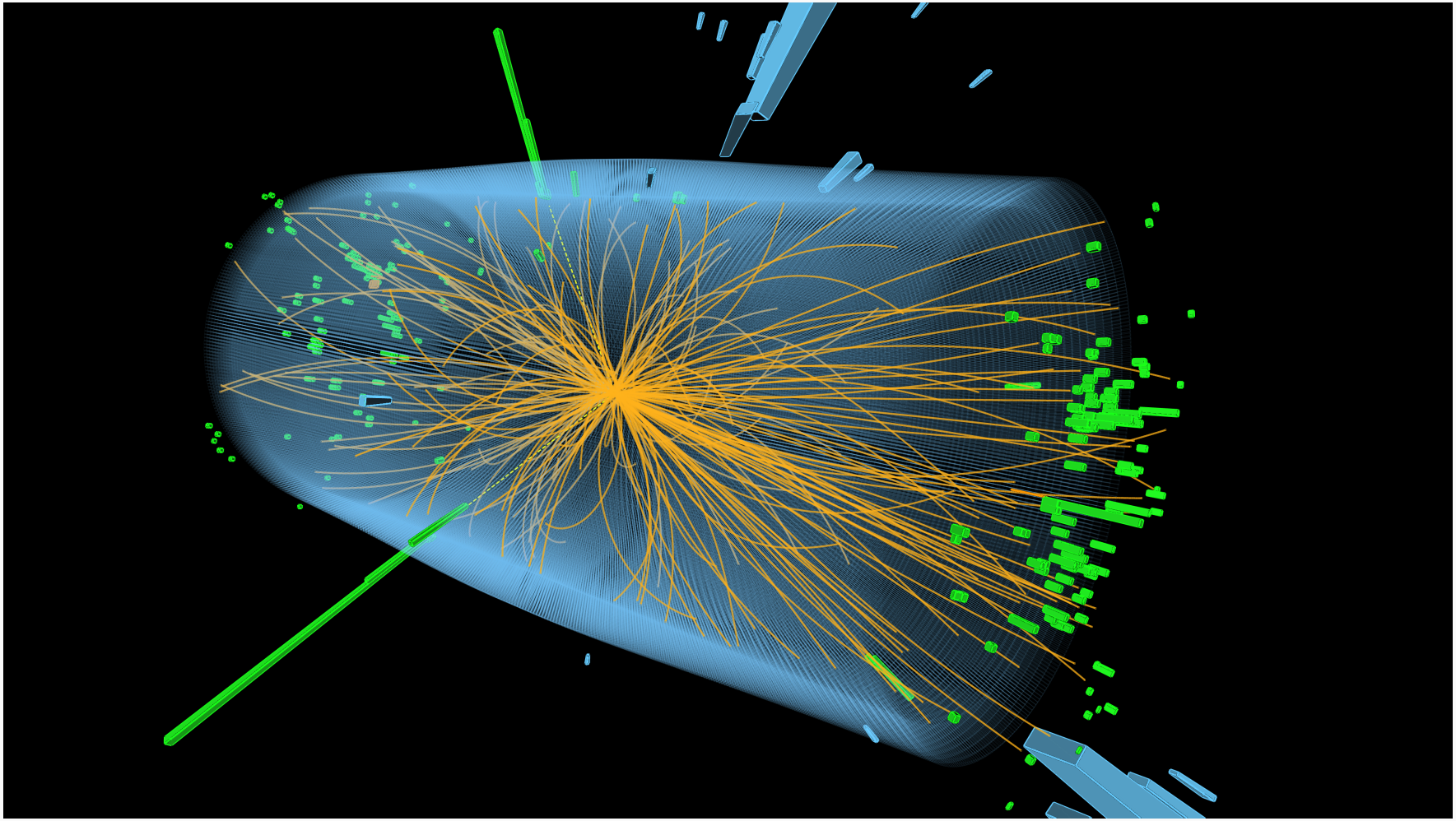
Phase-2: instantaneous luminosity will increase to $7.5 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$

For the experiments this (unprecedented) high luminosity means:

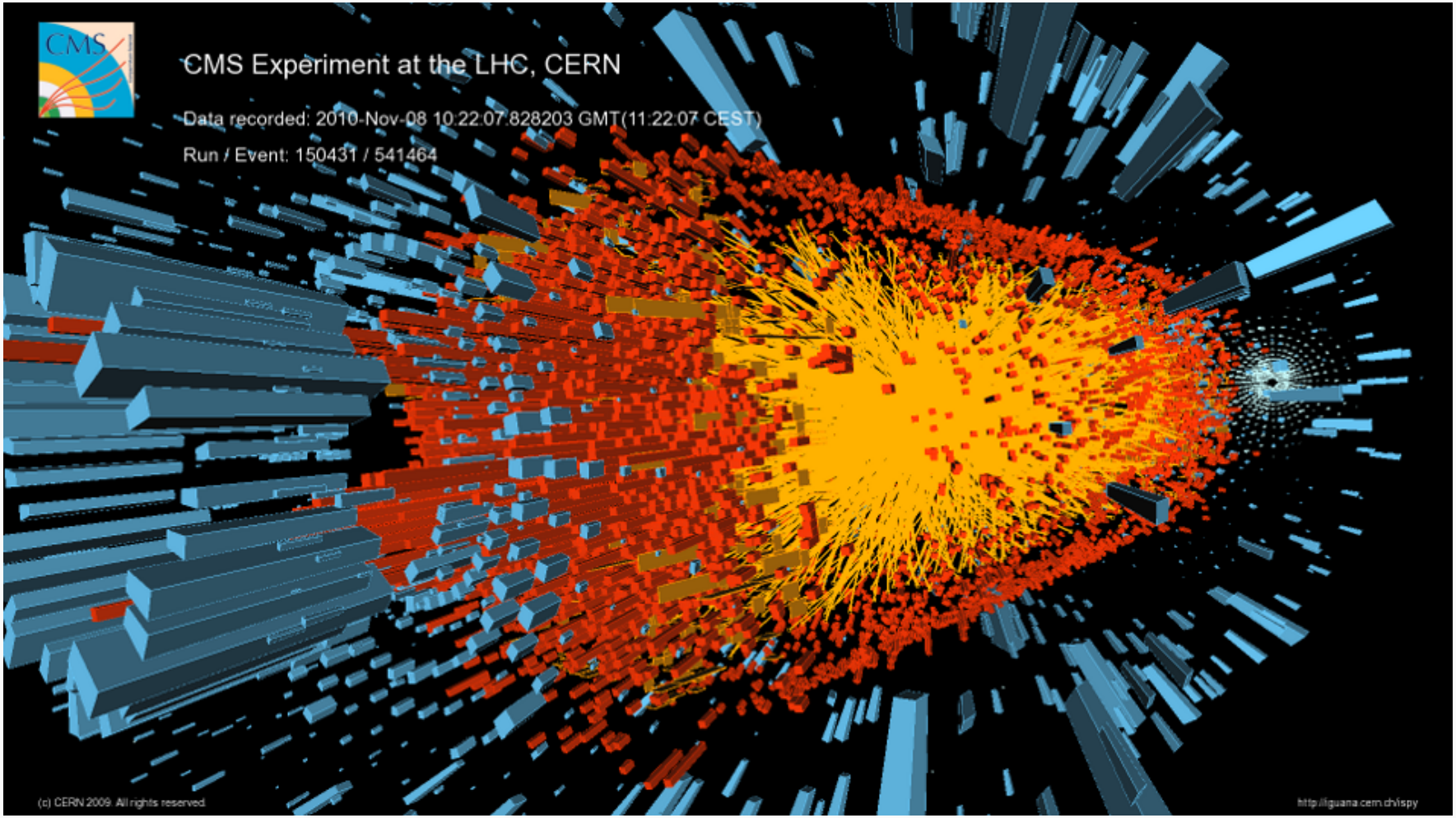
- Extended physics reach, and increased precision
- Much increased radiation levels -> harsher environment for detectors and electronics
- Increased number of simultaneous proton-proton collisions per bunch crossing -> more advanced pile-up mitigation strategies needed

http://hilumilhc.web.cern.ch

Higgs $\rightarrow$ two photons!



Higgs $\rightarrow$ two photons?



CMS Phase-2 Upgrade

CMS Phase-2 design goal
Maintain the current, excellent, performance in terms of efficiency, resolution, and background rejection for all physics objects.

