BlueField Platform Product Update

Q1 2024

Agenda

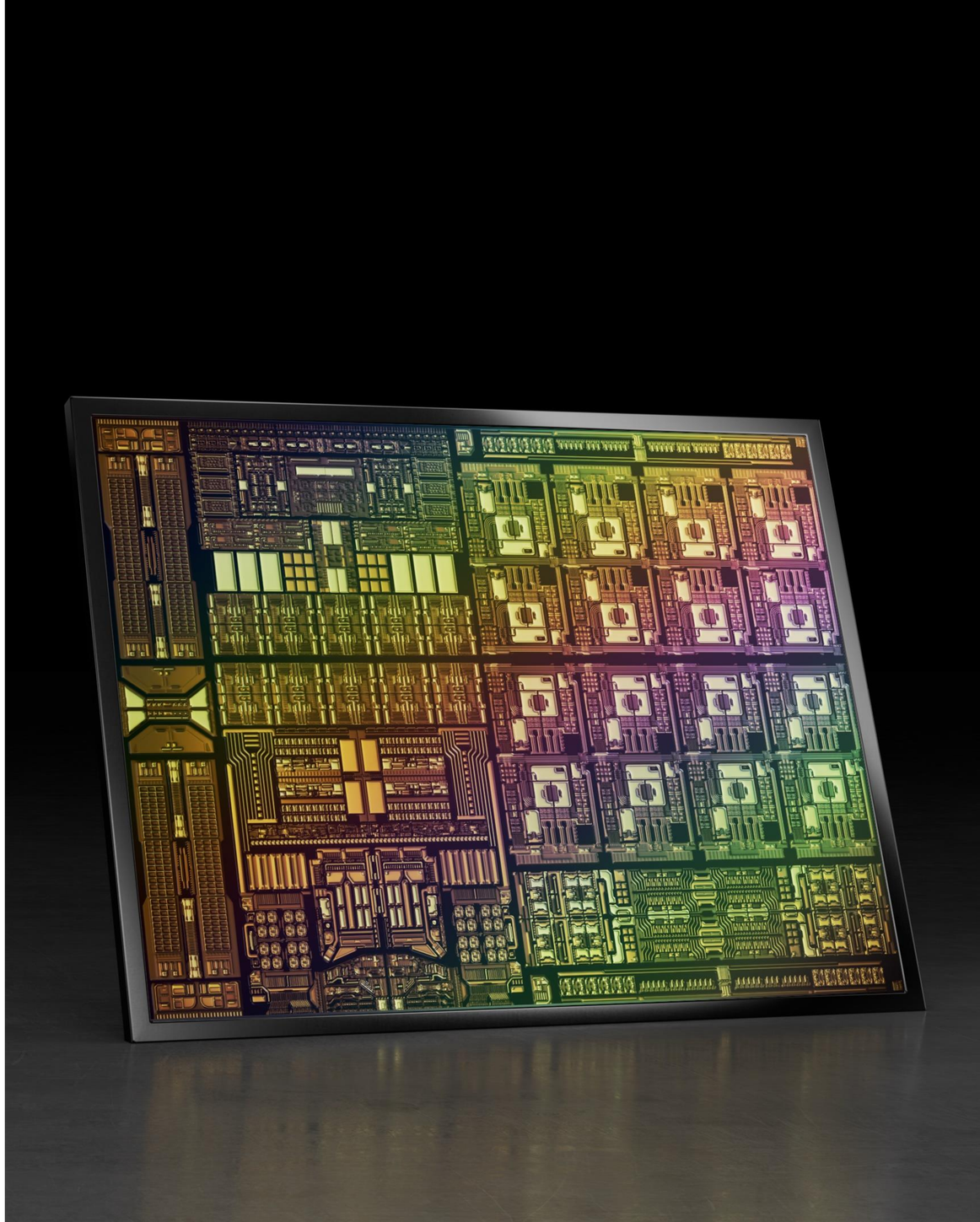
- Nvidia BlueField SoC
- Nvidia BlueField in AI Cloud Data Center
- Nvidia BlueField in Traditional Cloud Data Center
- Nvidia BlueField-3 Product Portfolio
- Nvidia BlueField-3 Management Solution
- Nvidia BlueField-2 Product Portfolio
- Nvidia Software Stack for BlueField Product

Nvidia BlueField SoC

NVIDIA BlueField-3 DPU

400Gbps Processor for data center networking, storage, and security

- Powerful CPU – 16x Arm A78 Cores
- Datapath Accelerator – 16x Cores, 256 Threads
- Offloads and Accelerates Data Center Infrastructure
- Isolates Application from Control and Management Plane



Industry-Leading Performance

Massive Advancements, Built for Cloud Scale

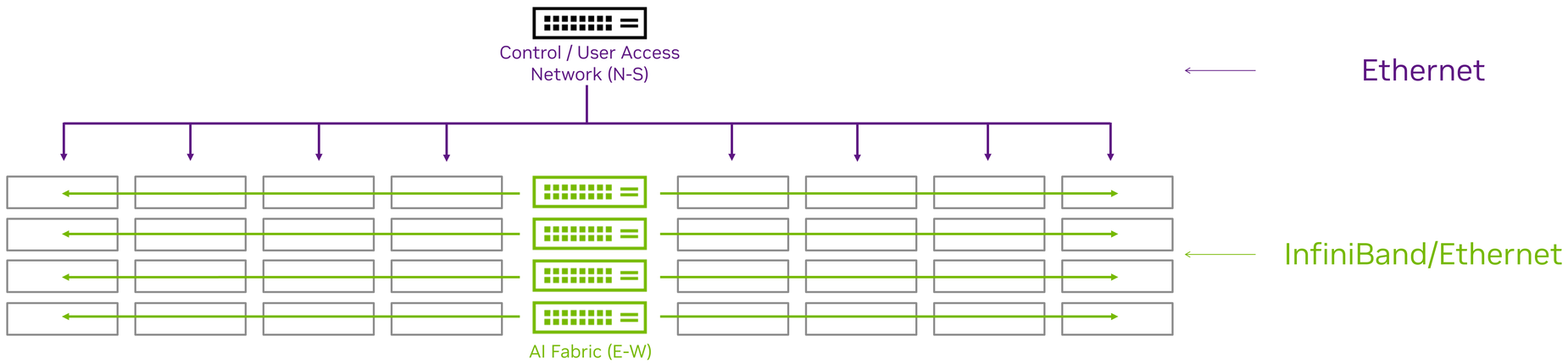
	BlueField-2	BlueField-3
Bandwidth	200Gb/s	400Gb/s
DPDK Max msg Rate	215Mpps	250Mpps
RDMA max msg rate	215Mpps	330Mpps
Compute	SPECINT: 70	SPECINT: 350
VirtIO Acceleration	40Mpps (*)	80Mpps (*)
IPsec Acceleration	100Gb/s	400Gb/s
TLS Acceleration	200Gb/s	400Gb/s
MACsec Acceleration	X	400Gb/s bi-dir
NVMe SNAP	5.4M IOPs @4K	12M-18MIOPS @4K

* Total packet rate for the sum of Tx and Rx

Nvidia BlueField in AI Cloud Data Center

Peak AI Performance Demands Optimized Networking

AI Workloads Require an AI Fabric

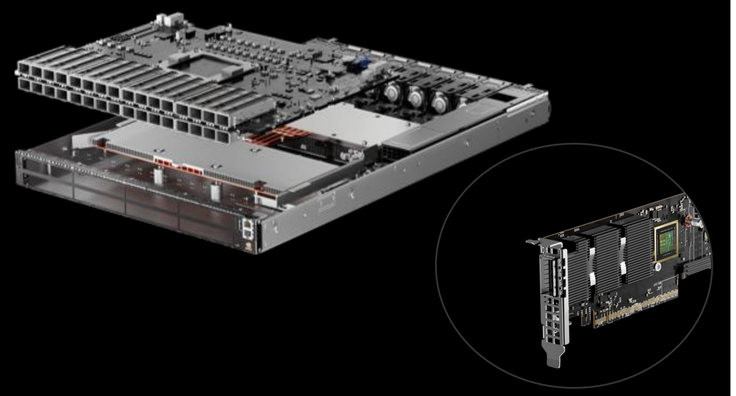


Control / User Access Network (N-S)		AI Fabric (E-W)
Loosely-Coupled Applications		Tightly-Coupled Processes
TCP (Low Bandwidth Flows and Utilization)		RDMA (High Bandwidth Flows and Utilization)
High Jitter Tolerance		Low Jitter Tolerance
Oversubscribed Topologies		Nonblocking Topologies
Heterogeneous Traffic, Statistical Multi-Pathing		Bursty Network Traffic, Predictive Performance

