

seL4 + TrustZone: Spanning both worlds

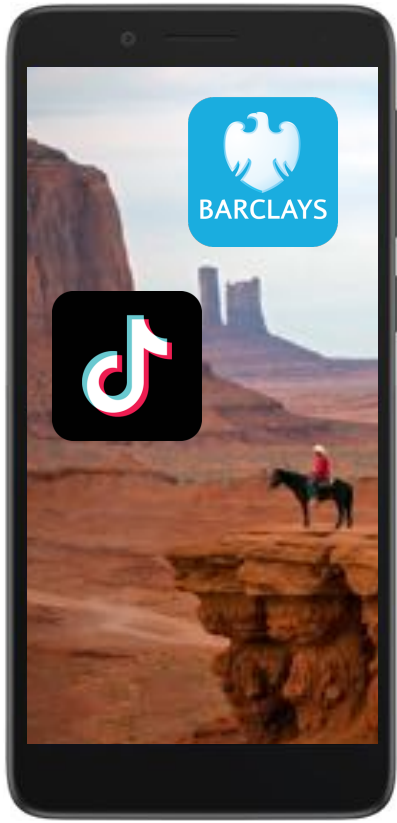
Nick Spinale
Arm Research

nick.spinale@arm.com

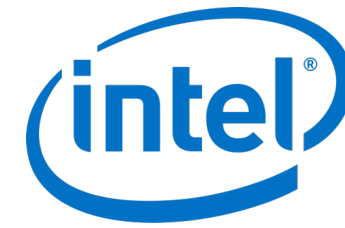
arm

Confidential Computing and Virtualization

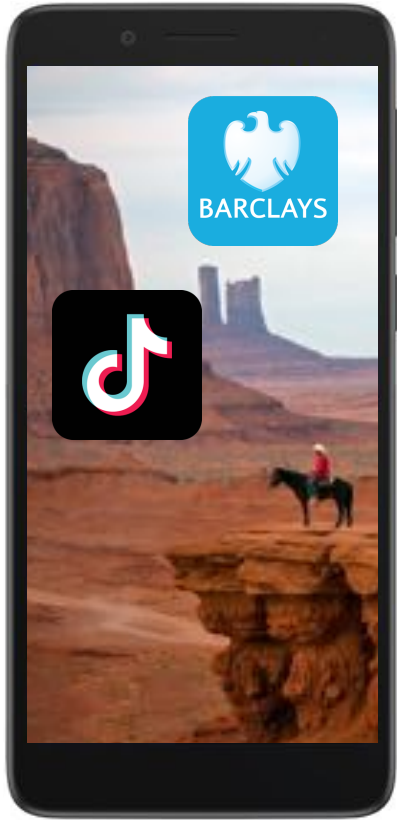
Confidential Computing



Protecting data ~~at rest in transit~~ in use



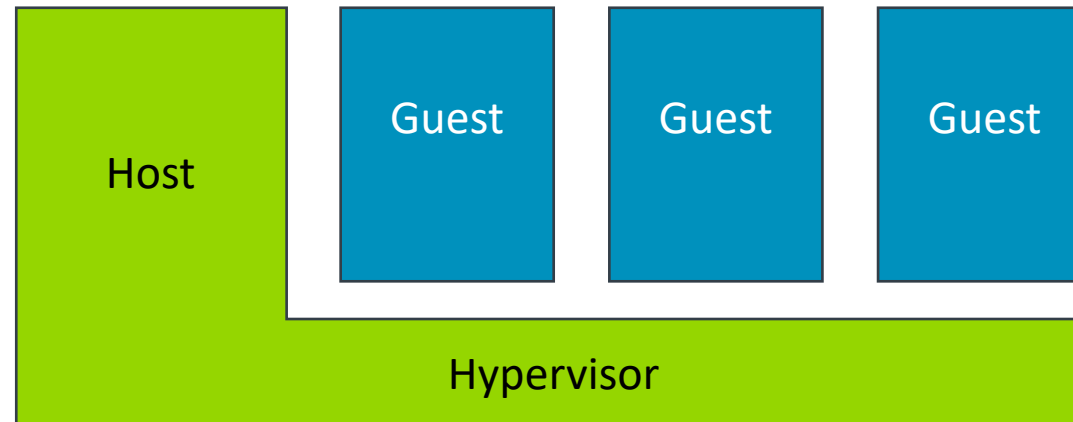
Confidential Computing



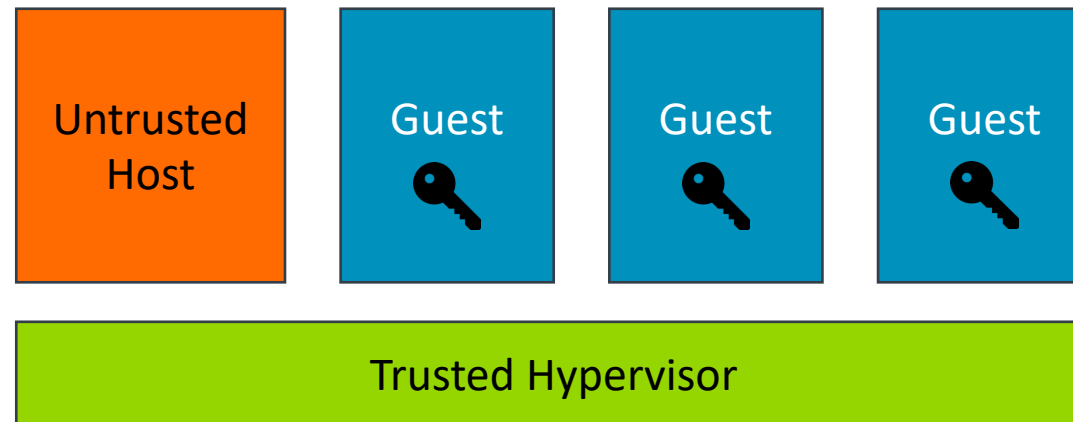
How far can we get with *software alone*?



