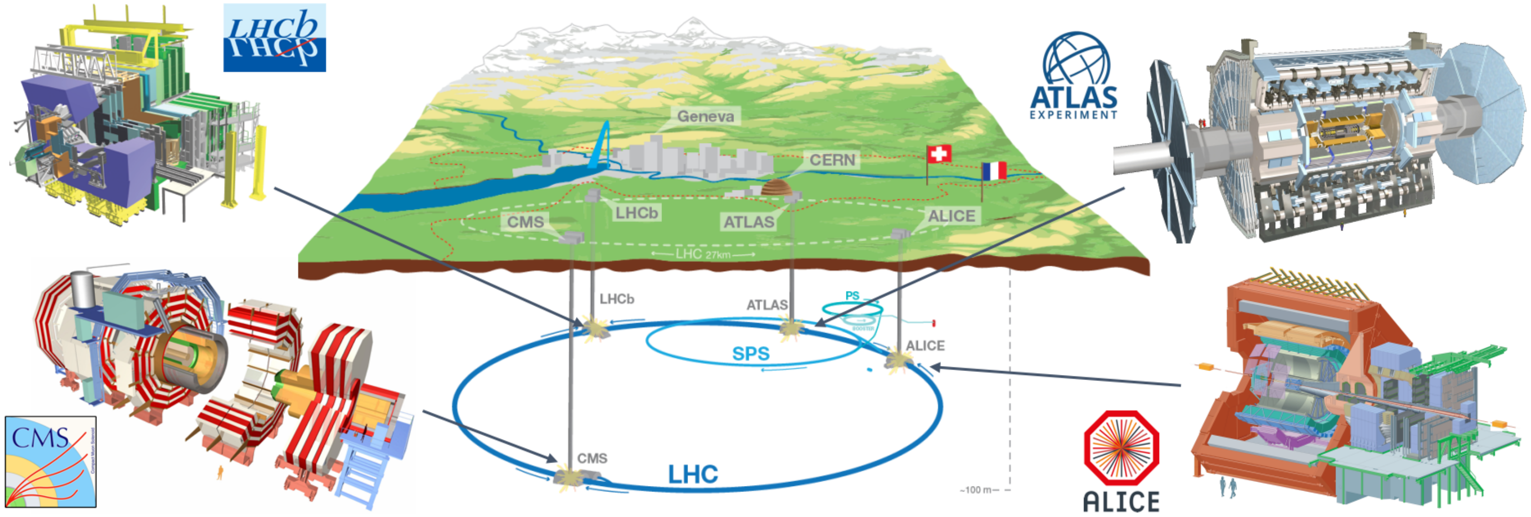
ATLAS TRIGGER AND DAQ CURRENT AND FUTURE

Streaming Readout Workshop (SRO XII)
3 December 2024, at Tokyo University

Kunihiro Nagano (KEK/IPNS, Japan)

素粒子原子核研究所
Institute of Particle and Nuclear Studies

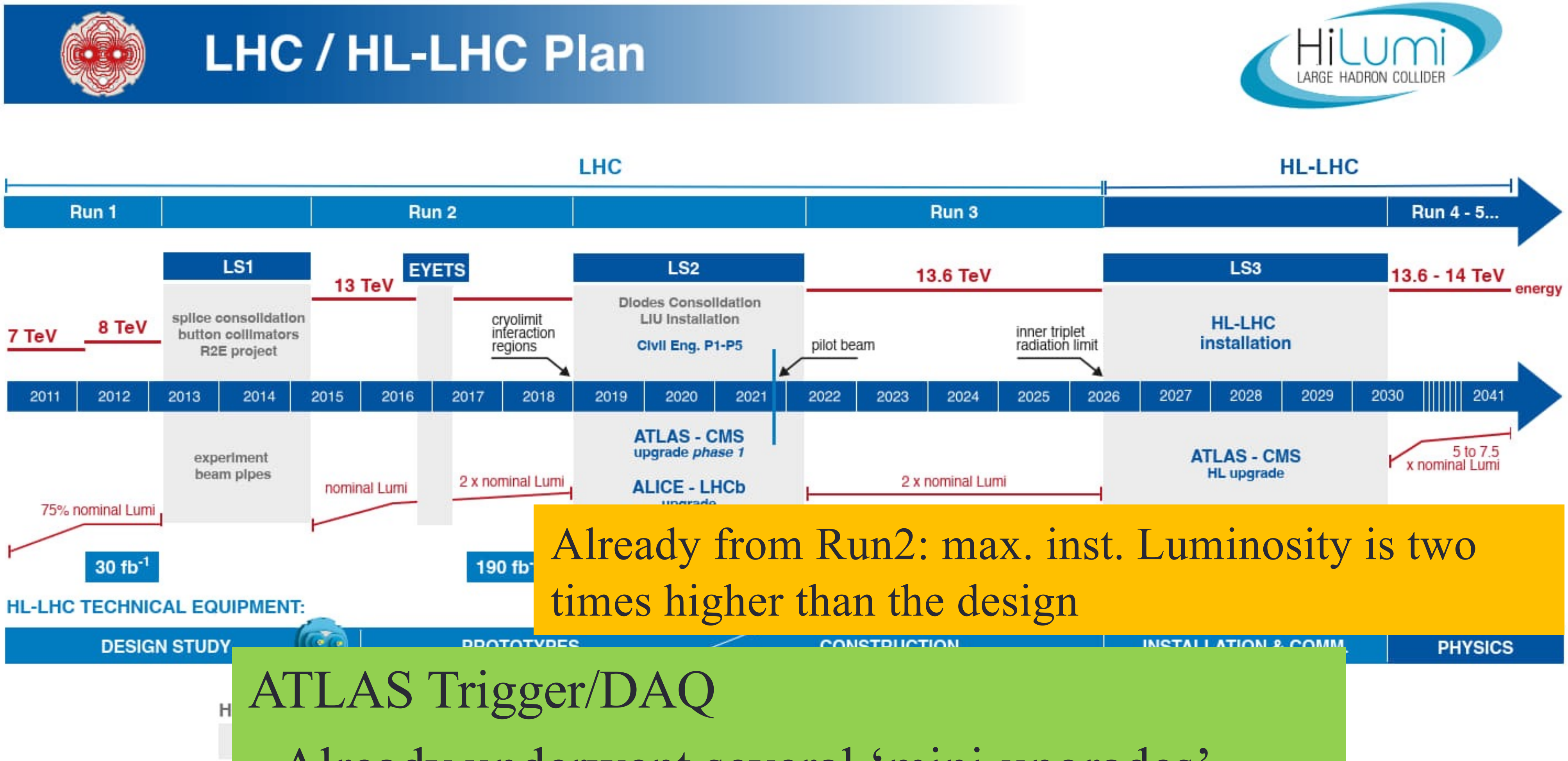
LHC



Google Logo
at First beam

- LHC starts from 2010
- 4 experiments(ATLAS, CMS, LHCb, ALICE)
- 2010 ～ 2012 : 【Run1】 E=7, 8 TeV
- 2015 ～ 2018 : 【Run2】 E=13 TeV
- 2022 ～ 2025 : 【Run3】 E=13.6 TeV
- 2029 ～ : 【HL-LHC】

LHC design：E=14 TeV, Lmax=1 x 10^{34} cm^2/s

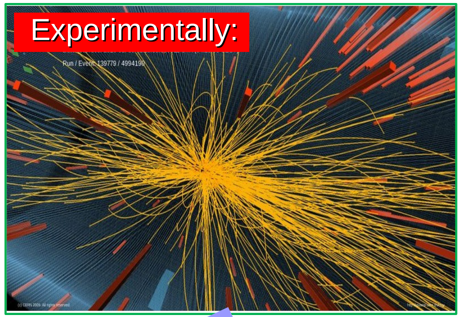


Already from Run2: max. inst. Luminosity is two times higher than the design

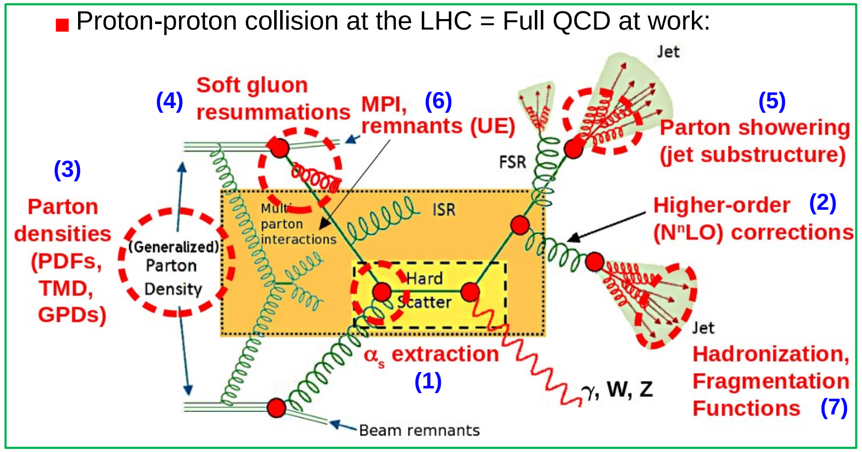
ATLAS Trigger/DAQ

- Already underwent several ‘mini-upgrades’
  - Good exercise / intro for HL-LHC Upgrade
- In full swing toward HL-LHC Upgrade

Physics at LHC



First time exploiting physics at TeV energy scale
- The origin of EW symmetry breaking: Higgs mechanism
- Direct searches for Physics beyond SM (BSM)

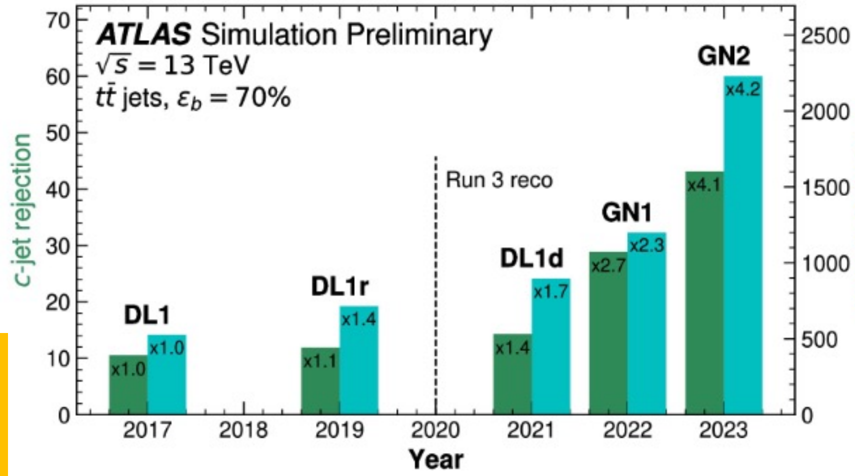


Only possible with advanced technologies not only for experiments/detectors but also physics

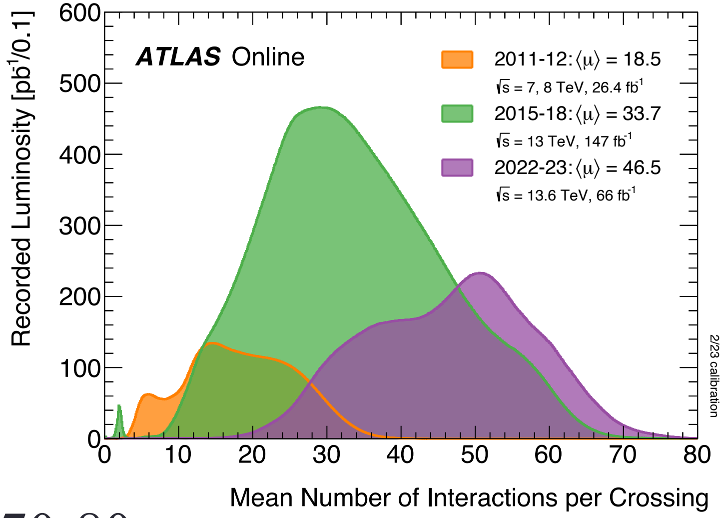
- Experimental
  - Radiation tolerance
  - Pileup, high luminosity (vs. trigger, r/o)
- Physics analysis
  - Data driven technique, ML, statistics …
- Physics/theory
  - QCD (higher order, jet algo, PDF etc.) ,,,

Remarkable advances particularly in hadronic objects reconstruction e.g. with ML/AI
-- b-jet tagging
-- Constituent (in large R jet) based W, top taggers
-- Jet calibrations
…

Extensively used as well in the trigger



Pileup

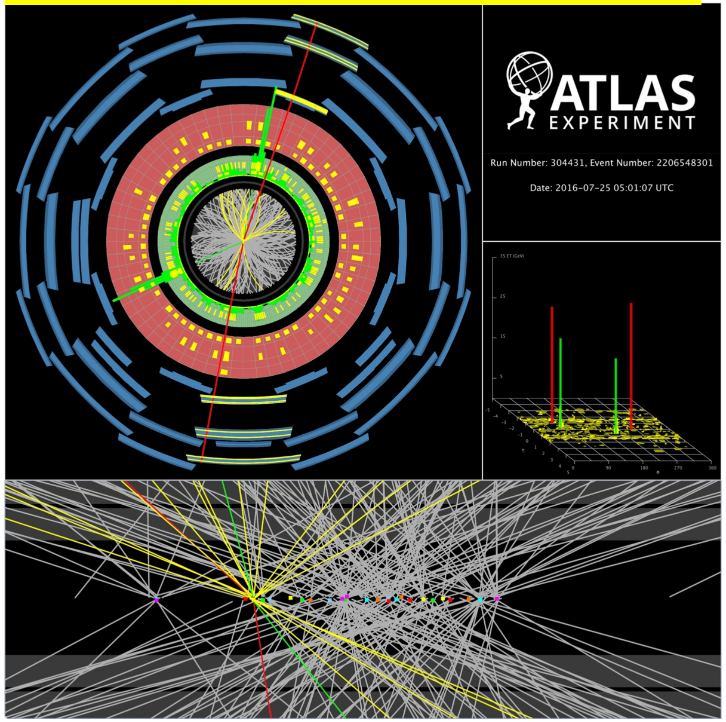


- Multiple pp interactions per 1 bunch crossing = called "pileup"
  - pp total cross section x 10^{34} cm^2/s >> nr. of bunch crossings / s
  - Run3 average nr. of pileups （<μ>） 47、up to 70-80

