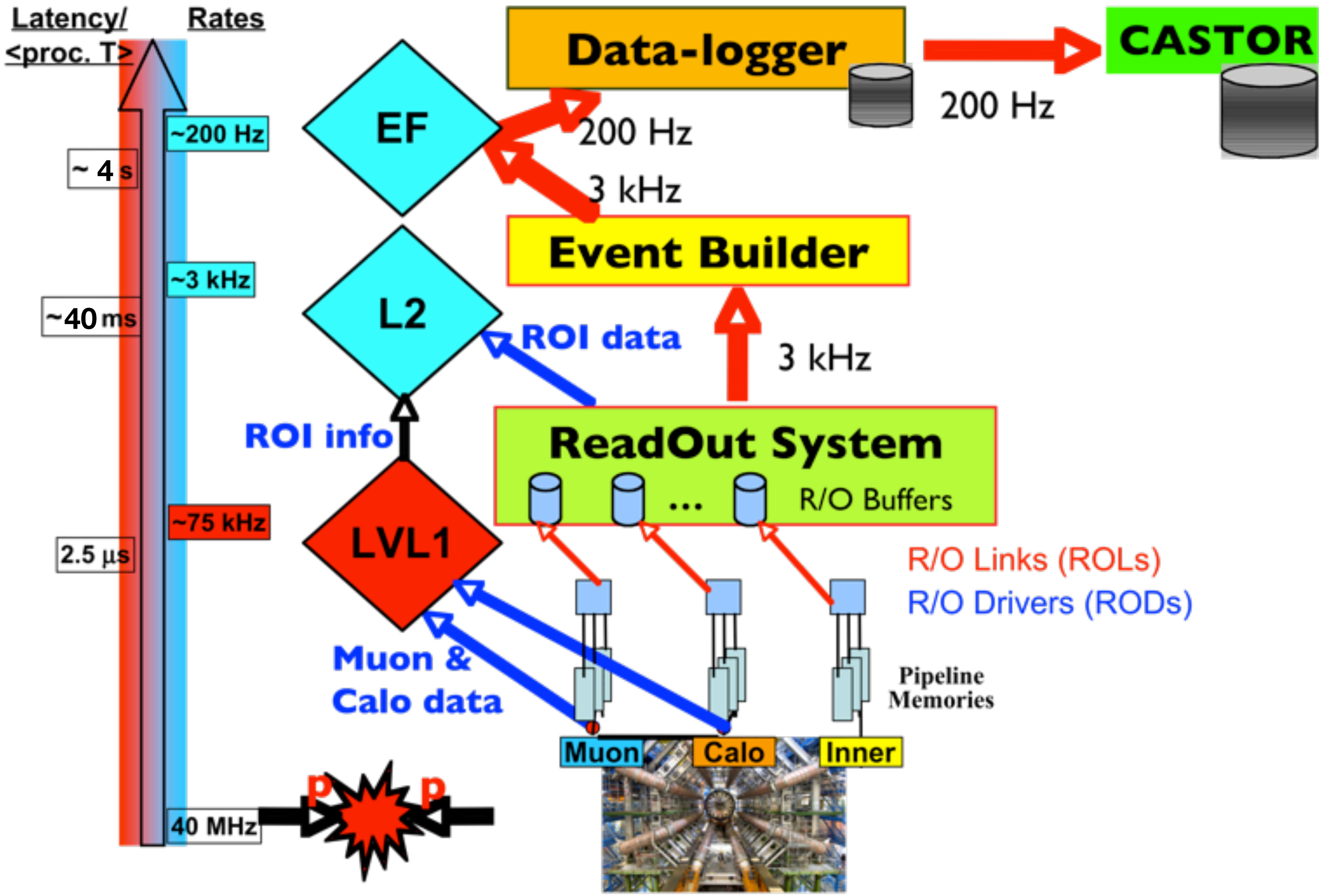
ATLAS online data quality monitoring

Cristóbal Cuenca Almenar, Yale University
for the online DQ team

Outline

- ATLAS Trigger and Data Flow
- Data Quality Monitoring Framework
  DQMF
- Data Quality Monitoring Display
  DQMD
- Data Quality Monitoring Configurator
  DQMC
- Conclusions

