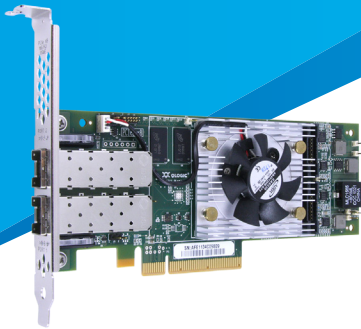


Optimize Server Virtualization with Cavium 10GbE Secure SR-IOV

Flexible, Secure, and High-Performance Network Virtualization with Cavium 10GbE SR-IOV Solutions



Cavium 8300 Series 10GbE Converged Network Adapters from Cavium offer a flexible and secure SR-IOV solution that is qualified with major server and hypervisor vendors and offers a seamless path to high-performance I/O virtualization for enterprise, public and private, cloud deployments.

BENEFITS OF CAVIUM SR-IOV

- **High Performance** – VMs have direct access to the network adapter
- **Low Latency** – VM network traffic bypasses the hypervisor and associated data movement
- **Efficient Traffic Switching** – VM-to-VM traffic is switched in the adapter hardware
- **Lower CPU Utilization** – Switching of Ethernet frames is offloaded from the hypervisor vSwitch to the adapter, reducing server CPU utilization
- **Reduced Load on External Switch Infrastructure** – VM-to-VM traffic does not burden the external switch and LAN infrastructure
- **Secure** – Cavium’s unique implementation provides complete isolation of VM network traffic

TECHNOLOGY SUMMARY

Consolidation driven by server virtualization is accelerating the move to 10GbE technology in data center networks. Virtualized environments require higher performance and greater bandwidth with the least overhead on server CPU cycles. Server virtualization technologies have made significant progress over the past decade on improving the efficiency of virtualizing and sharing processor, memory, and peripheral devices. However, achieving high performance with low overhead for I/O has remained a significant challenge, especially for 10Gb Ethernet and other high-speed I/O interconnects.

Software-based I/O virtualization solutions implemented in hypervisors suffer from significant performance degradation and require considerable host CPU cycles to work effectively, thereby limiting the adoption of the solutions. An emerging and highly efficient approach to addressing the performance issues of I/O virtualization is the use of Single Root I/O Virtualization (SR-IOV), a PCI special interest group (PCI-SIG®) standard. SR-IOV provides a set of hardware enhancements for a PCIe® device (such as a Cavium 10GbE Converged Network Adapter), which removes major hypervisor intervention for data movement from the virtual machines (VMs) to the network. With SR-IOV, a Cavium 8300 Series 10GbE Converged Network Adapter from Cavium can present itself as multiple virtual devices. Each of these virtual devices can be dedicated to a VM. The VM has direct access to the networking adapter, thereby bypassing a significant and resistive part of the hypervisor.

Cavium 8300 Series 10GbE Converged Network Adapters offer a flexible and secure SR-IOV solution that is qualified with major server and hypervisor vendors and offers a seamless path to high-performance I/O virtualization for enterprise, public and private, cloud deployments.

While SR-IOV enables significant performance benefits, there are a few limitations. For example, it requires special support in server hardware and hypervisors. In addition, SR-IOV increases the complexity in achieving common VM tasks, such as checkpointing and live migration as it bypasses the hypervisor. For I/O virtualization needs in deployments on legacy systems, the Cavium 8300 Series Converged Network Adapter also offers a server- and hypervisor-agnostic I/O virtualization technology called NIC Partitioning.

BACKGROUND

I/O virtualization is an emerging category of technologies aimed at resolving the performance, provisioning, and management problems of server I/O in highly virtualized and clustered data center environments. In the evolution of the virtualized data center, the issues of CPU and storage virtualization are actively being addressed, but are raising problems with the load, provisioning, and management of storage and network infrastructure. The technology evolution started with virtualized processors and memory. Next came VMs, then virtual storage, and finally I/O virtualization (IOV), where the I/O path from the server to the peripheral device is now virtualized.

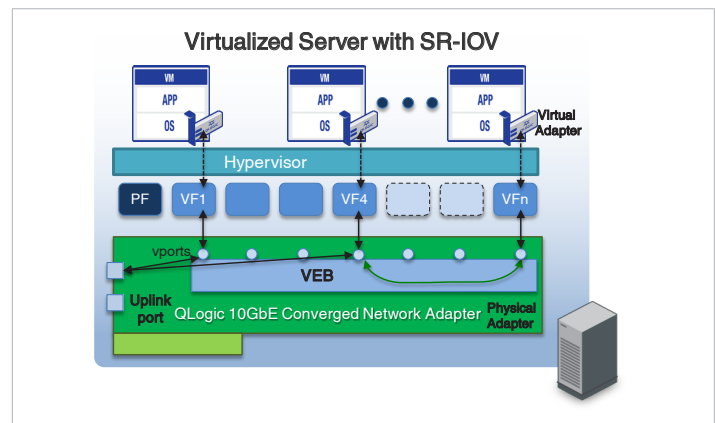
SR-IOV VIRTUALIZATION ARCHITECTURE

PCI-SIG has completed a suite of I/O virtualization specifications that in conjunction with system virtualization technologies allows multiple OSs running simultaneously within a single computer to natively share PCI Express® devices. With SR-IOV, the Cavium 8300 Series 10Gb Converged Network Adapter appears to the operating system (OS)/hypervisor as multiple physical functions (PFs), virtual functions (VFs), and an embedded Layer-2 switch.

