Controlling a large CPU farm using industrial tools

Alba Sambade Varela

On behalf of the LHCb Online group

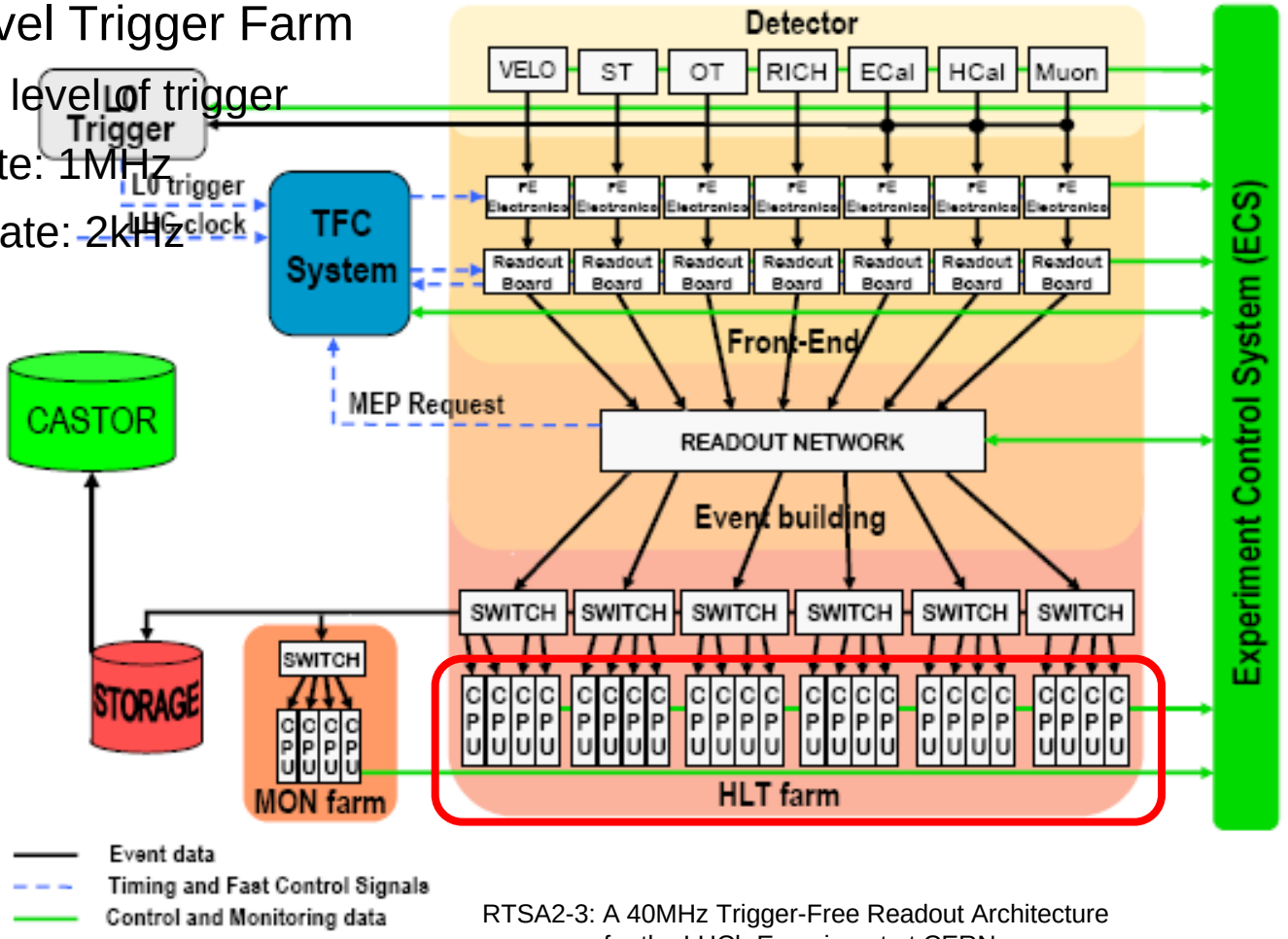
Real Time 2009

May 10-15 IHEP Beijing

ECS and data flow in LHCb

High Level Trigger Farm

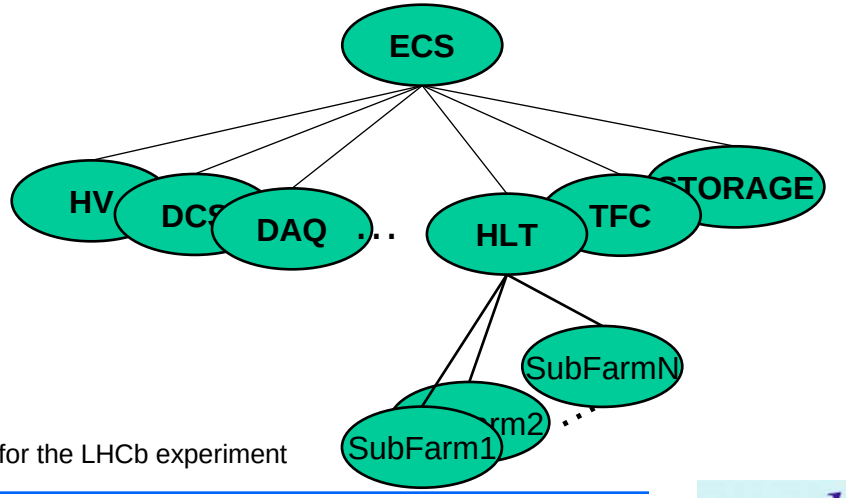
- Second level of trigger
- input rate: 1MHz
- output rate: 2kHz



RTSA2-3: A 40MHz Trigger-Free Readout Architecture for the LHCb Experiment at CERN

