Transition from SR-IOV to SIOV (PCISIG) for IO Virtualization

Session Name : **Virtualization and Orchestration**
Session Date and Time : **Aug 07, 2025 Santa Clara Convention Center**

Gordon Waidhofer, Senior Technical Staff Engineer, Architecture
Microchip Technology Inc.

Douglas Arens, Technical Staff, Applications
Microchip Technology Inc.

Agenda

- **Virtualization Architecture and Benefits**
- **What is SR-IOV?**
  - Current standard to achieve multi-tenancy
  - Multiple SR-IOV instances on Server attach to NVMe®
  - Objects (Namespaces) on SSD drive
- **What is SIOV?**
  - How is it the same?
  - How is it different?
    - Footprint
    - Performance
    - Software ecosystem
  - Are there any disadvantages?
    - Market preferences / adoption
- **What comes next?**

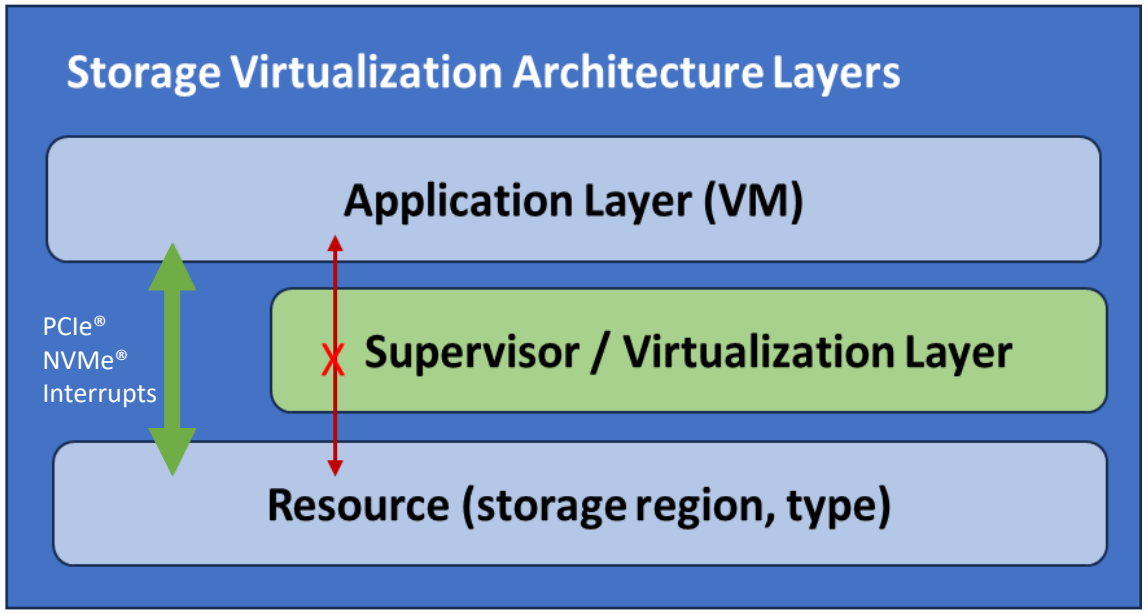
Virtualization Architecture and Benefits

Architecture Layers

- More direct-access from an Application (VM) to a resource
- Minimal supervisor-layer to initially configure access from Application to the Virtual resource.
- Then supervisor-layer gets out of the way.
- Utilize the direct-access as a superhighway for unencumbered data flow

Benefits of Virtualization

- ✓ Reduced capital and operating costs
- ✓ Minimized or eliminated downtime.
- ✓ Increased IT productivity, efficiency, agility and responsiveness
- ✓ Faster provisioning of applications and resources



Challenges of multi-tenant Virtualized Systems

- **Server Memory**
  - More VMs (tenants) require more RAM
- **Server Storage**
  - Large boot drive with separate VMs
  - More VMs (tenants) often require more Storage
- **Server Performance**
  - Each tenant takes its share of CPU and other resources
- **Share resources by as many VMs as possible**
  - Overcome limitations
- **Maintain application independence & isolation**
- **Maintain security while sharing system**
- **Migrate applications and namespace data**

