



# ACPI for RISC-V: Enabling Server Class Platforms

**Sunil V L**  
**Ventana Micro Systems**



# Agenda



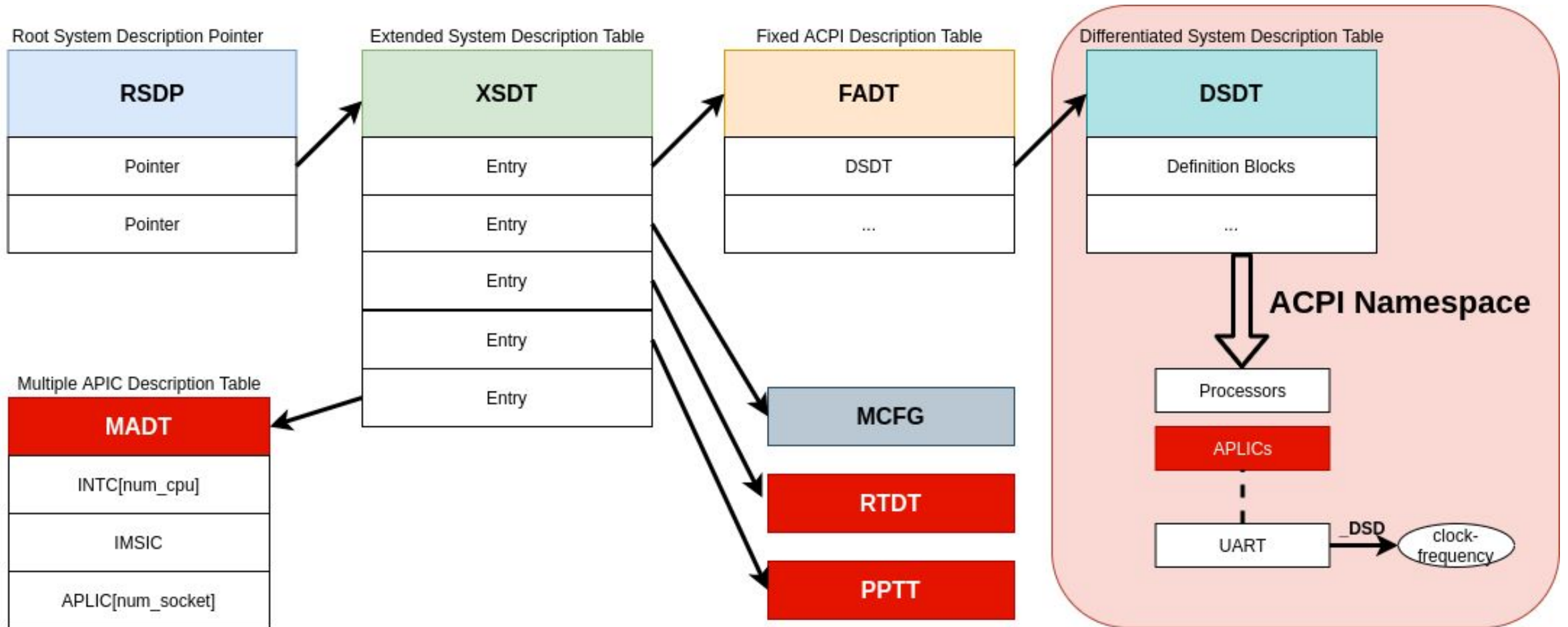
- Why ACPI for RISC-V Server Class Platforms?
- What it takes to enable ACPI for RISC-V?
- Proof of Concept
- Next Steps
- Q & A

# Why ACPI for RISC-V Server Class Platforms?



- Preferred choice of Enterprise Hardware OEMs, OS Vendors because of numerous reasons –  
<https://www.kernel.org/doc/html/latest/arm64/arm-acpi.html#why-acpi-on-arm>
- RISC-V OS-A Platform Specification mandates ACPI as the Hardware Discovery mechanism for server class platforms.
- Plan is to support ACPI for RV64 platforms with AIA Interrupt Controllers

# Enable ACPI for RISC-V : A bird's eye view.



# Proof of Concept



## QEMU

- **ACPI Tables** – RSDP, XSDT, FADT, DSDT, MADT, RTDT, MCFG
- **MADT:**
  - Per-hart INTC
  - IMSIC
  - Per-socket APLIC
- **DSDT:**
  - Processors
  - APLIC with \_MAT
  - Generic 16550a UART(PNP0501) with \_DSD method
  - Virtio