NVIDIA Networking Platforms

Accelerated Networking Solutions for the Era of AI

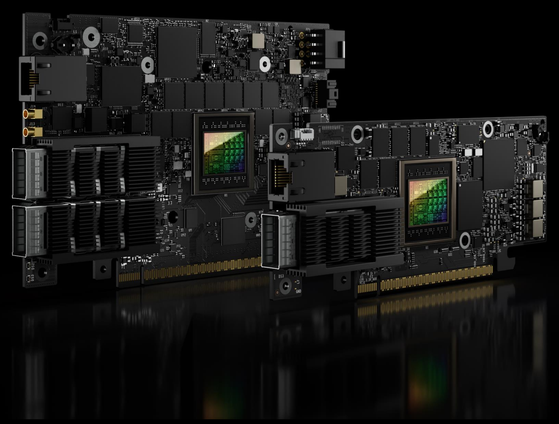
Quantum-2 InfiniBand



Supercomputing Networking Platform

AI Factories and Cloud-Native Supercomputing

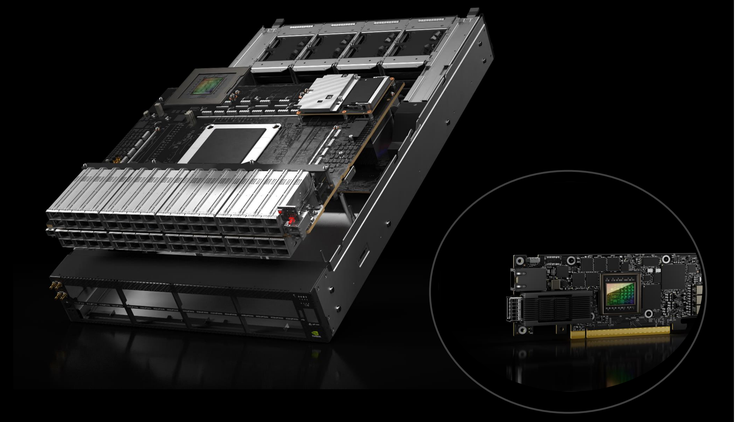
BlueField-3 Networking



Infrastructure Compute Platform

Offload, Accelerate, and Isolate Data Center Infrastructure

Spectrum-X Ethernet



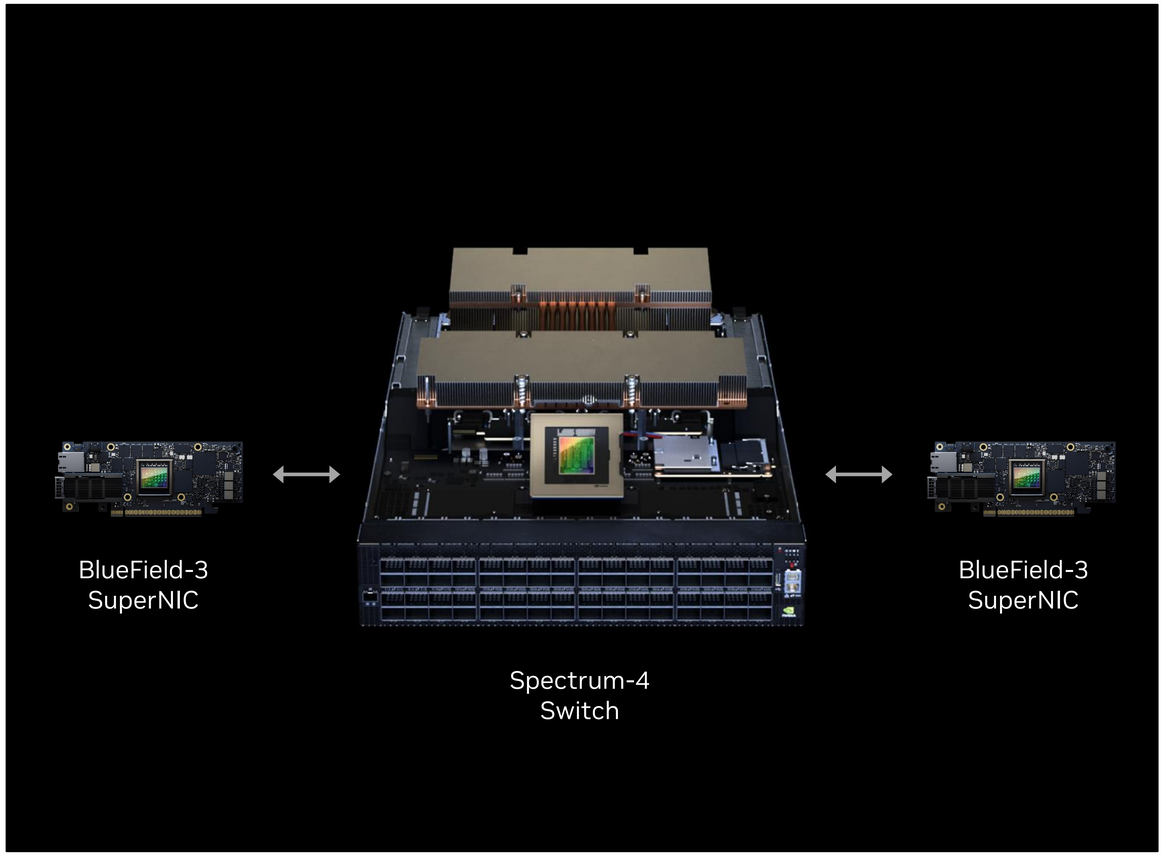
Hyperscale Networking Platform

Purpose-built Ethernet Networking for AI Clouds

NVIDIA Spectrum-X Networking Platform

Ethernet AI Performance Leadership

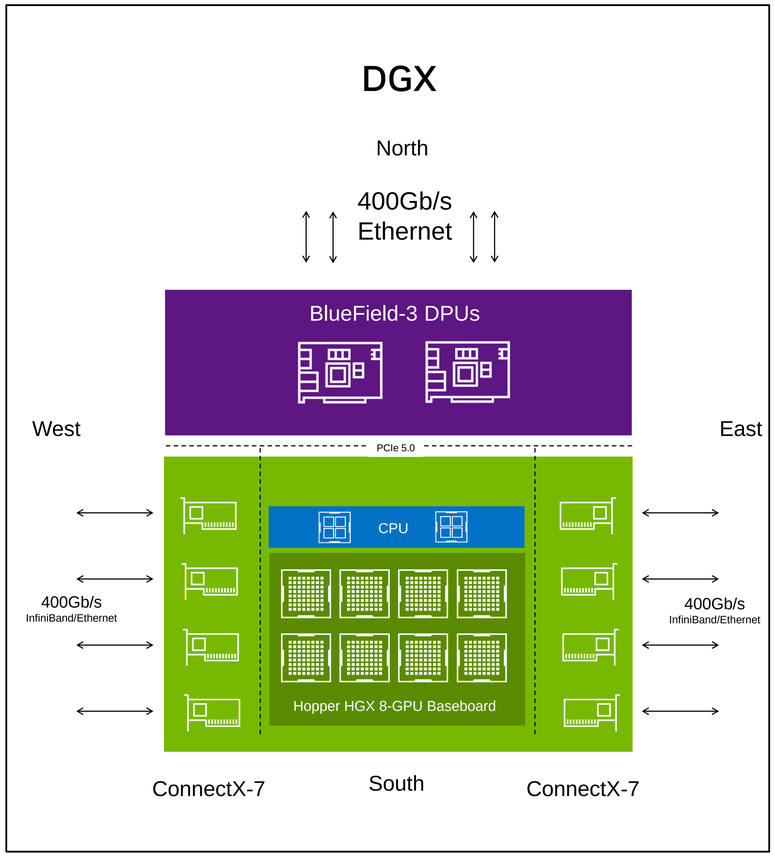
- World’s first Ethernet platform purpose-built for AI
- Specialized high-performance architecture
- Standard Ethernet connectivity
- Spectrum-X ROI is much higher than the network cost
- Future proof acceleration with NVIDIA workflow



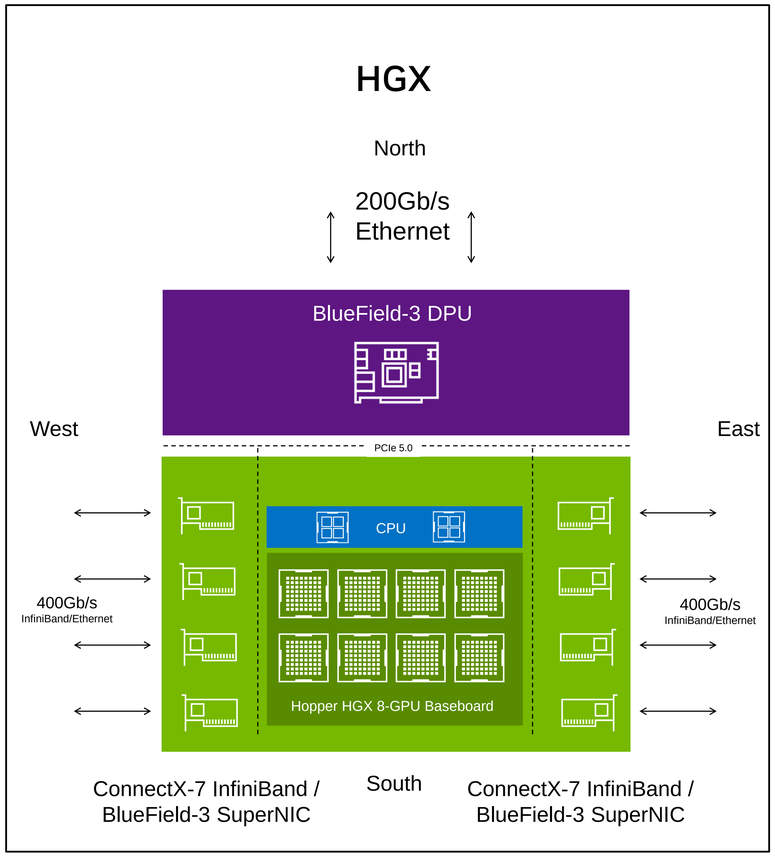
NVIDIA Spectrum-X Networking Platform
Purpose-built Ethernet Fabric for AI Clouds

NVIDIA Accelerated Systems for Generative AI

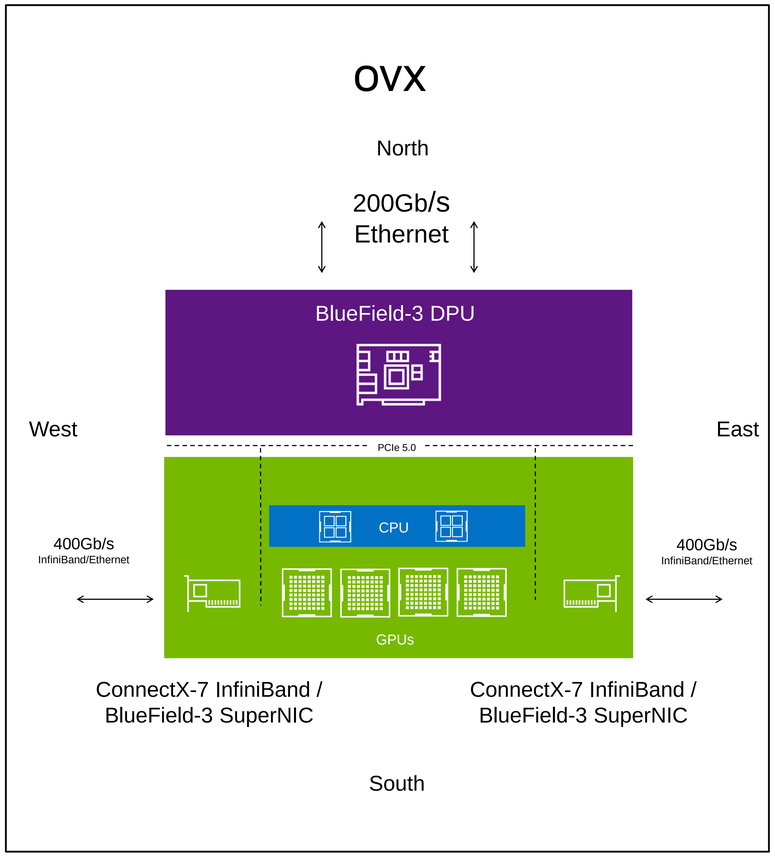
BlueField Powers GPU-Accelerated Systems for the Best of NVIDIA AI and NVIDIA Omniverse



