Andy Rose, Imperial College, London

Imperial College London

cea SACLAY

INFN Istituto Nazionale Fisica Nucleare Sezione di Pisa

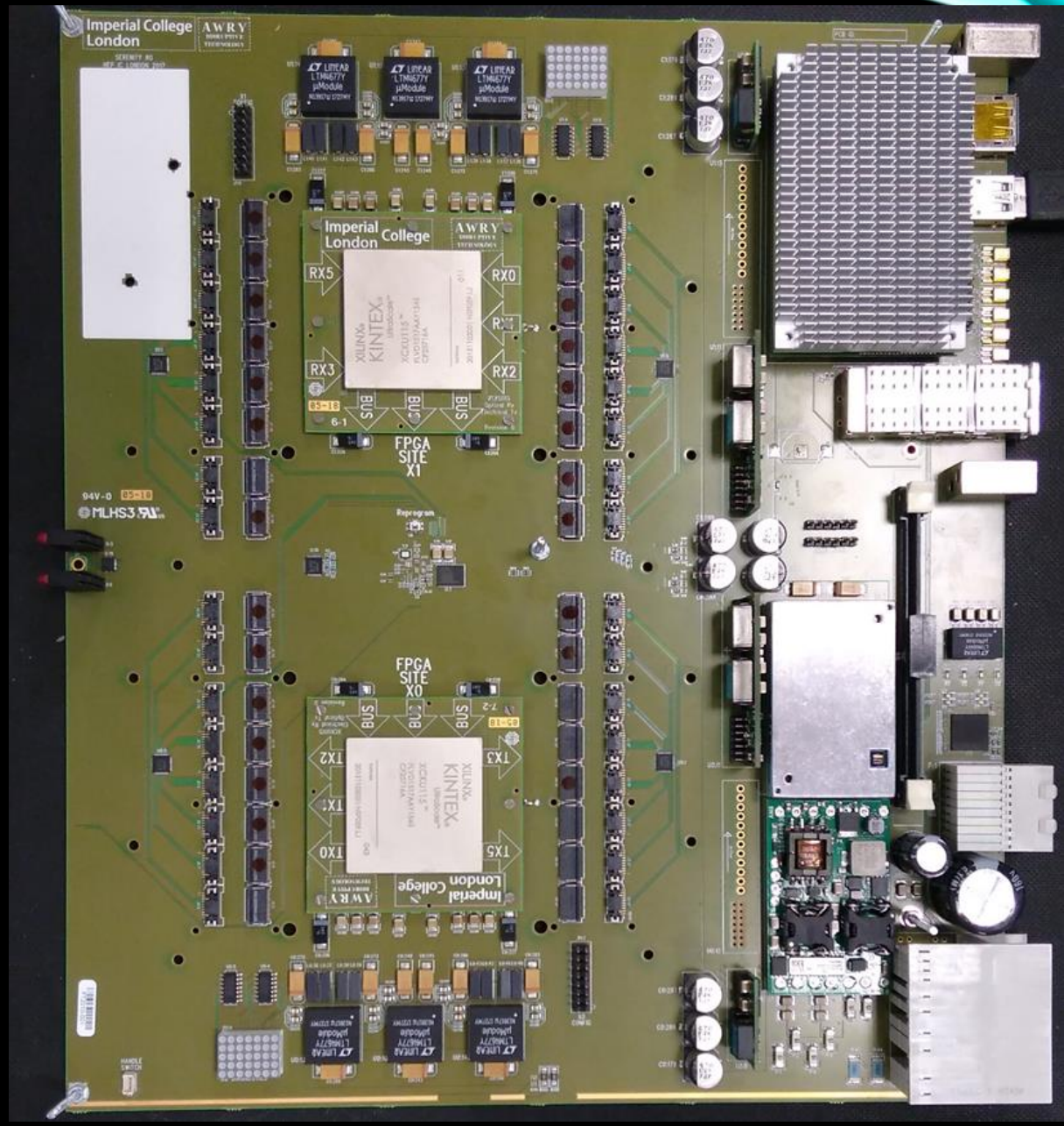
Science & Technology Facilities Council

tifr tata institute of fundamental research

KIT Karlsruher Institut für Technologie

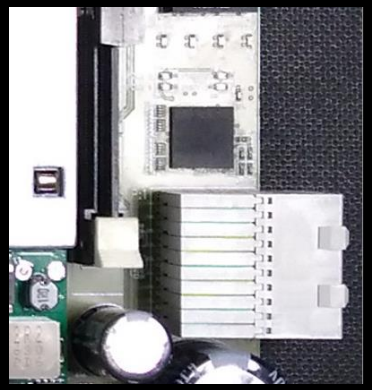
# WHAT IS SERENITY?

- ATCA Development Platform
- Carrier Card
  - Services - Power, Clocks, Optics, Interconnects, IPMC & CPU
- Daughter Cards
  - Data Processing FPGAs
- Firmware & Software
  - Generic, Flexible Infrastructure



# AN ASIDE: COTS COMPONENTS

ATCA low level control – IPMC
- Available from CERN
- Runs the commercial standard software



Ethernet
- Switch with integrated Gigabit Ethernet Phys
- AC coupled via capacitors
- Small form factor 1cm$^2$
- VLAN capable

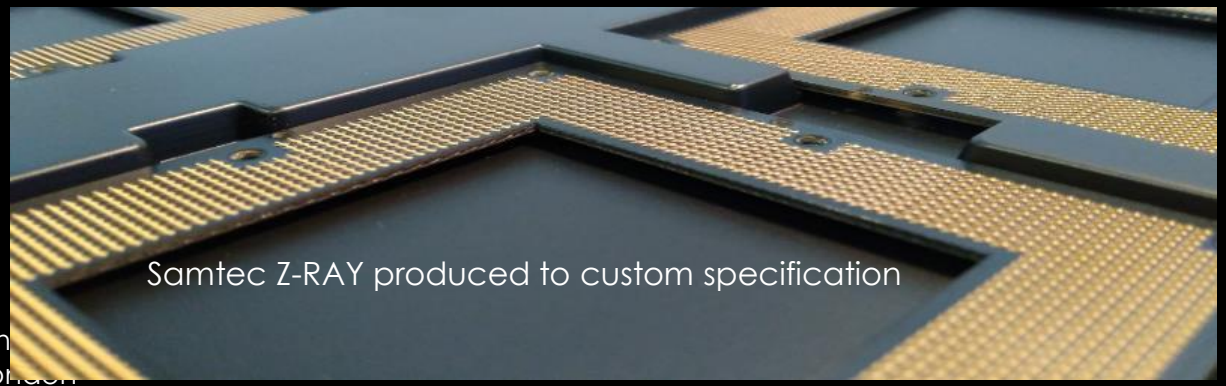
Standard Intel x86 COM-Express Type 10 CPU
- Running standard Centos Linux
- What sys-admins want, not necessarily what hardware engineers want!
- PCIe interface to FPGAs
- Clean separation of hardware, firmware and software

# WHICH PROBLEMS IS SERENITY MEANT TO SOLVE?

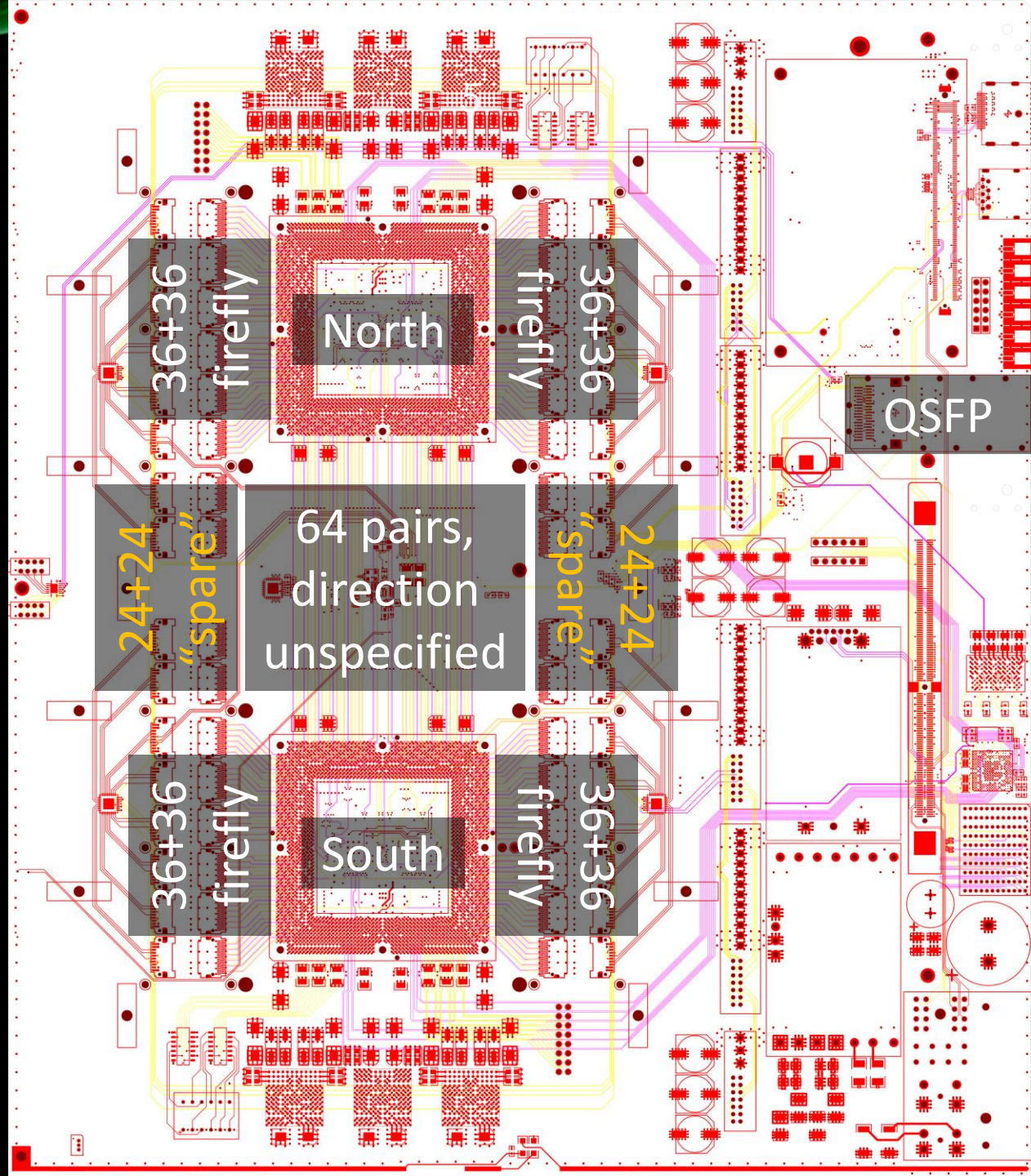
- Different projects want different FPGAs, different optical and electrical connectivity
  - Can one board make everyone happy?
  - If not, can we at least provide a rapid-prototyping platform for ATCA?
- Bulk of cost concentrated in FPGAs, bulk of potential failure modes in carrier
  - Can we decouple financial risk from production risk?

# HOW?

- FPGAs on Daughter-cards:
  - Freedom to choose your preferred family, package, generation, (vendor?)
- Freedom to choose your balance of optical and electrical connectivity
- Carrier testing done with FPGAs safely in their static-bags



Samtec Z-RAY produced to custom specification



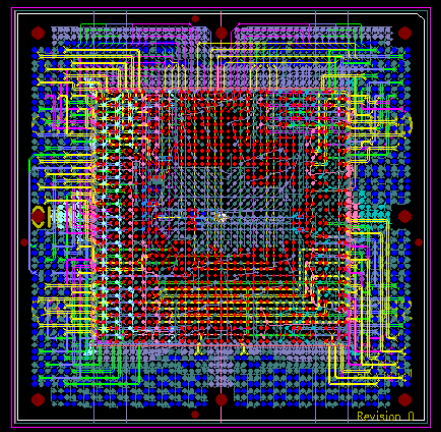

An
London

# WHAT WERE THE CONCERNS?

- We started the project with serious concerns over choice of ATCA, particularly:
  - Thermal management
    - Optics must be kept below 50C or longevity drops at shocking rate
  - Acoustic noise
- Others expressed concern over
  - Interposer signal integrity
  - Supplying power to FPGAs
  - Cooling FPGAs on daughter-cards
  - Limited prior experience with 16 or 25G links
  - Whether different institutes could really produce their own daughter-cards?