- **Physical Function (PF)** – One PF is exposed per physical port of the adapter. 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 in some cases for I/O from the parent partition itself. The PF is strictly owned by the native OS/hypervisor. From a security perspective, the PF is considered as a trusted function.
- **Virtual Function (VF)** – VFs are lightweight PCIe functions associated with a PF. VFs are associated with VMs and are limited to processing I/O streams, moving data from the VM to the network adapter. I/O traffic from the VF to the network adapter bypasses the hypervisor, improving performance and reducing latency. VFs do not support management of the physical device. The Cavium 8300 Series Converged Network Adapter supports up to 64 VFs per PF. From a security perspective, VFs are considered un-trusted functions.

- **Virtual Edge Bridging (VEB)** – A VEB is a basic Layer-2 Ethernet switch embedded into the adapter. There is one eSwitch per physical port. The VEB is capable of switching traffic between all VFs on a PF and between the PF and VFs. The embedded switch enables efficient routing of VM-to-VM traffic without burdening the external switch infrastructure and preserving host CPU resources.

Although from a standards perspective SR-IOV applies to both networking and storage I/O, most current implementations, including the Cavium implementation, are focused on networking I/O only.



CAVIUM 10GBE CONVERGED NETWORK ADAPTER SR-IOV

Features

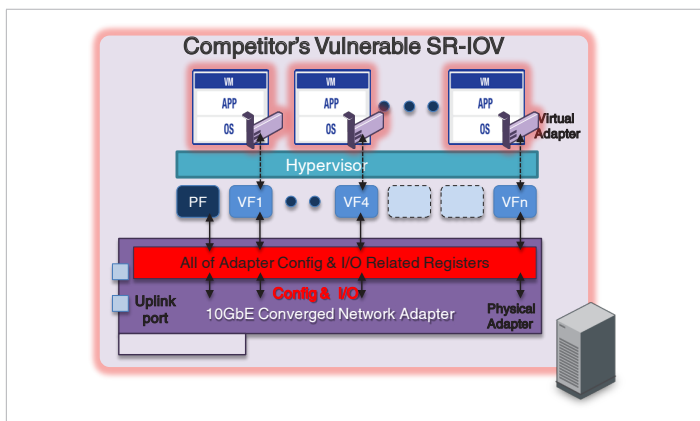
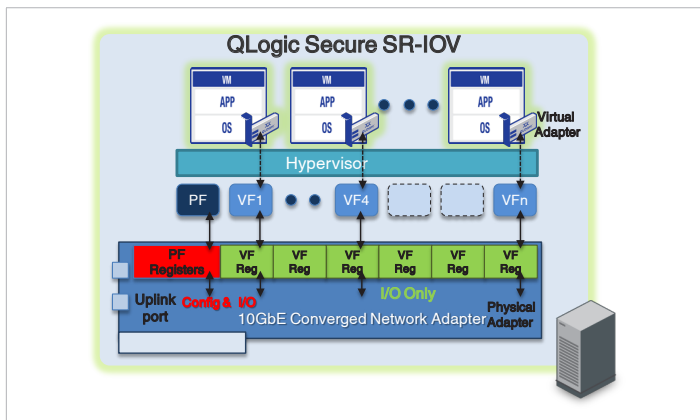
Cavium 8300 Series 10GbE Converged Network Adapters provide a fully featured implementation of networking SR-IOV that is compliant to the PCI Express Base Specification, revision 3.0, and Single Root I/O Virtualization and Sharing Specification, revision 1.1. It has the following key features:

- A maximum of two SR-IOV-capable PFs per adapter
- A maximum of 64 SR-IOV VFs per PF and up to 128 VFs per adapter; However, the hypervisor may impose a limit smaller than this number
- An adapter-embedded L2 switch (VEB) switches traffic between VFs and between the PF and VFs. The switch does not switch traffic between PFs
- The assignment of PFs to a hypervisor is supported. But some hypervisors like VMware® ESXi 5.1 do not allow this. VFs can only be assigned to VMs and not to the hypervisor
- All functions of the Converged Network Adapter—physical and virtual—can concurrently carry I/O traffic
- The hardware design of the Converged Network Adapter offers complete isolation of traffic between VFs
- For security, the Cavium implementation allows only the PF to have the capability to perform operations that affect the health and behavior of the entire adapter or physical port. VFs, which could be owned by tenants in public cloud deployment, are not allowed to perform any disruptive operations to protect against damage from a rogue driver—irrespective of whether it is a rogue or trusted driver attached to the VF