AI services: 400Gb/s InfiniBand/Ethernet (East-West)
Tenant networking: 400Gb/s Ethernet (North-South)



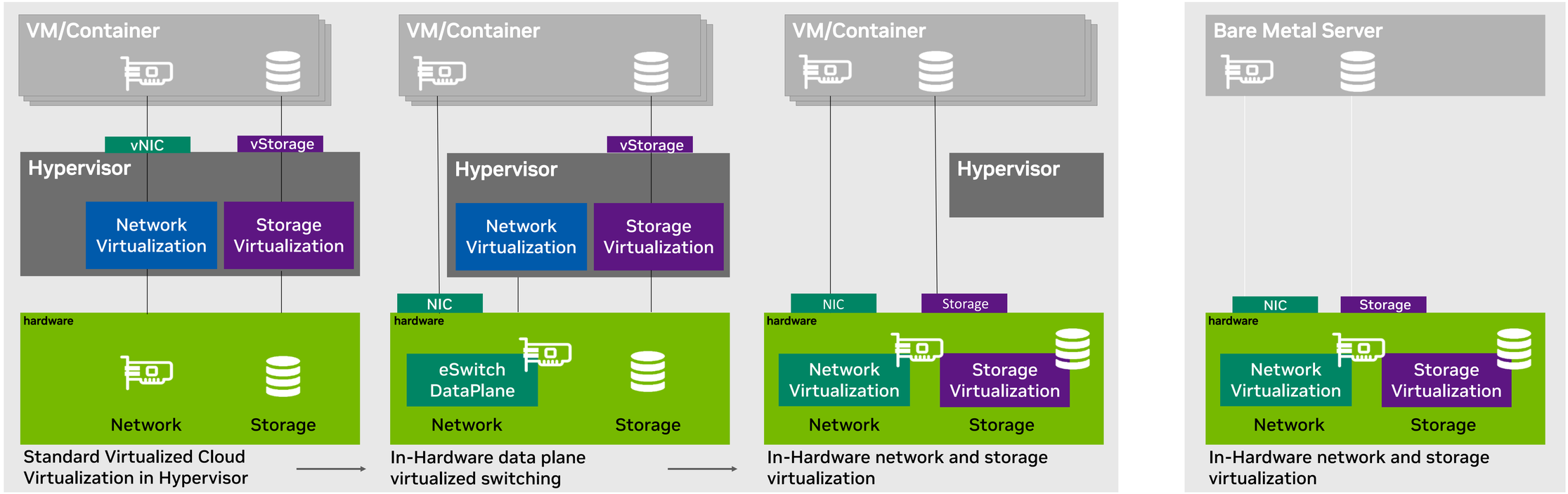
AI services: 400Gb/s InfiniBand/Ethernet (East-West)
Tenant networking: 200Gb/s Ethernet (North-South)



AI services: 200Gb/s InfiniBand/Ethernet (East-West)
Tenant networking: 200Gb/s Ethernet (North-South)

Nvidia BlueField
in Traditional Cloud Data Center

Path to Accelerated Virtualization



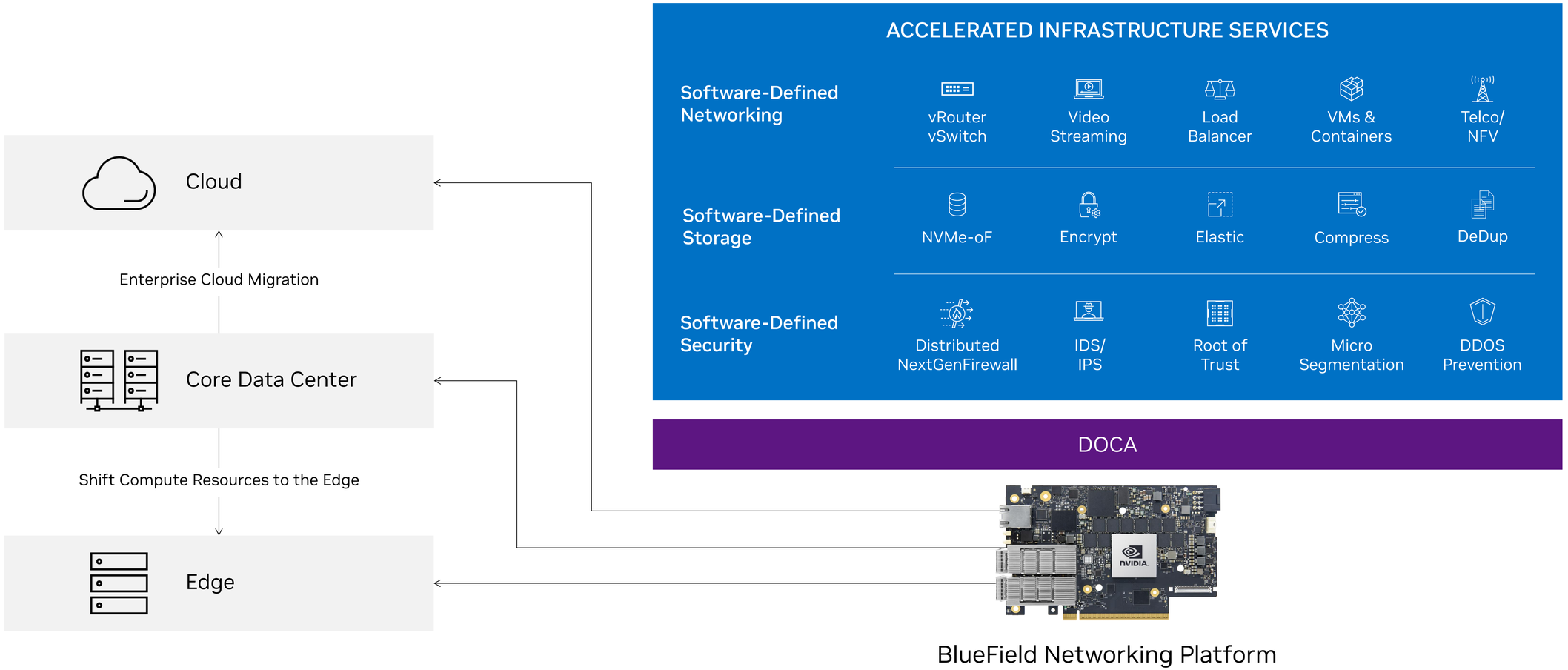
Freeing up CPU resources

Improve Security & Isolation

Virtualizing Everything

BlueField is a Cloud Services Compute Platform

NVIDIA BlueField Accelerates Data Center Infrastructure Services from Cloud to Edge



BlueField Networking Platform

Nvidia BlueField-3 Product Portfolio

NVIDIA BlueField-3

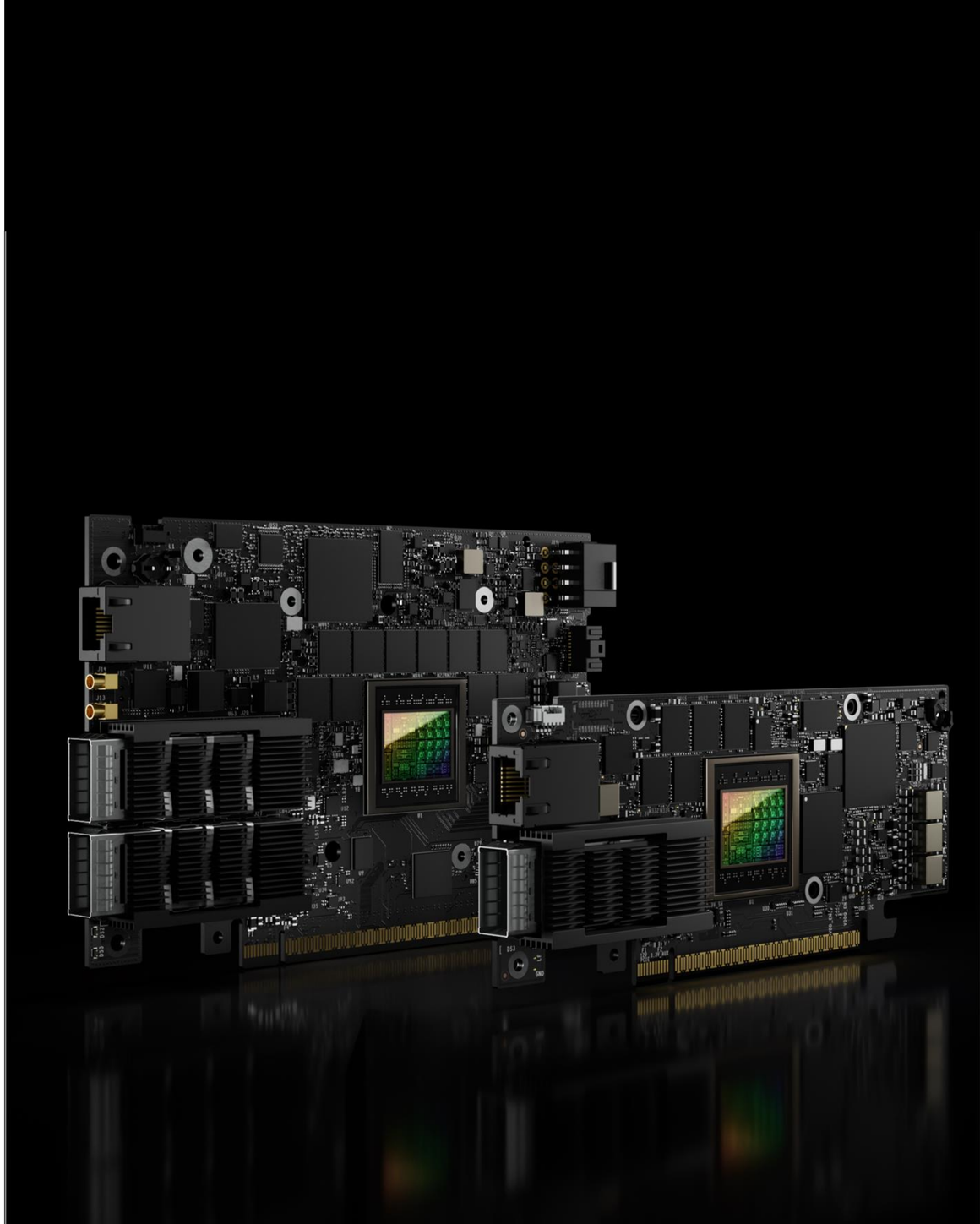
400Gb/s Infrastructure Compute Platform

400Gb Networking
RDMA/RoCE Accelerations
SDN/NFV Accelerations
Precision Timing

Programmable Engines
16 x 64-bit A78 Arm Cores
16 Hyperthreaded DPA Cores
Accelerated Pipeline