## EDK2

- **Integrated OpenSBI with AIA support**
- **ACPI enablement** (AcpiTableDxe, QemuFwCfg)
- **SMBIOS enablement**

## Linux

- **Basic ACPI enablement for RISC-V** (ACPIICA and ARCH specific ACPI)
- **ACPI based timer driver** (RTDT)
- **ACPI based INTC Driver**
- **ACPI based IMSIC driver**
- **ACPI based APLIC driver**
- **SMBIOS enablement**
- **Hart capabilities using SMBIOS table 44**

# How to communicate HART capabilities to OS?



- **ISA extensions** (rv64imafdc\*), **MMU type** (sv39, sv48) etc.
- SMBIOS specification has **Table 44** for these details. *But this makes SMBIOS mandatory along with ACPI to boot.*
- **Proposal**: Use “**Processor Properties Topology Table**” (PPTT) and add one more structure type to this table.

# Next Steps

- **Engineering Change Request (ECR)** – Review by UEFI ASWG
- **NUMA** – SRAT, SLIT
- **RAS** – APEI, Exception Delegation
- **Power/Performance Management (PCCT, CPPC)**
- **Watchdog**
- **ACPI Tools support**
  - Acpiview in edk2
  - Acpidump, etc.
- **Compliance Tests**



**THANK YOU!**

**Mailing List :**

**[tech-unixplatformspec@lists.riscv.org](mailto:tech-unixplatformspec@lists.riscv.org)**



# Supplementary Information

# ACPI Requirements for RISC-V Server Platforms



ACPI Table	Description	PoC Status
Root System Description Pointer (RSDP)		Done
Extended System Description Table (XSDT)		Done
Fixed ACPI Description Table (FADT)		Done
Differentiated System Description Table (DSDT)	ACPI Namespace – Processors, APLIC, peripherals etc	Done
<b>Multiple APIC Description Table (MADT)</b>	RISC-V INTC ( HART local interrupt controller) IMSIC, APLIC	Done
<b>RISC-V Timer Description Table (RTDT)</b>	timebase-frequency	Done
Memory-mapped Configuration space (MCFG)	PCIe	Done
Processor Properties Topology Table (PPTT)		TBD
Serial Port Console Redirection (SPCR)		TBD
Debug Port Table 2 (DBG2)		TBD

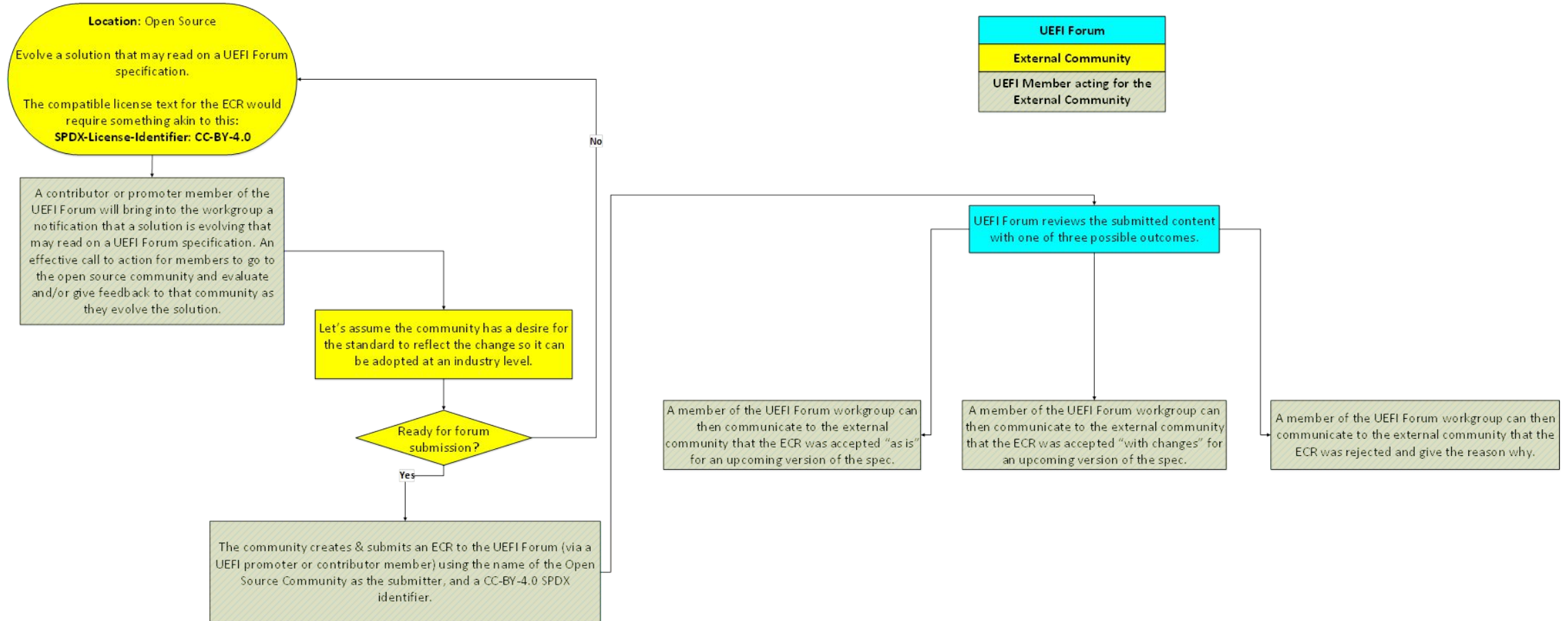
# ACPI Requirements for RISC-V Server Platforms (Contd.)



