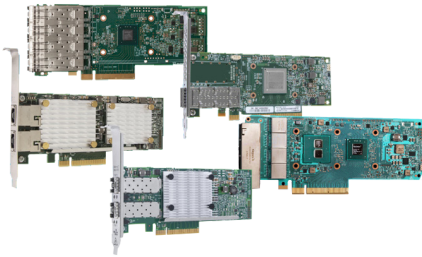


Concurrent NIC Partitioning and SR-IOV

Do More with Less in Server Virtualization!



- Reduces Capital and Operational Expenditures
- Eliminates Dependency on a Captive Switch
- Simplifies Deployment and Management
- Improves Ability to Scale the Business
- Delivers a Flexible I/O Virtualization Solution

Executive Summary

Virtualization of networks, storage, servers, and applications continues to grow in the data center. While virtualization provides many benefits, virtual deployment workloads have increased the demands placed on the network I/O infrastructure, which can significantly impact system and application performance.

Architectural improvements are occurring on the interface between the CPUs and the network interface hardware. Virtualization of I/O using Single Root I/O Virtualization (SR-IOV) for individual servers and switch independent NIC Partitioning (NPAR) are technology standards that can improve the scalability, flexibility, throughput, and latency performance of networking operations.

With Marvell FastLinQ® 10/20/25/40/50/100GbE 3400/8400 (578xx)/41000/45000 Series Intelligent Ethernet Adapters and Converged Network Adapters, NPAR and SR-IOV can be used independently or simultaneously to divide a physical adapter port into multiple virtual ports. This Marvell unique capability allows customers to concurrently have access to storage offloading, have bandwidth Quality of Service (QoS), use NIC teaming, and use SR-IOV VF's enhanced I/O performance.

The following operating systems support both SR-IOV and NPAR modes:

- Windows Server 2012 Hyper-V
- Windows Server 2012 R2 Hyper-V
- Windows Server 2016 Hyper-V
- Linux RHEL 6.4+ with KVM hypervisor
- Linux SLES 11.3+ with XEN hypervisor
- Citrix XenServer 6.2+ VMware ESXi 5.5+

NIC Partitioning – Hardware-Based QoS

The Marvell FastLinQ 10/20/25/40/50/100GbE 3400/8400 (578xx)/41000 /45000 Series Intelligent Ethernet Adapters and Converged Network Adapters NPAR technology is driving next-generation server I/O virtualization by allowing the seamless migration from 1GbE to 10GbE/25GbE/40GbE/50GbE/100GbE while preserving the bandwidth segregation that 1GbE brought to virtual server environments. NPAR or switch independent Network Partitioning divides a single physical Ethernet port into multiple¹ PCI physical functions or partitions using a flexible bandwidth capacity allocation.

1. Up to 8 physical functions on the 3400/8400 (578xx) series adapters and up to 16 physical functions on the 41000/45000 series adapters

This approach enables administrators to apply QoS to individual virtual machine or hypervisor service workloads, improving I/O performance while maintaining a low total cost of ownership (TCO). Cavium’s implementation maps multiple PCI functions to each physical port of a Cavium FastLinQ 3400/8400 (578xx)/41000/45000 Series Adapter.

NPAR presents the multiple PCI functions per device using standard PCI configuration space. Each function or partition is assigned unique MAC addresses. Each Cavium FastLinQ Converged Network Adapter port can support simultaneous networking (with Universal RDMA iWARP/RoCEv2/RoCEv2 on the 41000/45000) or storage protocols, enabling flexible bandwidth provisioning to applications. Unlike competitive solutions, these adapters can be assigned to simultaneously run both networking and storage protocols (on any port) with an implementation that is agnostic to the external Ethernet switch.

NPAR provides a non-captive (switch agnostic) solution to customers, thereby eliminating vendor lock-in.

