Level-1 Trigger Online Software

Simone Bologna[1], Tom Williams[2]

[1] *University of Bristol*
[2] *Rutherford Appleton Laboratory*

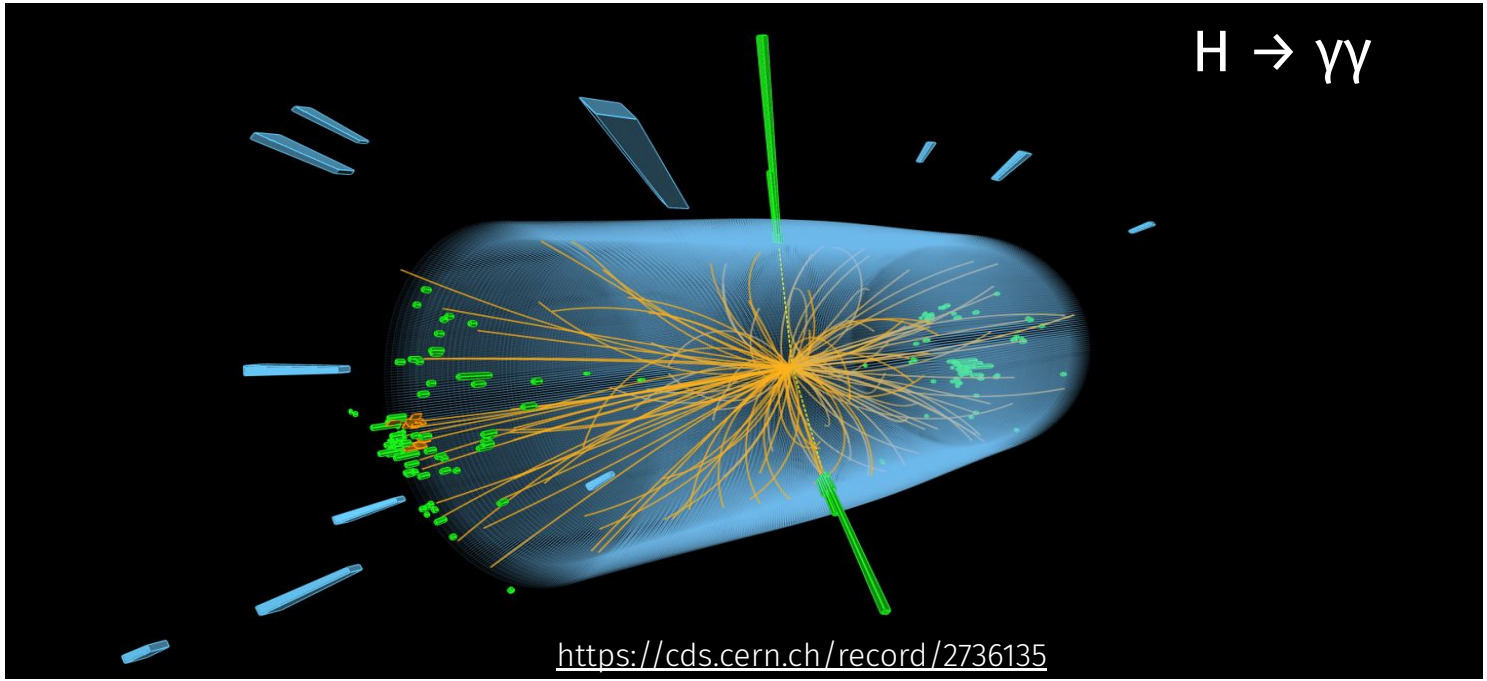
CMS induction course, 21/09/2021

Outline

- CMS level-1 trigger
- Online software tools
  - Control
  - Configuration management
  - Monitoring
- A glance into the future

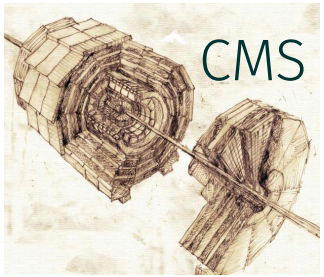
Goal of CMS

Record results of LHC collisions, to understand fundamental laws of physics



https://cds.cern.ch/record/2736135

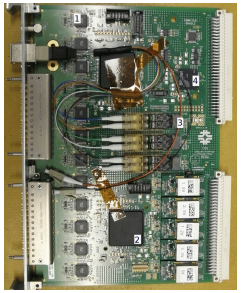
CMS data flow



40 MHz

40-80 TB/s

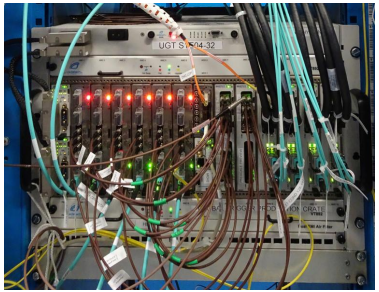
Readout



40 MHz

~ 1 TB/s

Accept/reject



Level-1 trigger

100 kHz

100 GB/s

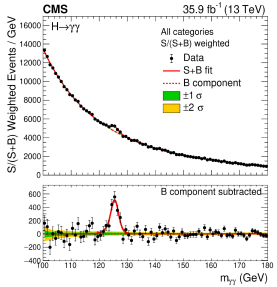
Event builder &
High-level trigger

1 kHz

~ 1-2 GB/s

300 nHz

100 papers/y



CMS data flow

40 MHz

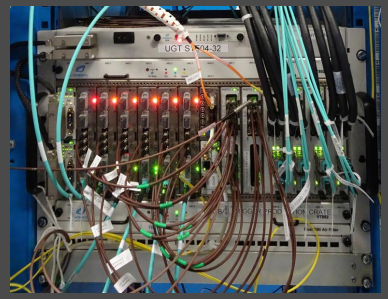
40-80 TB/s

Readout

40 MHz

~ 1 TB/s

Accept/reject



Level-1 trigger

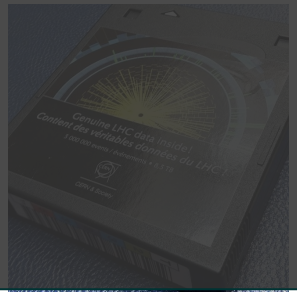
100 kHz

100 GB/s

Event builder &
High-level trigger

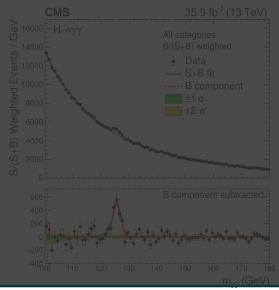
1 kHz

~ 1-2 GB/s



300 nHz

100 papers/y



Focussing on level-1 trigger in this talk

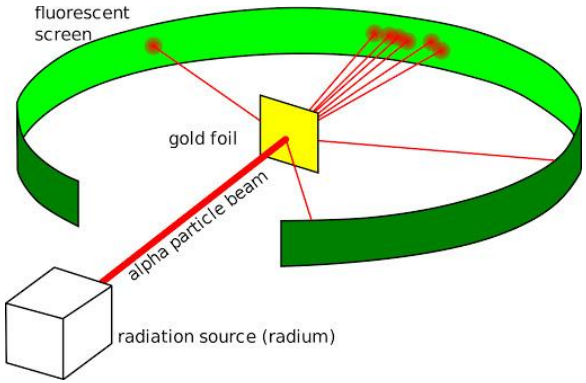
Triggers & online SW

- Trigger recognises interesting activity in the detector
- Online Software is responsible for …
  - Configuration
  - Control
  - Monitoring
- If an event isn't accepted by the trigger, it's **lost forever**
  - L1T online SW: Critical role in data taking — must be designed carefully to …
    - Ensure reliable, robust behaviour
    - Allow on-call experts to rapidly resolve problems
    - Maintain high uptime
      (in addition to other requirements)
  - Achieve the above despite complexity of the level-1 trigger (& upstream systems)

First, for contrast: Brief interlude on some trigger systems from previous generations of particle/nuclear experiments …

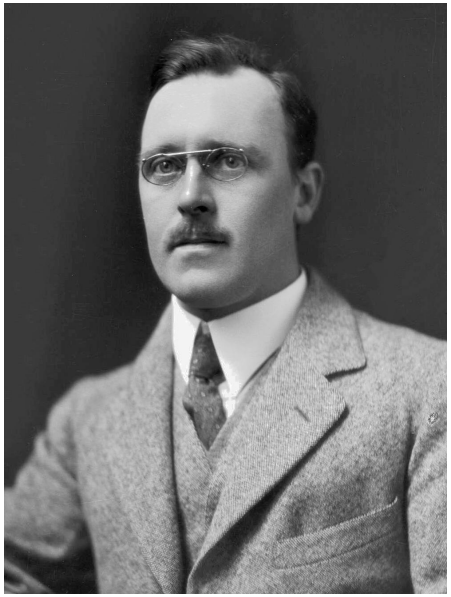
Triggers

- Rutherford scattering expt: Geiger and Marsden manually counted α particles scattered at given angles
- Acted as organic TDAQ (trigger & data acquisition) system
  - Low event rate



