Hybrid ADAS Development Environment for Architecture-specific Performance Estimation

Akira Takeda, Yuji Ishikawa, Tatsuya Mori,
Takeshi Kodaka, Yuji Okuda, and Takashi Yoshikawa

Toshiba Electronic Devices & Storage Corporation

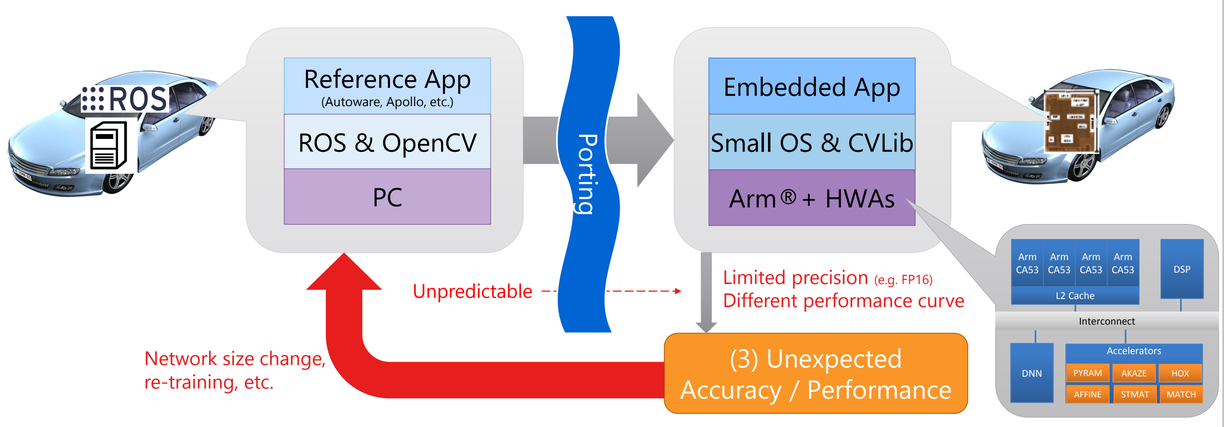
TOSHIBA

DESIGN AUTOMATION CONFERENCE 56

Motivation

(1) Prototype: running in real time on real vehicle

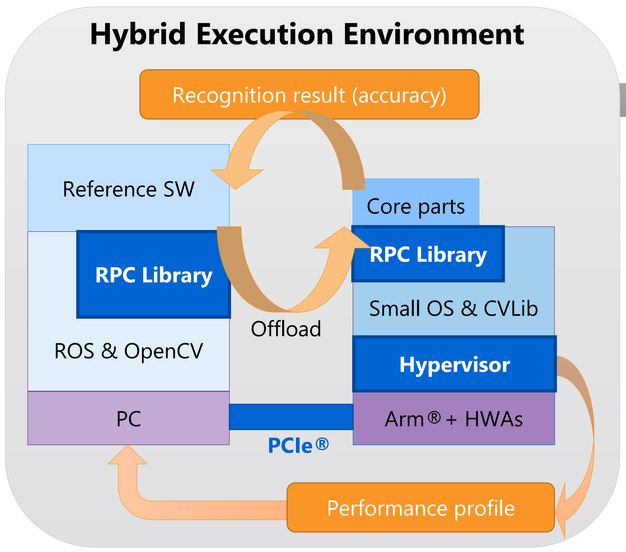
(2) Production: high efficient, high quality



Rich PC-based environment with capability of:
- **early estimation of accuracy and performance** on target SoC
  - **without porting** of overall application
- **real-time execution** (on real vehicle is possible)

Our Solution

"Rich PC" x "Profilable SoC" Hybrid achieves early estimation



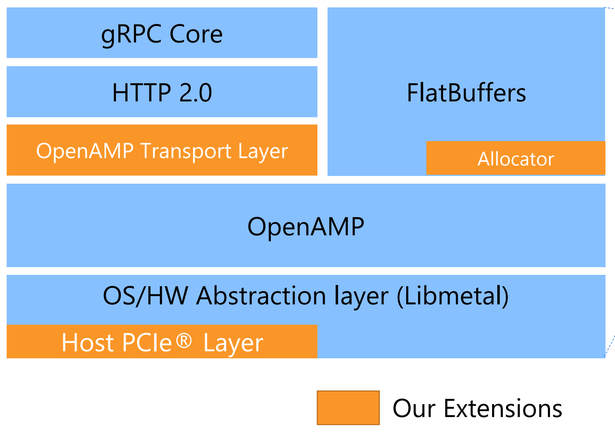
RPC Library
- **Fast** enough for real-time execution
- **Offload only core parts** of application
- Quick definition of core part I/Fs

