Mixed Software/Hardware-based Fault-Tolerance Techniques for Complex COTS System-on-Chip in Radiation Environments

Ivan Rodriguez-Ferrandez[1,2] Leonidas Kosmidis[1,2] Maris Tali[3] David Steenari[3]

[1]Universitat Politécnica de Catalunya

[2]Barcelona Supercomputing Center

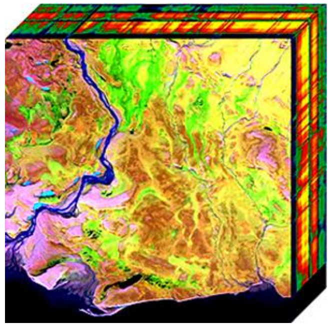
[3]European Space Agency

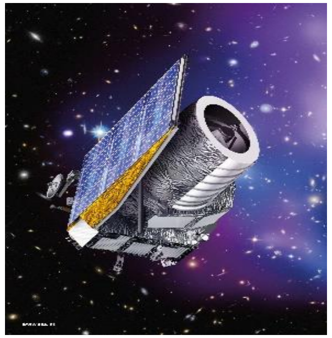
25/03/2025

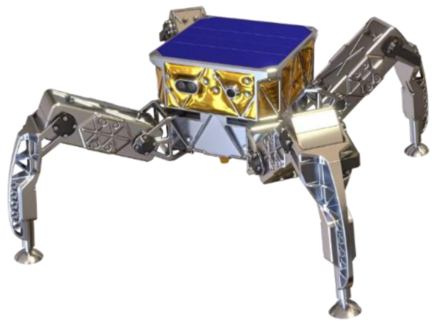
ACCEDE 2025

Background & Motivation

- Space systems are traditionally based on simple processors, with custom designs for space
- Increasing need for on-board computing in space systems
  - More complex tasks to be carried out autonomously
  - Bigger sensor sizes
  - Limited communication bandwidth and increased latency
- New Space domain imposes additional requirements
  - Software-defined satellites (sw not known at launch)
  - Low-cost, frequently based on COTS (i.e. non-space devices)
- Traditionally FPGAs are used when high performance is needed
  - High hardware and development cost
  - Long development lifecycle
- New processing technologies are considered for space







Images courtesy of ESA and SPACEBIT

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Background & Motivation

- Embedded GPUs are among the promising on-board processing technologies for future space systems
  - Ideal for image processing
  - Great performance per watt
  - Mature general purpose programming languages, easier to program than FPGAs
  - Fast software reconfiguration / update
  - High performance embedded ~10/15 Watts
  - Efficient for massively parallel, highly regular computations

NVIDIA Xavier, Image courtesy of NVIDIA

Image of the V1605B of AMD

Background & Motivation

- ESA-funded "Low Power GPU Solutions for High-Performance On-Board Data Processing" (GPU4S – GPU for Space) at BSC.
  - Objective: "**Study the availability, suitability, and applicability for space of existing LPGPU architecture**"
  - Outcomes:
    - Standard benchmarks defined: **OBPMark, OBPMark kernels, OBPMark ML**
    - **proved that GPUs have a very high performance for several on-board processing applications** in spacecraft compared to existing space solutions as well as COTS solutions (multicores).
    - Identified Limitation: Radiation Performance of GPUs
- 2 adoption paths were identified for GPUs in Space:
  - Long Term: Design Rad-Hard System-on-Chip based on GPU IP
  - **Short Term: Upscreen high-performance COTS Embedded GPU SoC components (OSIP PHD Co-funding)**

esa

GPU4S

esa

PhD study on COTS GPUs

METASAT

Rad-hard SoC Design Study

Research Question

- **Research Question:**
  - **Can we use automotive-grade GPU-SoC devices in OBP for space applications – through software and/or hardware FT-methods?**

- **Value proposition of research:**
- Enable the use of highly integrated low-power System-on-Chip (SoC) GPUs processors with accelerators for data processing (video, ML, etc), allowing **software-defined processing** in space systems.
- Provide **first results of radiation effects of key selected components**
- Provide **fault-detection and recovery mechanisms** (through SW and HW) for selected components

Device Selection

Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

Device Selection Benchmark

- OBPMark : Reuse building blocks to complex application scenarios
  - Optical Image processing, SAR image processing, CCSDS compression standards 121 and 122, CCSDS encryption
  - Implemented in C, OpenMP, CUDA and OpenCL
  - Code and data available at http://obpmark.org
  - ESA GPL v3 compatible license
  - Available to everyone, not only ESA member states
- Representative inputs, reference outputs and sequential C versions for validation
- Reuse OBPMark-Kernels/GPU4S Bench building blocks
- Additional benchmarking suite focused on Machine Learning:
  - OBPMark-ML

