NEPP Processor Enclave:
A Trek Towards Technology Road Mapping

Edward Wyrwas
SSAI, Inc supporting NASA Goddard Space Flight Center, Greenbelt, MD, USA

Acronyms

- Artificial Intelligence Machine Learning (AI/ML)
- Advanced Micro Devices, Inc. (AMD)
- American Society of Mechanical Engineers (ASME)
- Comma Separated Values (CSV)
- Computer Vision (CV)
- Defense Enterprise Common Parts Testing in Radiation (DECPTR)
- Devices Under Test (DUT)
- Error Code Correction (ECC)
- Extravehicular Mobility Unit (EMU)
- Frequently Asked Questions (FAQ)
- GlobalFoundries (GF)
- Graphics Double Data Rate Synchronous Dynamic Random Access Memory (GDDR)
- Graphics Processing Unit (GPU)
- Hypertext Markup Language (HTML)
- High Temperature Operating Life (HTOL)
- International Organization for Standardization (ISO)
- California Institute of Technology Jet Propulsion Laboratory (JPL)
- Key Performance Parameters (KPP)
- Linear Energy Transfer (LET)
- Missile Defense Agency (MDA)
- Mean Time Between Failures (MTBF)
- NASA Electronic Part and Packaging Program (NEPP)
- Natural Language Processor (NLP)
- Neural Processing Unit (NPU)
  - In this presentation, NPU refers to *any* standalone device or IP block that performs machine learning functions such as an Ai accelerator, deep learning processor, or neuromorphic processor
- Original Development Manufacturer (ODM)
- Printed Circuit Board Assembly (PCBA)
- Portable Document Format (PDF)
- Single Event Effects (SEE)
- Single Event Latchup (SEL)
- Single Event Upset (SEU)
- System on Chip (SOC)
- To Be Determined (TBD)
- Exploration Extravehicular Activity Services (XEVAS)

Agenda

- Purpose
- Data Scraping Existing Processor Radiation Data
- Survey of Scraped Datasets and Gap Analysis
- Related Efforts and Collaborations
- Future work
- Summary

Purpose

- Create a cyclic approach to manage our reduction of resources and yield the best bang-for-the-buck for each processor enclave test
  - Data Aggregation through Marketplace Analysis and Working Group surveys
  - Synchronizing Test Planning within the radiation effects community
  - Standardize Tests at the Top-Level (Design/Plans/Performance/Reporting)

- Test Campaigns can therefore bridge institutions and eliminate double-spend
  - Allows us to leverage everyone's data to complement or supplement our own tests, reducing or eliminating duplication altogether
  - Permits easier supply-chain navigation, procurements, and obsolescence planning

Data Aggregation to Minimize Duplication

- The inclusion of relevant radiation response data in the parts library will help the overall test community and “lift all boats”.
  - Reducing duplicative and unnecessary testing is a key tool in managing scarce beam hours and scarce funds to procure testing equipment and DUTs.
- Successful parts testing by one user means that the next user might choose to use that same part for another design — and not need to test it.
  - The government has a strong interest in collecting data to optimize and allocate highly constrained resources.
- Further, this provides opportunities for collaboration from the get-go, and offers avenues for long term sequential test campaigns.

This is a paraphrase from the MDA DECPTR FAQ

Why NASA Should Care

- Missions have been told to prioritize and optimize mass and volume allocations
  - Gateway discussions have indicated there will be reduced thermal dissipation and requires low power, low thermal load on subsystems
  - EMU/XEVAS have substantial constraints in power, mass, and form factor which is driving unique form factors for displays, which require graphics processor devices of the same constraints
  - Lunar Surface Habitats and Mars Vehicles electronics requirements have severe mass limitations and very long radiation exposure, but will require more complex devices to meet mission parameters
  - Artemis programs have indicated low power, high compute capability requirements

- COTS devices are becoming more prevalent in vendor subassemblies and our own designs. Additional SEE testing of product and reporting is required, yet funding and timelines will never permit comprehensive test campaigns.

