Summary of TDAQ week

M. Medinnis
DESY/ATLAS
9.6.2006

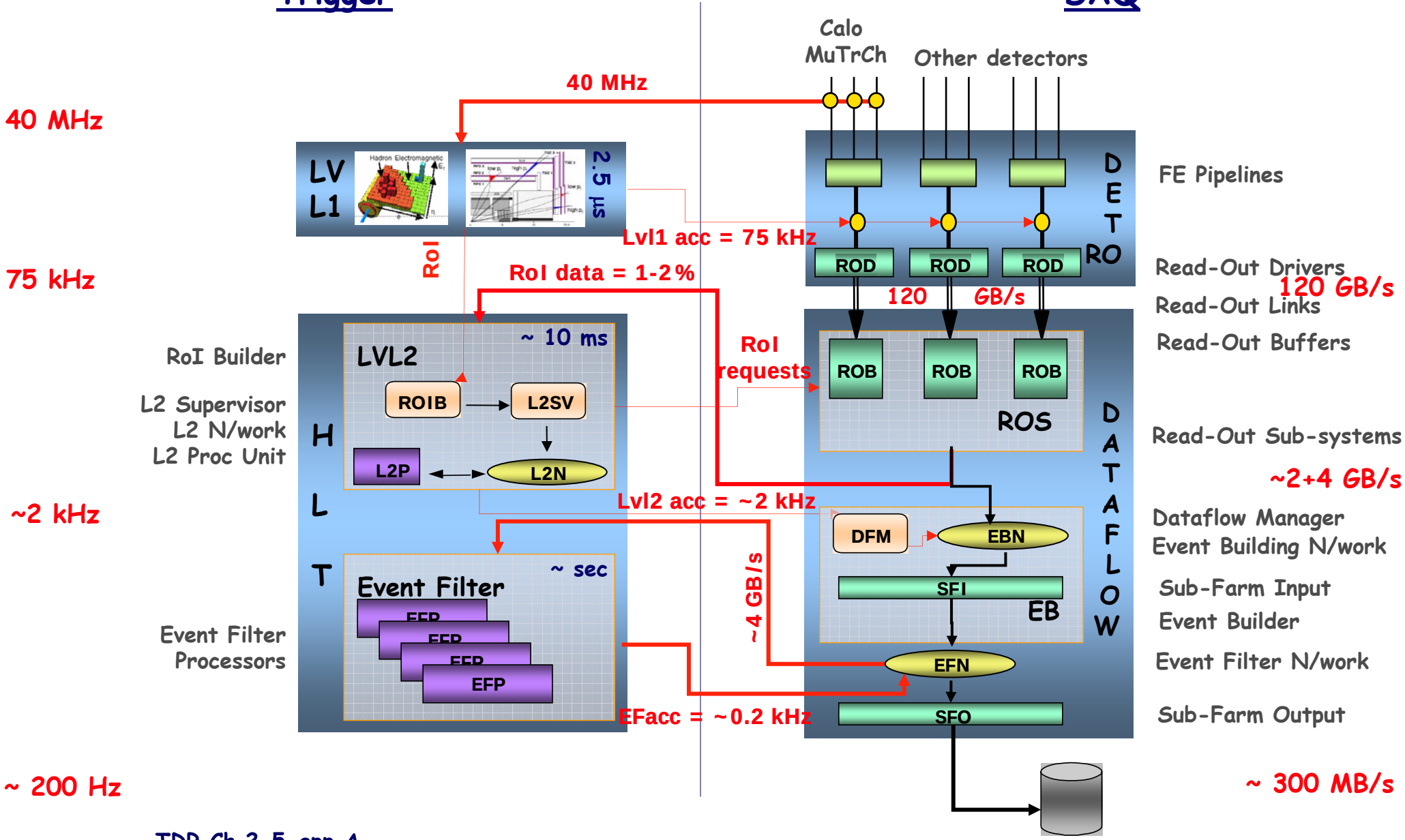
Contents

- Reminder of TDAQ structure
- General schedule
- DAQ infrastructure, control room
- L1 hardware
- Monitoring
- Global tests (pre-series & LST)
- Data flow & run structure issues

Lots more material was presented which I won't mention:
PESA (see Martin's talk), HLT steering, Trigger configuration,
Software management, Networking

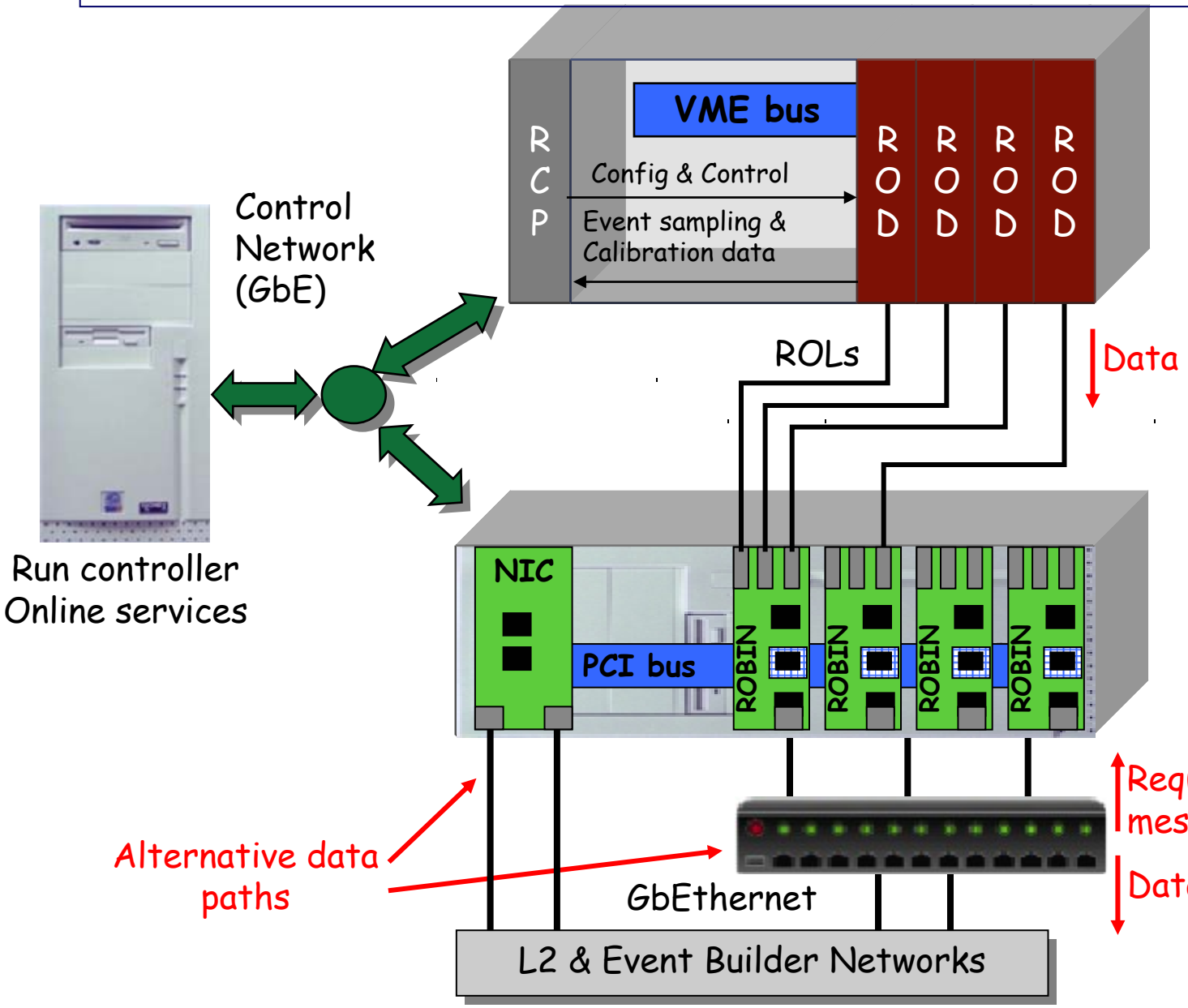
ARCHITECTURE

(Functional elements and their connections)



TDR Ch.2-5-app.A

ROD crate and ROS - Full system underground (USA15)



Read Out Driver Crates
(86 in final system)
Full DAQ functionality
--> can be used standalone

Read Out Links
(1600 in final system)
3 ROLs/ROB
--> can adjust for data load balancing

Read Out Systems
(150 in final system)
ROD-ROS association
--> full detector partitioning

I/O path
(PCI or N/W)
Main path is PCI
GbE as upgrade option

Schedule

Installation Milestones

From G. Mornacchi (tech. management)

on the global ATLAS installation schedule.

- Barrel calorimeters in final position
- EC calorimeters mechanically completed by June
- Muon Barrel continuing up to Fall 06
- ID barrel (SCT+TRT) in UX15 in July
- Barrel Toroid tests July-August
- Solenoid field mapping end June

Each sub-system and activity to present status and discuss evolution plan to meet ATLAS and TDAQ installation schedules.