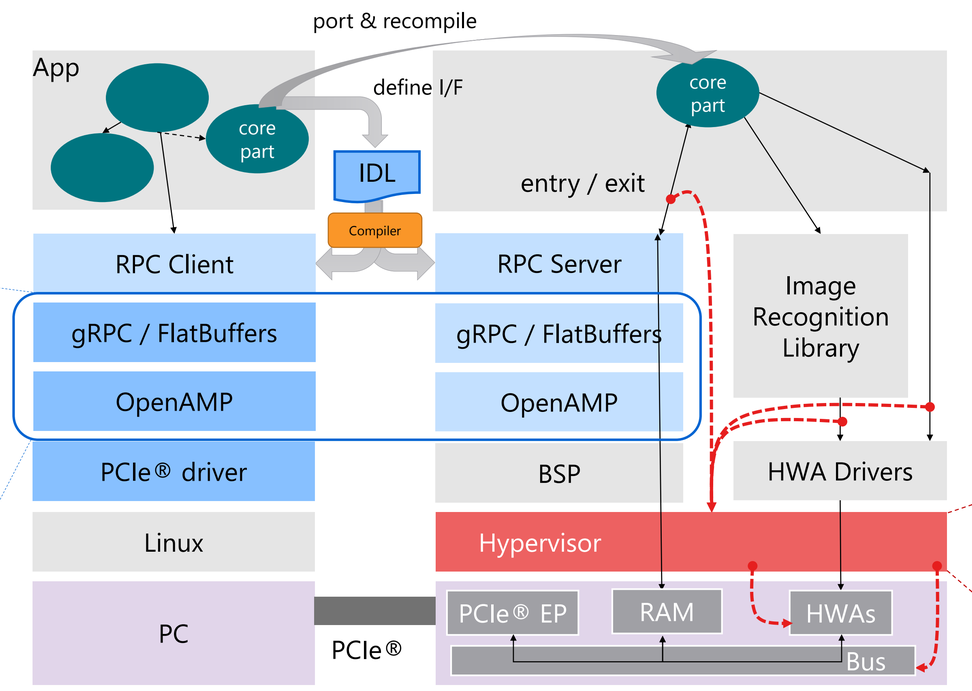
Hypervisor
- Specialized in **profiling** and **monitoring**
- Event-based & sampling-based profiling of SW functions, memory/bus, accelerators
- Error Detection/Trap

Software architecture: Two key software components

Zero-copy RPC over PCIe®

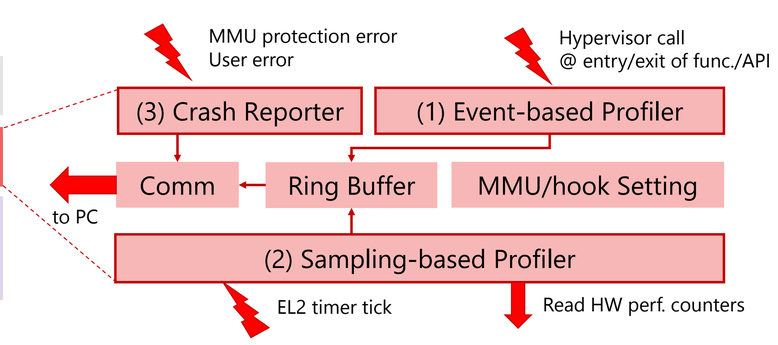
- Careful Integration of gRPC, FlatBuffers, OpenAMP zero-copy API, and direct I/O PCIe® driver
- New extensions to improve their cooperation