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

***Implementation contributed by this Thesis**

Device Selection Benchmark Evaluation

- We used OBPMark 1.1 to evaluate different devices and compare them with existing space systems
  - COTS
    - Nvidia Jetson TX2
    - Nvidia Jeton Xavier Industrial – Automotive Grade
    - Nvidia Jeton Xavier NX – Automotive Grade
    - Nvidia Jetson Orin – Automotive Grade
    - Zynq UltraScale+ ZCU 102 – Automotive Grade
    - Freescale LS1046A (ARM multicore)
    - Freescale LX2160A (ARM multicore)
    - AMD V1605B (AMD Embedded GPU)
  - Systems with Space Heritage:
    - iX5-100 – AMD GPU Radiation Tolerant COTS
    - HISAOR (OCE Technologies) – Radiation Tolerant
    - GR740 - Space Processor

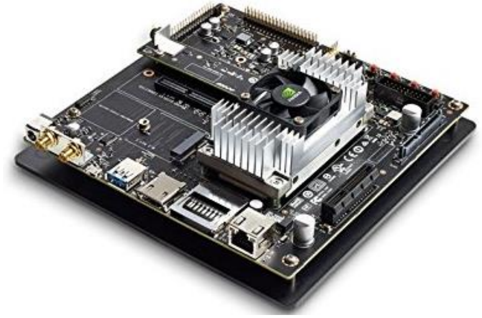
ZCU 102 courtesy of AMD/Xilinx

LX2160A Teledyne ev2

Device Selection Trade-off

- Based on our testing we reduce the selection to four platforms based on GPU performance.
  - NVIDIA TX2
  - NVIDIA Xavier – Automotive Grade
  - NVIDIA Orin – Automotive Grade
  - AMD V1605B
- TX2 discarded due lack of support (ended in 2020).
- In similar way, the AMD V1605B lacks official ROCm support for the use of the GPU
- At the time of the selection, Xavier has a good radiation performance, however Orin offers a more advanced automotive CPU.
- Considering these factors this Thesis focused on both the NVIDIA Xavier and NVIDIA Orin.

Nvidia Xavier NX


Nvidia TX2


Nvidia Orin AGX

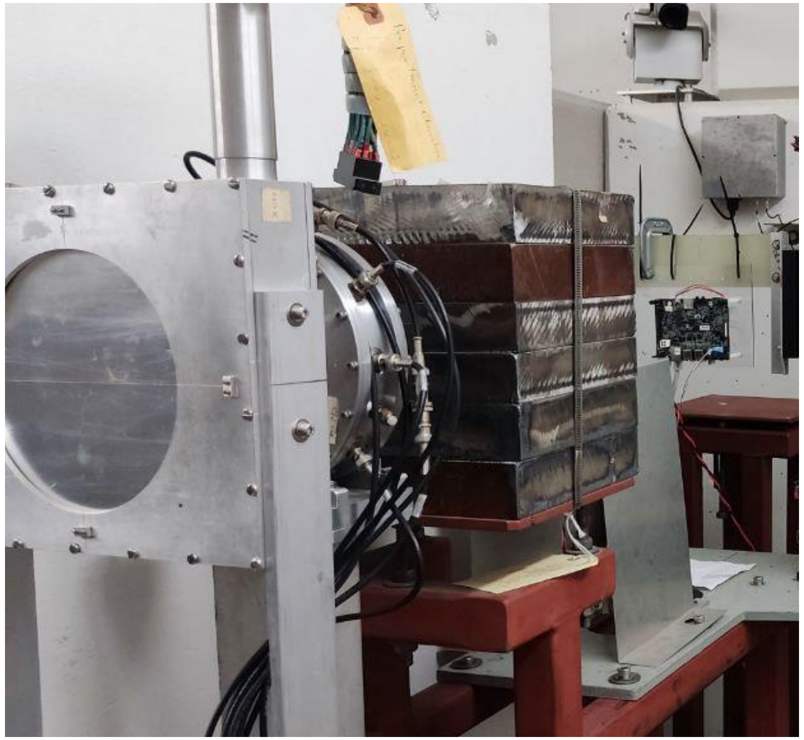
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Device Characterization

Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

Device Characterization

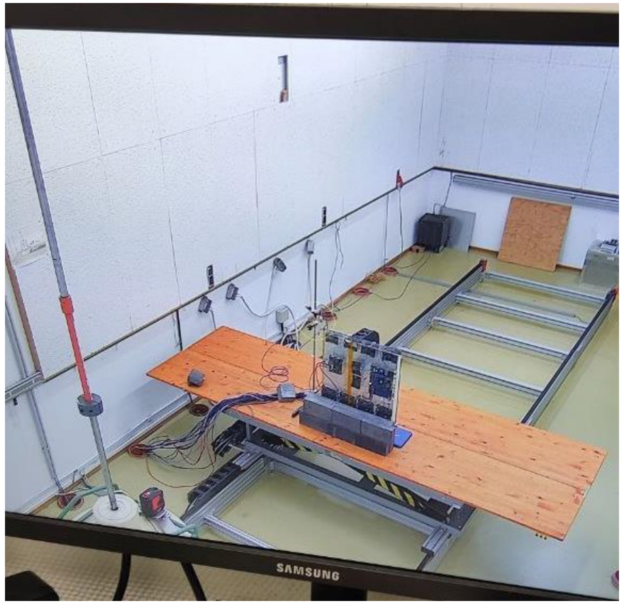
- To understand what will be the effects of the space environment on such device a characterisation campaign was performed.
    - This includes 3 proton tests
    - a TID test
    - a Heavy Ion Test
- All of these were performed to get insides of how the device failure modes happen:
    - SEU, SEFI for both the CPU and GPU subsystems
    - Maximum Irradiation tolerance (TID)
    - Detect possible any possible SEL on the system of module

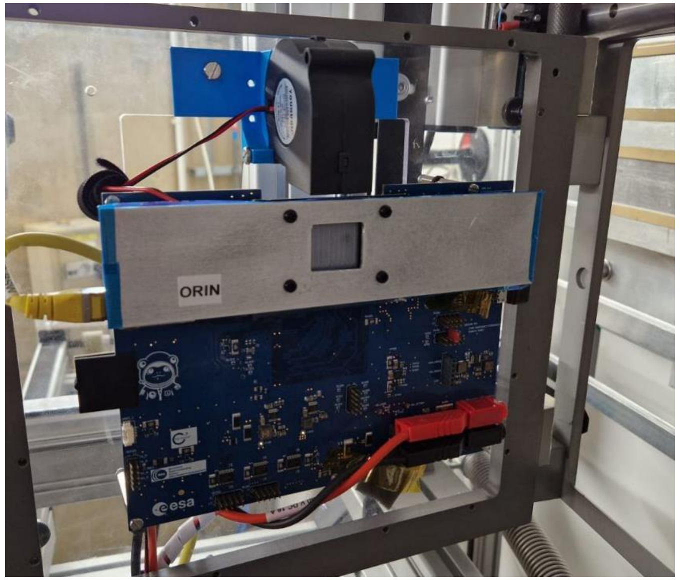


Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Device Characterization

- Proton Test
  - First Two test for evaluation of the NVIDIA Xavier & Orin Faliure modes
  - Third Proton test to evaluate the Software & Hardware mitigations.
- TID test
  - Evaluation of 12 devices to up to 50KRads test
  - 6 Devices unbiased and & biased
  - 1 Board with the hardware mitigations
- Heavy Ion test
  - Up to 35 LET tested
  - Both Xavier & Orin SoC tested

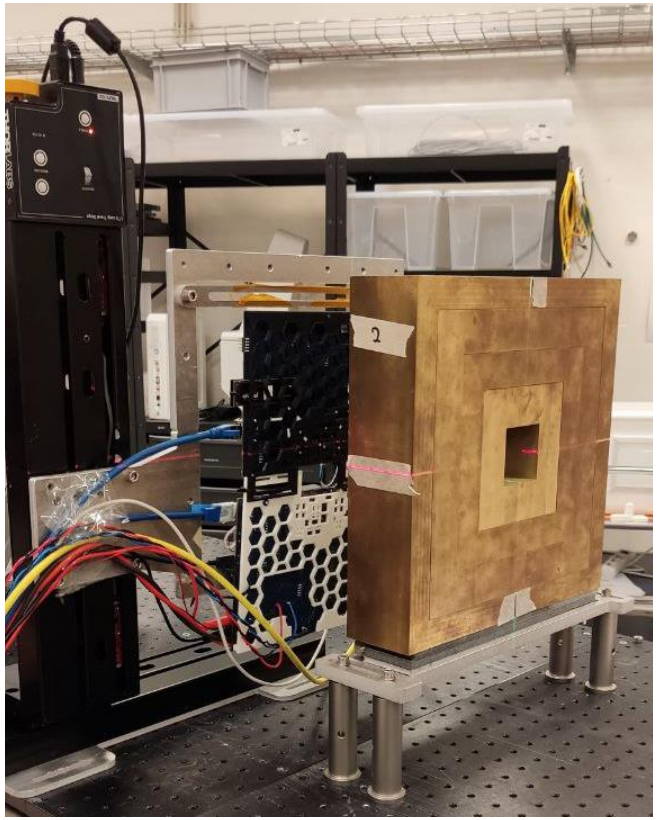




Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Radiation Test Summary

- Summarizing all the Radiation Data we get the following numbers:
  - Xavier
    - SEFI Cross-section (200 Mev): ~10^{-08}
    - SEFI Cross-section at (15 LET): ~ 4×10^{-4}
    - TID Bias Lower limit: 22 Krad
    - TID Bias Upper limit: 24 Krad
    - TID Unbiased limit: 50 Krad
    - SEL free below 15 LET.
    - Non-destructive SEL below 35.5 LET
  - Orin
    - SEFI Cross-section (200 Mev): ~7×10^{-11}
    - SEFI Cross-section at (15 LET): ~ 1×10^{-4}
    - TID Bias Lower limit: 37 Krad
    - TID Bias Upper limit: 39 Krad
    - TID Unbiased limit: 50 Krad
    - SEL free below 30 LET.



