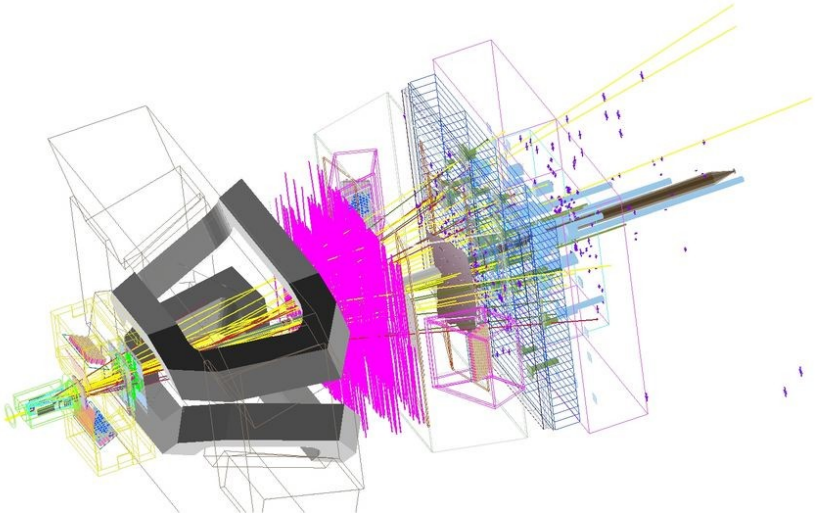
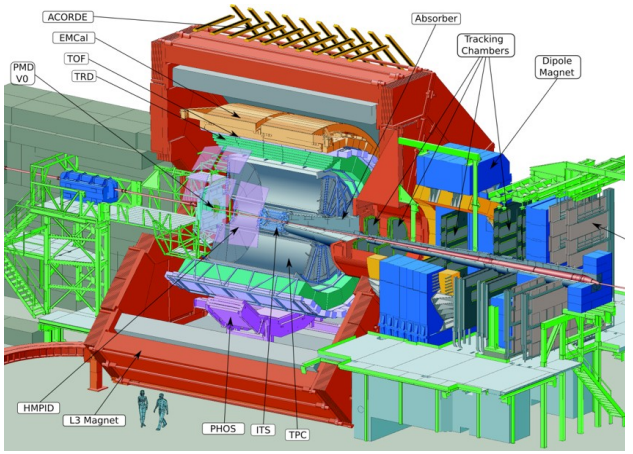
The PCIe40 card
and the importance of efficient production tests

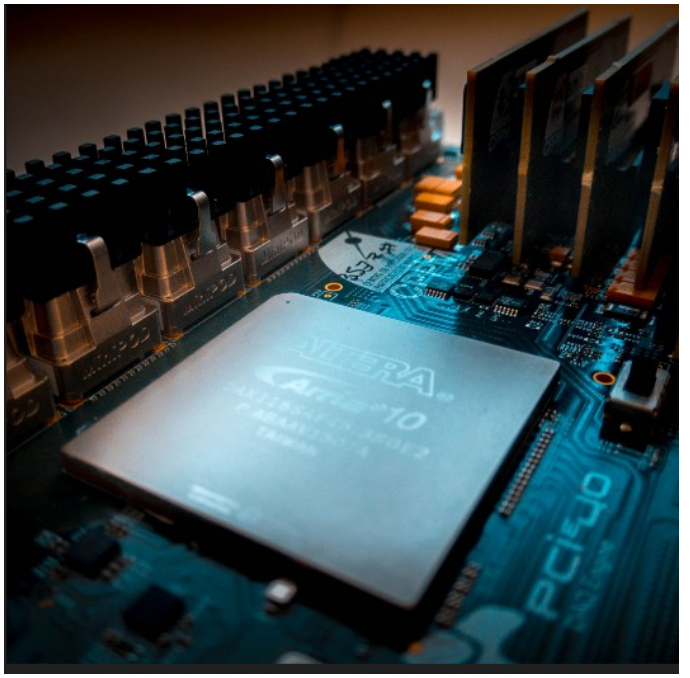




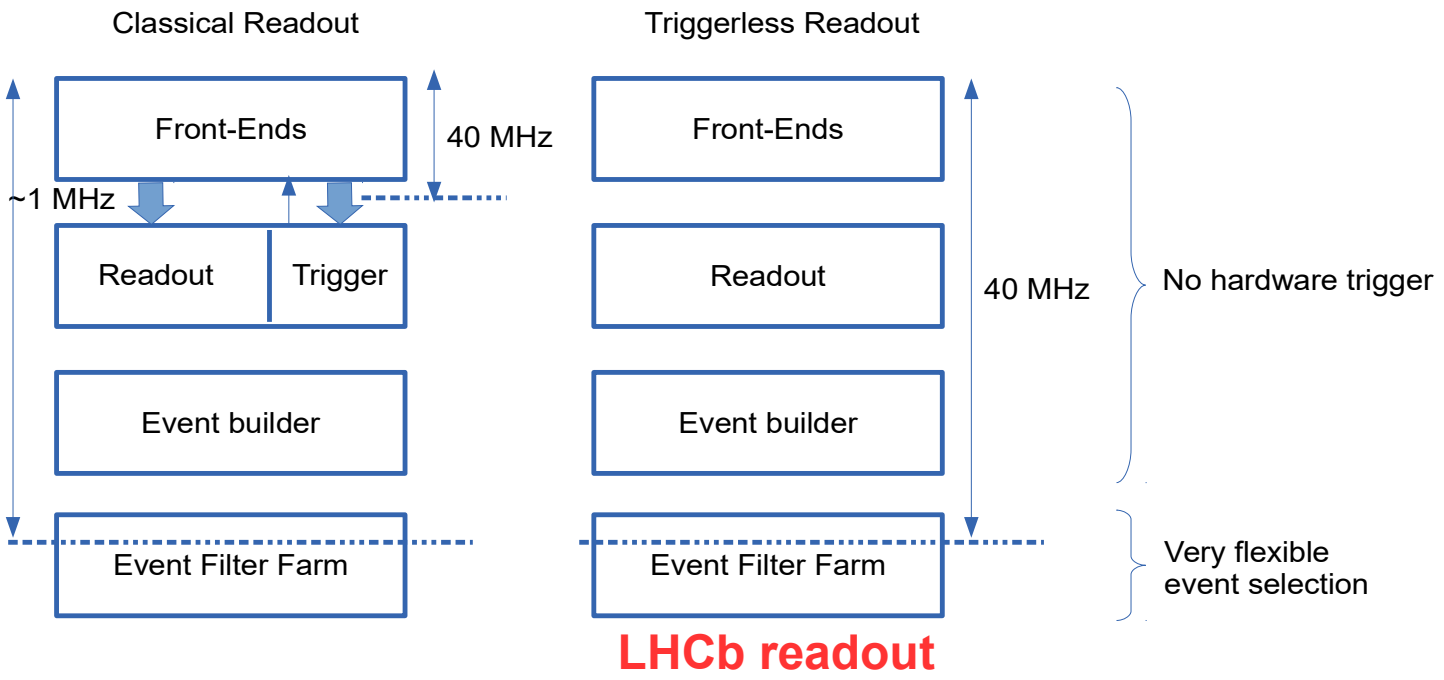
J.P. Cachemiche, on behalf of the LHCb collaboration

Outline

- The PCIe40 card
  - LHCb and ALICE Readout architecture
  - Card main features
  - Measurements
  - Production
- Testing to the limits



LHCb Upgrade key features

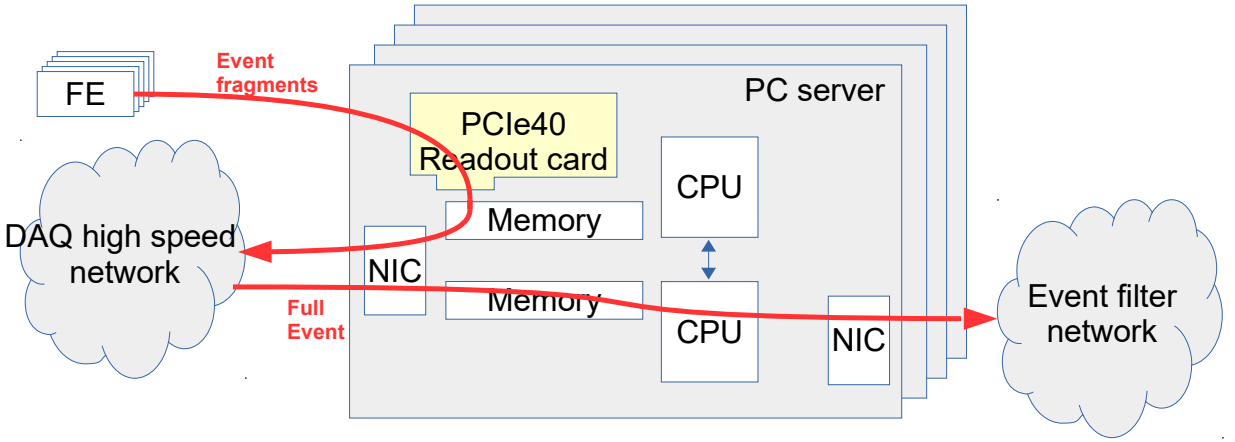


- LHCb uses a triggerless readout
- All event fragments routed at 40 MHz up to the farm

LHCb Upgrade key features

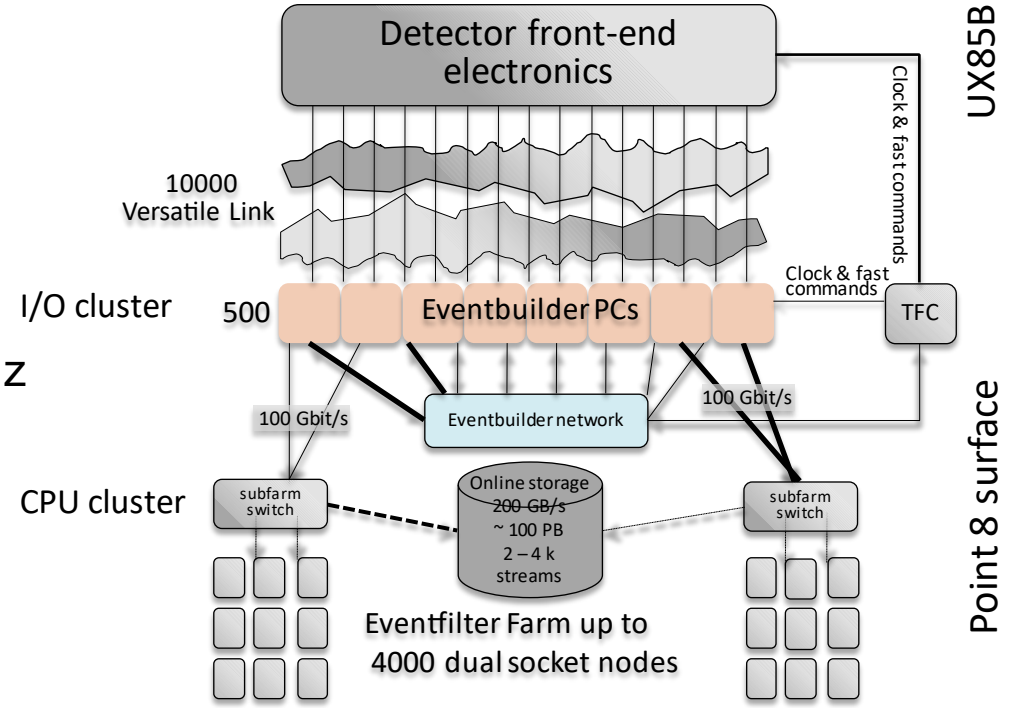
Principle

- Event building done by tightly coupled acquisition boards, CPUs and high speed network
- No intermediate back-end stage
  - Readout card implemented as a PCIe module
- Event building through servers in real time
  - Now possible due to internal CPU architecture evolution
- Event reconstruction with offline quality in real time
- Triggering replaced by filtering of reconstructed events



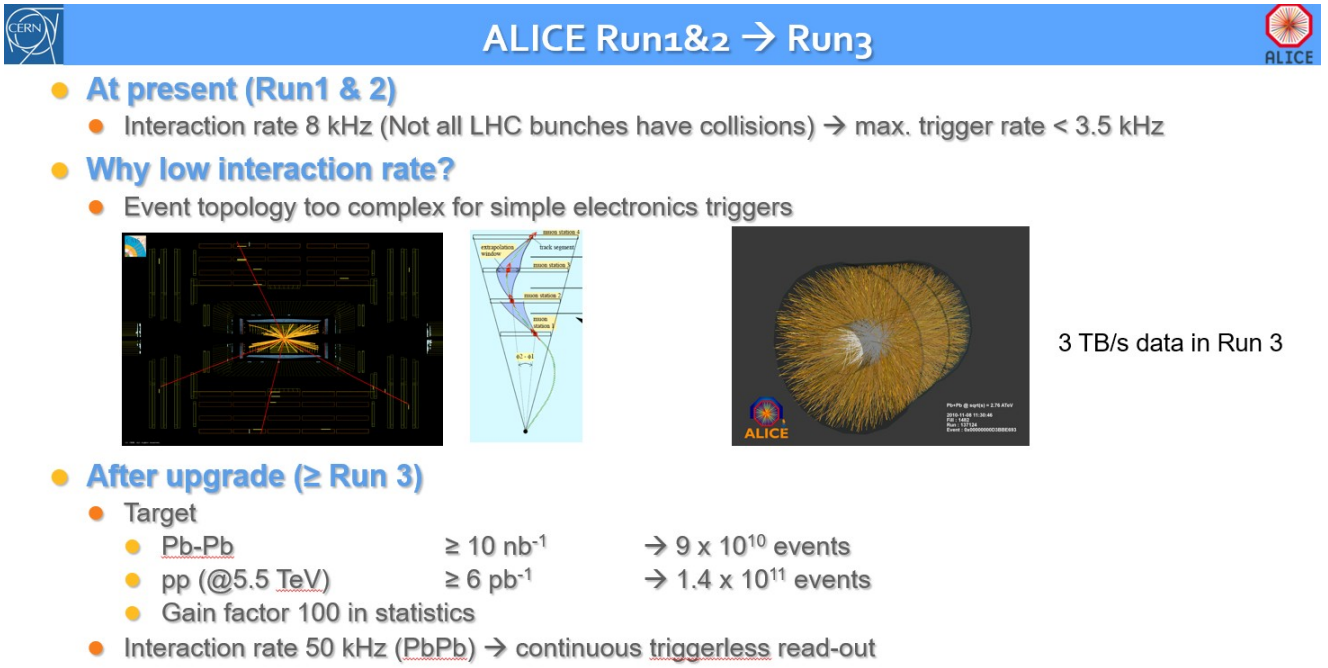
LHCb architecture

- Readout located on surface
  - Distance between FE and RO : ~350m
- ~ 10000 optical links
- ~ 500 readout boards
- ~ 50 TFC/ECS cards
- ~ 100 kBytes per event at 40 MHz
- ~ 32 Tb/s aggregate bandwidth
- ~ 4000 dual CPU nodes



Alice upgrade key features

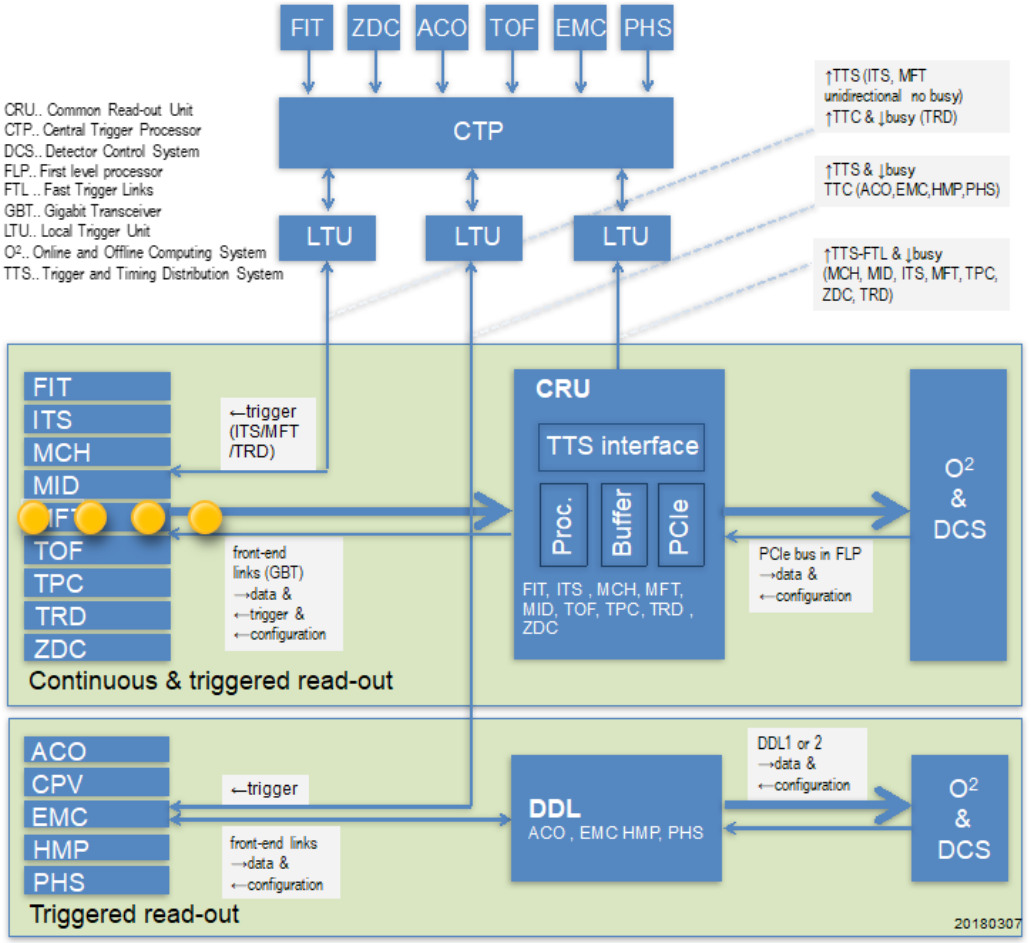
- Event topology too complex for electronics trigger
- 60% of events are kept
  - Continuous triggerless readout + Low interaction rate (50 kHz)
- CRU (Common Readout Unit) based on the PCIe40 card
- Acquires and compresses data on the fly



Courtesy Alex Kluge

ALICE architecture

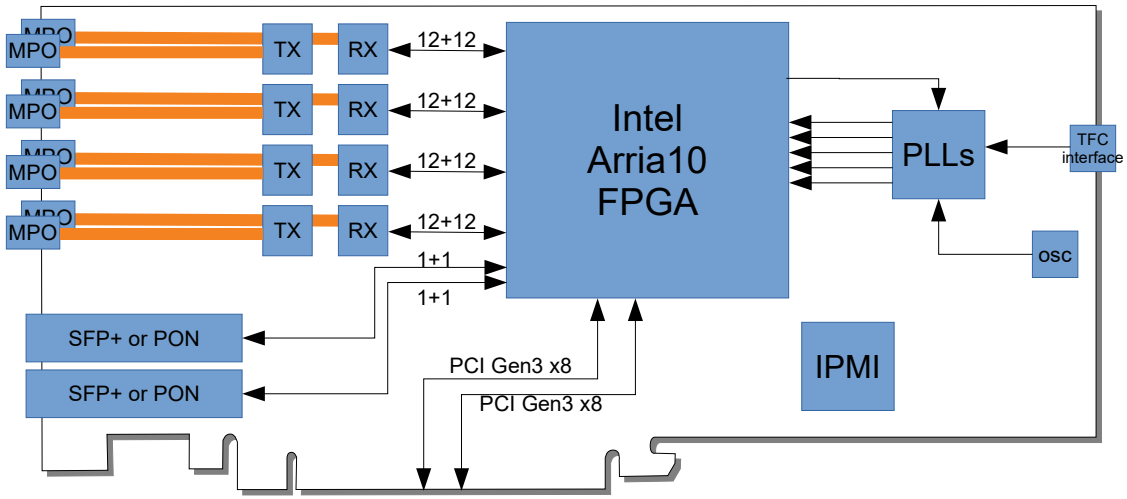
- Readout located on surface
  - Distance between FE and RO : ~120m
- ~ 9000 optical links
- ~ 540 readout boards
- ~ 68 MBytes per event at 50 KHz
- ~ 27 Tb/s aggregate bandwidth
- ~ 1500 GPU based event processing nodes



Courtesy Alex Kluge

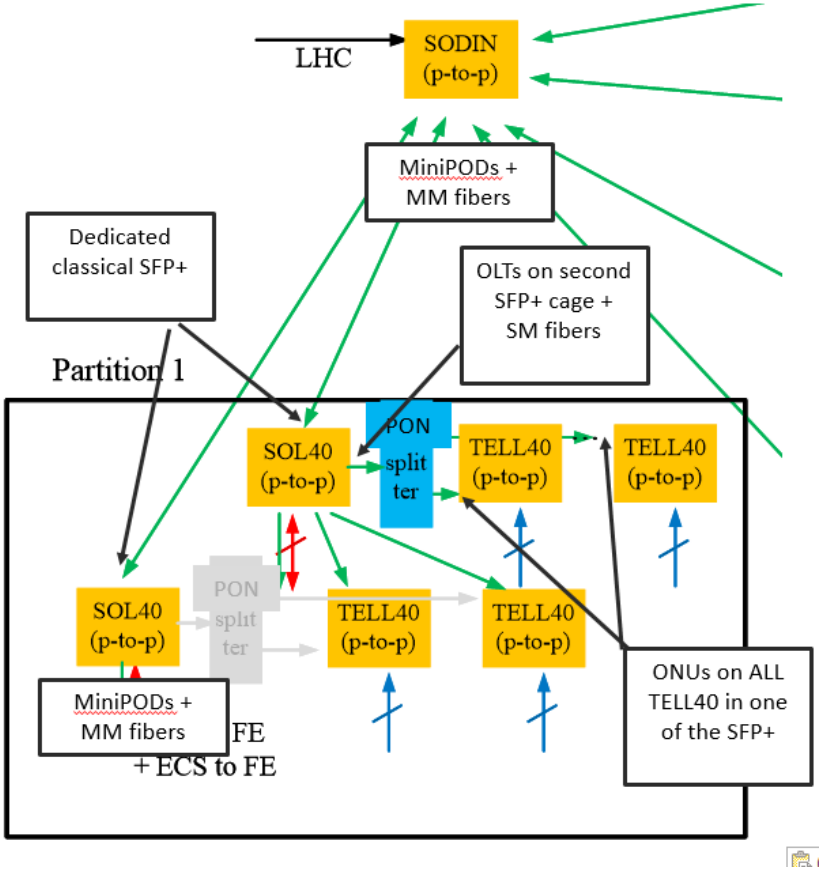
The readout board : PCIe40

- Features :
  - 1 large FPGA 1.15 million cells (Arria10 10AX115S3F45E2SG)
  - 48 bidirectional links running at up to 10 Gbits/s each (minipods)
  - 2 bidirectional links running at up to 10 Gbits/s devoted to time distribution (can use SFP+ or 10G PON devices)
  - Sustained 112 Gbits/s interface with CPU memory through PCIe
  - No on-board buffer memory : we use the PC memory instead
  - Remote reconfiguration of all the programmable devices
  - Fully instrumented: all voltages, currents and temperatures measured



Versatility

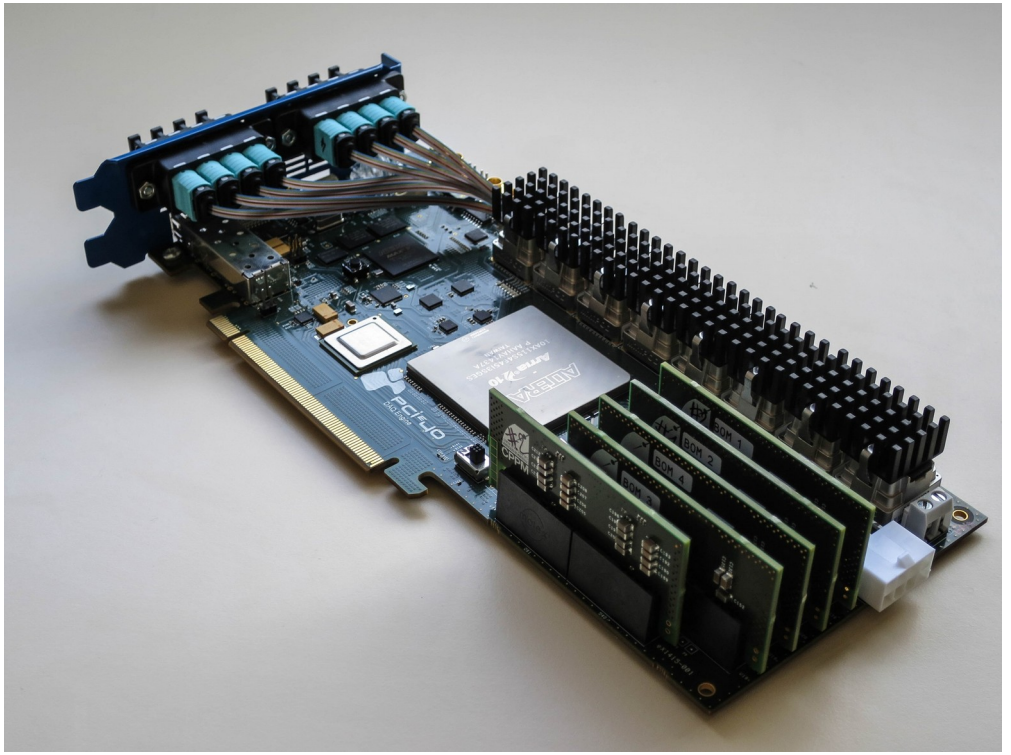
- Can be mapped over several functions by reprogramming the FPGA
- Different names for the same card in LHCb according to its programmation :
  - SODIN : Timing distibution and Fast Control
  - SOL40 : Slow control
  - TELL40 : Acquisition
- Minipods for interfaces with Front Ends
  - GBT protocol at 4.8 Gbits/s
- PON devices for TFC
  - 8B10B protocol at 3.2 Gbits/s