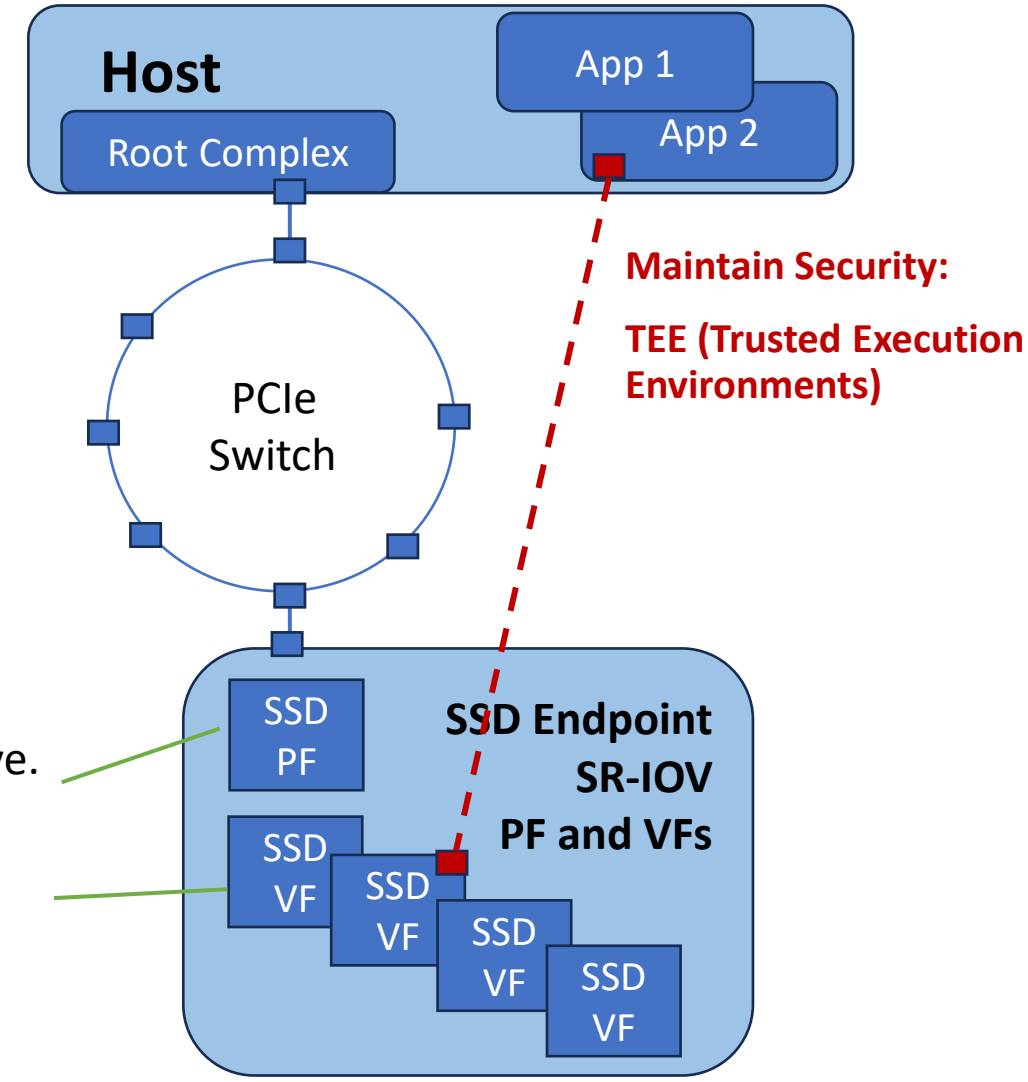
PCIe® SR-IOV (Single-Root I/O Virtualization)

USE CASE: Multifunction device at PCIe layer, better cost/features

SR-IOV allows a PCIe endpoint under a single root complex to appear as multiple PCIe functions (e.g. NVMe® Drives) to the hypervisor or the guest operating system.

Virtualization Benefits:

- ✓ It makes it possible to share an endpoint's resources (NVMe storage) among 255 tenants, which reduces the need for separate hardware per tenant and the resultant costs.



Physical Function – NVMe drive.

Virtual Function – Mapped by tenant. Also an NVMe drive.

