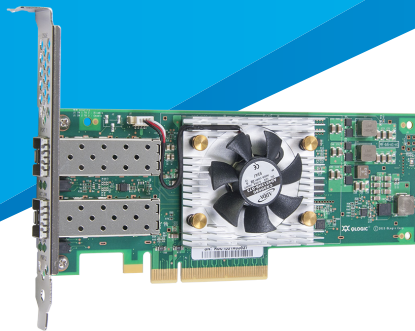


Network Function Virtualization Using Data Plane Developer's Kit

Enabling 25GbE to 100GbE Virtual Network Functions with Cavium FastLinQ Intelligent Ethernet Adapters



- DPDK addresses key scalability issues of NFV workloads
- Cavium FastLinQ 10|25|40|50|100GbE Adapters offer a robust set of capabilities addressing NFV requirements
- Cavium PMD enables 25GbE and 100GbE NFV-based platforms using DPDK
- The combination of a DPDK poll mode driver (PMD) and OpenStack creates the foundation for VNFs

EXECUTIVE SUMMARY

Network data centers are undergoing major transformations by introducing virtual network devices to provide the agility and efficiency required today.

Network Function Virtualization (NFV) provides similar capital expenditure (CAPEX) and operational expenditure (OPEX) savings for communications and networking workloads in the network data center, as server virtualization provides for compute and application workloads. A network function workload has different attributes than a compute workload that tests existing processing models, particularly as speeds reach toward 100Gbps. These requirements drive the need for libraries tailored to the data plane or traffic-forwarding component of these network functions. Data Plane Development Kit (DPDK) provides a scalable, proven, mature set of libraries for these workloads.

Cavium is a contributor to DPDK.org and has developed a PMD for Cavium FastLinQ Intelligent Ethernet Adapters, enabling DPDK support on Cavium adapters from 10Gb Ethernet (10GbE) to 100GbE. This complements other key features of Cavium FastLinQ Adapters for the telco market, including tunneling offloads, Network Interface Card (NIC) virtualization, and iSCSI offloads.

TELCO MARKET OPPORTUNITY

Just as cloud computing has been a simultaneous disrupter and enabler—disrupting traditional data center architectures while enabling many new market opportunities—some of the core underlying technology is enabling a similar opportunity in the telecommunications marketplace.

In the recent past, deployment of new services was done with purpose-built telecommunications equipment. With the rapid global expansion of mobile devices with roughly two billion new mobile phones shipped annually¹, carriers need to be much more responsive in deploying network services. Netflix, which drives roughly one-third of the traffic on the North American Internet during peak viewing hours², is currently launching a global expansion, announcing extension to 130 new countries earlier this year³, and carriers worldwide are bracing for the impact on their networks.

With these dynamics, carriers are now architecting their network data centers to enable the nimbleness necessary for new demands and to improve efficiency, striving to reduce CAPEX and OPEX. Carriers are looking to quickly deploy 25GbE and 100GbE technologies in the data center, and are increasingly looking at virtualization of network functions to aid in the efficiency improvements. This virtualization is commonly known as NFV.

¹ Gartner, 2015

² Spangler, 2015

³ Rubin, 2016

Cavium, with its new FastLinQ 45000 Series Adapters, offers key technologies for the data center, including 25GbE and 100GbE interfaces, and key enabling technology for these network workloads, as discussed in this paper.

ABOUT NFV

Virtualization has been a key component of enterprise data centers for more than a decade, as many application workloads moved from physical servers to virtual servers. NFV is the virtualization of traditional networking appliances, enabling the rapid deployment and redeployment of new network services while significantly reducing both OPEX and CAPEX. Examples include virtual load balancers, firewalls, intrusion detection systems, content delivery appliances, and Wide Area Network (WAN) accelerators.

As these network workloads are virtualized, compute workloads become constraints, particularly for appliances that are built for operation in 25GbE to 100GbE networks. Processing network workloads must be highly optimized to run at wire speed with minimum-sized (64 byte) Ethernet frames.