ACPI Table	Description	PoC Status
System Resource Affinity Table (SRAT)	NUMA	TBD
System Locality Information Table (SLIT)	NUMA	TBD
Boot Error Record Table (BERT)	RAS (APEI)	TBD
Error Injection Table (EINJ)	RAS (APEI)	TBD
Error Record Serialization Table (ERST)	RAS (APEI)	TBD
Hardware Error Source Table (HEST)	RAS (APEI)	TBD

<https://github.com/riscv-non-isa/riscv-acpi/blob/master/riscv-acpi-platform-req.adoc>

# ACPI ECR Process



# Proof of Concept – Changes

- **QEMU**

- [https://github.com/ventana-micro-systems/RISC-V-qemu/tree/acpi\\_v1\\_aia\\_v2](https://github.com/ventana-micro-systems/RISC-V-qemu/tree/acpi_v1_aia_v2)

- **Tianocore (EDK2)**

- [https://github.com/ventana-micro-systems/RISC-V-edk2-platforms/tree/acpi\\_v1\\_aia\\_v1](https://github.com/ventana-micro-systems/RISC-V-edk2-platforms/tree/acpi_v1_aia_v1)
- [https://github.com/ventana-micro-systems/RISC-V-edk2/tree/acpi\\_v1\\_aia\\_v1](https://github.com/ventana-micro-systems/RISC-V-edk2/tree/acpi_v1_aia_v1)

- **Linux**

- [https://github.com/ventana-micro-systems/RISC-V-Linux/tree/acpi\\_v1\\_aia\\_v1](https://github.com/ventana-micro-systems/RISC-V-Linux/tree/acpi_v1_aia_v1)