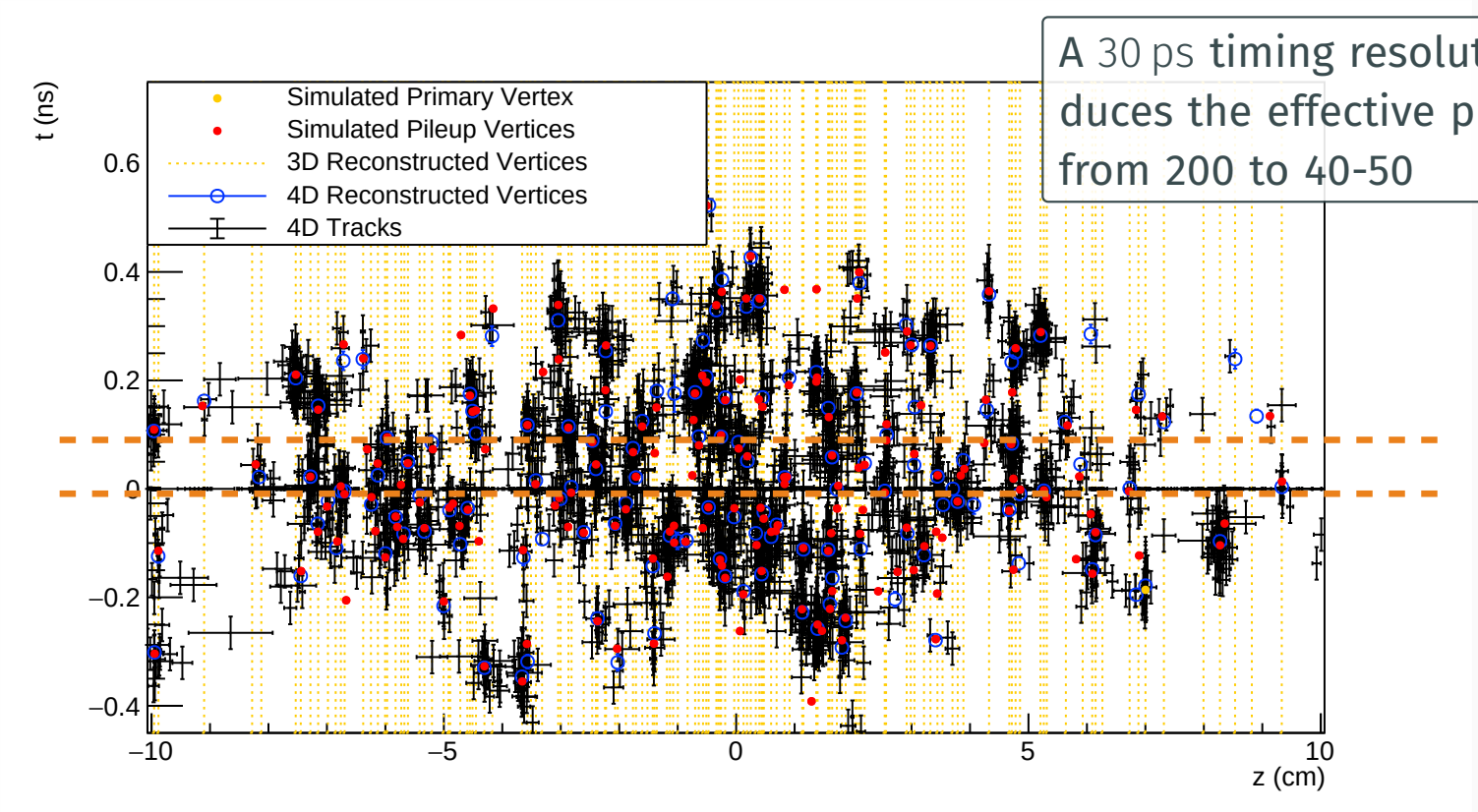
CMS Phase-2 Upgrade

CMS Phase-2 design goal
Maintain the current, excellent, performance in terms of efficiency, resolution, and background rejection for all physics objects.

Primary CMS Phase-2 approaches:

- Increased granularity, focusing on particle flow reconstruction algorithms in the high multiplicity environment
- Addition of precise timing information to several subdetectors
- Addition of a dedicated MIP timing detector
- Introduction of tracking information into the L1 trigger in order to tighten selection

CMS Phase-2 – Highlight: MIP timing detector



A 30 ps timing resolution reduces the effective pile-up from 200 to 40-50

Including hit-timing information helps distinguish primary vertices/collisions. Clustering, tracking and vertexing can be done in full 4D.

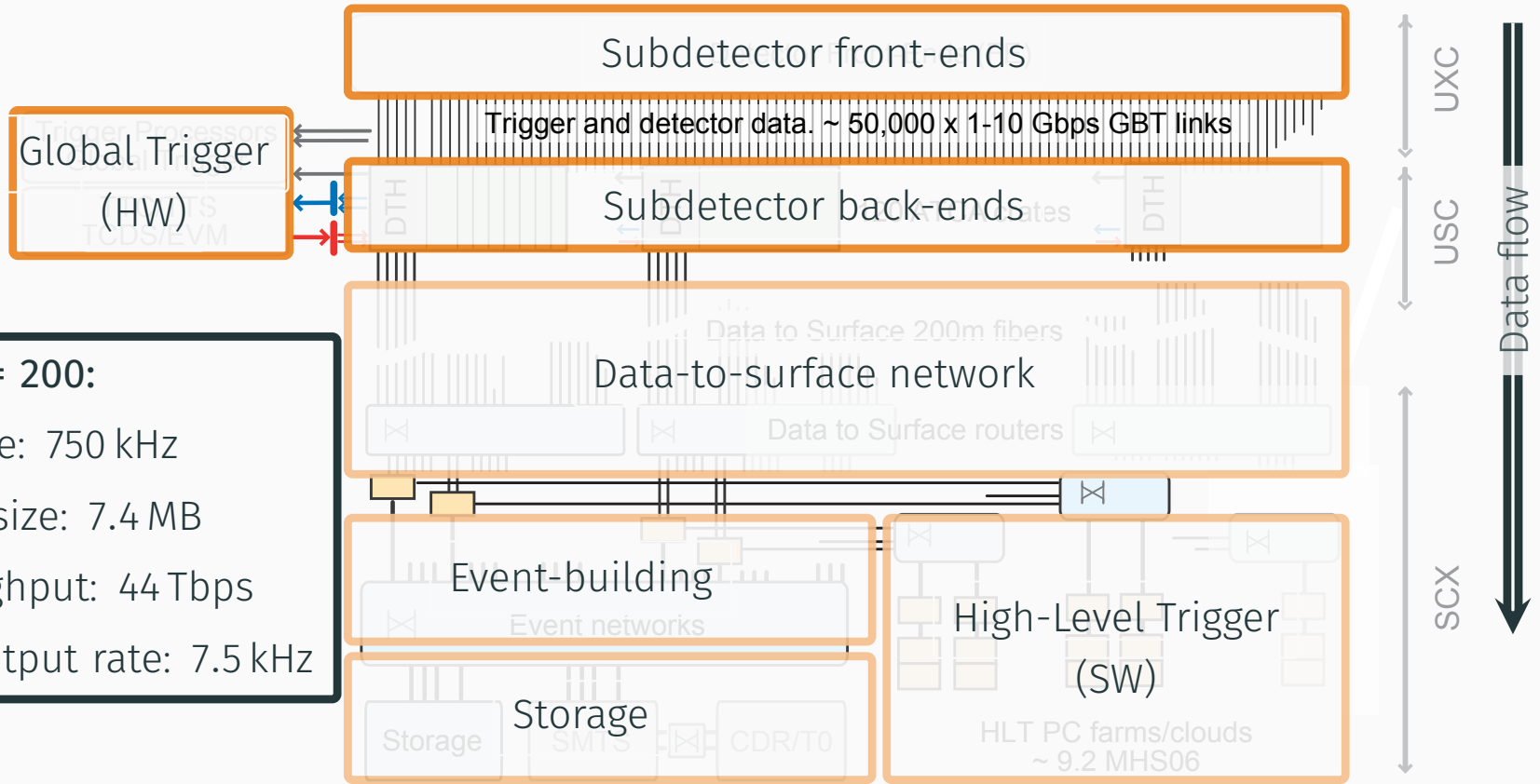
Requires distribution and monitoring of fine-grained timing information to all front-ends → Cross-experiment precision timing studies group formed

Note: numbers taken from preliminary studies, for demonstration purposes only

DAQ Design for CMS Phase-2

Phase-2 CMS baseline DAQ (DAQ4)

CMS trigger-DAQ architecture choice: Stick to the proven two-stage trigger approach, but at higher rate and throughput

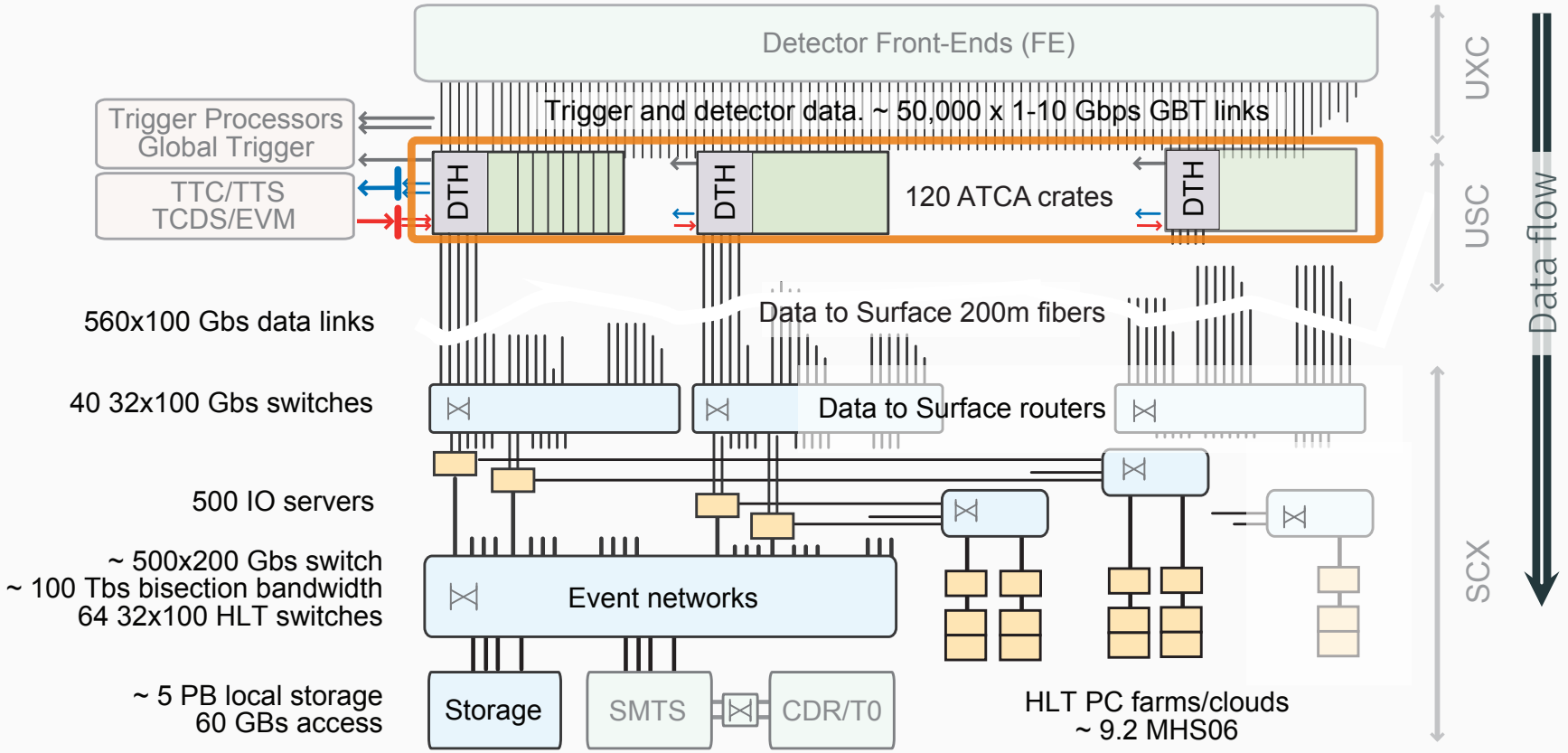


Pile-up = 200:

