Linux Plumbers Conference

Vienna, Austria | September 18-20, 2024

Updates on RISC-V Confidential Computing (CoVE) ISA, non-ISA

Ravi Sahita, Atish Patra
Rivos Inc.

RISC-V Ratification Status of Confidential Computing ISA spec

ISA: Supervisor Domains and CoVE Spec status

Depends on RISC-V H-extension, AIA (for IMSIC) and IOMMU (for device assignment)
ISA spec has completed 2 week TG review - there will be a public review phase after STABLE.
spec repos: https://github.com/riscv/riscv-smmtt/releases/download/v0.1/smmtt-spec.pdf

Qemu, Open SBI POC

Initial Smmtt implementation per latest version of the spec in OpenSBI and QEMU. (credit Gregor Haas)
Repo with build/run/debug instructions can be found at https://github.com/grg-haas/smmtt
Contains automated tests and CI for these new features. Emulation of IOMTT and Smsdia is pending.

Want to get feedback from community on the approach for:
- ISA for Isolation of Memory, IO Interrupts [and Device functions]
- non-ISA ABI between host and TSM

See longer discussion in these slides - https://static.sched.com/hosted_files/lsseu2024/2b/LSSEU24-RISC-V%20CoVE.pdf

RISC-V Ratification Status of Confidential Computing non-ISA specs

CoVE spec and TSM POC

ABI spec defines TSM ← → Host interface and TSM ← → TVM (Guest) interfaces.
v0.7 spec released as RC - completed 2 week TG review - will be revised to STABLE after comments addressed: https://github.com/riscv-non-isa/riscv-ap-tee/releases/download/v0.7/riscv-cove.pdf
there will be a public review phase after STABLE.

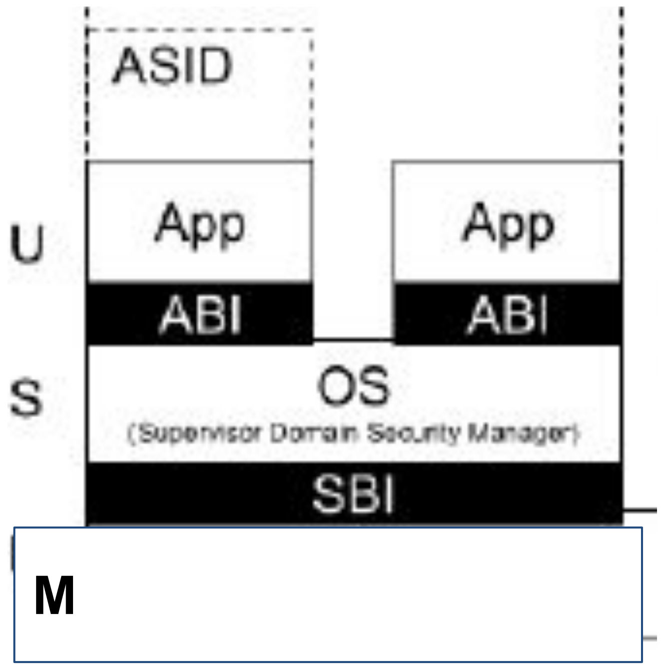
Existing open source Rust TSM implementation of the CoVE ABI, called Salus:
https://github.com/rivosinc/salus