Data Aggregation

- **Problem:** Sparse data exists across many sources and there are limited experts to assess, digest, and report back the value.

  Can we aggregate it?

- Technology Landscape Monitoring
  - Independent news
  - Technology journals
  - Conference papers
- **Solution:** Create a Rapid Analyses Framework using both recent and historical information pertaining to Processor Enclave technologies for both automated and manual analyses.

Data Source Identification

↓

Web Agent Scraper

↓

NLP Engine (keywords, summary)

↓

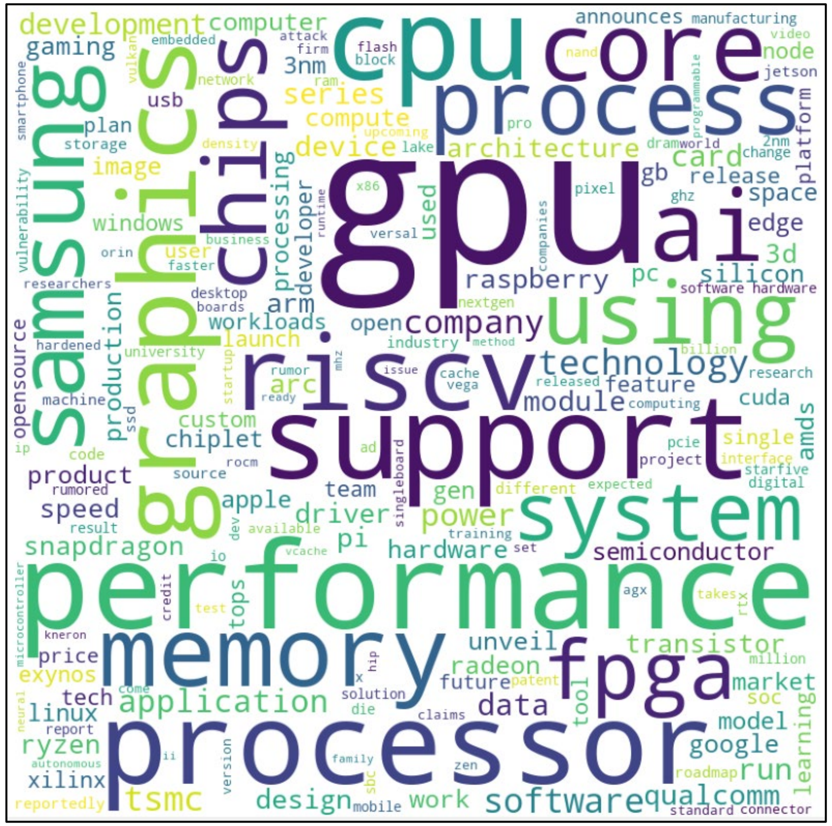
CSV File Creation

↓

Filters of Useless Words

↓

Analyses:
- Word Clouds
- Keyword Trends
- Test and Evaluation Preparedness Timelines



Availability of Community Data

- **Goal:** Track custodians of both ongoing and historical radiation and reliability tests pertaining to Processor Enclave data.

- **How:** By providing access to those providing data to the community
  - Information submitted by Processor Enclave Group members since FY22
  - Captured high-level information on over 80 devices

- **Example:** NVIDIA JETSON TX2
  - NEPP did its proton tests in 2018-19
  - NEPP guided dozens of OGA and industry through additional rounds of testing and deployments from 2019-23
  - In early 2024, another test was conducted by a team at the University of Trento, Italy (with NEPP and JPL providing support) to explore more AIML types of testing

Schema for a Limited Data Sampling

- Device info
- Test campaign(s), facility and beam-tune info
- URL to attachments hosted elsewhere (as copyright rules allow)
- A *mailto: URL* to contact the data custodian
  - Because every radiation community player has different relationships and different distribution levels
- Dynamic data plots (i.e., LET vs Cross Section)