LHC design ~25

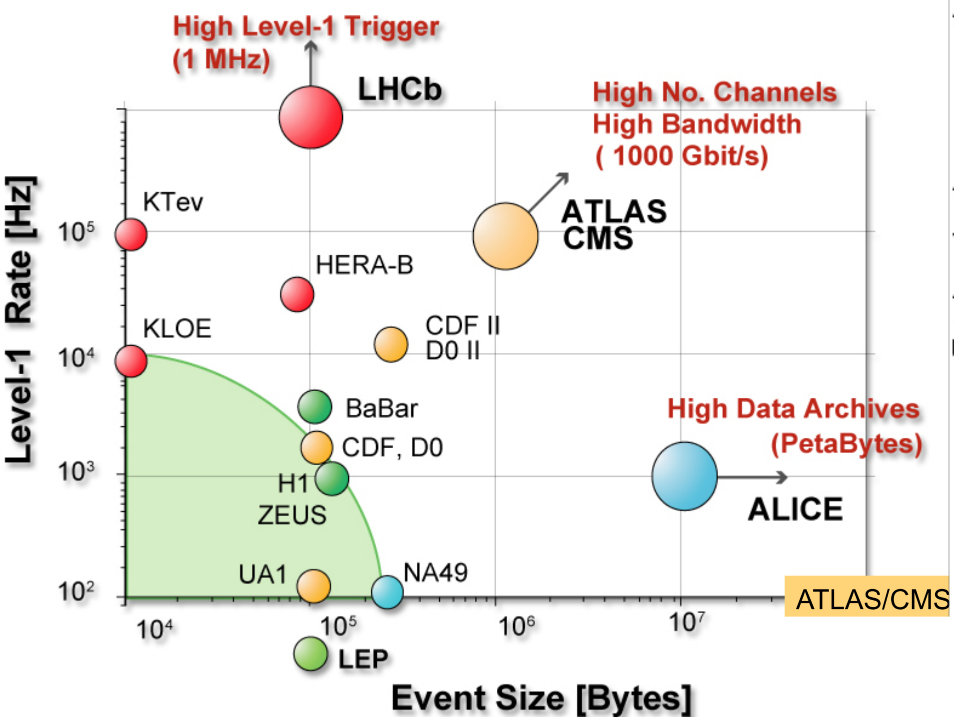
HL-LHC: ~200

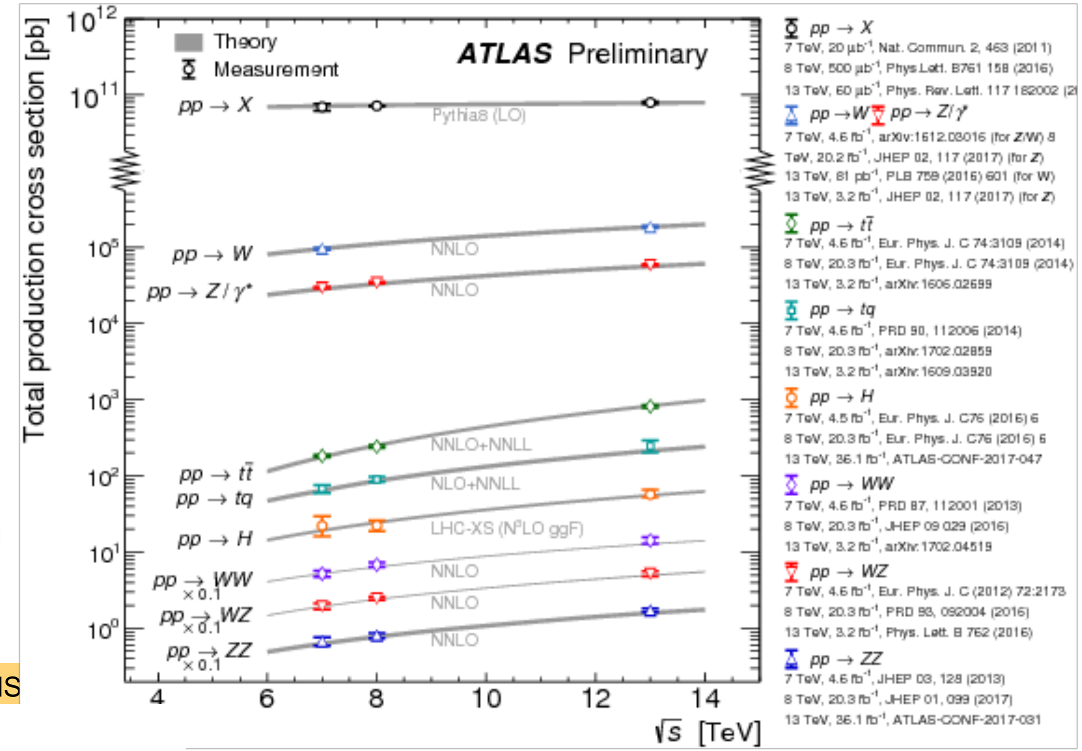
h→2e2μ candidate（25 vertices）



- A big challenge for experiments
  - Deterioration of detection/identification of "objects" (e.g. electron/photon, τ, b-jets, $E_T$miss)
  - Increase of hit rates
    → Heavy load for readout system
  - Increase of computing time, particularly for tracking
    → Challenge for Trigger/Computing

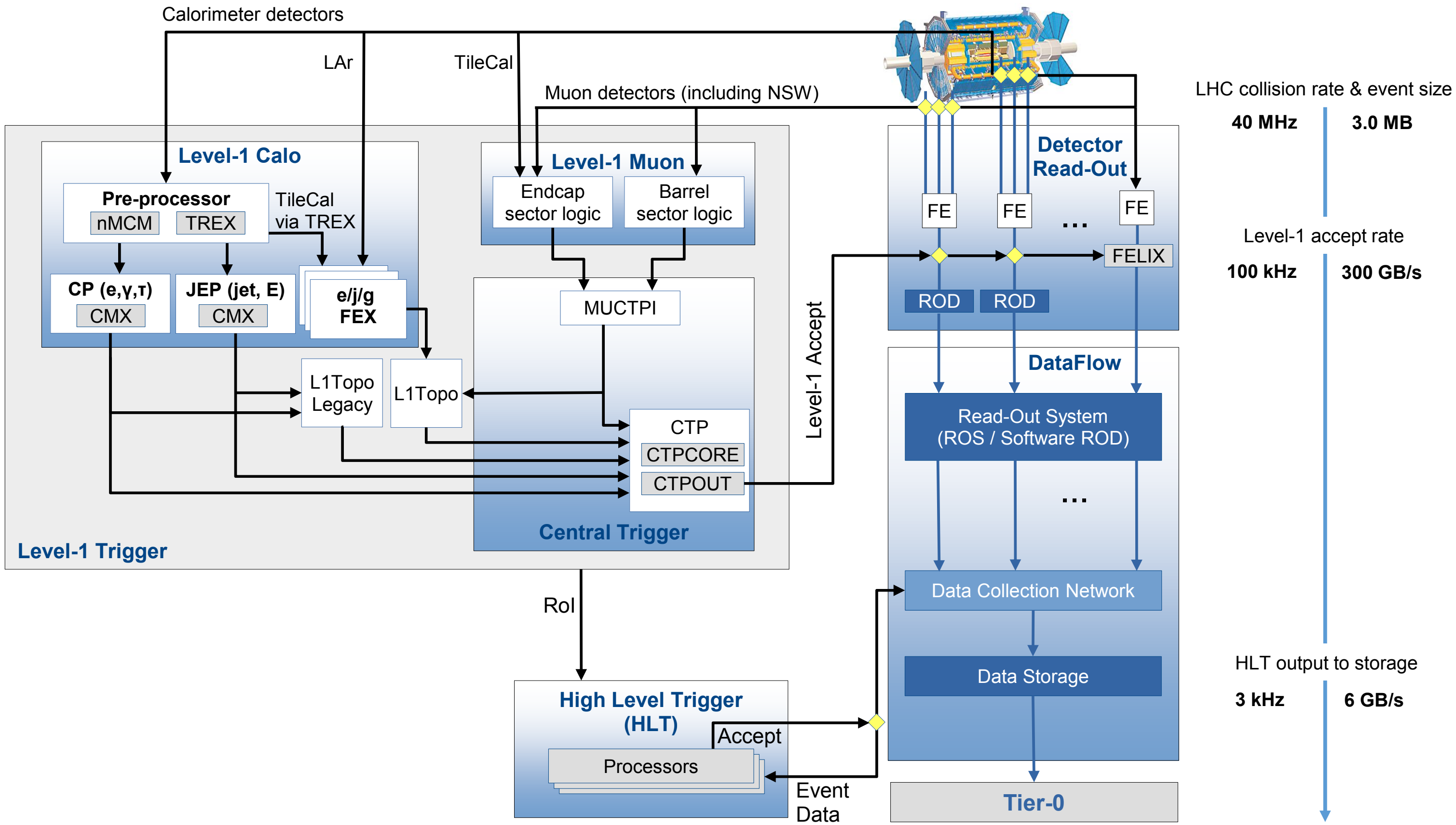
Trigger @ LHC





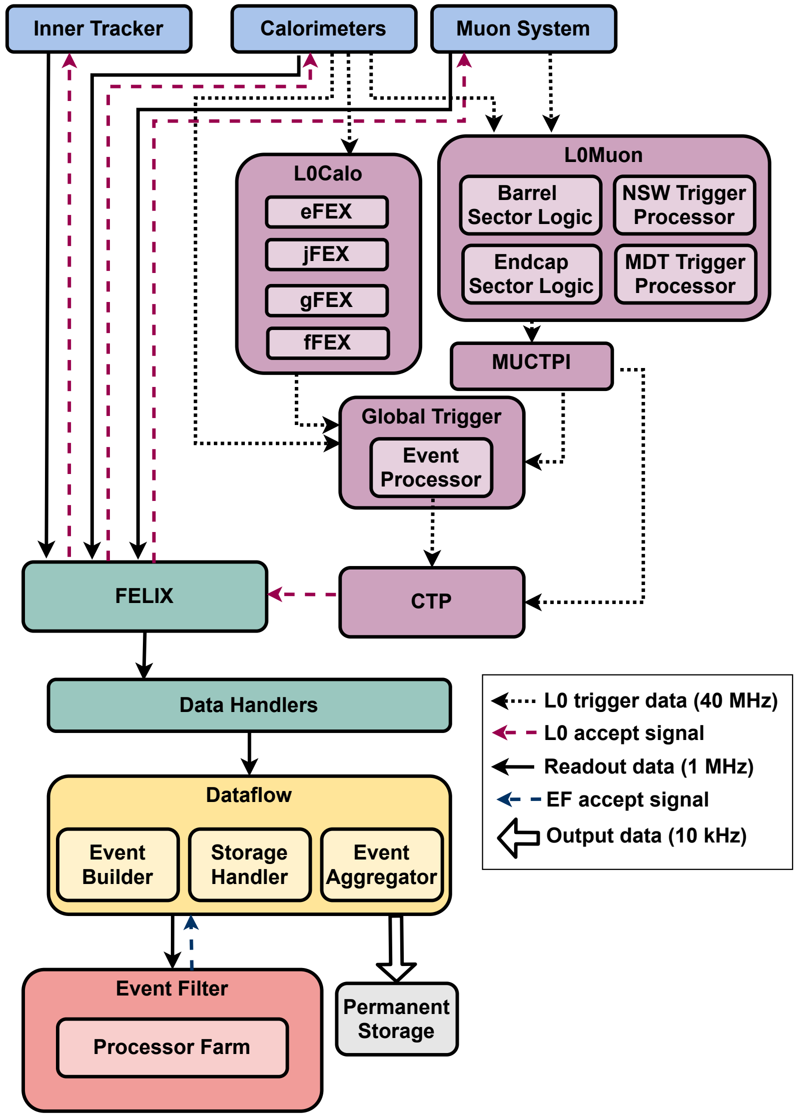
- 40 MHz bunch crossing (25ns time spacing) with ~2 MB event size
  - Large event rate reduction : 40 MHz → 100KHz for readout, further down to O(1 kHz) for recording
  - Deadtime free → pipelined online processing (at L1 – before readout)
- ATLAS Trigger: 2-level trigger system, L1 and HLT (High Level Trigger)
  - L1: fully time synchronous, with a fixed latency (2.5μs). Custom hardware.
  - HLT: processing at computing farm

