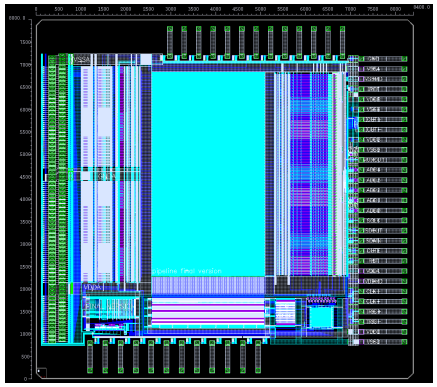
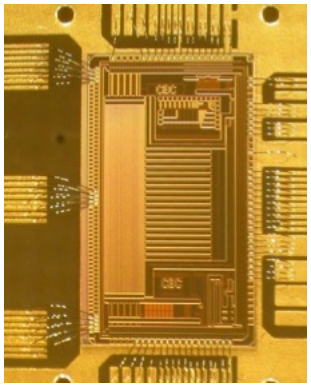
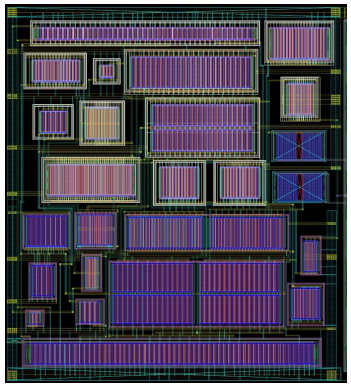


250 nm



130 nm



28 nm

ASIC development for HEP – past, present and future

April 2023

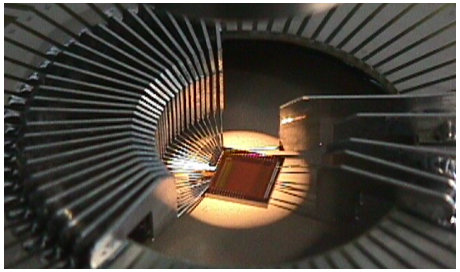
Geoff Hall

Imperial College London

Overview

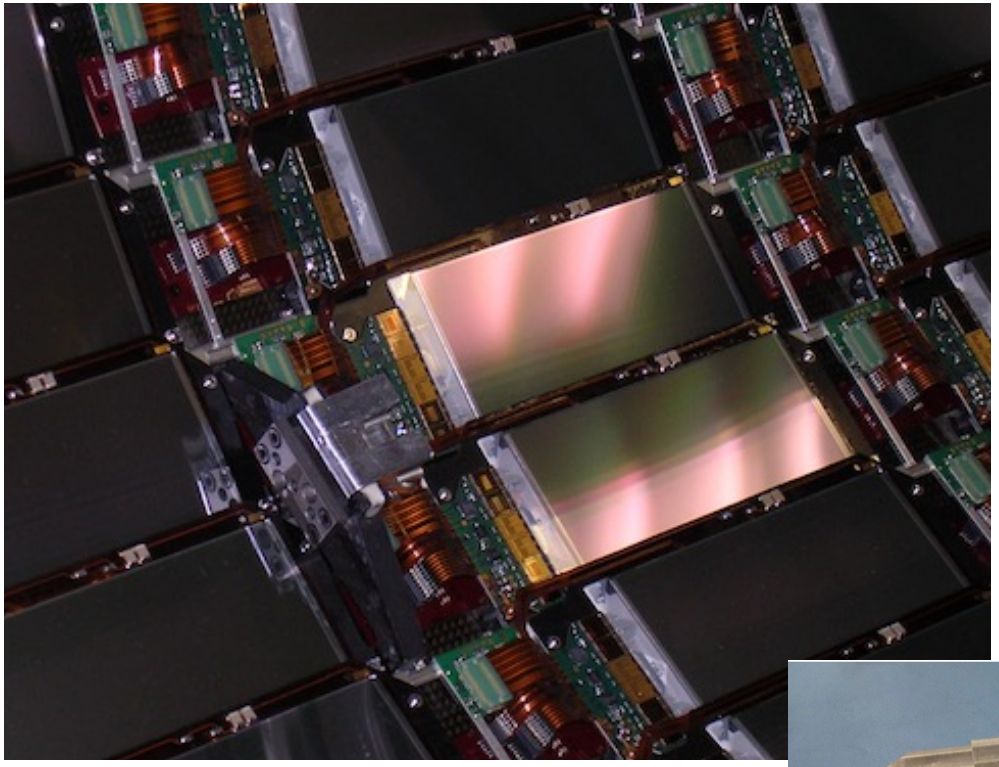
- ASICs are now a vital part of (almost?) all experiments
  - but what are ASICs?
  - most of us “know”, but seemingly sometimes have different expectations
- What can we learn from history?
  - the LHC era involved a step change in electronic developments
  - why – and will it continue in the same way?
  - history should include experience – and not only successes
- A survey of all ASICs is impossible – and probably not very interesting
  - focus on some selected “typical” case studies
    - given available time, necessarily with some bias – you can calibrate…
- Declaration of interest
  - Now, many years of involvement in this area, mainly for the CMS experiment
    - But I am not a chip designer or even, any longer, a hands-on user

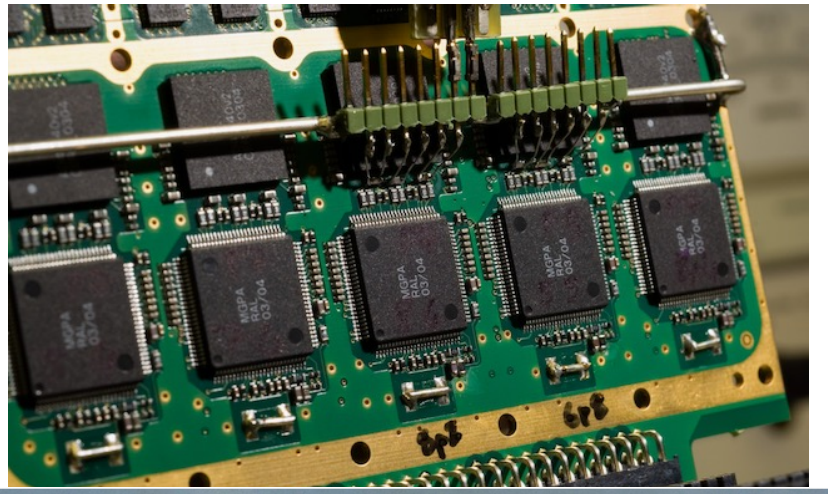
What is an ASIC and why do they matter?

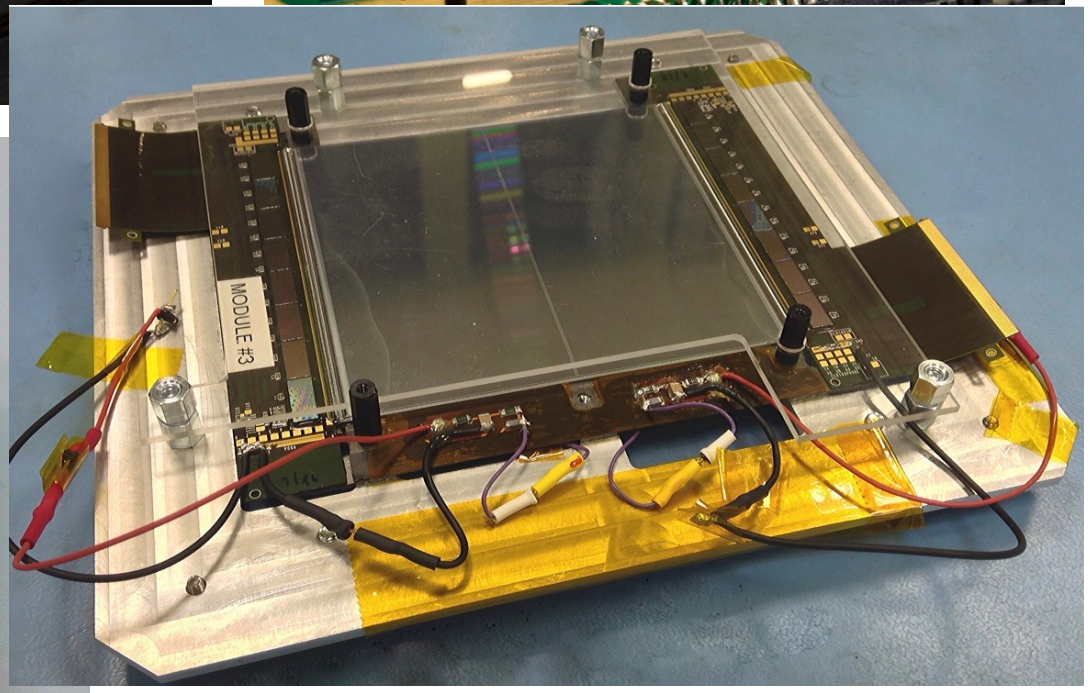


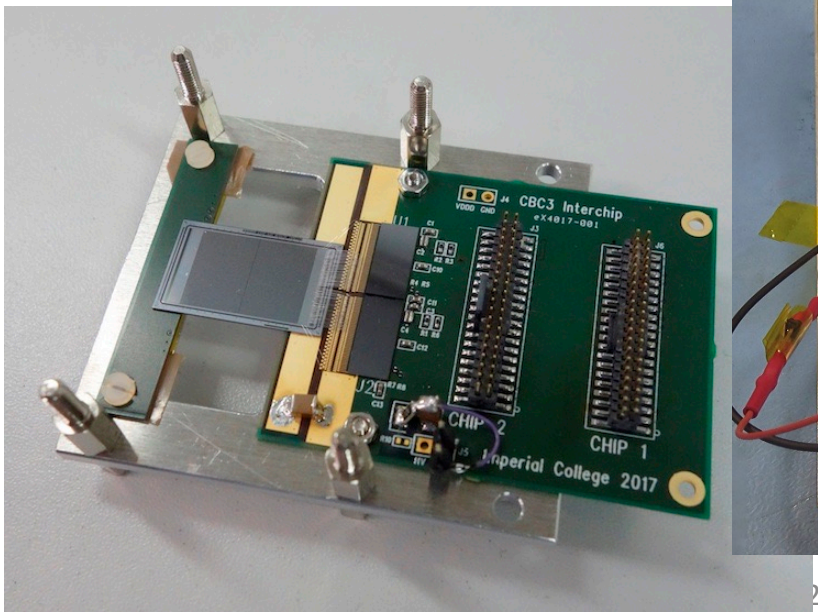
- Application Specific Integrated Circuit
  - customised electronic circuit for a well-defined requirement
    - this need not mean a single detector, or even experiment
  - generally manufactured in **CMOS processes** (Complementary Metal-Oxide-Silicon)
    - dominant electronic technology, with steadily shrinking minimum feature size
- Pros of ASICs
  - **can be optimised** for demanding requirements: size, power, functions, performance,…
  - **miniature**, so ideal for high density HEP – large numbers of channels
  - very **dependable manufacturing quality with low unit cost** on large scale
  - **radiation hardness** now understood, and can be excellent in commercial processes
- Cons of ASICs
  - **Big development investment** required in both time and cost
    - increasing as functionality (= complexity) increases
  - **Unchangeable** once complete, unless a lot of flexibility built-in (adds complexity)
  - **Substantial design and evaluation** requiring specialist skills (industry pays well!)