- L1A rate: 750 kHz
- Event size: 7.4 MB
- Throughput: 44 Tbps
- HLT output rate: 7.5 kHz

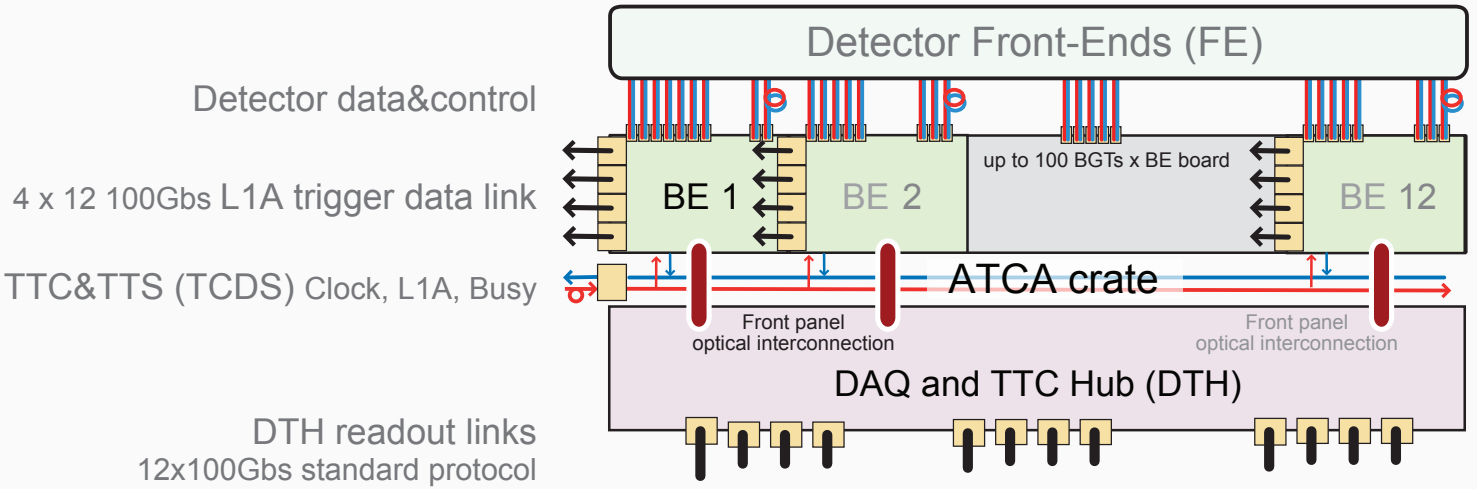
Phase-2 CMS baseline DAQ (DAQ4)

CMS trigger-DAQ architecture choice: Stick to the proven two-stage trigger approach, but at higher rate and throughput



CMS Phase-2 DAQ and TCDS Hub

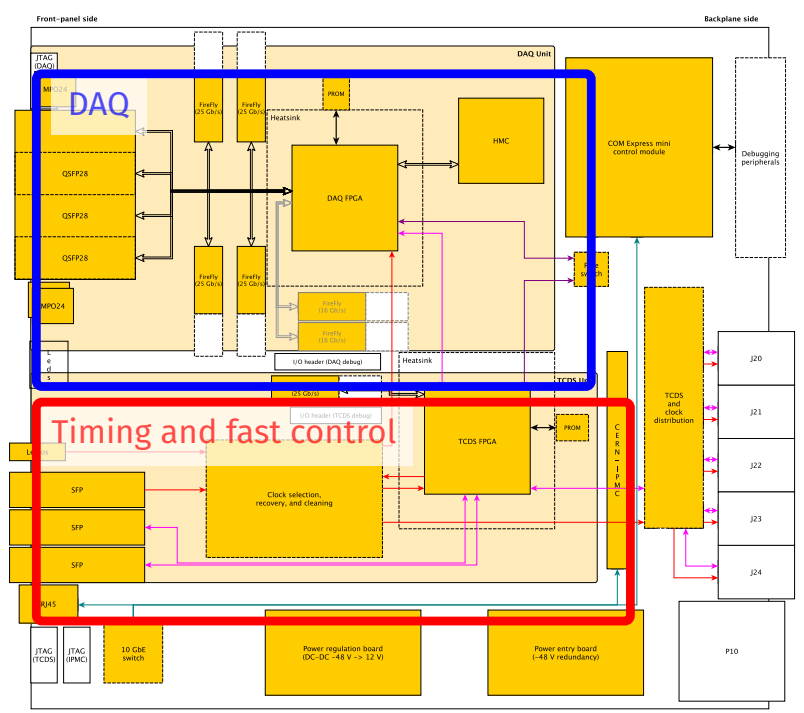
DAQ and Timing Hub (DTH)

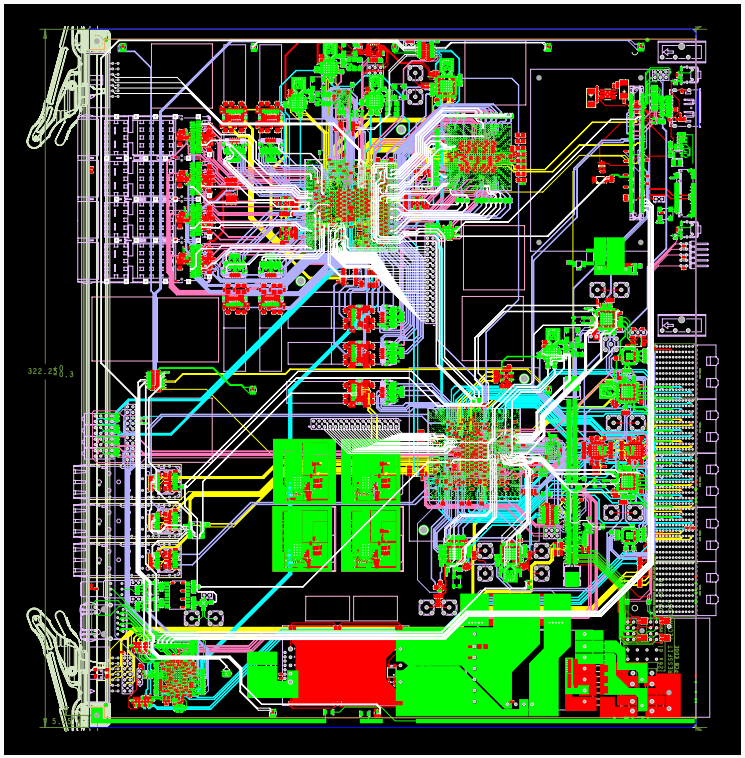


DTH: ATCA hub card housed in subdetector back-end crates

- Lives between the synchronous world of L1 trigger and detectors and the asynchronous DAQ world
- Distributes clock signals, synchronization commands, L1 trigger, and slow control from central systems to back-ends
- Receives event fragments from back-ends, concentrates and buffers these, and transmits them to the data-to-surface network
- Provides functionality for stand-alone ‘mini’ DAQ and TCDS
- Contains built-in support for DAQ and TCDS link testing and monitoring

DAQ and Timing Hub (DTH)

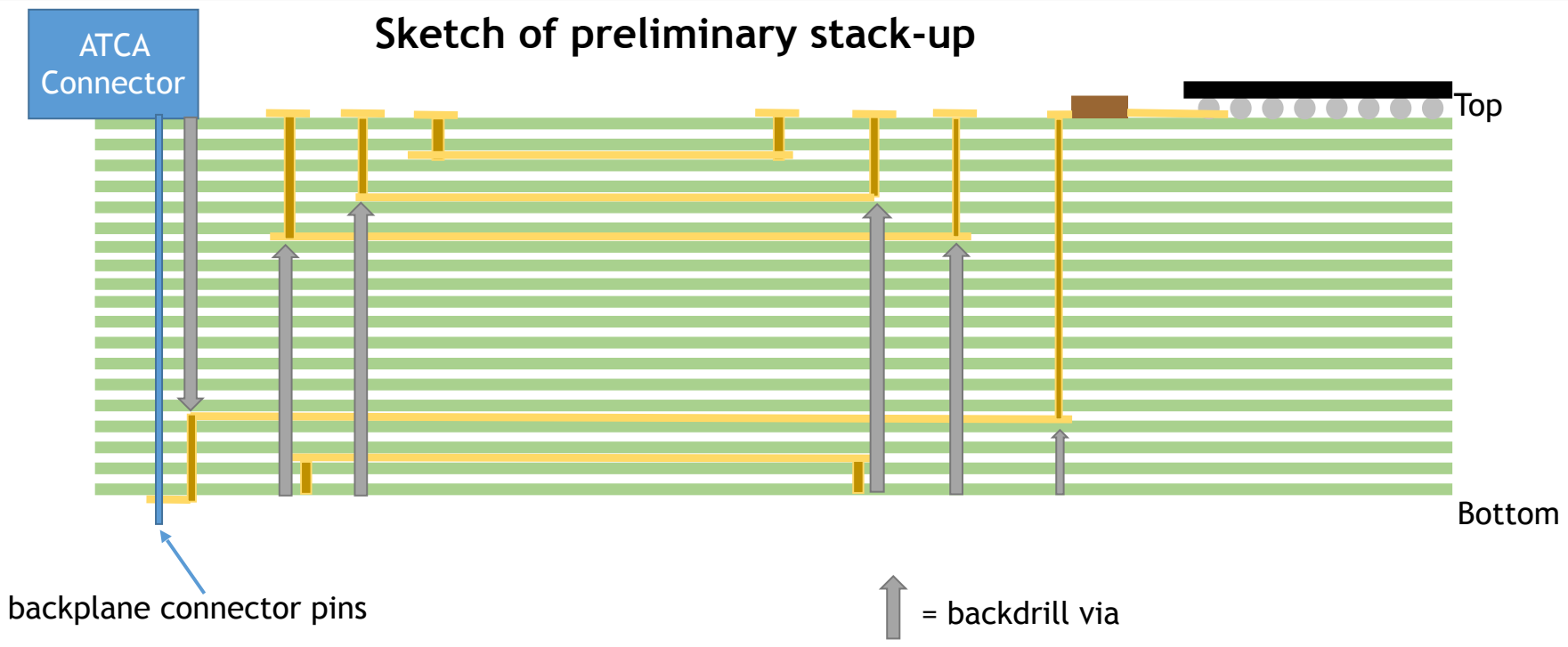




Prototype 1 has been designed as an ‘evaluation board’

