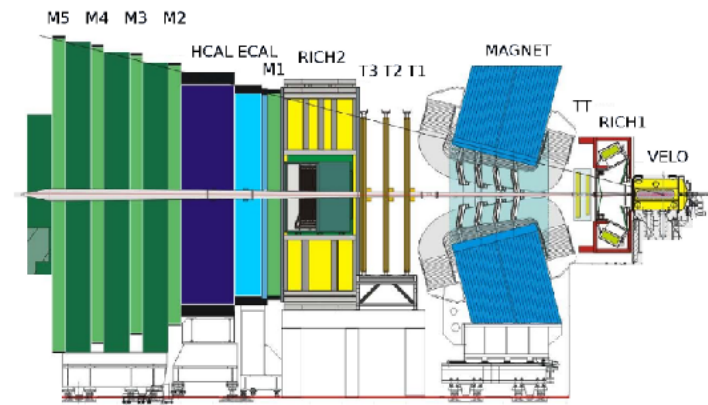




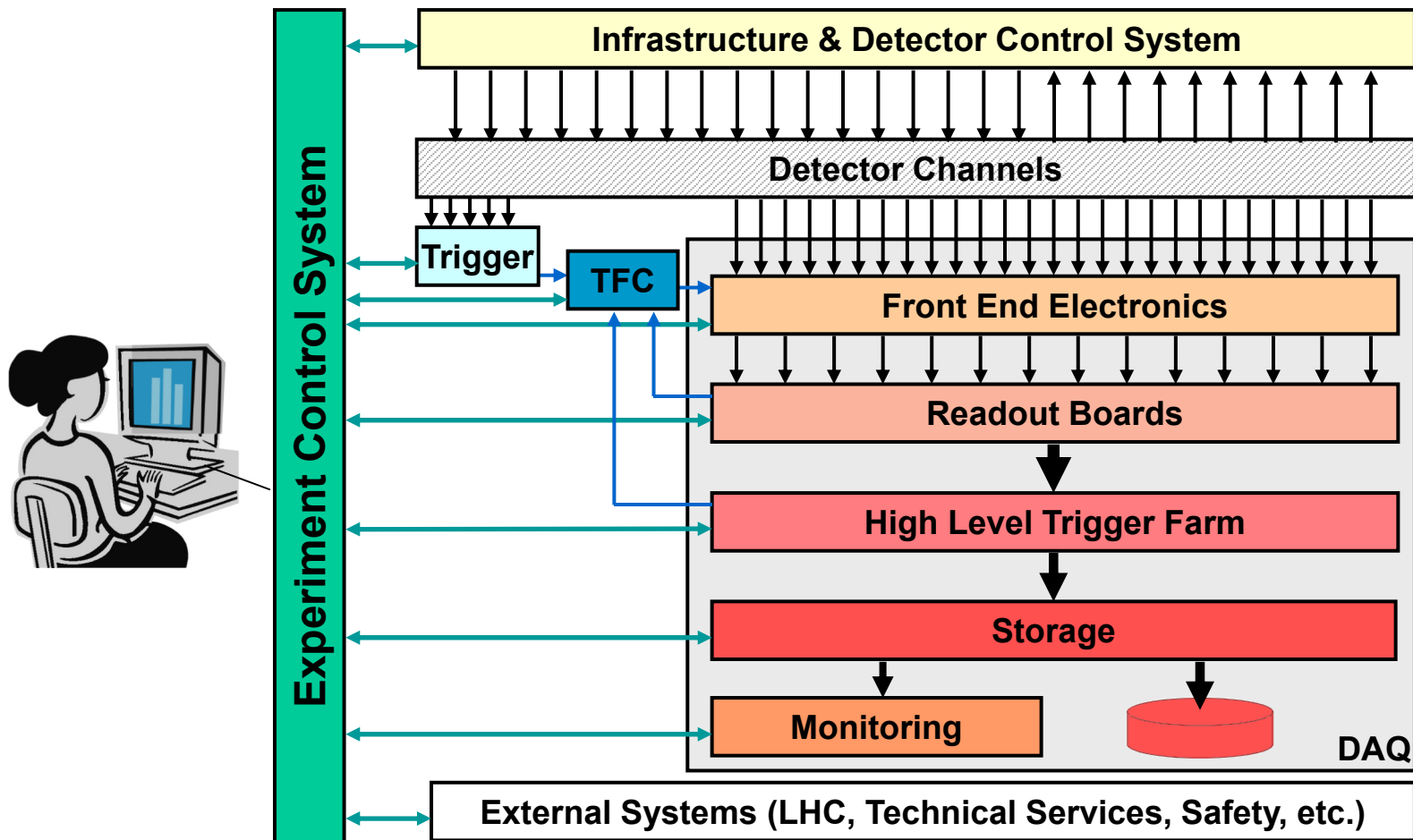
Configuring and Automating an LHC Experiment

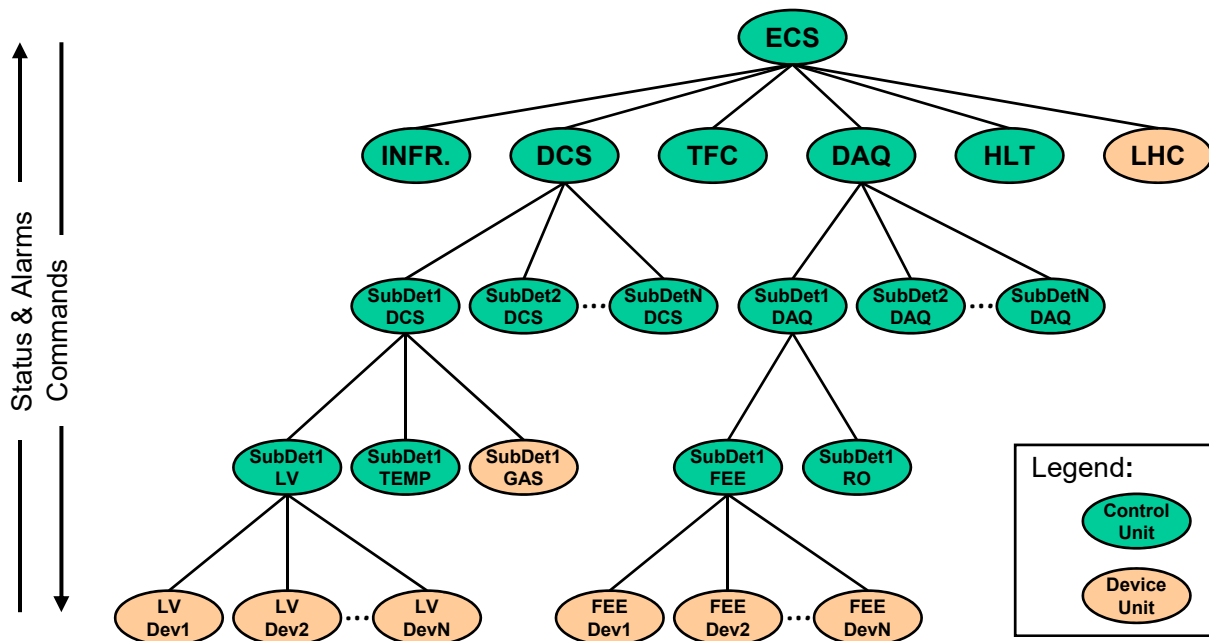
For faster and better Physics Output

Clara Gaspar, October 2017

- **LHCb's New Online Dataflow in Run II**
 - Introduced Prompt Calibration and Alignment
 - Aims:
 - | Better Trigger Efficiency
 - | Full offline-quality data directly out of the Online System
-> straight to Physics Analyses
 - | By Making better use of online farm resources
 - | Take advantage of LHC's duty cycle and shutdown periods
- **Represented a challenge for the control system**

■ Scope:





■ Implementation: (JCOP project)

- WinCC-OA
- SMI++

■ Deployment:

- Runs distributed over ~160 PCs (Virtual Machines)

■ Control Units are logical entities:

- Behave as a Finite State Machine / Rule Based system:
 - | Capable of Partitioning: Exclude/Include children
 - | Can take local decisions: Sequence & Automate Operations or Recover Errors

■ Device Units

- Provide the interface to the device (hardware or software)

- **Selects interesting events for Physics**
- **Runs distributed on the HLT Farm**

- **Farm Hardware**

- ~1600 nodes, ~ 50000 cores, heterogeneous
- Organized in 62 sub-farms, 24 to 32 nodes each

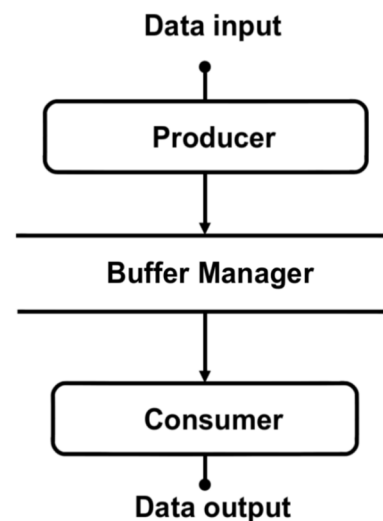
- **Farm Software**

- **Dataflow Pattern**

- Buffer Manager Concept

- **Dataflow Tasks**

- Based on Gaudi Online
- Integrated in Control System like any other Device via FSM states&actions

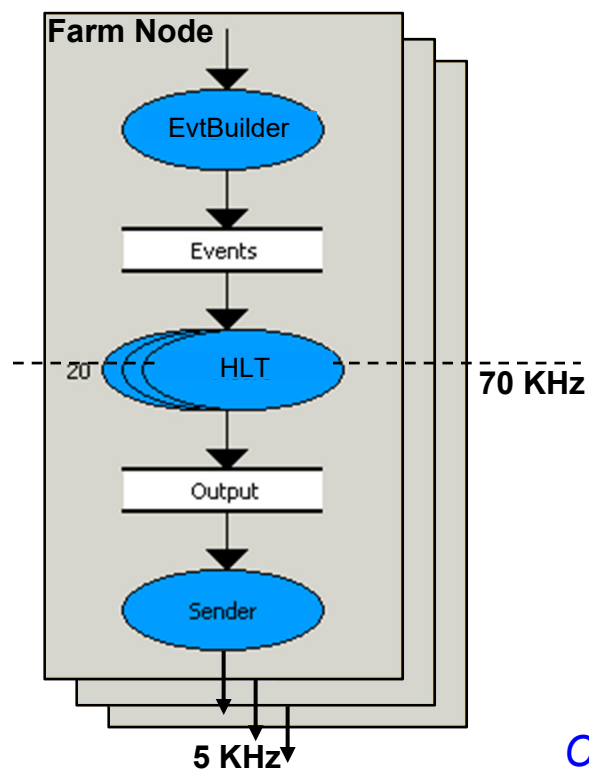


■ Resource Optimization in 2012 -> Deferred HLT

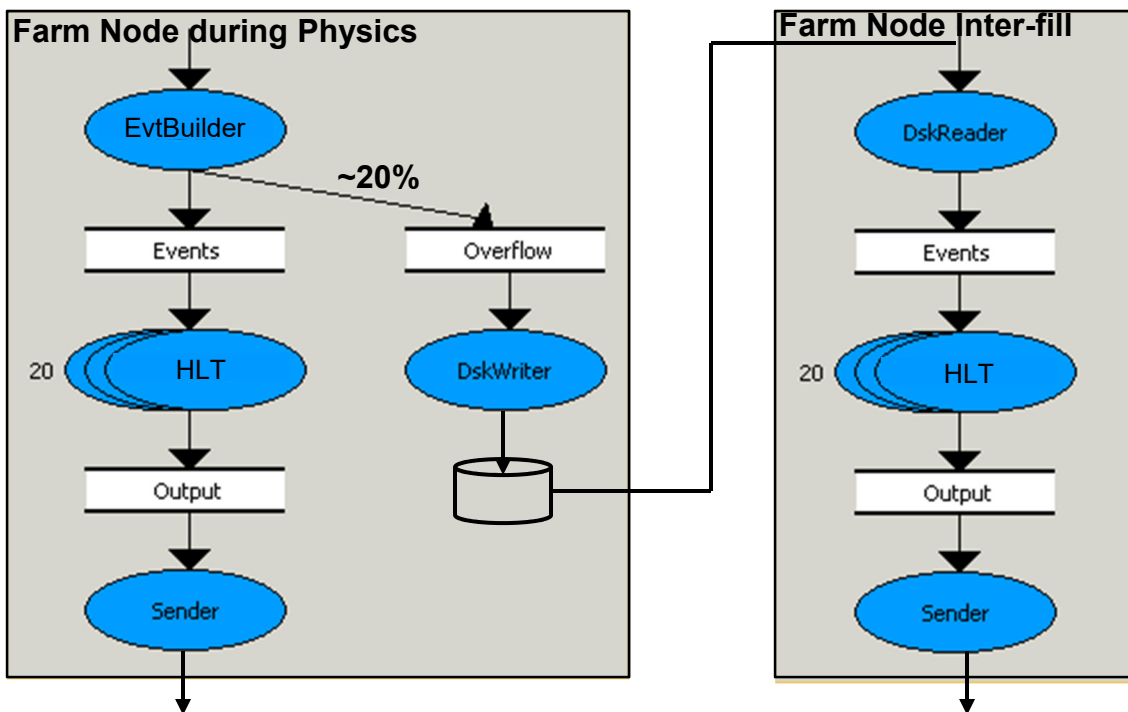
- Idea: Buffer data to disk when HLT busy / Process in inter-fill gap
- ➡ Reconfigure the whole farm at start/end of Physics

