Trigger, DAQ, & Online Planning and R&D needs

Super B Factory R&D Workshop 2008.2.14

S. Luitz and G. P. Dubois-Felsmann

Outline

- Design Principles
- Rate and Event Size Estimates
- FCTS and Data Flow
- Event Builder and Level-3 Trigger Farm
- Logging and Data Quality Monitoring
- Control Systems & Logbook
- Configuration Management
- Hardware Reuse, Cost and Labor Estimates
- R&D Projects
- Summary

Design Principles

- ## Open trigger
  - ### Just like BaBar
  - ### Recoil and rare-decay physics essentially demand this: the critical physics signatures are not distinctive at trigger level

- ## Commodity hardware (wherever possible)

- ## Uniform design across subsystems

Open Trigger?

- Yes: today's commodity hardware in can already handle event building and Level 3 at the resulting rates

  - Some concrete estimates below assume 2009 technologies at today's prices…

- The challenge: Level 1 rates and the corresponding front-end electronics

  - We need your help to define the requirements!

Rate and Event Size Estimates/Assumptions

- Event size
  - BaBar: ca. 35kByte average
  - SuperB: educated guess based on BaBar: 75kByte average
    - Possibly smaller cells in DCH
    - More samples needed to detect pile-up and overlapping events
    - Higher occupancy due to background events

    - But you need to define this (this week?)!

Rate and Event Size Estimates/Assumptions II

- **Trigger Rates –BaBar Open Trigger Philosophy**
  - **L1-Accept (dominated by e^+e^- interactions)**
    - From BaBar:
      - ~20nb (BaBar physics filter output) → 20kHz @ 1e36
      - ~50nb Bhabhas (BaBar acceptance) → 50kHz @ 1e36
      - 10% scaled backgrounds of BaBar → ~25kHz
      - Operational BaBar experience: 50% headroom desirable
    - Would require L1 and Front End DAQ to handle 150kHz L1-Accept rate
    - Bhabha veto could bring down L1-A rate (with headroom) to ca. 100kHz
    - Assume Bhabha veto for now
  - **L3-Accept**
    - Assume ~20nb + ~5nb for monitoring triggers → 25kHz @ 1e36 L3 accept rate
    - Moderately optimistic extrapolation from BaBar

CDR DAQ / Online design

- ## Heavily based on the existing BaBar system
  - ### Nice:
    - Build on the BaBar success
    - BaBar is the source of the most reliable performance and effort data available to us
    - Not a radically new design, we think we understand most issues
  - ### But keep in mind:
    - The design parameters are fairly challenging for an HEP DAQ system - particularly for the front-end electronics
    - There may be other ways to build such a system (e.g. with LHC-style technology)
    - Not a radically new design, we think we understand most of the issues but there may be some unforeseen surprises
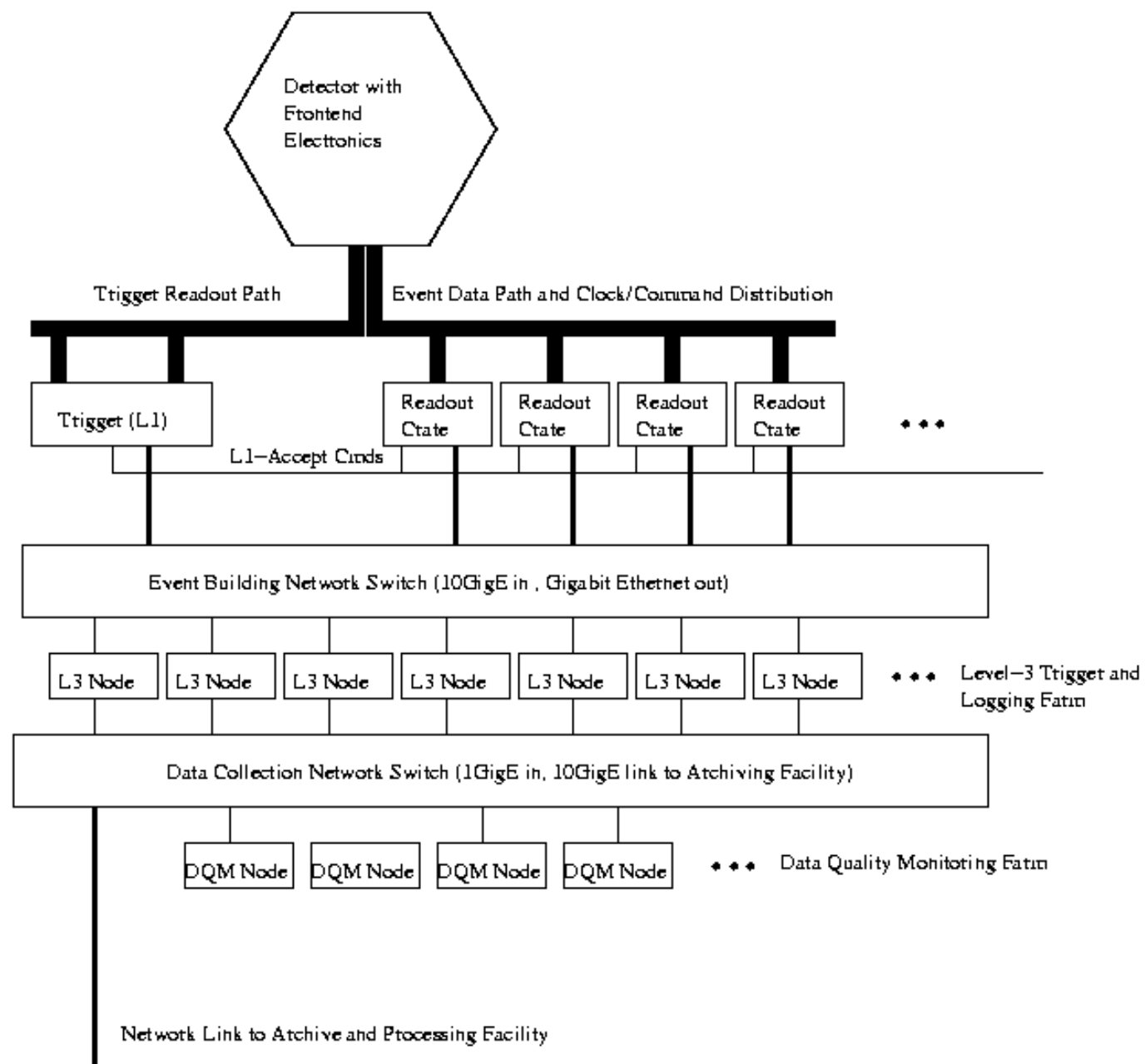
Looks like BaBar, doesn't it?

Trigger

- ## BaBar-like trigger structure:
  - ### Level-1
    - Pre-event-readout trigger based on drift chamber and calorimeter information
    - Level-1-Accept initiates event readout
  - ### Level-2 (optional)
    - Trigger in data path acting on partial event information
    - Currently not foreseen (event sizes are small!)
  - ### Level-3
    - Software trigger running on server farm, acts on complete events.
    - Refines Level-1 calculations, allows more complex operations
  - ### Level-4
    - "Offline" filter - may need to place upstream of permanent storage!

CDR DAQ and Fast Control and Timing System

- Design based on a modular DAQ system currently developed at SLAC (for other programs!)
  - FPGA/System-On-A-Chip
    - PPC general-purpose CPUs
    - VHDL-style "DSPs"
    - Various link options, including 10GigE and 2.5GBit/s fiberoptic serial links

- Similar to BaBar, but faster fiber links to front-end electronics (2.5GBit/s)

- Feature extraction as close to front-ends as possible
  - Proof-of-concept in revised BaBar DCH electronics

Fast Control and Timing system
## (trigger distribution and flow control)

- "No" mandatory dead-time between triggers allowed

  - Frontend readouts that can "keep up" with a (very short) minimal allowed command spacing OR

  - Trigger queuing and fully addressable front end readout buffers

Event Build and Level-3 Trigger Farm

- **BaBar-style network event builder and Level-3 farm**
  - 10GigE outputs of DAQ modules are fed into a commodity network switch
  - Event contributions are sent to Level-3 trigger farm nodes
    - Farm nodes connected to the switch by 1GigE
    - Take advantage of cheap (and possibly long) copper links. Copper 10GigE is limited to a few meters.
    - Use lightweight UDP/IP protocol with simple retransmit
    - Most of UDP protocol done by NIC
  - 100kHz x 75kByte $\rightarrow$ 7.5GByte/s or 60 GBit/s
    - Can be handled by current generation of network switches
    - The BaBar Cisco 6500-720 switch could handle this rate
  - Will require at least 120 farm nodes for networking reasons
    - Can't utilize 1GigE links more than 50% (or buffering is ineffective)
    - 8 CPU cores per node would allow ~5ms / event / CPU
  - Level-3 output rate: 25kHz $\rightarrow$ 210 Hz per node

DAQ Module

- Modular DAQ system design currently developed at SLAC for LSST, PetaCache, LCLS and other projects
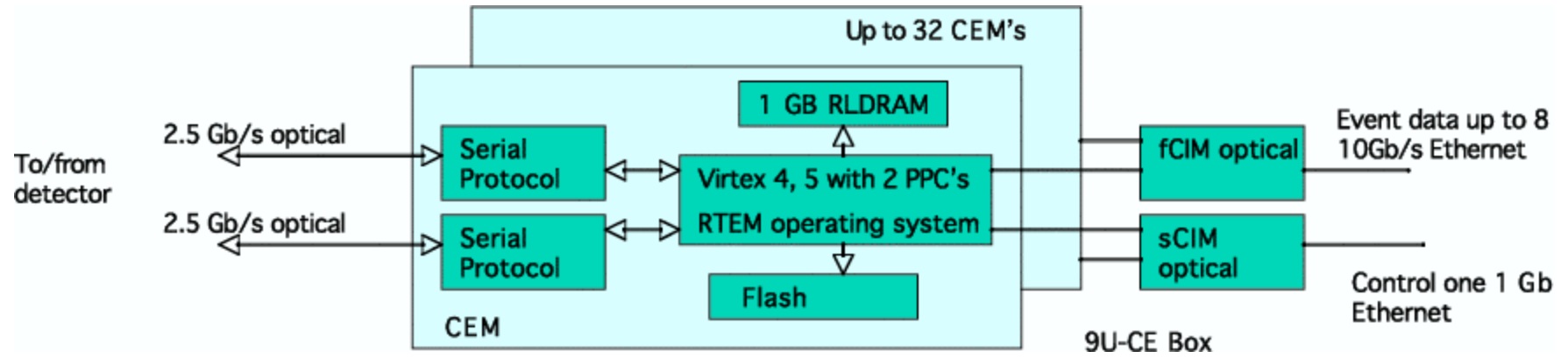
Logging

- Log to RAID-1 (mirrored) disks local to farm nodes, 2-4TByte/node

- 16MByte/s (210Hz x 75kByte) logging rate per node

- Send data from farm nodes to archival storage and further processing
  - Possible Level-4 before archival storage

- Total I/O requirement per node: 16MByte/s write, 2 x 16MByte/s read
  - Within reach of current disks

- Each farm node holds part of a run ("Runlet")
  - Runlets expected to be preserved, probably no further merging
  - Bookkeeping system will need to deal with this

Data quality monitoring

- **BaBar experience seems satisfactory**
- **Foresee additional farm of ca. 40 machines to run fast online data quality monitoring**
- **Alternatively data quality monitoring application could run on the L3 farm (e.g. using low-overhead resource-isolation technology).**
- **Histogram collection via DHP-like distributed histogramming mechanism**
- **GUI/Web based display**
  - Design this early in collaboration with detector systems!

Control Systems & Logbook

- **Experiment ("Run") Control**
  - Operator interface for DAQ
  - Sequences DAQ configuration, calibration, partitioning, etc.
  - Automate operations as much as possible
- **Detector ("Slow") Control**
  - Control detector and detector support systems
  - Monitor and record detector and environment conditions
  - Based on toolkit that provides interface to hardware
    - BaBar uses EPICS
    - Depending on hardware, a different system may be more appropriate
  - Should share a timebase with the FCTS (the event stream)
- **Unified operator interface / GUI**
- **Electronic Logbook**
  - Integrated with Experiment and Detector Control

Configuration Management & Release Control

- **All hardware and software configurations defined in configuration database**

- **Software release management and tracking**

  - What was running where at any given time in any part of the DAQ and Online system

  - With more "traditional software functionality" implemented in FPGAs, FPGA firmware is considered "software" and must be included in the configuration and release management scheme

Hardware Re-usability & Cost Estimates

- Most BaBar Online system hardware not reusable for Super B factory:
  - DAQ and Trigger – for obvious reasons
  - Computing & Networking
    - Need to take advantage of Moore's Law
    - Useful life of commodity computing 3-4 years
- Basis for online computing cost estimate
  - BaBar experience with 2 upgrades shows:
    - "Boxes don't get cheaper, they get faster"
    - "A farm node always costs $3k, now it has 8 cores"

Software Re-use?

- Concerns about re-using online software:
  - Tied to hardware
  - May be sub-optimal
    - As recent optimizations of feature extraction codes have shown
  - Very specific optimizations for BaBar operations
    - E.g. detector controls
  - In-depth expertise may no longer be available
  - Different expertise and tools available in LHC era
  - Documentation sub-optimal or non-existentConcepts and frameworks can (and should) be re-used
  - A lot of code will need to be written from scratch anyway

Labor Estimates

- ## 80-90 FTE-years (based on BaBar effort estimates)

- ## Given the challenges, we estimate the Super B Factory effort to be comparable to the initial BaBar online computing effort

- ## Apply lessons learned to development model

  - ### E.g. feature extraction code to be developed by "professionals"

  - ### More core effort for e.g. the slow controls system to avoid the "3 different drivers for the same hardware" syndrome

R&D Projects – further study

- **Detailed evaluation of channel counts, and per-channel DAQ requirements**
- Study backgrounds and occupancies and their effect on event sizes

R&D Projects – further study II

- Study attainable performance of a Level 1 Bhabha veto

- Assess cost of extending L1-Accept capacity to 150kHz
    - Headroom, no Bhabha veto, later lumi upgrades

- Investigate consequences of overlapping events, triggers and trigger queuing

- Queue modeling for data path to determine buffering requirements

R&D Projects – further study III

- Study adaptation of the current "Level-4" BaBar physics filter to Level-3 quantities
- Review LHC technologies for applicability:
  - Look at control system alternatives to EPICS
  - Concrete experience with ~100kHz DAQ (though expensive)

Summary

- "BaBar++" design of Online and DAQ seems feasible for Super B Factory
    - Front-end DAQ is major challenge and needs prompt R&D attention
    - Need a Level 1 trigger R&D effort integrated with tracking, EMC
    - More R&D in many other areas needed for more detailed design
- Hardware re-use very limited
- Software re-use needs careful consideration
    - In many cases a redesign with lessons learned probably the better approach
- All challenges can (probably) be solved with technology already available today – costs will continue to decline