Memory management and scheduling are key concerns when developing applications for NFV. Network workloads are characterized by large flows of packets that must be placed in memory for classification and dispensation. Physical network appliances have dedicated memory pools for these packets. Compute-centric virtualization acquires memory on an as-needed basis. Typical compute-centric operations interrupt the virtual machine (VM) when there is a new packet from the network, requiring a kernel context switch, and require locks to be acquired to ensure proper memory protection. In addition, data is copied from kernel space to user space. Underlying this is lock management and interrupt management that hinders the ability to scale to support processing of millions of packets per second. To achieve the per-packet instruction count required with a network workload, these operations must be avoided—simple optimization is not sufficient. Key operations requiring optimization include buffer management, ring management, queue management, flow classification, and network driver architecture.

DPDK

DPDK is a set of libraries and drivers developed over the past five years that address the requirements of the data plane, or forwarding engine, of virtualized network functions (VNFs) for fast packet processing. The data plane is where the packet handling is done in network devices. The DPDK

libraries use a run-to-completion scheduling model coupled with PMDs to eliminate the processing overhead of a scheduler and asynchronous interrupts. DPDK optimizes buffer management to eliminate data copies (zero copy). It also eliminates the use of locks for memory and workload management through optimized queue and ring management to achieve fast data plane performance. Portability is enabled via an environment abstraction layer (EAL) that hides the specifics of the underlying hardware and operating environment and provides a standardized programming interface to libraries.

The DPDK is open source software for Linux and FreeBSD licensed under a BSD license. A project to maintain and enhance the software is coordinated through dpdk.org.

CAVIUM FASTLINQ ETHERNET ADAPTERS

Cavium offers a comprehensive family of Ethernet adapters running at speeds from 10Gbps through 100Gbps. These adapters offer a number of advanced capabilities, including:

- Overlay network offloads for Virtual Extensible Local Area Network (VXLAN)
- Generic Network Virtualization Encapsulation (GENEVE) and Network Virtualization using Generic Routing Encapsulation (NVGRE)
- Single Root I/O Virtualization (SR-IOV) and switch-independent NIC partitioning (NPAR) support
- Storage offloads, including Fibre Channel over Ethernet (FCoE), iSCSI and, iSCSI Extensions for RDMA (iSER) support
- Remote Direct Memory Access (RDMA) support over Internet wide area RDMA protocol (iWARP), RDMA over Converged Ethernet (RoCE), and RoCEv2

CAVIUM PMD

Cavium has developed a PMD for use natively with FastLinQ Adapters and for use on a virtual function using SR-IOV while running DPDK. By using polling rather than software interrupts, the driver has lower overhead and much higher performance for the small packet workloads typical in virtualized network functions, such as firewalls and intrusion detection systems. This driver is also well suited for use with higher speed Ethernet appliances, such as those using 25GbE to 100GbE interfaces. When the PMD is used with SR-IOV, the hardware eSwitch on the FastLinQ Adapter is used in place of the software Open vSwitch (OVS). The Cavium PMD is available at dpdk.org and is supported on major Linux distributions. (See Figure 1.)