# CAN DIFFERENT GROUPS PRODUCE DAUGHTER-CARDS?

- IC - Xilinx KU115: Symmetric & Daisy Chained
- KIT - Xilinx KU15P
- TIFR - Xilinx VU9P
- Saclay – a clock-network analysis daughter card

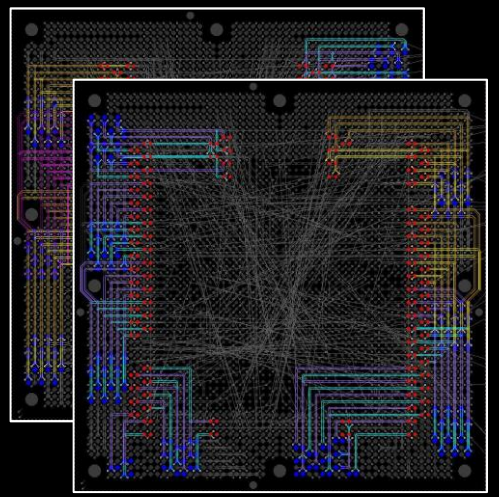


All optical
KU115, Imperial

Andrew Rose, Imperial College London

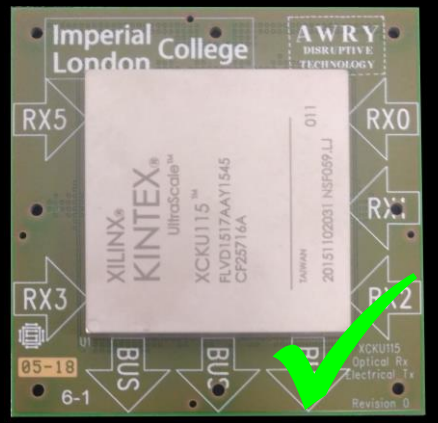
Mixed optical/electrical
KU15P, KIT

Clock-performance analyzer
CEA Saclay



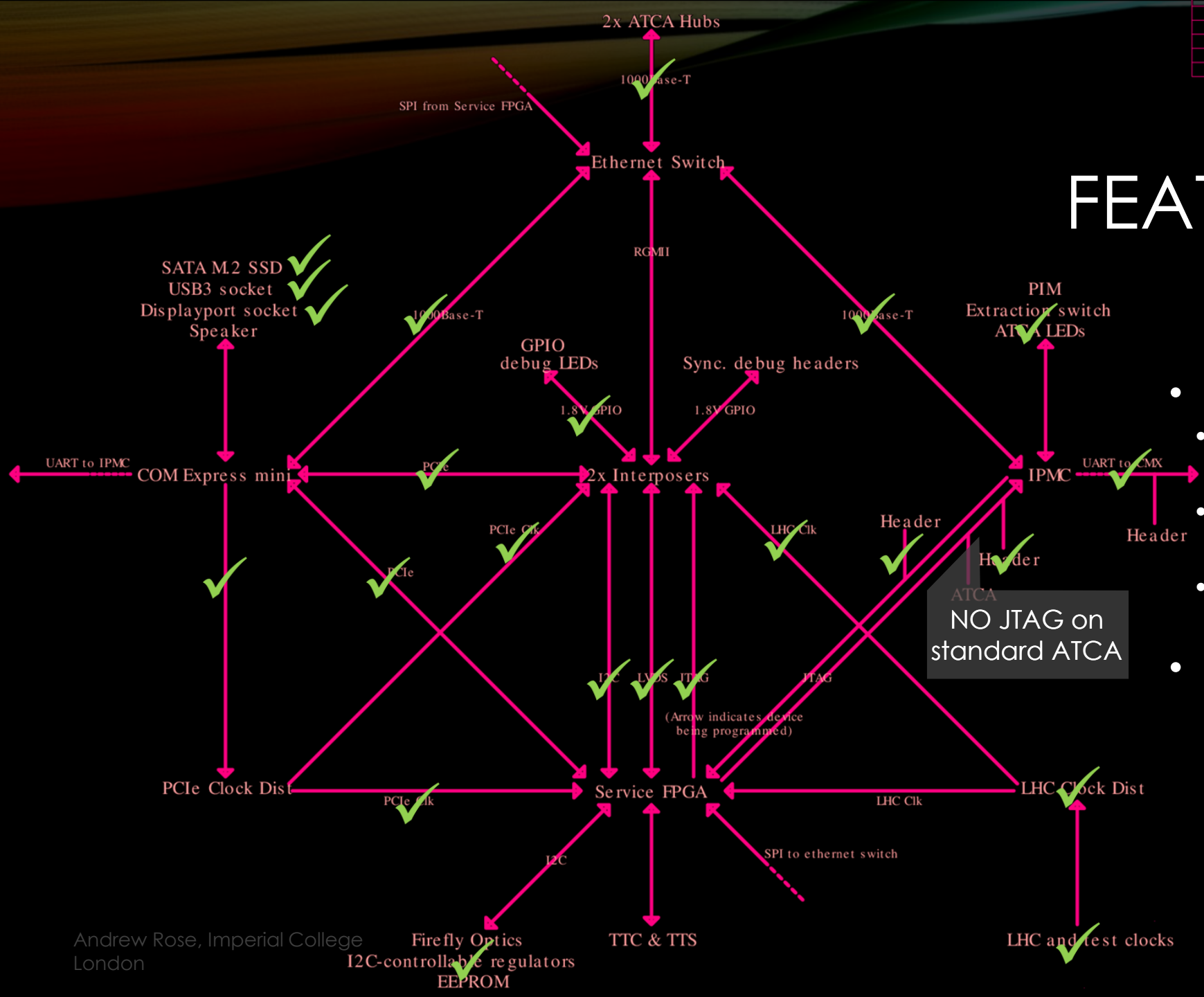
In progress, All optical
VU9P, TIFR

Daisy-chain, optical in
KU115, Imperial



Daisy-chain, optical out
KU115, Imperial

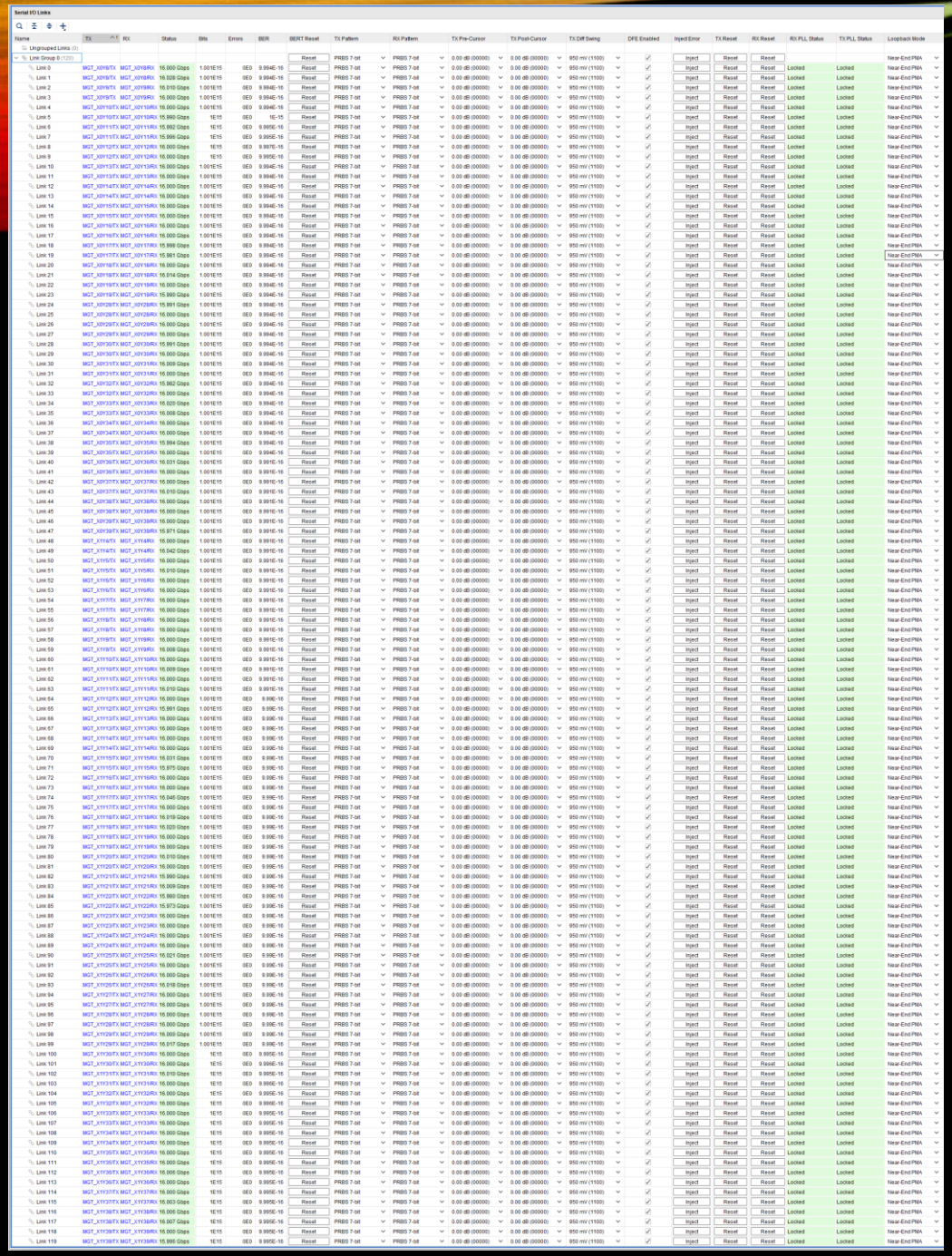
# FEATURE TESTING

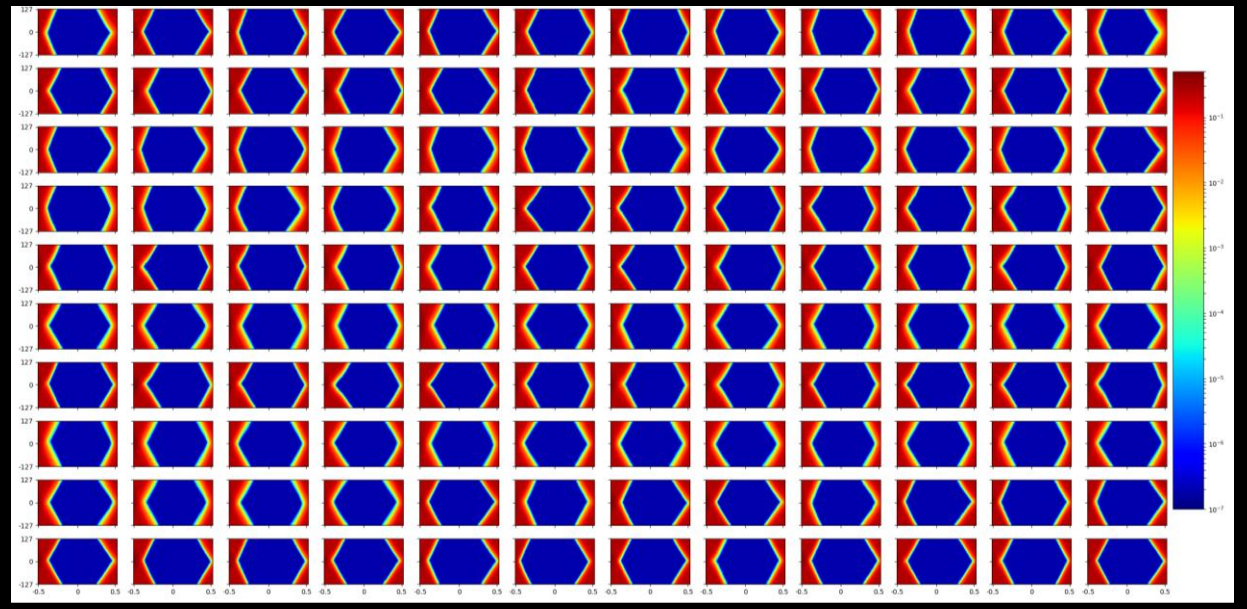



- All features tested except
  - TTC/TTS distribution
    - No DTH
  - SPI control of Ethernet switch
    - Switch works in default config.
  - RGMII to FPGAs
    - Using PCIe instead
- All tested features working

# LINKS - INTERNAL LOOP

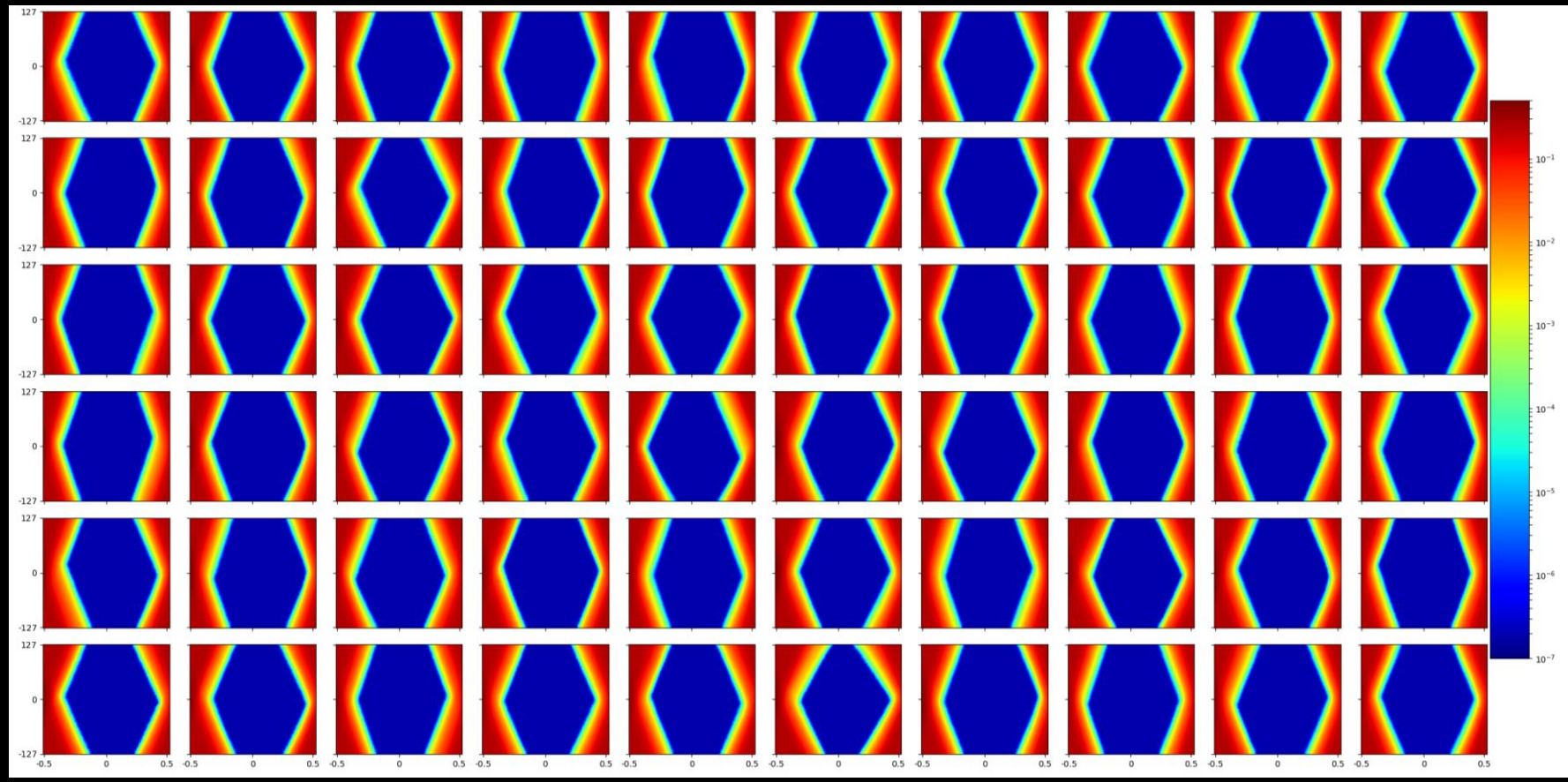
- 120 simultaneous IBERTs
  - 16Gbps, Near-end PMA loopback. Default settings
- Each link passed 1e15 bits - No errors
- 120Pb of data through silicon
- Indicated good clock and power stability