Sample Results Snapshot for AMD Radeon e9173

Device Name	**AMD Radeon e9173**
Processor type	GPU
GPU Processor type	GCN 4.0 Lexa
Bus Width (bits)	64 bit
Power (Watts)	35 W
Operating Voltage (V)	12V
Radiation Hardness (TD, DR, n, SEE, nSEE)*	SEU: characterized SEE: characterized SEL: > 6 MeV-cm^2/mg
Supply Chain	piece part
Technology Interest	14nm GF
Data Custodian	NEPP (Edward Wyrwas) – edward.j.wyrwas@nasa.gov
URL for publications	https://ntrs.nasa.gov/citations/20190031855 https://ntrs.nasa.gov/citations/20230009291
URL for references	https://github.com/golmschenk/ramjet/tree/experimental-gpu-configuration-simplified-infer
Other specific needs?	This is a low power GPU candidate for Artemis generation missions

Code or Library compatibility?	Vulkan, OpenGL, TensorFlow-ROCm version 2.3, OpenCL
previous missions that used this (if able to share)	n/a
any triage data from flight?	n/a
was it rebooted?	See test reports
were there errors?	See test reports
how were the errors captured?	See test reports
what was the powered-on or operational duty cycle?	Graphics, Math, AI/ML
power scalability? Can it operate at 3V when 20V nominal (as example)	12V only
documentation or workability with manufacturer/ODM	Potential, though EOL expected
is there a mitigated design for the device?	n/a
are there host or enclave chips required to operate?	Yes, microprocessor, PCIE
long term reliability? Test structure based or lot testing based? HTOL-based MTBF?	TBD
failure mechanisms applicable on the product level?	TBD
memory footprint definition - cache vs ram vs rom (ECC option?)	2GB GDDR5

AMD Radeon e9173 Graphics Processor

Data Aggregation Results

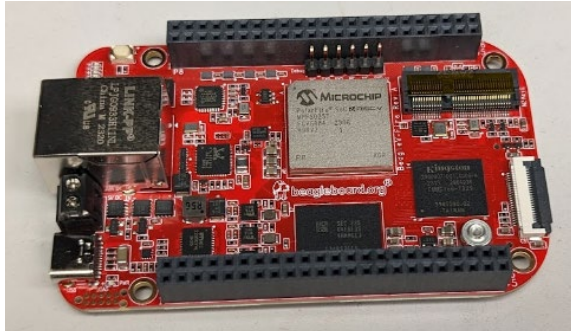
- These Key Performance Parameters (KPP) permit NASA to track State of the Art technologies.
  - Processor fabrication technologies (realistically <16nm)
  - Low Power NPU devices (>1 TOPS/W)
  - New cores: RISC-V and ARM (e.g., A72, M23, M4)
  - Automotive & ISO-26262 grade with fault handling and error correction
  - Graphics Rendering cores with extensive software framework capabilities (i.e., Vulkan) that can demonstrate real-time video streams with overlays

- The COTS piece-part devices we are presently testing are representative of these KPP. They enable NASA's future missions by performing tasks faster with less power. The technologies represented should have more inherent reliability as compared to their predecessors (e.g., design refinements to correct errata, process maturity, additional features like ECC) or competition (e.g., x86, PowerPC).
  - EdgeCortix Sakura (NPU)
  - Microchip PolarFire Icicle (RISC-V)
  - Texas Instruments TDA4VM (ARM A72, ISO-26262)