Spot the ASIC





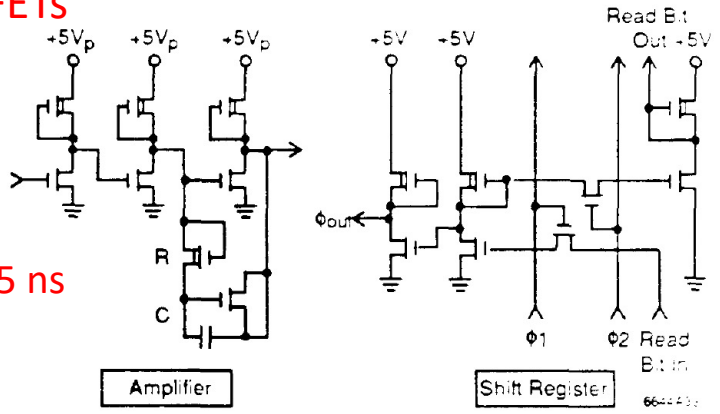




A very brief pre-history

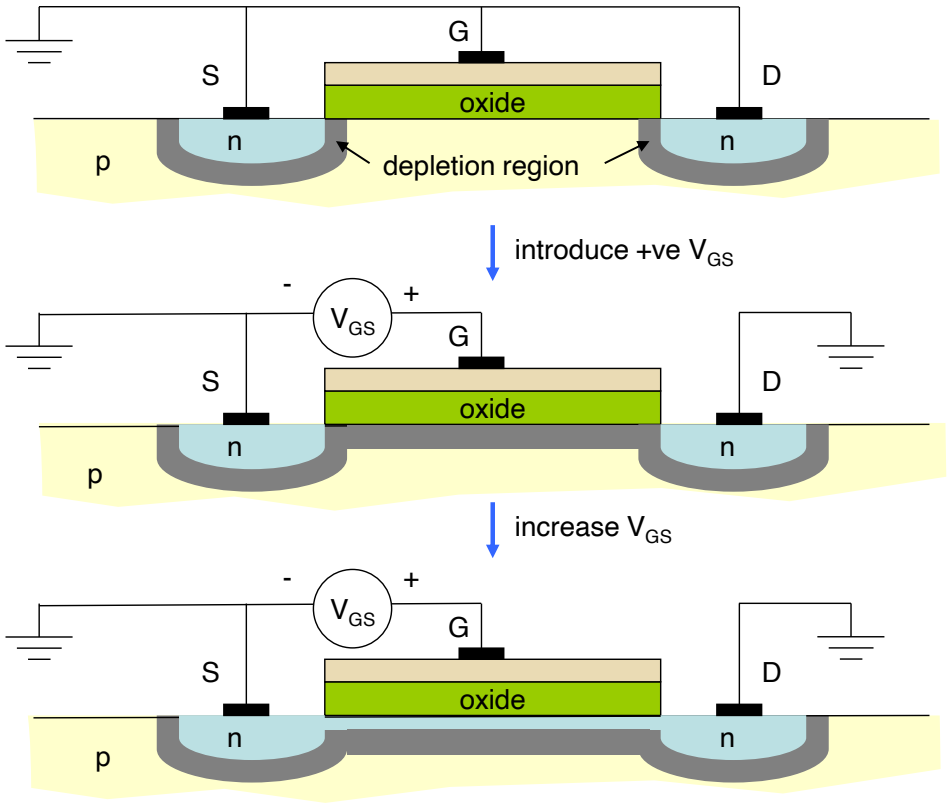
- 1984 - first HEP ASIC: Microplex at SLAC (California!) Mark-II silicon vertex detector
  - NMOS only : 128 channel - amplifier, Sample-Hold, DCS processing, multiplexing.
    - 34 mm^2 , 5µm university lab process, **14 mW/ch**. **Pioneers** learned from first principles!
- Late 1980s: MX3, MX7, CAMEX64 - LEP silicon vertex detectors
  - Commercial CMOS, initially ~3µm and later 1.5µm => 1-2 mW/channel
    - Amplifiers: integrators with switched capacitor filters. Switching noise injected during the amplifier reset subtracted due to its very reproducible behaviour. $t_{rise} \approx 400$ ns, $t_{int} \approx 1.8$ µs
- 1988: SVX ASIC for CDF (& L3) – memory & sparsification
  - amplifier, comparator, multiplexer, nearest neighbour logic, pedestal subtraction
    - 128 channels in 3 µm CMOS $t_{int} \approx 200$ ns, $t_{sample} \approx 0.5$ µs
- 1990: Amplex for UA2 Si pads - feedback resistors using FETs
  - 16 channel 3µm CMOS, precise control of non-linear R
    - more conventional RC filters implemented: $\tau_{peak} = 0.75$ µs
- Early 1990s – LHC developments began
  - originally 66 MHz beam crossing rate, later 40 MHz => $\tau \approx 25$ ns
    - almost x 10^6! - from 120 Hz at Mark-II in a decade
    - but 1-2 µm processes

Part of the Microplex circuit



MOSFETs – in principle and practice

- Transistor is simple device: layout and behaviour, especially digital
  - but also many good analogue properties, if needed



channel just forms at $V_{GS} = V_T$ = threshold voltage

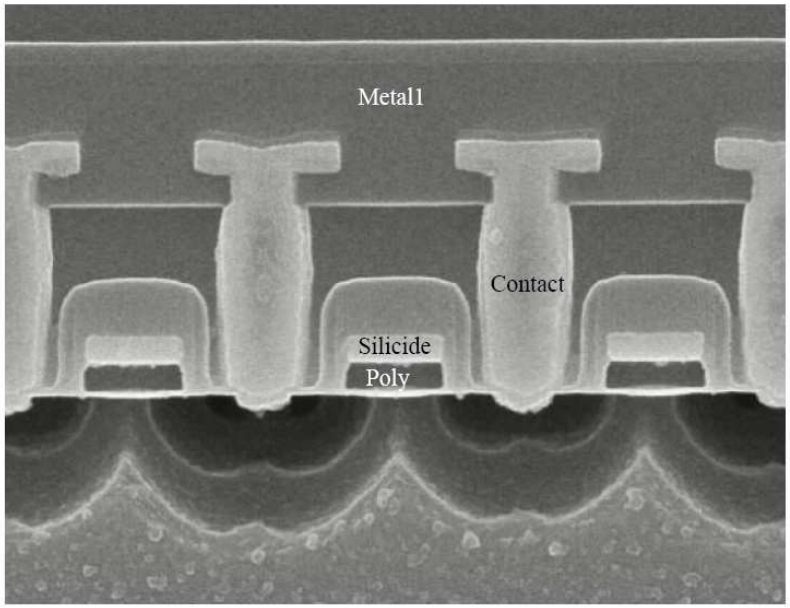


Figure 5.22 SEM image of the cross-section of three MOSFETs.

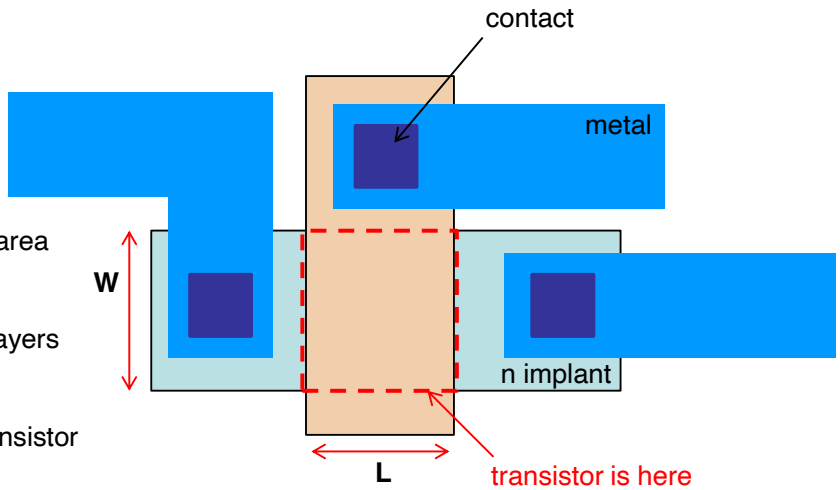
MOSFET design

designer draws the masks

NMOS transistor is formed where gate area crosses n implanted area

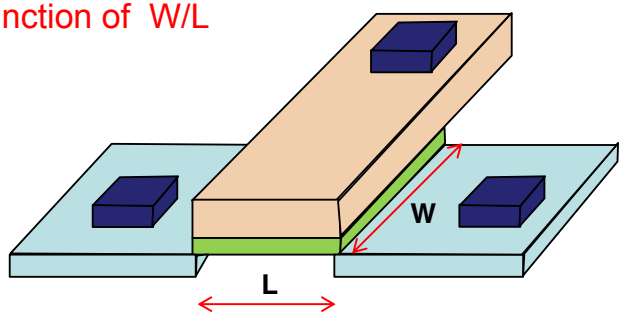
contact regions are defined and metal layers used for connections

only the width **W** and length **L** of the transistor are under the designer's control



transconductance is a rather simple function of W/L

In practice today, sophisticated computerised design software is used to lay out transistors often with libraries of frequently used circuits



plus a lot of complex simulation tools and checking to validate designs

It really was like this around 1985!

Today it would be rare to design an individual transistor.

Circuit properties mainly scale with feature size

L, W, t_{ox}, V =>
L/S, W/S, t_{ox}/S, V/S

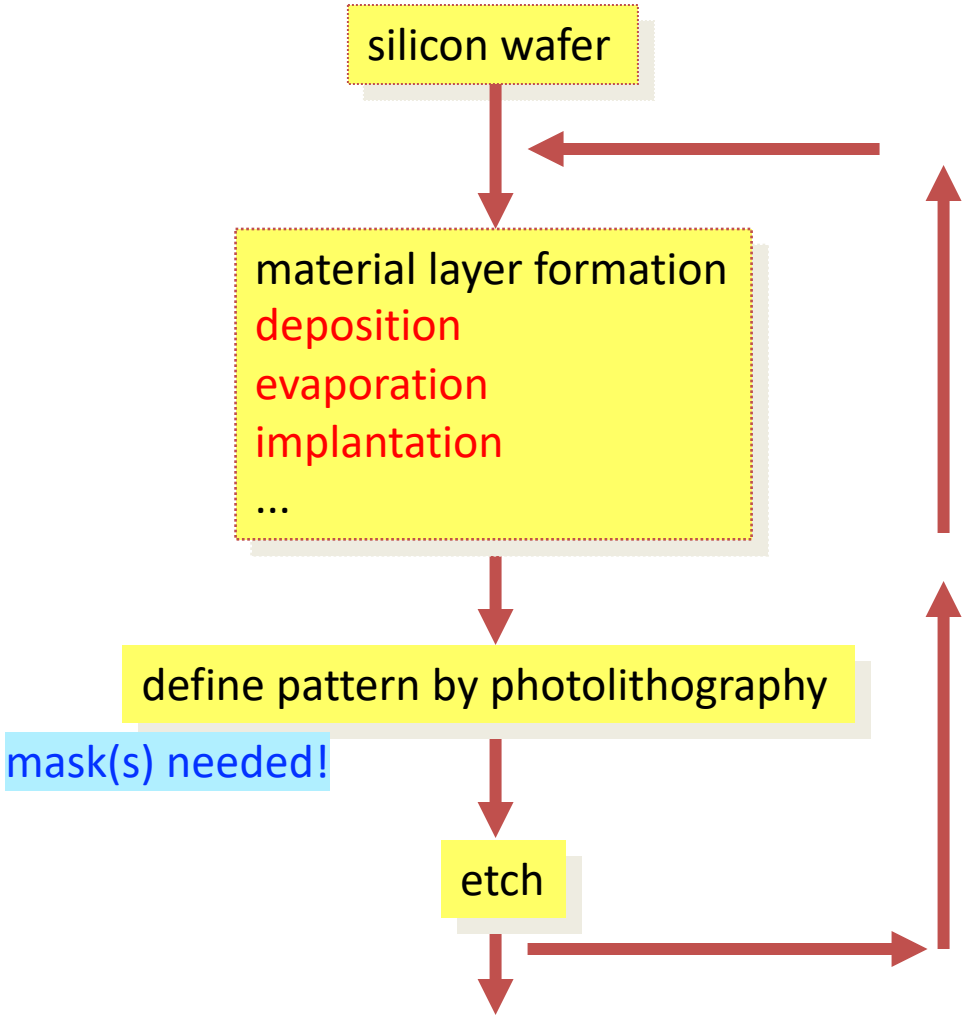
Table 6.3 CMOS scaling relationships.

