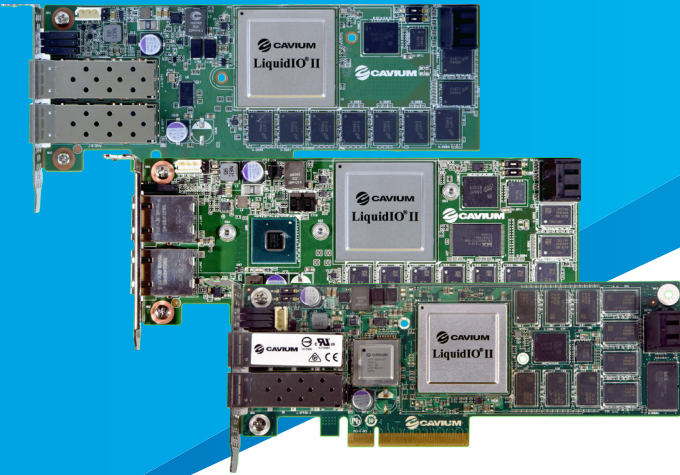


# LiquidIO® OVS Software Architecture



LiquidIO® Smart NICs from Cavium™ accelerate complex packet processing and preserve CPU cycles, allowing data centers to deploy more VMs on the server

## INTRODUCTION

Cavium™ LiquidIO® Smart NIC family enables data centers to rapidly deploy high-performance SDN applications for both installed and new infrastructures while enhancing server utilization, response times, and network agility. LiquidIO adapters are designed for deployment in data centers across market segments including public cloud, cloud service providers, enterprises, and private data centers.

Server virtualization technology has enabled data centers to deploy multiple virtual machines (VM) per server, giving them significant cost benefits. However, the complete benefits of server virtualization cannot be realized without the associated network virtualization, software-defined network (SDN) technology to enable flexible virtualization of the network. Enabling network virtualization in a server-based network architecture requires software-based switching and tunneling to decouple physical and logical networks with significant overhead to server CPU with reduced performance and increased network latency.

Data centers need to tackle these issues to meet their customer commitments on performance, bandwidth, and quality of service (QoS). They need a hardware solution that can offload SDN protocols and network processing from the main server CPUs (server based network) to an intelligent network adapter.

As data centers continuously evolve across different cloud segments, they need a flexible network adapter to efficiently handle additional network and security protocol offloads, unburden the host CPU, and provide differentiated services. This requires an adapter that is fully programmable, supports additional offloads in hardware, and eliminates host CPU utilization with improved network performance and processing.

### LiquidIO Smart NIC

LiquidIO Smart NIC family provides a proven, high-performance, programmable adapter platform for cloud service providers. It is offered in a standard server compliant half-height PCI Express form-factor and is supported by a feature rich Software Development Kit (SDK). LiquidIO Smart NICs provide out-of-box OVS solution that supports VXLAN, NVGRE, and GENEVE standard overlay tunneling protocols at 10G/25G performance.

The SDK allows customers and partners to develop high performance applications with additional security and overlay encapsulation, traffic management/QoS, deep packet inspection (DPI), TCP and SSL processing capability available on LiquidIO Smart NICs.

**Open vSwitch (OVS) Architecture**

Data centers today run a virtual switch on the compute nodes forwarding packets to the right virtual machine (VM) on each node. Open vSwitch (OVS) is the most popular and commonly used open source virtual switch today. LiquidIO Smart NIC fully offloads OVS version 2.6, eliminating the costly VM packet switching, which frees up host CPUs.

Cavium has developed efficient OVS FP (fast path) code for handling VXLAN, NVGRE, GENEVE, and packet switching under Cavium’s Simple Executive (SE) environment. OVS running on LiquidIO is transparent to the host and completely offloads OVS and overlay protocols to LiquidIO.

