NVIDIA DATA PROCESSING UNIT (DPU) ARCHITECTURE

Idan Burstein, DPU Principal Architect
DATA CENTER IS THE NEW UNIT OF COMPUTING

Monolithic
Software-Defined Datacenter
Disaggregated, Micro-Services, Scaled Out
GPU-Accelerated Computer
DPU-Accelerated Data Center Infrastructure

EVOLUTION OF THE DATA CENTER
INTRODUCING THE DATA PROCESSING UNIT

Software Defined Data Center Infrastructure-on-a-Chip

From Hardware Appliances

To Software Defined Infrastructure on CPU

To Software Defined Infrastructure on DPU

Microservices
East-West Traffic
Storage Access
Zero Trust Security
NAIVE MOVING WORKLOADS TO NIC CPUS DOESN’T WORK

18 DPU Cores Replace 12 Server CPU Cores —
No Gain in Performance or Efficiency
Not compatible for higher bandwidth without requiring significant system modification
DPU MUST INCLUDE HARDWARE ACCELERATION

Today's Environment

Offload, Accelerate, Isolate

Arm Cores Run Control Plane or Security Workloads Requiring Domain Isolation

DPU Accelerators and 8 Arm Cores Replace 20 to 120 CPU Cores — HUGE Efficiency Gain
NVIDIA DPU ROADMAP

Exponential Growth in Data Center Infrastructure Processing

BlueField-2
7B Transistors
9 SPECint*
0.7 TOPS
200 Gbps

BlueField-3
22B Transistors
42 SPECint*
1.5 TOPS
400 Gbps

BlueField-4
64B Transistors
160 SPECint*
1000 TOPS
800 Gbps

DOCA — ONE DEVELOPMENT ARCHITECTURE

* SPECint 2k17-rate
nVIDIA BlueField-3 DPU
First 400Gb/s Data Processing Unit

- 22 Billion Transistors
- 400Gb/s Ethernet & InfiniBand Connectivity (1-4 Ports)
- PCIe Switch Gen 3/4/5 x32+x4
- 400Gb/s Crypto / Security Acceleration
- 330M PPS, 80M PPS at scale of millions of flows
- 18M IOP/s Elastic Block Storage
- 300 Equivalent x86 Cores
- 42 ARM SPECINT2k17
- 128b DDR5-5600
NVIDIA DOCA
Enabling Broad BlueField Partner Ecosystem

Software Development Framework for BlueField DPUs
Offload, Accelerate, and Isolate Infrastructure Processing
Support for Hyperscale, Enterprise, Supercomputing and Hyperconverged Infrastructure
Software Compatibility for Generations of BlueField DPUs
DOCA is for DPUs what CUDA is for GPUs
# BLUEFIELD-3 PROGRAMMABLE ENGINES

## ARM
- 16 Arm A78 cores
- Fully programmable OS
- Apps/services, service chaining
- Control Path / Slow Path
- Memory to Memory Accelerators

## Datapath Accelerator
- 16 cores, 256 threads
- Programmability through DOCA
- Heavy multi-threading application acceleration

## ASAP²
- Programmable packet processor flow pipeline
- Flow table based
- Data Path
Server Class CPU subsystem
- Data center operating system control plane
- Isolated memory subsystem optimized for networking

NIC subsystem
- Isolated boot domain, real time OS
- Accelerating data path at line rate

PCIe subsystem
- Flexible EP/DP assignment, PCIe switching, NTB, p2p communication, optimized for IO

Data acceleration
- Accelerating ARM workload

DPA – Data Path Accelerator | RDMA – Remote Direct Memory Access | ASAP² – Accelerated Switching and Packet Processing
DPU ACCELERATED SWITCHING AND PACKET PROCESSING
Programmable Data Path | Software-Defined Orchestration

Accelerated
- Virtio-Net/Other Emulation
- OoS & scheduling
- Telemetry and statistics
- Micro Segmentation
- Encryption (Ipsec/MACsec)
- Tunneling (VXLAN/GRE)
- NAT
- Routing
- ACL