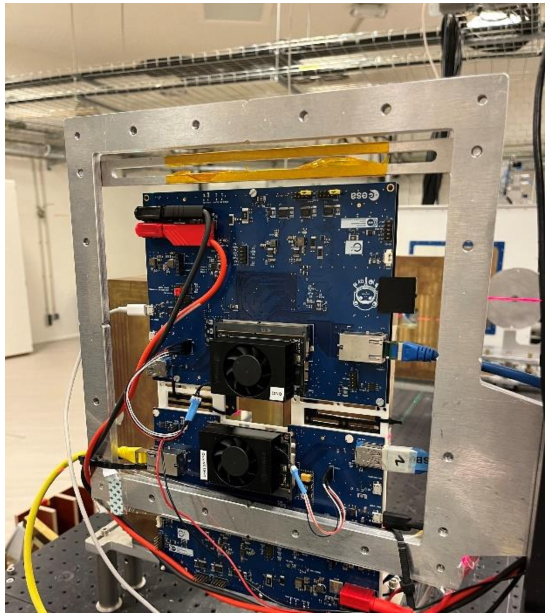
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

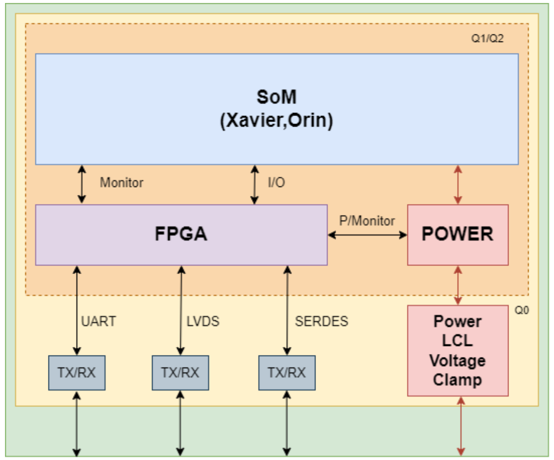
Software & Hardware Mitigations

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Software & Hardware Mitigations

- The software part,
  - more focus on the GPU error mitigations
  - Issues from radiation that stop the execution of the GPU code
- Hardware part
  - Mitigations on the interconnection with the rest of the spacecraft
  - Find Tested COTS devices to be use on the peripherals of the module
    - Boot secuencing, Power, etc