Confidential Computing and Virtualization



Confidential Computing and Virtualization

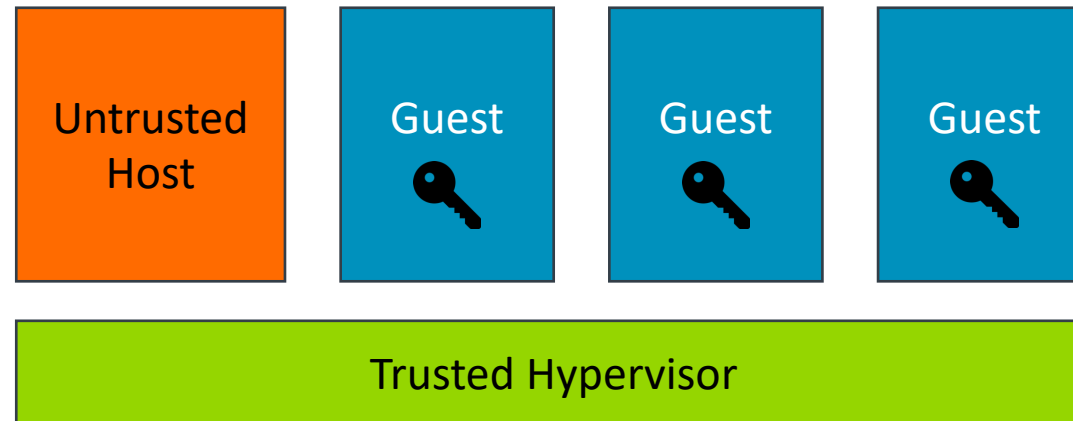


arm

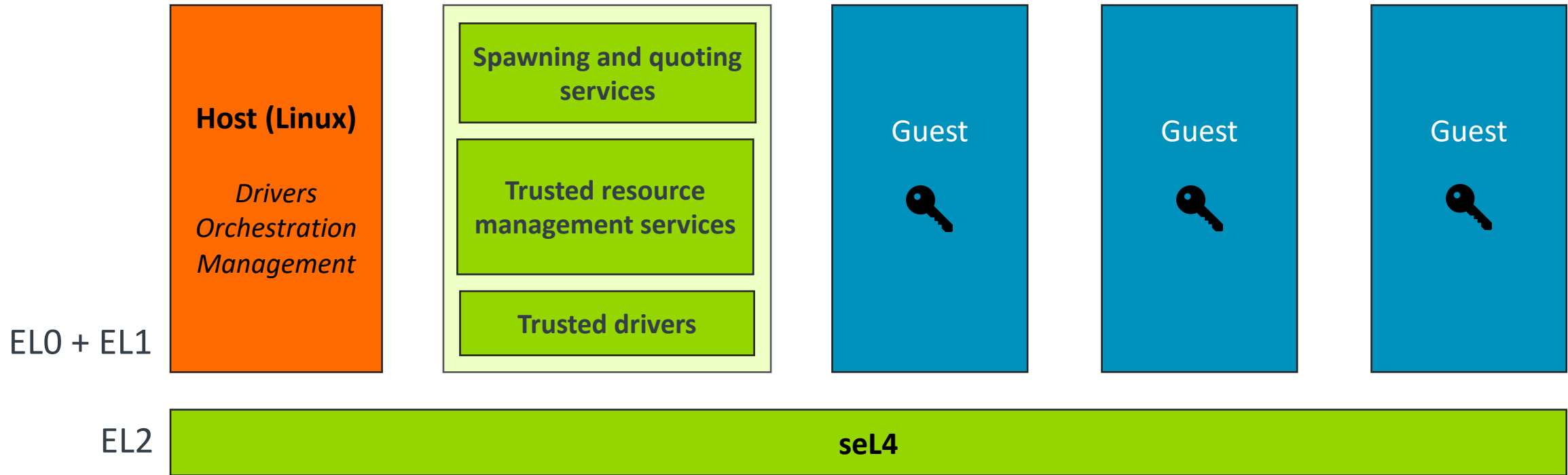
IceCap

<https://gitlab.com/arm-research/security/icecap/icecap>

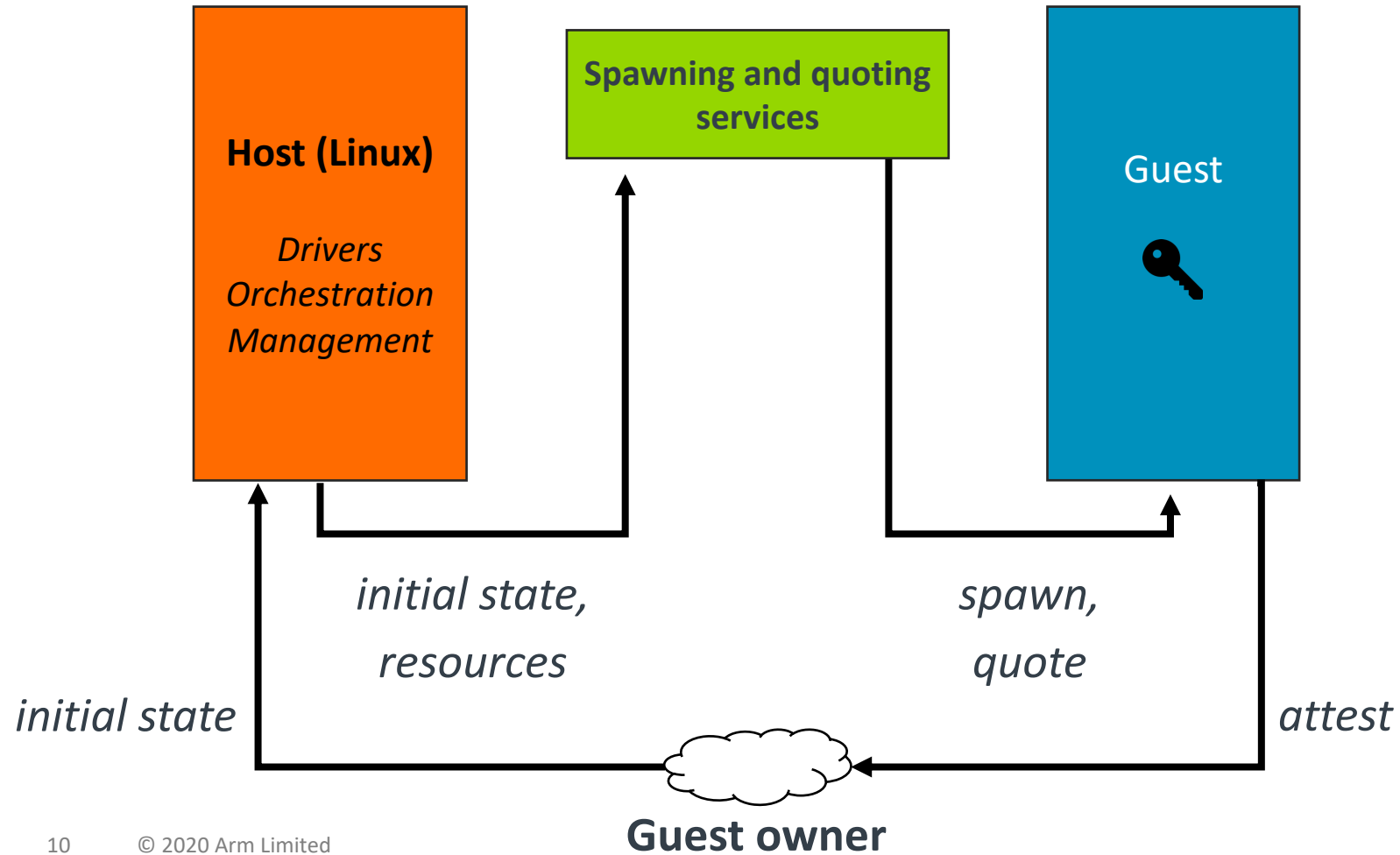
IceCap



IceCap



IceCap: Attestation

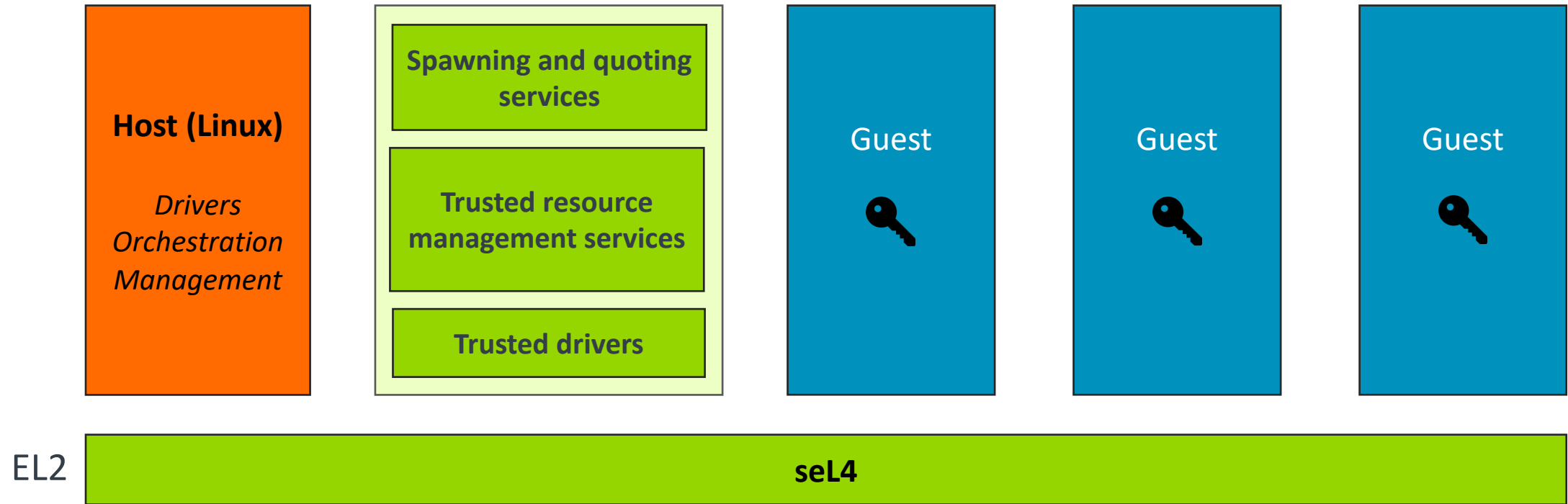


IceCap: Extended CapDL

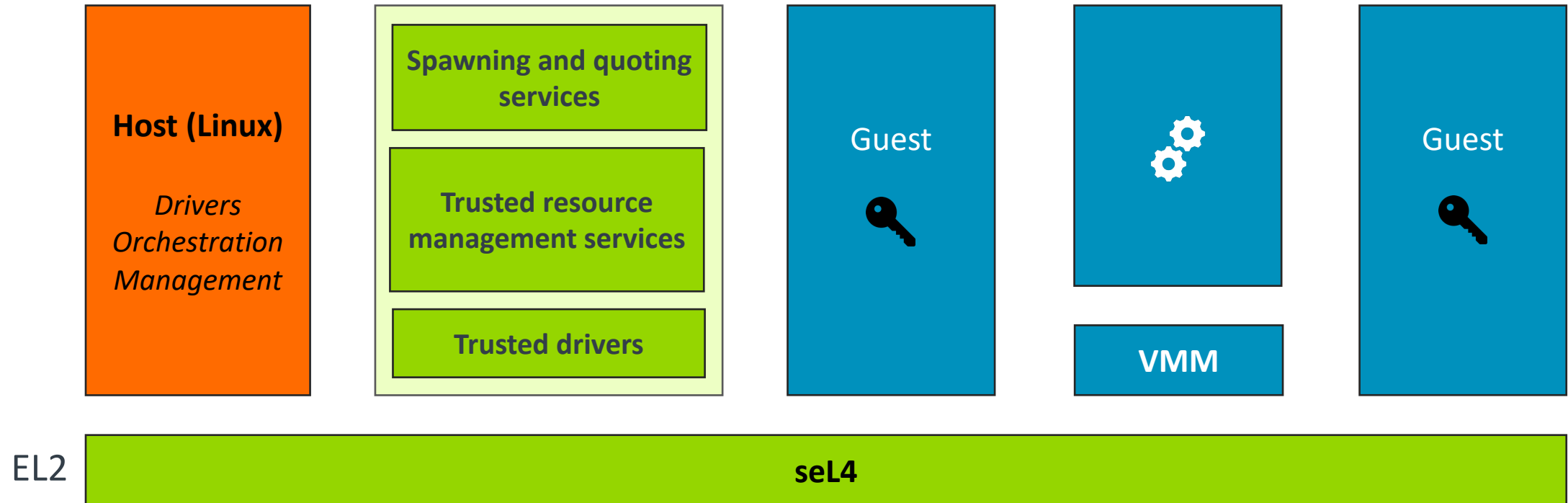
```
arch aarch64
```

```
objects {  
    extern host_shared_memory_region (4k)  
    extern timer_endpoint = ep  
    extern timer_wait = notification  
  
    guest_primary_thread = tcb (...)  
    guest_elf_0001 = frame (4k, fill: [...])  
    ...  
}  
  
caps {  
    guest_cnode {  
        0x1: timer_endpoint (W, badge: ...)  
        ...  
    }  
    ...  
}
```

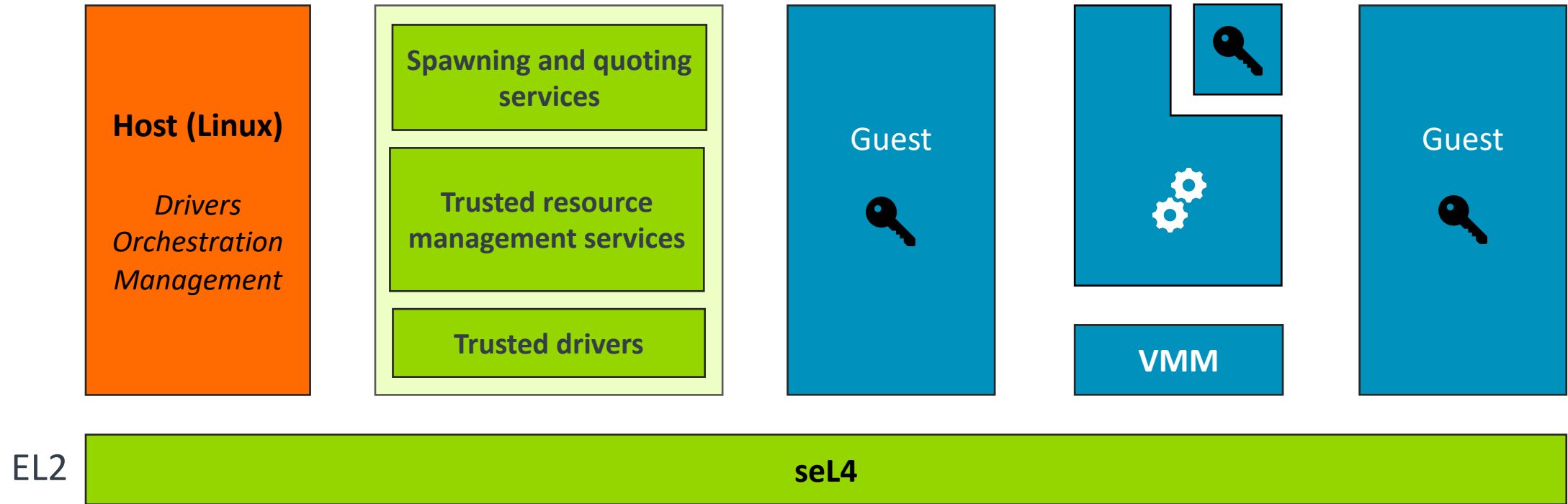
IceCap



IceCap

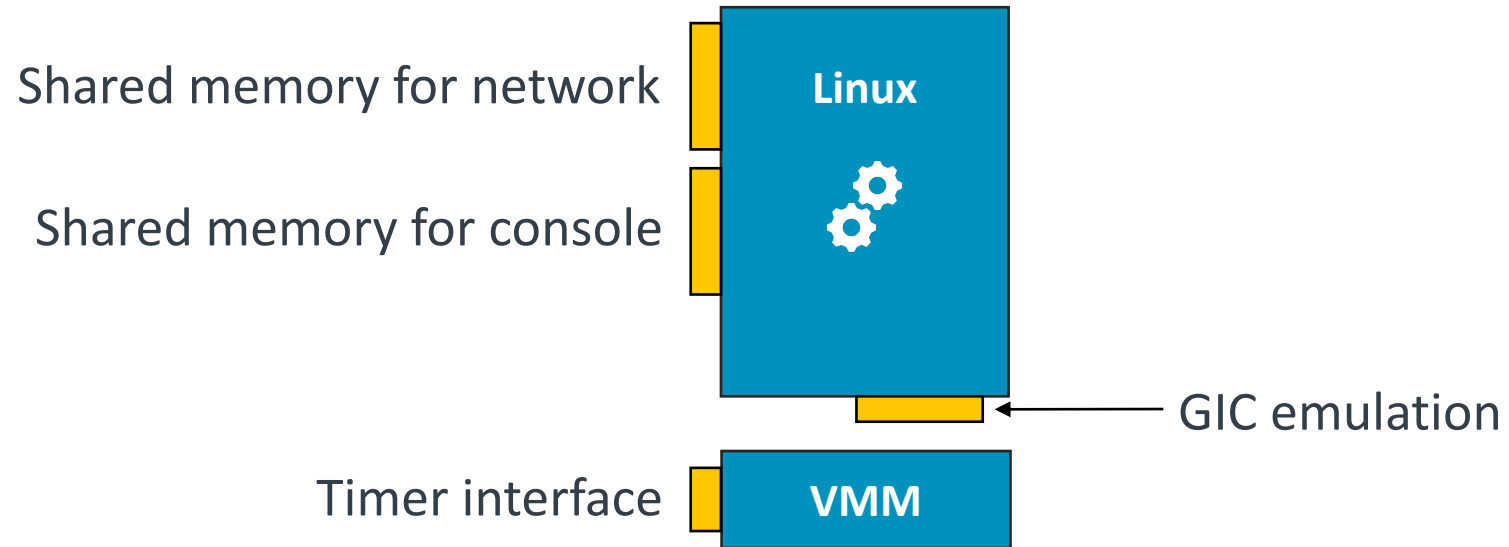


IceCap



IceCap VMM

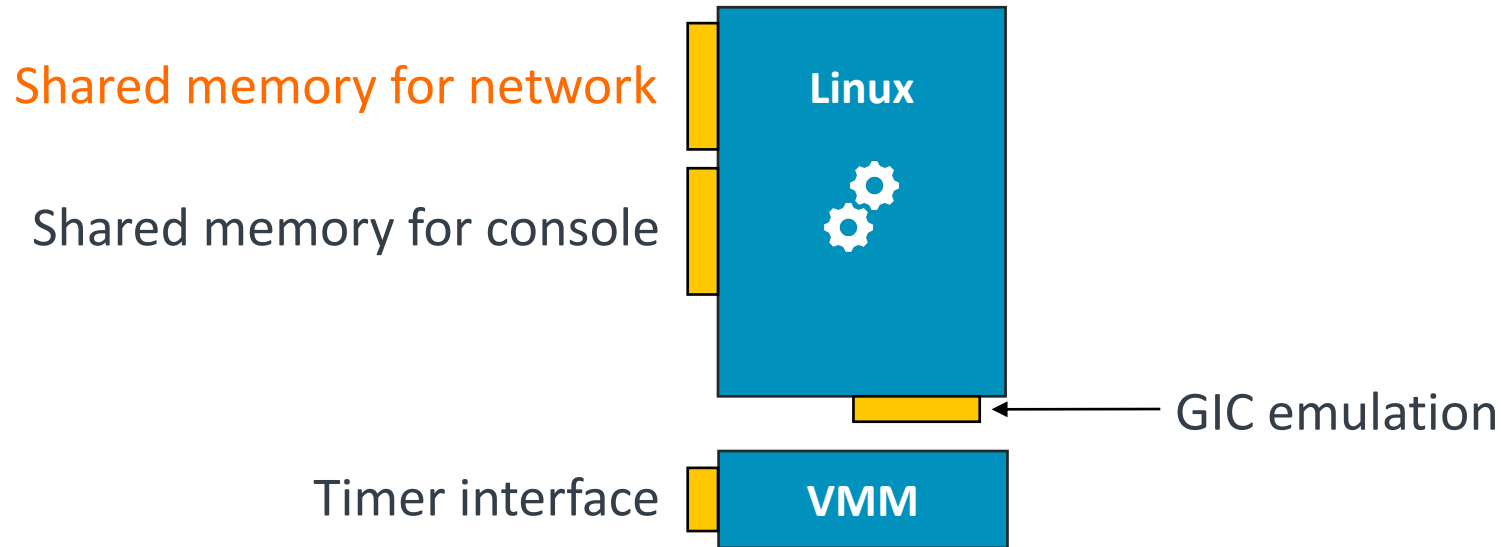
Only VM exits are for interrupt injection and GIC emulation
<1kLOC (Rust)



IceCap VMM: Preliminary observations

No benchmarks yet

- Preliminary observations suggest host-guest network performance in the neighborhood of AWS Firecracker (open source VMM for KVM used in AWS Lambda)¹

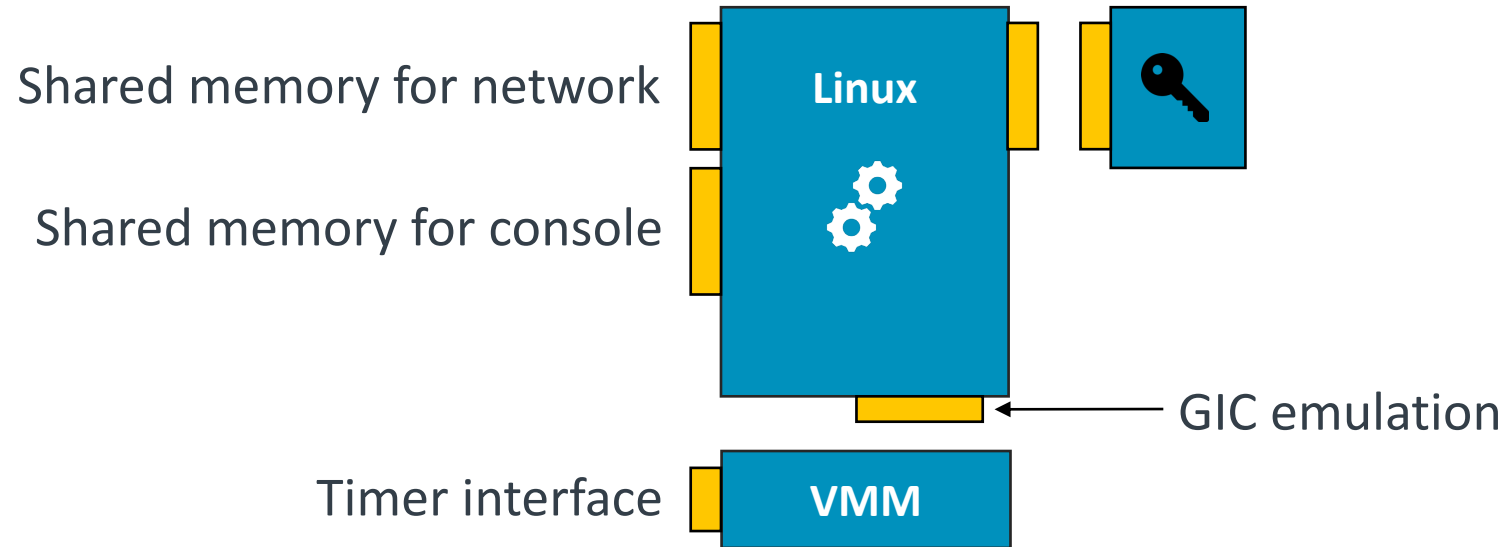


¹ iperf host-guest on Raspberry Pi 4:

- Firecracker: 2.67 Gbit/s
- IceCap: 2.48 Gbit/s

IceCap VMM

Guest may subdivide further



IceCap: Source code

seL4 userland written entirely in Rust (only C is seL4, libsel4, and CapDL)