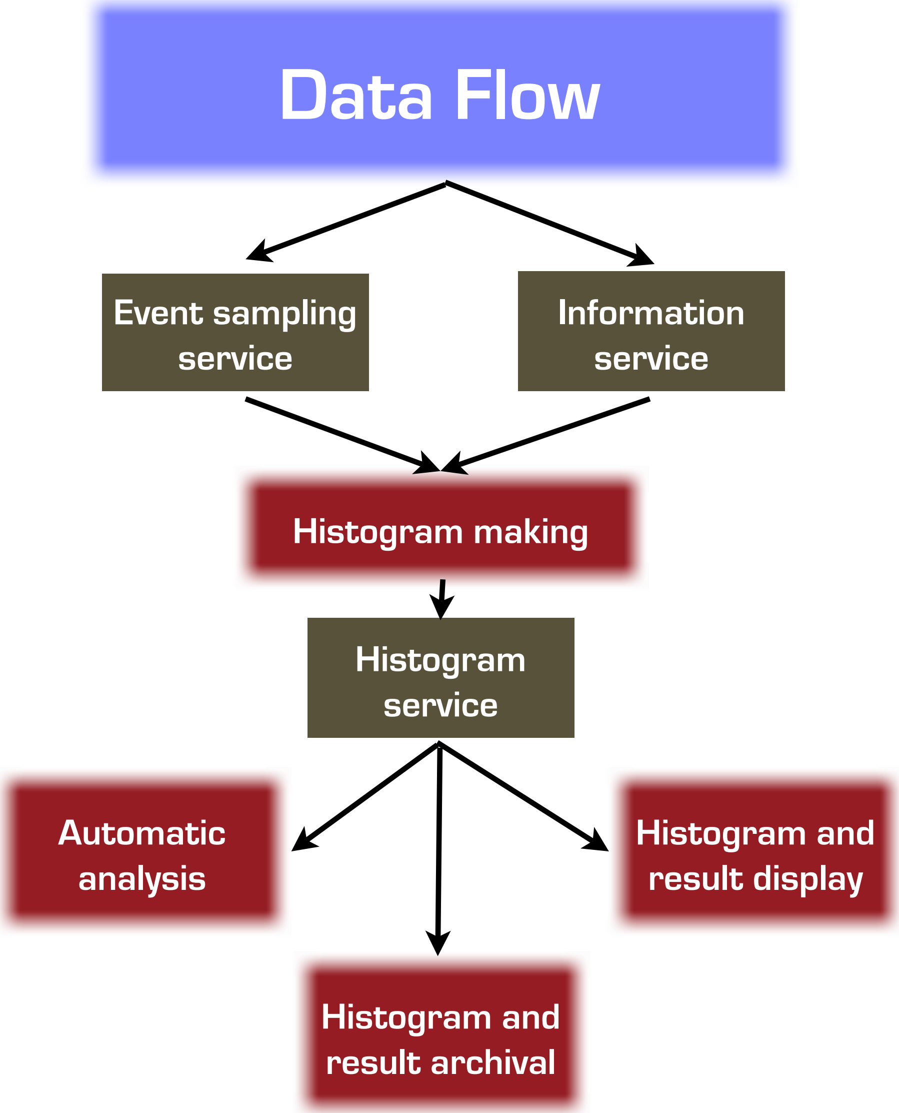
Trigger and Data Flow



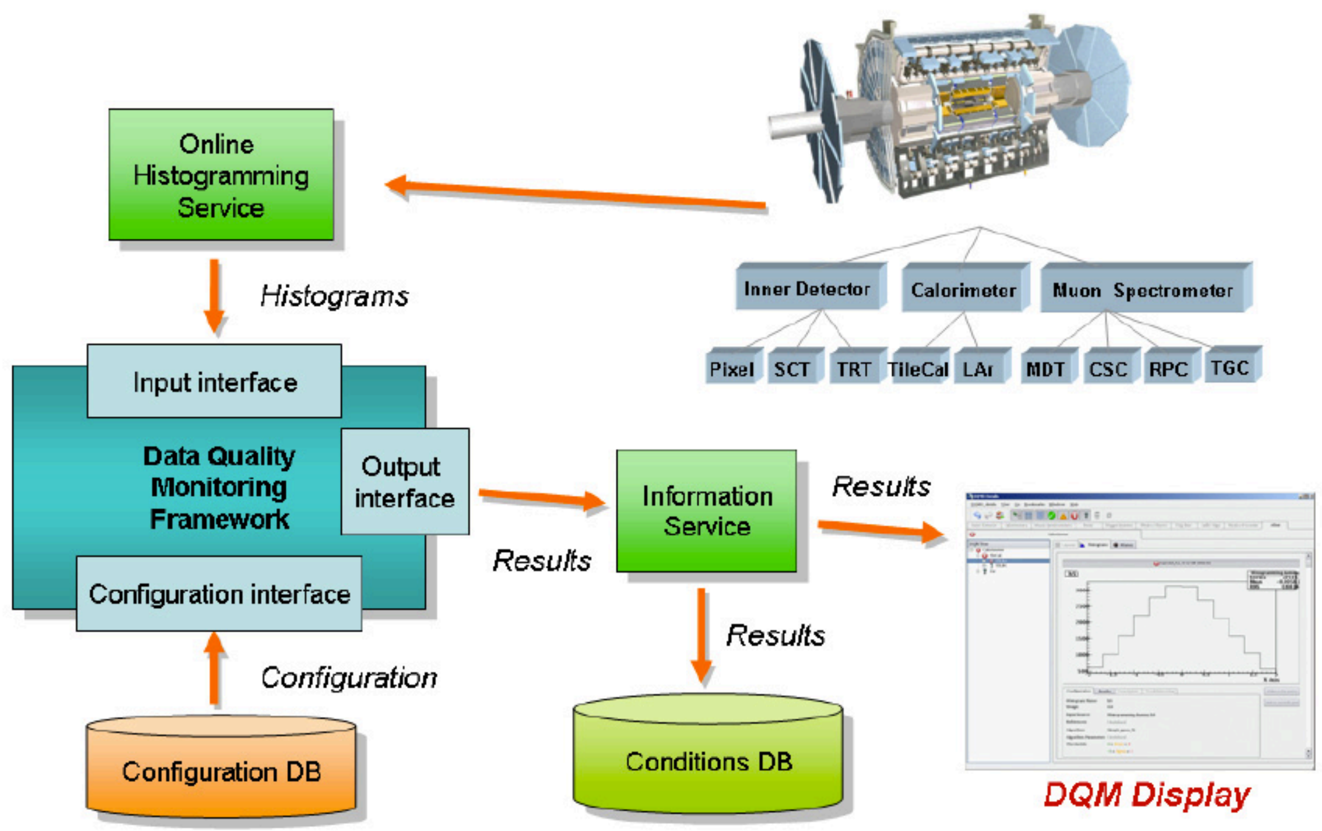
- ATLAS Data Acquisition system has a **complex architecture** to deal with high rates and large data throughput
- **Online data quality monitoring** samples data from all subsystems at all stages of Trigger/ Data Flow

Monitoring infrastructure

- Many monitoring applications run online: **diversity of monitoring needs**
  - analyze data and produce **histograms**
  - analyze **operational** conditions
  - **automatic** checks
  - automatic data **archiving**
  - **visualize** locally and remotely
- About 32 **dedicated** machines
- Access to data and operational conditions at all stages of Data Flow and trigger



DQMF: the framework



Online Histogramming Service

Histograms

Input interface

Data Quality Monitoring Framework

Output interface

Configuration interface

Configuration

Configuration DB

Results

Information Service

Results

Results

Conditions DB

Inner Detector

Calorimeter

Muon Spectrometer

Pixel

SCT

TRT

TileCal

LAr

MDT

CSC

RPC

TGC

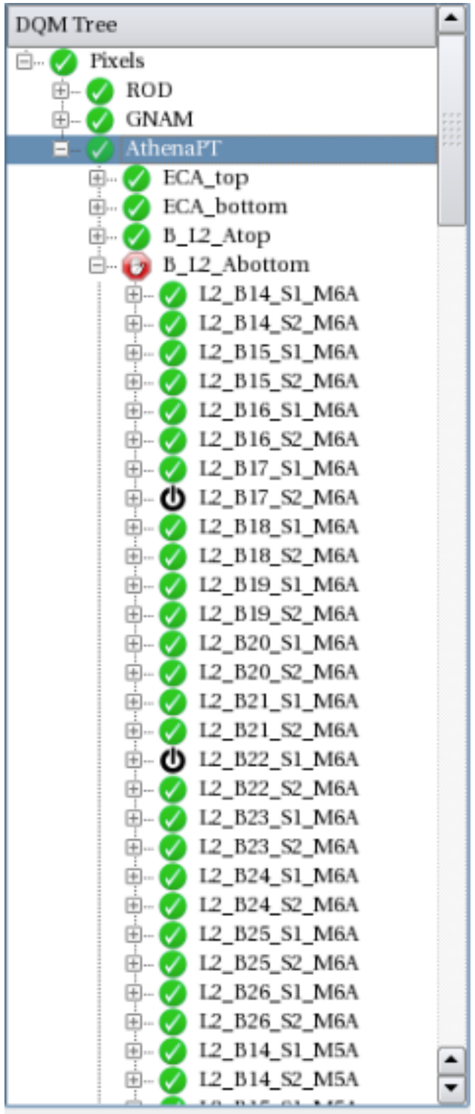
DQM Display

- DQMF components:
  - **DQRegions** and **DQParameters**: define tree
  - **DQAgents**: core application
  - **DQAlgorithms** run on input to produce **DQResults**