Experiment Control System

- ECS is in charge of the configuration, operation and supervision of all the online components in LHCb.
  - Industrial SCADA system: PVSS
  - FSM package
    - Definition in terms of hierarchies of Finite State Machines.
  - Distributed Information Management System (DIM)

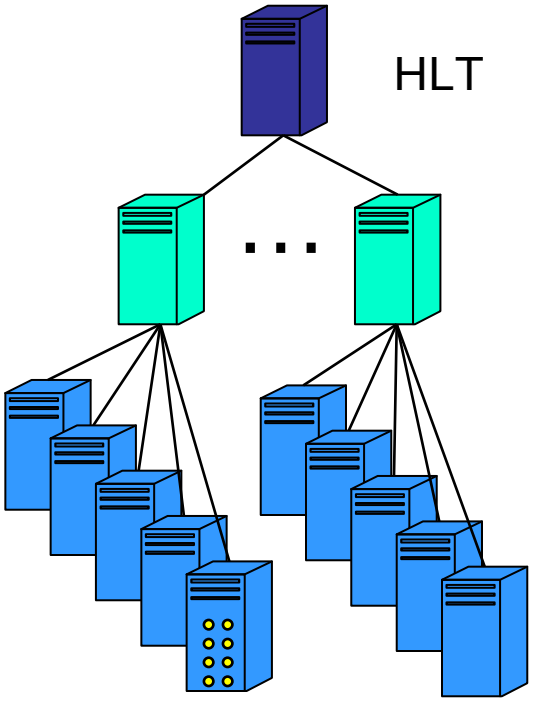


See talk CMS1-1
An Integrated Control System for the LHCb experiment

What do we have to control?