- **Example of steps in global ATLAS schedule (to be refined):**

1. Combined calorimeter cosmic run — *July 06*
2. Installation of Muon-Barrel Read-Out electronics — *Late summer 06*
3. Installation of Inner Detector Read-Out electronics — *Autumn 06*
4. Full Barrel integration — *~End 06*
5. Completion of installation of Read-Out electronics (End Caps) — *Spring 07*
6. Data taking with complete detector — *July 07*

Plenary sessions

Monday pm
- Schedule / Point-1

Tuesday am
- LVL1 / DAQ-HLTI

Wednesday am
- DAQ-HLTI

Thursday am
- HLT

Friday am
- Global issues / Wrap-up

Parallel sessions

Tuesday pm
- LVL2 review / Pre-series
- LVL1-RoIB / LST / TDRC

Wednesday pm
- PESA Algos / Network&Controls
- ROB-ID / Conf&Transitions
- TDIB
  - *BBQ*

Thursday pm
- HLT Conf&Steer / MDT calib
- Run structure

DAQ infrastructure, control room

DAQ/HLT-I installation schedule

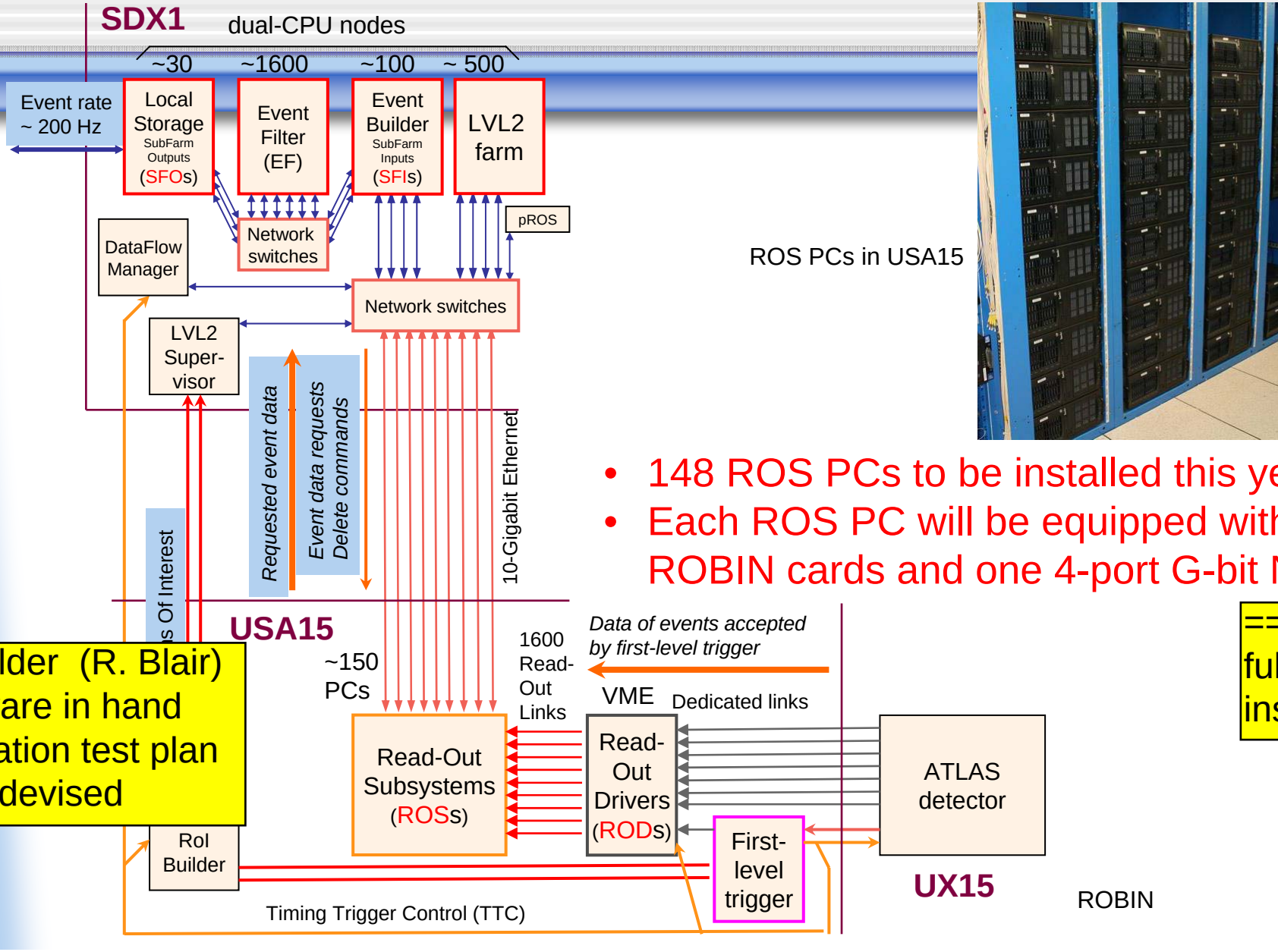
From D. Francis

on DAQ infrastructure installation

- In today's picture:
  - June (End): SDX1 upper level ready
    - Racks in final position & power (critical point for SDX1 activities)
    - Cooling later (July)
  - End June: Start cable installation in SDX1 upper level
  - July (Mid): ROS installation complete
  - July (End): Start installation of 32 SFIs, 12 DFMs, 2 L2SVs & 1pROS
  - Sept. (End): ROS standalone commissioning complete
  - Oct.: slice of DAQ/HLT-I system
    - 153 ROSs; 32 SFIs, 12 DFMs, 2 L2SVs, 1pROS, associated switches, Online infrastructure
    - Use Pre-series L2 and EF racks
  - Jan. '07: 1 L2 rack & 3 EF racks
- Caveat
  - Picture driven by purchasing schedules
  - Picture does not take into account Holidays or in general man power

From B. Gorini
on readout system (ROS) status

Overview



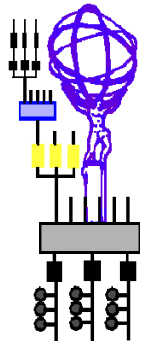
ROS PCs in USA15

- 148 ROS PCs to be installed this year
- Each ROS PC will be equipped with 1-4 ROBIN cards and one 4-port G-bit NIC

==> essentially the full system will be installed this year

RoI Builder (R. Blair)
- hardware in hand
- integration test plan being devised

Location & amount of Online Infrastructure at P1



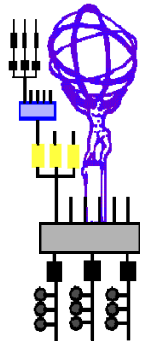
- *4 racks as of 2006 all in SDX1*
- *At present:*
  - 5 online machines delivered, to be setup for commissioning
  - 4 LFS delivered, to be setup for commissioning
- *Plan foresees to buy all Online Infrastructure HW in 2006 (20 + 32 machines)*
- *The present Online rack foresees:*
  - 5 Online PCs
  - 8 Monitoring PCs
  - 2 CFS (4 U; we don't need 8 CFS)
  - 1 LFS (2 U)
  - 1 control switch
- *Technical Specifications:*
  - Draft exists for "Monitoring PC" and LFS
  - They are similar and compatible

From G. Lehman on Online Infrastructure

How much hardware is needed to provide DAQ services -- specifically run control, data base servers, monitoring, error logging, messaging service ...

Allocation of applications to processors.

Monitoring Capability: Online PC



- ***Baseline specification for the Monitoring PC (is being developed):***
  - **Dual core dual-processor PIV CPU ~3GHz**
  - **4 to 8 GB of memory**
  - **300 GB disk space**
  - **Gigabit Ethernet link**
- ***One such PC is able to handle:***
  - **Up to 3 Gbytes of monitoring data in total**
  - **Up to 3000 monitoring data providers**
  - **From 2 to 10 KHz of overall update frequency**

Example

- ***2000 applications, producing 10 histograms each with the average histogram size of 10 Kbytes and update frequency 0.1 Hz:***
  - **1 PC is sufficient**
- ***200 applications, producing 100 histograms each with histogram size of 100 bytes and with update frequency 1 Hz***
  - **2 PC are required (overall update rate is 20 KHz)**

From A. Nisati
TDIB - TDRC Wrap up

TDIB == Trigger DAQ Institute Board
TDRC == Trigger DAQ Resources Committee

TDRC - 16 May 2005 - cont'd

- All UTP cables for 2007 ordered, Fibres out for quote
- Switch Procurements (Chassis + Pizza Boxes)
  - Plan finalised
  - Shopping list produced for 2006 being checked, then order
- HLT Processors
  - Specification for 2006 purchase drafted, now being tuned
  - Start tender by end June for ~120 PC's (+ Monitoring?) + options for ~200 with delivery in November
- Proposed schedule for other HLT Processor orders
  - ~120 PC's early in 2007 (to be purchased by Japan)
  - In early 2007 buy any more needed for first LHC running
  - Buy bulk of HLT PC's late in 2007 to be ready for running in 2008
- Recalled need for effort for Installation + Commisioning
  - Estimate of effort requirements to be produced

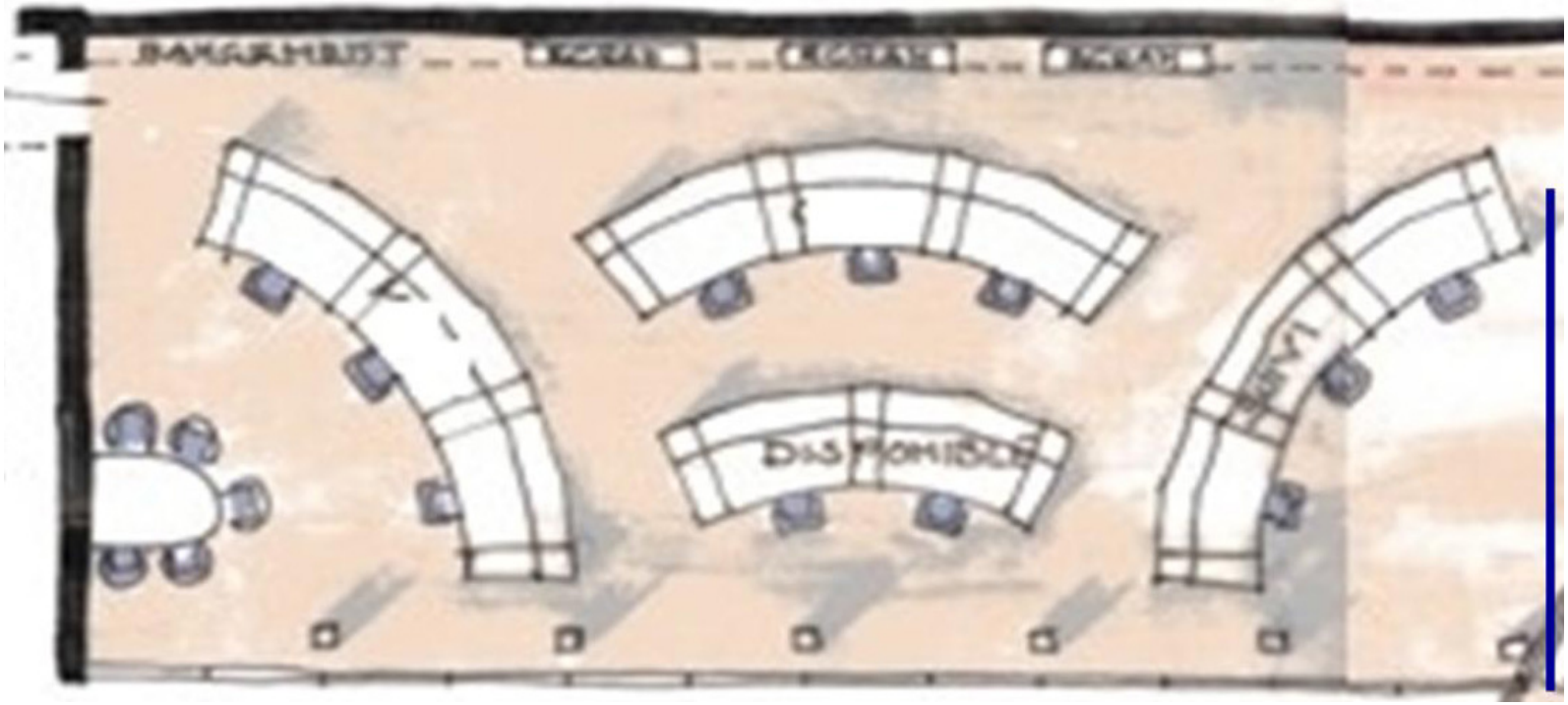
From Armen Vartapetian on the control room

