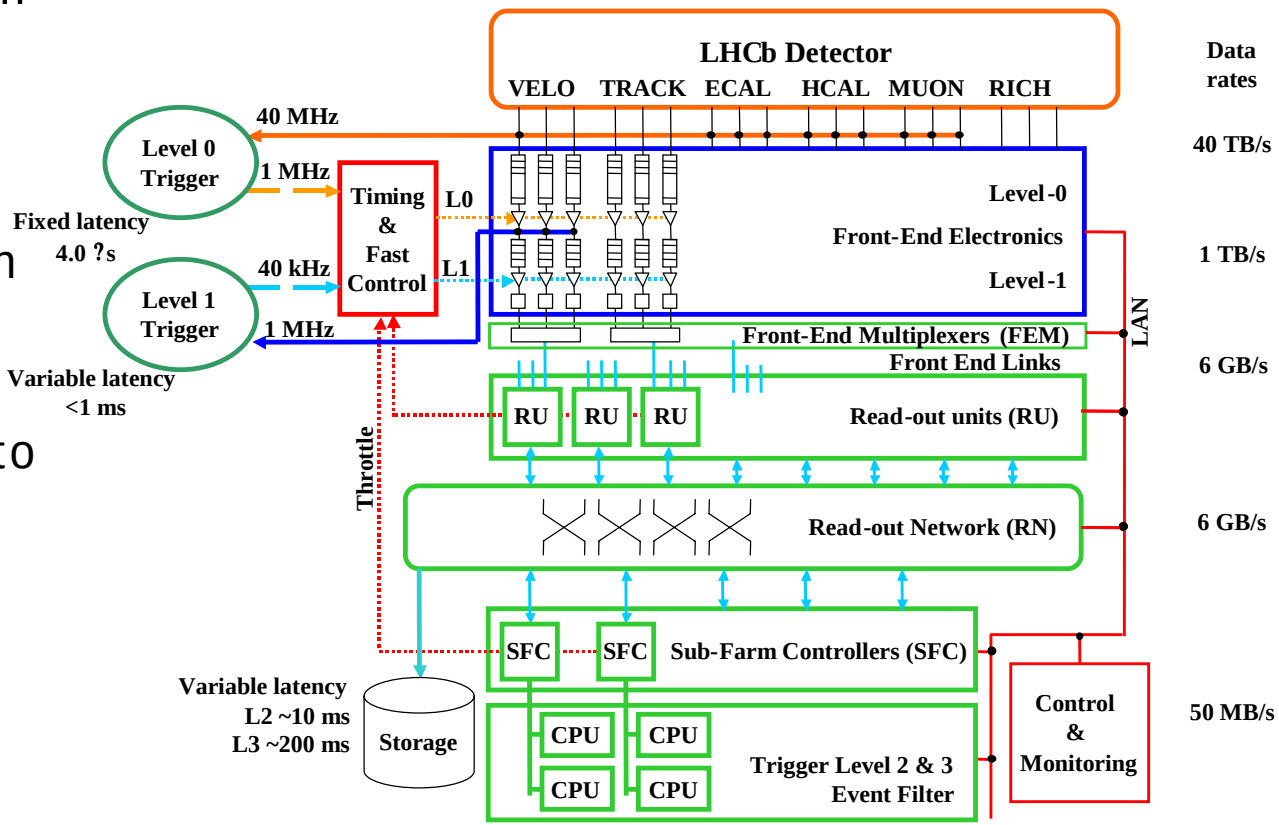
Status Report on the LHCb Online System

LHCC Referees Meeting
14 May 2001

Clara Gaspar & Beat Jost
Cern / EP

Introduction

- The LHCb Online System is composed of
  - Dataflow System responsible for timing and trigger distribution and the coherent data movement from the front-end Electronics to the data-store.
  - Experiment Controls System (ECS) controlling and monitoring the operational conditions of the entire experiment



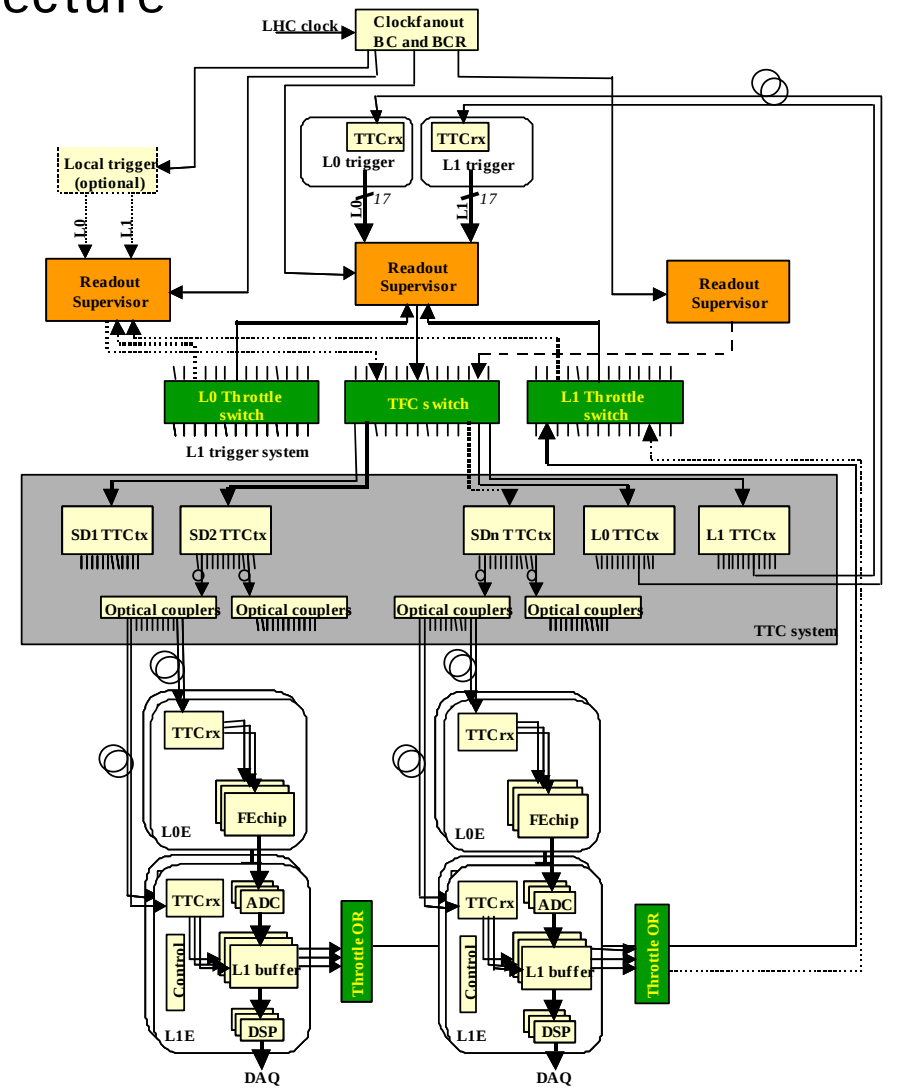
Outline

- Dataflow System Status
  - Timing and Fast Control (TFC)
  - Front-End Multiplexing/Readout Unit (FEM/RU)
  - Event Building (EB) and switching fabric (Readout Network, RN)
  - Event-Filter Farm (EFF)
- ECS Status
  - SCADA Framework
  - Interfacing to Electronics
  - Test-Beam DAQ Controls
- Summary
- Planning towards TDR