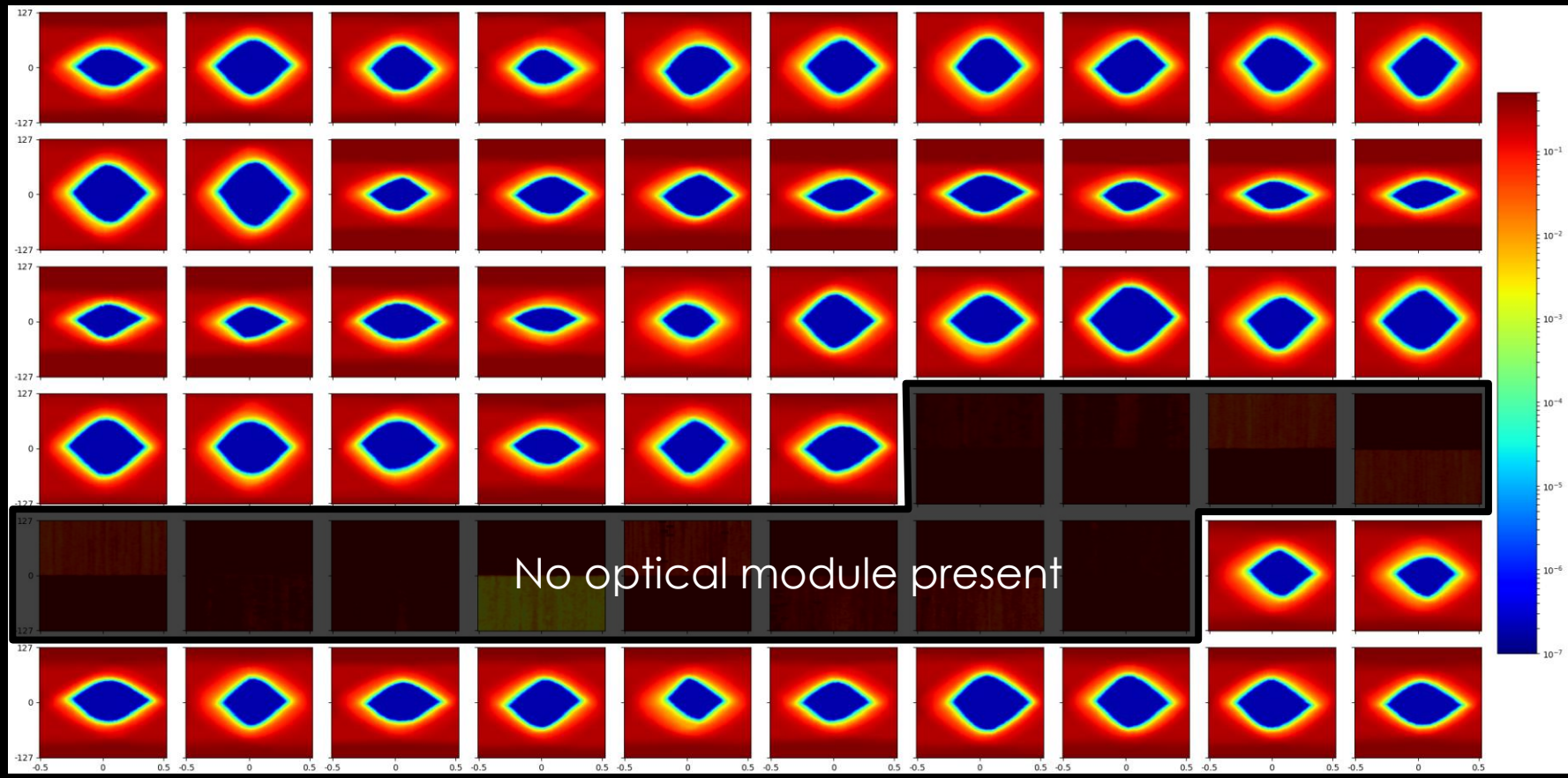
# COPPER LINKS – EYE DIAGRAMS

- Inter-interposer bus
- Two independent FPGAs
- 16Gbps, DFE disabled, No Pre- or Post-Cursor
- Each link passed 8e14 bits – No errors



# OPTICAL LINKS – EYE DIAGRAMS

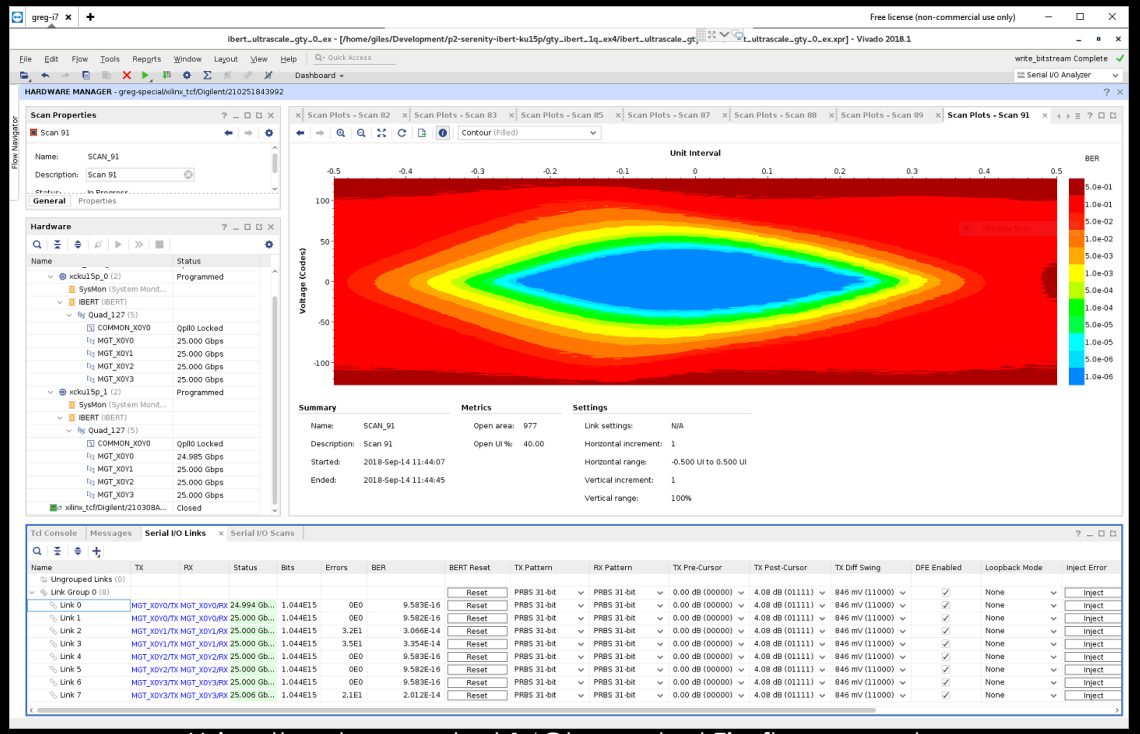
- Firefly optics
- Two independent FPGAs
- 16Gbps, DFE disabled, No Pre- or Post-Cursor
- Default optical module settings
- 10m optical fibre
- Each link passed 8e14 bits – No errors




FPGA 1 → daughter-card → interposer → motherboard → firefly → MTP → 10M optical cable → MTP → firefly → motherboard → interposer → daughter-card → FPGA 2

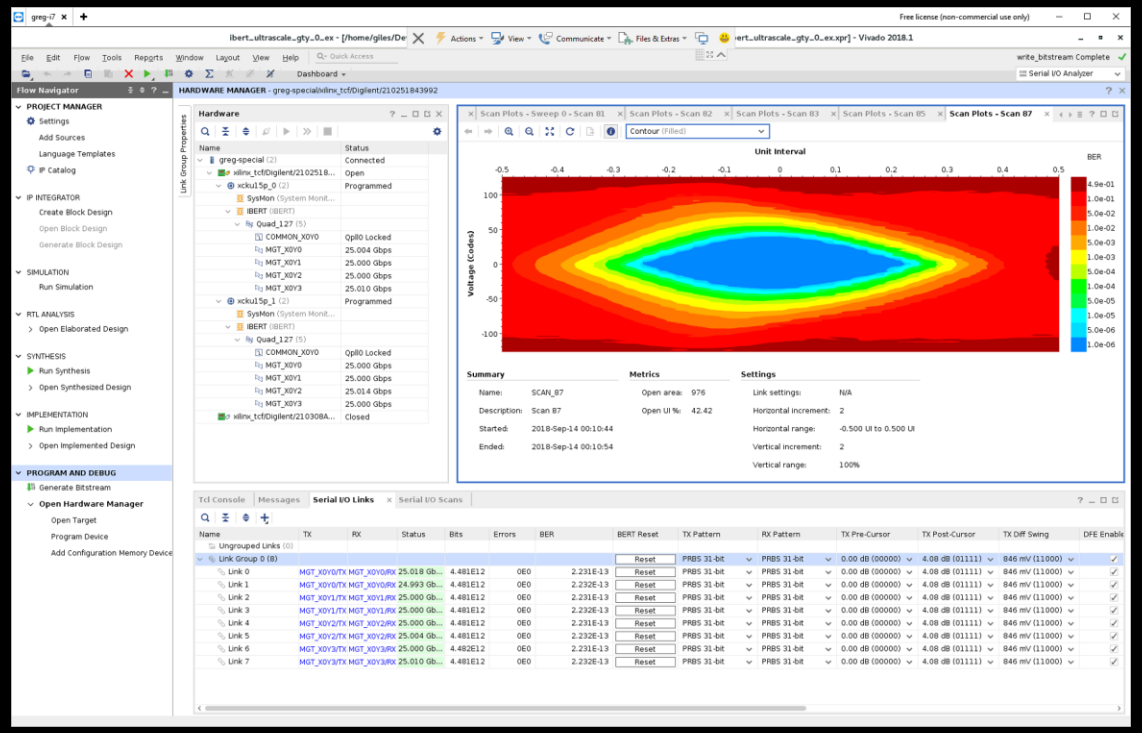
# 25GBPS INITIAL RESULTS

Firefly over copper 25Gbps PRBS31

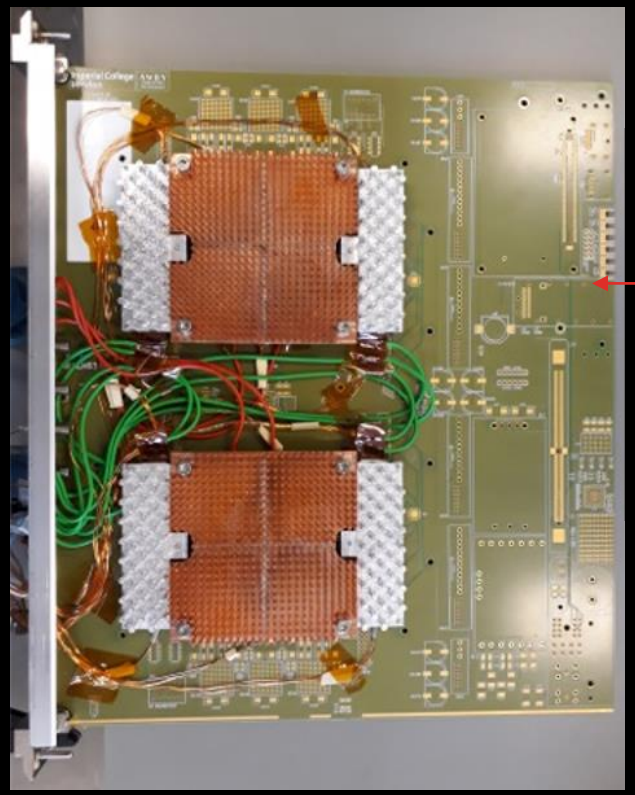


Using the deprecated 14Gbps-rated Firefly connector

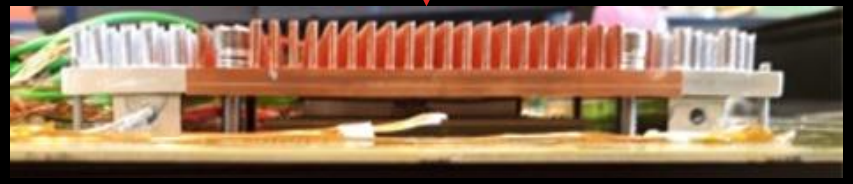
Inter-interposer bus 25Gbps PRBS31

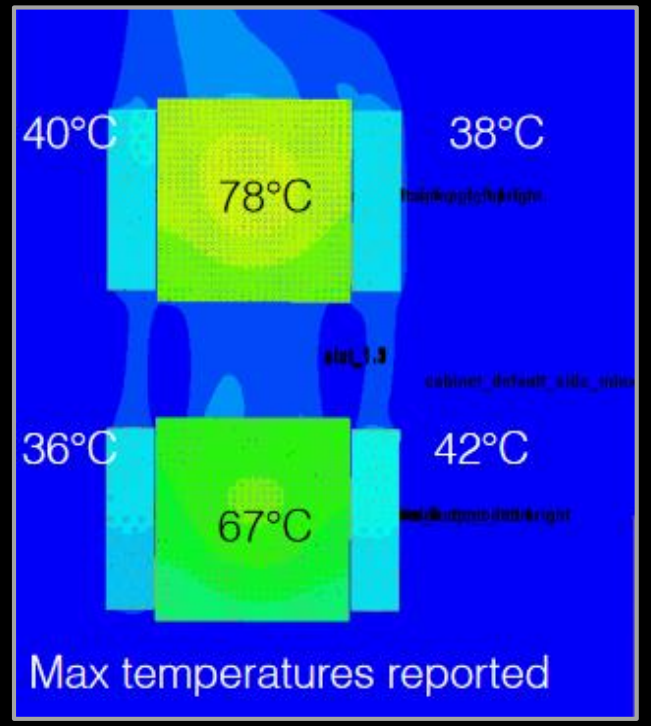


# THERMAL & MECHANICAL TESTS



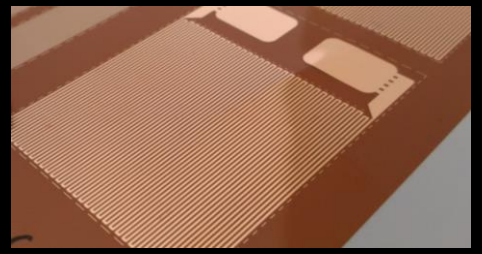
- Thermal simulations
- Physical thermal studies at CERN
- Mechanical component design, studies into stress on FPGA solder balls and stress on PCBs under way at IC

