

New Strategies at CMS ASET Seminar for the HL-LHC



On Behalf of the Phase II Timing Layer and Trigger Working Groups

MIP Timing Detector

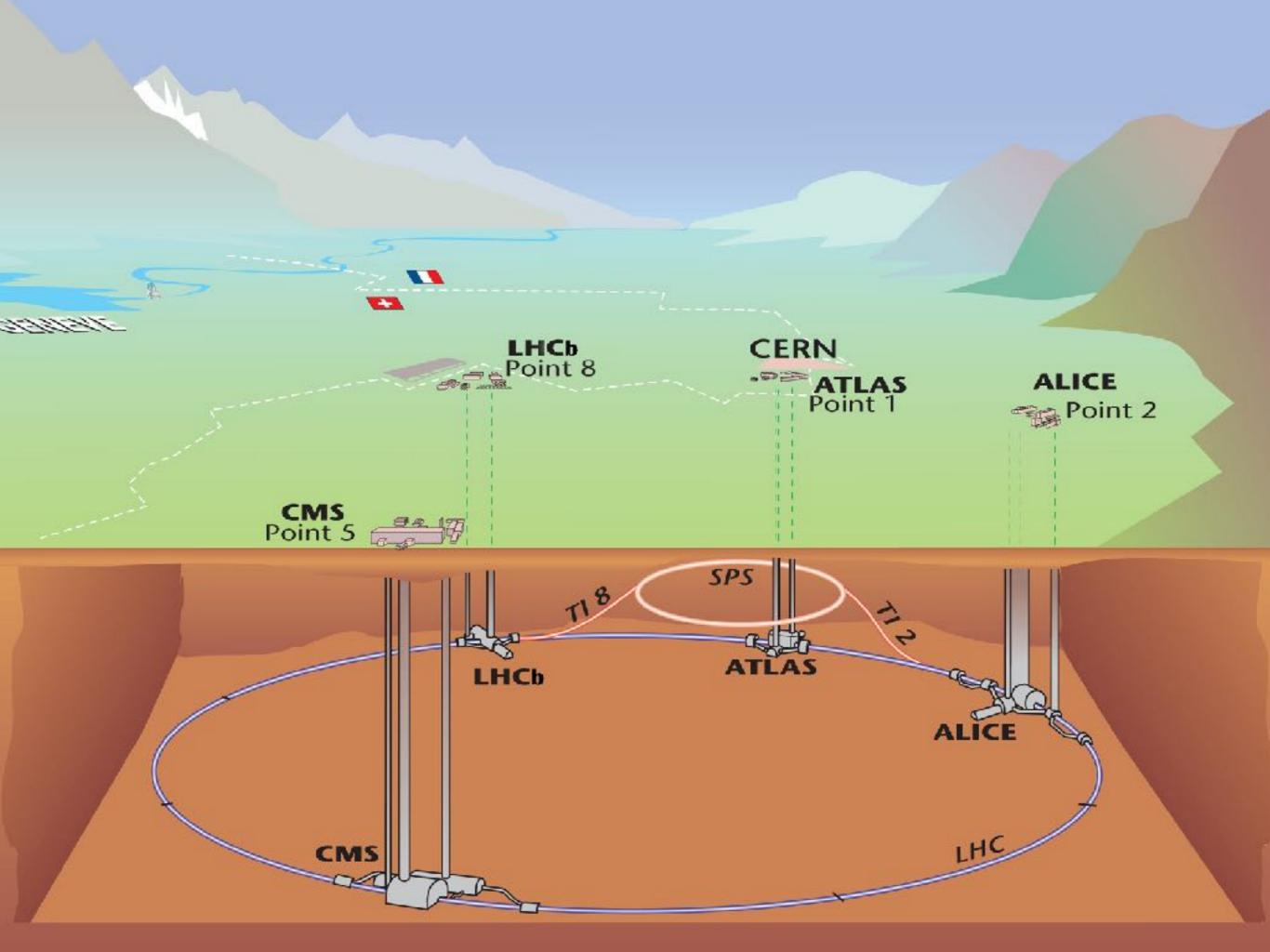
Brown, Caltech, Florida State, Notre Dame, Princeton, MIT, UVa, FNAL, INFN, Milano Bicocca, Roma1, Padova, Trieste, LIP, ETH Z, Sacaly

Trigger

Vienna, Beijing, Cyprus, Tallinn, LLR, NKUA, Ioannina, TIFR, Bologna, Padova, Warsaw-IEP, Warsaw-INS, CIEMAT, Bristol, RAL, ICL, FNAL, Boulder, Boston, MIT, UIC, TAMU, Davis, Northwestern, Florida, Rice, UCLA, Wisconsin, Rutgers, Princeton



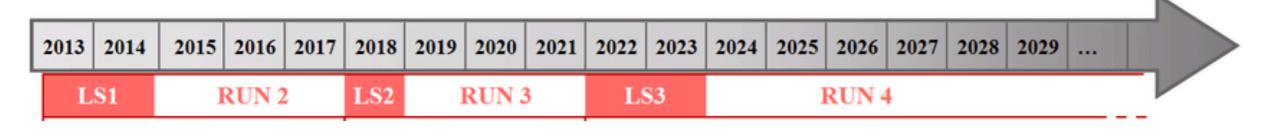






The LHC plans a program of Increased Luminosity over the next 10 years

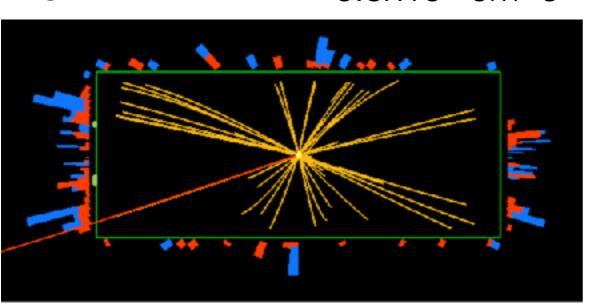
More data should lead to more precise measurements and searches with finer sensitivity

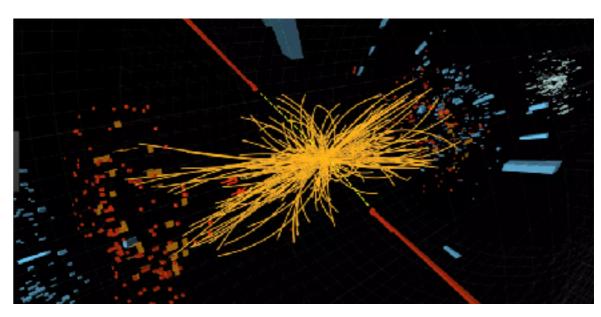


~0.5x10³⁴cm⁻²s⁻¹

2016

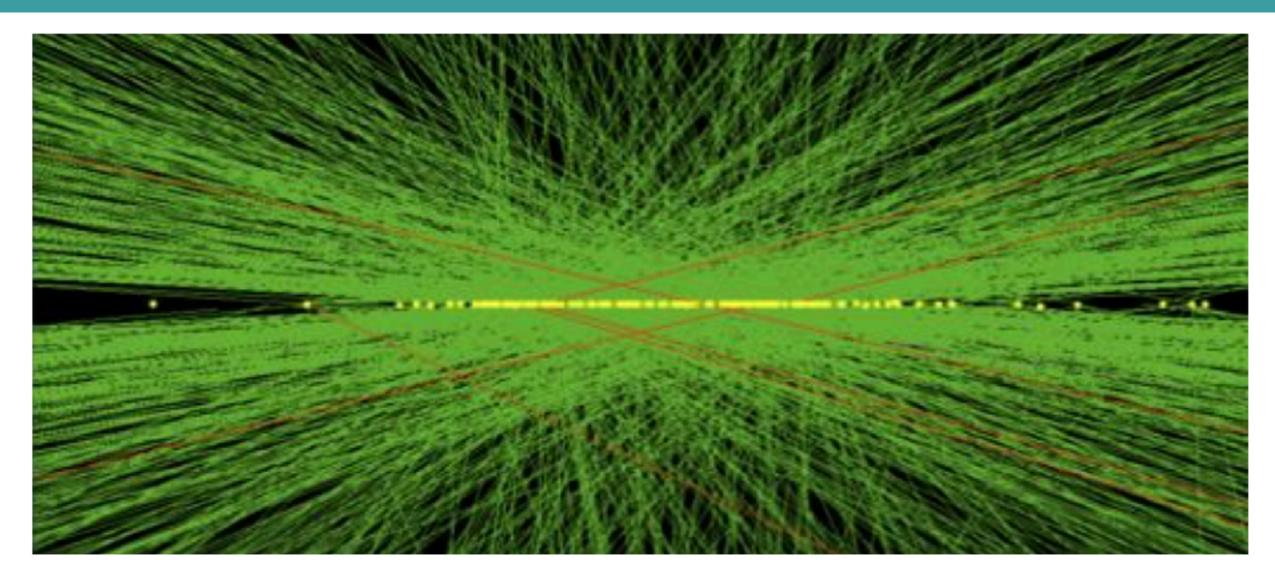
2011





2x10³⁴cm⁻²s⁻¹

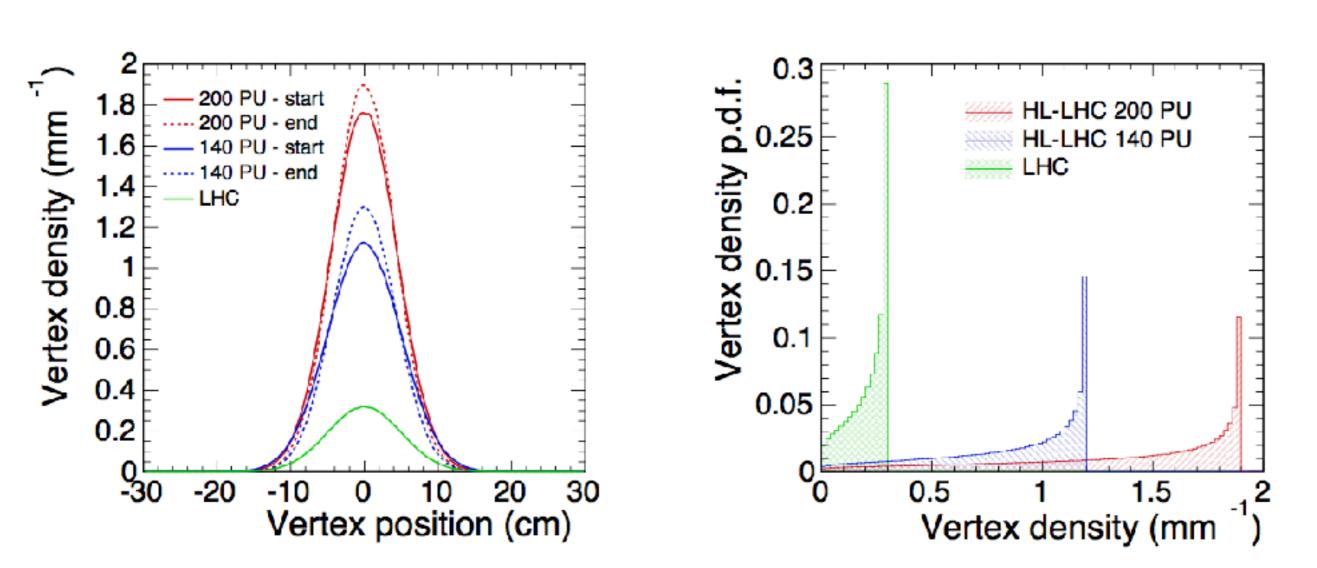
Challenges of the High Luminosity LHC (HL-LHC)



- Due to the increased instantaneous luminosity, the HL-LHC represents a significant challenge for Event Reconstruction and **Primary Vertex** identification
- While thousands of tracks, vertices and calorimeter clusters must be accurately reconstructed most of the tracker hits and calo deposits are due to Pile Up ... very little of this information is interesting for physics analysis! 5