- **Distributed** and **scalable** framework to monitor data quality both online and offline
- The **configuration** determines what tests are performed on what histograms
- DQMF **reads** configuration, **inputs** histograms, **executes** tests, **produces** results and **writes** to the DB
- Tools are provided for easy and fast **visualization** of the results, locally and remotely

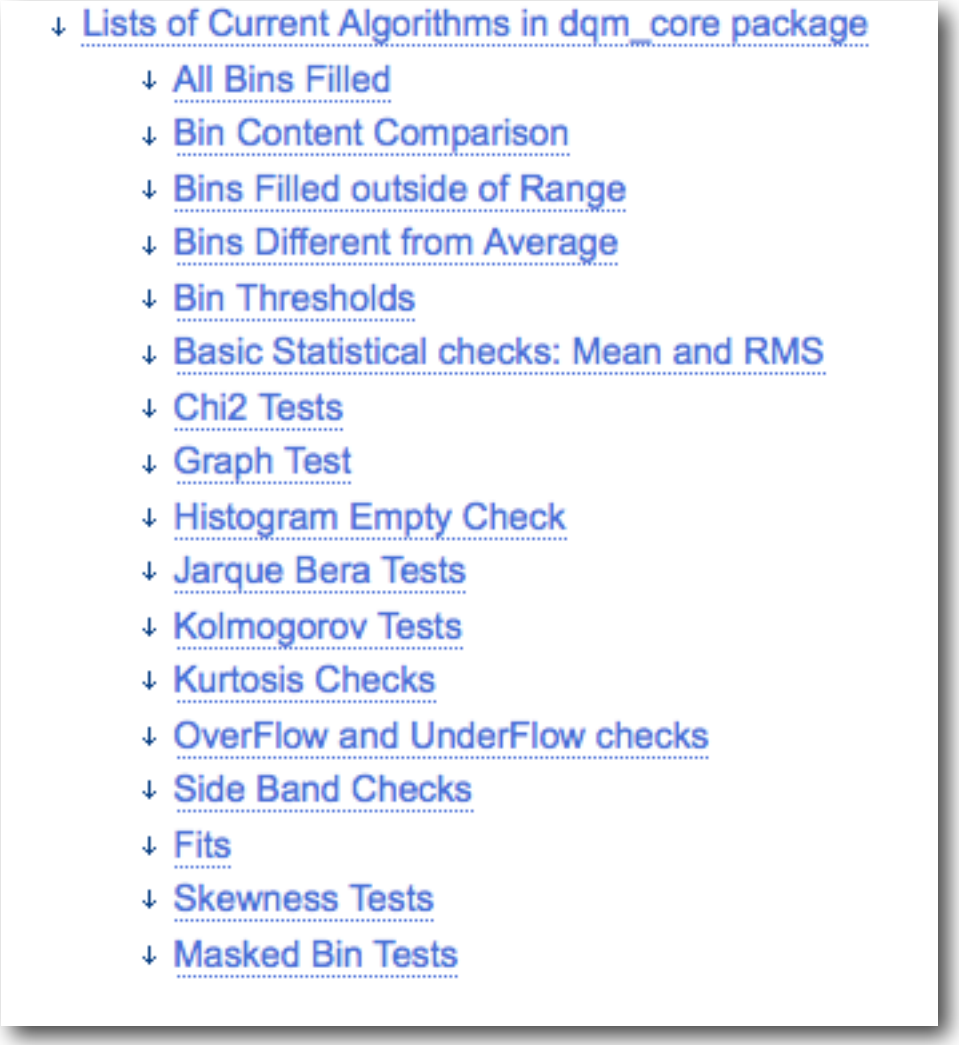
DQMF: Data Quality tree

- Single data quality tests are handled by **DQParameters**
- Each DQParameter specifies
  - what **input** histogram(s) to use
  - what **algorithm** to apply (DQAlgorithm)
  - the **thresholds** to define good or bad result (DQResult)
- DQParameters are grouped in **DQRegions**
- DQRegions also have **DQResults** associated
  - the mechanism to combine the results of the subparameters is specified in the configuration
- DQRegions can be grouped in mother DQRegions, thus creating a **DQ tree**



```
DQM Tree
⊟ ✓ Pixels
    ⊞ ✓ ROD
    ⊞ ✓ GNAM
    ⊟ ✓ AthenaPT
        ⊞ ✓ ECA_top
        ⊞ ✓ ECA_bottom
        ⊞ ✓ B_L2_Atop
        ⊟ ⛔ B_L2_Abottom
            ⊞ ✓ L2_B14_S1_M6A
            ⊞ ✓ L2_B14_S2_M6A
            ⊞ ✓ L2_B15_S1_M6A
            ⊞ ✓ L2_B15_S2_M6A
            ⊞ ✓ L2_B16_S1_M6A
            ⊞ ✓ L2_B16_S2_M6A
            ⊞ ✓ L2_B17_S1_M6A
            ⊞ ⏻ L2_B17_S2_M6A
            ⊞ ✓ L2_B18_S1_M6A
            ⊞ ✓ L2_B18_S2_M6A
            ⊞ ✓ L2_B19_S1_M6A
            ⊞ ✓ L2_B19_S2_M6A
            ⊞ ✓ L2_B20_S1_M6A
            ⊞ ✓ L2_B20_S2_M6A
            ⊞ ✓ L2_B21_S1_M6A
            ⊞ ✓ L2_B21_S2_M6A
            ⊞ ⏻ L2_B22_S1_M6A
            ⊞ ✓ L2_B22_S2_M6A
            ⊞ ✓ L2_B23_S1_M6A
            ⊞ ✓ L2_B23_S2_M6A
            ⊞ ✓ L2_B24_S1_M6A
            ⊞ ✓ L2_B24_S2_M6A
            ⊞ ✓ L2_B25_S1_M6A
            ⊞ ✓ L2_B25_S2_M6A
            ⊞ ✓ L2_B26_S1_M6A
            ⊞ ✓ L2_B26_S2_M6A
            ⊞ ✓ L2_B14_S1_M5A
            ⊞ ✓ L2_B14_S2_M5A
```

DQMF: Data Quality Algorithms



- ↓ Lists of Current Algorithms in dqm_core package
  - ↓ All Bins Filled
  - ↓ Bin Content Comparison
  - ↓ Bins Filled outside of Range
  - ↓ Bins Different from Average
  - ↓ Bin Thresholds
  - ↓ Basic Statistical checks: Mean and RMS
  - ↓ Chi2 Tests
  - ↓ Graph Test
  - ↓ Histogram Empty Check
  - ↓ Jarque Bera Tests
  - ↓ Kolmogorov Tests
  - ↓ Kurtosis Checks
  - ↓ OverFlow and UnderFlow checks
  - ↓ Side Band Checks
  - ↓ Fits
  - ↓ Skewness Tests
  - ↓ Masked Bin Tests