Hardware design

PCIe40 prototype

- First prototype developed in 2016
- 24 copies manufactured for both the LHCb and Alice collaboration
  - Used as « mini DAQ » for debugging front-end cards
  - Programmed to provide acquisition, ECS and TFC in a single firmware



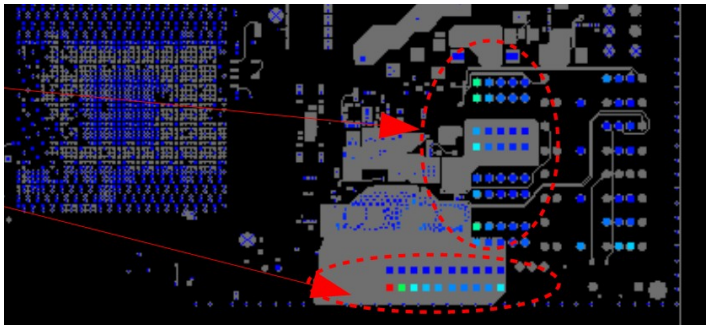
Preparing the final module

Power consumption of large FPGAs very high

- Up to 52 A on the core !
- Power consumption
  - FPGA estimated at ~ 80 W
  - Card estimated at ~ 150 W with Engineering Sample
  - Limited thickness for the stackup

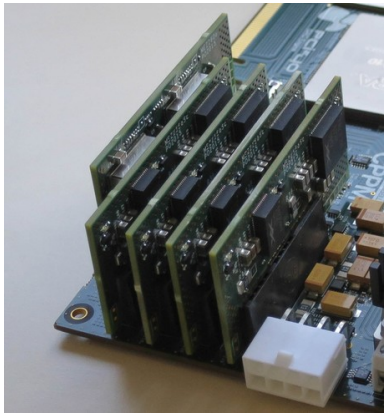
Refining of current flow simulations

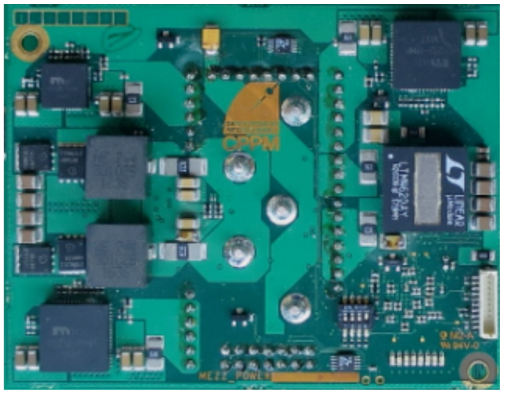
- Simulations of current flow showed dangerous hot spots at full load
  - Power planes have been redesigned and vias placement has been optimized
- Current flow through power mezzanine connections not symetric

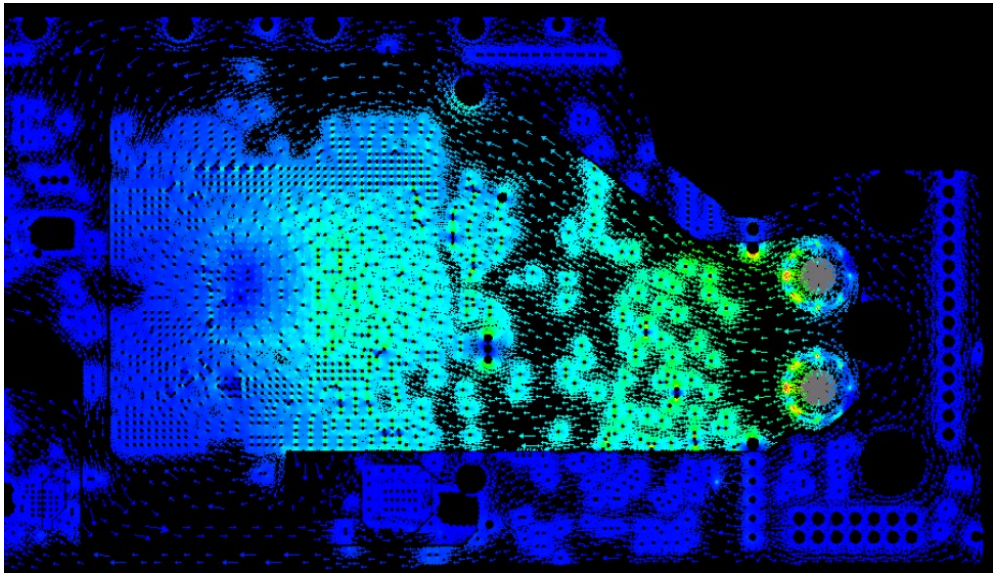


Preparing the final module

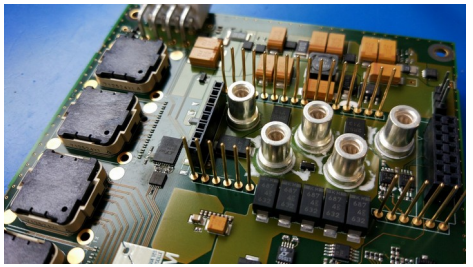
Replacement of the 5 vertical mezzanines by a single flat one







Current flow between mezzanine and FPGA with new design



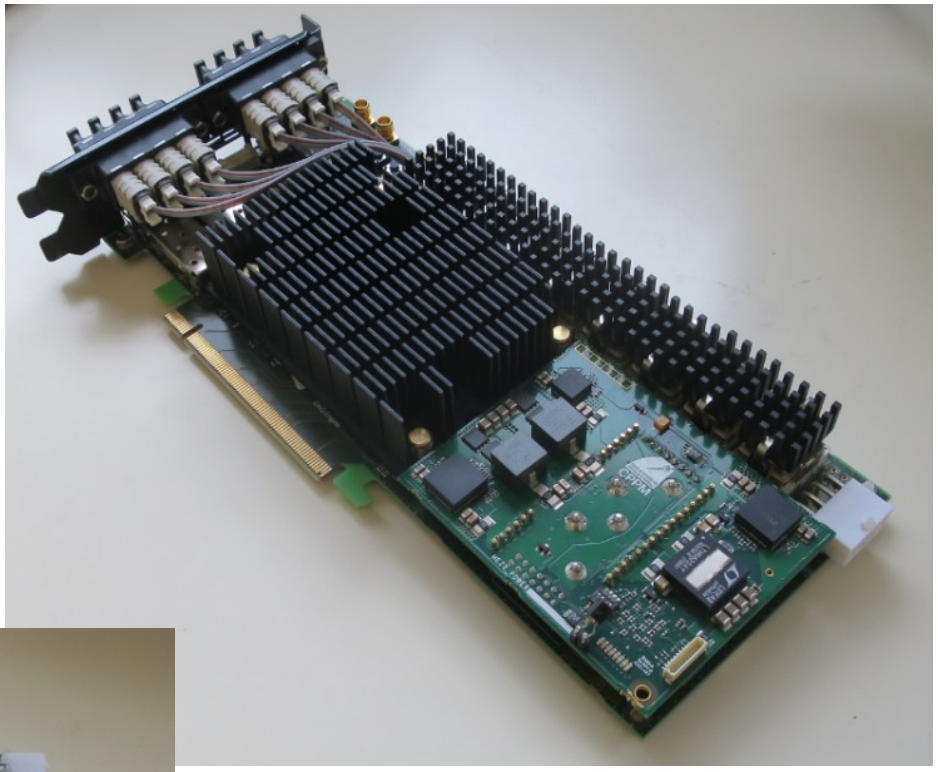
Optimizations

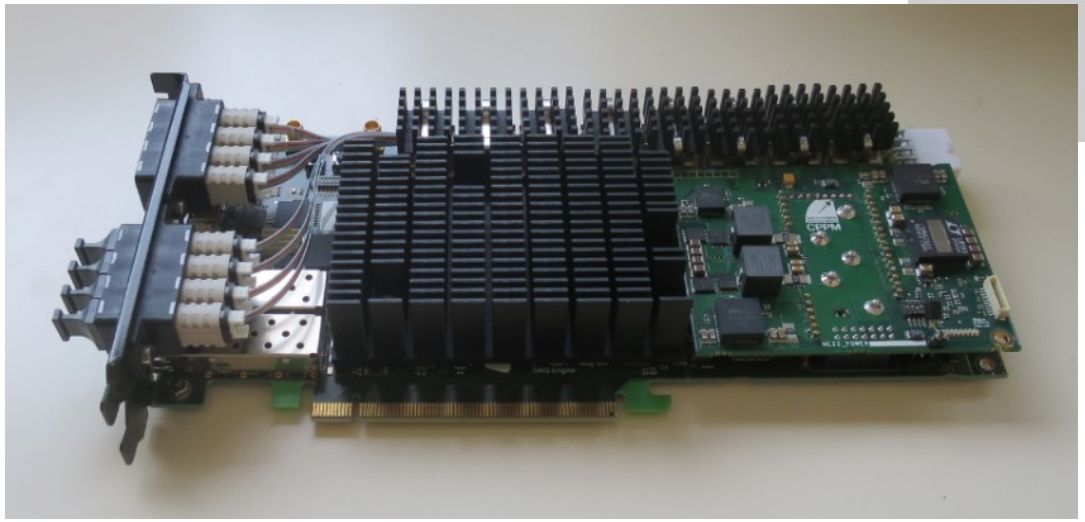
Many improvements

- Cost savings
  - Removal of expensive components (PCIe bridge, Serial Flash and corresponding power supply)
  - One additional SFP+ or PON cage added → less TFC/ECS modules
- Performance improvement
  - Use of new PLLs with a very low jitter compared to previous ones
- Reliability
  - Complete redesign of the power supply due to buggy DCDC converters
  - Optimisation of current flows → avoids local over-heating in the PCB
    → Single power mezzanine now horizontal for symmetrical current flow
  - Improvement of power sequencing to ease maintenance and guaranty a longevity of the module → manages now power down
  - Optimization of decoupling → less noise
  - Heat sink redesign for better cooling
- New functionalities
  - Programming speed multiplied by factor 4 with a new embedded USB Blaster II
  - Serial flash for identificating modules during production
  - IPMI management : allows the system to adjust the fan speed in function of the temperature or automatically cut the power supply if temperature is too high

Final module

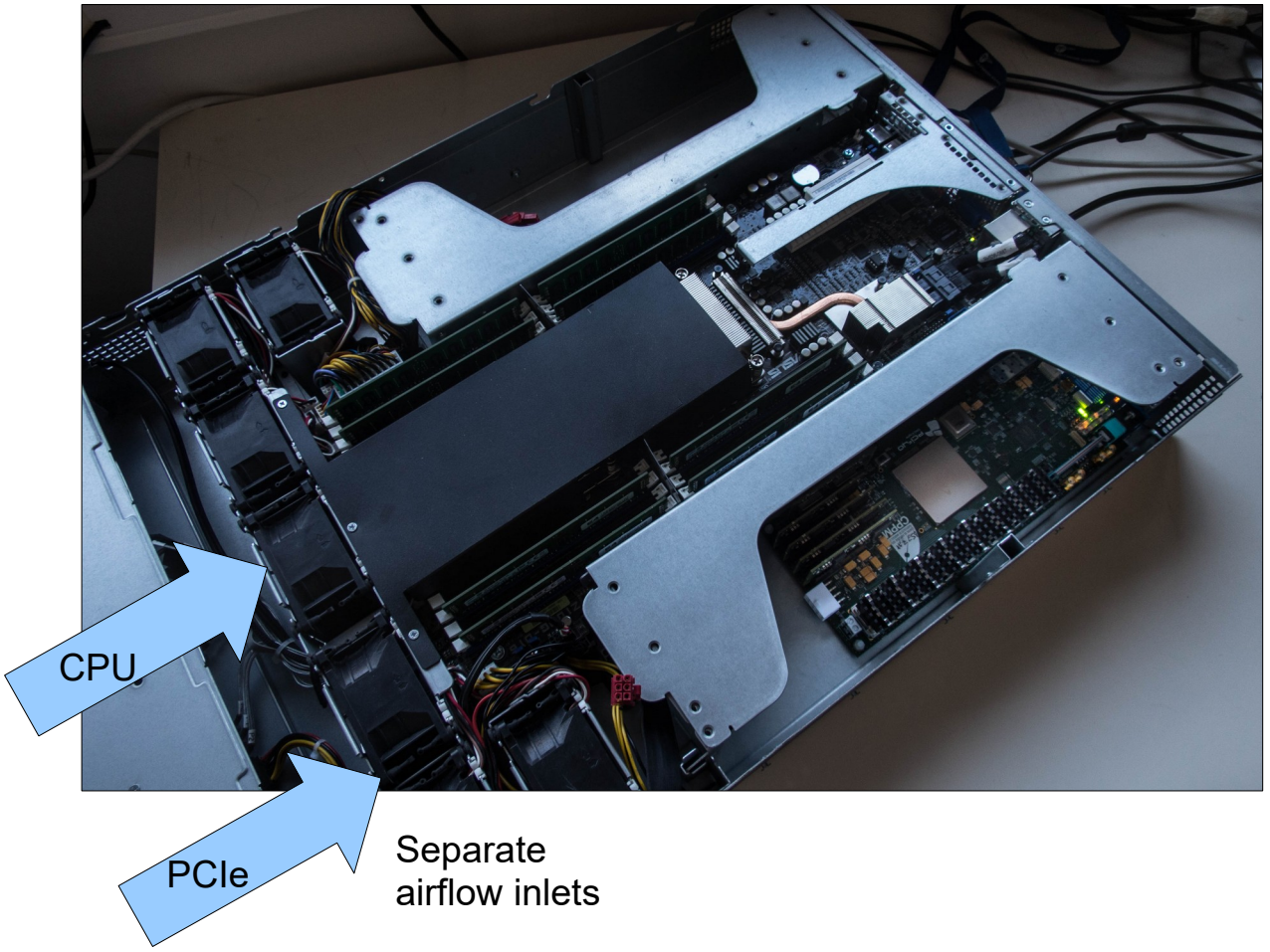
- Two first modules validated end 2017
- Early duplication by Alice of 28 modules to speed up first production





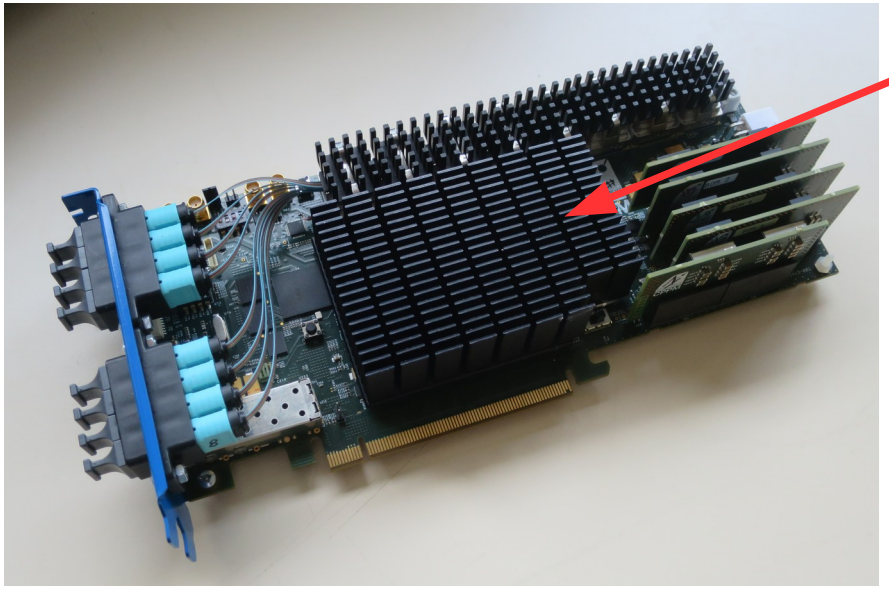
Cooling

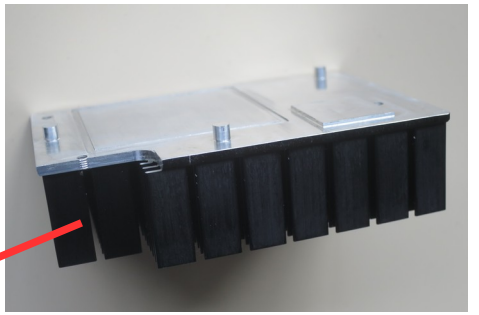
- PC environment not as well defined as xTCA systems
- Very well cooled PC server has been selected



Cooling solution

Use of a custom passive cooling





Custom passive heatsink

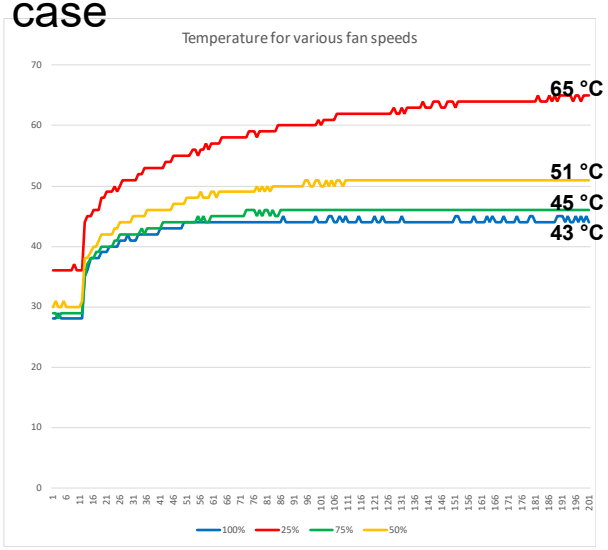
Power consumption and cooling

Power consumption and cooling

- Push the module at the limit of power dissipation
- Principle:
  - Use a « heating function» replicated thousands of times to get an FPGA occupancy of 89%
  - Inject a clock with programmable frequency between 10 MHz and 600 MHz
- Automatic power off if the FPGA temperature overpasses 82°C
- Vary the speed of server fans (25%, 50%, 75%, 100%)
- Measure voltages, currents and temperature in each case

Results obtained with ASUS server

- 2 cards on same side
- Passive cooling seems sufficient



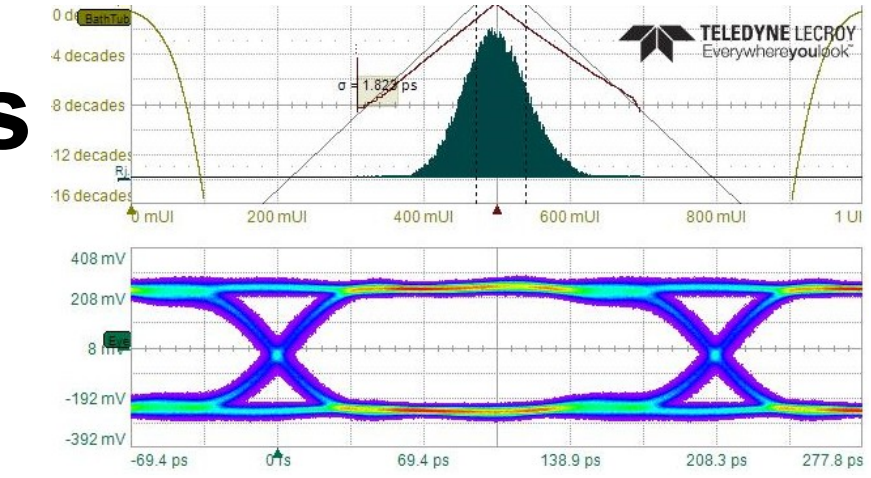
FPGA temperature for several fan speeds in ASUS server

Links measurements

BER << 10^{-16}

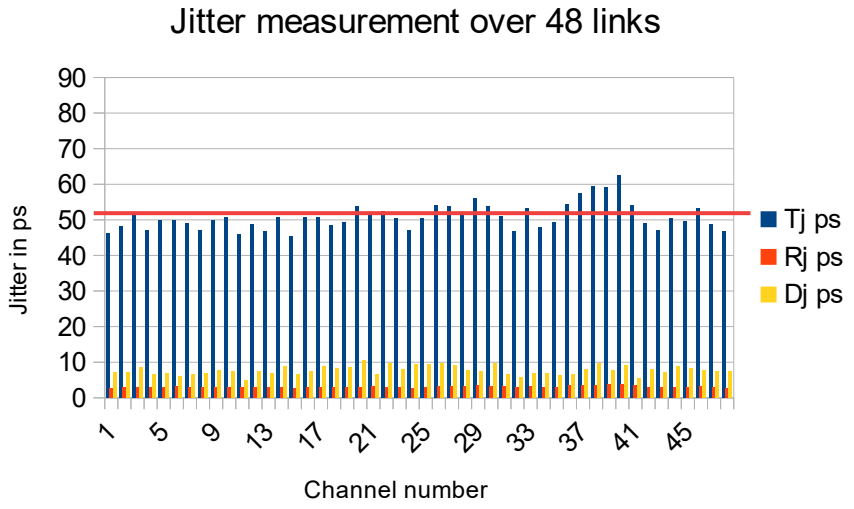
Jitter

- Final card jitter improved vs prototype
  Total jitter goes from 51 ps → 38 ps

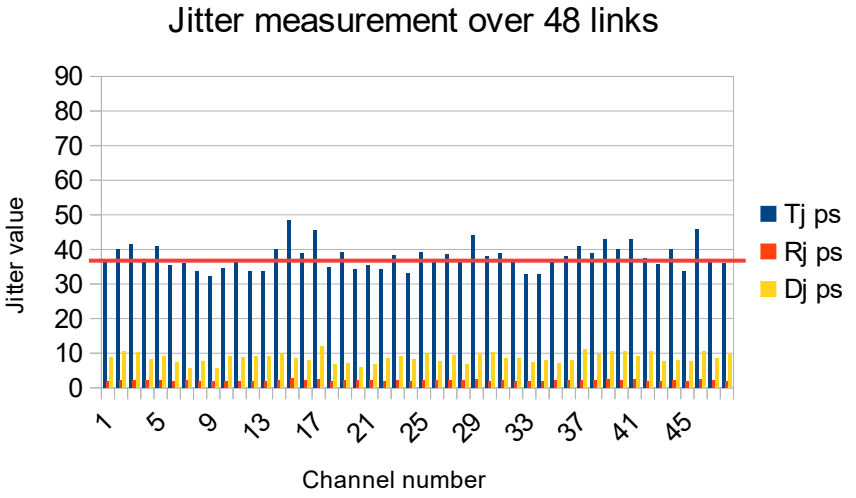


SDA Jitter	Tj(1e-12.0)	Rj(sp)	Dj(sp)	Pj	ISI	DCD	DDj	BitRate
Lane1	36.902 ps	1.995 ps	8.847 ps	6.89 ps	6.902 ps	4.795 ps	11.418 ps	4.80942 Gbit/s
SDA Eye	EyeHeight	EyeOne	EyeZero	EyeAmpl	EyeWidth	EyeCross	EyeAvgPwr	
Lane1	424.5 mV	248.9 mV	-247.7 mV	496.5 mV	192.1 ps	45.52 %	73 µV	