MirageOS (OCaml unikernel) ported to IceCap

Open source: gitlab.com/arm-research/security/icecap/icecap



GitLab

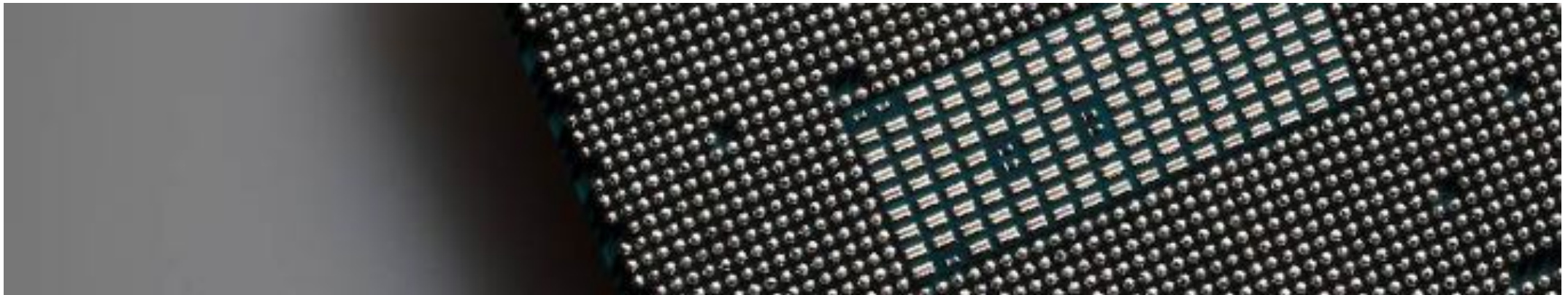


IceCap: Big Kernel Lock

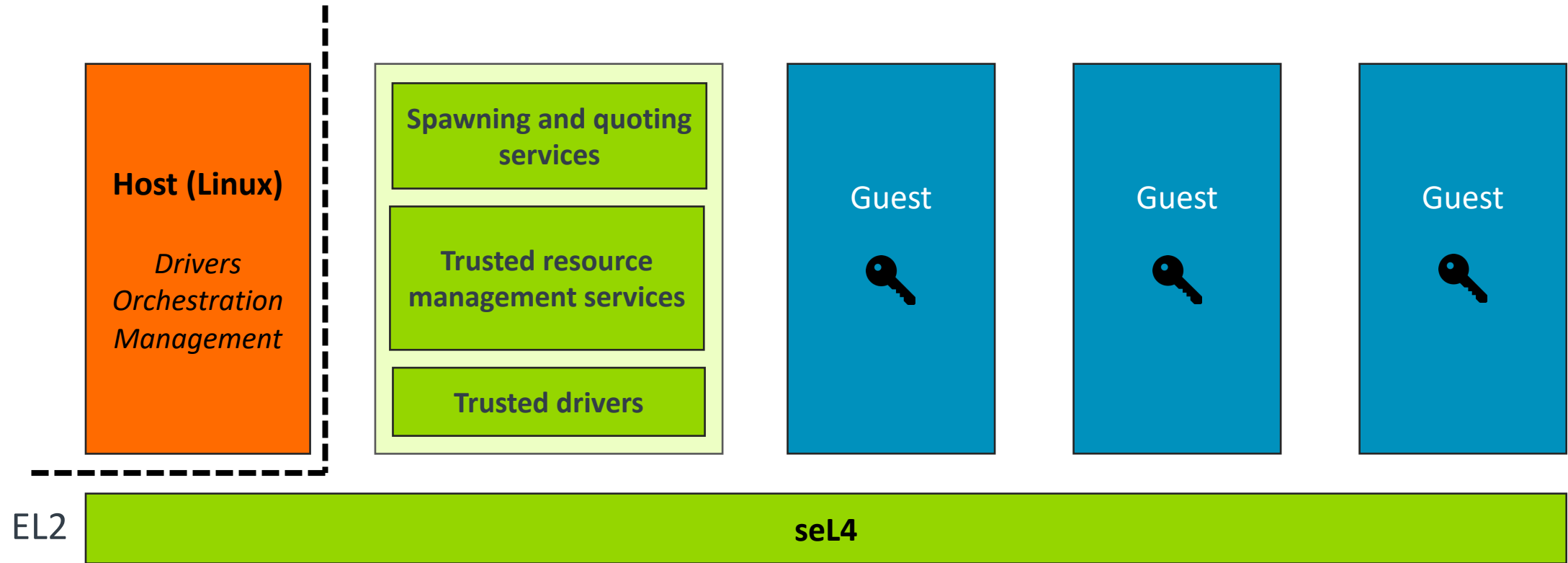
seL4 can only run on one core at a time

Effects performance and availability on some types of hardware platforms

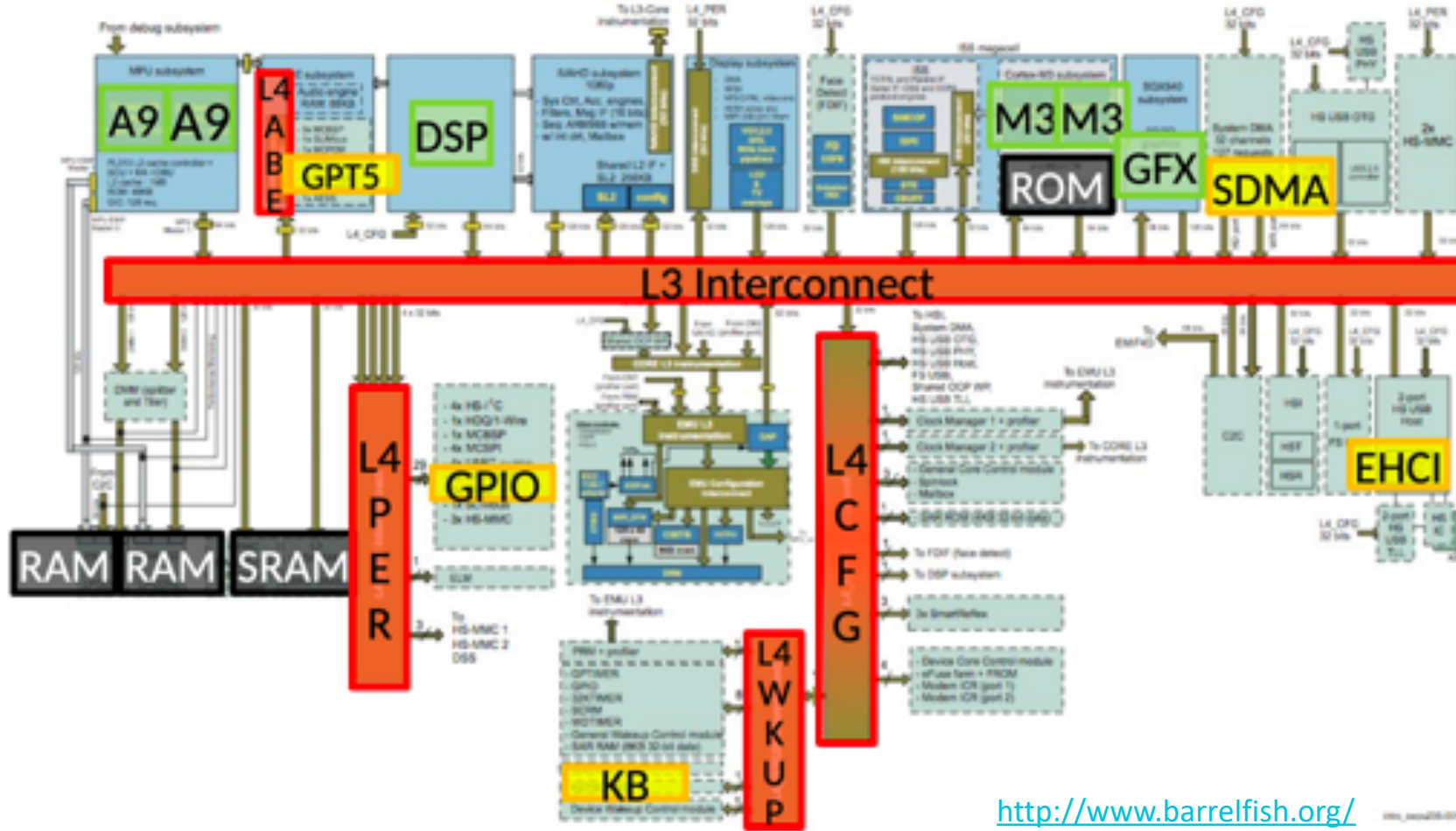
Interrupt mitigation + paravirtualized interrupt controller