Picture of an early TDAQ system

Hans Geiger
Post-doc



Ernest Marsden
PhD student

Triggers

- Human brains are not designed for boring, repetitive tasks
- Machines do not get bored!
  - And can process readings from detectors quite a bit faster than us
- Idea already considered in 1919
  - “visual or audible methods of counting are quite trying on the nerves ... A self-recording device would therefore be an obvious improvement”
    - Kovarik on automatic recording device for the Rutherford scattering experiment
      Phys. Rev. 13, 272, 1st April 1919

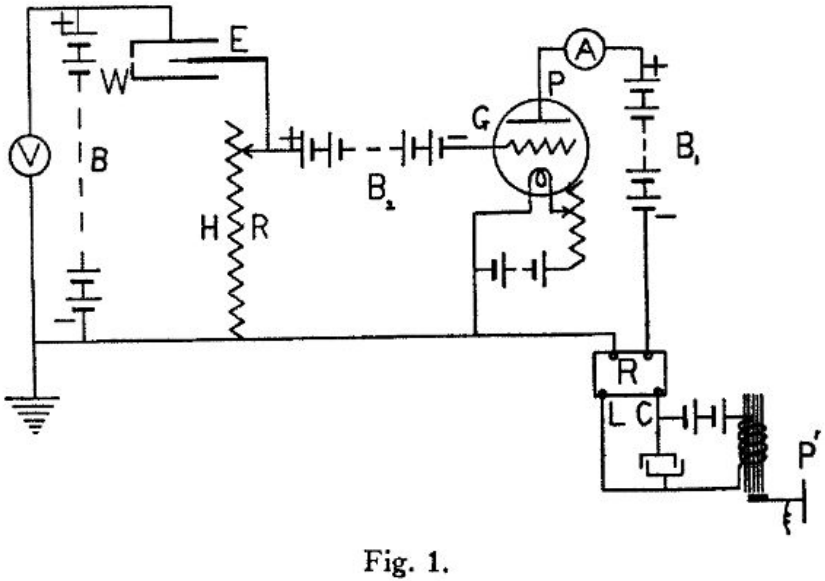


Fig. 1.

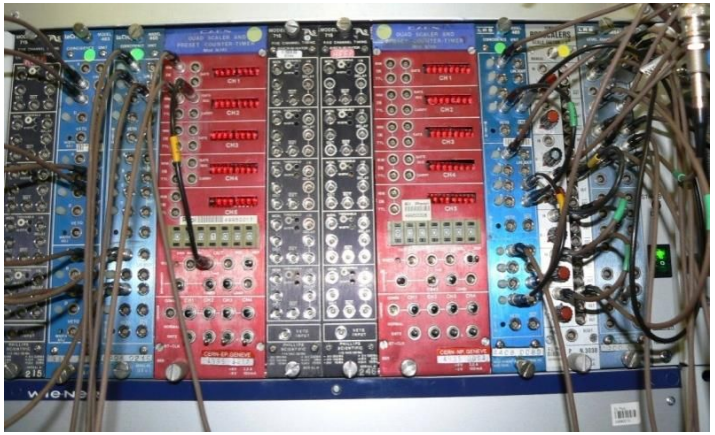
Recording system proposed by Kovarik

Phys. Rev.13, 272 , 1st April 1919

Automated triggers

- More sophisticated TDAQ systems developed as experiment goals became more ambitious
- Early systems: Relatively simple, based on coincidences and thresholds of analogue signals
- UA1 & 2 @ SppS, 1981 — first example of experiment with broad physics program and generic triggers
  - 250kHz → ~ 5Hz, three level trigger
- Complex trigger → sophisticated control & monitoring software

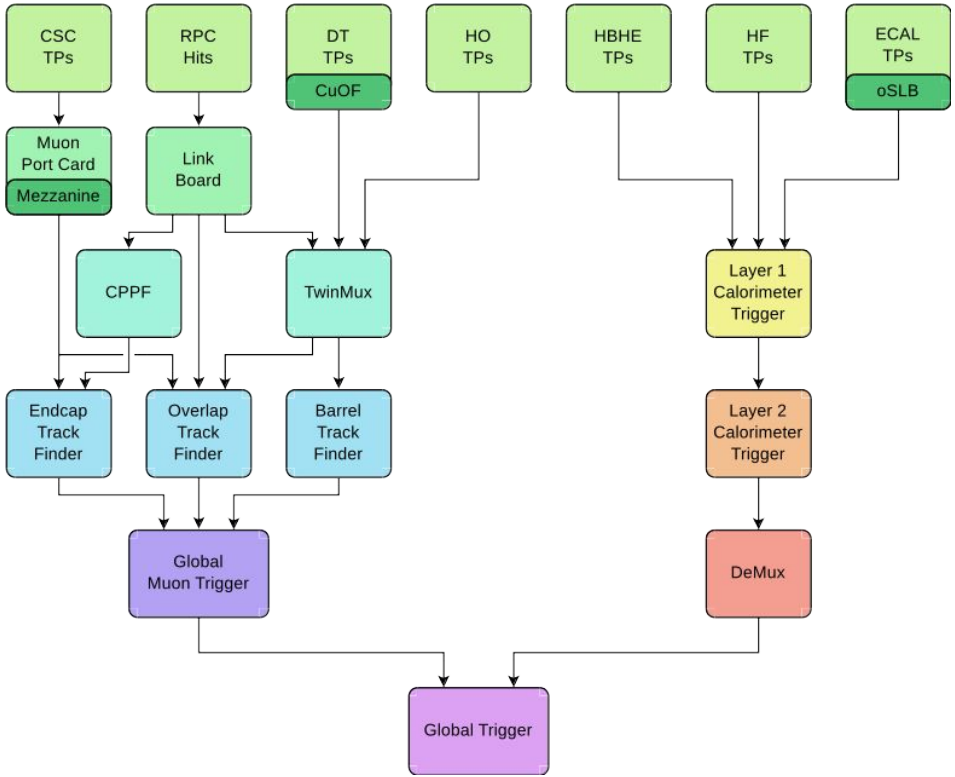
NIM crate



Signal discriminator

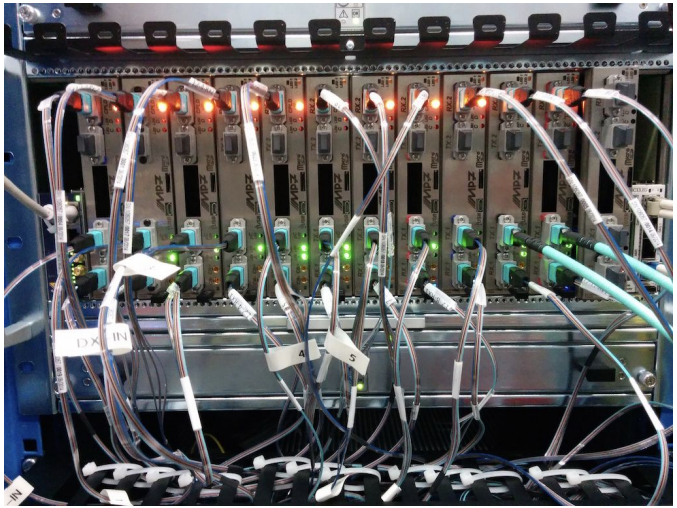
CMS level-1 trigger

- ~ 150 boards
- ~ 3000 optical links
  - Up to 10Gbps
- Complex system
  - Input from 5 detector systems
    - 9 subsystems → different data-processing algorithms (calo clusters, local μ reco ...)
  - 5 different board designs



CMS level-1 trigger

- Configuration settings
  - Must be easy to edit and manage
- Status continually monitored
  - Intelligent analysis
    - Concise error summaries & useful suggestions
  - Rich user interfaces
- ✓ **Rapidly solve problems**



CMS level-1 trigger

- Configuration settings
  - Must be easy to edit and manage
- Status continually monitored
  - Intelligent analysis
    - Concise error summaries & useful suggestions
  - Rich user interfaces
  - ✓ **Rapidly solve problems**
- Online software is crucial
  - Design & implementation far from simple

Example scenario: Anomalous signals from ECAL electronics

- CMS running stably
- In the middle of the night, L1T rate suddenly increases to 40MHz
  ⇒ *can't record data until problem fixed*

➢ Shift crew phone on-call expert

➢ To solve this, they need to quickly …

- Identify which of the several hundred trigger paths have high rates & which don't (e.g. electron paths)
- Optical links: Check whether there are any CRC errors, and check for anomalous data values