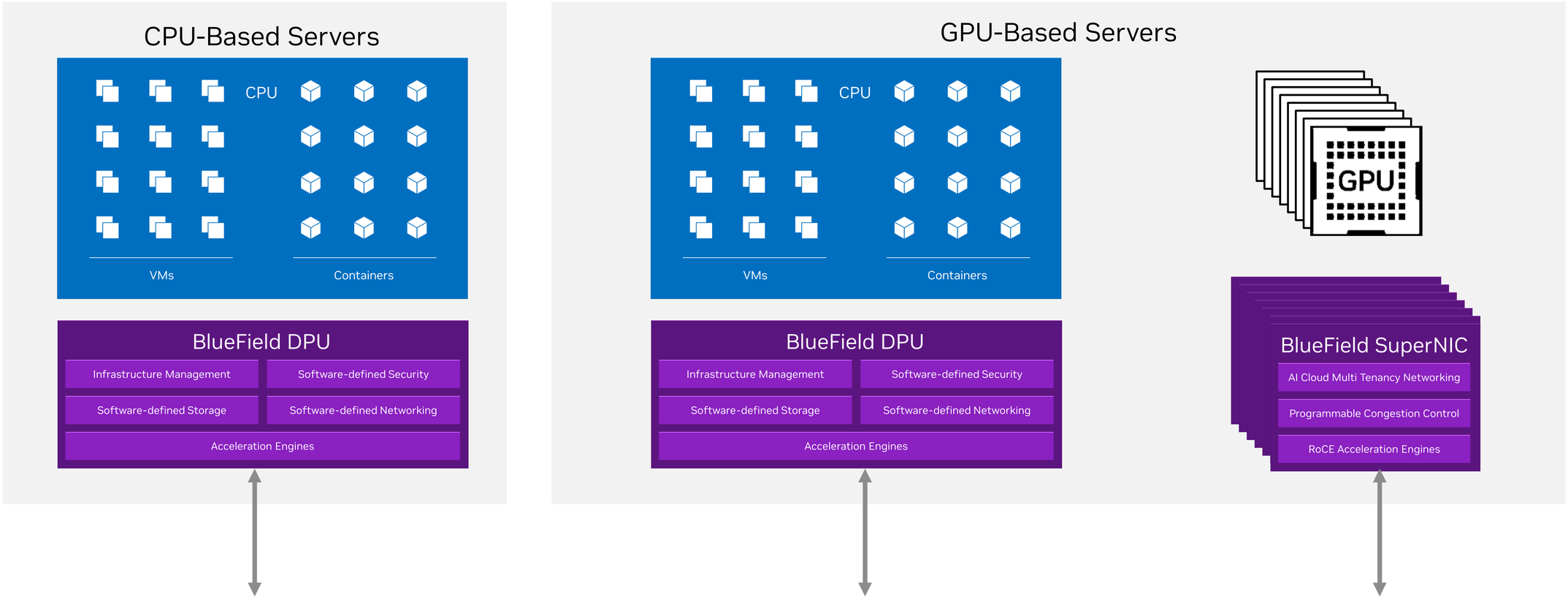
Zero-Trust Security
Platform Security
Crypto Accelerations
Zero-Trust Infrastructure

Composable Storage
Storage Disaggregation
NVMe-oF, NVMe/TCP
Storage Encryption



NVIDIA BlueField-3 – One Platform, Two Products

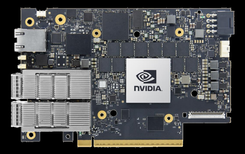
BlueField-3 DPUs and SuperNICs



BlueField-3 Portfolio

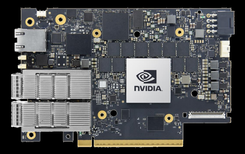
DPUs and SuperNICs

B3220



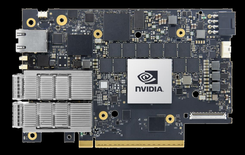
2P x QSFP112, 200G

B3210E



2P x QSFP112, 100G

B3240



2P x QSFP112, 400G

- Ethernet & InfiniBand
- 16 x Arm A78 @ 2.0/2.2Ghz
- 32GB DDR5
- >75w TDP
- FHHL

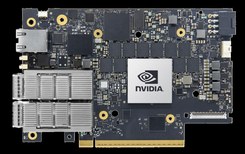
BlueField-3 DPUs

Advanced Programmability, Powerful Computing

B3140H

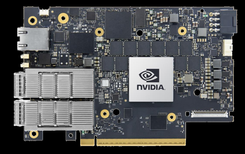
1P x QSFP112, 400G

B3220L



2P x QSFP112, 200G

B3210L



2P x QSFP112, 100G

- Ethernet & InfiniBand
- 8 x Arm A78 @ 2.0Ghz
- 16GB DDR5
- <75w TDP
- FHHL/HHHL

BlueField-3 SuperNIC

Lean, Power Efficient

NVIDIA BlueField-3 Networking Platform

Advanced Infrastructure Computing Platform for Powering Generative AI Clouds

	BlueField-3 DPU	BlueField-3 SuperNIC
Mission	> Cloud infrastructure processor > Offload, accelerate, and isolate data center infrastructure > Optimized for N-S in GPU-class systems	> Accelerated networking for AI computing > Best-in-class RoCE networking > Optimized for E-W in GPU-class systems
Shared Capabilities	> VPC network acceleration > Network encryption acceleration > Programmable network pipeline > Precision timing > Platform security	
Unique Capabilities	> Powerful computing > Secure, zero-trust management > Data storage acceleration > Elastic infrastructure provisioning > 1-2 DPUs per system	> Powerful networking > AI networking feature set > Full-stack NVIDIA AI optimization > Power-efficient, low-profile design > Up to 8 SuperNICs per system