Dataflow System Status

Timing and Fast Controls

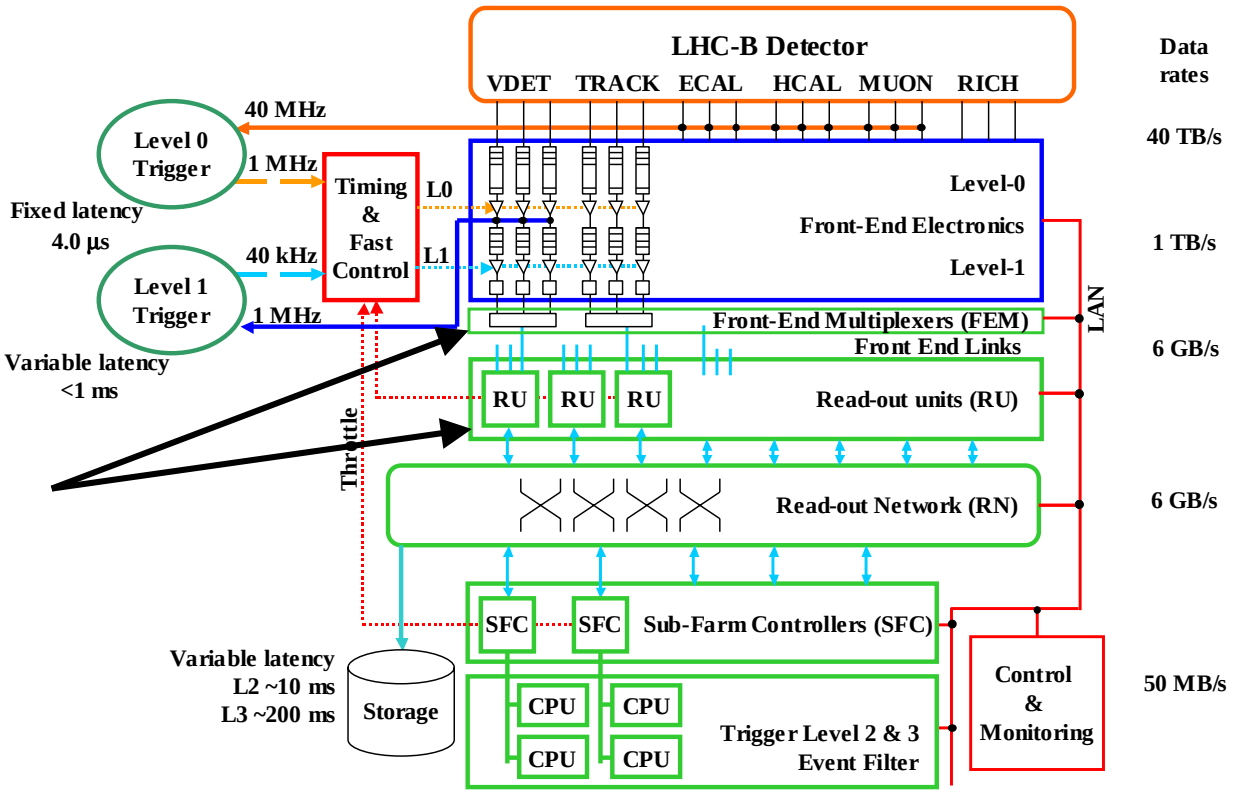
Architecture



- Two Reviews held with ext. reviewers
  - General architecture and design of switches (Oct 2000)
  - Specifications and Design of Readout Supervisor (RS) (April 2001)
  - Both successful
- Hardware development
  - Switch Prototype (almost) ready
  - RS prototype by October 2001
- Test bed for TFC Components
  - Verified that TTC system can issue broadcasts at Level-0 rate ~1.11MHz (measured rate ~1.6 MHz using TTCvi)
  - Will measure jitter etc. with switch prototype soon.

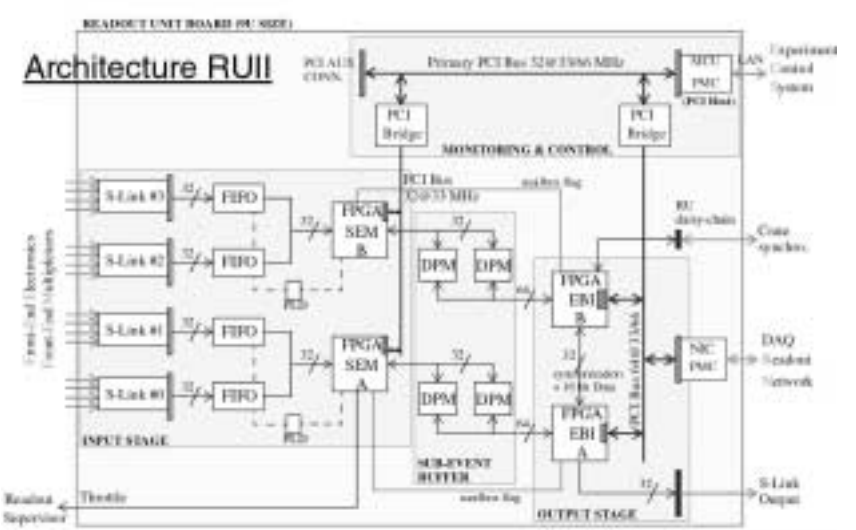
FEM/RU Complex

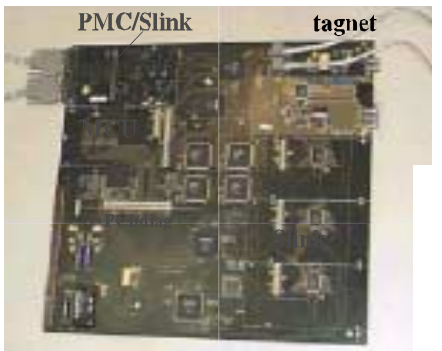
- Currently two candidate solutions
  - FPGA-based
    - Conventional technology using standard components
    - Prototype exists
  - Network Processor-based
    - Fully programmable
    - very flexible
- Review for technical assessment end of July 2001
- Decision on baseline solution end of August



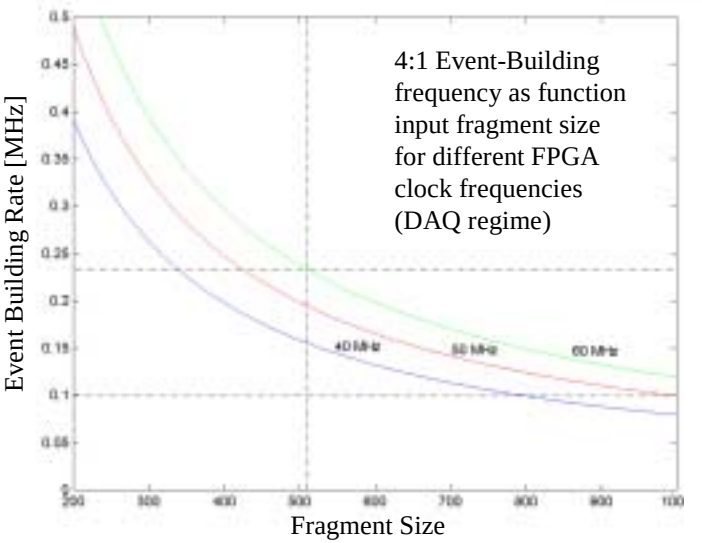
FEM/RU Complex – FPGA RU

- New Architecture
  - Better performance, fewer chips
- Second prototype finished
- Tests ongoing