Parameter	Scaling
Supply voltage (VDD)	S
Channel length (L_{min})	S
Channel width (W_{min})	S
Gate-oxide thickness (t_{ox})	S
Substrate doping (N_A)	S^{-1}
On current (I_{on})	S
Gate capacitance (C_{ox})	S
Gate delay	S
Active power	S^3

Manufacture

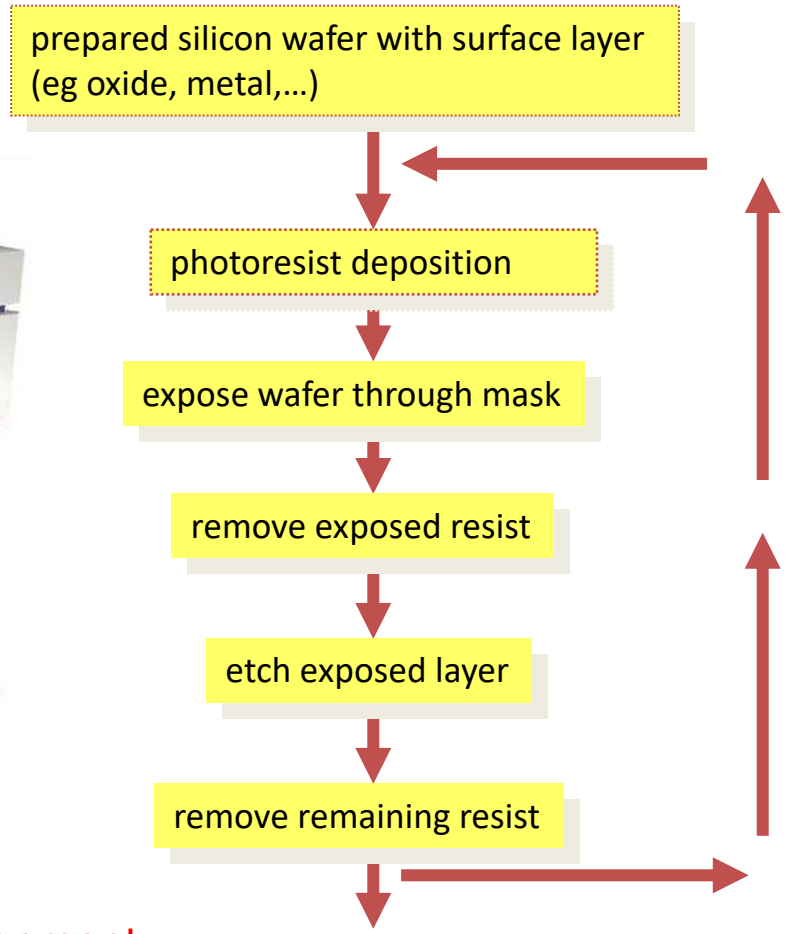
- Repeat basic step many times
  - Select from:

- Wafer preparation
- Oxidation
- Lithography
- Etching
- Diffusion
- Ion implantation
- Metallisation
- Deposition
- Thermal treatment
- …



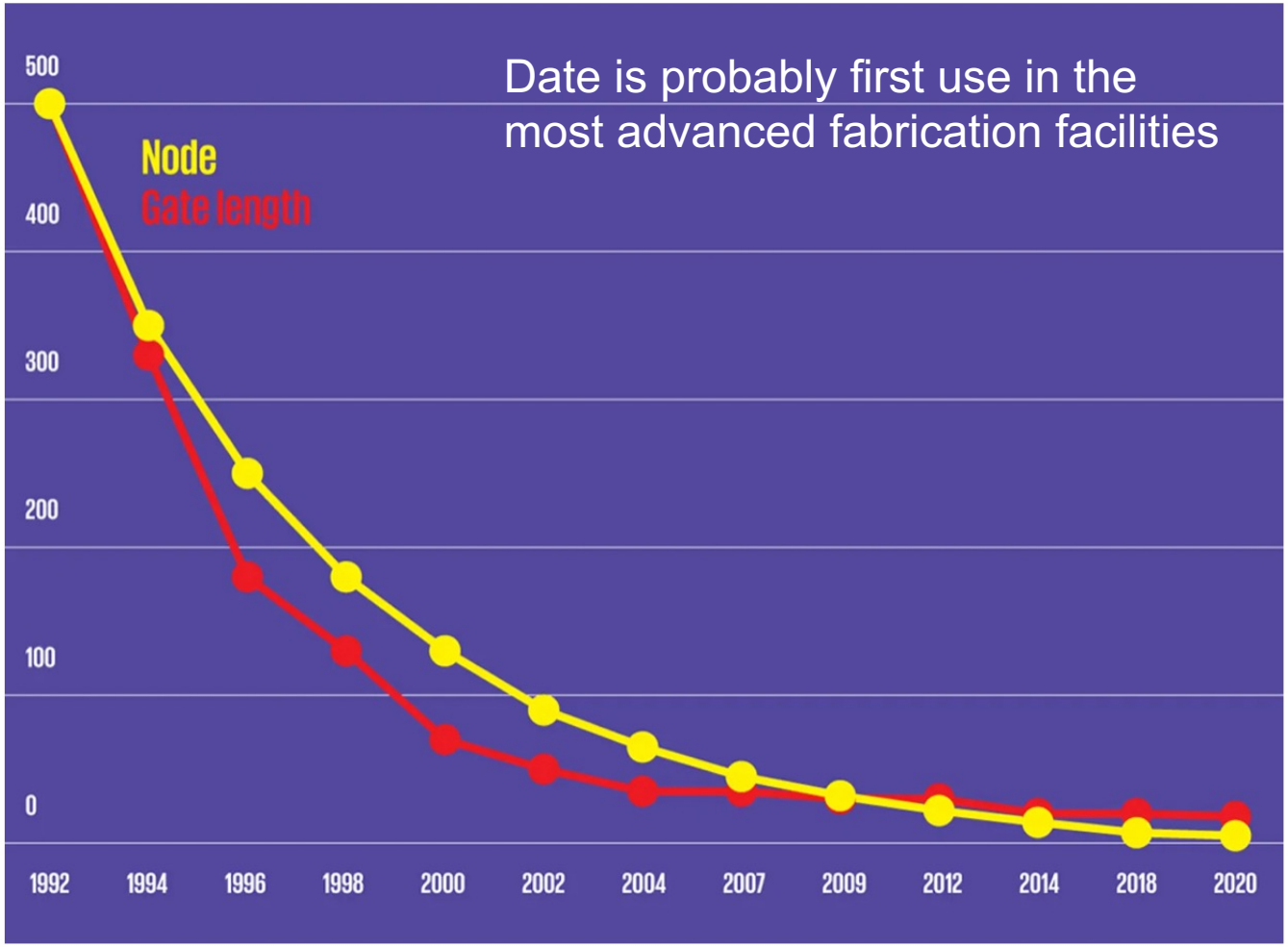
e.g. Lithography

- One of the most crucial steps:
  - nm features require ultra-short wavelengths, and high precision handling



Avoid people & wafer handling in super-clean environment using super-expensive, highly optimised equipment

The outcome

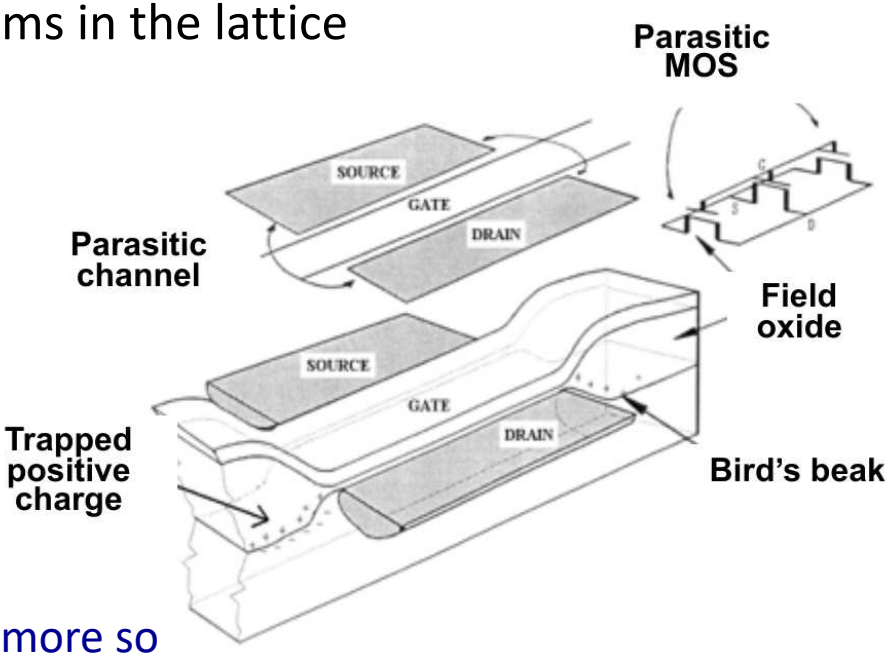


SOURCES: STANFORD NANOELECTRONICS LAB, WIKICHIP, IEEE INTERNATIONAL ROADMAP FOR DEVICES AND SYSTEMS 2020

- In 1998 we were designing the APV25 in 0.25 µm
  - finalised 2000/2001
- We started designing the CBC chip in 2010, in 130 nm
  - completed CBC3.1 in 2018/2019
- Today HEP is designing in 65 nm
  - Some are testing the water with 28 nm
- Note the lag to HEP

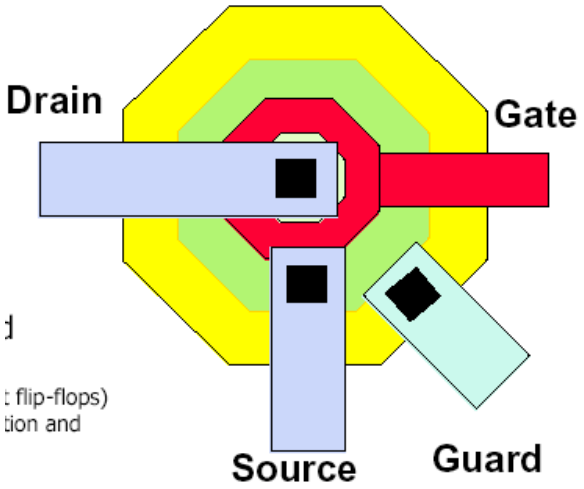
Radiation effects

- CMOS electronics is mostly sensitive to trapped oxide charge
  - originally observed in satellite electronics, attributed to gate oxides
    - generated by charged particles, x-rays, …
  - shifts threshold voltage and on-off transistor operation, increases noise
- Because current flows in extremely thin layer under the gate CMOS is virtually immune to displacement of atoms in the lattice
- Subtle effects not ruled out
  - e.g. dose rate dependence
  - creation of parasitic devices
    => destructive **latch-up**
  - gates are not the only oxide locations
- Single Event Effects
  - up to now less important, but becoming more so

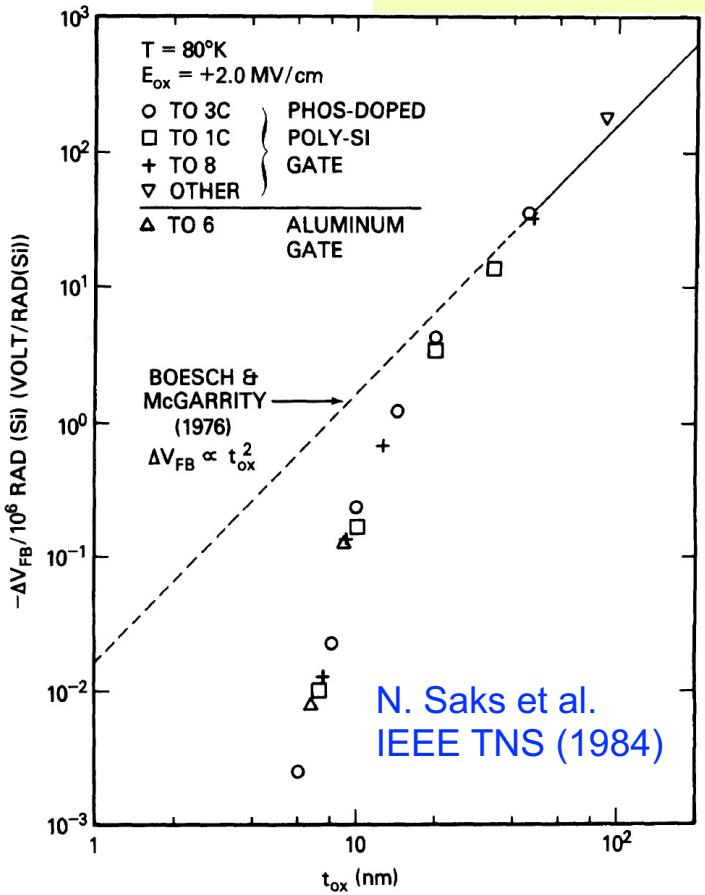


Radiation behaviour of commercial CMOS

- Evidence of radiation tolerance in mid-1980s, improving with thinner oxides
  - tunnelling of carriers reduces trapped oxide charge
  - but confusing results from commercial chip evaluations
- negative effects attributed to leakage paths around NMOS transistors
  - *cured in CERN 1997 with enclosed gate geometry*



ΔV_{FB}/Mrad vs t_{ox}



How did we get to where we are today?