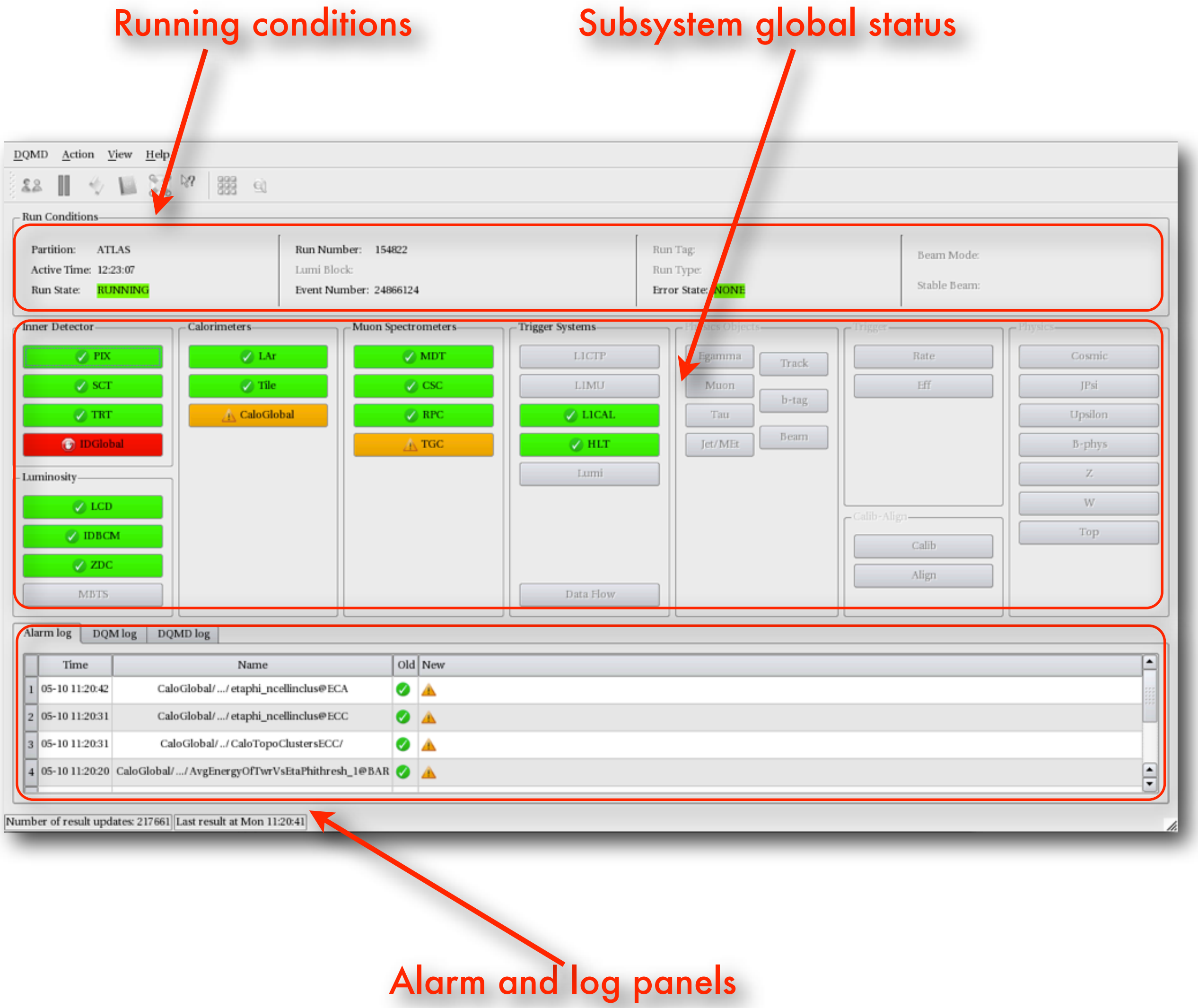
- A **dedicated library** has been built with the algorithms that can be added to the configuration of the DQParameters
- Many of these algorithms are **generic** enough
  - some parameters need to be specified, for example the bin number that is checked
- Also, some optional parameters allow extra **versatility**, for instance
  - require enough statistics
  - further checks, if failure

DQMF: DQ agents and DQ results

- **DQ agents** are the applications that run the checks online
- **Input** and **output** are implemented as **plug-ins**
- DQ agents can **read** from IS servers or root files with histograms
- Different **outputs** can be configured: IS servers, root files, Conditions data base
- The configuration specifies the **checks** and the **thresholds**
    - the agent runs a given algorithm and publishes a result
- **DQ results** consist of a **colored** tag and any output that the algorithms might want to attach:
    - If some areas of the detector are disabled, then the corresponding dq results will be **black**
    - otherwise, results might be
        - **green** (good), **yellow** (warning), **red** (bad) or **gray** (undefined)

DQMD: the data quality GUI

- Application for easy **visualization** of Data Quality status of each subsystem
- Main panel provides **overview**
- **Alarms** and **Log** tabs have been added for enhanced control
- **One button** per **subsystem**
- Clicking on each button brings up a **detailed panel** with the **subsystem** data quality tree, histograms, resuts
- Clicking on **Alarm** or **Log** entries brings up a detailed panel with further info



Running conditions

Subsystem global status

Alarm and log panels

Partition: ATLAS
Active Time: 12:23:07
Run State: RUNNING

Run Number: 154822
Lumi Block:
Event Number: 24866124

Run Tag:
Run Type:
Error State: NONE

Beam Mode:
Stable Beam:

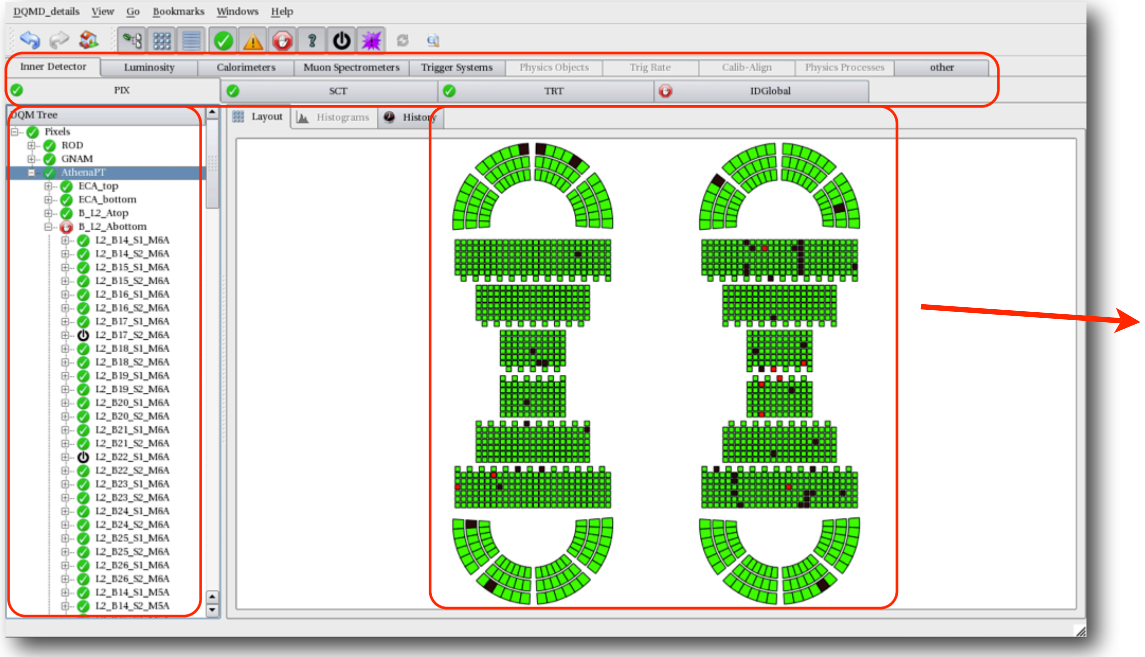
	Time	Name	Old	New
1	05-10 11:20:42	CaloGlobal/…/etaphi_ncellinclus@ECA	✓	⚠
2	05-10 11:20:31	CaloGlobal/…/etaphi_ncellinclus@ECC	✓	⚠
3	05-10 11:20:31	CaloGlobal/../CaloTopoClustersECC/	✓	⚠
4	05-10 11:20:20	CaloGlobal/…/AvgEnergyOfTwrVsEtaPhithresh_1@BAR	✓	⚠