B3220 DPU

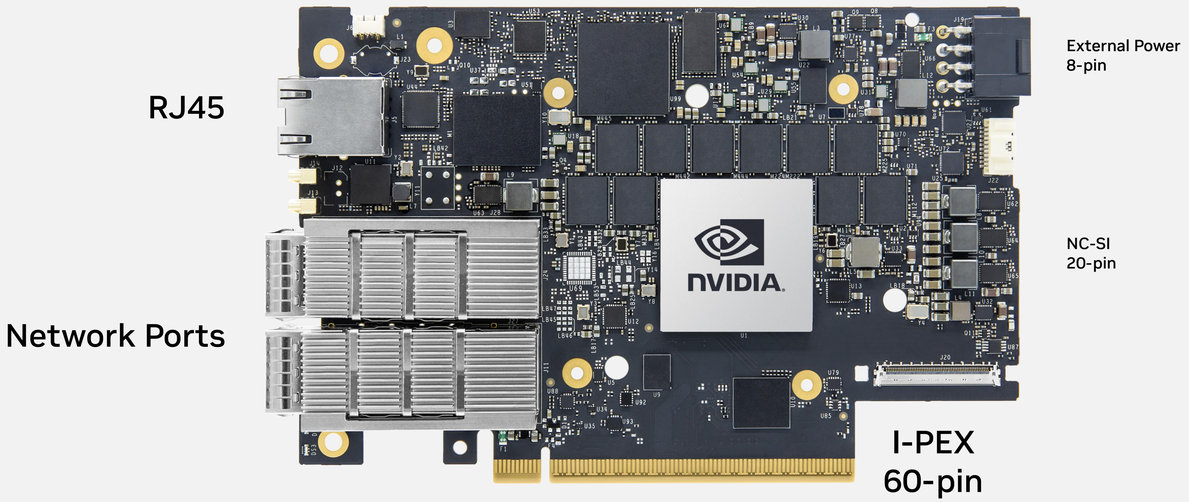
400Gb/s Infrastructure Compute Platform

Specifications

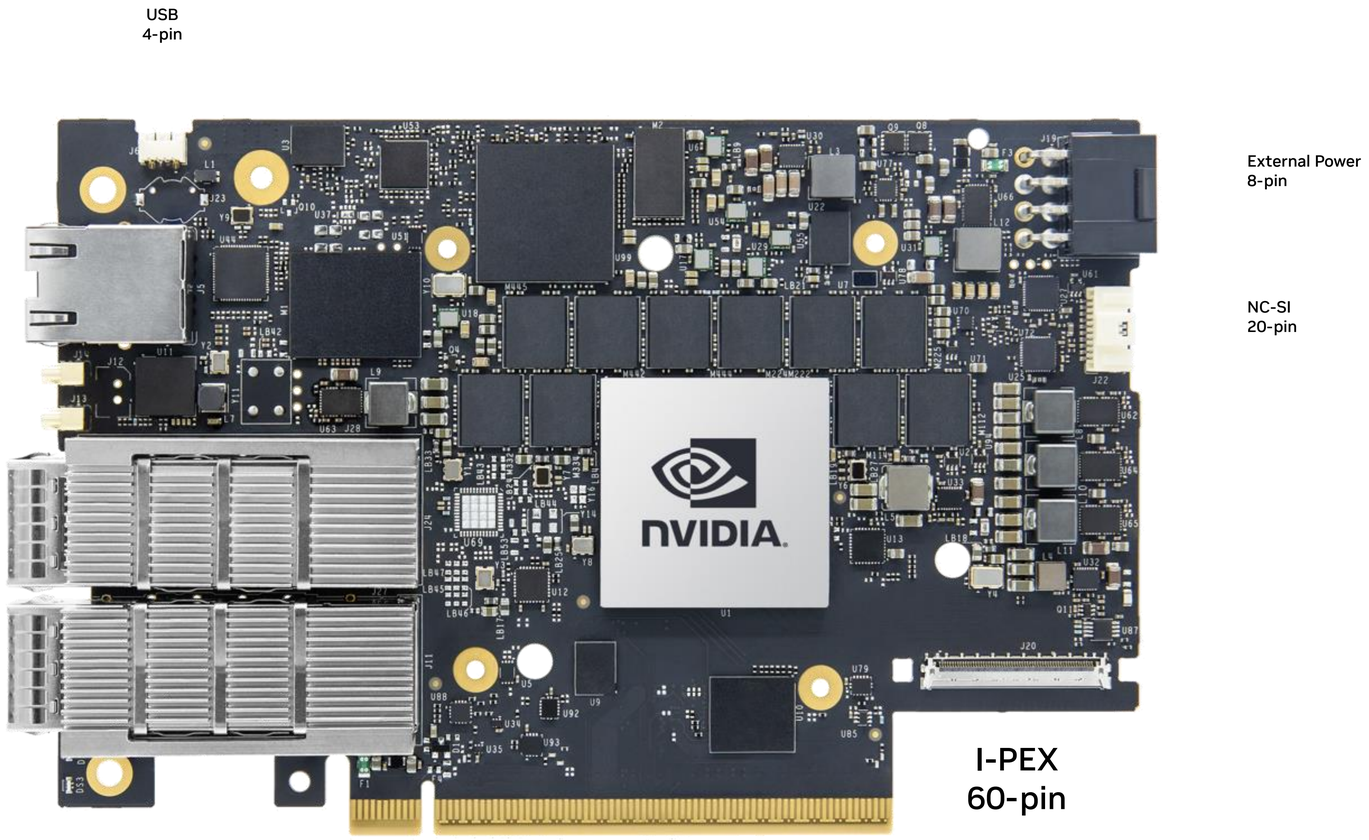
Total Bandwidth	400Gb/s
Network Speed	2 x 1/10/25/50/100/200
Interface Type	QSFP112
Programmable Compute	16 x Arm A78 @ 2.2Ghz
Memory	32GB DDR5
Host Interface	Gen5 x16 + x16
TDP Power w/ Active Cables	>75W
Thermal solution	Passive
Form factor	FHHL

Ordering Part Numbers

Crypto Disabled	900-9D3D4-00NN-HA0
Crypto Enabled	900-9D3D4-00EN-HA0



B32XX Interfaces



B3140H SuperNIC

Network Accelerator for AI Compute Fabric

Specifications

Network Speed	1 x 400Gb/s
Interface Type	QSFP112
Programmable Compute	8 x Arm A78 @ 2.0Ghz
Memory	16GB DDR5
Host Interface	Gen5 x16
TDP Power w/ Active Cables	<75W
Thermal solution	Passive
Form factor	HHHL

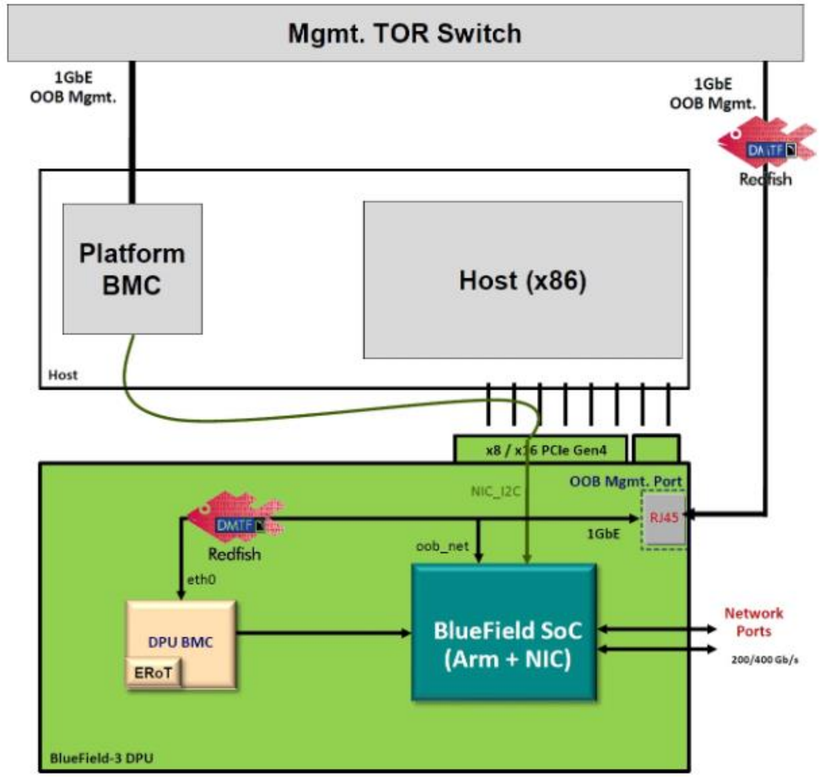
Ordering Part Numbers

Crypto Disabled	900-9D3D4-00NN-HA0
Crypto Enabled	900-9D3D4-00EN-HA0

Nvidia BlueField-3 Management Solution

BlueField-3 Management Solution

- 1GbE OOB connection is required to manage BlueField cards
- Multi-card deployment is simplified with the BlueField management solution.
- Manage all BlueField cards in a server through a single OOB management:
  - Minimize cabling overhead
  - Redfish based management of all DPUs and SuperNICs - consolidated schema
  - Arm telemetry & containers provisioning will be in-band over the high-speed port

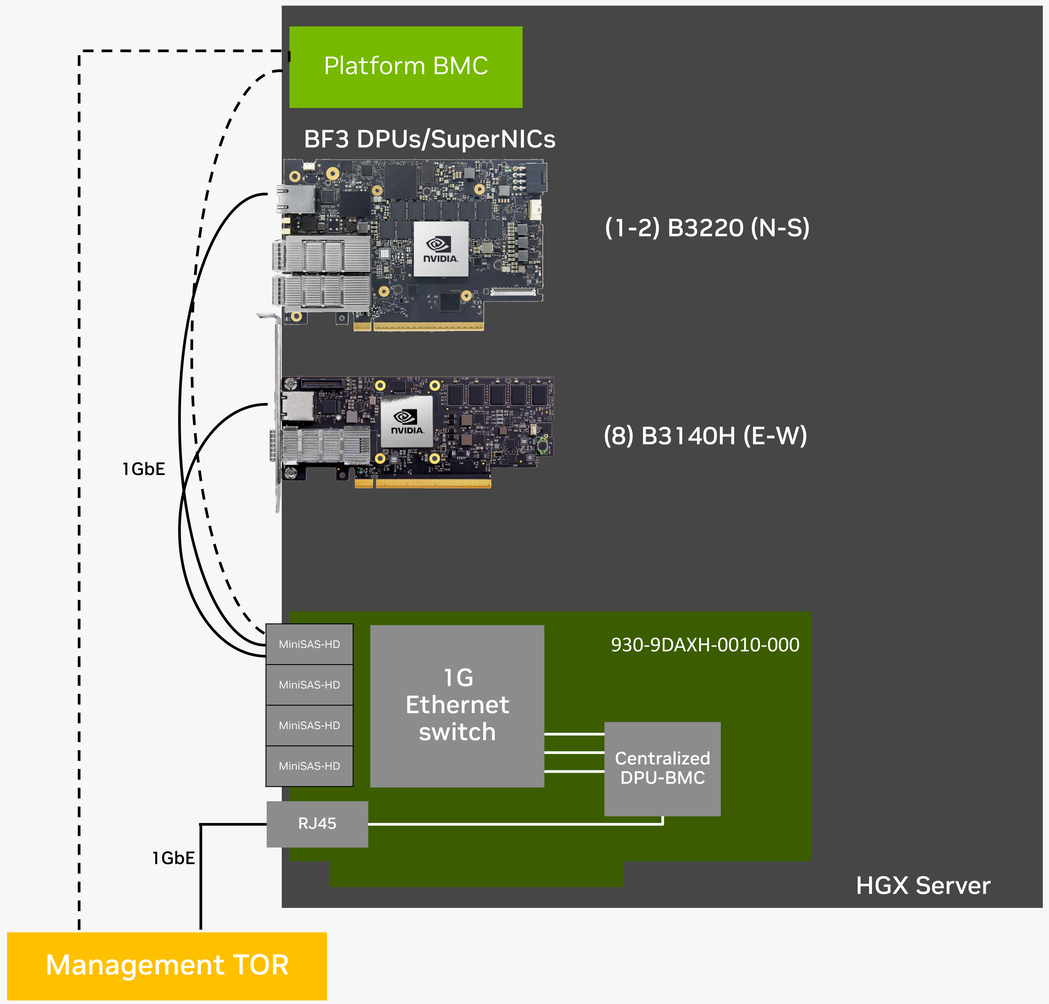


NVIDIA BlueField DPU Management and Initial Provisioning - NVIDIA Docs

BlueField-3 Management Controller

1GbE-Based Solution

- Ethernet-based solution
  - Auxiliary management card with integrated 1G Eth switch
  - PCIe HHHL form factor
  - Centralized BlueField BMC
  - All BlueField's OOB connect to multiple card
  - All BlueField's OOB connections are external to server
- BlueField management controller:
  - Enables up to 12 x 1GbE connections
  - Enables platform BMC tunneling (optional)
  - Use Redfish schema to manage all DPUs and SuperNICs
- Two options to connect platform BMC:
  - Directly to management ToR switch
  - Use tunneling option via BlueField management card