physical layout
desk top environment

Introduction

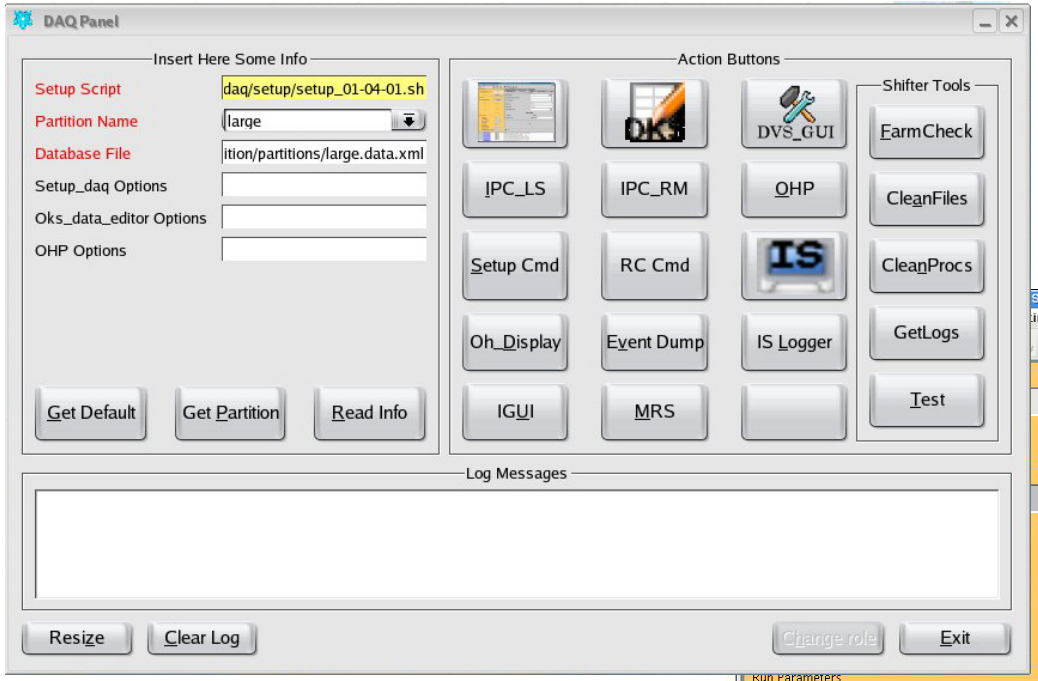
- We are at the initial phase of the ATLAS Control Room (ACR) design and construction
- First stage: extensive testing, gaining experience with a small scale prototype
- Exercise with installation and running of the available applications for the ACR desktop
- Use the experience with the prototype to continue development of the hardware towards the final setup
- Discussions on Control Room (CR) model – finalize our view on different aspects of the ACR design and operation

ACR Layout

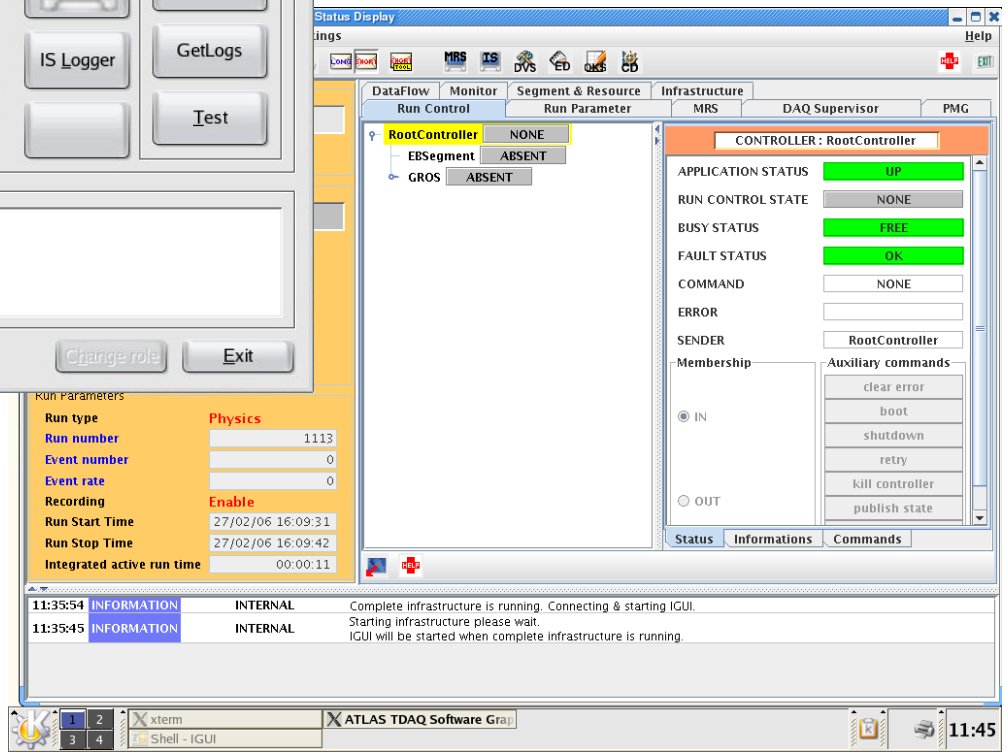


- Proposal of the ACR layout
- 11 stations with 4 and more screens each
- Several large wall displays

DAQ Applications



All the DAQ services are consolidated inside a single panel (DAQ panel)