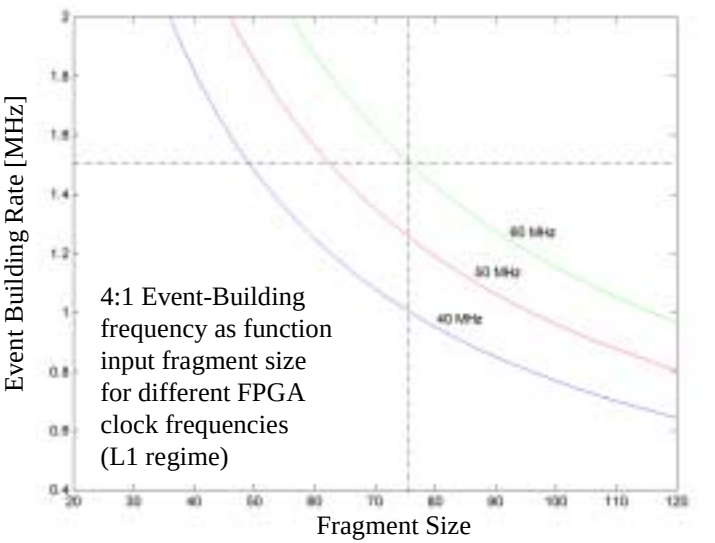




- Performance simulations show sufficiency for FEM/RU and Level-1 application



4:1 Event-Building frequency as function input fragment size for different FPGA clock frequencies (DAQ regime)



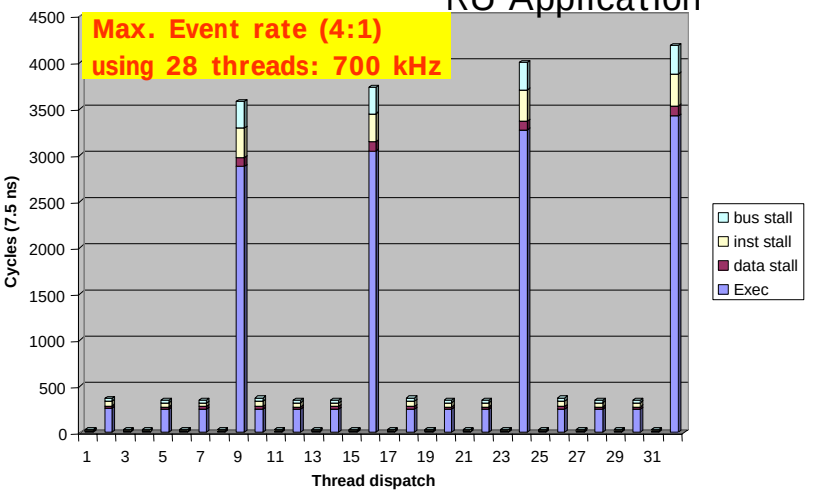
4:1 Event-Building frequency as function input fragment size for different FPGA clock frequencies (L1 regime)

FEM/RU Complex – Network Processor RU

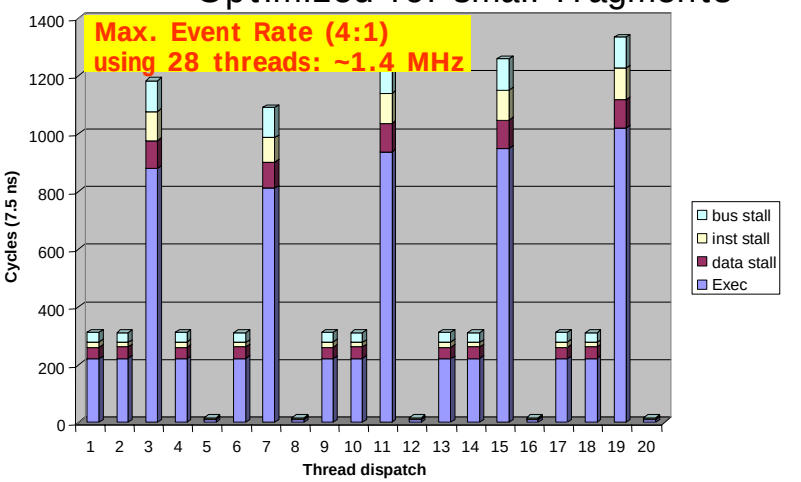
- Network Processors are a new technology gaining very much in momentum in the switch industry. All major chip manufacturers are working on them (IBM, Intel, Siemens, …)
- Target market are switch manufacturers using them as input stage of high-speed switches.
- LHCb software written, debugged and simulated
- Performance simulated and found sufficient for FEM/RU application and prob. also as L1-RU
- Hardware reference kit ordered (functionally equivalent to RU but on several boards)
  - ➢ Real performance measurements
- Also some interest in Atlas

Simulated Performance (n:1 event-building)

