PhD Project:

Trustworthy Real-Time Virtualization for Heterogeneous Platforms

José **Martins**

Advisor :

Prof. Sandro Pinto

Co-Advisors:

Prof. Adriano Tavares &

Prof. Marko Bertogna

FCT Fundação para a Ciência e a Tecnologia

CENTROALGORITMI

UNIVERSIDADE DO MINHO

AGENDA

01 PhD Proposal

02 Virtualization & Mixed-Criticality

03 Bao Hypervisor

04 Bao WIP

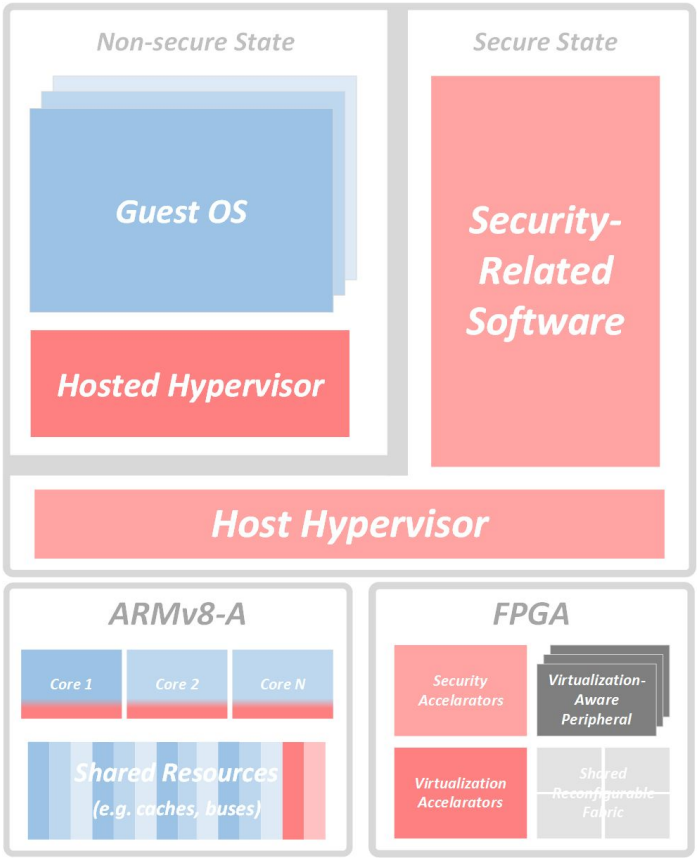
05 Master’s Thesis Proposals

PhD PROPOSAL

This project proposes the development of an holistic VIRTUALIZATION infrastructure targeting MIXED-CRITICALITY, CYBER-PHYSICAL systems deployed on heterogeneous platforms.

Main Goals:

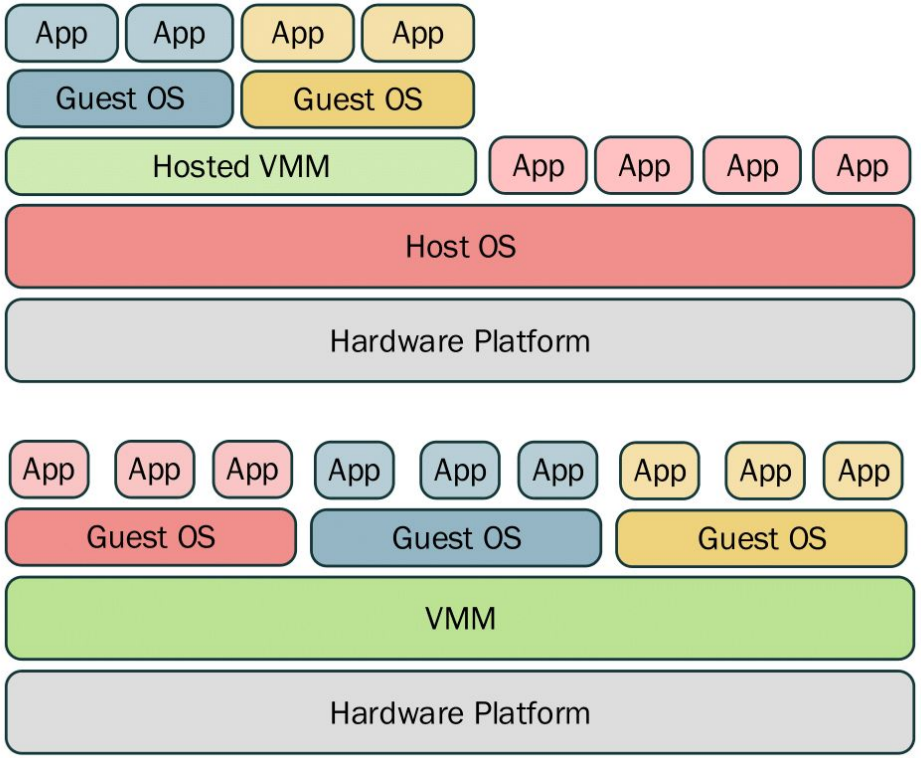
- Identify **synergies between existing virtualization and security hardware extensions** (e.g., Arm's VE and TrustZone) and implement a dual-layer hypervisor architecture embodying these ideas.
- **Survey and identify new microarchitectural vulnerabilities** (e.g. timing side channels) on existing COTS hardware and existing embedded hypervisors. Propose software approaches to mitigate or minimize and achieve higher virtual-machine microarchitectural isolation.
- Analyze **drawbacks and overheads of current existing virtualization extensions** having in mind embedded virtualization requirements such as real-time and SWaP-C. Propose alternatives and prototype them using open architectures such as RISC-V.
- Investigating the possible **benefits of decentralized virtualization** by taking advantage of the reconfigurable fabric in heterogeneous platforms to implement virtualization and security functionality.



VIRTUALIZATION

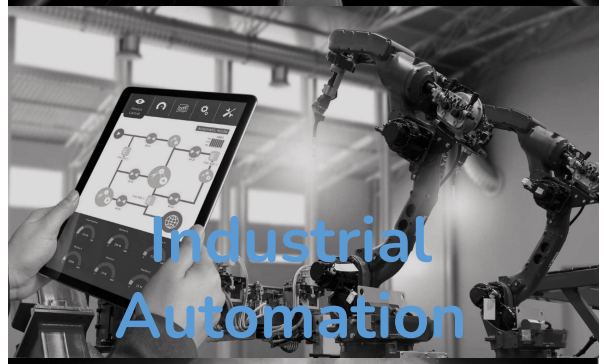
Allows the execution of multiple Operating Systems in the same hardware platform.
*An **Hypervisor or VMM** is to an OS, as an OS is to a process.*

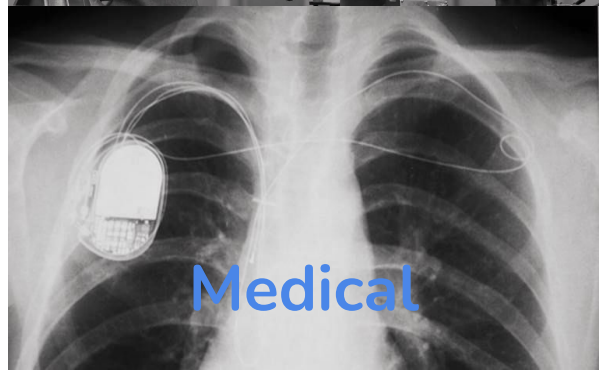
- Main functions:
  - Resource Management
  - Abstraction
  - Protection/Isolation
- The hypervisor provides a Virtual Machine (VM) abstraction for guest OSs.
- Used extensively in:
  - servers (load balancing, power management)
  - desktops (cross-platform, systems development)



MIXED-CRITICALITY / CYBER-PHYSICAL

- Cyber-Physical Systems
- Growing complexity & connectivity
- Multiple Subsystems:
  - Heterogeneous Software Stacks (RTOS, GPOS)
  - Different Criticality Levels
- Size, Weight, Power, Cost constraints -> Consolidation





CYBERSECURITY

Hackers remotely kill Jeep's engine on highway

PUBLISHED TUE, JUL 21 2015•5:18 PM EDT | UPDATED WED, JUL 22 2015•8:41 AM EDT

NEWS

Siemens: Stuxnet worm hit industrial systems

Tech

Barnaby Jack Could Hack Your Pacemaker and Make Your Heart Explode

VIRTUALIZATION & MIXED-CRITICALITY

Consolidation

Size
Weight
Power
Cost

Low Engineering Cost

Full-Virtualization allows direct porting of guest OSs
Add sec/saf mechanisms at hypervisor layer