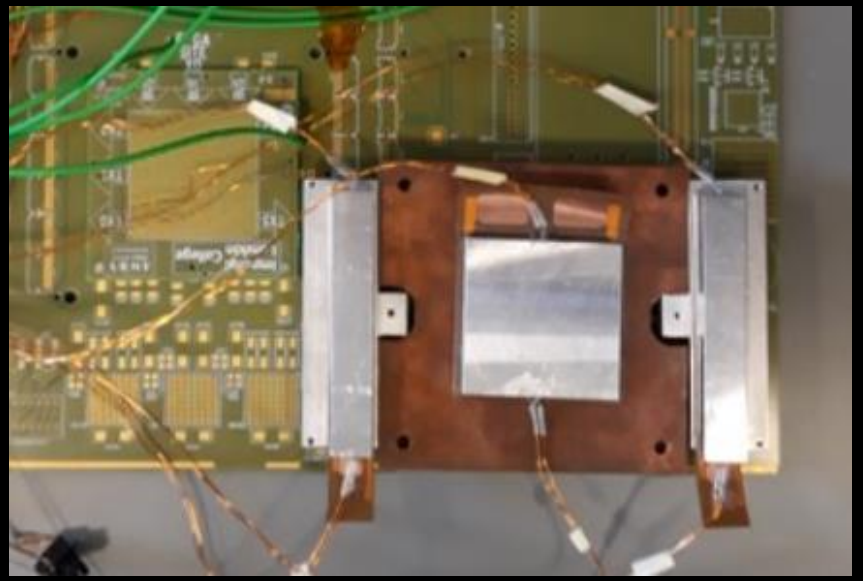


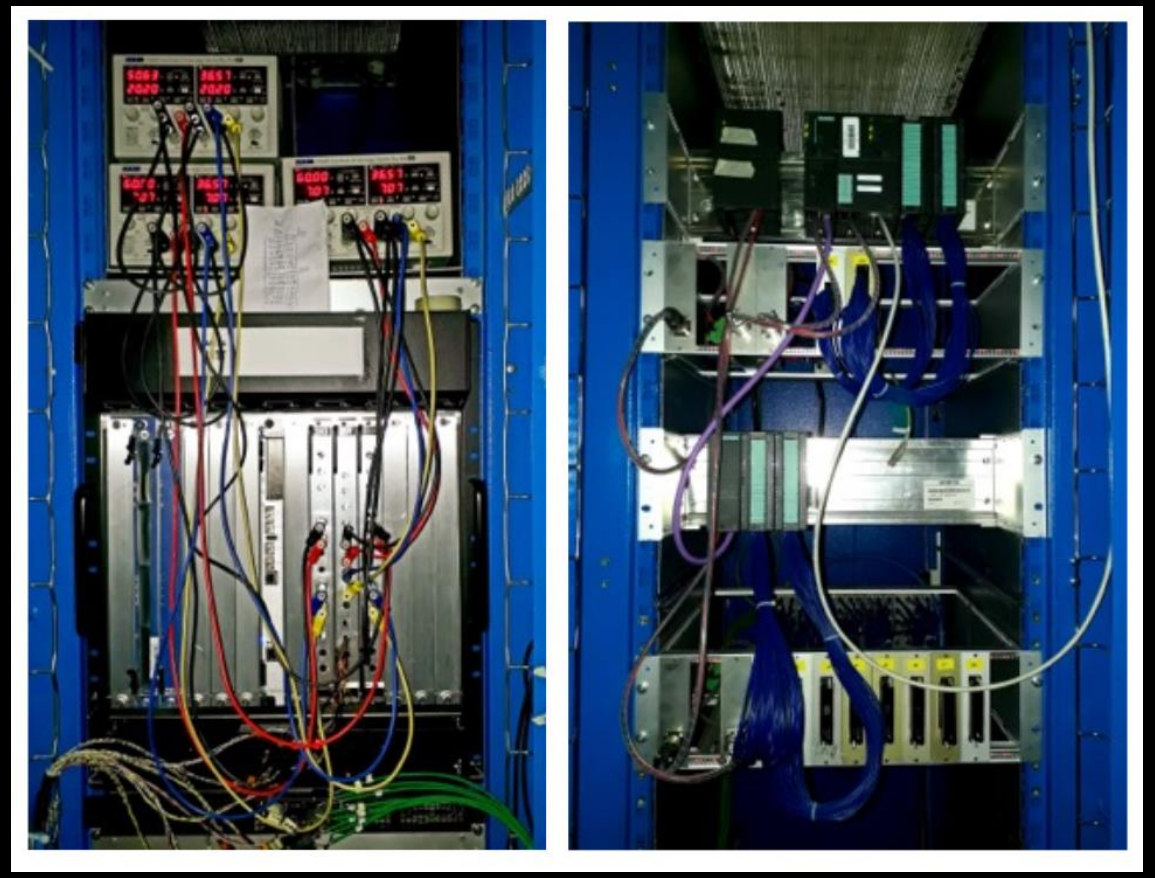


# TEST STAND AT CERN

- Kapton heaters
- Comtel Crate
  - Front-Back Airflow

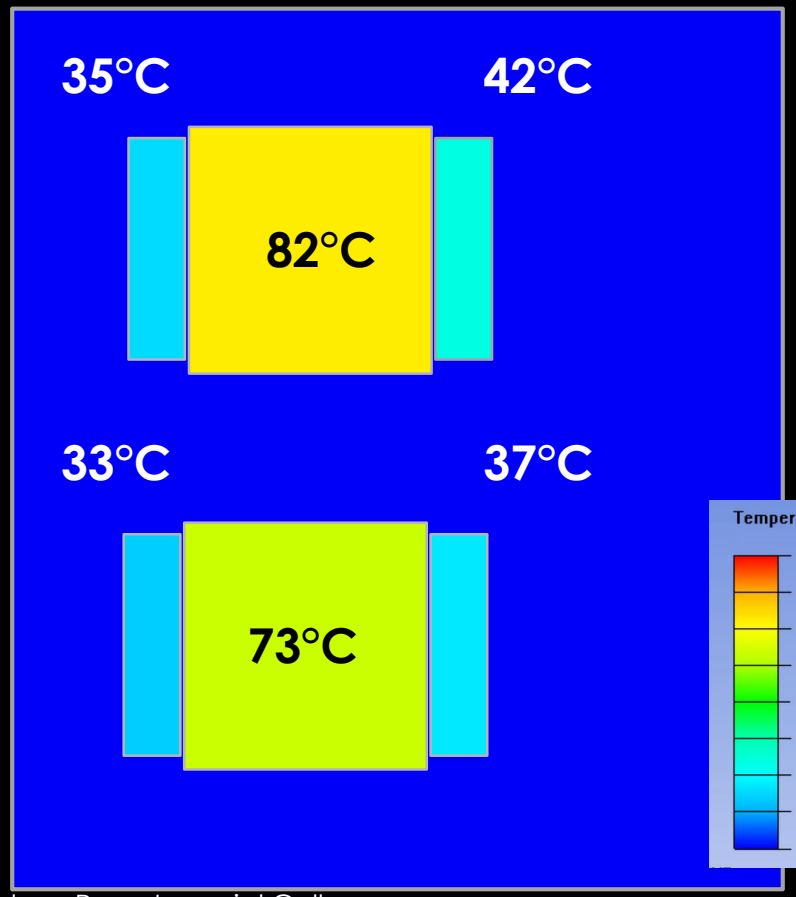




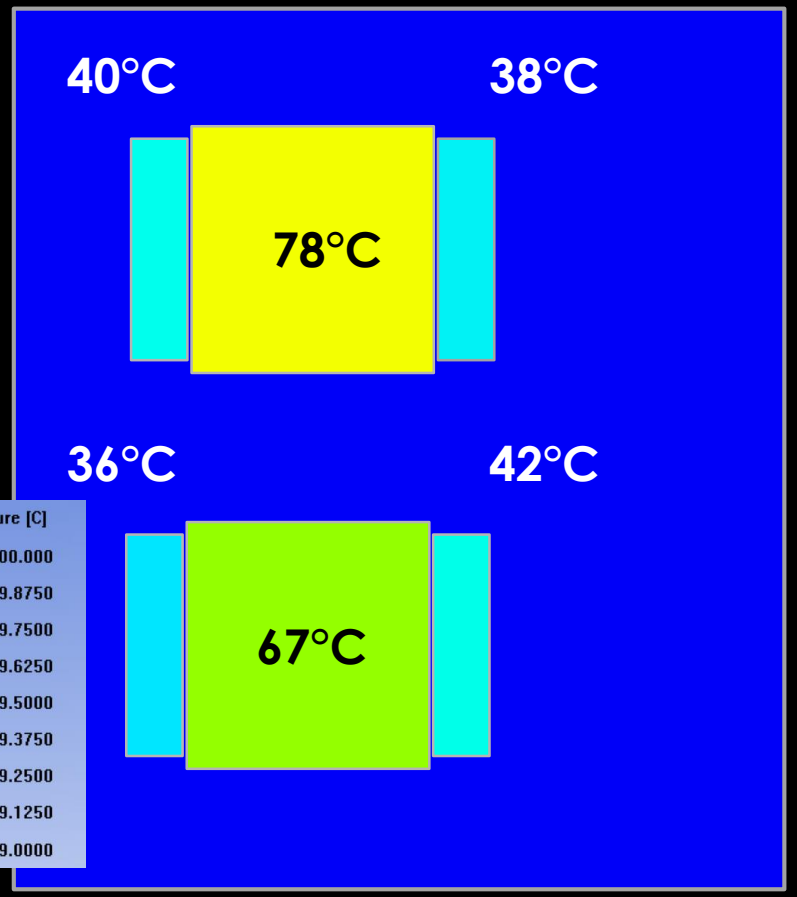


# THERMAL TEST RESULTS

**Measurement**, Fan Speed 15




**Simulation**, Fan Speed 15

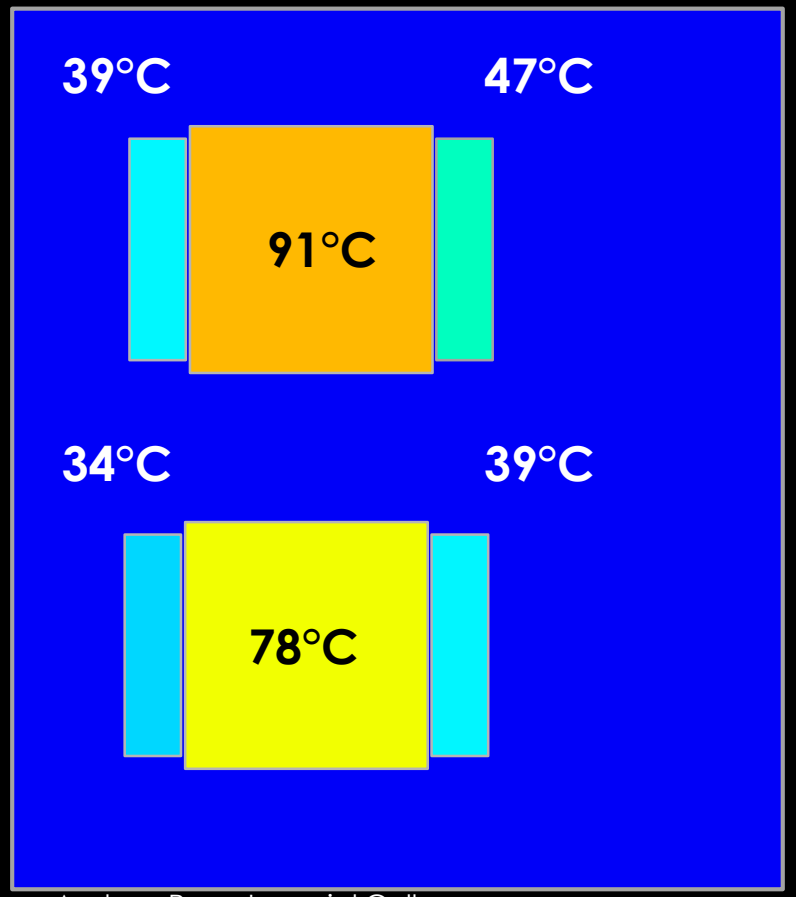



- Standard CERN crate made by Schroff, Vertical Airflow
  - **FPGAs** @ **90W** each
  - Max temperature **100 °C**
  - **Optics** @ **10W** each
  - Max temperature **50 °C**
  - May increase to **20W** so need margin
- Simulation by Ansys
- KIT also running simulations with Mentor FloTHERM for comparison