IceCap: Protecting guests



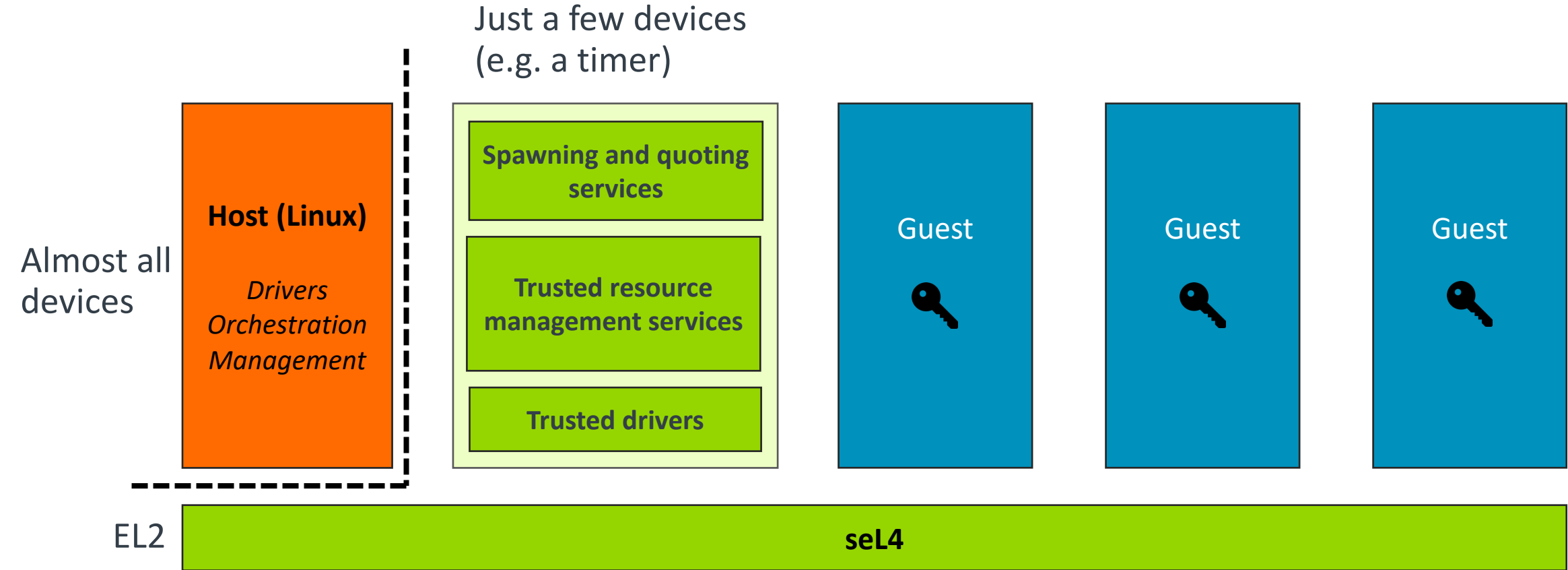
IceCap: Protecting guests



<http://www.barrelfish.org/>

<https://entropy2018.sciencesconf.org/data/cock.pdf>

IceCap: Protecting guests

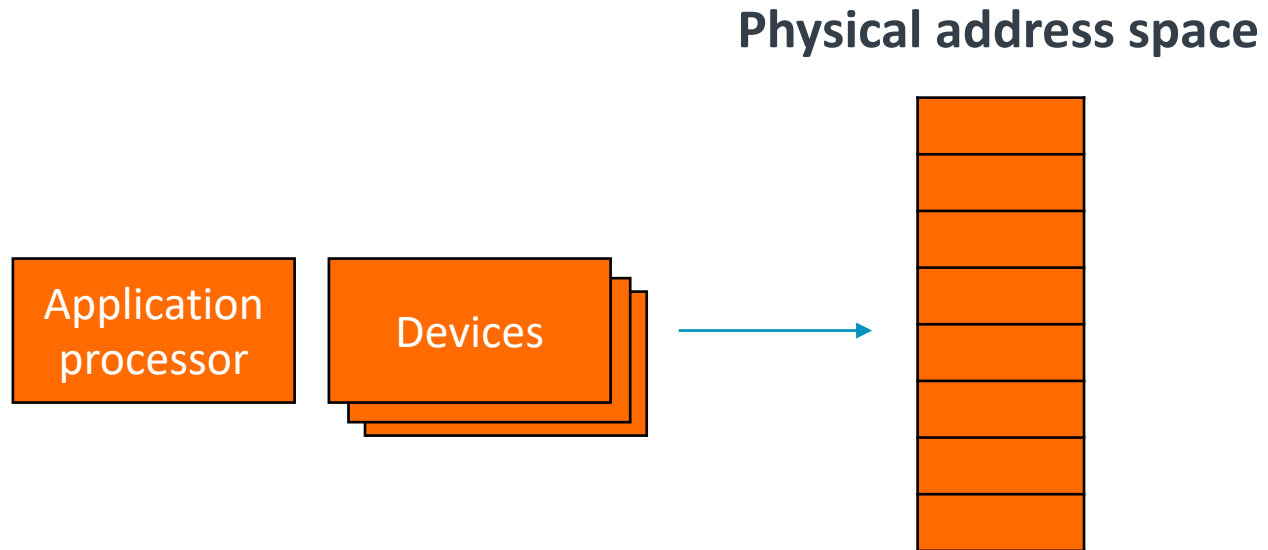


arm

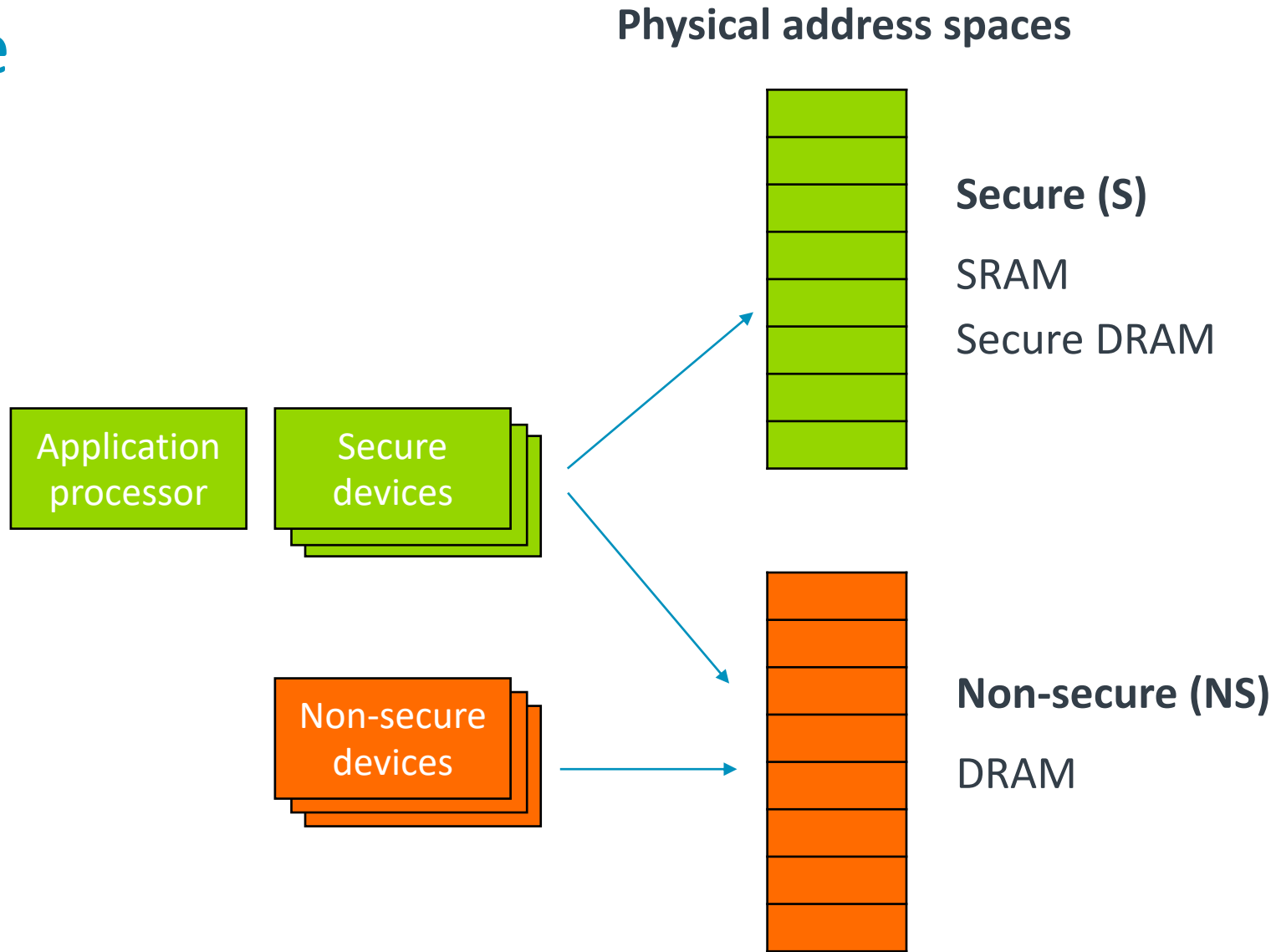
Arm TrustZone™

TrustZone

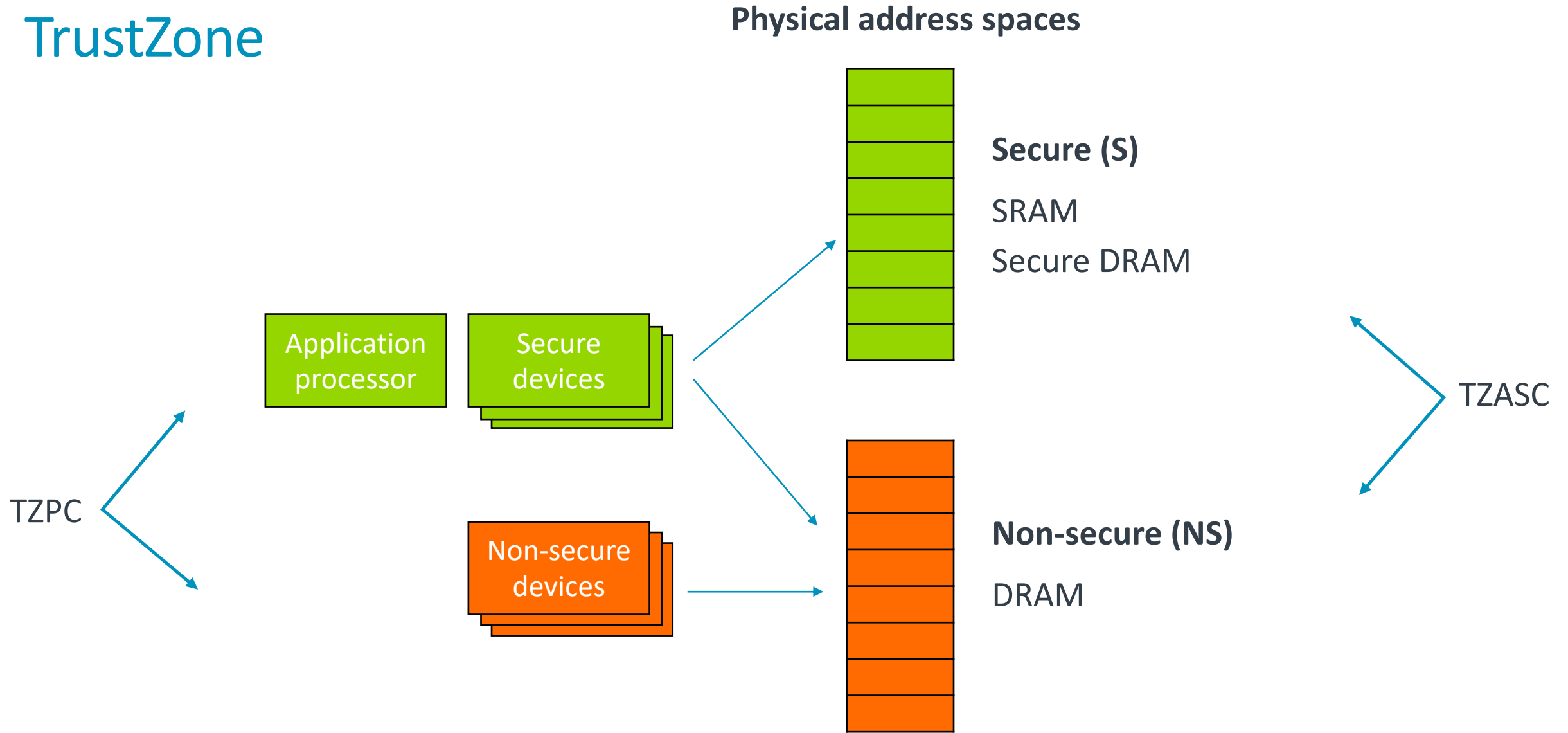
Without TrustZone



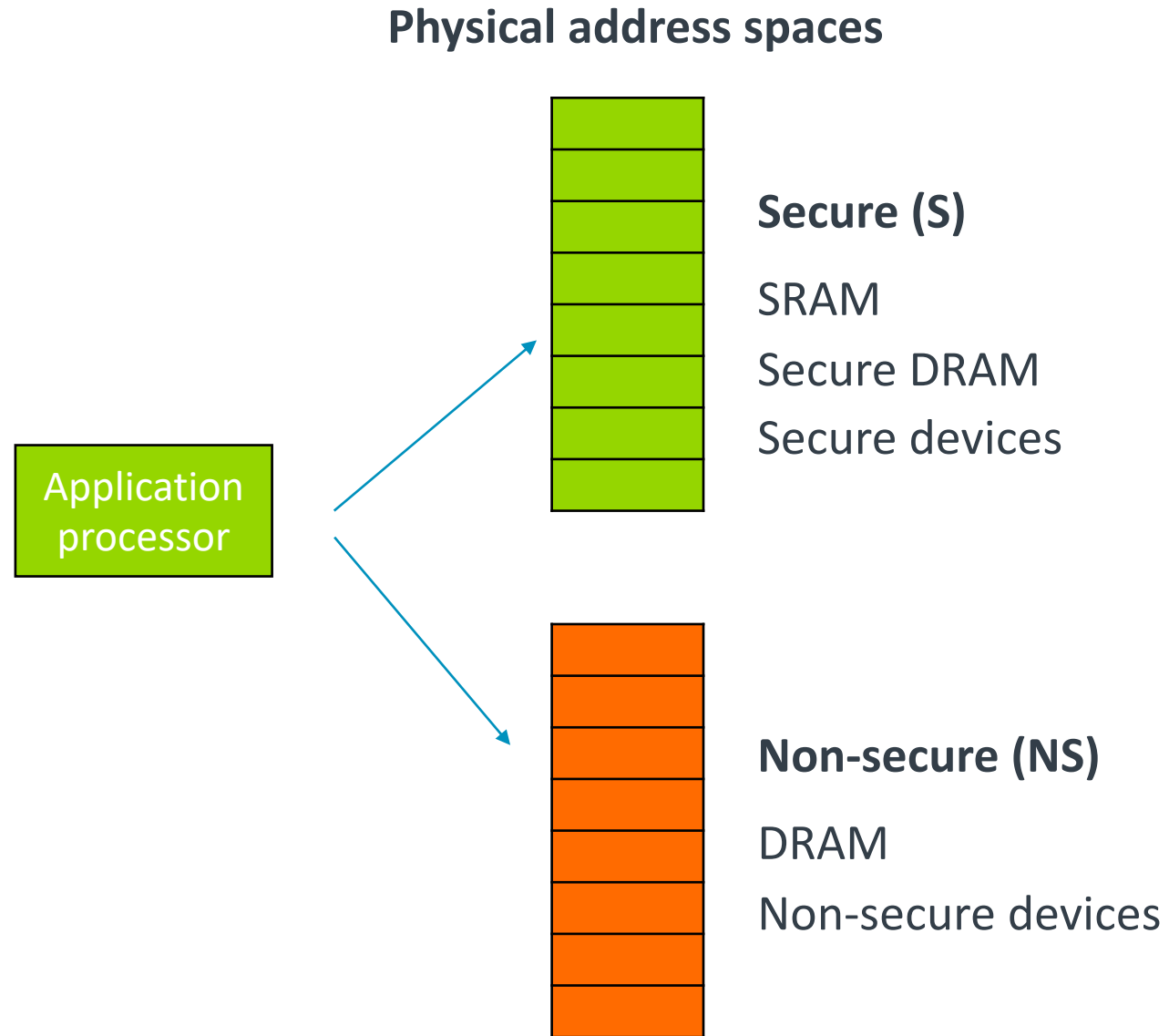
TrustZone