# ACPI Enabled RISC-V Linux booting on Qemu



```
Shell> fs0:\Image root=/dev/vda2 rootwait console=ttyS0 earlycon initrd=\initramfs.cp
EFI stub: Booting Linux Kernel...
EFI stub: Using DTB from configuration table
EFI stub: Exiting boot services and installing virtual address map...
PROGRESS CODE: V03101019 I0
[ 0.000000] Linux version 5.14.0-rc4-00046-gfd541c2c4495-dirty (sunil@sunil-ThinkPad-T490) (riscv64-linux-g
nu-gcc (Ubuntu 10.2.0-5ubuntu1-20.04) 10.2.0, GNU ld (GNU Binutils for Ubuntu) 2.34) #102 SMP Mon Aug 30 19:46
:02 IST 2021
[ 0.000000] OF: fdt: Ignoring memory range 0x80000000 - 0x81200000
[ 0.000000] Machine model: riscv-virtio,qemu
[ 0.000000] earlycon: ns16550a0 at MMIO 0x0000000010000000 (options '')
[ 0.000000] printk: bootconsole [ns16550a0] enabled
[ 0.000000] efi: EFI v2.70 by EDK II
[ 0.000000] efi: SMBIOS 3.0=0xbfb46000 ACPI 2.0=0xbfb75018 RNG=0xbfb47f18 MEMRESERVE=0xbe9def18
[ 0.000000] efi: seeding entropy pool
[ 0.000000] OF: fdt: Ignoring memory block 0x81000000 - 0x81020000
[ 0.000000] OF: fdt: Ignoring memory range 0x81020000 - 0x81200000
[ 0.000000] ACPI: Early table checksum verification disabled
[ 0.000000] ACPI: RSDP 0x00000000BFB75018 000024 (v02 BOCHS )
[ 0.000000] ACPI: XSDT 0x00000000BFB75F18 00004C (v01 BOCHS BXP 00000001 01000013)
[ 0.000000] ACPI: FACP 0x00000000BFB75B18 00010C (v05 BOCHS BXP 00000001 BXP 00000001)
[ 0.000000] ACPI: DSDT 0x00000000BFB74018 000F55 (v02 BOCHS BXP 00000001 BXP 00000001)
[ 0.000000] ACPI: APIC 0x00000000BFB75C98 00023C (v03 BOCHS BXP 00000001 BXP 00000001)
[ 0.000000] ACPI: RTD 0x00000000BFB75098 0000C2 (v02 BOCHS BXP 00000001 BXP 00000001)
[ 0.000000] ACPI: MCFG 0x00000000BFB75A98 00003C (v01 BOCHS BXP 00000001 BXP 00000001)
[ 0.000000] ACPI: SPCR 0x00000000BFB75818 000050 (v02 BOCHS BXP 00000001 BXP 00000001)
[ 0.000000] Zone ranges:
[ 0.000000] DMA32 [mem 0x0000000081200000-0x00000000bfffffff]
[ 0.000000] Normal empty
[ 0.000000] Movable zone start for each node
[ 0.000000] Early memory node ranges
[ 0.000000] node 0: [mem 0x0000000081200000-0x00000000be847fff]
[ 0.000000] node 0: [mem 0x00000000be848000-0x00000000be8b0fff]
[ 0.000000] node 0: [mem 0x00000000be8b1000-0x00000000be9d6fff]
[ 0.000000] node 0: [mem 0x00000000be9d7000-0x00000000be9d9fff]
[ 0.000000] node 0: [mem 0x00000000be9da000-0x00000000be9defff]
[ 0.000000] node 0: [mem 0x00000000be9df000-0x00000000be9e0fff]
[ 0.000000] node 0: [mem 0x00000000be9e1000-0x00000000bea94fff]
[ 0.000000] node 0: [mem 0x00000000bea95000-0x00000000bea95fff]
[ 0.000000] node 0: [mem 0x00000000bea96000-0x00000000bfb15fff]
[ 0.000000] node 0: [mem 0x00000000bfb16000-0x00000000bfb69fff]
[ 0.000000] node 0: [mem 0x00000000bfb6a000-0x00000000bfb75fff]
[ 0.000000] node 0: [mem 0x00000000bfb76000-0x00000000bfb79fff]
[ 0.000000] node 0: [mem 0x00000000bfb7a000-0x00000000bfffffff]
[ 0.000000] Initmem setup node 0 [mem 0x0000000081200000-0x00000000bfffffff]
[ 0.000000] SBI specification v0.3 detected
[ 0.000000] SBI implementation ID=0x1 Version=0x9
[ 0.000000] SBI TIME extension detected
[ 0.000000] SBI IPI extension detected
[ 0.000000] SBI RFENCE extension detected
[ 0.000000] SBI SRST extension detected
[ 0.000000] SBI v0.2 HSM extension detected
[ 0.000000] SMBIOS 3.0.0 present.
[ 0.000000] DMI: , BIOS
[ 0.000000] riscv: ISA extensions acdfhimsu
[ 0.000000] riscv: ELF capabilities acdfim
```

```
[ 0.000000] percpu: Embedded 19 pages/cpu s39464 r8192 d30168 u77824