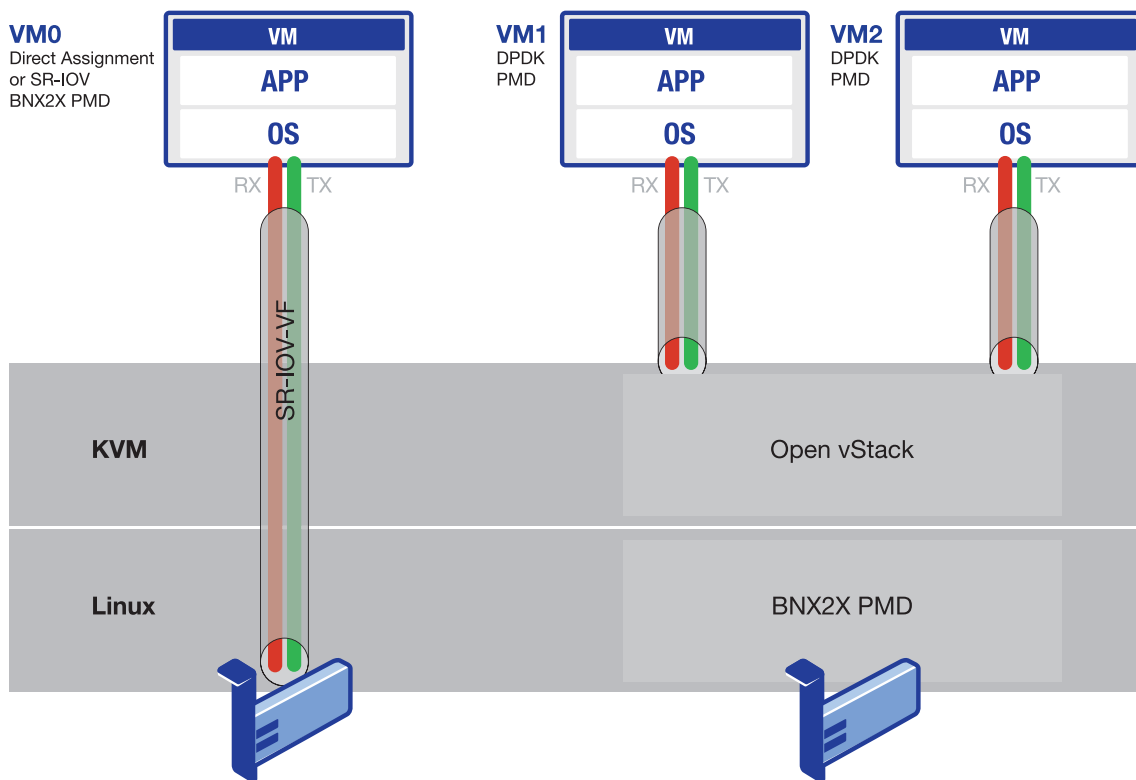


Figure 1. FastLinQ BNx2X PMD

CAVIUM AND DPDK.ORG

Cavium works with dpdk.org to maintain DPDK and is the official maintainer of the Cavium FastLinQ PMD driver. This activity includes support for the PMD on current releases of DPDK and testing for compatibility with upcoming releases of DPDK. The current release of DPDK and Cavium drivers can be found at dpdk.org.

DPDK AND OPENSTACK

OpenStack is the orchestration solution that is commonly used in network data centers and has several components: compute, storage, dashboard, and networking. For networking connectivity under Neutron, FastLinQ Adapters support a comprehensive list of virtualization and multi-tenant services, including SR-IOV and Virtual Extensible LAN (VXLAN) offloads for the most sophisticated enterprise data centers, as well as private, public, and hybrid cloud deployments. With the Neutron networking component, Cavium offers several optimizations beyond DPDK, including VXLAN offloads and OVS integration. Similar to other NFV environments, OpenStack also benefits from the same optimizations from DPDK. The Cavium PMD allows the use of optimized networking for OpenStack using DPDK, as well as virtual network appliances running on the OpenStack Nova compute platform. (See Figure 2.)

CAVIUM OPENSTACK INTEGRATION

FastLinQ Adapters offer a number of key capabilities for the network data center. On the basic NFV workload, the DPDK capability is key. There may be circumstances where network virtualization is required and a solution like NPAR or SR-IOV may be of benefit. On the networking front, support for virtual networking or tunneling with a protocol such as VXLAN allows more flexible network deployment, with logical networks tunneled over the underlying physical network. Here, Cavium FastLinQ tunnel offloads save critical system resources to allow greater NFV capacity. With the use of network virtualization, integration with a DPDK-aware virtual switch also optimizes the solution. iSCSI is frequently used in telco environments, and the use of a high-performance adapter with iSCSI offload capability reserves more system resources for packet handling with the virtualized network functions. In content delivery networks, object-based storage may be used. Here, FastLinQ Adapters have been tuned to work with Ceph to provide object storage support. (See Table 1.)