Standard HLT

1 MHz

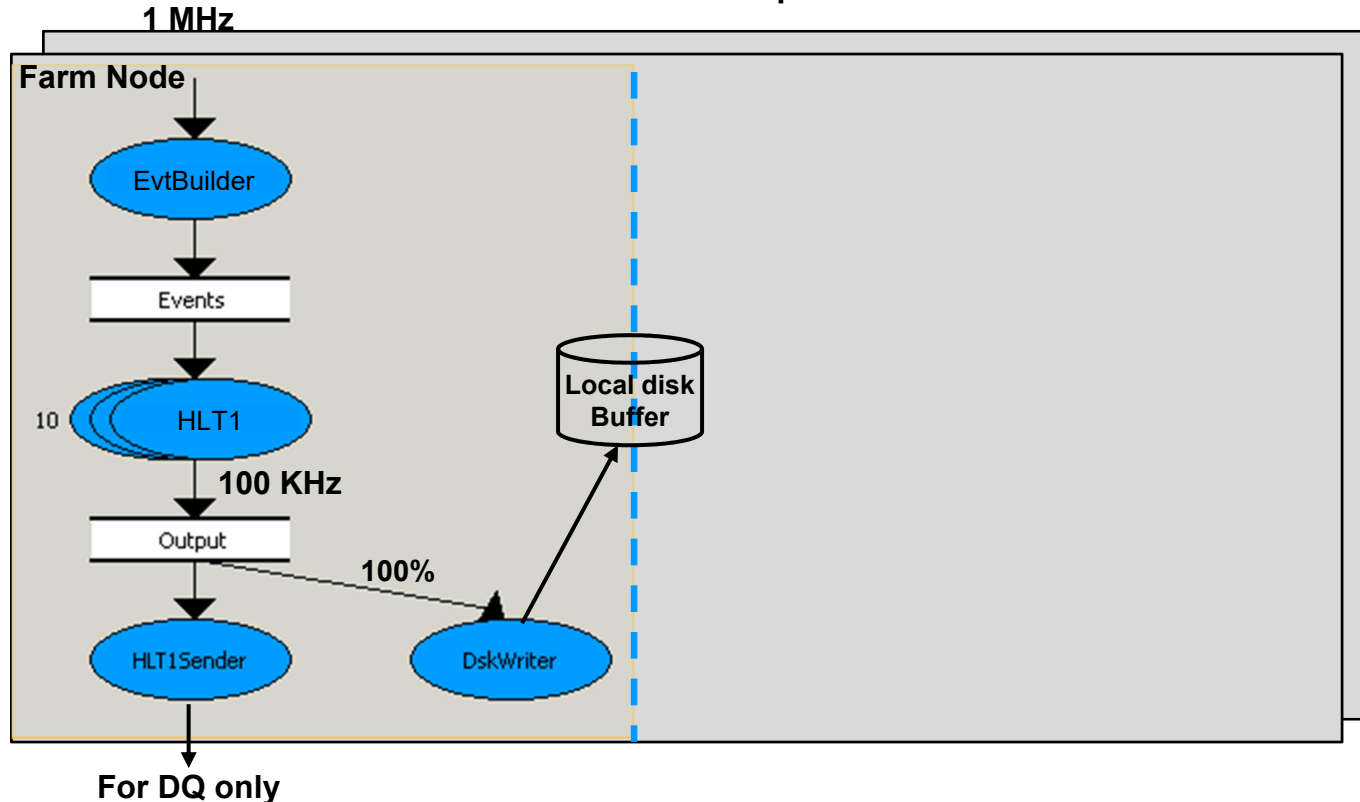


Deferred HLT



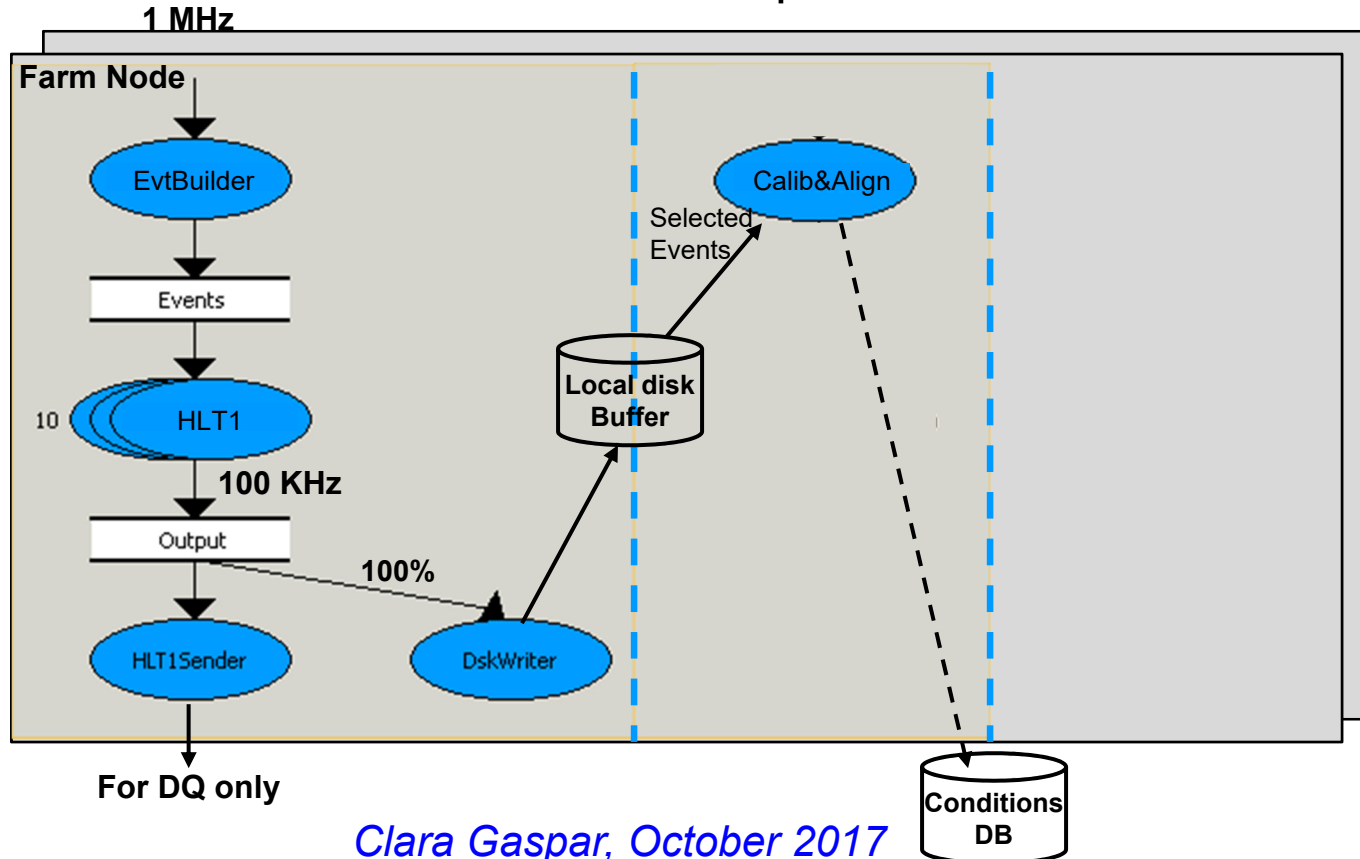
■ Even better Resource Optimization in 2015 -> Split HLT

- Idea: Buffer ALL data to disk after HLT1 / Perform Calibration & Alignment / Run HLT2 permanently in background
- ➡ Each node has 3 concurrent/independent dataflow slices



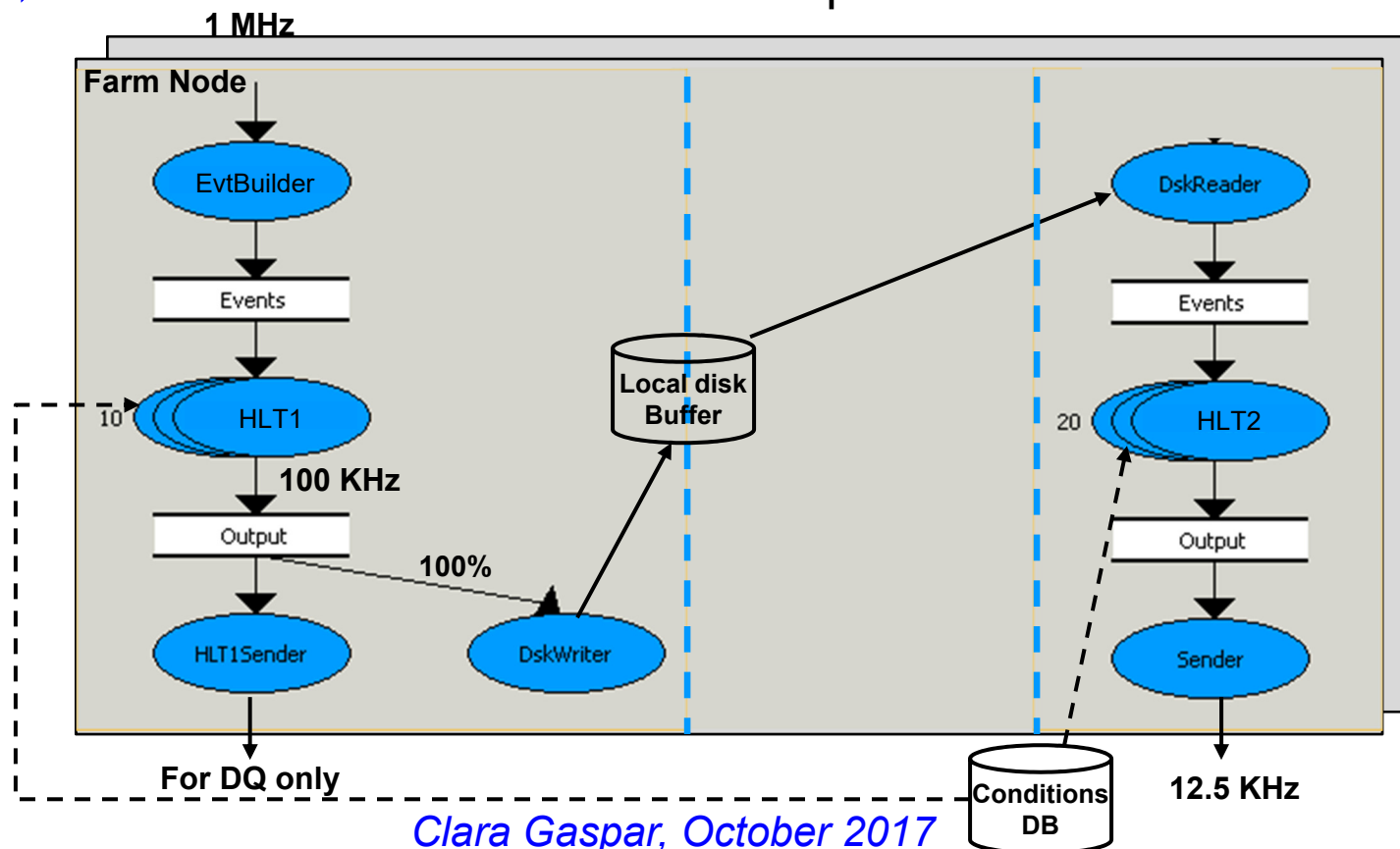
■ Even better Resource Optimization in 2015 -> Split HLT

