

NVIDIA DATA CENTER PROCESSING UNIT (DPU) ARCHITECTURE

Idan Burstein, DPU Principal Architect



DATA CENTER IS THE NEW UNIT OF COMPUTING





DPU-Accelerated Data Center Infrastructure



NAIVELY MOVING WORKLOADS TO NIC CPUS DOESN'T WORK

Traditional Server - 30 Total Cores



Shift CPU Workload to DPU Cores

Server with Non-Accelerated DPU Offload - 36 Total Cores



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



DPU Accelerators and 8 Arm Cores Replace 20 to 120 CPU Cores – HUGE Efficiency Gain

Software-defined Storage



Arm Cores Run Control Plane or Security Workloads Requiring Domain Isolation



NVIDIA DPU ROADMAP Exponential Growth in Data Center Infrastructure Processing





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

2024

5



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

2x370M PPS, 2x40M PPS at scale of millions of flows

18M IOP/s Elastic Block Storage

300 Equivalent x86 Cores

16C A78 ARM 42 SPECINT2k17-rate

128b DDR5-5600

DATA PATH ACC	CELERATOR
CONNECTX-7	
PCIe GEN 5.0	
	1 IT M



NVIDIA DOCA Enabling Broad BlueField Partner Ecosystem

Software Development Framework for BlueField DPUs

Software Compatibility for Generations of BlueField DPUs

Offload, Accelerate, and Isolate Infrastructure Processing

Support for Hyperscale, Enterprise, Supercomputing and Hyperconverged Infrastructure

DOCA is for DPUs what CUDA is for GPUs

PLATFORM INFRASTRUCTURE	l
CANONICAL	l
🔍 Red Hat	l
vm ware [®]	l
Orchestration	
Orchestration Management	



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	

ASAP ²	Programmable packet processor flow pipeline	
	Flow table based	
	Data Path	



DOCA

orm

Infrastructure Applications

Containers

DOCA Framework

Open APIs and Services

NVIDIA BlueField DPU BlueField Operating System



NVIDIA DPU SYSTEM ARCHITECTURE

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

PCle subsystem

Flexible EP/RP assignment, PCIe switching, NTB, p2p communication, emulated devices, optimized for IO

Data acceleration

Accelerating ARM workload





DPU ACCELERATED SWITCHING AND PACKET PROCESSING Programmable Data Path | Software-Defined Orchestration

Accelerated

- Virtio-Net/Other Emulation
- QoS & scheduling
- Telemetry and statistics \checkmark
 - Micro Segmentation / IPS / IDS / WAF
 - Encryption (Ipsec / MACsec)



 \checkmark

 \checkmark

Tunneling (VXLAN / GRE)





- Routing
- ACL

Software Defined eSwitch management **Connection Establishment Key Association**

Monitoring & Stats

 \checkmark

 \checkmark

 \checkmark

 \checkmark

BlueField



≥ NVIDIA.

100G OVS-DPDK - VXLAN & CONNECTION TRACKING Faster Performance | Lower CAPEX





BLUEFIELD-2 100G IPSEC TCP PERFORMANCE Faster Performance | Lower CAPEX



- BlueField-2 P-series 100GbE Single port
- Intel(R) Xeon(R) CPU E5-2687W v4 @ 3.00GHz
- Host OS: RHEL 8.3 (Ootpa)



- Each SW core does : 21.2 / 14.2 = ~1.5G
- To reach ~90G in SW we need ~60 cores
- To reach ~90G w/ DPU we need ~10 cores

- SW cores do : 21.2 / 14.2 = ~1.5G
- With DPU SW cores do : 88.9 / 10.3 = -9G-
- Core performance increased by magnitude 6X



DPU ACCELERATED STORAGE PROCESSING

Programable Data Path | Software Defined Orchestration



≥ NVIDIA.

STORAGE NVME-OF PERFORMANCE Latency Determinism | IOPs Efficiency



RoCE Maintains the latency determinism of local SSD access

RoCE IO Processing @ 4KB is x4-5 Lower

壑 NVIDIA.



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









HIGHER APPLICATION PERFORMANCE

Up to 35% App Performance, MPI Collectives Offload



32 servers, Dual Socket Intel® Xeon® 16-core CPUs E5-2697A V4 @ 2.60 GHz (32 processes per node), NVIDIA BlueField-2 HDR100 DPUs and ConnectX-6 HDR100 adapters, NVIDIA HDR Quantum Switch QM7800 40-Port 200Gb/s HDR InfiniBand, 256GB DDR4 2400MHz RDIMMs memory and 1TB 7.2K RPM SATA 2.5" hard drive per node.

Courtesy of Ohio State University MVAPICH team and X-ScaleSolutions



DPU ISOLATES GEFORCE NOW CLOUD GAMING Isolated and Secured Infrastructure | More Concurrent Users

GeForce NOW Pod



Game	Graphics Driver	Video		
SDDC	Security	Telemetry		
NAT DDOS Reverse Proxy				
Ethernet NIC				



📀 NVIDIA.

Traditional AI-on-5G Platforms

Heterogeneous, Mixed Programming Models



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

No need to move data back and forth from accelerators to host memory

Domain specific acceleration for 5G, AI, Network Security

Independent DPU time domain (5T for 5G)

Fully optimized data path

DPU-GPU Accelerated 5G

Homogeneous, Common Programming





.....

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
- NVIDIA DPU & DOCA is a computing platform with rich stack optimized ideal for AI, bare metal cloud, cloud supercomputing, storage, gaming, 5G wireless, and more
- NVIDIA is committed to line rate performance every generation.