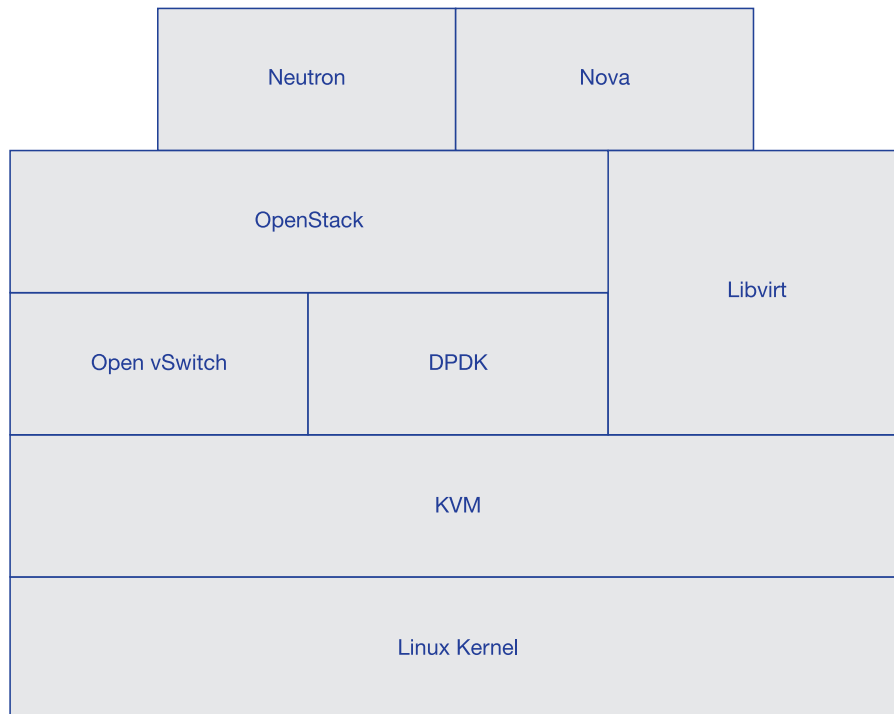


Figure 2. OpenStack and DPDK

Table 1. OpenStack Integration

Compute	Nova	SR-IOV Integration	10GbE I/O (PCI Pass Through) for VNF
			Docker Containers Acceleration
		DPDK Support for NFV	
Networking	Neutron	VXLAN Offload	
		OVS Integration	
		DPDK Support for NFV	
Storage	Cinder / Swift	High Performance Ceph	iSCSI Caching Tier
		iSCSI Offload	
Management	Horizon	ML2 Plugins	
		Python-based installers	SR-IOV
			VXLAN

DEPLOYING FASTLINQ ADAPTERS WITH NFV

As customers develop and deploy NFV solutions using FastLinQ Adapters, they typically use DPDK-specific PMD building blocks in the deployment. The PMD is a use space library that allows the NFV user space application to directly address the hardware and avoid kernel mode context switches and data copies that would impede performance for the network function being deployed. Customers typically use OpenStack to orchestrate the solution, enabling the setup and teardown of the underlying infrastructure and the virtual network devices as business conditions dictate.

SUMMARY

Network data centers are undergoing major transformations by introducing virtual network devices to provide the agility and efficiency required today. The DPDK addresses key scalability issues for communications and networking workloads when using NFV. Users of Cavium FastLinQ 10/25/40/50/100GbE Adapters receive a robust set of capabilities addressing NFV requirements. Cavium PMD enables the industry's first 25GbE and 100GbE NFV-based platforms using DPDK. Customers looking to deploy DPDK on OpenStack or KVM using Linux-based virtual network appliances can build compelling solutions using Cavium FastLinQ Adapters.

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.



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 © 2016 - 2017 Cavium, Inc. All rights reserved worldwide. Cavium, FastLinQ, and QConvergeConsole are registered trademarks or trademarks of Cavium, Inc. All other brand and product names 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.