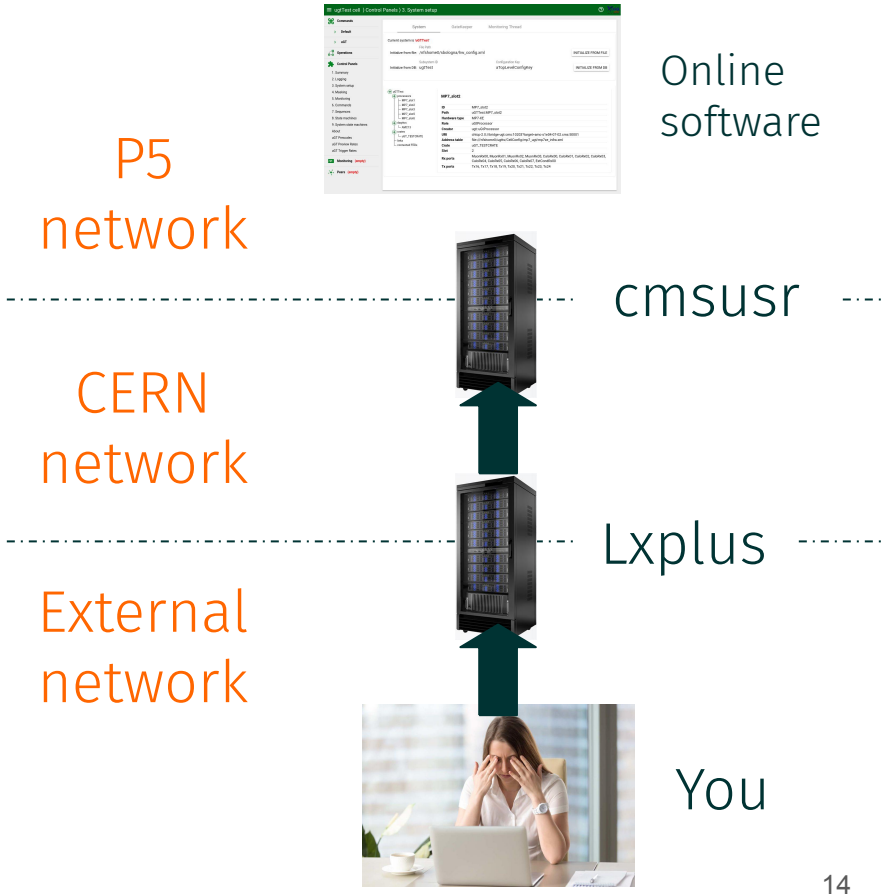
Losing data for every second of delay

Level-1 online software: Key components

- Configuration and control
  - SWATCH & SWATCH cell
  - Level-1 configuration editor
- Monitoring
  - L1page
  - SWATCH & SWATCH cell
  - Data Quality Monitoring (DQM)
    - Not part of online software itself, but essential for operating trigger

Where are the tools?

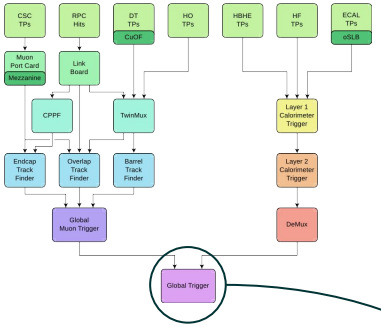
- Online software, together with all software required to run CMS, is installed in a separate network
  - Network not directly reachable from outside
- Internal resources accessible via tunnel
  - Much safer and easier to protect



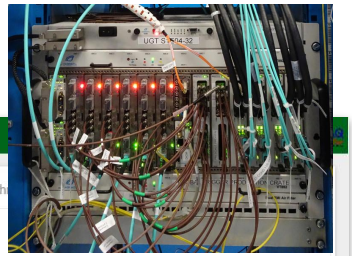
Configuration & control

SWATCH & SWATCH cell

- SWATCH = SoftWare for Automating the conTrol of Common Hardware
- Centrally-developed library that provides unified software interface to control, configure and monitor trigger hardware
- Each subsystem is managed by a SWATCH cell



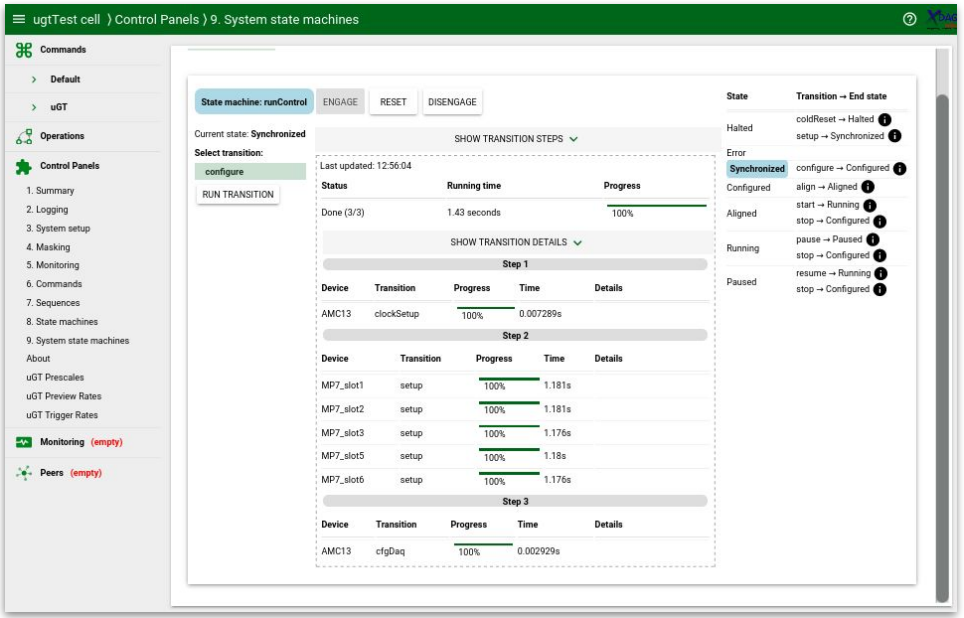
Crate setup



SWATCH

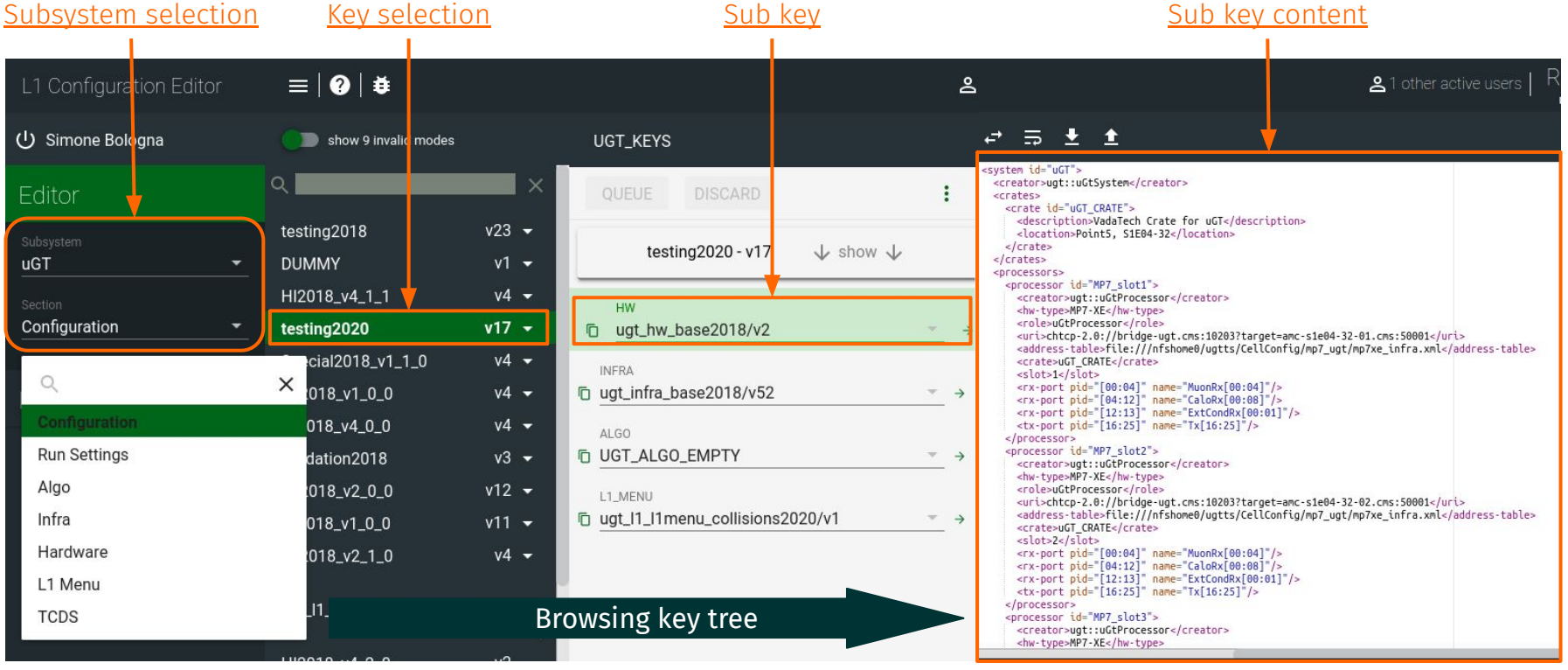
- Individual subsystems can be configured from a browser via the SWATCH cell web panels
  - E.g. for manual tests/debugging
- Each SWATCH cell can also be controlled by a central cell, that coordinates the entire trigger
  - I.e. 1 central cell instance for whole L1T
  - Configuration settings taken from central database & edited via L1CE