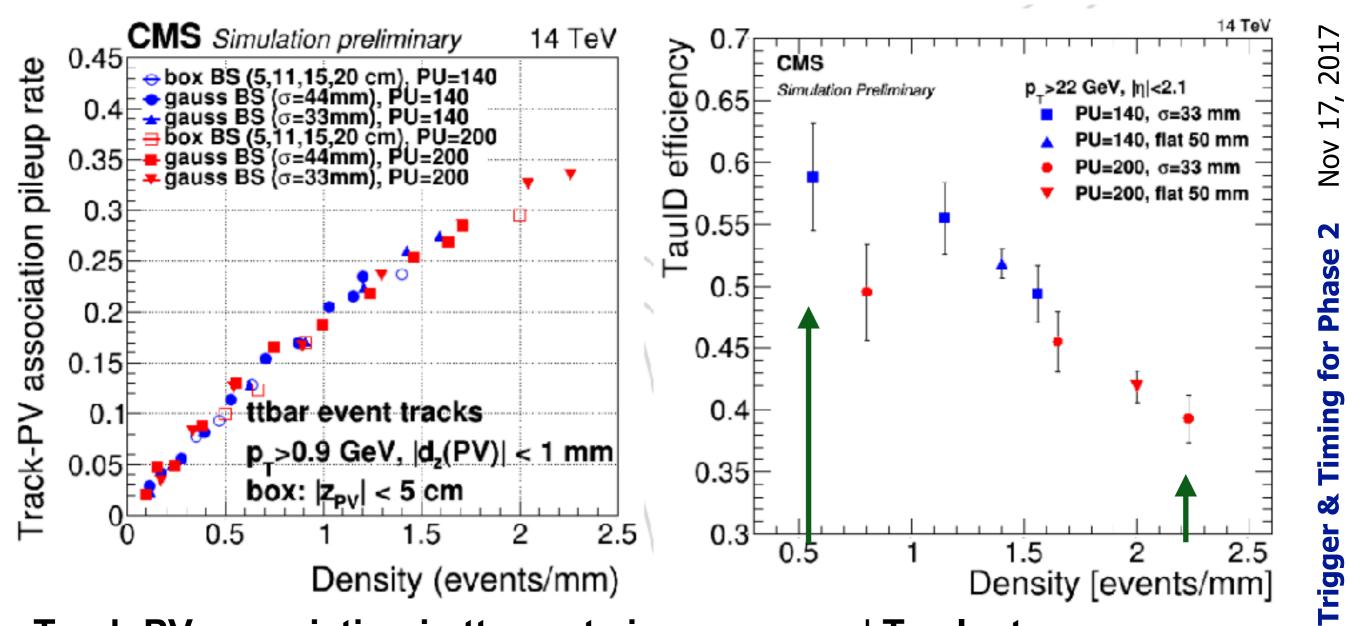


The Effects of High Pileup in Reconstruction



- Compared to current running conditions the HL-LHC will see a substantial increase in pileup line-density
- ► LHC Peak Vertex Density is 0.3–0.4 mm⁻¹, whereas at 140 and 200 pileup events, the peak vertex density is 1.3 mm⁻¹ and 1.9 mm⁻¹

The Effects of High Pileup in Reconstruction



- Track-PV association in tt events increases and Tau lepton
 Identification efficiency decreases as a function of vertex density
- This has a major effect on key SM and BSM analyses which are critical to the future success of CMS
 - Good Particle-Vertex association is essential to maintaining a good physics program at CMS
 7



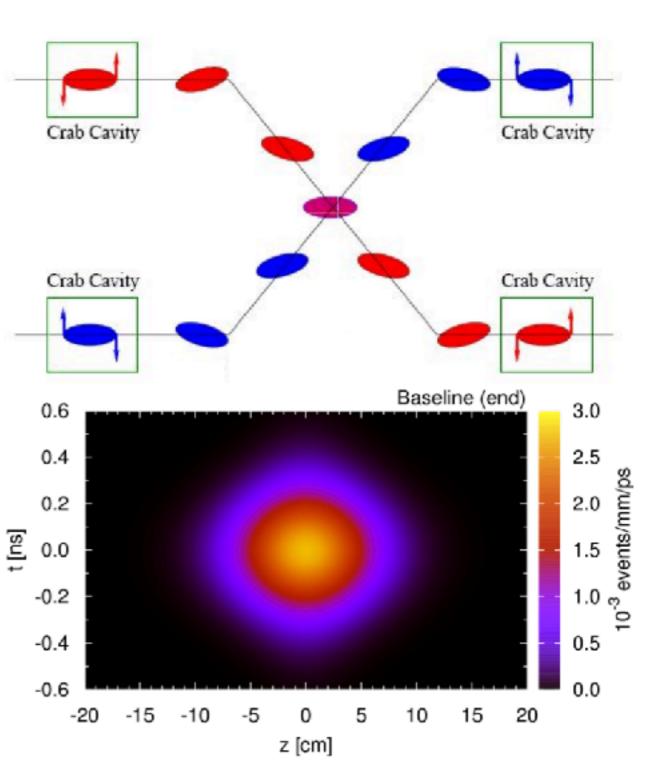
I.Ojalvo

Time Structure of Crossing Bunches

What can be done?

- During collisions, bunch crossing operates over a discrete time interval
- When bunches overlap entirely there is maximum pileup density as well as maximum spread in z-axis
- Normally CMS sees only the integral of this process over time
- Need to discriminate between vertices over an RMS of ~180 ps

8









Nov 17,

N

Timing for Phase

Frigger &

I.Ojalvo

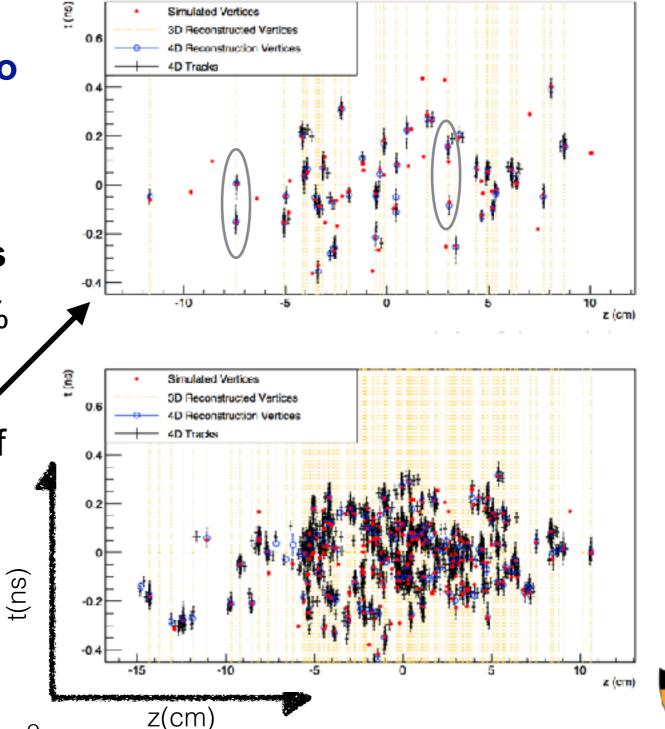
Effects of Timing Information on Track Reconstruction

Including a single layer of MIP Timing information to nominal Track Identification techniques

 Track Minimally Ionizing Particles (MIPs) not just by position but also by Time

Simulated time of flight resolution $\sigma_T = 30 \text{ ps}$

- 15% merged vertices reduce to 1.5%
 Purity of vertices recovered!
- In 50 PU figure, ample separation of previously merged vertices apparent at -7.3cm and 3cm
 - Separation of previously merged vertices notably present throughout 200 PU!



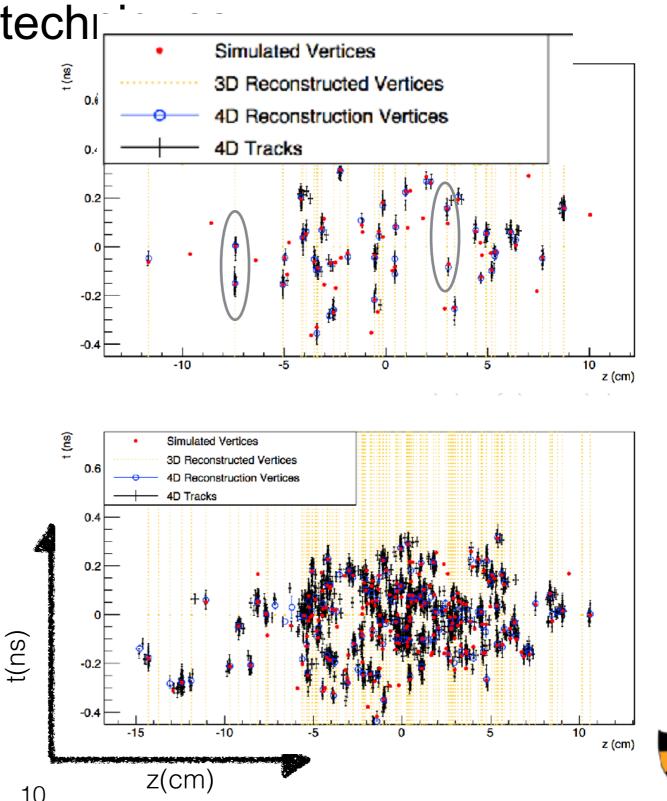
Effects of Timing Information on Track Reconstruction



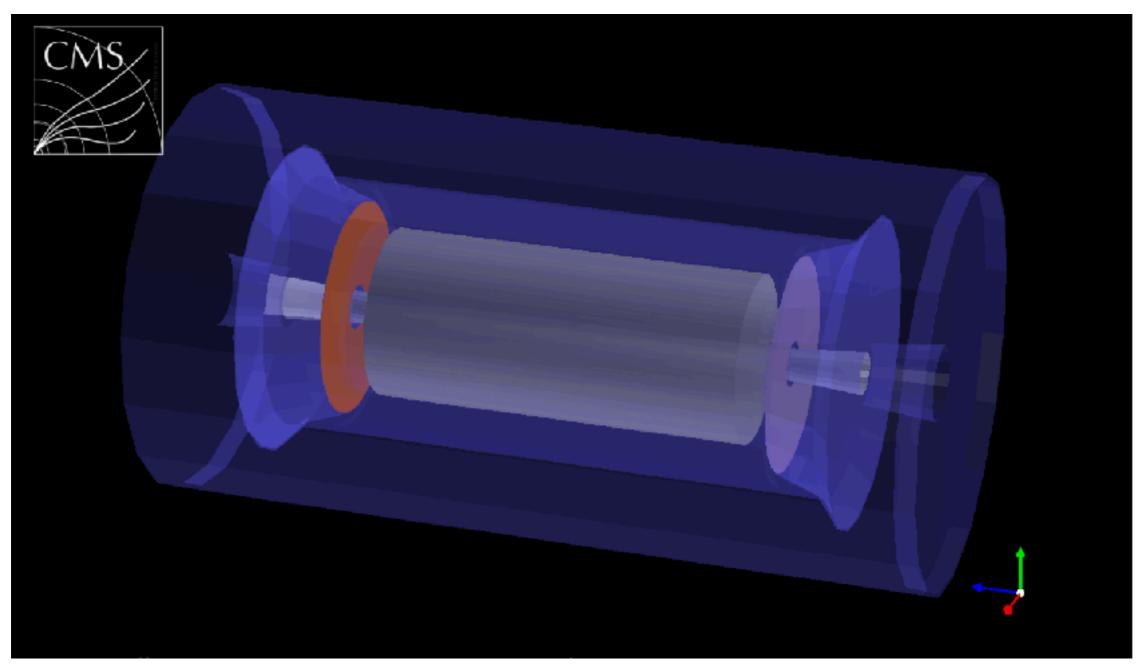
Including a single layer of MIP Timing information to

nominal Track Identification tech

- Time of flight resolution ~30 ps
- 15% merged vertices reduce to ~1.5%
 - Purity of vertices recovered!
- In 50 PU ample separation of previously merged vertices apparent at -7.3cm and 3cm
 - Separation of previously merged vertices notably present throughout 200 PU!



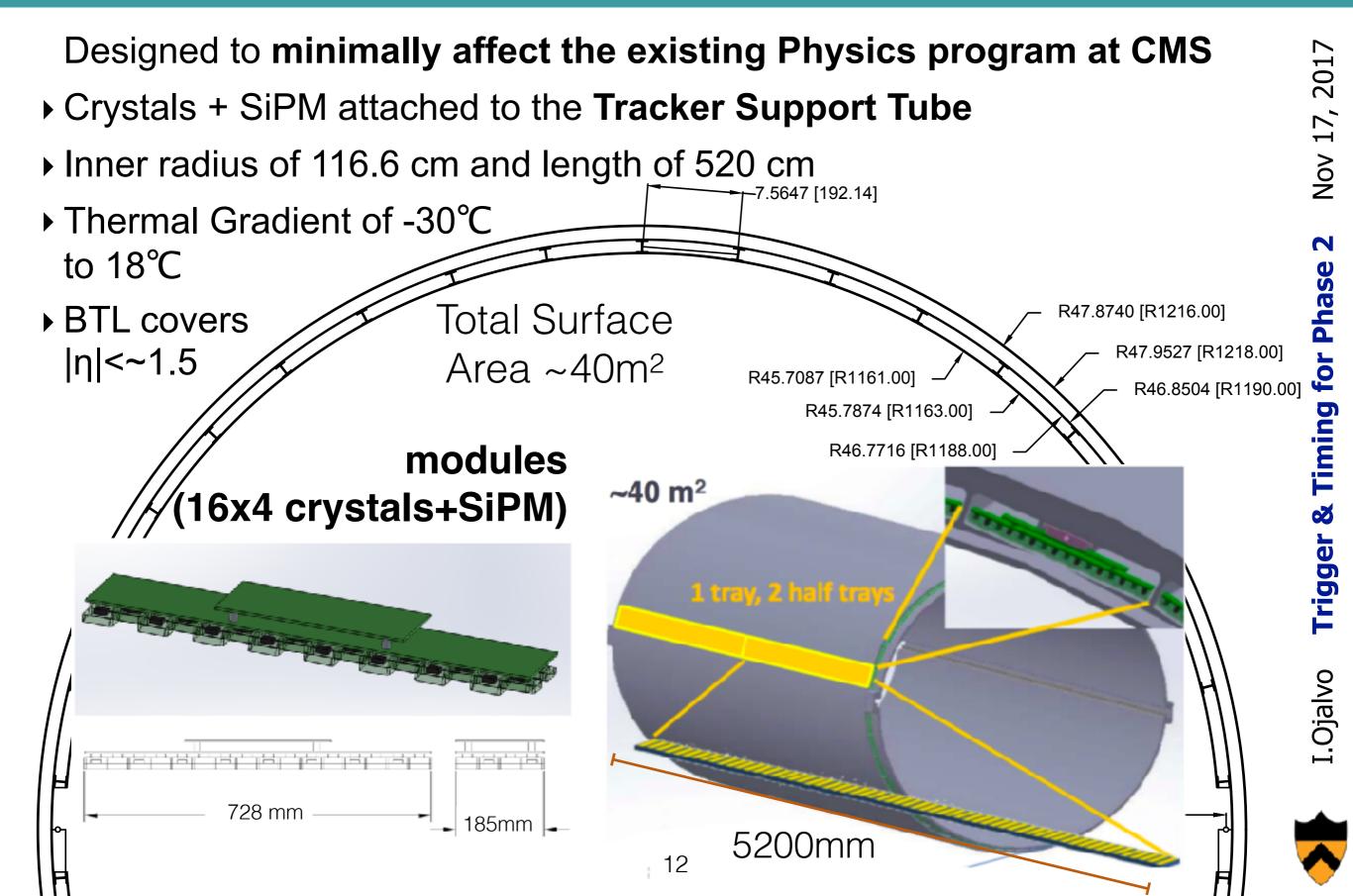
MIP Timing Detector Layout



- Additional thin MIP Timing Detector between tracker outer layer and ECAL Front End cooling plates
 - Just outside the tracker, Acceptance: $|\eta| < 3.0$ and $p_T > 0.7$ GeV
- Designed to have limited or negligible effect on tracker performance