Measurements at reception stage for a PRBS31 pattern running at 4.8 Gbits/s



Prototype

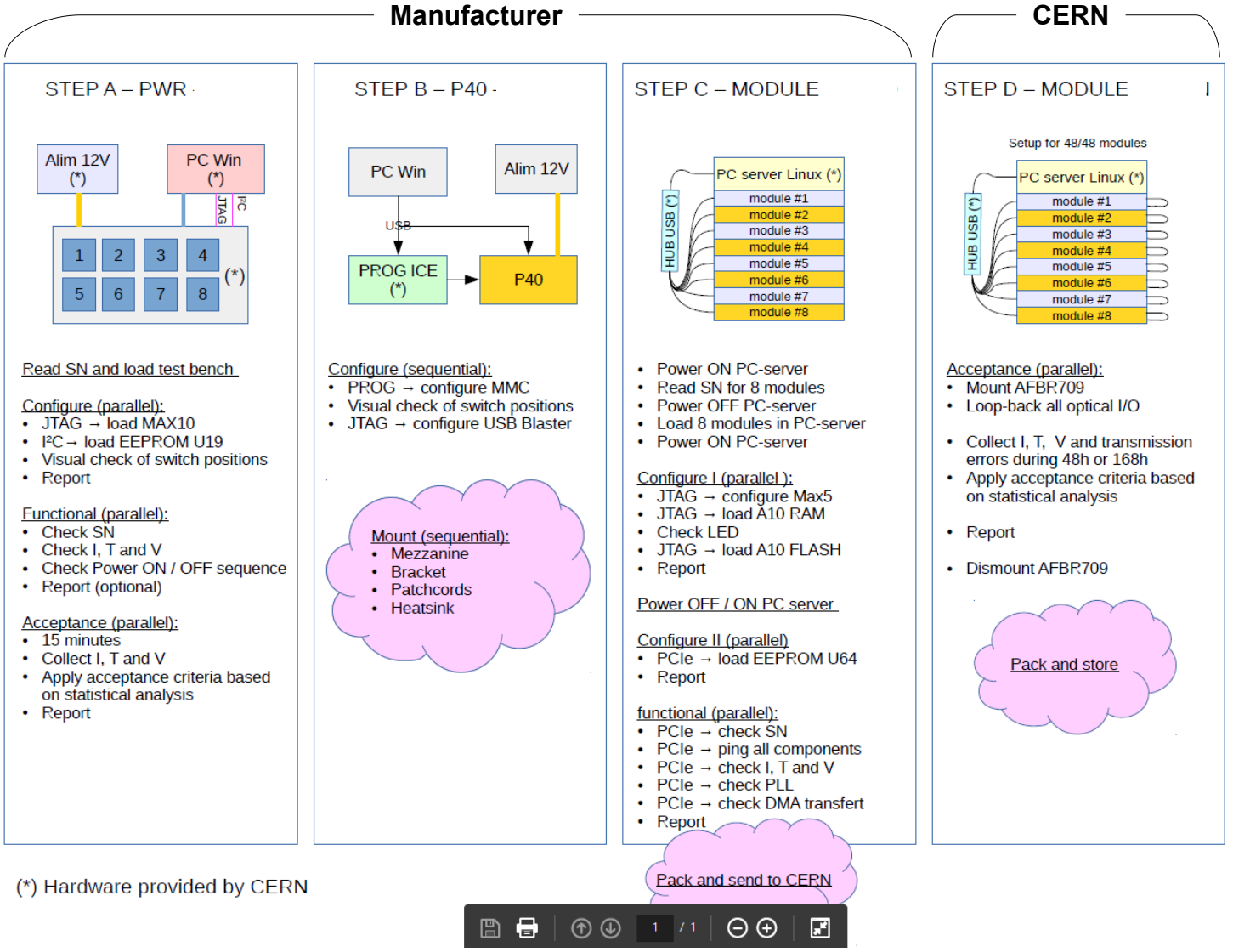


Final card

Production

Testing methodology

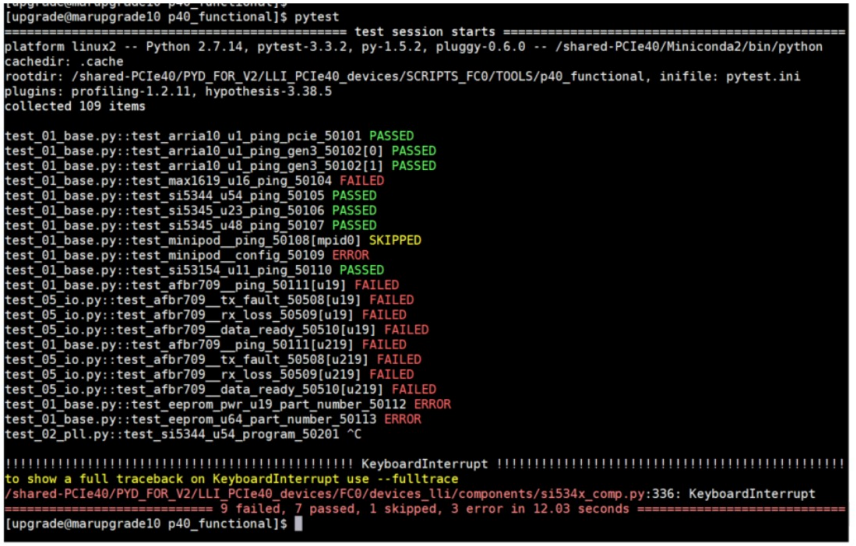
4 steps



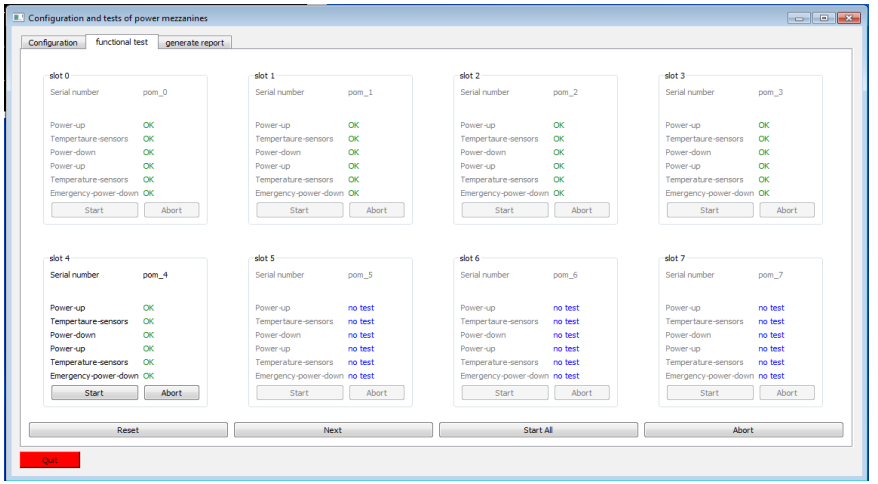
Production tests

Run in assembly company

- Based on Pytest
  - Very flexible command line testing tool
  - Able to test target sub-set of components
  - Object oriented design
  - Can be driven by a GUI or ncurse
- Fully tests the board
  - ~ 146 unitary tests ran in a few minutes on 8 cards at a time
  - Check the operation of all the devices on the modules
  - Measure voltages, currents, temperatures, frequencies, etc.
  - Produces test reports for each module
- Centralized management of reports
  - Reports directly sent to CERN data base



Expert interface

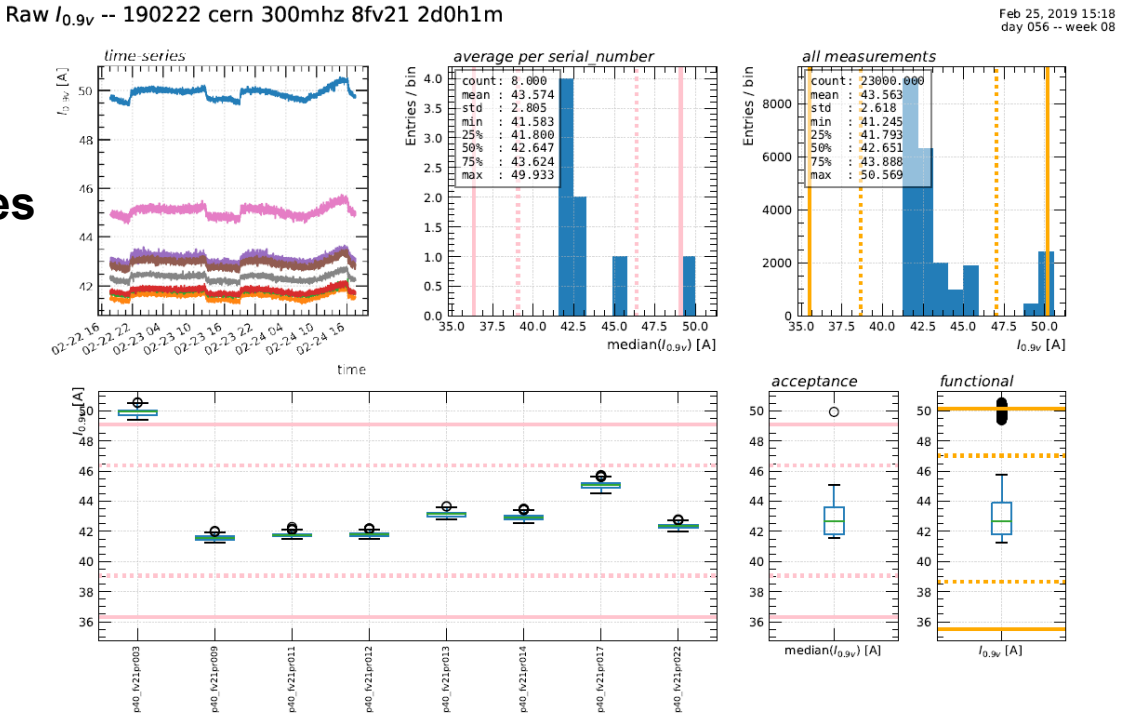


Example of operator interface

Acceptance tests

Run at CERN

- Duration 24 or 168 hours
  Allow to eliminate **early failures**
- Rely on Pytest
- Possible post processing of results
- Logged in data bases



FPGA core current measurement on 8 cards

Testing to the limits

Not everything is pink

Which firmware for testing ?

Target design

- Up to 100% occupancy
- Average clock 250 MHz
- Average toggle rate : 50%

How to test the design at maximum load ?

- Final firmware was not ready (will it be one day ?)
- Since then there is a preliminary one but, very difficult to handle
    - Requires WIN-CC
    - Complicate initialization
    - Not scriptable
- Fixed configuration

We decided to emulate it

- Not a perfect emulation because many unknown features
- But scalable design allowing to explore the limits set when designing the card

Firmware emulation

LLI (Low Level Interface)
+ programmable load emulation

- n*80 blocs of 16 random pattern generators

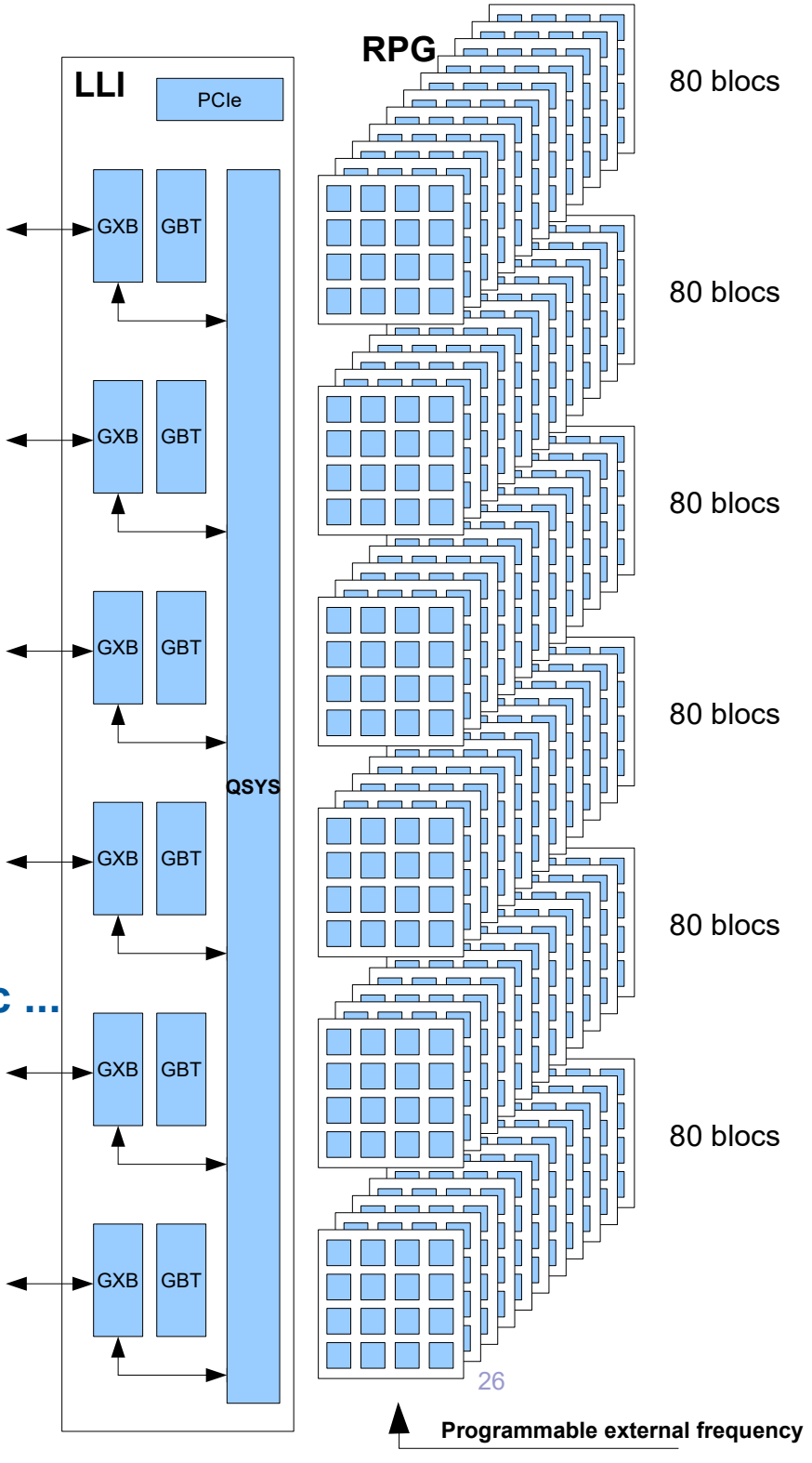
Number of Macro blocks	Number of blocks	Block size	Individual RPG size	Total number of RPG	FPGA occupancy
0	80	16	128 bits	0	14%
1	80	16	128 bits	1280	27%
2	80	16	128 bits	2560	39%
3	80	16	128 bits	3840	52%
4	80	16	128 bits	5120	65%
5	80	16	128 bits	6400	78%
6	80	16	128 bits	7680	89%

- Programmable frequency injected in random pattern generators
  - SI5344 PLL embedded on the card

Initial goal: checking power supply, cooling, etc ...

Extended to checking errors at full load

- BER tests made with GBX internal PRBS generators and checkers
  - TTK-like test by addressing GXB registers by software

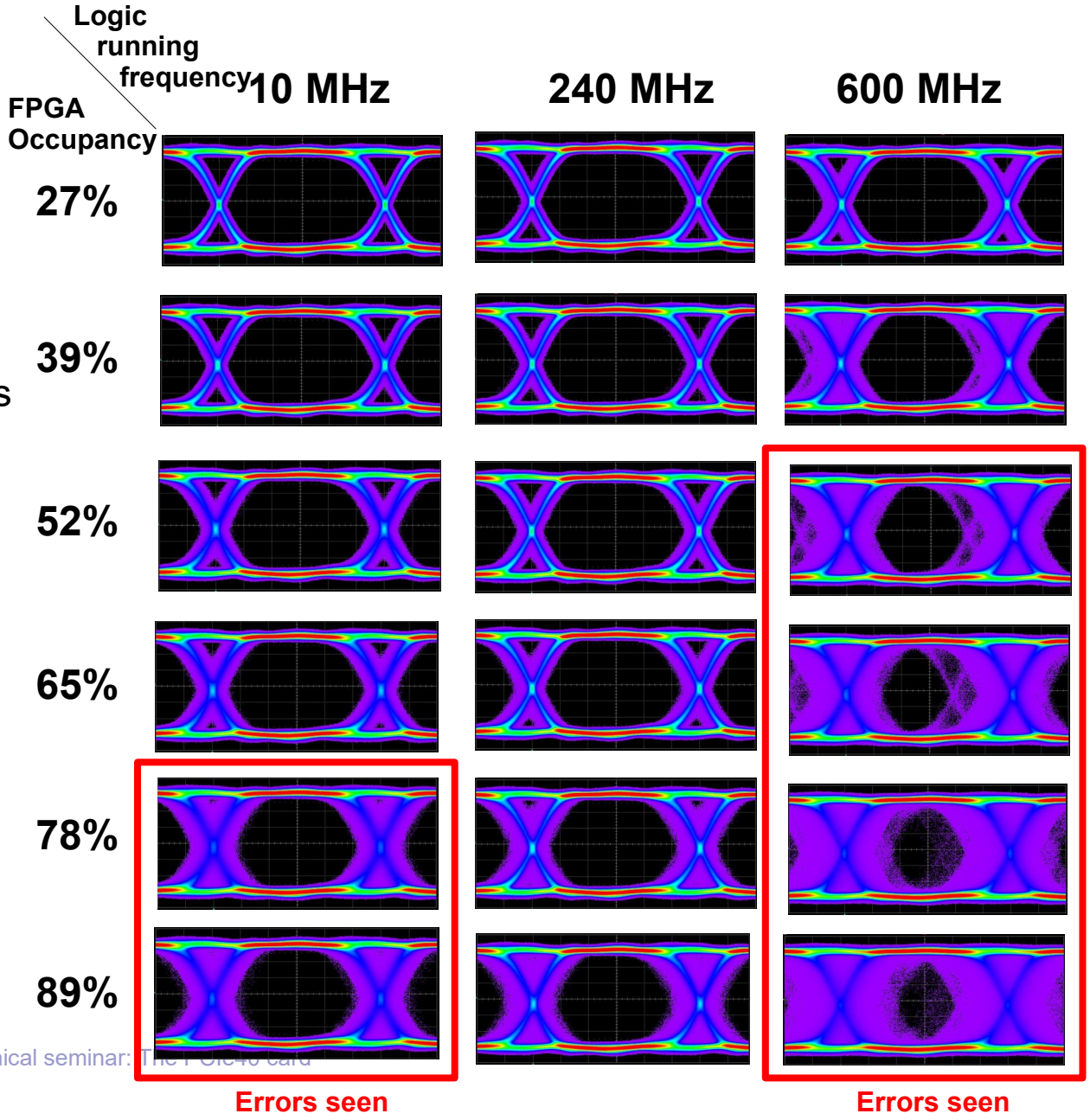


Error checking vs occupancy

Conditions

- 48 links at 4.8 Gbits/s
- PRBS31
- Internal loopback mode
- Load emulator made of pseudo-random generators
- Increasing occupancy by varying the number of random generators
- Programmable clock frequency for internal logic
  - Design runs at 10 240 or 600 MHz

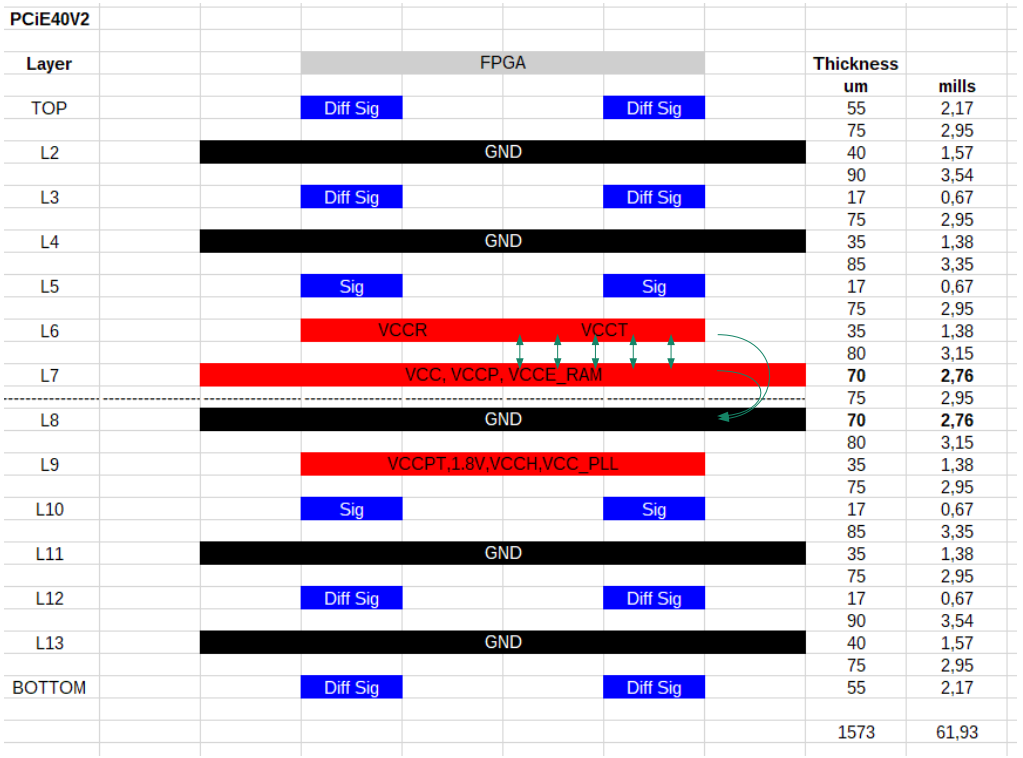
Unexpected errors seen at high occupancy!



Probable cause

VCCR/VCCT plane and VCC plane proximity

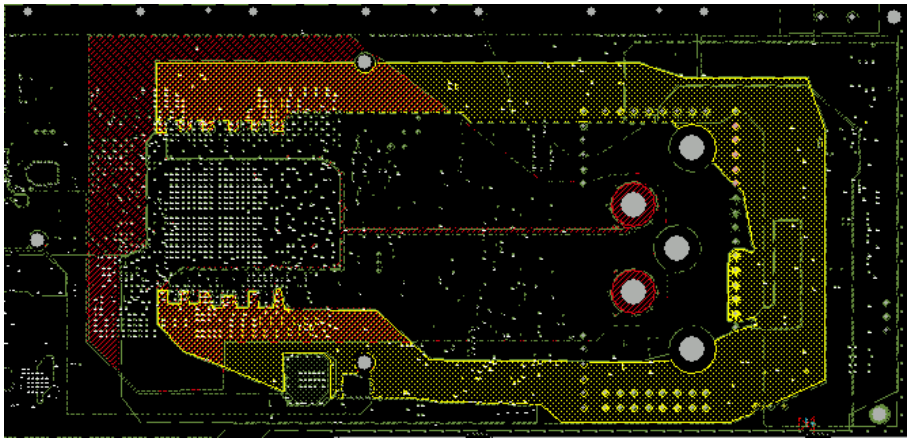
- Both capacitive and inductive effects



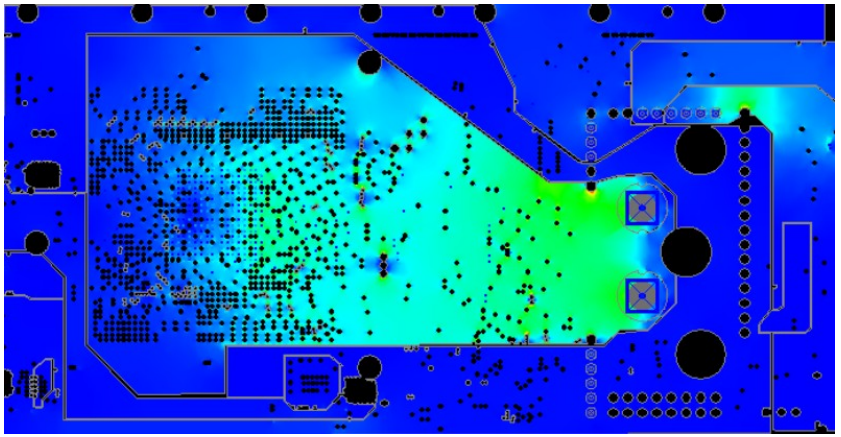
Probable cause

Overlap mostly of VCCT plane and VCC

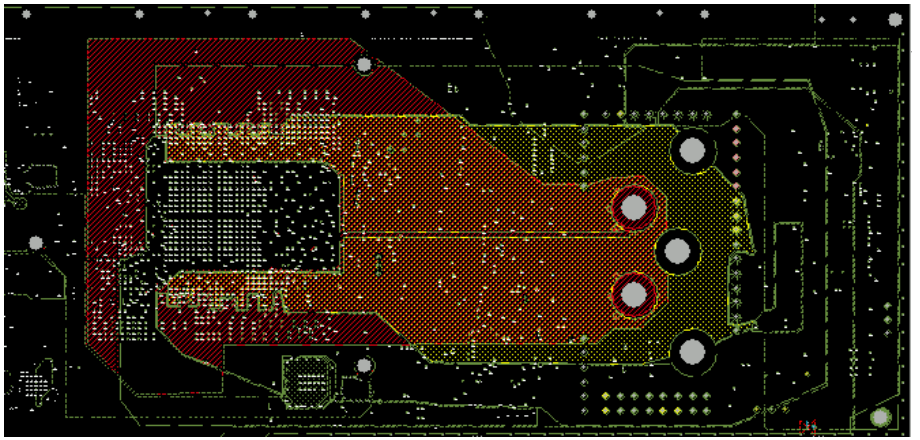
- Partial overlap between VCCR and VCC
  - Weak because in an area with nearly no current
- Large overlap between VCCT and VCC
  - Strong : high currents here



VCCR overlap (orange)



Current density in VCC (0.9V)