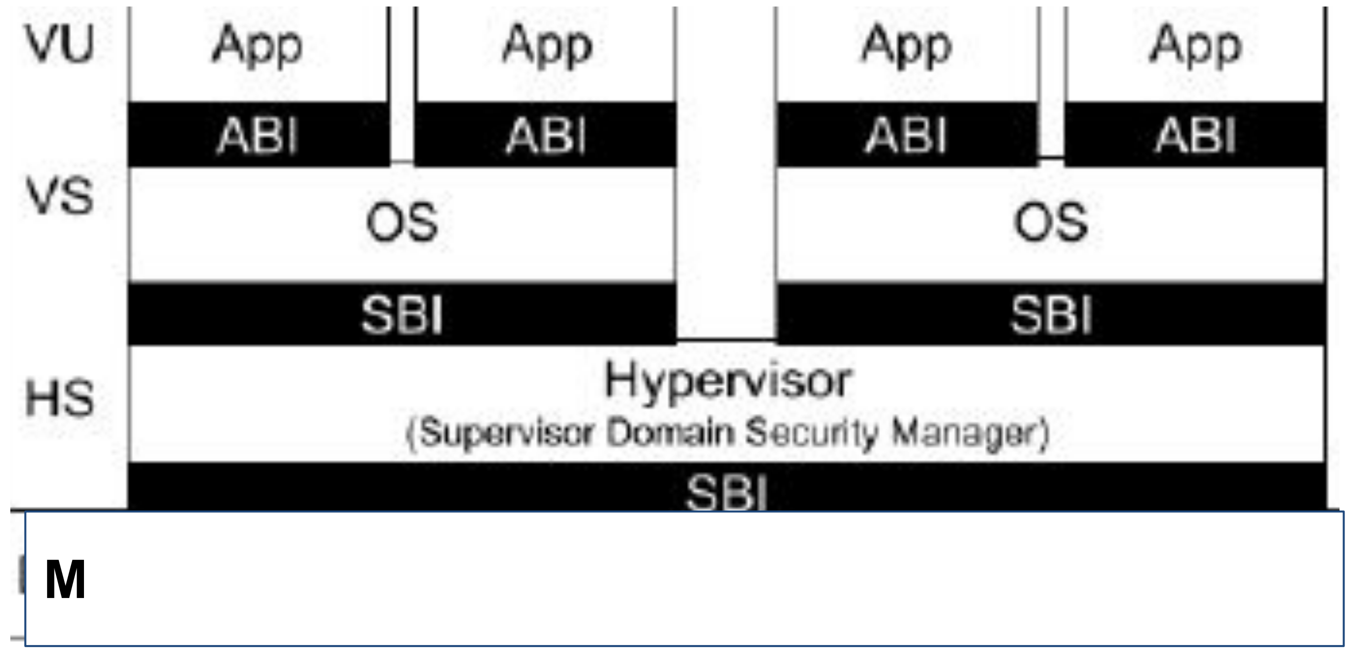
CoVE-IO spec

v0.2 released:
https://github.com/riscv-non-isa/riscv-ap-tee-io/releases/download/v0.2.0/riscv-cove-io-v0.2.0.pdf
Related discussion during PCIe authentication BOF

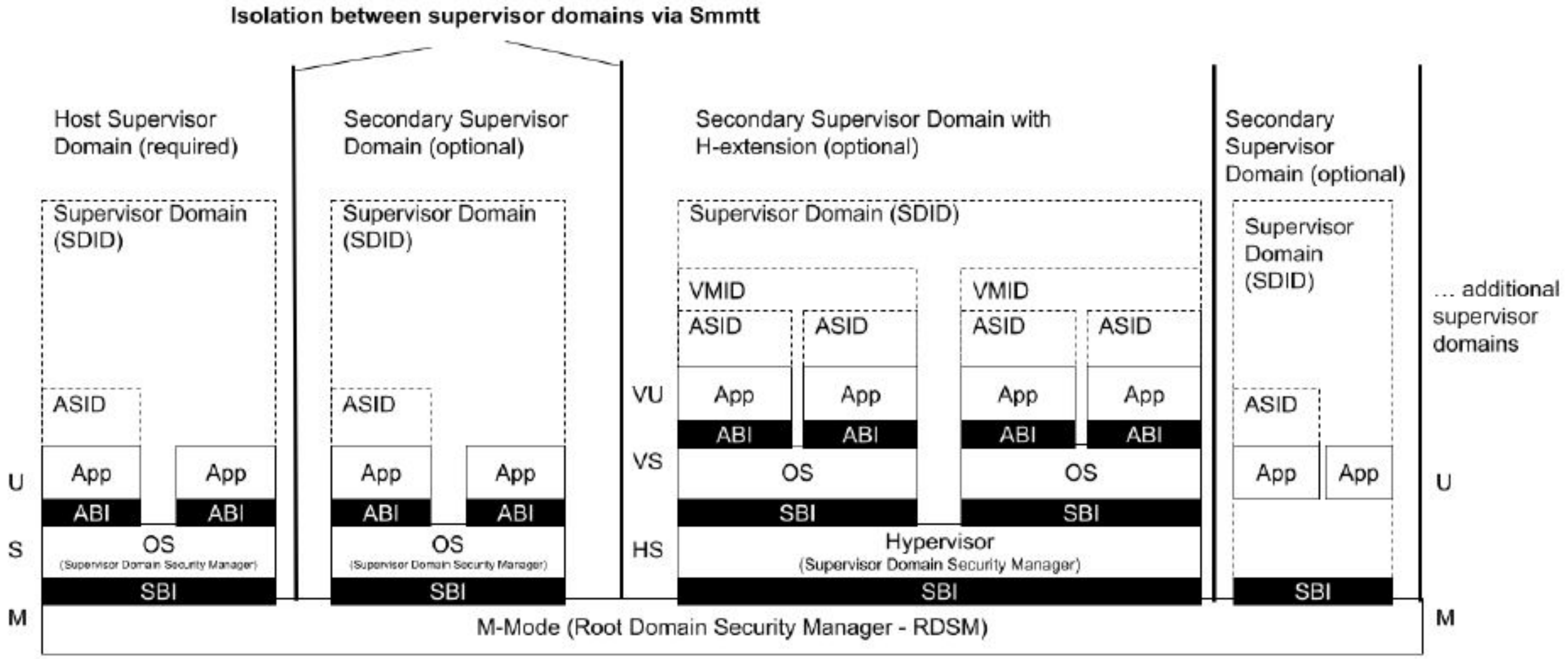
See longer discussion in these slides - https://static.sched.com/hosted_files/lsseu2024/2b/LSSEU24-RISC-V%20CoVE.pdf