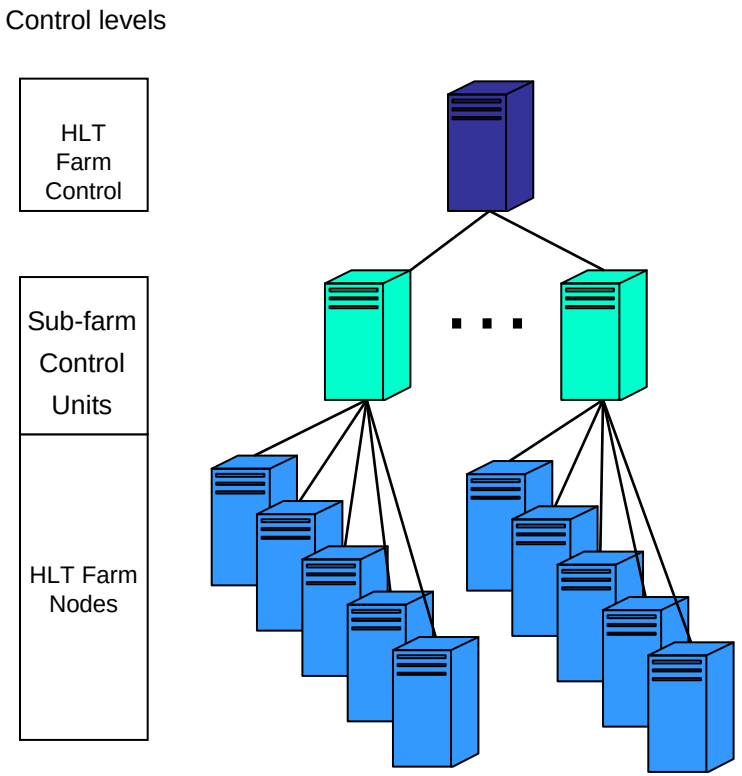
High Level Trigger CPU Farm

- Composed out of up to 2000 boxes (= nodes)
- grouped into 50 subfarms (50 racks)
  - up to 40 nodes/subfarm
    - 8 cores/node *
    - 1 HLT algorithm running/core

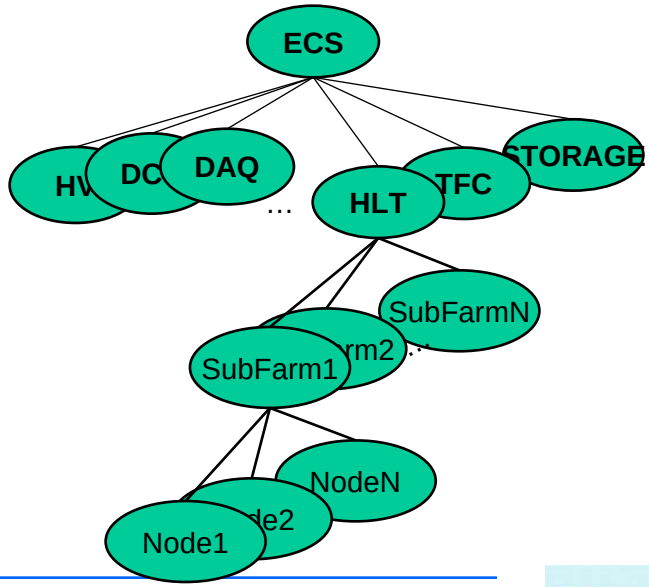


* Nowadays

HLT Top level control

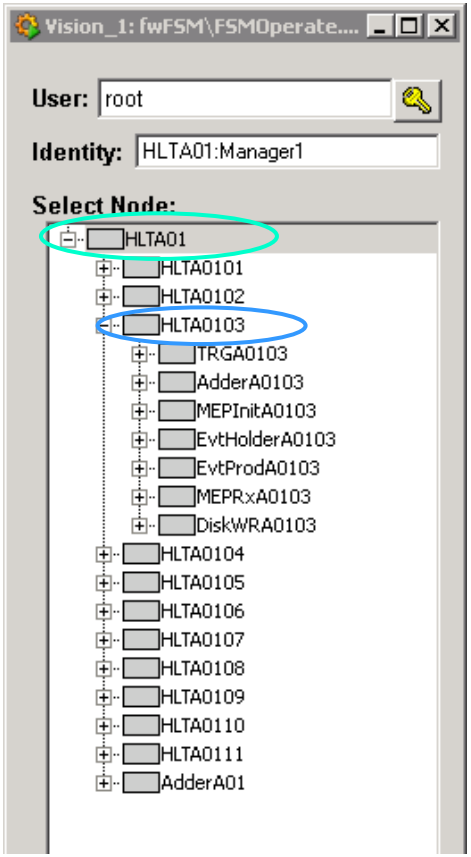


- HLT control PC
  - Running main HLT project
- Control PC per sub-farm
  - Running its own control project
  - Corresponds to a control unit in ECS