Evolution of PCIe® Virtualization Technologies

SR-IOV – Single Root IO Virtualization

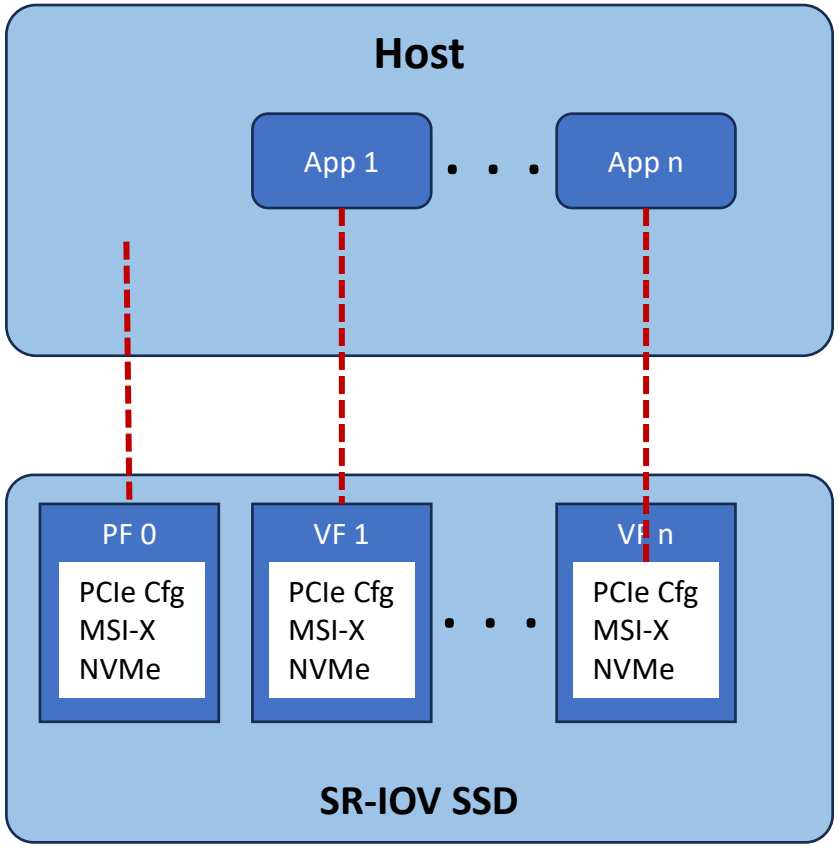
- Part of the PCIe spec for over 10 years
- Rich ecosystem, took quite a while
- Addressed by NVMe® spec
- The drive can be used without a VMM (Hypervisor) because the drive handles all device and configuration registers
- Awkward to scale above 255 virtual functions (VFs). ARI (Alternate Route Interpretation), which would raise the limit, isn't widely supported.
- More drive complexity and engineering
- VF configuration is standardized

SIOV – Scalable IO Virtualization

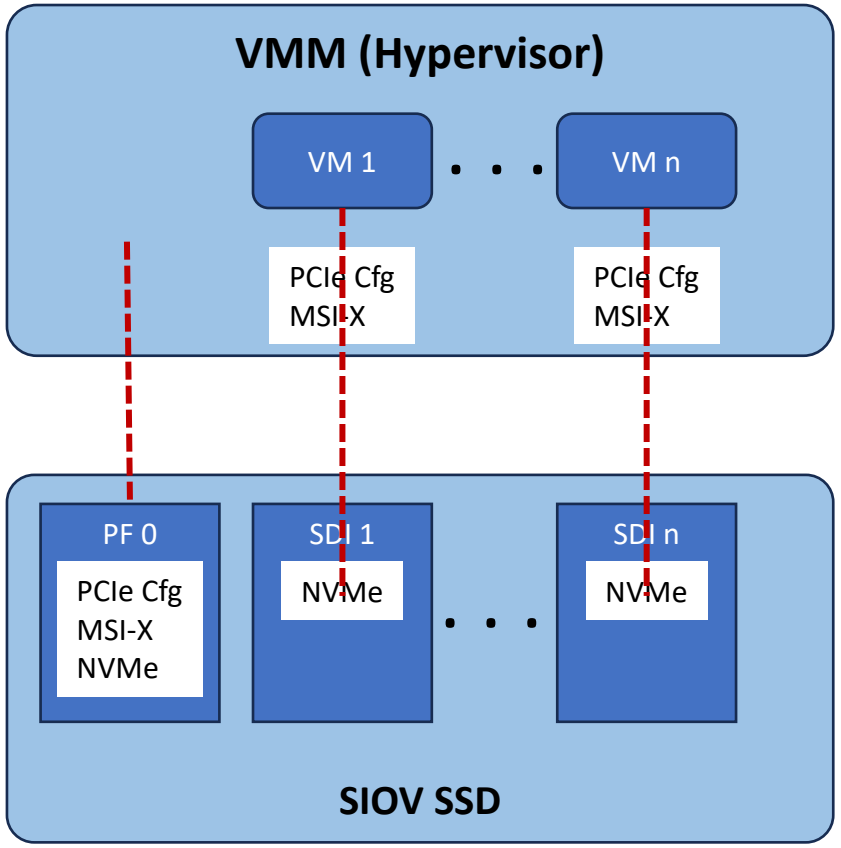
- Will be in next PCIe spec, proposed by Intel in 2017
- Early days for an ecosystem
- TODO for NVMe spec
- The drive works in partnership with a VMM (Hypervisor) because the drive handles only the registers used for I/O
- Discussions are happening to support more than 255 SDI (Scalable Device Interface) with methods simpler than ARI
- Less drive complexity and engineering
- SDI configuration will be implementation-specific at first, hopefully NVMe and other standards will follow-up with SIOV for consistent configuration

