Management of the LHCb Online Network Based on SCADA System

Guoming Liu[*†], Niko Neufeld[†]

* University of Ferrara, Italy
† CERN, Geneva, Switzerland

2009
ICALEPCS

- Introduction to LHCb Online system

- LHCb online network

- Network management based on SCADA system

- Summary

LHCb online system

❑ LHCb is one of the large particle physics experiments on LHC at CERN

❑ Online system is one of the infrastructures for LHCb, providing IT services for the entire experiment

❑ Three major components:

  ➢ Data Acquisition (DAQ)

    Transfers the event data from the detector front-end electronics to the permanent storage

  ➢ Timing and Fast Control (TFC)

    Provides fast clock and drives all stages of the data readout of the LHCb detector between the front-end electronics and the online processing farm

  ➢ Experiment Control System (ECS),

   Controls and monitors all parts of the experiment

LHCb online system

Guoming Liu

LHCb Online Network
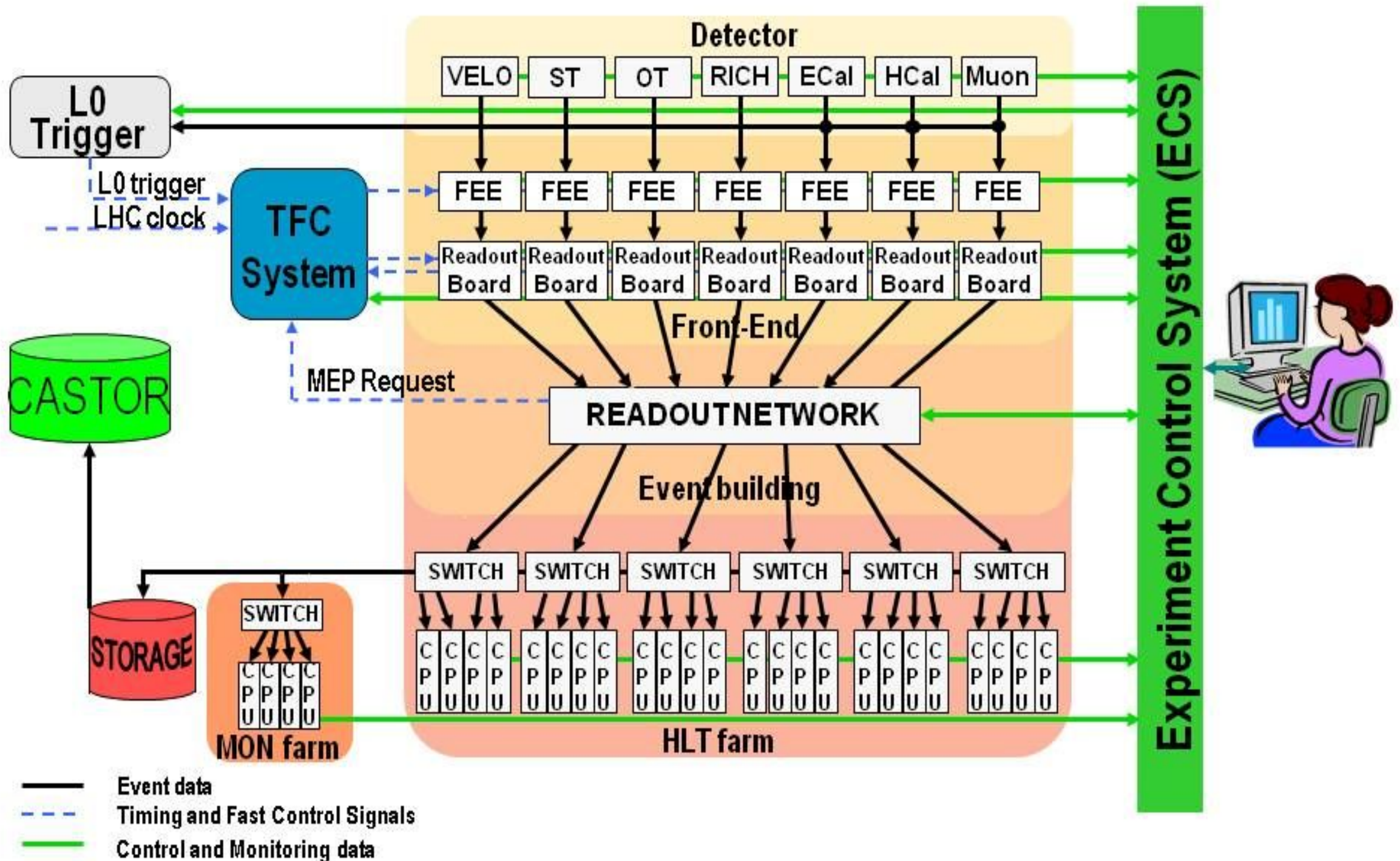
❑ Two dedicated networks:

 ➢ **Control network**: general purpose network for experiment control system

   Connects all the Ethernet devices in LHCb
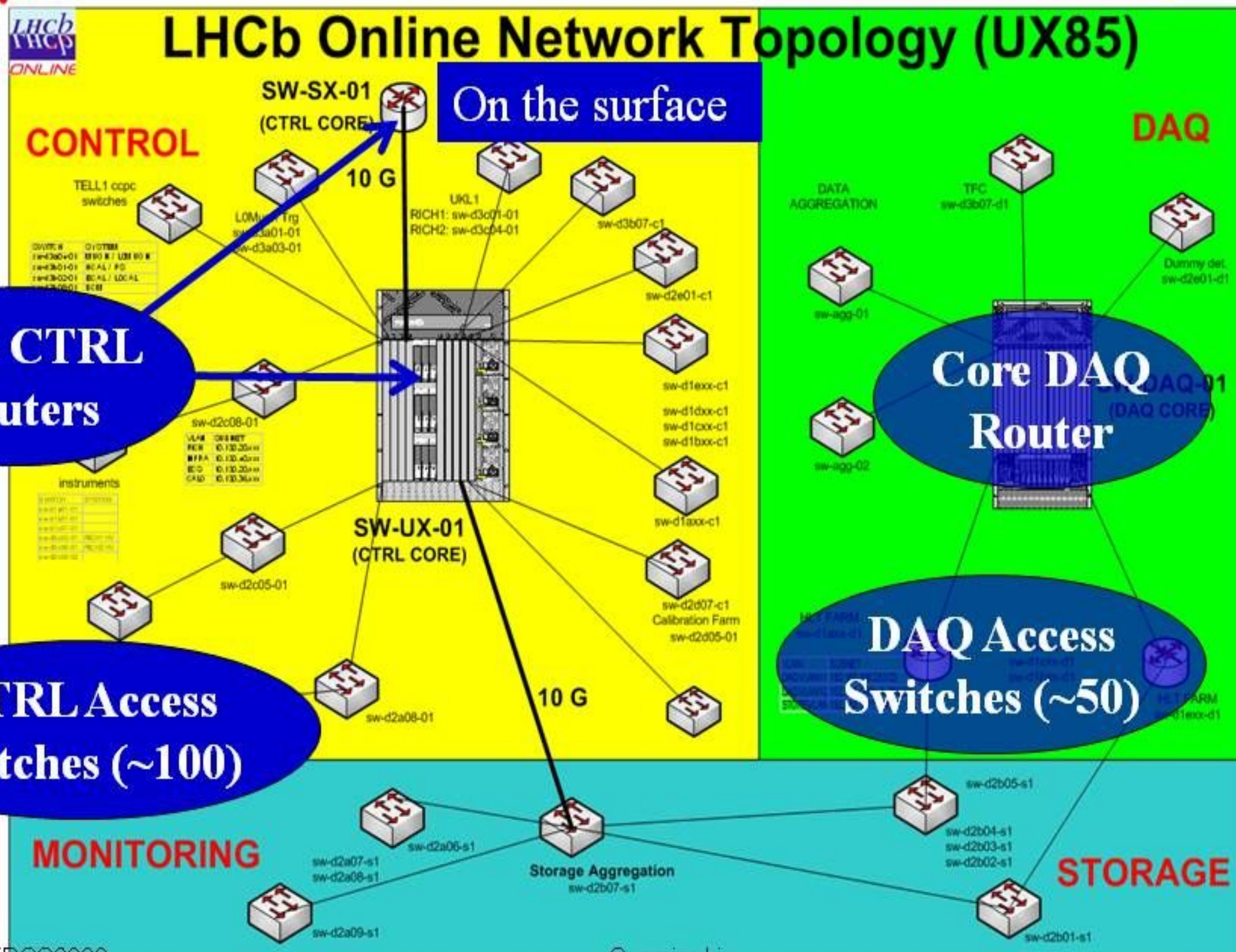
 ➢ **Data network**: dedicated to data acquisition