BlueField-3 Management Controller

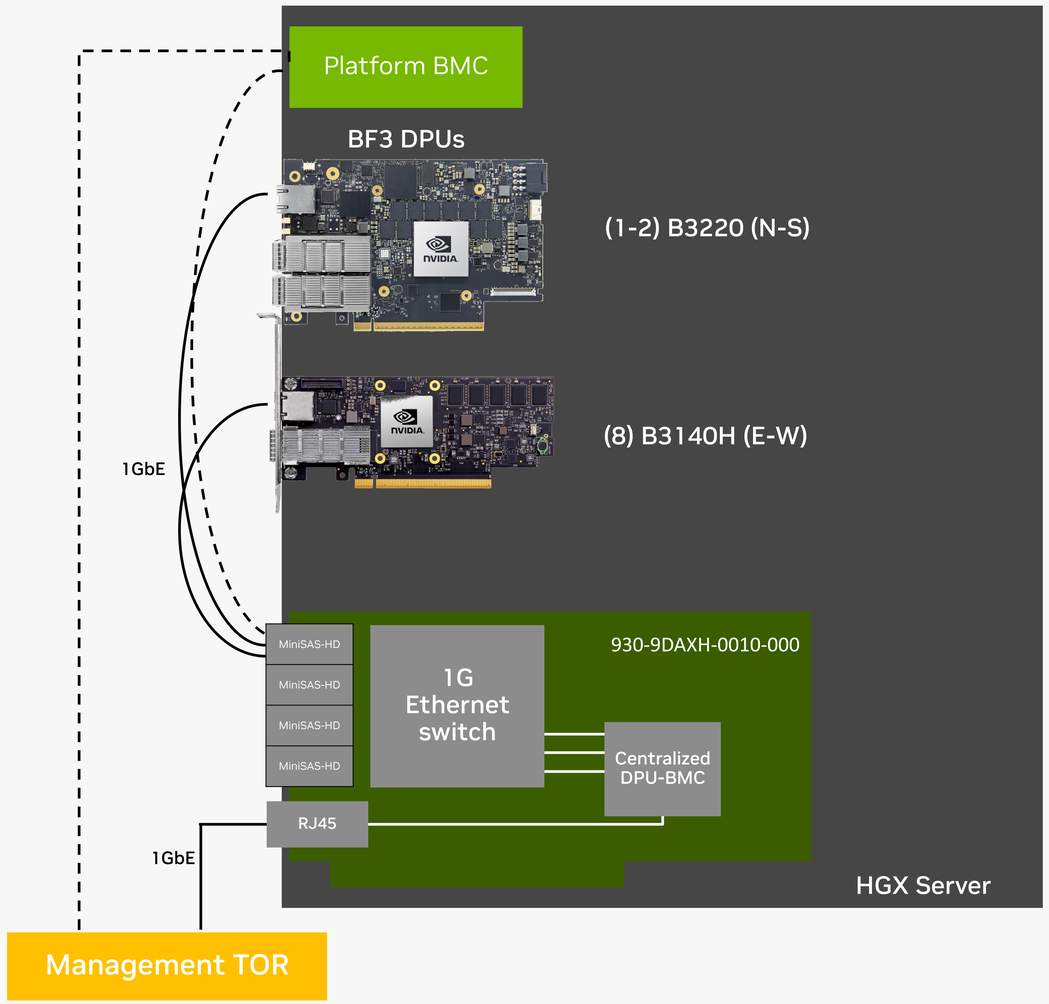
1GbE-Based Solution

Specifications

Network Speed	1GbE
Interface Type	12 x RJ45
Cables length	70 cm (fits 6U server connectivity)
Host Interface	Gen4 x8
Max Power	<25W
Thermal solution	Passive
Form factor	HHHL

Ordering Part Number

BlueField Management	930-9DAXH-0010-000



BlueField-3 Portfolio

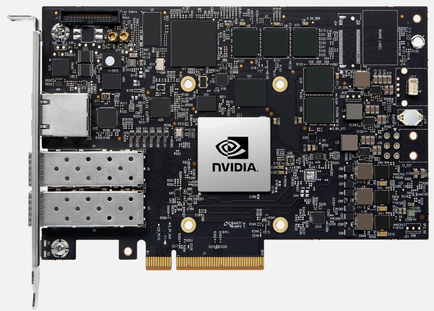
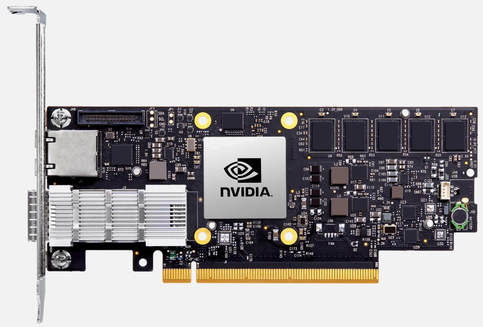
Ordering Information

Model	Network Speed	Target Usage	Arm Cores	Host Interfaces [PCIe]	Memory	Ordering Part Number (OPN)	
						No Crypto	With Crypto
B3210E	2 x 100Gb/s	DPU	16 x ArmA78 @2.0GHz	Gen5 x16 + x16	32GB DDR5	900-9D3B6-00SC-EA0	900-9D3B6-00CC-EA0
B3220	2 x 200Gb/s	DPU	16 x ArmA78 @2.2GHz	Gen5 x16 + x16	32GB DDR5	900-9D3B6-00SV-AA0	900-9D3B6-00CV-AA0
B3240	2 x 400Gb/s	DPU	16 x ArmA78 @2.2GHz	Gen5 x16 + x16	32GB DDR5	900-9D3B6-00SN-AB0	900-9D3B6-00CN-AB0
B3140H	1 x 400Gb/s	SuperNIC	8 x ArmA78 @2.0GHz	Gen5 x16	16GB DDR5	900-9D3D4-00NN-HA0	900-9D3D4-00EN-HA0
B3220L	2 x 200Gb/s	SuperNIC	8 x ArmA78 @2.0GHz	Gen5 x16	16GB DDR5	900-9D3B4-00SV-EA0	900-9D3B4-00CV-EA0
B3140L	1 x 400Gb/s	SuperNIC (IL-1)	8 x ArmA78 @2.0GHz	Gen5 x16	16GB DDR5	900-9D3B4-00PN-EA0	900-9D3B4-00EN-EA0
B3210L	2 x 100Gb/s	SuperNIC	8 x ArmA78 @2.0GHz	Gen4 x16	16GB DDR5	900-9D3B4-00SC-EA0	900-9D3B4-00CC-EA0
B3220SH	2 x 200Gb/s	Storage Controller	16 x ArmA78 @2.0GHz	Gen5 x16 + x16	48GB DDR5	900-9D3C6-00SV-DA0	900-9D3C6-00CV-DA0
MB312	-	Multi-BF Management	-	Gen4 x 8	-	930-9DAXH-0010-000	-

Nvidia BlueField-2 Product Portfolio

BlueField-2 Portfolio

Comprehensive Portfolio to Support Entry Level & Premium Boards

2P X 25G - 2.5GHZ - HHHL ENTERPRISE	2P X 25G - 2.5GHZ - FHHL CSP/CRISP/MONTEREY	2P X 100G - 2.75GHZ - FHHL CSPS/CRISP/MONTEREY	1P X 200G - 2.0GHZ - HHHL HPC/AI
			

VMware Project Monterey :

MBF2H532C-AECOT, - NVIDIA BlueField-2 P-Series DPU 25GbE Dual-Port SFP56, integrated BMC, PCIe Gen4 x8, Secure Boot Enabled, Crypto Enabled, 32GB on-board DDR

MBF2H536C-CECOT, NVIDIA BlueField-2 P-Series DPU 100GbE Dual-Port QSFP56, integrated BMC, PCIe Gen4 x16, Secure Boot Enabled, Crypto Enabled, 32GB on-board DDR

Nvidia Software Stack for BlueField Product

DOCA Software Packages

Introducing naming and format

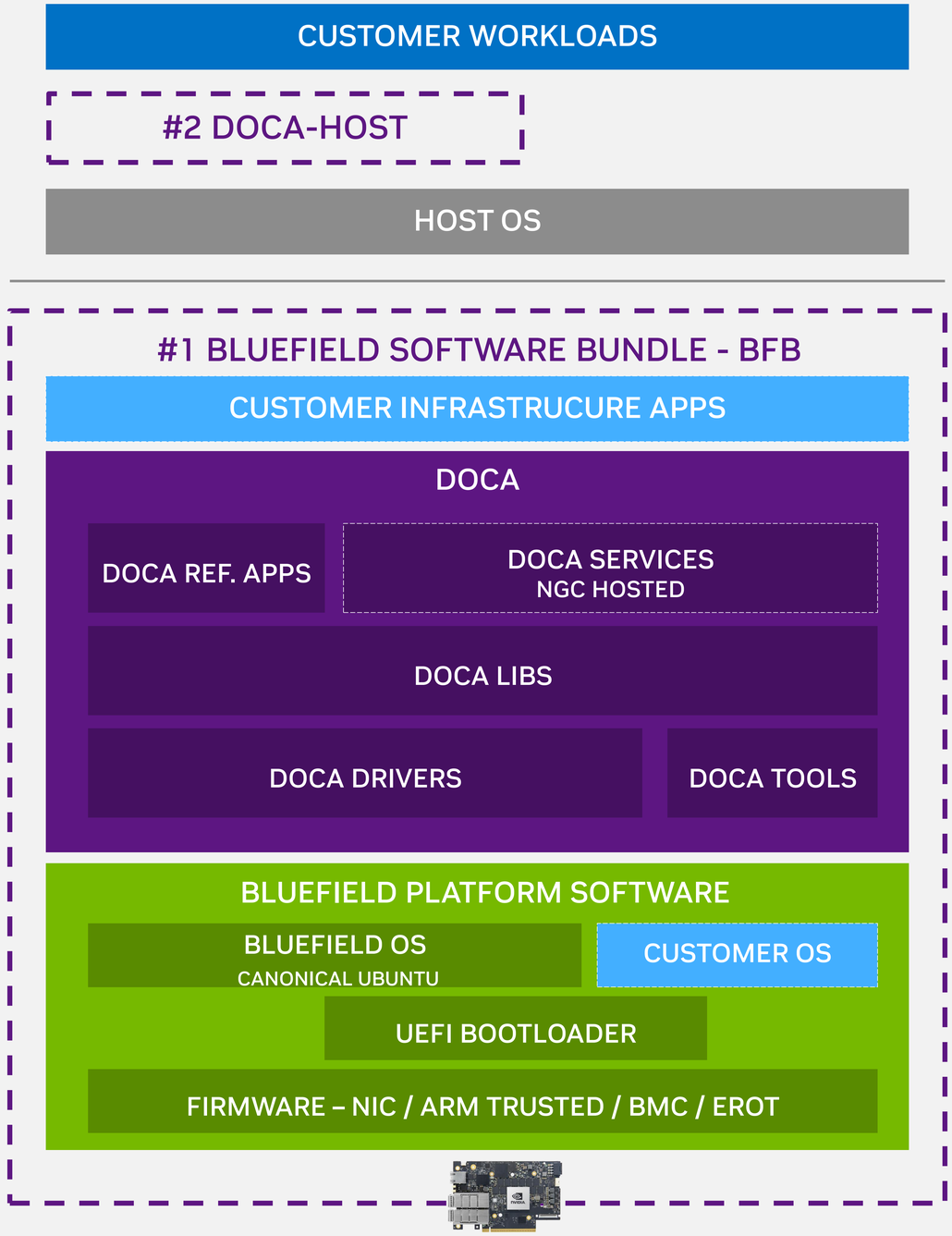
Package #1 BlueField Software Bundle

- Installable Image .bfb/.iso with upgrade tool
- BlueField DPU and SuperNIC
- Includes
  - DOCA runtime components for BlueField Arm cores
  - BlueField OS and Platform Firmware
  - DOCA SDK package can be added