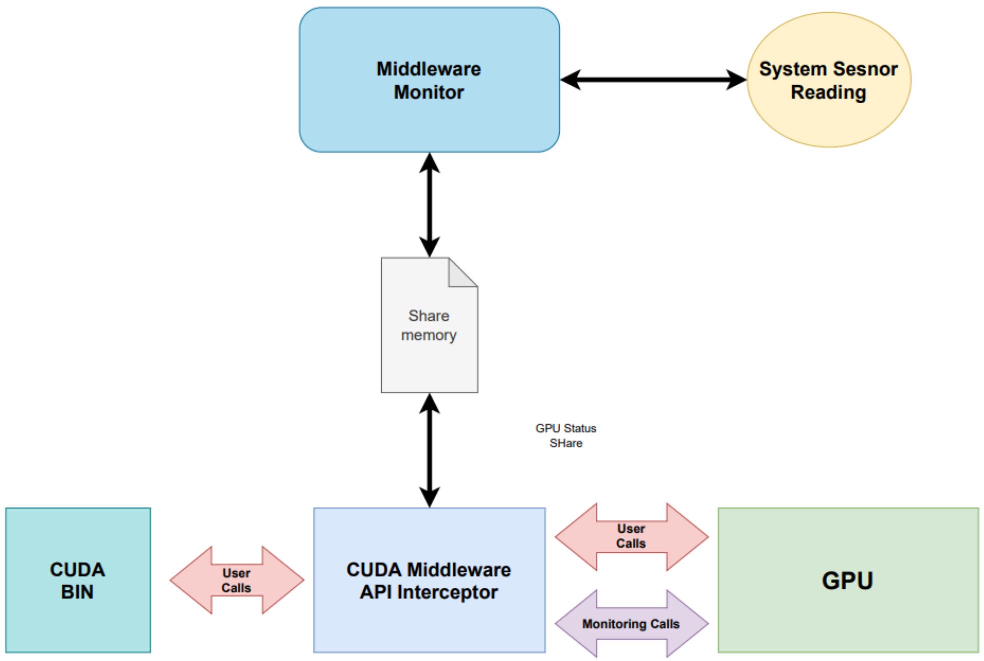




Barcelona Supercomputing Center
Centro Nacional de Supercomputación

SW Mitigations: GPU Middleware

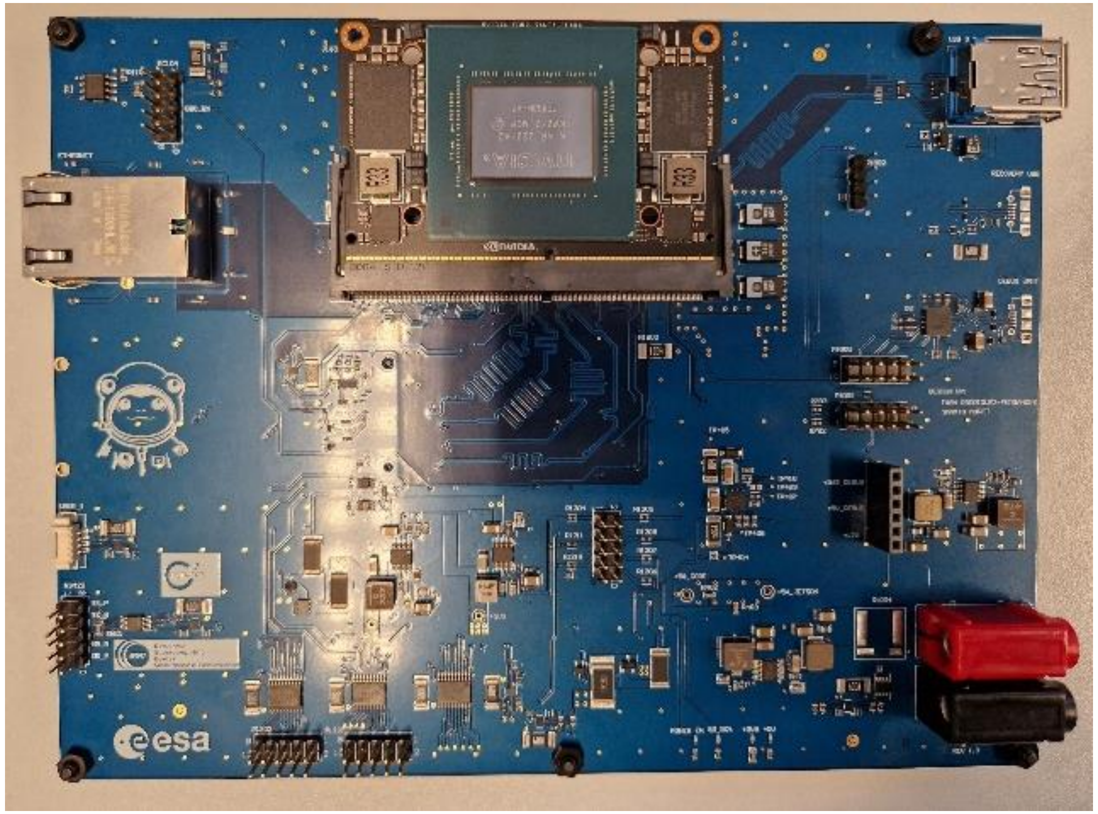
- The other open point was to create a watchdog to monitor the GPU execution
- This is required to detect halts of the GPU due to radiation effects
- This effect has been observed both on persistent kernels as well as on normal kernels
- To minimize code modifications a CUDA call capture approach was chosen based on library preloading to allow any CUDA code to be used without modification
- Using the calls and a state machine to tract the status of the GPU



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

HW Mitigations: COTS Base Carrier Board

- Goals
  - Provide a reliable platform to carry the GPU
  - Isolate the interfaces directly coming from the COTS device
  - Provide a demonstrator to be tested under irradiation conditions
- NVIDIA Jetson Carrier Board
  - Tested up to 50 Krads (biased)
  - Used on Proton and Heavy Ion test to evaluate the critical parts



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Conclusions & Outcomes

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Conclusions & Outcomes

1. General Outcomes
    a) Better understanding of how to do Radiation testing for complex systems
2. Software Outcomes
    a) A GPU middleware was developed for tackling the issue of faults on the GPU
3. Hardware Outcomes
    a) Nvidia Xavier/Orin family Radiation evaluation.
    b) System level protection means for COTS automotive/industrial GPU platforms.
    c) Reference PCB for future use in space missions.

