SAND2021-12324O

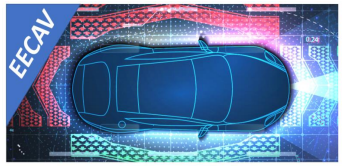
Technical Area II: Chips: Architecture, Safety and Security

Moderator:
Rob Aitken
(Arm Research)

Organizing Team Members for TAII
Rob Aitken (Arm), Jace Mogill (USCAR), Yorie Nakahira (CMU), Craig Vineyard (Sandia)

Workshop on Energy Efficient Computing for Automated Vehicles (EECAV)

May 11 - 12, 2021



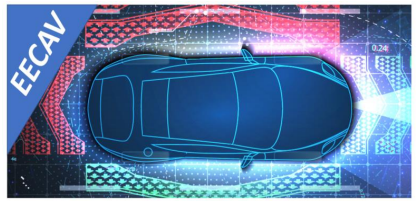
arm

Sandia National Laboratories

Carnegie Mellon University

USCAR
UNITED STATES COUNCIL FOR AUTOMOTIVE RESEARCH LLC

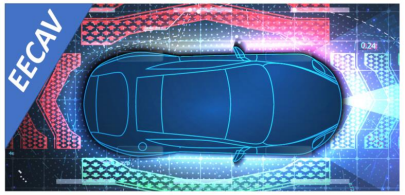
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525



Technical Area II:
Architecture, Safety and Security

Why This Area?:

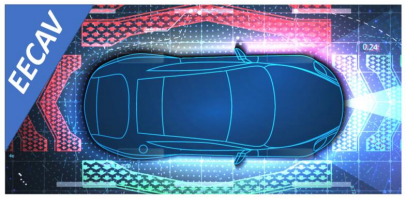
The computational power needed for autonomous driving, and the technology needed to compute with high energy efficiency, starts with the Chip. However, the chips comprising a CAV system must exist within an architecture that provides not only a structure but also an anchor point for software and for communication within and outside the system itself. This architecture needs to meet the power, performance and physical constraints imposed by the vehicle, but it must also be designed from the ground up to incorporate, safety, security and foster energy efficiency. Technical Area II focusses on the computer architecture for energy-efficient AV computation, with a focus on Safety and Security.



EECAV R&D Challenges Identified for Technical Area II

1. How do we best define and use **distributed, heterogenous multiprocessor systems** (CPUs, GPUs, Neural accelerators, etc.) to support the algorithms needed for AV computation?
2. What is the right **network/interconnect** for the above? Bandwidth limitations, security considerations
3. How much **memory** (addressable and storage) at what bandwidth is needed/possible, and where is it located in the system?
4. **Thermal/energy management** (architectural, not physical) to control the total in-cabin computing power of 350W or less?
5. How to maintain **reliable computer performance** with the wide thermal range of the computer in vehicles (AEC-Q100 Grades 1, 3)
6. What are novel opportunities for optimizing **functional safety of digital logic**? Fail operational vs fail safe
7. How does **security** fit into computation for highly automated driving? What hardware support needs to be there? Quantum-safe crypto, threat model, perception errors, OTA update vs on-car. Homomorphic encryption too far out there
8. How to retain **reconfigurability** while improving energy efficiency? (can be varying levels of reconfigurability – neural engine weights vs embedded FPGA)
9. How to decouple image signal processing from the machine learning process? Humans process the imaging in a completely different way from the pixelated frame images and computer vision/ML approach. The key computing problem for AV is handling the tasks of "**computer vision and ML**" (including perception)
10. Can we build an **analog processing system** that through redundancy, tuning, and/or algorithmic error tolerance can (probably) approach or achieve ASIL B safety levels?
11. What are the **ideal sensors** to use for AV? What part of the electromagnetic spectrum should the sensors be based on so as to minimize the computational "Demand" for AV? How to interpret images in order to minimize required computational resources.
12. Where do **advanced packaging techniques** fit in?

Did we miss anything?

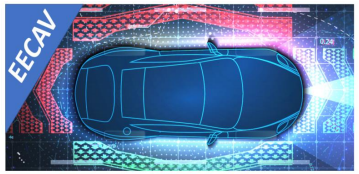


Tech Area II: Detailed Feedback on Specific R&D Problems

R&D Problem 1:

"How do we best define and use distributed, heterogenous multiprocessor systems (CPUs, GPUs, Neural accelerators, etc.) to support the algorithms needed for EECAV?"

- Architectures focusing on bringing raw data to a centralized compute area, even with smart sensors.
- Need to satisfy system constraints including
  - Functional safety
  - High energy efficiency, high performance, high throughput
  - Reliability, robustness, resilience, fault tolerance, security
  - Reprogrammability, upgradeabilty, reusability across multiple vehicles
- How high a priority is R&D funding for heterogeneous distributed systems?
- What are key challenges?
- Have we missed anything specific in this area?