- Idea: Buffer ALL data to disk after HLT1 / Perform Calibration & Alignment / Run HLT2 permanently in background
- ➡ Each node has 3 concurrent/independent dataflow slices



■ Even better Resource Optimization in 2015 -> Split HLT

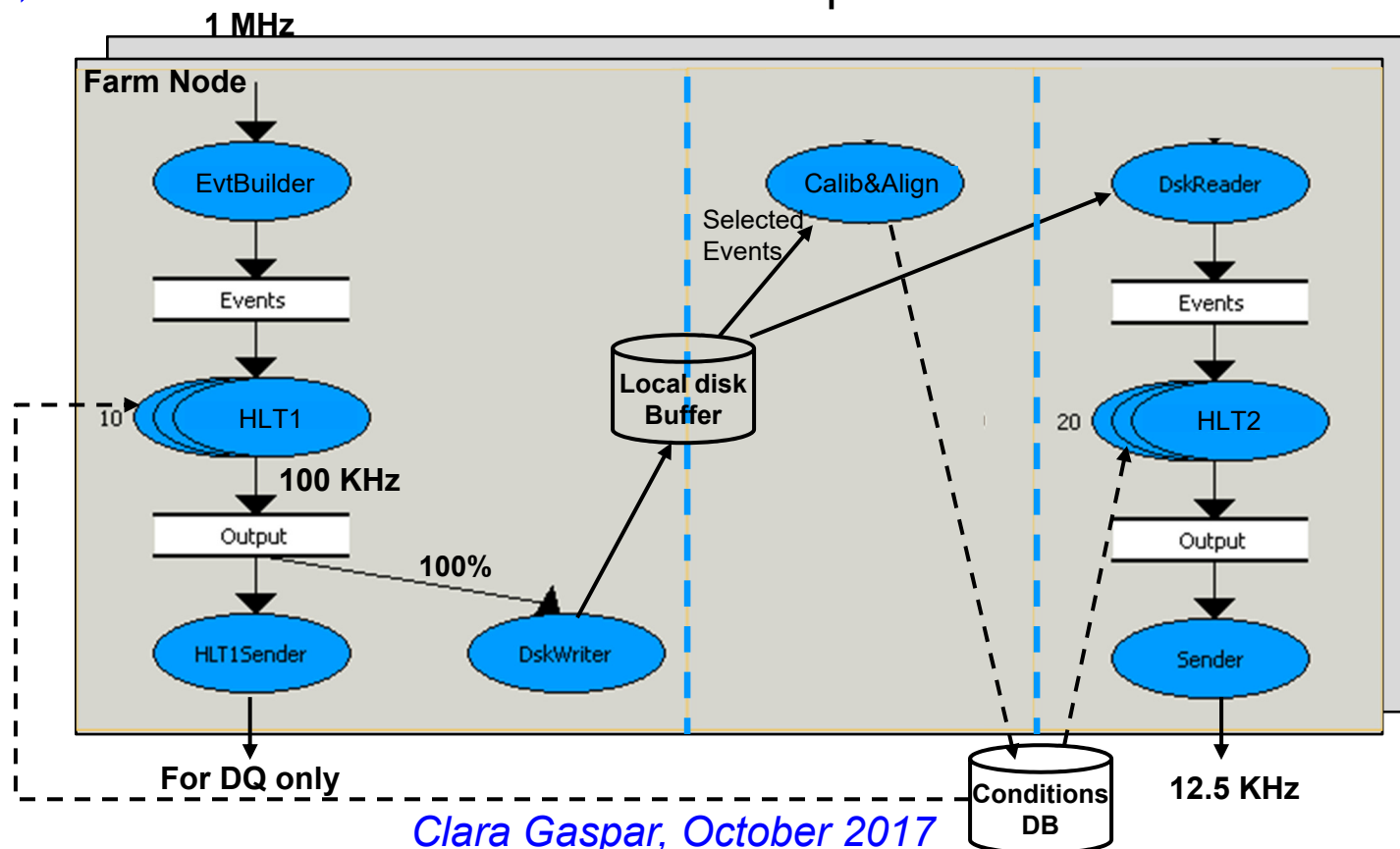
- Idea: Buffer ALL data to disk after HLT1 / Perform Calibration & Alignment / Run HLT2 permanently in background
- ➔ Each node has 3 concurrent/independent dataflow slices

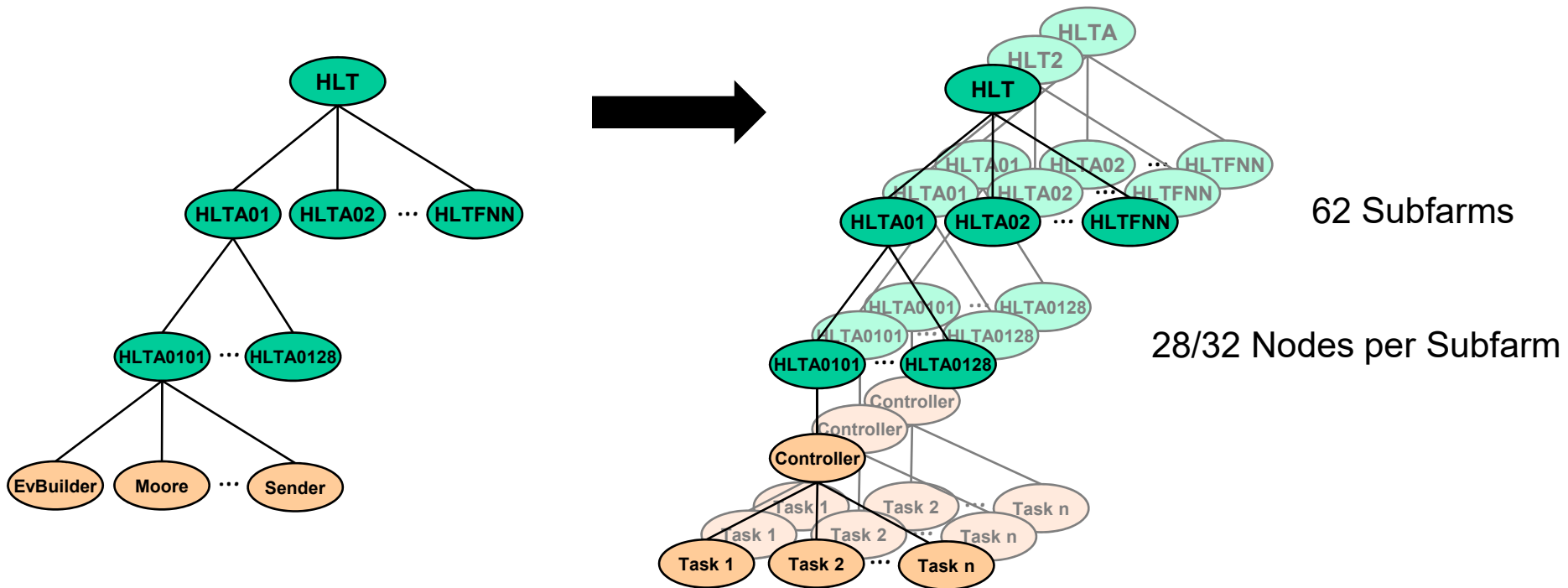


Clara Gaspar, October 2017

■ Even better Resource Optimization in 2015 -> Split HLT

- Idea: Buffer ALL data to disk after HLT1 / Perform Calibration & Alignment / Run HLT2 permanently in background
- ➔ Each node has 3 concurrent/independent dataflow slices





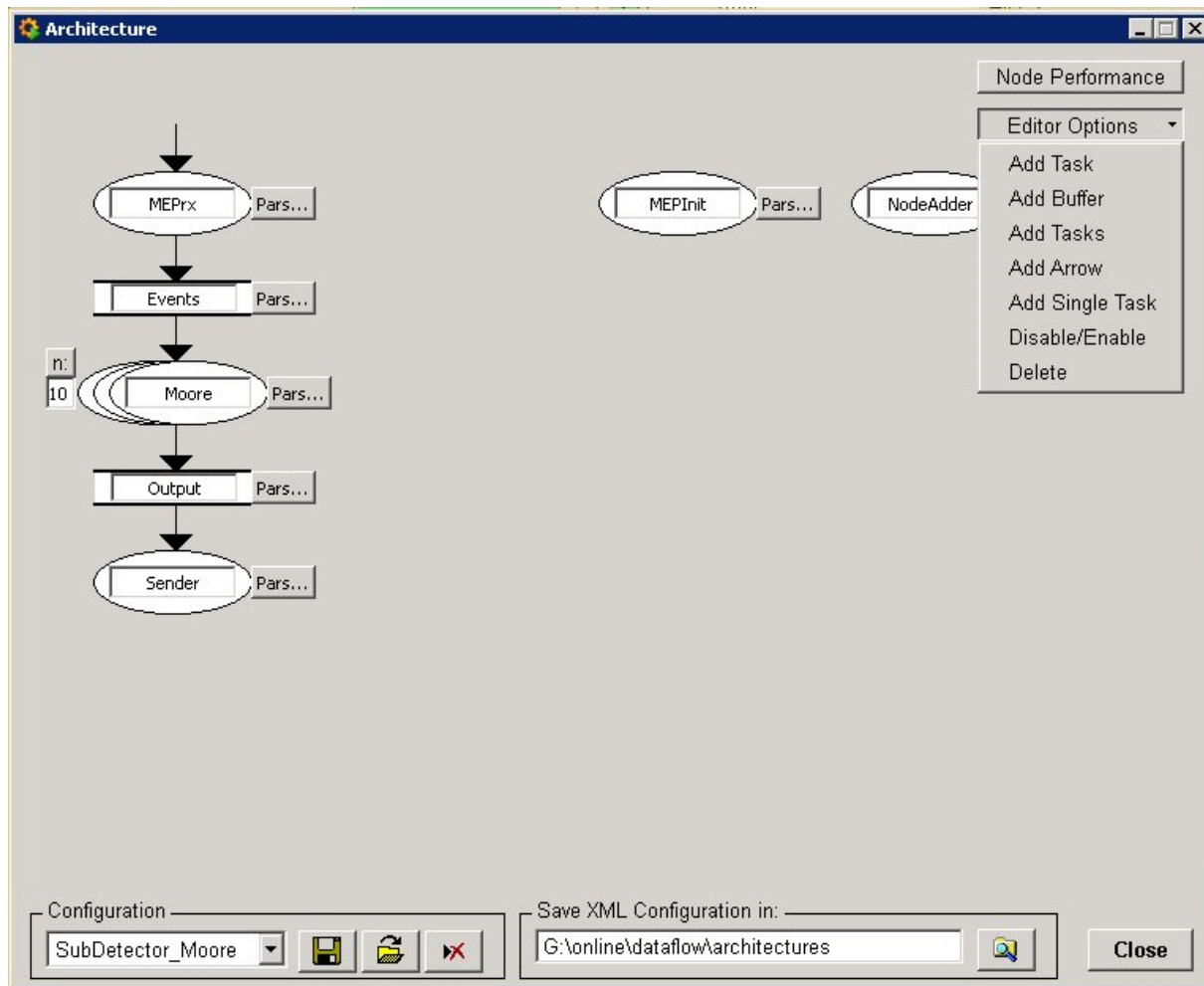
■ Control Tree

- From 1 to 3 Independent Trees (actions, states and partitioning)

■ Dataflow Tasks:

- In the past were Hardwired in the Control Tree
- Now Dynamic via a Controller which receives an “Architecture”

