FPGA AND VIRTUALIZATION TECHNOLOGY IN DPDK TO ACCELERATING AND SCALING THE CLOUD NETWORKING

Zhang Tianfei

Rosen Xu
AGENDA

Part 1: FPGA and OPAE
- Intel® FPGAs and the Modern Datacenter
- Platform Options and the Acceleration Stack
- FPGA Hardware overview
- Open Programmable Acceleration Engine (OPAE)

Part 2: FPGA, OPAE and DPDK
- DPDK Overview
- FPGA in cloud Networking
- FPGA Acceleration on DPDK
- DPDK for vSwitch FPGA Acceleration
Part 1: FPGA and OPAE
Data Movement and Processing Explosion

- Hyper-connected World
- Big Data Processing and Analytics
- High-Performance Computing Demands
- 5G Wireless

Markets
- Government
- Enterprise
- Cloud
- Communications

Infrastructure
- Network
- Storage
- Compute

Workloads
- Security
- Big Data Processing and Analytics
- Video processing and transcode
- Artificial Intelligence & Machine Learning
- Packet processing

† Source: "Gartner Says 8.4 Billion Connected "Things" Will Be in Use in 2017, Up 31 Percent From 2016", 2/7/2017, http://www.gartner.com/newsroom/id/3598917 (Table 1 - IoT Units Installed Base by Category, 2020 column – Grand Total, including consumer+business units)
ACCELERATION ASSIST TO PROPEL DATA INSIGHT & OPERATIONAL EFFICIENCY

**Workload Optimization:**
Ensure Intel® Xeon® CPU cores serve highest value processing

**Efficient Performance:**
Improve performance/watt

**Real-Time:**
High bandwidth connectivity and low-latency parallel processing

**Developer Advantage:**
Code re-use across Intel FPGA data center products

The Intel® Xeon® processor with FPGA acceleration can reduce TCO and solve new problems
INTEL® FPGA DATA CENTER FORM FACTORS OPTIONS
Enabled By The Acceleration Stack for Intel® Xeon® CPU with FPGAs

PCIe Acceleration Cards

- System flexibility with Intel Xeon CPU SKU options
- Dedicated local memory
- Can be slotted into 1U servers

Server Platform Option with In-Package FPGA

- Coherent interface benefits software developers
- Superior performance for bandwidth & latency sensitive applications

Choose the Intel FPGA form factor matched to your application needs

1^UPI = Intel® Ultra Path Interconnect
2^HSSI = High Speed Serial Interface
THE INTEL® APPLICATION DEVELOPER ADVANTAGE

Acceleration Stack for Intel® Xeon® CPU with FPGAs – Enhanced Performance, Simplified

- Saves developer time to focus on unique value-add of their solution
- Enables unprecedented code re-use across multiple Intel FPGA form-factor products
- World’s first common developer interface for Intel FPGA data center products
- Optimized and simplified hardware and software APIs provided by Intel
- Enables easier development and deployment of Intel FPGAs for workload optimization

The stable and optimized foundation for building your Intel FPGA-accelerated solution
ACCELERATION STACK FOR INTEL® XEON® CPU WITH FPGAS
Enhanced Performance, Simplified

Dynamically Allocate Intel® FPGAs for Workload Optimization

Rack-Level Solutions

Simplified Application Development

User Applications

Leverage Common Frameworks

Industry Standard SW Frameworks

Fast-Track Your Performance

Acceleration Libraries

Workload Optimization with Less Effort

Intel Developer Tools

Leverage Common Frameworks

Intel® Parallel Studio XE, Intel FPGA SDK for OpenCL™, Intel Quartus® Prime

Common Developer Interface for Intel FPGA Data Center Products

Intel Acceleration Engine with OPAE Technology, FPGA Interface Manager (FIM)

Intel® delivers a system-optimized solution stack for your data center workloads

OpenCL and the OpenCL logo are trademarks of Apple Inc. used by permission by Khronos
Logos and names provided for illustrative purposes only. Current availability may be different.
OPEN PROGRAMMABLE ACCELERATION ENGINE (OPAE) TECHNOLOGY
Simplified FPGA Programming Layer for Application Developers

- Consistent cross-platform API
- Minimal software overhead and latency
- Supports virtual machines and bare metal platforms
- Open source code licensing and developer community
  - Intel FPGA drivers being upstreaming to Linux kernel
  - Intel FPGA userspace drivers have merged into DPDK

Start developing for Intel FPGAs with OPAE today: http://01.org/OPAE

**ASE requires Acceleration Functions written in RTL and a properly installed RTL simulator: Synopsys* VCS-MX, Mentor Graphics* ModelSim-SE*/QuestaSim

Supports: Red Hat Enterprise Linux* 7.3 w/ kernel 4.7, Intel® Xeon® Processors v4 or newer
ACCELERATION ENVIRONMENT
Common Developer Interface For Intel FPGA Data Center Products

- **CPU**
  - User Application & Libraries
  - Intel® Acceleration Engine with OPAE\(^1\) Technology
  - Hypervisor & OS
- **FPGA**
  - Accelerator Function (Developer created or provided by Intel)
  - FPGA Interface Manager (FIM)

Optimized and simplified hardware and software APIs provided by Intel

---

1. OPAE = Open Programmable Acceleration Engine
2. UPI = Intel® Ultra Path Interconnect
3. HSSI = High Speed Serial Interface

Supports: Red Hat Enterprise Linux® 7.3 w/ kernel 4.7, Intel® Xeon® Processors v4 or newer
FPGA INTERFACE MANAGER (FIM) OVERVIEW

Device memory organized in Device Feature List data structure

Supported features exposed through Device Feature List
FPGA INTERFACE MANAGER (FIM) DETAILS

- FPGA Management Engine
  - Provides: power and thermal management, error reporting, partial reconfiguration, performance reporting, and other infrastructure functions.
  - Each FPGA has one FME, accessible through the physical function.

