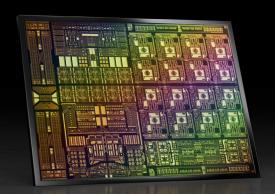# NVIDIA BLUEFIELD-3 DPU

Programmable Data Center Infrustructure on-a-Chip

The NVIDIA® BlueField®-3 data processing unit (DPU) is the 3rd-generation data center infrastructure-on-a-chip that enables organizations to build software-defined, hardware-accelerated IT infrastructures from cloud to core data center to edge. With 400Gb/s Ethernet or NDR 400Gb/s InfiniBand network connectivity, BlueField-3 DPU offloads, accelerates, and isolates software-defined networking, storage, security, and management functions in ways that profoundly improve data center performance, efficiency, and security.

Providing powerful computing, and a broad range of programmable acceleration engines in the I/O path, BlueField-3 is perfectly positioned to address the infrastructure needs of the most demanding applications, while delivering full software backward compatibility through the NVIDIA DOCA™ software framework.

BlueField-3 DPUs transform traditional computing environments into secure and accelerated virtual private clouds, allowing organizations to run application workloads in secure, multi-tenant environments. Decoupling data center infrastructure from business applications, BlueField-3 enhances data center security, streamlines operations and reduces total cost of ownership. Featuring NVIDIA's in-network computing technology, BlueField-3 enables the next generation of supercomputing platforms, delivering optimal bare-metal performance and native support for multi-node tenant isolation.

PORTFOLIO

> 1, 2, 4 ports with up to 400Gb/s connectivity

> 16GB on-board DDR5 memory

> Form factors: HHHL, FHHL

> M.2 / U.2 connectors options for direct attached storage

> 1GbE out-of-band management port

BlueField-3 DPU – 2x 200Gb/s FHHL form factor

Key Software-Defined, Hardware-Accelerated Applications

Cloud Networking
Cloud overlay, SDN acceleration, NAT, load balancer, NFV, video streaming

Storage
NVMe™ over Fabrics (NVMe-oF™), NVMe/TCP™, elastic storage, hyper converged infrastructure (HCI), encryption, data integrity, data deduplication, de-compression, erasure coding/RAID

Security
Distributed next-generation firewall, IDS/IPS, root of trust, micro-segmentation, DDOS prevention

HPC / AI
Cloud-native supercomputing, multi-tenancy and security, communication accelerations

Telco and Edge
Cloud RAN, virtualized edge gateways, VNF acceleration, edge microservers

Features

Network and Host Interfaces

Network Interfaces

> Ethernet - 1, 2, 4 ports with up to 400 Gb/s connectivity
> InfiniBand - Single port of NDR (400Gb/s), or dual ports of NDR200 / HDR (200Gb/s)

PCI Express Interface

> 32 lanes of PCIe Gen 5.0
> PCIe switch bi-furcation of up to 16 downstream ports
> Non-transparent bridging (NTB) support

Compute and Memory

Arm CPU Cores

> Up to 16 Armv8.2+ A78 Hercules cores (64-bit)
> 8MB L2 cache
> 16MB LLC system cache

Programmable Datapath Accelerator

> 16 cores, 256 threads
> Programmability through DOCA
> Heavy multi-threading applications acceleration

DDR DIMM Support

> Dual DDR5 5600MT/s DRAM controllers
> 16GB on-board DDR5
> ECC error protection support

Hardware Accelerations

Security

> Secure boot with Public key accelerator (PKA) root-of-trust
  > Secure firmware update
  > Flash encryption
  > Cerberus compliant

> Functional isolation layer
> Regular expression (RegEx) matching processor
> MACsec/IPsec/TLS data-in-motion encryption
  > AES-GCM 128/256-bit key
> AES-XTS 256/512-bit data-at-rest encryption
> Connection tracking for stateful firewall
> Public key accelerator (PKA)
  > RSA, Diffie-Hellman, DSA, ECC, EC-DSA, EC-DH
> True random number generator (TRNG)

Storage

> BlueField SNAP - Elastic block storage - NVMe™ and VirtIO-blk
> NVMe-oF™ and NVMe/TCP™ acceleration
> Decompression engine
> Erasure coding for RAID implementation
> M.2 / U.2 connectors for direct attached storage

Networking

> RoCE, Zero Touch RoCE
> ASAP² - Accelerated Switch and Packet Processing® for SDN and VNF acceleration
> Single Root I/O Virtualization (SR-IOV)
> VirtIO acceleration
> Overlay network acceleration
  > VXLAN, GENEVE, NVGRE
> Programmable flexible parser: user defined classification
> Connection tracking (L4 firewall)
> Flow mirroring, sampling and statistics
> Header rewrite
> Hierarchical QoS
> Stateless TCP offloads

HPC/AI Accelerations

> HPC / AI All-to-All engine
> NVIDIA GPUDirect
> NVIDIA GPUDirect Storage (GDS)
> HPC MPI Tag Matching

Advanced Timing and Synchronization

> IEEE 1588v2 (any profile)
> G.8273.2 Class C
> PTP hardware clock (PHC)
> Line rate hardware timestamp
> SyncE
> G.8262.1 (eEEC)
> Configurable PPS In and PPS Out
> Time triggered scheduling
> Time-based SDN acceleration

Boot Options

> Secure boot (RSA authenticated)
> Remote boot over Ethernet
> Remote boot over iSCSI
> PXE and UEFI

Management

> 1GbE out-of-band management port
> NC-SI, MCTP over SMBus, and MCTP over PCIe
> PLDM for Monitor and Control DSP0248
> PLDM for Firmware Update DSP026
> I²C interface for device control and configuration
> SPI interface to flash
> eMMC memory controller
> UART
> USB

Connection®
we solve IT®

Contact an Account Manager for more information.

Business Solutions
1.800.800.0014

Enterprise Solutions
1.800.369.1047

Public Sector Solutions
1.800.800.0019

www.connection.com/nvidia

C1619321 0122

nVIDIA®