**LIQUIDIO OVS ARCHITECTURE**

LiquidIO Smart NICs feature MIPS-based general-purpose processing cores. These are configured to run in hybrid mode, where a few cores run Linux and rest of the cores run in the Cavium Simple Executive (SE) environment. Simple Executive is Cavium’s bare-metal application environment that leverages the underlying processing architecture to accelerate packet processing in a run-to-completion model. LiquidIO OVS architecture utilizes Simple Executive environment for offloading overlay protocols and switching functionality while using Linux OS for configuration, management, and exceptions.

LiquidIO OVS solution has two processing paths: FP (fast path) running in SE environment and SP (slow path) running on Linux. Linux OS running on LiquidIO cores are responsible for slow path processing to handle packets not handled by fast path.

Figure 1 presents a block diagram of OVS full offload on LiquidIO. The following sections further describe these blocks.

**Slow Path (SP)**

Linux handles protocol processing that does not require tight cycle budgets using a standard OS packet path, called a “slow path” (SP). The SP processing is further divided into two parts: user space and kernel space. User space consist of OVS vswitchd process, which is responsible for:

- Implementing OpenFlow 1.3 protocol and interacting with the OpenFlow controller
- Performing all management functions, such as collecting flow/ port statistics, creating, deleting, and expiring flows, and servicing interface events
- Interacting with OVS database server (ovsdb\_server) and saving its runtime information (such as virtual ports, interfaces and their statistics) in the database