PCI-SIG SR-IOV Standard

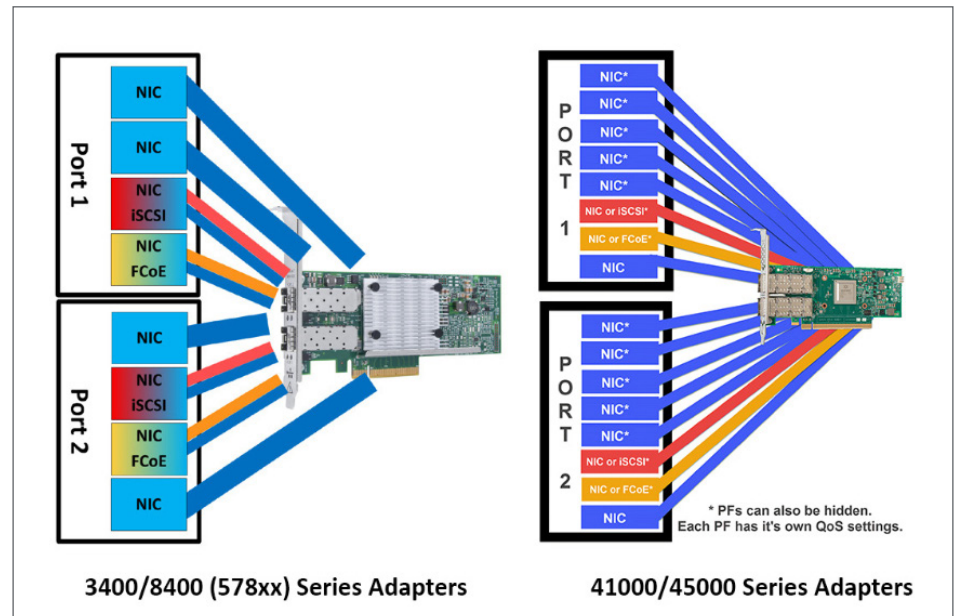


Figure 1. NPAR

PCI special interest group (PCI-SIG®) has completed a suite of I/O virtualization specifications that work in conjunction with system virtualization technologies to allow multiple OSs running simultaneously within a single computer to natively share PCI Express® devices. I/O virtualization technology aims to resolve the performance, provisioning, and management problems of server I/O in highly virtualized and clustered data center environments.

The SR-IOV specification designates how a hardware device can expose multiple “light-weight” Virtual Functions (VFs) for use by VMs. VFs are focused primarily on data movement and are associated with a Physical Function (PF). The PF is equivalent to the regular PCIe® physical function device’s fully featured PCI functions. The PF is responsible for arbitration relating to policy decisions (such as link speed or MAC addresses in use by VMs in the case of networking) and for I/O from the parent partition itself.

For example, the Marvell FastLinQ 3400/8400 Series Adapters support up to 64 VFs per port in single PF mode and NPAR multiple PF mode. In SF mode, up to 64 VFs can be assigned per port along with both storage (iSCSI-Offload and FCoE-Offload) personalities enabled. In NPAR mode, administrators can assign the port's VFs [via the various Marvell management utilities or the pre-boot UEFI HII or Comprehensive Configuration Management (CCM) utility] to any PF (that does not have a storage personality enabled), in groups of 8 where the total number of VFs on the port's PFs is equal to or less than 64. The Marvell FastLinQ 3400/8400 (578xx) Series Adapters support up to 128 VFs per adapter. The Marvell FastLinQ 41000 Series Adapters support up to 192 VFs per adapter. The Marvell FastLinQ 45000 Series Adapters support up to 240 VFs per adapter.