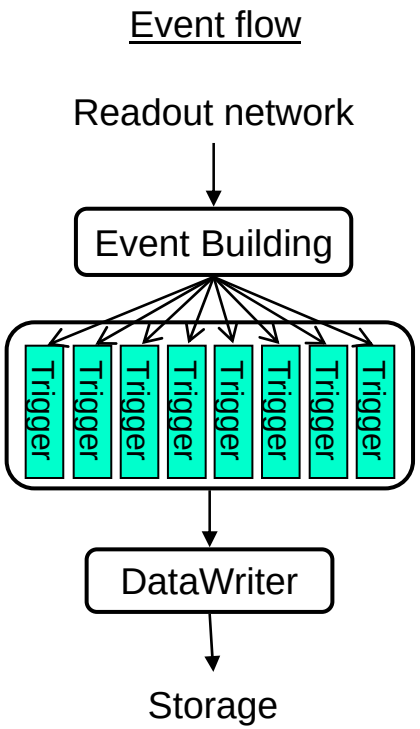


HLT low level control

- Logical aggregation
  - Functional algorithms
- Division of HLT node by tasks



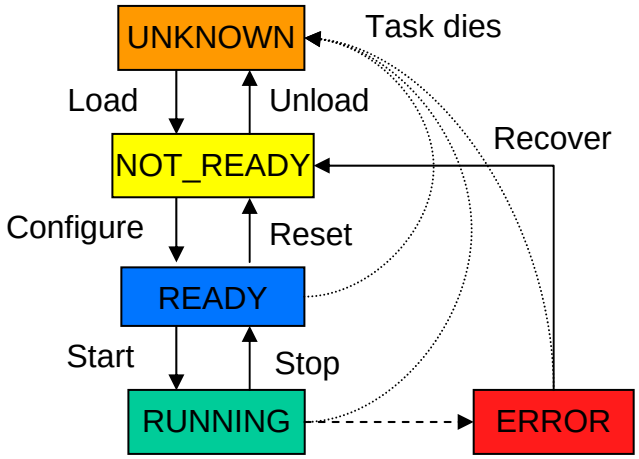
Task architecture on HLT node



- Each node in the EFF runs one event-builder process.
  - It distributes the assembled events to trigger processes.
- As many trigger-processes as there are CPU cores.
  - Compute trigger decision and declare accepted events.
- Each node also runs one instance of the data-writer.
  - Sends accepted events to Storage system.
- Algorithms implemented with GAUDI.
  - GAUDI: data processing experiment independent framework.
  - Same software used as for offline analysis.

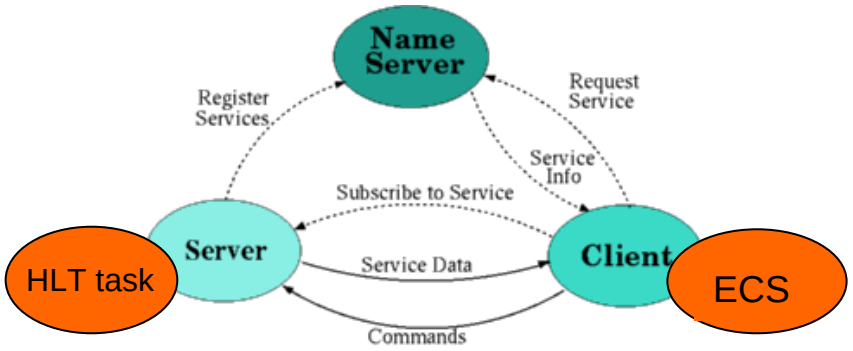
HLT Tasks control

- Algorithms are treated by the ECS as hardware devices.
  - Integrated as Device Unit (DU)
- DU behavior modeled with Finite State Machines
  - Common state diagram for all algorithms
  - Transitions mapped to Gaudi transitions
    - Command parameters