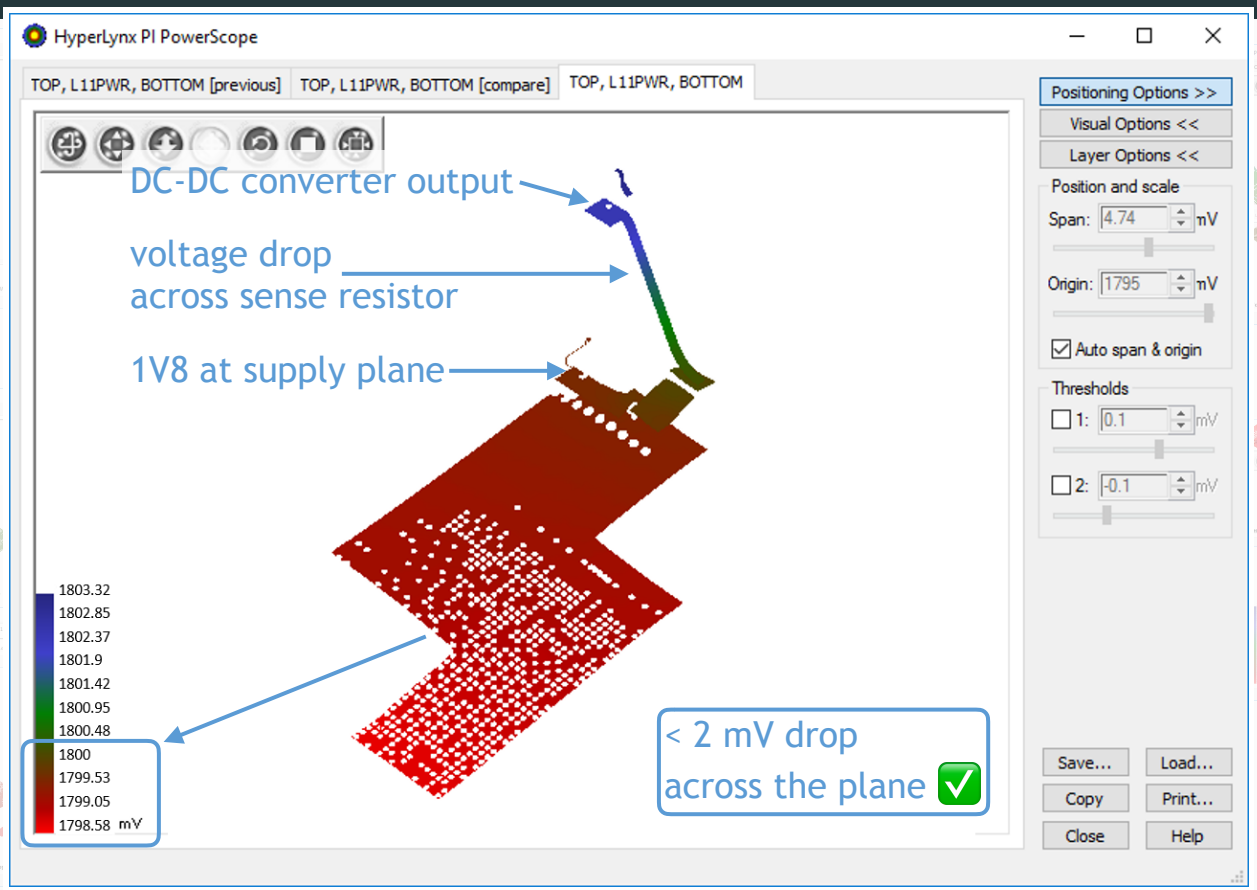
- DAQ unit: baseline design, aimed at validating optics, memory, and FPGA
  Data concentrator from 16×16/25 Gbit/s custom protocol to 4×100 Gbit/s TCP/IP
- TCDS part: instrumented to show proof-of-principle, and to validate components
- In general: over-instrumented for debug access and with humidity and temperature sensors (on front and back)

DTH - PCB design



- 20 layers of I-Tera MT40 ($Dk = 3.45$ up to 25 Gbit/s)
- All high-speed layers sandwiched between ground layers
- Combination of via technologies: compromise between signal integrity, PCB symmetry, and cost

DTH - Power integrity



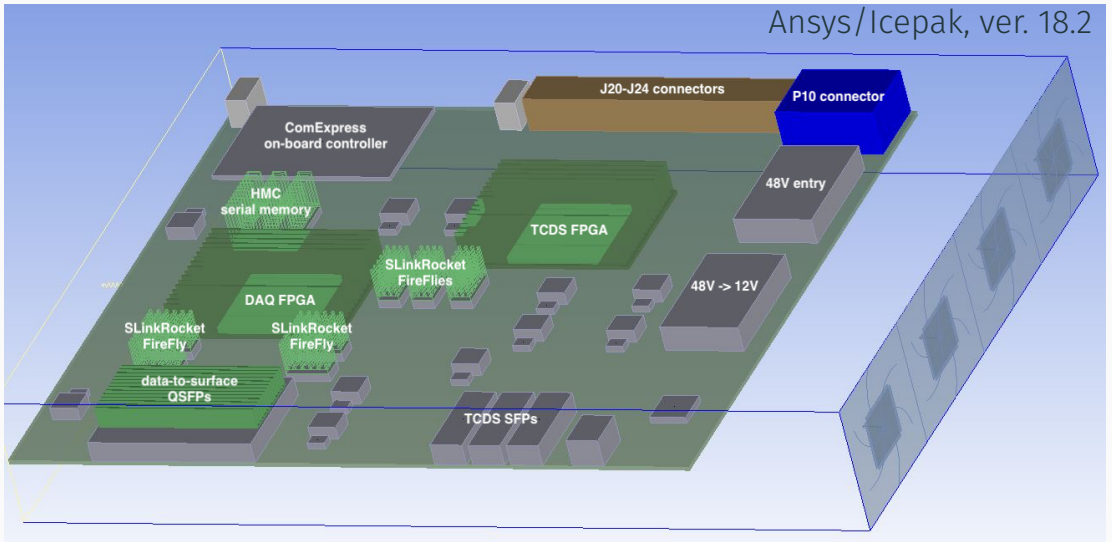
- All supply nets have been simulated
- Based on simulation, the routing was optimized based in several places to reduce current density and to assure power integrity
- Very useful exercise, but requires careful ‘merging’ of simulation results with designer experience

DTH - Thermal aspects

Aim of thermal simulation: verify design for component specs and longevity

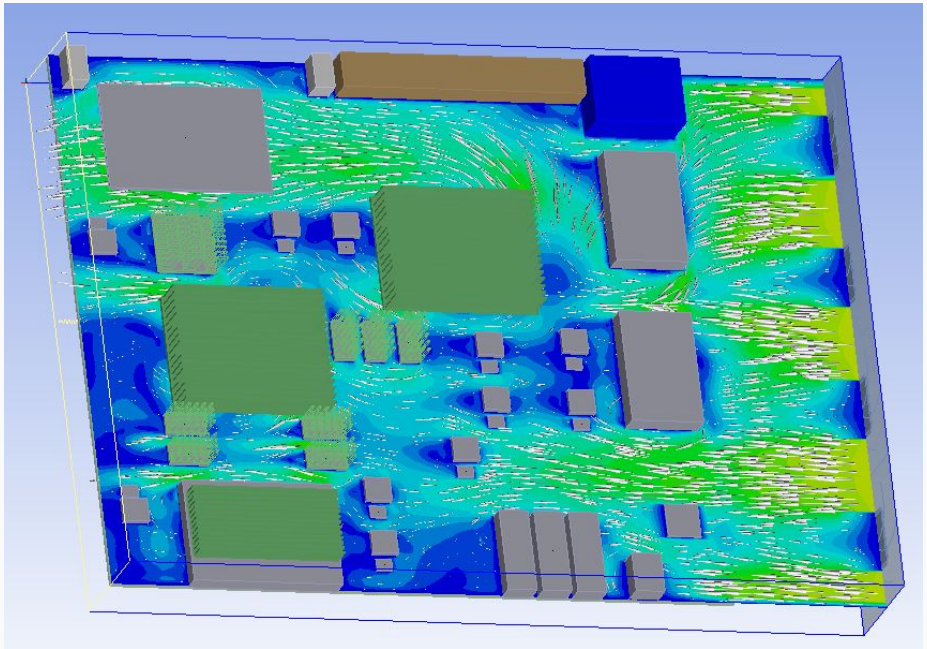
- Model includes only those components with high power dissipation ($\gtrsim 3\,\mathrm{W}$) or that significantly shape/obstruct the airflow
- Airflow generated by perfect fans according to ATCA specs