Mechanical Structure (Barrel Timing Layer)



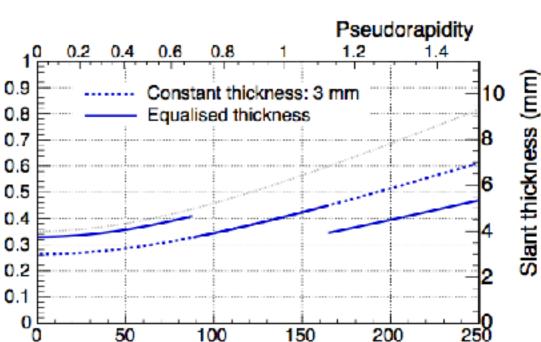
Fundamental Cell Design (Barrel Timing Layer)

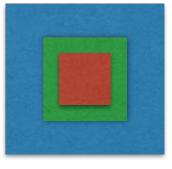
Thin Crystals with SiPM affixed to the back (thin detector!)

- 1 detector cell will be either 11×11 or 12×12 mm² coupled to SiPMs with a size of ~4×4 mm²
- In order to reduce the material budget in front of ECAL the crystal thickness will vary based on pseudo rapidity (position in the detector)
- Optimal Crystal width 3.75 to 2.4 mm
- Amount of light produced by a charged particle should not decrease with η
- Slant thickness traversed by particles scales like 1/sin(θ)
- ▶ Readout ASIC will be connected directly to the SIPM
 - Readout chip (TOFPET2) requires precision better than 20ps power consumption less than 20mW per channel

LYSO:Ce crystals read-out with silicon photomultipliers (SiPMs)







z (cm)



13

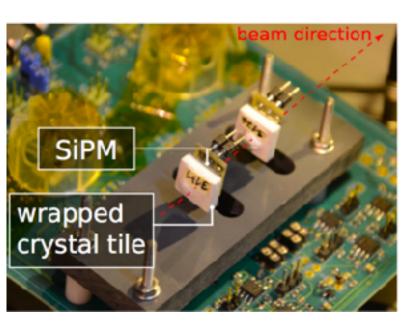
Slant thickness

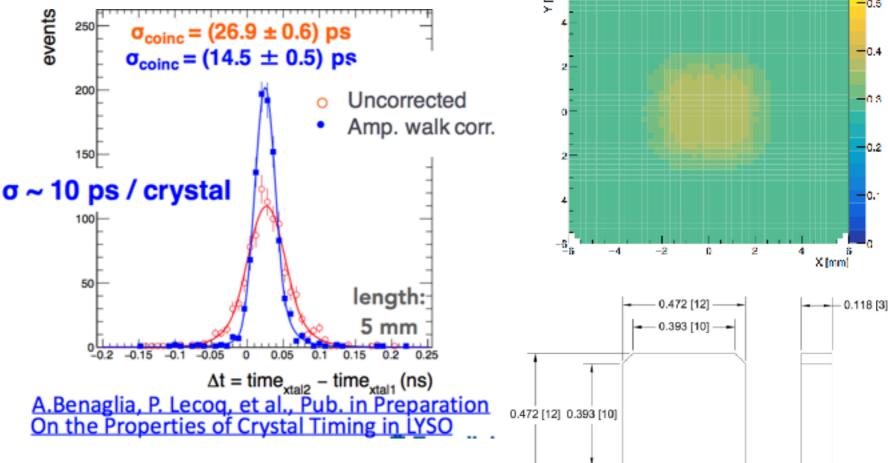


LYSO:CE crystals

- short decay time, high light output, excellent Scintillator
- In Test Beams, sensors with similar geometry as for the BTL have been proven capable of achieving MIP Timing resolution better than 30ps
 - In these devices a MIP was detected with 100% efficiency
 - Capable of withstanding 100MRad with minimal transparency loss

Overall, performance is simulated well

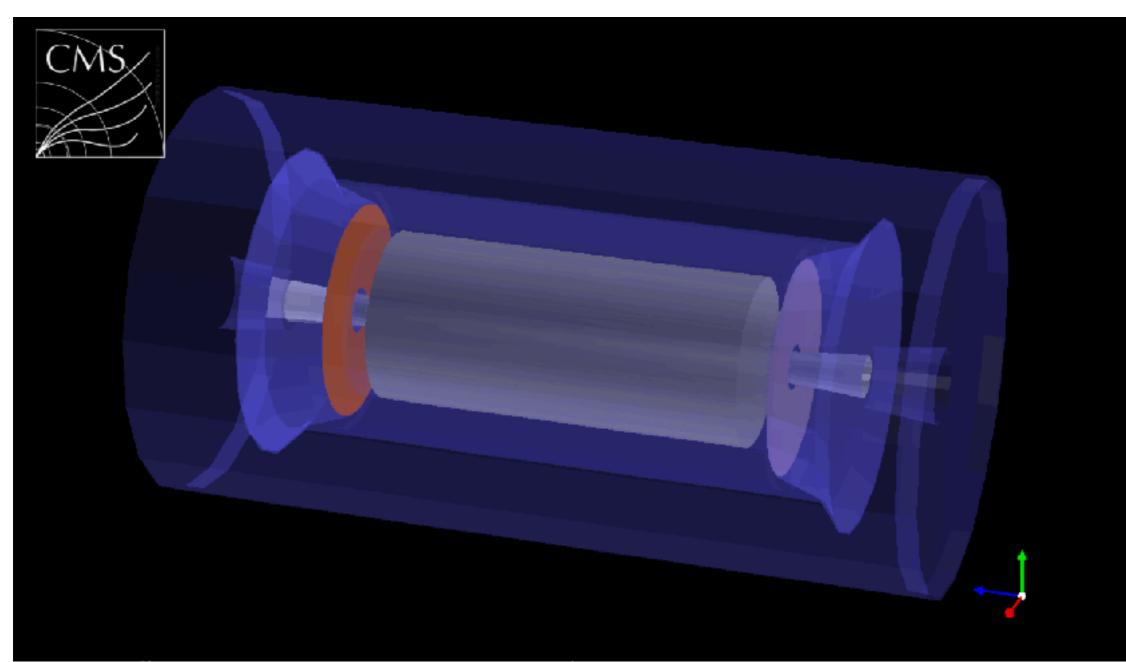




eant5, Preliminary Simulation - LSO Ce tile 12x12x3 mm² - SIPM 4x4 mm

Endcap Timing Layer

Sits in front of the HGCAL Detector, covers 1.6 < |η| < 2.9
 Due to its position a later installation could be possible

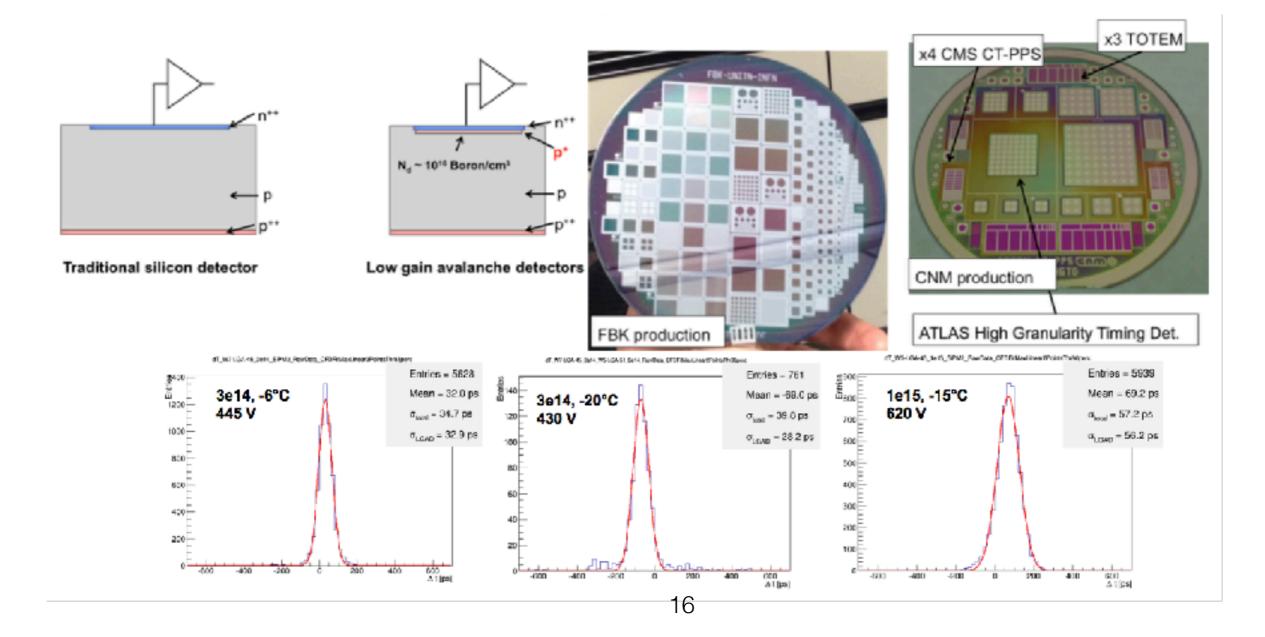






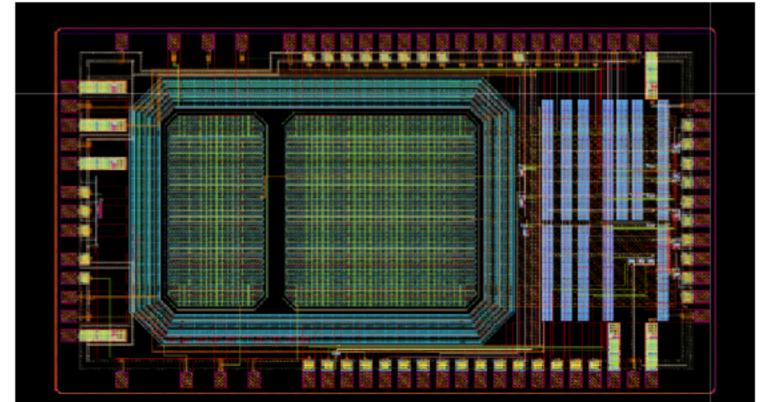
Low Gain Avalanche Devices (LGADs)

- Ultra Fast Silicon Detectors
- Optimize silicon sensor to increase dV/dt (gain), reduce shot noise, decrease landau fluctuations (thinning to 50 um)
- < 30ps resolution achieved up to a fluence of 3e14 n. eq.</p>



Detector Readout

- To with stand the HL-LHC occupancy the cell size is chosen such that the occupancy will only be a few percent
- Due to proximity to sensors, requires minimal EMI
- Requires low-power, radiation hard and high bandwidth preamps



- Current TOFPET2 chip is a 64 channel ASIC based on CMOS 110 nm technology (radiation tolerant up to ~ 10 MRad) for the Barrel
- Dedicated Read-Out Chip being developed for the Endcap





Test Beam CERN Aug. 30 to Sep. 06

BTL Configuration Testing

- Study reproducibility of signal amplitude and time resolution with many channels
 - Using sensors + module assembly
- Integrate TOFPET2 DAQ with test beam infrastructure
- Crystals: 16 11x11x2 mm³ (lat faces unpolished)
- SiPM: 16 6x6 mm² HPK 50 um (S13360-6050CS)

Single Sensor Testing

- (Re-)Test 9-in-1 SiPM array from FBK
 - Improved alignment with Photek, aluminization of area between SiPMs

Study uniformity with different surface states:

- naked, rear Teflon-wrapping, full Teflon-wrapping, Vikuiti reflector, black paint on back of SiPM
- Study new SiPMs from FBK:

Device, Size, Manufacturer

6x6 mm² HPK 50 um

6x6 mm² FBK NUV-HD 30 um

5x5 mm² FBK NUV-HD 20 um

4x4 mm² FBK NUV-HD 25 um ¹⁸



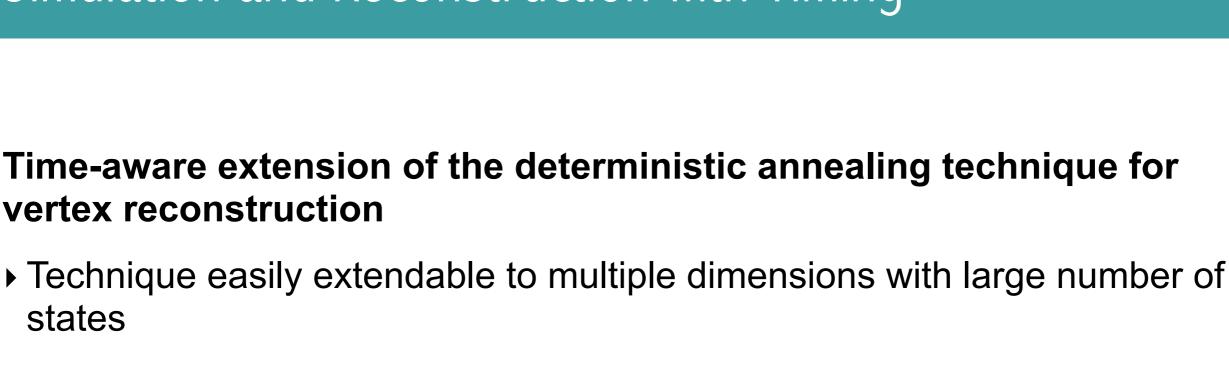






Reconstruction and a Few Performance Results





- Significant impact on b-tagging, pile up jet ID, ETMiss calculations and **Isolation (especially for Taus)**
- Further extended to tie individual photons reconstructed in the calorimeters to a collision vertex
 - First determine a reference time for EM shower

states

Then assign it in a straight trajectory to the vertex with corresponding zposition (includes time-of-flight corrections)



Time-aware extension of the deterministic annealing technique for vertex reconstruction



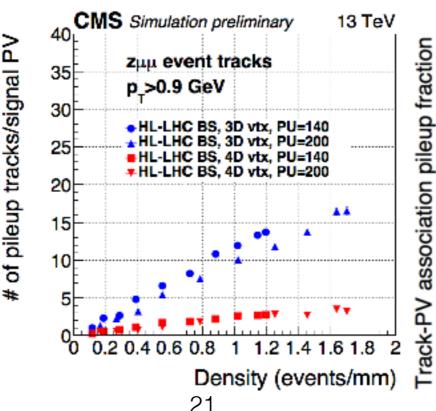
Reduction on Pileup Tracks associated to Hard Interaction

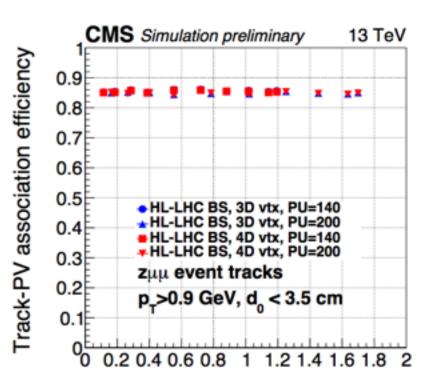
- Reducing the number of tracks from Pile
 Up associated to the Hard Interaction is an
 essential step to reducing the effects from a
 High PU environment
- In Run II, a cut of |Δz(track,PV)| < 1mm was commonly used
- With Timing, augment with a selection of [Δt(track,PV)] < 60ps equivalent to ~3σ Δt(track, PV)

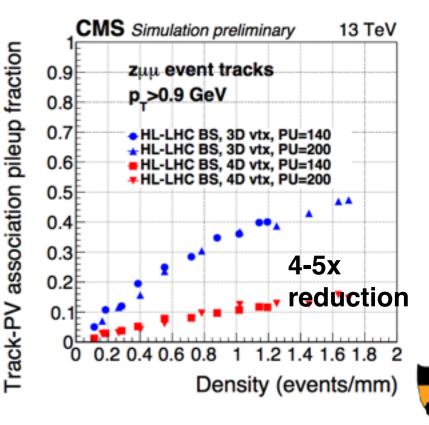


 True Track-PV efficiency preserved

Pile Up Track-PV association reduced with PV density

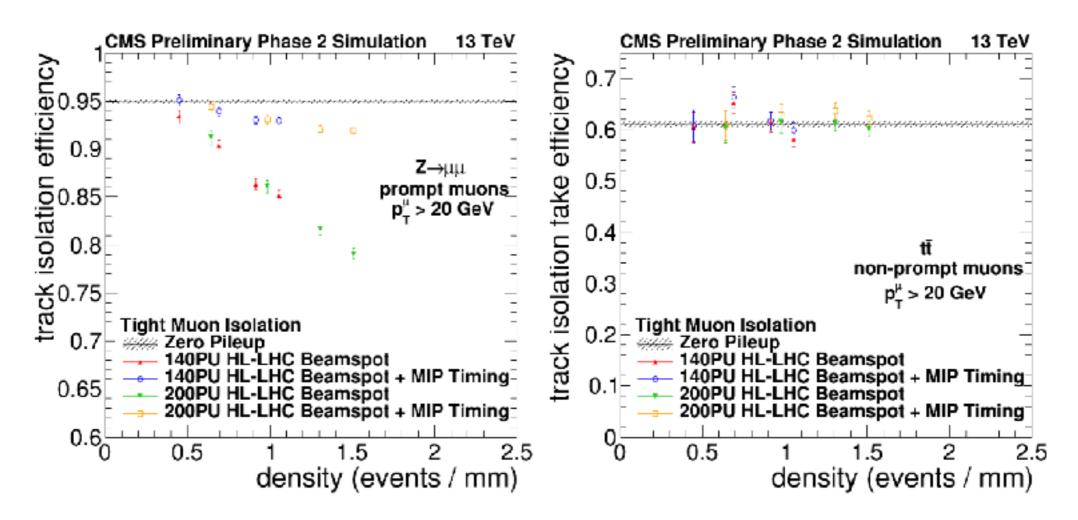






Muon Isolation

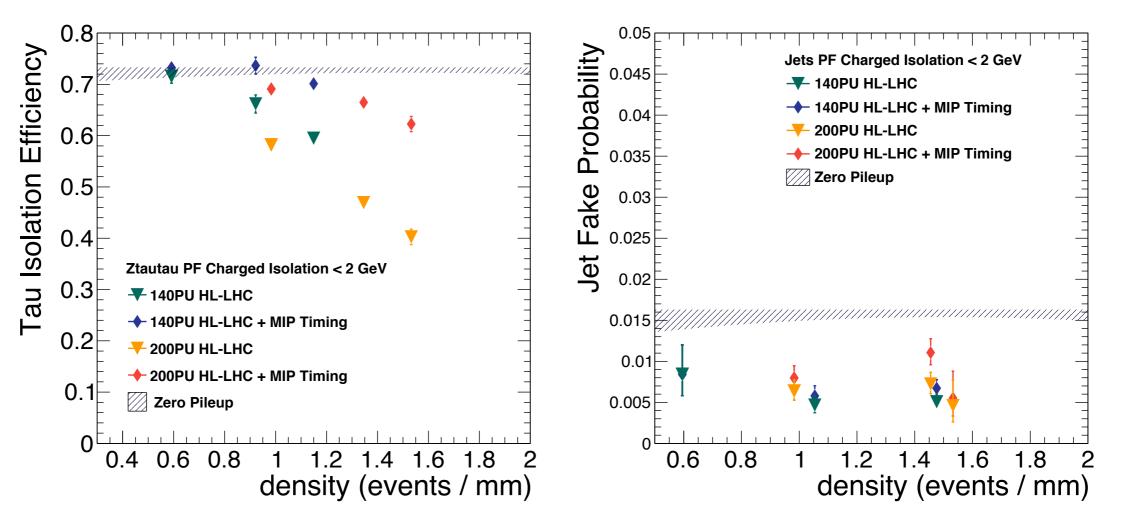




- Charged Particles comprise the largest fraction of hadronic activity in the event
- Charged Isolation sums are the most important contribution to isolation sums in the context of identifying true isolated leptons and photons
 Number of PU tracks associated to the PV are reduced with the MTD
- Showing Tight WP -> Relative Isolation of < 0.05</p>



Tau Isolation



- Due to their composite nature (1 prong, 1 prong + π⁰, 3 prong, 3 prong + π⁰) Taus are very susceptible to the effects of pileup
- Even for Run II Tau fake probability is high compared to signal efficiency
 Ability to distinguish between PU hadrons and True hadrons is
 essential to performing H->TauTau analysis with HL-LHC
- Performance is measured using PF Charged Isolation with PV-matched Charged Hadrons with and without timing



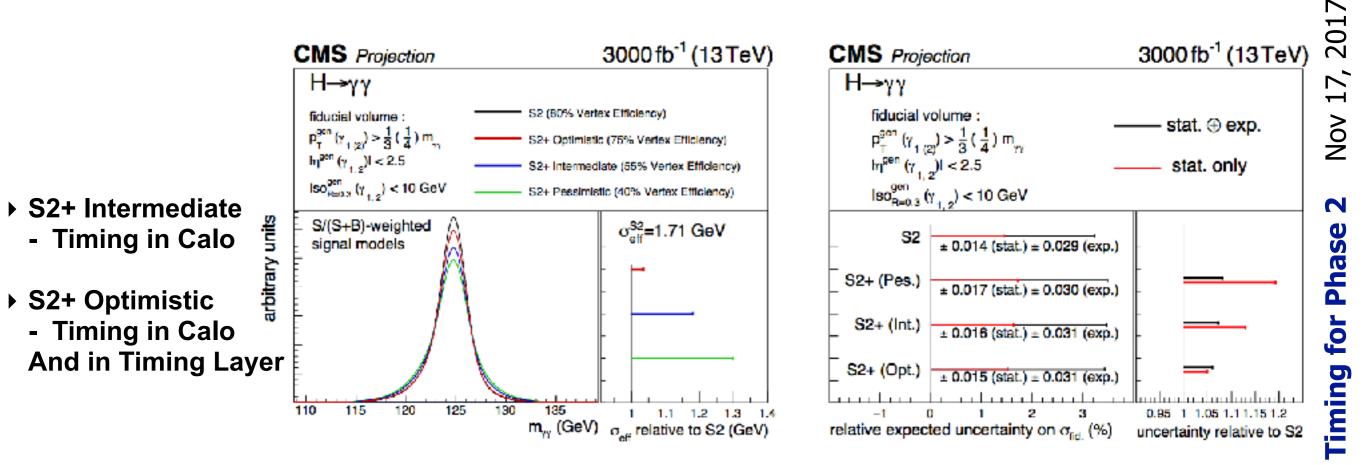
H->Gammagamma



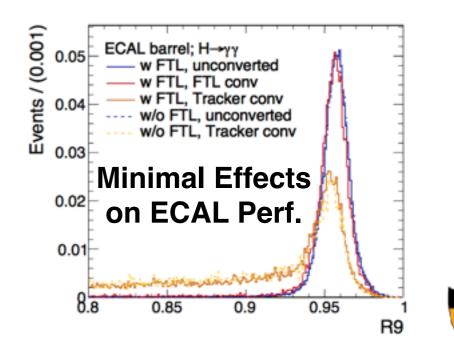
Ø

Trigger

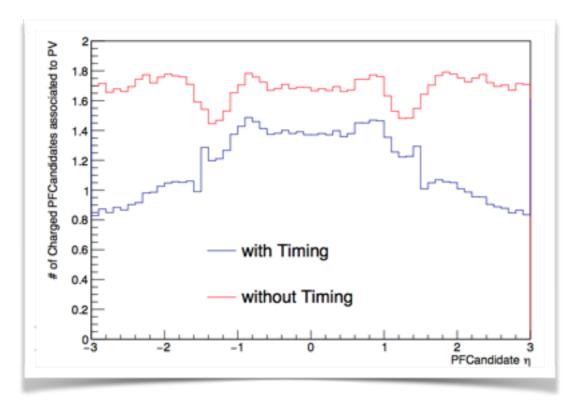
.Ojalvo

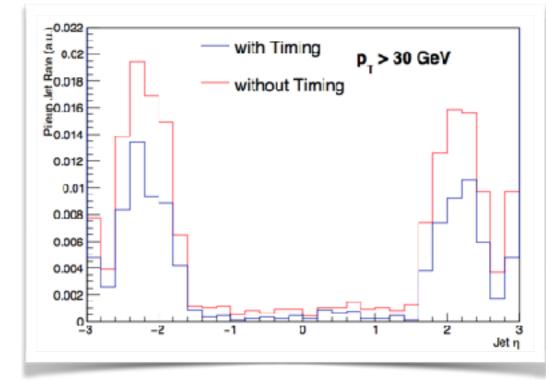


- Improvements correspond to approximately 30% increase in total integrated luminosity
- Primary vertex efficiency is 75% -> Nearly recovering current running conditions
- Photons which convert in the MTD volume have similarly narrow showers to those which are unconverted up to the ECAL face