Number of result updates: 217661 Last result at Mon 11:20:41

DQMD: layouts

Doubled tiered tabs link to each subsystem tree

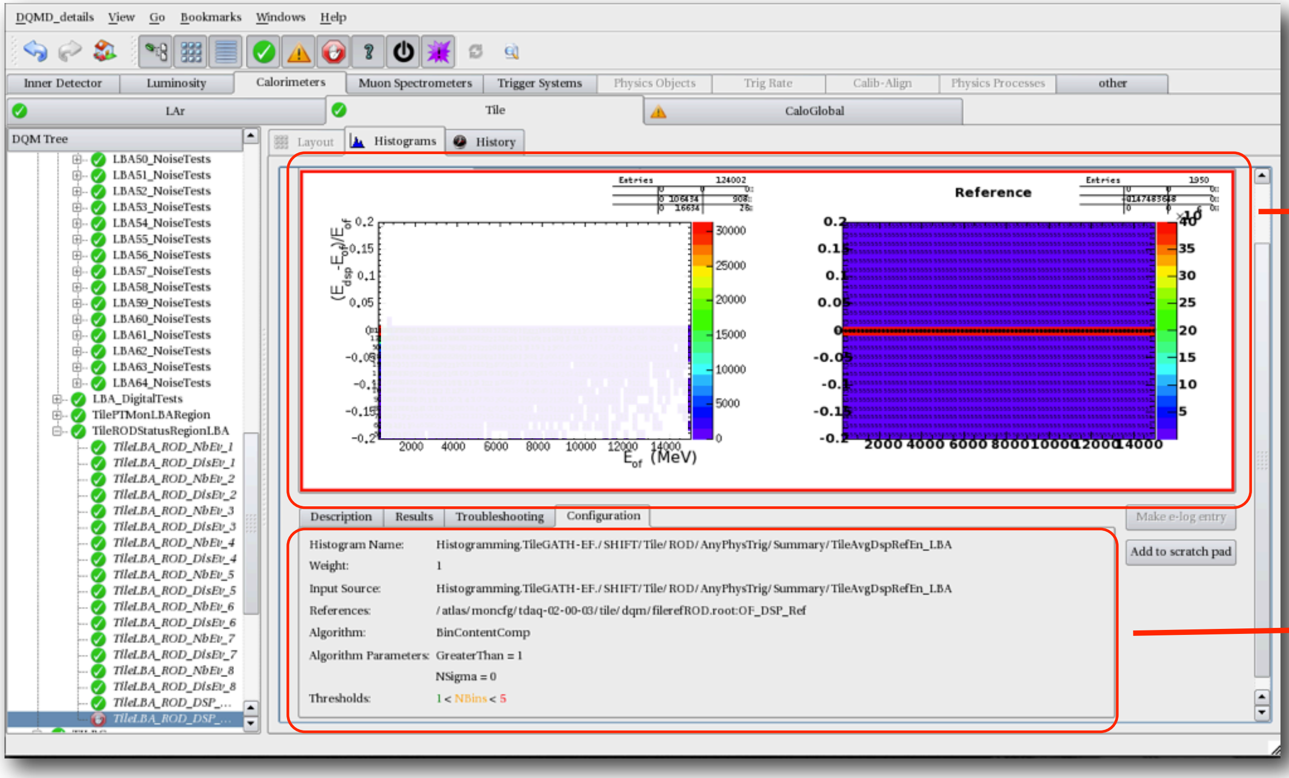
Data quality tree



Layout representing the data quality hierarchy

- Subsystems define a data quality hierarchy that can be browsed thru a **tree** or with a **layout**
- Data quality layouts allow for **easier understanding** of the status of the subsystems and **faster navigation** to problematic regions
- Layouts are defined and **configured** together with the structure and the tests
- Clicking on each part of the layout brings the appropriate **result**

DQMD: histograms and configuration



Histogram Name: Histogramming.TileGATH-EF./SHIFT/Tile/ROD/AnyPhysTrig/Summary/TileAvgDspRefEn_LBA
Weight: 1
Input Source: Histogramming.TileGATH-EF./SHIFT/Tile/ROD/AnyPhysTrig/Summary/TileAvgDspRefEn_LBA
References: /atlas/moncfg/tdaq-02-00-03/tile/dqm/filerefROD.root:OF_DSP_Ref
Algorithm: BinContentComp
Algorithm Parameters: GreaterThan = 1
NSigma = 0
Thresholds: 1 < NBins < 5