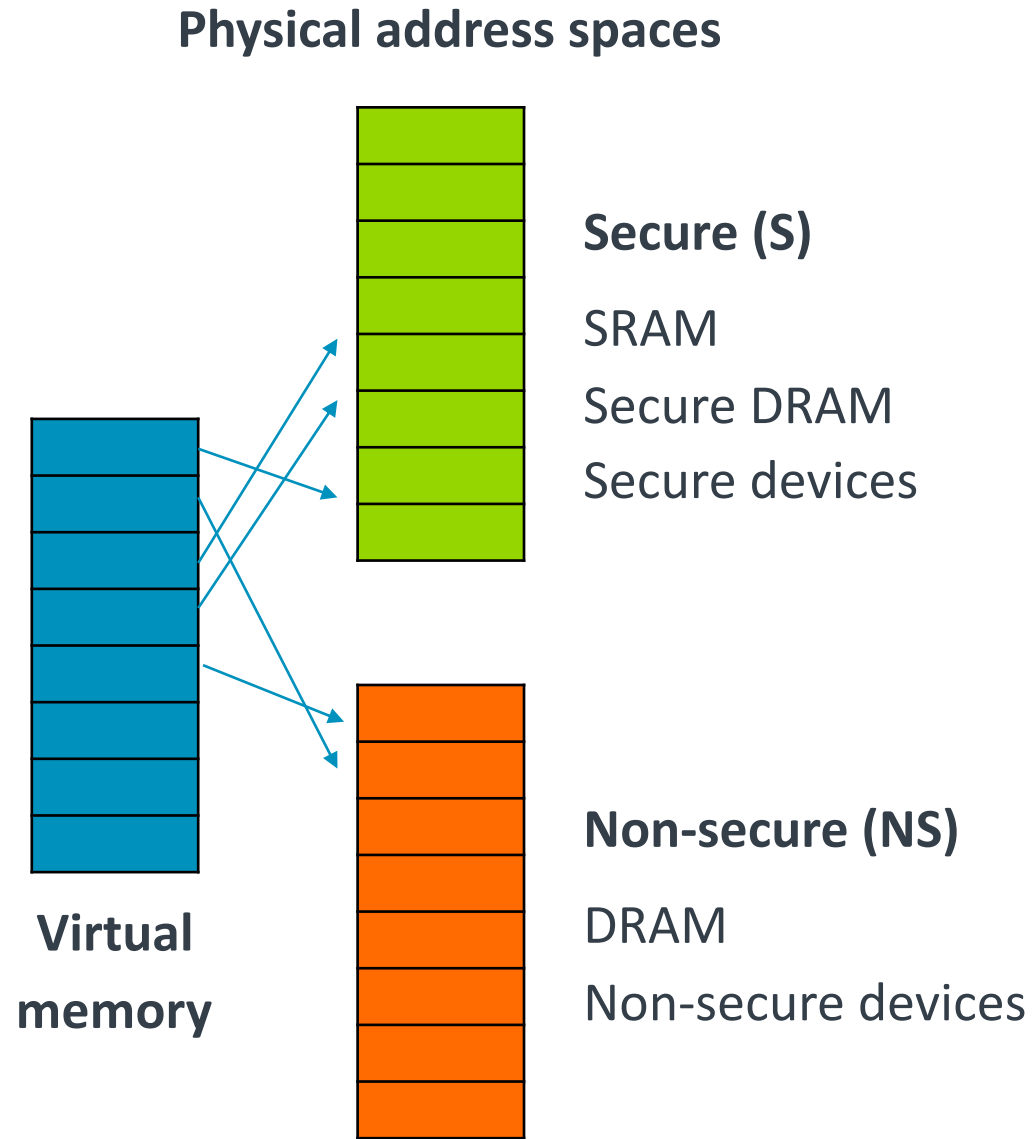
TrustZone



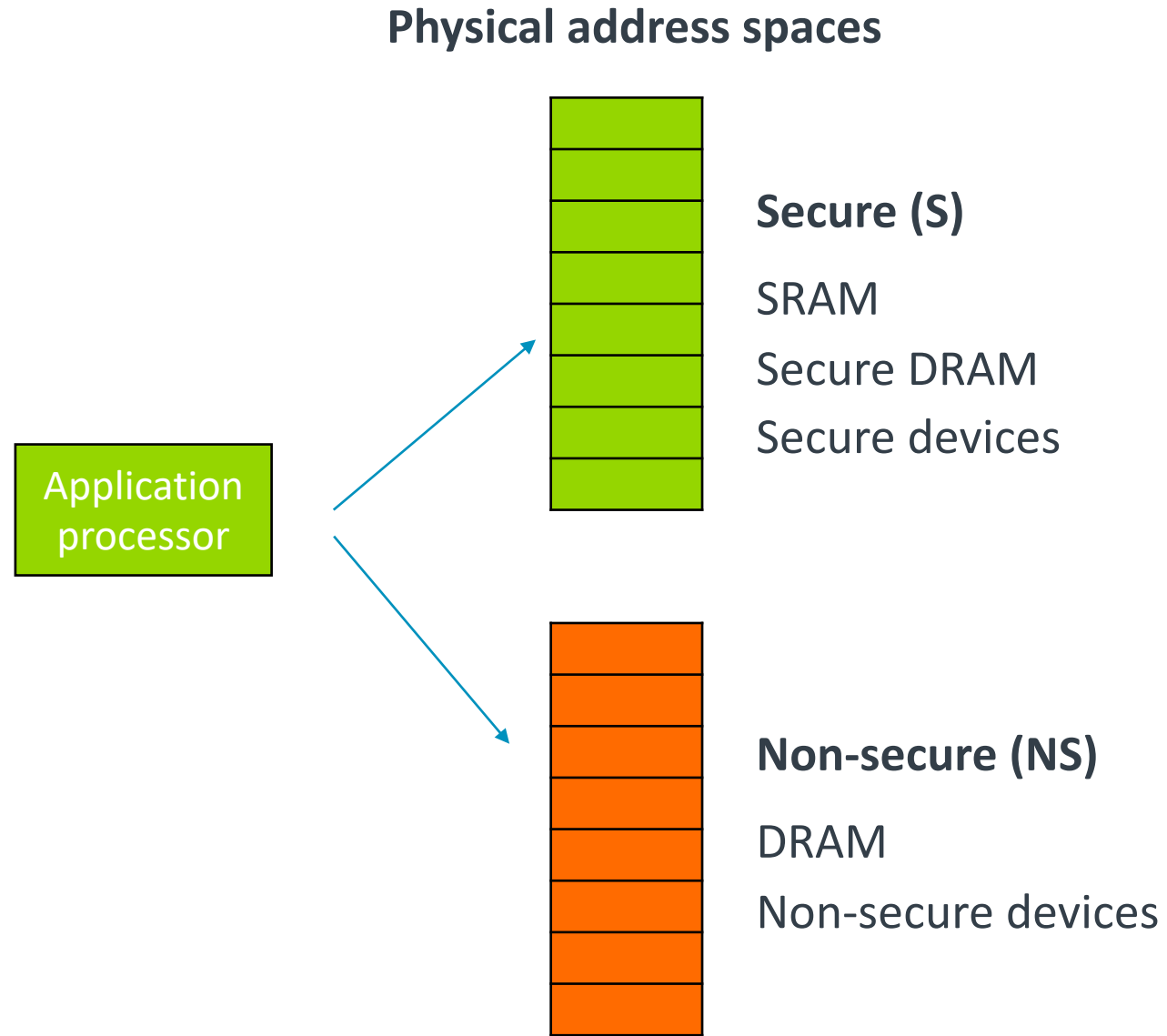
TrustZone



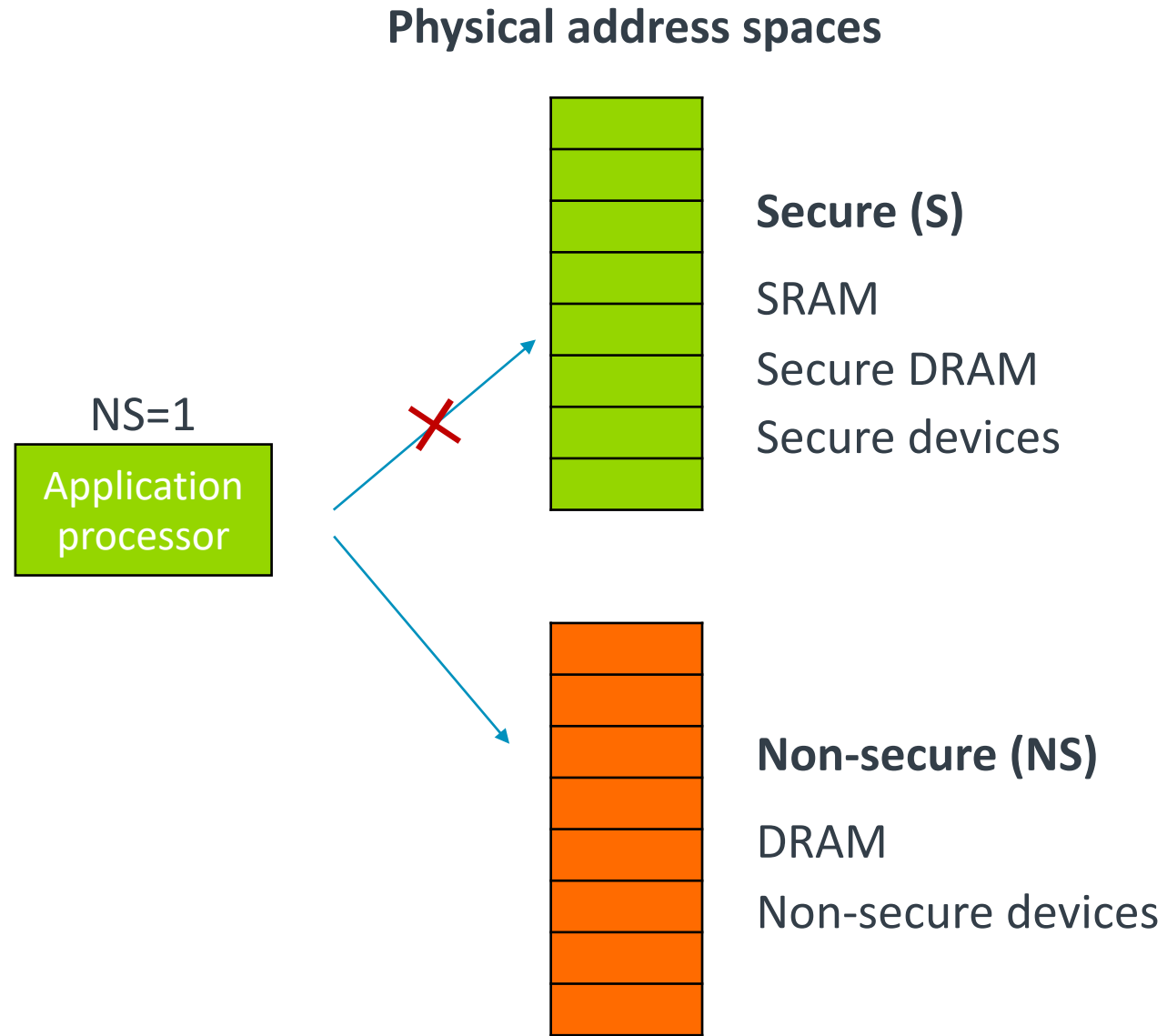
TrustZone



TrustZone

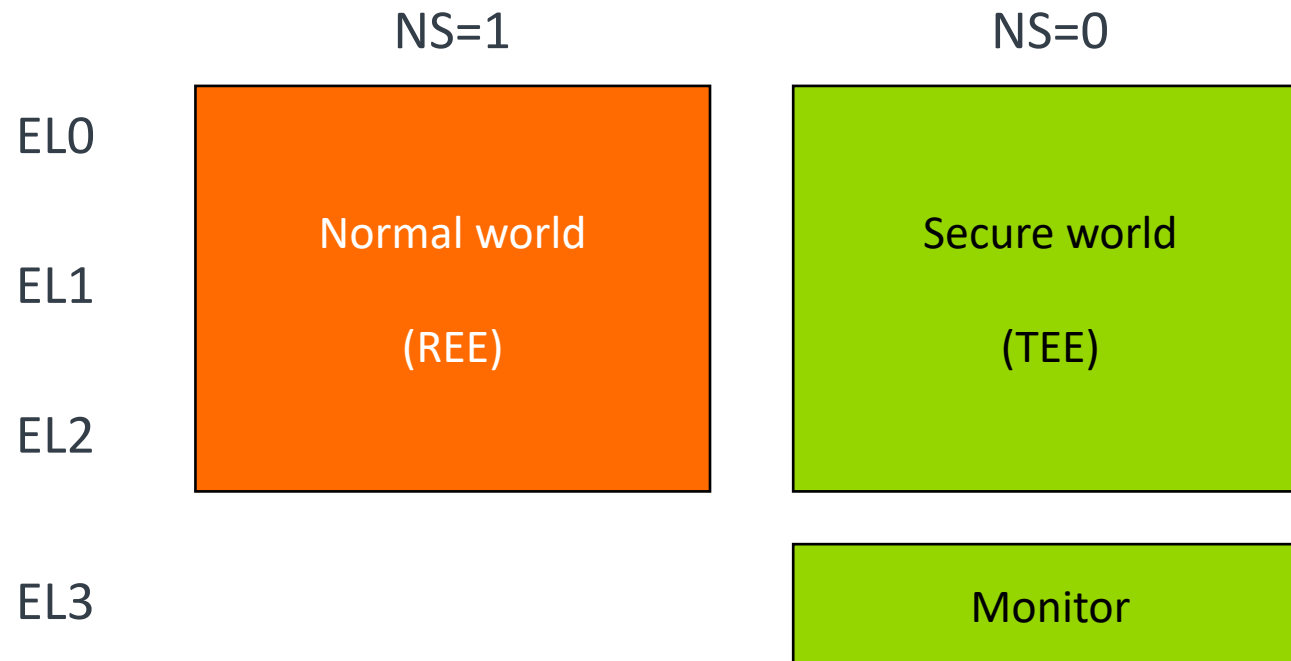


TrustZone



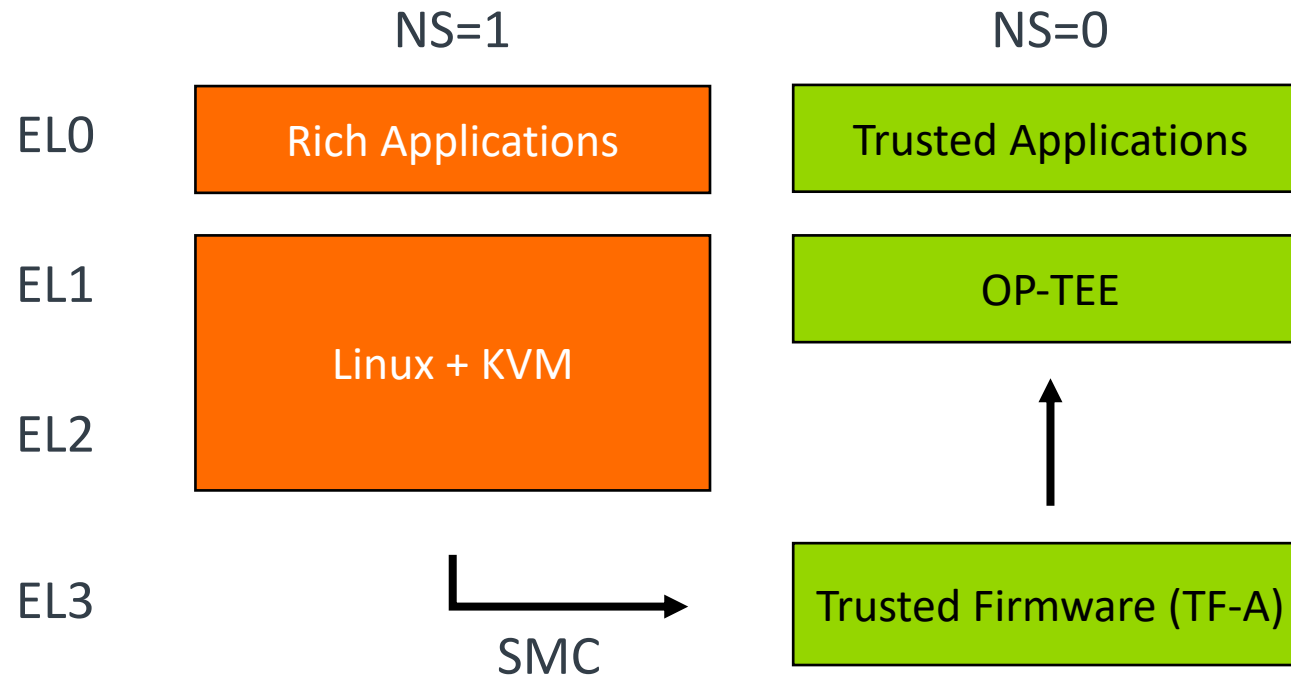
TrustZone: Typical firmware design

Typical two-world layout



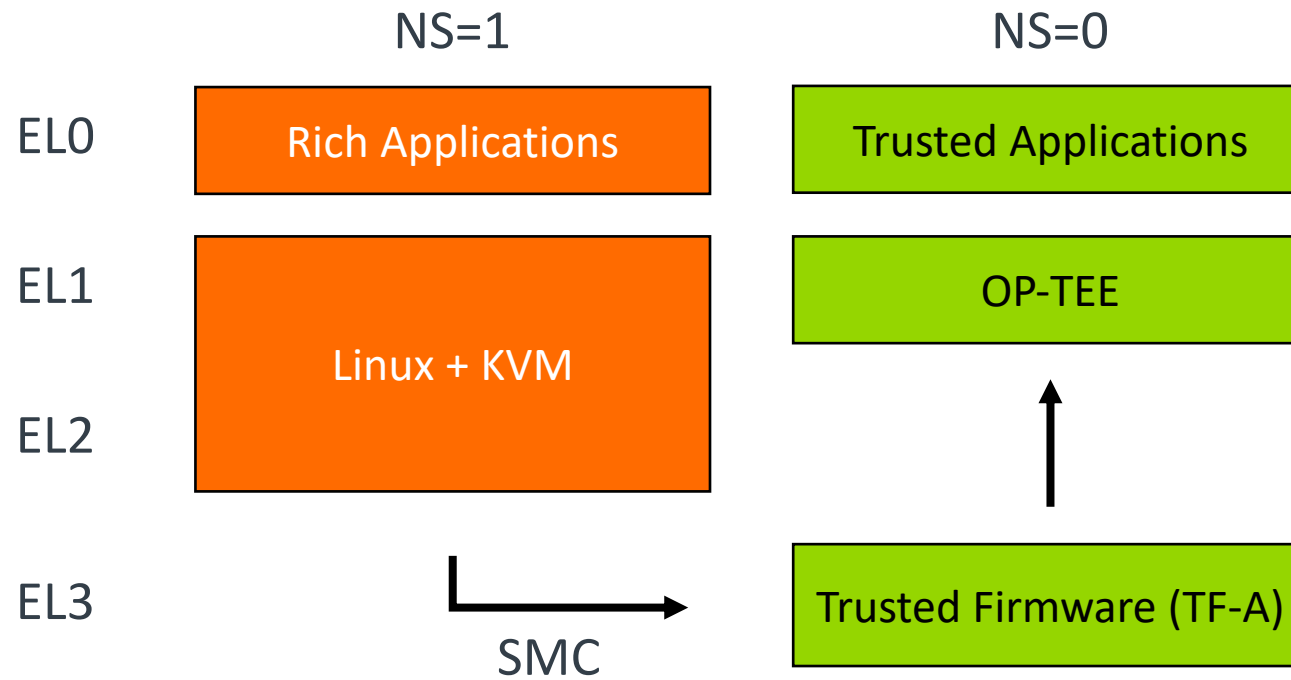
TrustZone: Typical firmware design

Typical two-world layout