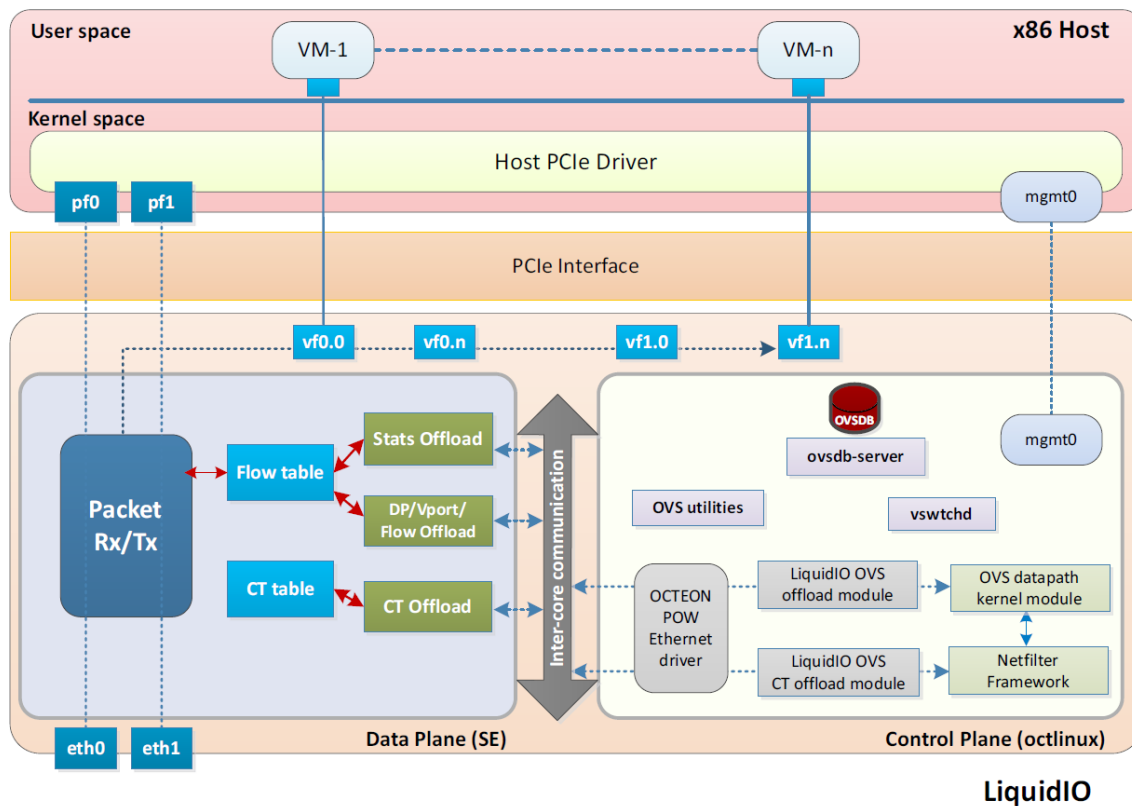


Figure 1. LiquidIO Open VSwitch (OVS) Block Diagram

- Maintaining primary flow table to store the flows and associated actions sent by OpenFlow controller or configured using ovs-ofctl tool
- Communicating with the kernel module for pushing the missing flows in the datapath

**Fast Path (FP)**

The fast path (FP) is designed to be in sync with the slow path OVS components using Cavium's Intern-Component Message Layer (ICMSG), which helps synchronize the slow path OVS configuration with the fast path OVS offload component. Adding and deleting flows and virtual ports in OVS kernel module within Linux are synchronized with FP using the Inter-component Messaging described in ICMSG section of this document.

**Flow Handling in Fast Path (FP)**

The FP is responsible for matching and forwarding transmit and receive packets based on the flow table. Figure 2 shows the packet's transmit and receive path. A flow doesn't exist in the FP flow table in two specific cases:

- This is the first packet of the flow or
- The flow has been deleted from the flow table after expiration. This happens when no traffic for that particular flow has been exercised for a long duration.

Let's look at the example of a packet originating from the network destined to a VM on the host.

Upon packet ingress, OVS running on the fast path cores read the packet from the network. FP looks for flow action in its flow table, which is in sync with SP OVS cache. If there is no entry configured for this flow, the packet is sent to LiquidIO Linux core for slow path processing.

Cavium OVS slow path checks its flow table for the matching flow action. If an OpenFlow controller is configured, the slow path will communicate with OpenFlow controller to get the flow action. The flow and its action are pushed to the fast path flow table. The packet is also sent to the fast path core. The FP now performs the actions defined in the flow table and sends the packet to its proper destination (VM).

Let's consider the case when a flow exists in the FP flow table for a packet transmitted from the VM destined to the network.

The packet originates from VM and is sent via the LiquidIO VF assigned to that VM. This packet traverses the LiquidIO OVS offload fast path.

OVS running in SE finds the flow action for the packet and sends the packet to the egress port (network interface) as per flow action defined.