Isolation & Fault Containment

Sandboxed Environments

Performance

Hardware virtualization ext.
Low mode crossing frequency

Low-latencies (interrupt, boot, ..)
Determinism/Predictability
Freedom-from-interference

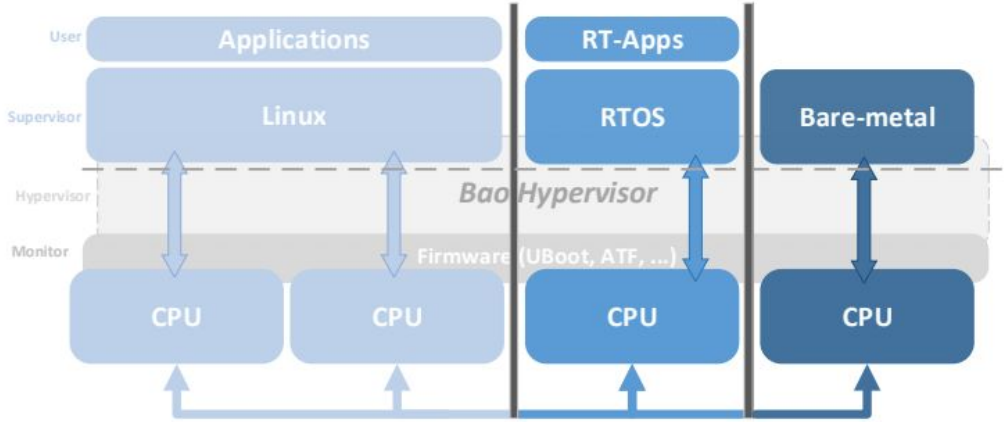
Real-time & Safety

Small TCB
Side-channel Protection
TEE support

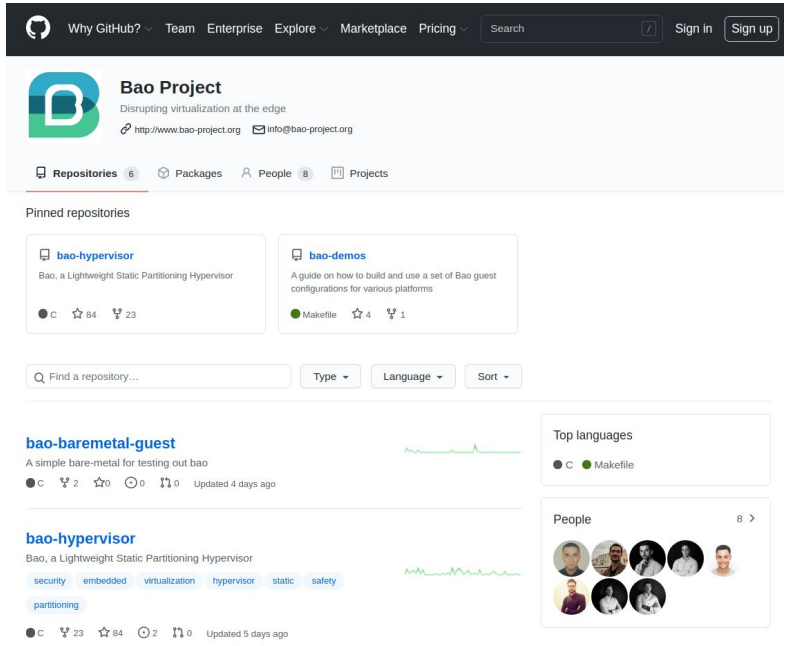
Security

BAO HYPERSIVOR

- Type-1 / Bare-metal
- SMP
- Needs HW Virtualization Extensions
- Static Partitioning:
  - 1:1 vCPU-to-pCPU mapping (no scheduling)
  - Static memory assignment
  - Device Pass-through (needs IOMMU)
  - Hardware interrupts
- Small TCB (~7 KSLoC)
- Arm & RISC-V
- Open-source (checkout bao on github!)
- Main goal: ISOLATION