[ 0.000000] Built 1 zonelists, mobility grouping on. Total pages: 254015
[ 0.000000] Kernel command line: fs0:\Image root=/dev/vda2 rootwait console=ttyS0 earlycon initrd=\initramf
s.cp
[ 0.000000] Unknown command line parameters: fs0:\Image
[ 0.000000] Dentry cache hash table entries: 131072 (order: 8, 1048576 bytes, linear)
[ 0.000000] Inode-cache hash table entries: 65536 (order: 7, 524288 bytes, linear)
[ 0.000000] Sorting __ex table...
[ 0.000000] mem auto-init: stack:off, heap alloc:off, heap free:off
[ 0.000000] Virtual kernel memory layout:
[ 0.000000] fixmap : 0xfffffcef000000 - 0xfffffcef000000 (2048 kB)
[ 0.000000] pci io : 0xfffffcef000000 - 0xfffffcef000000 ( 16 MB)
[ 0.000000] vmemmap : 0xfffffcef00000000 - 0xfffffcef00000000 (4095 MB)
[ 0.000000] vmemmap : 0xfffffcef00000000 - 0xfffffcef00000000 (4095 MB)
[ 0.000000] vmalloc : 0xfffff00000000000 - 0xfffff00000000000 (65535 MB)
[ 0.000000] lowmem : 0xfffffe0000000000 - 0xfffffe03ee000000 (1006 MB)
[ 0.000000] kernel : 0xfffffff800000000 - 0xfffffff800000000 (2047 MB)
[ 0.000000] Memory: 990560K/1030144K available (7671K kernel code, 4894K rdata, 4096K rodata, 2187K init,
3909K bss, 39584K reserved, 0K cma-reserved)
[ 0.000000] SLUB: Hwalign=64, Order=0-3, MinObjects=0, CPUs=8, Nodes=1
[ 0.000000] rcu: Hierarchical RCU implementation.
[ 0.000000] rcu: RCU debug extended QS entry/exit.
[ 0.000000] Tracing variant of Tasks RCU enabled.
[ 0.000000] rcu: RCU calculated value of scheduler-enlistment delay is 25 jiffies.
[ 0.000000] NR_IRQS: 64, nr_irqs: 64, preallocated irq: 0
[ 0.000000] riscv-intc: riscv_intc acpi init: 64 local interrupts mapped
[ 0.000000] riscv-imsic: imsic acpi init: mapped 255 interrupts using 8 handlers
[ 0.000000] random: get_random_bytes called from start_kernel+0x4fa/0x730 with crng_init=0
[ 0.000000] clocksource: riscv_clocksource: mask: 0xffffffffffffff max_cycles: 0x24e6a1710, max_idle_ns:
440795202120 ns
[ 0.000304] sched clock: 64 bits at 10MHz, resolution 100ns, wraps every 4398046511100ns
[ 0.017446] Console: colour dummy device 80x25
[ 0.022229] ACPI: Core revision 20210604
[ 0.034741] Calibrating delay loop (skipped), value calculated using timer frequency.. 20.00 BogoMIPS (lpj=
40000)
[ 0.036002] pid_max: default: 32768 minimum: 301
[ 0.040564] Mount-cache hash table entries: 2048 (order: 2, 16384 bytes, linear)
[ 0.041520] Mountpoint-cache hash table entries: 2048 (order: 2, 16384 bytes, linear)
[ 0.134753] ASID allocator disabled
[ 0.138377] rcu: Hierarchical SRCU implementation.
[ 0.145077] Remapping and enabling EFI services.
[ 0.157591] smp: Bringing up secondary CPUs ...
[ 0.221514] smp: Brought up 1 node, 8 CPUs
[ 0.243537] devtmpfs: initialized
[ 0.257631] clocksource: jiffies: mask: 0xffffffff max_cycles: 0xffffffff, max_idle_ns: 764504178510000 ns
[ 0.259345] futex hash table entries: 2048 (order: 5, 131072 bytes, linear)
[ 0.270129] NET: Registered PF_NETLINK/PF_ROUTE protocol family
[ 0.281675] ACPI: bus type PCI registered
[ 0.318698] HugeTLB registered 2.00 MiB page size, pre-allocated 0 pages
[ 0.322998] wait_for_initramfs() called before rootfs_initcalls
[ 0.357475] ACPI: Added _OSI(Module Device)
[ 0.358011] ACPI: Added _OSI(Processor Device)
[ 0.358203] ACPI: Added _OSI(3.0 _SCP Extensions)
[ 0.358392] ACPI: Added _OSI(Processor Aggregator Device)
[ 0.358774] ACPI: Added _OSI(Linux-Dell-Video)
[ 0.358966] ACPI: Added _OSI(Linux-Lenovo-NV-HDMI-Audio)
```