Package #2 DOCA-HOST

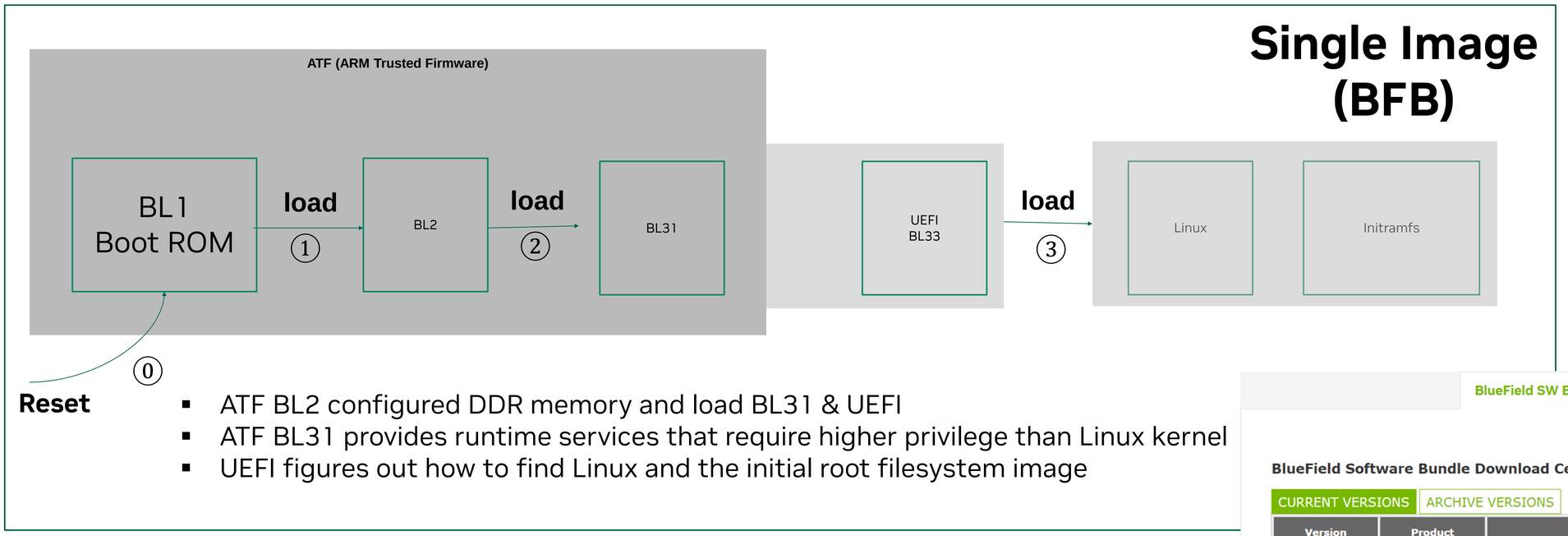
- Standard Linux package, installable on server itself
- BlueField & ConnectX
- Includes
  - DOCA Runtime Drivers & Libs & Reference Apps
  - DOCA SDK package can be added
  - DOCA Install profiles per usage type

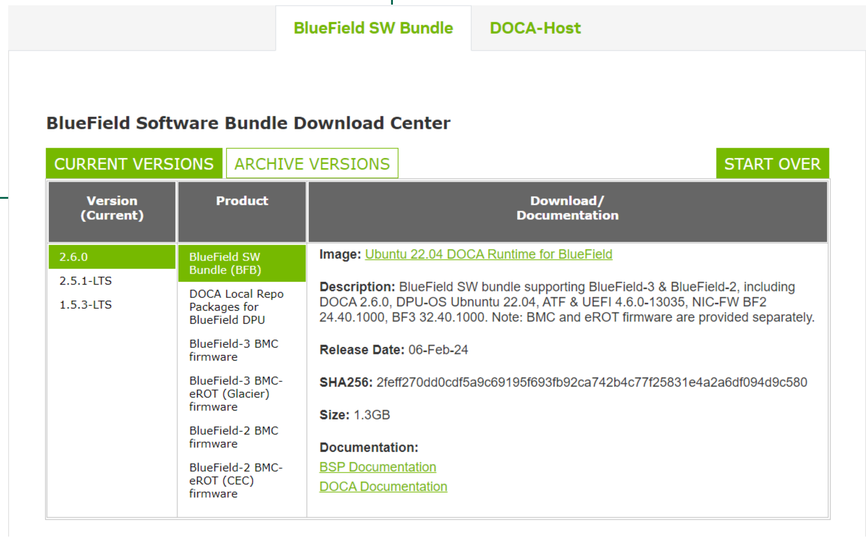
Delivered in this release | Downloaded From NGC | (Optional) Added By Customer



BlueField BFB

- Typical USB or PCIe boot scenario. All binaries are packed into a single image





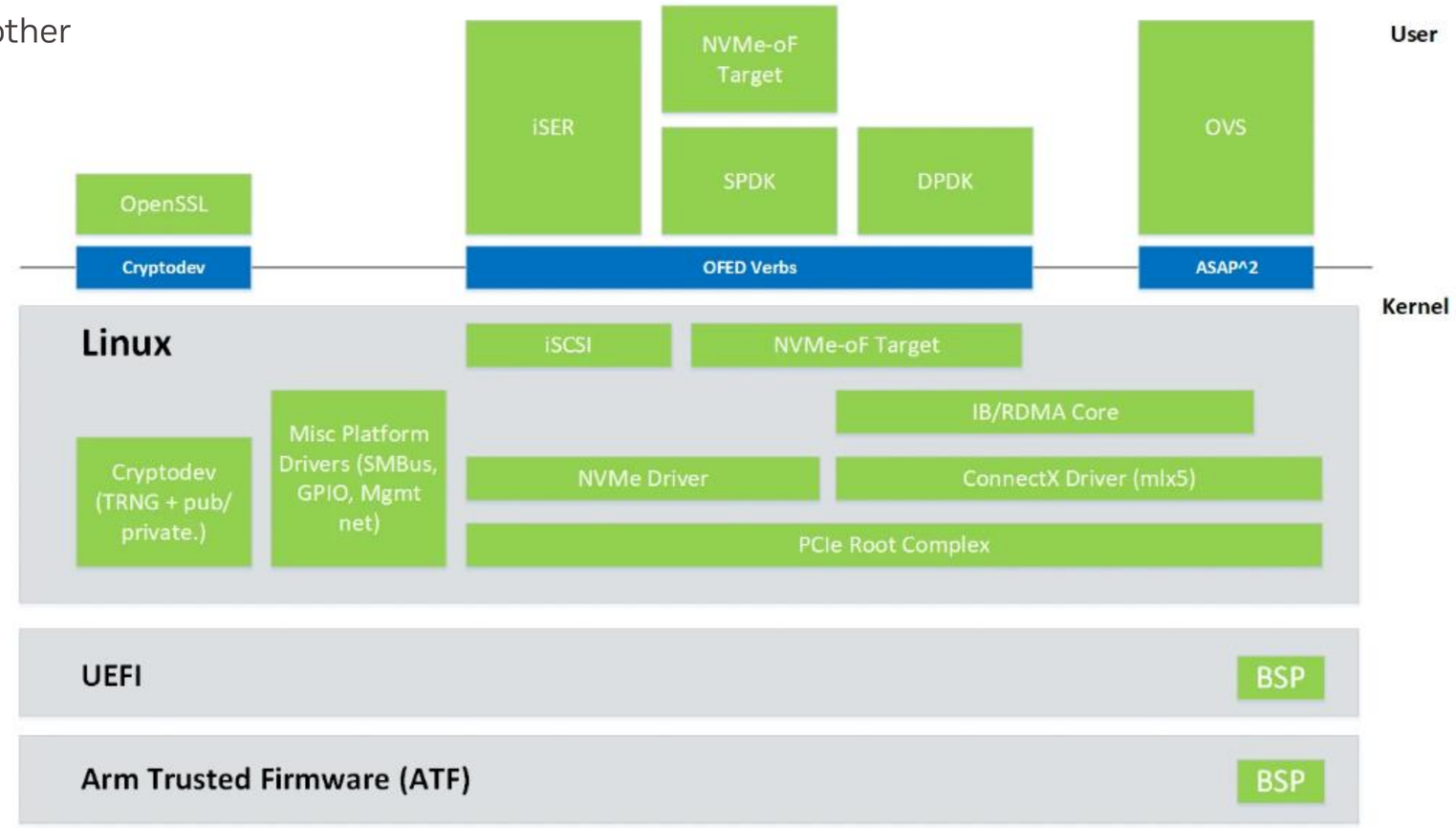
- Other configurations are also possible
- like loading Linux from eMMC, mounting a rootfs from eMMC or NVME, etc.