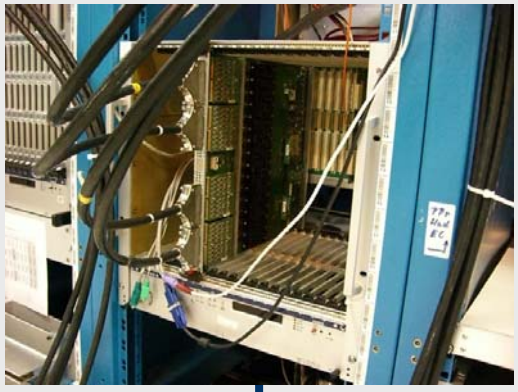


Place to start the run control window

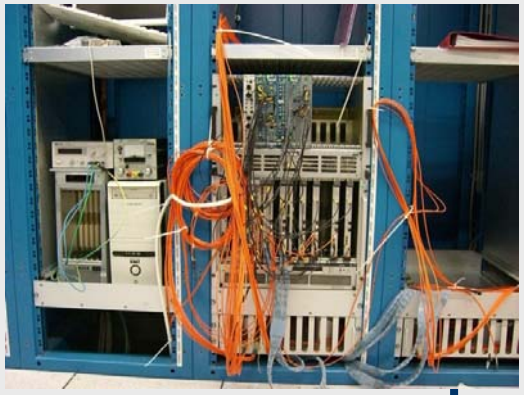
L1 hardware

Status

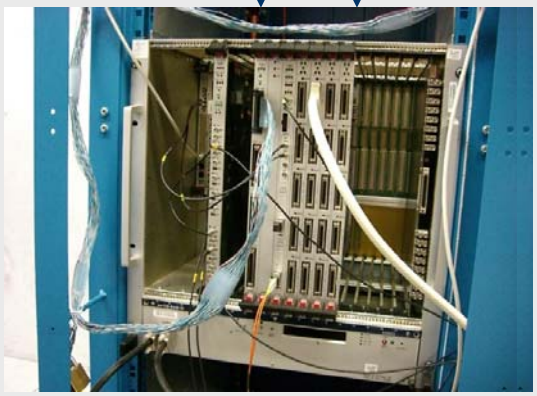
Calorimeter trigger



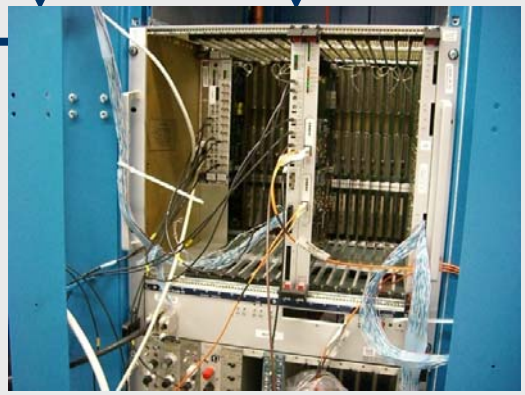
Muon Barrel trigger



Muon Endcap trigger



CTP



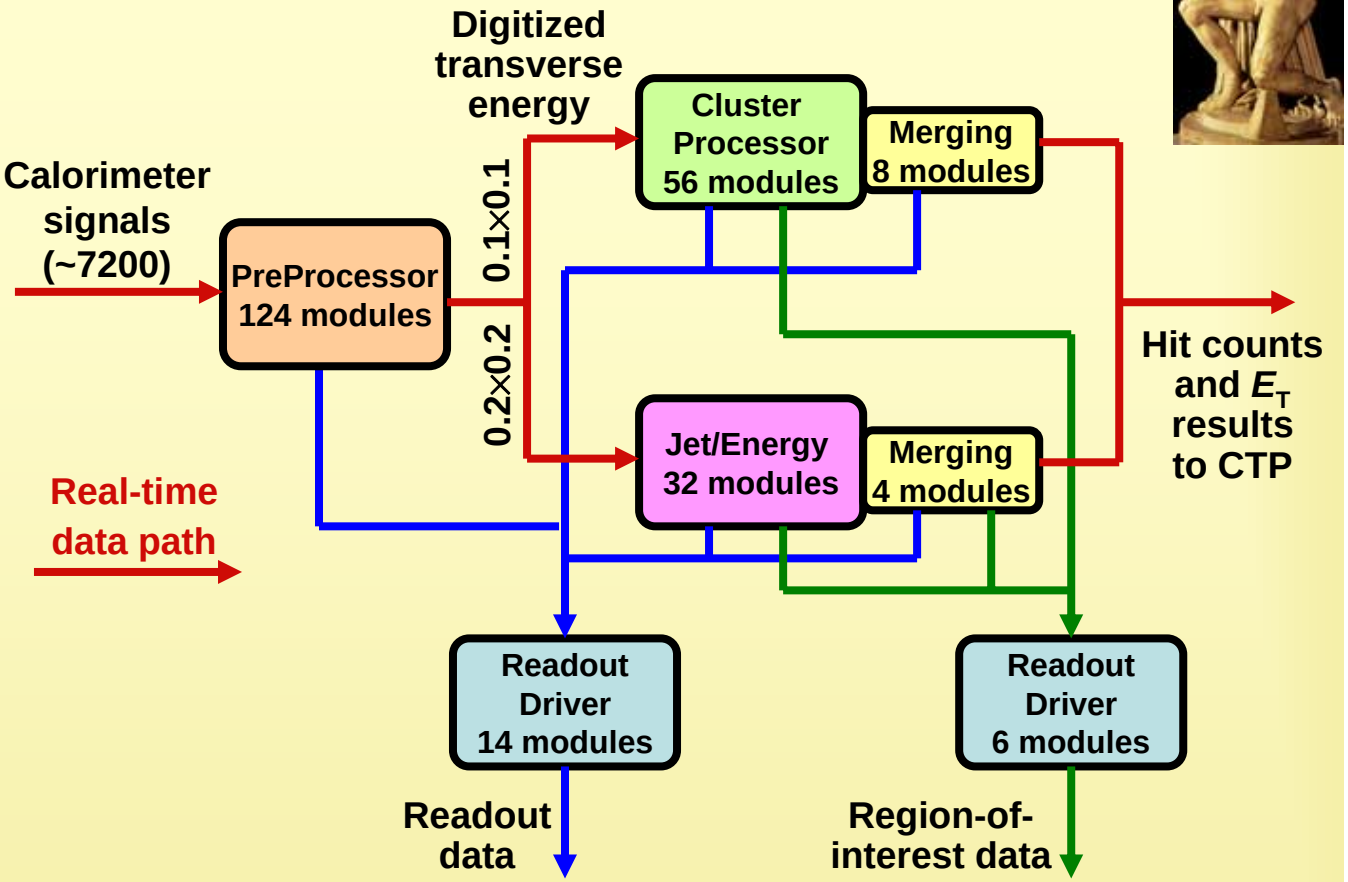
MuCTPI

From E. Eisenhandler
on L1Calo Trigger status

Overview of the trigger system

Features

- Real-time data path: <1 μs fixed-latency pipeline
- Many processing stages
- Massive parallelism: >300 Gbyte/s
- Multi-purpose modules (CMM, ROD, backplane)
- Heavily based on FPGAs



Five main types of custom 9U modules

PPM	CPM	JEM	CMM	ROD
Heidelberg	*UK*	*Mainz/ Stockholm*	*UK*	*UK*

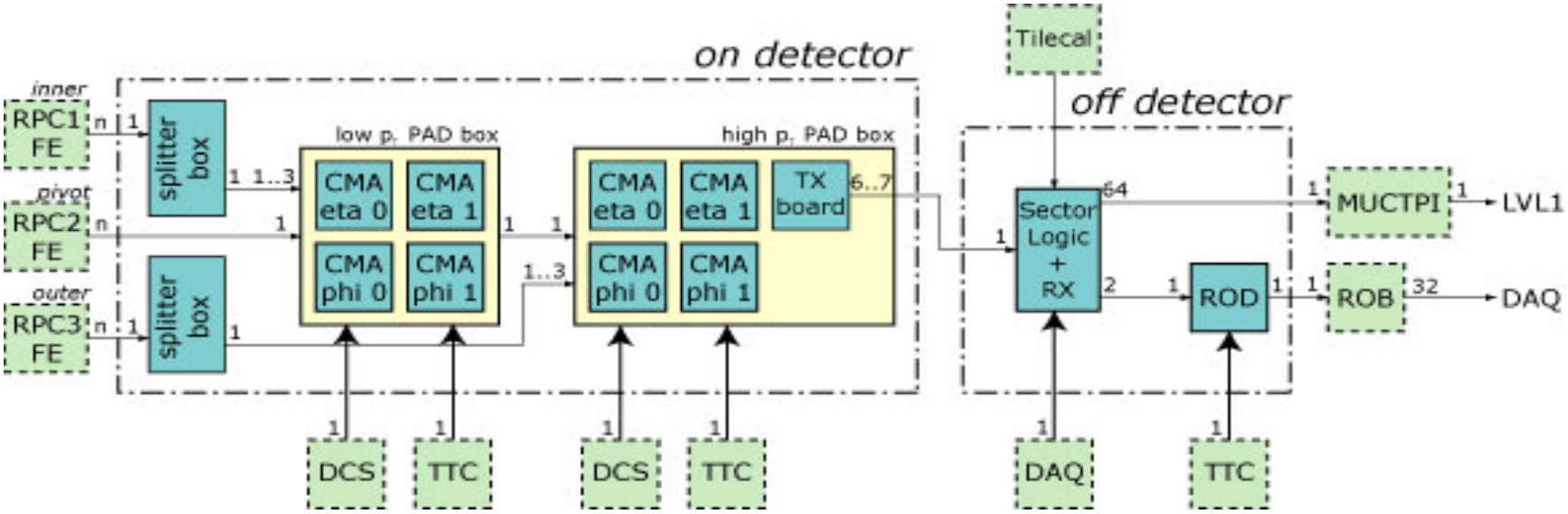
L1Calo Trigger status

Total required

- PPM (160) – prototyping – time table critical
- CPM (65) – production underway
- JEM(21) – in preproduction
- CMM (21) – production underway
- ROD (26) – testing prototypes, begin production in July – time table critical

from F. Pastore
on L1 Muon trigger

LVL1 muon in the BARREL

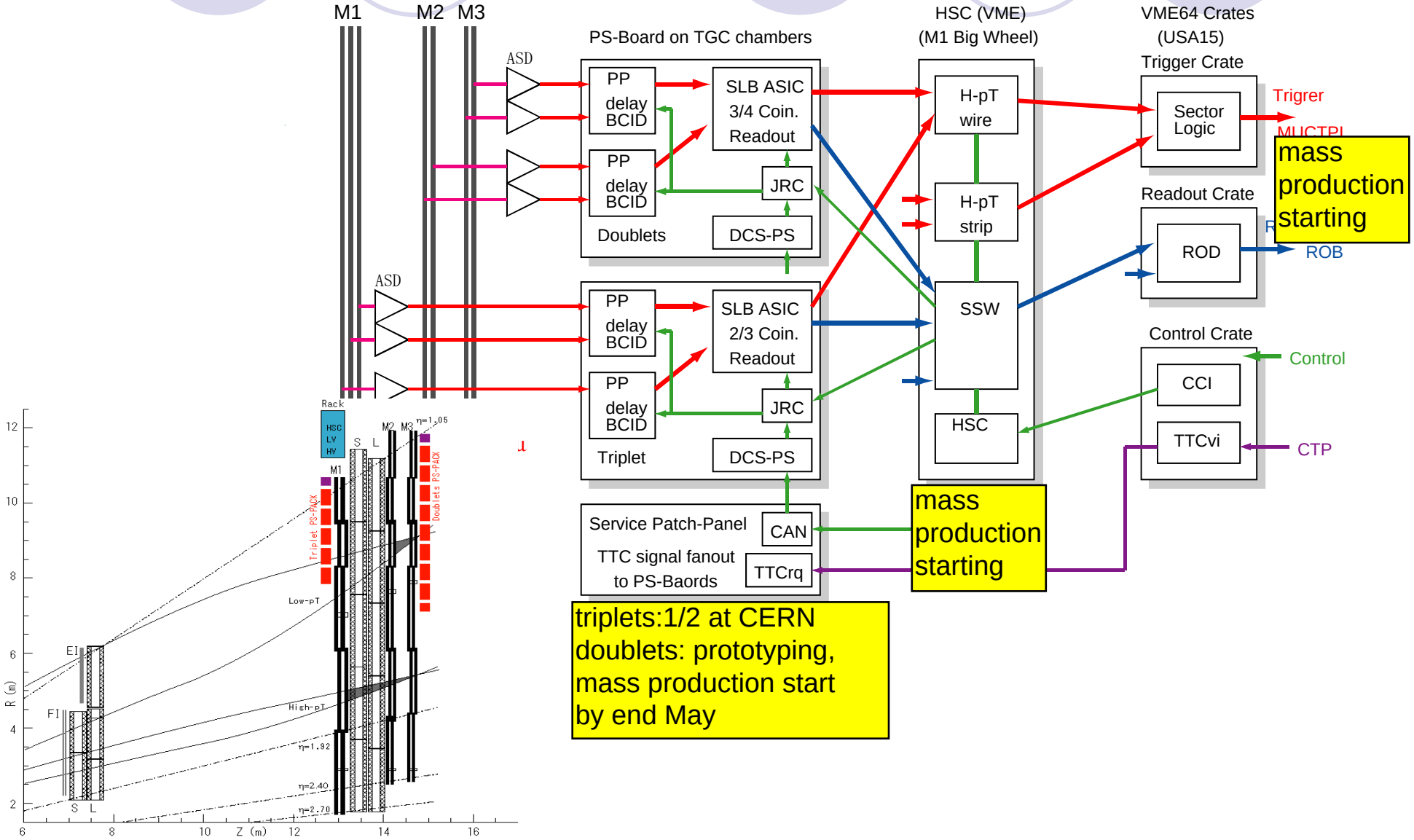


- On-detector electronics fully produced
  - Test and Installation on chambers is smoothly going on
- Off- detector electronics
  - RX/SL board: first prototype is under test in Rome
  - ROD board: integration tests are starting next June