   Performance critical

LHCb Online Network

□ Two geographic parts: surface and underground
 Connected by two 10G links

Guoming Liu
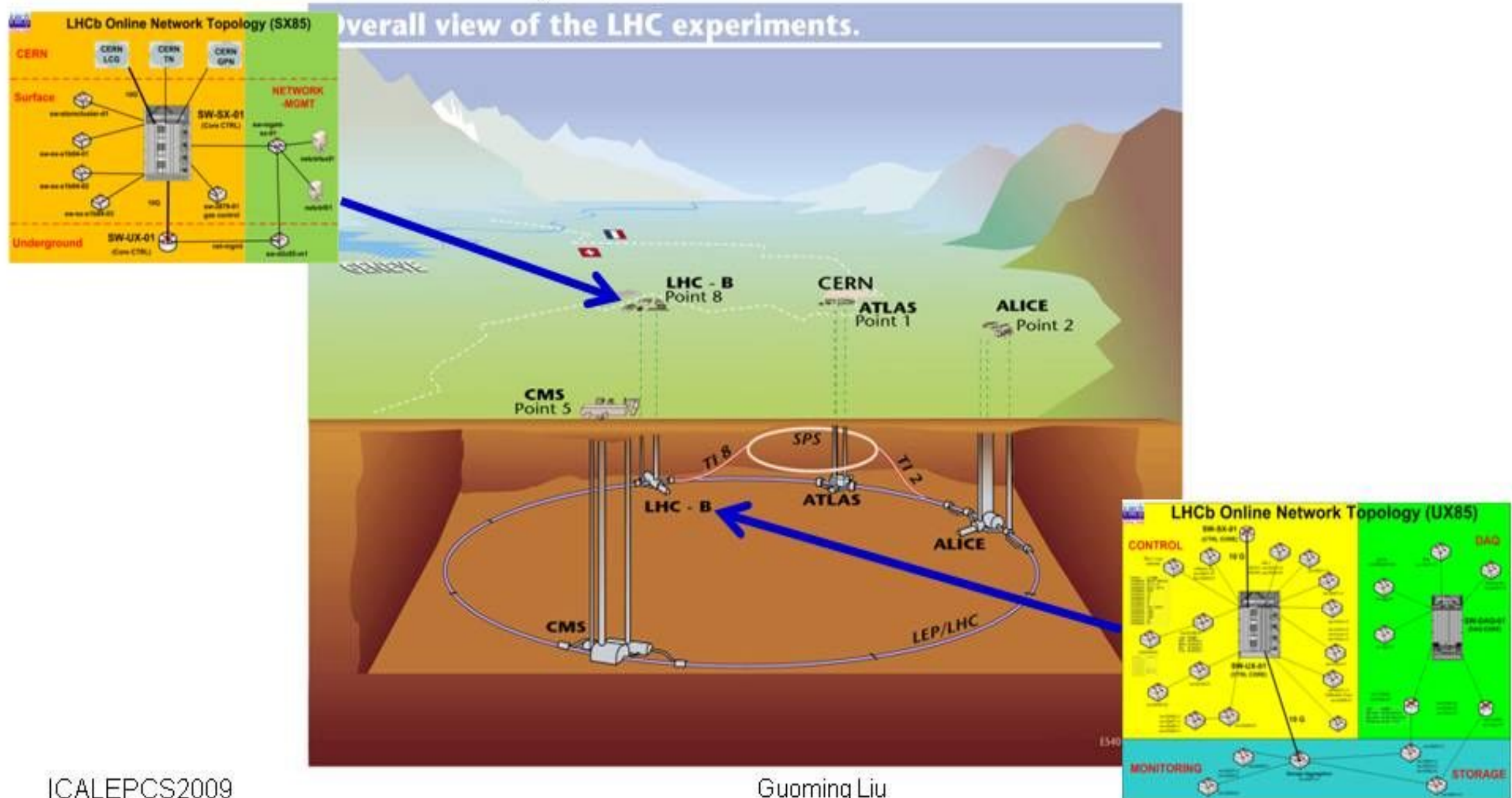
❑ Motivation

  ➤ This large network needs sophisticated monitoring
  ➤ Integration into LHCb ECS coherently
  ➤ Provides homogeneous interfaces for non-expert shift-crew

Commercial network management software?

  ➤ Expensive
  ➤ Integration?