Subsystem configuration page



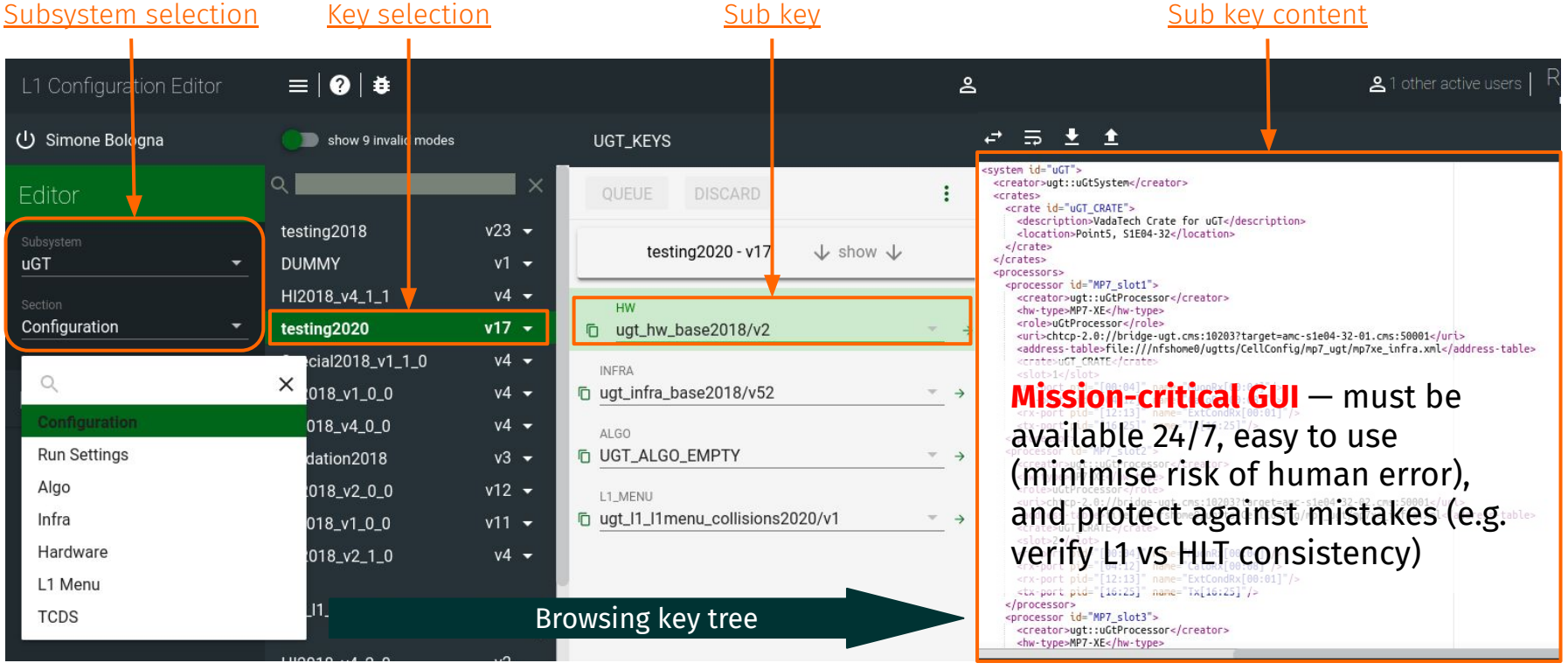
Level-1 Configuration Editor

- Configuration settings are stored in a database, with specific configuration chosen by selecting key from shift crew web interfaces before the start of a data-taking run
  - Top-level key made of subkeys (2 per subsystem), each containing XML



Level-1 Configuration Editor

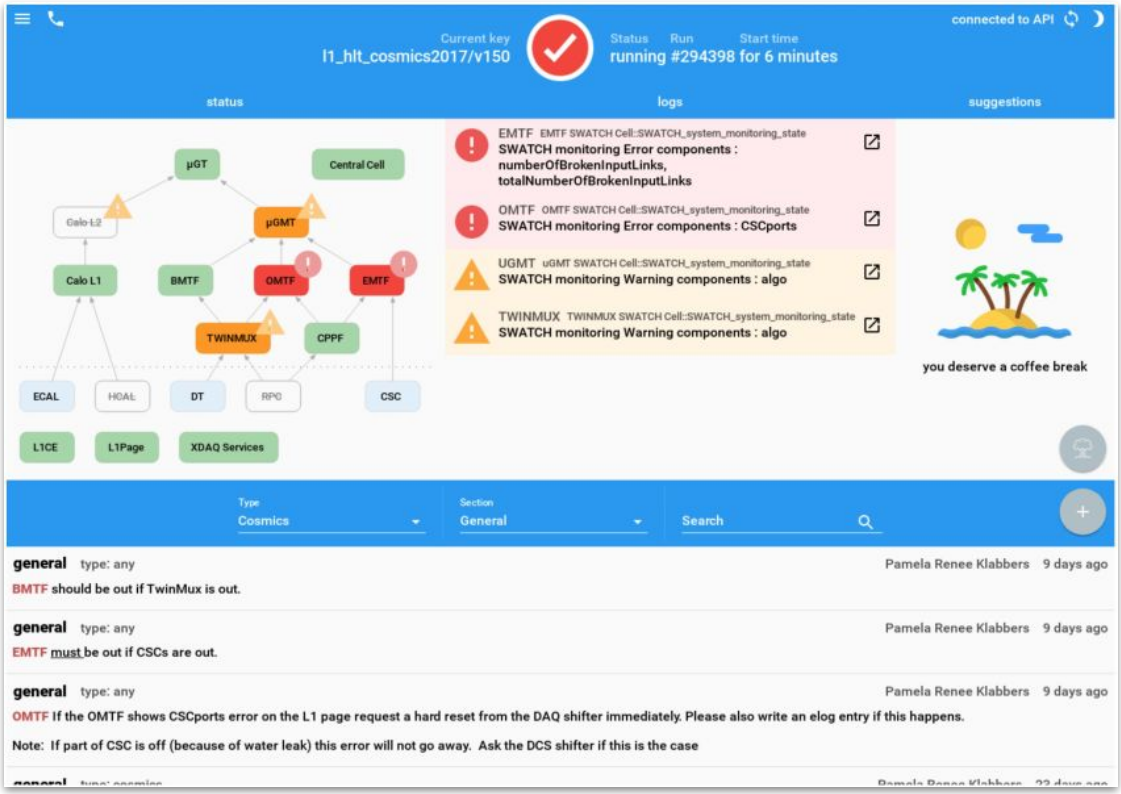
- Configuration settings are stored in a database, with specific configuration chosen by selecting key from shift crew web interfaces before the start of a data-taking run
  - Top-level key made of subkeys (2 per subsystem), each containing XML