Cosmic tests will use RX/SL demo boards, and ROD emulators (r/o via VME).
Production of RX/SL and ROD planned to finish in Spring 2007.

from F. Pastore
on L1 Muon trigger

LVL1 Muon trigger in the Endcaps



From S. Ask
on L1 Central trigger

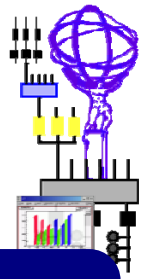
- The LVL1 central trigger work is progressing in USA15, without significant problems
- Starting ROS / HLT integration next, followed by setting up the cosmic ray trigger from the RPCs
- In a couple of weeks L1A will start to be distributed
- **Cosmic events triggered by a full ATLAS trigger chain might appear soon !**

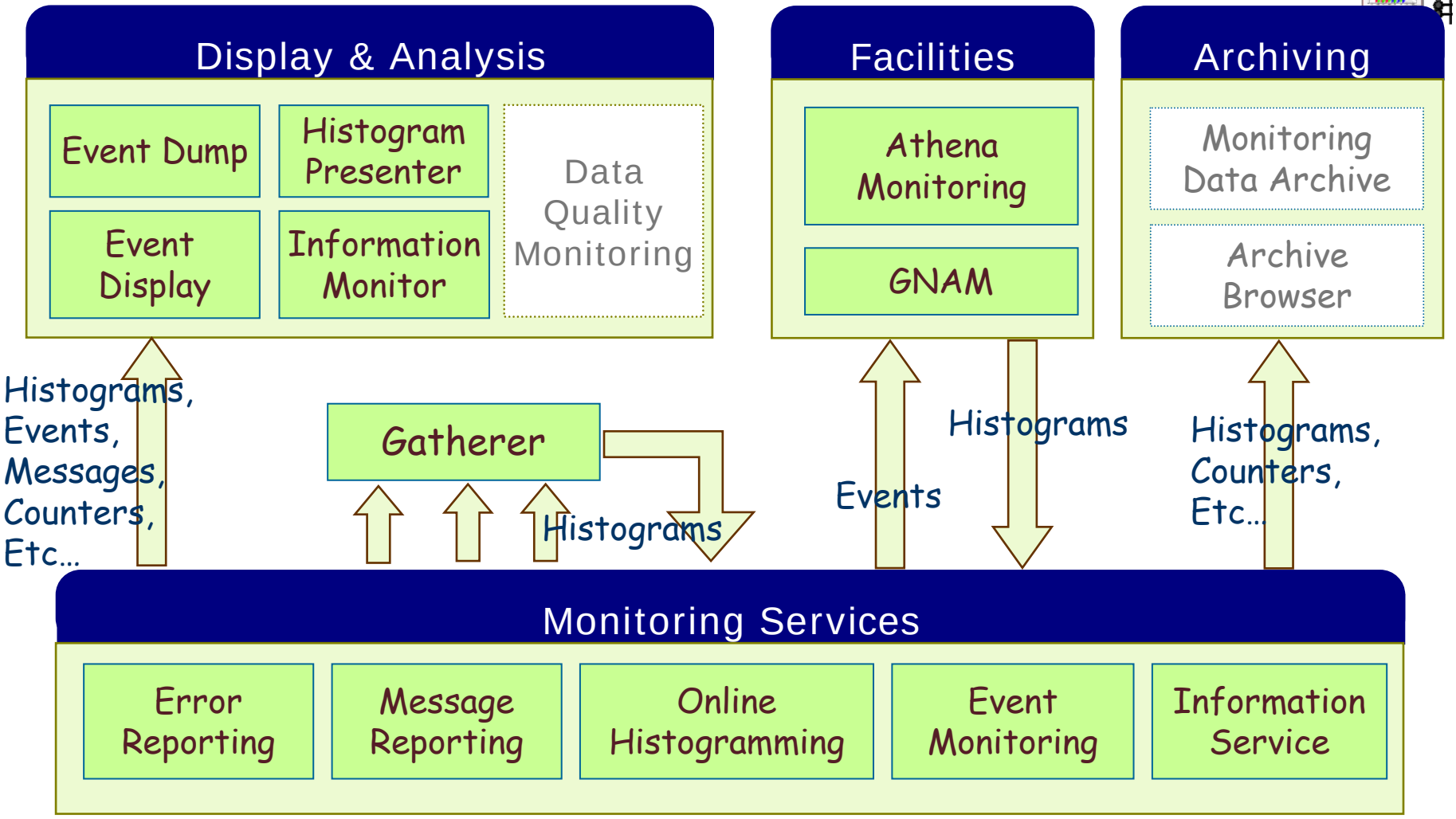
CTP: final version operational
MUCTPI: demonstrator version operational (sufficient for cosmics.) final board in autumn
TTC: most components available (except an i/f module -- prototype forseen for July)