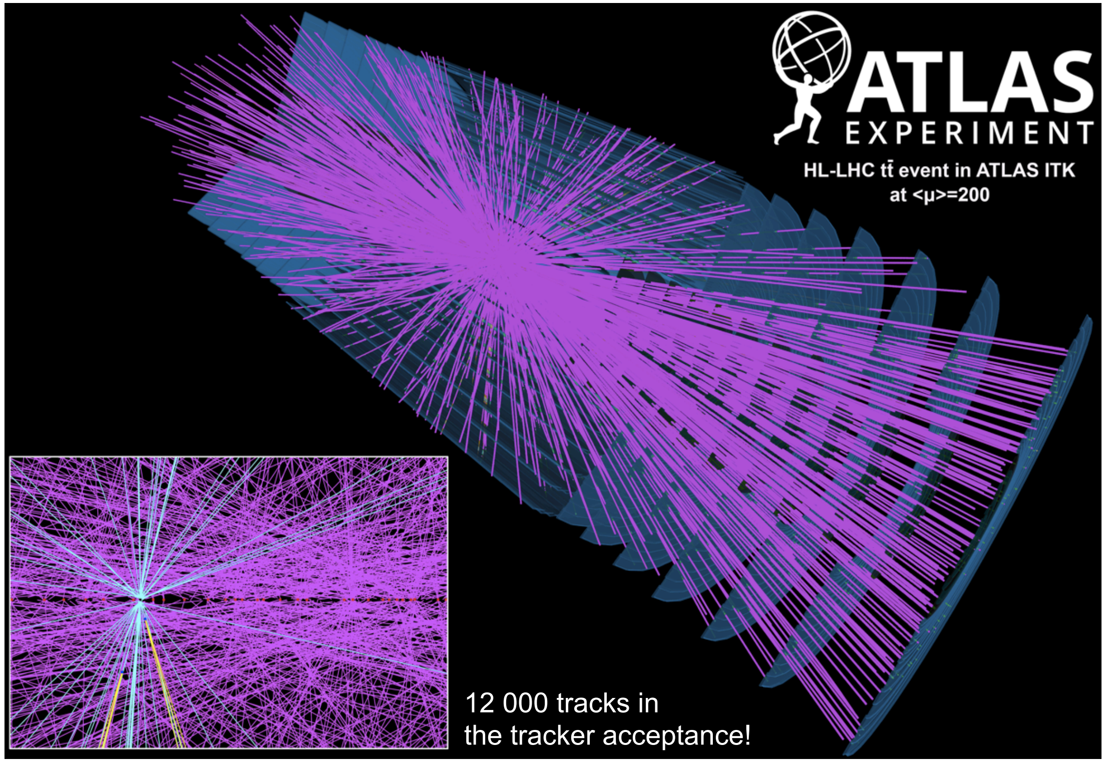
ATLAS Trigger DAQ [Run3]



- Already at L1, there is L1Topo that calculates correlation between objects (e.g. mass between 2-muons)
- Inner tracker information is available only from HLT
- HLT: “Region of Interest (RoI)” concept →bandwidth saving
  - e/mu/τ/b-jet/jet : HLT readouts only the areas identified by L1

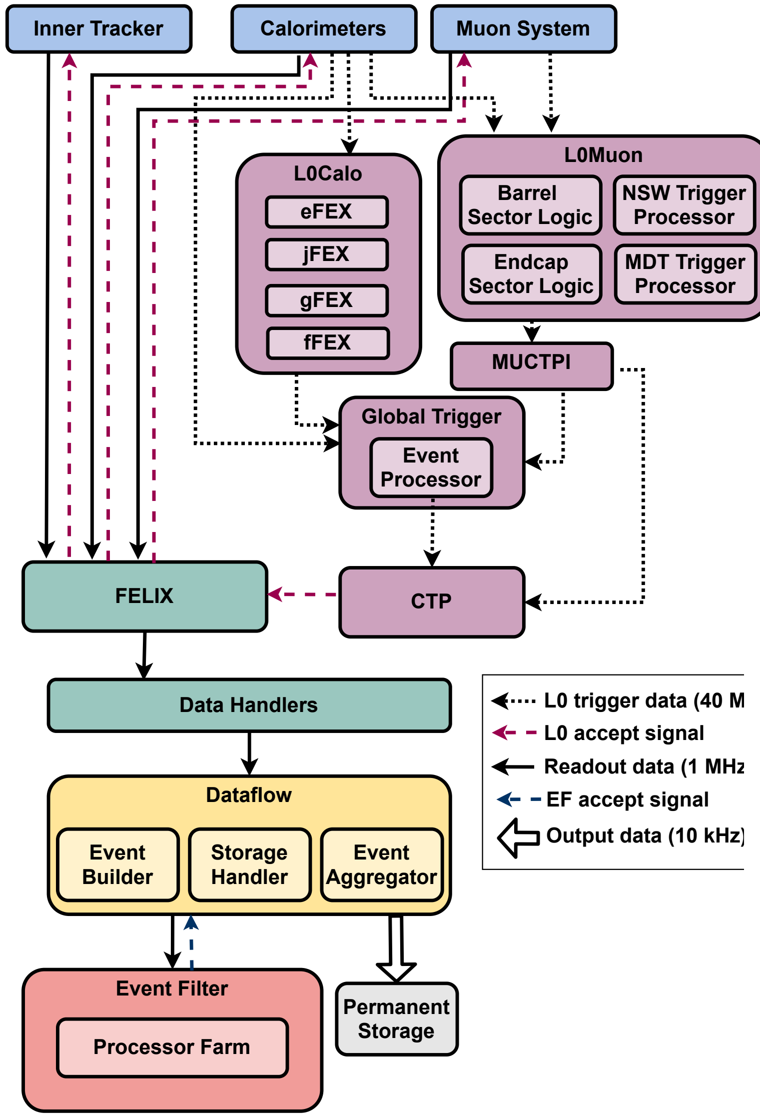
ATLAS Trigger DAQ [HL-LHC]





- Latency gets 4 times longer → all FEs will be replaced
- A major upgrade for Trigger and DAQ
  - L0: 100 kHz → 1 MHz
    - Latency : 2.5 → 10 μs
  - EF: ~3 kHz → 10 kHz

ATLAS Trigger DAQ [HL-LHC]



Tracking – a key to control pileup

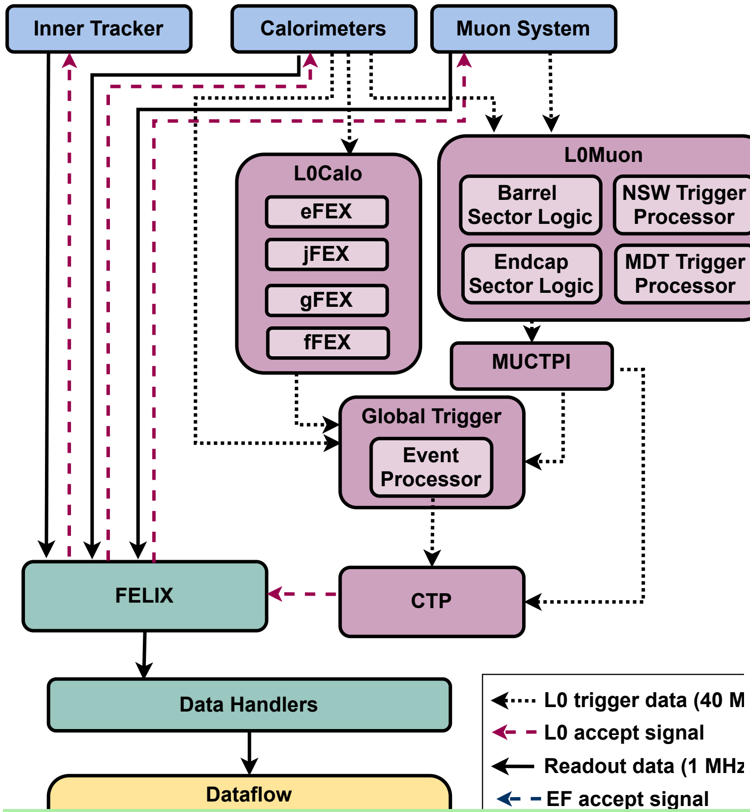
HL-LHC trigger menu（TDR)

Single lepton, di-lepton
- Threshold lower than Run3: 28→22 GeV

Hadronic（$E_T$miss, multi-jets）
- Threshold similar to Run3

Trigger Selection	Run 1 Offline p_T Threshold [GeV]	Run 2 (2017) Offline p_T Threshold [GeV]	Planned HL-LHC Offline p_T Threshold [GeV]	L0 Rate [kHz]	After regional tracking cuts [kHz]	Event Filter Rate [kHz]
isolated single e	25	27	22	200	40	1.5
isolated single μ	25	27	20	45	45	1.5
single γ	120	145	120	5	5	0.3
forward e			35	40	8	0.2
di-γ	25	25	25,25		20	0.2
di-e	15	18	10,10	60	10	0.2
di-μ	15	15	10,10	10	2	0.2
$e-\mu$	17,6	8,25 / 18,15	10,10	45	10	0.2
single τ	100	170	150	3	3	0.35
di-τ	40,30	40,30	40,30	200	40	0.5†††
single b-jet	200	235	180	25	25	0.35†††
single jet	370	460	400			0.25
large-R jet	470	500	300	40	40	0.5
four-jet (w/ b-tags)		45†(1-tag)	65(2-tags)	100	20	0.1
four-jet	85	125	100			0.2
H_T	700	700	375	50	10	0.2†††
E_T^{miss}	150	200	210	60	5	0.4
VBF inclusive			2x75 w/ ($\Delta\eta > 2.5$ & $\Delta\phi < 2.5$)	33	5	0.5†††
B-physics††				50	10	0.5
Supporting Trigs				100	40	2
Total				1066	338	10.4

ATLAS Trigger DAQ [HL-LHC]



HL-LHC trigger menu（TDR)

Single lepton, di-lepton
- Threshold lower than Run3: 28→22 GeV

Hadronic（$E_T$miss, multi-jets）
- Threshold similar to Run3

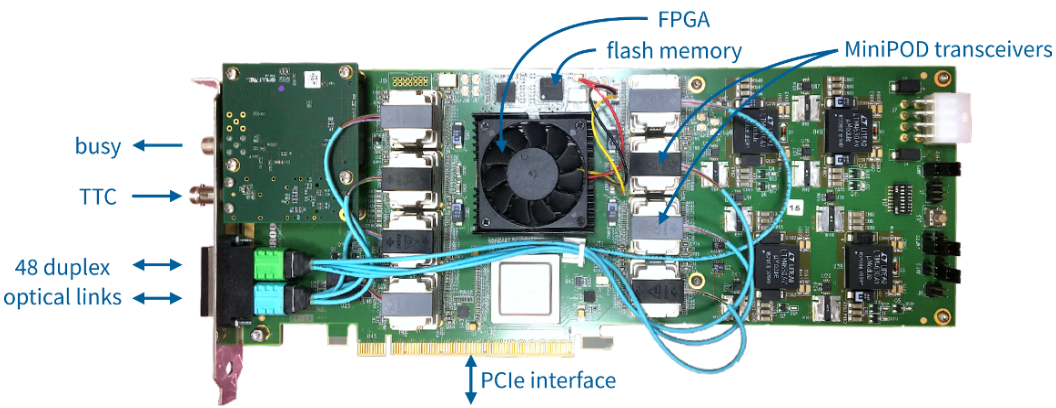
Trigger Selection	Run 1 Offline p_T Threshold [GeV]	Run 2 (2017) Offline p_T Threshold [GeV]	Planned HL-LHC Offline p_T Threshold [GeV]	L0 Rate [kHz]	After regional tracking cuts [kHz]	Event Filter Rate [kHz]
isolated single e	25	27	22	200	40	1.5
isolated single μ	25	27	20	45	45	1.5
single γ	120	145	120	5	5	0.3
forward e			35	40	8	0.2
di-γ	25	25	25,25		20	0.2
di-e	15	18	10,10	60	10	0.2
di-μ	15	15	10,10	10	2	0.2
$e-\mu$	17,6	8,25 / 18,15	10,10	45	10	0.2
single τ	100	170	150	3	3	0.35
di-τ	40,30	40,30	40,30	200	40	0.5†††
B-physics				50	10	0.5
Supporting Trigs				100	40	2
Total				1066	338	10.4

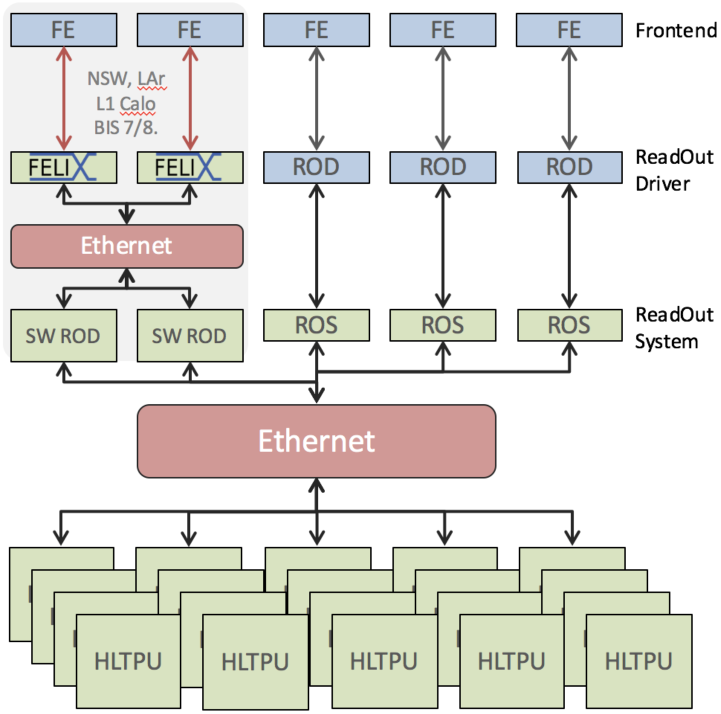
- Toward HL-LHC
  - Trigger – tracking (commodity acceleration, ML tracking)
  - DAQ – readout (FELIX)

Tracking – a key to control pileup

Readout

- FELIX to harmonize detector readout system
  - To reduce the number of custom electronics components and design effort by leveraging commercial products (network i/f cards, servers, and network switches), resulting in granting greater flexibility in maintenance, upgrades, and customization
  - Commodity servers equipped with PCIe FELIX cards
    - Data received over point-to-point optical links and routed to peers via ethernet
    - The primary peer on this network is SWROD
  - FELIX card (FLX-712)
    - 16-lane PCIe Gen3
    - Xilinx Ultrascale FPGA (XCKU115-FLV-1924)
    - 8 Avago Minipod transceivers (TX and RX)
    - Optical links protocols: GBT (rad-hard standard developed at CERN) 4.8 Gb/s, and FULL mode (9.6 Gb/s)
    - Being used a since Run3 for certain subsystems



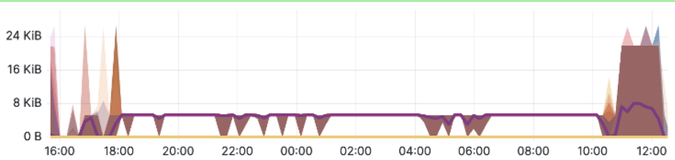


Readout

LAr digital processing blade (FULL mode)

(a) Max. Throughput ~ 40 Gb/s

(b) Rate ~ 100 kHz



(c) Max. Message size 22 kB

Figure 4: Examples of FELIX readout of LDPB: Throughput, Rate and Message Size.