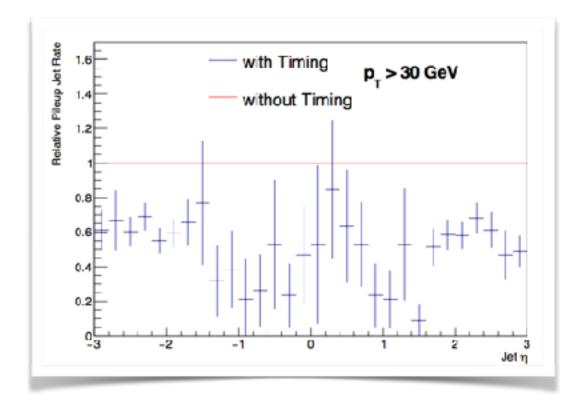


Jets and PF Charged Hadron Multiplicity





- Track Vertex association with Timing requires Δz<1mm and Δt<90ps
- Multiplicity of Charged Particles significantly reduced with Timing
- Addition of timing reduces pile up jet probability by 35 to 50%







Trigger System at CMS





Trigger System at CMS

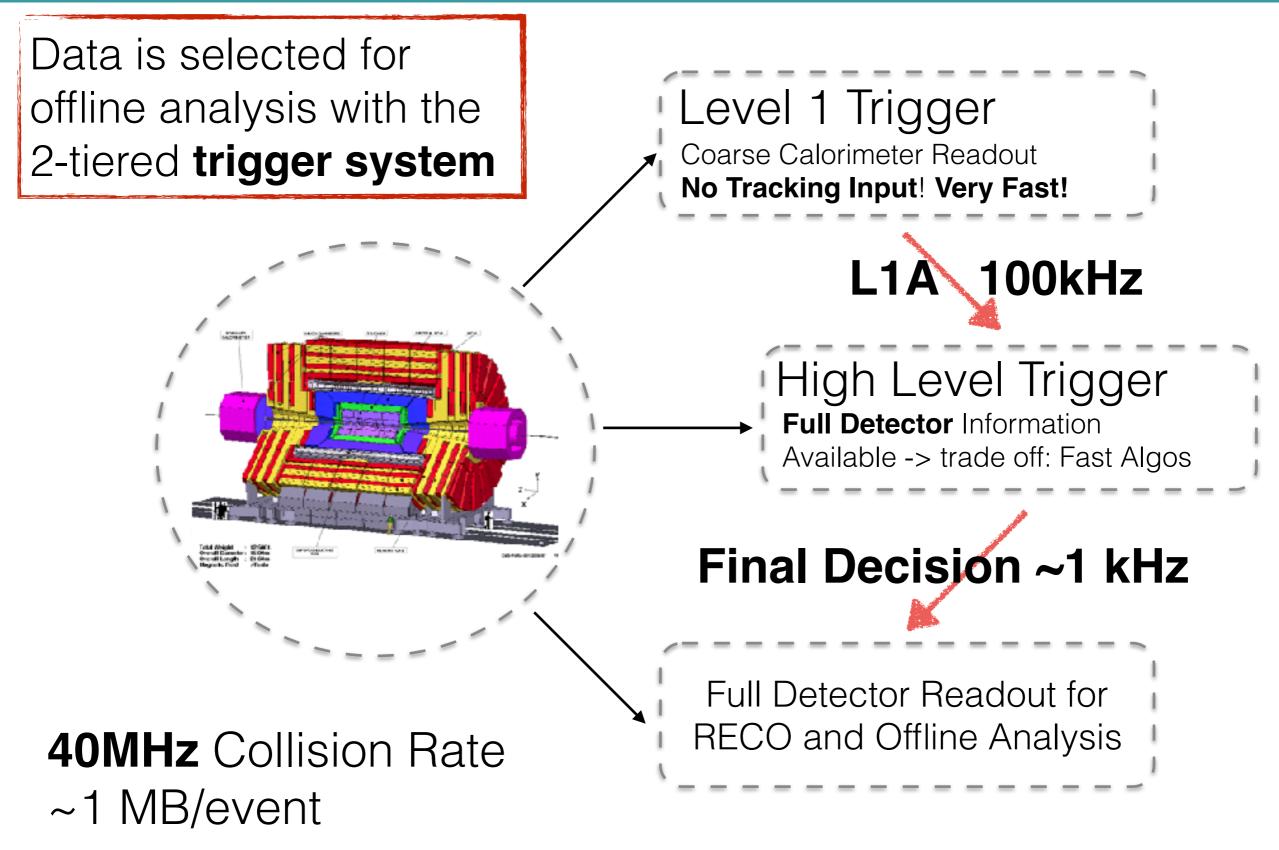


2017

Vov 17,

N

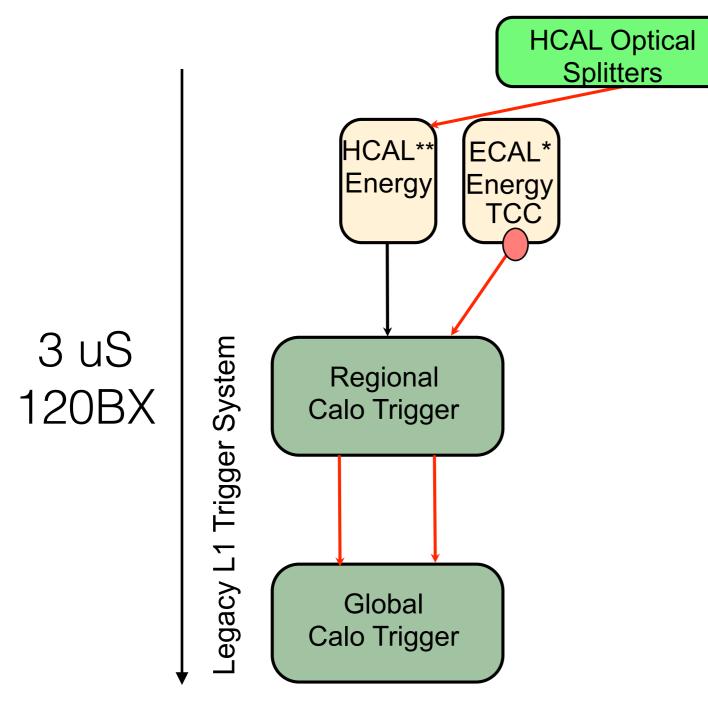
Trigger & Timing for Phase





I.Ojalvo

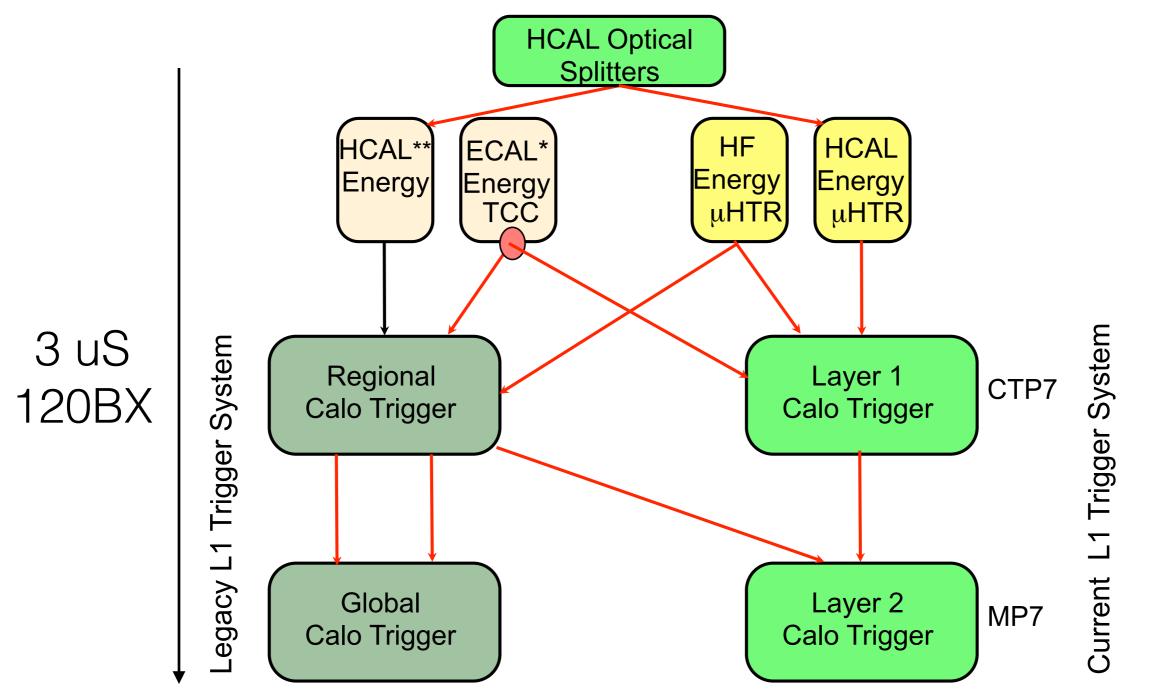
Calorimeter Trigger Design Run I System



- Coarse inputs from the Calorimeters
- Generate a trigger within 3 uS at a max rate of 100kHz
- Upgraded system installed in parallel, extra optical links



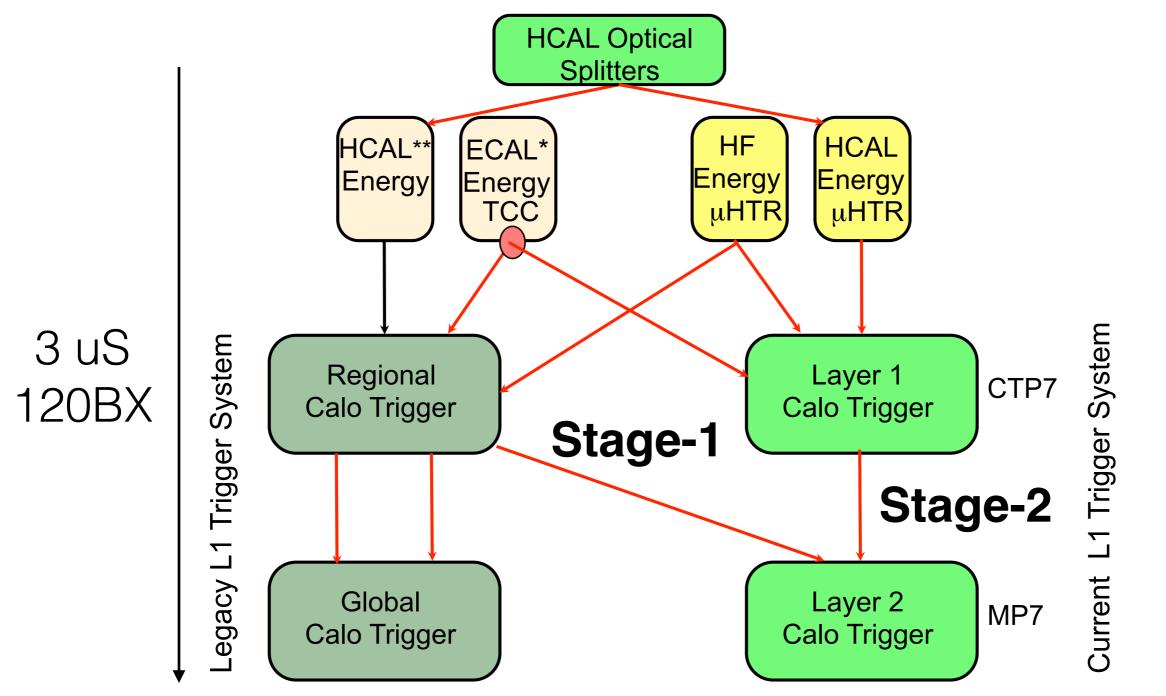
Calorimeter Trigger Design Current (Run II) System



 Trigger system upgrade policy has been to have the new system be installed mostly in parallel while we commission the new system
 Much needed safety for an essential system!



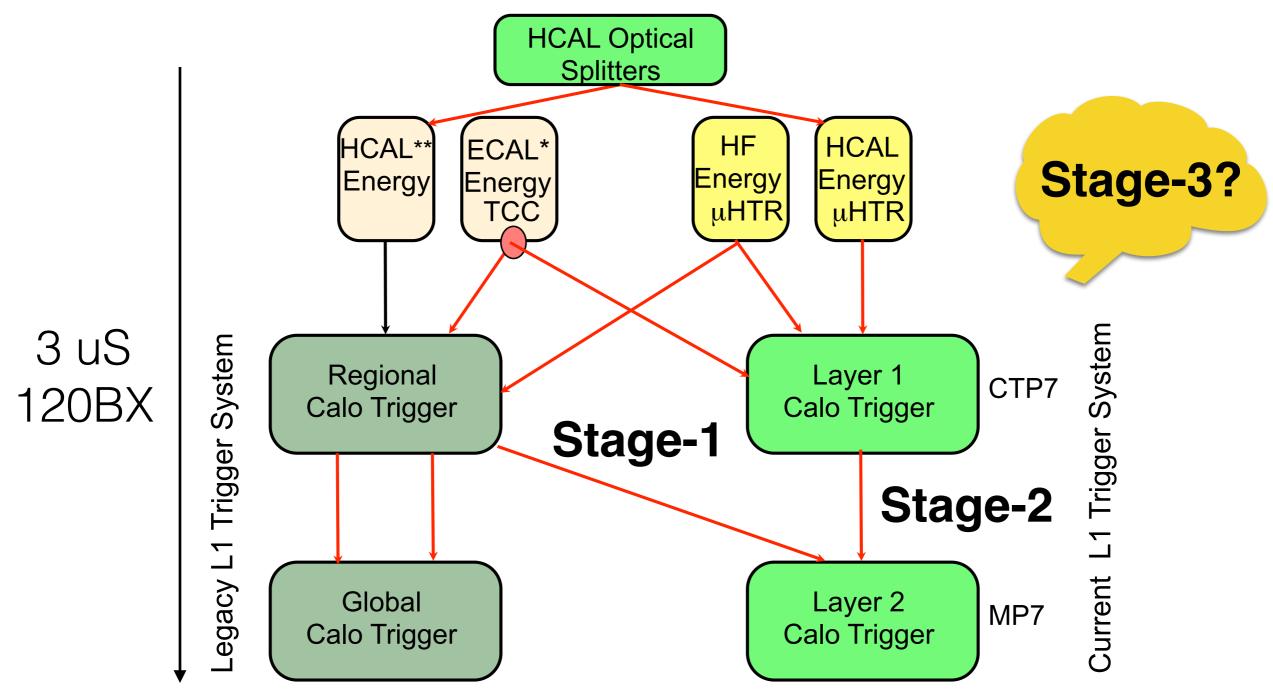
Calorimeter Trigger Design Current (Run II) System



 Trigger system upgrade policy has been to have the new system be installed mostly in parallel while we commission the new system
 Much needed safety for an essential system!