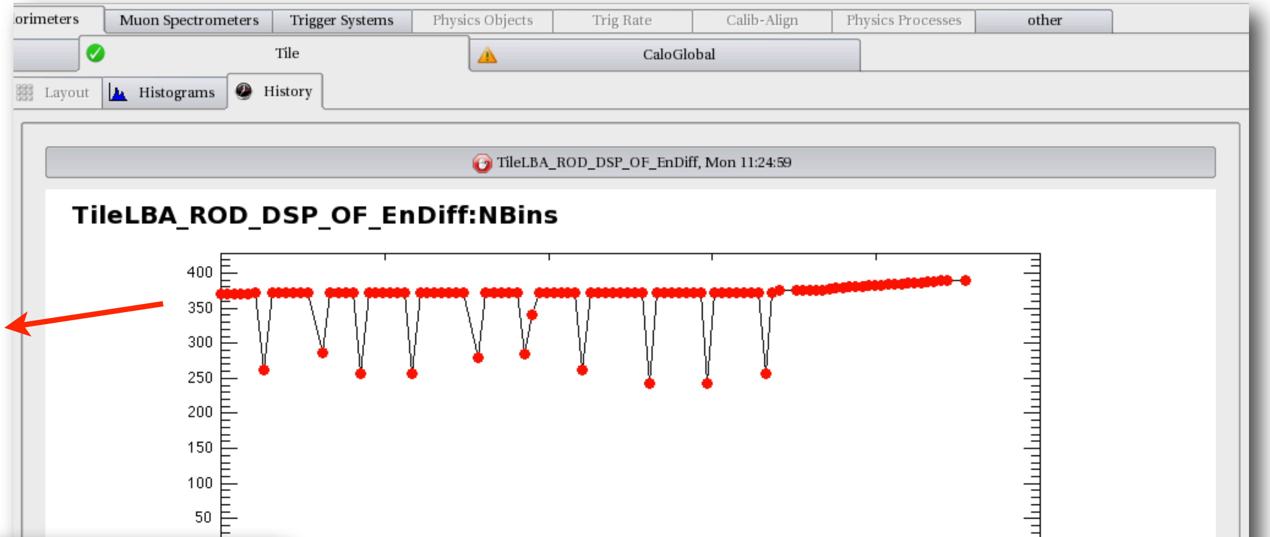
online histogram and reference, for comparison

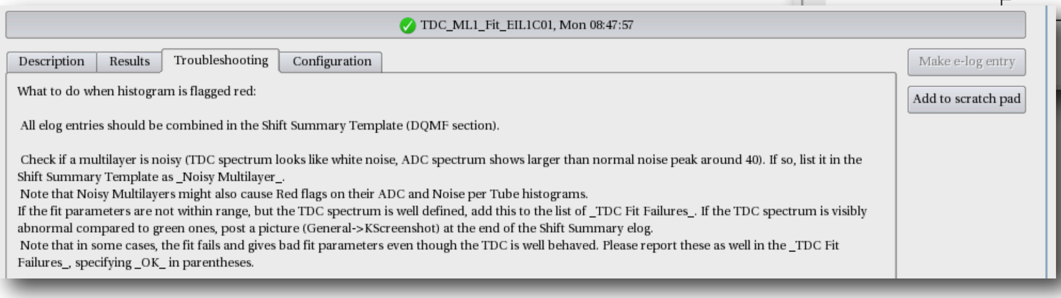
Configuration info: input algorithms, thresholds

- The **configuration** specifies the input location, the checks and thresholds to be applied
- A new **result** is produced every time a histogram is updated
- If results bring new histograms attached, these are also displayed

DQMD: History, description and troubleshooting

- **Time evolution** of the results is also displayed in an adjacent tab.
- One **graph** per parameter, **colored dots** according to result status

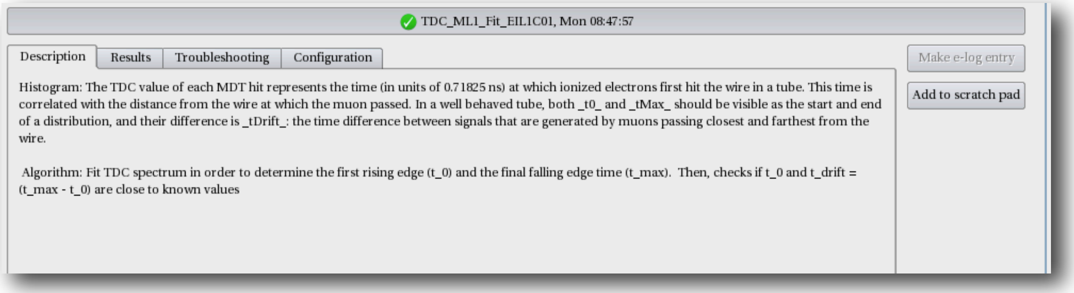




What to do when histogram is flagged red:

All elog entries should be combined in the Shift Summary Template (DQMF section).

Check if a multilayer is noisy (TDC spectrum looks like white noise, ADC spectrum shows larger than normal noise peak around 40). If so, list it in the Shift Summary Template as _Noisy Multilayer_.
Note that Noisy Multilayers might also cause Red flags on their ADC and Noise per Tube histograms.
If the fit parameters are not within range, but the TDC spectrum is well defined, add this to the list of _TDC Fit Failures_. If the TDC spectrum is visibly abnormal compared to green ones, post a picture (General->KScreenshot) at the end of the Shift Summary elog.
Note that in some cases, the fit fails and gives bad fit parameters even though the TDC is well behaved. Please report these as well in the _TDC Fit Failures_, specifying _OK_ in parentheses.



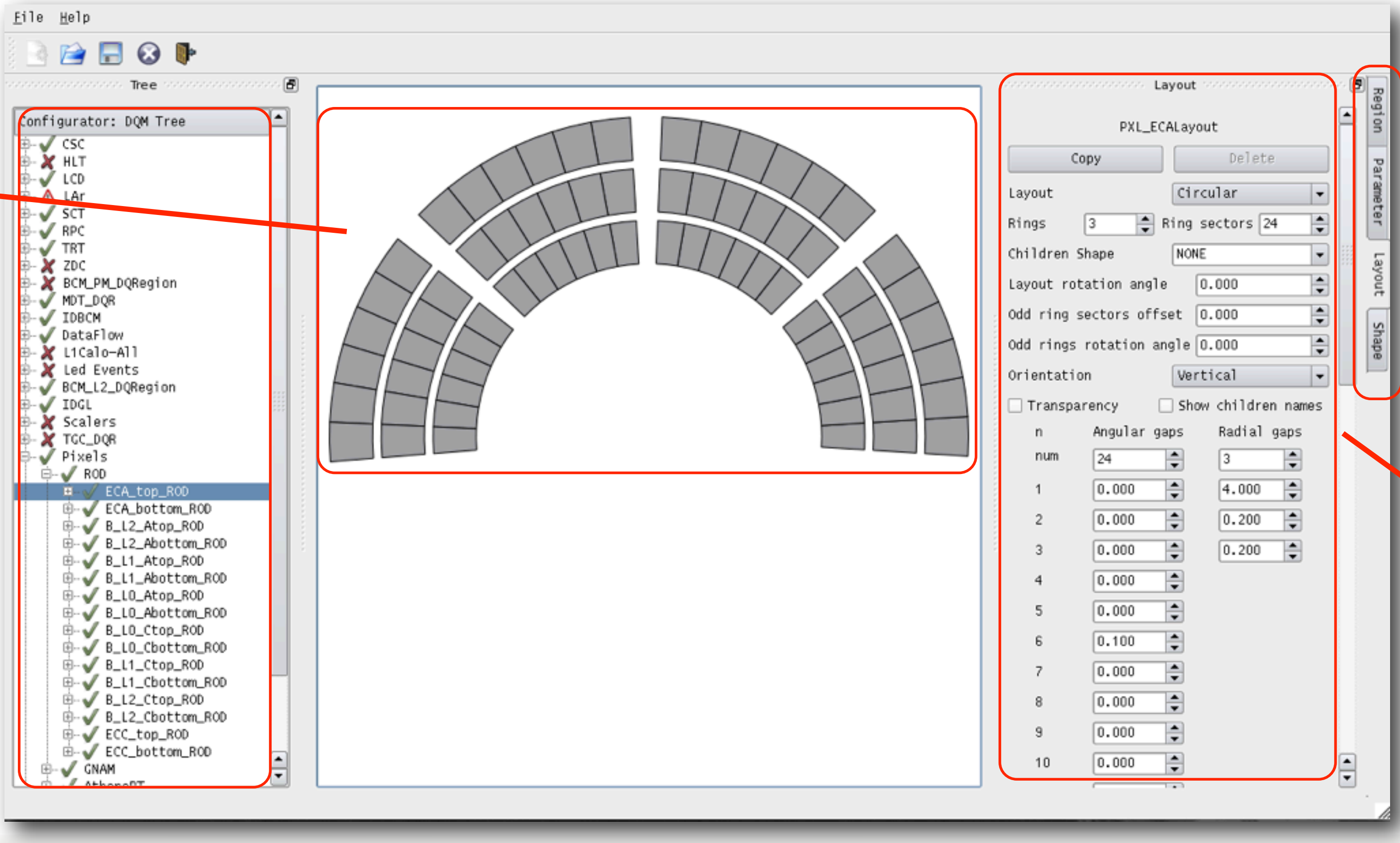
Histogram: The TDC value of each MDT hit represents the time (in units of 0.71825 ns) at which ionized electrons first hit the wire in a tube. This time is correlated with the distance from the wire at which the muon passed. In a well behaved tube, both _t0_ and _tMax_ should be visible as the start and end of a distribution, and their difference is _tDrift_: the time difference between signals that are generated by muons passing closest and farthest from the wire.