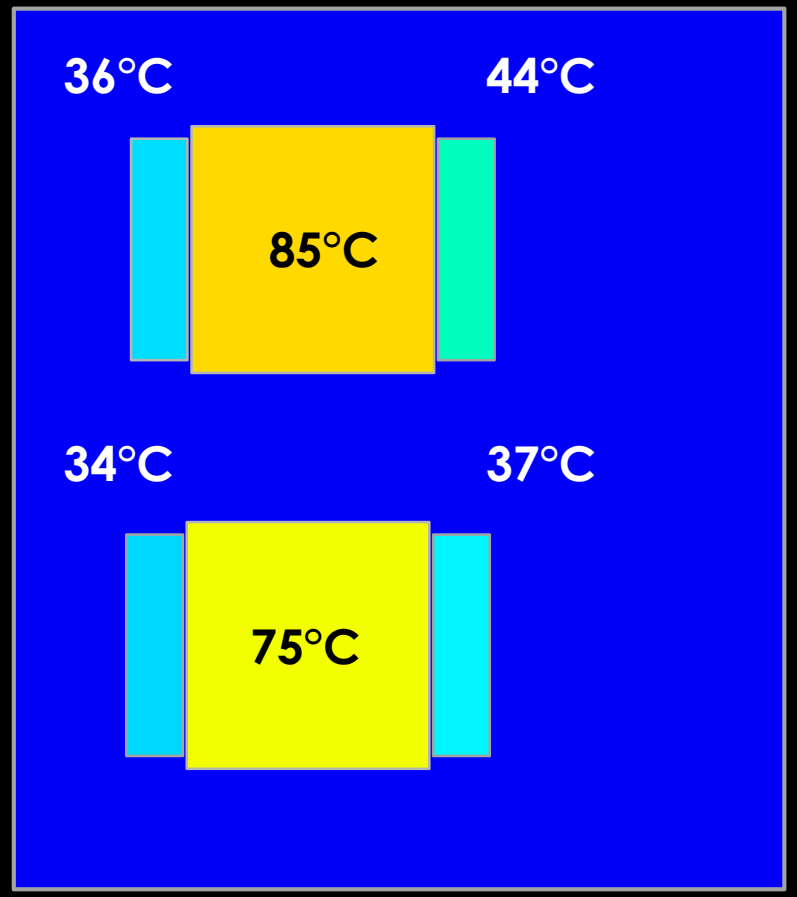
Pisa

# FAN SPEED

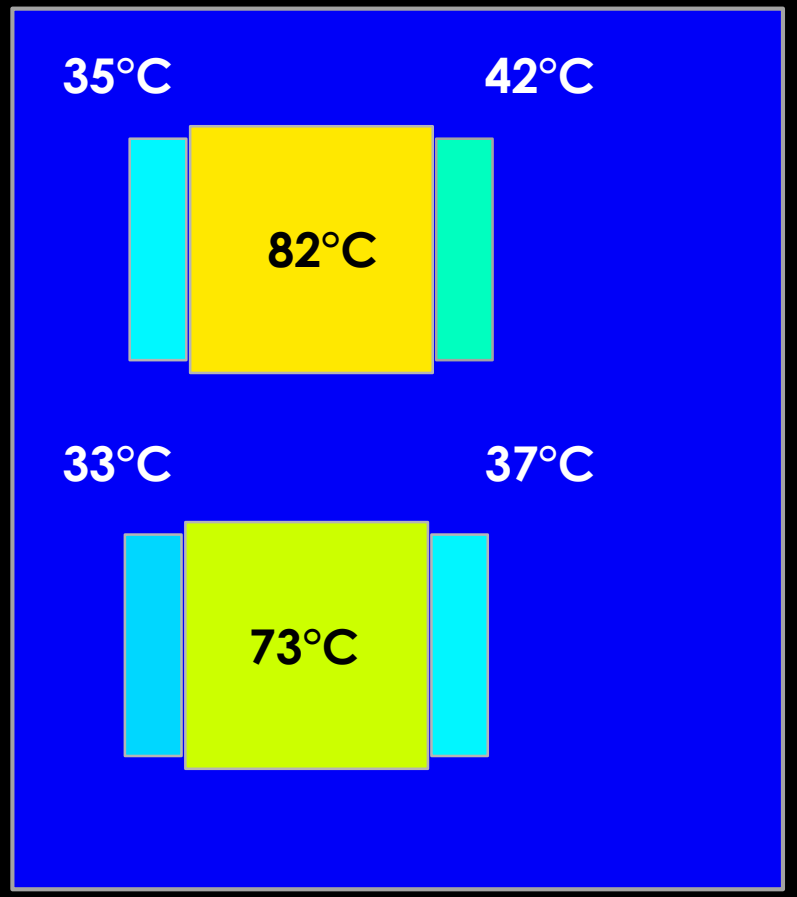
Measurement, Fan **Speed 10**




Measurement, Fan **Speed 13**



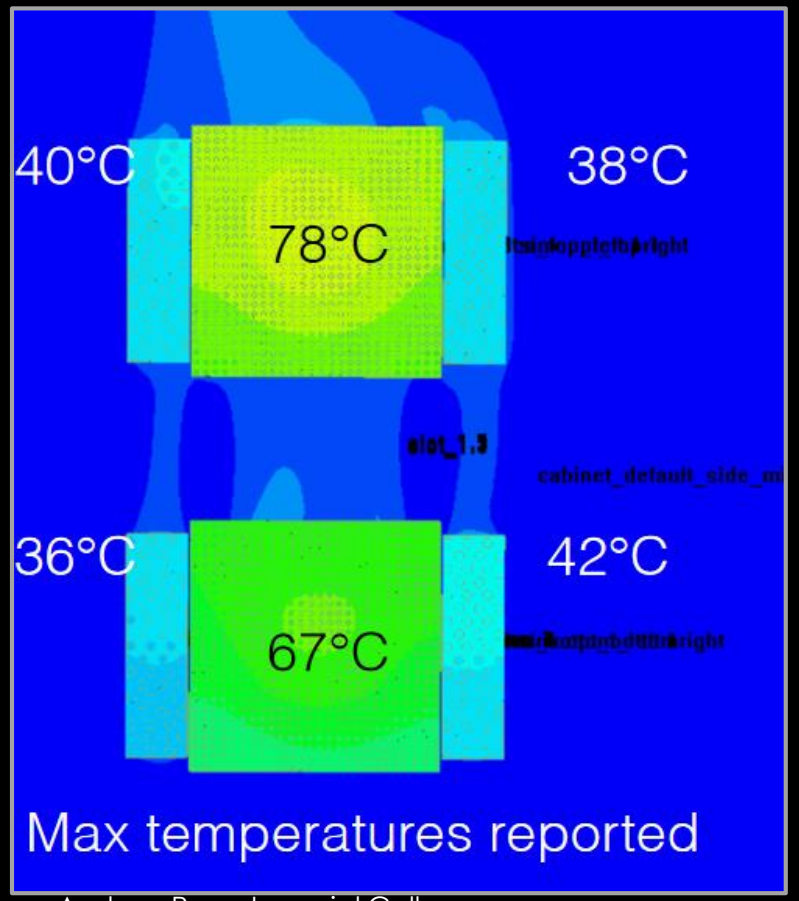

Measurement, Fan **Speed 15**

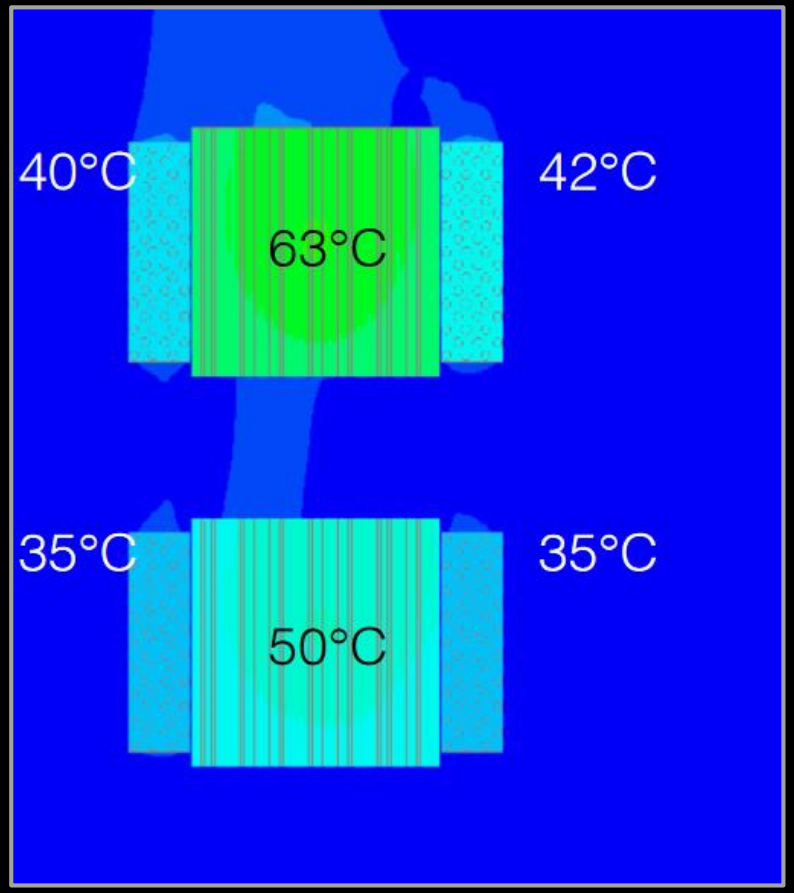


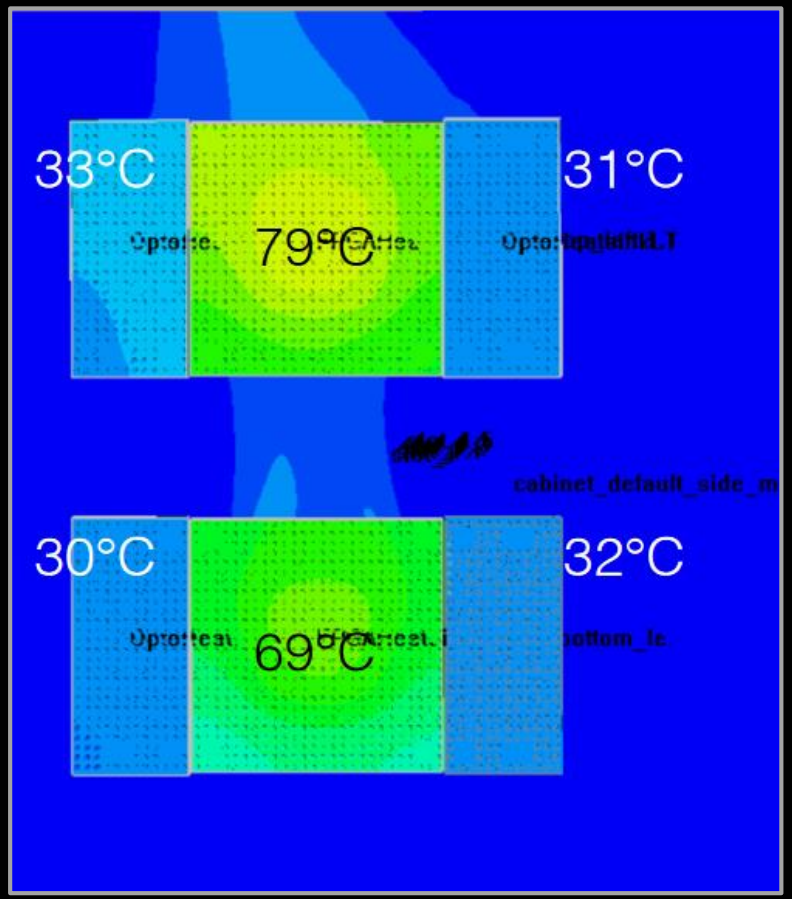
Pisa

# HEATSINK TYPE

- Simulations suggest significant improvement with custom heatsink.
- Validation of simulation with measurement under way



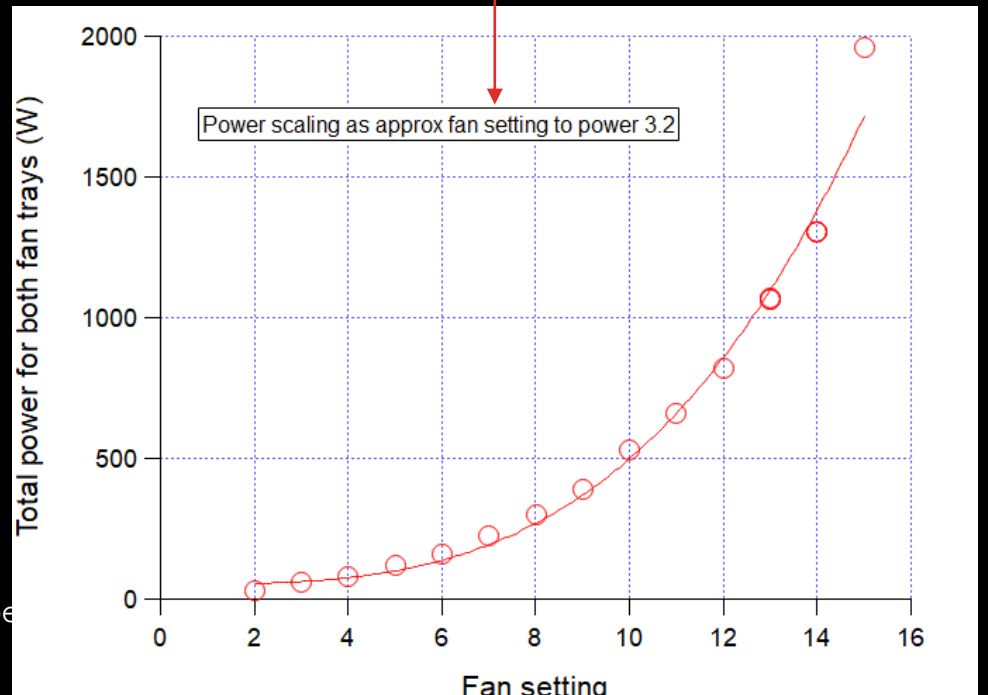



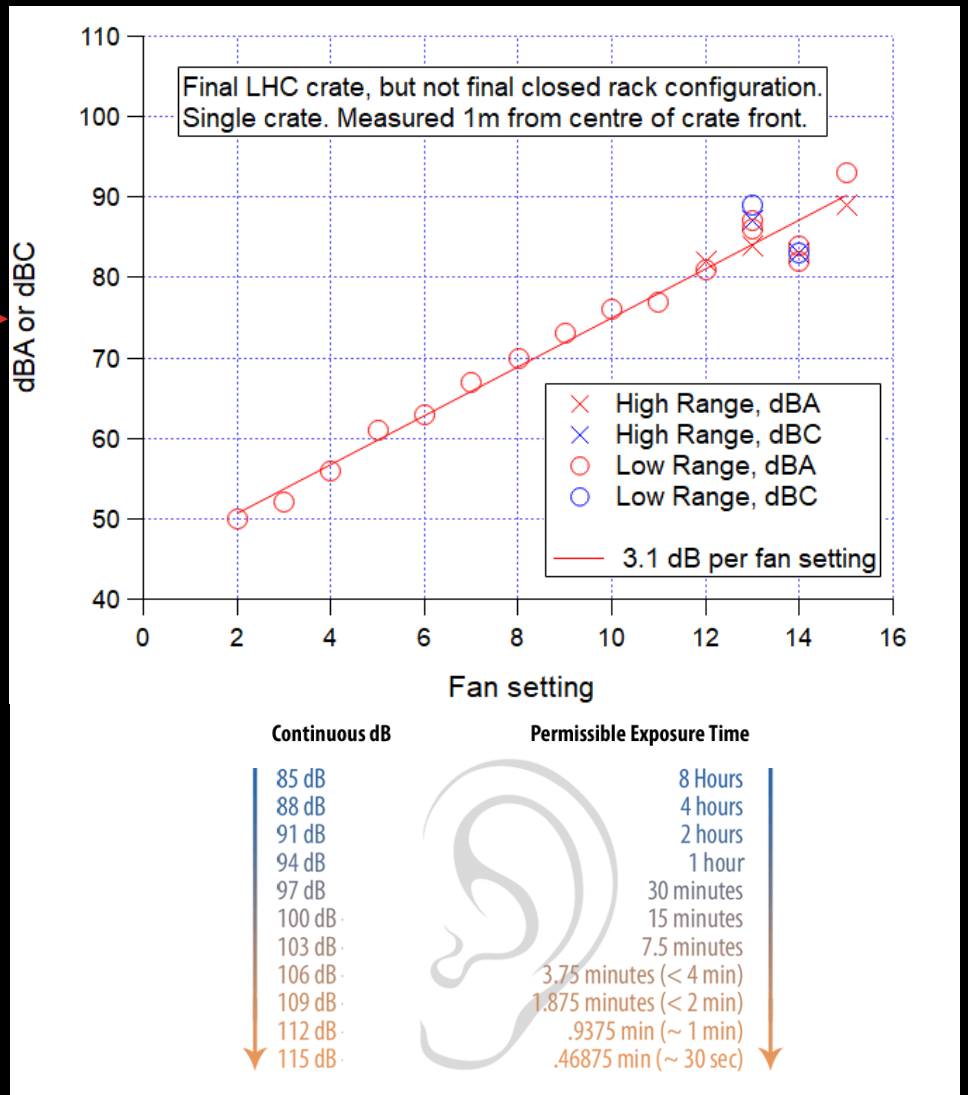

All Simulation, Fan Speed 15

Pisa, IC

# CRATE NOISE & POWER

- Measurements of the running ATCA system still a cause for concern regarding
  - Aural (acoustic) health and safety
  - Power-consumption required to cool system