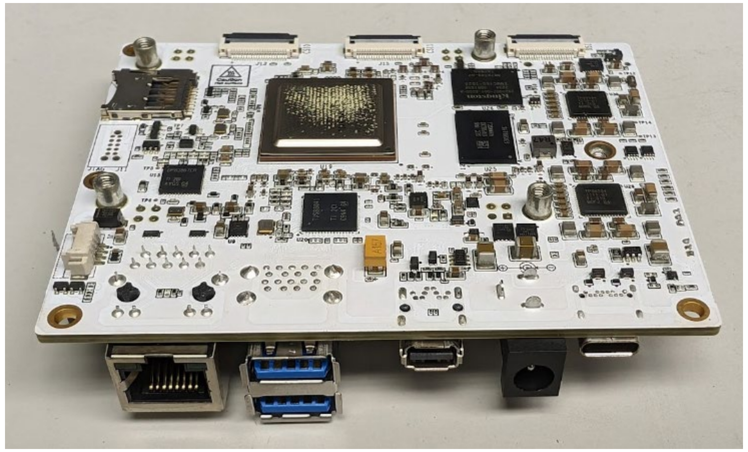
Present Work

- Synergistic testing using personnel at JPL and GSFC
  - Texas Instruments TDA4VM (Dual ARM 64-bit Cortex A72)
    - High reliability, Automotive grade, Fault Tolerant
  - Microchip PolarFire Icicle (RISC-V E51 + U54)
    - Could be Radiation Hardened

- Rapid Roll-Out of Test
  - Reuse NEPP and JPL test bench software
  - Explore OS loaded-driver dependencies
  - 3D Printed fixtures for easy PCBA swap
  - Standard footprint headers for easy integration of peripherals
  - Both test vehicles have networking support