- ❑ Supervisory layer
  - ➢ PVSS II: commercial SCADA system
  - ➢ JCOP: Joint Control Project for LHC experiments
- ❑ Front-end Processes:
  - ➢ SNMP
  - ➢ sFlow
  - ➢ syslog
- ❑ Data communication
  - ➢ DIM: Distributed Information Management
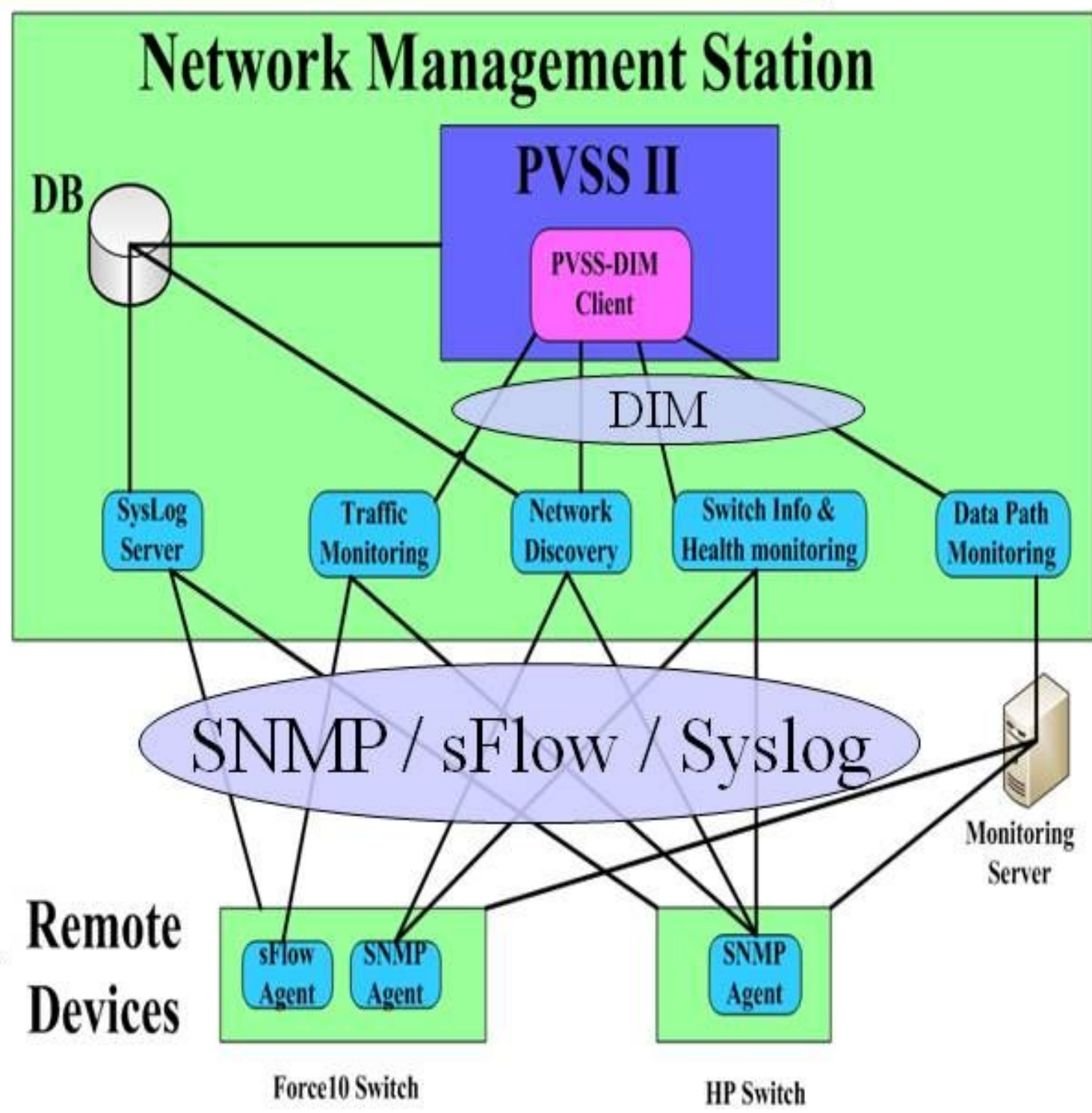
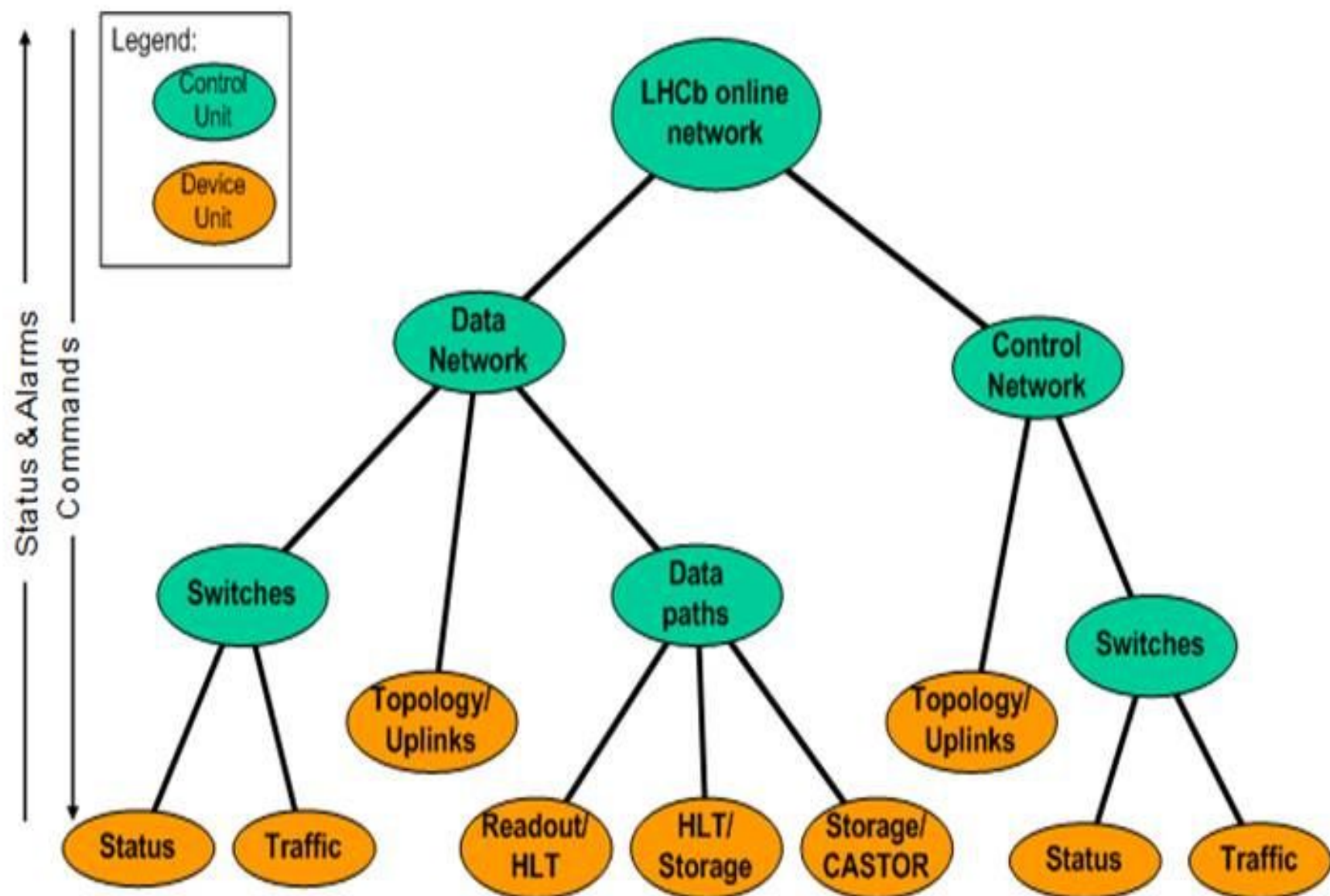❑ **All behaviors are modeled as Finite State Machines (FSM)**