Monitoring

From S. Kolos on monitoring

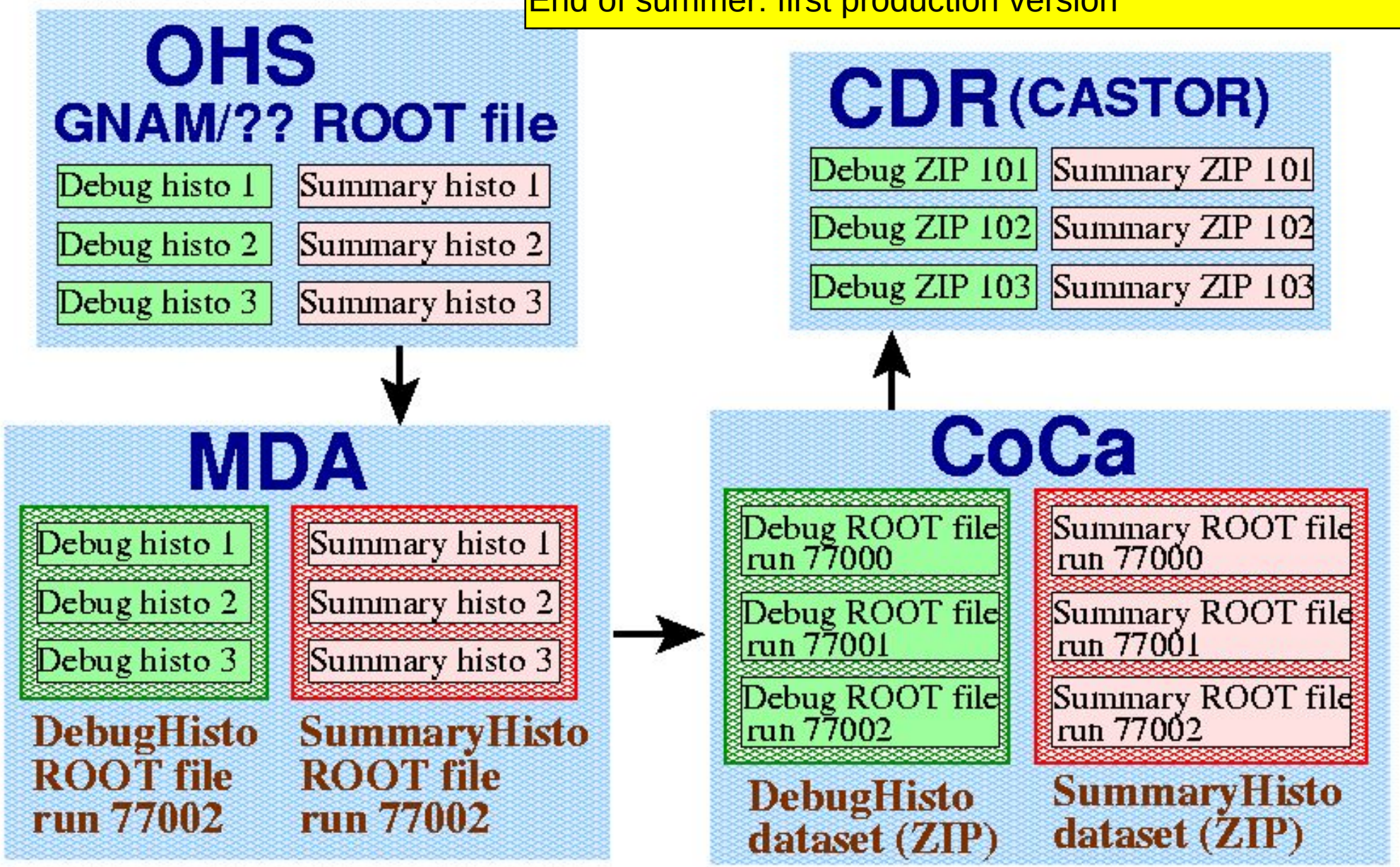
Monitoring Framework





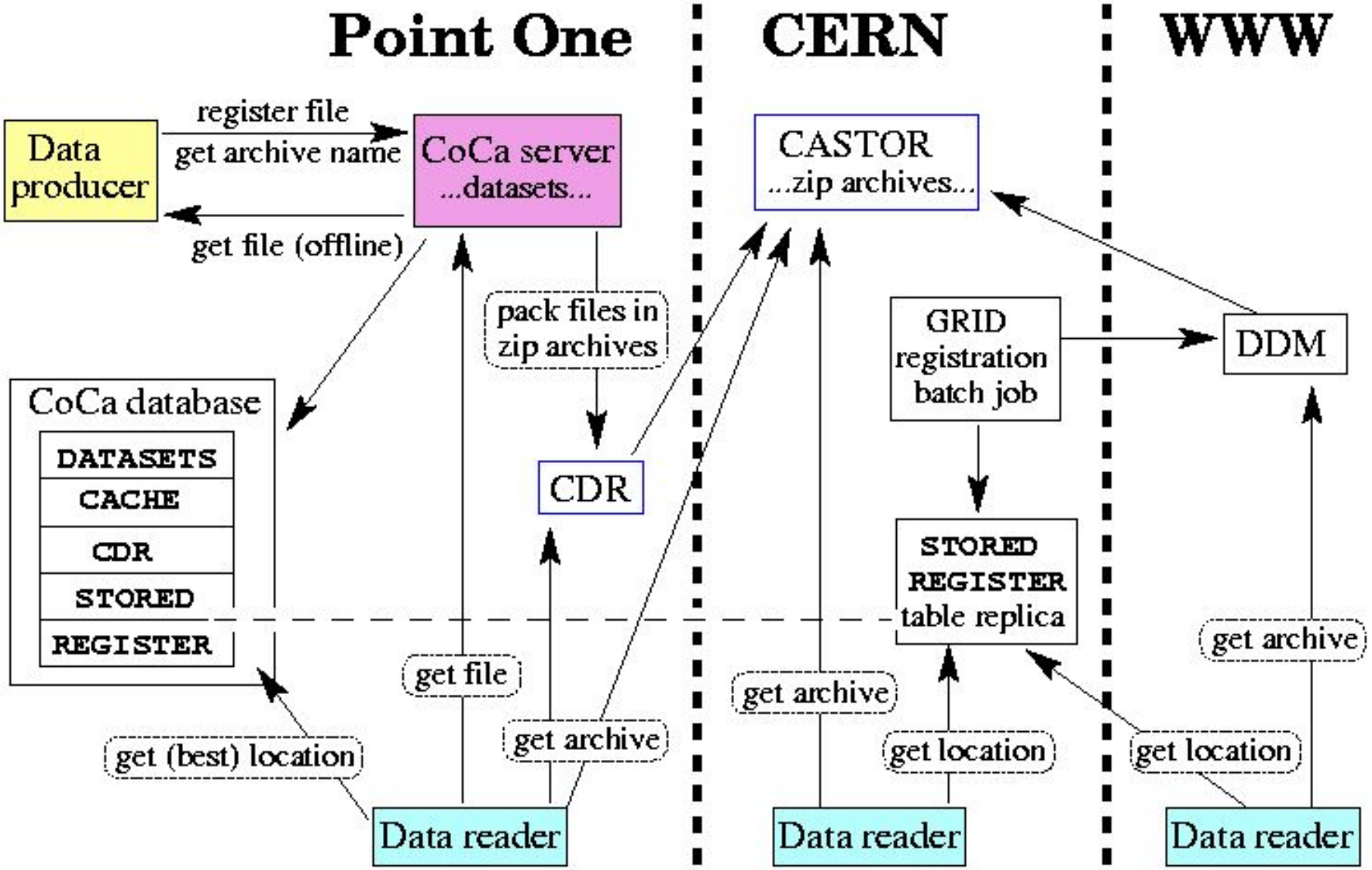
MDA and CoC

F. Zema
Monitoring and Data Archiving (MDA)
COllection and CAche (COCA)
Status: design completed (apart from some details)
End of summer: first production version



CoCa diagram

Access modes forseen for users at the experiment, CERN and via the web

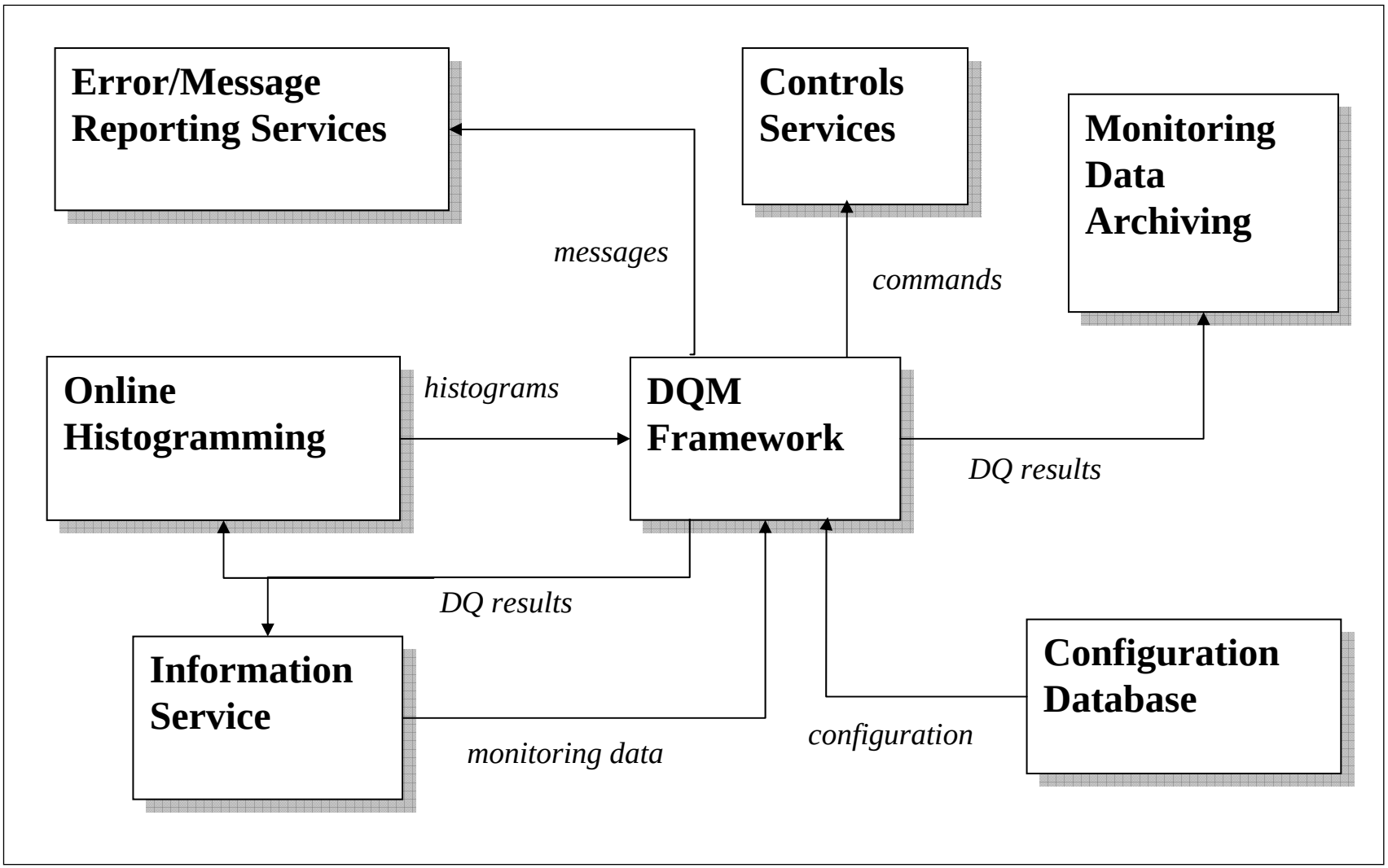


Being designed (Lisbon): oNline Object extended Database browsEr (NODE) for viewing histos

DQM

A. Corso-Radu: Data quality monitoring
- analyze monitoring info, assess data quality
- issue alarms, messages, emails to shifters when problems found
- to be used online and at Tier-0

Plan to have "prototype-0" in the autumn TDAQ release



Global tests (pre-series & LST)

pre-series exploitation

Talk by Gokhan Unel

Large effort involving modeling, testing on a slice, extrapolating to full system.

Includes long-term (e.g. 36 hour) runs with "realistic" shift crews

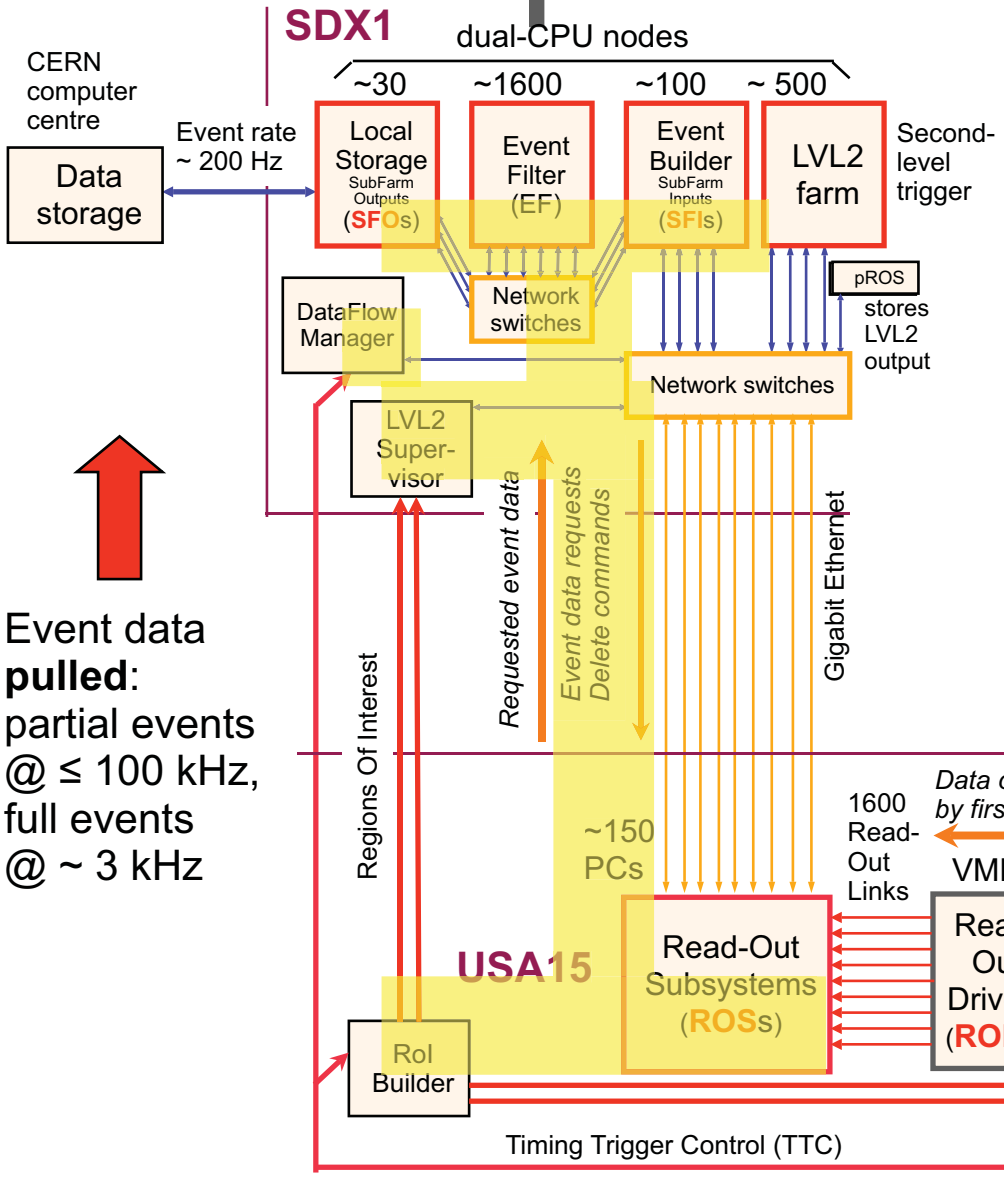
Some bugs discovered and fixed.