NSW readout efficiency (GBT)

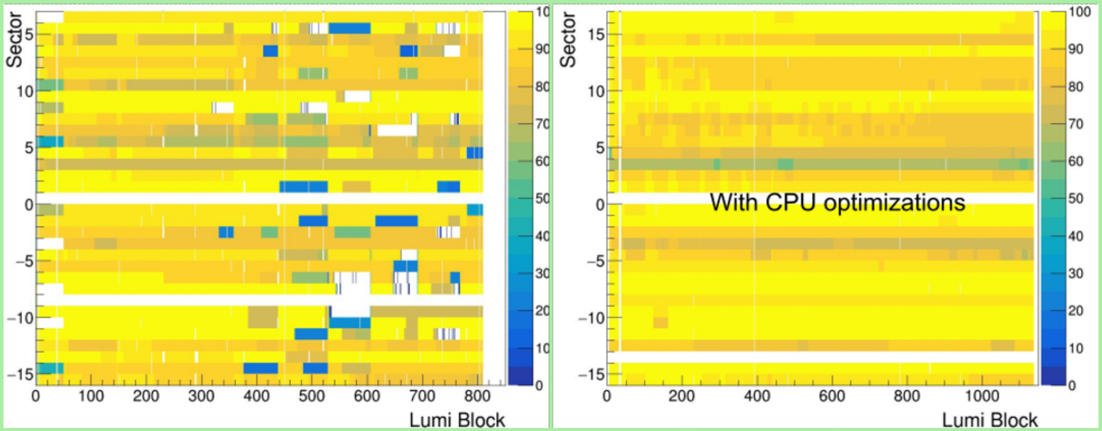


Figure 6: NSW readout efficiency before and after CPU optimizations to correct for late packet arrival.

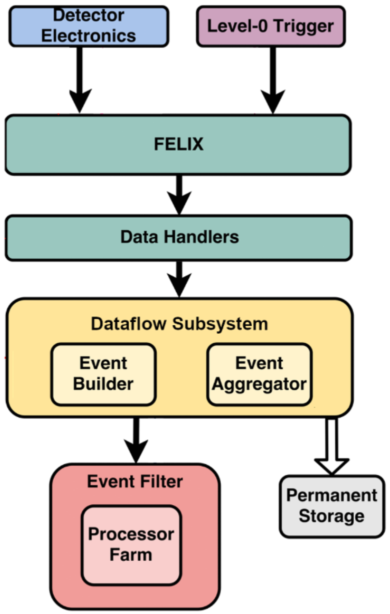
→ Significant improvement during commissioning and early data-taking to allow 100 kHz readout with data losses due to FELIX negligible

Readout

- FELIX for HL-LHC
  - Maximum link speed : up to 25 Gb/s (current: 10 Gb/s)
  - Supports: GBT, lpGBT, 64b/67b-encoded Interlaken
  - Two prototypes FLX-181 and FLX-182
    - AMD Versal Prime FPGA
    - 24 FireFly optical links
    - 16-lane PCIe Gen4
- All ATLAS readout will use FELIX + Data Handler
  - ~14000 optical links with 25 Gb/s per link
  - Data Handler – evolve from SWROD

(a) FELIX card prototype: FLX-181

(b) FELIX card prototype: FLX-182



FELIX in the wild

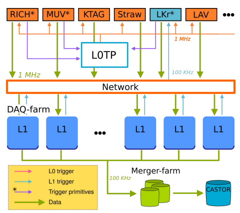
More information in backup slides

ProtoDUNE SP

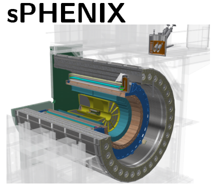
- 2 MHz readout via FELIX
- 15360 channels, 55 GB/s throughput
- Data taking started in 2018

NA62

- Kaon physics experiment
- At CERN SPS
- Readout with FELIX



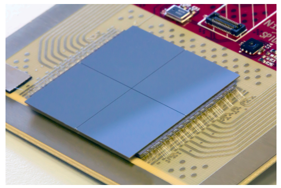
sPHENIX



- Located at RHIC, BNL
- 3 subdetectors readout with FELIX
- Streaming and triggered readout

SPIDR4

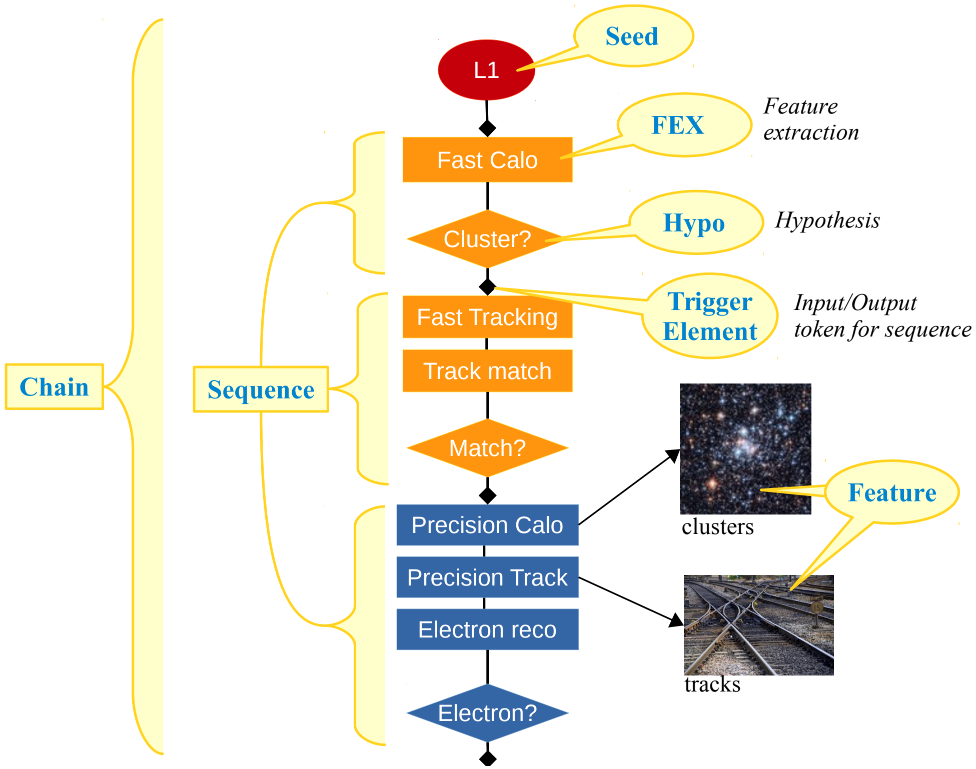
- Readout of Timepix4 pixel sensor
- Modified version of FELIX firmware and software



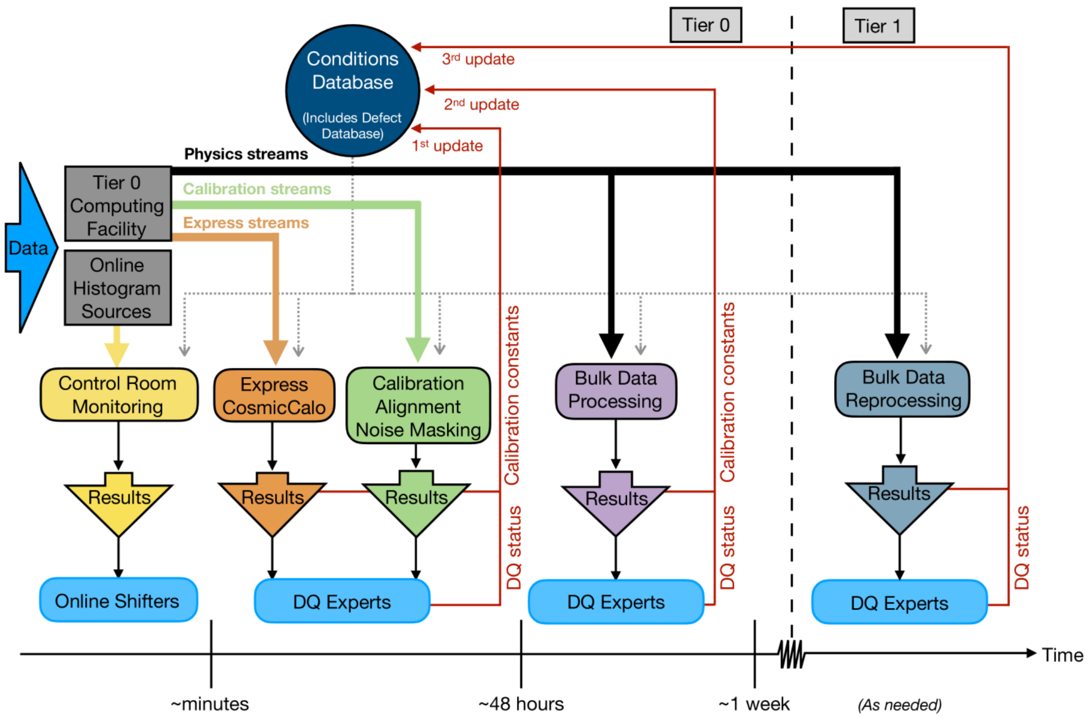
Online processing

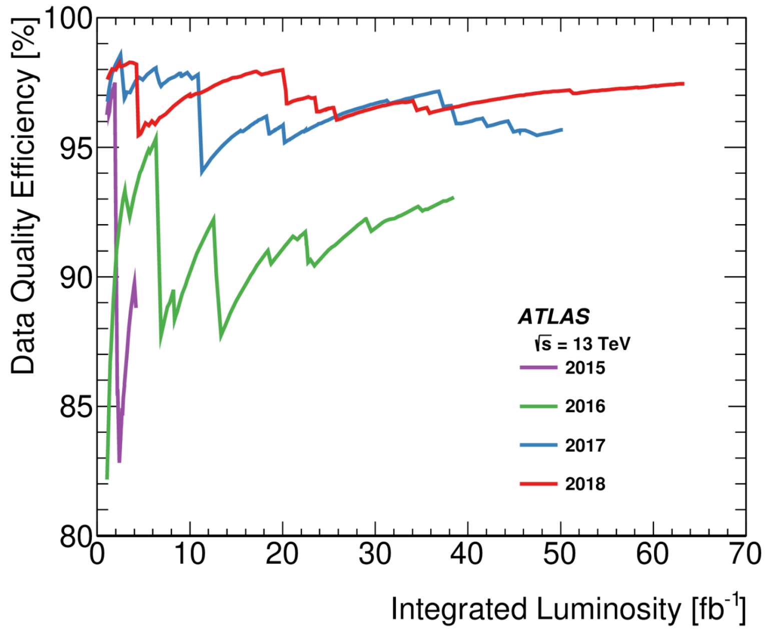
- Reconstruction of physics objects : e/γ, τ, b-jet, jet (large-R), μ, E_{Tmiss}
- Trigger logics and prescales ("Menu")
  - A "Chain" – one logic (e.g. electron E_T>28 GeV)
  - Menu – collection of O(1000) chains with prescales
- Trigger-specific functionalities (different to offline)
  - HLT rejects events ("early rejection")
  - HLT does regional reconstruction ("RoI")
  - Need to record execution history ("navigation")
- Events are recorded in streams
  - Physics: Prompt, Delayed
  - Calibration, Express (DQ), Debug
  - Trigger Level Analysis (TLA)

Only for illustration
(was - up to Run2)