Existing RVI Priv ISA Modes





Priv ISA Extension - Supervisor Domains

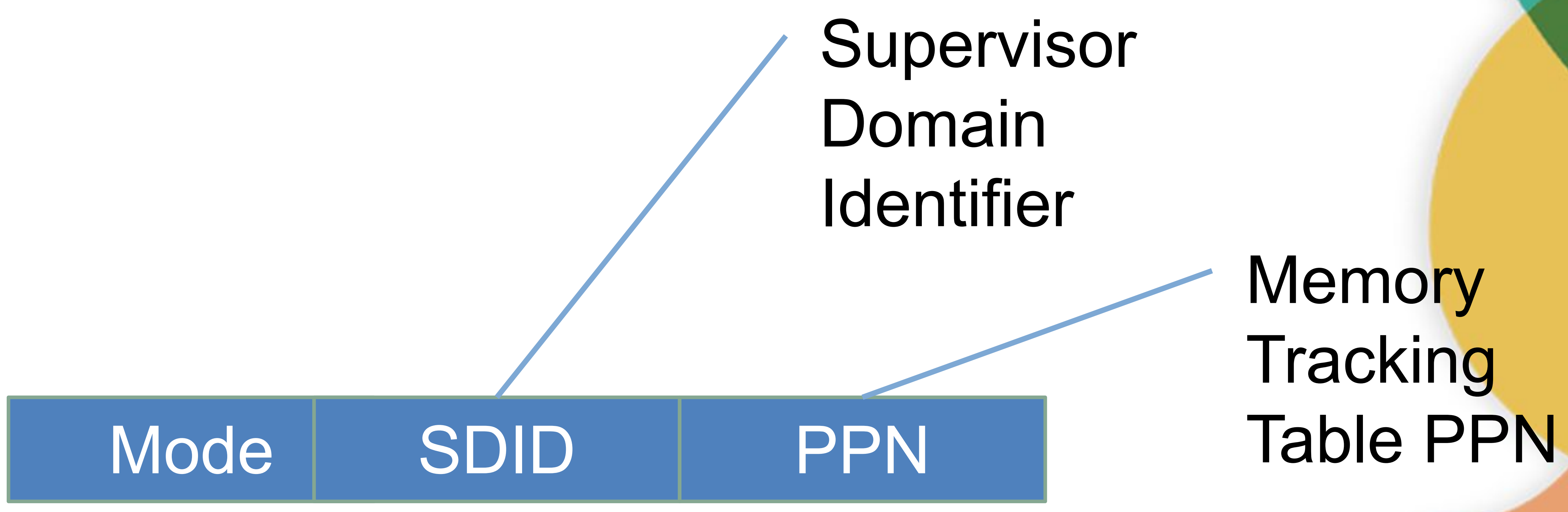

Isolation between supervisor domains via Smmtt
Host Supervisor Domain (required)
Secondary Supervisor Domain (optional)
Secondary Supervisor Domain with H-extension (optional)
Secondary Supervisor Domain (optional)
Supervisor Domain (SDID)
Supervisor Domain (SDID)
Supervisor Domain (SDID)
Supervisor Domain (SDID)
VMID
VMID
ASID
ASID
ASID
ASID
ASID
ASID
ASID
... additional supervisor domains
VU
VS
HS
U
S
M
U
M
App
ABI
OS
(Supervisor Domain Security Manager)
SBI
Hypervisor
(Supervisor Domain Security Manager)
M-Mode (Root Domain Security Manager - RDSM)

Priv ISA extension approach - Supervisor Domains aka “Smsdid”

- per-hart CSRs to manage Supervisor Domain Identifier assignment
    - to manage access-control properties on harts (extends VMID, ASID)
    - Introduces new physical memory isolation programmed via a HW Memory Tracking Table
    - works with legacy PMP

- M-mode SD fence instructions
    - MFENCE.SPA & MINVAL.SPA

- M-mode CSRs msdcfg used to configure other assignments for SD
    - QoS, Debug, Trace, Interrupt controller