Algorithm: Fit TDC spectrum in order to determine the first rising edge (t_0) and the final falling edge time (t_max). Then, checks if t_0 and t_drift = (t_max - t_0) are close to known values

- Extra information is linked to each histogram:
  - a detailed **description** of the histogram and the tests
  - instructions on **actions** to be taken if problems appear

DQMC: data quality monitoring configurator

- Easy-to-use **graphical interface** to generate the layouts linked to the data quality tree
- Reads from and writes to standard **configuration** files
- Any parameter of the layout shown in the central panel can be modified



Configurator: DQM Tree

- ✓ CSC
- ✗ HLT
- ✓ LCD
- ⚠ LAr
- ✓ SCT
- ✓ RPC
- ✓ TRT
- ✗ ZDC
- ✗ BCM_PM_DQRegion
- ✓ MDT_DQR
- ✓ IDBCM
- ✓ DataFlow
- ✗ L1Calo-All
- ✗ Led Events
- ✓ BCM_L2_DQRegion
- ✓ IDGL
- ✗ Scalers
- ✗ TGC_DQR
- ✓ Pixels
  - ✓ ROD
    - ✓ ECA_top_ROD
    - ✓ ECA_bottom_ROD
    - ✓ B_L2_Atop_ROD
    - ✓ B_L2_Abottom_ROD
    - ✓ B_L1_Atop_ROD
    - ✓ B_L1_Abottom_ROD
    - ✓ B_L0_Atop_ROD
    - ✓ B_L0_Abottom_ROD
    - ✓ B_L0_Ctop_ROD
    - ✓ B_L0_Cbottom_ROD
    - ✓ B_L1_Ctop_ROD
    - ✓ B_L1_Cbottom_ROD
    - ✓ B_L2_Ctop_ROD
    - ✓ B_L2_Cbottom_ROD
    - ✓ ECC_top_ROD
    - ✓ ECC_bottom_ROD
  - ✓ GNAM

PXL_ECALayout

Layout: Circular

Rings: 3 Ring sectors: 24

Children Shape: NONE

Layout rotation angle: 0.000

Odd ring sectors offset: 0.000

Odd rings rotation angle: 0.000

Orientation: Vertical

☐ Transparency ☐ Show children names

n	Angular gaps	Radial gaps
num	24	3
1	0.000	4.000
2	0.000	0.200
3	0.000	0.200
4	0.000	
5	0.000	
6	0.100	
7	0.000	
8	0.000	
9	0.000	
10	0.000	

Layout linked to the highlighted region in the tree

Tree as displayed on DQMD

Tabs to access the different features of the layout

Layout parameters

Operational experience

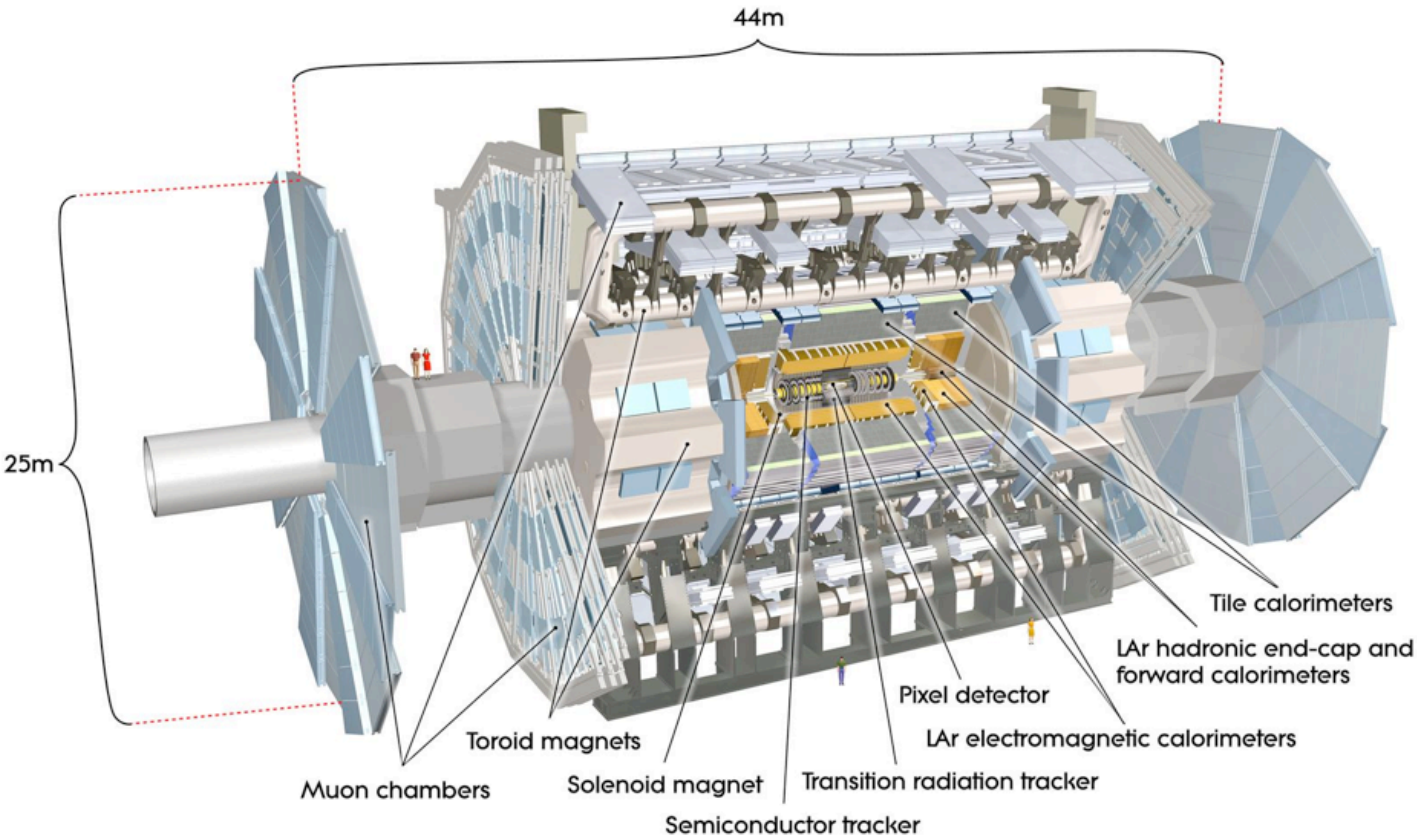
- So far, the framework handles
  - **20** DQAgents
  - more than **75000** DQParameters
  - more than **15000** DQRegions
  - more than **150000** new DQResults per minute
- This is only for the Data Quality **framework**
  - event sampling, information extraction and histogram generation and publishing use other processes and resources
- System experts and shifters feedback has resulted on **many upgrades,** specially in visualization tools
- **DQMC** used by most systems to generate layouts
- **DQMD** always used in most desks in the ATLAS control room to ensure good data quality taking and chase down any issues that might arise