With NPAR, each physical port is partitioned into multiple physical functions on the PCIe bus. Marvell's unique concurrent NPAR and SR-IOV technology allows the physical ports

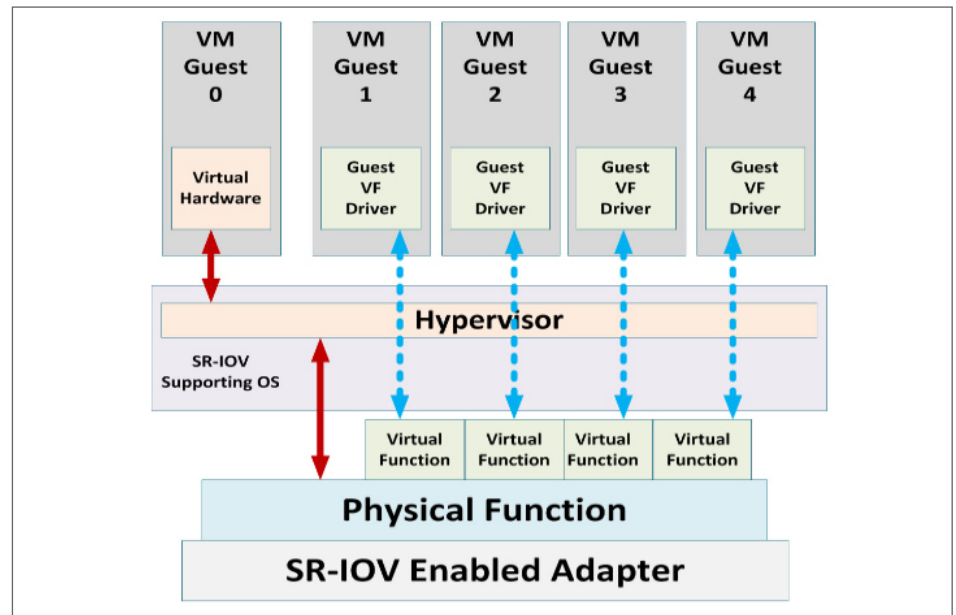


Figure 2. Diagrammatic Representation of SR-IOV

to be further partitioned into multiple virtual functions. This feature allows administrators to deploy NPAR in both bare metal (non-virtualized) OSs and virtualized OSs. In contrast, SR-IOV is primarily targeted toward virtualized platforms. Combining the two allows them to share their unique benefits— independent PFs with or without storage personalities, host teaming of selected PFs (for host applications and select vSwitches) while being able to use SR-IOV VFs to select VMs through the other PFs of the same physical port with VM teaming of SR-IOV VFs when needed.

Currently, a single VM can have no more than eight VFs assigned to it.

Virtual Server Architectures Leveraging Concurrent NPAR and SR-IOV

Figures 3 and 4 show examples of network architectures for virtual server deployments on Windows Server 2012/2012R2/2016 Hyper-V.

- Host applications (such as Live Migration) can use a partition directly (pNIC) or through a Host Team (tNIC) or through a “shared” vSwitch connection (vNIC).
- Hyper-V vSwitches can use a partition directly or through a Host Team.
- VM vNICs can use a vSwitch connection or SR-IOV VF connection.
- VMs can team vNIC connections from any source (from a vSwitch or VF).
- The partitions Maximum BW and Relative (Minimum) BW Weight QoS settings apply to all transmit traffic (Host, VM vSwitch, and VM VF) passing through it.
- A partition with a Host Team (both SF and NPAR mode) or with Storage enabled (NPAR mode only) cannot use SR-IOV VFs; otherwise, the PF and multiple VF connections can share the same partition.