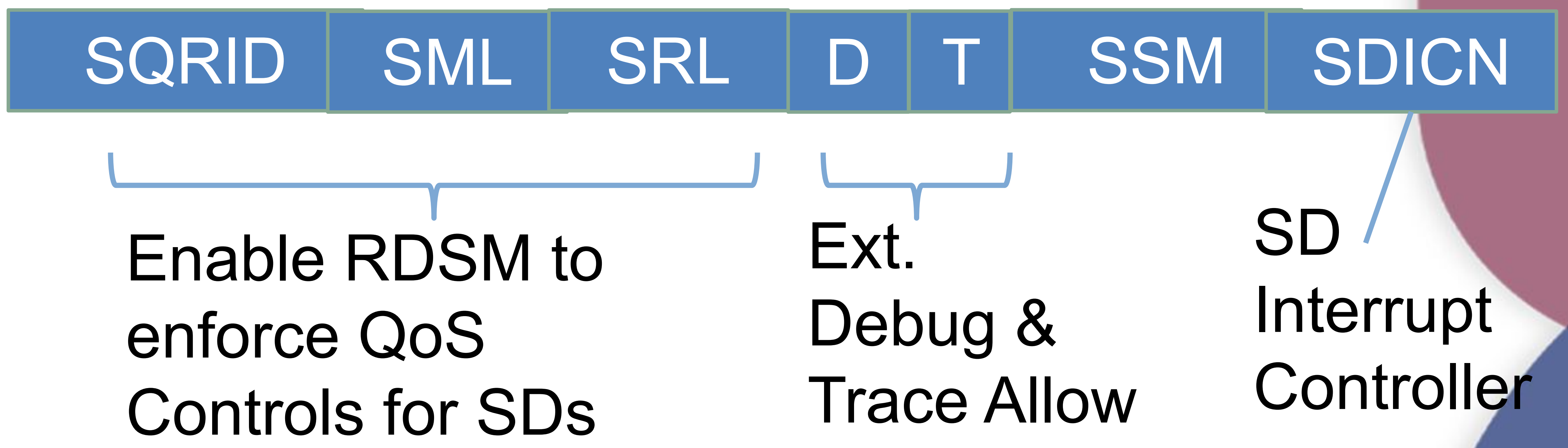
Mode	SDID	PPN

Supervisor Domain Identifier

Memory Tracking Table PPN

Memory Tracking Table Pointer register

M-mode Supervisor domain config register



SQRID	SML	SRL	D	T	SSM	SDICN

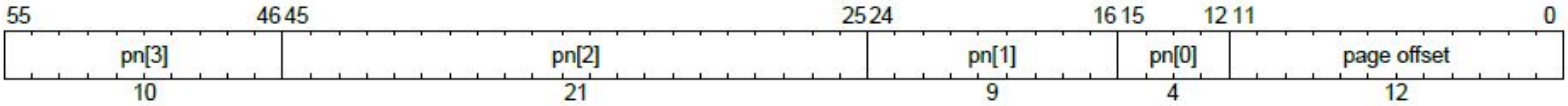
Enable RDSM to enforce QoS Controls for SDs