Calibration/Monitoring

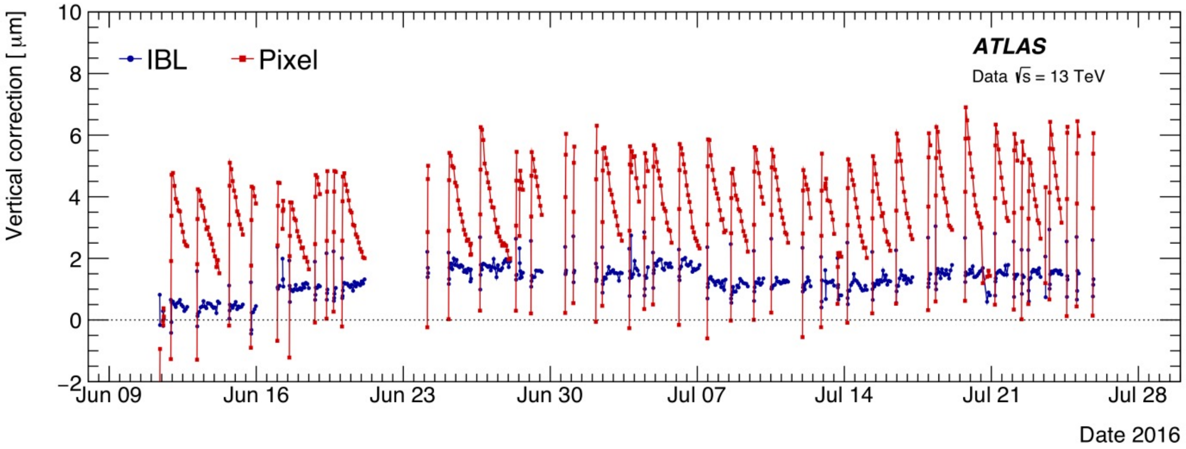


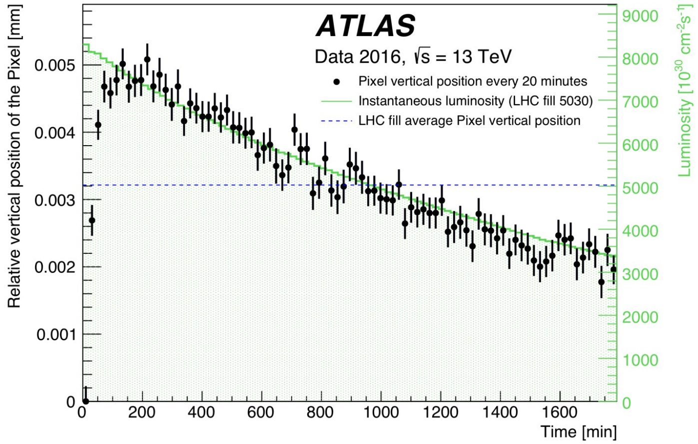


DQ flag is set for every Luminosity-Block (LB; ~1 min.) -> Good Run List (GRL) for physics

- Calibration (alignment, noise etc.) data are recorded with dedicated triggers (zero-bias, random, noise-burst triggers etc.) into dedicated streams
- First-pass calibrations ready within 48 hours → bulk reconstruction of physics data
- Online (trigger) uses the best-known (usually, latest) calibration constants, some exceptions are online-measured luminosity, beamspot (average x,y position of beam), dynamic pedestal subtraction of L1 calorimeter, etc.
- The data quality (DQ) is monitored at data-taking (online) and with express stream (before 1st pass reco)

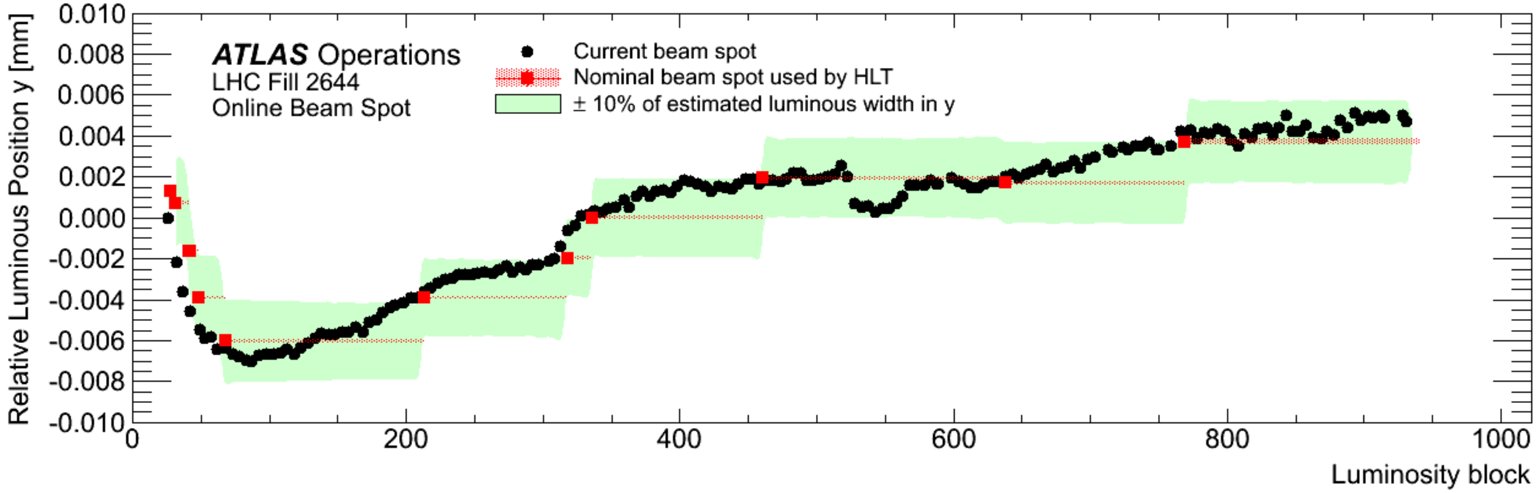
c.f. for perfect alignment of trackers





- Not only run-by-run, time-depending (during run) alignment is corrected at offline (for physics analysis) which is not done for trigger. (Trigger impact is negligible)

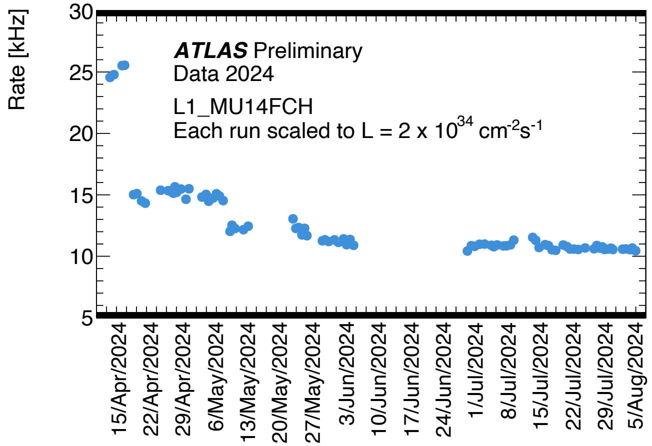
Online beamspot measured/used in trigger

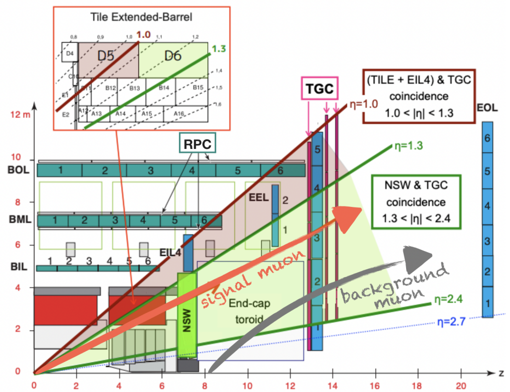


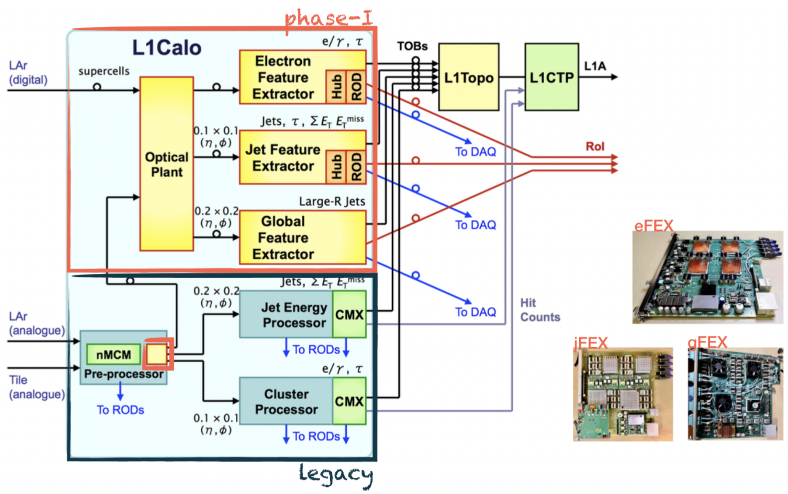
- HLT beamspot is updated if a shift (~5%) is observed
  - To balance between accuracy and stability

L1 Upgrade [Run 3]

- L1 Muon
  - Coincidence patterns to utilize newly installed detectors (NSWs), resulting in better discrimination against fake muon backgrounds
  - ~12 kHz L1 rate reduction
  - Activated in all sectors since this year
    - 'Monitored mode' in the commissioning period, activated sectors by sectors once validated with real data
- L1 Calo
  - New digital readout (SuperCell) providing higher granularity and resolution
    - Better resolution power particularly for electrons
  - All new trigger electronics (Feature Extractors)
    - eFEX: isolated e/photon, tau
    - gFEX: jets, tau, missing Et
    - jFEX: large radius jets, missing Et
  - Legacy system has been running for 2 years data taking in parallel to commissioning of the new system
    - Gradually migrated to the new triggers once fully validated with real data
    - Last 'legacy' triggers disabled recently in this year

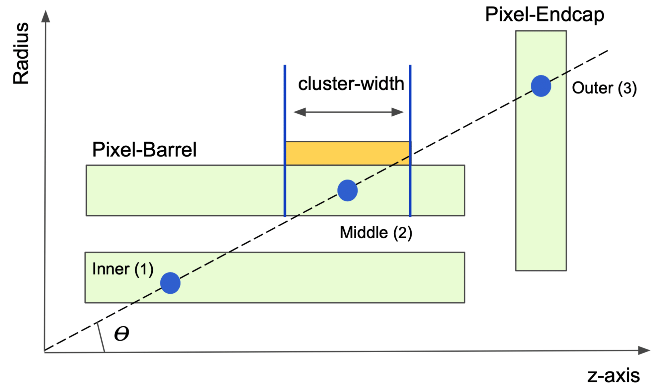


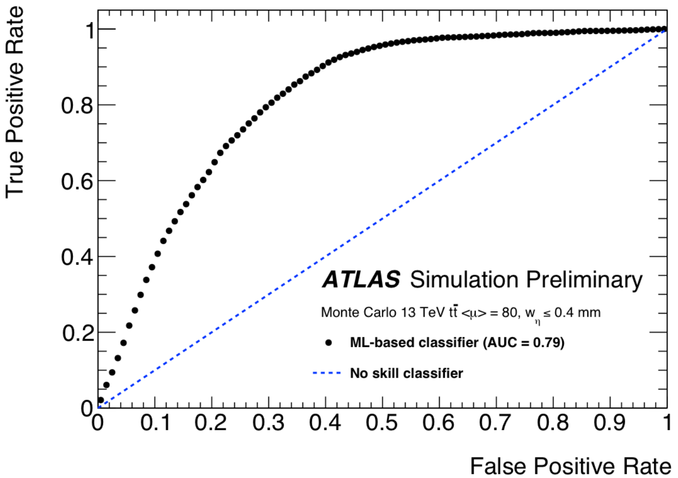




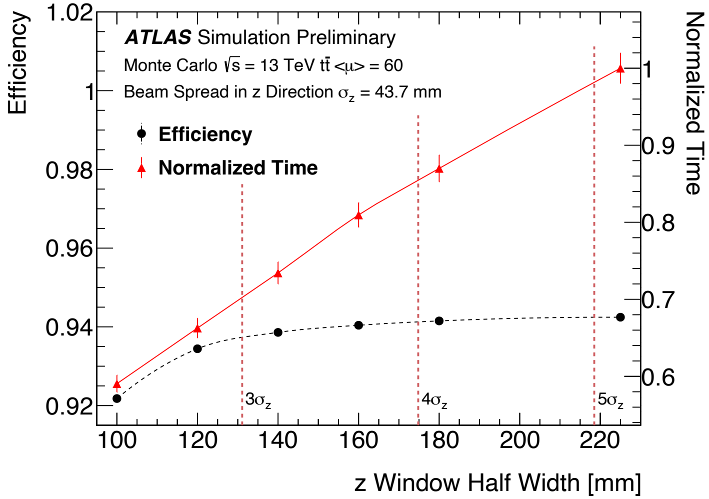
Full-scan tracking at HLT [Run3]