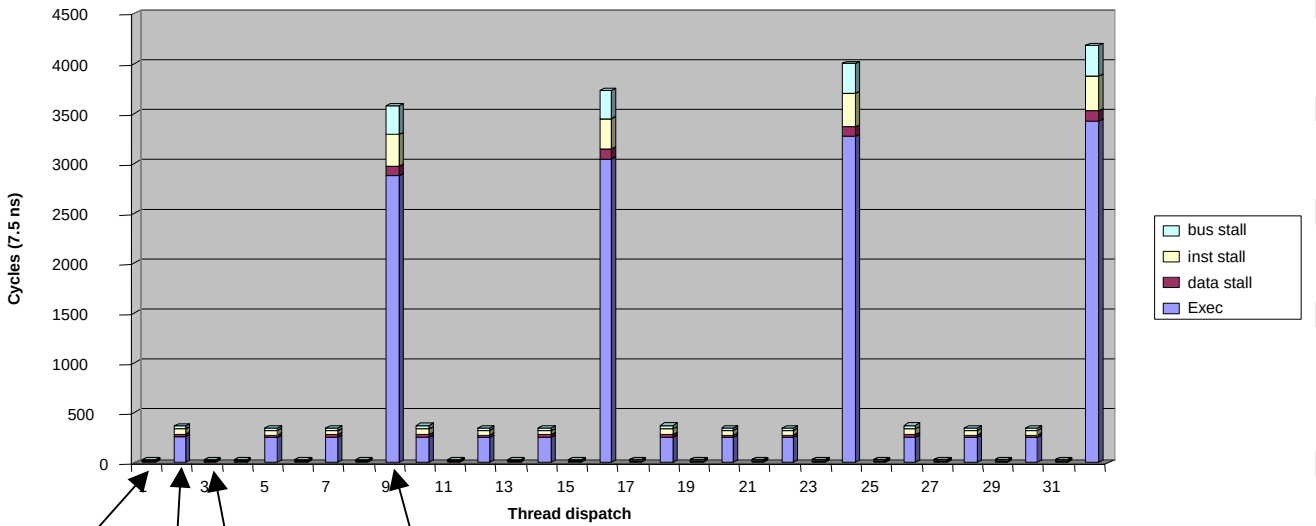
RU Application



Optimized for small fragments



Interpretation of Timing Plots



Input Handling
Fragment #1
Event #1

Input Handling
Fragment #2
Event #1

Book-keeping
Fragment #1
Event #1

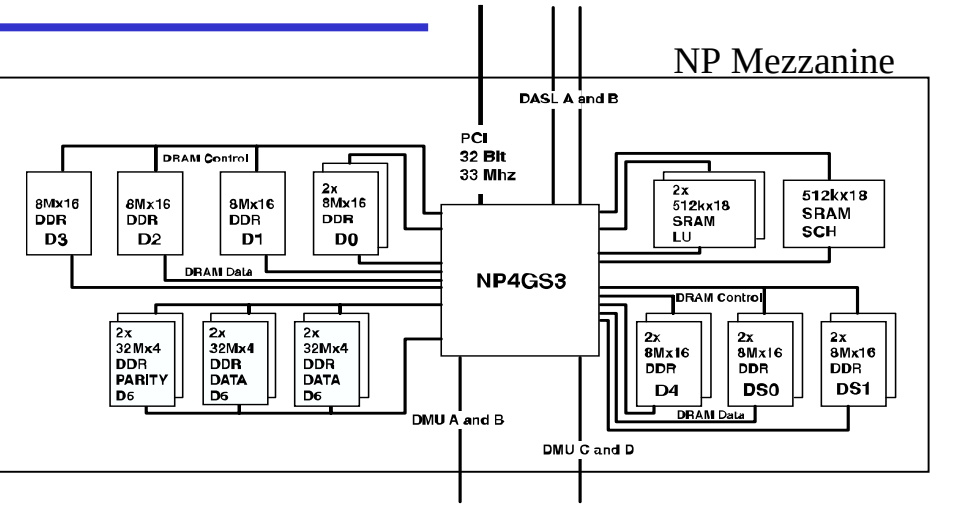
Event-Building
For Event #1

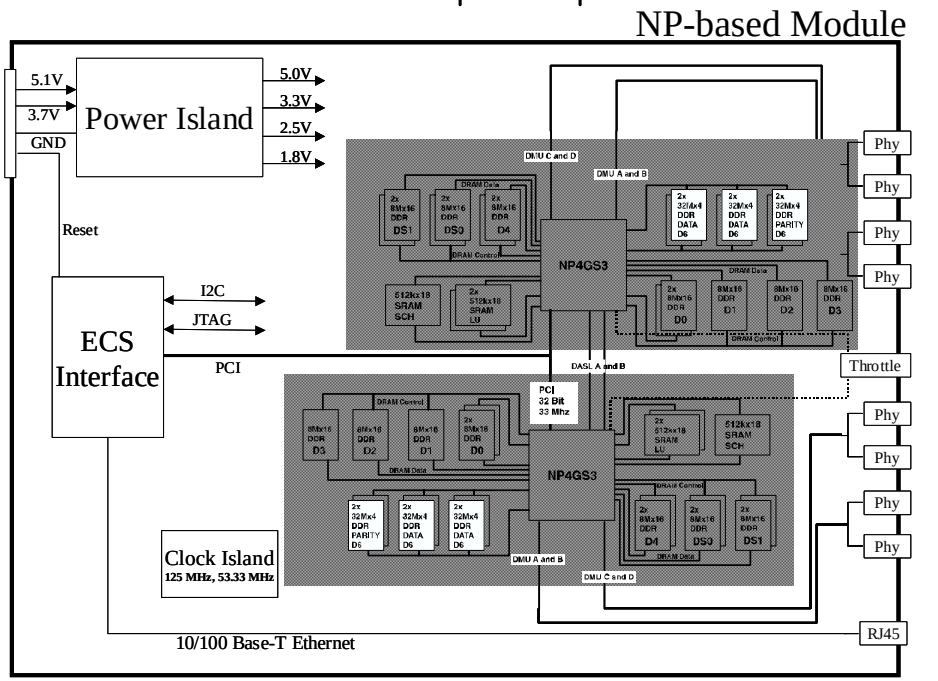
All to be verified with real hardware!

- 4:1 Multiplexing/Event-Building
- Input handling takes almost no time
- Book-keeping takes ~400 cycles à 7.5 ns per fragment
- Event-Building takes ~3800 cycles for 4 Fragments (incl. book-keeping)
- Time per fragment incl. Book-keeping:
  ~(4*400+3400)/4 =
  ~1250 cycles = 9.4µs or ~100 kHz
- There are 28 threads running in parallel, hence 2.8 MHz of fragments can be treated. At 4 fragments per event this leads to ~700 KHz of events
- For small fragments this rate is ~5.6 MHz of Fragments or ~1.4 MHz of Events (4:1)

A Network Processor based Readout Unit

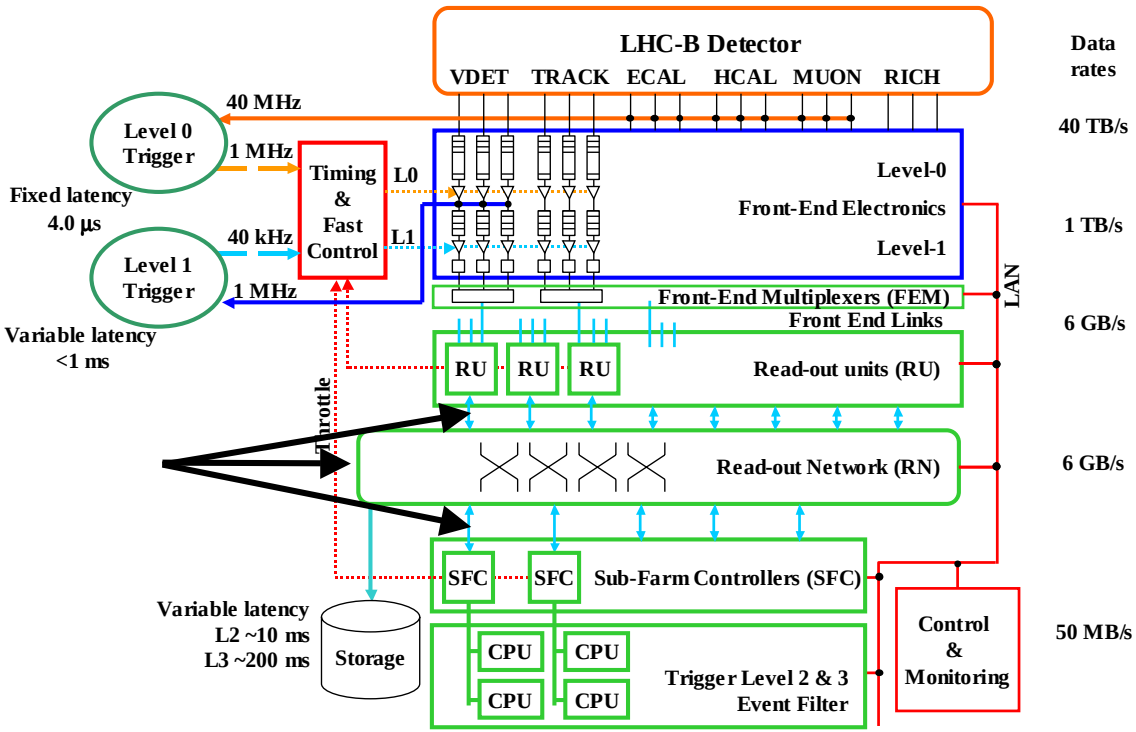
- Possible implementation scenarii being studied. One example could be a mezzanine card containing the NP+memories together with a simple mother board.
- Very generic board. Can be used as (depending on the software loaded)
  - FEM/RU (n:1 Multiplexer/EB)
  - 4x4 switch
  - 4:4 event-builder (input stage of SFC)
- One single module for all applications in LHCb and only one network technology (Gb Ethernet)
- Negotiations started on external design/production of boards





Event-Building and Readout Network

- No deviation from Full Readout protocol (Push-through, i.e. all data are transferred to farm CPU)
- Baseline technology is xGb Ethernet (limit to 1 Gb/s at RU/SFC)
- We have setup a 2x2-port development test bed but also have access sometimes to a 32x32 test bed of CMS
- Restarting simulation effort to confirm scalability
- Looking into alternative topologies to minimize number of ports (cost)
- Studied smart NIC (Network Interface Card)
  -> sufficient performance up-to ~90 kHz L1-rate

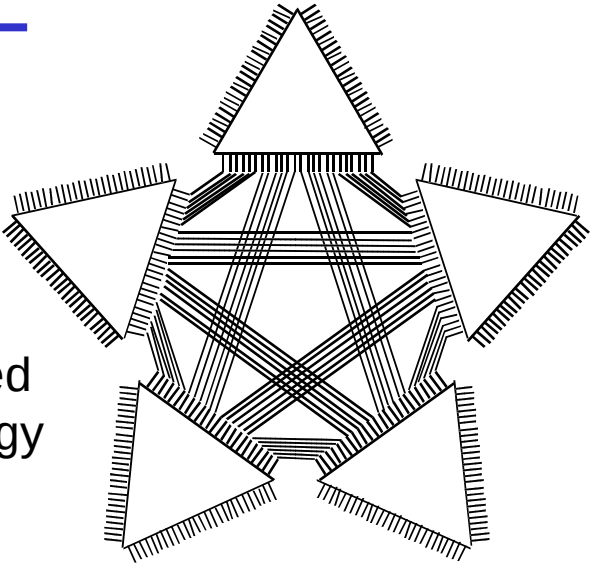


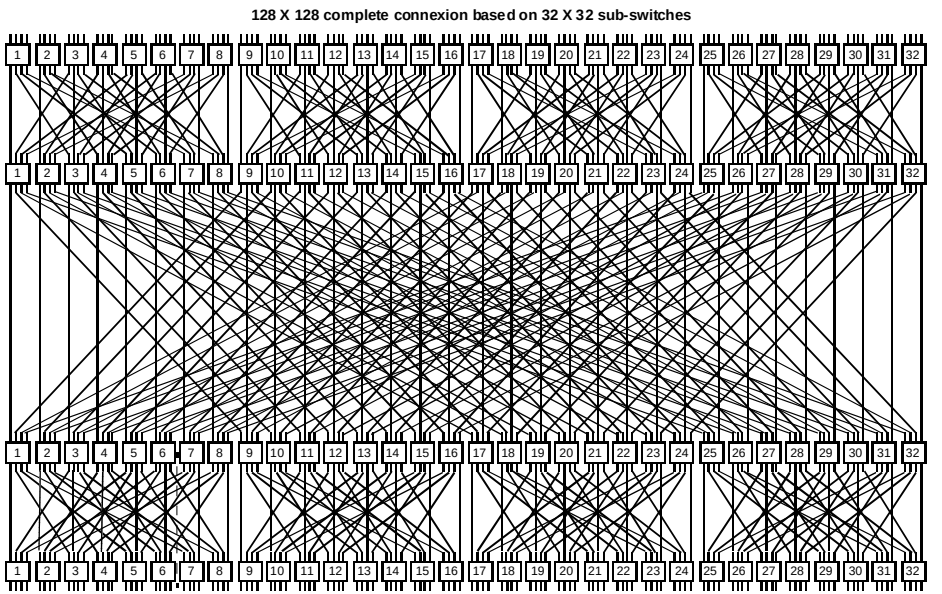
Readout Network - Alternative Topology

- Relies on the fact that traffic is very unidirectional
- Working assumption: 4x4 switch as building block
- Mixing source/destination on same module
- Still full bandwidth available for event-building

'Final' network:
270 ports with 90 4x4 modules

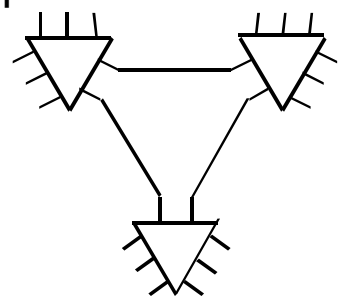
➢ 30% fewer Modules compared to **Banyan** topology





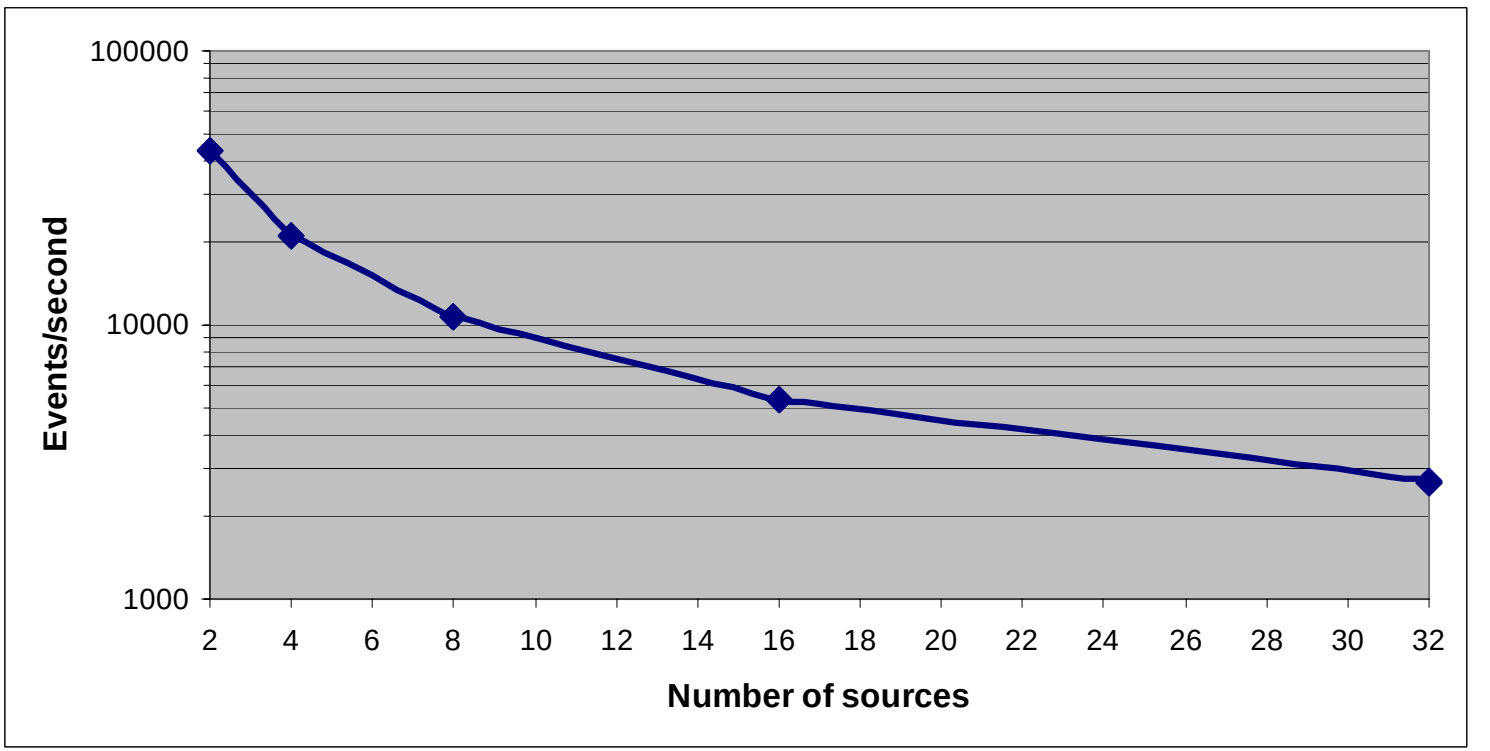
Basic Building Block

➢ 3 4x4 switches to make an 18 port-switch



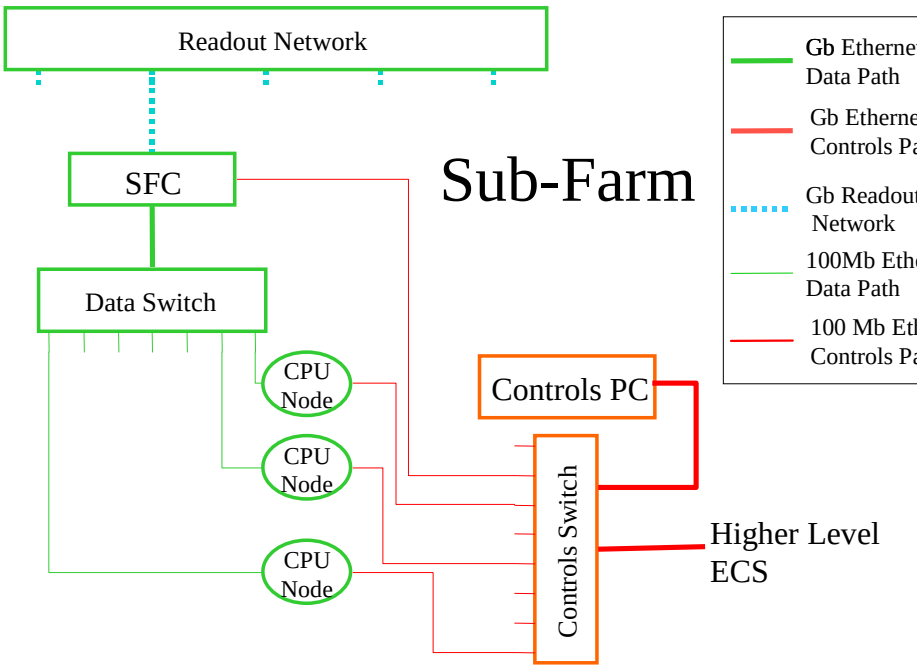
Smart NIC performance

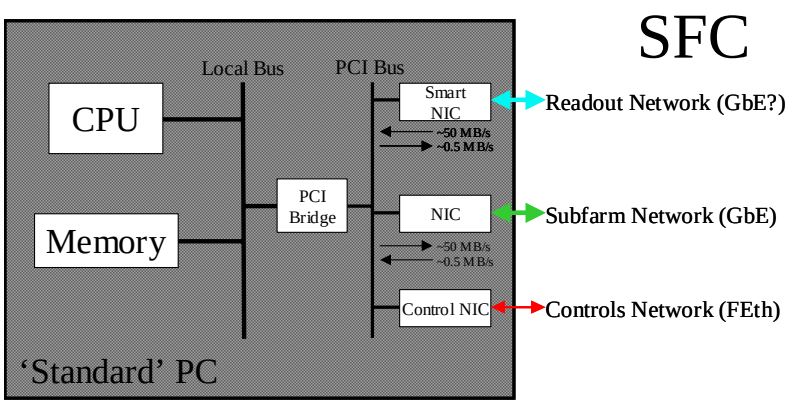
- Measurements emulating the sending of data fragments from several sources and event-building at the destination using two smart NICs.
- Results show that ~90 kHz from 100 sources can be sustained (Code not yet optimized)



Sub-Farm and SFC Architecture

- So far very limited effort invested. Ideas…
- Each sub-farm is an autonomous entity
- It will have a strictly separated controls and data path, reflected throughout (individual CPU nodes, SFC)
- Typically 10-20 CPUs per SFC and per controls PC
- Possibly the controls PCs will have another hierarchical level
- Should scale very nicely…



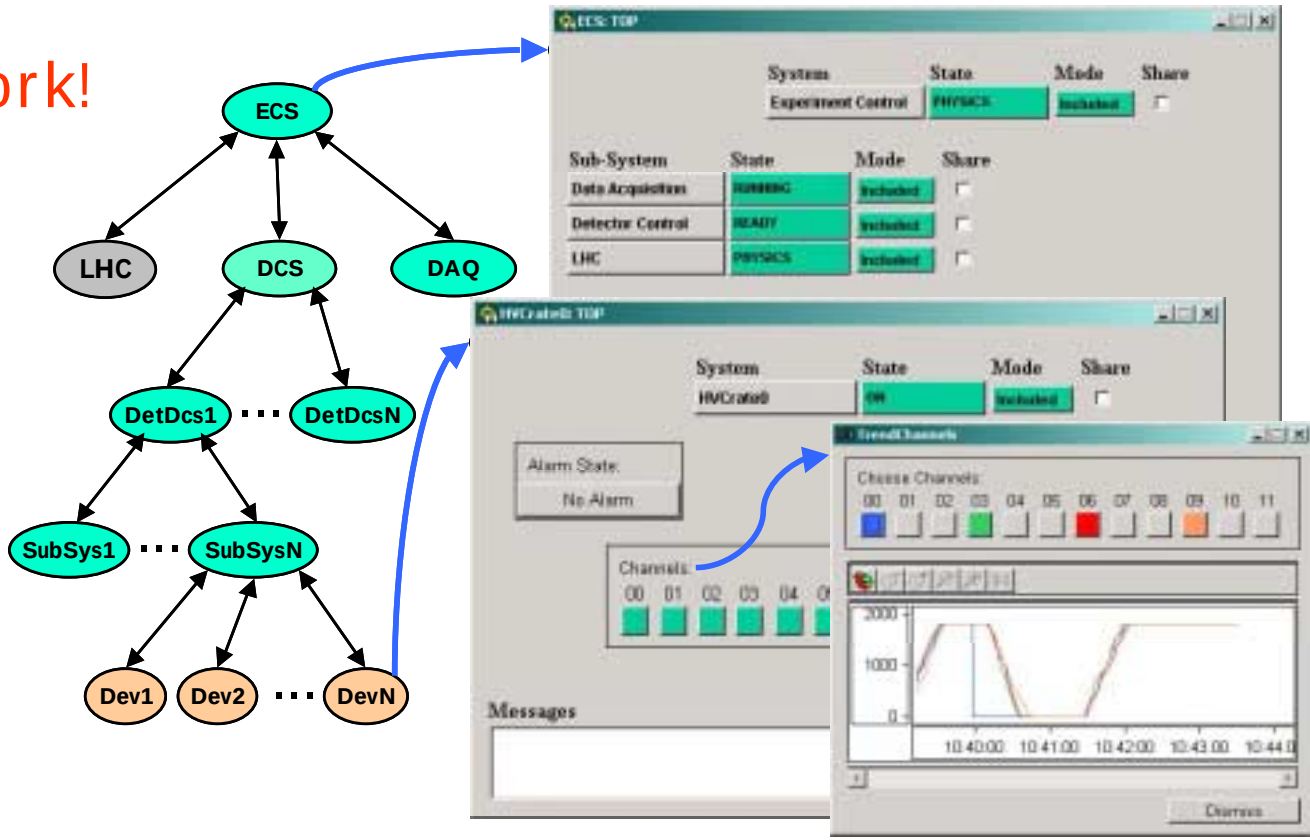


Experiment Control System (ECS) Status

LHCb Controls Strategy

- LHCb will have an integrated Experiment Controls System, governing the operational state of the all components of the entire experiment (Detector, DAQ, environment,…)
- **AIM:**
  No Duplication of work!

- The ECS is based on a hierarchical architecture, that is supported and enforced by a common software framework (first version available)
- Components are modeled as Devices, devices are, e.g.
  - HV/LV Power Supplies
  - RUs
  - L2/3 CPU node
  - L2/3 Algorithm



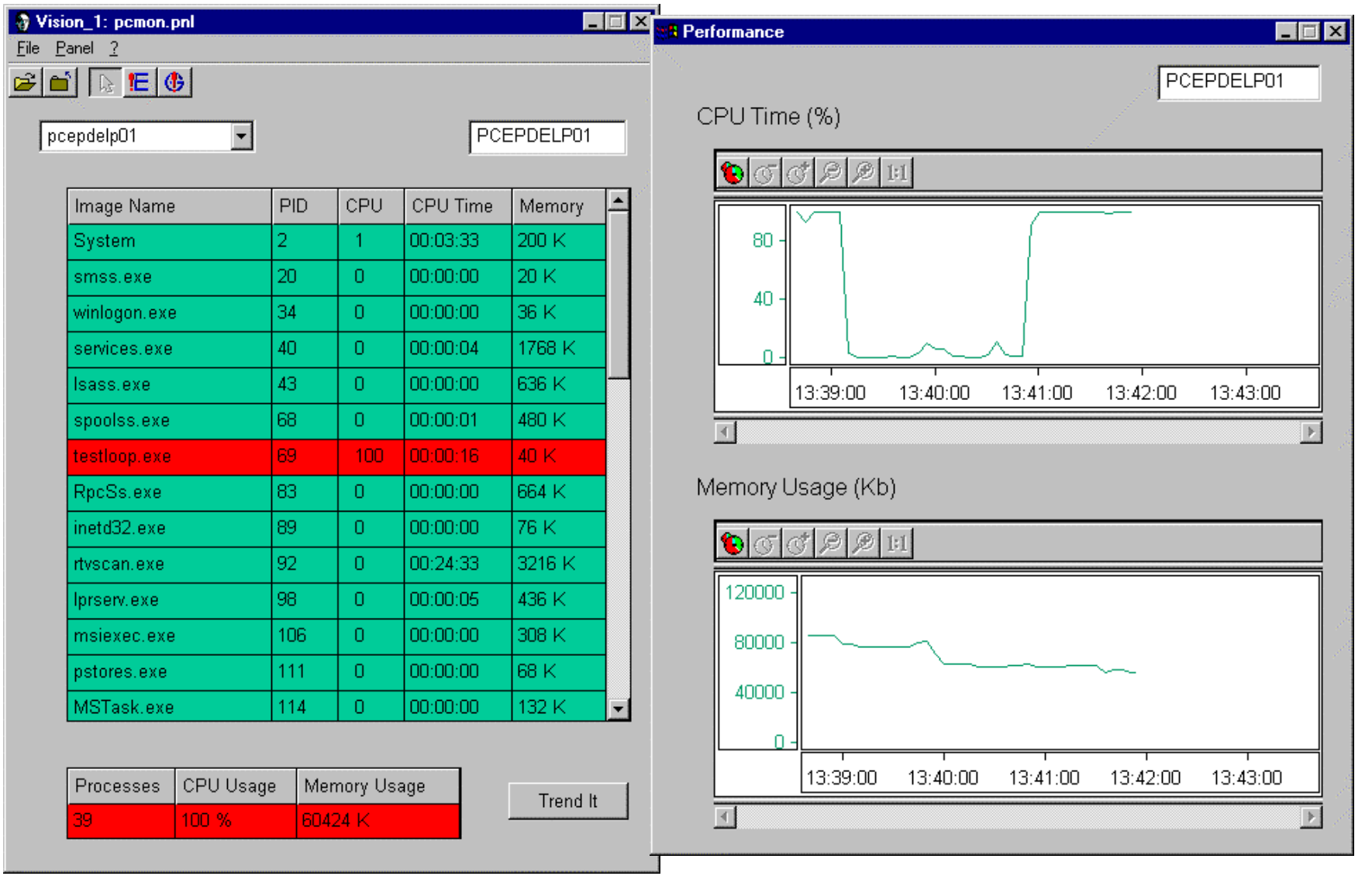
Software and Tools

Basic software infrastructure is a commercial SCADA (Supervisory Controls And DAq) system (PVSS II), common to all LHC experiments and others…

- Joint COntrols Project (JCOP)
  - SCADA (acquisition, licensing)
  - Architecture Working Group
  - Controls Framework
    - Including Common Components for HV, LV, etc.
  - Data Interchange Working Group
  - Gas Working Group
  - GIF
  - Rack Control

- LHCb Specific Activities
  - Test Beam Upgrade to PVSS-based run-controls
  - Interface to electronics
  - Eventually, interfacing to “Slow Controls” devices

Example: CPU Monitoring using SCADA

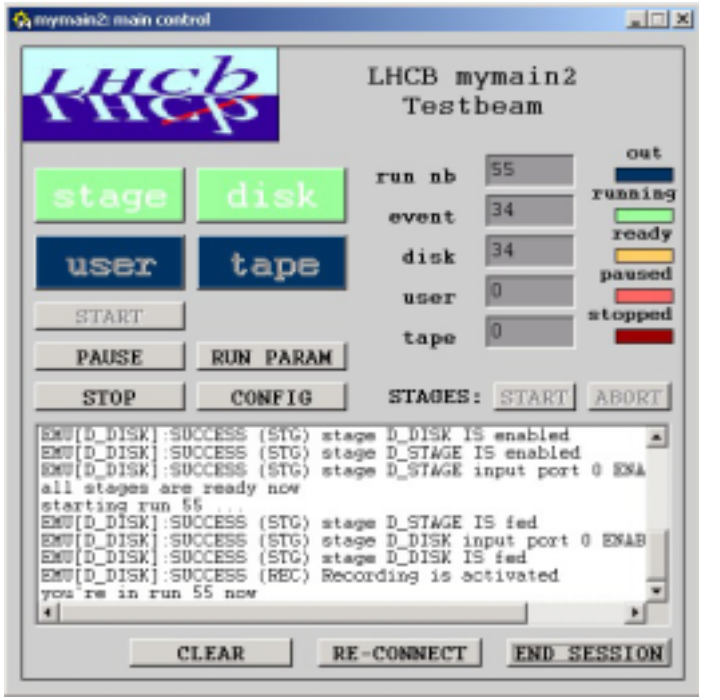


Test-Beam Controls upgrade

Replace original run-control by PVSS-based run control

- Prove of concept of general idea of integrated controls system (run control & slow control)

New Test Beam Run-Control Panel

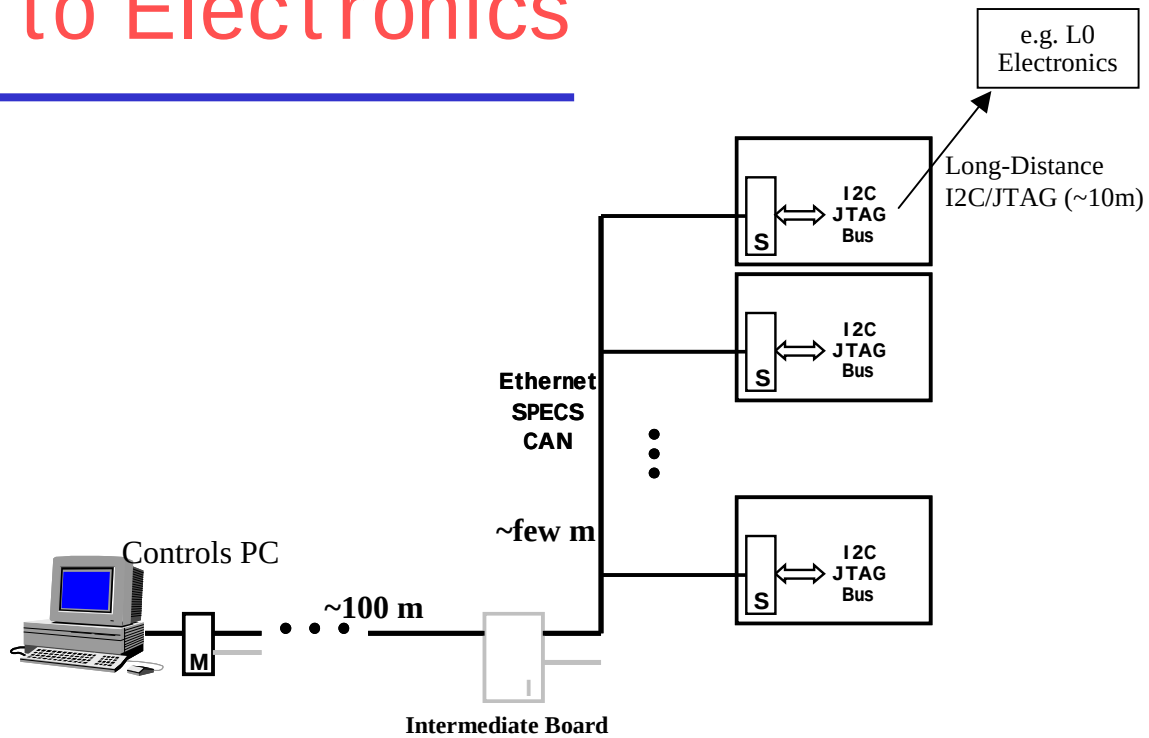


Add other ('slow controls') devices when and where needed

- Data from the PS
- High/Low Voltage
- Gas
- ...

ECS Interfaces to Electronics

- LHCb standardized on 3 interfaces to Electronics (board-based)
  - Credit-Card PCs for electronics boards in the barracks
  - SPECS (evolution of SPAC) and CAN-ELMB (Atlas) for areas with radiation (SEU), but no total-dose problem, provided they are SEU-hard (Sep '01).
- For front-end chips usually I2C or JTAG is used for controls and monitoring.
- I2C and JTAG will be generated over long-distance (~10 m) from SPECS/CAN-ELMB



S (slave) can be
- CC-PC (not in cavern)
- SPECS
- CAN-ELMB (Atlas)

Summary

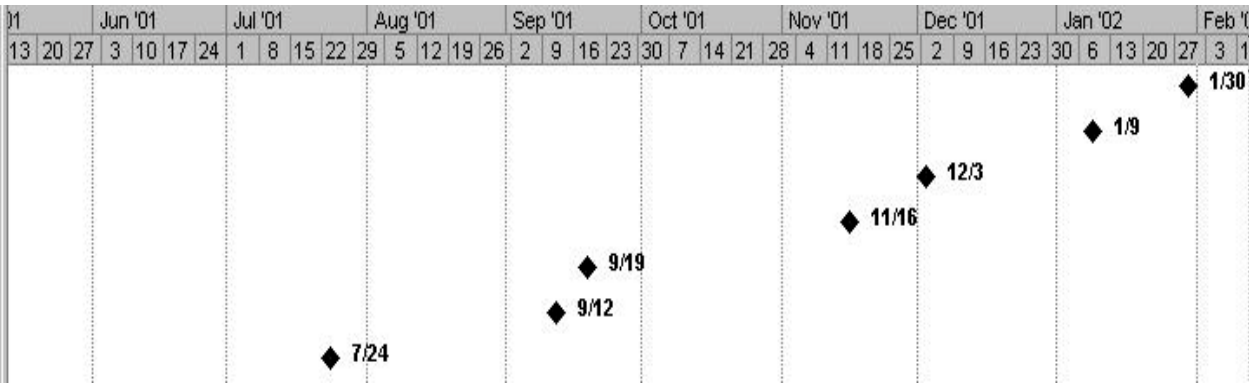
- **TFC Test-bed set up**
  - Verified that TTC system can issue Level-1 decisions at required rate (1.11 MHz) in test-bed
  - TFC switch prototype almost ready
  - Readout Supervisor designed, prototype in October '01
  - Design reviews done
- **FEM/RU**
  - Second prototype of FPGA-based RU ready and being simulated/tested
  - Investigating Network Processors (-> uniform module for entire LHCb DAQ)
- **Event-Building**
  - Looking at alternative topologies for Readout Network
  - Restarting simulation effort of RN
  - Measured performance of "smart" NICs, ok.
  - Test-bed set up for development and performance measurements
- First ideas on CPU farm
- Overall review of Dataflow subsystem planned for September

- ECS
  - Integration of Dataflow controls and Slow controls in a single Experiment Controls System
  - Hierarchical architecture defined, first version of software framework and tools available
  - Hardware interfaces to (board-level) electronics standardized
  - Proof of concepts in test-beam

Planning towards the TDR

TDR Planning

	Task Name
1	LHCC Session
2	Submission to LHCC Secretariat
3	Approval by LHCb TB
4	TDR ~Final Draft. Submission to LHCb TB
5	Intermediate Report to LHCb TB
6	Review of Dataflow System
7	Review of RU/FEM solutions



- Working backwards from the LHCC session in January 2002
- Make use only of the foreseen LHCb weeks (no special TBs, etc.)
- TDR Format:
  - ‘Terse’ main text with supporting notes (à la TP)