- Full support for leading hypervisors supporting SR-IOV, including VMware® ESXi 5.1+, Windows Server® 2012, and Linux®-based hypervisors

Secure SR-IOV

Cavium 8300 Series Converged Network Adapter hardware implements the concept of trusted and un-trusted functions. PFs are trusted functions and VFs are un-trusted. Only trusted functions have the capability to do operations that affect the operation of the entire adapter or physical port. Firmware restricts control commands issued from a VF if the command adversely affects the operation of other VFs or PFs. Strict cross-function isolation of adapter registers are visible through PCI BARs.



NIC SR-IOV – State of the Union

With all the benefits that SR-IOV has to offer, the current state of SR-IOV is that the technology is nascent. The server support list is small but growing and non-Linux hypervisors have just begun supporting it. Significant challenges remain for Live Migration and Snapshots. So while the adapter SR-IOV ecosystem is fully developed, SR-IOV adoption will continue to gain traction as the ecosystem matures.

SR-IOV Features	Windows	VMware
SR-IOV Support	Windows Server 2012	VMware ESXi 5.1
Total Number of VFs	No OS limit	128
VM Guest OS Support with SR-IOV	W2012	W2008/2012
Live Migration with SR-IOV	✓	✗
Advanced Virtual Edge Bridging	✓	✗
Data Center Wide Management	✓	✗
Extensive OS CLI Support for SR-IOV	✓	✗

Alternative Solutions

SR-IOV is a standard recognized by most OS and hypervisor vendors. It enables virtualization of I/O data paths at the server level. Cavium offers NPAR as a standard way of addressing the customer needs of today for the network partitioning function while the SR-IOV technology is being adopted. Cavium NPAR is an OS-, platform-, and switch-agnostic solution available today. NPAR together with SR-IOV provides a comprehensive solution for the heterogeneous enterprise data center.

Features	Cavium Networking SR-IOV	Cavium NIC Partitioning (NPAR)
Ability to Partition a 10GbE Port into Multiple Partitions	✓	✓
Maximum Number of Partitions or Functions Available	128	8
Server Hardware Platform Agnostic	✗	✓
OS Agnostic	✗	✓
Quality of Service (Transmit Bandwidth Control)	✗ (lack of support in leading hypervisors)	✓

CAVIUM ADVANTAGE

Cavium 8300 Series Converged Network Adapter provides significant benefits in its SR-IOV implementation as compared to the competition.

SR-IOV Features	Cavium	Emulex	Intel
Number of PFs	2	2	2
Number of VFs per PF	64	32	32
Secure SR-IOV Implementation	✓	✗*	--

*Derived based on the Emulex BE3 SR-IOV security advisory at <http://www.emulex.com/downloads/sr-iov.html>.

ABOUT CAVIUM

Cavium, Inc. (NASDAQ: CAVM), offers a broad portfolio of infrastructure solutions for compute, security, storage, switching, connectivity and baseband processing. Cavium's highly integrated multi-core SoC products deliver software compatible solutions across low to high performance points enabling secure and intelligent functionality in Enterprise, Data Center and Service Provider Equipment. Cavium processors and solutions are supported by an extensive ecosystem of operating systems, tools, application stacks, hardware reference designs and other products. Cavium is headquartered in San Jose, CA with design centers in California, Massachusetts, India, Israel, China and Taiwan.

Related Video: QLogic KnowHow – SR-IOV Configuration for Cavium 8300 Series Adapters



Click on the Image to Watch a YouTube Video



Follow us:      

Corporate Headquarters Cavium, Inc. 2315 N. First Street San Jose, CA 95131 408-943-7100

International Offices UK | Ireland | Germany | France | India | Japan | China | Hong Kong | Singapore | Taiwan | Israel

Copyright © 2013 - 2017 Cavium, Inc. All rights reserved worldwide. QLogic LLC (formerly QLogic Corporation) is a wholly owned subsidiary of Cavium, Inc. Cavium and QLogic are registered trademarks or trademarks of Cavium Inc., registered in the United States and other countries. All other brand and product names are registered trademarks or trademarks of their respective owners.

This document is provided for informational purposes only and may contain errors. Cavium reserves the right, without notice, to make changes to this document or in product design or specifications. Cavium disclaims any warranty of any kind, expressed or implied, and does not guarantee that any results or performance described in the document will be achieved by you. All statements regarding Cavium's future direction and intent are subject to change or withdrawal without notice and represent goals and objectives only.