Monitoring

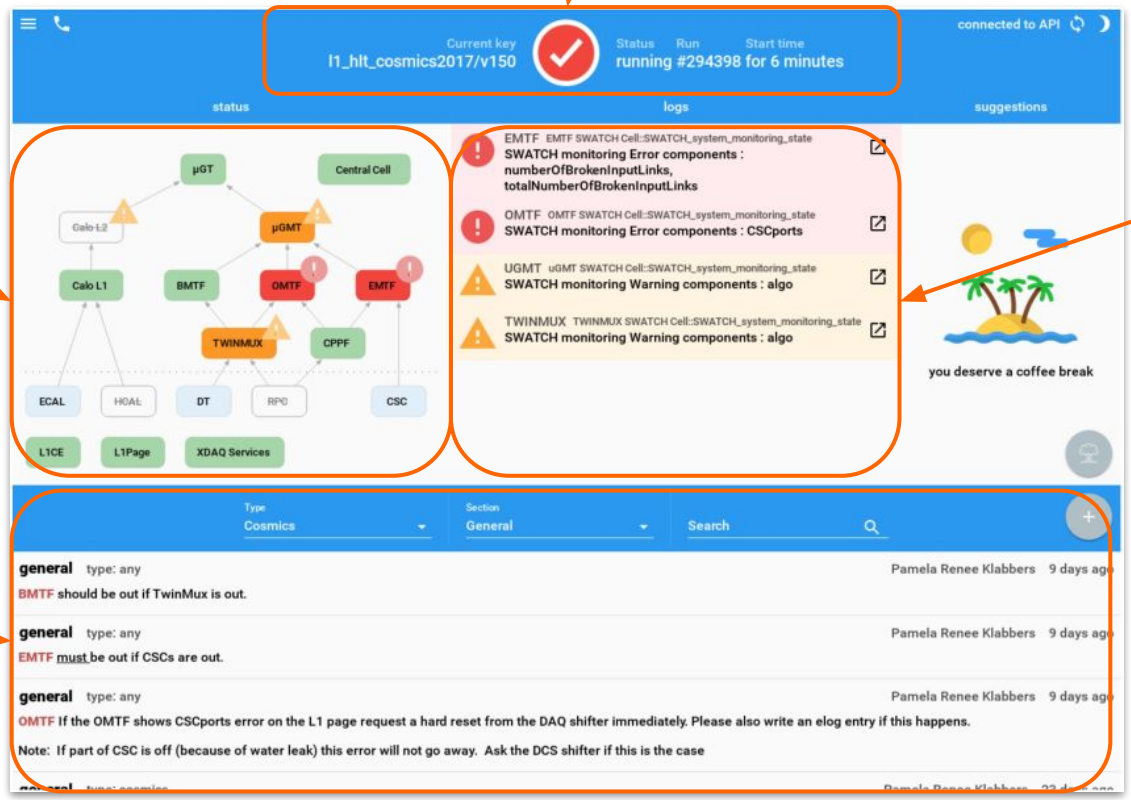
Level-1 page

- Main dashboard for L1T status



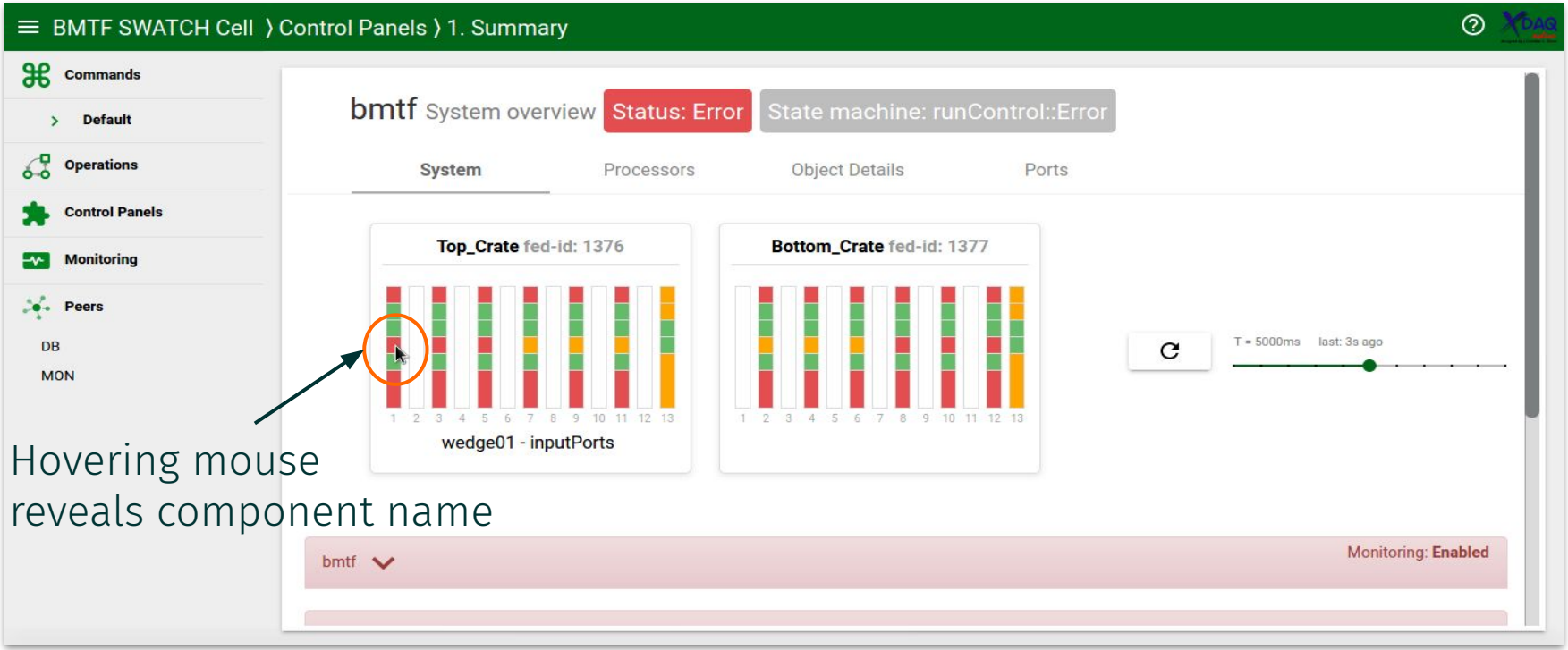
Level-1 page

- Main dashboard for L1T status



SWATCH cell: Monitoring

SWATCH → generic interface for subsystem/board monitoring info.