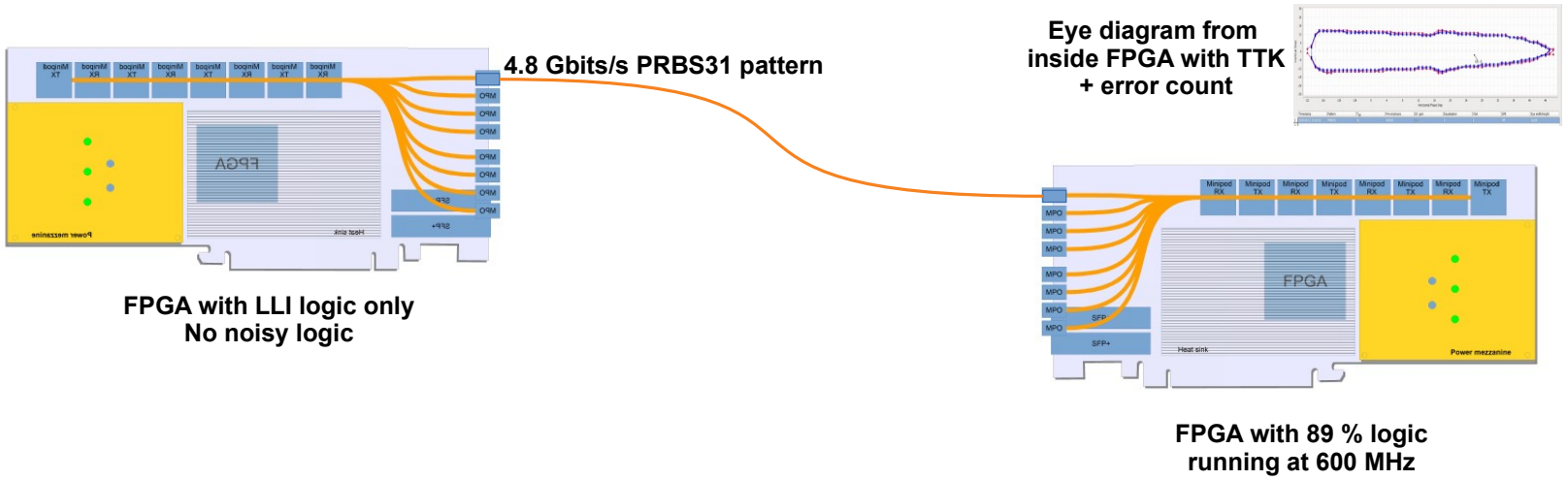
VCCT overlap (orange)

Checking hypothesis

If weak overlap between VCCR and VCC receiving side should not be affected

Verification

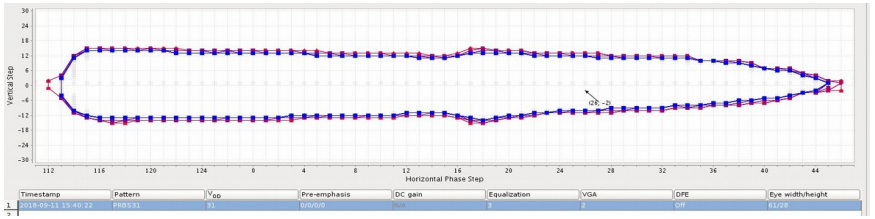
- Use of two cards:
  - Emitting card not loaded → emission of a quiet signal
  - Receiving card fully loaded (89%) + injection of frequencies 10, 240 and 600 MHz
- Check error count with the Transceiver Tool Kit in Quartus
- Check eye diagram from inside of receiving FPGA



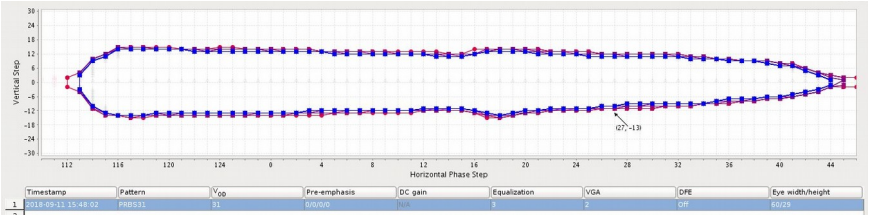
Receiving side

Measurement results

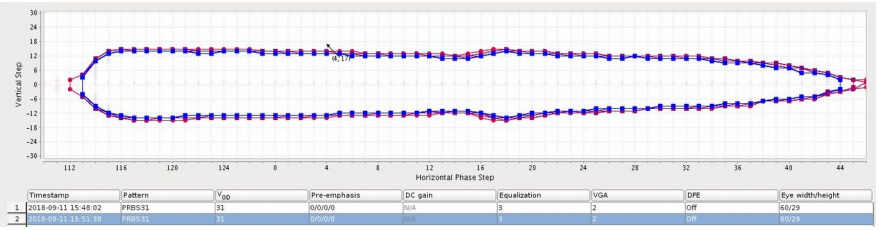
- **No error, even at full load**
  (89% occupancy, 600 MHz)
- No degradation of eye diagram



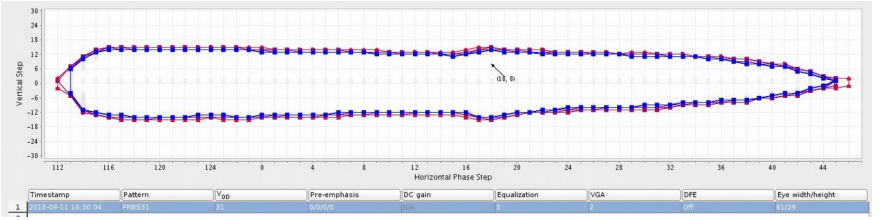
Reference measurement : LLI only - no load



89% occupancy – running frequency = 10 MHz



89% occupancy – running frequency = 240 MHz

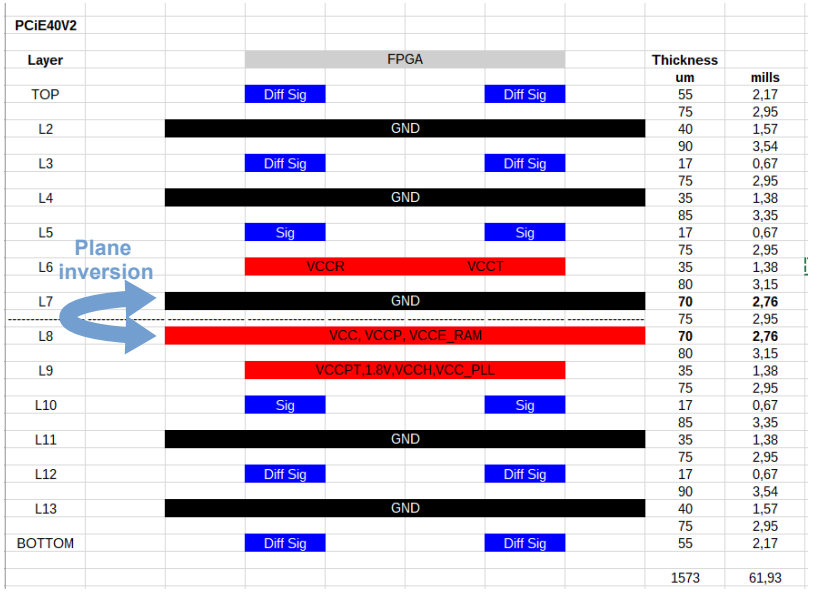


89% occupancy – running frequency = 600 MHz

Corrective action

Invert power and ground planes

- 3 possibilities
- chosen 2 and 3



① **Invert GND and VCCT/VCCR**
Inconvenient : coupling between VCC (0.9) and VCCPT/1.8V/H

② **Invert VCC (0.9V) and VCCT/VCCR**
Invert bottom GND and VCCPT/1.8V/VCCH
Inconvenient : diff signals over cut power plane

③ **Invert VCC (0.9V) and VCCT/VCCR**
Invert top GND and VCCT/VCCR and
Invert bottom GND and VCCPT/1.8V/VCCH
Inconvenient : more diff signals over cut power plane

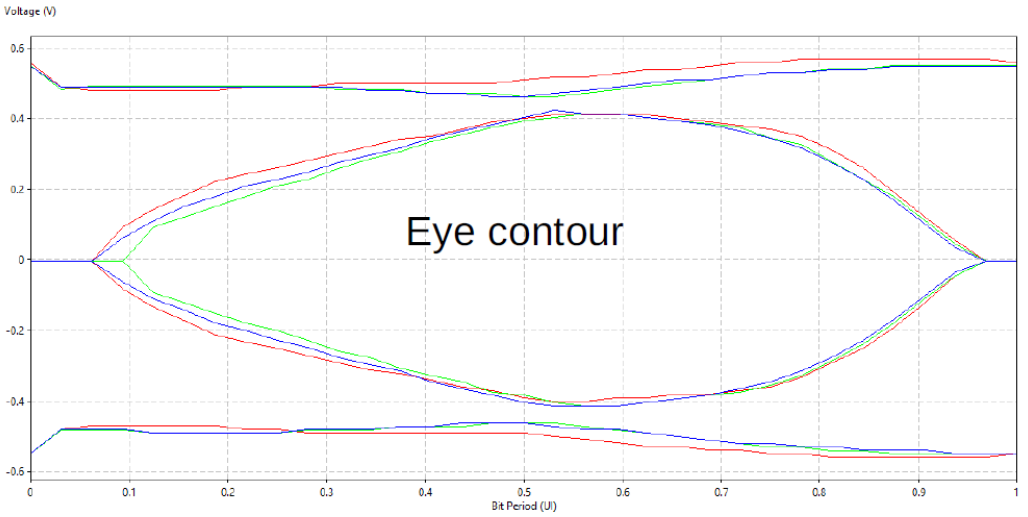
Diff signals simulation

Simulation to mitigate the risk of signal integrity issue if solution 2 and 3 are chosen

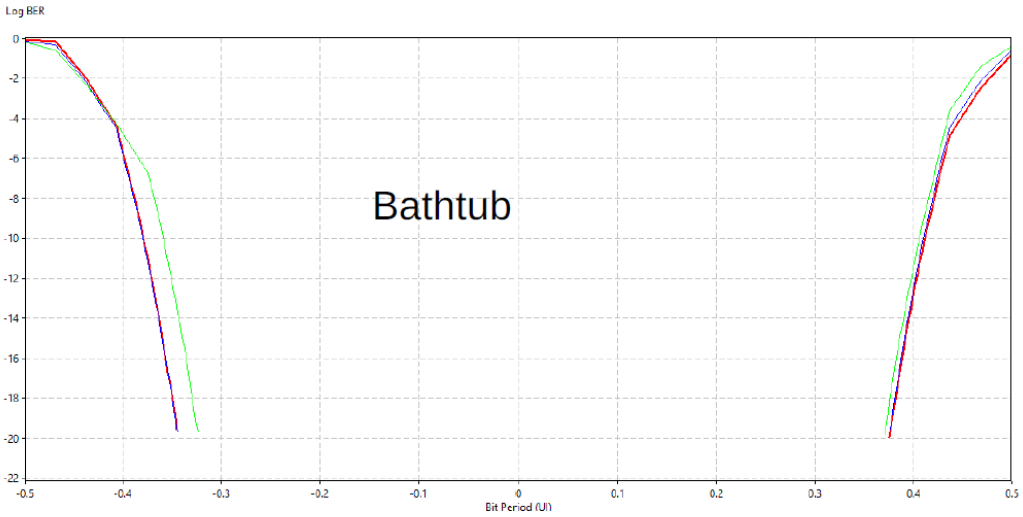
3 cases tested with same diff track:

- **Initial**: diff signal between 2 continuous GND planes
- **Simple plane inversion**: diff signal between continuous GND plane and power plane (5 cuttings)
- **Improved plane inversion**: power plane rearranged (2 cuttings only)

→ Negligible loss



10 Gbps simulation results with Sigrity



Error measurements

Building of tools to measure errors at the 3 maximum occupancy rates

- Conditions:
  - Serial links: PRBS31 at 4.8 Gbits/s
  - Test duration 100s per frequency
  - Internal serial loop back
  - Measurement of total number of errors on all 48 links
- **There exists a quite wide domain of operation without error**
- Correct operation range decreases with FPGA occupancy

SOL40

TELL40
CRU

Frequency	Time	Average temperature	Average current	Errors found
10	100	51	5693	0
20	100	48	5769	1083694
40	100	47	6395	0
80	100	48	7768	0
120	100	49	9156	0
160	100	50	10552	0
200	100	51	11931	0
240	100	52	13076	0
280	100	53	14721	0
320	100	54	16130	0
360	100	55	17516	0
400	100	57	18900	0
440	100	58	20337	0
480	100	59	21778	0
520	100	61	23218	683
560	100	62	24674	1365
600	100	63	26114	24595

65%

Frequency	Time	Average temperature	Average current	Errors found
10	100	45	5323	188961960
20	100	45	5708	1182867725
40	100	46	6523	0
80	100	47	8174	0
120	100	48	9828	0
160	100	49	11495	0
200	100	51	13141	0
240	100	52	14808	0
280	100	53	16478	0
320	100	55	18162	0
360	100	56	19847	0
400	100	58	21551	0
440	100	59	23229	162
480	100	61	24997	1960
520	100	62	26665	6432
560	100	63	28342	86571
600	100	65	30110	2927490

78%

Frequency	Time	Average temperature	Average current	Error found
10	100	51	5905	432194662
20	100	47	6050	127539359177
40	100	48	6983	0
80	100	49	8952	0
120	100	50	10927	0
160	100	51	12915	0
200	100	52	14879	0
240	100	54	16868	0
280	100	55	18863	0
320	100	57	20881	0
360	100	59	22895	0
400	100	60	24923	271
440	100	62	26955	2863
480	100	63	28997	2693403
520	100	64	30941	11665646
560	100	66	33051	354496016
600	100	68	35212	7818417758

89%

Domain of operation around 40 MHz

Same measurements with a finer granularity (5 MHz)

- Peak of errors centered on 25 MHz
- No resurgence of errors after 30 MHz

	Frequency	Time	Average temperature	Average current	Errors found
	5	100	41	5267	72
	10	100	42	5896	0
	15	100	43	6537	0
	20	100	44	7137	205117
	25	100	45	7660	18604767359276
	30	100	47	8399	305779605112
	35	100	48	9152	0
SOL40	40	100	48	9699	0
	45	100	49	11730	0
	50	100	48	11071	0
	55	100	49	11759	0
	60	100	52	12465	0
	65	100	54	13153	0
	70	100	55	13826	0
	75	100	55	14508	0
	80	100	54	15189	0
	85	100	53	15857	0

Modified cards

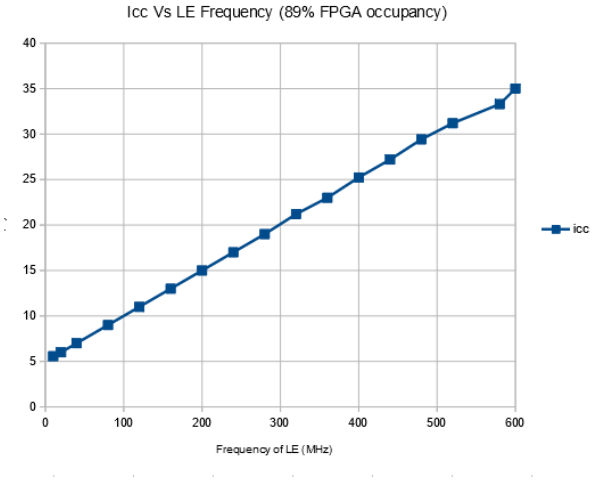
Same results with 2.1 (initial), 2.2.2 and 2.2.3 modified cards !

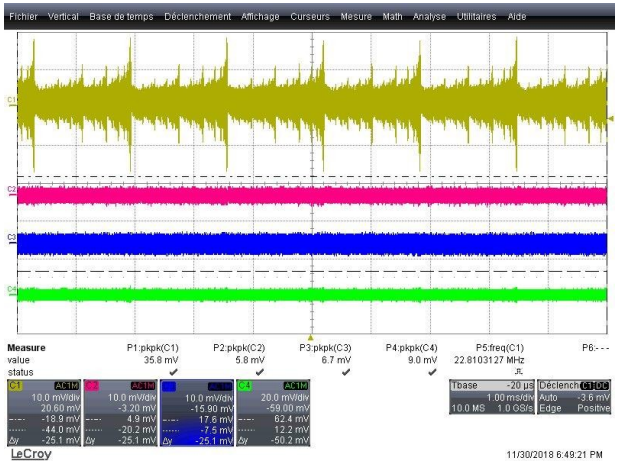
Decoupling ?

Significant noise on VCC

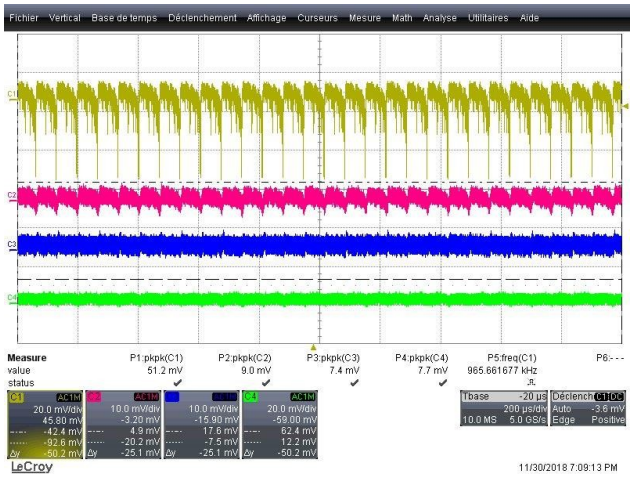
- **High frequency** cyclic noise
- Bursts of **low frequency** noise
- Proportional to injected frequency

- VCCR and VCCT preserved

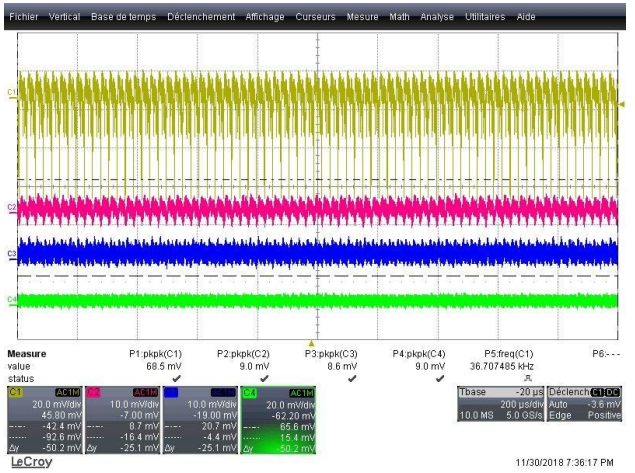




Ripple noise at 10 MHz



Ripple noise at 240 MHz

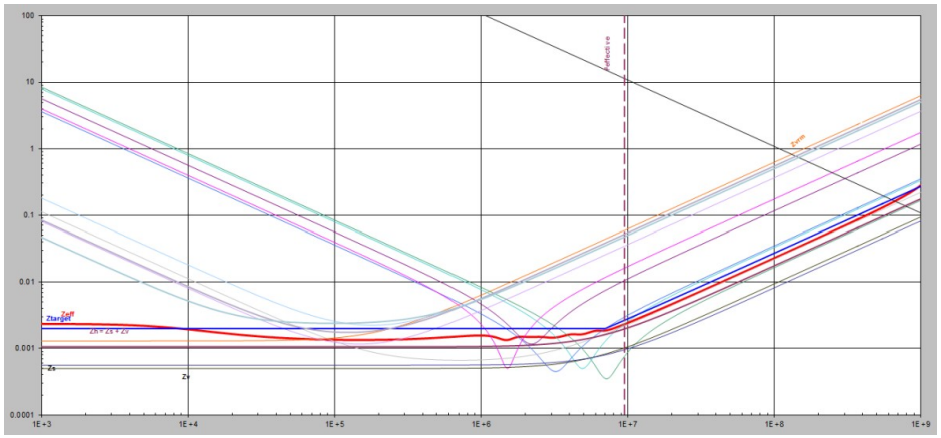


Ripple noise at 600 MHz

Decoupling simulations and measured ripple noise

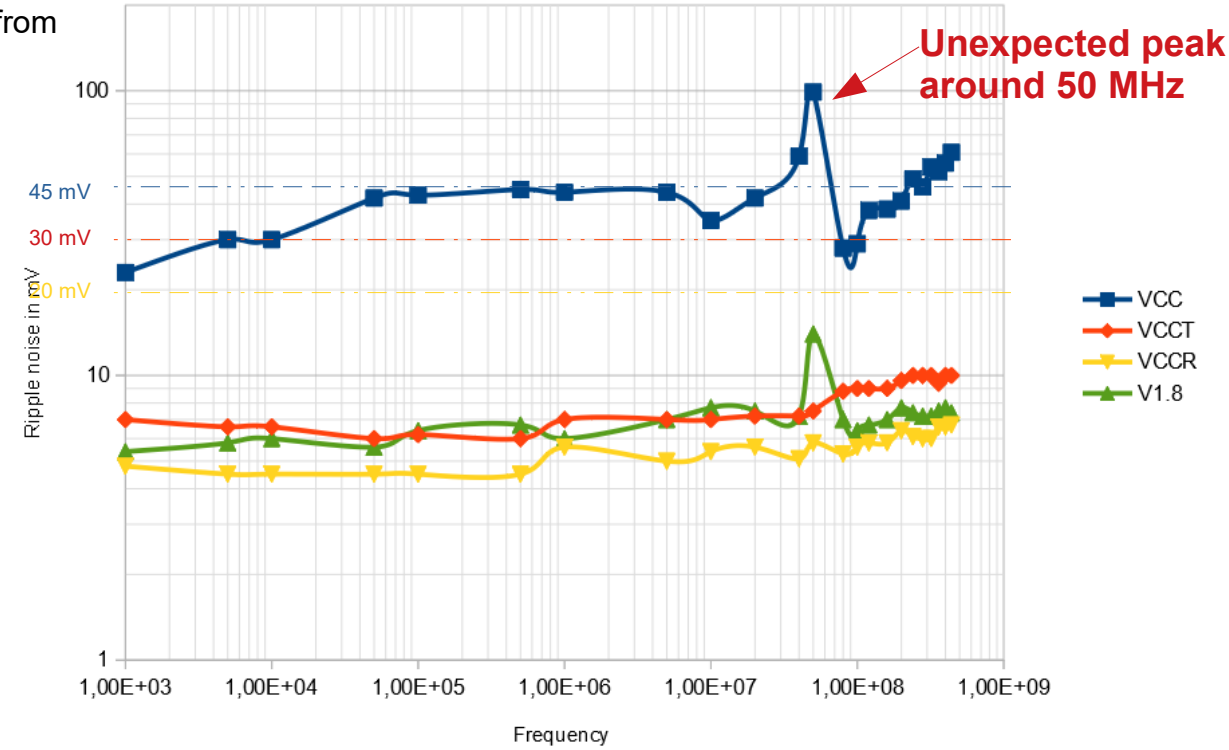
Simulations

- Made with Intel PDN tool
- Impedance below Ztarget except very low frequencies
- Simulations cross checked with similar results with:
  - Sigrity: takes into account board geometry and exact BOM
  - Ansys by a CERN expert from Alice (Michel Morel)



Measured ripple noise

- Authorized max values:
  - VCC : 45 mV
  - VCCT: 20 mV
  - VCCR: 30 mV
  - 1.8V : 54 mV



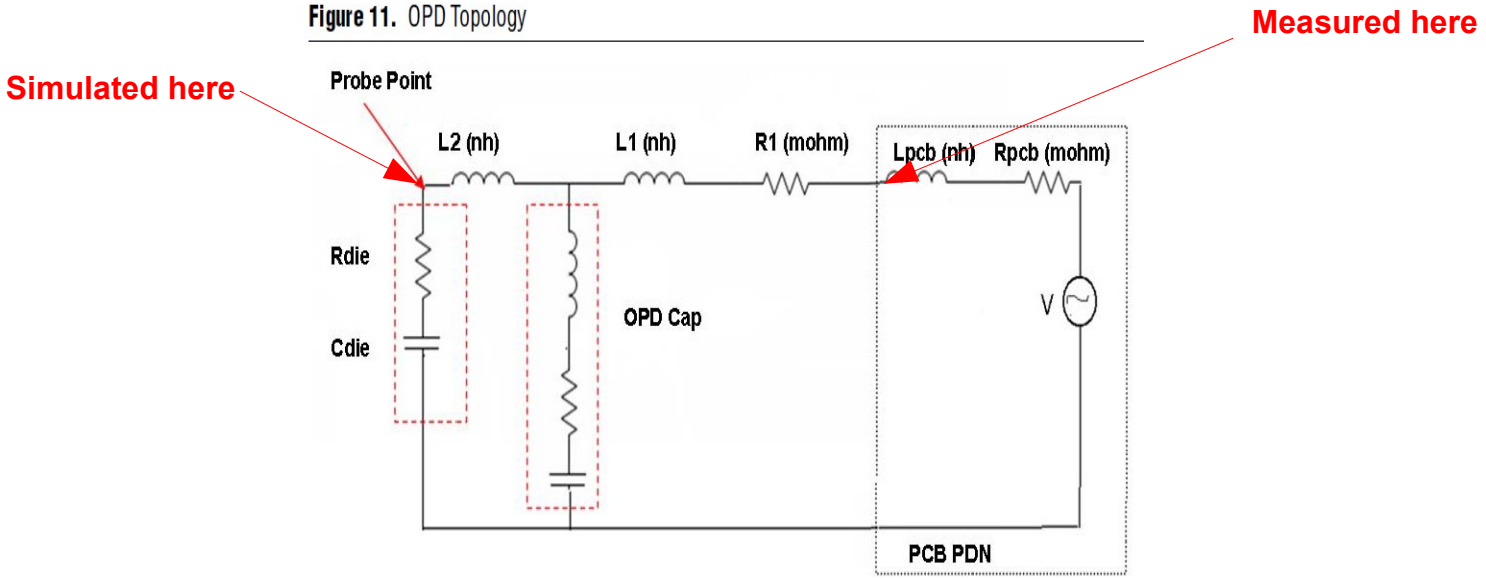
➡ VCC above threshold

➡ VCCR and VCCT OK

PDN simulation tools accuracy

PDN tool from Intel does not modelize accurately the leading inductance of capacitors

- Lots of approximations
- Use of external tools for determining values (lead inductances, etc ...)