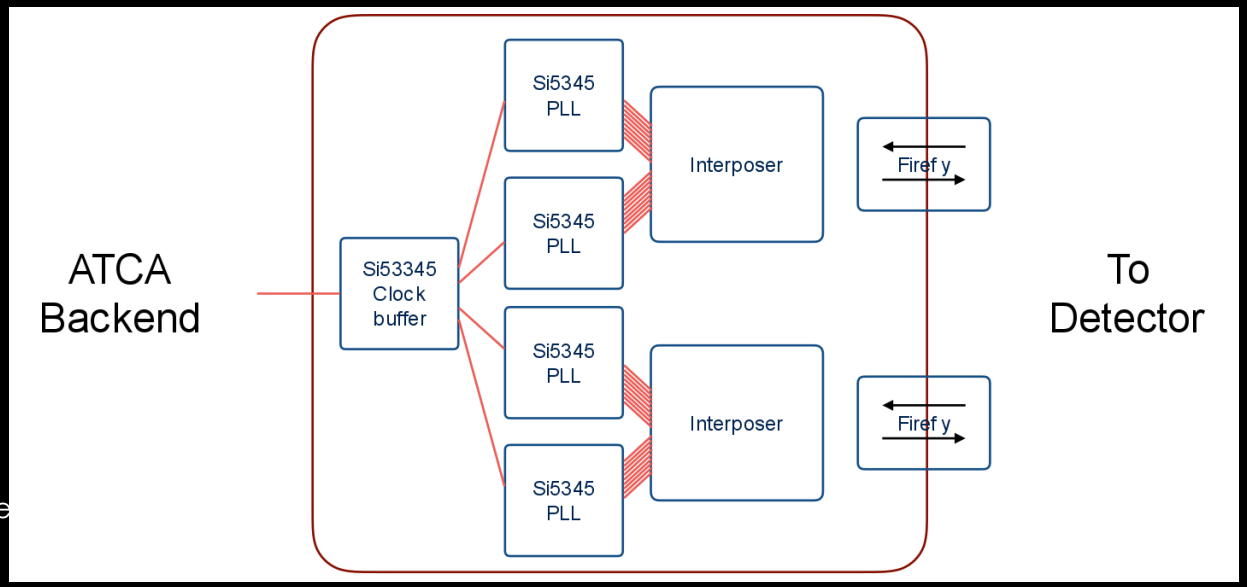


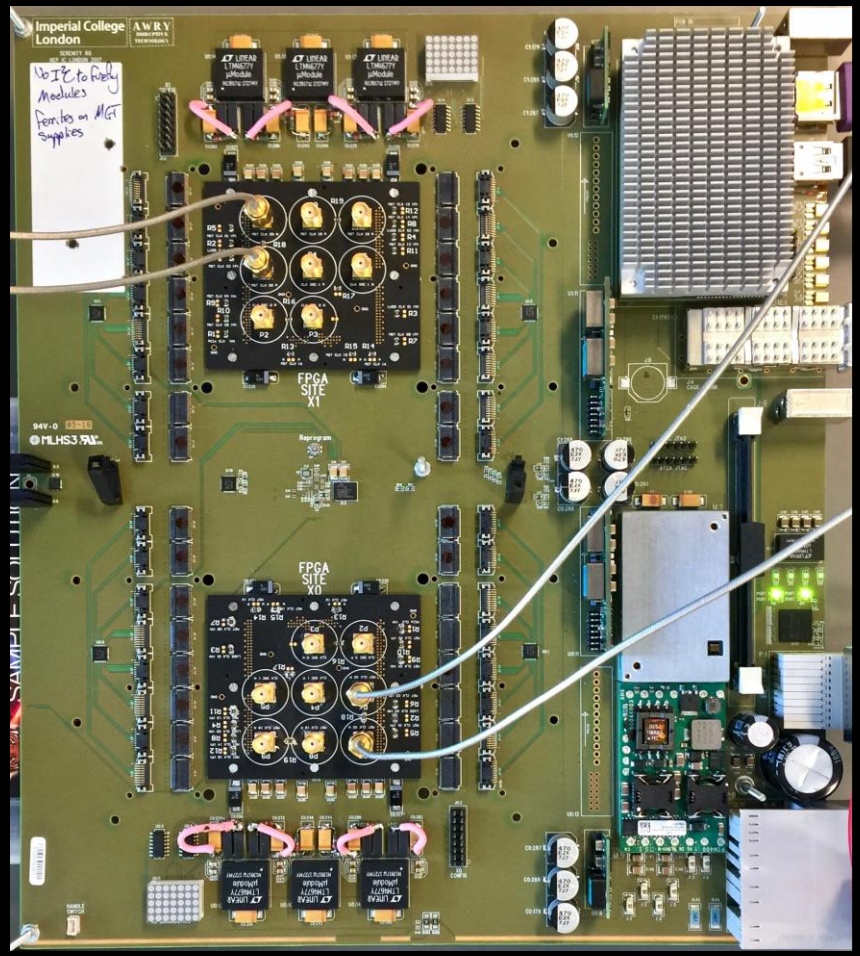



Saclay

# CLOCK-DISTRIBUTION TESTS

- Serenity is TDR proposal for CMS HGC DAQ, trigger, clocking and control board.
- HGC requires precision timing distribution
- CEA Saclay have been testing whether Serenity's clocking performance meets HGC requirements



Saclay

# CLOCK-DISTRIBUTION TESTS

- 320.624MHz test clock (HGC precision-timing frequency)
- External clock-source, RMS = 1.3ps
- Serenity tested at 20GS/s (Eval board at 40GS/s)

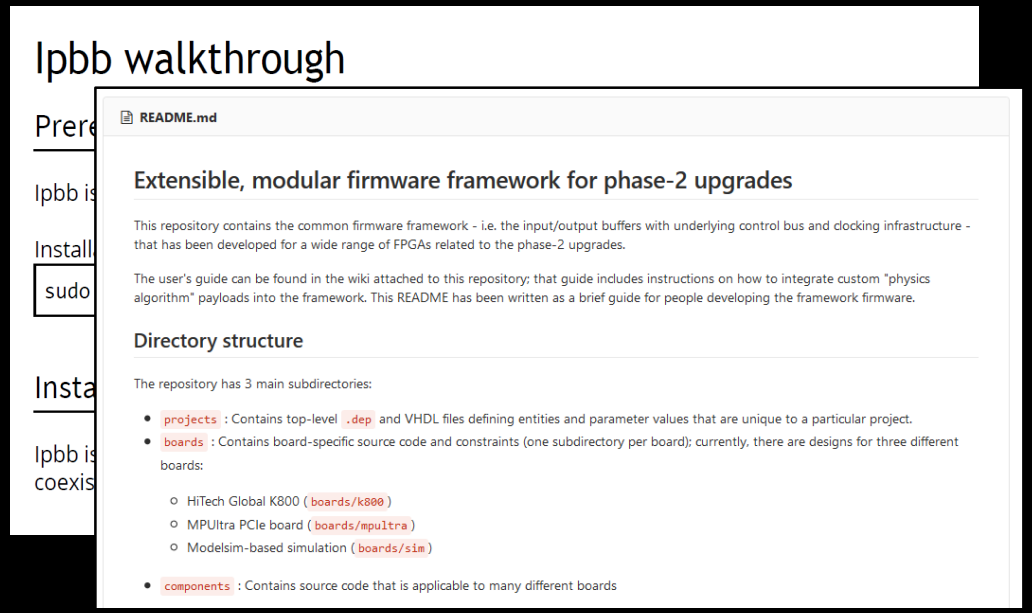
| Measurements by CEA Saclay | SI5444 Eval board for comparison (4 outputs) | Serenity North LHC clock (9 outputs) | Serenity South LHC clock (9 outputs) | Serenity Overall |
|---|---|---|---|---|
| Random Jitter | 0.9ps | 1.6ps | 1.4ps | 1.8ps |
| Deterministic Jitter | 1.9ps | 3.2ps | 3.1ps | 4.2ps |
| RMS | 1.3ps | 2.3ps | 2.1ps | 2.8ps |

- Based on these results, Serenity's contribution to system jitter is minimal and it can be considered a "pure clock distribution" node

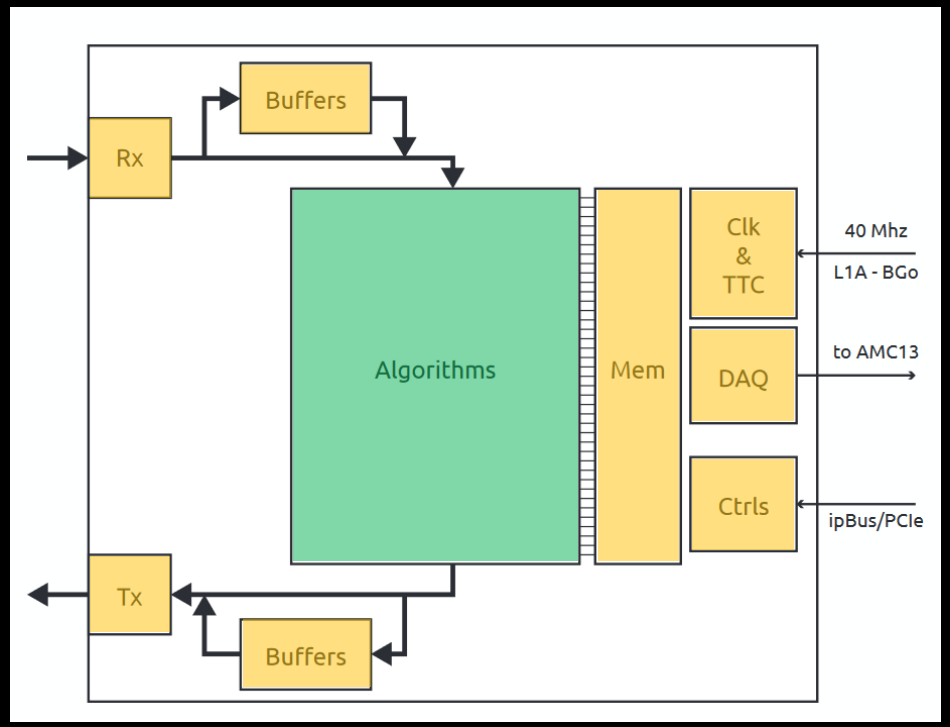
RAL+TIFR

# FIRMWARE

- Ipbus-over-PCIe worked out-of-the-box
- Have built upon the success we had with MP7 model in Phase-1 upgrades
  - Infrastructure firmware – EMP
  - Build Tool - IPBB



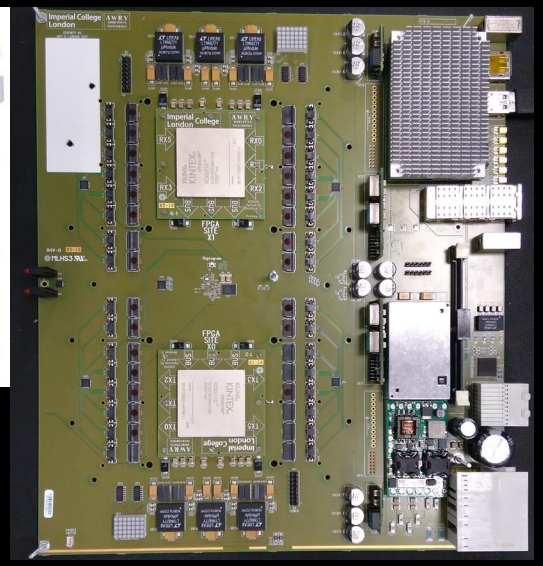

- Infrastructure separated from payload for algorithms



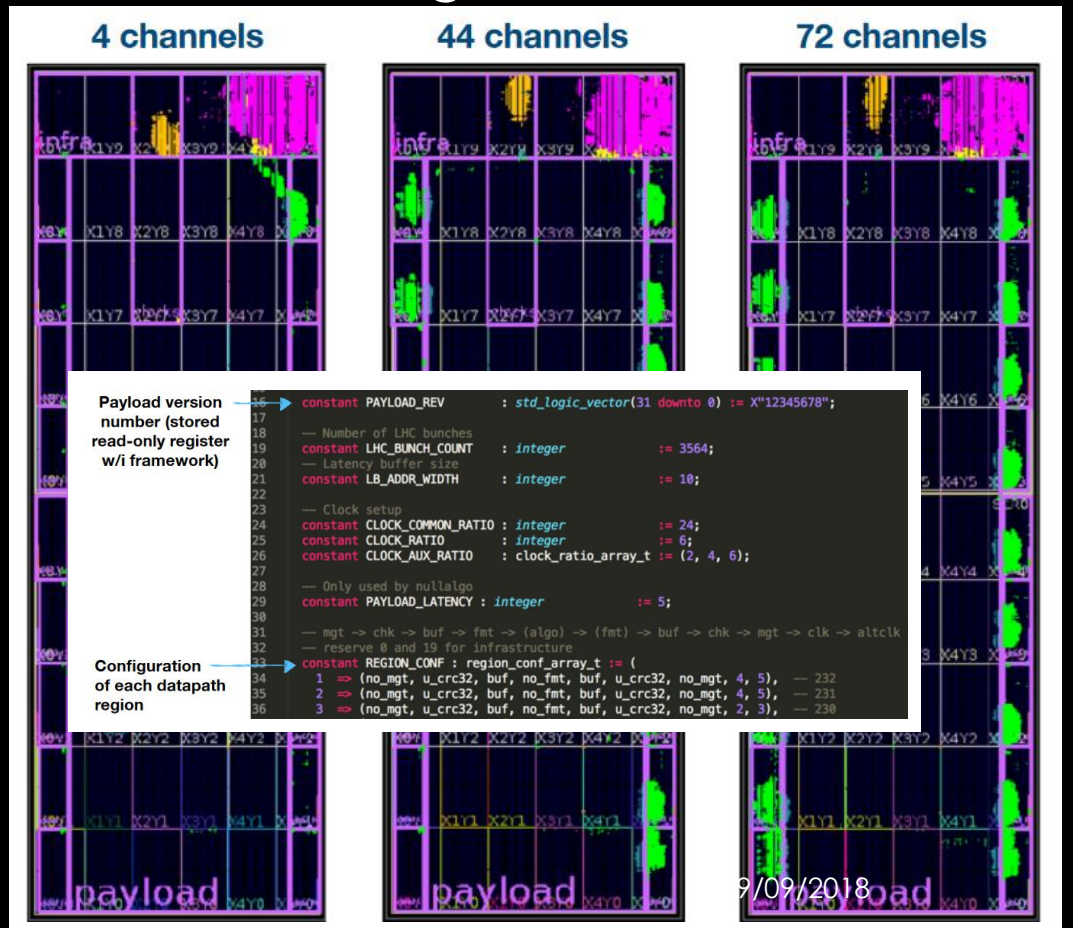
RAL+TIFR

# EMP FIRMWARE

- Different boards



- Different configurations



# CONCLUSIONS

- Serenity has succeeded in its initial aims of allowing multiple types of FPGA and user-specified connectivity, and also in mitigating risk
  - Different institutes have successfully produced their own daughter-cards, each with different generations & models of FPGAs
- Concerns over interposer signal integrity and supplying power to FPGAs seem unfounded
- Performance at both 16G and 25G looks good
- Cooling of both optics and FPGAs on daughter-cards looks manageable
  - General concerns over power and noise of LHC-scale systems built using ATCA remain

