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

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

- 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

---

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
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 database
- 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, results
- Clicking on **Alarm** or **Log** entries brings up a detailed panel with further info
DQMD: layouts

- 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

- 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

Online histogram and reference, for comparison

Configuration info: input algorithms, thresholds
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.

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
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

- Muon chambers
- Toroid magnets
- Solenoid magnet
- Semiconductor tracker
- Transition radiation tracker
- Pixel detector
- LAr electromagnetic calorimeters
- LAr hadronic end-cap and forward calorimeters
- Tile calorimeters
ATLAS online data quality monitoring

Cristóbal Cuenca Almenar

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