Software Defined
- eSwitch management
- Connection Establishment
- Key Association
- Monitoring & Stats
- IDS/IPS/WAF

Bare Metal Host

Data Path

Control Path

BlueField

Virtual PCIe Switch
NVMe
IO Processors
RDMA/TCP Pipeline
ASAP2 Pipeline
InfiniBand/Ethernet

Network

BlueField DPU

APM
Hypervisor
SDS
SDN

Virtio-Net/Other Emulation
IDM / IPS / WAF
- NAT
- Routing
- ACL
- Virtual PCIe Switch
- NVMe
- IO Processors
- RDMA/TCP Pipeline
- ASAP2 Pipeline
- InfiniBand/Ethernet
- APM
- Hypervisor
- SDS
- SDN
100G OVS-DPDK – VXLAN & CONNECTION TRACKING

Faster Performance | Lower CAPEX

Throughput
114B, 64 connections
~60X

Throughput
114B, 250K connections
~20X

Server Cores Utilization

BlueField-2 P-series DPU, 100GbE Single Port Card | AMD EPYC 7742 64-Core Processor
DPU ACCELERATED STORAGE PROCESSING
Programable Data Path | Software Defined Orchestration

**Accelerated**

- ✔ NVMe / Virtio-Block Emulation
- ✔ Data Reduction, Erasure Coding
- ✔ Data at Rest Crypto & Integrity
- ✔ RDMA – RoCE / InfiniBand, TCP
- ✔ Data-in-Flight Encryption

**Software Defined**

- ✔ LVM / RAID control plane
- ✔ Key Association
- ✔ QoS Monitoring & Stats
- ✔ Namespace/ controller management

---

Bare Metal Host

Data Path

Control Path

BlueField DPU

Virtual PCIe Switch

NVMe

Net

IO Processors

RDMA/TCP Pipeline

ASAP2 Pipeline

InfiniBand/ Ethernet
STORAGE NVME-OF PERFORMANCE

Latency Determinism | IOPs Efficiency
DPU ENABLES CLOUD-NATIVE SUPERCOMPUTING
Multi-Tenancy with Zero-Trust Security

Collective offload with UCC accelerator
Smart MPI progression
User-defined algorithms
1.4X higher application performance
IMPROVING NON-BLOCKING MPI PERFORMANCE

44% Faster for MPI iAlltoall, 36% Faster for MPI iAllgather

Source: Courtesy of Ohio State University MVAPICH Team and X-SCALE Solutions
DPU ISOLATES GEFORCE NOW CLOUD GAMING

Isolated and Secured Infrastructure | More Concurrent Users

GeForce NOW Pod

Game Engine 40ms

Encode 5ms

10ms

User Input 10ms

Decode 5ms

Render 10ms

<100ms

Game Graphics Driver Video

SDDC Security Telemetry

NAT | DDOS | Reverse Proxy

BlueField-2 DPU

Driver Accelerator Streamer Accelerator

SDDC Security Telemetry

NAT | DDOS | Reverse Proxy

Ethernet NIC
DPU ENABLES FULLY INLINE 5G NETWORK PROCESSING
Offload, Isolate, Accelerate 5G Infra

Accelerate 5G or AI – Fully fungible – Fully programmable
Support for CUDA, DOCA – Toolchains, SDKs, Libraries
Secure, Isolated, Accelerated data processing
Domain specific acceleration for 5G, AI, Network Security
Fully optimized data path
SUMMARY

NVIDIA DPU Enables the Data Center as the New Unit of Computing

- The CPU can no longer do it all
- Must offload & isolate server infrastructure tasks to a DPU
- Effective DPU must offer hardware acceleration and security isolation
- To enable such effective DPU, need to develop broad software eco-system to utilize hardware acceleration across variety of disciplines (e.g. HPC, AI/ML, Storage, Networking, Security) - DOCA
- BlueField-3 includes new data path accelerator, supports 400 Gb/s with offloads
- NVIDIA DPU architecture & DOCA is a computing platform with rich stack optimized ideal for AI, bare metal cloud, cloud supercomputing, gaming, 5G wireless, and more