- Defines the dataflow / task layout on each (farm) node



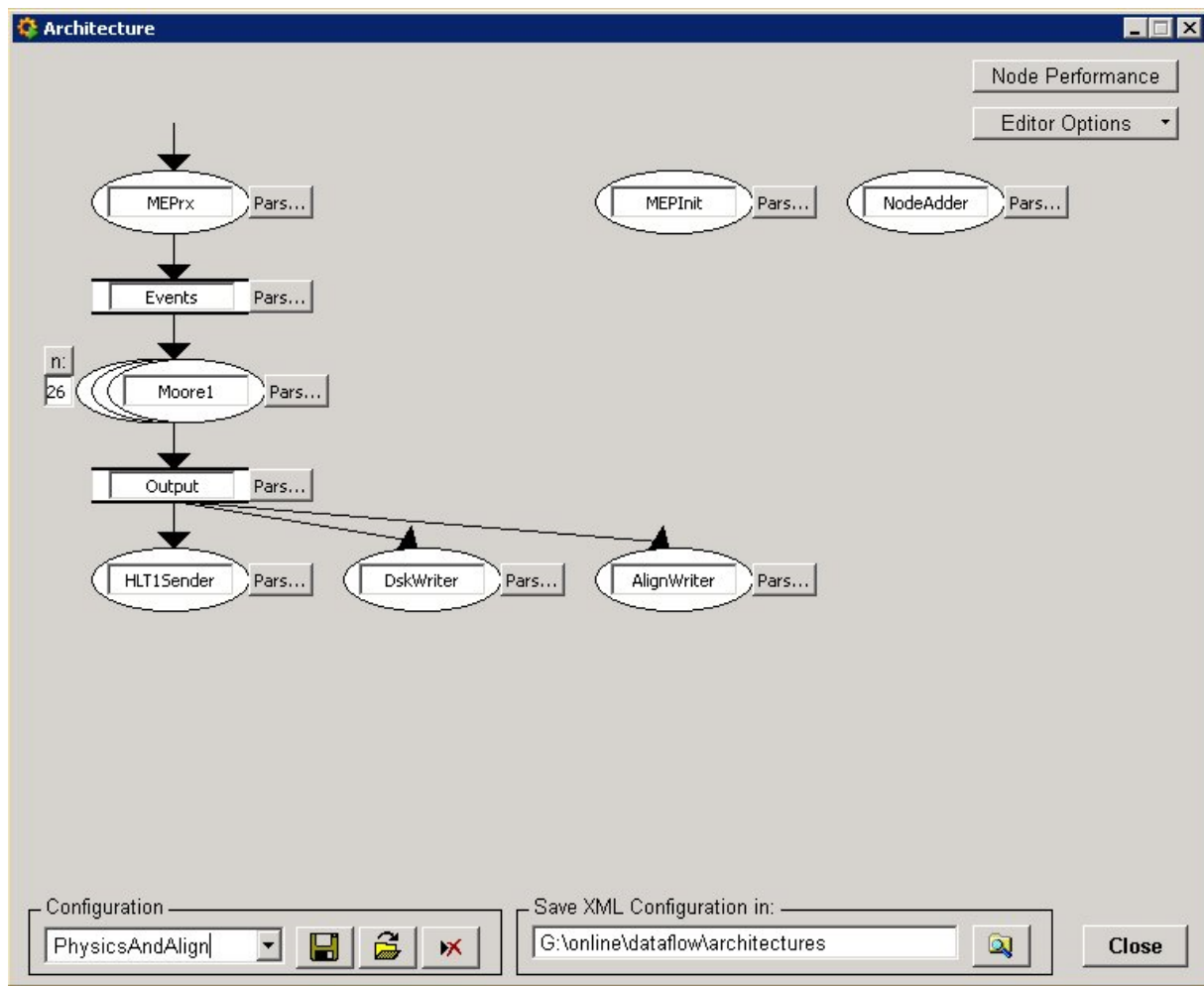
- Graphic Editor:

- Examples:

- | Old-style HLT
 - | HLT1
 - | HLT2
 - | Calibration, Alignment, etc.

- Graphic Used at Run Time

- Defines the dataflow / task layout on each (farm) node



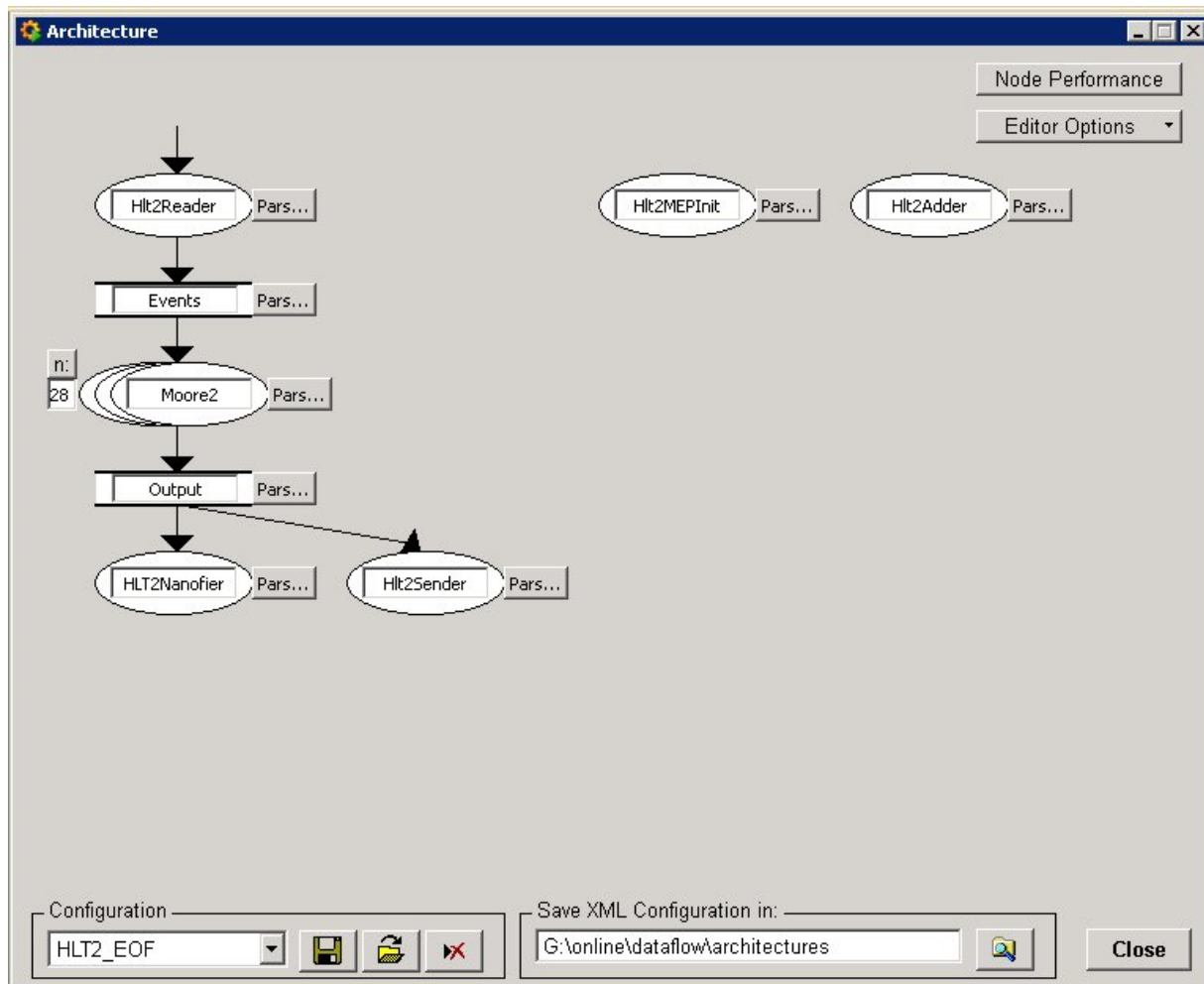
- Graphic Editor:

- Examples:

- | Old-style HLT
 - | HLT1
 - | HLT2
 - | Calibration, Alignment, etc.

- Graphic Used at Run Time

- Defines the dataflow / task layout on each (farm) node



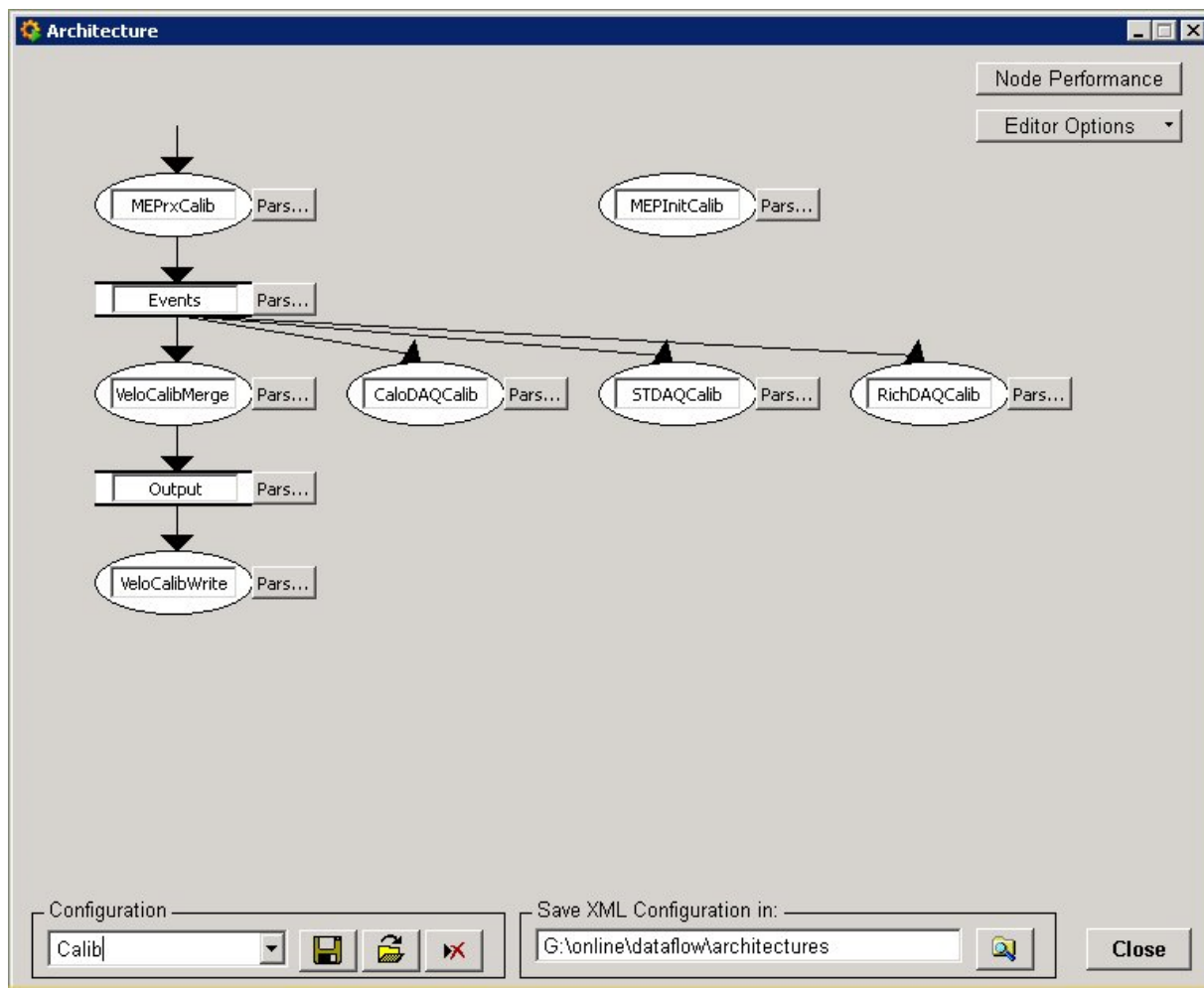
- Graphic Editor:

- Examples:

- | Old-style HLT
 - | HLT1
 - | HLT2
 - | Calibration, Alignment, etc.