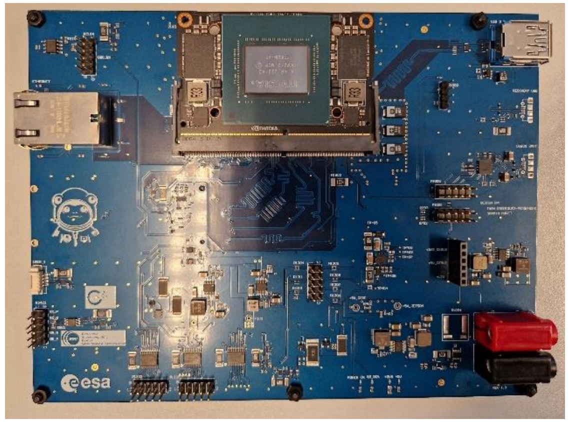
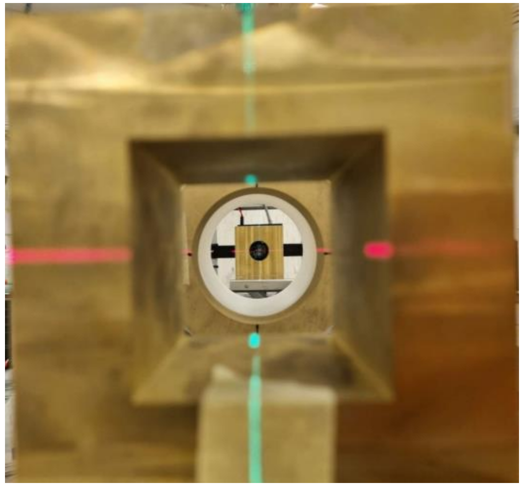
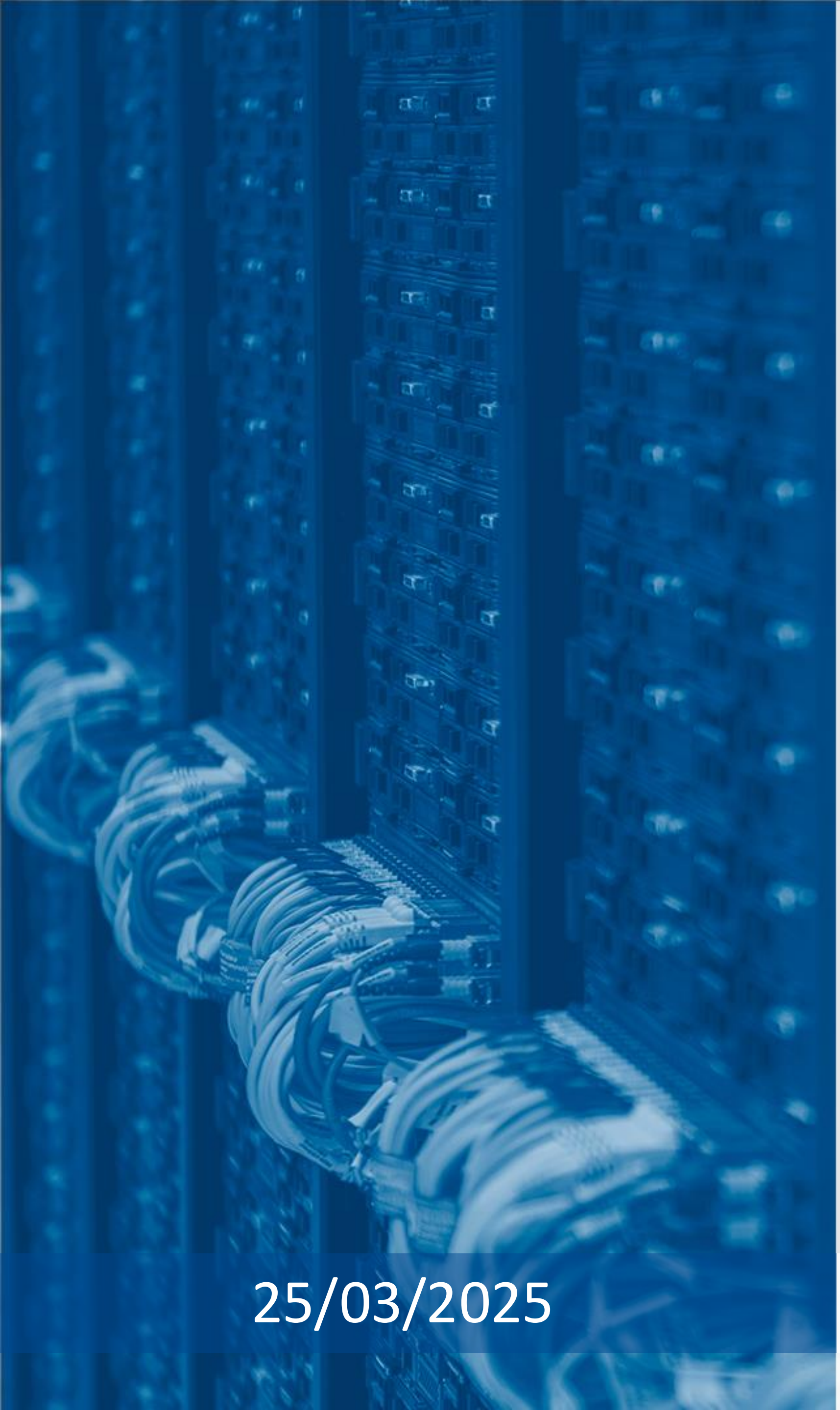