Evolution of PCIe® and NVMe® Virtualization Technologies

SR-IOV – Single Root IO Virtualization



SIOV – Scalable IO Virtualization



Drive and hypervisor partnership

All SDIs reside in PF0's BARs.

Each VF and SDI have a unique RID (Requester ID) for DMA routing and security.

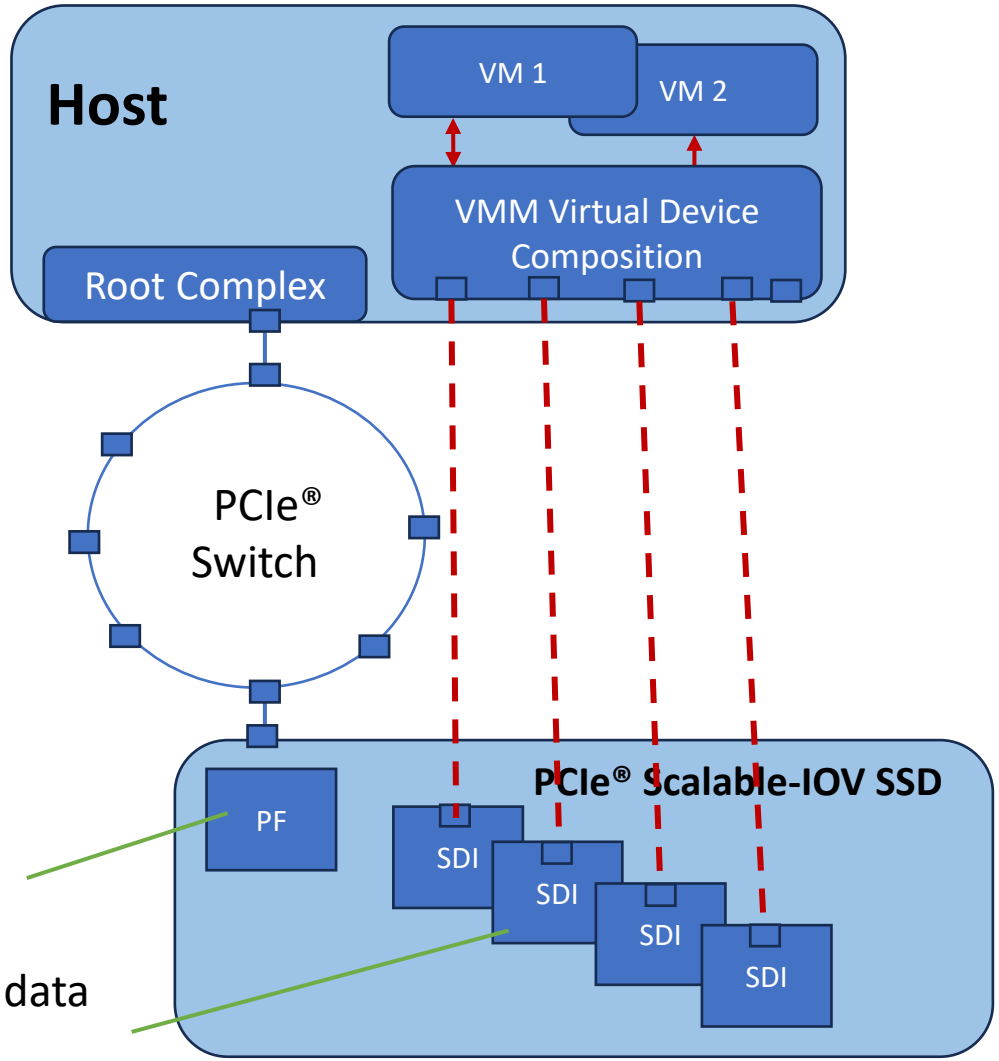
PCIe® SIOV (Scalable I/O Virtualization)

USE CASE: Highly Scalable Multi-tenant device

PCIe Scalable IOV improves upon SR-IOV to allow a single SSD under a single root port to virtualize ~1000 separate storage devices for each host in hyper-scale datacenters