Calorimeter Trigger Design Current (Run II) System



 Trigger system upgrade policy has been to have the new system be installed mostly in parallel while we commission the new system
 Much needed safety for an essential system!



Current Boards Phase 1



31 Rx and 12 Tx Frontpanel Optical 10G links on MiniPODs

36 Tx and Rx Frontpanel Optical 10G links on CXP Modules **A 3 GTH** Back-plane Tx/Rx links

University of Wisconsin - Madison

- Virtex 7 used as main processor board
 - Half a Million Logic Cells, Up to 500MHz Clock Frequency, MultiGigabit transceivers
- Large amount of I/O
- CTP7 uses Embedded Linux for ancillary functions
- Firmware storage on uSD cards
- 30 boards installed at CMS



CTP7 Main Processor





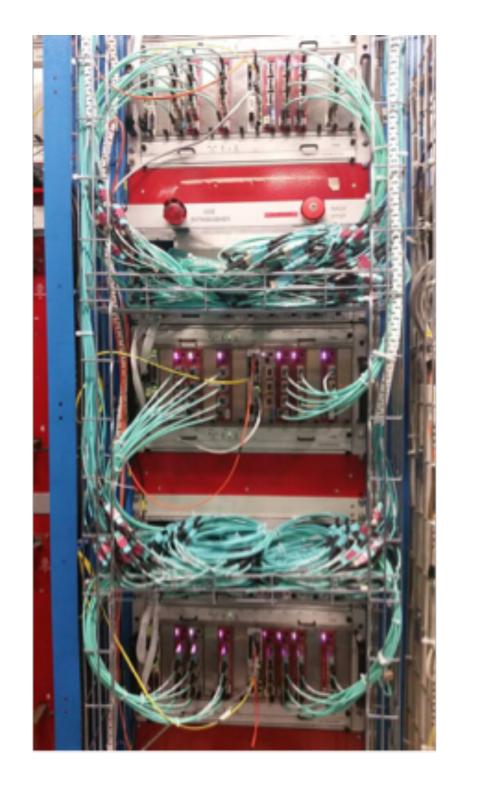
Installation

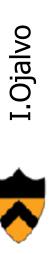


Nov 17, 2017

Trigger & Timing for Phase 2









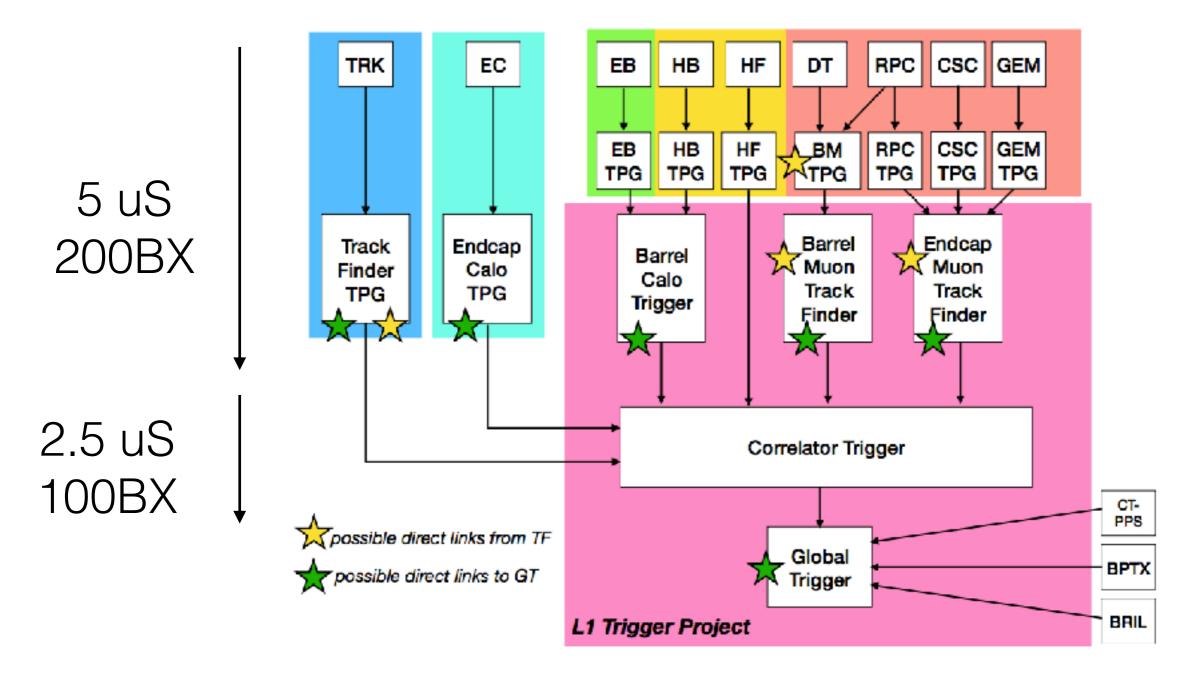
CMS Trigger System for HL-LHC





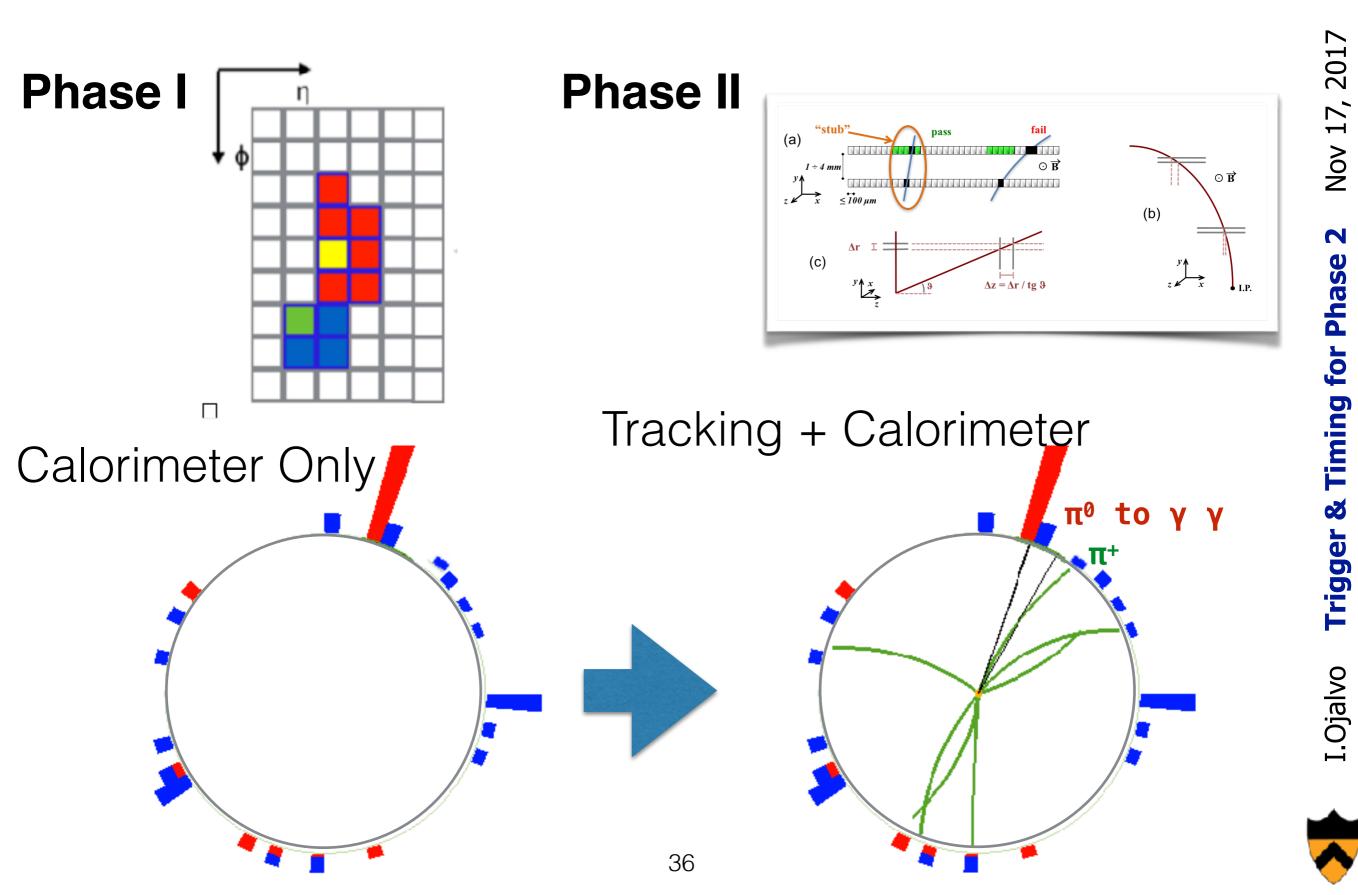
Trigger Design Phase-II (HL-LHC) Upgrade System

Tracking system for input, improved ECAL granularity, HGCAL
Generate a trigger within 12.5 uS at a max rate of 750kHz





Algorithms for Phase II







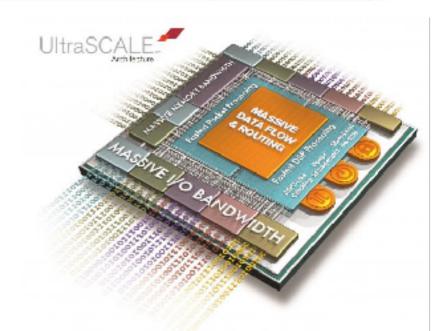
Overview on main technologies

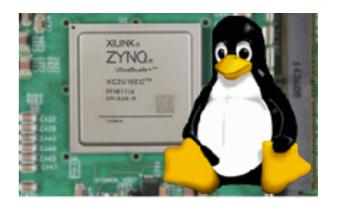
Phase II Trigger Upgrade is attempting to bring more offline reconstruction to Level 1 while increasing flexibility

What is needed?

- Large, multi-purpose Ultrascale+ class FPGAs
 - Industry FPGA size appears to be more than doubling with each generation
 - Programmable Systems are excellent for improving algorithms over time!
- Multi-Gigabit Transceiver Links 16 to 25 Gb/s
- Advanced Telecommunications Architecture (ATCA) Form Factor
 - Better Form Factor for board routing
- IPMI and Embedded Linux Solutions 37

	KINTEX?	KINTEX	VIRTEX.	VIRTEX
Logic Cells (LC)	478	1,161	1,995	4,407
Block RAM (BRAM (Mbits)	34	76	68	115
DSP-48	1,920	5,520	3,600	2,880
Peak DSP Performance (GMACs)	2,845	8,180	5,335	4,268
Transceiver Count	32	64	96	104
Peak Transceiver Line Rate (Gb/s)	12.5	16.3	28.05	32.75
Peak Transceiver Bandwidth (Gb/s)	800	2,086	2,784	5,101
PCI Express Blocks	1	4	4	6
100G Ethernet Blocks		2	•	7
150G Interlaken Blocks		1	•	9
Memory Interface Performance (Mb/s)	1,866	2,400	1,866	2,400
VO Pins	500	832	1,200	1,456

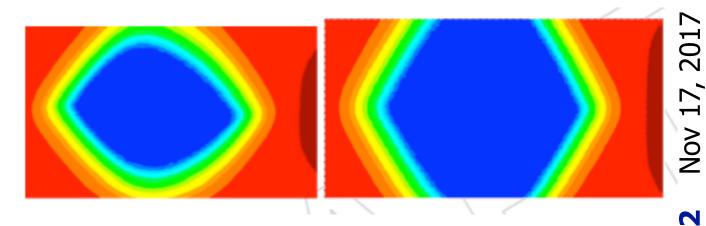




R&D Program

Key technologies/challenges :

- ▶FireFly Modules, Multi-Gbit transceivers/optics
- ►ATCA form factor
- ▶Embedded Linux
- ▶Large RAM
- Power delivery and thermal management
- System level integration & maintenance