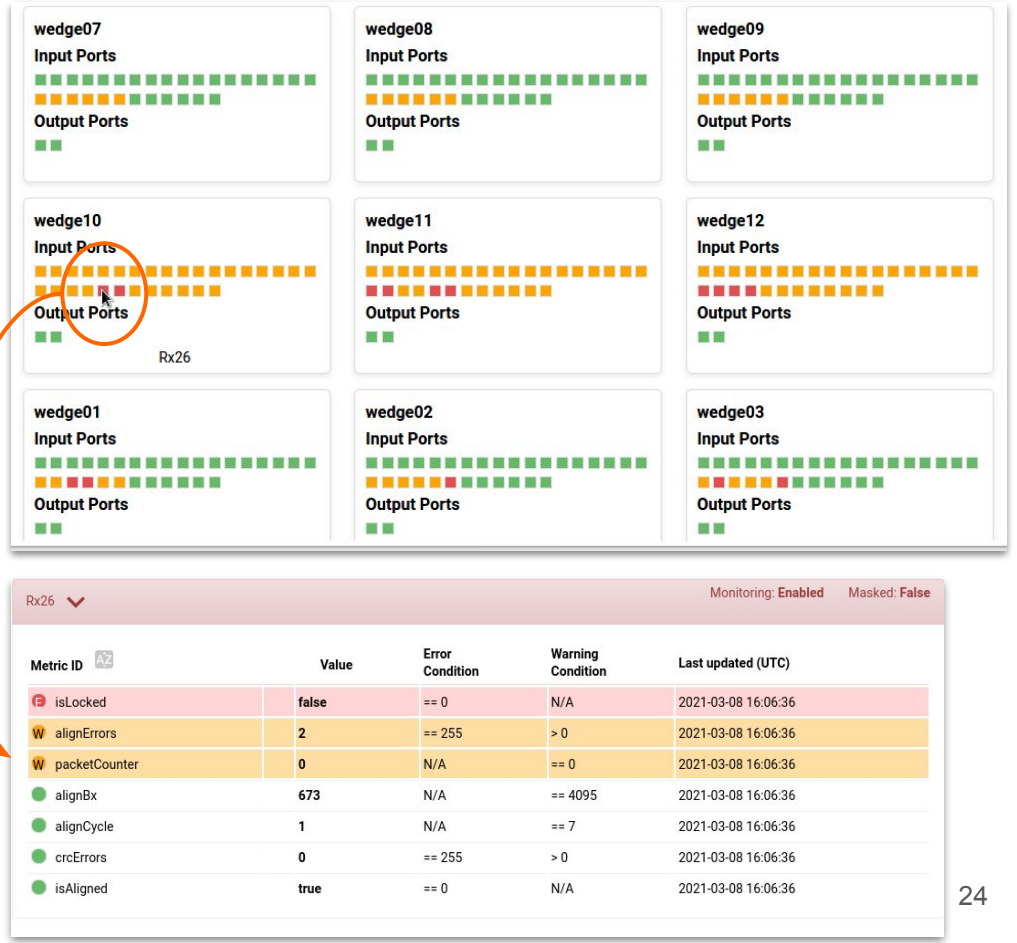
Hovering mouse reveals component name

SWATCH cell: Monitoring

I/O port dashboard

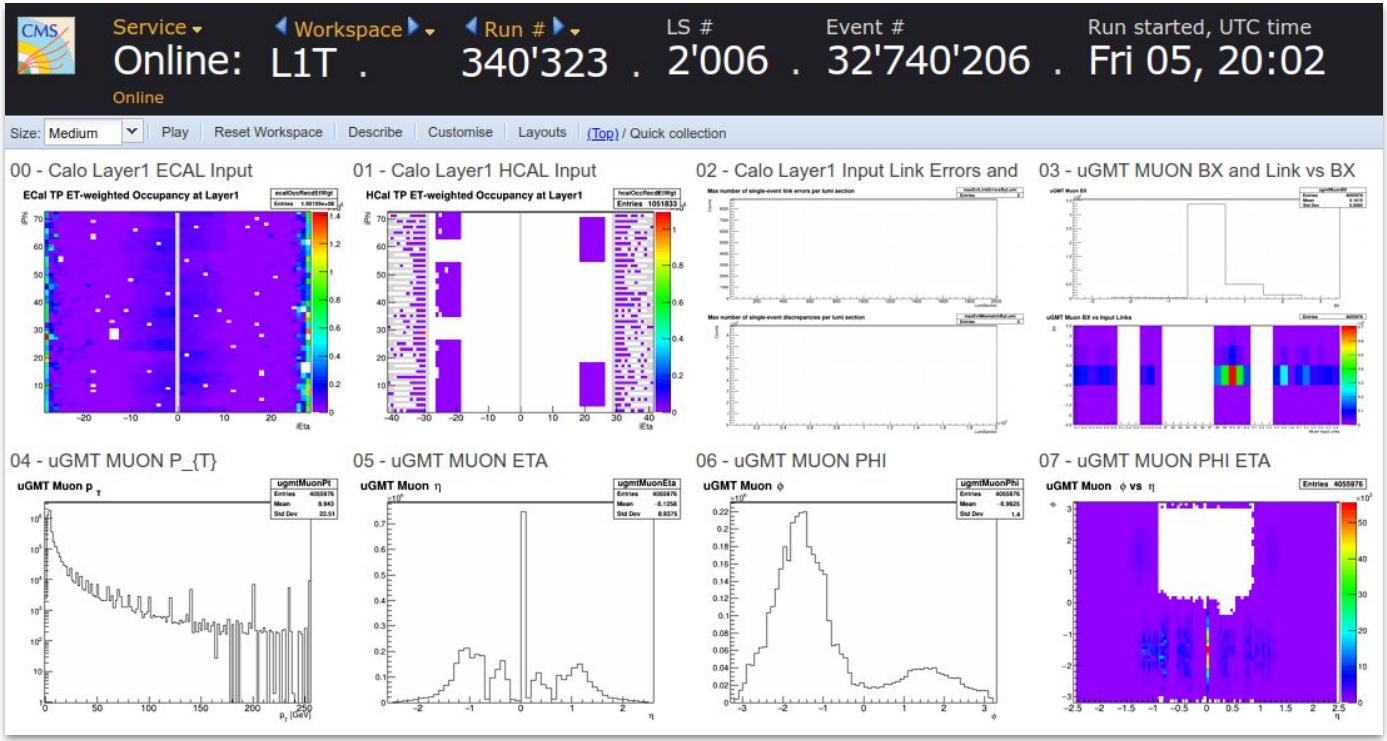
Highlights problematic components

Gives immediate access to in-depth info so that on-call experts can rapidly diagnose and solve the problem



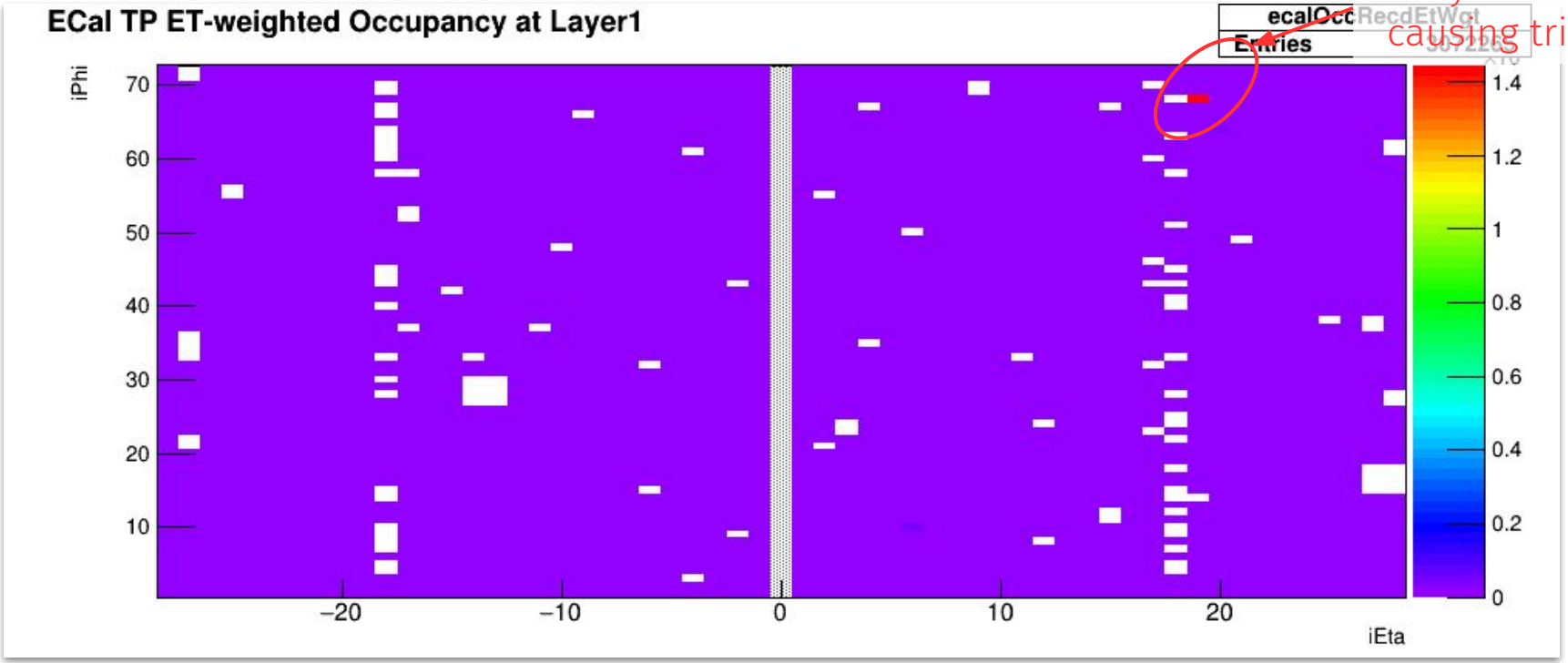
Data Quality Monitoring

DQM ⇒ dashboard for prompt feedback on the quality of the collected data



Data Quality Monitoring

E.g. we sometimes get noisy calorimeter towers ⇒ high L1 rates



Updating the monitoring system

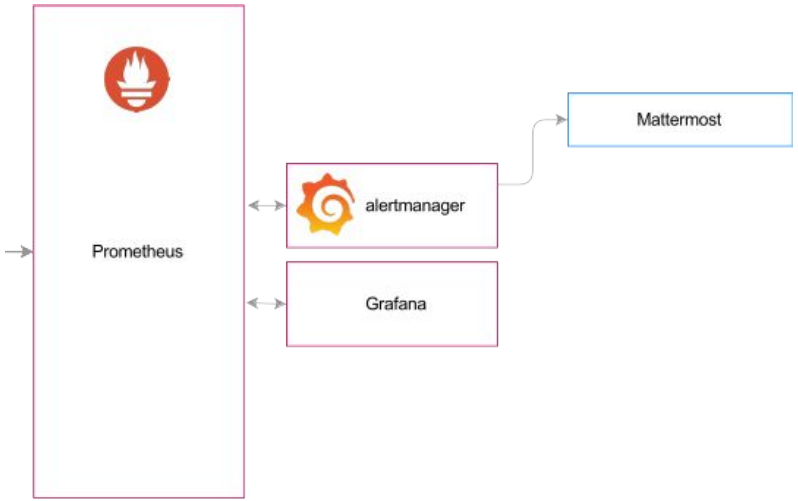
- Software tools shown so far got us through Run 2, but continuous improvement with new tools essential to further reduce downtime and further increase reliability

- Over last few years, have been integrating new industry-standard tools
- Goals
  - Improve monitoring dashboards
  - Develop richer and smarter alerting
  - Reduce overall stress on experts
    - E.g. by suppressing warnings that we know on-call team doesn't need to react to

New monitoring tools

- All metrics fed into a central Prometheus monitoring DB instance
- Alertmanager for alerting
- Grafana for system-wide dashboards