- Sponsored R&D (US SSC detector R&D 1988?, CERN DRDC 1990) played a big role
  - much credit to the wise people who devised those programmes
    - but was there any choice? SSC/LHC detectors were unbuildable
  - **serendipity**: mass commercial electronics era -> internet age
    - 1980s-1990s: RISC processors, PCs, Apple desktops, WWW, modems, mobile phones, laptops,…
  - historians can explain the sequence but all were driving greater miniaturisation
    - many "invisible" components, like connectors, packaging, batteries, not just ICs

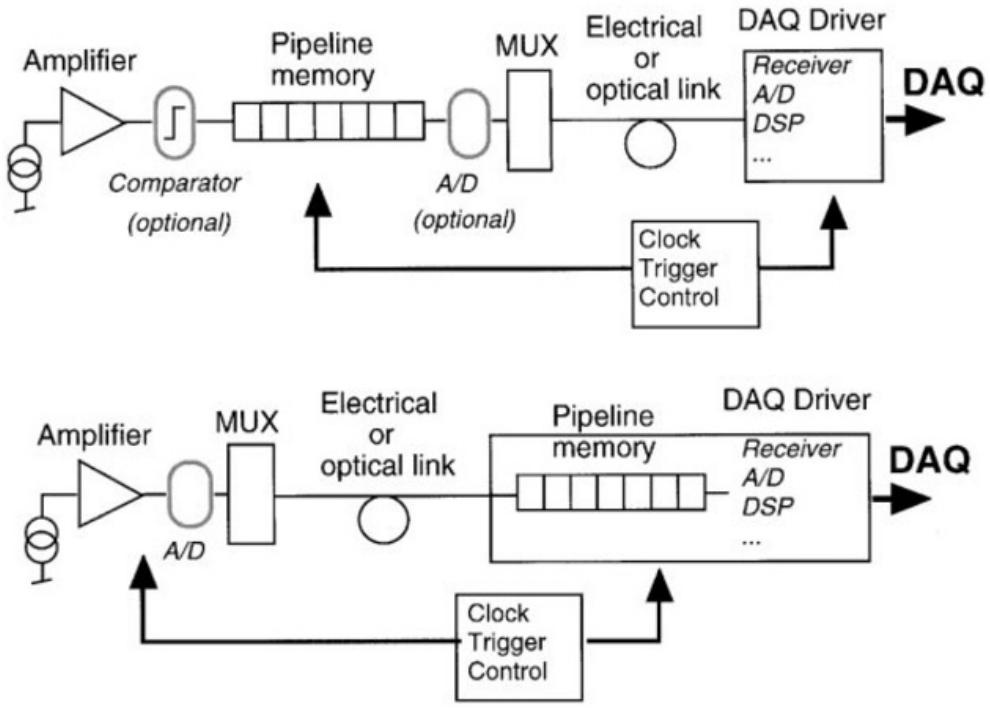
Several crucial areas - **commercially driven** - not by HEP

- Integrated circuit electronics
  - emergence of **accessible ASIC processes**, and design tools
  - significant investment in **training** of new generation of engineers as designers
- Optoelectronics
  - practically non-existent before LHC, and fortuitous **technology co-evolution**
- Off-detector electronics
  - **FPGAs** evolved dramatically, coincidentally, from mid-1990s

LHC ASIC requirements

- Developments initially driven mainly by trackers and ECALs
  - Amplifiers, data conversion and storage, data transfer, ….
  - Followed by clock and control distribution, and full system development

- Many possible design choices
  - Preferences and prejudices
    - e.g. analogue/digital/binary
    - latency requirements
    - performance
  - on- or off-detector
  - fraction of commercial parts
  - radiation tolerance
  - power
  - cost issues
  - innovation vs risk



a couple of possible variants

Evolution of CMS silicon µstrip tracker ASICs

- Began in 1991 in RD20 DRDC project
  - amplifier designed, with several variants investigated
  - pipeline concept already proven, implemented in steps
  - several prototype chips in supposed rad-hard 1.2 µm technology, 32- then 128-channel
    - plus a couple of non-rad hard ASICs, for detector prototyping (NB important!)
    - MSGCs were part of the detector concept – fortunately Si electronics was initially applicable
    - By 1997: APV6 might have been a candidate for final system
    - many irradiation studies of single transistors (noise) and chip performance
  - Along the way, some important innovations
    - programming chip parameters via serial command interface (I2C) using software control
      - **hugely beneficial in accelerating** chip configuration and evaluation
      - primitive (by today's standard) signal processing on chip (e.g. deconvolution) to save power
- Bad luck – which turned out to be good luck – in September 1997
  - Foundry move of 1.2 µm process **reduced** radiation tolerance (marginal anyway)
  - Unexpected **positive** CERN results on "standard commercial" 0.25 µm radiation tolerance
    - In 1998 switched process – which was a turning point – for **APV25** (Sept 1999)

Short summary of history

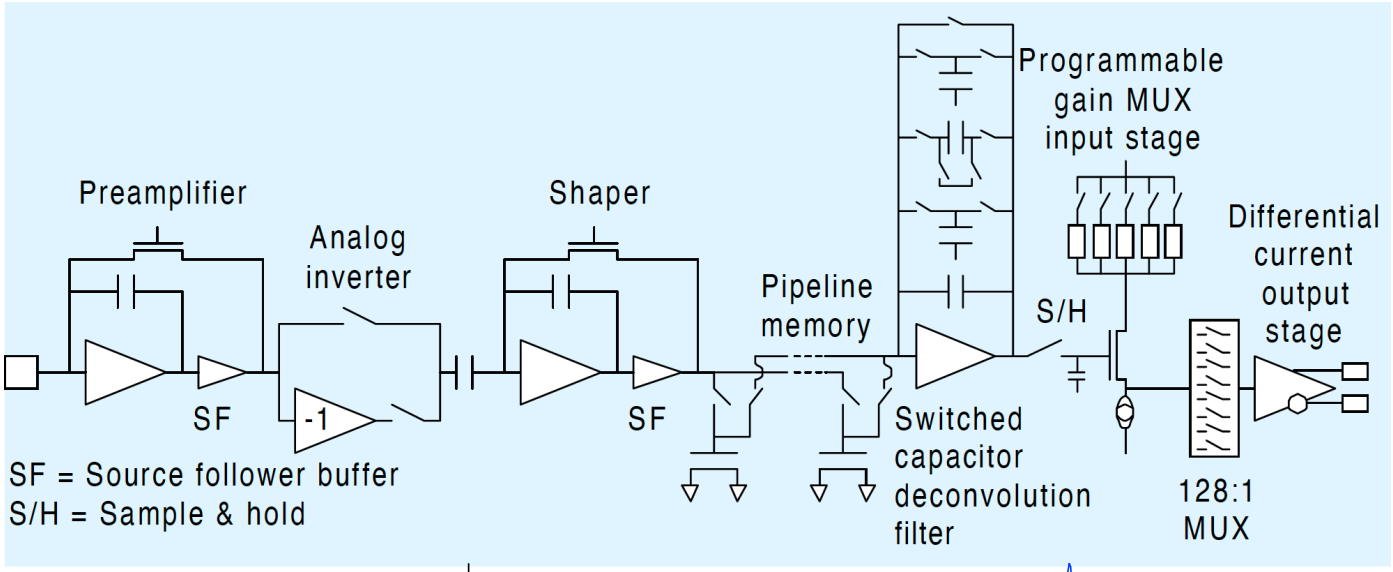
- Process switch to IBM 0.25 µm was incredibly successful
  - engineers from RAL, Imperial & CERN transferred design quickly
    - well characterised process: designs matched simulations
  - big gains from scaling 1.2 µm to 0.25 µm
  - radiation hardness was well in excess of requirements

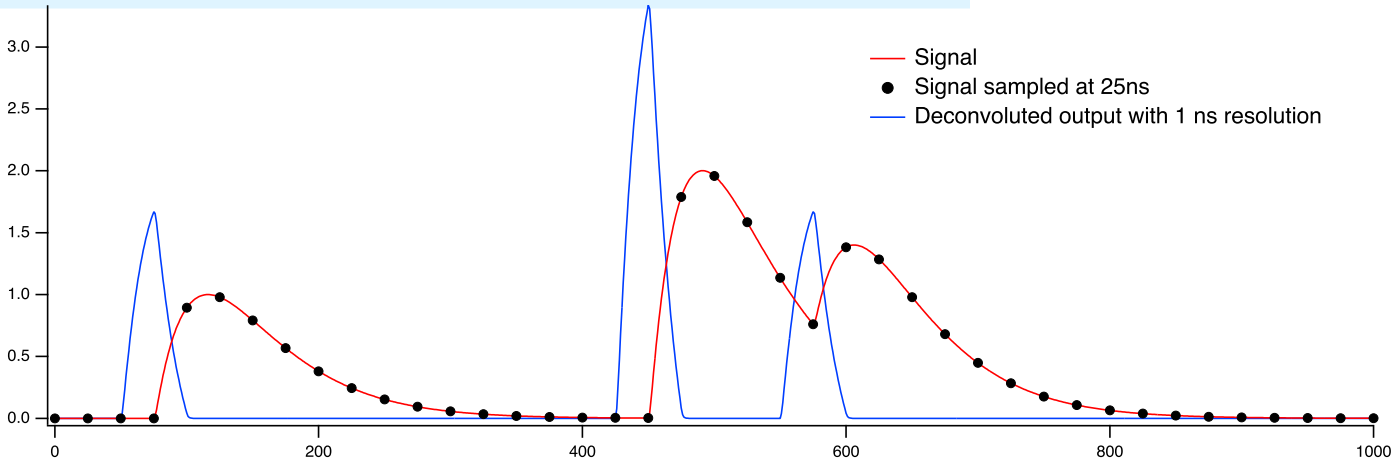
- It was obvious that **all** tracker ASICs should use the technology (others took note)

- It wasn't quite the end of the story
  - unexpected yield variations were experienced at the early production stage
  - weak points in the (not quite standard) manufacturing process were identified
    - by IBM specialists - and fixed – we would not have been able to
    - production quality was subsequently excellent – very high yield of good chips

APV25 architecture

- Only signal processing is the analogue deconvolution filter
  - forms weighted sum of three consecutive samples to “reshape” sampled pulse



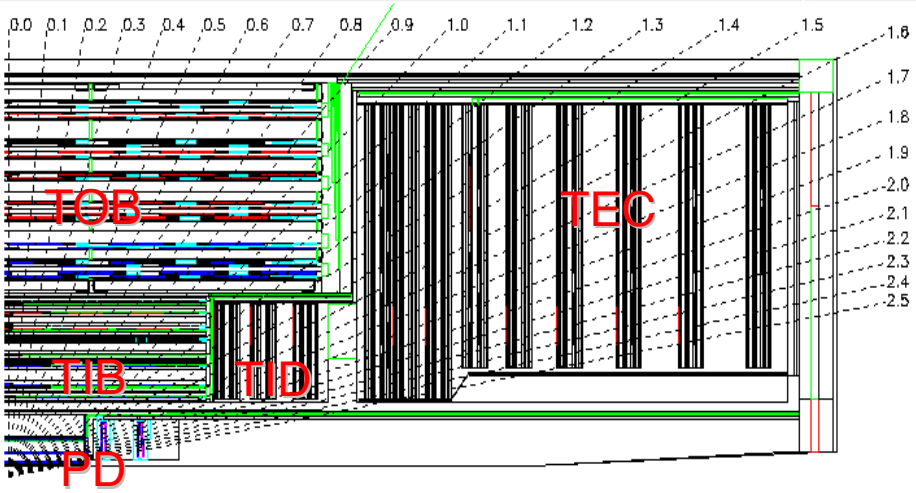


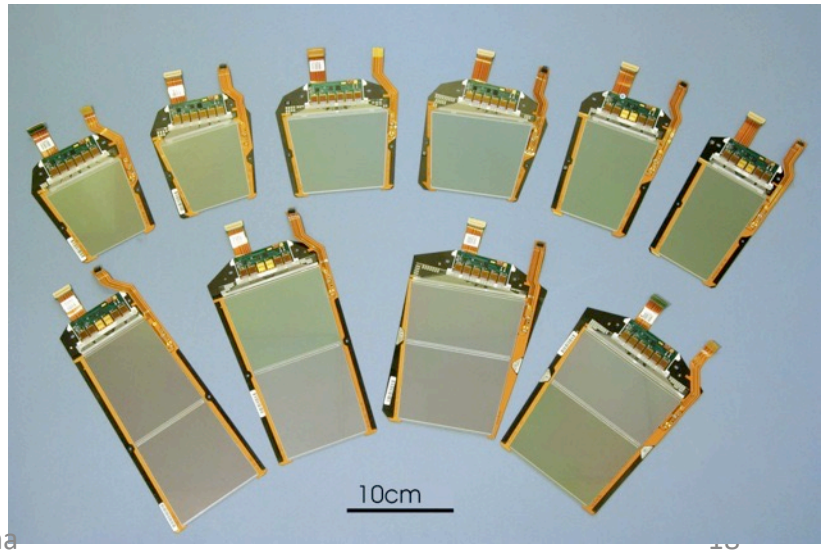
CMS Tracker and its sub-systems

Pixel upgrade 2017 to 4-layer barrel and 3-layer endcaps

- Two main sub-systems: Silicon Strip Tracker and Pixels
  - pixels quickly removable for beam-pipe bake-out or replacement
  - SST not replaceable in reasonable time