- Currently studying different levels of description for thermal behaviour
- Intend to keep the thermal model as simple as possible
- Target is to have reasonable temperature accuracy, tuneable to reality
- Impact of neighbour cards will be studied (as a function of neighbour dissipation)

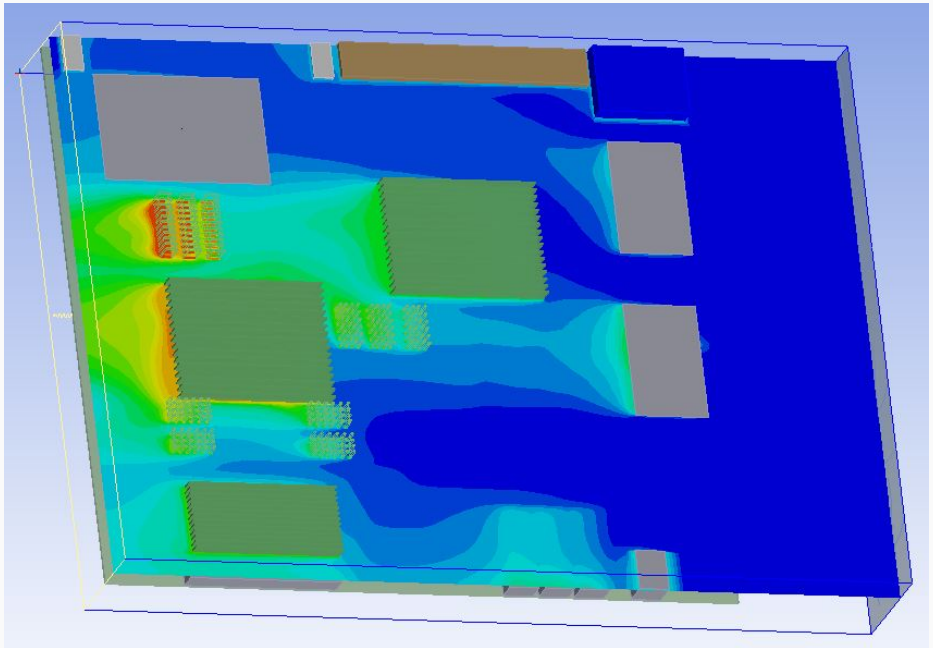


DTH - Power and thermal aspects

Airflow



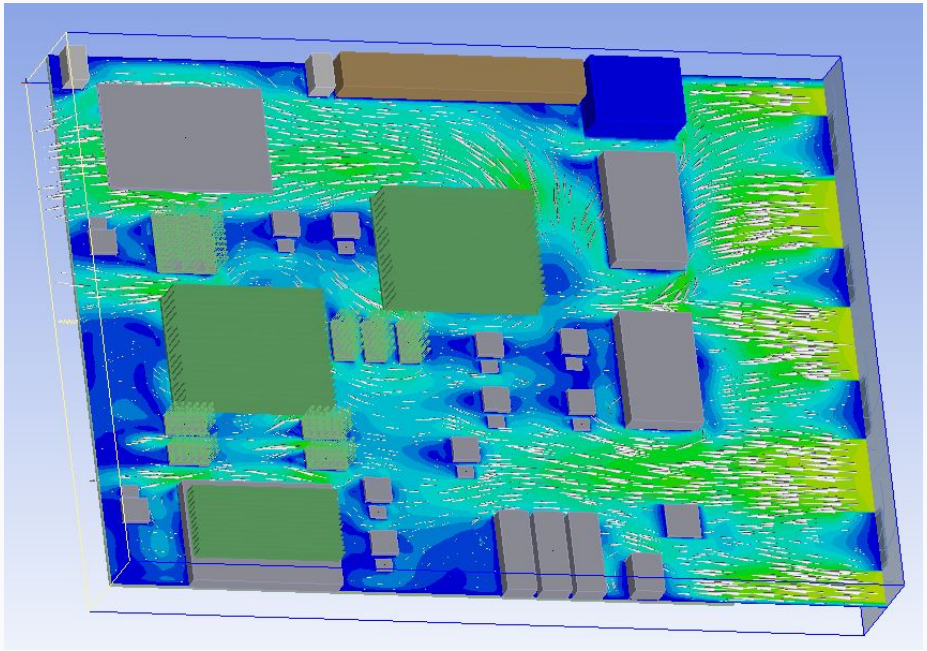
Temperature



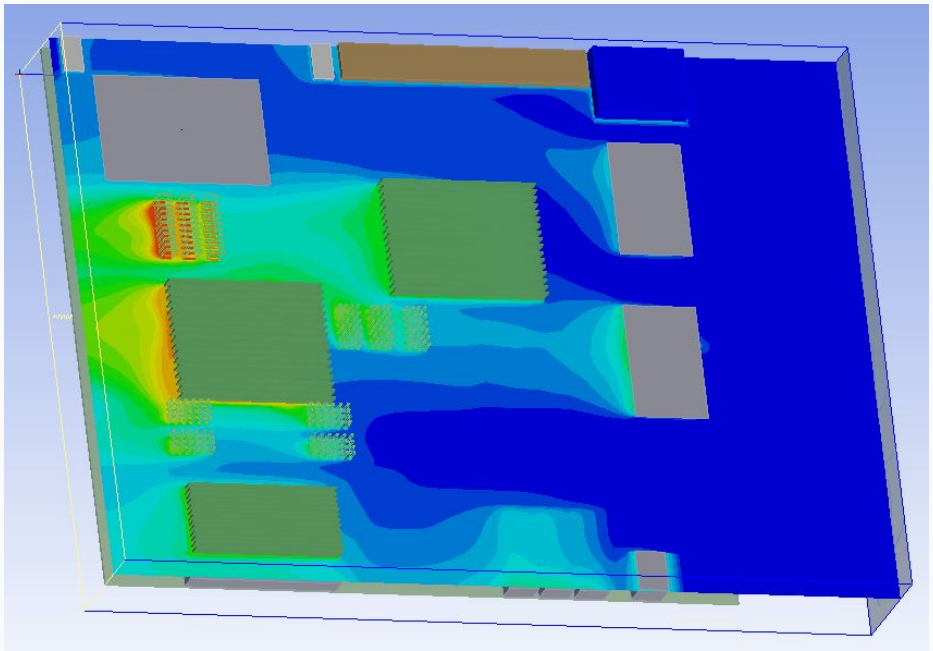
- Airflow and temperature behaviour follow basic physics, but the details quickly become quite intricate
- Started a collaboration with other back-end developers to pool experience in thermal design and simulation
- Current situation shows there is ample room for improvement both in thermal modeling and in DTH thermal design

DTH - Power and thermal aspects

Airflow



Temperature



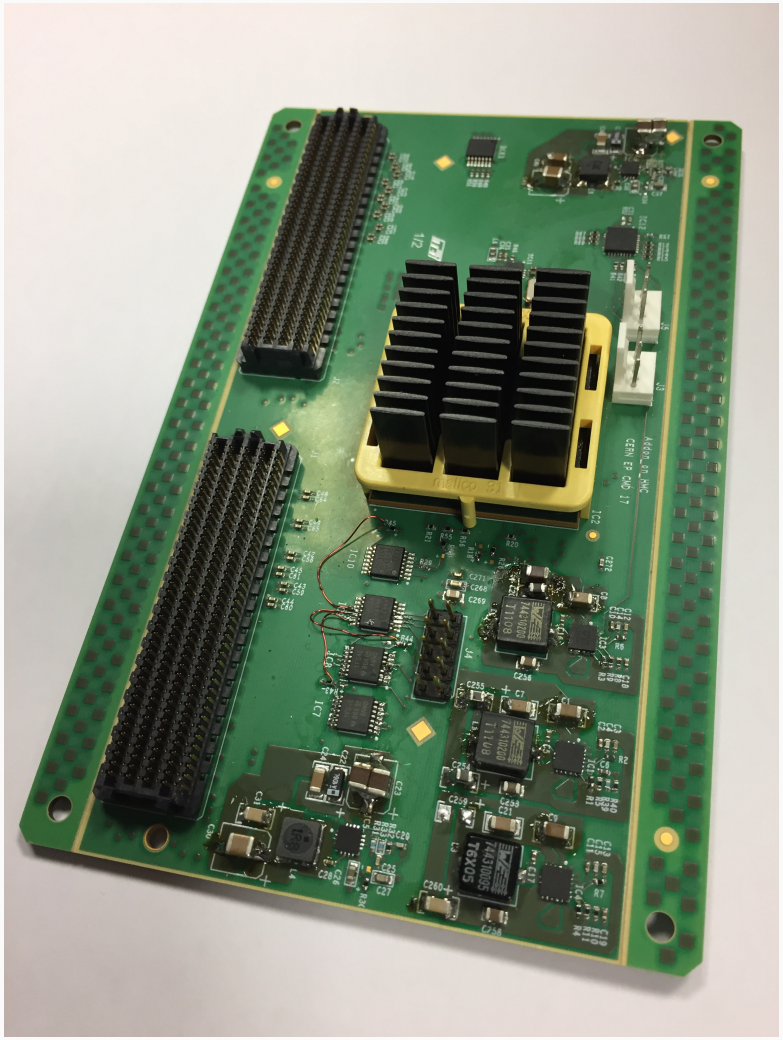
Intended approach:

- Use DTH P1 to improve thermal model, and to tune to reality
- Then use simulation to optimize the thermal design for future DTH versions

