

### **DQM Database Needs**

#### Data Quality Monitoring Alert Management Data request/ Local DQM Processing Raw data Data analysis TP Broadcast Data request/ Processing service **Trigger primitives** Storage Data TR display Data request/ Processing ۲ Trigger records Data Filter Trigger Trigger decisions / Trigger Trigger completed primitives Data request / response Data files Data request / Offline Raw data **Data collection** Detector Readout response electronics Trigger primitives and calib/dbg data stream

Imperial College London

### Imperial College London





Native format data

e.g. raw data,





Native format data

e.g. raw data,





# What Data To Store? London

### **1.** Monitoring plots

Plots summarising the sampled data. Fully interrogatable; accompanied by metadata describing e.g. time, detector unit.

### **2. Monitoring results**

Derived from monitoring plots. Can take a range of forms, from a boolean flag (e.g. an alert) to another plot. Flexible category.



Monitoring plots produced in last cold box run.

## How Much?

- In our most recent cold-box runs, our heat map plots were ~1 MB each. We generated 2 per plane (⇒ 6 per APA) every 10s.
- Scaling that up to a full FD module of ~150 APAs, a nominal data rate might be 90 MB/s.
- Not all this data necessarily needs to be kept long-term. Some sensible archive downsampling scheme is intended (but not yet implemented/tested), informed by storage budget.

#### • This is all still very fluid.

- Our current sampling rates are arbitrary (and likely to fall).
- We have many more systems to hook in whose monitoring processes have to run in parallel inside the same CPU budget.
- Our plots (and other metrics) are being defined and redefined as we go along.

**Imperial College** 

London

# How Will It Be Used? London

#### **1. Live Inspection**

Users of the web display need to be able to navigate through all the subsystems and units of the detector to view monitoring plots and results as they're generated in real time.

#### 2. Historical record

Users also need to be able to go back and use the monitoring plots/results to assess the detector state at a given time in the run history, just as they would live (though maybe not with the same time resolution).

A lot of information to represent, and a lot of potential questions users may be trying to answer, so important that users are able to access and display the data in a configurable way (e.g. show me variables X, Y and Z from systems A, B and C).

## Access & Support

- The vision for the DQM display is to be entirely webbased, so that shifters can easily view and query it from anywhere in the world with minimal prerequisite setup.
- Access would be gated through CERN/FNAL SSO.
- A smaller group of DQM experts would be responsible for DB management.
- The DQM will be critical for good data collection, so 24/7 support seems a requirement during runtime both on the software and the hosting infrastructure. The long-term provision of that support is a conversation we're yet to have.

### What We Already Have

#### Analysis AnalysisSource Parameter PK UUID Id PK UUID Id char 50 Name char 50 TextName char 250 Description Factor float Data parameters Database Analyse Pannel AnalysisParameter PK UUID Id PK UUID Id PK UUID Id FK UUID IdData char 50 Name FK1 UUID IdParameter FK UUID IdAnalysisSource char 250 Description FK2 UUID IdAnalyse Running bool DataDisplayAnalyse float Degree string Name AnalysisPannel int Interval PK UUID Id string Description PK UUID Id char 50 Type FK UUID IdSource FK UUID IdPannel FK UUID IdAnalyse FK UUID IdDataDisplay AnalysisResult PK UUID Id FK1 UUID IdDataPath DataDisplay DataType FK2 UUID IdAnalysisParameter FK1 UUID IdDataType DataDisplayData PK UUID Id FK2 UUID IdData char 50 Decision PK UUID Id Name char 50 float Confidence FK3 UUID IdSamplingProfile FK UUID IdSource PlottingType char 50 FK4 UUID IdAnalyse FK UUID IdDataDisplay char 200 Description int PlotLength DataPath Data PK UUID Id PK UUID Id FK UUID IdData FK UUID IdSource char 256 Path SamplingProfile RententionTime WriteTime int timestamp PK UUID Id string name string 30 DataStorage char 50 Name char 200 Description char 50 Technique DataSource float Factor PK UUID Id char 50 Source char 200 Description Web Display Data

#### Imperial College London

- This DB schema was drawn up by Yann Donon to serve the needs of the remote platform.
- Currently an instance hosted at CERN.

#### • Subject to change!

- ASP.NET infrastructure proving v. difficult for us to maintain.
- Looking at rebuilding this end of the system with django.

## Summary

Imperial College London

- The DQM DB needs to store monitoring plots and results covering the full run history of the experiment.
  - No benchmarking on data volumes here yet, but expect <<< volumes from main dataflow.</li>
- The DB needs to allow users to interrogate its contents through the web display along many dimensions (data type, detector unit, channel, time, etc.).
- We have a design for such a DB, but we're currently in the process of rebuilding it from the bottom up!