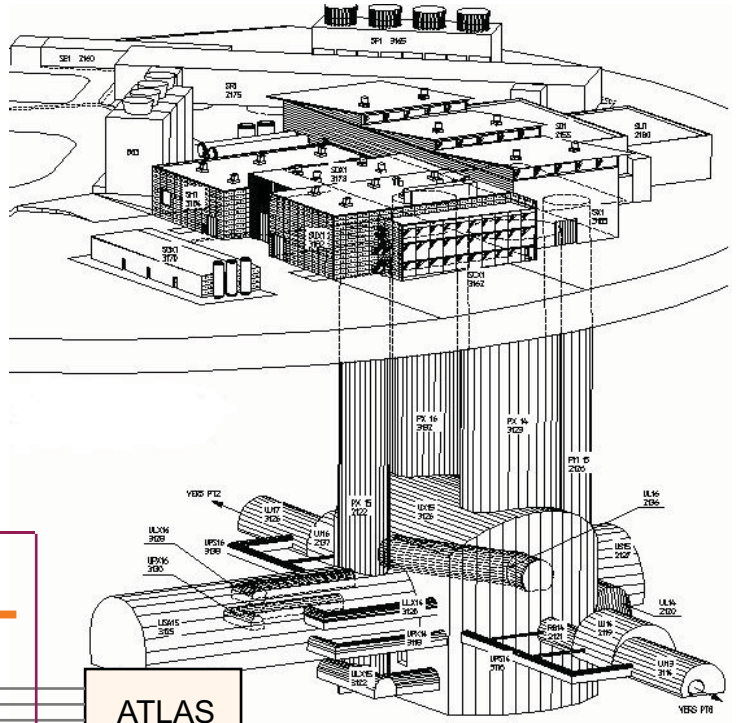
a short

by pre-series enthusiasts:

Alina, Andre, Andrea, David, Erden, Giovanna, Giuseppe, Gokhan, Hegoi, Imma, Jim, Kostas, Kris, Louis, Marc, Per, Sarah, Serge ...

The pre-series test bed



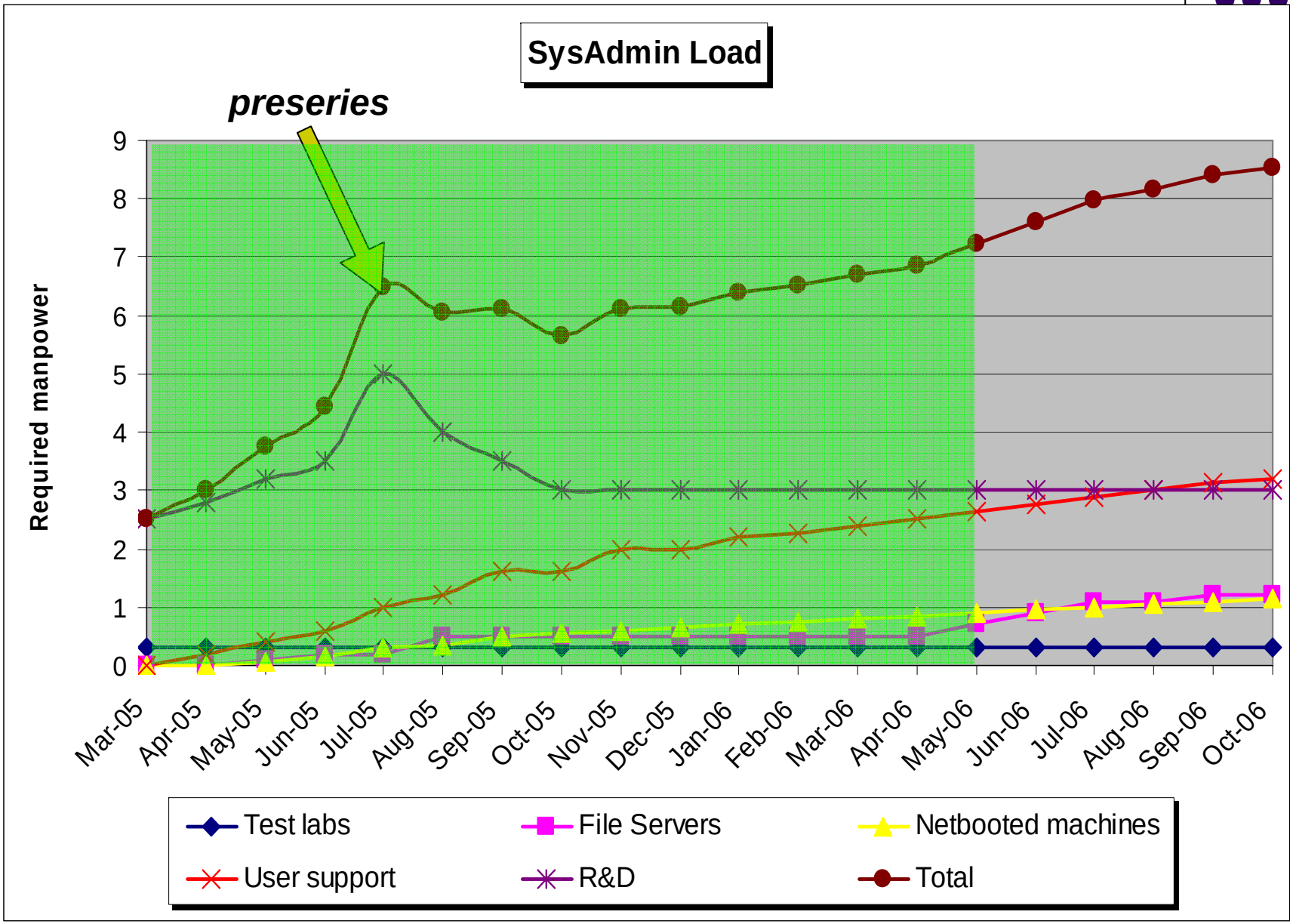


Event data **pulled**:
partial events
@ ≤ 100 kHz,
full events
@ ~ 3 kHz

Event data **pushed** @ ≤ 100 kHz,
1600 fragments of ~ 1 kByte each

a vertical slice: pre-series

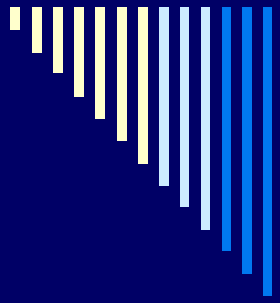
From Marius Leahu
on system administration



From D. Burckhart Chromek on Large scale test plans

Large Scale Tests 2006 (LST06)

- CERN IT LXBATCH farm
- Agreement between Atlas & IT (CERN) – request for up to (O)1200 nodes granted
  - **Farm size growing in steps: 400, 800, 1200**
- 4 weeks, Q4 2006
  - ~ November 2006 for the tests on the farm
  - **Last LST before experiment start-up**
- >= 1GBytes RAM, >= 2.4 GHz dual nodes
- IT support:
  - Operating system installation
  - DAQ/HLT-I distribution via RPMs
  - Farm monitoring
  - Oracle Server