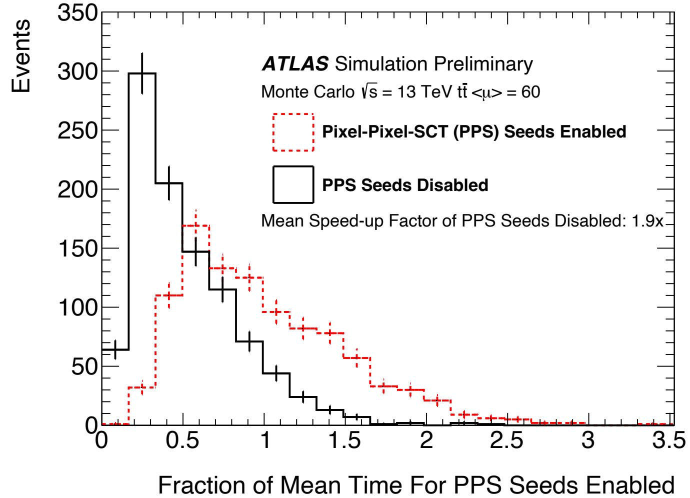
- ML predictor in seeding



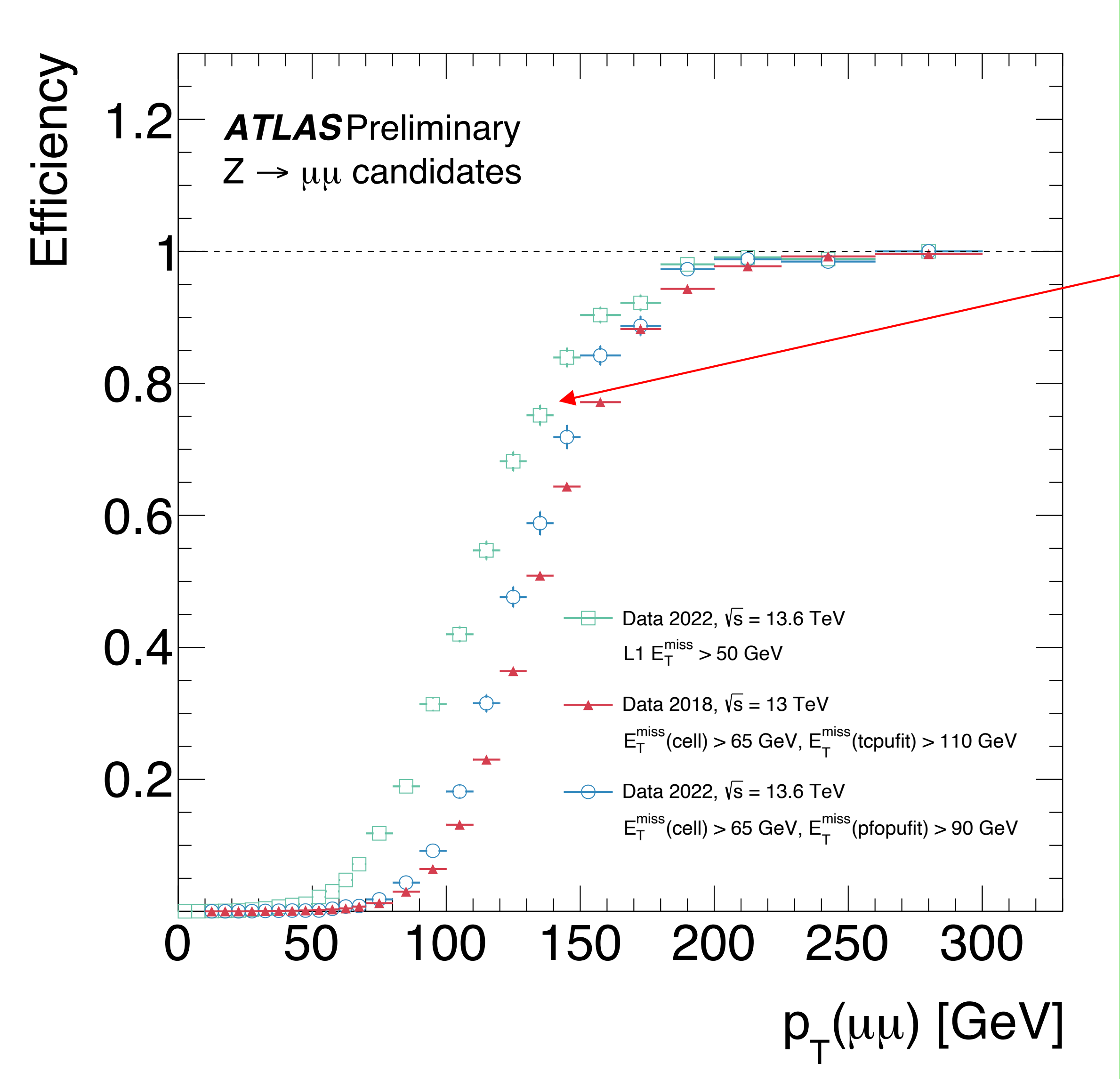


- Seeding strategy/window optimization





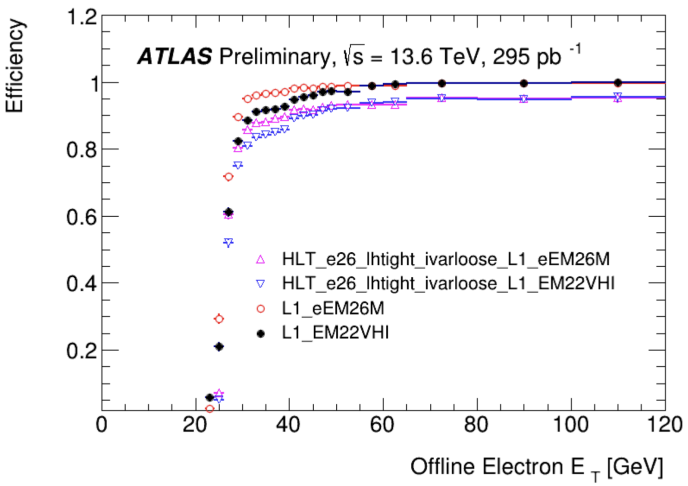
Full-scan tracking at HLT [Run3]



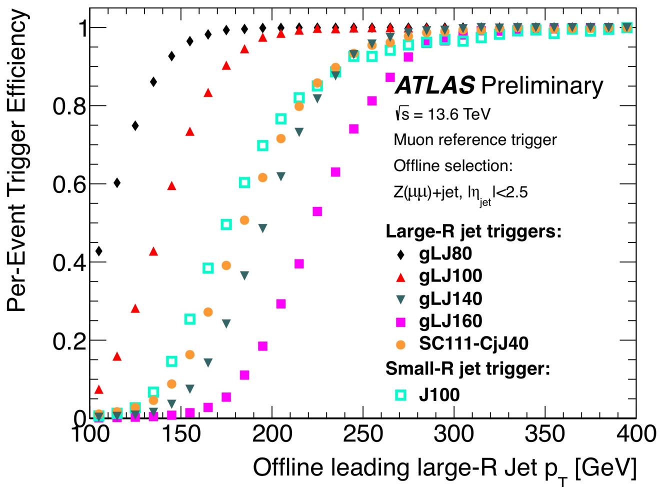
Improvements in ETmiss trigger

Trigger Performance [Run3]

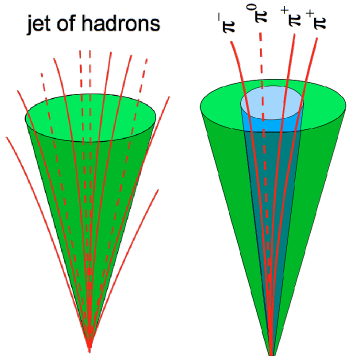
- Electron
  - Improved efficiency with L1 eFEX

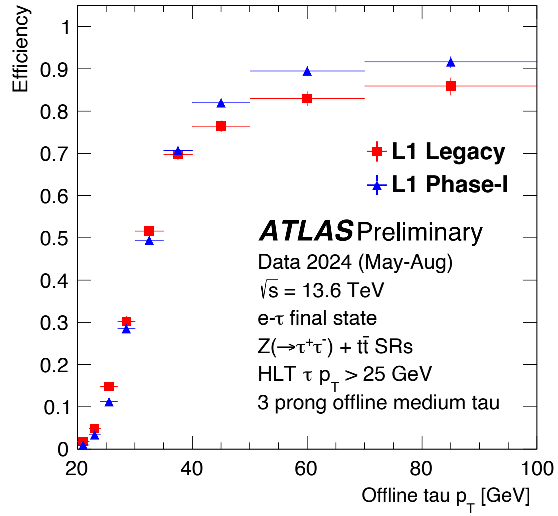


- Large-R jets
  - New opportunities at L1



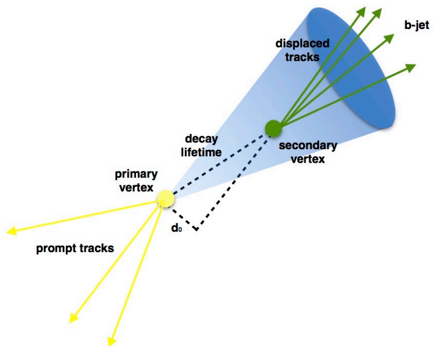
- Tau
  - L1 eFEX/jFEX combined, improved isolation
  - Calo and Track info combined at HLT via ML (RNN) algorithms

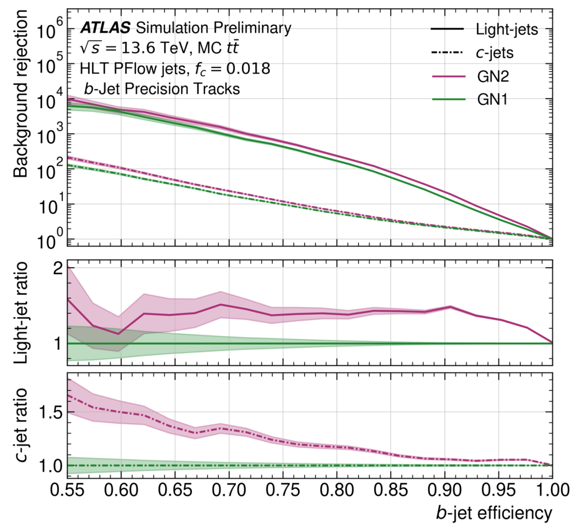




- Jets
  - Multi-jets utilize full-scan tracks, objects much close to Particle Flow (Pflow) objects at offline

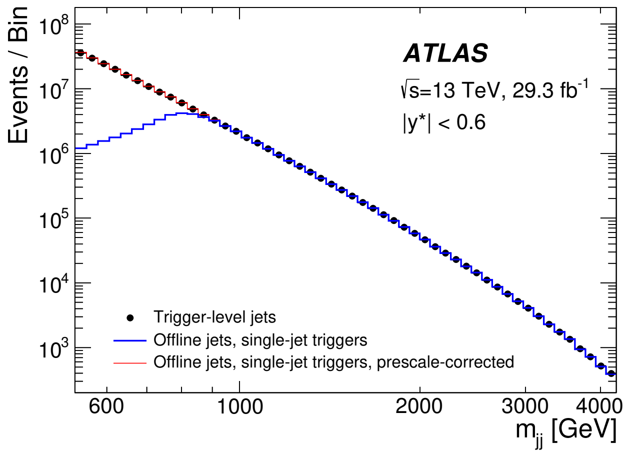
- b-jets
  - Utilizes Pflow jets and tracks and vertex with ML (DNN) algorithms

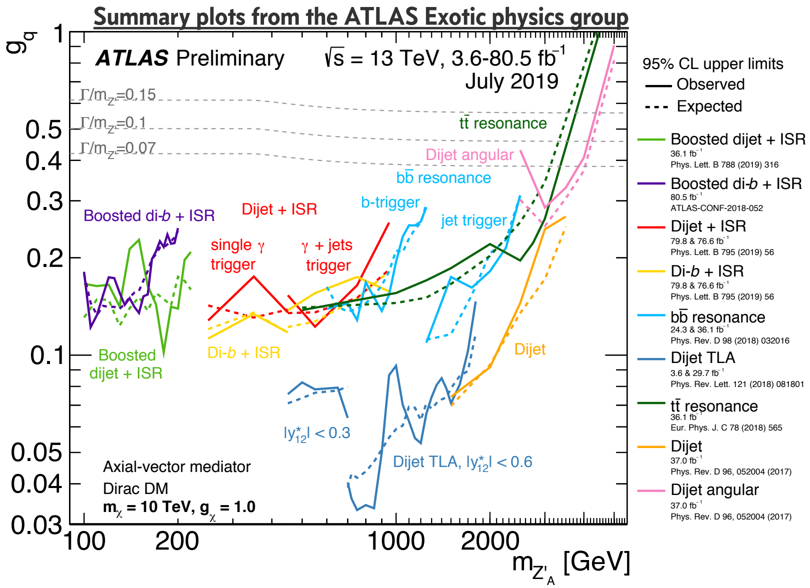


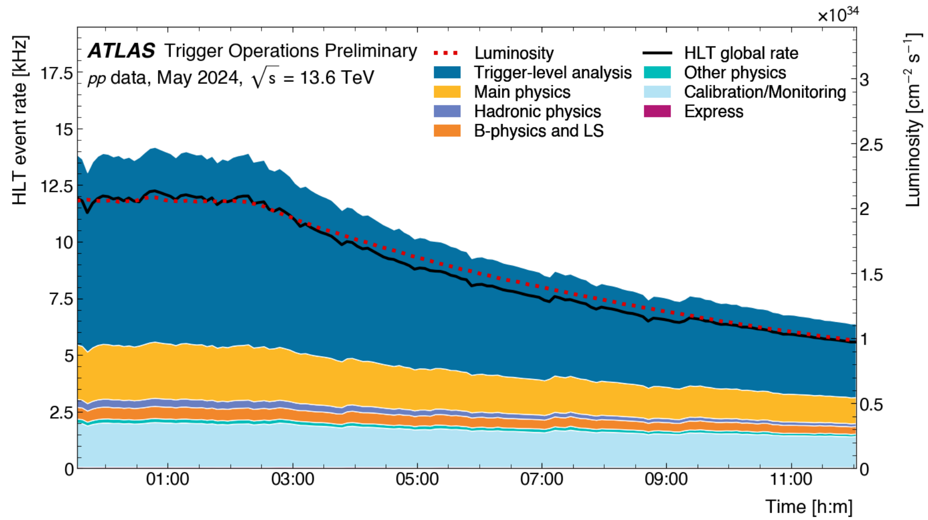


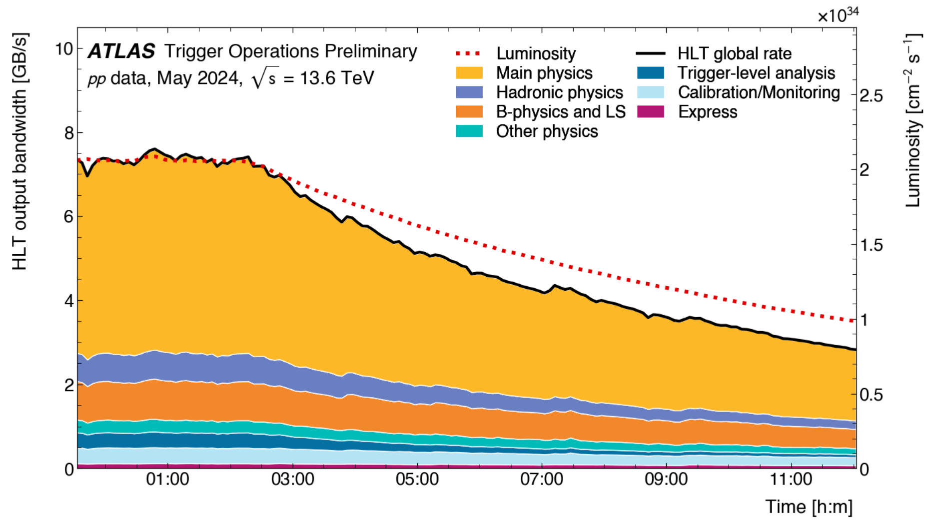
Trigger level analysis (TLA)