BeagleV-Fire Single Board Computer

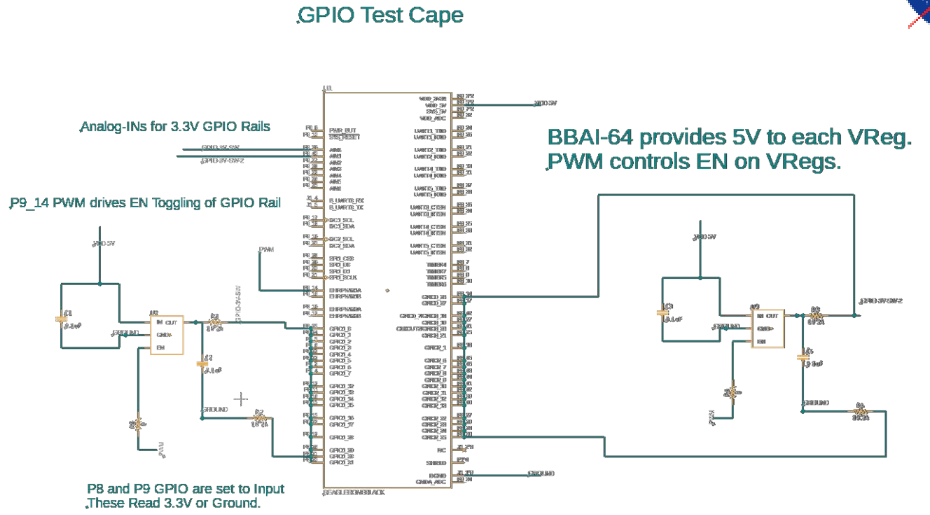


BeagleBone AI-64 Single Board Computer

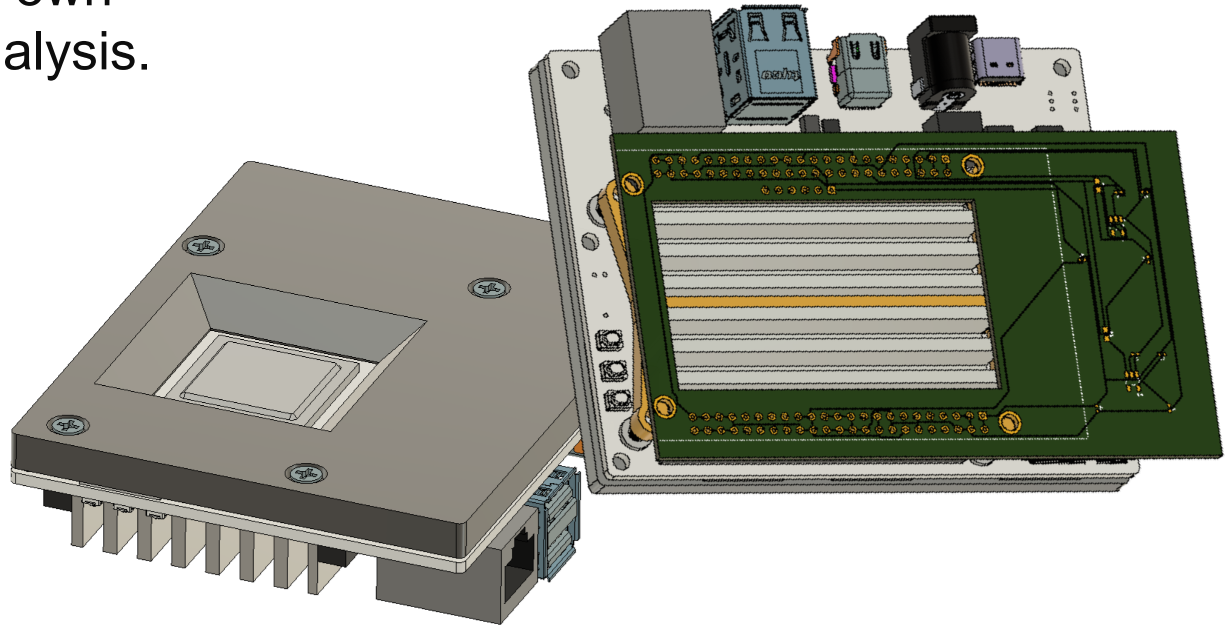
Simple Loopback Tests with Hardware

Expanding our test suite to do more than a singular mathematics routine or AIML program

- Using in-house resources, we can create test specific daughter cards to add circuitry to the COTS board (e.g., voltage regulators, watchdog microcontroller)
- The device is configured to send and receive its own signals permitting easier failure mode effects analysis.



- Test Circuit for BeagleBone Single Board Computer:
  - Hardware Pin Mapping for PWM, Analog Inputs, GPIO set to Input
  - Voltage regulators with current limit to reduce latch-up condition
  - Daisy-chained GPIO which senses a square wave created by the PWM on the DUT
- New heatsinks were fabricated to reconfigure thermal dissipation paths and expose the device to the radiation source