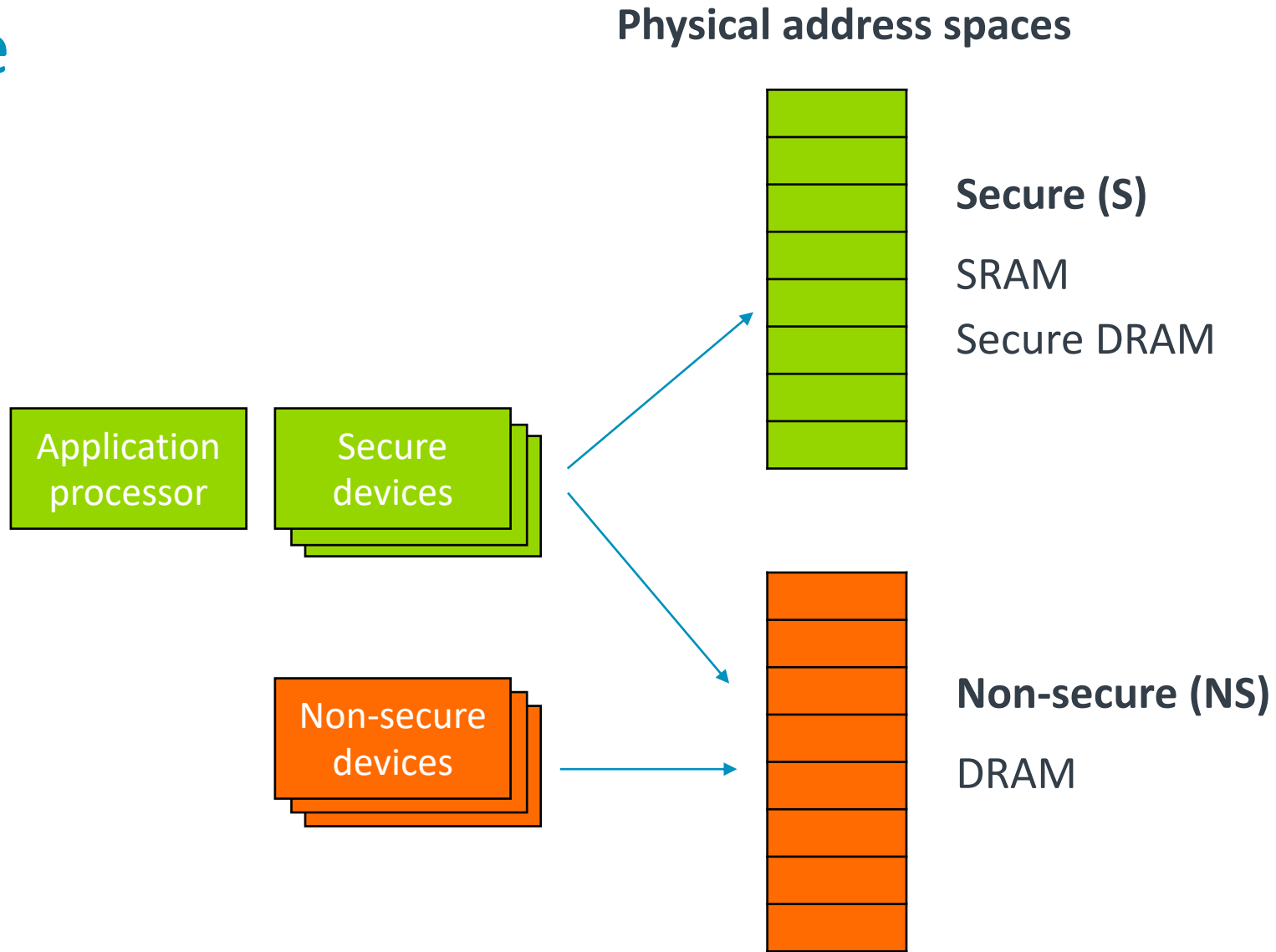


TrustZone: Typical firmware design

Coarse *world switch* minimizes the attack surface of the TEE



TrustZone

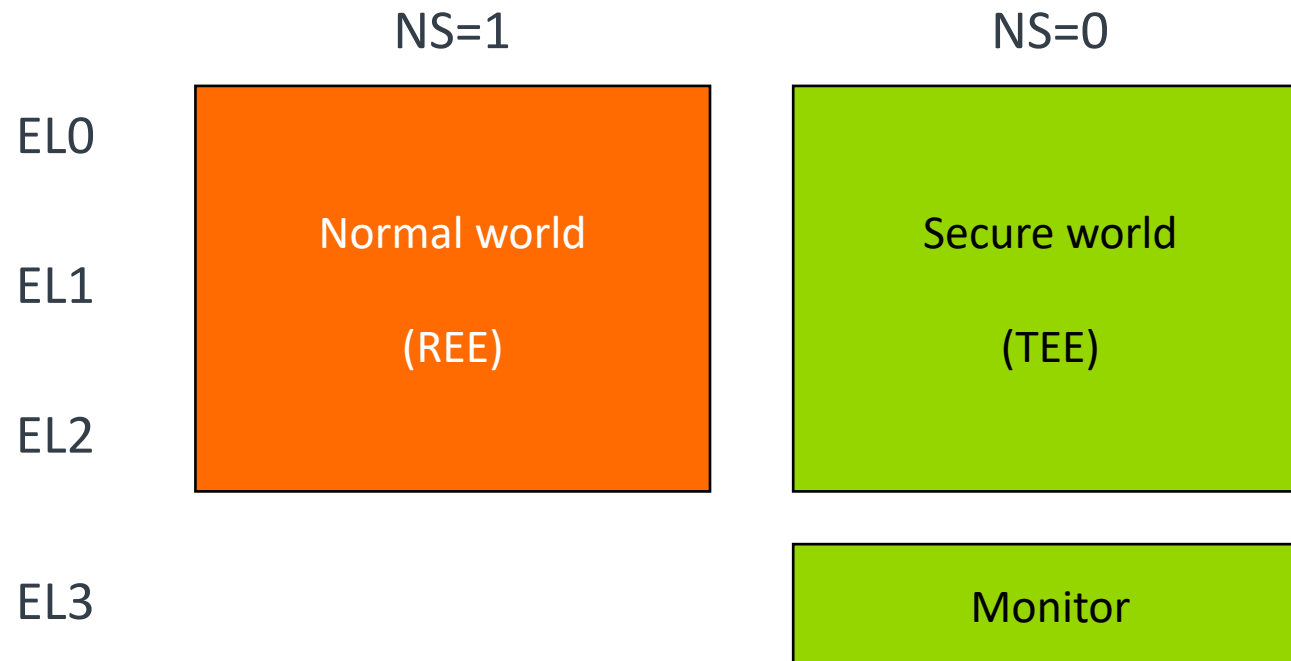


arm

seL4 + TrustZone

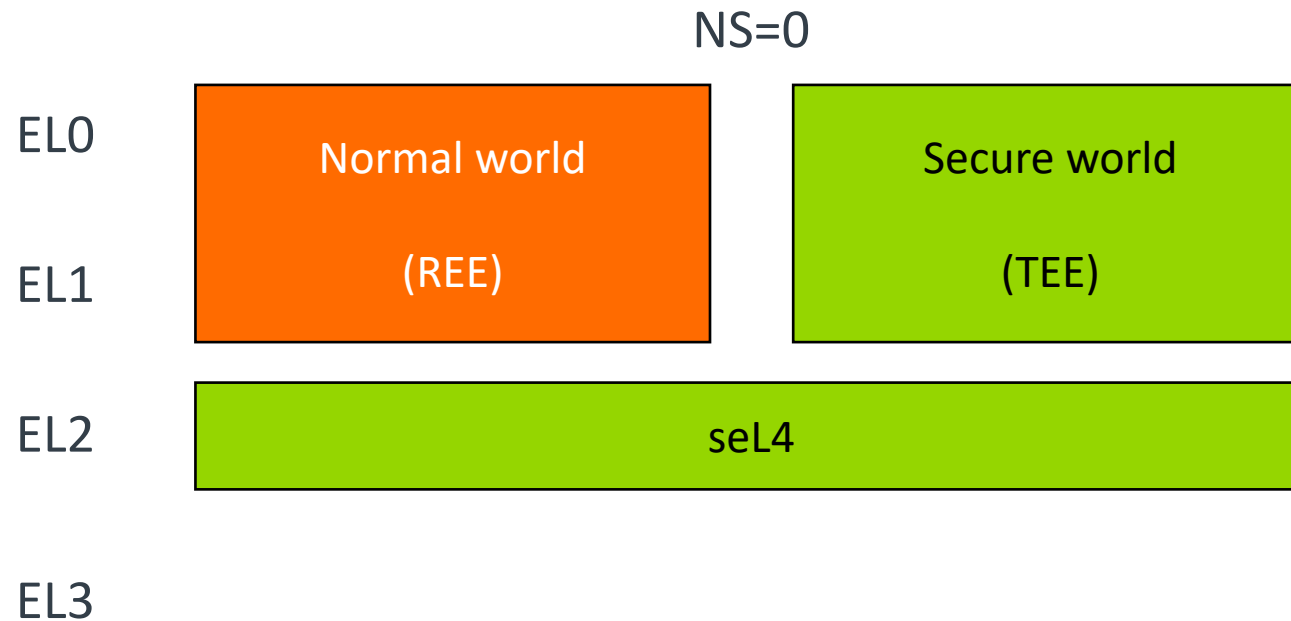
seL4 + TrustZone

Typical two-world layout



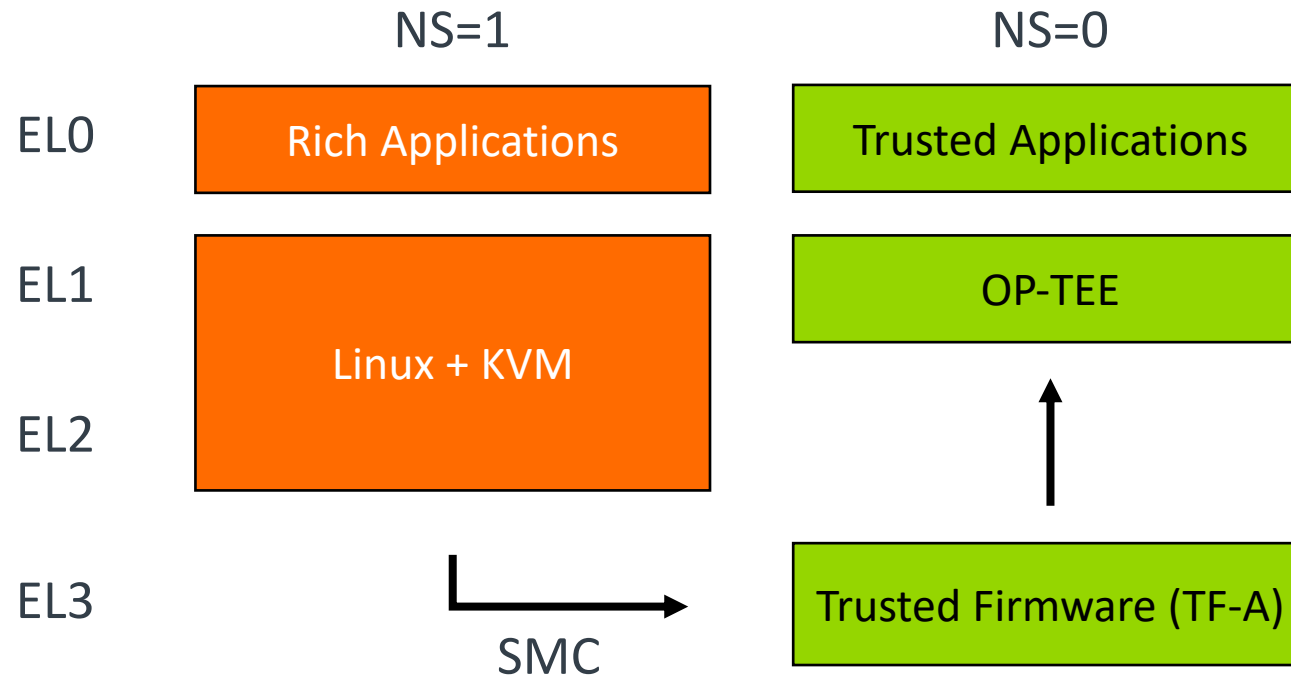
seL4 + TrustZone

A high-assurance hypervisor can isolate the REE to EL1 using just stage-2 translation tables