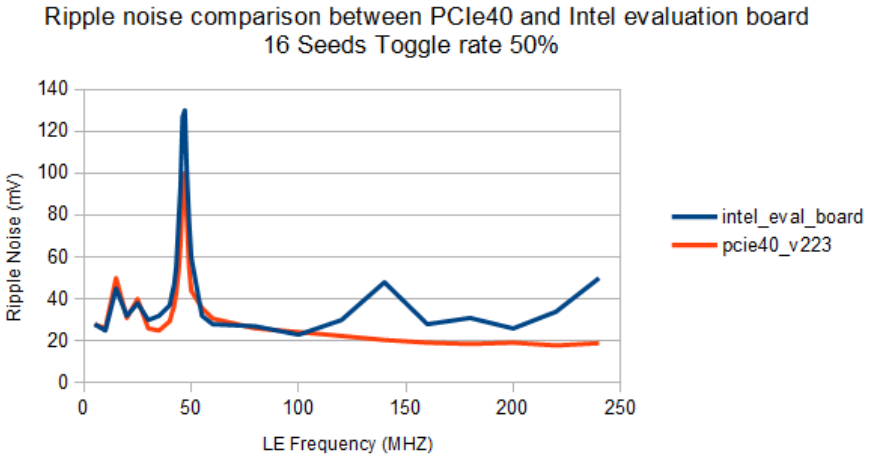
Sigrity or Ansoft do not modelize the FPGA on-package decoupling caps, nor the on-die caps

- Intel does not provide any model

Comparison with Intel SDK

Porting of a simple version of firmware on an Intel SDK (RPG only)

- Ripple noise measurements and comparison

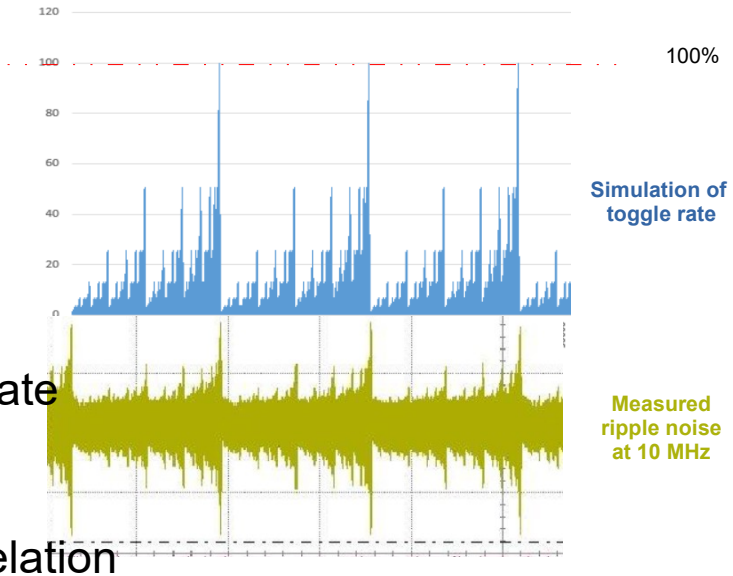


- Same peak around 50 MHz and even more ripple noise on high frequencies
  - Decoupling certainly not the reason

Noise bursts causes

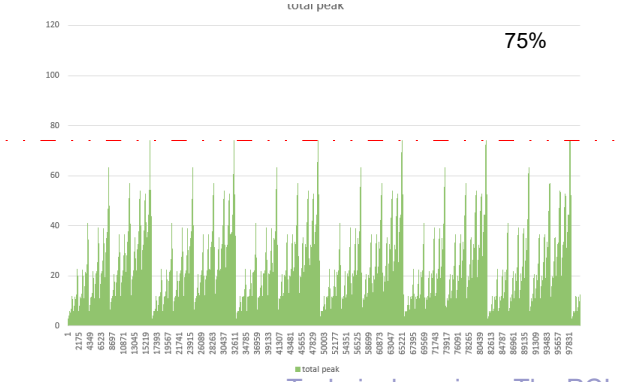
Correct firmware emulation ?

- Random pattern theoretically gives 50% of toggle rate in average ...
- … but locally can be much higher than this.
- Simulations of toggle rate showed an obvious correlation with observed noise
- Average toggle rate 12.5% with peaks at 100% !



Toggle rate simulation and ripple noise measurement comparison

- new tests made with 16 different seeds
  - Toggle rate amplitude decreases: 75%
  - Average toggle rate 20%
  - **Nearly no more errors at high speed**, but errors around 20 MHz remained

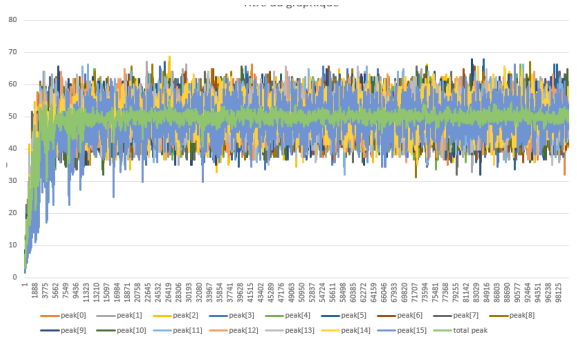


Frequency	Time	Average temperature	Average current	Errors found
10	100	44	5393	914342
20	100	42	5899	750598840
40	100	43	7139	0
80	100	46	9698	0
120	100	46	12251	0
160	100	48	14824	0
200	100	50	17384	0
240	100	52	19986	0
280	100	54	22584	0
320	100	57	25218	0
360	100	59	27861	0
400	100	61	30527	0
440	100	63	33210	0
480	100	65	35919	0
520	100	67	38652	0
560	100	70	41384	0
600	100	72	44139	992

BER at 89%
16 different seeds

Fixing the issue

Improvement of RPG by adding more feed backs



Simulated toggle rate (total contribution in green)

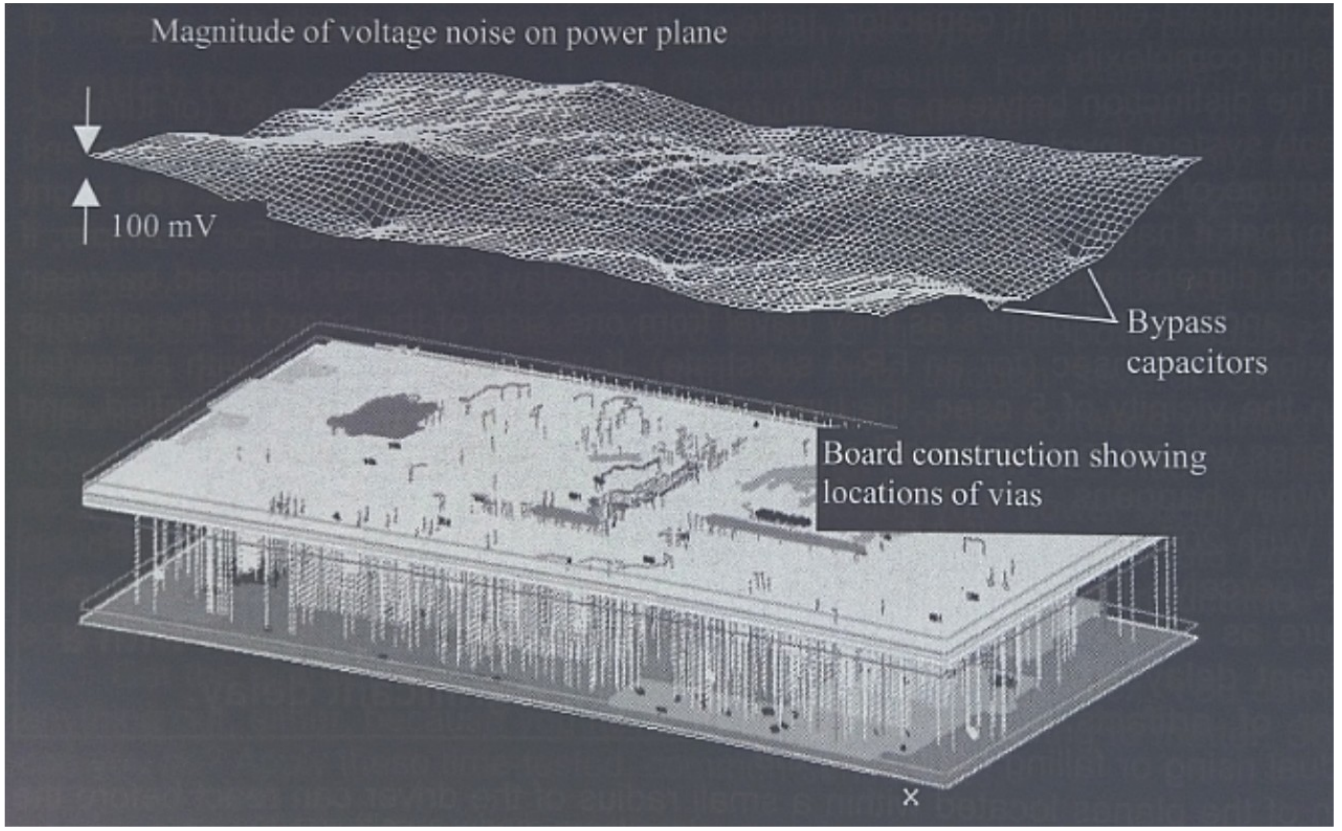
	Frequency	Time	Average temperature	Average current	Errors found
	5	100	41	5278	216
	10	100	43	5929	0
	15	100	44	6562	0
	20	100	45	7180	2462
	25	100	46	7701	18503123109715
	30	100	47	8419	37593657128
	35	100	48	9162	0
SOL40	40	100	49	9753	0
	80	100	54	15173	0
	120	100	53	20619	0
	160	100	58	26114	0
	200	100	62	31746	0
TELL40 CRU	240	100	67	37486	0
	280	100	72	43351	0
	320	100	78	49683	0

Limited measurement because of current drawn

- Real toggle rate of 50%
- More current drawn
  - Maximum reached at 320 MHz

Power plane resonance ?

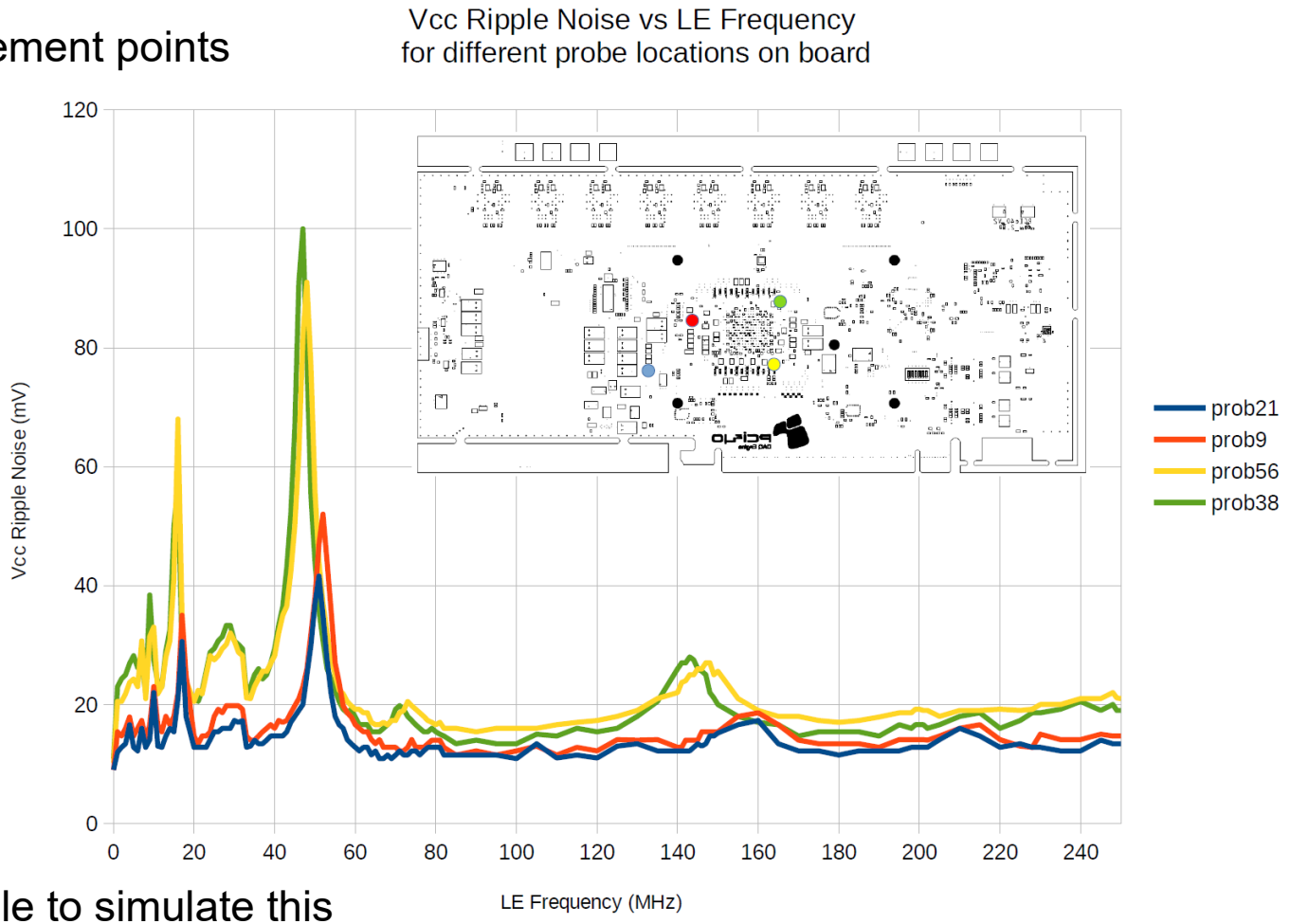
- Peaks at ~50 MHz, errors at ~25 MHz
- Resonance could happen if injected frequency is reflected on board edges



Power plane resonance ?

Geographical measurements to check this hypothesis

- 4 measurement points



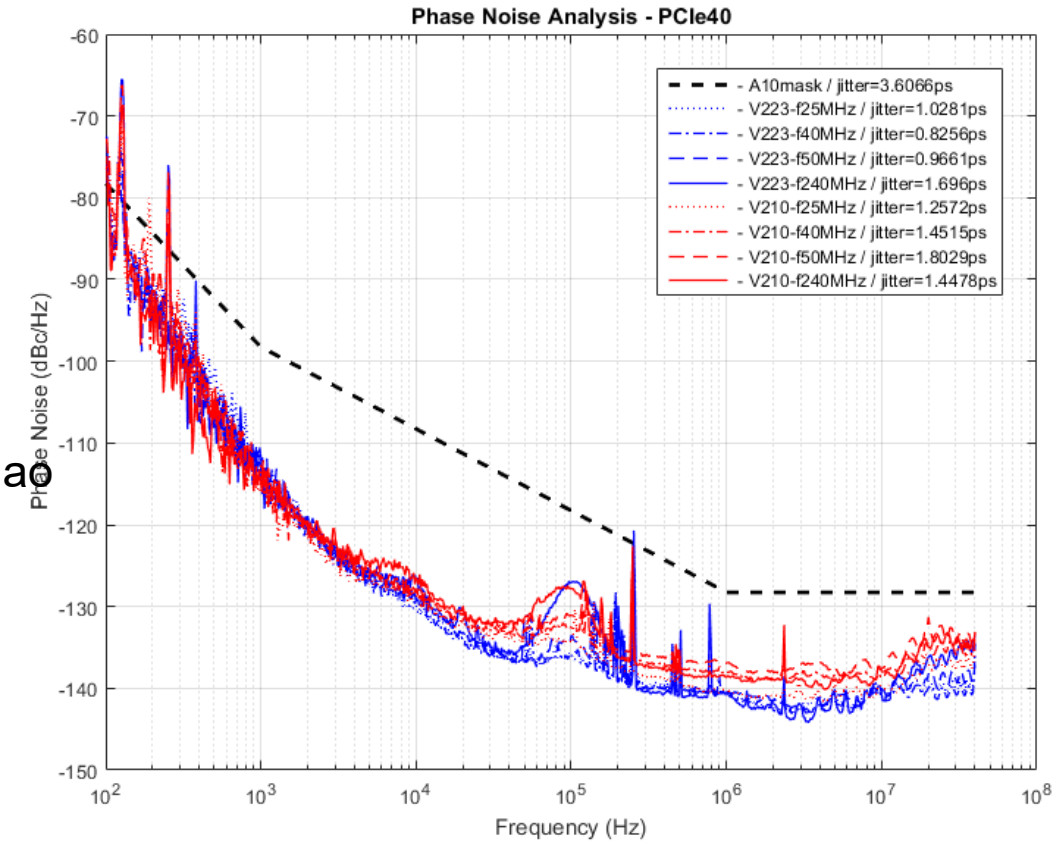
- No tool able to simulate this

PLL phase noise ?

Errors detected mainly on the emitting side: PLLs suspected

- *Phase noise on external PLLs could generate jitter on TX lines*

  - To check this we measured the phase noise of PLLs feeding the refclks
  - 4 frequencies injected:
    - 25 MHz
    - 50 MHz
    - 40 MHz
    - 240 MHz
  - Two cards tested:
    - V210
    - V223
  - Measured with Eduardo Brandao on an Agilent E5052B Signal Source Analyzer

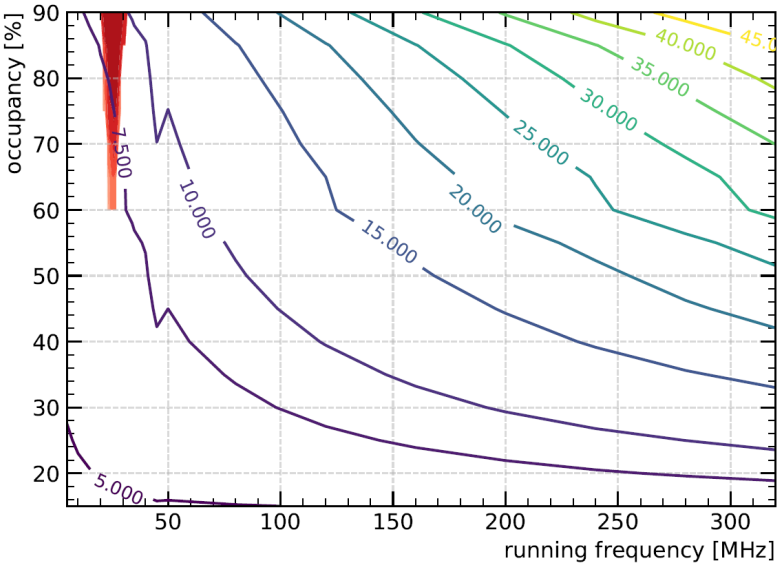


Results

- Phase noise within spec

Decision to go in production

No visible impact on frequencies used by LHCb and Alice

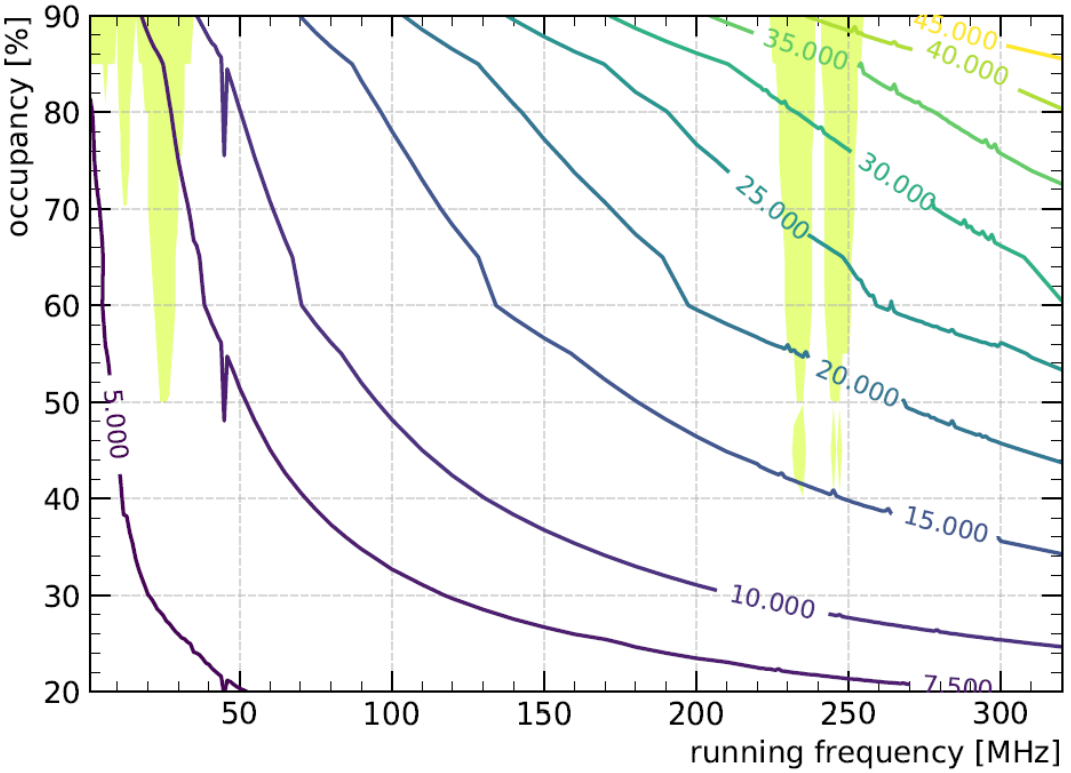


Possible alternatives in case we would work in the critical frequencies

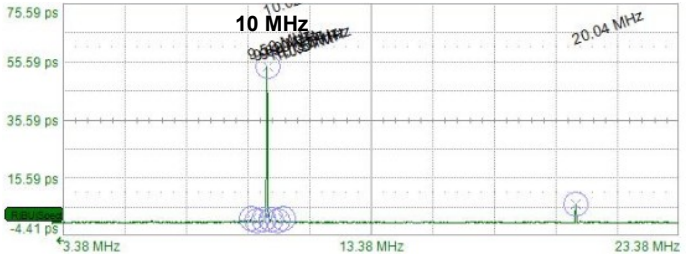
- Spreading the toggle rate over several shifted-phase clock domains
- Increase clock frequency

Refining the error space

Measurements with a finer granularity: new type of errors !



- Rightmost peaks clearly indentified as beats between the 240 MHz refclk and the core logic injected clock



Jitter noise spectrum when injecting a 230 MHz clock

FPGA internal PLLs testing

Change many PLL related parameters to see if errors disappear or increase

- Internal vs external feed back
- Bandwidth

Check other possible sources of beat

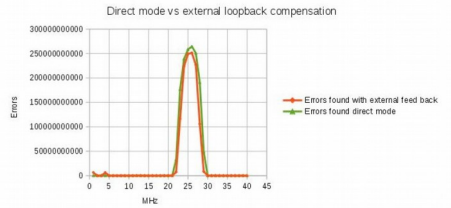
- PCIe DMA at 250 MHz

- *No significant change until ...*

Change external feed back by internal one in fPLL

Replaced loopback compensation mode by direct mode in fPLL

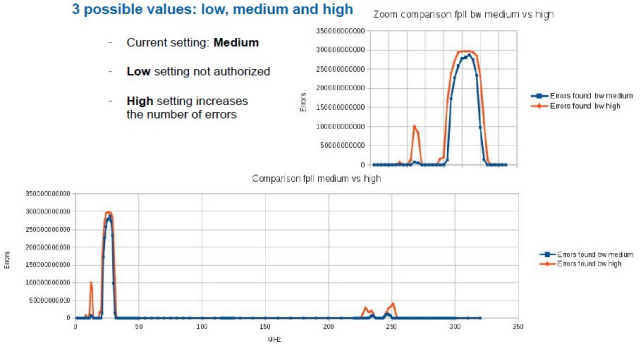
- External loopback circulates in the fabric
  - Could be subject to noise
- No visible effect.