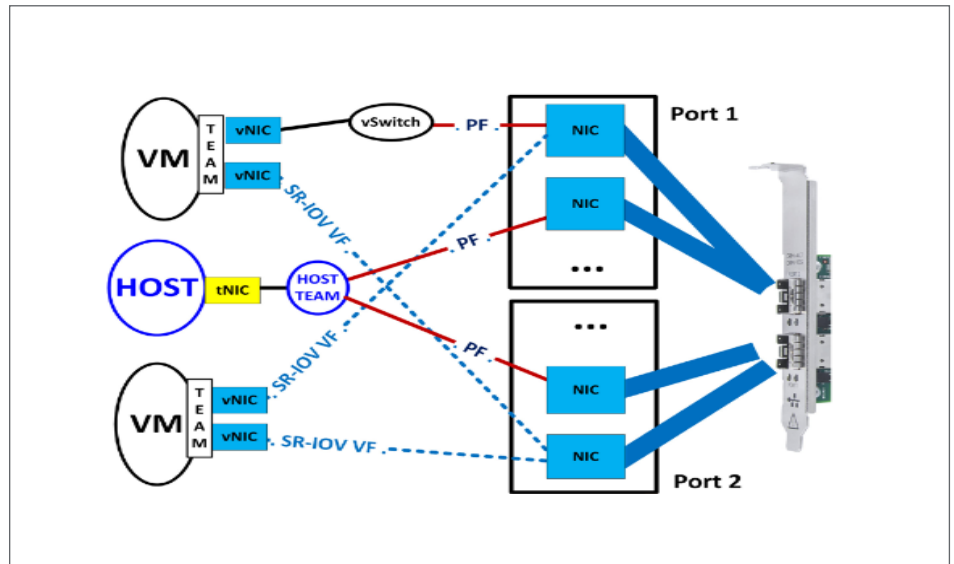


Figure 3. Marvell FastLinQ Enables Concurrent Hypervisor and VM Teams with a Single Adapter

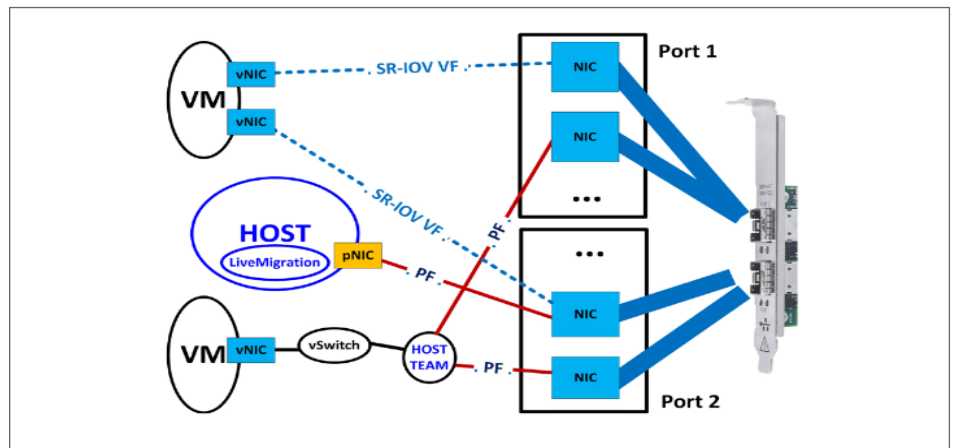


Figure 4. Marvell FastLinQ Enables QoS Across All Hypervisor and VM Services with a Single Adapter

For example (Figure 5), if you wanted to provide redundant load balanced 10GbE/25GbE/40GbE/50GbE/100GbE paths for all traffic (Host Live Migration/Management/etc, VM vSwitched, and VM SR-IOV VF) on a virtualized server (such as with Hyper-V) using different VLANs, you would need at least six separate ports (without NPAR mode).

- One pair of ports would need to be teamed together (at the Host level) and could be used by the various Host applications (each application could be on its own VLAN).
- Another pair of ports would need to be teamed together (at the Host level) and used by the VM vSwitched traffic that does not need the advantages of SR-IOV—with each vNIC connection pair possibly being on its own VLAN.
- The final pair of ports would not be teamed together (at the host level) but would instead use SR-IOV VF connection pairs to the desired VMs (which would team those vNIC connection inside the VM to provide load balancing and failover for those traffic pathways) and each of these pairs could be on their own VLAN.