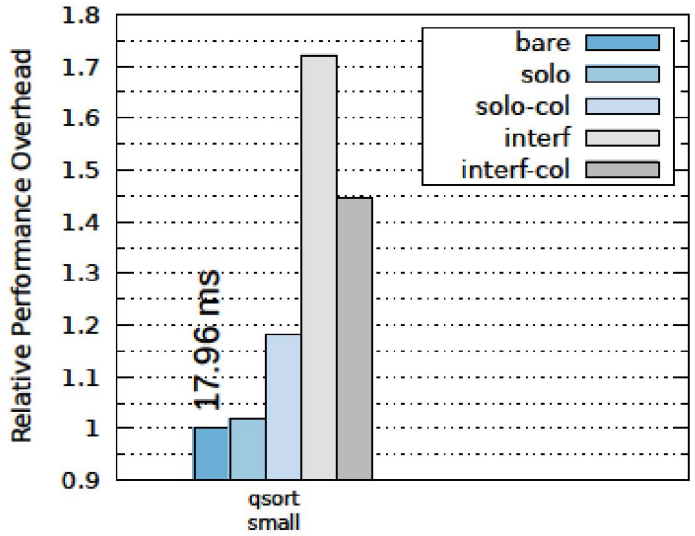


Paper: *Bao: A Lightweight Static Partitioning Hypervisor for Modern Multi-Core Embedded Systems*