FPLL loop bandwidth influence

3 possible values: low, medium and high

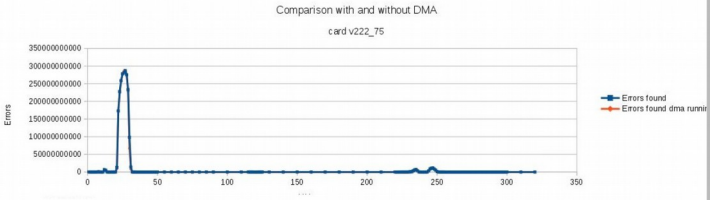
- Current setting: **Medium**
- **Low** setting not authorized
- **High** setting increases the number of errors



Run full DMA while running BER

Check that 250 MHz of PCIe does not interfere with 240 MHz of refclks

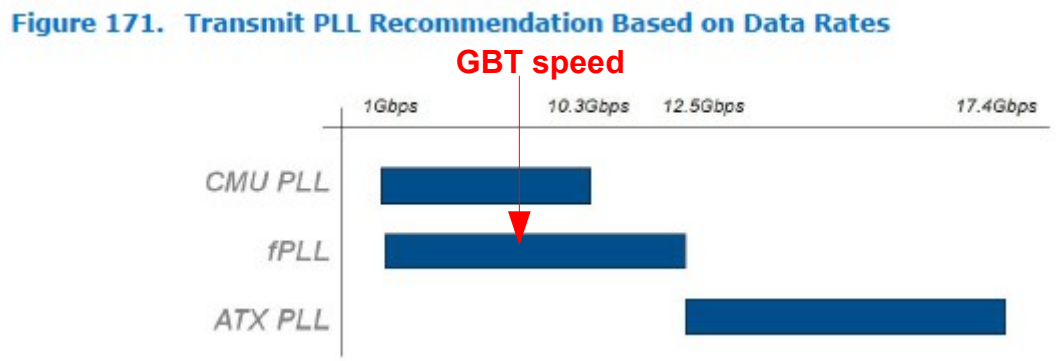
- Run BER when injecting all frequencies from 1 to 320 MHz
- Run DMA full speed
- Compare results with and without DMA
  - No difference



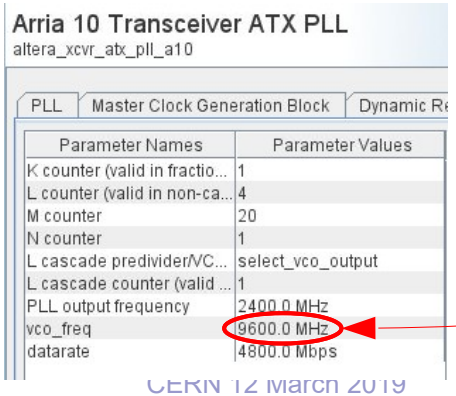
Replacement of fPLL by ATX PLLs

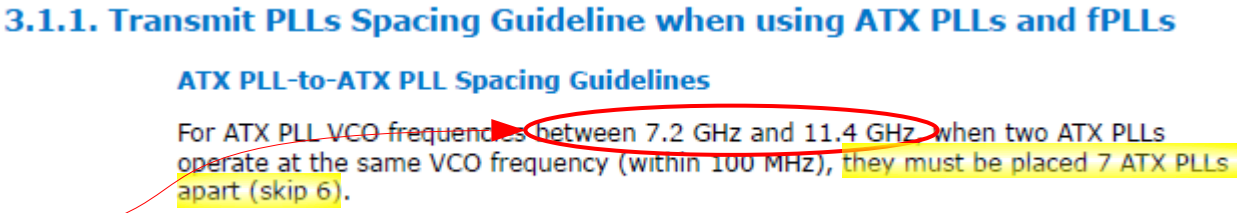
Although not recommended because not compatible with two rules given by Intel

- Use of PLL type:



- Spacing

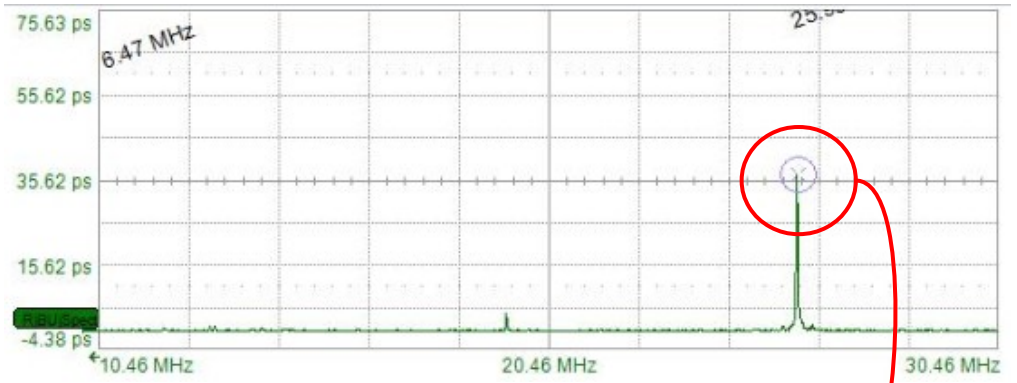


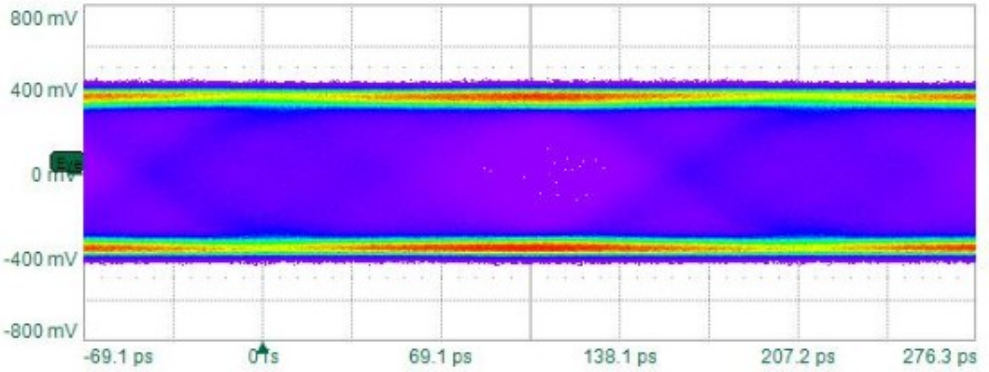


No more errors !

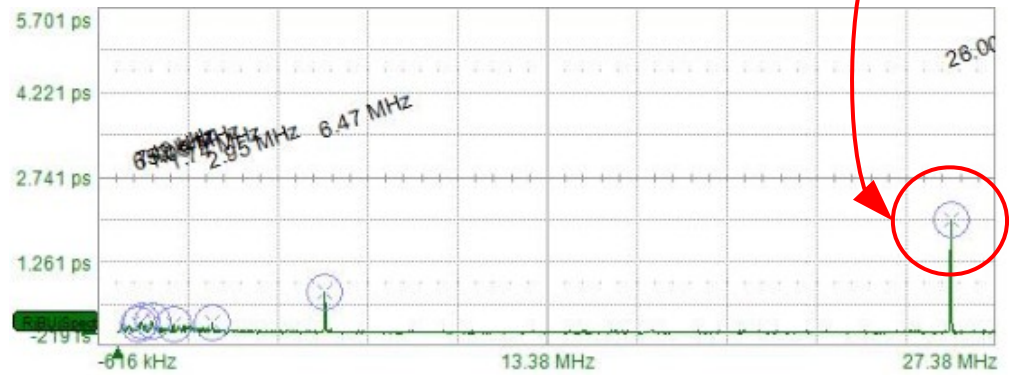
Worst case 1: 26 MHz injection

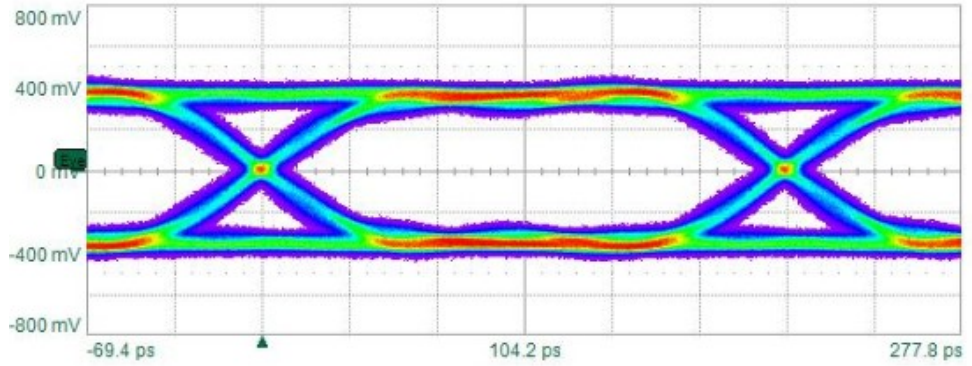
- **fPLL design**, jitter contribution of the 26 MHz core clock = **35 ps**





- **ATX PLL design**, jitter contribution of the 26 MHz core clock = **2.2 ps**

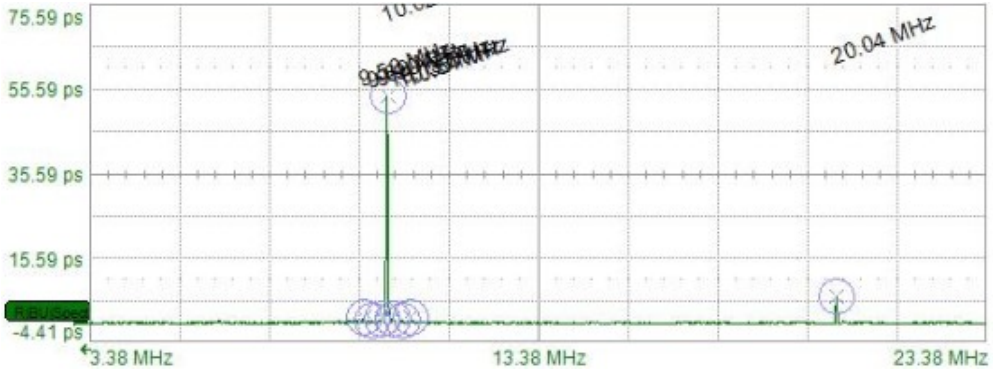


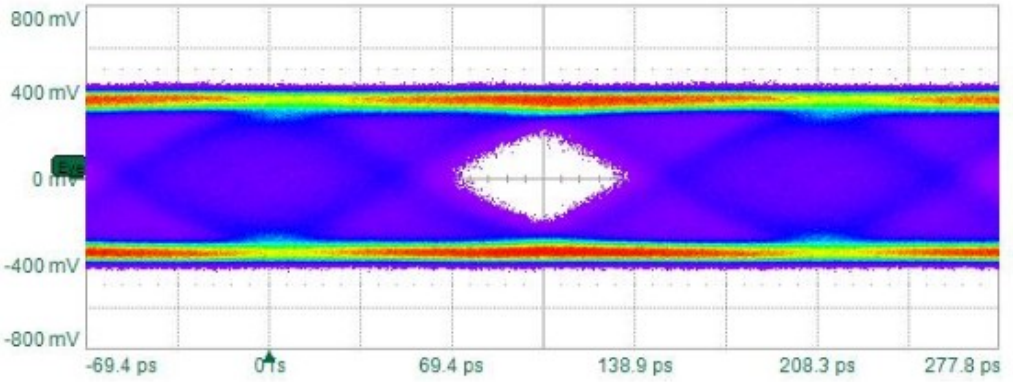


No more errors !

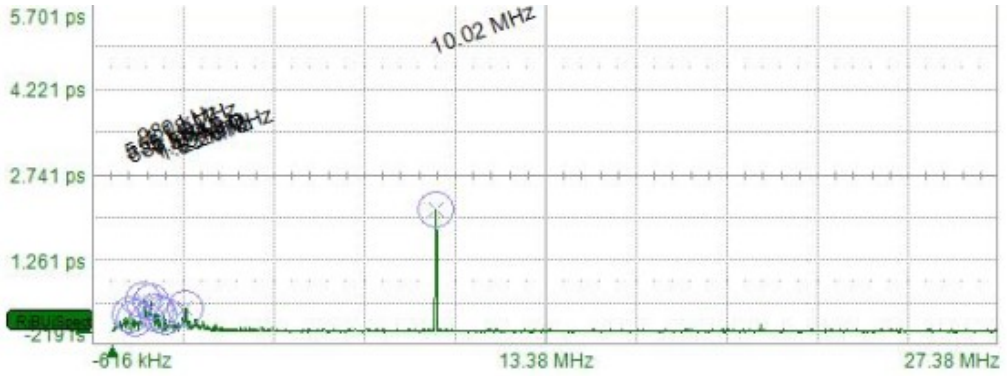
Worst case 2: 230 MHz injection (beats at 10 MHz)

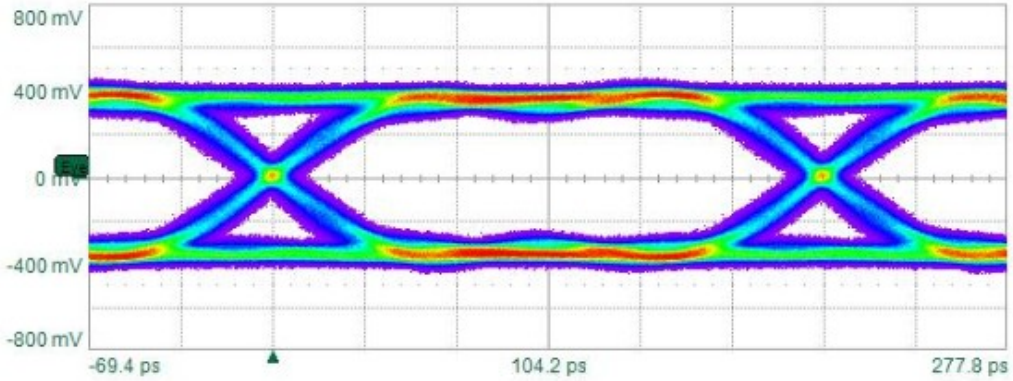
- **fPLL design**, jitter contribution of the 26 MHz core clock = **56 ps**





- **ATX PLL design**, jitter contribution of the 26 MHz core clock = **2.2 ps**





Why does it works with ATX PLLs ?

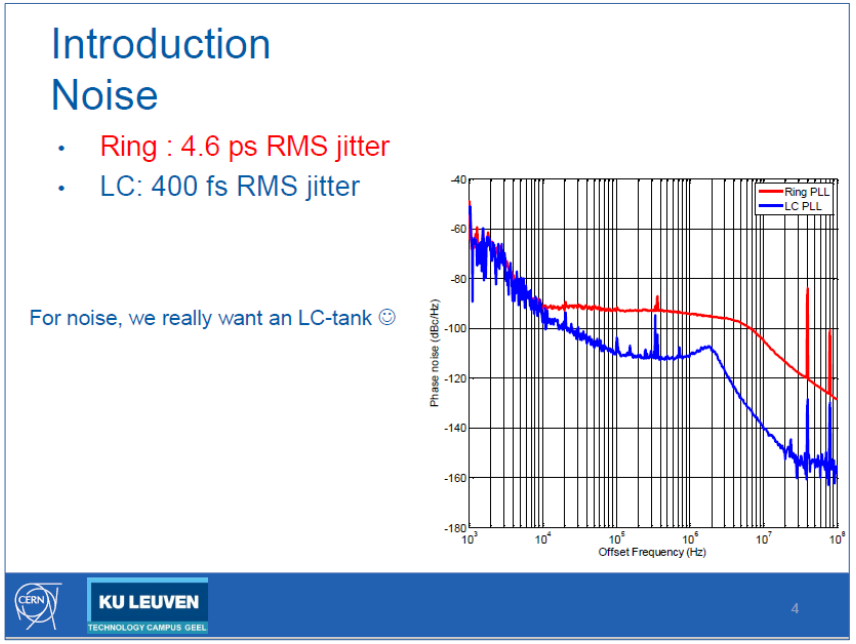
Not the same technology

- fPLLs are ring oscillator based VCO PLLs
- ATX PLLs are LC tank based

  - The second type is much more robust to noise

    see :

    - 2.56-GHz SEU Radiation Hard LC-Tank VCO for High-Speed Communication Links in 65-nm CMOS Technology
      (Jeffrey Prinzie , Student Member, IEEE,
      Jorgen Christiansen, Paulo Moreira,
      Michiel Steyaert, Fellow, IEEE,
      and Paul Leroux, Senior Member, IEE)
    - Phase Noise and Jitter in CMOS Ring Oscillators
      (Asad A. Abidi, Fellow, IEEE)

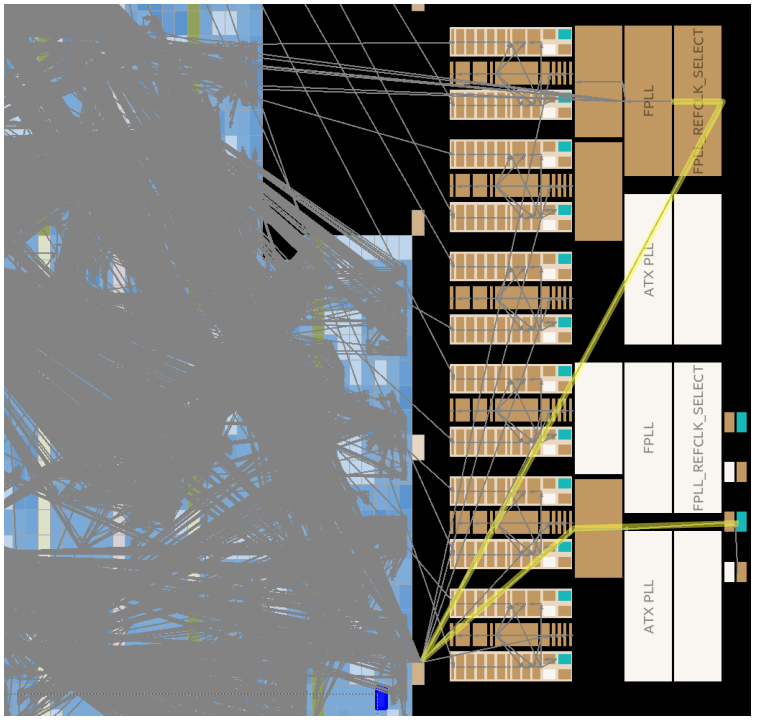


Ring oscillator vs LC tank PLLs noise comparison
Jeffrey Prinzie (CERN)

Further investigation with Intel

Finally find out to be a firmware issue !

- Refclk went into the matrix before going on fPLL !



- Due to residual constraint placed for allowing first versions of Quartus to converge (formerly crashed with a clock tree congestion message)
- Same topology when ATX PLL is instantiated but ATX PLL more robust

fPLL design results after modification

Frequency In MHz	Total jitter In ps	Random RMS	Deterministic p-p	Core current In A
10	37.23	2.167	6.762	5.953
20	37.23	2.169	6.728	7.209
25	37.47	2.170	6.963	7.710
40	37.81	2.170	6.728	9.735
80	37.72	2.170	7.200	15.072
120	37.62	2.170	7.109	20.367
160	37.59	2.169	7.076	25.814
200	37.55	2.170	7.031	31.206
240	37.35	2.168	6.865	36.499
280	37.45	2.169	6.938	41.956
320	37.37	2.168	6.877	47.462

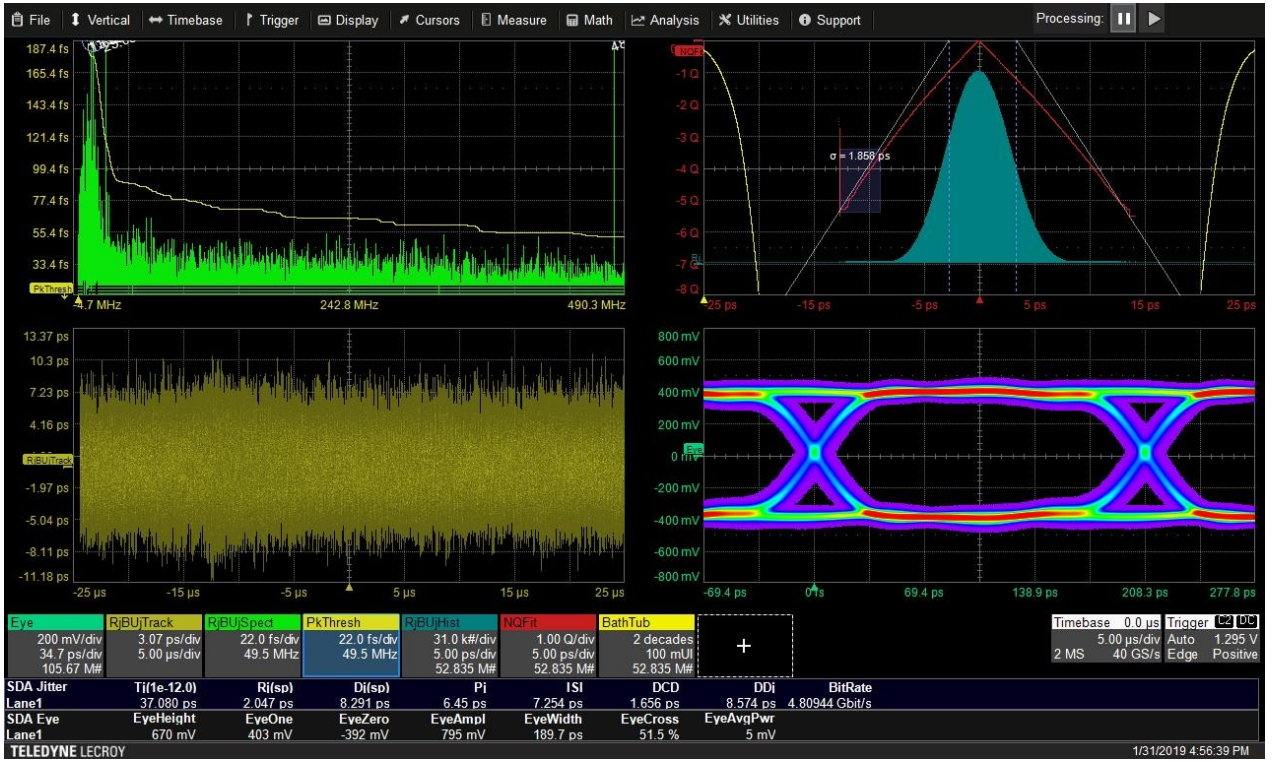
Worst case: 90% occupancy, 25 MHz core clock

- Very stable over the full spectrum:
  - Total jitter 37 ps
  - Random: 2 ps RMS
  - Deterministic: 8.2 ps p-p

Test run during 2 days

- Over 48 links
- On 16 cards
- 300 MHz injected clock (45 A on the core)
- 10 m optical loopback

➡ **No errors**

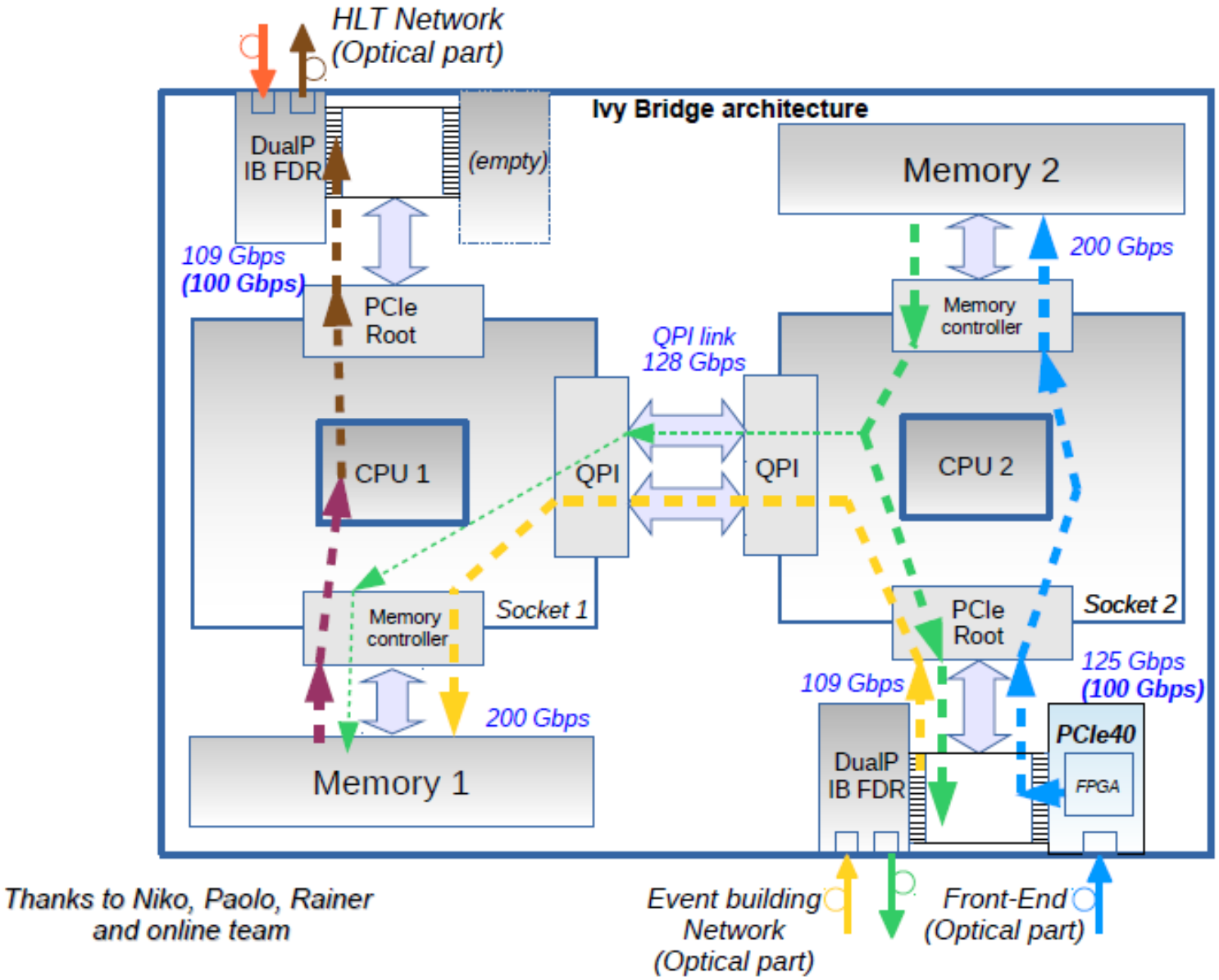


Conclusion

- Cards adressing many needs in our community
  - Large acquisition capability, high processing power
  - Powerful interface between dedicated Front-Ends and commercial computer CPUs
  - Flexible enough to be used in many ways (3 functions in LHCb: DAQ, ECS and TFC, can fit ALICE needs as well, also selected for the readout of the µ3E experiment)
  - Lots of effort spent for optimizing the card for production (automatic testing, long time acceptance testing, automatic recording)
- Importance of testing a card to its limits
  - Most of debugs are led with minimum firmwares
  - High currents in high ends FPGAs raise new problems
- The PCIe40 card has been exhaustively tested
- Many lessons learned
  - Better understanding of power plane geometry effects
  - Better understanding of decoupling
  - Limits of simulation tools
- Cards are now in production
  - 700 cards for LHCb, 550 cards for ALICE
  - Available end of this year