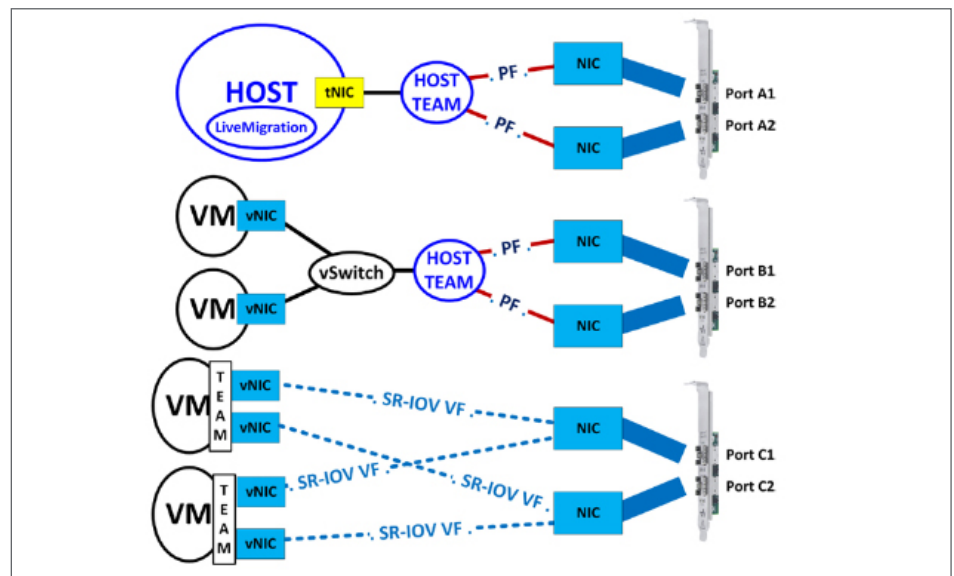


Figure 5. SR-IOV without NPAR Adapter

With Marvell FastLinQ Concurrent NPAR and SR-IOV mode (Figure 6), all of this could be done by simply using a dual port adapter (each port with multiple shared partitions).

- One pair of partitions (port¹ partition¹ + port² partition¹) could be teamed together (at the Host level) and be used by the Host applications (each application on its own VLAN).
- Another pair of partitions (port¹ partition² + port² partition²) could be teamed together (at the Host level) and be used by the VM vSwitched traffic that does not need the advantages of SR-IOV – with each vNIC connection pair possibly being on its own VLAN.
- The final pair of partitions (port¹ partition³ + port² partition³) could use SR-IOV VF connection pairs to the desired VMs (which could use teaming of those vNIC connections at the VM level to provide load balancing and failover for those traffic pathways) and each of these connection pairs could be on their own VLAN. Additionally,

Universal RDMA-Offload (iWARP/RoCEv1/RoCEv2) and/or FCoE-Offload and/or iSCSI-Offload could be used on the remaining partition pairs or share the first or second teamed partition pair to provide storage connections with MPIO-ed load balancing and failover protection. And each partition can have its own QoS settings.

This would reduce the number of NIC ports needed, the number of switch ports consumed, and the associated cabling—this does not even consider the power savings

