# Handling/surviving with radiation effects in large scale complex systems (HEP experiment)

Jorgen Christiansen
PH-ESE

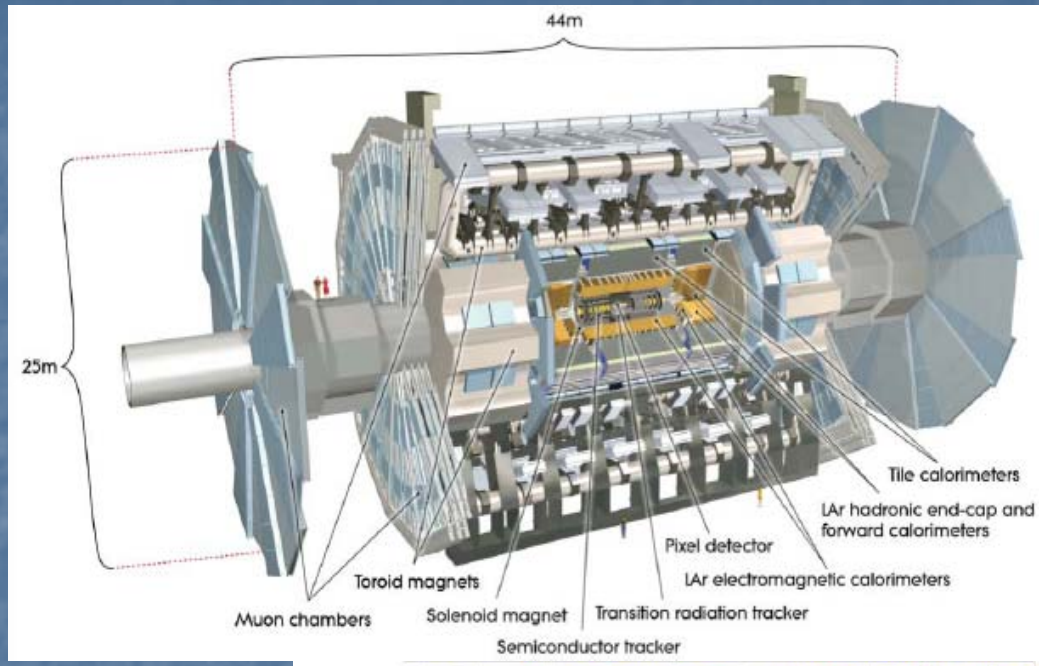
# Outline

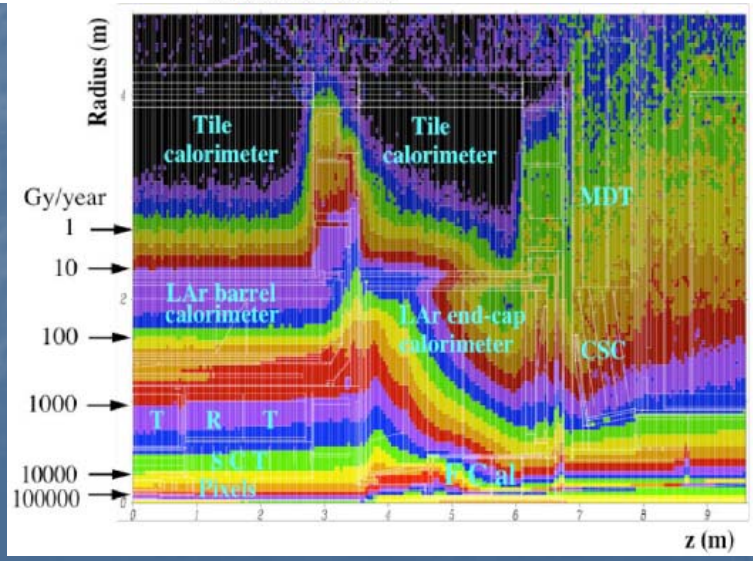
- Large scale system
- Risks
- Radiation “culture” and reviews
- Zones
- Failure types and recovery
- System Architecture, Simplify, redundancy
- Classification of systems and functions
- COTS, CPU’s, FPGA’s, ASIC’s
- Accumulating effects: TID, NIEL
- SEE
- Some examples
- The wall
- What to avoid
- Summary
- Source of information

# Large scale HEP system: ATLAS

- ~100M detector channels
- ~400k front-end ASIC's
  - ~40 different ASIC types
    (large majority in same technology)
- ~50k front-end hybrids
- ~20k optical links
- ~300 KW of power for on-detector electronics
- ~220 readout/trigger crates
- ~2000 computers
- Radiation (10 years): Pixel layer
  TID: ~1MGy
  NIEL: ~1x10$^{15}$ 1Mev n cm$^{-2}$
  SEE: ~1x10$^{15}$ >20Mev h cm$^{-2}$
- ~Thousand man-years invested in design, test, radiation qualification, production, installation, and final commissioning of electronics !
- Other experiments
  - CMS: Similar to ATLAS
  - LHCb: ¼ size, forward, "open", similar levels
  - ALICE: ¼ size, Much lower levels, SEE initially underestimated

Typically the same within factor 2 - 4 inside experiment. Order of magnitude differences in cavern !



# What risks to take ?

- Risk analysis
  - What local failures provokes what system  failures ?
    - This is often more complicated than initially thought
  - Can a given failure destroy other equipment
  - Hard and soft failures
  - How long does it take to recover from this ?
- Fold in radiation induced failure types and rates
  - Zero risk do not exist
  - Very low risk will be very expensive
  - Anything else than low risk will be unacceptable

# Radiation “culture”

- In large projects all project managers and electronics designers must be educated in this so they are capable of taking appropriate measures and feels responsible for this.
  - When ever making major/small design decisions all designers must continuously ask the question: what about radiation effects ?.
  - Speak common language on this “new” field.
- Expertise, advice and design “tricks” for radiation hard design must be exchanged between people in project (and from other projects and communities: Space, Mill, etc.).
  - Common/standardized solutions can/must be identified.
    - Use common rad hard technology and libraries
    - Use common ASIC components (e.g. particular link interface)
    - Use common COTS components (e.g. particular FPGA)
    - Use common design tricks (ECC, triple redundancy)
    - Common and standardized test and qualification methods.
    - Etc.
  - A radiation working group can be a means to obtain this.
    (can be termed electronics/radiation/? coordination group/board/committee , ,)
- Make everybody aware that designs will be reviewed and must be tested
- Make everybody aware of importance of radiation hardness policy and develop(enforce) consensus that this is for the best for us all.
- Clear guidelines: radiation levels, safety factors,testing methods, design methods, etc.

# Reviews

- Design reviews can be an excellent way to verify/enforce/encourage that radiation effects have been properly taken into account and tested for in each component/sub-system.
- By exchanging reviewers across sub-system boundaries a very useful exchange of ideas/approaches is accomplished.
  - A lot of this useful exchange actually then happens outside the review context
  - We are all here to make a full system that works well.
- Reviews should mainly be seen as an additional help to the designers to make radiation tolerant and reliable designs.
  - If used as a “police force” then it can be very difficult to get real potential problems to the surface.
- Radiation effects and consequences must be a permanent part of the review “menu”.
- Reviews
  - Global system architecture
  - Sub-system architectures
  - Designs (ASIC’s, boards)
  - Production readiness reviews (in principal too late to resolve problem)

# Radiation "zones"

- High level zones: 1MGy – 1KGy, $10^{15}$ – $10^{11}$ >2oMev h $cm^{-2}$ , 10 years (trackers)
  - Everybody knows (must be made to know) that radiation has to be taken into account and special systems based on special components needs to be designed/tested/qualified/etc.
  - TID, NIEL, SEE - Estimate of soft failure rates -> TMR
  - ASIC's
- Intermediate zones: 100Gy – 1kGy, $10^{11}$ – $10^{10}$ >2oMev h $cm^{-2}$ (calorimeters and muon)
  - ASIC's
  - Potential use of COTS
    - Radiation tests for TID, (NIEL), SEE (Use available tests if appropriate)
    - Special design principles for SEE (e.g. Triple Modular Redundant, TMR in FPGA's)
- Low level zones: < 100Gy, <$10^{10}$ >2oMev h $cm^{-2}$ (cavern),
  - Extensive use of COTS
  - TID and NIEL not a major problem
  - SEE effects can be severely underestimated
- Safe zones: < ? (LHC experiments: counting house with full access)
  - One has to be very careful of how such zones are defined. Do not confuse with low rate zones !!!

# Failure types to deal with

- Hard failures
  - Requires repair/spares
  - TID, NIEL, SEB, SEGR, (SEL)
- Soft/transient failures:
  - Requires local or system recovery
  - Can possibly provoke secondary hard failures in system!.
  - SEU
  - (SEL if recovery by power cycling)
    - Can/will be destructive because of high currents and overheating
    - Over current protection to prevent “self destruction”.

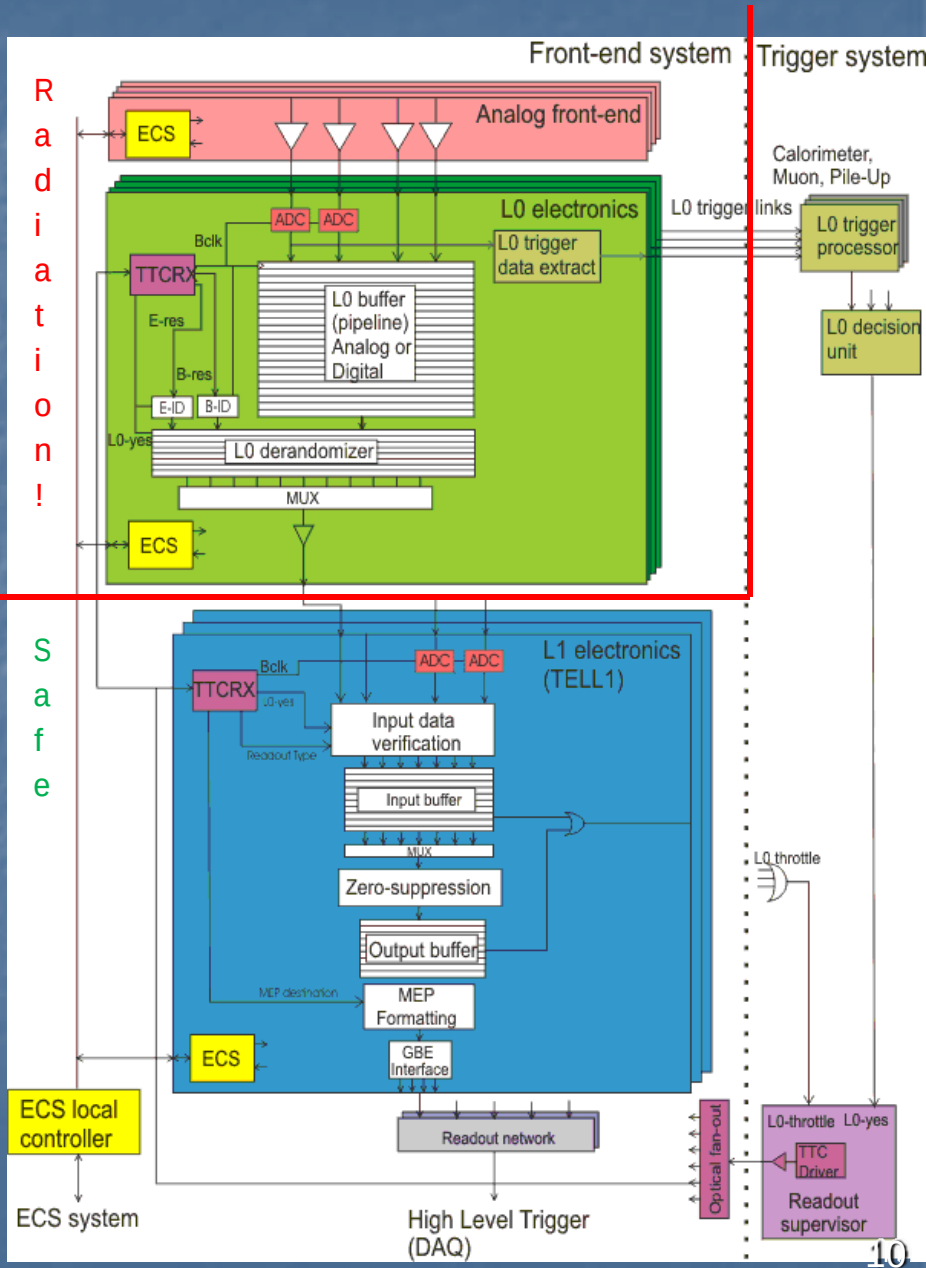
# Recover from problem

- Hard failures (TID, NIEL & SEE: SEB, SEGR,SEL)
  - Identify failing component/module
    - Can be difficult in large complex systems
  - Exchange failing component (access, spares, etc.)
  - Effects of this on the running of the global system ?
- Soft failures
  - Identify something is not working as intended -> Monitoring
  - Identify cause
  - Re-initialize module or sub-system or whole system
    - Simple local reset (remotely !)
    - Reset/resync of full system (if errors have propagated)
    - Reload parameters and reset
    - Reload configuration (e.g. FPGA's)
    - Power cycling to get system out of a deadlock (a la your PC)
- How much "beam" time has been lost by this ?

# System Architecture

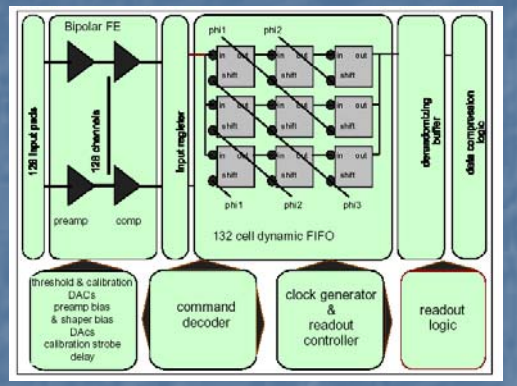
- Detailed knowledge (with clear overview) of system architecture is required to estimate effect of hard and soft failures on global system
- Which parts are in what radiation environment.
- Which parts can provoke hard failures
- Which parts can make full system malfunction for significant time ( seconds, minutes, hours, days, months)
  - Keep in mind total quantity in full system
- Which parts can make an important functional failure for significant time but global system still runs
- Which parts can make an important function fail for short time

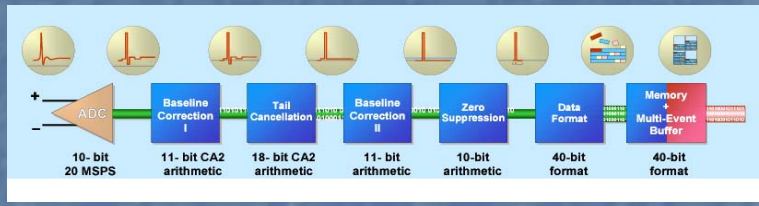
- Short malfunction in non critical function
- Central functions in safe area's



# Simplify, but not too much

- At system level:
  - Clear and well defined system architecture
    - Identify particular critical parts of system
  - It is much easier to predict behavior of simple systems for what can happen when something goes wrong.
    1. Simple pipelined and triggered readout
       - Require lots of memory and readout bandwidth (IC area and power)
    2. Complicated DSP functions with local data reduction/compression.
       - May require much less memory and readout bandwidth but is much more complicated.




  - Keep as much as possible out of radiation zones (within physical and cost limitations)
  - Do not simplify so much that key vital functions are forgotten or left out. (e.g. protection systems, power supply for crate, vital monitoring functions, built in test functions, etc.)
- At sub-system/board/chip level:
  - Nice fancy functions can become a nightmare if they malfunction (SEU's)

# Redundancy

- Redundancy is the only way to make things really fail safe
  - At system level (for all kinds of failures)
  - ……..
  - For memories: ECC (beware of possible multi bit error)
  - At gate level (for SEU)
    - Triple redundancy implementation: TMR
      - Significant hardware and power overhead
      - Certain functions can in some cases not be triplicated (Central functions in FPGA's)
      - How to verify/test that TMR actually works as intended
- Comes at a high cost (complexity, money, , ,) and can potentially make things "worse" if not carefully done.
- Use where really needed.

# Classification of equipment

- Personnel safety:
  - Must never fail to prevent dangerous situations
    - Redundant system(s)
  - Acceptable that it reacts too often ?
- Equipment safety:
  - Much of our equipment is unique so large scale damage must be prevented at “all cost”. Can it actually be reproduced ?
- Critical for running of whole system
  - Whole system will be un-functional if failure
  - Central system (Power, DAQ, etc.) or critical sub-system (e.g. Ecal, Tracker, , )
- Critical for running of Local system
  - Local system will be affected for some time but whole system can continue to run
  - Limited part of sub-detector (e.g. tracker with “redundant” layers)
- Uncritical
  - Loose data from a few channels for short time.

# Classification of functions

- Protection system
  - Have independent system for this
- Powering
  - Can affect significant parts of system
  - Can possibly provoke other serious failures
  - Must be detected fast
- Configuration
  - Part of system runs without correct function.
    - Can possibly occur for significant time if not detected
    - Continuously verify local configuration data.
      Send configuration checksum with data, etc.
  - Can possibly provoke other serious failures
- Control (e.g. resets)
  - Part of system can get desynchronized and supply corrupted data until detected and system re-synchronized.
    - Counters
    - Buffers
    - State machines
  - Send synchronization verification data (Bunch-ID, Event-ID, etc.)
- Monitoring (Separate from protection system)
  - Problematic situation is not detected, and protection system may at some point react.
  - Wrong monitoring information could provoke system to perform strong intervention (double check monitoring data)
- Data path
  - A single data glitch in a channel out of 100M will normally not be important

# Using COTS (Commercial Of The Shelf)

- For use in “low” level radiation environments, But what is “low” ?.
- Internal architecture/implementation normally unknown
  - Often you think you know but the real implementation can actually be quite different from what is “hinted” in the manual/datasheet
    - E.g. internal test features normally undocumented, but may still get activated by a SEU.
    - Unmapped Illegal states, , , , ,
    - Small difference in implemenation can make major difference in radiation tolerance.
- Assuring that all units have same radiation behavior
  - Problem of assuring that different chips come from same fab. and process/design has not been modified (yield enhancement, bug fix, second source, etc)
    - Chips: Buying all from same batch in practice very difficult
    - Boards/boxes: “Impossible”
- TID/NIEL: Particular sensitive components may be present but this can relatively ease be verified.
- Major problem for use in “low” radiation environments will be SEE effects:
  - Rate ?
  - Effect ?
  - Recovery ?
- PS: Radiation qualified space/mill components extremely expensive and does often not solve the SEU problem.

# Use of CPU's

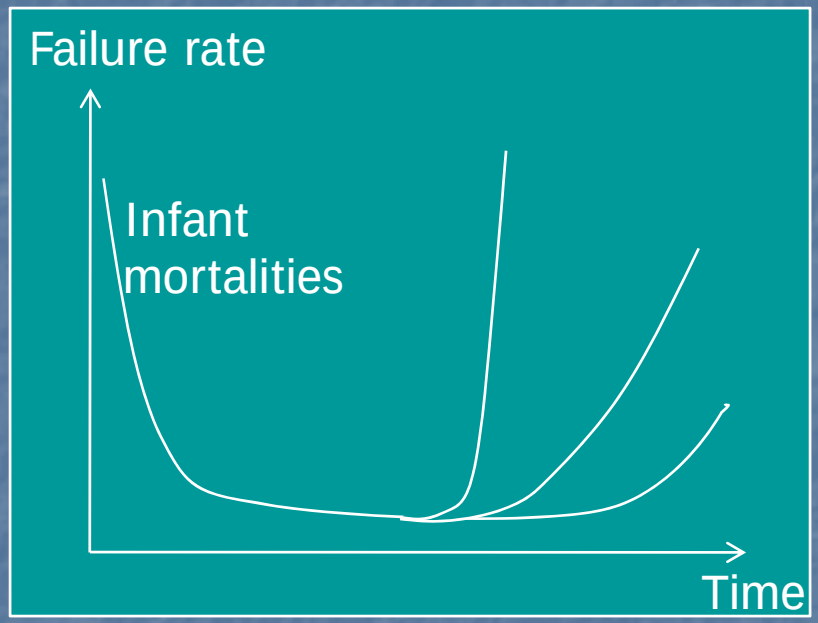
- Containing CPU's: PC, PLC, , ,
  - Basically any of piece of modern commercial electronics equipment contains micro controllers !.
- Contains: Program memory, Data memory, Cache, Buffers, Register file, registers, pipeline registers, state registers, ,
  - Different cross-sections for different types of memory
  - How large a fraction of this is actually "functionally sensitive" to SEU ?
- Complicated piece of electronics where radiation effects (SEU) are very hard to predict !.
  - Operating system
  - User code
- Schemes to possibly improve on this:
  - Special fault tolerant processors (expensive and old fashioned)
  - Error Correcting Code (ECC) memory (and Cache). What about multi bit flips ?.
  - Watch dog timers
  - Self checking user software
    - Does not resolve problem of operation system
    - Does not resolve problem of SEU in key control functions of CPU (e.g. program counter, memory)
- Do NOT use in systems where "frequent" failures can not be tolerated.

# ASIC's versus COTS

- ASIC's
  - High time, manpower and money costs
  - Cost effective if many parts needed: ~100k
  - Needed anyway for very specific functions (e.g. FE chips for experiments)
  - Can be made very rad hard (if done correctly)
    - Use of appropriate technology, libraries, tools and design approach (TMR)
  - Option: Flexible ASIC for use in similar applications (link chip, interface, etc.)
    - Requires a lot of communication across projects/groups/departments/experiments/ / /
- COTS
  - Lots of "nice and very attractive" devices out there.
  - "Cheap" but VERY problematic (and may therefore finally become very expensive).
  - Use with extreme care.
- FPGA's with tricks (mitigation)
  - Chose appropriate FPGA technology (common platform across projects)
    - Special space/mill versions, Antifuse, FLASH, RAM with scrubbing , , + TMR
  - Be careful with hidden functions: SEFI (test, programming, ,)
  - Be careful with intelligent CAE tools that may try to remove TMR redundancy (synthesis)

# TID and NIEL

- In principle this could be seen as early wear-out of electronics.
  - But much earlier than expected
  - Starts in small hot regions or in particular batches of components.
    - Use available spares to repair
    - Functions may possibly “recover” after some time without radiation because of annealing
  - But will then quickly expand like a virus
    - Spares run out and component obsolescence problems then gets a serious problem.
- Before systems clearly “hard fail” they may be seen as unreliable/unstable (and no evident cause of this can be found)
  - Decreased noise margins on signals
  - Timing races
  - Increased leakage currents -> power supplies marginal
  - Etc.

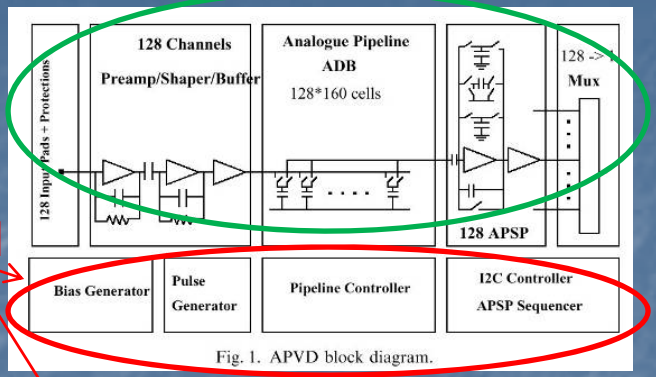


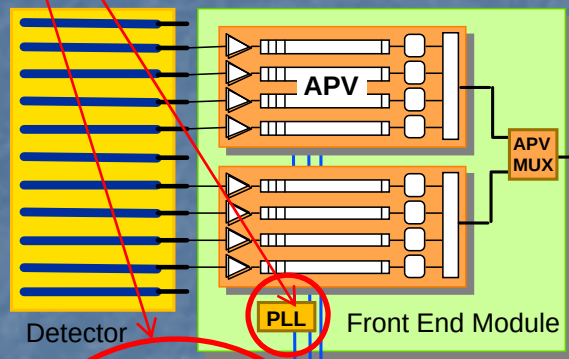
# SEE

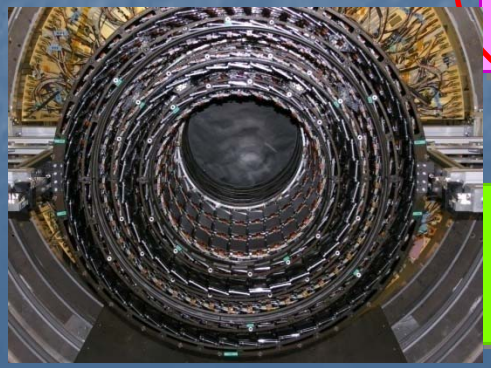
- Systems have sporadic failures that happens randomly and at no well identified locations.
  - One may possibly have a few particular sensitive parts.
- “Normal” (non rad induced) sporadic failures are identified and corrected during system and sub-system tests and final commissioning of full system.
  - Already indentifying cause of sporadic failures at this level and resolve them is very difficult.
- SEE failures will only be seen when radiation present (after system commissioning) and then it is often too late to react as resolving this may require complete system re-design.
  - May possibly be bearable during first initial low luminosity but then .....
  - We have not yet had any of our large scale systems running under such radiation conditions.
  - Surprises may(will) be ahead of us even if we have done our best to be immune to these.
- They hit you at exactly the time you do NOT want system failures to occur.

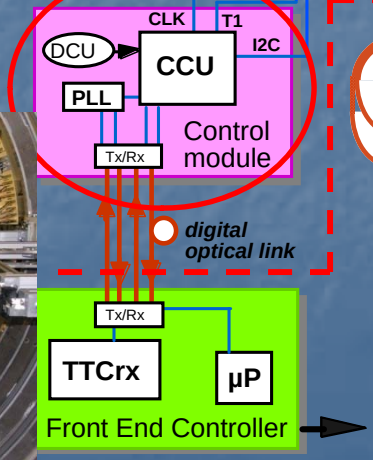
# CMS tracker front-end

- Uncritical (SEU)
  - Data path
  - (Monitoring)
    - Single corrupted read can be redone
- Critical (SEU)
  - Configuration/control
  - Data path control
    - Buffer controller
    - Readout controller
  - Synchronization
    - PLL
    - Resets
    - Event tag identifiers
- Vital (but not seen)
  - Power
  - Protection (thermal)



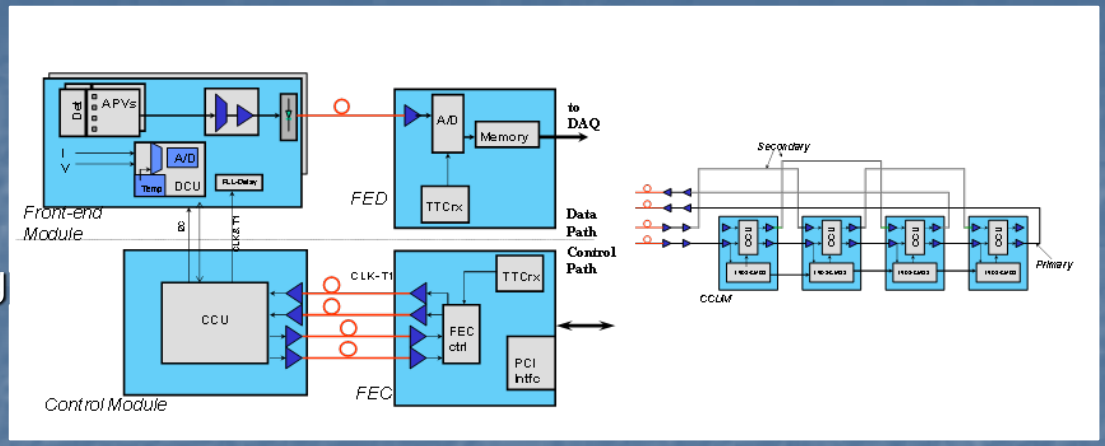



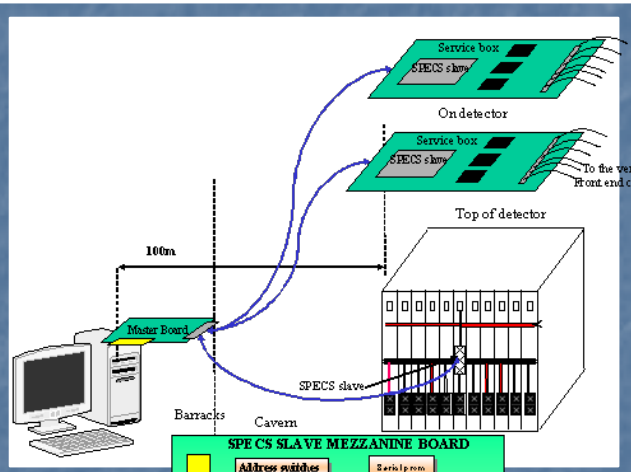



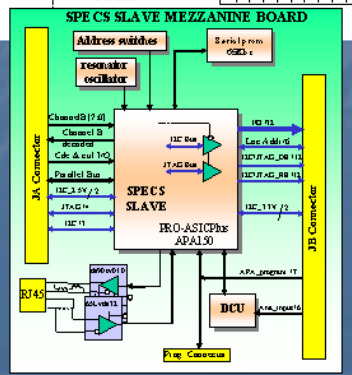


# Control/monitoring

- CMS tracker: CCU, DCU
  - Inside detector
  - Radiation hard: ~1MGy
  - SEU immune
  - Redundant token-ring routing
  - ASIC's based
- LHCb: SPECS
  - At periphery of detector
  - Radiation tolerant: ~1kGy
  - SEU "immune"
  - FPGA implementation with well chosen FPGA family (radiation tested) and TMR
- ATLAS general monitoring: ELMB:
  - In cavern and muon systems
  - Radiation tolerant: ~100Gy
  - SEU sensitive, but self checking functions
  - Processor with watchdog and special software.

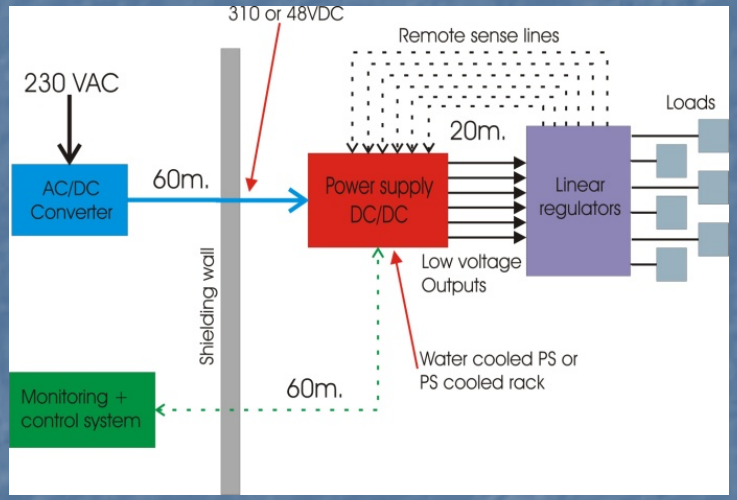


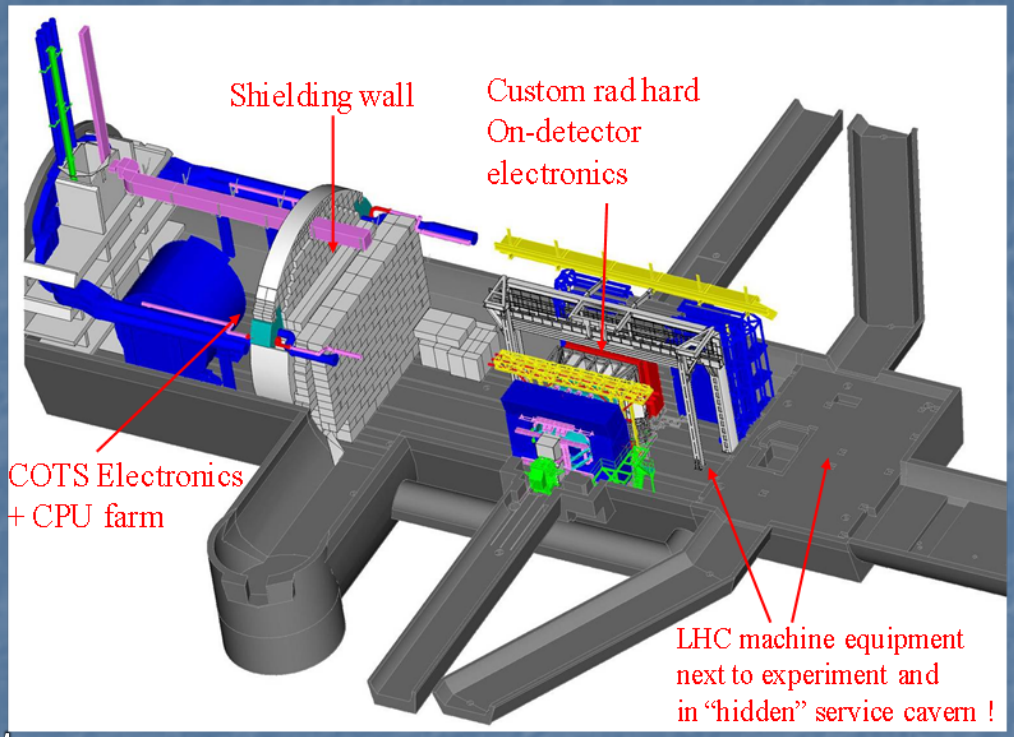
# Power supplies

- Worries:
  - Destroy our electronics if over voltage/current
  - Destruction of power supply itself
  - Loose control/monitoring requiring power cycling
- Wiener: Keep intelligence out of radiation zone
  - Radiation zone
    - DC/DC converter: 400VDC – LV (De-rated HV power transistors tested)
  - Outside radiation:
    - Control and monitoring (CPU)
    - PFC (Power Factor Correction): CPU and high power devices
  - Cost
    - Cabling -> relatively few “high power” channels
    - Limited control and monitoring
- CAEN: Keep HV out of radiation zone
  - Radiation zone:
    - 48v DC – LV, possibility of many low/med power channels.
    - Local processor with dedicated software (SEU !)
  - Outside radiation
    - AC/DC (48V)
    - PFC compensating filter
    - (rad tolerant version required for high power systems as power loss of 48v on ~100m to large)
- Extensive use of rad tolerant low voltage drop linear regulator developed by STM in collaboration with CERN
- Custom:
  - Atlas calorimeter



# The Wall (not by Pink Floyd)

- LHCb experience:
  - Physical (thin) walls does not make problem disappear.
    - What you do not see, you do not worry about.
    - When you see it, it is too late.
    - Lots of concrete needed to give effective radiation shielding.
  - Political/organisatorial walls does not make things better




- All participants in global project (experiments + machine + ?) must be aware of the potential problems.
  - Extensive exchange of information/experience
  - Key part of project requirements
  - Reviews

# What to avoid

- Underestimate the problem
- Safety systems in radiation
- Forget that local errors can propagate to the system level and make the whole system fall over (very hard to verify in advance for complex systems)
- Assume that somebody else will magically solve this.

- Complicated not well known electronics (black box) in radiation environment
  - Computers, PLC, Complicated Communication interfaces , ,
- High power devices in radiation zones
  - SEE effects can become “catastrophic”
- Particular known weak components
  - Some types of opto couplers, etc.
- Uncritical use of complicated devices (e.g. FPGA’s)

# Summary

- Expensive
  (If forgotten then much more expensive)
- Takes significant efforts across all systems, levels and hierarchies.
  - Radiation culture, policy and reviews
- No easy solutions in this domain
- “The Devil is in the details”

- Still waiting for unexpected problems when final beam, despite huge efforts done in the experiments in this domain.

# Source of information

- Experiment web pages:
  - LHCb:
    - Electronics coordination: http://lhcb-elec.web.cern.ch/lhcb-elec/
    - Radiation policy: http://lhcb-elec.web.cern.ch/lhcb-elec/html/radiation_hardness.htm
    - Reviews (many): http://lhcb-elec.web.cern.ch/lhcb-elec/html/design_review.htm
  - ATLAS Radiation policy: http://atlas.web.cern.ch/Atlas/GROUPS/FRONTEND/radhard.htm
- Elec 2005 lectures: https://webh12.cern.ch/hr-training/tech/elec2005___electronics_in_high.htm
- FPGA's: http://klabs.org/fpgas.htm
- Summer student lecture: http://indico.cern.ch/conferenceDisplay.py?confId=34799
- Google it on internet.