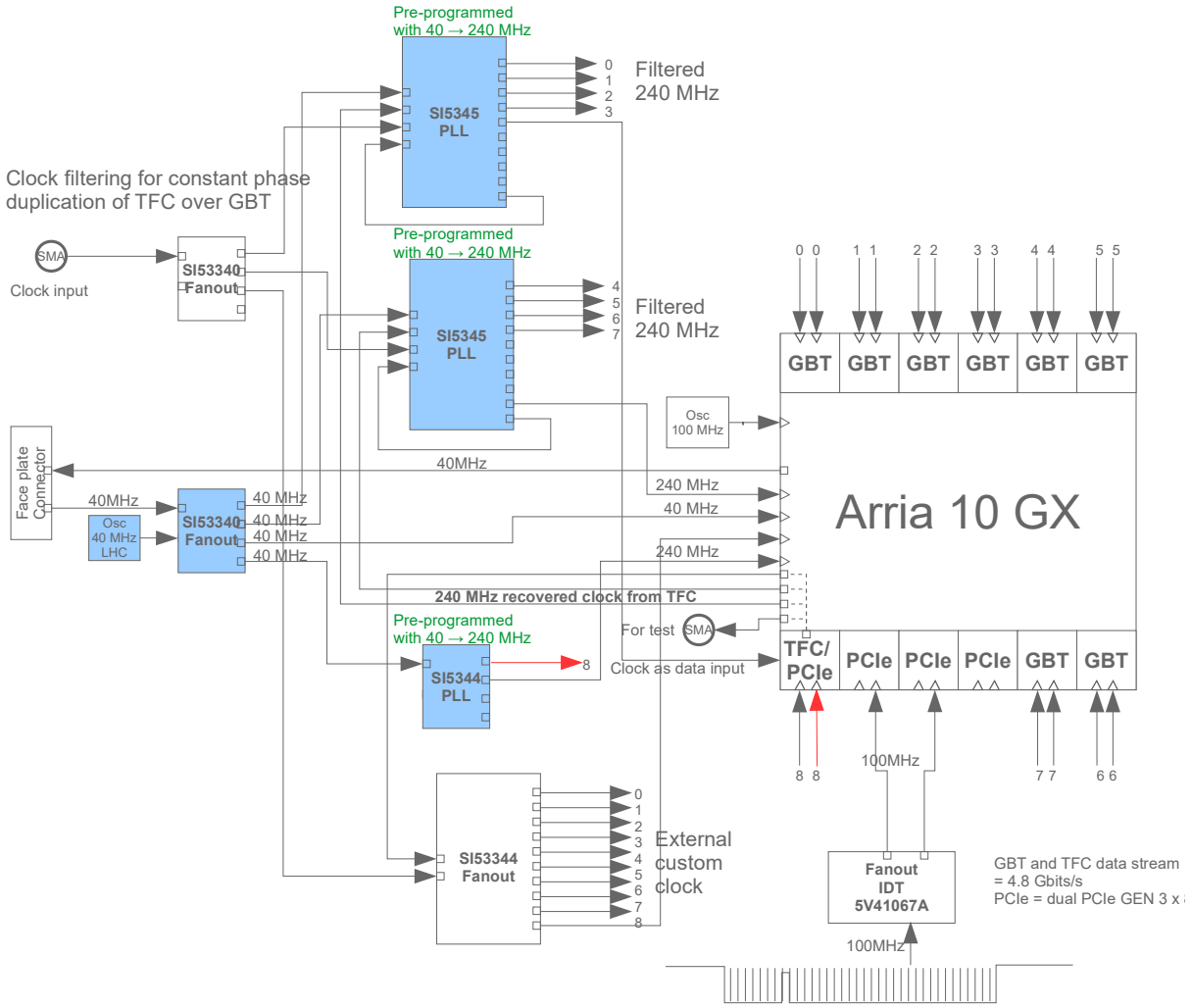
More information

Data path in the computer

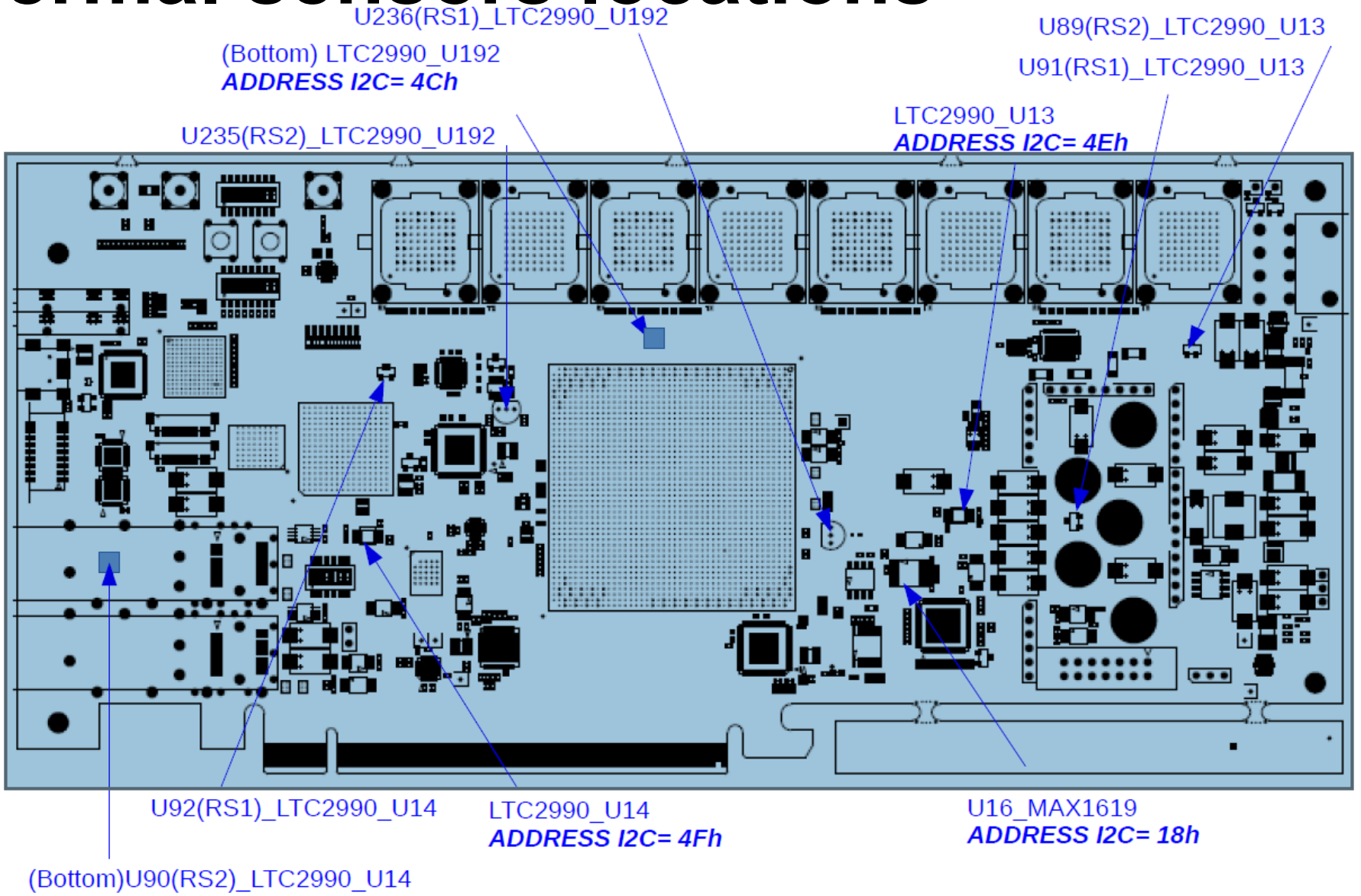


Clock distribution

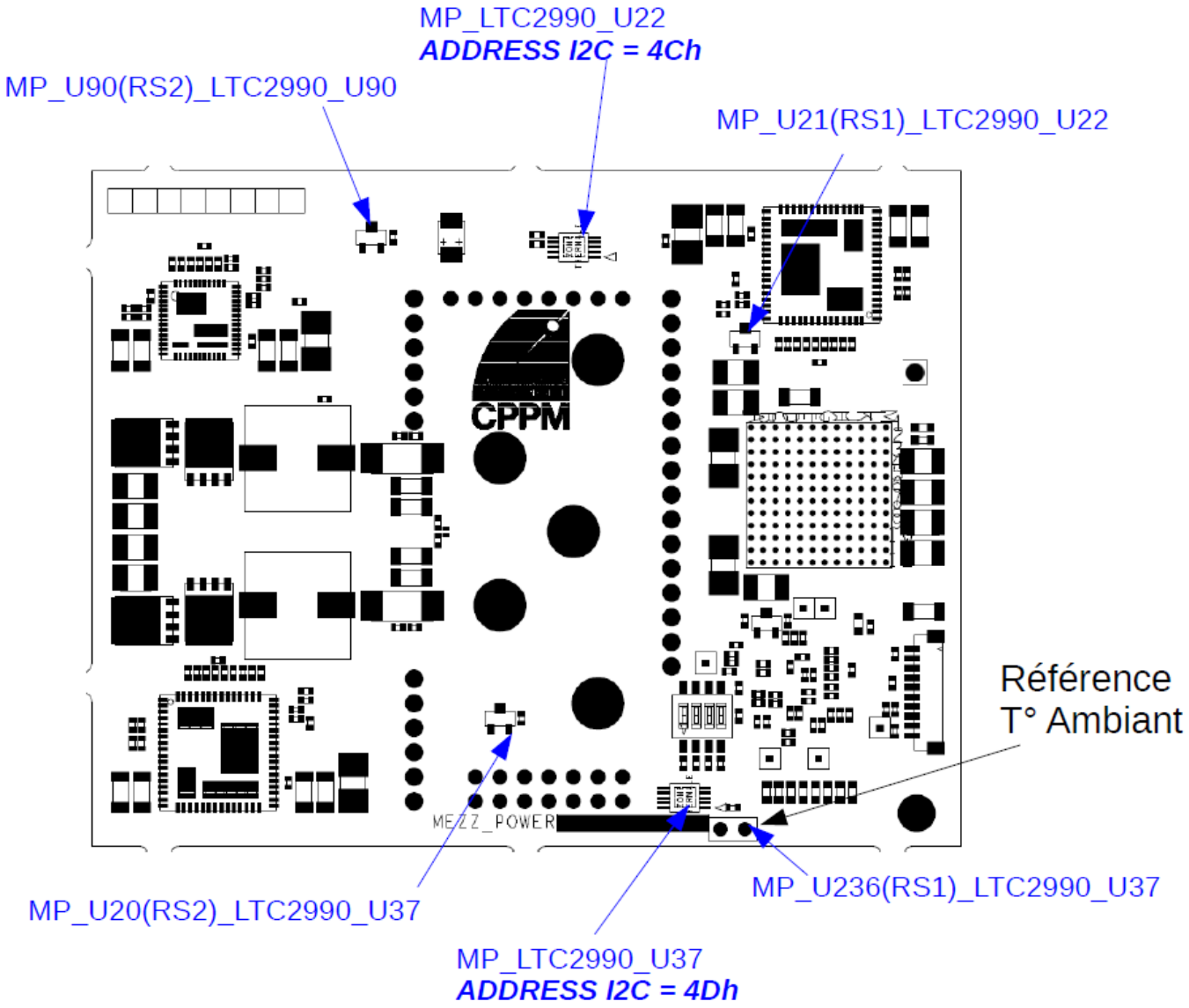
Clock Tree PCIe40V2



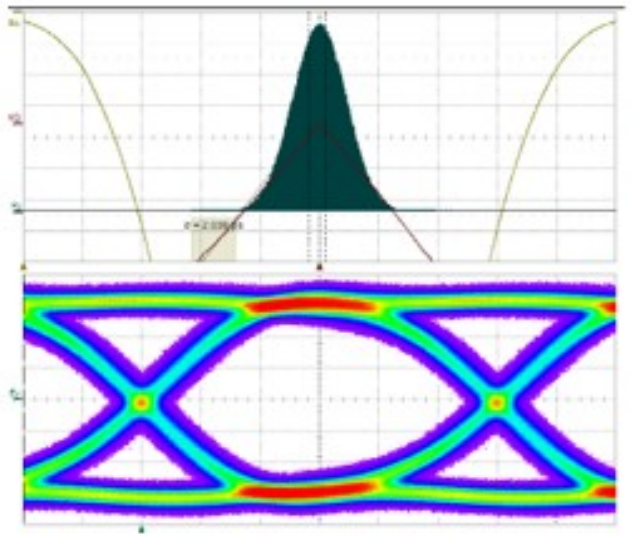
Thermal sensors locations

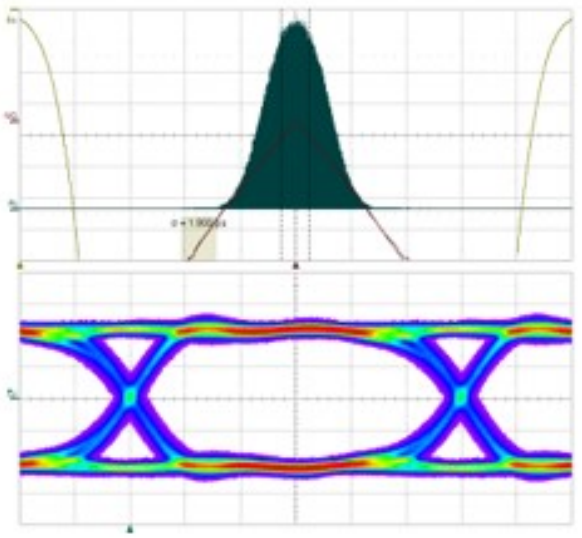


Thermal sensors locations



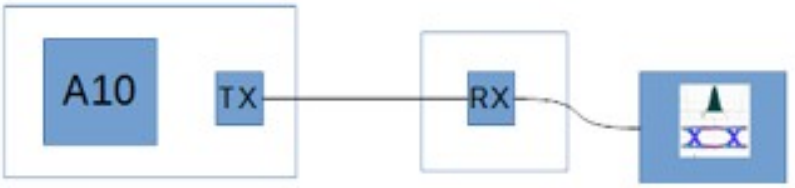
Eye diagrams



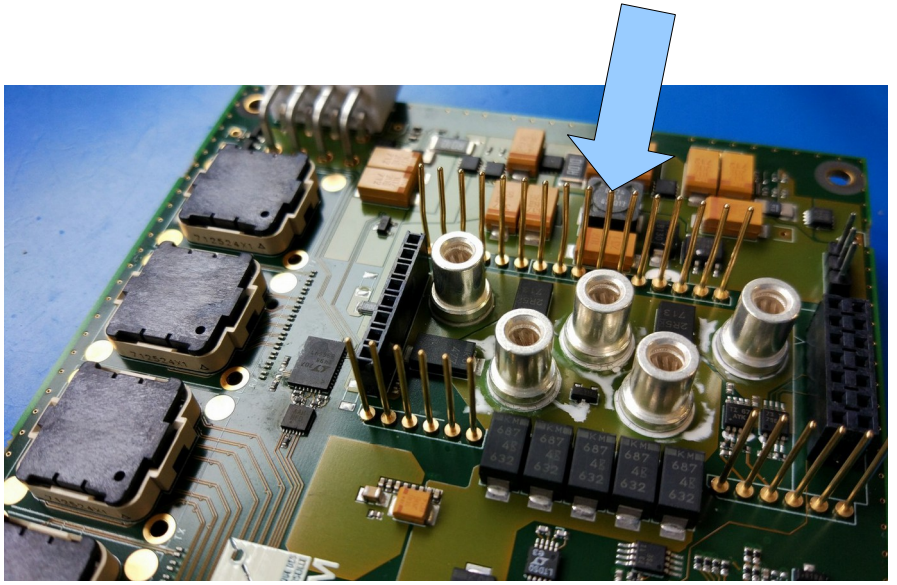


Measurements at 10.0Gbit/s
Total jitter ~ 36.82ps
Random Jitter ~ 2.22ps
Deterministic Jitter ~ 5.6ps

Measurements at 5.0Gbit/s
Total jitter ~ 37ps
Random Jitter ~ 2.1 ps
Deterministic Jitter ~ 3.44 ps



Mezzanine connector

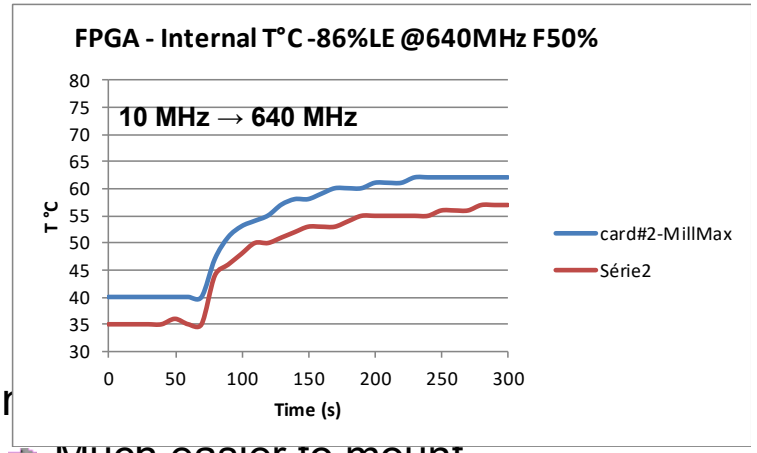


Two choices : Samtec or Millmax

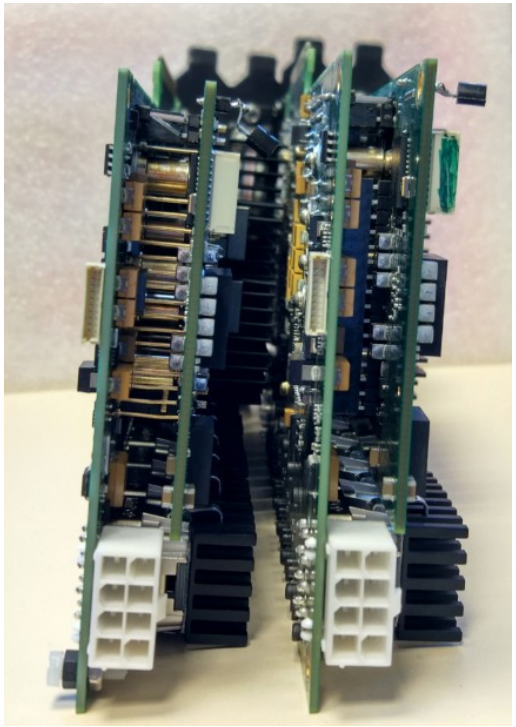
- Samtec : classical « full » connectors
- Millmax « transparent » connectors to let the air flow under the mezzanine

Cooling tests made with both solutions

- Counter intuitive results : Millmax card hotter than Samtec one (~5 to 6°C)
  - Venturi effect ?

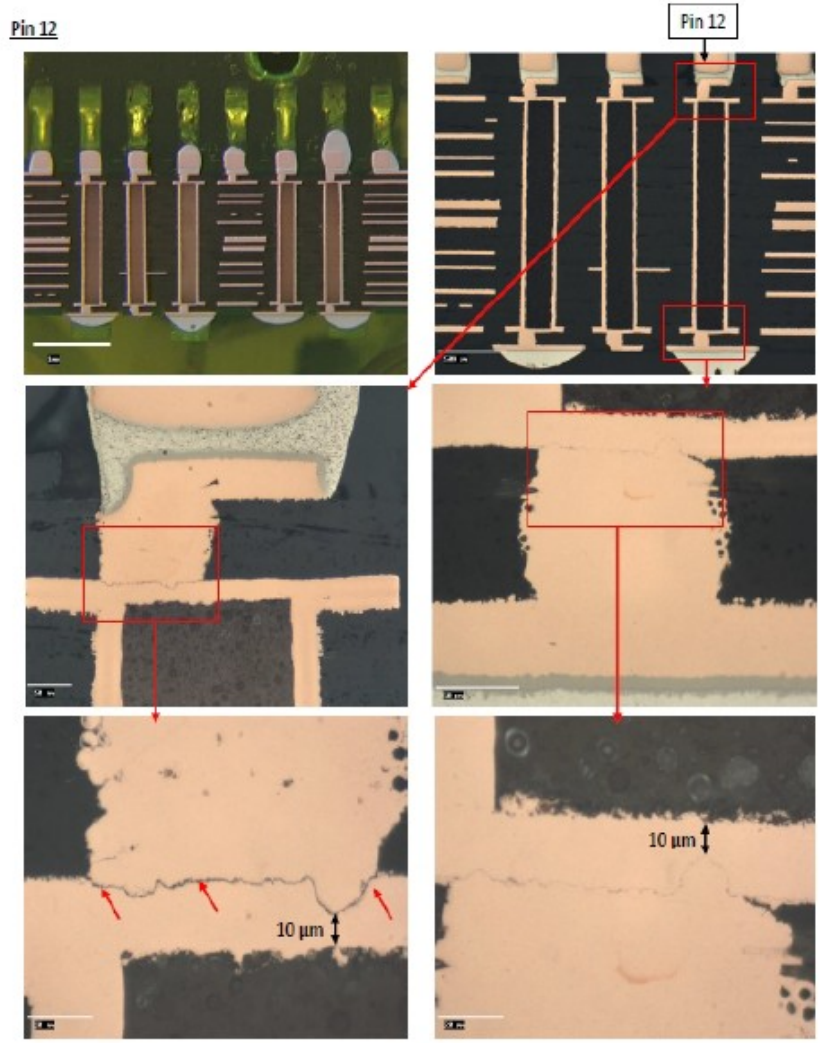


- Fir
  - Much easier to mount



The PCB episode

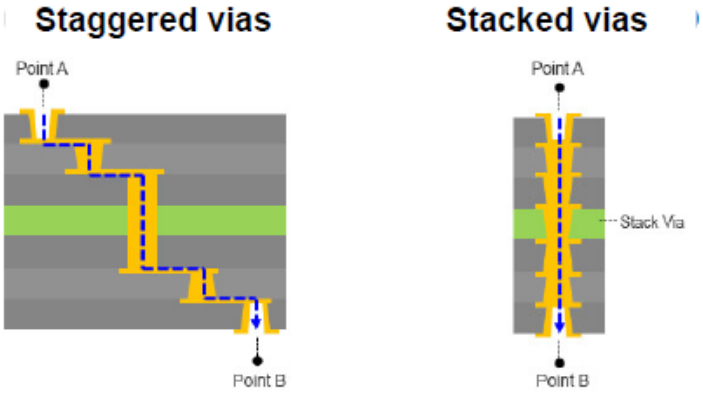
- First batch of 6 MiniDAQ2 almost failed. Three boards survived but would die soon.
- After a long investigation, the issue was localized on the PCB. It was due to micro-cracks in the so-called stacked vias.
- A new board with a PCB from a different manufacturer was delivered Feb 15, 2017.
- After an extensive campaign of tests we concluded that the board is fully functional.