DTH - Feature: HMC

Serves as buffer memory for TCP/IP data-to-surface connection

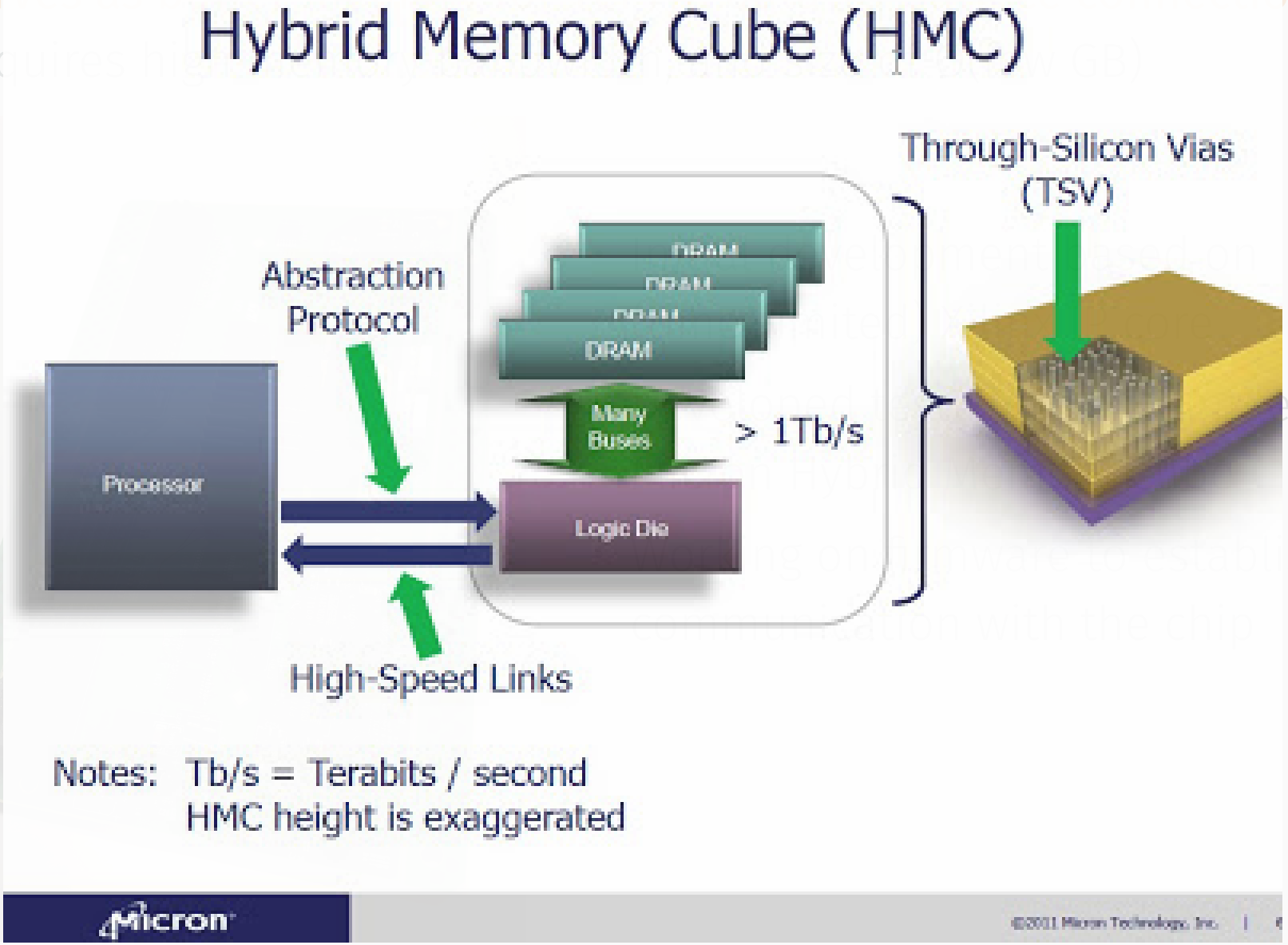
Requires high memory bandwidth, and size of O(few GB)



- Initial development based on (time-limited) Xilinx IP core
- Developed test FMC for samples of Micron Hybrid Memory Cube
- Working on firmware to establish communication with the chip

DTH - Feature: HMC

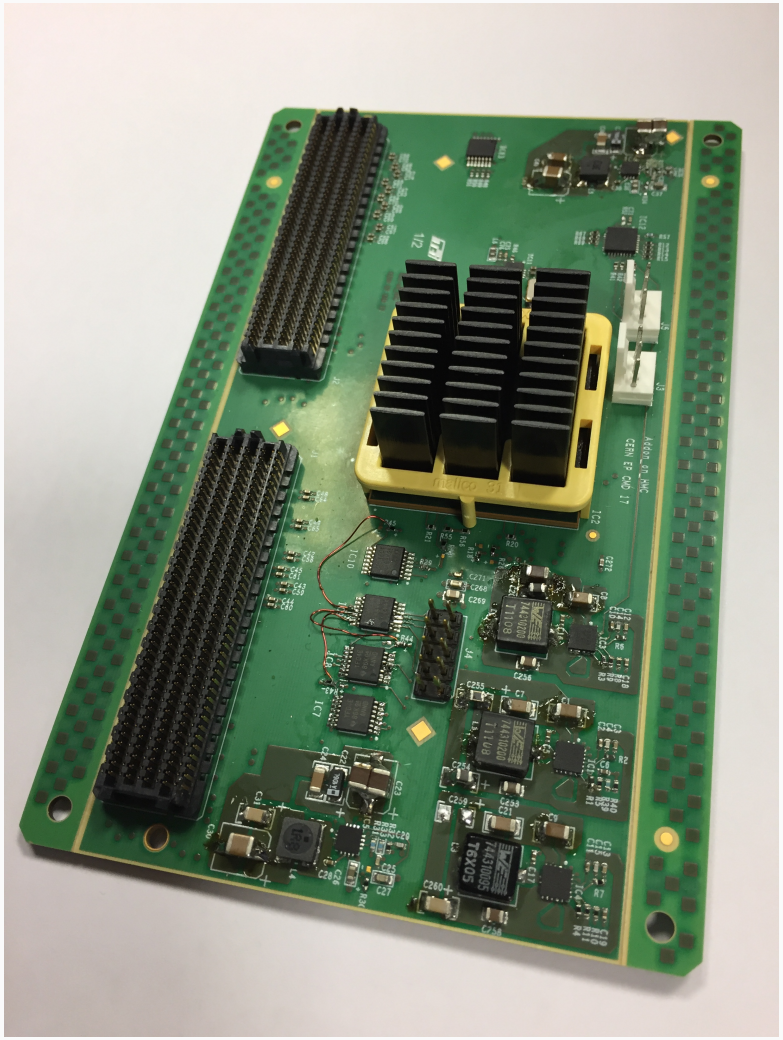
Collaboration initiated by Intel and Micron



DTH - Feature: HMC

Serves as buffer memory for TCP/IP data-to-surface connection

Requires high memory bandwidth, and size of O(few GB)

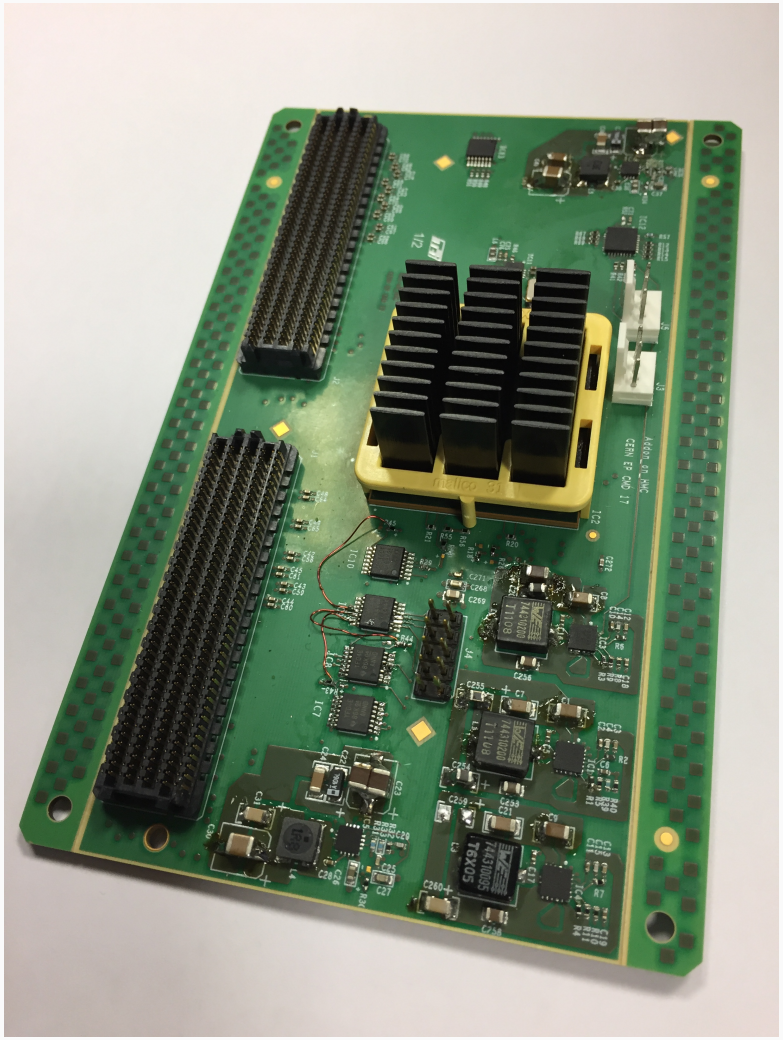


- Initial development based on (time-limited) Xilinx IP core
- Developed test FMC for samples of Micron Hybrid Memory Cube
- Working on firmware to establish communication with the chip

DTH - Feature: HMC

Serves as buffer memory for TCP/IP data-to-surface connection

Requires high memory bandwidth, and size of O(few GB)



- Initial development based on (time-limited) Xilinx IP core
- Developed test FMC for samples of Micron Hybrid Memory Cube
- Working on firmware to establish communication with the chip
- Already out of fashion/out of grace, and out of production (by next year)?