BeagleBone AI-64 Single Board Computer

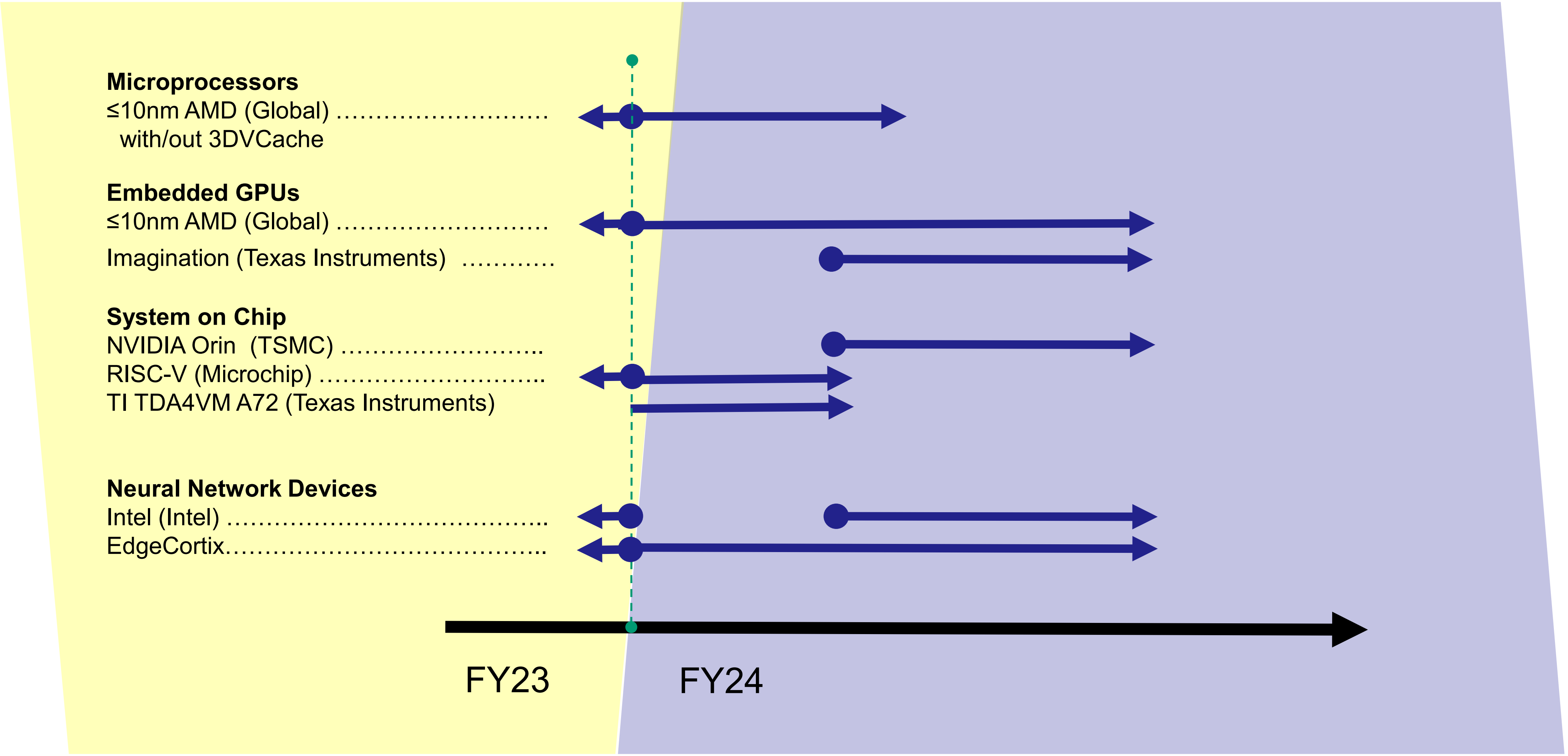
Collaborations

- Aerospace Corporation – AMD, NVIDIA, NPU
- ASME Additive Manufacturing working groups
- Boeing Inc – NVIDIA
- Edge Cortix – Low Power NPU
- Flex Logic – Low Power Embedded NPU
- Johns Hopkins University – AMD, NVIDIA
- Kupros Inc – 3D printable thermal and electrical conductors
- Orbits Edge – Board-level shielding for compute
- NASA ARC – Flight scenarios, CV application
- NASA Avionics Architecture Community of Practice
- NASA GSFC – Generative Design Team
- NASA JSC – Displays, Graphics rendering application
- Troxel Aerospace – AMD, NVIDIA, Middleware
- University of Georgia – NVIDIA TX2i
- University of Trento, Italy (Paolo Rech) – NVIDIA TX2

Processor Enclave telecon

- 1st and 3rd Fridays each month @ 1100 eastern
- Contact edward.j.wyrwas@nasa.gov

Future Work


Microprocessors
≤10nm AMD (Global) ……………………….
with/out 3DVCache
Embedded GPUs
≤10nm AMD (Global) ……………………….
Imagination (Texas Instruments) …………
System on Chip
NVIDIA Orin (TSMC) ……………………..
RISC-V (Microchip) ………………………..
TI TDA4VM A72 (Texas Instruments)
Neural Network Devices
Intel (Intel) …………………………………..
EdgeCortix…………………………………..
FY23
FY24

Summary

- With a reduction of funding and other resources, data sharing has become prominent across institutions to reduce gaps. However, the integrity and structure of that data is paramount to make it usable across different radiation test campaigns.
- A three-part cyclic approach is necessary to assure data can be aggregated, thus allowing test campaigns to bridge institutions and eliminate double-spend.
    1. Aggregating data
    2. Synchronizing test planning
    3. Standardizing tests at the top-level
- Utilizing this approach allows us to leverage everyone's data to complement or supplement our own tests, reducing or eliminating duplication altogether, and permitting comprehensive testing to take place.

End

This work has been sponsored by

- NASA Electronic Parts and Packaging (NEPP) Program
- NASA Office of Safety & Mission Assurance

Thanks is given to the NASA Goddard Space Flight Center's Radiation Effects and Analysis Group (REAG) for their technical assistance and support.