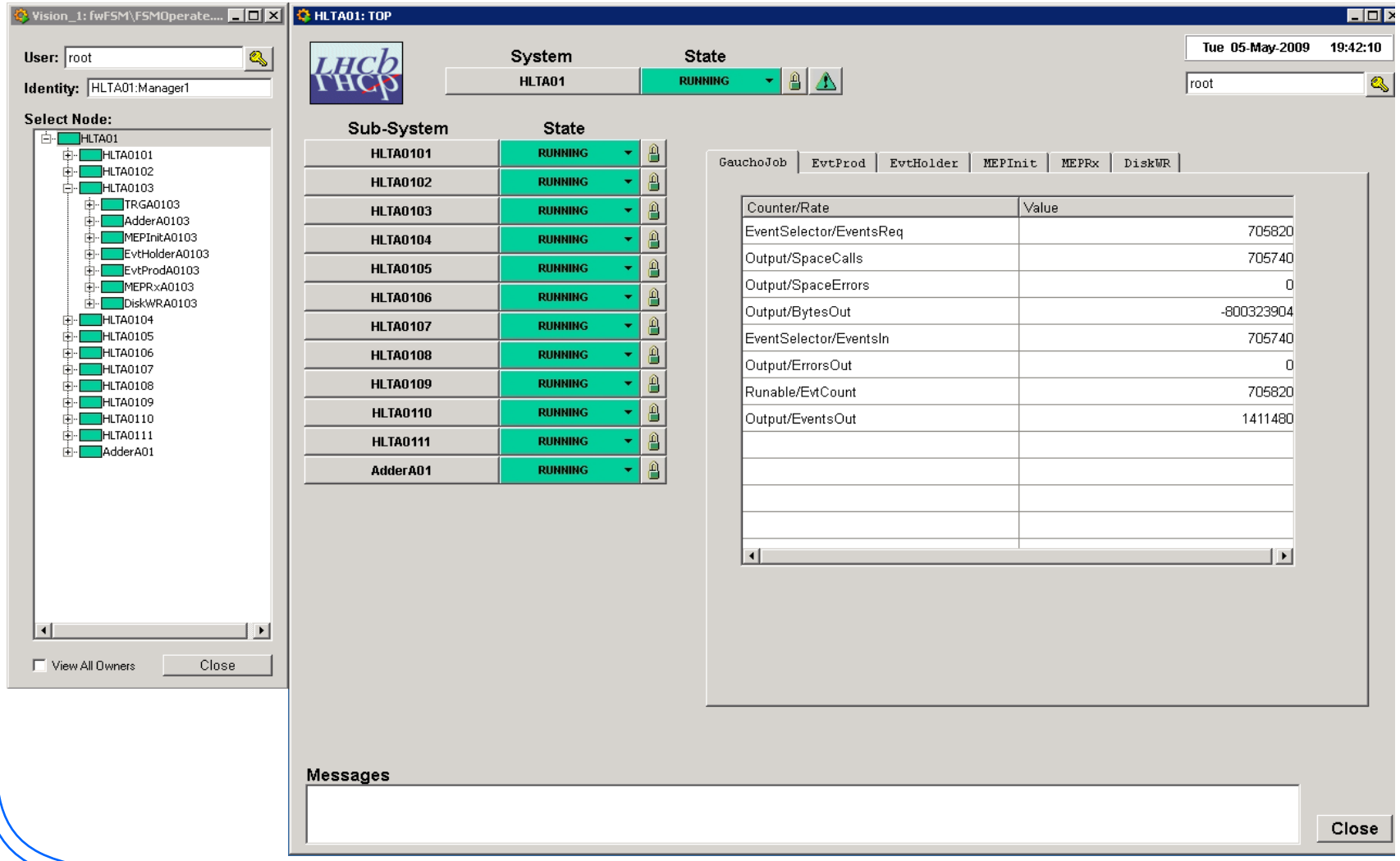
Communication layer

- Communication via Distributed Information Management System (DIM)
  - Communication mechanism based on client-server paradigm
- Task behaves as a DIM server
  - Publishes services
    - Algorithm status
    - Counters
  - Receives commands



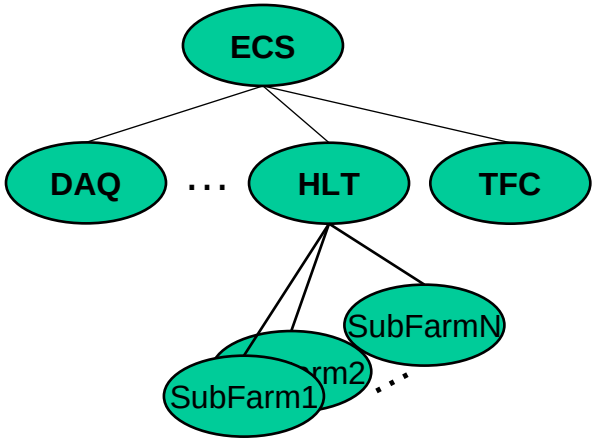
- Controls based on HLT naming convention
  - Sub-farm (row and index)
  - UTGID (task name)