- Graphic Used at Run Time

- Defines the dataflow / task layout on each (farm) node



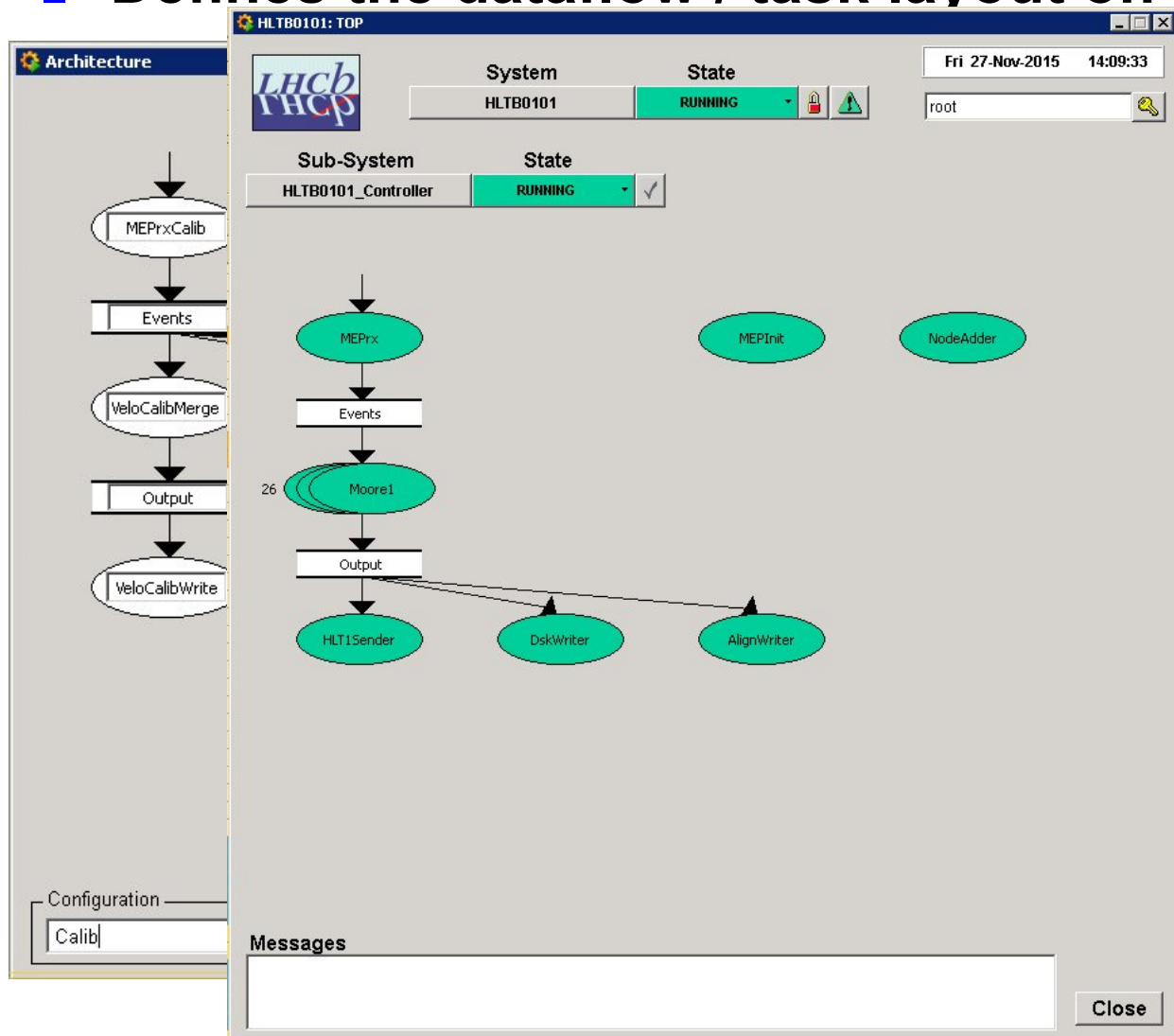
- **Graphic Editor:**

- **Examples:**

- | Old-style HLT
 - | HLT1
 - | HLT2
 - | Calibration, Alignment, etc.

- **Graphic Used at Run Time**

- Defines the dataflow / task layout on each (farm) node



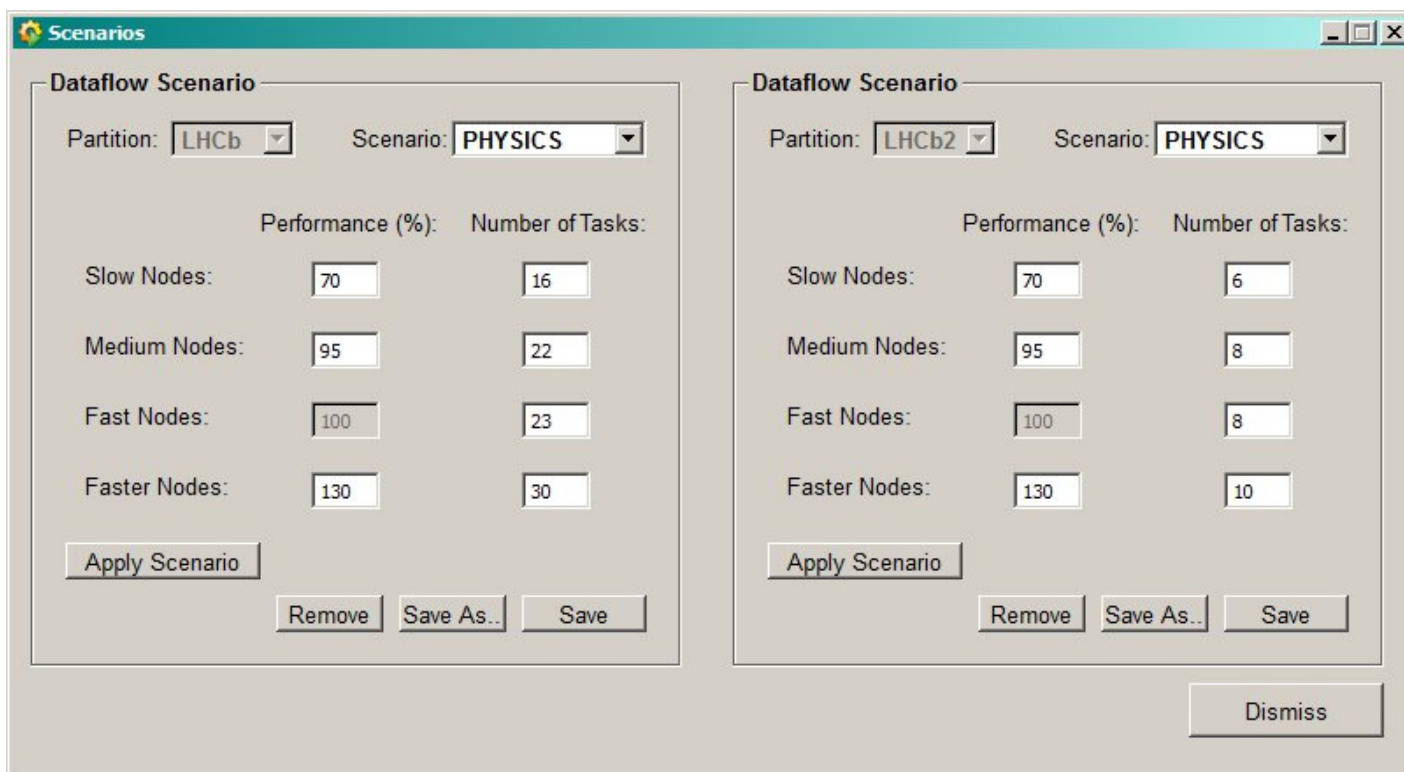
- Graphic Editor:

- Examples:

- | Old-style HLT
 - | HLT1
 - | HLT2
 - | Calibration, Alignment, etc.

- Graphic Used at Run Time

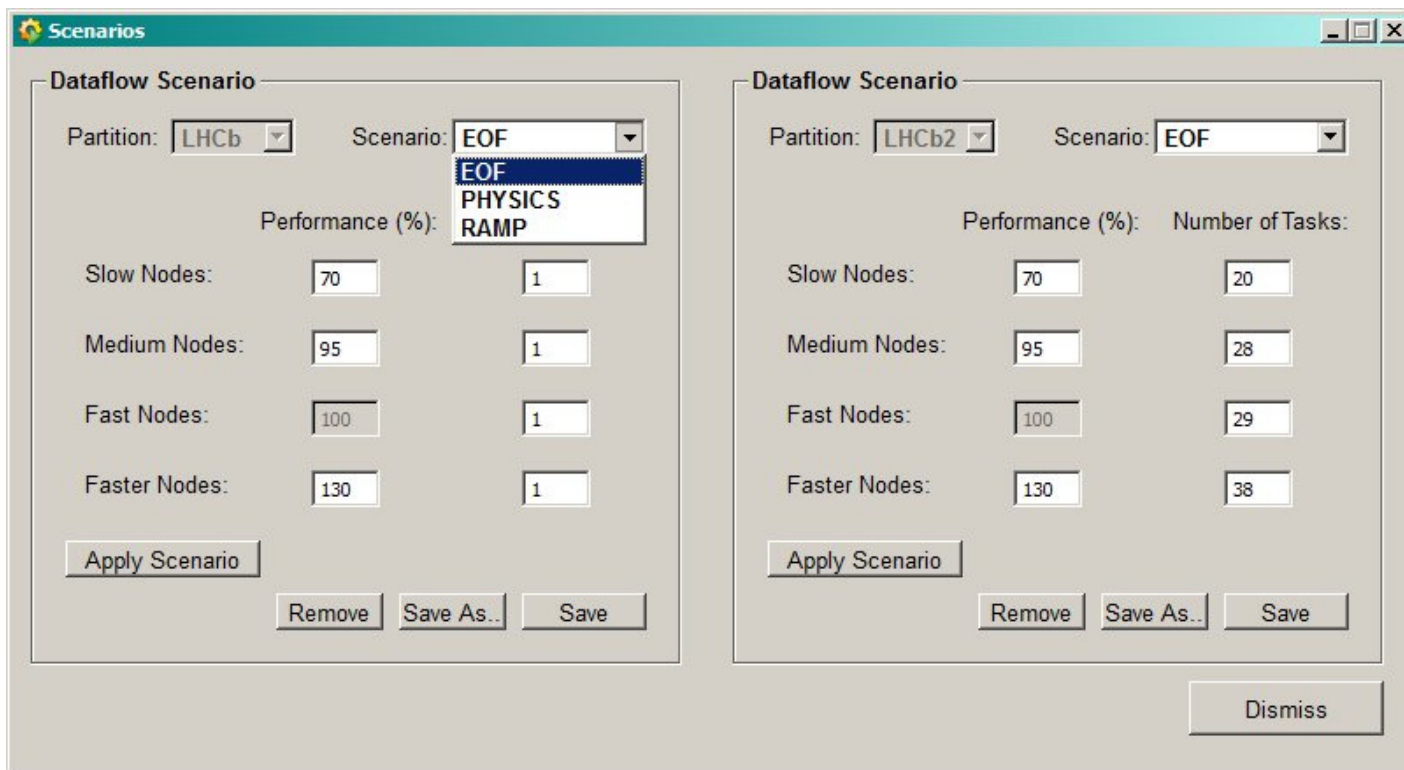
- Defines number of HLT1 & HLT2 tasks for load balancing (initially hardwired in the Architecture)
 - Depending on the LHC/LHCb State and type of node
 - Can be applied at RunTime and takes effect within a second



The screenshot shows the 'Scenarios' application window with two panels for configuring dataflow scenarios. Each panel has a 'Partition' dropdown (set to 'LHCb' and 'LHCb2' respectively) and a 'Scenario' dropdown (both set to 'PHYSICS'). Below these, there are two columns: 'Performance (%)' and 'Number of Tasks'. The 'Performance (%)' column has four rows: 'Slow Nodes' (70), 'Medium Nodes' (95), 'Fast Nodes' (100), and 'Faster Nodes' (130). The 'Number of Tasks' column has four rows: 'Slow Nodes' (16), 'Medium Nodes' (22), 'Fast Nodes' (23), and 'Faster Nodes' (30). At the bottom of each panel are buttons for 'Apply Scenario', 'Remove', 'Save As..', and 'Save'. A 'Dismiss' button is located at the bottom right of the window.

Node Type	Performance (%)	Number of Tasks
Slow Nodes	70	16
Medium Nodes	95	22
Fast Nodes	100	23
Faster Nodes	130	30

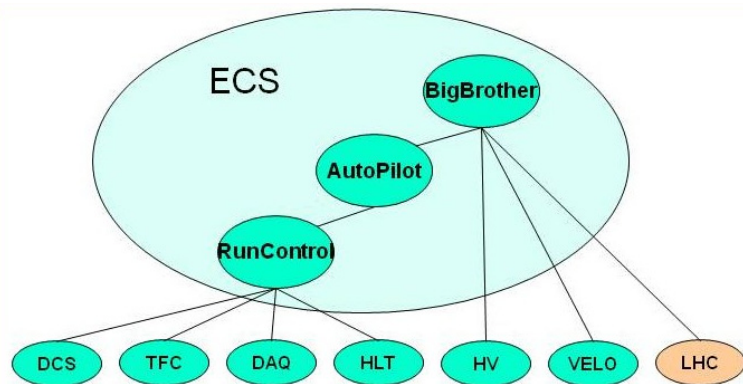
- Defines number of HLT1 & HLT2 tasks for load balancing (initially hardwired in the Architecture)
 - Depending on the LHC/LHCb State and type of node
 - Can be applied at RunTime and takes effect within a second



The screenshot shows the 'Scenarios' application window with two panels for configuring dataflow scenarios. The left panel is for 'LHCb' and the right panel is for 'LHCb2'. Both panels show a 'Dataflow Scenario' configuration with a dropdown menu for 'Scenario' (currently set to 'EOF') and a table of node types and their counts. The 'Performance (%)' column is also present. The 'LHCb' panel has a 'Performance (%)' column with values 70, 95, 100, and 130. The 'LHCb2' panel has a 'Performance (%)' column with values 70, 95, 100, and 130. The 'Number of Tasks' column is also present. The 'LHCb' panel has a 'Performance (%)' column with values 70, 95, 100, and 130. The 'LHCb2' panel has a 'Performance (%)' column with values 70, 95, 100, and 130. The 'LHCb' panel has a 'Performance (%)' column with values 70, 95, 100, and 130. The 'LHCb2' panel has a 'Performance (%)' column with values 70, 95, 100, and 130.