# ACPI Enabled RISC-V Linux booting on Qemu



```
0.359243] ACPI: Added_OSI(Linux-HPI-Hybrid-Graphics)
[0.379605] ACPI: 1 ACPI AML tables successfully acquired and loaded
[0.391307] ACPI: Interpreter enabled
[0.391495] ACPI: Using RISC-V AIA model for interrupt routing
[0.392472] ACPI: MCFG table detected, 1 entries
[0.427340] riscv-aplic-acpi APLIC001:00: probing via ACPI
[0.442969] riscv-aplic-acpi APLIC001:00: 53 interrupts forwarded to MSI base PPN 0x0000000028000000
[0.449204] ACPI: PCI Root Bridge [PCI0] (domain 0000 [bus 00-ff])
[0.451473] acpi PNP0A08:00: _OSC: OS supports [ExtendedConfig ASPM ClockPM Segments MSI HPX-Type3]
[0.459694] acpi PNP0A08:00: _OSC: platform does not support [LTR]
[0.461305] acpi PNP0A08:00: _OSC: OS now controls [PME PCIECapability]
[0.463963] acpi PNP0A08:00: ECAM area [mem 0x30000000-0x3fffffff] reserved by PNP0C02:00
[0.466348] acpi PNP0A08:00: ECAM at [mem 0x30000000-0x3fffffff] for [bus 00-ff]
[0.467717] ACPI: Remapped I/O 0x0000000030000000 to [io 0x0000-0xffff window]
[0.471303] PCI host bridge to bus 0000:00
[0.471701] pci_bus 0000:00: root bus resource [mem 0x40000000-0x7fffffff window]
[0.472101] pci_bus 0000:00: root bus resource [io 0x0000-0xffff window]
[0.472394] pci_bus 0000:00: root bus resource [mem 0x40000000-0x7fffffff window]
[0.472947] pci_bus 0000:00: root bus resource [bus 00-ff]
[0.475757] pci 0000:00:00.0: [1b36:0008] type 00 class 0x060000
[0.482512] pci_bus 0000:00: resource 4 [mem 0x40000000-0x7fffffff window]
[0.482885] pci_bus 0000:00: resource 5 [io 0x0000-0xffff window]
[0.483235] pci_bus 0000:00: resource 6 [mem 0x40000000-0x7fffffff window]
[0.484681] ACPI: PCI: Interrupt link GSI0 configured for IRQ 32
[0.485141] ACPI: PCI: Interrupt link GSI1 configured for IRQ 33
[0.485492] ACPI: PCI: Interrupt link GSI2 configured for IRQ 34
[0.486226] ACPI: PCI: Interrupt link GSI3 configured for IRQ 35
[0.492958] vgaarb: loaded
[0.495301] SCSI subsystem initialized
[0.498018] ACPI: bus type USB registered
[0.499354] usbcore: registered new interface driver usbfs
[0.500203] usbcore: registered new interface driver hub
[0.500768] usbcore: registered new device driver usb
[0.503546] Registered efivar operations
[0.520089] clocksource: Switched to clocksource riscv_clocksource
[0.523429] pnp: PnP ACPI init
[0.528491] system 00:00: [mem 0x30000000-0x3fffffff window] could not be reserved
[0.530000] pnp: PnP ACPI: found 2 devices
[0.561793] NET: Registered PF_INET protocol family
[0.564785] IP ident's hash table entries: 16384 (order: 5, 131072 bytes, linear)
[0.571802] tcp_listen_portaddr_hash hash table entries: 512 (order: 2, 20480 bytes, linear)
[0.572369] TCP established hash table entries: 8192 (order: 4, 65536 bytes, linear)
[0.573048] TCP bind hash table entries: 8192 (order: 6, 262144 bytes, linear)
[0.573977] TCP: Hash tables configured (established 8192 bind 8192)
[0.576182] UDP hash table entries: 512 (order: 3, 49152 bytes, linear)
[0.576950] UDP-Lite hash table entries: 512 (order: 3, 49152 bytes, linear)
[0.579885] NET: Registered PF_UNIX/PF_LOCAL protocol family
[0.585389] RPC: Registered named UNIX socket transport module.
[0.585766] RPC: Registered udp transport module.
[0.585932] RPC: Registered tcp transport module.
[0.586133] RPC: Registered tcp NFSv4.1 backchannel transport module.
[0.587619] kvm [1]: hypervisor extension available
[0.588012] kvm [1]: using Sv48x4 G-stage page table format
[0.588304] kvm [1]: VMID 14 bits available
[0.594783] workingset: timestamp bits=62 max_order=18 bucket_order=0
```

```
0.617585] NFS: Registering the id_resolver key type