- TLA: record trigger level objects only → high rate (relaxed prescale)
  - Can extend physics reach at low p_T/mass regions



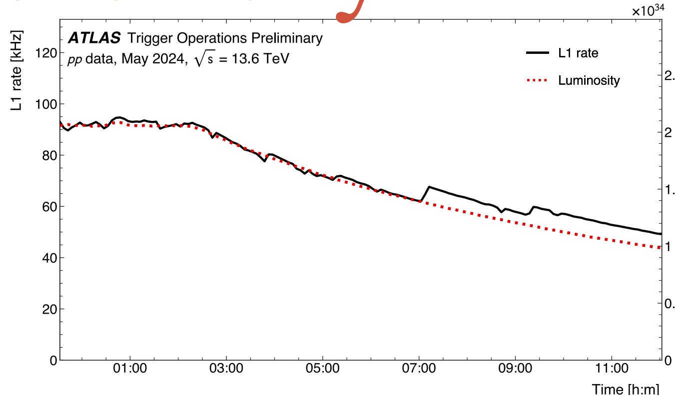


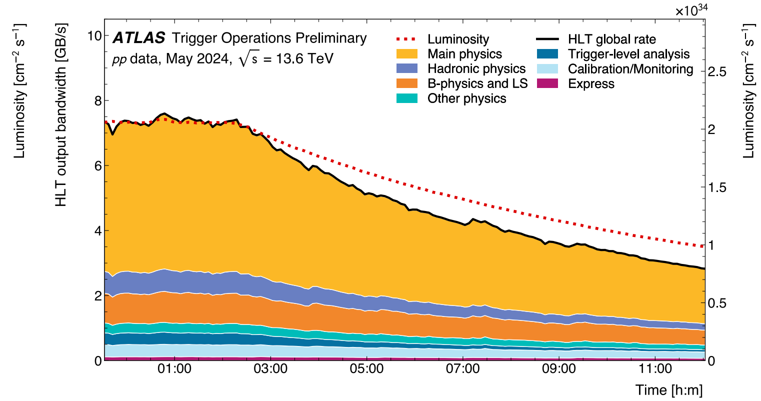




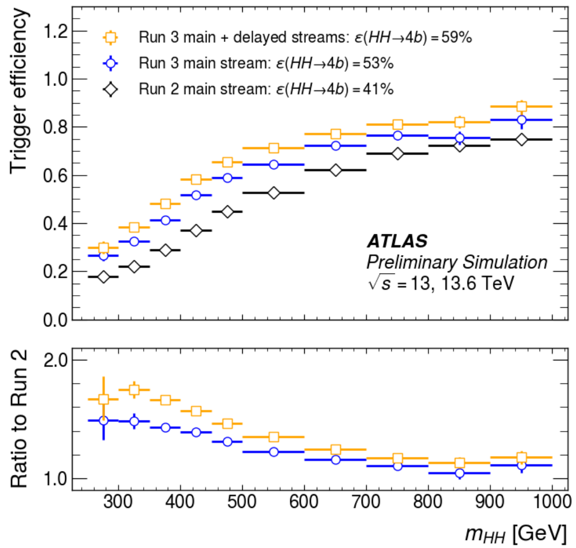
Enhancing more physics sensitivity

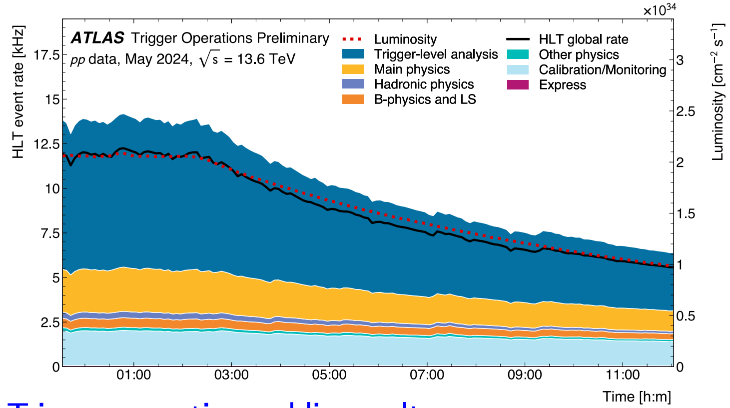
- Trigger menu optimized with physics coordination to achieve
  - Maximum physics output (physics priorities, high efficiency for rare processes, supporting triggers for calibration/efficiency measurements, …)
  - Keep L1 rate to 95 kHz to minimize the deadtime from readout
  - Total bandwidth within 8 GB/s
  - -> Menu vs. luminosity/time : constantly fine-tuned according to running conditions and new developments





- Delayed stream
  - Larger recording bandwidth that undergo event reconstruction only when offline resource allows
  - B-physics and hardonic signatures (such as HH->bbbb, bbττ)



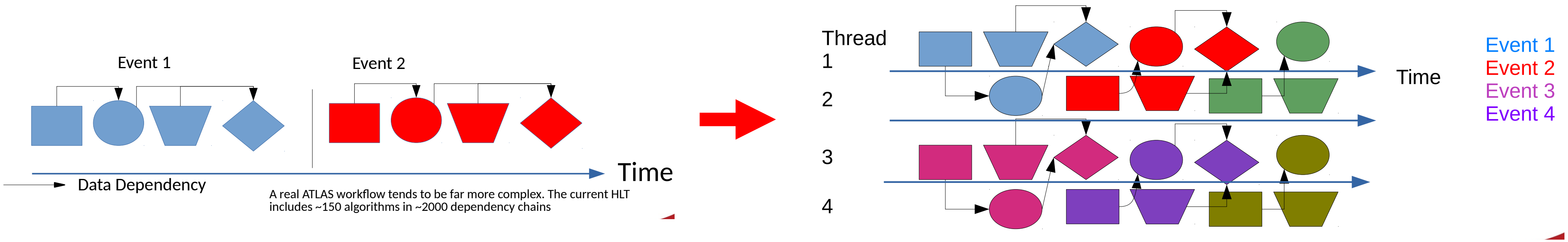


TLA: high rates with small bandwidth usage

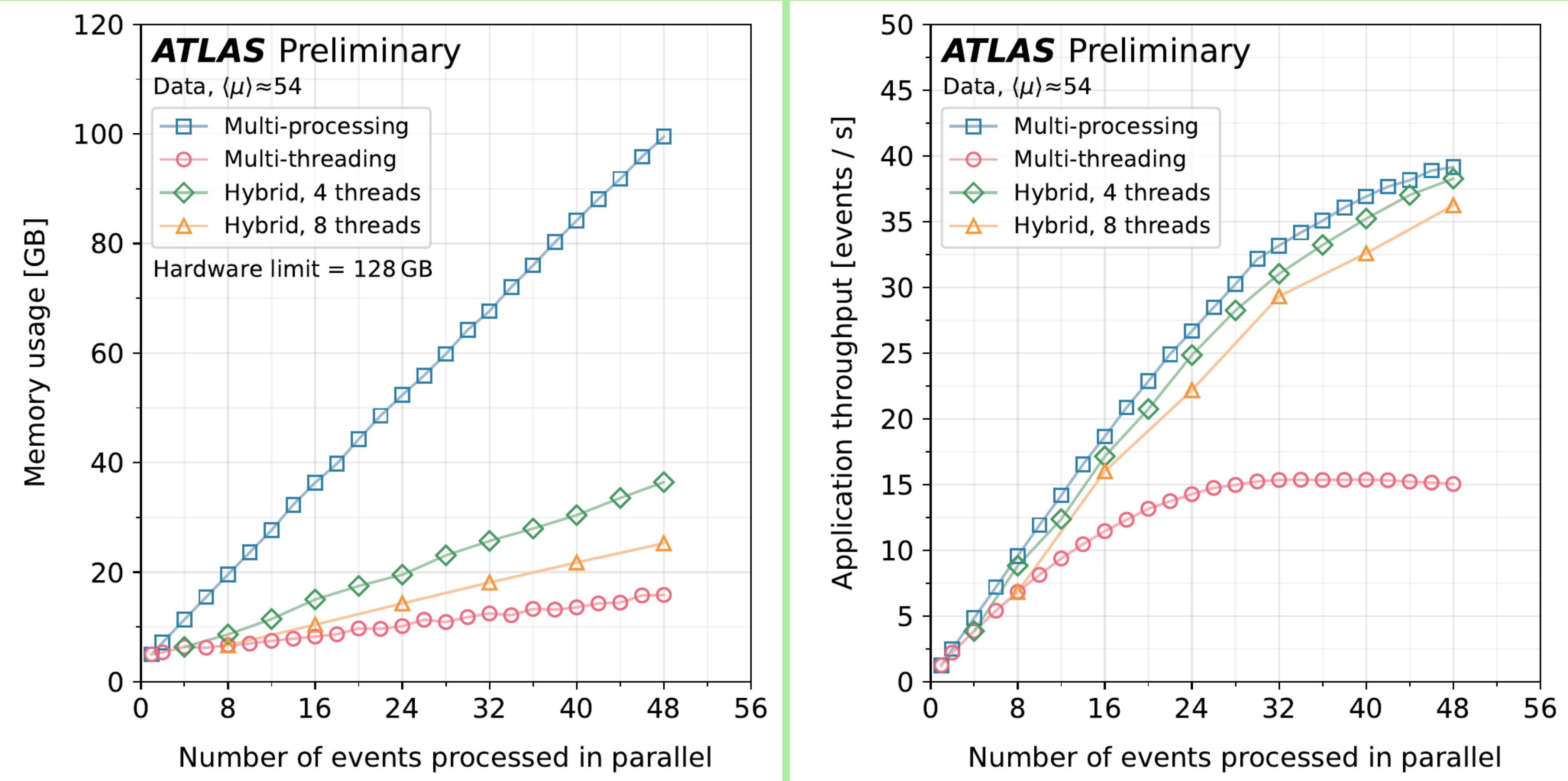
Trigger operation public results

Trigger software

- Processing power demand has been increasing (e.g. full-scan calo/track), resulting in that HLT processing power becomes a bottleneck
  - “Luminosity leveling” operation
- Almost complete redesign and rewrite of HLT framework together with significant updates in all HLT software
  - ATLAS software framework adopted to support multi-theading (AthenaMT)
    - Multi-processing based on Intel’s Threading Building Blocks (TBB)
      - Excellent memory sharing across computing cores ← doubling memory is much expensive than doubling core
  - HLT framework takes full advantage of the AthenaMT scheduler
    - Up to Run2: HLT framework is one single algorithm steering all trigger algorithms

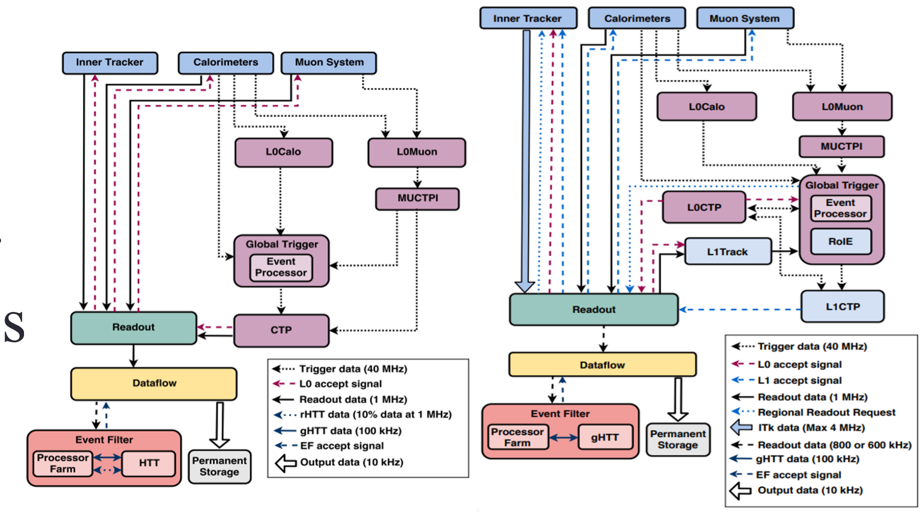


Trigger software

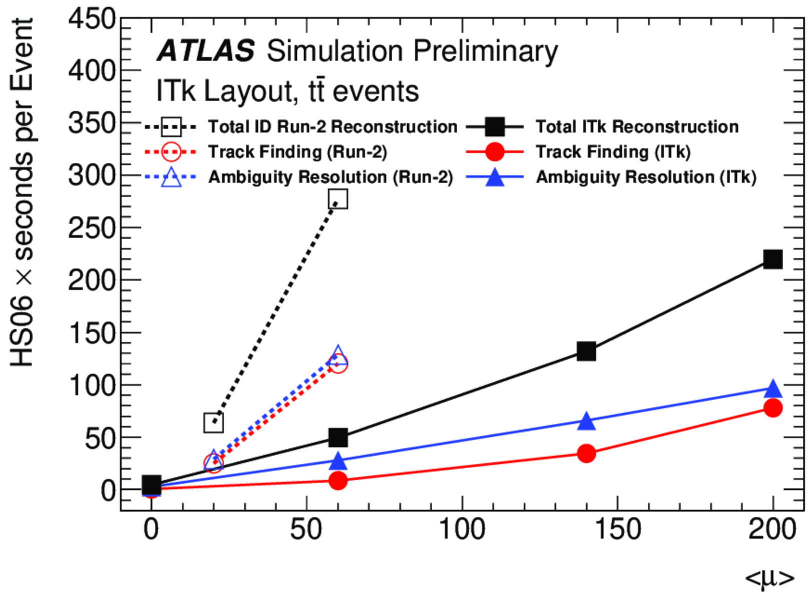


ATLAS Track Trigger

- Hardware track trigger
  - Once tried in Run2 ("FTK") : the aim was for 100 kHz (all L1) full-detector tracking with custom-made electronics → project canceled
  - Initial HL-LHC TDR includes an option of hardware track trigger at Event Filter (with possible migration to L1), which was recently updated to either software-only or heterogeneous computing (e.g. commodity acceleration)
    - Tracking software was largely optimized, resulting in a significant reduction in processing time