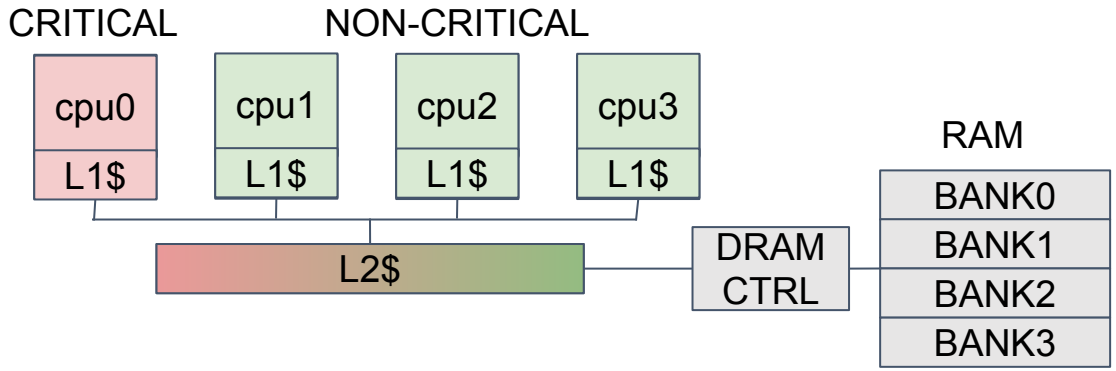


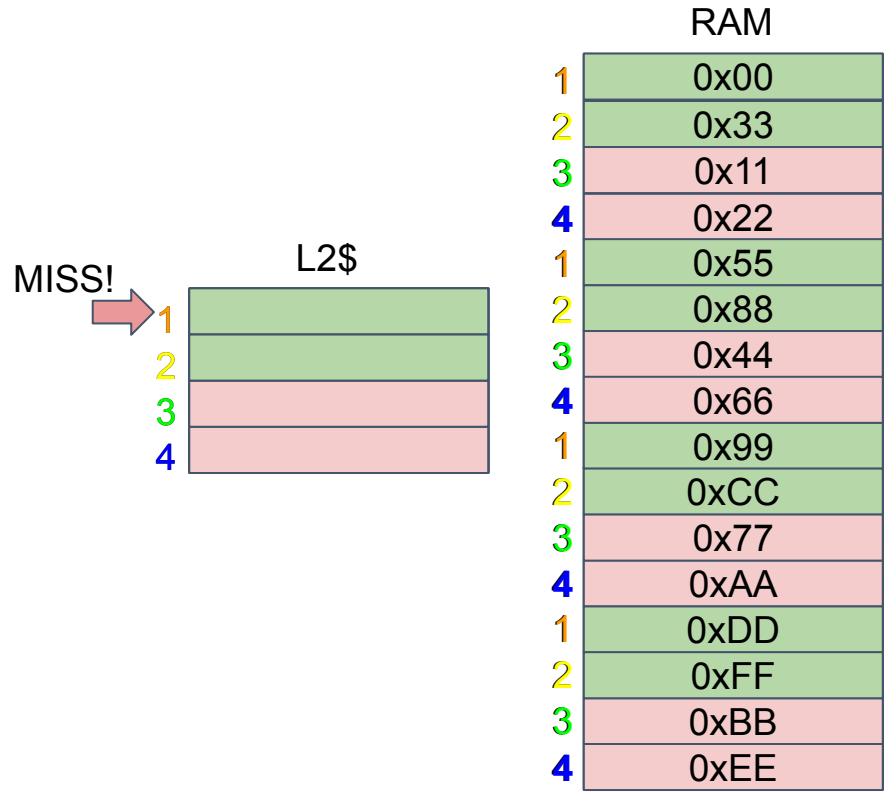
AVOIDING INTERFERENCE & SIDE-CHANNELS

- Cache Partitioning (e.g. coloring) solves (most) of cache interference problems.
- But what about the rest of the memory hierarchy?

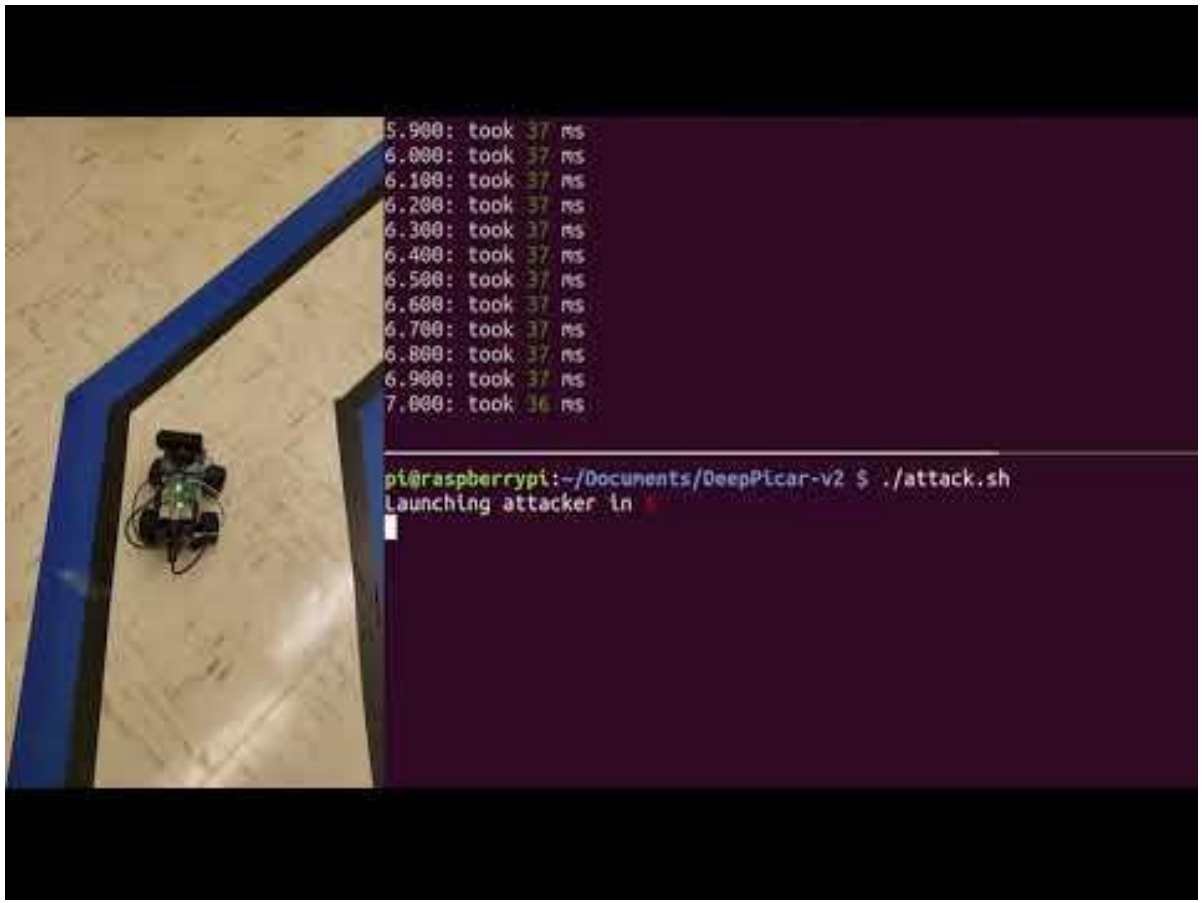


- DRAM bank coloring
- Memory throttling
- Hardware mechanisms?