Microstrip tracker	Pixels (pre-2017)
~210 m^2 of silicon, 9.3M channels	~1 m^2 of silicon, 66M channels
73k APV25s, 38k optical links, 440 FEDs	16k ROCs, 2k olinks, 40 FEDs
27 module types	8 module types
~34kW	~3.6kW (post-rad)

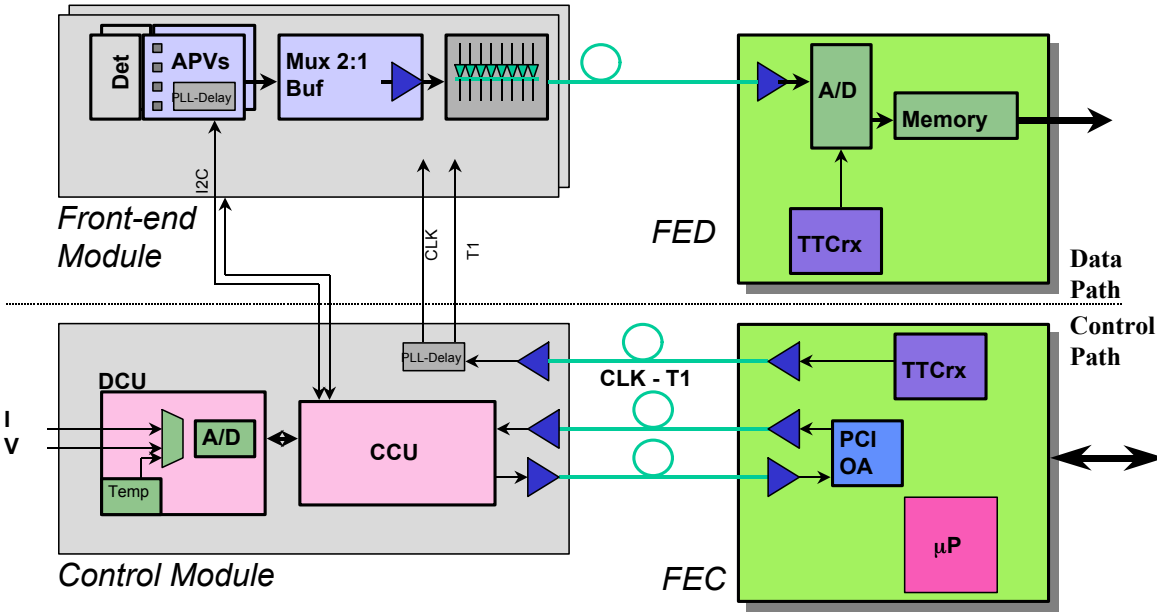




Tracker electronic system

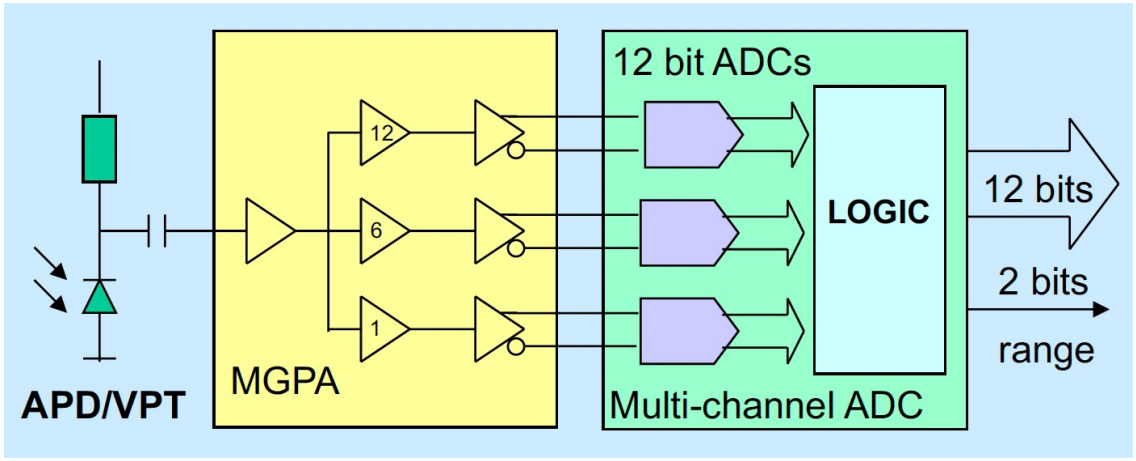
- For a working system, several other important tracker ASICs were needed
  - CCU – distribute clock and trigger, and slow control signals and data
  - DCU – monitor internal detector parameters: currents, temperature, voltage
  - PLL – recover clock and trigger
  - DLL – fine tuning of delays in system
  - APVMUX – multiplex APV25 2:1
  - Laser driver and optical receiver

We don't talk much about these ASICs but they're **vital** and should not be completely forgotten

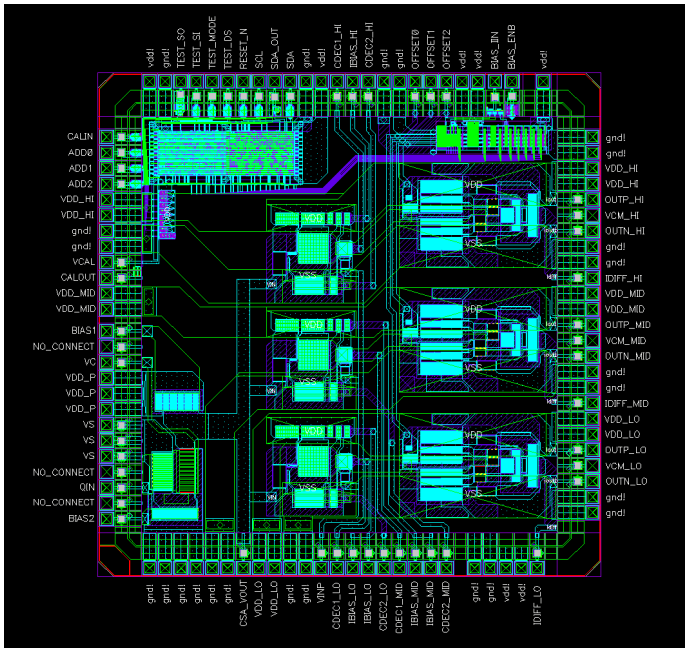


An interlude

- In 2002, tracker effort and developments deployed to revise ECAL system
  - data and control optical links, control hardware
  - 0.25 µm ASIC expertise to design MGPA FE ASIC, new ADC, digital logic ASIC
    - entirely new implementation in different technology than previous attempts



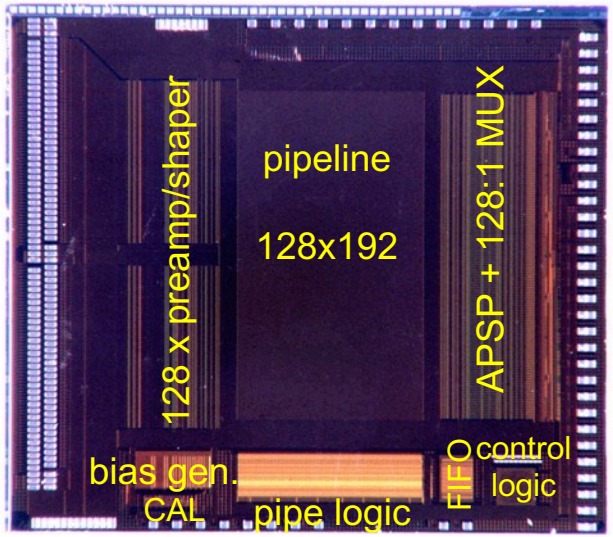
MGPA



- Despite demanding requirements, successfully carried out in less than two years
  - of course, profiting from prior extensive conceptual system design, so work focused on implementation
  - collaboration – including with industry

CMS Tracker FE ASIC evolution into upgrade era

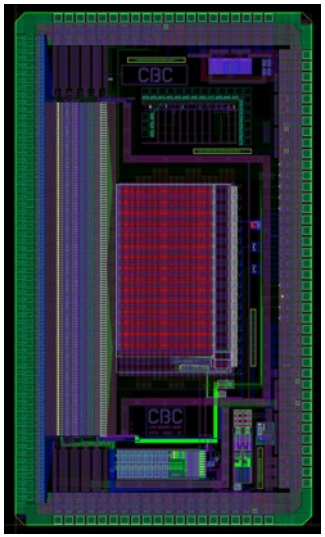
- 1999: APV25 0.25µm
  - 7 mm x 8mm (128 chan)



programmable settings
analogue data
~4 µs latency
wire-bondable
pulse-shaping choice

2011: CBC 0.13µm

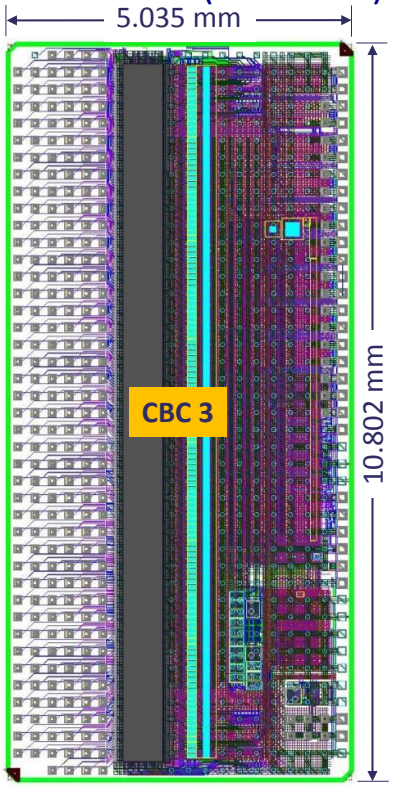
7mm x 4mm (128 chan)