DTH - Feature: DAQ input/output

- Firmware development setup for SlinkRocket point-to-point protocol from back-end to DTH
- Firmware development setup for 100 Gbit/s TCP/IP
- Evaluation setup for choice between single-mode and multi-mode optical link technology for Data-to-Surface network

DTH - Feature: DAQ input/output

Data from back-end to DTH: SlinkRocket

One very useful product of these exercises: resource usage estimates (!) for IP cores delivered to back-end developers

SlinkRocket and CRC implementation studies (on KU15P, preliminary)

	Resource	Utilisation	Utilization (%)
SlinkRocket (without generator)	LUT	7198	1.34
	LUTRAM	3858	5.02
	FF	4286	0.40

CRC	Resource	Utilisation	Utilization (%)
16-bit	LUT	857	0.16
	FF	542	0.05
32-bit (Ethernet)	LUT	2027	0.39
	FF	570	0.05
32-bit (iSCSI)	LUT	2027	0.39
	FF	569	0.05

DTH - Feature: DAQ input/output

Data from back-end to DTH: SlinkRocket

One very useful product of these exercises: resource usage estimates (!) for IP cores delivered to back-end developers

SlinkRocket and CRC implementation studies (on KU15P, preliminary)

	Resource	Utilisation	Utilization (%)
SlinkRocket (without generator)	LUT	7198	1.34
	LUTRAM	3858	5.02
	FF	4286	0.40

- May need optimization, given the number per back-end FPGA
- Thinking about a way to 'share' one generator implementation for multiple SlinkRockets

DTH - Feature: DAQ input/output

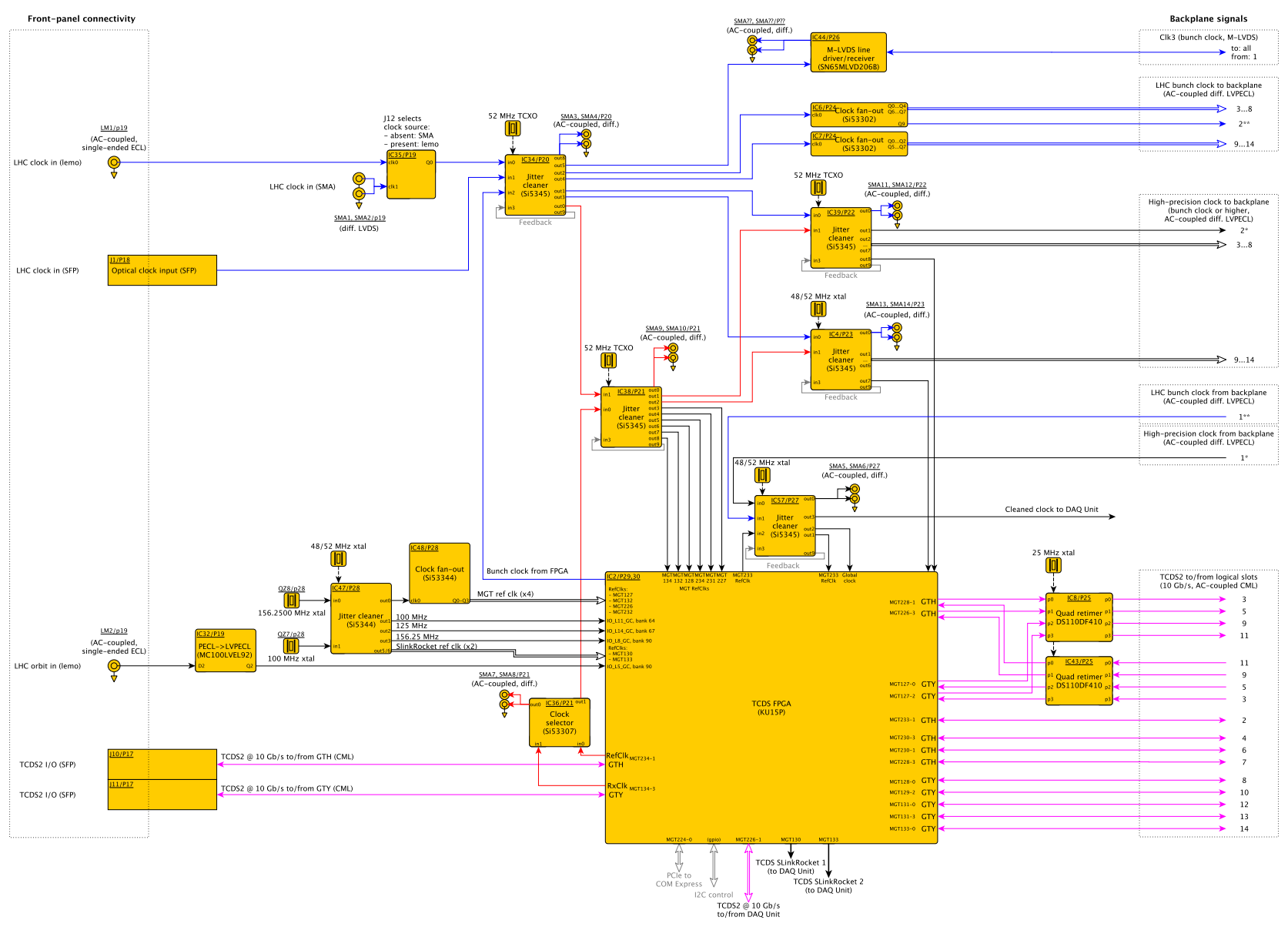
Data from back-end to DTH: SlinkRocket

One very useful product of these exercises: resource usage estimates (!) for IP cores delivered to back-end developers

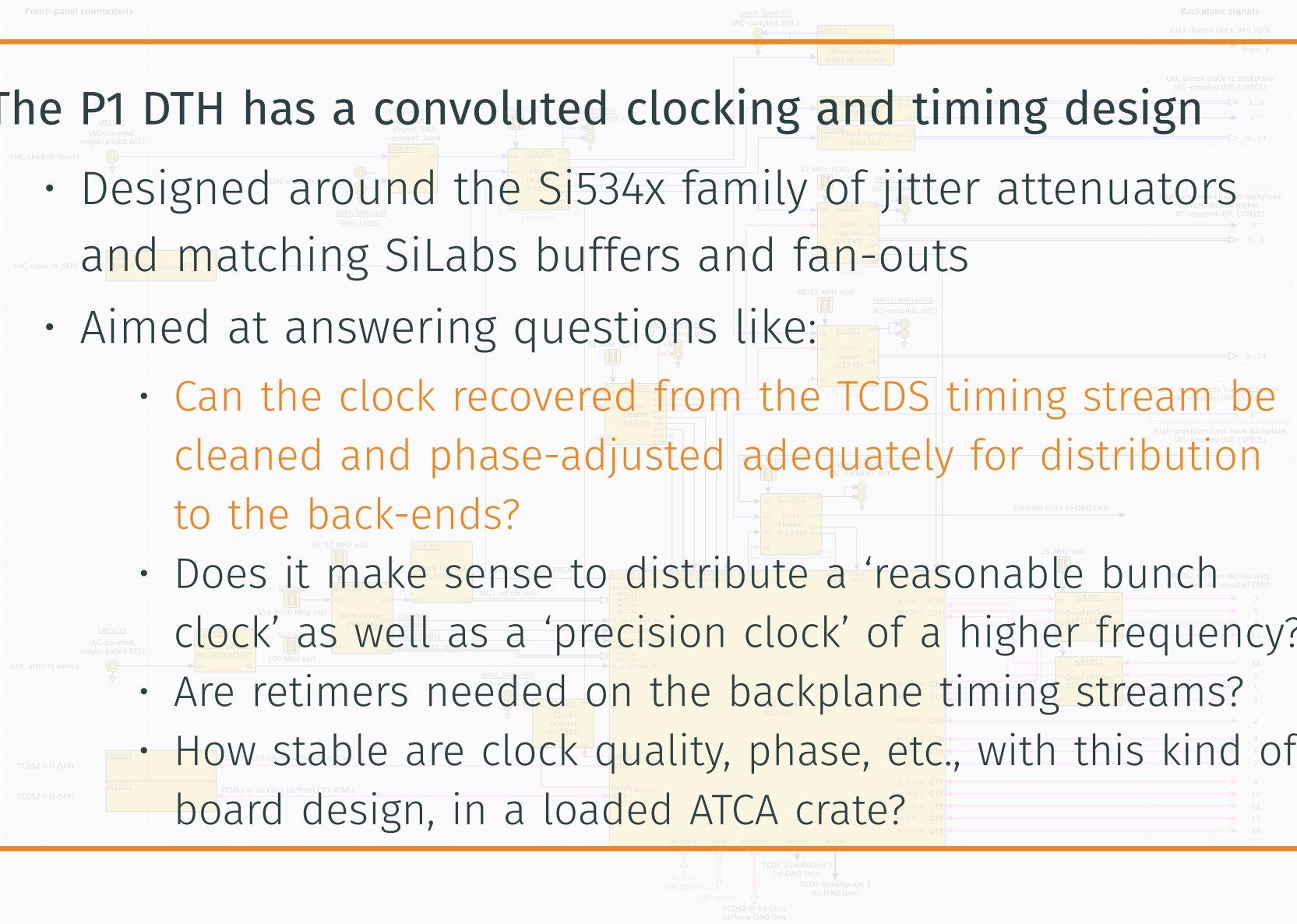
- Small enough not to worry for the moment
- May use 16-bit CRC on the SlinkRocket senders, and use 32-bit CRC on the data-to-surface for CPU compatibility

CRC	Resource	Utilisation	Utilization (%)
16-bit	LUT	857	0.16
	FF	542	0.05
32-bit (Ethernet)	LUT	2027	0.39
	FF	570	0.05
32-bit (iSCSI)	LUT	2027	0.39
	FF	569	0.05