Sub-Farm Control GUI



Sub-farm control

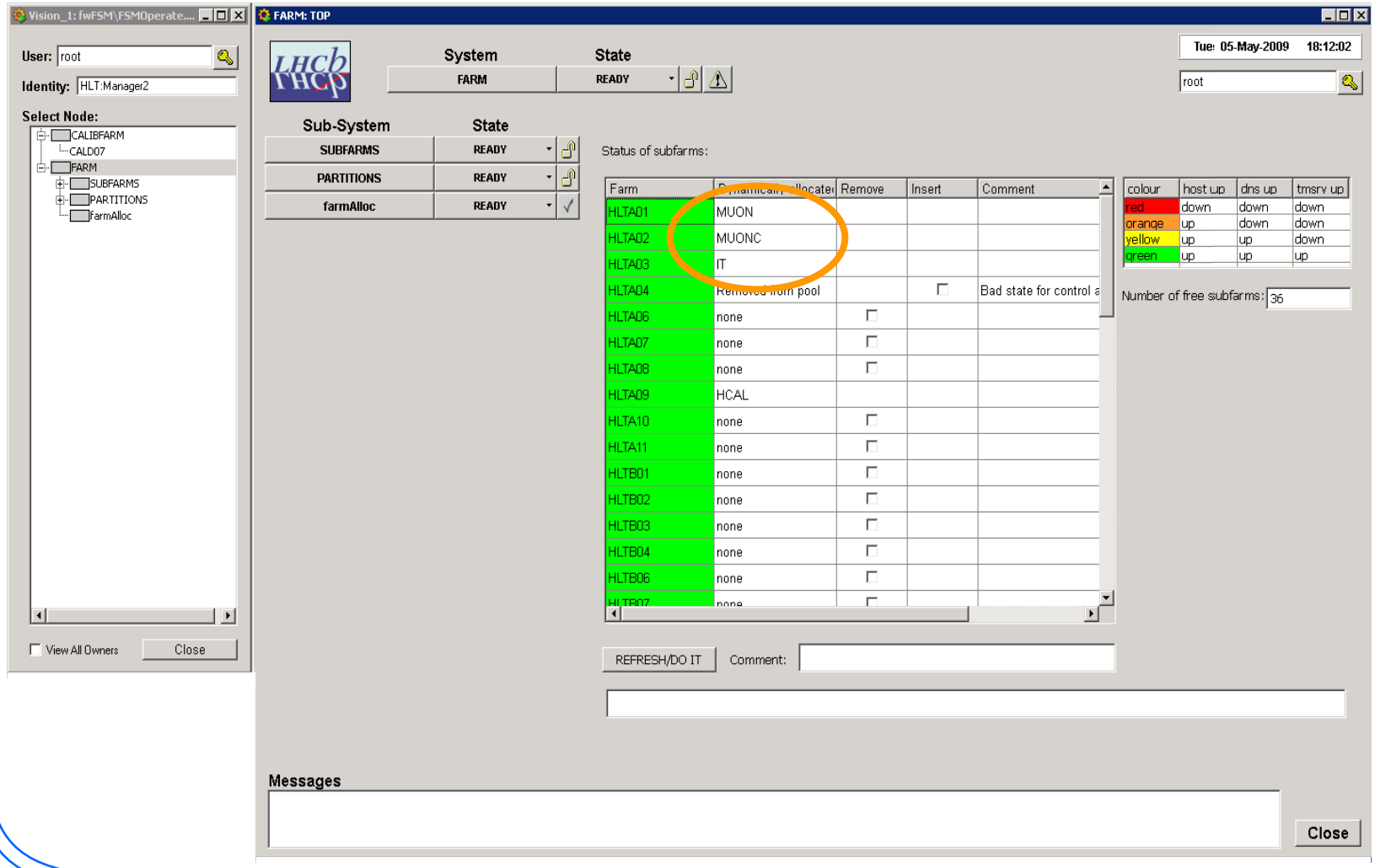
- Hierarchical ECS main control + dynamic allocation
  - Easy to include new farms into global system
- Single sub-farm controls as framework package + base on naming convention
  - Easy to duplicate farm control system.



Partitioning

- Several sub-detectors (many teams)
- Different possible configurations (commissioning stage).
  - Possibility to modify dynamically the readout components included to control.
  - Different running modes (Physics, Calibration, Cosmics, etc).
    - Run_type parameter sent with “Configure” command
- Pool of sub-farms → Dynamic allocation
  - Different instances of run control running in parallel (readout partitions)

Farm pool controls



Conclusions

- HLT control completely defined and integrated into global LHCb control system.
- Implemented with same toolkit (PVSS & FSM) used through the ECS.
  - Keeps Homogeneity
- Fully configurable at real time.
- Automatic control of processes for the shift operator.