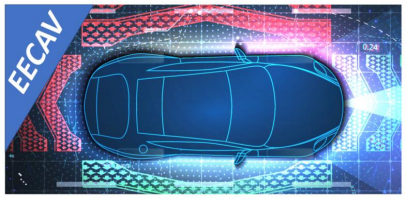
Tech Area II: Detailed Feedback on Specific R&D Problems

R&D Problem #2

"What is the right network/interconnect for energy-efficient computation for AV?"

- What are bandwidth needs and limitations, on-chip, in-car and off-car?
- How do existing protocols/standards fit?
- How will security be maintained, monitored and enforced?
- How will functional safety concerns be ensured? When does the system need to fail safe vs fail operational?

- Do you agree networking and interconnect is a high priority area for R&D funding?
- What are key challenges?
- Have we missed anything specific in this area?

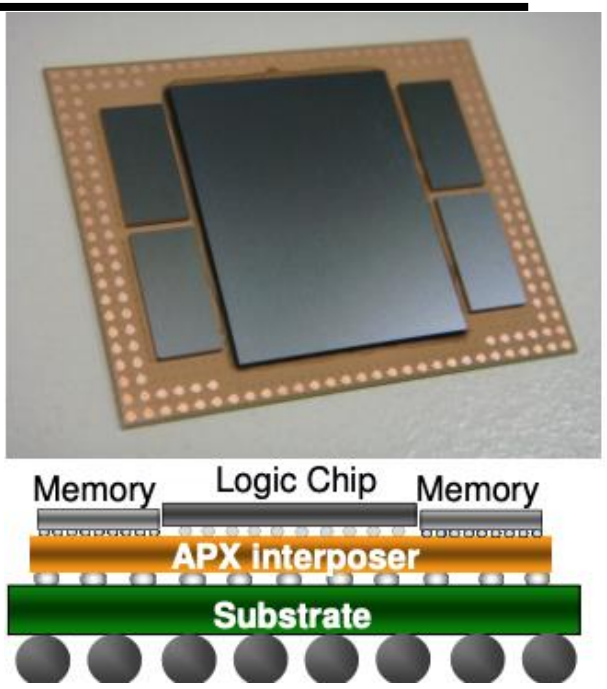


Tech Area II: Detailed Feedback on Specific R&D Problems

R&D Problem 3:

"How much memory (addressable and storage) at what bandwidth is needed/possible, and where should it be located in the system?"

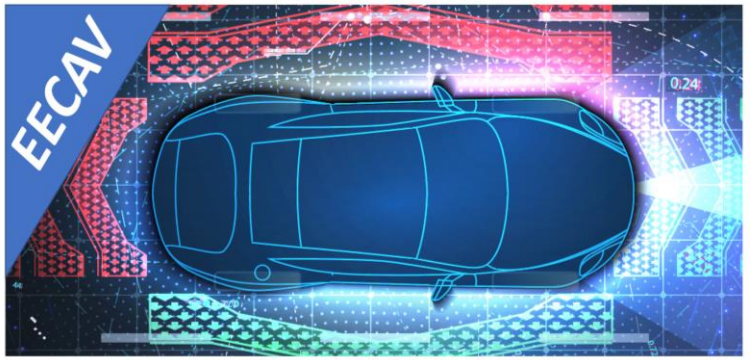
- How do approaches like HBM (High Bandwidth Memory) fit?
- How much coherence is needed/required across the system?
- How might compute in/near memory be used?
- Where could non-Von Neumann approaches help?
- Safety/security phrased in the form of a question

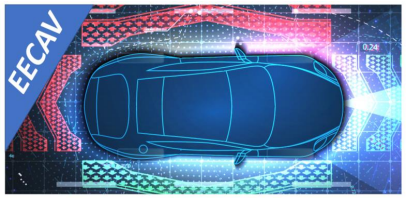


Source: Kyocera

- How high a priority should be placed on memory systems when considering AV computation R&D funding?
- What are key challenges?
- Have we missed anything specific in this area?

Thank You for Your Feedback on Tech Area II!





Tech Area II: Detailed Feedback on Specific R&D Problems

Backup question:

"What is the role of V2X and X2V communication in CAV? Can/should significant computation be offloaded from the vehicle?"

- System-centric vs car-centric view
- Would such a system be more like air-traffic control or stop lights?
- What commonalities would (need to) exist between jurisdictions?

- Should this topic even be considered when considering AV computation R&D funding?
- What are key challenges?
- Have we missed anything specific in this area?