PCIe® Scalable I/O Benefits:

- ✓ Storage is more scalable, hardware-assisted IO paired with software flexibility. Scalable 20-bit ID (PASID, Process Address Space) between VM and SDI.
- ✓ DMA and interrupt remapping
- ✓ VMM Directed IO to NVME® queue level at Device
- ✓ Separation of fast path I/O from slow path (configuration, reset)



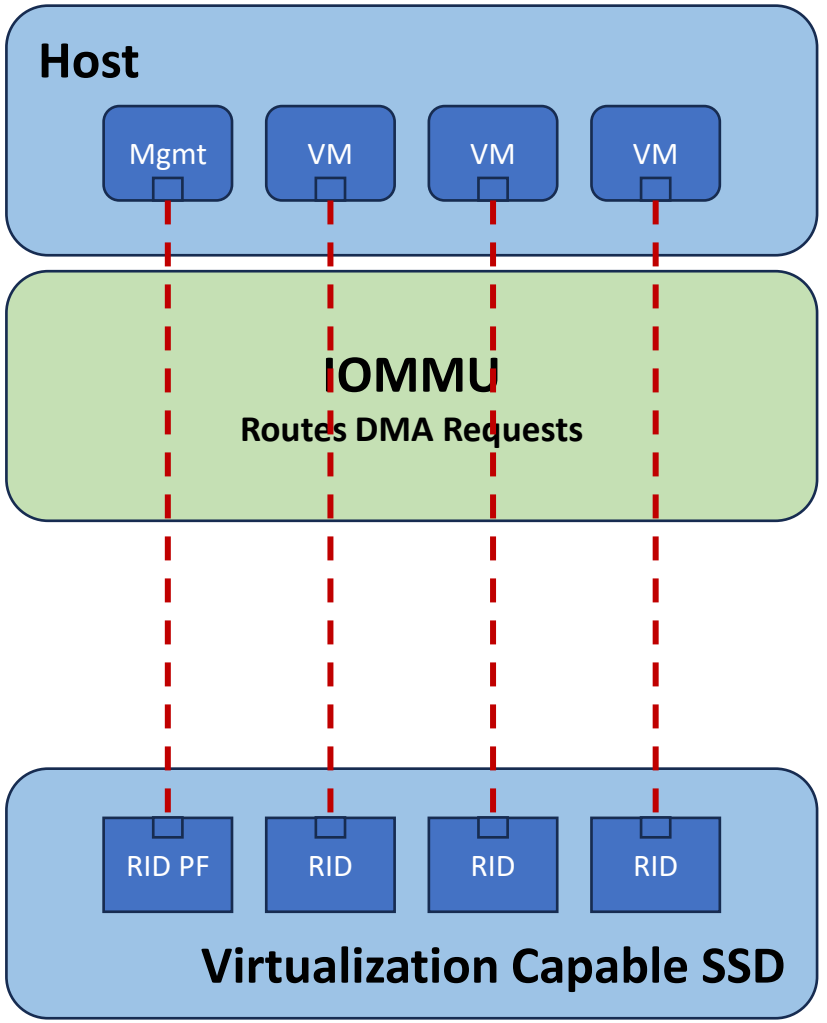
Physical Function- Real NVMe device, configuration, access

Scalable Device Interface. Exchange of data (direct-addressing to queue level)

What's Next? – Common Challenges

- Limitations
  - PCIe® end points are usually limited to 256 RIDs, not enough for future virtualization
  - RIDs (Requester IDs) route DMAs and maintain security.
  - SR-IOV VFs and SIOV SDIs have this limit

- Moving beyond those limitations:
  - Devices (NVMe® drives) with internal PCIe switches appear as multiple endpoints, more than 255 RIDs total
  - Configure PCIe switches to route multiple PCIe buses (devices) to a single endpoint (ARI, SIOV)
  - SIOV opens the door for SDIs to share RIDs, so more than 255 SDI in one endpoint (PASID)
  - Need coordinated evolution of the ecosystems



What's Next? – Another Wave of Improvements

- **Hardware Improvements**
  - More tenants per platform (density, compact)
  - Larger SSD drives to utilize fewer PCIe® slots
  - More HW automation (performance)
  - Bigger IOMMU, PASID-ready
- **Standards Improvements**
  - NVMe® address of SIOV like was done for SR-IOV
  - More evolution of PCIe® SIOV
- **Software Improvements**
  - Focus on customer application needs
  - Hypervisors enhanced for SIOV approach

Thank you !

Visit Microchip Booth # 115