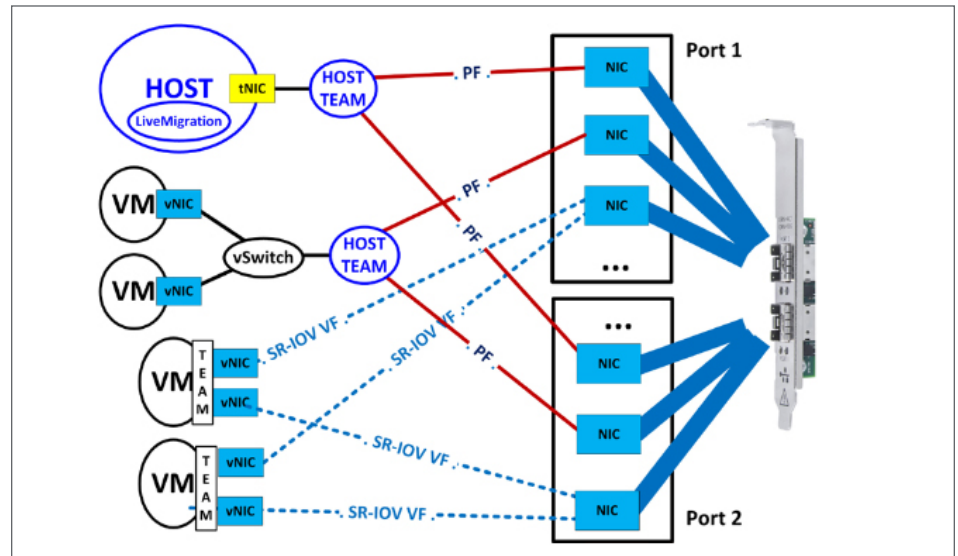


Figure 6. SR-IOV with NPAR

and QoS benefits

Key Benefits

Here are some of the ways NPAR with SR-IOV deliver value to an enterprise and/or MSP data center:

NPAR Benefits:

- Reduced Capital and Operational Expenditures
 - Enables hardware consolidation of multiple slower speed ports into a single higher speed port
 - Allows the creation of multiple partitions (both PF and VF) for flexible bandwidth provisioning on FastLinQ adapters
 - Provides lower adapter, cabling, switch port, and management costs.
- Eliminates Dependency on a Captive Switch
 - Removes dependency on an external Ethernet switch to regulate and manage bandwidth
 - Offers customers the freedom to choose an Ethernet switch of their choice.
- Simplified Deployment and Management

- Provides concurrent support for Ethernet, Universal RDMA (41000/45000), FCoE, and iSCSI protocols, which appear as discrete functions similar to native OS, minimizing deployment disruptions
- Eliminates the need for OS changes to implement flexible bandwidth provisioning
- Quality of Service (QoS)
 - Provides maximum bandwidth QoS of all traffic through a specific PF
 - Provides relative (minimum) bandwidth weight QoS of all traffic through a specific PF
 - Provides DCB-ETS minimum bandwidth weight QoS of all traffic through a physical port

SR-IOV Benefits:

- Improved Ability to Scale the Business
 - Reduces I/O emulation overhead
 - Assigns I/O hardware directly to VMs
 - Allows the dedication of bandwidth (QoS) for VMs and associated applications
 - Conserves PCIe slots in all server environments
 - Enables high I/O performance in virtualized environments

SR-IOV Over NPAR Combined Benefits:

- All of the previous two benefits plus
 - Allows cohabitation of host teaming, storage, VM vSwitching, and SR-IOV VFs on significantly less physical ports
 - Quality of Service for a group of SR-IOV virtual functions associated with a physical function—Seamless co-existence of teaming and Direct I/O (SR-IOV) on the same physical adapter port
 - Reduces the number of adapter and switch ports required to host a high availability high performance virtual server deployment by half

Summary

Most OS and hypervisor vendors support SR-IOV. It enables virtualization of I/O data paths at the server level. Marvell is the only solution provider that offers SR-IOV and NPAR operating concurrently as a standard way of addressing the customer needs of today for the network partition in function. Marvell's NPAR is the OS-, platform-, and switch-agnostic solution available today. NPAR in combination with SR-IOV provides a comprehensive solution for the heterogeneous enterprise data center while delivering CAPEX savings by reducing the total number of physical ports required.



To deliver the data infrastructure technology that connects the world, we're building solutions on the most powerful foundation: our partnerships with our customers. Trusted by the world's leading technology companies for 25 years, we move, store, process and secure the world's data with semiconductor solutions designed for our customers' current needs and future ambitions. Through a process of deep collaboration and transparency, we're ultimately changing the way tomorrow's enterprise, cloud, automotive, and carrier architectures transform—for the better.

Copyright © 2020 Marvell. All rights reserved. Marvell and the Marvell logo are trademarks of Marvell or its affiliates. Please visit www.marvell.com for a complete list of Marvell trademarks. Other names and brands may be claimed as the property of others.