DTH - Feature: Clock recovery and distribution



DTH - Feature: Clock recovery and distribution

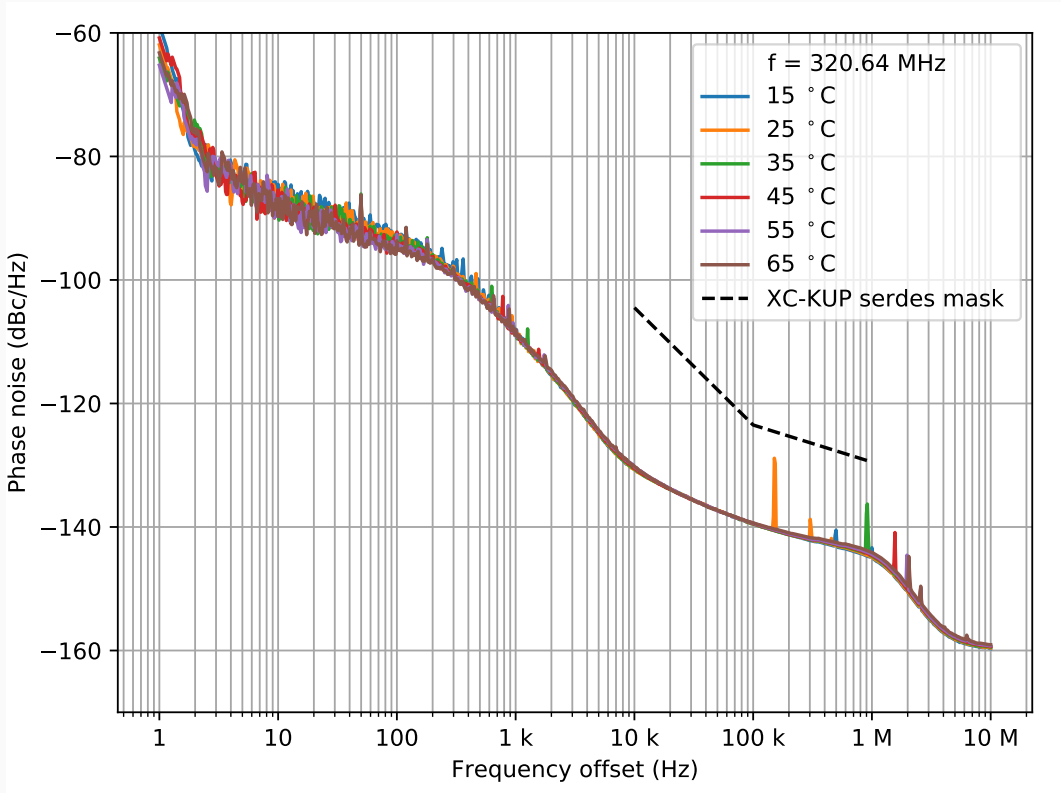


The P1 DTH has a convoluted clocking and timing design

- Designed around the Si534x family of jitter attenuators and matching SiLabs buffers and fan-outs
- Aimed at answering questions like:
  - Can the clock recovered from the TCDS timing stream be cleaned and phase-adjusted adequately for distribution to the back-ends?
  - Does it make sense to distribute a ‘reasonable bunch clock’ as well as a ‘precision clock’ of a higher frequency?
  - Are retimers needed on the backplane timing streams?
  - How stable are clock quality, phase, etc., with this kind of board design, in a loaded ATCA crate?

DTH - Component validation

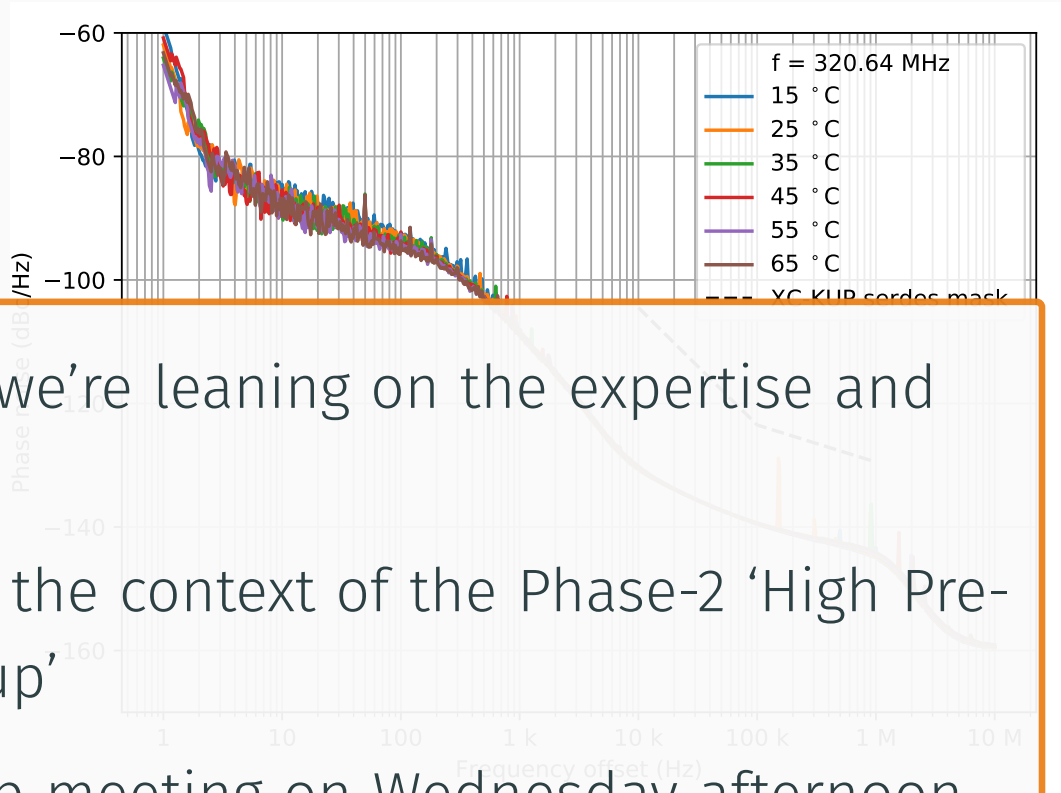
- Si5345 on SiLabs evaluation board
- PLL bandwidth set to 1 kHz
- Input clock: 10 MHz from SRS FS725
- Output clock: integer multiples of 40.08 MHz LHC bunch clock up to the lpGBT reference clock frequency



- First tests with SiLabs jitter attenuators indicate less temperature sensitivity than expected
  $\rightarrow$Would imply no need for TCXOs (over crystals)
- To be verified with real (DTH P1) hardware

DTH - Component validation

- Si5345 on SiLabs evaluation board
- PLL bandwidth set to 1 kHz
- Input clock: 10 MHz from SRS FS725
- Output clock: integer multiples of 40.08 MHz
- Input clock: 10 MHz from GPS reference clock frequency



For design and validation, we're leaning on the expertise and experience in CERN EP-ESE

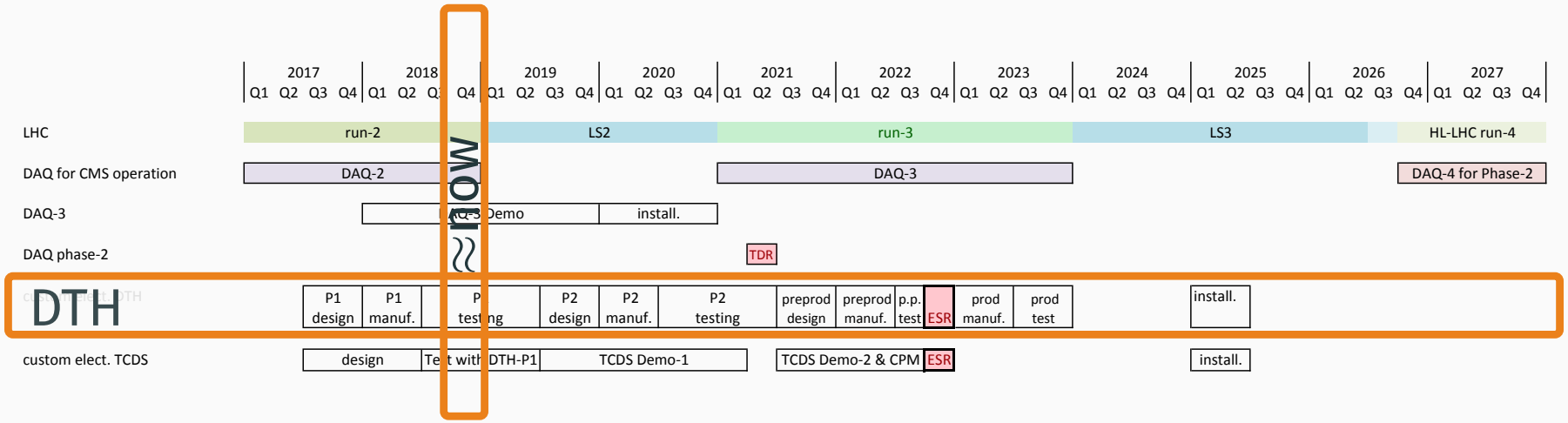
Evaluation work is done in the context of the Phase-2 'High Precision Timing Working Group'

See also the Working Group meeting on Wednesday afternoon

- First tests with SiLabs jitter attenuators indicate less temperature sensitivity than expected
  →Would imply no need for TCXOs (over crystals)
- To be verified with real (DTH P1) hardware

Next steps

CMS DAQ and DTH development timeline



(From 'The Phase-2 Upgrade of the CMS DAQ Interim Technical Design Report', CERN-LHCC-2017-014)

R&D should converge before the Technical Design Report in Q2-2021. Most notably:

- DTH design and timing performance validation
- Replacement of HMC for TCP/IP buffer memory

DTH development timeline

We're planning to:

- Q4 2018: receive first DTH P1 board(s) for initial testing
- Q1/Q2 2019: small production of P1 boards, functional testing within CMS DAQ group and with early-adopter back-end developers
- Q3 2019: present first results from real hardware at TWEPP'19
- Early 2020: support development in test and integration setups based on DTH prototype
- ...
- Early 2025: install the final system at the experiment
- Early 2026: commission the final system with subsystems
- Mid 2026: commission with collisions

Thank you