# WHAT NEXT?

- 20 Revision 1.1 cards in assembly now
- A number of cards destined for “time-share test-stands” at CERN
  - Allow systems to test “full-stack” algorithms from source to sink
  - Also, horizontal integration with CERN DTH & online software when they become available
- Remaining cards destined for collaborating institutes

# THANKS FOR LISTENING!
# ANY QUESTIONS?

CEA Saclay: Ozgur Sahin, Pierre-Anne Bausson

IC: Andrew Rose, Duncan Parker, Greg Iles

INFN Pisa: Giacomo Fedi, Piero Giorgio Verdini, Andromachi Tsirou

KIT: Luis Ardila, Matthias Balzer, Thomas Schuh

RAL: Tom Williams, Alessandro Thea, Kristian Harder

TIFR: Shashi Dugad, Raghu Shukla, Irfan Mirza

University of Ioannina: Stavros Mallios

Imperial College London

cea SACLAY

INFN Istituto Nazionale Fisica Nucleare Sezione di Pisa

Science & Technology Facilities Council

tifr tata institute of fundamental research

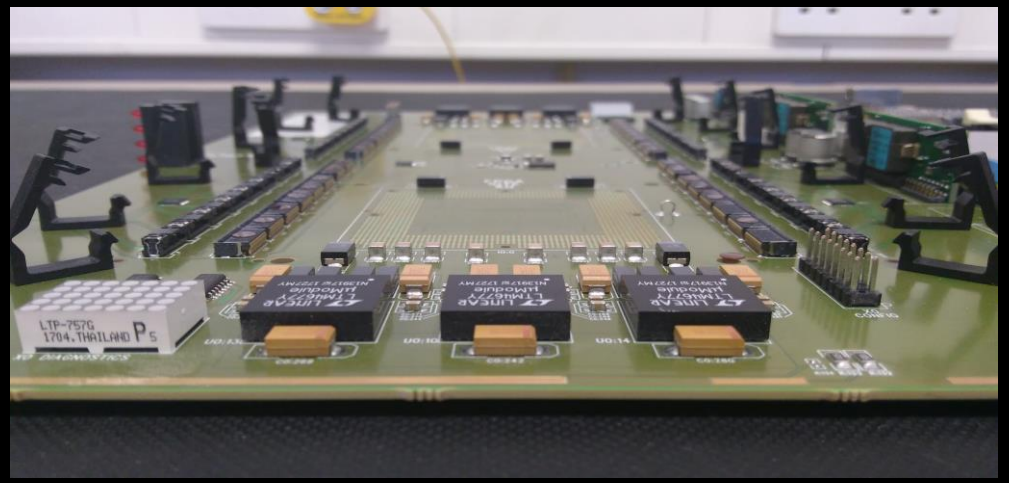
KIT Karlsruher Institut für Technologie

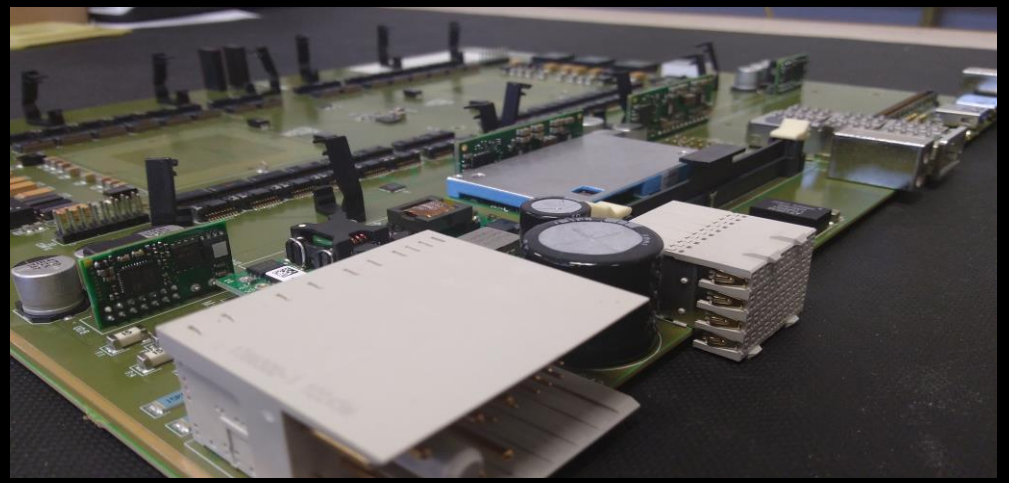
# DISCLAIMER

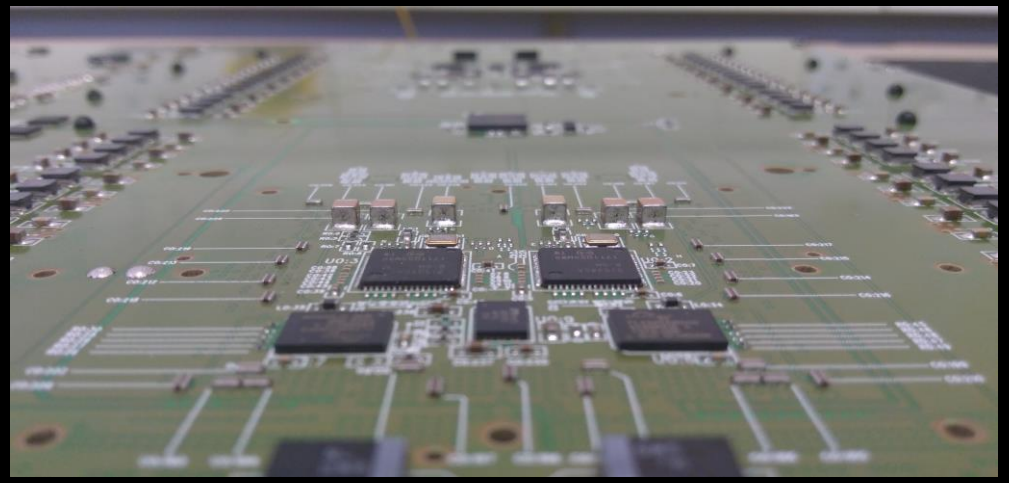
I do not own, nor do I assert any right of ownership over, the Serenity logo.
It is just a cool logo from a cool TV series (Firefly)
and subsequent film (Serenity).
I urge you to go and buy them both: They're a lot of fun.