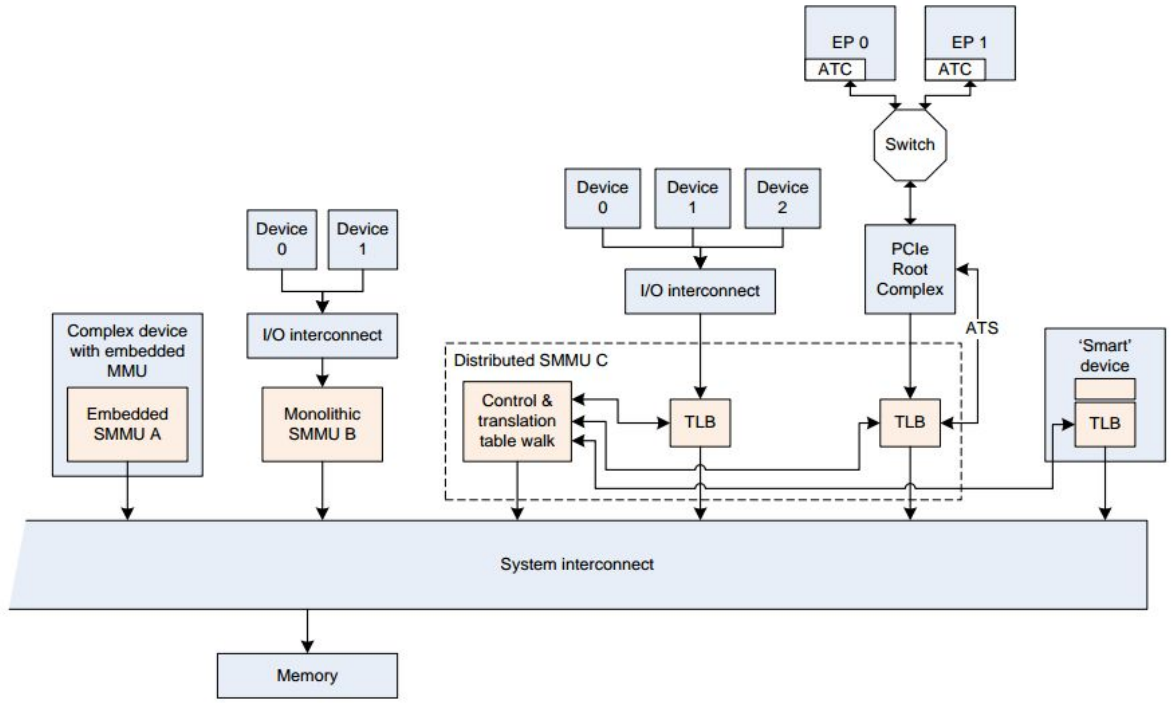
AVOIDING INTERFERENCE & SIDE-CHANNELS



DeepPicar: A Low-cost Deep Neural Network-based Autonomous Car
Michael G. Bechtel, Elise McEllhiney, Minje Kim, Heechul Yun

IOMMU

- **Heterogeneity and limitations:**
  - Non-existent (need hypervisor mediation)
  - Limited number of simultaneous active devices (SMMUv2, solved in v3).
  - Platform-specific: IOMMU, master ID programming, …
- **Mitigations (e.g. memory throttling) not currently extended to SMMU.**
- **Shared micro-architectural structs:**
  - TLBs
  - PTWs
- **SMMUv3.2 adds MPAM. But…**