Routing

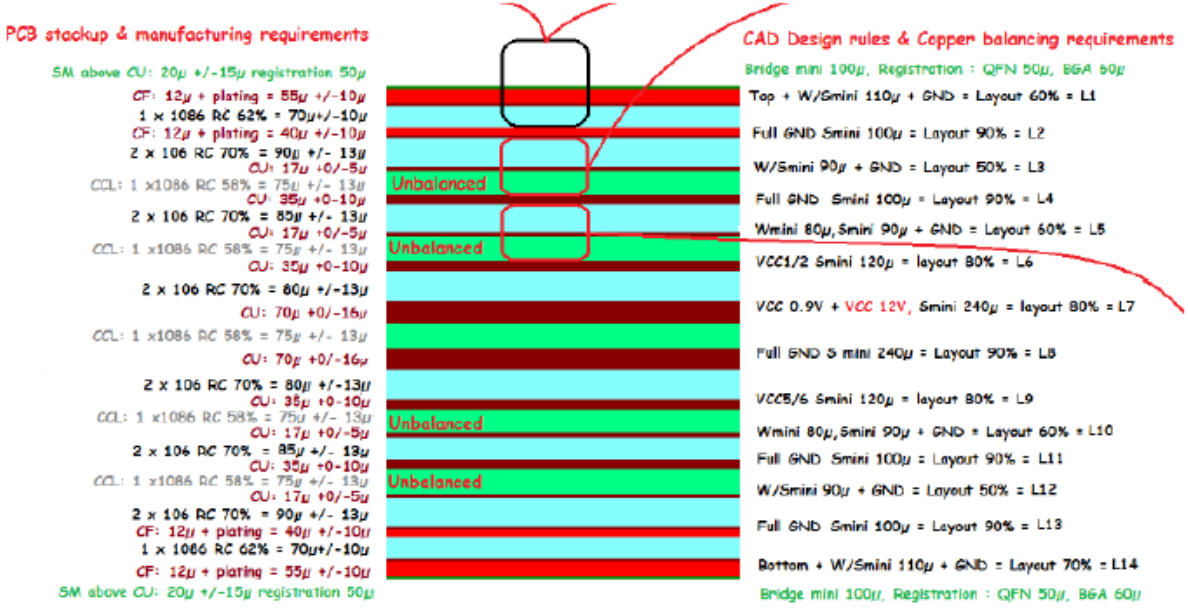
Use of staggered vias instead of stacked vias

- Slight degradation of signal integrity
- But more subcontractors able to manufacture the card



Stackup

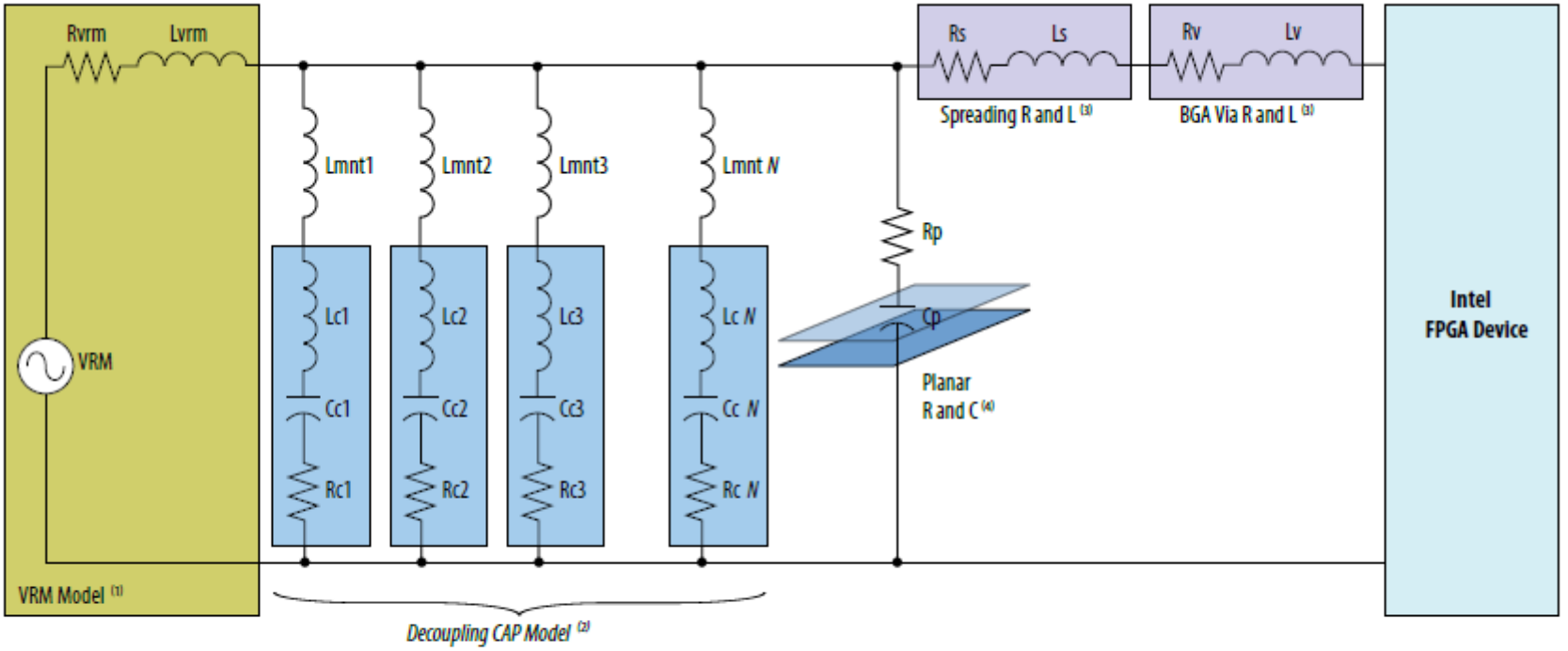
- 14 layers
- 70µ thick planes for power
- HR408 high speed PCB
- More than 10000 vias among which 67% are microvias
- ~ 1750 components



Decoupling

Principle

The PDN impedance profile is the impedance-over-frequency looking outward from the device.



Notes:

1. You can define or change VRM parameters in the Library sheet of the PDN tool.
2. You can define or change Decoupling Caps parameters in the Cap Mount, X2Y Mount, and Library sheets of the PDN tool.
3. R* and L* are parasitic resistances and inductances from BGA balls and PCB traces and connections.
4. Represents PCB layers dedicated to power and ground planes.

$F_{EFFECTIVE}$ in PCB Decoupling

The PCB PDN cutoff frequency ($F_{EFFECTIVE}$) calculated by the PDN tool depends on the design trade-offs made on the PCB. The role of $F_{EFFECTIVE}$ is analyzed for both OPD and non-OPD packages.

Non-OPD Scenario

Figure 9 shows a simple topology for a rail without on-package decoupling.

Figure 9. Non-OPD Topology

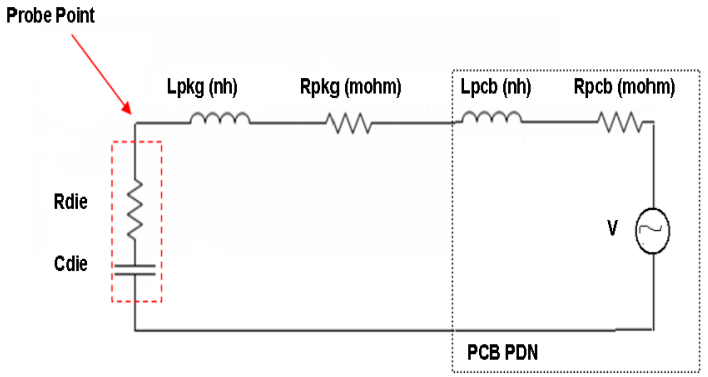
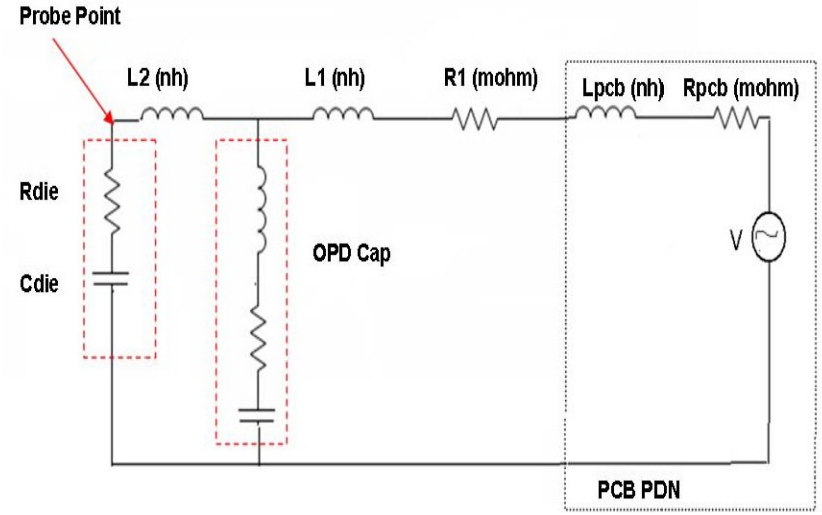


Figure 11. OPD Topology



$$F = 1/2\pi\sqrt{(Lpkg + Lpcb) * Cdie}$$

Figure 10. Non-OPD Topology Frequency Response

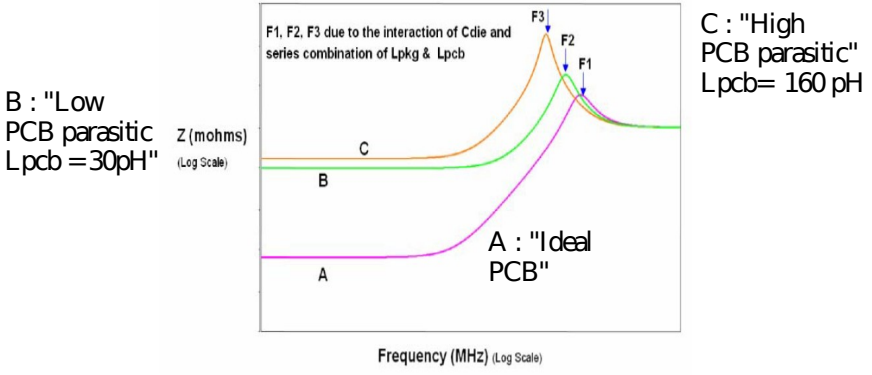


Figure 12. OPD Frequency Response

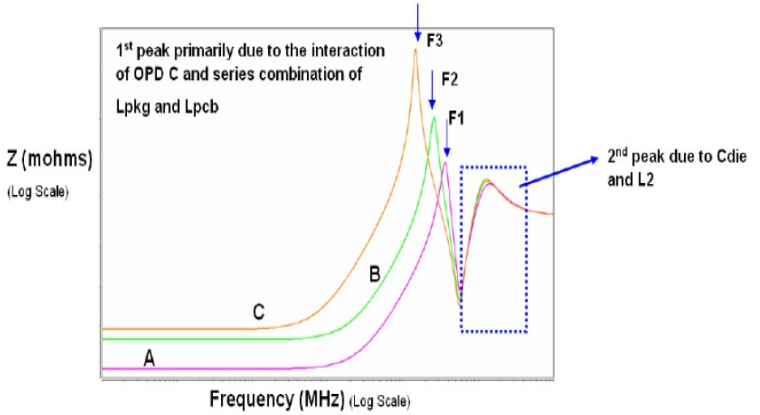
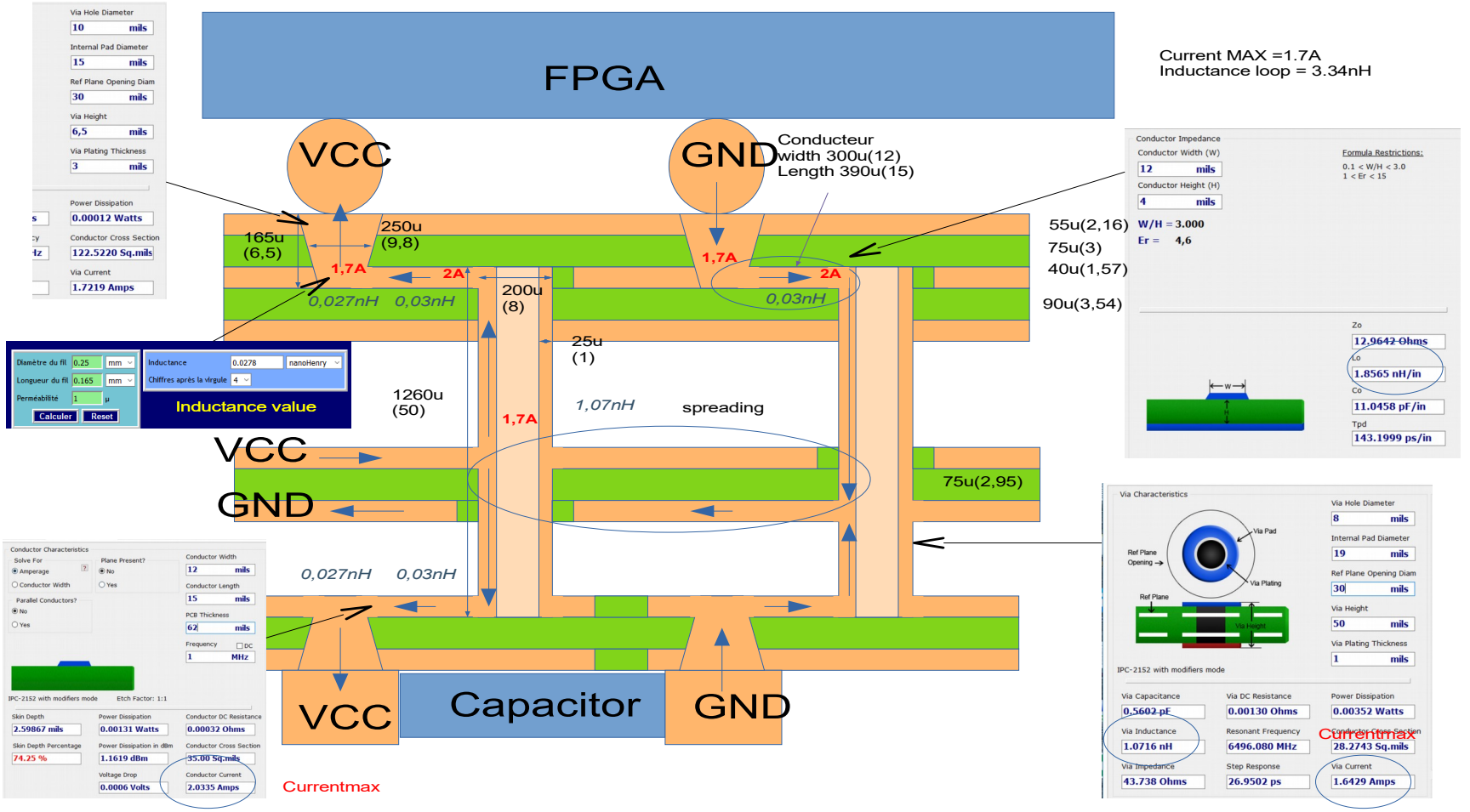


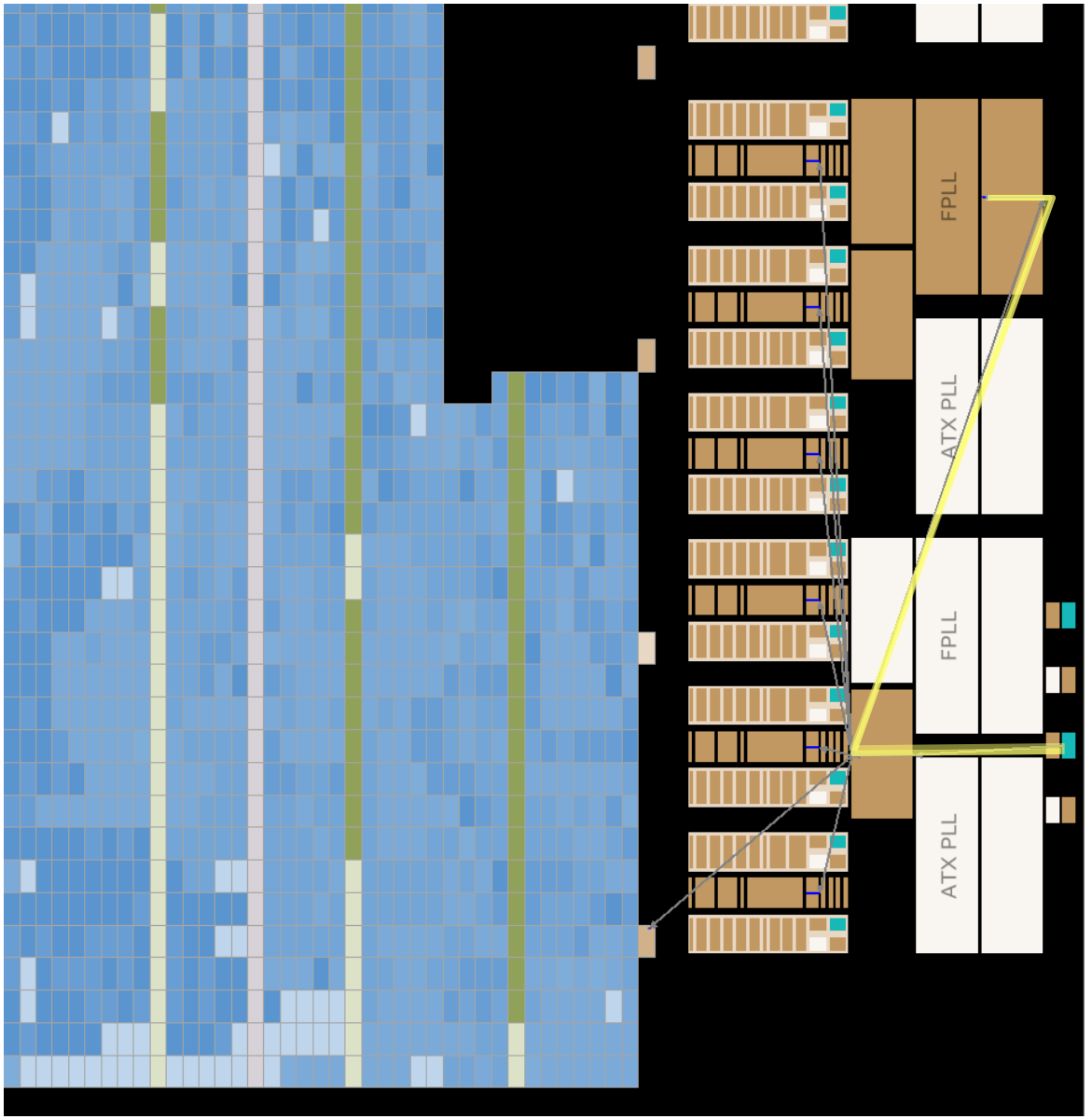
Figure 12 shows the simulated waveforms for three scenarios:

- Z-profile for "Ideal PCB"—The purple waveform (A) with resonance frequency F1
- Z-profile for "Low PCB parasitics"—The green waveform (B) with resonance frequency F2
- Z-profile for "High PCB parasitics"—The orange waveform (C) with resonance frequency F3

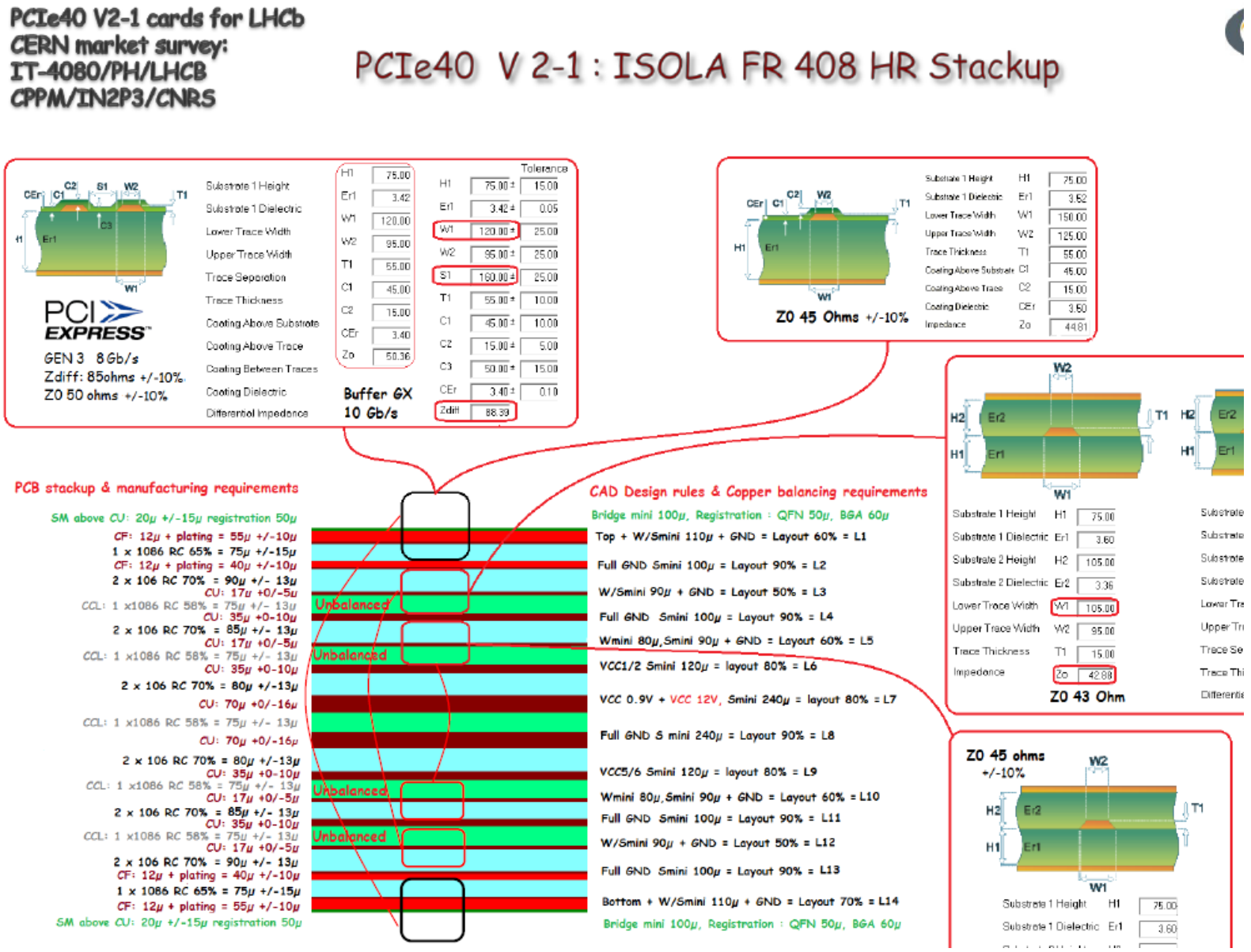
PDN parameters estimation



Correct clock routing



Stackup


PCIe40 V2-1 cards for LHCb
CERN market survey:
IT-4080/PH/LHCB
CPPM/IN2P3/CNRS
PCIe40 V 2-1 : ISOLA FR 408 HR Stackup
PCI EXPRESS
GEN 3 8 Gb/s
Zdiff: 85ohms +/-10%
Z0 50 ohms +/-10%
Substrate 1 Height
Substrate 1 Dielectric
Lower Trace Width
Upper Trace Width
Trace Separation
Trace Thickness
Coating Above Substrate
Coating Above Trace
Coating Between Traces
Coating Dielectric
Differential Impedance
Buffer GX
10 Gb/s
Tolerance
Z0 45 Ohms +/-10%
PCB stackup & manufacturing requirements
SM above CU: 20µ +/-15µ registration 50µ
CF: 12µ + plating = 55µ +/-10µ
1 x 1086 RC 65% = 75µ +/-15µ
CF: 12µ + plating = 40µ +/-10µ
2 x 106 RC 70% = 90µ +/- 13µ
CU: 17µ +0/-5µ
CCL: 1 x1086 RC 58% = 75µ +/- 13µ
CU: 35µ +0-10µ
2 x 106 RC 70% = 85µ +/- 13µ
CU: 17µ +0/-5µ
CCL: 1 x1086 RC 58% = 75µ +/- 13µ
CU: 35µ +0-10µ
2 x 106 RC 70% = 80µ +/-13µ
CU: 70µ +0/-16µ
CCL: 1 x1086 RC 58% = 75µ +/- 13µ
CU: 70µ +0/-16µ
2 x 106 RC 70% = 80µ +/-13µ
CU: 35µ +0-10µ
CCL: 1 x1086 RC 58% = 75µ +/- 13µ
CU: 17µ +0/-5µ
2 x 106 RC 70% = 85µ +/- 13µ
CU: 35µ +0-10µ
CCL: 1 x1086 RC 58% = 75µ +/- 13µ
CU: 17µ +0/-5µ
2 x 106 RC 70% = 90µ +/- 13µ
CF: 12µ + plating = 40µ +/-10µ
1 x 1086 RC 65% = 75µ +/-15µ
CF: 12µ + plating = 55µ +/-10µ
SM above CU: 20µ +/-15µ registration 50µ
Unbalanced
Unbalanced
Unbalanced
Unbalanced
CAD Design rules & Copper balancing requirements
Bridge mini 100µ, Registration : QFN 50µ, BGA 60µ
Top + W/Smini 110µ + GND = Layout 60% = L1
Full GND Smini 100µ = Layout 90% = L2
W/Smini 90µ + GND = Layout 50% = L3
Full GND Smini 100µ = Layout 90% = L4
Wmini 80µ,Smini 90µ + GND = Layout 60% = L5
VCC1/2 Smini 120µ = layout 80% = L6
VCC 0.9V + VCC 12V, Smini 240µ = layout 80% = L7
Full GND S mini 240µ = Layout 90% = L8
VCC5/6 Smini 120µ = layout 80% = L9
Wmini 80µ,Smini 90µ + GND = Layout 60% = L10
Full GND Smini 100µ = Layout 90% = L11
W/Smini 90µ + GND = Layout 50% = L12
Full GND Smini 100µ = Layout 90% = L13
Bottom + W/Smini 110µ + GND = Layout 70% = L14
Bridge mini 100µ, Registration : QFN 50µ, BGA 60µ
Z0 43 Ohm
Z0 45 ohms
+/-10%

3D model