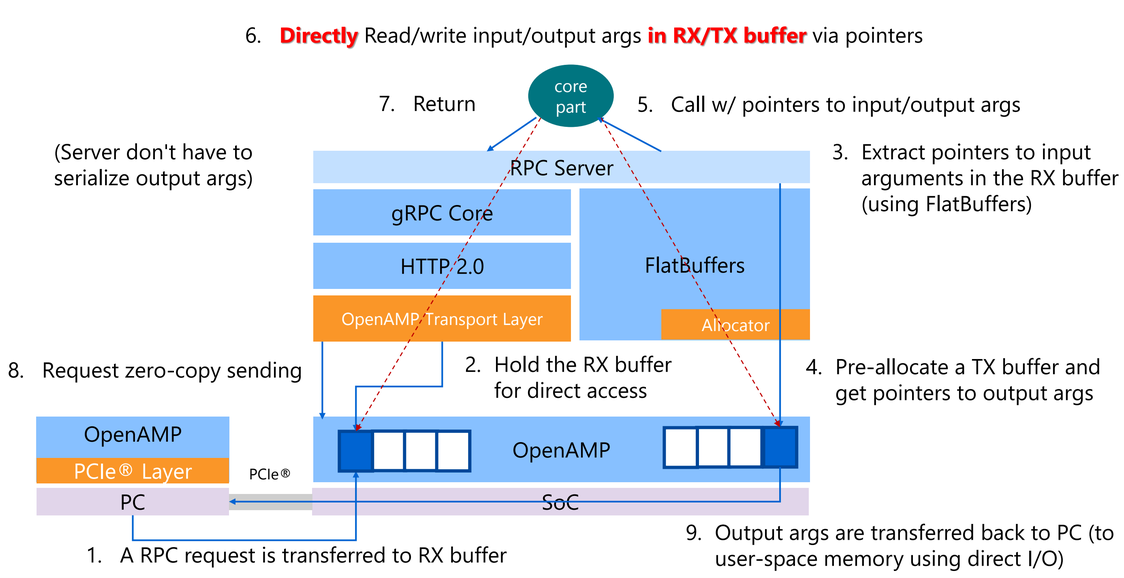


Lightweight Hypervisor

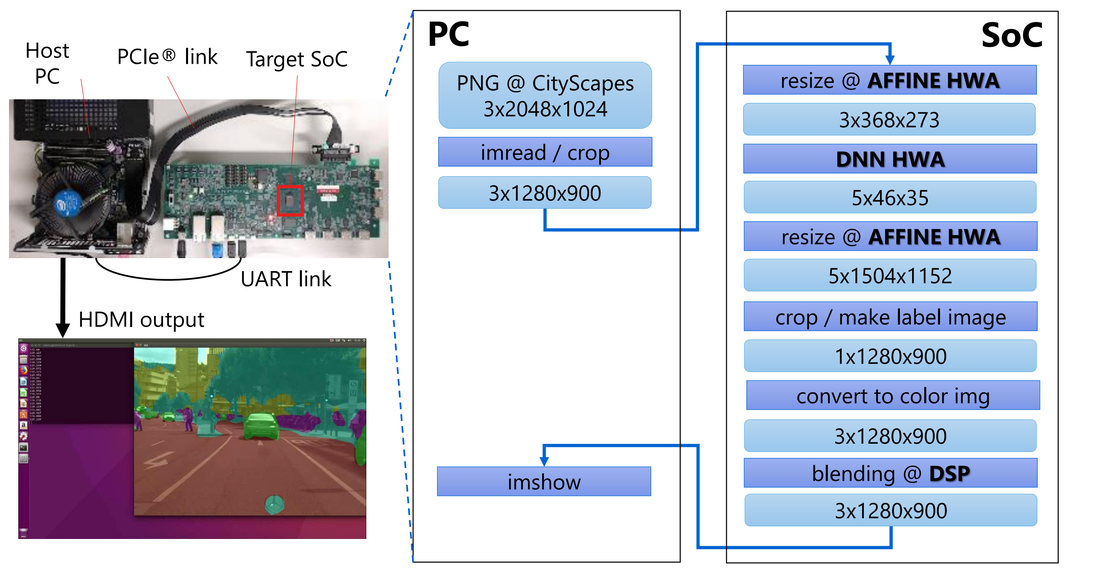
- Dedicated to profiling and monitoring
  - No support for multiple VM
- Use Arm® AArch64 virtualization features
  - Second state translation of MMU, interrupt virtualization, etc.
- Hook insertion at run-time
  - Can be applied for pre-compiled binary app



Zero-copy RPC: Direct accesses to RX/TX buffers via pointers

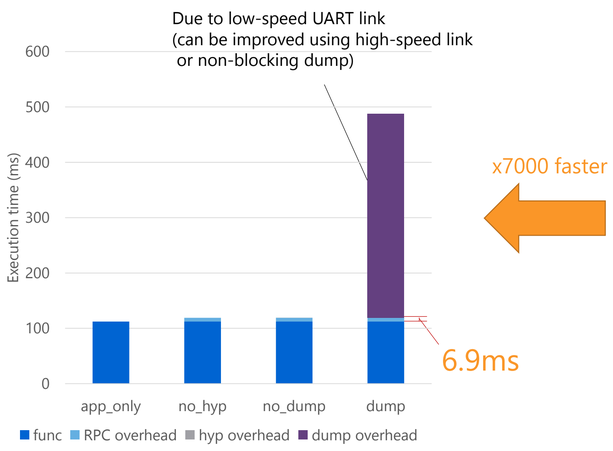


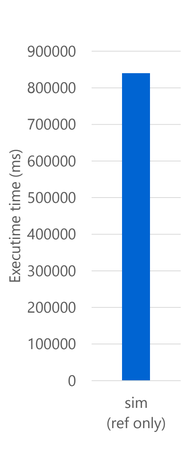
Experiment setup: 5-class DNN semantic segmentation



Experiment result

Overhead





Visualized profile

Conclusion

Difficulty of early accuracy and performance estimation for embedded ADAS applications
- Due to limited precision and different performance curve of target SoC

Our solution: PC x SoC hybrid execution environment
- Achieves low-overhead accuracy and performance estimation
  - Offloading hot-spots of application to target SoC
  - Profiling and monitoring on target SoC
- Proposed techniques
  - Zero-copy RPC over PCIe®
  - Lightweight hypervisor specialized in profiling and monitoring

Experiment
- Our solution has enough capability for early estimation using DNN semantic segmentation
  - achieves 6.9ms offload latency @1280x900 image which is small enough to run application in real-time
  - probes user SW, HWAs, and bus behavior on target SoC with tiny overhead