Node Type	Performance (%)	Number of Tasks
Slow Nodes	70	1
Medium Nodes	95	1
Fast Nodes	100	1
Faster Nodes	130	1

Buttons: Apply Scenario, Remove, Save As.., Save, Dismiss



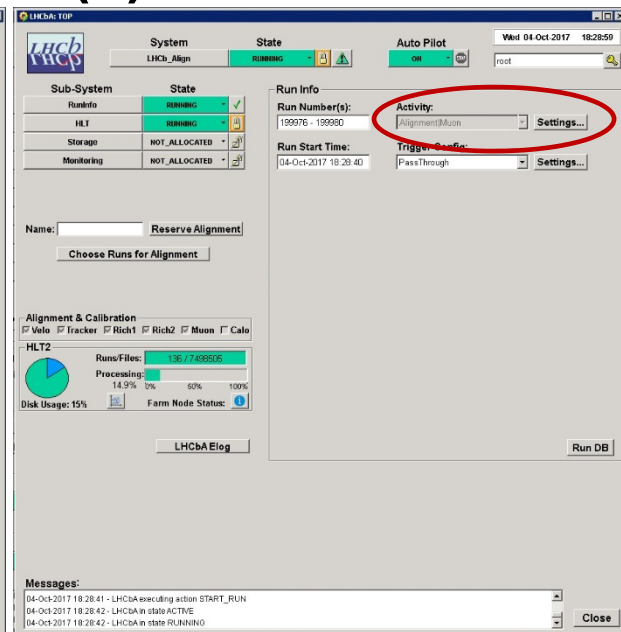
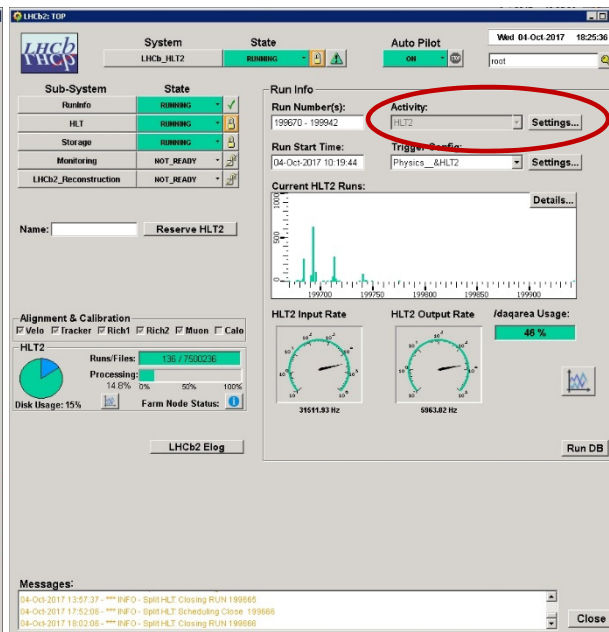
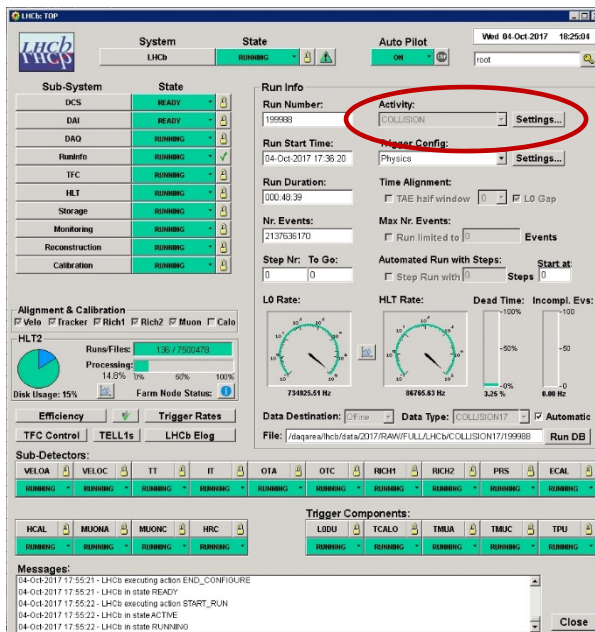
Run Control driven by an Activity:

- Defines the set of parameters (“recipe”) which is applied by each sub-system.

- Example: “PHYSICS|LEAD”

- Contains, for example, the “Architecture”

Three Farm Slices -> Three Run Control(s):



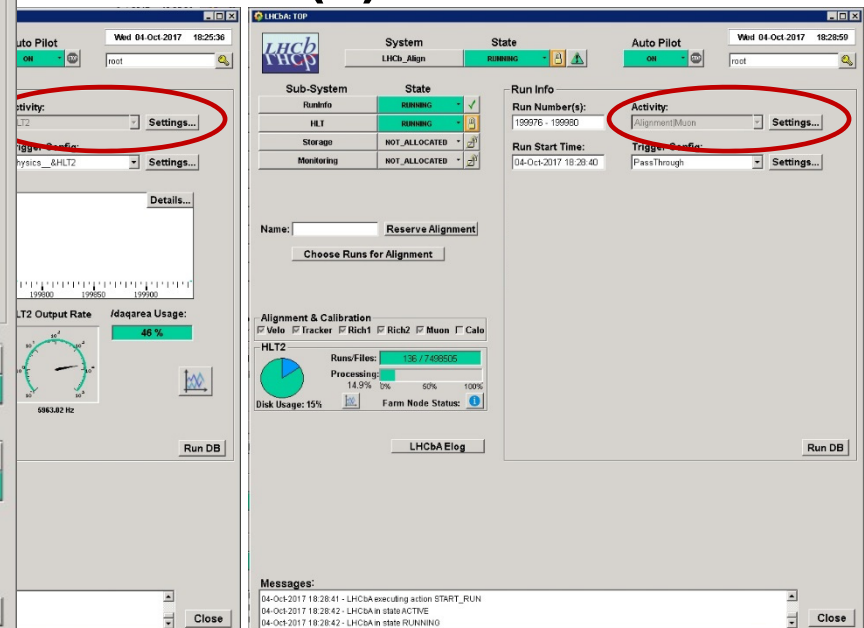
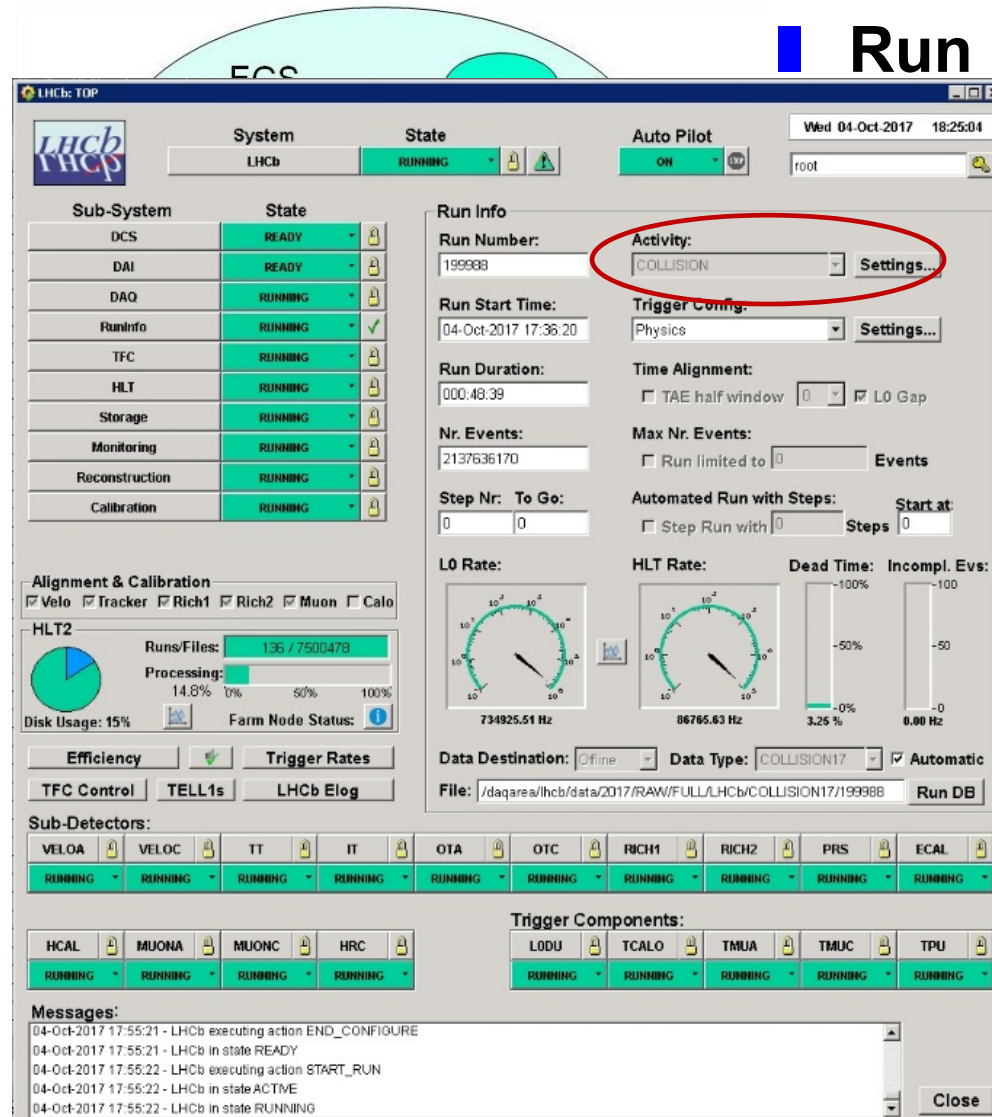
Run Control driven by an Activity:

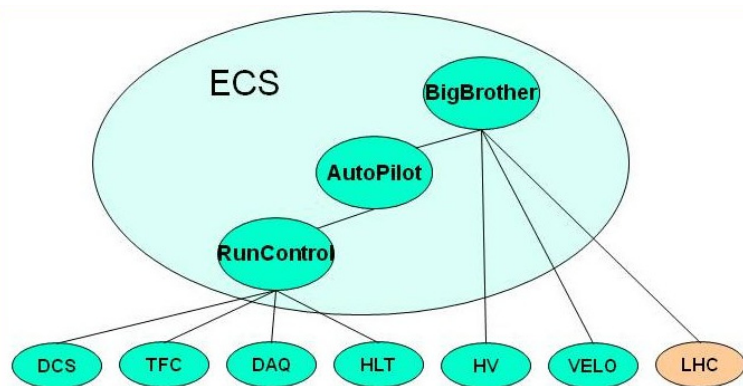
defines the set of parameters (“recipe”) which is applied by each sub-system.