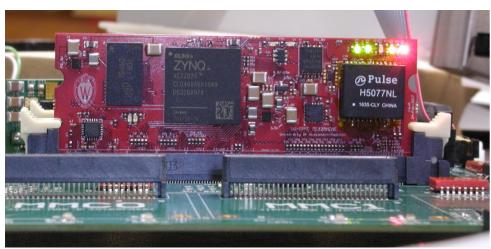


Eye diagrams for 26Gb/s transceivers

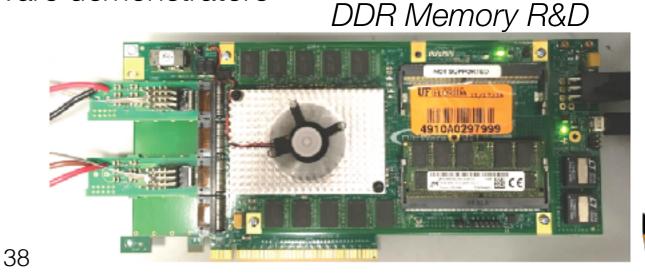


Hardware demonstrators





IPMS - IPMI for ATCA





Timing for Phase

ø

Trigger

I.Ojalvo

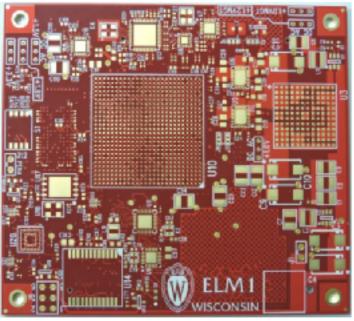
Development Boards



Advanced Processor Development Board (APd1)

- Targets a single high-end Xilinx FPGA (C2104 package)
- Allows up to 96 optical link pairs
- Supports speeds up to 16 or 25 Gb/s or a mix of both using Samtec Firefly Modules
- Carry two mezzanines: IPMI controller, Embedded Linux endpoint
 both based on the Xilinx ZYNQ platform
- Partial prototyping taking place on the Controller Development Board (CDB)
- The CDB will also be the initial test platform for 10 Gb/s Ethernet, on-board Ethernet switching and persistence storage on SSDs

Embedded Linux Mezzanine



ATCA Service Card (Core Infrastructure)

- Modular Processing Path supports different FPGAs
- Processing sites support up to 96 links at 26 GB/s
- Primary goal reduce manufacture risk given the high

value of the FPGAs

 Interconnect is a new technology and will need careful evaluation











CMS

2017

Nov 17,

N

Trigger & Timing for Phase

- HLS is an automated design process that interprets algorithm specification at a high abstraction level and creates digital hardware/ RTL code that implements that behavior
- HLS significantly accelerates design time while keeping full control over the choice of optimal architecture exploration, proper level of parallelism and implementation constraints
- Reduces overall verification effort
- Several HLS options in use: product / vendor:
 - Catapult-C / Calvpto Design Systems
 - BlueSpec / BlueSpec Inc.
 - Symphony C / Synopsys
 - MaxCompiler / Maxeler
 - Cvnthesizer / Cadence
 - HDL Coder / MathWorks (Matlab)
 - OpenCL (Intel/Altera)
 - Vivado HLS / Xilinx

Pioneered @ CMS by A. Svetek + Team at UW Madison

Xilinx Vivado HLS

Vivado HLS C/C++ libraries contain functions and constructs that are optimized for implementation in an FPGA.

Using these libraries helps to ensure high Quality of Results (QoR)

final output is a high-performance design that makes optimal use of the FPGA resources.

Vivado HLS also provides additional libraries to extend the standard C/C++ languages:

- <u>Arbitrary precision data type (e.g. 5-bit unsigned integer: ap_uint<5>)</u>
- Fixed-point data type (e.g 18-bit integer with 6 bits above binary point:
 - ap_fixed<18,6,AP_RND >
- Half-precision (16-bit) floating-point data types
- Math and video operations, Xilinx IP functions (FFT, FIR)

Radar Design (1024x64 floating-point QRD)	RTL Approach (VHDL)	Vivado Hls
Design Time (weeks)	12	1
Latency (ms)	37	21
Resources:		1
- BRAMS	273	38
• FFs	29,686	14,263
• LUTs	28,512	24,257

Realistic? Let's evaluate!

UW-Madison actively studying Xilinx Vivado HLS tools, in a joint effort with with **Princeton**, and TIFR. Utilizing a test platform using a CTP7 Virtex-7 board to evaluate HLS algorithms in actual hardware. Promising results so far,







HLS caveats, traps and pit-falls

It's a disruptive technology

it implies **change in the methodologies**, in the design processes, and to some extent, in the skills required.

Not all C/C++ coding styles are equal in terms of QoR, and there is still the potential for ending up with poor quality RTL when the C++ code is not suited for HLS

Required resource estimates provided by HLS tools might not be always reliable and need careful checking

Good style not only requires an understanding of the underlying hardware architecture of an algorithm, so that it is reflected in the C++ design, but also an understanding of how HLS works.



Example Algorithm with HLS

Example: HLS algorithm to compute "Pile Up" Level as part of CMS Trigger Calorimeter Logic

```
#include <stdio.h>
    #include "ap int.h"
 2
    #define NR_CALO_REG ( (6 + 7) * 2 * 18) // 468
 4
    #define PUM_LEVEL_BITSIZE (9)
 5
 6
 7 // helper function to count number of bits set in "bitString"
 80 ap_uint<PUM_LEVEL_BITSIZE> popcount(ap_uint<NR_CALO_REG> bitString)
 9
10
        ap uint<PUM LEVEL BITSIZE> popcnt = 0;
        loop popcnt: for (int b = 0; b < NR CALO REG; b++)</pre>
110
12
        {
13
    #pragma HLS unroll
14
            popcnt += ((bitString >> b) & 1);
15
        }
16
        return popent;
17 }
18
19⊖ ap_uint<PUM_LEVEL_BITSIZE> UCT_pum_level_impl3
            ap_uint<10> region_et[NR CALO REG],
20
21
            ap uint<10> pum thr)
22
    #pragma HLS PIPELINE II=6 // target clk freq: 250 MHz (~6 clks/BX)
23
    #pragma HLS ARRAY_RESHAPE variable=region_et complete dim=1
24
25
26
        ap uint<NR CALO REG> tmp = 0; // important: do var init
27
28⊝
        loop_pum: for (int idx = 0; idx < NR_CALO_REG; idx++)</pre>
29
        {
30
    #pragma HLS UNROLL // fully unroll the loop
310
            if (region et[idx] > pum thr)
32
                tmp.set bit(idx, true);
33
            else
34
                tmp.set bit(idx, false); // !! The only difference with impl2 !!
35
        }
36
37
        return popcount(tmp);
38 }
```

- Fully ANSI compliant C impl.
- Vivado HLS compiler guided by the user with **#pragma** directives
- HLS impl. is significantly easier to validate compared to traditional HDL approach. It also produces better results compared to HDL in several studied cases.

FPGA LUT count	2996
V7690T LUT [%]	~ 0.6
Latency in clk cycles @ 250 MHz	3

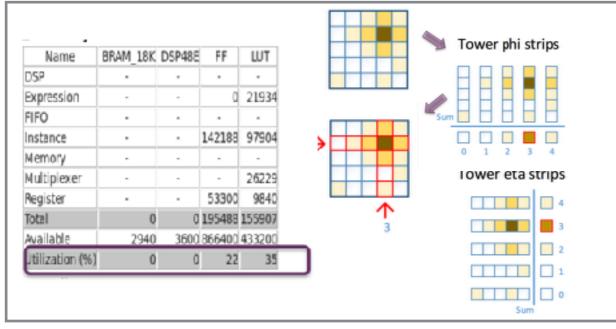


Moving Forward: HLS Development

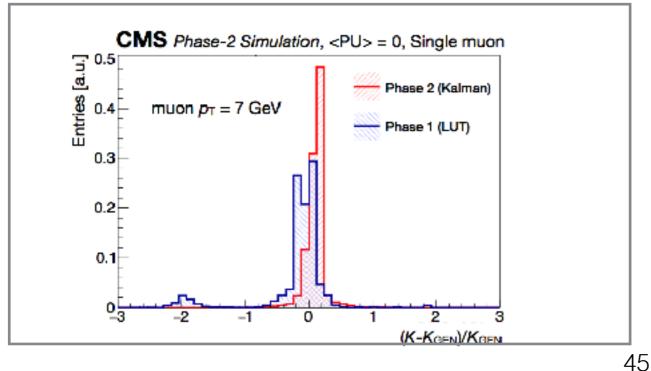


A number of algorithms are being implemented in Vivado HLS

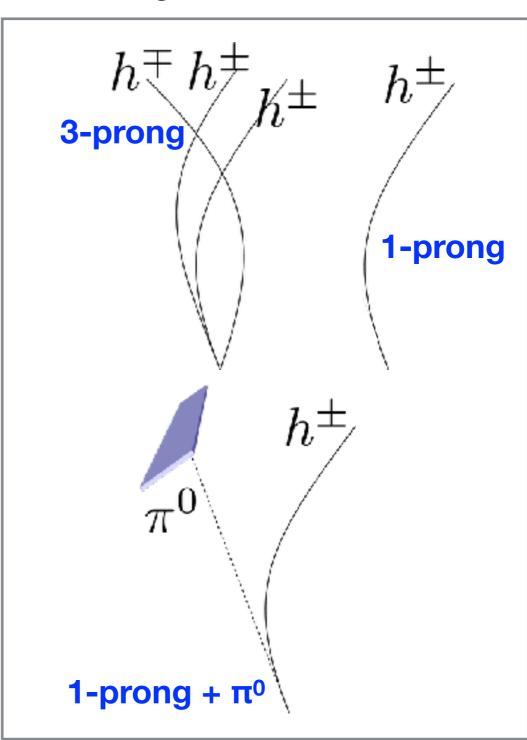
Cluster Producer



Kalman Filter Muon Reconstruction



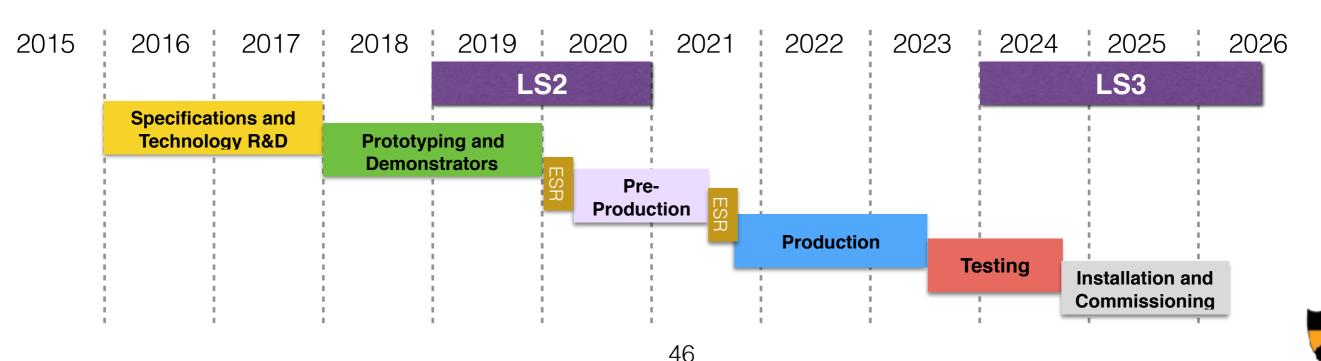
Tau Finding HPS@L1



A few Key Benchmarks

- Spring 2018 Completion of emulation of core algorithms and Benchmark performance baselines for representative sample of core trigger algorithms
- Fall 2018 Test Setups of 1st-generation demonstration processors; allows local comparisons with emulators
- Spring 2019 Benchmark performance baselines for all core trigger algorithms

Spring 2020 Phase-2 Technical Design Report

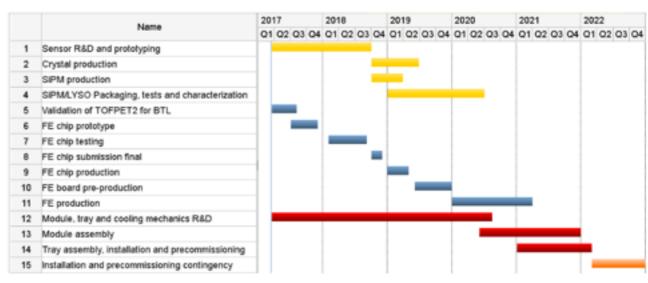


MTD: On the road to the HL-LHC



Installation

- Barrel Timing Layer is estimated to be completed by first quarter of 2022
 No interference with Tracker Installation
- Endcap Installation complete by 2023
- Cost-model driven by ECAL barrel for the scintillating crystals and readout electronics
- Detector support structure and cooling estimate driven by tracker cost model
- 7.5 MCHF for the barrel and 5.7 MCHF for the endcap



	Item	Cost per item	Full cost
		[CHF]	[CHF]
Barrel	Crystal	9	31 36 k
	SiPM	5	1419k
	FE Electronics	6/channel	1610k
	BE electronics		288k
	lpGBT	50/link	250k
	power supplies		568k
	power cables		20k
	Support materials	_	150k
	Total		7492k
Endcaps	Silicon Sensors	10	1555k
•	FE Electronics lpGBT		2194k
	Bump Bonding	$\langle \langle \rangle$	414k
	BE electronicss		589k
	power supplies		435k
	cables		100k
	Support structure	<u> </u>	280k
	Total	> / / /	5667k





Both Projects Presented today represent an ideological change for CMS

MIP Timing Detector

- Assigning timing to tracks, Primary Vertices and calorimeter deposits will give CMS a handle to differentiate between individual p-p interactions
 - Offers an important and necessary evolution of Reconstruction
 Techniques at CMS

Trigger System

- Adding tracking information to the Level 1 Trigger System will bring the Level 1 trigger System Closer to the Offline System!
- Timing Layer and Trigger studies, design and schedule are advancing well at CMS! Still much work to be done, look for updates!