Image of the test PCB



Thank You

Ivan.rodriguez@bsc.es



Mixed Software/Hardware-based Fault-Tolerance Techniques for Complex COTS System-on-Chip in Radiation Environments

Ivan Rodriguez-Ferrandez[1,2] Leonidas Kosmidis[1,2] Maris Tali[3] David Steenari[3]

[1]Universitat Politécnica de Catalunya

[2]Barcelona Supercomputing Center

[3]European Space Agency

25/03/2025

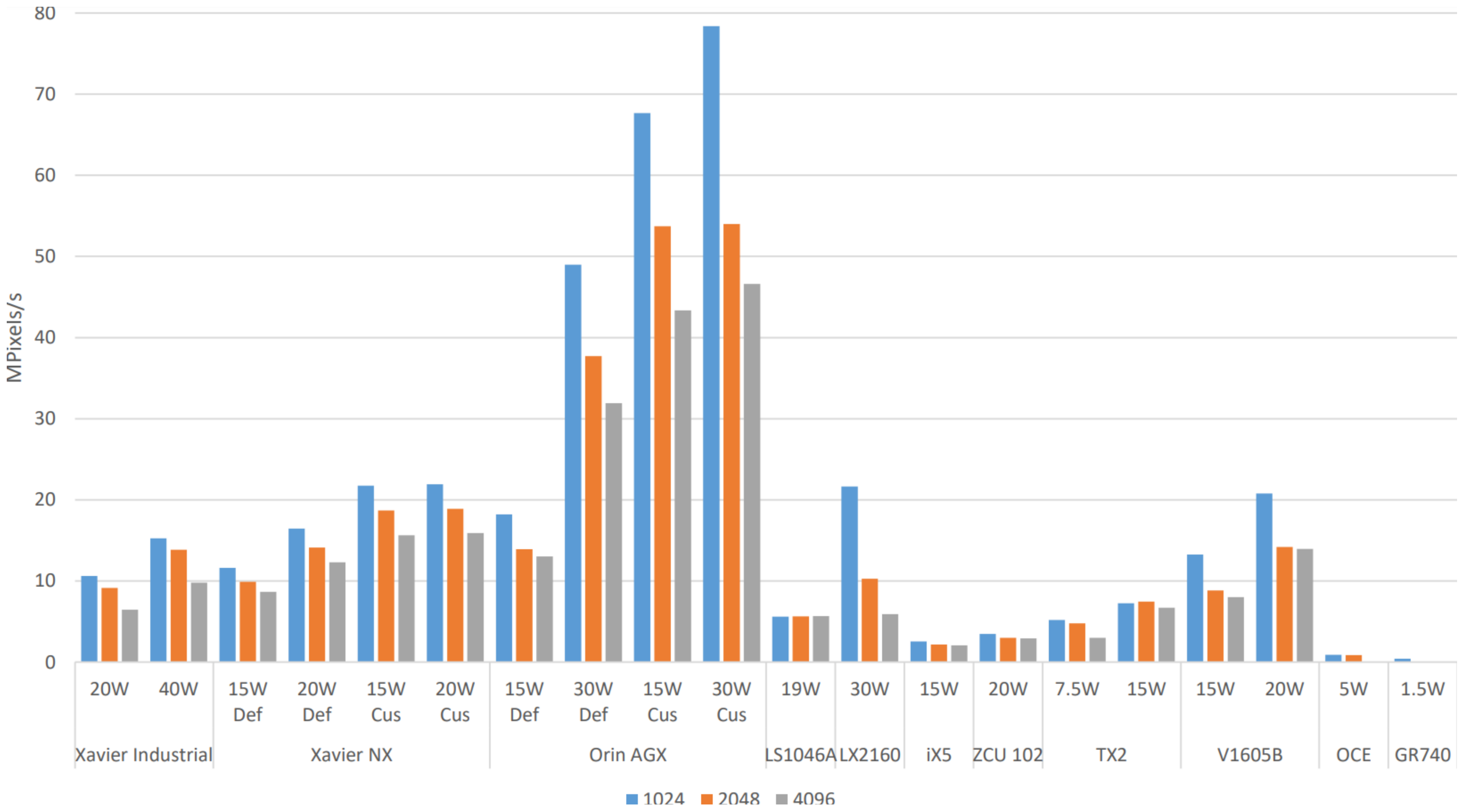
ACCEDE 2025

Research Methodology

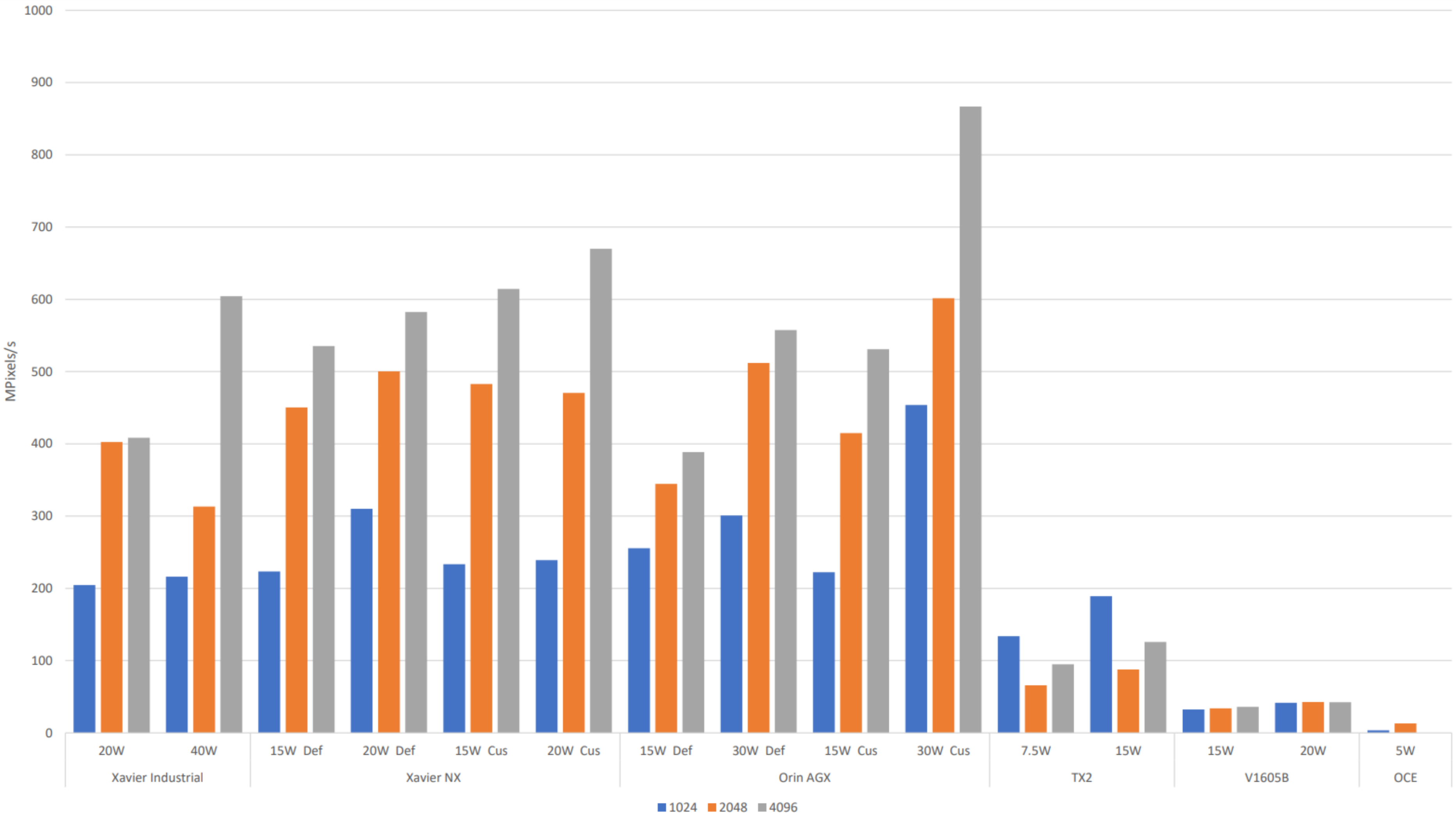
Device Selection & Benchmarking → Device Characterisation → FT Methods Implementation → FT Methods Characterisations → Conclusions

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Device Selection Benchmark Evaluation: Multicore



Device Selection Benchmark Evaluation: GPU



Research Question

- Challenges in using automotive SoC components is multi-disciplinary:
- This class of devices are **minimally tested under radiation**
  - …only what is specified under ISO 26262 without testing of SEL (some of them).
- The **detailed hardware architecture is proprietary** and is most of the time not fully known
- Sometimes **reliance on closed-source drivers/libraries (like CUDA) and complex high-level OSes (Linux)**

- **Therefore, we are focusing on the following technical tasks to meet the challenges:**

1. Computational performance benchmarks (through OBPMark) – *to select the best performing components*
2. Radiation assessment and mitigations – *to increase availability in flight*
3. Hardware integration
4. Software integration

Barcelona Supercomputing Center
Centro Nacional de Supercomputación