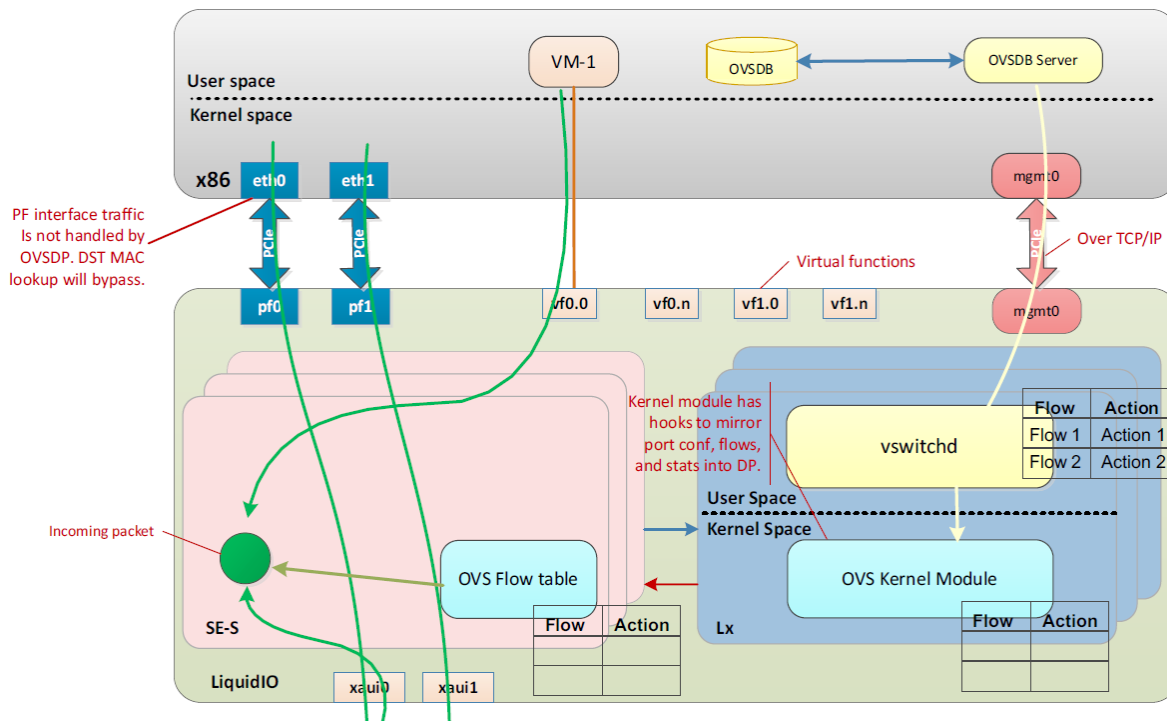


Figure 2. Packet Flow for Flow Miss in FP

Cavium OVS FP offload provides encapsulation/decapsulation support at line rate throughput for widely deployed VXLAN, GENEVE, and NVGRE overlay protocols. OVS FP cores isolate the VM traffic based on the MAC address and or VXLAN ID in the tunnel header. With SR-IOV attached VM, the FP uses this information to switch the packet to the VMs.

#### Inter-component Messaging Layer (ICMSG)

ICMSG is responsible for providing the communication mechanism between Linux slow path and Simple Executive fast path cores. This mechanism uses both shared memory as well as SSO (scheduling, synchronization, and order). The following architectural features are provided by IML:

- Software APIs to register for message deliveries based on message type (op-code)
- Order preservation of unordered operations within message groups
- Bi-directional mode of operation with support for semantics of remote procedure call
- Framework to aggregate multiple messages to reduce overhead of sending/receiving multiple messages individually
- Support synchronous and asynchronous modes of communication

#### OpenFlow

OpenFlow is a communications protocol that gives access to the switch's forwarding plane over the network. In simpler terms, OpenFlow allows software running on switches to determine the path of network packets through the switch network. This separation of the control from the forwarding allows for more sophisticated traffic management than is feasible using access control lists (ACLs) and routing protocols. OpenFlow is an SDN enabler.

Cavium OVS can be remotely controlled by OpenFlow 1.3-compliant controllers tested with industry leading OF controllers like OpenDaylight, Floodlight, Nox, and Beacon.

#### SOFTWARE DEVELOPMENT KIT (SDK)

LiquidIO Smart NICs support a feature-rich SDK that includes GNU tool-chain (compiler, debugger, and profiling tool). Developers use the LiquidIO SDK to develop various specialized network and security offloads available with LiquidIO Smart NIC including OVS and IPsec. The SDK is flexible, allowing customers and partners to develop additional custom and proprietary application-specific offloads for their data center needs or to add new features to Cavium's production-quality software packages.

SDK also includes out-of-box applications in binary and source packages. Source packages can be modified for customizations or for adding extra proprietary features to the existing packages. Well-documented APIs are available as part of SDK to develop custom applications.

#### CONCLUSION

Cavium OVS software provides the ability to accelerate the OVS data path by using LiquidIO Smart NICs. OVS control plane and data plane are completely offloaded to LiquidIO Smart NIC. Thus, LiquidIO Smart NICs accelerates complex packet processing and preserves CPU cycles, allowing more VMs to be deployed on the server.

#### 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 © 2017 Cavium, Inc. Cavium, QLogic, All rights reserved worldwide. Cavium, OCTEON, and LiquidIO are registered trademarks or trademarks of Cavium Incorporated, 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.