binary data,
6.4 µs latency
wire-bondable

2019: CBC3.1 0.13µm

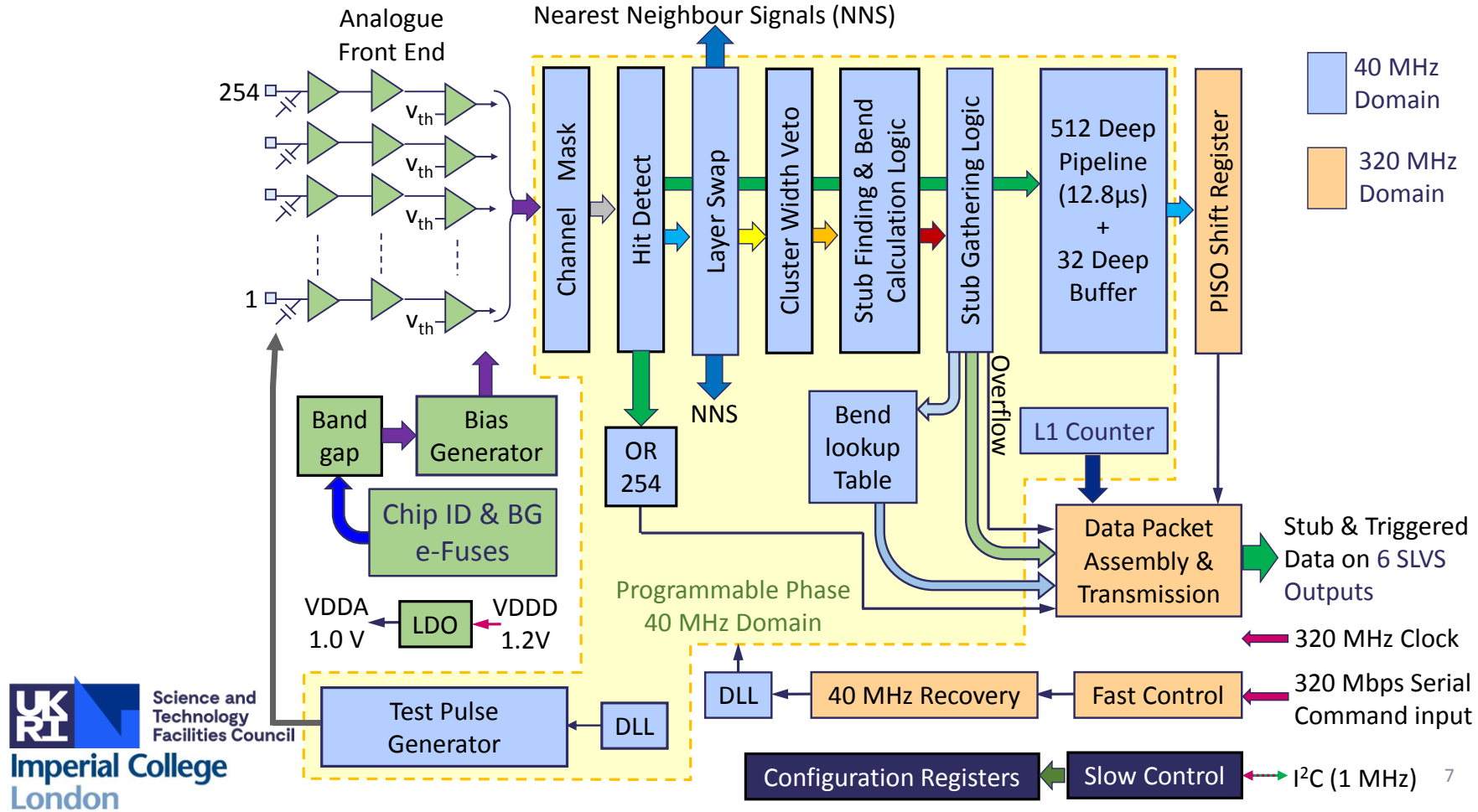
11mm x 5mm (254 chan)



bump-bondable
12.8 µs latency
cluster & correlation logic
many other features!

CBC architecture

- Much more complex than APV25
  - but not the most complex tracker ASIC!



Science and Technology Facilities Council
Imperial College London

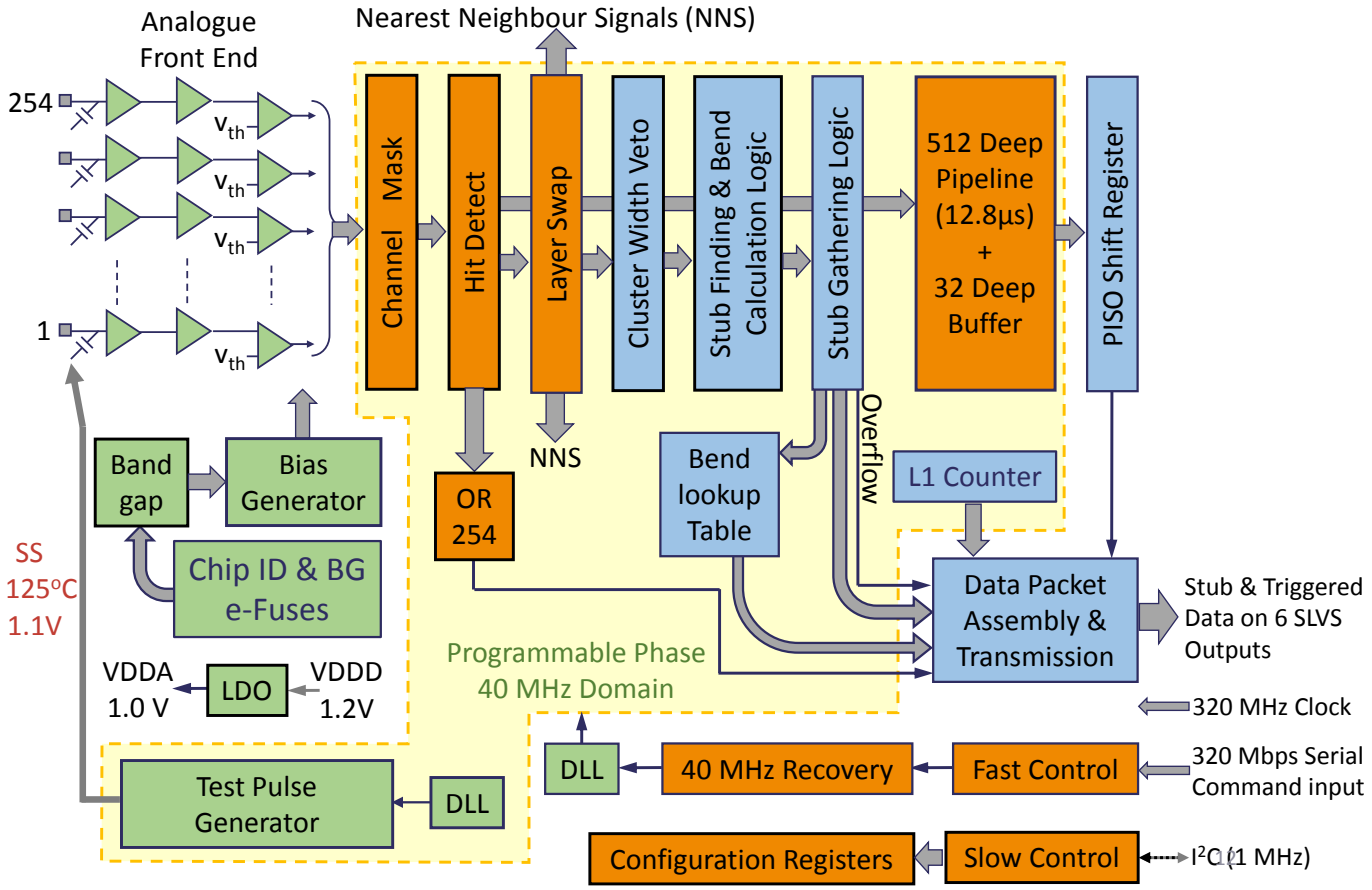
Some aspects of design work

- A lot of functionality to verify – not all displayed here
  - analogue performance, logic, interfaces, timing, … many simulations
    - several design methods – custom/synthesised/…

Analogue Circuits
Analogue simulations @
-- Process corners
-- Temp -30°C, 0°C, 30°C
-- Voltage 1.2V ±10%
*Down to 1V for Front End

Custom Digital Circuits
Analogue simulations @
-- Process corners
-- Temp -30°C, 0°C, 30°C
-- Voltage 1.2V ±10%

Synthesised Digital Circuits
Digital simulations @
Process corners FF, TT, SS
Temperature -55°C, 25°C, 125°C
Voltage 1.3V, 1.2V, 1.1V



Science and Technology Facilities Council
Imperial College London

Another example

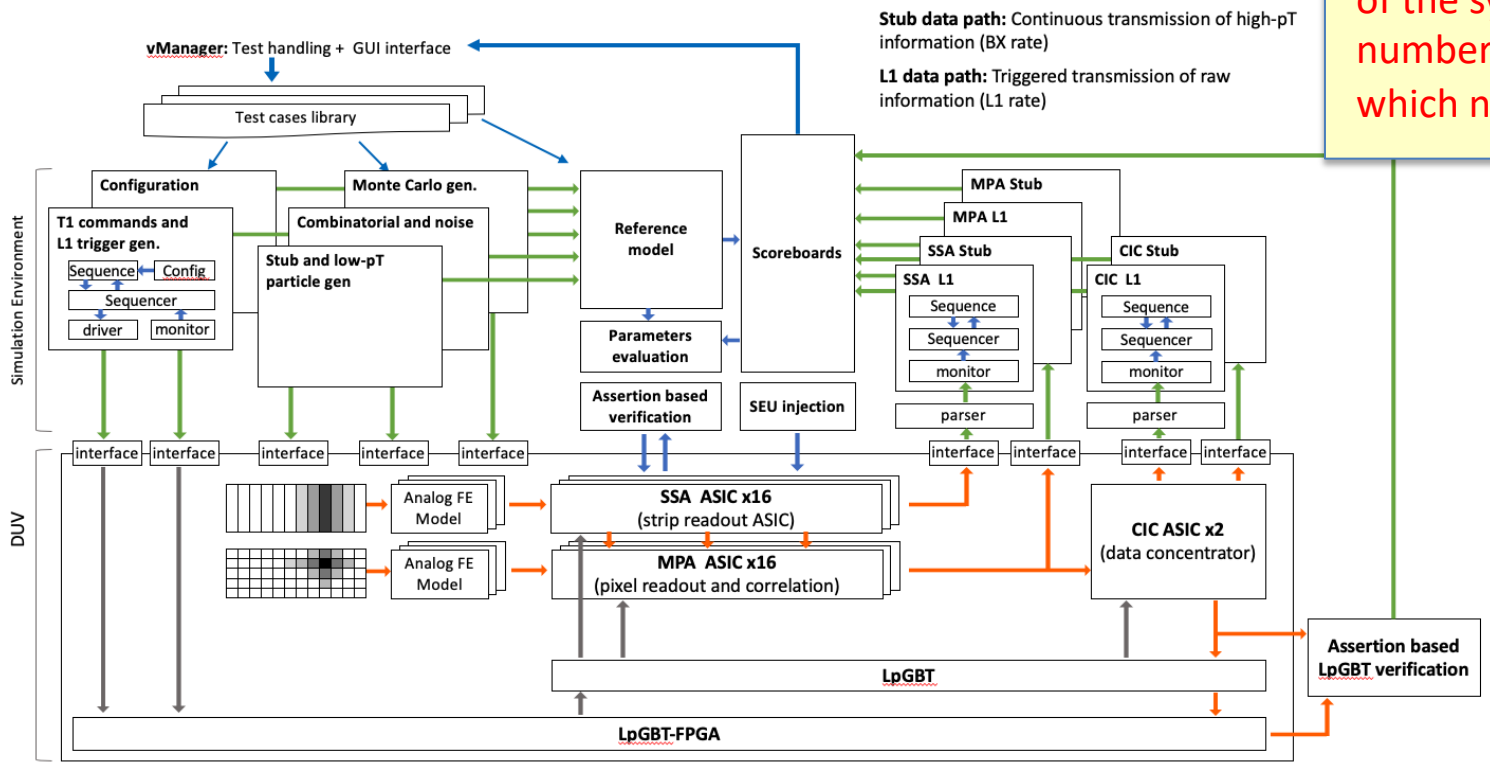
Full Readout chain verification

Simone Scarfi

Incorporating lpGBT and off-detector lpGBT FPGA

To verify stub data, L1 data paths and fast command synchronization

PS-module case



Notice the complexity of the system and the number of elements which need verification

ASICs are everywhere…

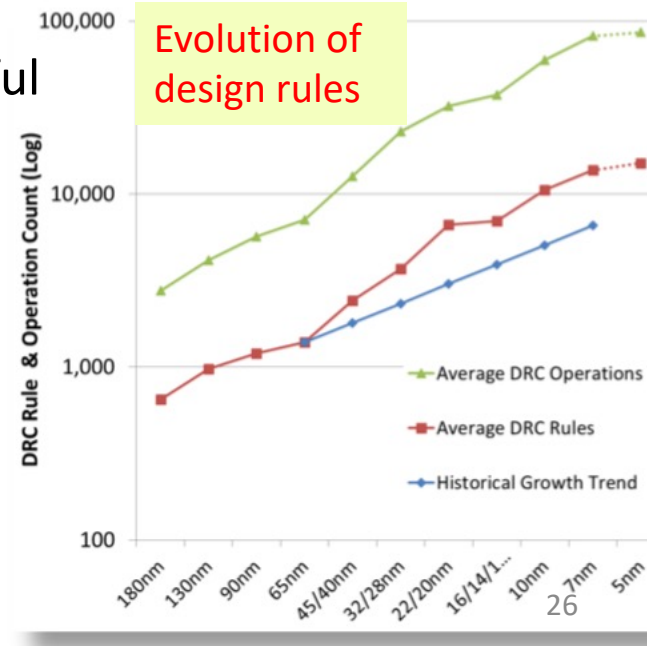
- It's probably easier to identify the few projects which don't use ASICs
  - there may be some, in non-accelerator applications?
  - although not all are developed from scratch –
    - reuse could be a useful trend… how?
- But the new ASICs are **much more complex** than the first LHC generation
  - many more built-in features
    - some functions that were separate chips, such as clock recovery
  - and new functionality, such as **time measurement**
  - external connections for control and data transfer, or assembly
    - how long before **optical elements** are included? or **wireless connectivity**?
- Inevitably, development and qualification times have increased
  - more engineers and applied physicists are needed
    - most with specialist skills
  - many things to go wrong, and higher costs at stake
- What can be done to control this?