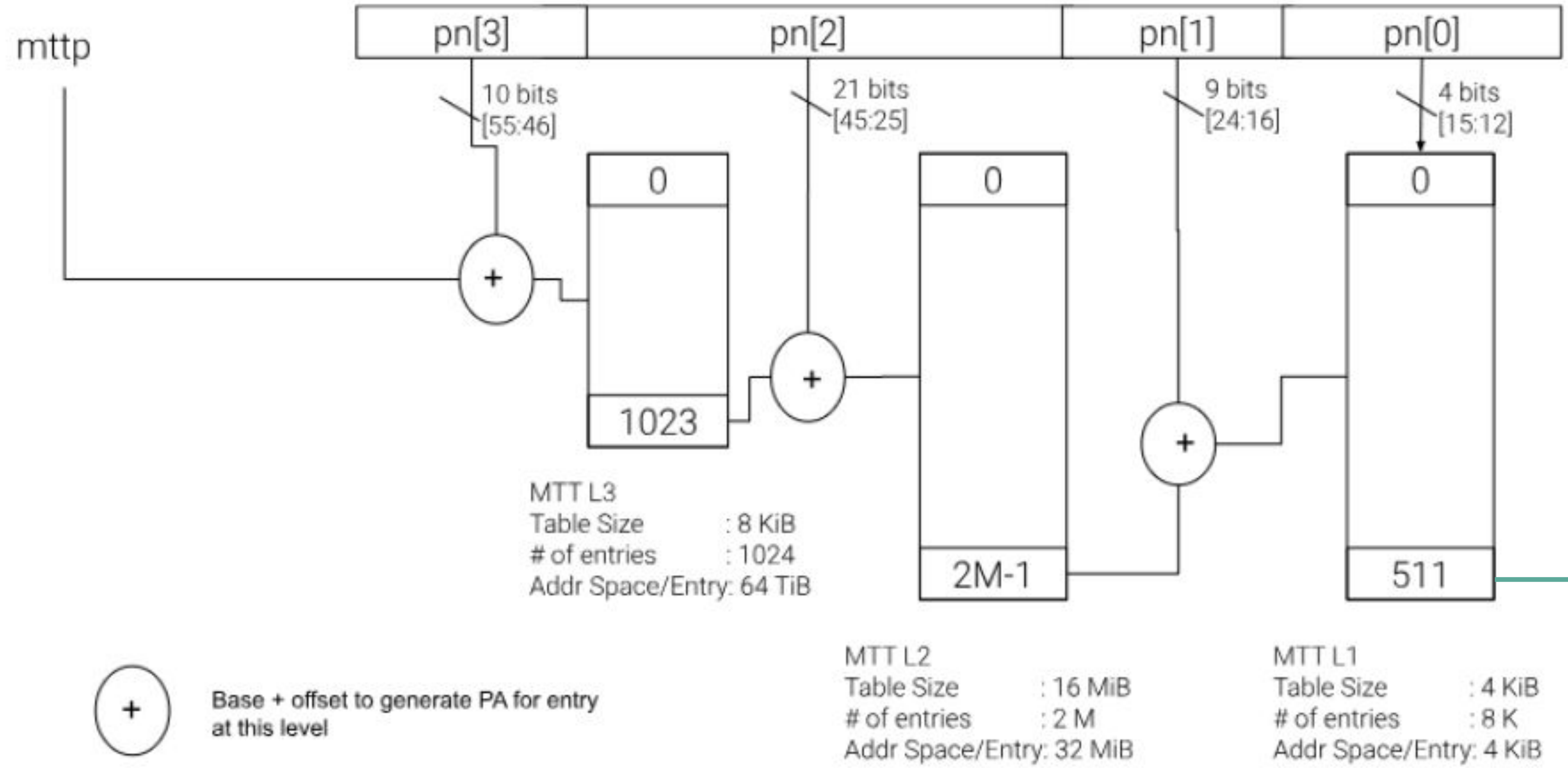
Ext. Debug & Trace Allow

SD Interrupt Controller

Supervisor Domains memory isolation approach - “Smmtt” Extension

RV64 SPA

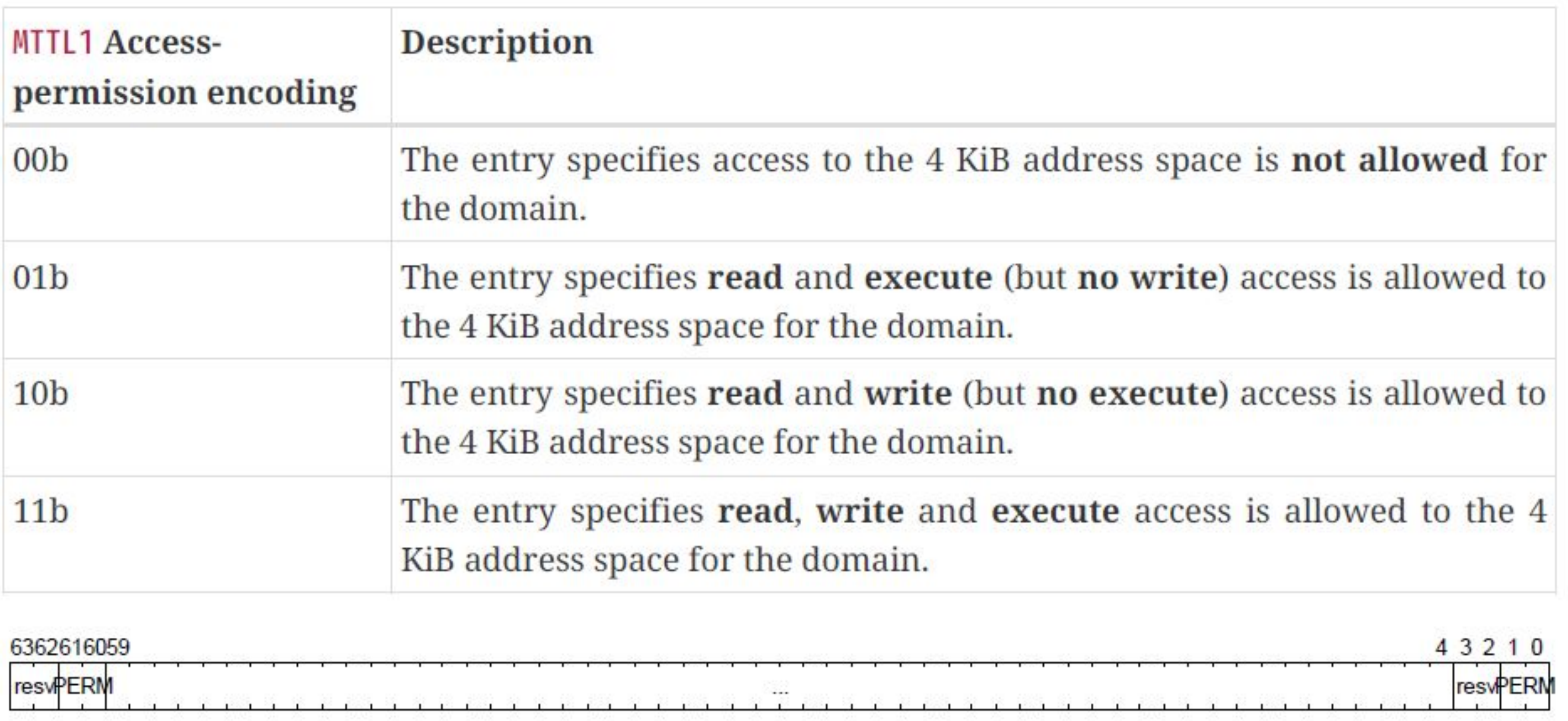




MTT L3
Table Size : 8 KiB
# of entries : 1024
Addr Space/Entry: 64 TiB

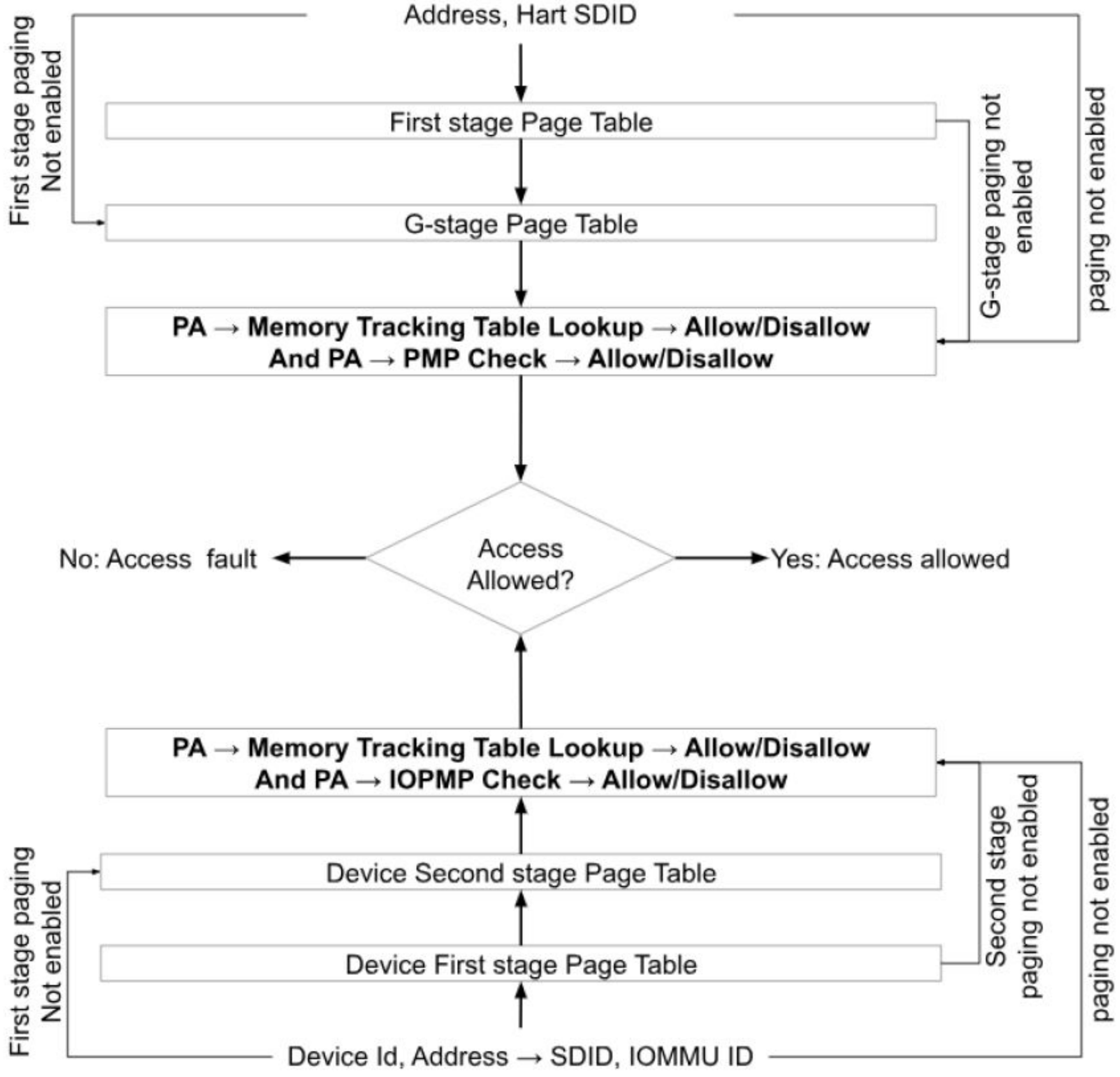
MTT L2
Table Size : 16 MiB
# of entries : 2 M
Addr Space/Entry: 32 MiB

MTT L1
Table Size : 4 KiB
# of entries : 8 K
Addr Space/Entry: 4 KiB



MTTL1 **Access-permission encoding**	**Description**
00b	The entry specifies access to the 4 KiB address space is **not allowed** for the domain.
01b	The entry specifies **read** and **execute** (but **no write**) access is allowed to the 4 KiB address space for the domain.
10b	The entry specifies **read** and **write** (but **no execute**) access is allowed to the 4 KiB address space for the domain.
11b	The entry specifies **read**, **write** and **execute** access is allowed to the 4 KiB address space for the domain.