- Port
  - Interface between the static FPGA fabric (FIM) and a partially reconfigurable region containing an Accelerated Function Unit (Accelerator Function).
  - Controls communication from SW and exposes features such as reset and debug.

- Accelerated Function Unit
  - Attached to a port and exposes a MMIO region for accelerator-specific control registers.
FPGA INTERFACE MANAGER (FIM) – VIRTUALIZATION SUPPORT

Supports PCIe SR-IOV function to create virtual functions (VFs) which can be used to assign individual accelerators to virtual machines.

Diagram:

- **FPGA PCIe Device**
  - FME
  - Port #0
  - Port #1
  - Port #2

- **Host**
  - PCIE PF
  - FME

- **Virtual Machine #1**
  - PCIE VF#1
    - Port #0
    - AFU
  - PCIE VF#2
    - Port #1
    - AFU

- **Virtual Machine #2**
  - PCIE VF#3
    - Port #2
    - AFU
OPAE USERSPACE DRIVER INFRASTRUCTURE

DPDK Application

Non-DKDK Application

DPDK Framework

standalone Framework

OPAE hardware layer API

ifpga base code

Map Card base code

other FPGA

opae_fpga_adapter API (Enumeration)

opae_manager API (Management)

opae_accelerator API (Acc Access)

opae_bridge API (Acc control)
OPAE User space Driver Architecture (Bare metal case)

- FPGA Management by OPAE Kernel Driver (Upstream in progress)
- DPDK in VM with Normal PMD
Summary

- OPAE is powerful and open software stack for FPGA to accelerating applications.
- OPAE can offer two kinds of drivers:
  - user space driver solution: has merged into DPDK 18.05
  - kernel driver solution: upstreaming now
- Start developing for Intel FPGAs with OPAE today: http://01.org/OPAE
Part 2：FPGA, OPAE and DPDK
DPDK Framework – BIG Picture

Data Plane APIs

VNF
Event Handler
DPI RegEx
User stack
IPsec

Run time
Crypto
Classifier
Compression

Physical
Virtualization
Container

* Other names and brands may be claimed as the property of others.
FPGA in cloud Networking

- **Opportunities**
  - Enhancing Performance: Provide NIC ASIC liked performance
  - Changing dynamically: Flexible enough for adding new feature

- **Problems**
  - Longer design cycle than software: Compilation, Analysis & Synthesis, Fitter(Place & Router), Assembler, Timing
  - Online upgrade affecting business: PCIe rescan and driver reprobe

* Other names and brands may be claimed as the property of others.
Partial Reconfiguration (PR)

- With Partial Reconfigure (PR) parts of Bit Stream, FPGA not only provides one kind of accelerator but also provides many types of accelerators at the same time
  - Hot upgraded
  - Resources time-shared
  - Fault tolerance

- How DPDK fully support FPGA?
  - Which type of DPDK Device can provide FPGA PR?
  - How can we bind DPDK Driver to FPGA Partial Bit Stream?
FPGA Acceleration on DPDK - Scope

- Rawdev probed as PCI Driver takes FPGA Configuration (Download/PR)
- 2 scans: FPGA PCI Device Scan (1st Scan) and AFU Scan (2nd Scan)
- OPAE Provides Common lib and API for low level FPGA management & accelerator access

* Other names and brands may be claimed as the property of others.
FPGA Acceleration on DPDK - Architecture

DPDK Application

DPDK Device PMD

FPGA Mgmt (partial reconfiguration)

Network Path (NIC)

librte_rawdev

librte_ethdev

fpga bus

Eth driver (AFU)

DPDK FPGA Device Driver

Bare metal Application

Custom User space driver

FPGA Accelerator

OPAE High Level APIs

OPAE Hardware Layer API

Intel fpga base code

VFIO

User Space

Kernel Space

crypto
event

* Other names and brands may be claimed as the property of others.
DPDK for vSwitch FPGA Acceleration

OVSDP

vswitchd

SW Data path

INIT SYSTEM

FLOW SETUP

DPDK

ethdev

rte_flow

rte_tm

rte_mtr

port_representors

AFU

PMD

Flow+ actions

HQoS

metering/policing

switch port representation

FPGA BUS

vSwitch AFU

FPGA Rawdev Driver

OPAE

FPGA

I/O

Rate Limit

Classify

FIB and Action

Meter

HQoS

DMA

* Other names and brands may be claimed as the property of others.
Summary

- FPGA BUS is in DPDK 18.05
- Start developing for DPDK with OPAE: http://dpdk.org/doc/guides/rawdevs/ifpga_rawdev.html