What can go wrong?

- Plenty – and we have many **examples** (for later discussion?)
  - many issues were overcome, so can be avoided in future
    - but it's rare that the same problems (**not usually mistakes**) occur the same way
  - it can be difficult to find enough effort, or anticipate where it is needed most
    - not always where expected – e.g. software readiness
    - it's traditional (and not wrong) to ask about radiation tolerance, but not so many other things…
  - many of the problems are "small" – but with bigger repercussions
    - but often hard to find, and need time to resolve
  - manufacturing is not usually at fault – but has natural quality variations
- Nevertheless, today designs are remarkably successful
  - Quality of design tools
  - Quality of manufacture
  - Experienced designers
  - Prototyping via MPWs
  - Careful evaluation

At a price, e.g.:

NB log scale



slide from Marcus French RAL 2022

ASIC Design skills and approaches changing

The Issue:

- Its not just us – the whole industry is facing the same challenge
- Must work with validated IP to manage risk
- Strict QA is demanded at all stages

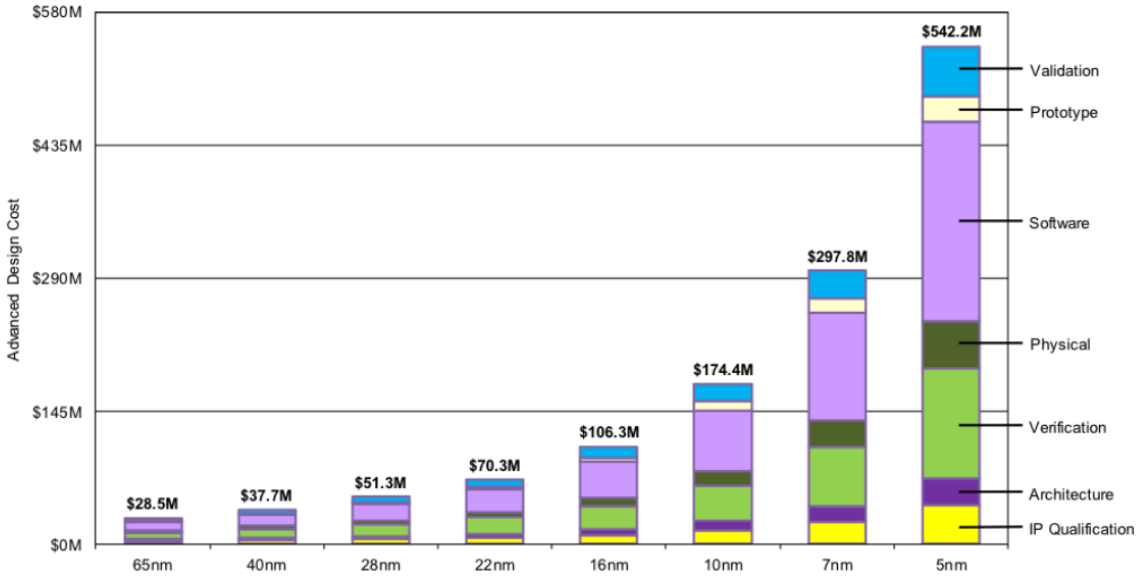
Consequences:

- Increased stress on design teams
- The software is becoming more specialised requiring committed teams dedicated over longer periods
- More effort required for the un-sexy verification and other tasks
- Links with National Labs and CERN are now mandatory

We simply can't work in the 'old' ways…

Even if not using such advanced processes, the same challenges are present – note breakdown by task

Design cost vs technology node



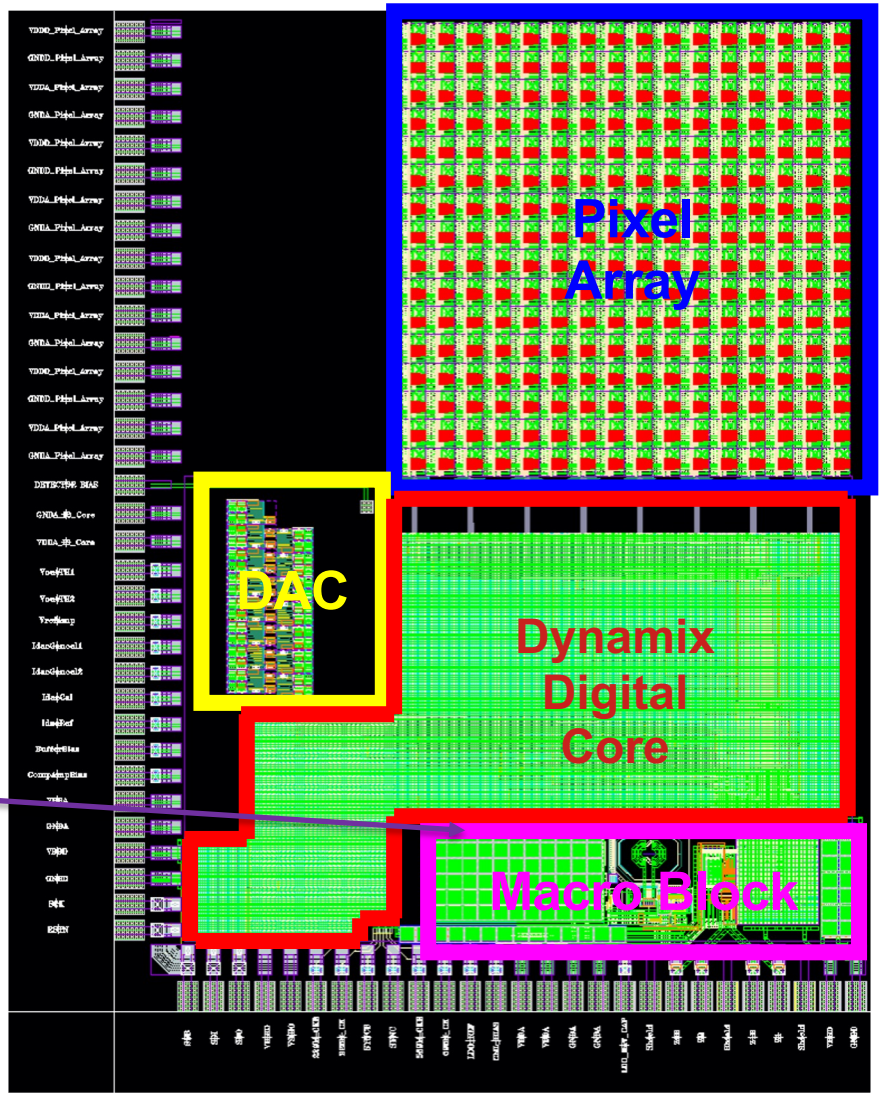
From: Semiconductor Engineering

Thanks to Mark Prydderch, RAL

Full circle…

- A starting point in late 1980s:
  - pixel detectors for x-rays with RAL TD
    - a lot of effort, but not much to show for it
    - with hindsight, a lot was learned

- 2022: floor plan of a 65 nm prototype
  - pixel array for future synchrotrons
  - high frame rate, high dynamic range
  - ADC at pre-amp stage
  - 100 µm pixels
  - segmented scalable architecture
- The **Macro block** contains
  - 14 Gbps serialiser for data output
  - 7 GHz PLL
  - 500 MHz LVDS receiver
  - DAC consisting of 12 10-bit IDACs



Some conclusions

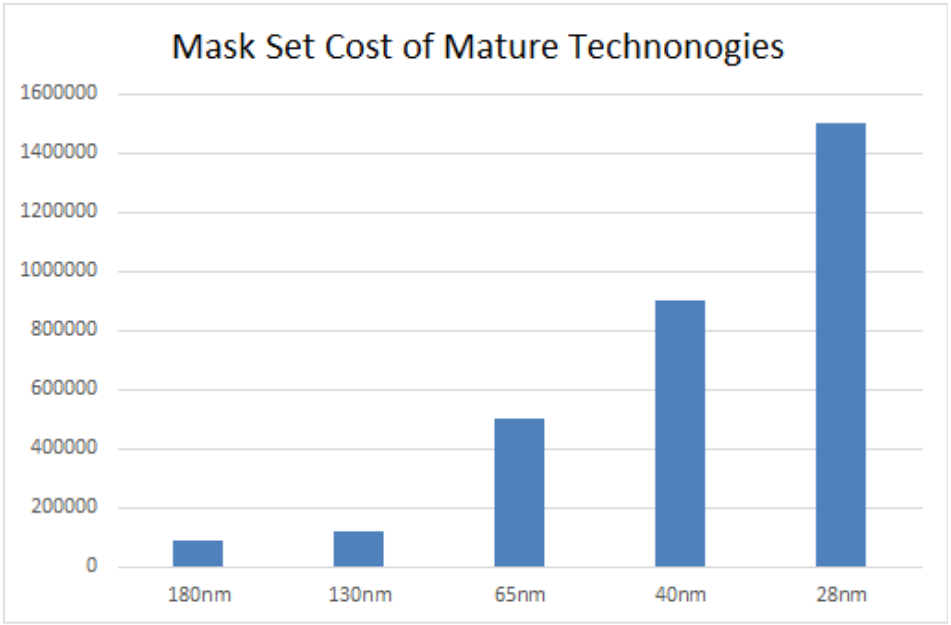
- The LHC fortuitously coincided with the internet era
  - explosion of relevant technology development – high volume, low unit cost
    - much of it - but not all - accessible to small users (such as HEP)
  - enabled implementation of increasingly complex functionality ASICs
    - now being exploited for HL-LHC era
- BUT greater complexity = longer development times and more resources
  - higher risk and less flexibility for small projects
- How to manage this for the next generation?
  - What are the needs after HL-LHC, and for which projects?
  - MPW access and the need to train engineers should be beneficial
- Positive, but not to be taken for granted
  - commercial motivations don't necessarily match scientific interests

- Further reading for those interested
  - G. Hall and A. A. Grillo *ASICs for LHC intermediate tracking detectors* NIM A1050 (2023), 168115 doi:10.1016/j.nima.2023.168115 (part of a special issue on ASICs in HEP)
  - Chris Miller *Chip Wars* Simon and Schuster (2022)
  - Michael Riordan & Lillian Hoddeson *Crystal Fire* W. W. Norton (1998)

Backup

NRE: Non-recurrent engineering

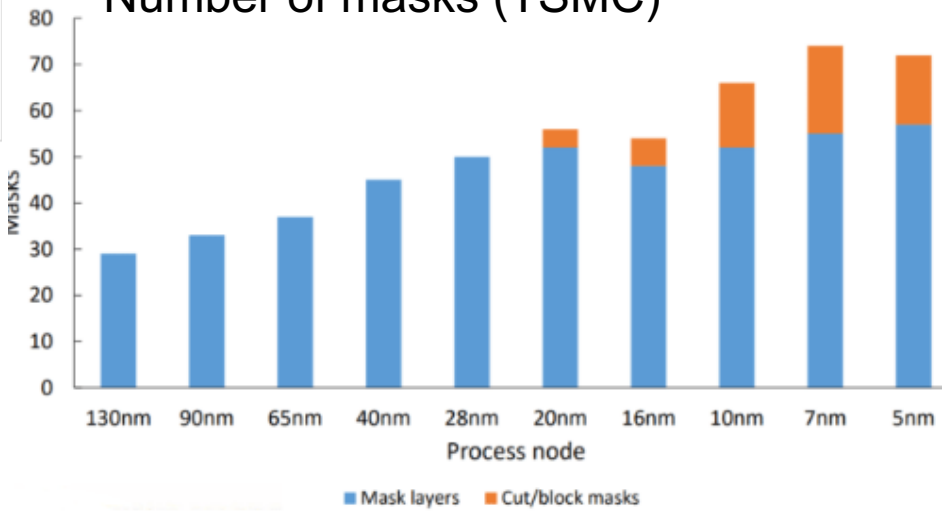
- Dominated by mask costs



source: friends of AnySilicon

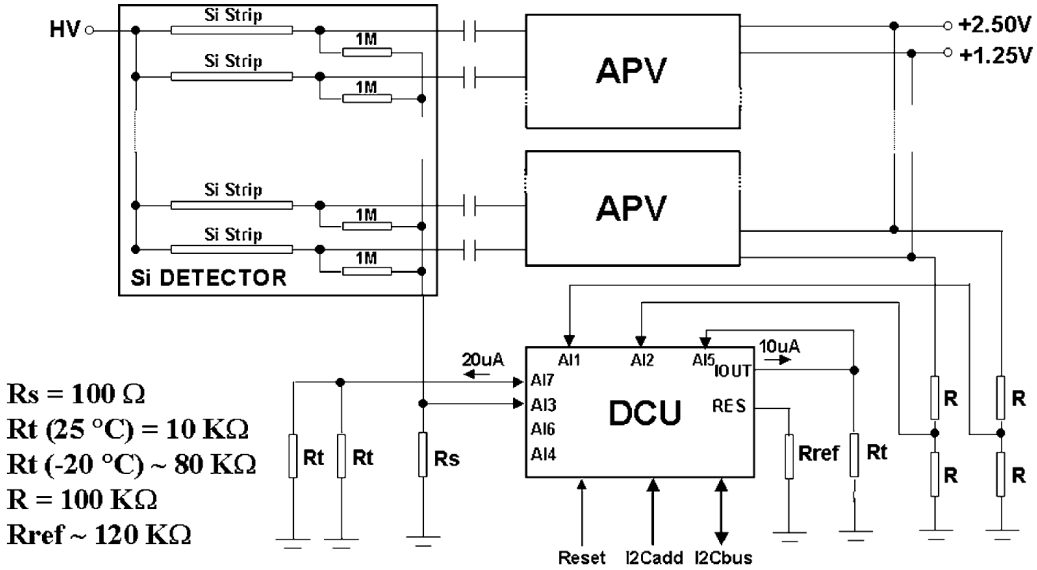
Design rules and manuals scale in a similar way