Supervisor Domains - Approach for Hart and IO side access-control



Address, Hart SDID
First stage paging Not enabled
First stage Page Table
G-stage Page Table
G-stage paging not enabled
paging not enabled
PA → Memory Tracking Table Lookup → Allow/Disallow And PA → PMP Check → Allow/Disallow
No: Access fault
Access Allowed?
Yes: Access allowed
PA → Memory Tracking Table Lookup → Allow/Disallow And PA → IOPMP Check → Allow/Disallow
Device Second stage Page Table
Device First stage Page Table
First stage paging Not enabled
Second stage paging not enabled
paging not enabled
Device Id, Address → SDID, IOMMU ID

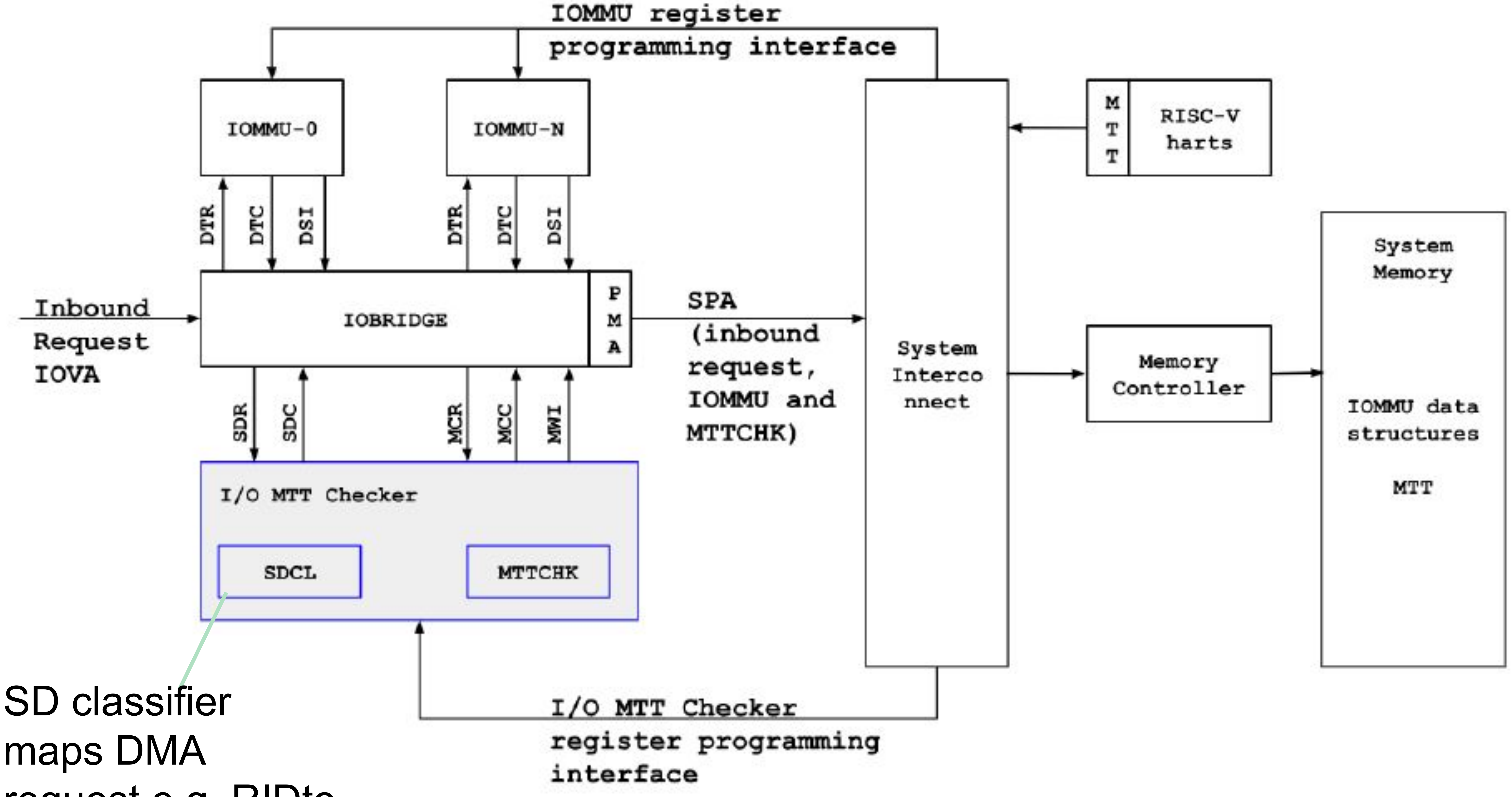
Supervisor Domains “IO-MTT” Approach

Security Objective - DMA from the devices and the IOMMU linked with a SD must adhere strictly to the access protections encoded in the MTT of the respective SD.

Supervisor domains may be granted control over DMA-capable devices by assigning IOMMU instances to the SD.

Using the MTT, RDSM enforces that the IOMMU memory-mapped programming regions are access-restricted to the SD the IOMMU is assigned to.