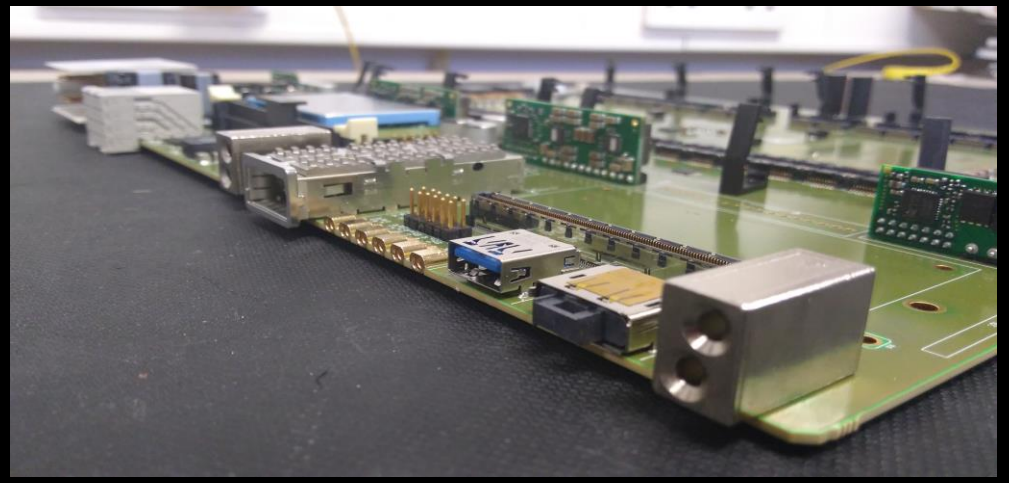
# SPARES & BACKUP

# HARDWARE IN HAND

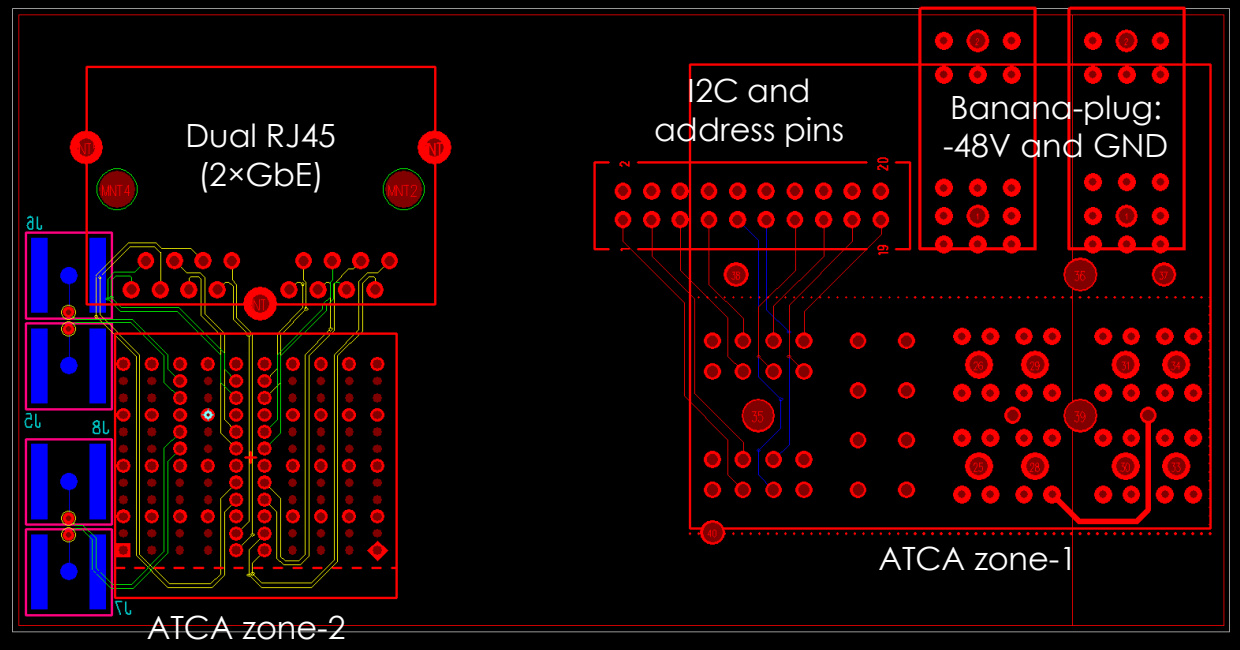


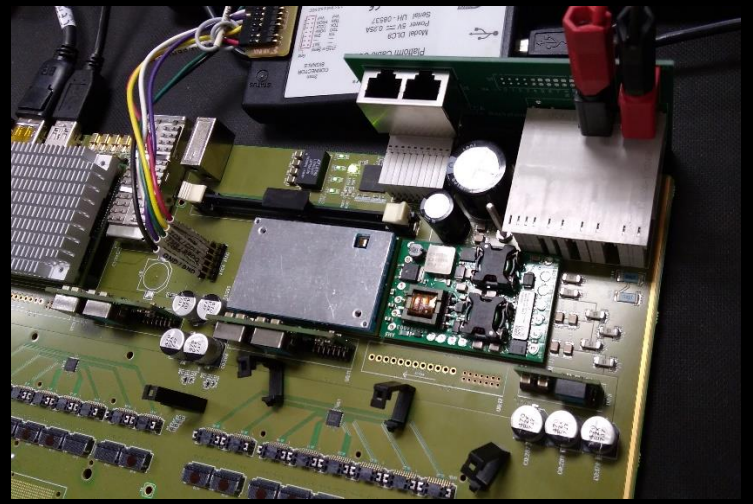




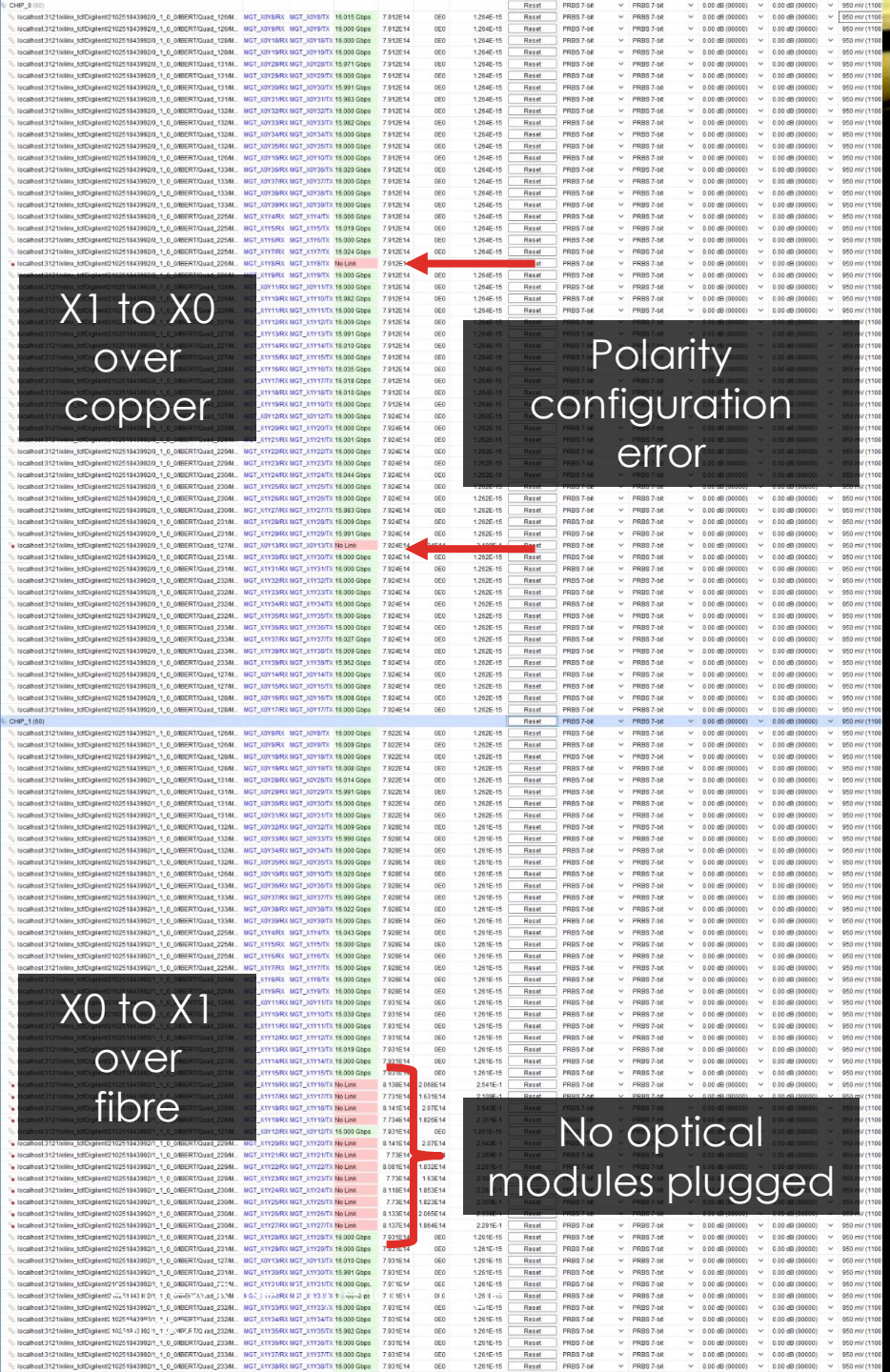


# MICRO-BACKPLANE



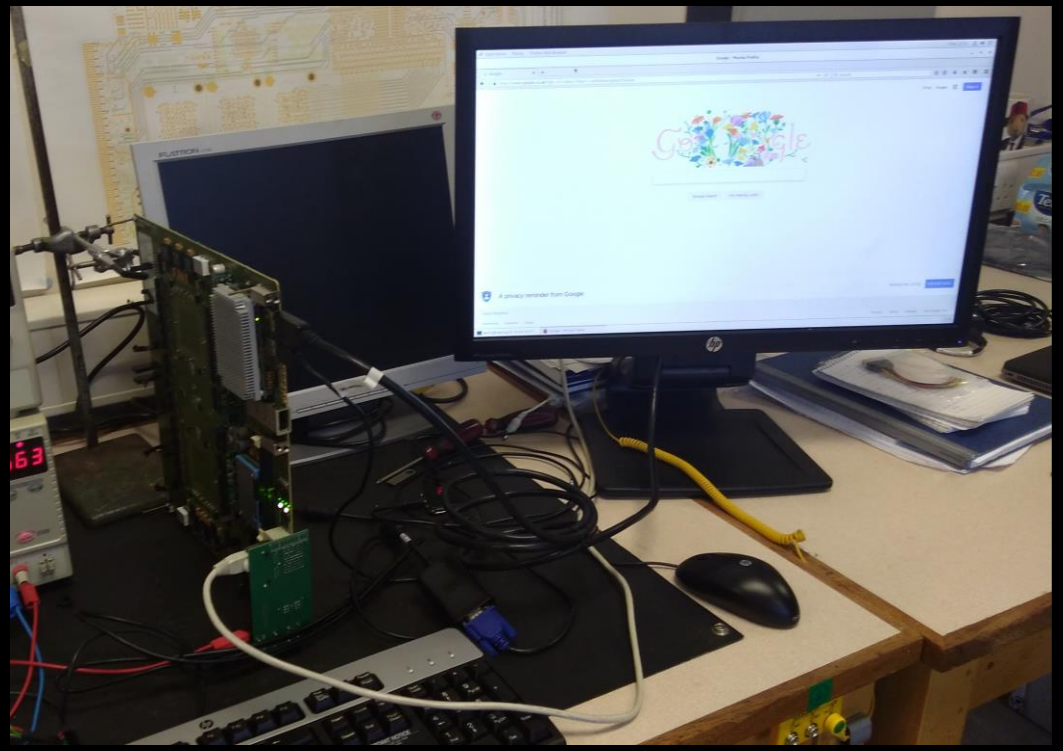
# OFF-CHIP LINK IBERTS




- 120 simultaneous IBERTs
  - 16Gbps
  - 60 through fireflys, 60 over copper
  - Default settings
- With highlighted caveats, each link passed 8e14 bits
  - No errors
- 1board, 85Pb between FPGAs

# CARRIER TESTING

- Com-express & GBE
  - HD display-port, USB, SSD, Centos-7, GBE† all working
- Artix-7 Service FPGA
  - JTAG, Flash, Heartbeat LED, Clocks, PCIe to Com-express all working
- IPbus-over-PCIe
- The existing MP7 I2C-over-Ipbus and JSM JTAG-over-IPbus infrastructure also work unchanged



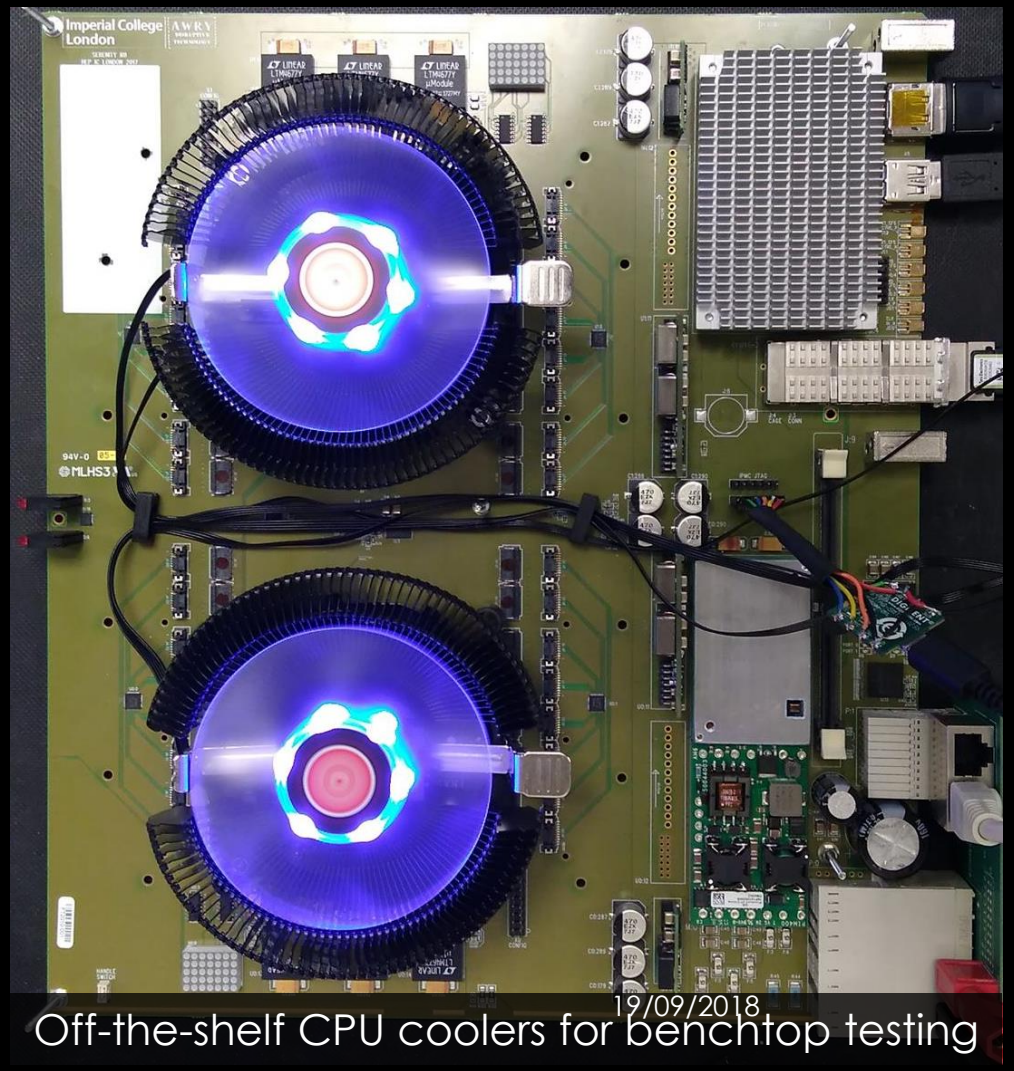


†Com-express→On-board GbE switch→ATCA zone-2 connector→Micro-backplane→RJ45→Outside world

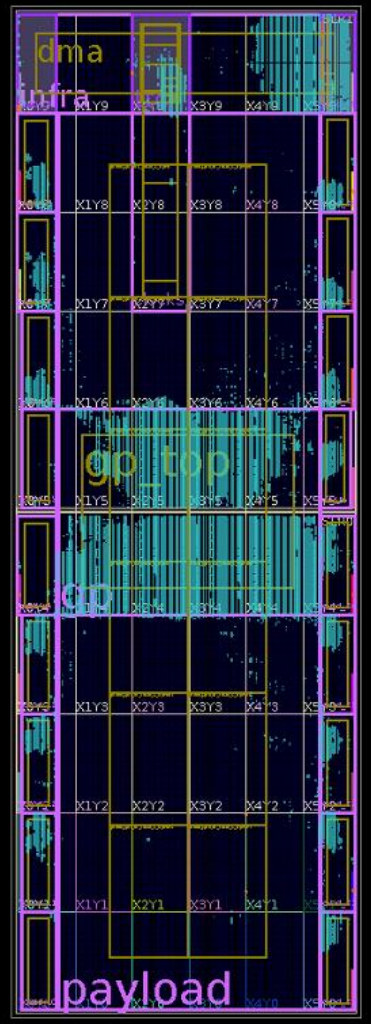
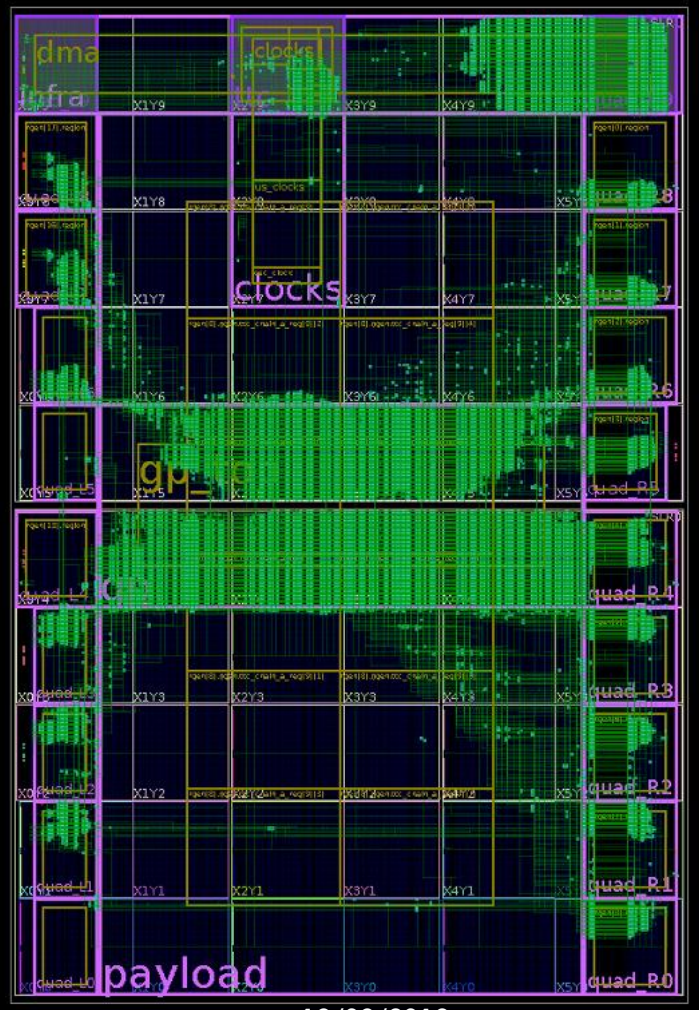
# COMMUNICATION VIA PCIE

- JTAG to both KU115s via the Artix
  - Digilent USB-debugger via Artix GPIO
  - XVC-over-Ipbus-over-PCIe
  - Directly over Ipbus
- Heartbeat LEDs
- Debug LED arrays
- Programming Flash
- PCIe & system clock

```
00:1f.3 SMBus: Intel Corporation Atom Processor E3800 Series SMBus Controller (rev 11)
01:00.0 Serial controller: Xilinx Corporation Device 7021
02:00.0 Serial controller: Xilinx Corporation Device 8031
03:00.0 Serial controller: Xilinx Corporation Device 8031
04:00.0 Ethernet controller: Intel Corporation I210 Gigabit Network Connection (rev 03)
```




Off-the-shelf CPU coolers for benchtop testing

# EMP FIRMWARE: EXAMPLE

- Many large Xilinx parts are now multi-die
  - Super Logic Region
  - How do algorithms developed in MP7 port to UltraScale?
- HT Track Trigger
  - Geometric Processor Component
  - Concern over inter SLR routes
  - Quickly port to KU115
  - VU9P or KU15P also available.
  - Only 20% of inter-die routes used.