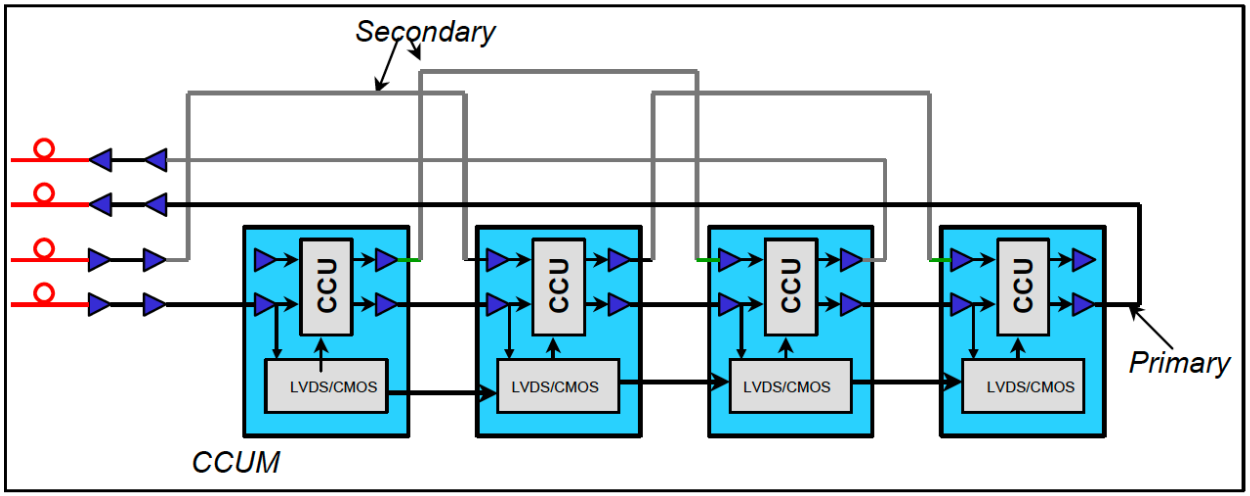
Number of masks (TSMC)



APV development history

1992	amplifier and transistor test structures in Harris
1993	**APV3** – 32 chans preamp/shaper/pipeline/APSP
1994/5	**APV5** – 128 channels + mux
1996	**APV6** – APV5 + bias generator + I2C interface
1997/8	**APVD** – DMILL version
1998	**APVM** – APV6 version for MSGCs
1999	**APV25s0** – 1st 0.25 μm version
2000	**APV25s1** – final version – the one we use today
2001	volume production launched
2002	low yield problems – long story eventually understood and process tweaked
2003	volume production resumed ~ 85% average yield
2005	production finished

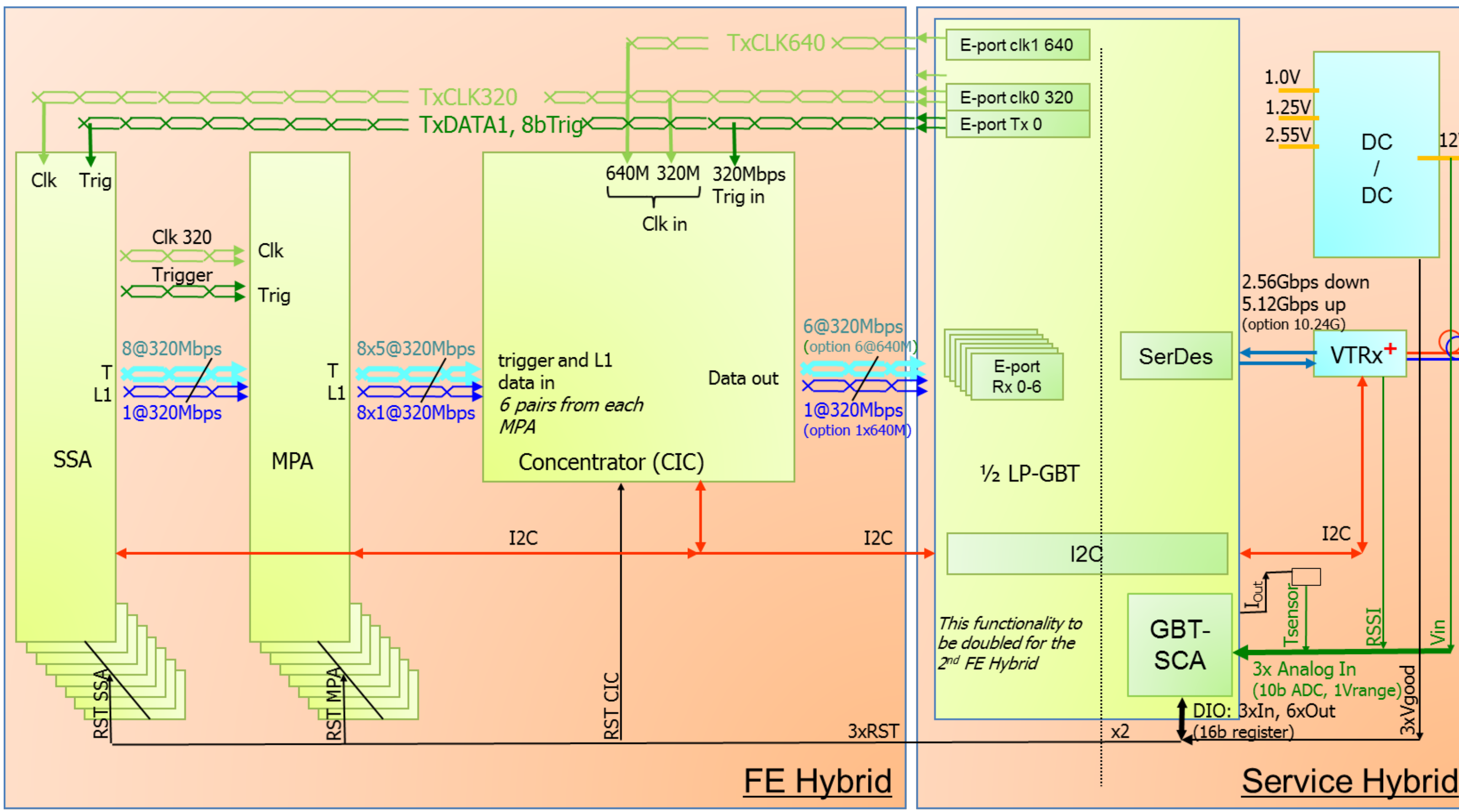

HV
Si Strip
Si Strip
Si Strip
Si Strip
1M
1M
1M
1M
Si DETECTOR
APV
APV
+2.50V
+1.25V
20uA
10uA
AI1
AI2
AI5
IOUT
AI7
AI3
AI6
AI4
DCU
RES
Rs = 100 Ω
Rt (25 °C) = 10 KΩ
Rt (-20 °C) ~ 80 KΩ
R = 100 KΩ
Rref ~ 120 KΩ
Rt
Rt
Rs
Rref
Rt
R
R
R
R
Reset
I2Cadd
I2Cbus

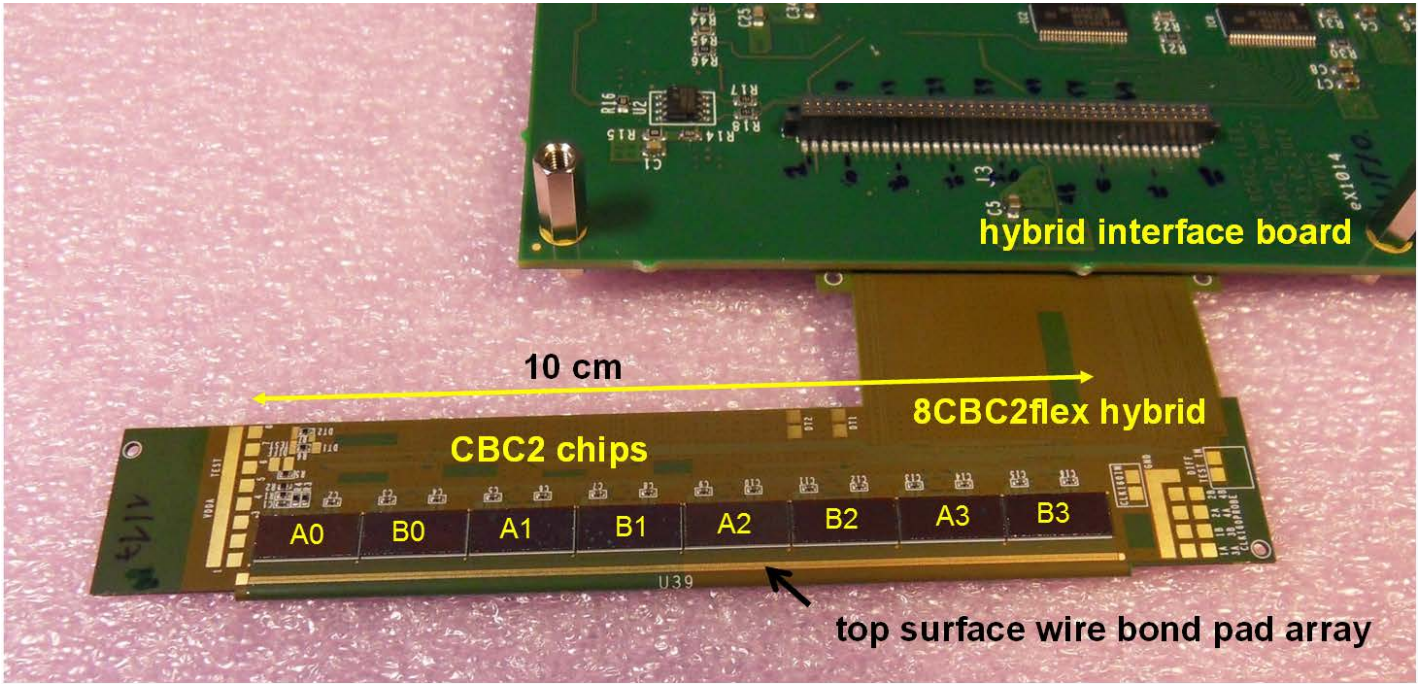

Secondary
CCU
CCU
CCU
CCU
LVDS/CMOS
LVDS/CMOS
LVDS/CMOS
LVDS/CMOS
Primary
CCUM

2. MPA-SSA-CIC

- 65nm CMOS (TSMC)
- ASICs are in design stage
  - MPA & SSA: final verification stage
    - MPA-Light precursor available since 2015
  - CIC: Digital system level modelling stage

Due for submission June 2017




hybrid interface board
10 cm
8CBC2flex hybrid
CBC2 chips
A0
B0
A1
B1
A2
B2
A3
B3
top surface wire bond pad array