Intel’s Hyperscale-Ready Infrastructure Processing Unit (IPU)

Brad Burres, Intel Fellow

Co-Authors: Dan Daly, Mark Debbage, Eliel Louzoun, Christine Severns-Williams, Naru Sundar, Nadav Turbovich, Barry Wolford, Yadong Li
Major Advantages of IPUs

Separation of Infrastructure & Tenant
Guest can fully control the CPU with their SW, while CSP maintains control of the infrastructure and Root of Trust

Infrastructure Offload
Accelerators help process these task efficiently. Minimize latency and jitter and maximize revenue from CPU

Diskless Server Architecture
Simplifies data center architecture while adding flexibility for the CSP
Infrastructure Workloads Migrating to IPU

Hypervisor responsible for all infrastructure services

Infrastructure tasks move to IPU or into the network

Free up cycles for more VMs or applications

Host

Hypervisor

Network

Storage

Mgmt. Security Monitor

VM

VM

VM

VM

VM

Foundational NIC

VM

VM

VM

VM

VM

Hypervisor

Network

Storage

Mgmt. Security Monitor

IPU

VM

VM

VM

VM

VM

Network

Storage

Mgmt. Security Monitor

IPU

Network

Storage

Mgmt. Security Monitor
Mount Evans
Intel’s 200G IPU

Hyperscale Ready
- Co-designed with a top cloud provider
- Integrated learnings from multiple gen. of FPGA sNIC/IPU
- High performance under real world load
- Security and isolation from the ground up

Technology Innovation
- Best-in-Class Programmable Packet Processing Engine
- NVMe storage interface scaled up from Intel Optane Tech
- Next Generation Reliable Transport
- Advanced crypto and compression accel.

Software
- SW/HW/Accel co-design
- P4 Studio based on Barefoot
- Leverage and extend DPDK and SPDK
- Enable broad adoption of IPUs
Mount Evans
Architectural Breakdown

Network Subsystem

Compute Complex

High-Level Breakdown

PCIe SerDes

VF 00 VF 01 ...
VF 00 VF 01 ...
VF 00 VF 01 ...
VF 00 VF 01 ...
VF 00 VF 01 ...

PF PF PF PF

PCIe Gen 4 x16 (SR-IOV, S-IOV)

Arm Neoverse N1 Cores

RDMA NVMe LAN

Packet Processing Pipeline

Inline Crypto Traffic Shaper

200G Ethernet MAC

56G Ethernet Serdes

System Level Cache

LP DDR4 LP DDR4 LP DDR4

Lookaside Crypto & Compression Engine

Management Complex
Mount Evans

Network subsystem

- Support for up to 4 host Xeons with 200Gb/s full duplex
- High-performance RDMA running with ROCEv2 & Reliable Transport Protocol
- NVMe device interface with inline AES-XTS and VM QoS for efficient software backend
- Programmable packet pipeline with QoS and telemetry capabilities supporting 200Mpps
- Advanced transmit scheduling capabilities
- Inline IPSec for high scale connection at wire speed
Mount Evans
Compute Complex

- Up to 16 ARM Neoverse N1 Cores @ up to 3GHz for infrastructure apps such as the Storage initiator backend, TLS proxy, vSwitch and other applications.
- Large System Level Cache backed by three channels of LP/DDR4
- Lookaside Crypto and Compression Engine for Host or Compute Complex use
- Dedicated management processor providing secure boot, life cycle management and overall manageability
Mt. Evans - Packet Processing

**Leadership P4 programmable pipeline**
- Support complete vSwitch + beyond fully in hardware
- Pipeline composition via recirculation and chained operations without sacrificing performance
- Programmable Parser, Exact Match, Wildcard Match, Range Match, LPM, Meters, Statistics, Modifier

**Packet Processing at scale**
- @scale classification for > 10M entries backed by DDR
- Support pipeline driven operations like flow auto-add and aging

**Tightly coupled with the Compute Complex**
- Large L1 caches, optionally backed in compute cache, designed to meet hyperscale performance challenges
- Multi-TB cross-sectional BW between the network subsystem and the compute complex
- Broad metadata capabilities, including handoff to software
**Scale out Storage Architecture**

- Control Plane
  - Create New Virtual SSD
  - Flexible Storage
    - NVMeoTCP
    - NVMeoRDMA
    - Ceph
    - BYOS

- Storage
  - Create Virtual NVMe Device
  - Virt + NVMe protocol layer
  - QoS / Load balance
  - HW crypto, compress, CRC

- Hardware
  - PCIe
  - CPU

- Software
  - Customer’s Software
  - Standard OS
  - NVMe – SSD1

- Virtualized Network

- Shared Storage

- SSD1 – 250GB

- Storage Service

- Create New Virtual Volume - SSD1

---

Hot Chips 2021
Mt. Evans – in Depth Security Strategy

A powerful element, a university, a turn.
Mt. Evans – in Depth Security Strategy
Mt. Evans – in Depth Security Strategy

System Security

Isolation and Recovery
Mt. Evans – in Depth Security Strategy

- System Security
- Isolation and Recovery
- Performance
- Flexibility

Diagram showing various components like PCIe Gen 4 x16 (SR-IOV, S-IOV), RDMA, NVMe, LAN, Packet Processing Pipeline, Inline Crypto, Traffic Shaper, 200G Ethernet MAC, 56G Ethernet Serdes, and Management Complex.
Mt. Evans – in Depth Security Strategy

- System Security
- Isolation and Recovery
- Performance
- Flexibility
- Survivability & Uptime
Mt. Evans – in Depth Security Strategy

System Security
Isolation and Recovery
Performance
Flexibility
Survivability & Uptime
Future proof
Security Compliance
Thank you!

Fig. 4
Rays, spots, errors and kisses...