RDSM configures SDCL to map device requests to SD (and MTT)



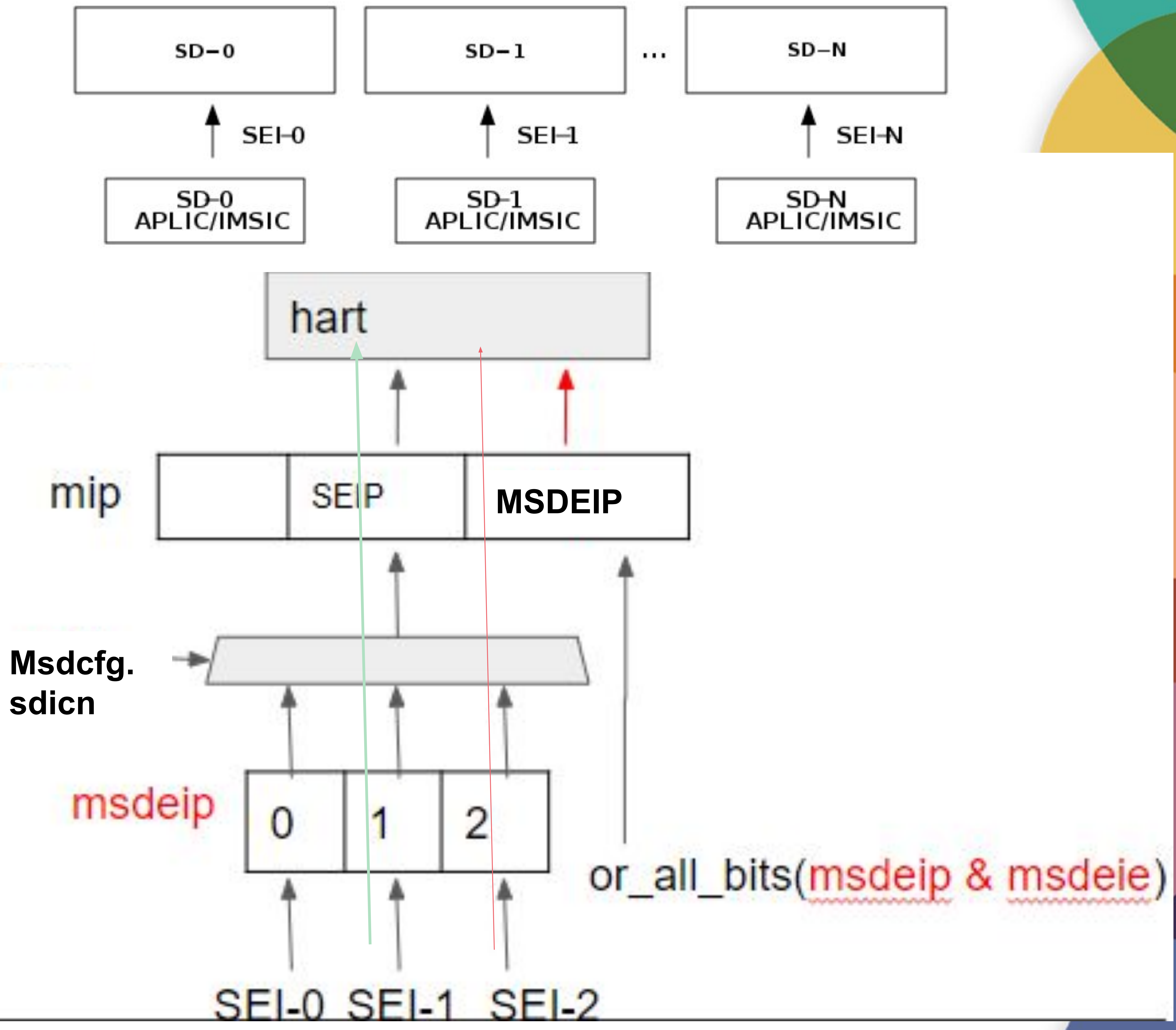
SD classifier maps DMA request e.g. RIDto MTT

Supervisor Domains Interrupt Isolation approach - “Smsdia” Extension

- **Security Objective** – RDSM must enforce integrity of interrupt delivery to the Supervisor Domain
- Smsdia **depends on RISC-V AIA**

- RDSM uses the **msdcfg.sdicn** to associate an interrupt controller with the SD
- RDSM uses MTT to enforce exclusive SD access to assigned interrupt controller
- RDSM uses CSRs **msdeip** and **msdeie** to get **MSDEIP** notification to M-mode when SD is not active.

Once an implemented interrupt controller is selected for SD, the H/S mode CSR interaction remains the same as defined in AIA.



Supervisor Domains - Summary

Also see :

- Smsqosid in the spec for QoS monitoring ID assignment to SD
- Smsdextdbg, Smsdexttrc - controls for external debug and trace allowance for SD

Longer discussion in these slides - https://static.sched.com/hosted_files/lsseu2024/2b/LSSEU24-RISC-V%20CoVE.pdf