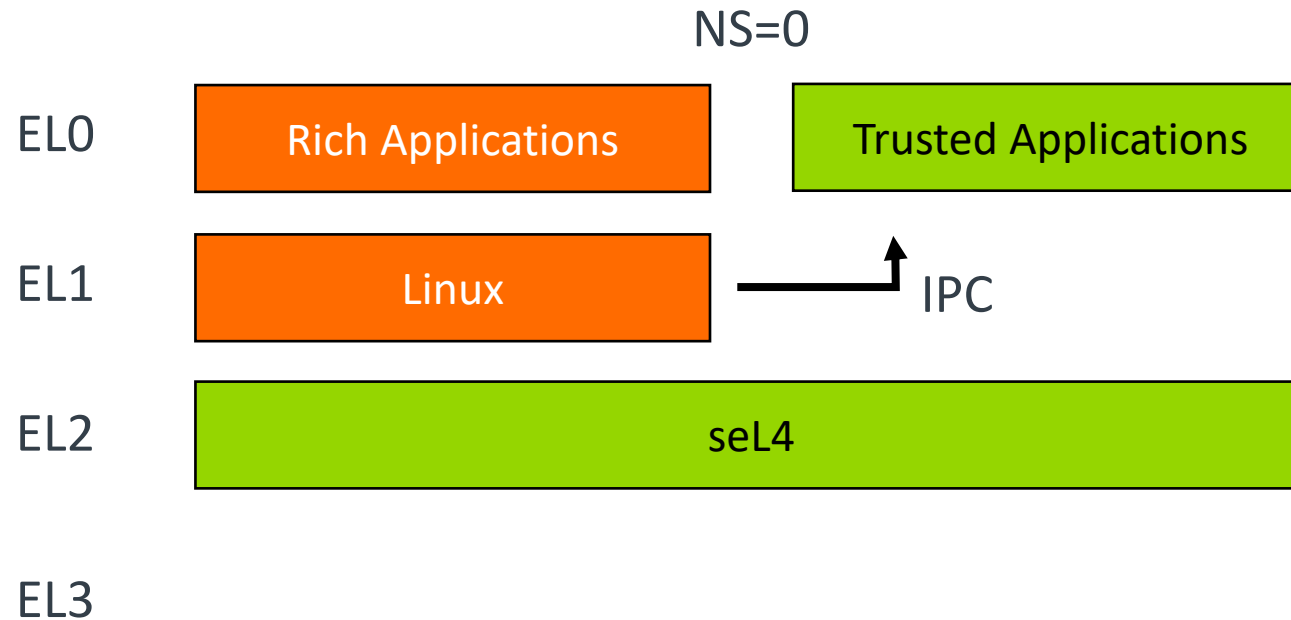
seL4 + TrustZone

Typical two-world layout



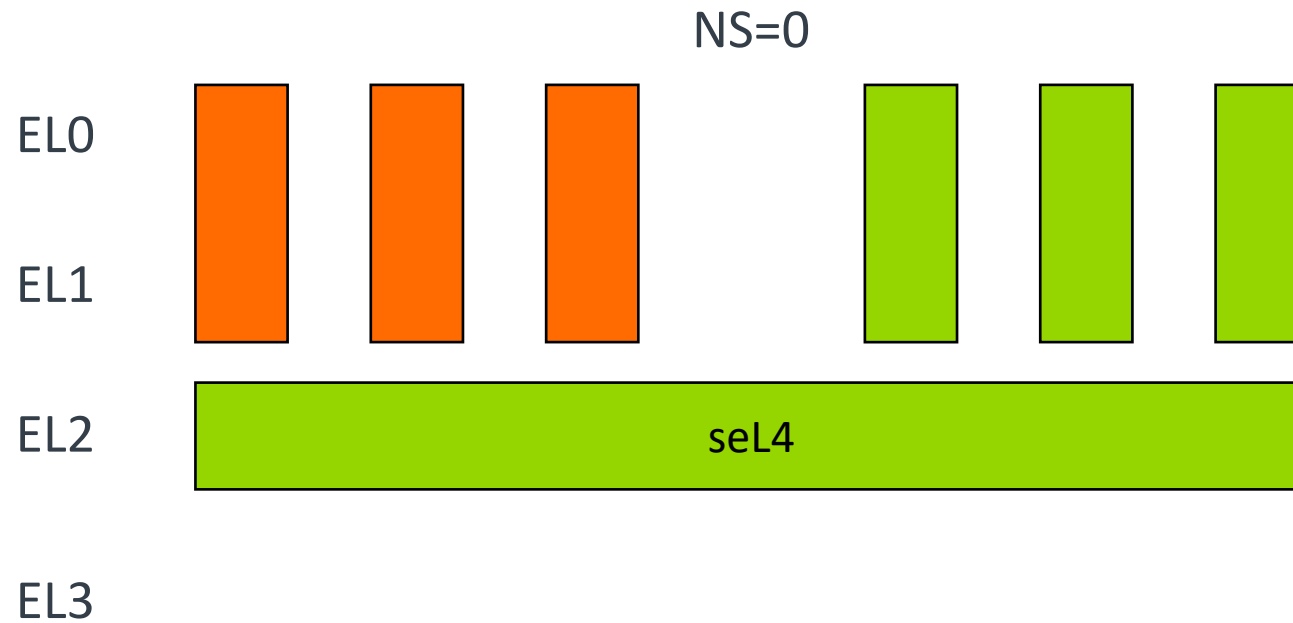
seL4 + TrustZone

seL4-based analog of typical two-world layout



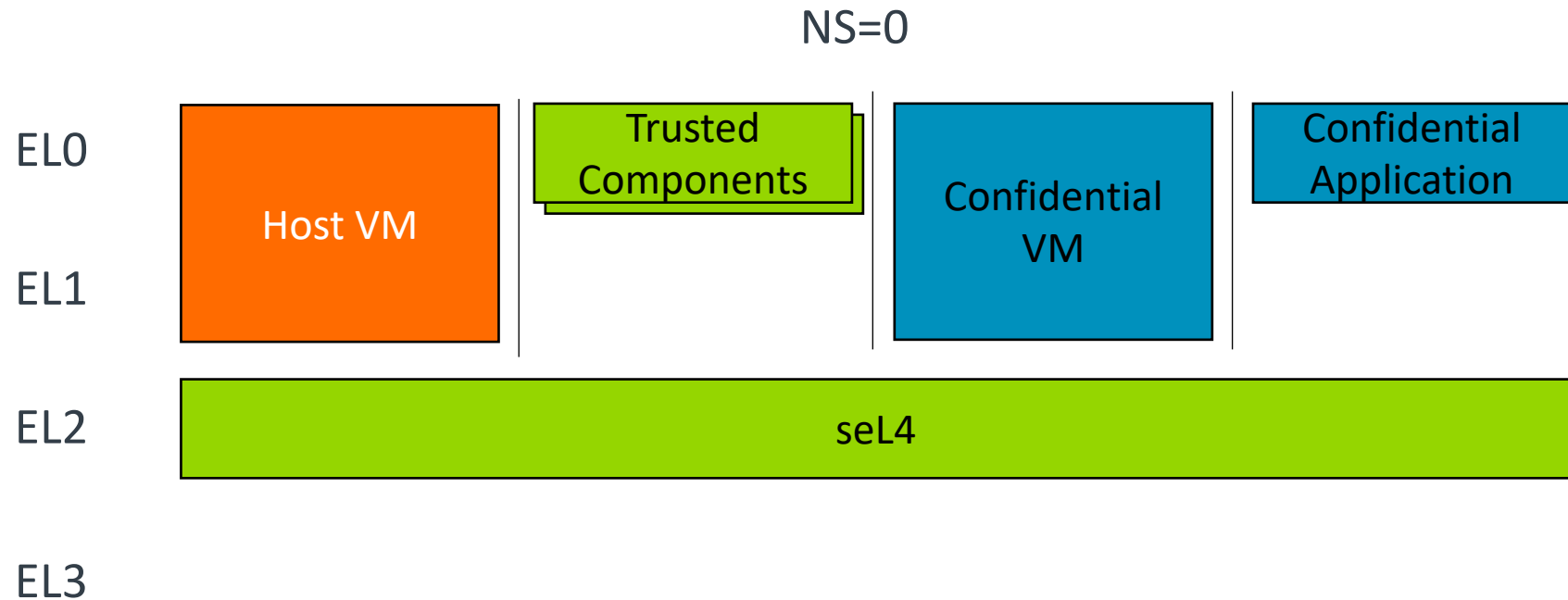
seL4 + TrustZone

seL4 can isolate with more granularity

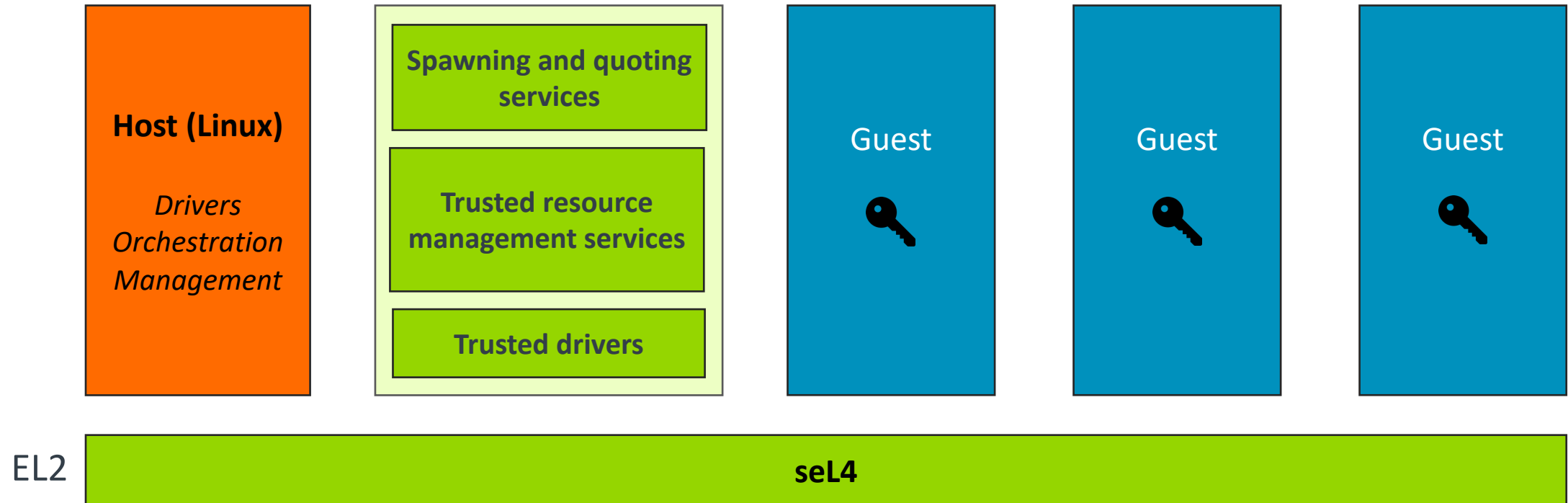


seL4 + TrustZone

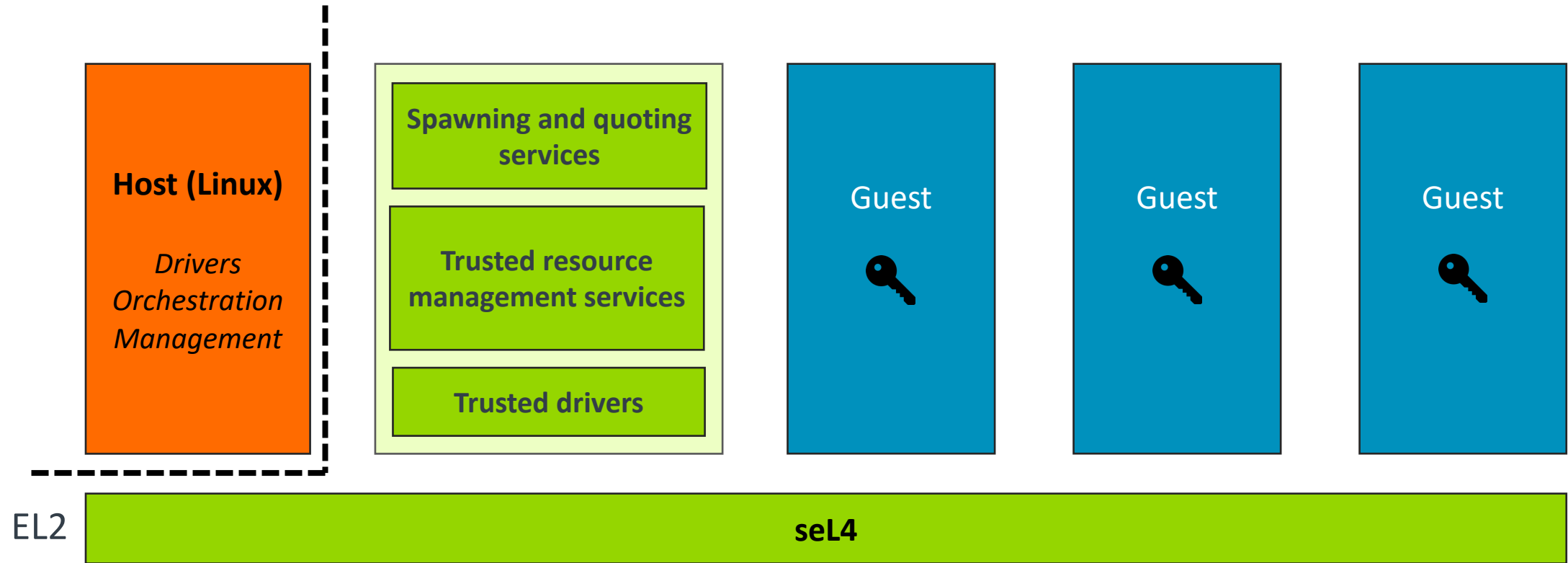
seL4 can isolate with more granularity



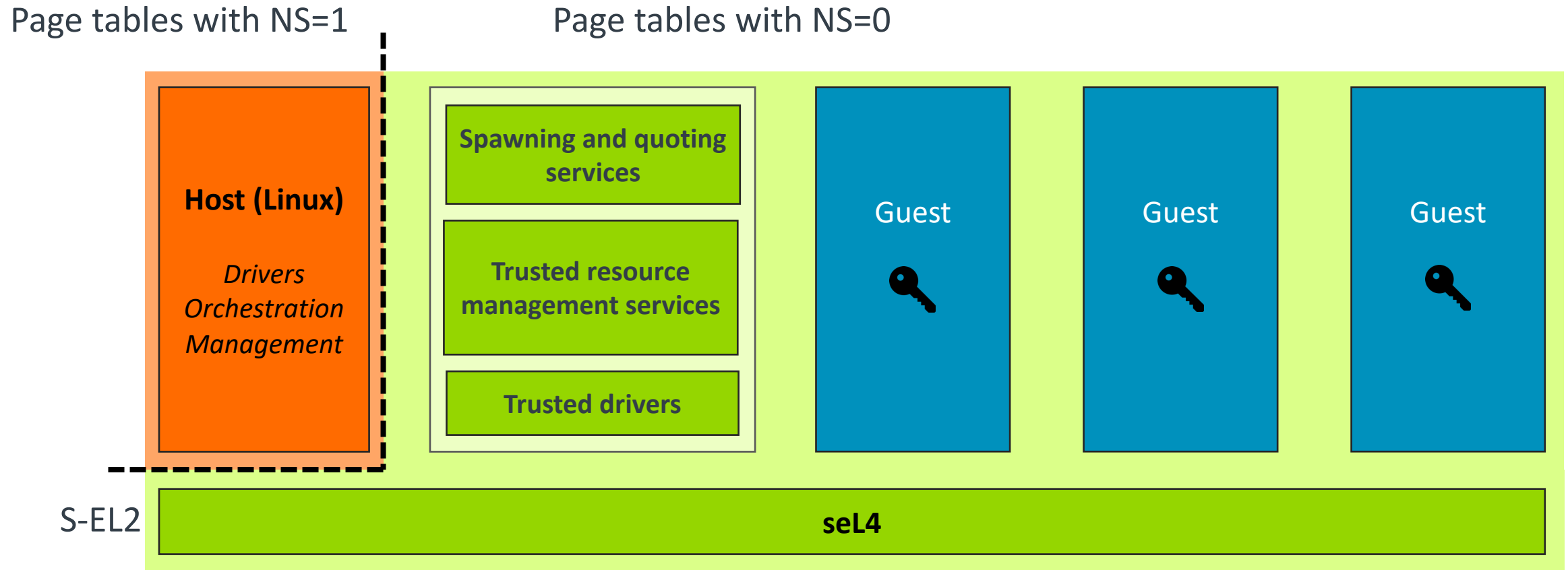
IceCap



IceCap



IceCap + TrustZone



seL4 + TrustZone: Awaiting Armv8.4-SecEL2

Specified in 2017

Expected to be available in silicon by early 2022

FEAT_SEL2, Secure EL2

FEAT_SEL2 permits EL2 to be implemented in Secure state. When Secure EL2 is enabled, a translation regime is introduced that follows the same format as the other Secure translation regimes.

This feature is not supported if EL2 is using AArch32.

This feature is mandatory in Armv8.4 implementations that implement both EL2 and Secure state.

The [ID_AA64PFR0_EL1.SEL2](#) field identifies the presence of FEAT_SEL2.

For more information, see:

- [Virtualization](#) on page D1-2318.
- [The VMSEv8-64 address translation system](#) on page D5-2534.