Conclusions

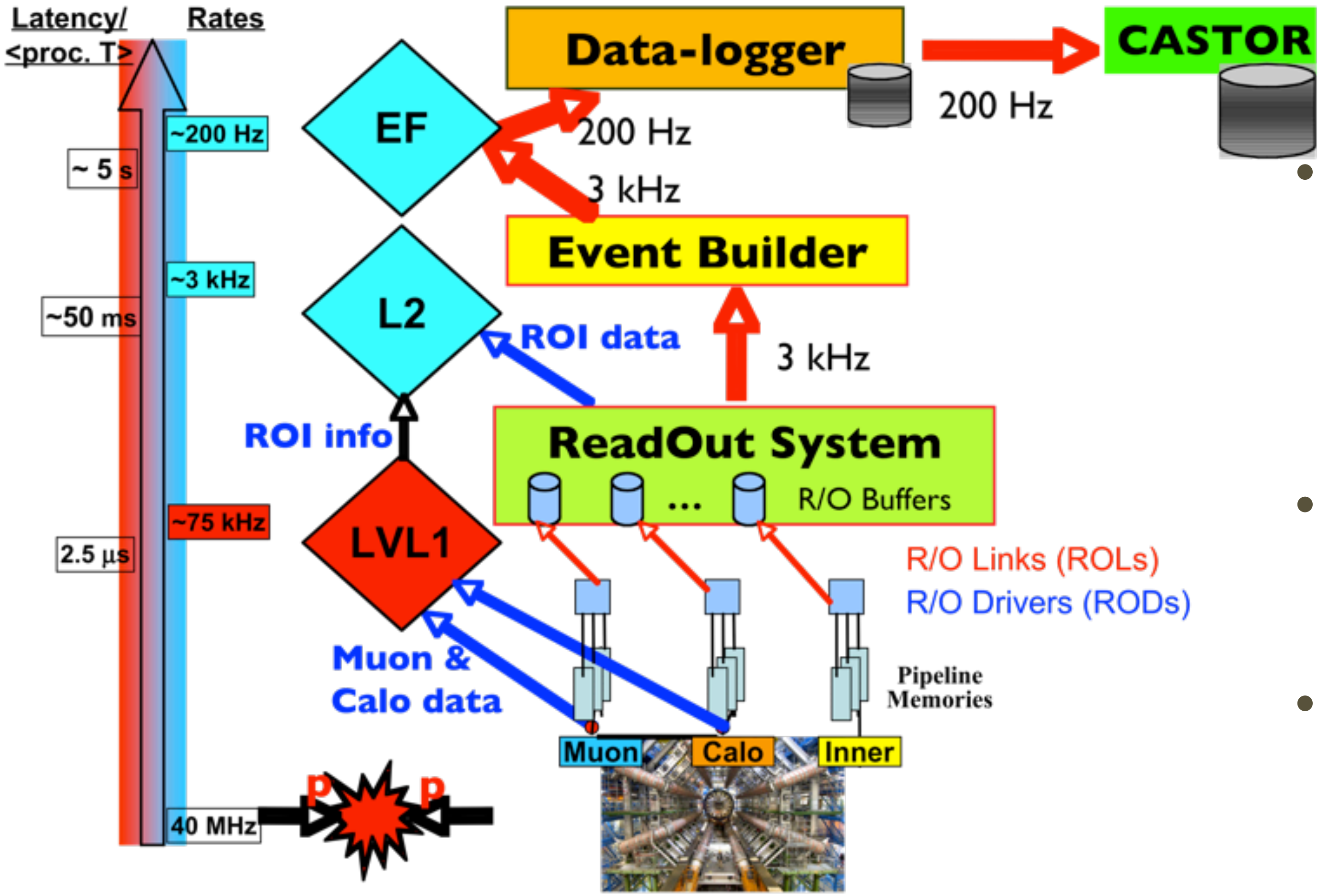
- **Several talks** and posters about **ATLAS** Data Acquisition, Trigger and Monitoring systems. See:
  - Martin zur Nedden on Trigger Monitoring
  - Claudia Borer on TDAQ overview
- Data Quality Monitoring Framework has been successfully **commissioned**, proving to be able to meet the stringent **ATLAS requirements**
- The framework together with the applications provided have proven very useful to **ensure good data quality**
  - DQMD has become one of the main tools used in the ATLAS control room
- This same framework is reused offline
  - based on those flags, **good run lists** generated for first physics results
- DQMF is **actively being used** to assure good data taking with collisions runs at 7 TeV

backup

The ATLAS detector


44m
25m
Tile calorimeters
LAr hadronic end-cap and forward calorimeters
Pixel detector
LAr electromagnetic calorimeters
Toroid magnets
Muon chambers
Solenoid magnet
Transition radiation tracker
Semiconductor tracker

Trigger and Data Flow



- **Three-level trigger architecture** to achieve a final rate of 200Hz, from the 40MHz collision rate
- **Data Flow** responsible for collecting data fragments, serve them to trigger processors and send them to mass storage
- Full event data only available at Event Filter level
- **Online data quality monitoring** samples data from all subsystems at all stages of DF