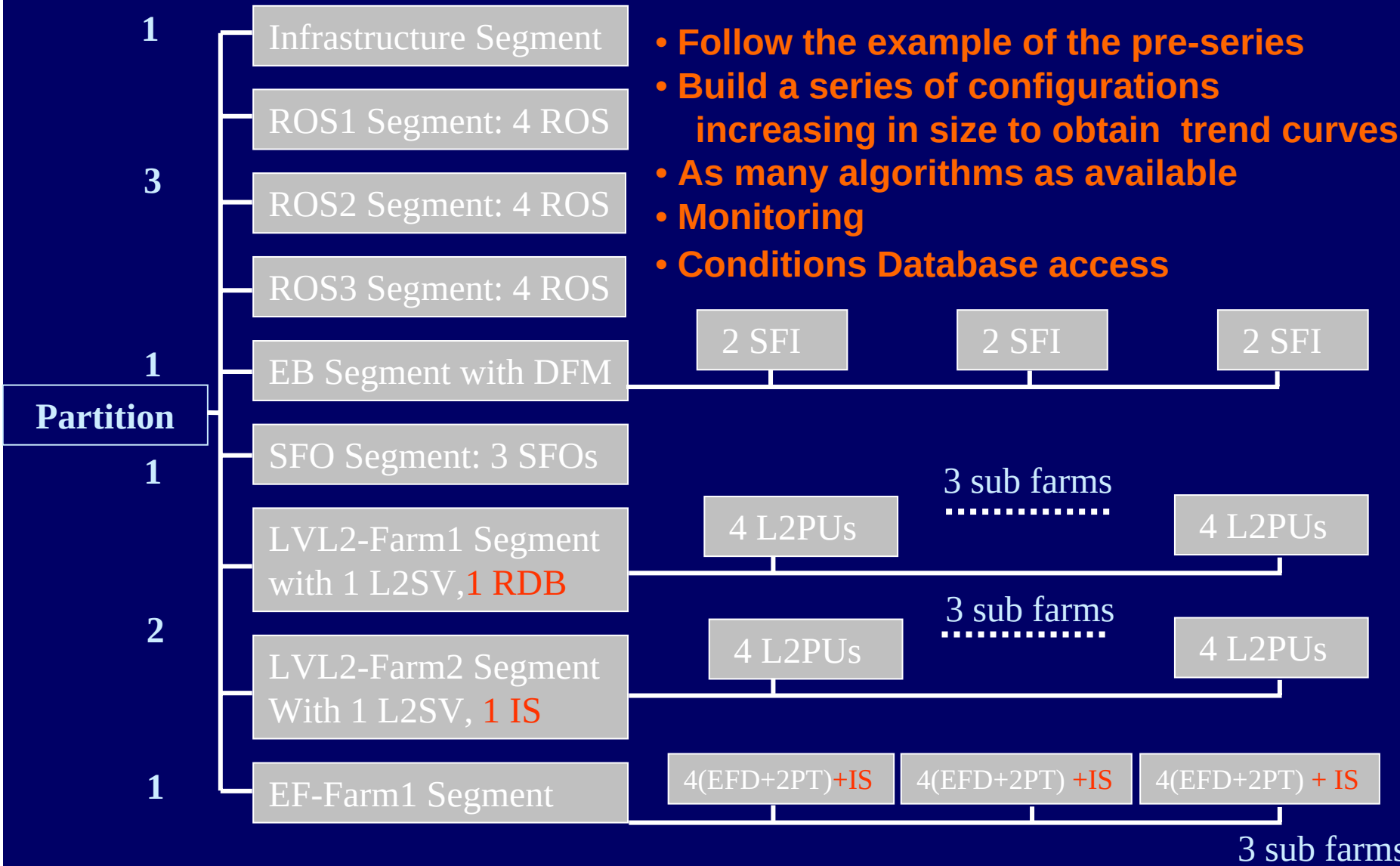


Aims

- Verify the functionality of the integrated DAQ/HLT software on a large scale
  for Atlas start-up,
  including algorithms and online databases
- Reliability
- Fault tolerance
- Selected sub-system and component tests
- *No experiment specific hardware*
- *Only one network connection per node*

Integrated tests: experiment-like structure Example:

The 'largest' partition (pre-series)



- **Follow the example of the pre-series**
- **Build a series of configurations increasing in size to obtain trend curves**
- **As many algorithms as available**
- **Monitoring**
- **Conditions Database access**

Each grey box represents a segment

Data flow & run structure issues

A. Lankford
on Detector calibration

Detector Calibration Model

ATLAS detector calibration model – preliminary subdetector requirements – R. Hawkings & F. Gianotti, 1 Mar 05

- A snapshot

Calibration requirements classified by location:

- ROD calibration — Only place where raw data is seen
- HLT calibration — ID, MDT alignment, MDT t0 calibration ...
- Calibration before prompt reconstruction
- Offline calibration

The first three impact TDAQ.

Several dedicated output streams foreseen, e.g.
- Generic tracks, pT > 2 (ROI based) for ID
- Electrons, pT > 20 (ROI based) for LAr
- Dileptons (Z -> ll) (full duplicate event)

Lots of data! 40-50 MB/sec requested

In addition dedicated LAr calib runs -- 4 hours monthly

From S. Gadomski

Streaming Study Group and it's mandate

ATLAS Data Streaming, chaired by Hans von der Schmitt (a "task force" setup by the ATLAS management)

http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/COMMISSIONING/streaming.html

Members: J.-F.Arguin (LBNL), S.Binet (LBNL), S.Gadomski (Bern), I.van Vulpen (NIKHEF), M.Shapiro (LBNL), I.Hinchliffe (LBNL), W.Verkerke (NIKHEF), J.Cranshaw (ANL), RD Schaffer (LAL), D.Malon (ANL), R.Hawkings (CERN), R.Jones (Lancaster), T.Lecompte (ANL), Z.Ren (Academia Sinica), G.Watts (Washington), S.Klous (NIKHEF), H.von der Schmitt (MPI) (chair)

ex officio:F.Gianotti (CERN), G.Polesello (INFN), Ch.Bee (CPPM), L.Mapelli (CERN), N.Ellis (CERN), D.Barberis (CERN+Genova), D.Quarrie (LBNL)

A study group for ATLAS Data Streaming has been established to answer outstanding questions in data streaming. The tasks of the group are:

- to do quantitative studies as required to answer questions such as
  - which streams are generated **at Point1** and at Tier0
  - overlapping vs. exclusive streams
  - streams and merging at the ESD and AOD level
- to prepare a recommendation to the COB on data streaming, due T&P week of May 29th.

Streaming “at the Event Filter”

This means streaming at Point 1, at the SFO.

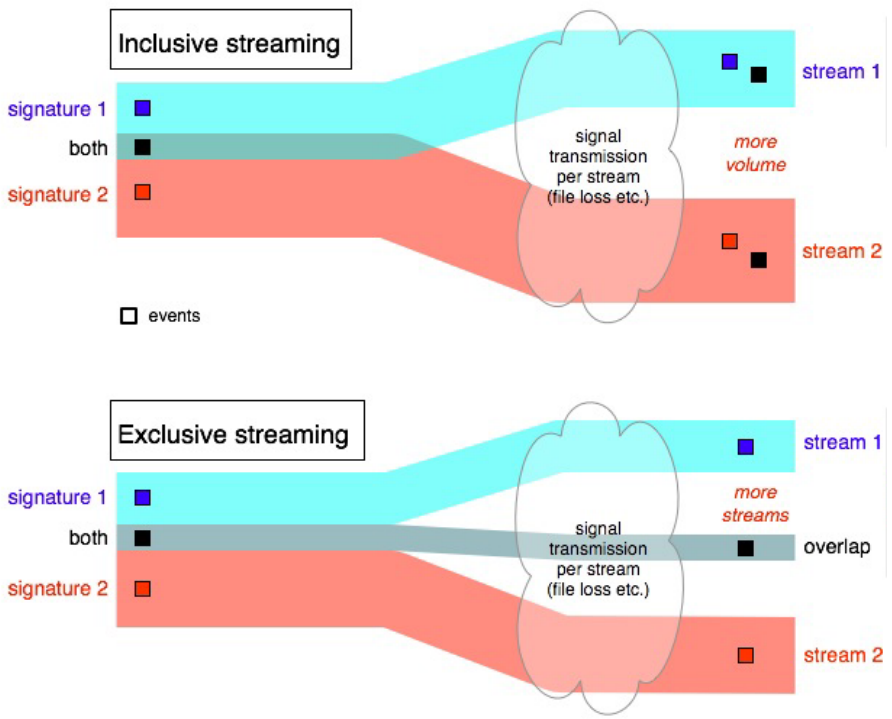
There seem to be roughly two points of view

1. Keep things like in TDRs (one bulk physics stream (like D0), several calibration streams, express and debug streams).
   - easier, KISS principle
   - offline database experts don’t care about the data files
2. Like above, but in addition split the bulk physics stream into several smaller streams along the PESA slices (muon, e-gamma, B-physics, jet tau and Etmiss)
   - may be useful for selective reprocessing of data, particularly at the beginning, when calibration and reconstruction will need to evolve quickly

No consensus, but a very clear majority in favour of variant 2.

Other ongoing discussions

1. Exclusive and inclusive streaming
   - Concerns raw data and the ESD + first pass AOD produced by ATLAS “centrally”, normally at the Tier 0.
   - Data samples produced further down stream, under control of physics groups, will be inclusive.
   - Both models can work in principle. It is a question of book keeping in relevant data bases of the offline.
   - No consensus either. A little tense at the moment.



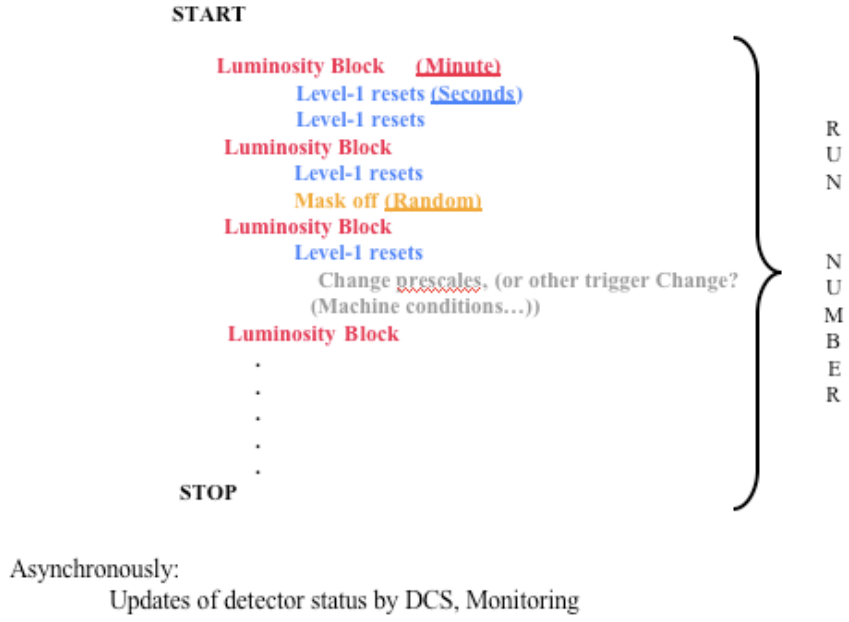
2. Existence, purpose and content of the express stream
   - The discussion is pending.

From G. Mornacchi
on Run structure --
report on the ongoing discussion in a working group set up in Feb'06

For discussion

- **Run structure**
  - **Run number, START, STOP**
  - **Luminosity block as unit of stability**
    - Extend its scope
    - Information integrated offline at Tier-0 with other time stamped data
  - **Keep the run going (constant run number)**
    - Recover automatically failures in DAQ, read-out
    - Prescale changes, local failures trigger change in luminosity block
    - Use "Pause"
  - **Change of run number**
    - Actions requiring major reconfiguration
    - Human decision (no longer worth to take data)
  - **Remove the concept of the checkpoint**

Length of order minutes, changes at regular intervals
DQ can be specified at this level of granularity



Asynchronously:
Updates of detector status by DCS, Monitoring

- **STOPless run**
  - **Mechanism to remove "things", leaves human "free will" unhampered though**
  - **Automatic:**
    - Triggered by BUSY On (removes source of the busy from read-out)
    - Triggered by a "FATAL" message generated by e.g. a ROD controller
  - **Can one "put things" back?**
    - If a ROD is involved (in one way or another) this looks not possible
  - **Can one just act on a source of error with an action less drastic than "remove"?**