From *Arm SMMU Architecture specification, version 3*

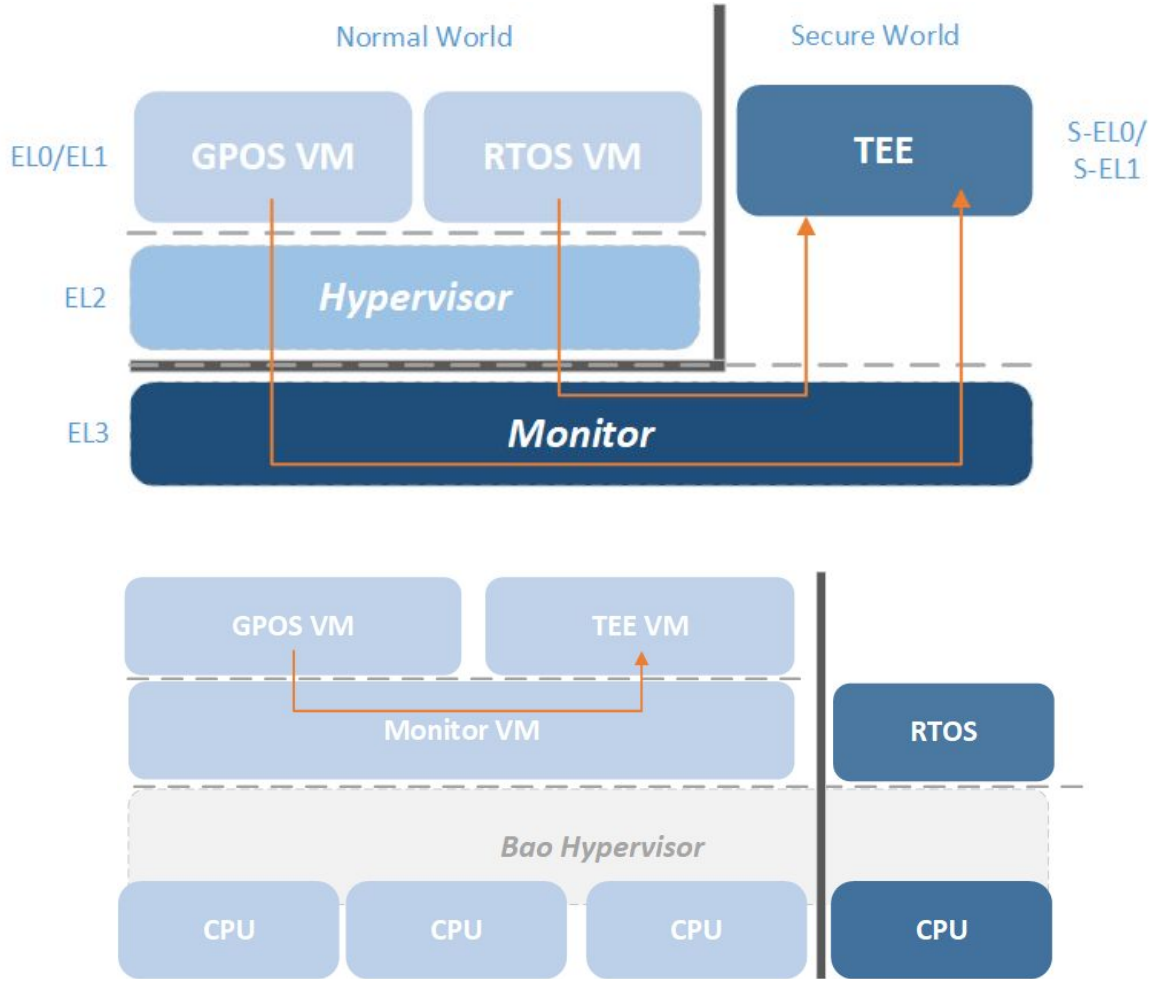
RISC-V HYPERVISOR EXTENSION

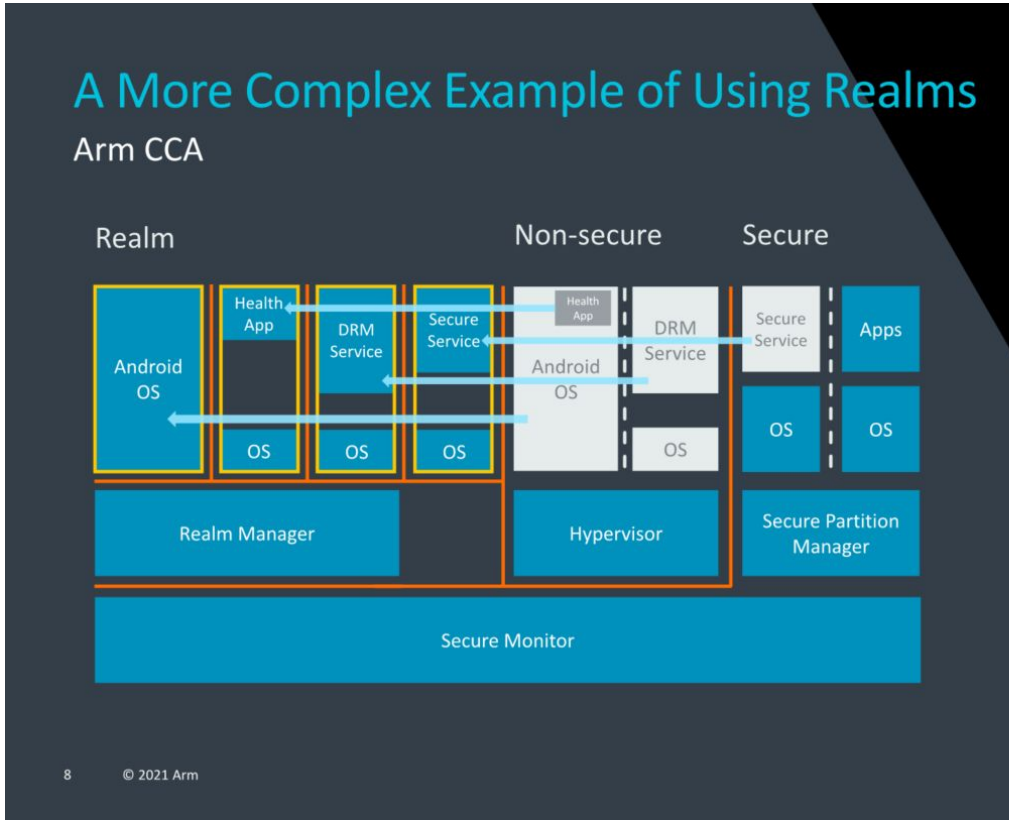
- First public RTL implementation (only QEMU was available) of the RISC-V hypervisor extension specification.
- Implemented in the Rocket Core using the Chisel Language
- Used by the KVM community for validation of the specification.
- To be used in future research to develop custom extension and hardware modules to assist Bao.

Presented at RISC-V Summit 2020 (video available on youtube)

Paper: *A First Look at RISC-V Virtualization from an Embedded Systems Perspective*

arm TEE SUPPORT





MASTER'S THESIS PROPOSALS

1. VirtIO infrastructure:
   - Study and select transport
   - Backend repository (virtio-console, virtio-net, virtio-block, ...)
2. Trusted Execution Environments using VM-stacking
3. Memory throttling, DRAM bank partitioning and Arm MPAM support.
4. SMMUv3 (Arm IOMMU) support and DMA attack and interference analysis.
5. *Breaking Bao*: attack surface and vulnerability analysis (e.g. using fuzzing).

OTHER OPTIONS

- Bao infrastructure Improvements:
  - Error handling and recovery
  - Automated TDD framework
  - Port using memory-safe languages (Rust)
  - Bao debug/tracing framework

- Implement security functionality:
  - explore secure TrustZone and Realms in Arm processors
  - at hypervisor level (secure Boot, ASLR, CFI/DFI, Arm PAuth support)
  - at VM level (VM introspection, enclaves, secure boot, secure updates, attestation)

- Device sharing/virtualization (related to VirtIO):
  - GPU
  - PCIe
  - ....

- RISC-V hypervisor hardware extensions:
  - Implement and explore ePMP and sPMP for partitioning virtualization
  - Hardware interference mitigations (see BRU implemented in Chisel)
  - Design/implement isolation specification/modules (e.g. IOMMU)

- MISC:
  - VM-level schedulers
  - Develop hypervisor ports and drivers for new hardware
  - New guest support (Zephyr, NuttX, MirageOS Unikernels)
  - Interdisciplinary project using Bao

ANY IDEAS ?!

THANK YOU FOR YOUR TIME!

QUESTIONS?

jose.martins@dei.uminho.pt