Clearly visible in current upgrade studies

- W[±]H+E_T^{miss} Search sensitivity at high mass decreases when going from 140 to 200 PU
- From [CERN-LHCC-2015-19, LHCC-G-165] many analyses using Taus, Jets and E_T^{miss} are degraded as Pile Up increases, even with other detector upgrades

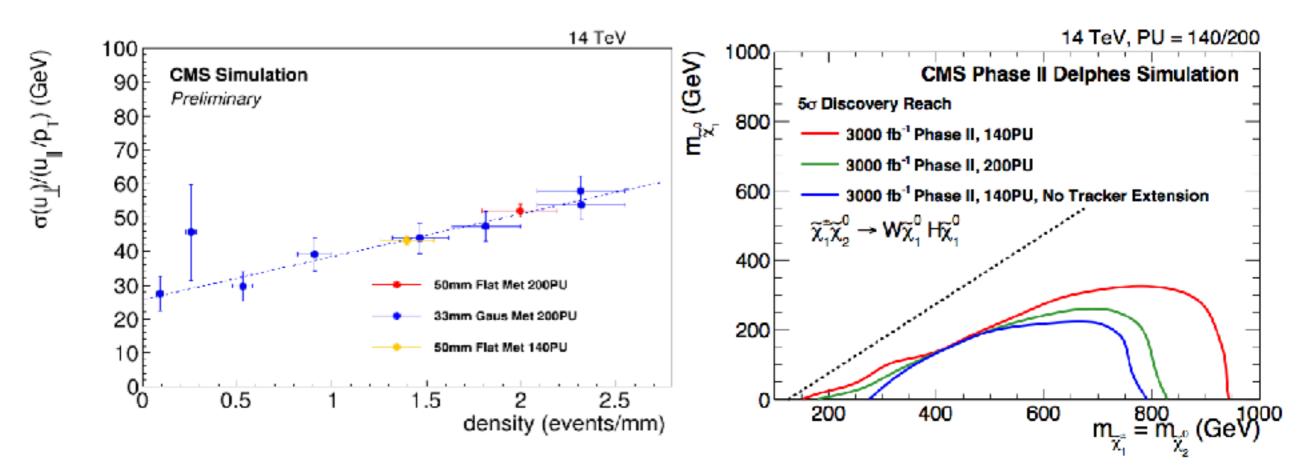




Table 2.2: Radiation levels in the BTL after 4000 fb⁻¹ of integrated luminosity for the CMS barrel region at different η .

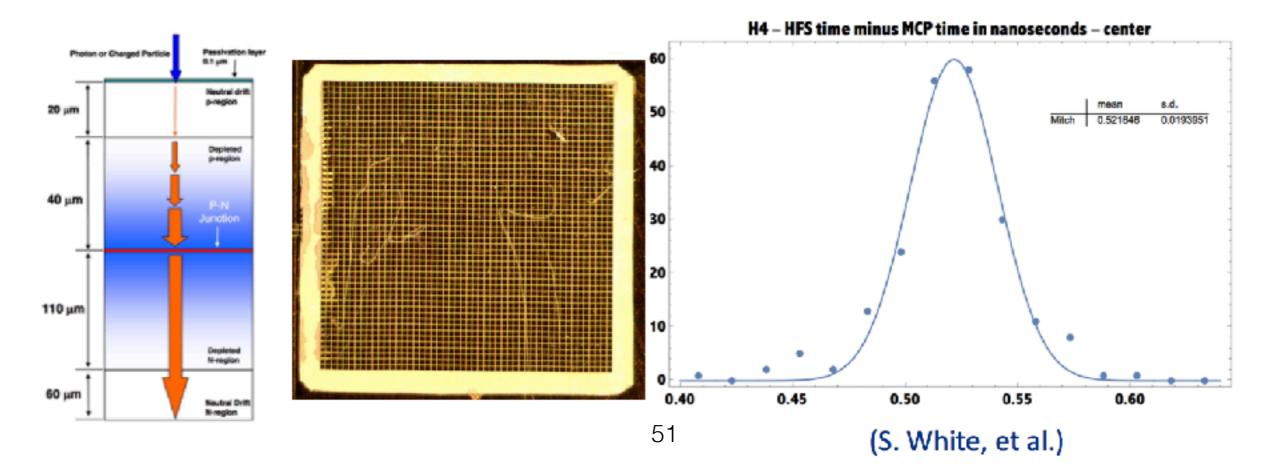
η	R	Z	1 MeV neq	Charged hadrons	Dose
	[cm]	[cm]	[cm ⁻²]	[cm ⁻²]	[kGy]
0	116.5	0	$1.7 imes 10^{14}$	$1.3 imes 10^{13}$	16
1.15	116.5	170	$1.9 imes 10^{14}$	$1.6 imes 10^{13}$	21
1.45	116.5	240	2.0×10^{14}	1.7×10^{13}	25

Crystals expected to be able to withstand 1000 kGy

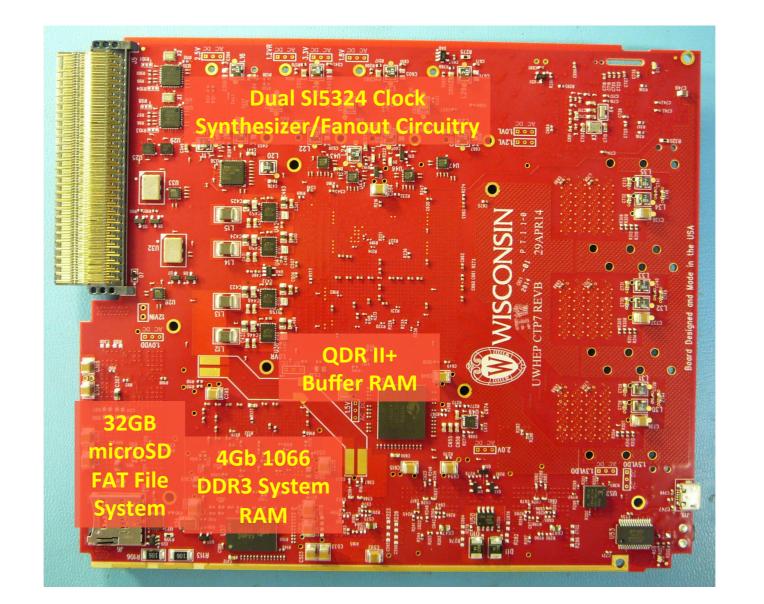


Timing Optimized Deep Depleted APDs

- Medium-gain (300x) APD read out with capacitatively coupled mesh
- Gain and drift regions overlap
 Mesh helps to stabilize E-field over device
- 20 ps resolution achieved for large sensors (8x8mm2), high capacitance
- 12×12mm² cross-section coupled to a 5×5mm² SiPM
- Crystal thickness varies from 3.75mm at $|\eta| < 0.7$ to 3.0 mm and to 2.4 mm at |eta|>1.1
- Radiation hardness and segmentation requirements not yet met!
 - Studies on-going... but not to be considered a viable option until these important milestones are met



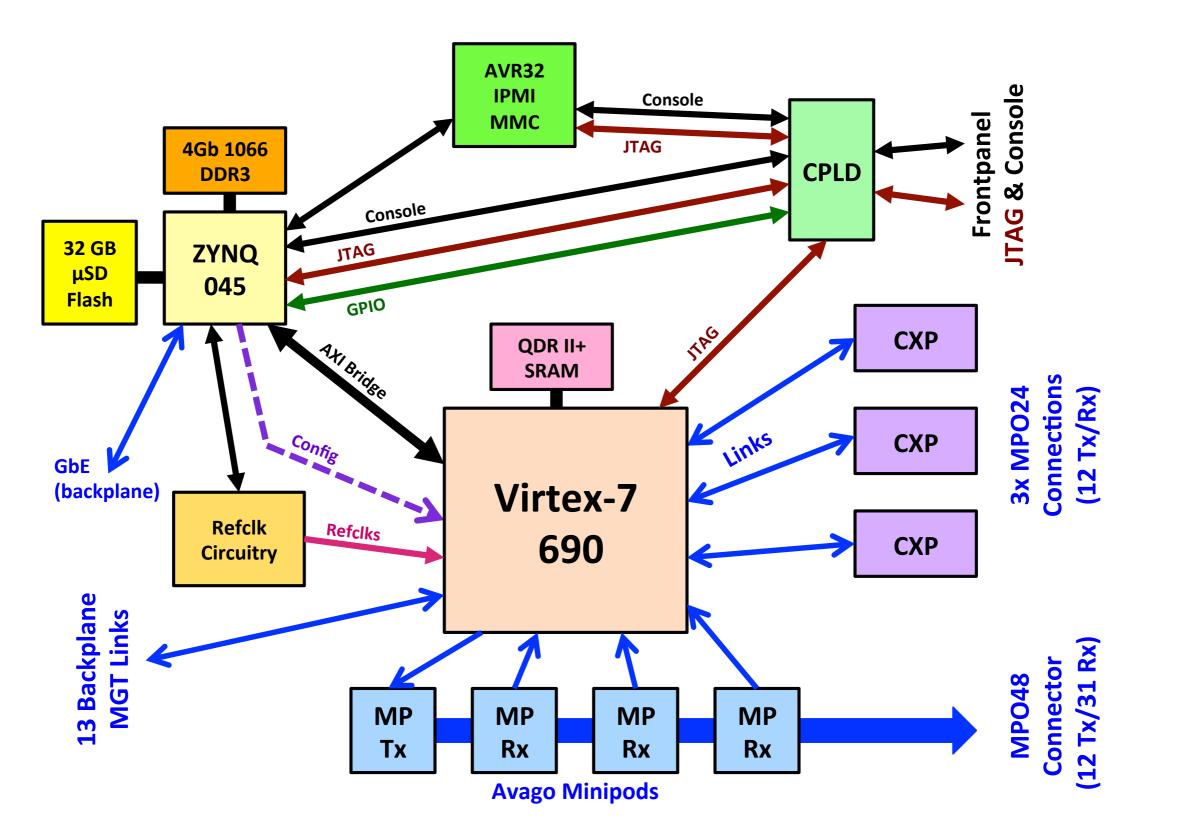






CTP7

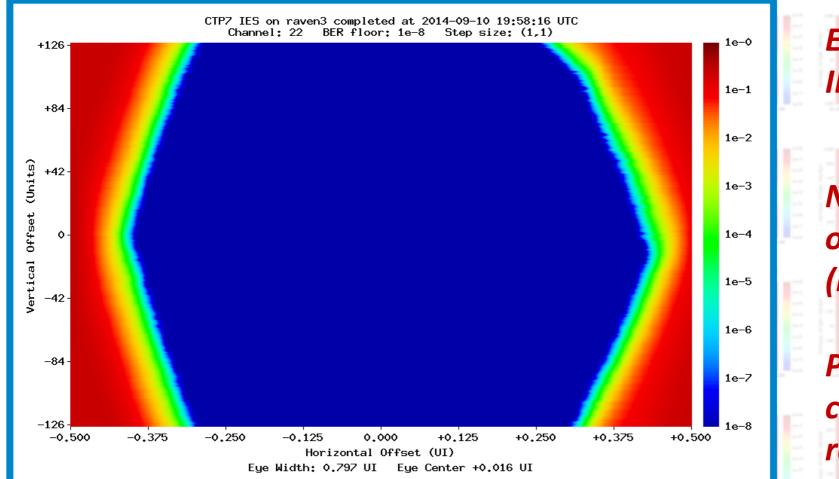






Track Trigger





Eye Scan Capability of IBERT ported to ZYNQ

Non-invasive scans taken on live operational data (not PRBS test patterns)

Parallel multi-channel capability for faster results

Programmable engine can be used for a wide range of applications:

- Quantitative link characterization and parameter optimization
- Operational troubleshooting

Trend analysis and forecasting for preventive maintenance
 Scalable architecture with scanning and image rendering on same ZYNQ platform for fewer bottlenecks in applications with high link counts

54

Current Boards Phase 1





31 Rx and 12 Tx Frontpanel Optical 10G links on MiniPODs



13 GTH Back-plane Tx/Rx links

36 Tx and Rx Frontpanel Optical 10G links on CXP Modules

University of Wisconsin - Madison



Links

- Replaces physical cable with TCP/IP connection to Xilinx toolset
- Useful where JTAG cable not easily employed
- Primitive protocol, well established technology found in ISE tools
- Does not require JTAG infrastructure in uTCA crate
- CTP7 Implementation:
 - ZYNQ connects to V7 JTAG pins through programmable logic
 - Cable Server runs on ZYNQ embedded CTP7 Linux system
 - Allows remote debug of Virtex-7 firmware via LAN connection to ZYNQ



Support PC



Xilinx JTAG Cable







On the road to the HL-LHC



2017 Nov 17, N **Trigger & Timing for Phase** I.Ojalvo



Trigger

Spring 2017 Completion of emula