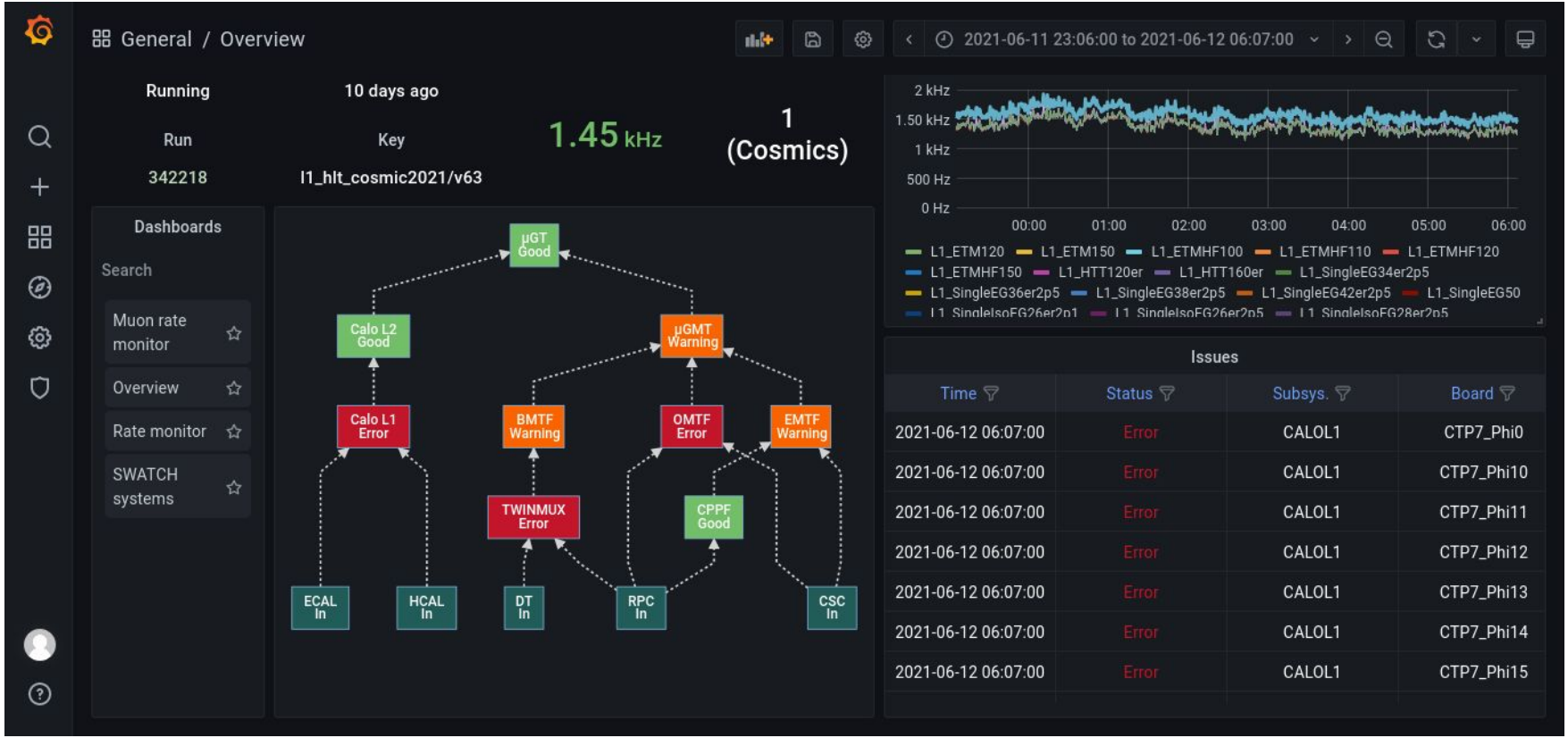
- Early testing and commissioning of the new infrastructure on OpenShift cluster (Kubernetes customisation)
  - Components are containerised for easy deployment
- Commissioning to production environment started in June



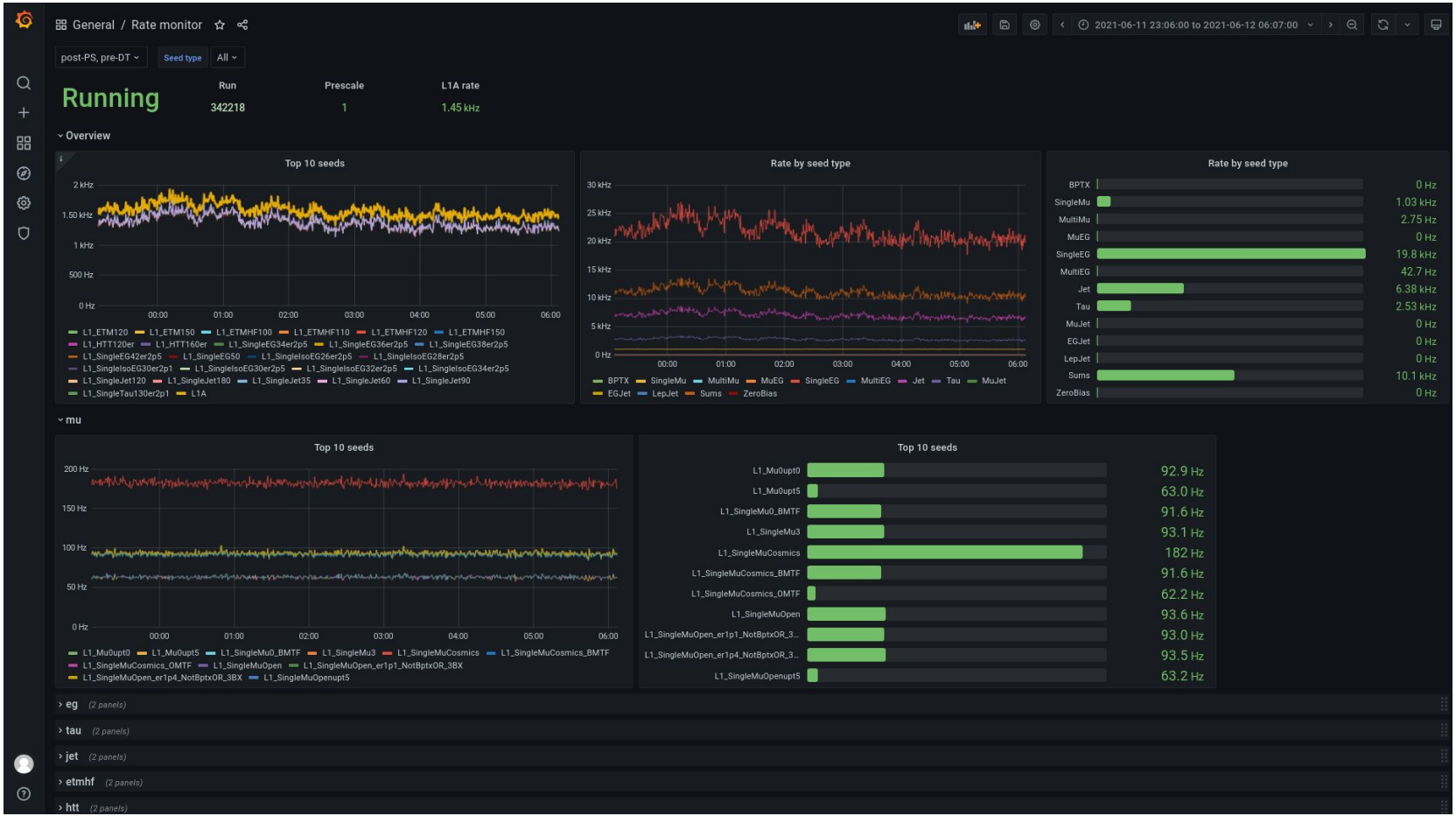
kubernetes OPENSHIFT docker

New monitoring dashboards

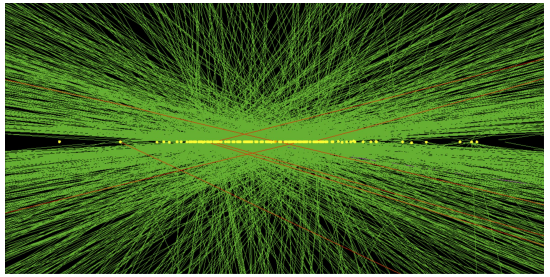
- Deployed new monitoring dashboards over the Summer