[0.619047] Key type id_resolver registered
[0.619443] Key type id_legacy registered
[0.620222] nfs4filelayout_init: NFSv4 File Layout Driver Registering...
[0.620826] nfs4flexfilelayout_init: NFSv4 Flexfile Layout Driver Registering...
[0.622818] 9p: Installing v9fs 9p2000 file system support
[0.625817] NET: Registered PF_ALG protocol family
[0.626603] Block layer SCSI generic (bsg) driver version 0.4 loaded (major 251)
[0.627191] io scheduler mq-deadline registered
[0.627780] io scheduler kyber registered
[0.767418] Serial: 8250/16550 driver, 4 ports, IRQ sharing disabled
[0.792016] 00:01: ttyS0 at MMIO 0x10000000 (irq = 75, base_baud = 115200) is a 16550A
[0.794247] printk: console [ttyS0] enabled
[0.794247] printk: console [ttyS0] enabled
[0.795126] printk: bootconsole [ns16550a0] disabled
[0.795126] printk: bootconsole [ns16550a0] disabled
[0.806646] [drm] radeon kernel modesetting enabled.
[0.807233] random: fast init done
[0.808904] random: crng init done
[0.859128] loop: module loaded
[0.876415] virtio_blk virtio1: [vda] 3987820 512-byte logical blocks (2.04 GB/1.90 GiB)
[0.900198] vda: vda1 vda2 vda3
[0.911498] libphy: Fixed MDIO Bus: probed
[0.921868] e1000e: Intel(R) PRO/1000 Network Driver
[0.922092] e1000e: Copyright(c) 1999 - 2015 Intel Corporation.
[0.922984] ehci_hcd: USB 2.0 'Enhanced' Host Controller (EHCI) Driver
[0.923355] ehci-pci: EHCI PCI platform driver
[0.924599] ehci-platform: EHCI generic platform driver
[0.924985] ohci_hcd: USB 1.1 'Open' Host Controller (OHCI) Driver
[0.925388] ohci-pci: OHCI PCI platform driver
[0.925917] ohci-platform: OHCI generic platform driver
[0.928180] usbcore: registered new interface driver uas
[0.928780] usbcore: registered new interface driver usb-storage
[0.930169] mousedev: PS/2 mouse device common for all mice
[0.932086] sdhci: Secure Digital Host Controller Interface driver
[0.932395] sdhci: Copyright(c) Pierre Ossman
[0.932754] sdhci-pltfm: SDHCI platform and OF driver helper
[0.933979] usbcore: registered new interface driver usbhid
[0.934248] usbhid: USB HID core driver
[0.936915] NET: Registered PF_INET6 protocol family
[0.946774] Segment Routing with IPv6
[0.947824] sit: IPv6, IPv4 and MPLS over IPv4 tunneling driver
[0.952431] NET: Registered PF_PACKET protocol family
[0.954550] 9pnet: Installing 9P2000 support
[0.955979] Key type dns_resolver registered
[0.959566] debug_vm_pgtable: [debug_vm_pgtable ]: Validating architecture page table helpers
[0.994543] EXT4-fs (vda2): INFO: recovery required on readonly filesystem
[0.994873] EXT4-fs (vda2): write access will be enabled during recovery
[1.025424] EXT4-fs (vda2): recovery complete
[1.030235] EXT4-fs (vda2): mounted filesystem with ordered data mode. Opts: (null). Quota mode: disabled.
[1.031010] VFS: Mounted root (ext4 filesystem) readonly on device 254:2.
[1.034004] devtmpfs: mounted
[1.067445] Freeing unused kernel image (initmem) memory: 2184K
[1.068772] Run /sbin/init as init process
```

# References

- <https://github.com/riscv-non-isa/riscv-acpi/blob/master/riscv-acpi-platform-req.adoc>
- <https://github.com/riscv-non-isa/riscv-acpi/wiki/ACPI-ASWG-ECR-Process>
- <https://lists.riscv.org/g/tech-unixplatformspec>
- <https://uefi.org/acpi>
- <https://github.com/riscv/riscv-smbios/blob/master/riscv-smbios.adoc>



# Acknowledgements

- Anup Patel – Western Digital
- Atish Patra – Western Digital
- Abner Chang – HPE
- Rahul Pathak – Ventana Micro Systems
- Kumar Sankaran – Ventana Micro Systems
- Mayuresh Chitale – Ventana Micro Systems
- RISC-V Platform HSC community