DOCA Software Framework | NVIDIA Developer | NVIDIA Developer

BlueField BSP

Board Support Package

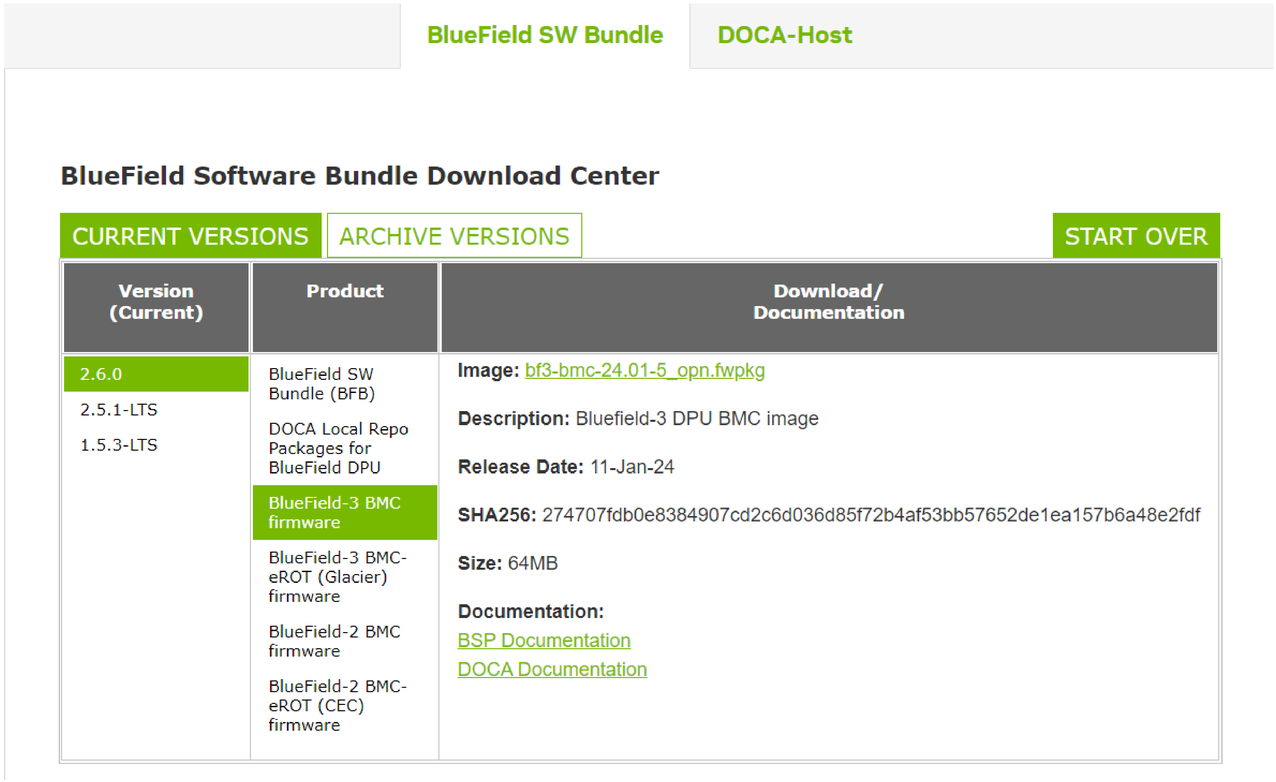
- NVIDIA BlueField DPU BSP v4.6.0 - NVIDIA Docs
- BlueField BSP includes the bootloaders and other essentials for loading and setting software components.



BlueField BMC

Board Management Controller

- NVIDIA BlueField BMC Software v24.01 - NVIDIA Docs
- Support for IPMI 2.0 (v1.1) Standards
  - Thermal control – access to all relevant temperature sensors, fan control
  - System management – power state control, power on/off, reboot/reset
  - Environmental monitoring – voltage/current/power
  - Serial over LAN (SOL)
  - RMCP/RMCP+
  - Event log management
  - Event alerting
  - VLAN support
- Support for DMTF Standards
  - Redfish specification (DSP0266)
  - Network Controller Sideband Interface (NC-SI) Specification (DMTF DSP0222)
- Support for BMC image update



BlueField SW Bundle | DOCA-Host

BlueField Software Bundle Download Center

CURRENT VERSIONS | ARCHIVE VERSIONS | START OVER

Version (Current)	Product	Download/ Documentation
2.6.0 2.5.1-LTS 1.5.3-LTS	BlueField SW Bundle (BFB) DOCA Local Repo Packages for BlueField DPU BlueField-3 BMC firmware BlueField-3 BMC-eROT (Glacier) firmware BlueField-2 BMC firmware BlueField-2 BMC-eROT (CEC) firmware	**Image:** bf3-bmc-24.01-5_opn.fwpkg **Description:** Bluefield-3 DPU BMC image **Release Date:** 11-Jan-24 **SHA256:** 274707fdb0e8384907cd2c6d036d85f72b4af53bb57652de1ea157b6a48e2fdf **Size:** 64MB **Documentation:** BSP Documentation DOCA Documentation

BlueField Firmware

- NVIDIA BlueField-3 DPU NIC Firmware Release Notes v32.40.1000 - NVIDIA Docs
- NVIDIA BlueField-2 DPU Firmware Release Notes v24.40.1000 - NVIDIA Docs
- NVIDIA Networking Firmware Downloads

NVIDIA BlueField DPU Firmware	
Product Line	Network Protocol
NVIDIA BlueField-3	InfiniBand/Ethernet
NVIDIA BlueField-2	InfiniBand/Ethernet

NVIDIA DOCA Stack

Comprehensive Acceleration Framework, SDK, Compilers, Services, and Tools

Containerized services

Applications

Networking | Security | Storage | HPC/AI | Telco | Media

DOCA

DOCA Services

- Orchestration
- Telemetry
- Firefly
- SDN/OVS-DOCA/P4 RT/HBN
- DPU Management

DOCA Libs

Crypto	App Shield	Rivermax	HPC/AI	GPUNetIO
Flow Gateway, Firewall	Comm Channel	Storage	PCC	DPA

DOCA Drivers

Networking ASAP² DPDK XLIO	Security DPDK SFT DPDK Crypto Inline Crypto	Storage SPDK SNAP Erasure Coding XTS Crypto	UCX	UCC
			RDMA	FlexIO

Secure, performance accelerated

Open foundation

[DOCA Documentation v2.6.0 - NVIDIA Docs]

DOCA LTS

New LTS release

Once a year based on DOCA October GA release

Lifespan - 3 years
After which no update supported

What's Included in DOCA LTS?

LTS of all DOCA components

Critical bugs and Security vulnerability fixes

All lifespan

Host OS (not DPU OS) minor version upgrade
(Ubuntu, CentOS, Rocky)

1st year only

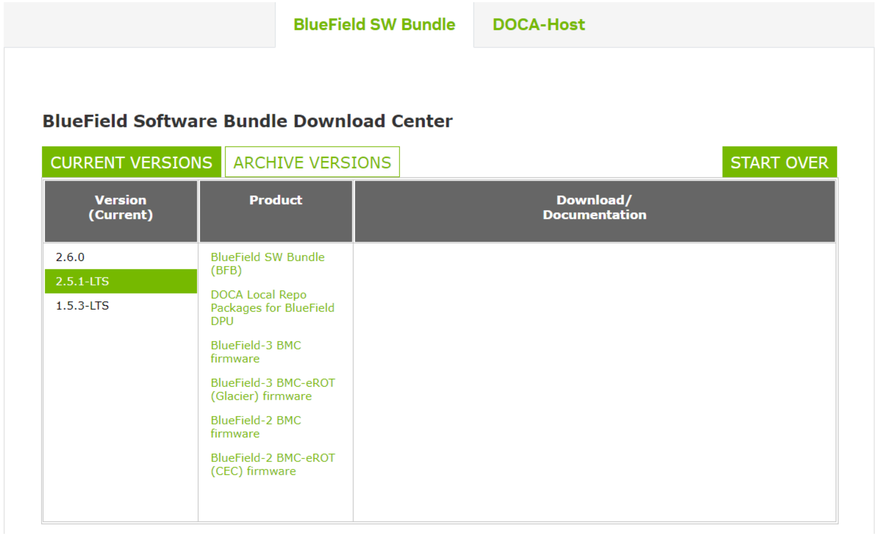
What's NOT Included in LTS?

New features

BlueField DPU OS updates

Host OS Major kernel updates

[DOCA Software Framework | NVIDIA Developer | NVIDIA Developer]



DOCA-HOST

DOCA Profiles and Supported Devices

DOCA Software Framework | NVIDIA Developer | NVIDIA Developer

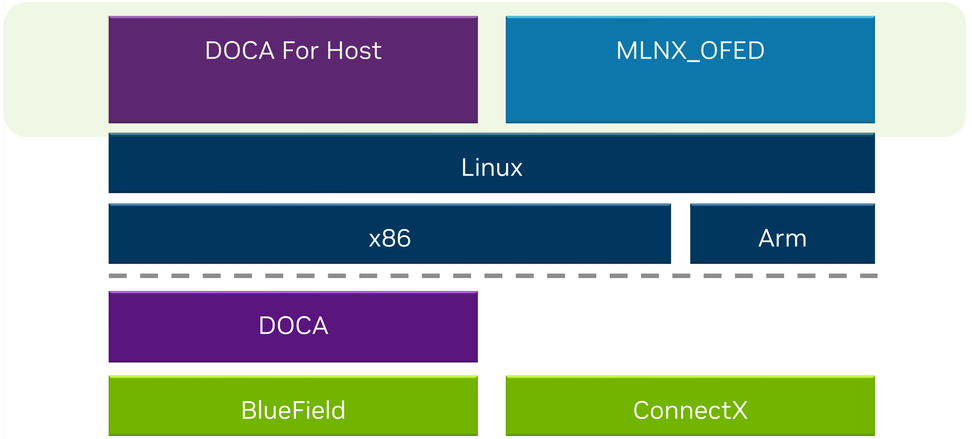
BlueField SW Bundle | DOCA-Host

DOCA-Host Software Download Center

CURRENT VERSIONS | ARCHIVE VERSIONS | START OVER

Version (Current)	OS distribution	OS distribution Version	Architecture	Download/ Documentation
2.6.0	Ubuntu	RHEL/Rocky 9.3	Select an Operating System from previous column	
2.5.1-LTS	UOS	RHEL/Rocky 9.2		
1.5.3-LTS	Sles	RHEL/Rocky 9.1		
	RHEL/CentOS/Rocky	RHEL/Rocky 9.0		
	Oracle Linux	RHEL/Rocky 8.9		
	OPENEULER	RHEL/Rocky 8.8		
	KYLIN	RHEL/Rocky 8.6		
	Debian	RHEL/CentOS/Rocky 8.5		
	CTYUNOS	RHEL/CentOS 8.4		
	BCLINUX	RHEL/CentOS 8.3		
	ALINUX	RHEL/CentOS 8.2		
		RHEL/CentOS 8.1		

Until DOCA 2.5:



2024+: