The ATLAS Pixel B-Layer Replacement Program

K. Einsweiler, LBNL

Goals for the ATLAS B-Layer Replacement:

- Describe goals, understanding that may only succeed in addressing some of them.

Corresponding technical issues:

- Sketch out key R&D areas, and ties to SLHC R&D program.

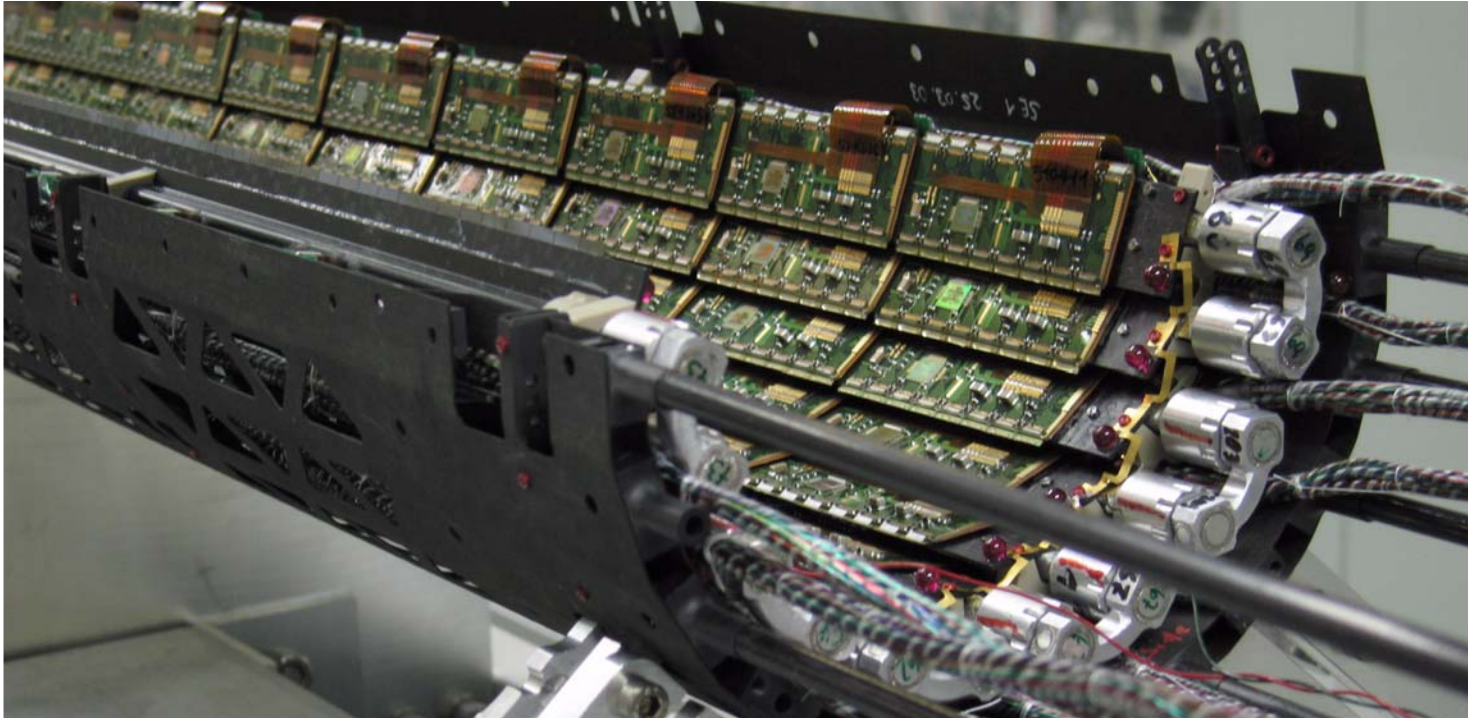
Schedule:

- Goal is to be ready for installation by Fall 2012.

Present ATLAS Pixel B-Layer

Innermost layer of the ATLAS Pixel Detector:

- Historically called B-layer. Sensor radius is 50mm, and layer consists of 22 carbon-carbon staves (11 evaporative cooling circuits), each supporting 13 modules:



- Total of 286 modules (16%) with 20 degree tilt angle, 13.2M channels, active area roughly 0.29 m^2, worst-case end-of-lifetime power load as high as 2.4kW.
- Features: two data fibers/module at 80Mbit/s each, all cooling connections on C-side (historical), operation to 10^{34} luminosity with 99% single hit efficiency.
- Staves are mounted inside carbon-fiber half-shells, which clamp to form the layer.

B-Layer Replacement Concept:

Justification:

- With nominal luminosity profile, expect B-layer performance to start degrading after 2-3 years at LHC design luminosity or about 300 fb^{-1} (10^{15} NIEL dose, 50MRad ionizing dose). Expect reduced efficiency and modest reduction in point resolution.
- The performance of the B-Layer has a large impact on ATLAS physics performance, particularly for B-tagging. On the timescale of next several years, expect that improvements in technology should allow construction of a B-Layer with improved segmentation, greater radiation hardness, and reduced material.
- Propose to prepare an upgraded B-Layer for installation during the Winter 2012/ 2013 shutdown, after roughly 4 full years of ATLAS operation.

Alternatives:

- Minimal “upgrade” would be to build essentially identical modules and staves. Expect to acquire enough 0.25µ wafers for FE/MCC to do this if necessary.
- Preferred scenario involves improvements to sensor/electronics design, to module and mechanics geometry, and operations with greater occupancy and total dose.
- Hope for beampipe radius reduction from 29mm to something in the range of 17-25mm (R. Veness, Liverpool meeting). With corresponding reduction in clearances, could optimistically achieve a 30mm B-layer radius.
- Consider either single small R layer, or double layer with R about 30mm and 55mm.

Goals for B-Layer Replacement Program

Principal goals are:

- **Reduction of material**, required to take full advantage of point resolution. This could use a combination of improved power distribution (reduced electrical services) and improved active fraction for the basic modules (closer to 90% rather than the present 71%). Present best estimate for pixel layer now is 2.5% X_0 per layer. Would like to target between 1.5% X_0 and 2.0% X_0 per layer.
- **Improvement of segmentation**, useful to cope with higher occupancy and provide improved point resolution in one or both measurement coordinates. Ideally, would like to reduce pixel area by a factor 2-3. What is the optimal aspect ratio ?
- **Increased radiation tolerance**, both for higher instantaneous luminosity and for higher total dose tolerance. Set nominal goal of a factor 3 increase, leading to instantaneous rate of 1×10^8 $cm^{-2}s^{-1}$, and a total dose of 3×10^{15} neutron equivalent. This is an intermediate step to SLHC, and would be consistent with operation at 30mm radius and the present LHC design luminosity up to SLHC (2016) period.
- **Improved layout geometry:** consider an ambitious geometry upgrade, with an inner measurement at 30mm and an outer measurement at roughly the present B-layer location, based on a highly-integrated double layer.

Corresponding Technical Issues and R&D

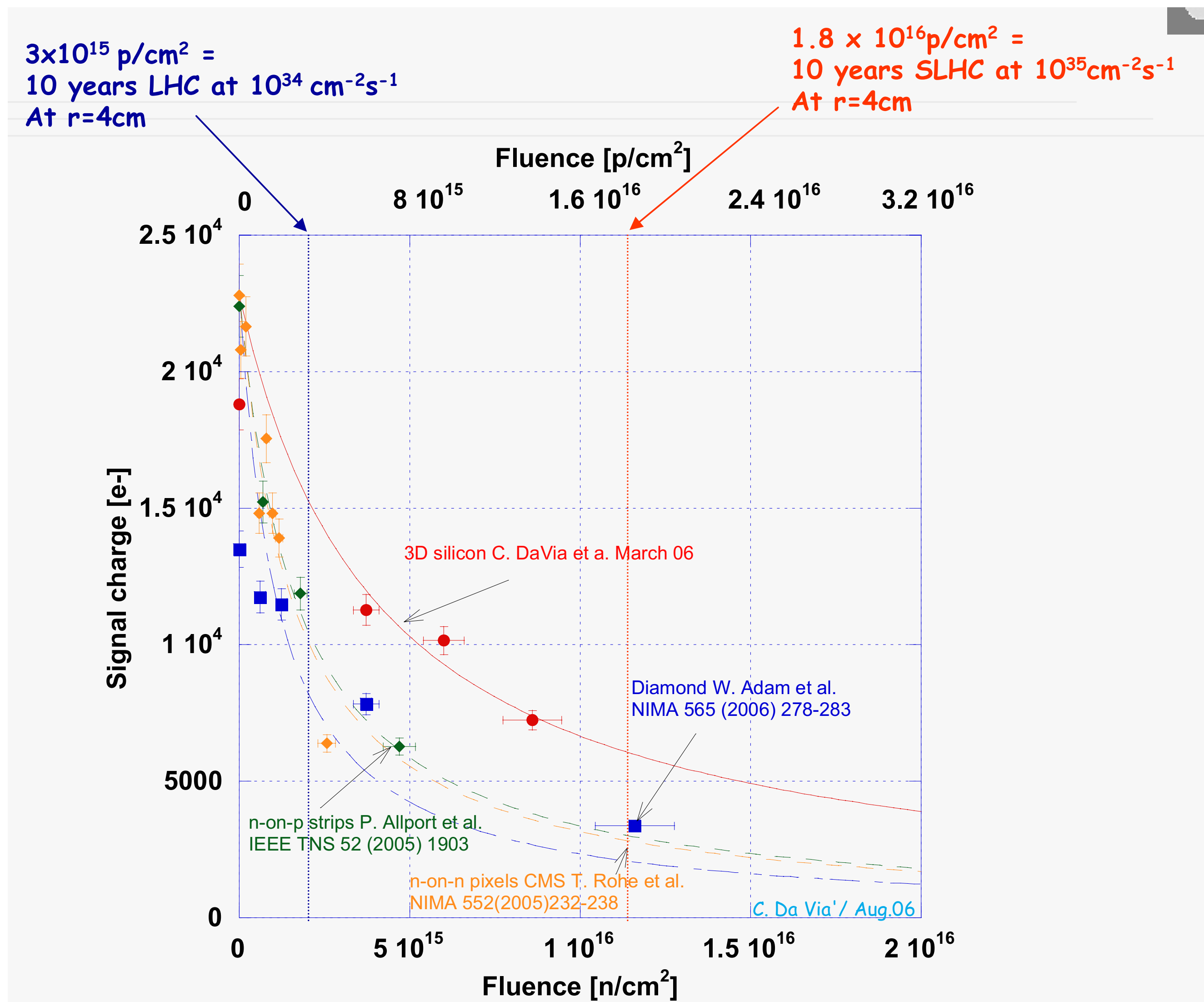
Services (11 cooling circuits and 286 module connections):

The B-Layer replacement should be compatible with existing services at the PP1 interface (end of pixel package at +/- 3.5m). **Need to evaluate the real limitations.**

- **HV bias distribution:** limited to about 1KV maximum. However, silicon sensor performance after high doses is limited by trapping, and improvements saturate at about 2-3 V/μ, so the services are well-matched to this limit.
- **LV distribution:** currents cannot be significantly increased with present services and linear regulator approach at PP2. However, implementation of a DC-DC converter approach on-detector could allow providing significantly more current at the 1.2V that will be typical of next generation micro-electronics designs.
- **Cooling infrastructure:** will act as a constraint on any changes. Maximum pressure and mass flow constraints down to PP1 need to be evaluated. Ideally, run colder to limit leakage currents, but evaporative systems with lower temperatures imply higher worst-case pressures (C_2F_6 about 30bar, CO_2 about 60bar).
- **Fiber infrastructure:** all multi-mode, and a mixture of rad-hard SIMM and commercial GRIN fiber. Tests indicate bandwidth is OK to above 1Gbit/s.
- **From PP1 inwards:** should change service panels and/or opto-boards. Not trivial since the 2-hit staging re-design placed the “staged” parts (middle disk and L1) on the ISP, instead of the B-layer as originally designed. Parts that would need changing are distributed, so lots of work required !

Sensors:

- Basic goal: **increase total dose tolerance by a factor 3**. Baseline would be to use the present sensor design, but it would be marginal, or would require more aggressive FE electronics designs (sensors limit present B-layer lifetime):



Compilation figure from recent talk by Cinzia Da Via.

Predicts 15Ke signal after 10^{15} NIEL for n+ on n sensors (agrees with our irradiations).

Same sensors would have only about 7Ke signal left after dose of 3×10^{15} NIEL.

Note 3D sensors (70μ gap) predicted to have 13Ke after same dose, and still 7Ke after 10^{16} NIEL !

- The dose increase will produce a linear **increase in the leakage current** at the present sensor temperatures. Ideally, run at least 10C colder (-15C to -20C).
- **Current barrel cooling is marginal:** thermal effects may define detector lifetime near 10^{15} dose. If sensor thermal runaway occurs, must use HV current limit to control leakage power, decreasing sensor bias voltage and signal. Note 3D sensor leakage power lower (lower bias voltage), could possibly keep C3F8 cooling fluid.
- The **depletion voltage will also increase** significantly from the 400V value we find with our present planar sensors at 10^{15}. Will almost certainly have to operate partially depleted if a planar geometry is used.
- **Real issue is charge trapping**: not yet amenable to improvement by defect engineering, and which can only be modestly reduced by increasing bias voltages.
- Propose **minimum signal size in the range of 8-10Ke** at end of lifetime is necessary for B-layer replacement to achieve 99% single-hit efficiency.
- Operating a 100M+ channel system requires very low occupancy (10^{-6} to 10^{-8} hits/channel/crossing) leading to a threshold of about 3Ke. In addition, the limits on peaking time and time slewing in practical designs further increase this by 1-2Ke.
- The critical operation parameter is the in-time threshold. The in-time threshold is defined such that hits from threshold to 100Ke fall within 20ns window. The present detector has an in-time threshold of about 5Ke for nominal biassing conditions.
- Full efficiency near pixel boundaries (within about +/- 5μ), where charge sharing is important, requires an in-time threshold below about 50% of the mean signal value.

Micro-electronics process and layout Issues:

- First IBM 0.13μ test chip built and tested in 2004. Have adopted CMOS8RF (0.13μ) process with LM metallization as baseline. MOSIS is official IBM prototyping vendor, then expect to pass through CERN for engineering and production runs.
- Next generations (CMOS9 = 90nm, and CMOS10 = 65nm) available now or later this year, likely to present significant challenges for low-power analog design due to off currents and tunneling currents. For now, stay with 130nm CMOS.
- RF processes from IBM are second generation CMOS, with full analog characterization, high-quality analog design kit, and useful technology options. CMOS8RF_LM is a Cu process, with 8 metal layers (up to 3 thick metal layers).
- Other useful options include: LP (low power) FETs with higher thresholds (like 0.25μ) and very low off-currents, triple-well NMOS for substrate isolation of critical devices, and the fact that it is a 1.5V process (max V = 1.6V).
- Prices under CERN frame contract are attractive, and roughly double those for 0.25μ (mainly due to increased mask costs).

Need to define layout rules !

- CMOS8RF design kit is now based on very complex BSIM4 models, where proper modeling of annular devices ("gate all around") looks very difficult.
- Tentative choice: linear devices, with NMOS guard rings, for custom layout. Artisan standard cells for synthesis of digital blocks. ***Significant risks above 100MRad !!!***

FE Design Requirements/Goals:

Radiation environment and occupancy:

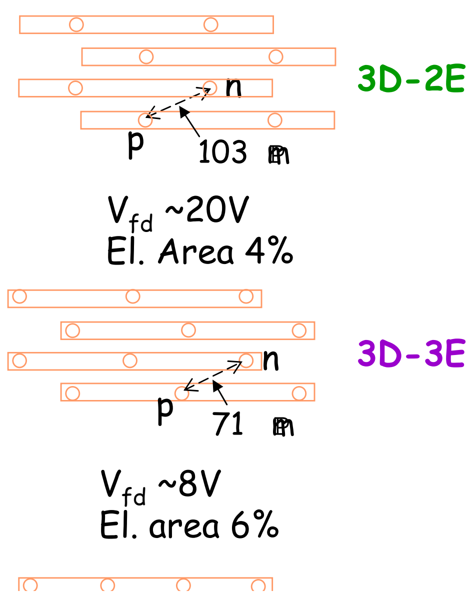
- Goal is 3 times FE-I3, or about 150MRad total dose, and 1×10^{8} particles/cm^{2}/sec instantaneous rate. This would allow operation at 10^{34} and 30mm radius.

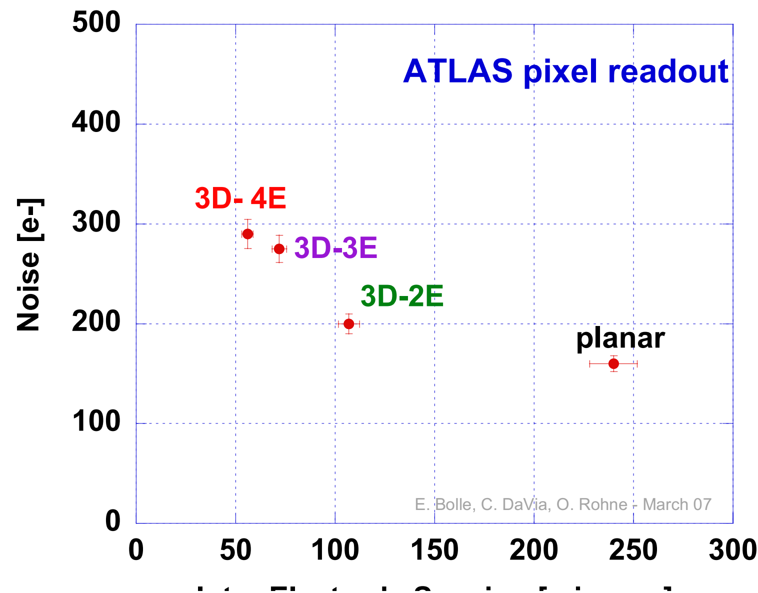
Analog front-end parameters:

- Goal is current consumption of less than 20μA per pixel. Still discussing goal for power supply voltage: try to meet specs at 1.2V, but allow operation up to 1.5V ?
- Noise should be less than about 400e and threshold dispersion less than about 200e, leading to a total threshold "variation" of about 500e worst case.
- Want to achieve peaking time of roughly 20-30ns, although real specification is for the total timewalk performance of the analog front-end and discriminator of an overdrive (charge above threshold to be within 20ns) of less than about 1.5Ke.
- Double pulse resolution should be about 400ns to avoid contributing significantly to hit losses at maximum occupancy (less than 0.5%).
- Ideally, would maintain a charge measurement of best quality consistent with other performance constraints.
- A TOT implementation is very limited by double pulse time specification unless timestamping frequency is increased (requires more power). A "flash ADC", or multi-threshold discriminator, consumes lots of area and provides very limited measurement range, requires speed/power optimization.

Sensor-related parameters:

- Nominal sensor definition is CDet=400fF (best estimate for our current production modules), but appears to be somewhat larger for 3D sensor case:





Recent lab measurements using FE-I3 by E. Bolle and O. Rohne, plot from Cinzia DaVia.

For 200-250µ pixel, would estimate ENC of about 200e ?

- Assume worst case of 100nA leakage current (50x200µ pixel after $3x10^{15}$ NIEL, annealed, at about 0 C operating temperature), should be conservative.
- Note: in the case of FE-I3, a modular design was used to deal with larger CDet for "special" pixels, in which a "modular" input FET was used, and a "modular" IP bias multiplier was used. Metalization to select x2 or x3 for the input FET W/L, and x2 or x4 for the bias multiplier in each cell. Could optimize for up to about 1000fF with acceptable performance (worst case has almost double the total analog current !)
- Similar technique could be used to adapt to planar versus 3D sensors.
- Propose **minimum signal size in the range of 8-10Ke** at end of lifetime.

Geometry parameters:

- Fix pitch to 50μ. Little justification for reducing the 50μ pixel pitch, provided we keep analog readout ability. This pitch allows more flexibility in sensor technology (interpixel isolation at high dose, or 3D implementations, are more challenging at smaller pitches), and does not require any bumping technology development.
- Work to minimize length in the other direction, with a goal of 200μ, and a more realistic target of 250μ. These values should be used for simulation studies to evaluate the benefits.
- **Goal is to increase the module live fraction** (fraction of surface area which is covered by active sensor area) as much as possible.
- For a design with a single row of much larger FE chips, optimized for minimum inactive circuitry at the periphery, and using "active edge" technology in the sensor, a module active fraction above 90% could be achieved.
- Maximum die size for FE chip is 19.5mm x 21.0mm in the target process, which is 5 times our present die area. An intermediate size of about 16mm wide (about 64-80 columns) x 19mm long (about 320 rows) is a better target. This would give about 20-25K channels per FE chip, and an active area of about 260mm^2 (4 chips would cover present 10cm^2 active area module tile). Could also integrate single-chip modules directly onto a highly integrated stave structure with power management.
- Scaling die size by two in both directions raises many FE design issues (analog bias distribution, clock and power distribution, readout architecture).

Analog Front-End Development:

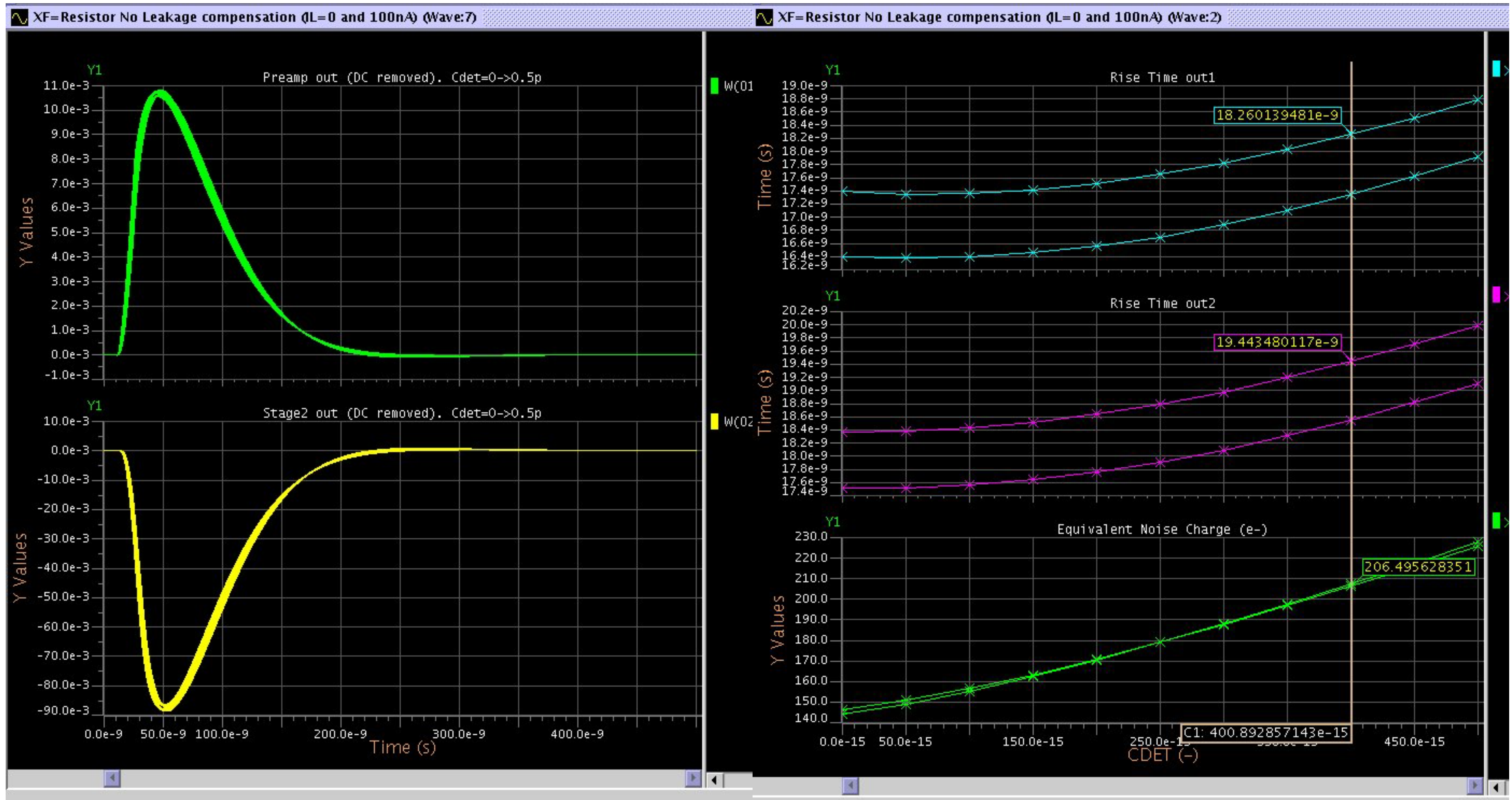
Overall Concept (designer is Abder Mekkaoui):

- Use a charge preamplifier AC-coupled to a closed loop 2nd stage. The 2nd amplifier is DC-coupled to a comparator with global threshold setting and local per pixel threshold tuning capability based on 5-bit DAC.
- This is similar to what was used in FE-I3, except the second stage was DC-coupled, and a 7-bit threshold adjust was used (this was probably overkill).
- After initial design studies, decided to pursue 2 designs through prototype stage.

Design 1 ("new, fixed shaping design"):

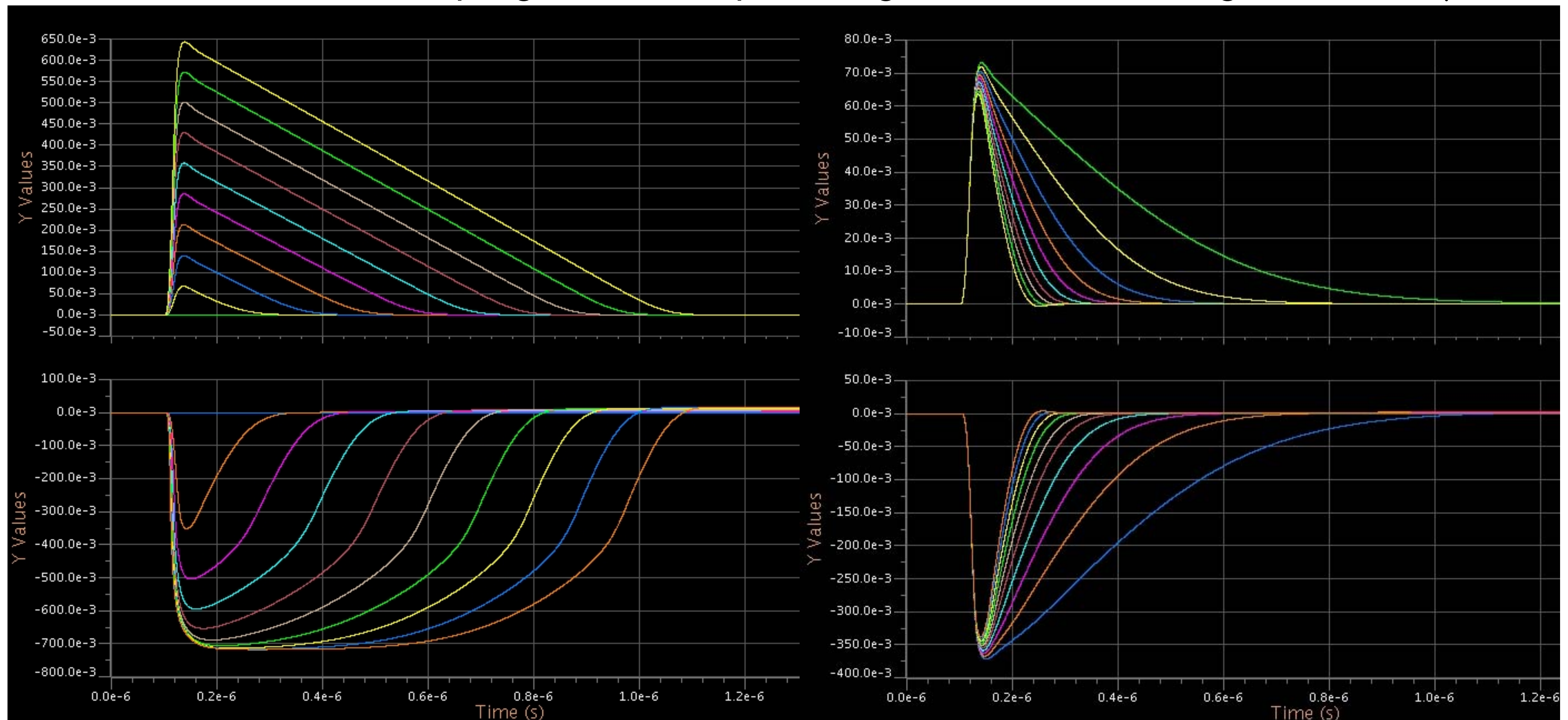
- Uses a triple-well NMOS for the input device. The second stage uses a PMOS input. Total current in 2 stages is about 18µA at nominal biassing.
- The preamplifier uses an OP poly resistor of about 1.5MΩ for the feedback, with a large feedback cap of about 50fF.
- There is no dedicated leakage compensation - the DC input current shifts the preamp output, which is absorbed by the AC-coupled second stage, resulting in a modest loss of dynamic range. Negligible change in noise up to 100nA leakage.
- The return to baseline time is fixed to about 200ns, and the charge measurement must be made with a "flash ADC' approach to provide perhaps a 3-bit measurement. Design is optimized for simplicity and performance without TOT.

- Noise is predicted to be about 200e for standard 400fF load, with about 20% contribution from resistor.
- Curves on left below show preamplifier response for leakage of 0nA and 100nA, and various CDet values to 500fF.
- Curves on right below show preamp and second stage risetime (18-20ns) and ENC (140e to 230e) for CDet values from 0fF to 500fF:



Design 2 ("new, TOT design"):

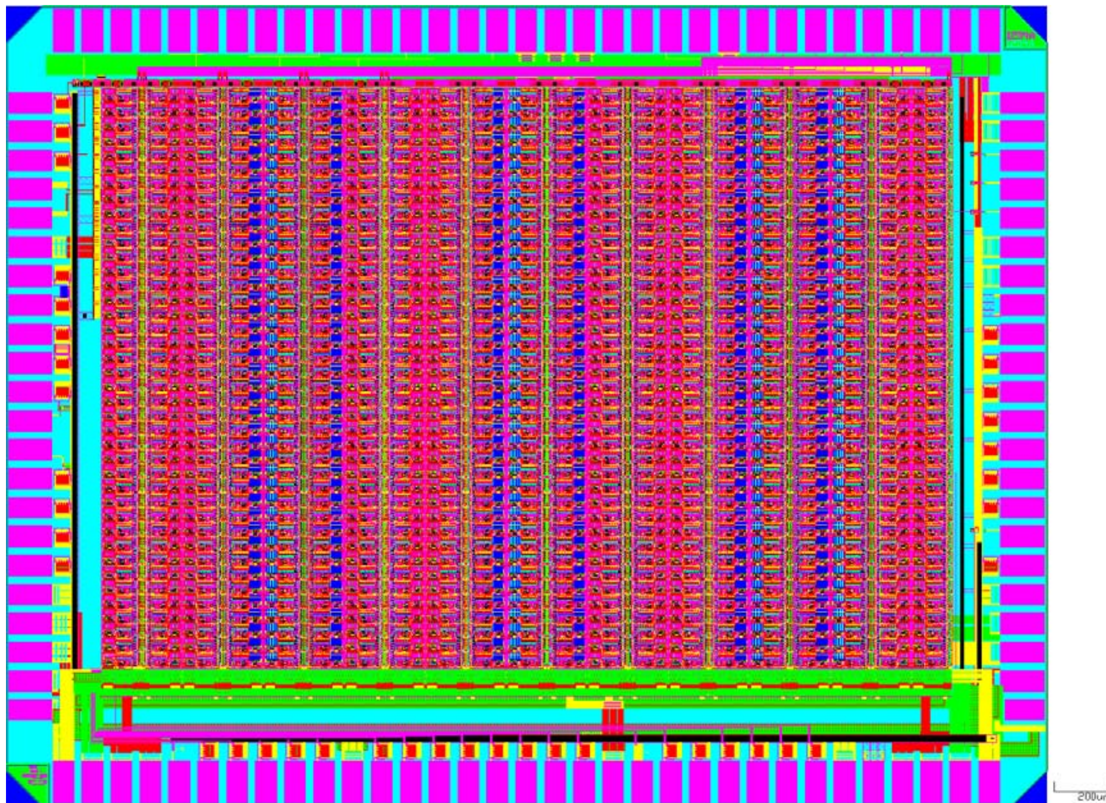
- Uses the same preamp design, but replaces the resistor in the feedback by two FETs, and an active, differential, leakage compensation circuit.
- In this case, the noise depends on the leakage current, varying from about 150e for no leakage to about 200e for 100nA worst-case leakage (for CDet = 400fF).
- Feedback current is programmable, providing a limited TOT range of order 1μs.



- Left plot is preamp/second stage for charges from 5Ke to 50Ke (20Ke is 500ns TOT). Right plot is for charge of 5Ke and feedback from 2nA to 20nA.

New Test Chip submitted in Feb 2007:

- Contains three minor variations on two basic designs, arranged in 21 columns of 40 pixels each. A simple shift-register readout is used, with an external gate.
- Implements a total of 14 bits of control per pixel, 5-bit TDAC (threshold) and 5-bit FDAC (feedback), plus 3 control functions (injection and masking).
- A total of 15 external bias current inputs are provided. Some columns have additional load capacitances and leakage injection circuits.
- Total die size is 3.6mm x 2.8mm, with 110 I/O pads (MANY test input/outputs):



Overall FE Integration Issues:

Standard Analog Blocks:

- Need to convert and update critical analog blocks, such as current-mode DACs, current references, charge injection circuits, LVDS-like I/O pads, general I/O pads and integrated ESD protection. **Much of this is Common 0.13μ Infrastructure !**
- Need to resolve issues of accurate voltage distribution within analog matrix, and related problem of accurate bias current distribution within matrix.

Architecture:

- Need to develop more efficient readout architecture for high-occupancy environment. This will involve using L1 trigger information at an earlier stage, to minimize data transfer requirements.
- In general, also want to move circuitry from the periphery to the "active" part of the pixel matrix, in order to optimize active fraction. This will involve more complex heirarchy between pixel and column-pairs, with distributed storage blocks.

SEU Tolerance:

- Need to develop both SEU-tolerant static storage cells and SEU-tolerant state machines and other clocked logic. Need to evaluate real upset cross-sections.
- In-pixel circuits are hand-optimized layout, peripheral circuitry can be more conventional place and route. Significant effort is required.

Power Distribution - reduce material and complexity:

- **Present power distribution scheme has a high service burden**. It requires 4 wires per low voltage, and a high current capacity to bring it to the module, .
- In general in our system, we pay a very **high price for the high modularity** we have implemented, with full services going to each module. This has often proved useful, but it may be too high a price to pay in the future. Two more aggressive approaches are under investigation, though combinations are also possible.
- **DC-DC converter scheme:** bring in power at a higher voltage and lower current (roughly 4 times the target voltage), and then convert to low voltage/high current close to the module using a DC-DC converter. Since the tracker is immersed in a 2T B field, the most promising technology for DC-DC conversion uses switched capacitors. Requires a specialized process to manage relatively high voltages and high currents in a high-radiation environment. Power efficiencies of 70-80% appear possible at switching frequencies of about 5 MHz. See talk of M. Garcia-Sciveres.
- **Serial powering scheme:** uses local linear and shunt regulators to allow operation of the FE electronics from a single LV supply, and also allows connecting modules in series using a single constant current supply. This approach is a significant departure: it is a “ground-less” system, where all control signals must be AC-coupled. Prototyped in pixel community with very encouraging results.
- Both schemes require more work at system level, and are best implemented in a highly integrated “stave”, where powering infrastructure is closely integrated with mechanics and cooling. **More integration => more engineering !!!**

Opto-links and Off-detector Electronics:

- Present B-layer uses two 80Mbit/s links (both edges of 40MHz clock used). To cope with an occupancy which is 3 times higher, **need bandwidth for optical links of roughly 500 Mbit/s**.
- Need to implement clocking in the range of 4-16 times 40MHz, either using a faster down-link (but distribution of clocks in range 160-640MHz requires more power) or using local frequency multiplication (but large-scale synchronization harder).
- Present location of opto-links is preferred (radius is about 15-20cm), but would need to work on good LVDS-like driver/receivers with 1.2V or 1.5V supply voltage.
- See talk of G. Darbo on bandwidth limits of present electrical (twisted pairs)/optical (SIMM/GRIN fiber) infrastructure. First studies indicate OK up to about 1Gbit/s.
- An issue, affecting the present detector as well, is limited drive voltage available from DSM processes for operating the VCSELs required for data transmission off of the detector (2.5V supply is marginal, even with best VCSEL technology). May require VCSEL driver development in more specialized processes in the future.
- Present off-detector electronics processing capability per ROD (number of modules or equivalently number of Mbit/s per ROD) is limited by the S-link bandwidth.
- In present B-layer, one opto-board on-detector (6/7 160 Mbit/s links) is processed by a dedicated 9U ROD module due to 1.6 Gbit/s S-link. The bandwidth would need improvement. Would benefit from **S-links operating at 5-10 Gbit/s**.
- **Basic elements are OK - mainly (common) electronics design issues !**

Schedule

Overall Schedule:

- Began development in 2004, after completing existing FE-I3 production pixel chip.
- Model for B-Layer Replacement assumes a lifetime of about 300 fb-1, after which significant signal loss in sensors begins to have an impact.
- Assume such luminosities could be accumulated over high-luminosity runs in 2010-2012, and therefore natural schedule is to replace in 2012/2013 shutdown.
- The replacement operation is complex, and will involve significant work around highly activated components (and risk of damage to other elements of detector).
- Minimum estimated replacement time is about 6 months, and realistically, it could be closer to 8 months. This is probably not a standard “annual” shut-down.

Provisional Milestones:

- Define key parameters by Oct 07 (pixel and FE chip geometry), hold R&D Review
- Test submissions in early 07 and 08, and finalize sensor choices (e.g. planar versus 3D), then build engineering run of full-scale chip by early 09.
- Evaluate modules built with prototype FE chips and sensors in 09.
- Overall B-Layer Replacement TDR in July 09, FE Electronics PRR in Mar 10.
- Already an aggressive schedule for significant development, so fall-back is always to go ahead with something quite close to the present B-layer design.