New dashboards — rate monitoring



Phase-2 upgrade



The challenge

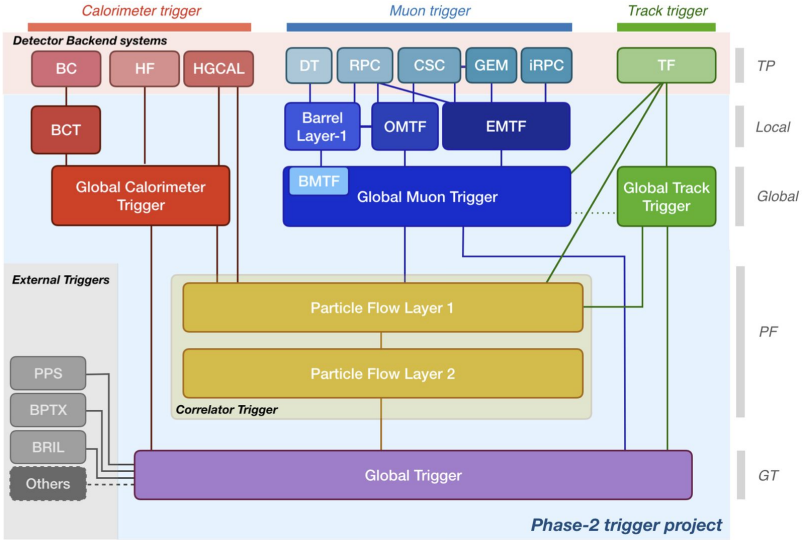
- 5-fold increase in collision rate, more pile up, but retain the same thresholds in the level-1 trigger

The solution

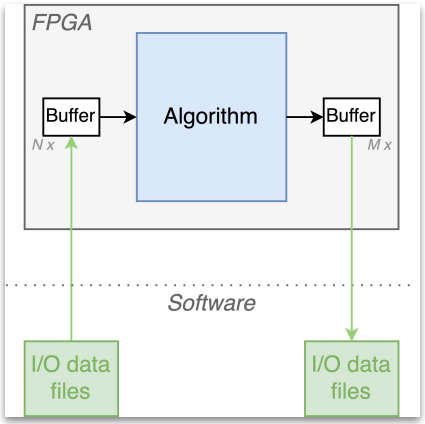
- Re-build system using latest tech.
  - 256 boards hosting UltraScale+ FPGAs
  - Connected by O(10k) links, up to 25Gbps

➢ Consume more data

  - Calo. & muons: Finer granularity
  - Particle trajectories from tracker



Phase-2 upgrade

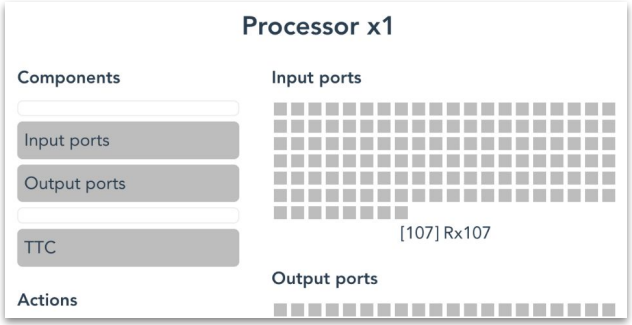


Focus over the next year: Implementing & validating all components of design from TDR

1. Test reco. & selection firmware in single boards
2. Test slices spanning multiple layers of system

⇒ Will need to control & monitor multiple boards — like in current system at P5

- Developing key components of phase-2 online software
  - Building on existing codebase, where applicable
  - Taking advantage of new tools & libraries
  - Informed by feedback from current system
    ⇒ *Design the best system possible*
- Use for slice tests in coming months



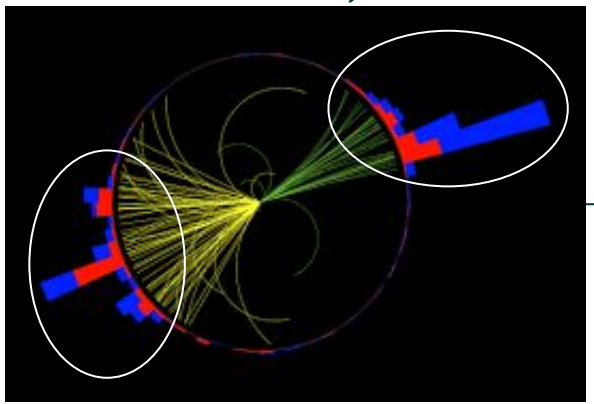
Wrap-up

- Level-1 trigger is a complex system and a fundamental component of CMS
- Configured, controlled and monitored through the L1T online software
  - A sophisticated set of libraries and tools — L1CE, SWATCH, L1Page & DQM
- Renovating online software with modern, industry-standard tools
  - On-going work, but already useful for on call personnel
- Phase-2 upgrade
  - More complex than current system (# boards, # links, # subsystems)
  - Next generation of online software under development
- Interested in contributing?
  - Email: simone.bologna@cern.ch and tom.williams@cern.ch

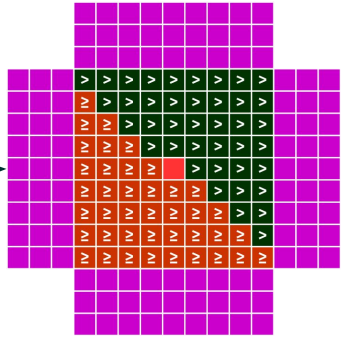
Backup

Example triggers

Hadron jets



L1T, fixed-area cone

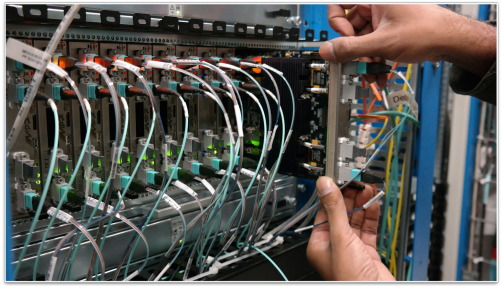


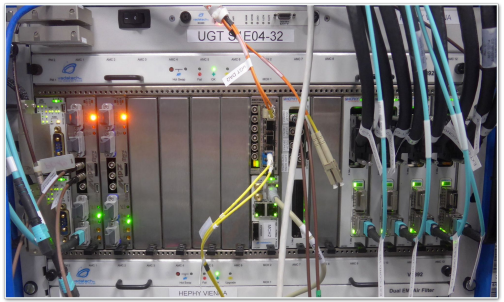
HLT, anti-kt 4

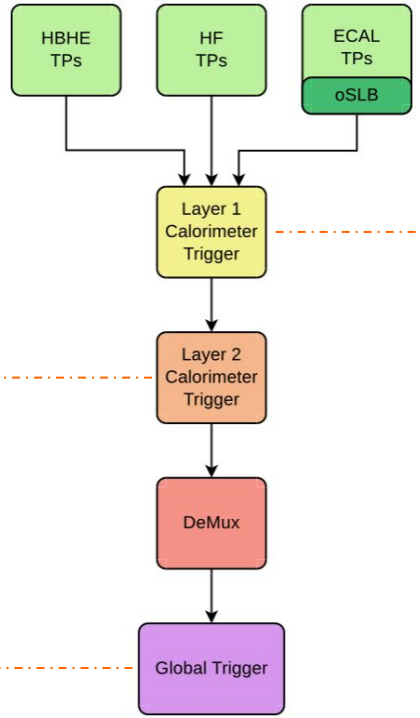
$$d_{ij} = \min(p_{\mathrm{T},i}^{-2}, p_{\mathrm{T},j}^{-2}) \frac{\Delta R_{ij}^2}{R^2} \qquad \Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$$

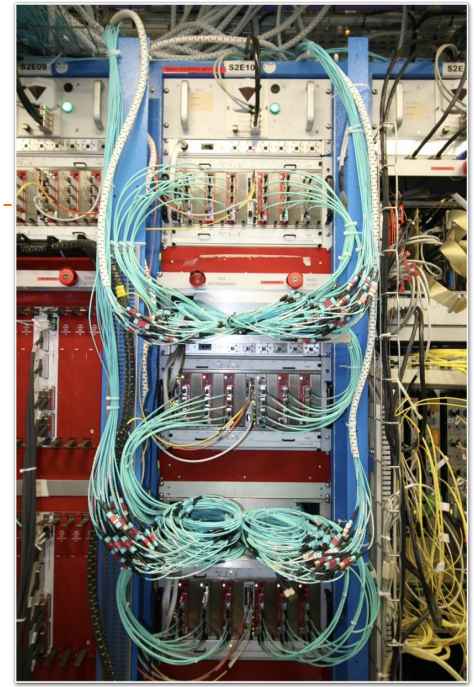
$$d_{iB} = p_{\mathrm{T},i}^{-2}$$

Level-1 trigger: Calorimeter chain







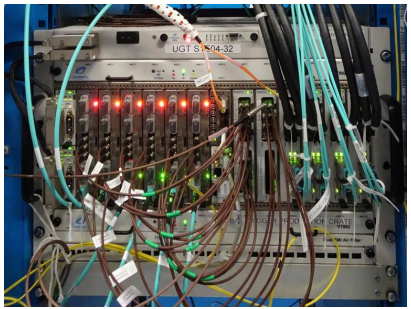


CMS trigger

- Level-1 trigger, HW-based
- High-level trigger, SW based
- Successive selections of finer granularity

L1T

HLT



Level-1 trigger

High-level trigger