❑ **Hierarchical structure**: status/command propagated

➢ **Device Units:**

  ✓ Device Description

  ✓ Device Access

  ✓ Based on PVSS II datapoint: Alarm Handling, Archiving, Trending etc.

➢ **Control Units**

  ✓ Abstract behavior modeling

  ✓ Represent the associated sub-tree

Network Monitoring System

The major items under monitor

❏ Physical topology

  ➢ Discovery of the network topology based on the Link Layer Discovery Protocol (LLDP)

  ➢ Discovery of the network nodes: based on the information in switches (ARP, MAC forwarding table)
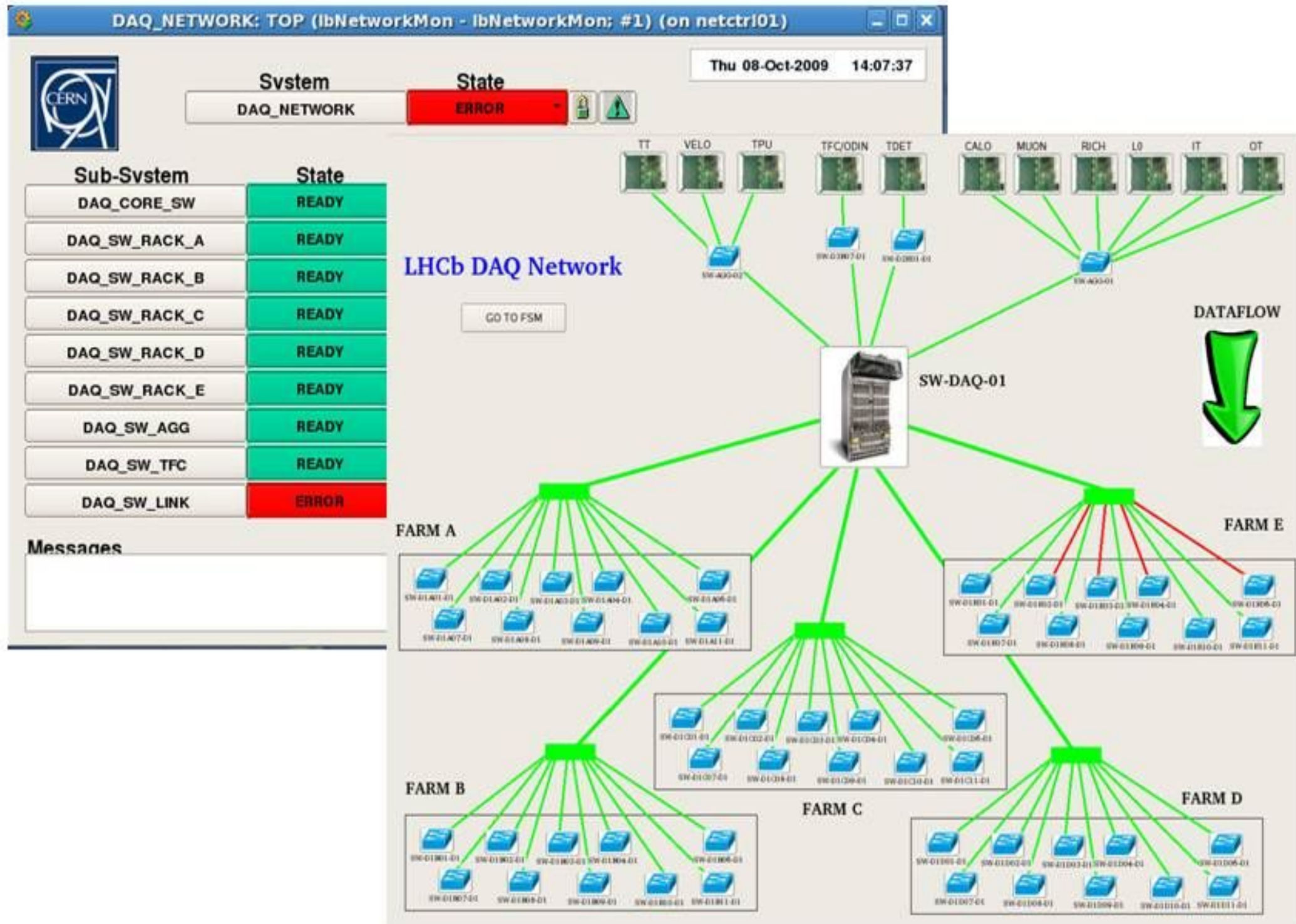
❏ Traffic