<https://developer.arm.com/documentation/ddi0487/latest/>

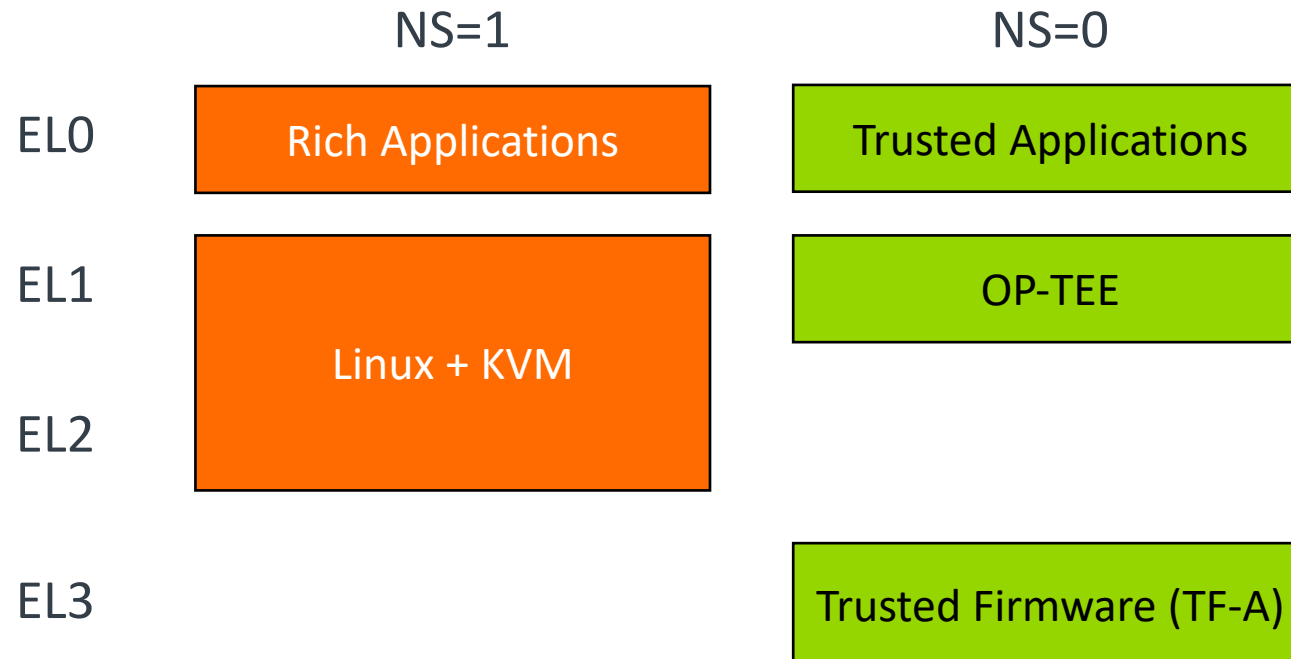
arm

seL4 + TrustZone

Beyond IceCap

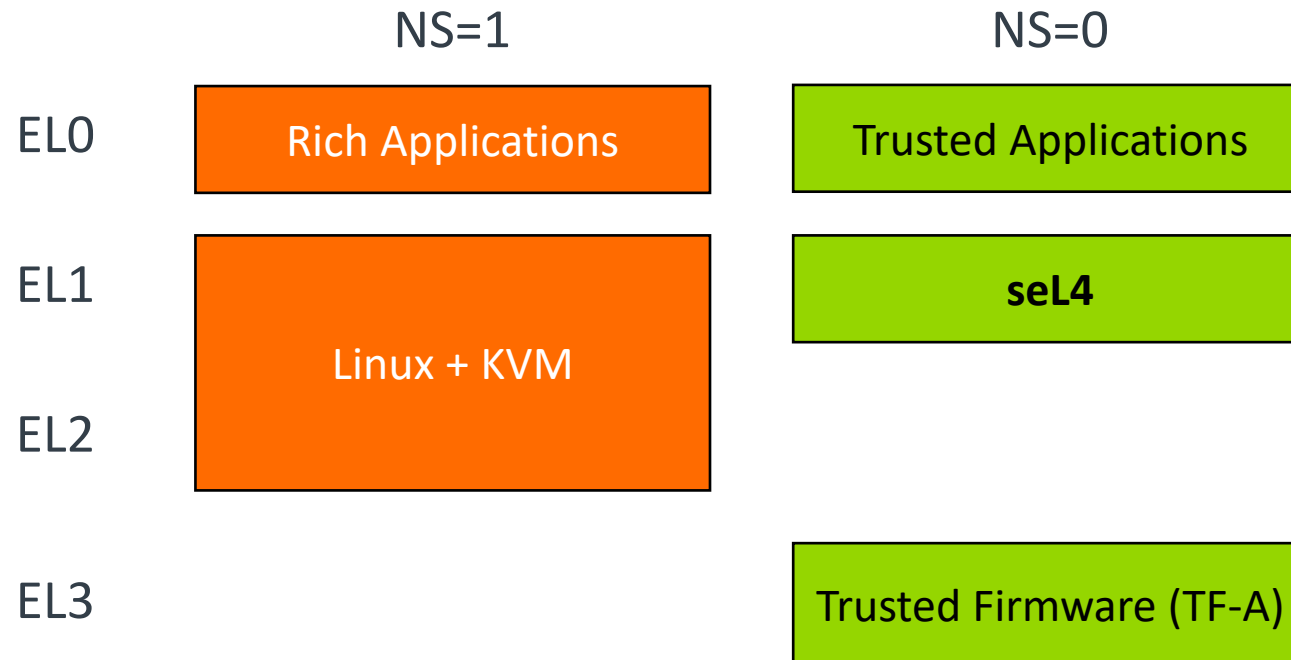
seL4 + TrustZone

Typical two-world layout



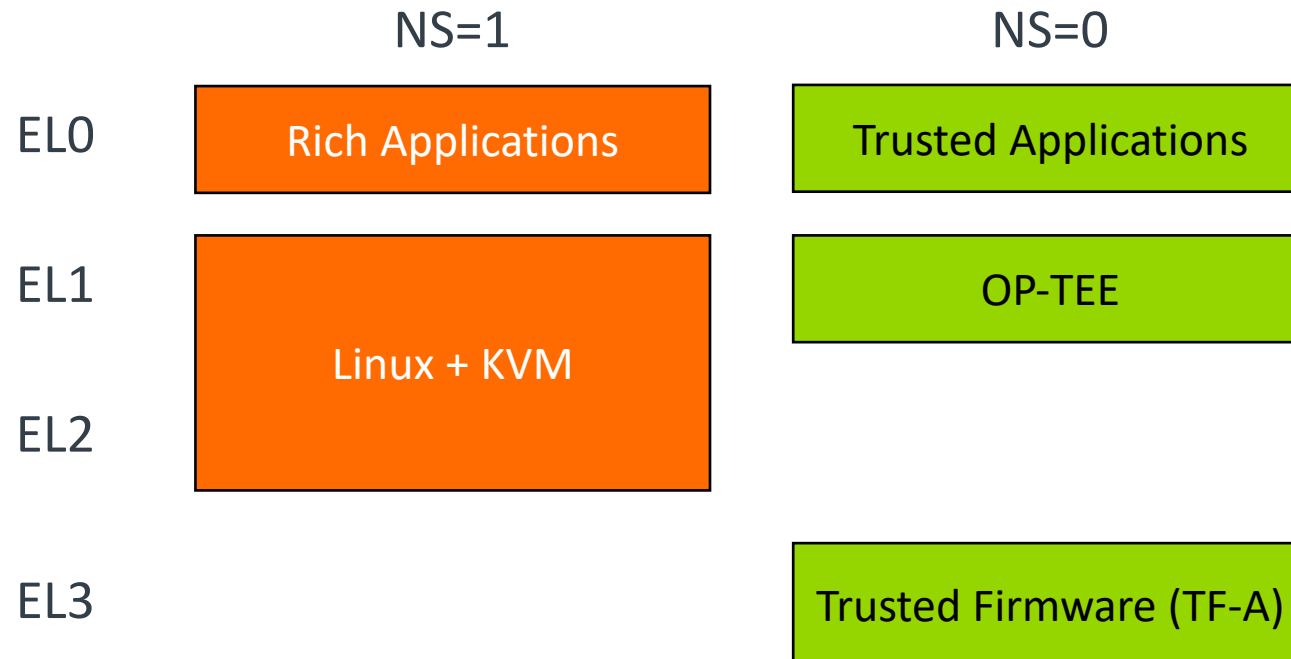
seL4 + TrustZone

Typical two-world layout with seL4 as the Trusted OS



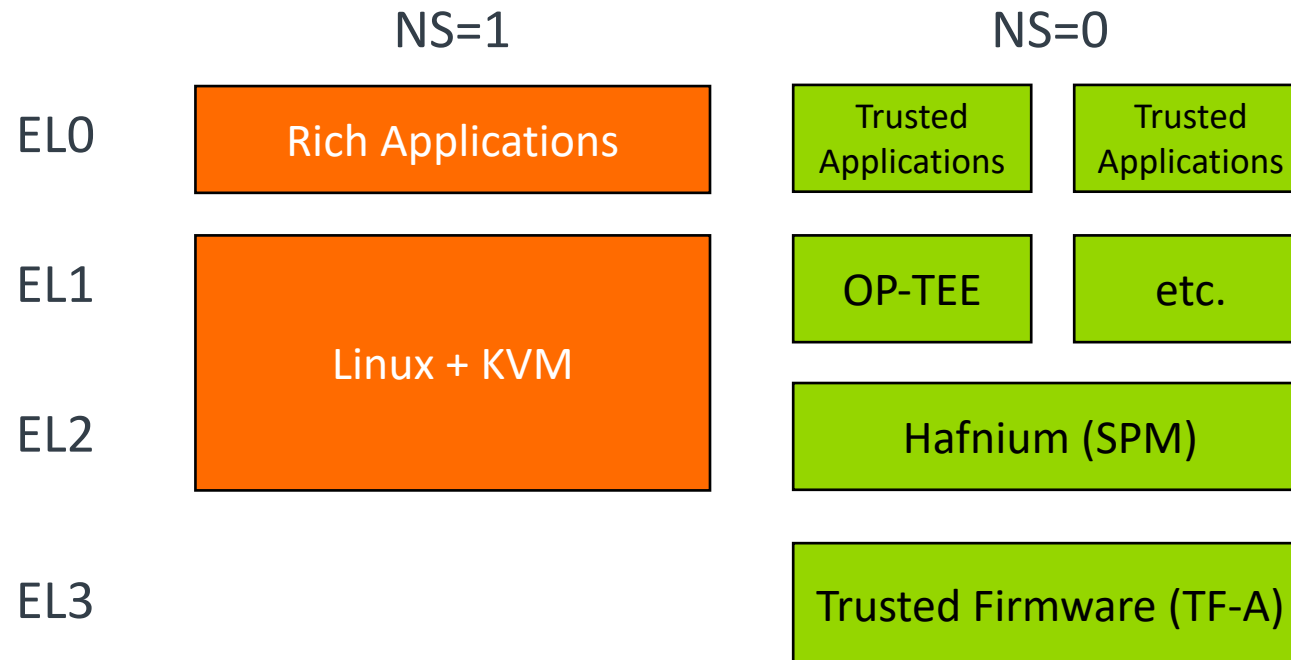
seL4 + TrustZone

Typical two-world layout



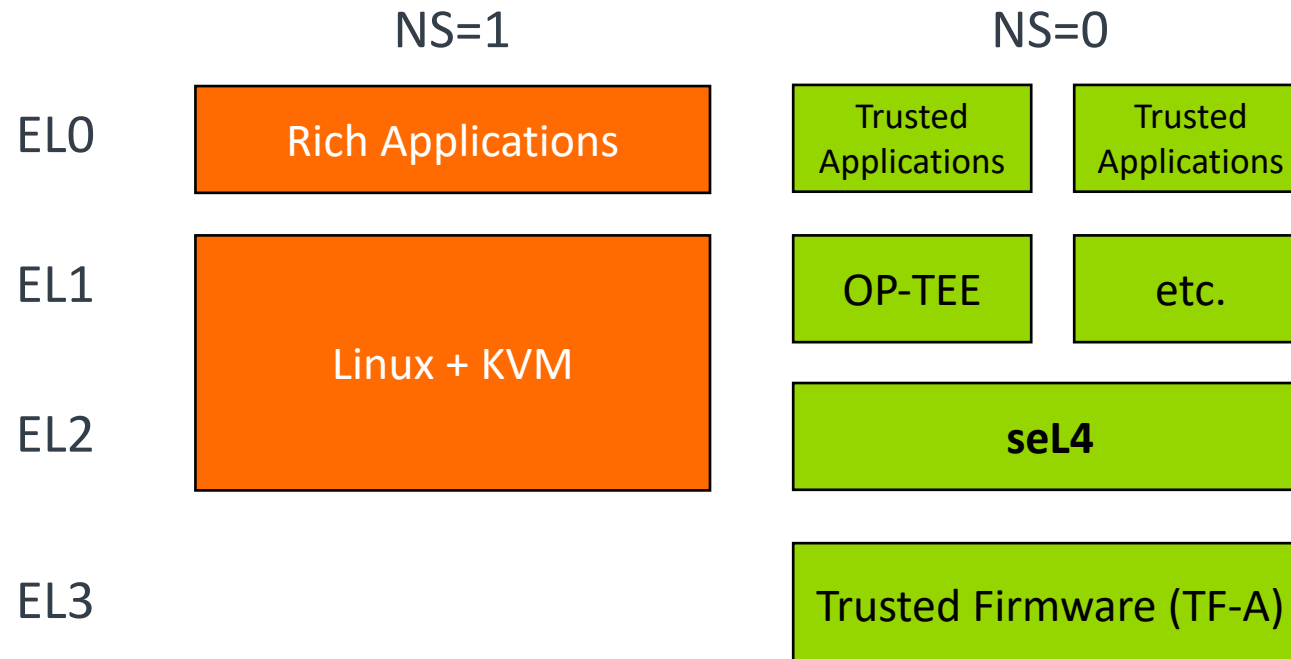
seL4 + TrustZone

Typical two-world layout with a Secure Partition Manager (SPM)



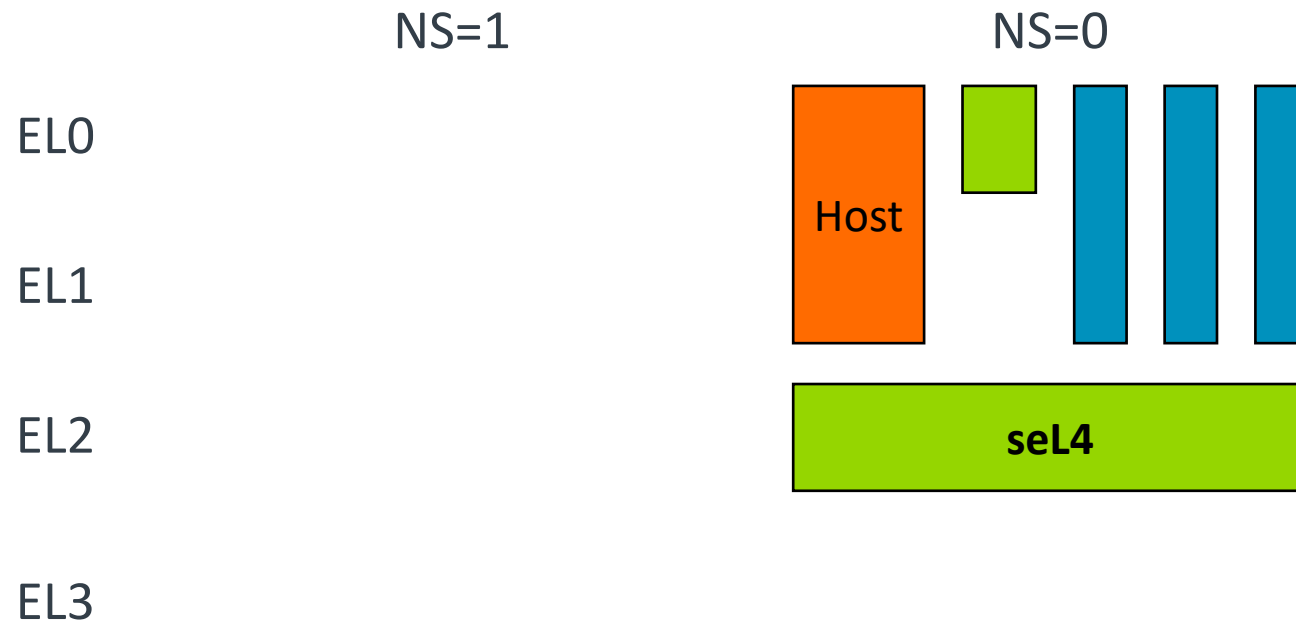
seL4 + TrustZone

Typical two-world layout with seL4 as a SPM



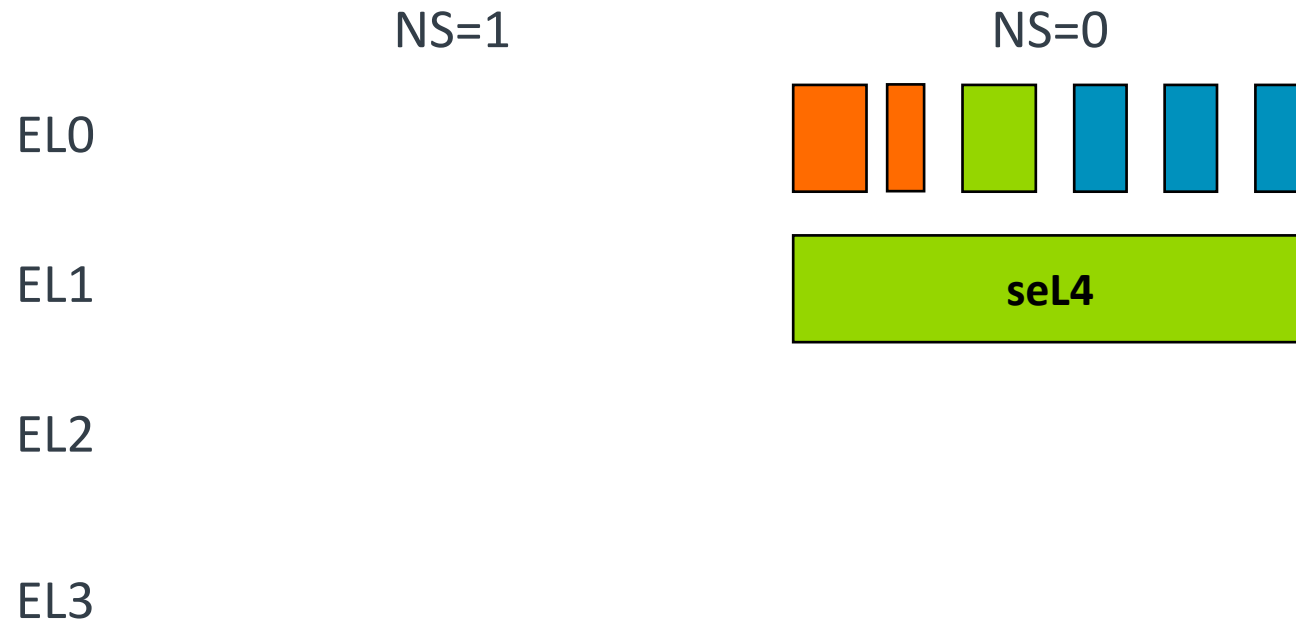
seL4 + TrustZone

IceCap in context



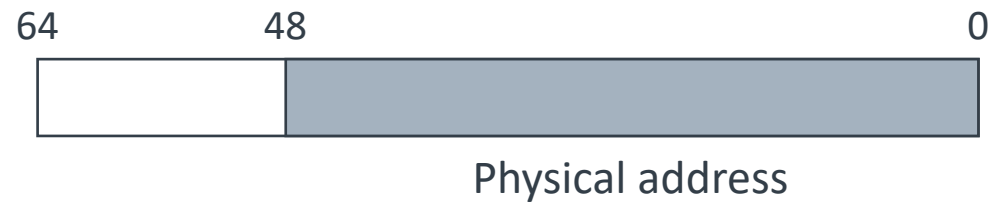
seL4 + TrustZone

Typical seL4-based system, now making use of TrustZone



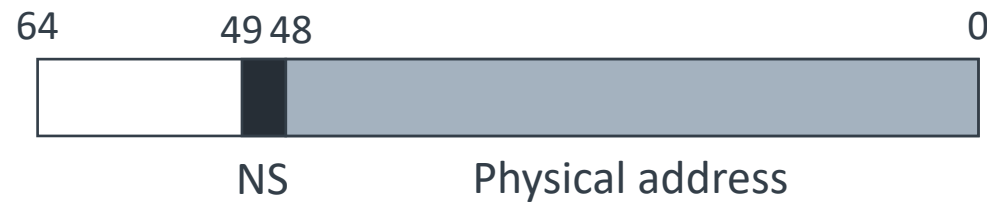
seL4 + TrustZone: Adding support to the kernel

Untyped memory



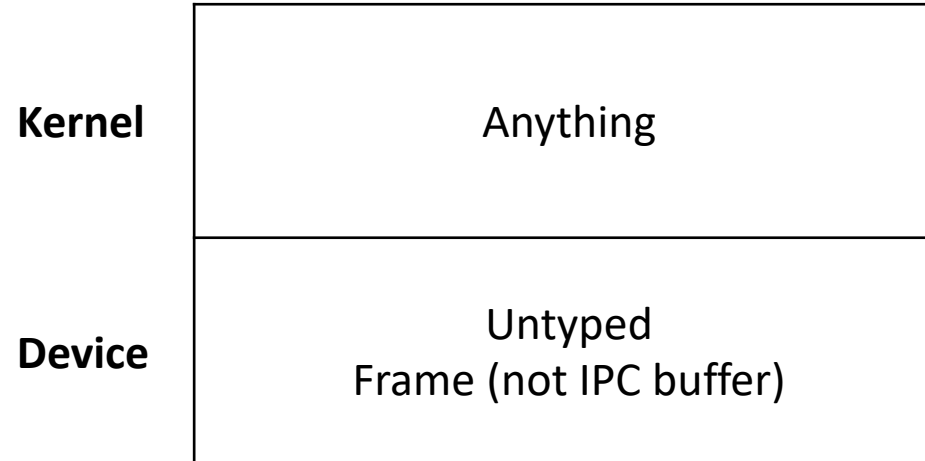
seL4 + TrustZone: Adding support to the kernel

Untyped memory must span both the secure and non-secure physical address spaces



seL4 + TrustZone: Adding support to the kernel

Untyped memory



seL4 + TrustZone: Adding support to the kernel

Untyped memory must span both the secure and non-secure physical address spaces

	Secure	Non-secure
Kernel	Anything	Untyped Frame
Device	Untyped Frame (not IPC buffer)	

seL4 + TrustZone: Adding support to the kernel

S-EL1 kernel requires ++200/--80

S-EL2 kernel requires more invasive changes

- Stage-1 and stage-2 translation tables are architecturally distinct on AArch64, but seL4 currently does not distinguish between the two
- Stage-2 translation tables lack NS bit

Discussion at <https://sel4.discourse.group>, to result in RFC

arm

Nick Spinale

nick.spinale@arm.com

<https://gitlab.com/arm-research/security/icecap/icecap>

Thank You

Danke

Merci

谢谢

ありがとう

Gracias

Kiitos

감사합니다

धन्यवाद

شكراً

תודה