Example: “PHYSICS|LEAD”

Contains, for example, the “Architecture”

Run Control(s):

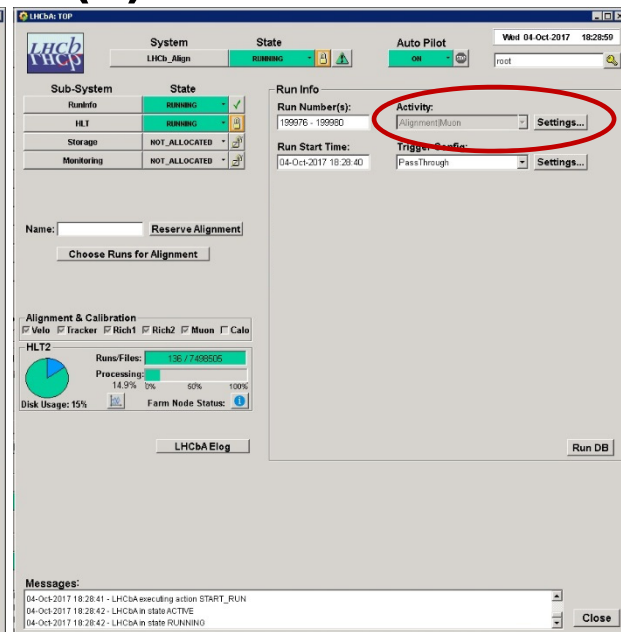
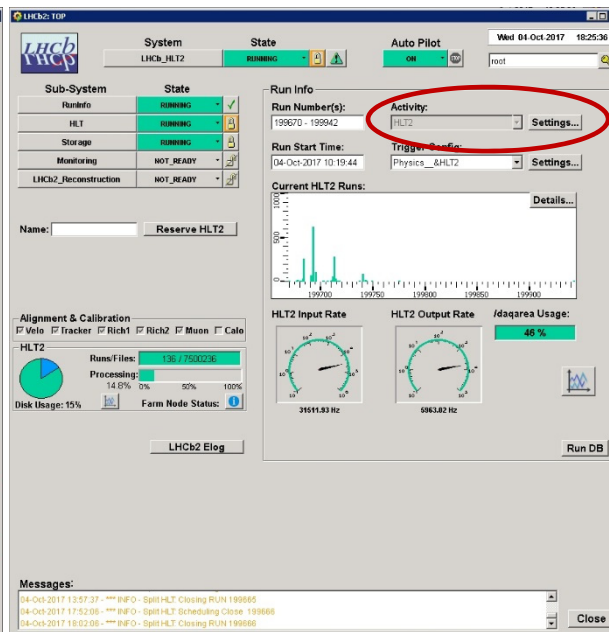
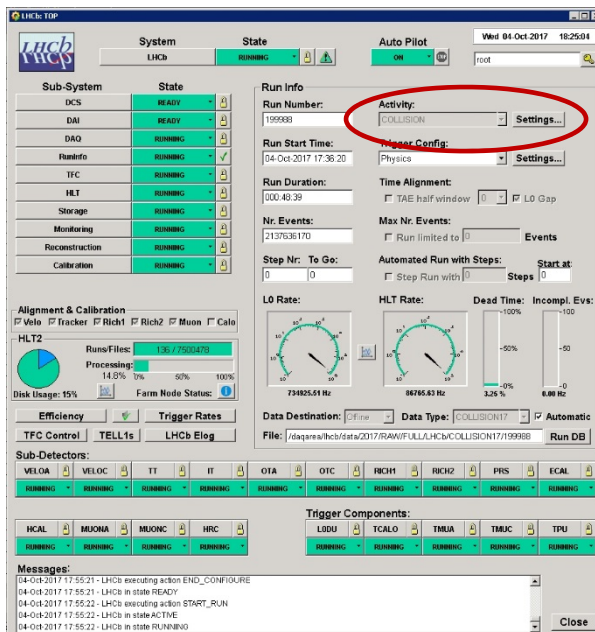


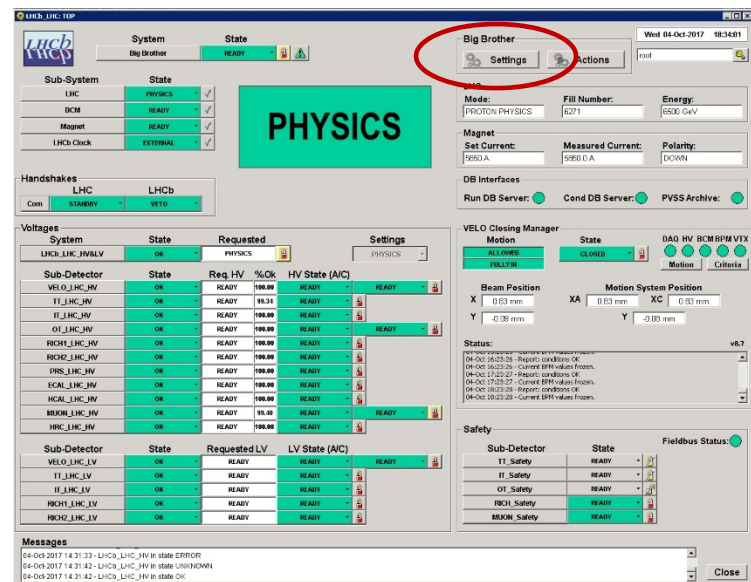


Run Control driven by an Activity:

- Defines the set of parameters (“recipe”) which is applied by each sub-system.
- Example: “PHYSICS|LEAD”
- Contains, for example, the “Architecture”

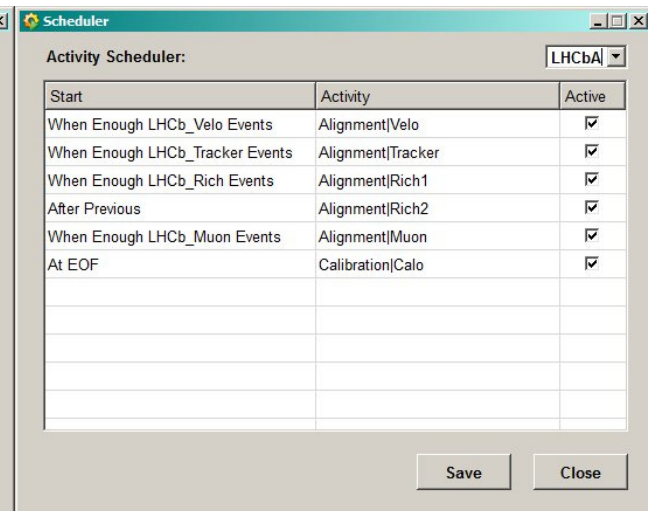
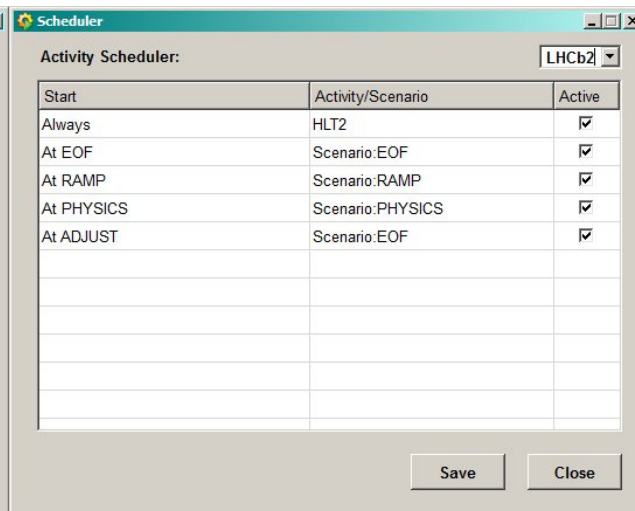
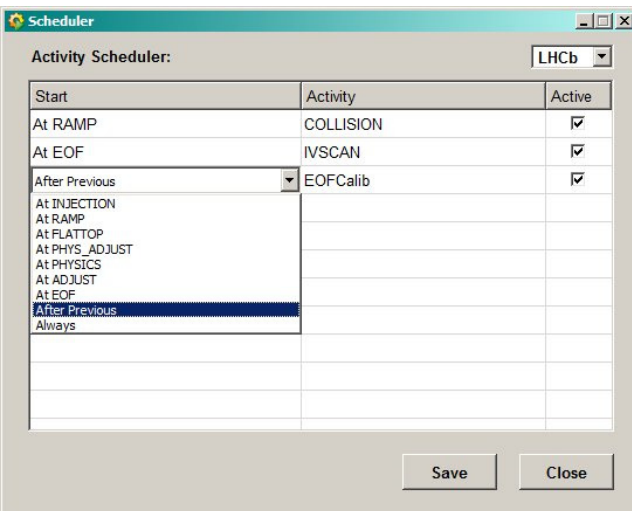
Three Farm Slices -> Three Run Control(s):

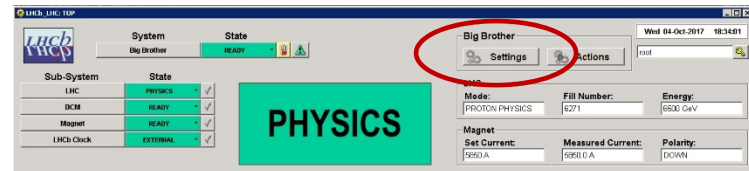




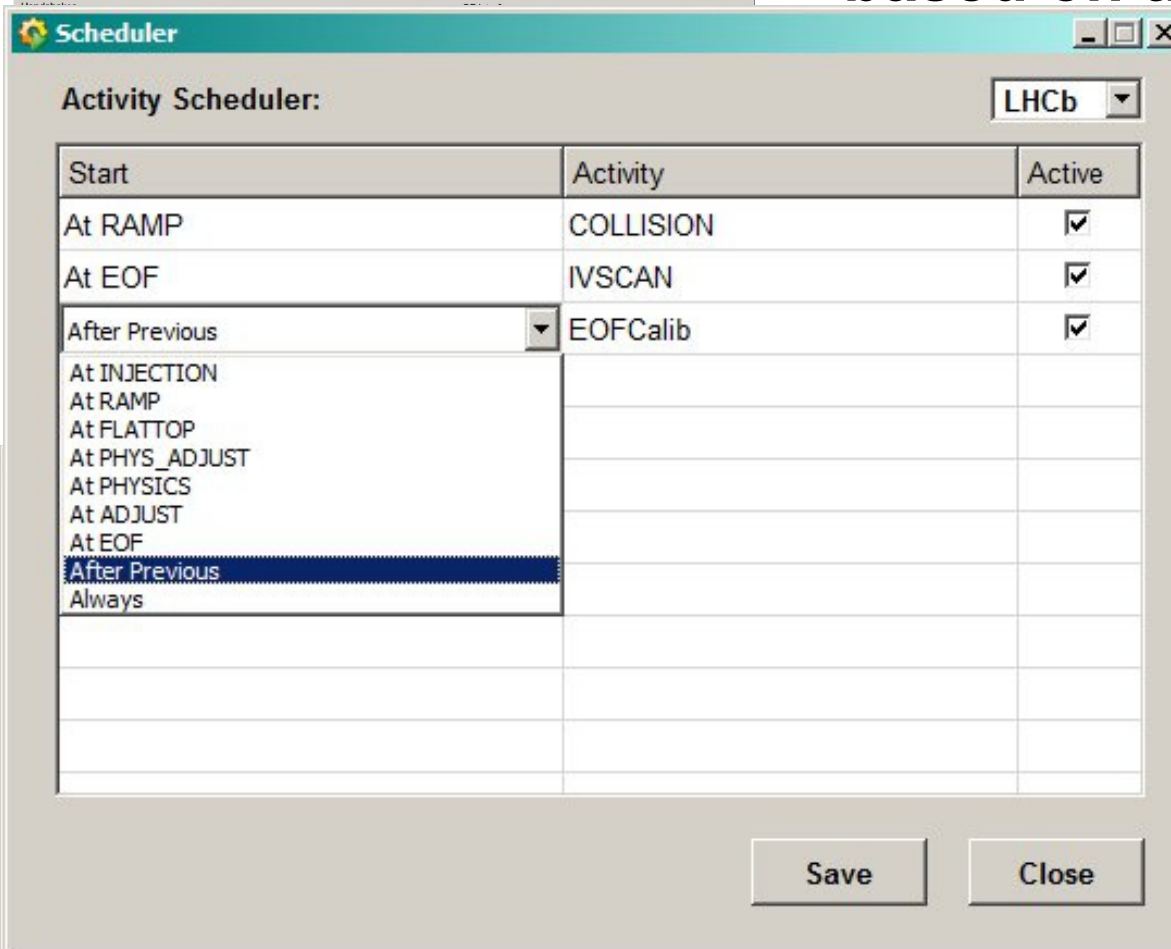
■ **Big Brother automates operations based on the LHC State, driven by Scheduler(s):**

- Define sequences of Activities/Scenarios
- Complete Automation of each Run Control with very few lines
- 3 lines to automate main Run Control