  ➢ Octet / packet counters

  ➢ Discard/Error counters

  ➢ . . .

❏ Switch status: CPU/Memory, temperature, power supply , . . .

❏ Data Paths for DAQ

Network Monitoring Snapshot(2): traffic

Guoming Liu

❑ The network management system has been implemented based on the commercial SCADA system PVSS II and the framework JCOP

❑ It provides sophisticated monitoring of the network which are essential for our operation, i.e. switch status, traffic

❑ It provides the homogenous operation interface and intuitive display as well

❑ Currently only monitoring is provided, some control commands of switches to be integrated

Thanks for your attention!

Backup

Guoming Liu

❑ SNMP: Simple network management protocol
Used for general network monitoring, configuring

❑ sFlow:
  ➢ A sampling mechanism to capture traffic data
  ➢ Based on hardware.
  ➢ Two kinds of sFlow samples: flow samples and counter samples.

Used on the core switch to collect traffic counters: SNMP too slow, and consumes high CPU/Memory

❑ Syslog: event notification messages
  ➢ Three distinct parts: priority, header and message.
  ➢ The priority part represents both the facility and severity of the message.

❑ Syslog can collect some information not covered by SNMP

❑ Syslog server is setup to receive the syslog messages from the network devices and parse the messages.

Alarm information:

➤ Hardware: temperature, fan status, power supply status

➤ System: CPU, memory, login authentication etc.

❑ All the messages with the priority higher than warning, will be sent to PVSS for further processing

❑ Monitoring the status of the routing using "ping"/"arping"

❑ Three stages for the DAQ:

1. From readout board to HLT farm

2. From HLT Farm to the LHCb online storage

3. From the online storage to CERN CASTOR

Detector

| VELO | ST | OT | RICH | ECal | HCal | Muon |

L0 Trigger

FEE ... FEE

Readout Board ... Readout Board

Front-End

READOUT NETWORK 1

Event building

SWITCH ... SWITCH

CASTOR

3

STORAGE

2

SWITCH

MON farm

HLT farm

—— Event data
- - - Timing and Fast Control Signals
—— Control and Monitoring data