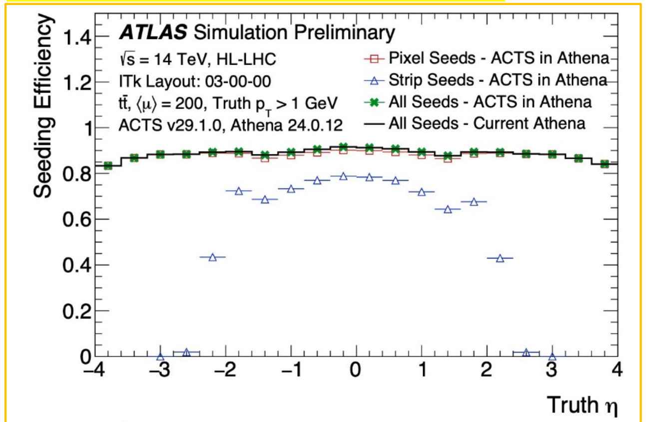
(a) (b)

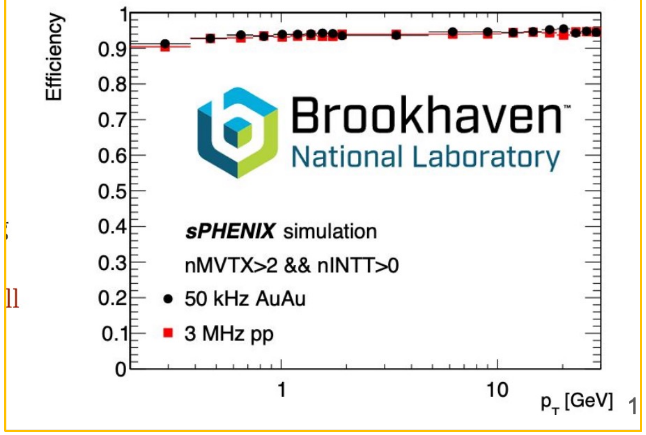


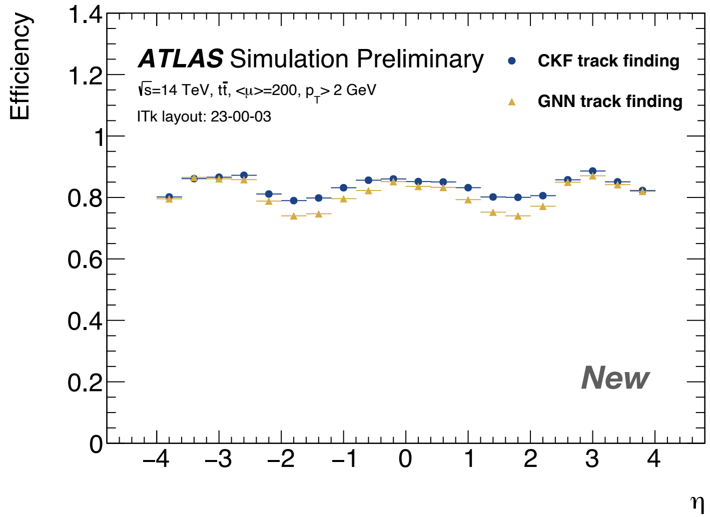
Trigger Tracking [HL-LHC]

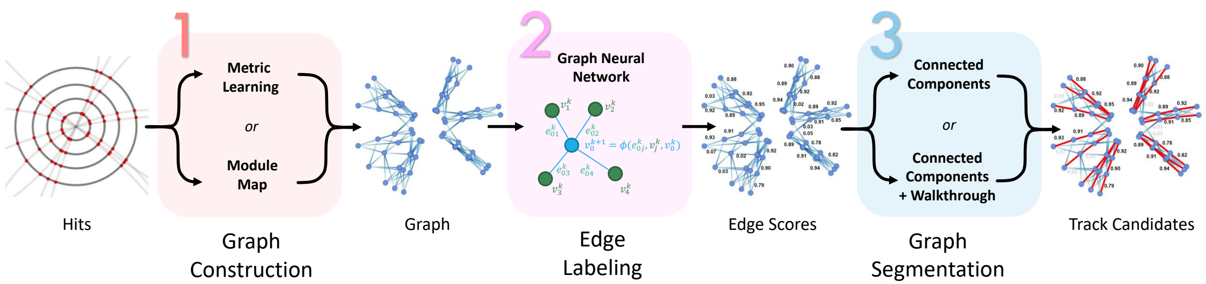
L. Coelho at CTD2023

- A Common Tracking Software "ACTS"
  - Experiment-independent toolkit with modern language (C++)
    - ATLAS, sPHENIX, ePIC, ALICE, NA60+, CEPC, R&D platform
  - Extending to parallelization → on GPU
- Exploring CPU, GPU, FPGA based solutions
  - Various R&Ds ongoing toward 'technology choice' decision in the next year
  - Including state-of-the-art tracking algorithm such as using GNN which is getting know to be suitable for tracking

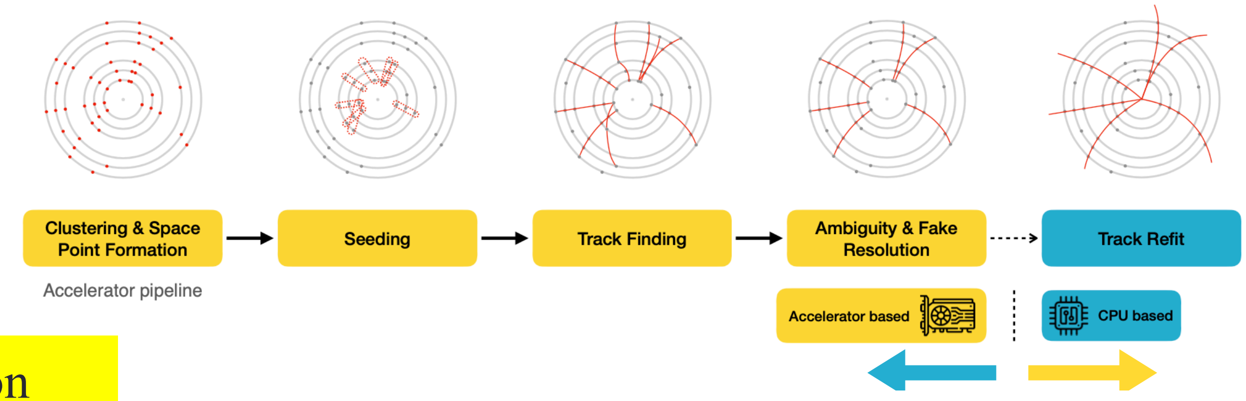




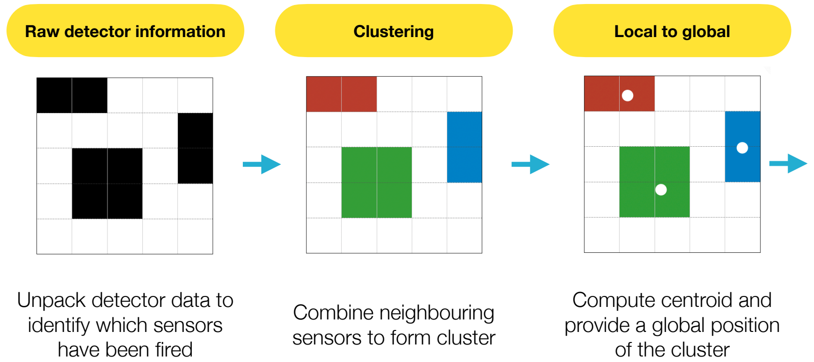




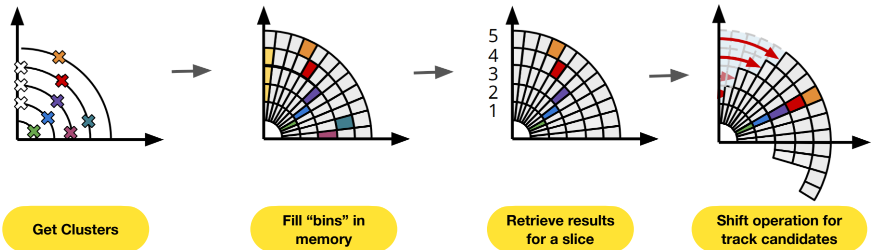
ex.) FPGA Tracking

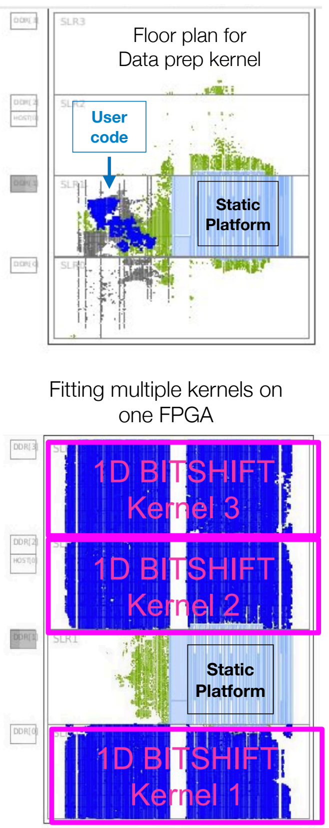


Data Preparation

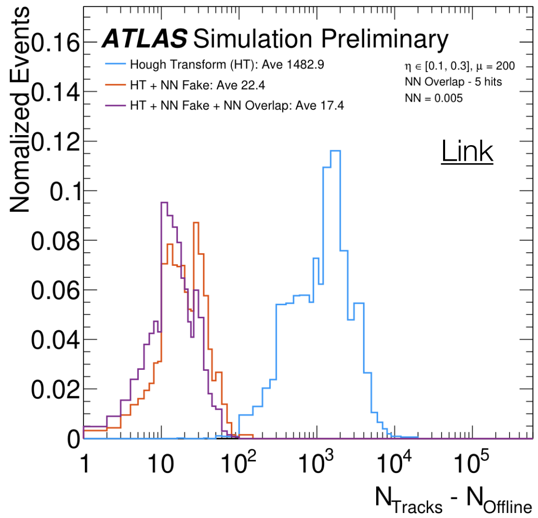


Seeding/Finding : 1d bitshift



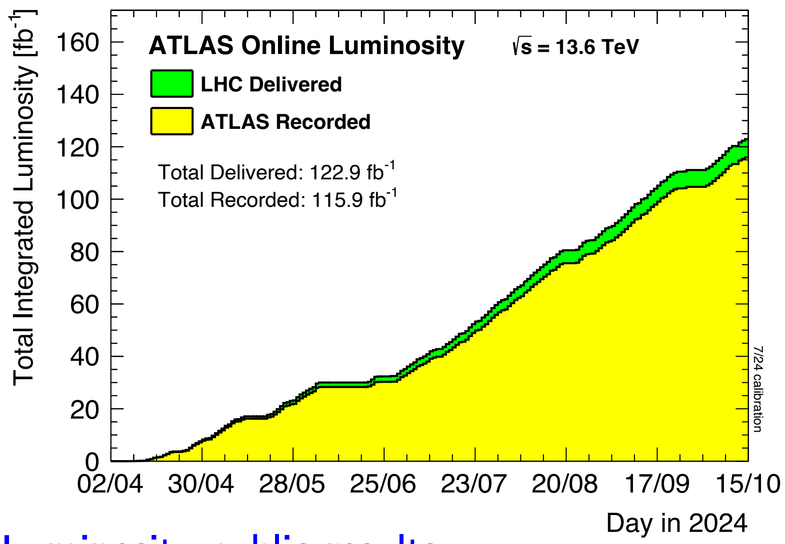


Ambiguity&Fake reduction using ML



Summary

- Completion of Run3 ("Phase-I") Upgrade
  - Full integration of NSW in L1 Muon
  - L1Calo full upgrade
- This year was an important milestone
  - More than 100 /fb collected! (record)



Luminosity public results

- And, already looking at the future
  - ~200 interactions per bunch crossing (pileup)
  - New inner detector trackers (all silicon)
  - Various R&Ds for increased usage of GPUs and FPGAs both at software and hardware levels
  - Large number of ML techniques already deployed and further being exploited