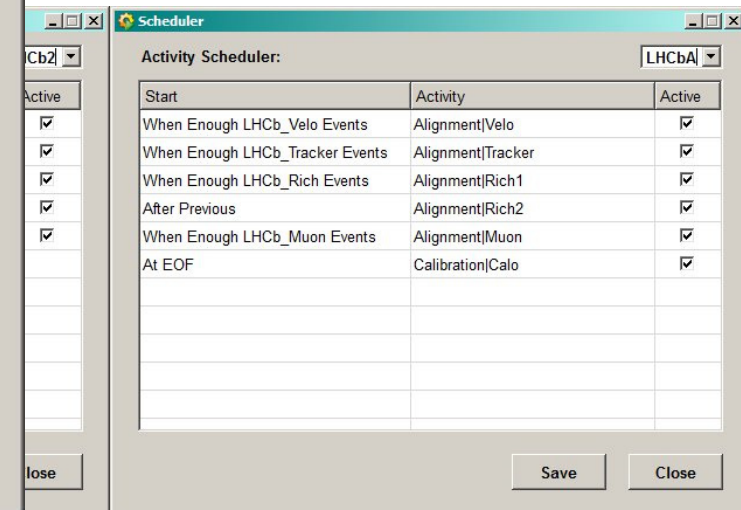


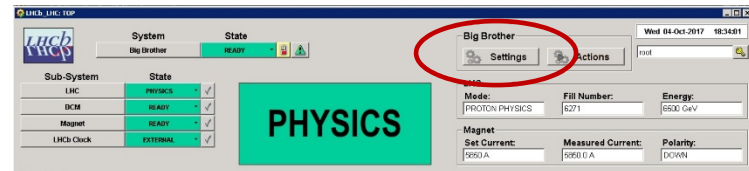


■ **Big Brother automates operations based on the LHC State, driven by**

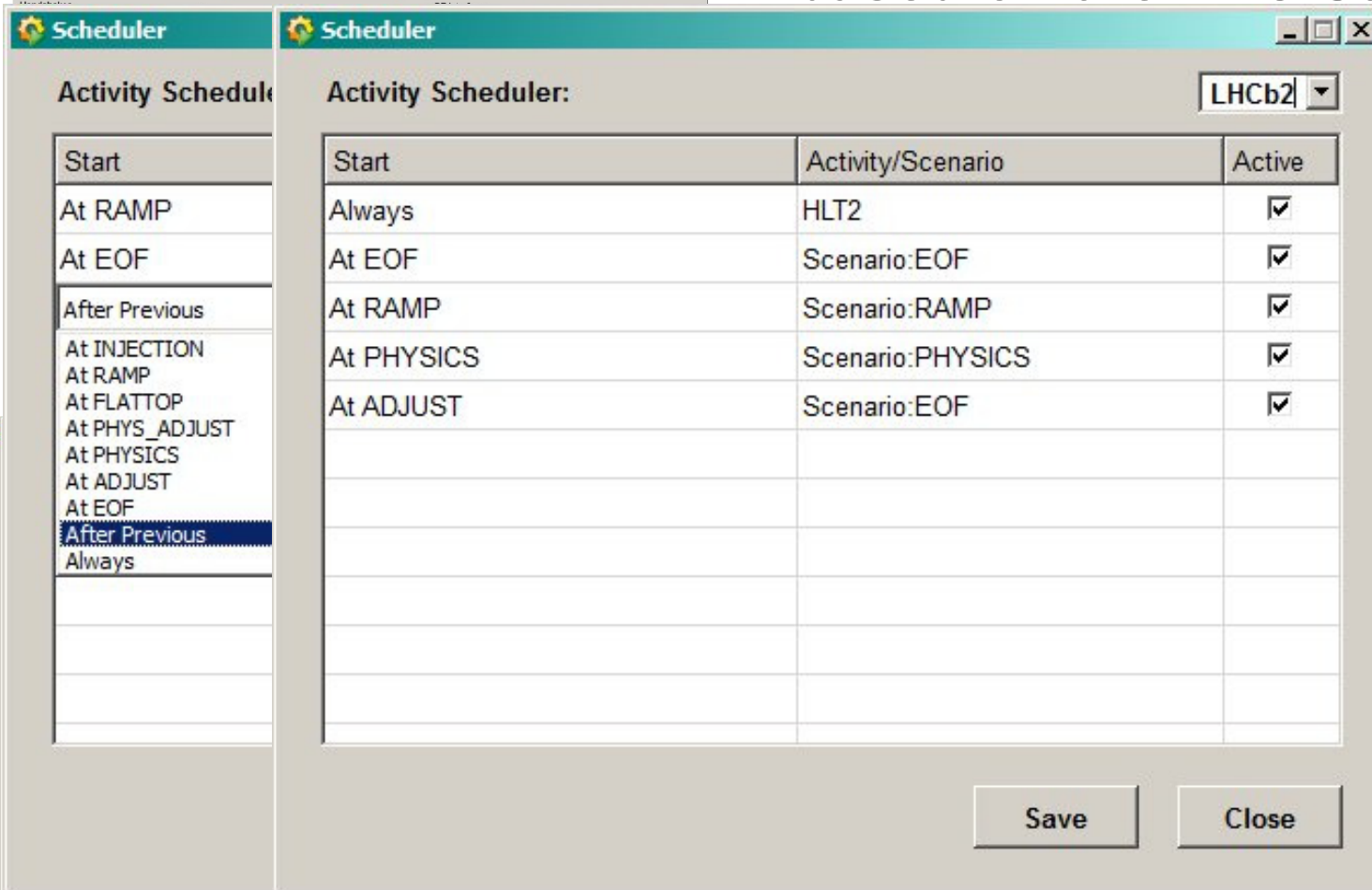


sequences of Activities/Scenarios
automation of each Run
in very few lines
Automate main Run Control

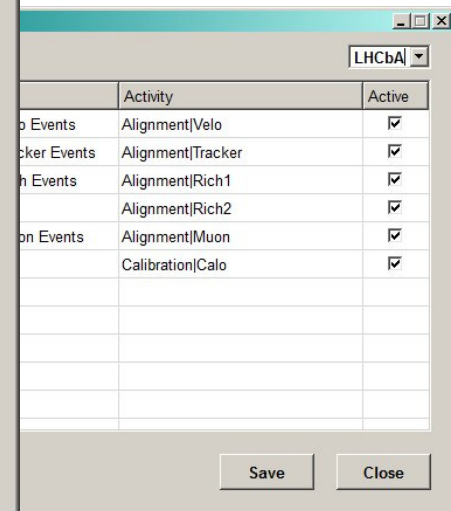


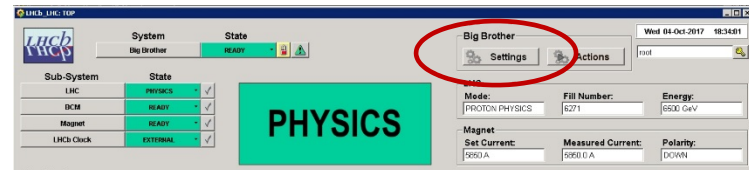


- Big Brother automates operations based on the LHC State, driven by



Activities/Scenarios
of each Run
are
in Run Control





- Big Brother automates operations based on the LHC State, driven by

Scheduler

Activity Scheduler:

Start	Activity	Active
At RAMP	Alignment Velo	☒
At EOF	Alignment Tracker	☒
After Previous	Alignment Rich1	☒
At INJECTION	Alignment Rich2	☒
At RAMP	Alignment Muon	☒
At FLATTOP	Calibration Calo	☒
At PHYS_ADJUST		
At PHYSICS		
At ADJUST		
At EOF		
After Previous		
Always		

Scheduler

Activity Scheduler:

Start	Activity	Active
Always		
At EOF		
At RAMP		
At PHYSICS		
At ADJUST		

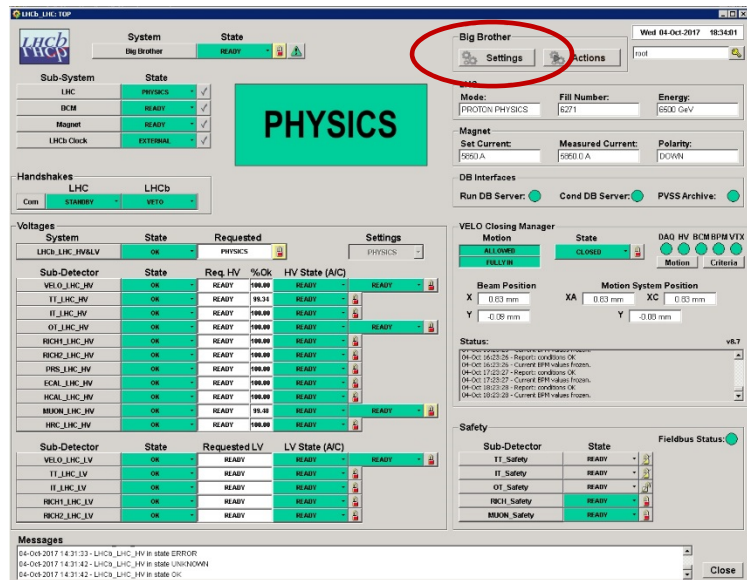
Scheduler

Activity Scheduler:

LHCbA

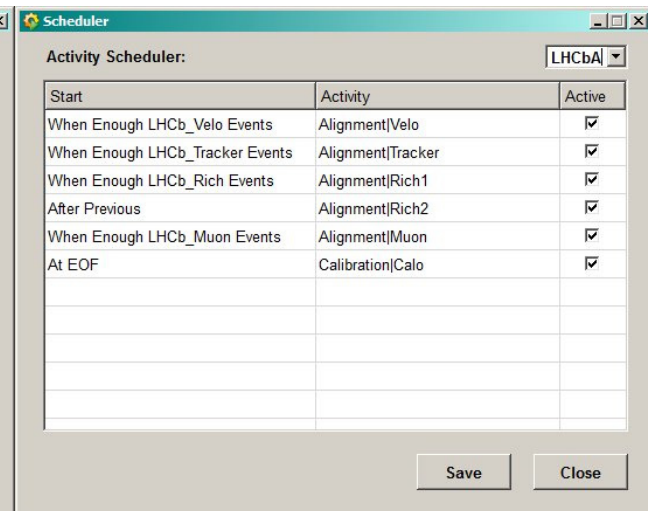
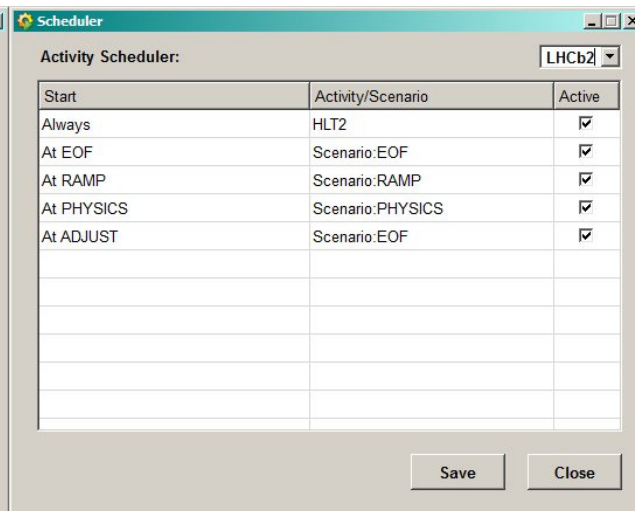
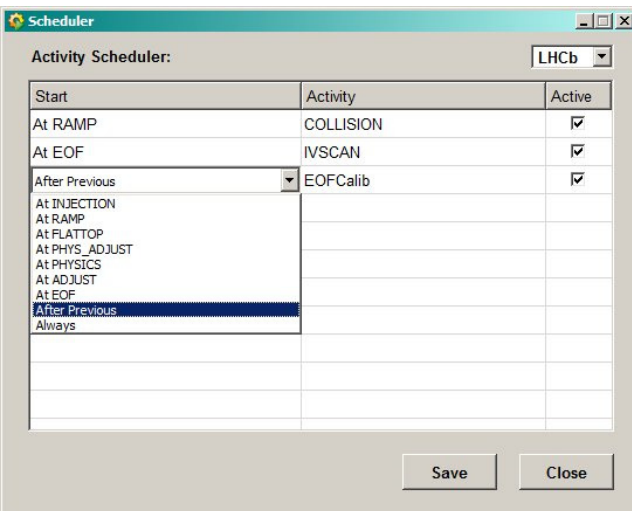
Start	Activity	Active
When Enough LHCb_Velo Events	Alignment Velo	☒
When Enough LHCb_Tracker Events	Alignment Tracker	☒
When Enough LHCb_Rich Events	Alignment Rich1	☒
After Previous	Alignment Rich2	☒
When Enough LHCb_Muon Events	Alignment Muon	☒
At EOF	Calibration Calo	☒

Save Close



■ **Big Brother automates operations based on the LHC State, driven by Scheduler(s):**

- Define sequences of Activities/Scenarios
- Complete Automation of each Run Control with very few lines
- 3 lines to automate main Run Control



- **LHCb's New Online Dataflow represented an Operational challenge, but due to:**
 - The modular architecture and flexibility of the original design
 - Some new concepts (architectures, scenarios, schedulers)
- **It is now:**
 - Seamlessly Integrated
 - Completely automated
- **And the farm is kept as busy as possible**
 - Also offline simulation (concurrently) when still free resources.

