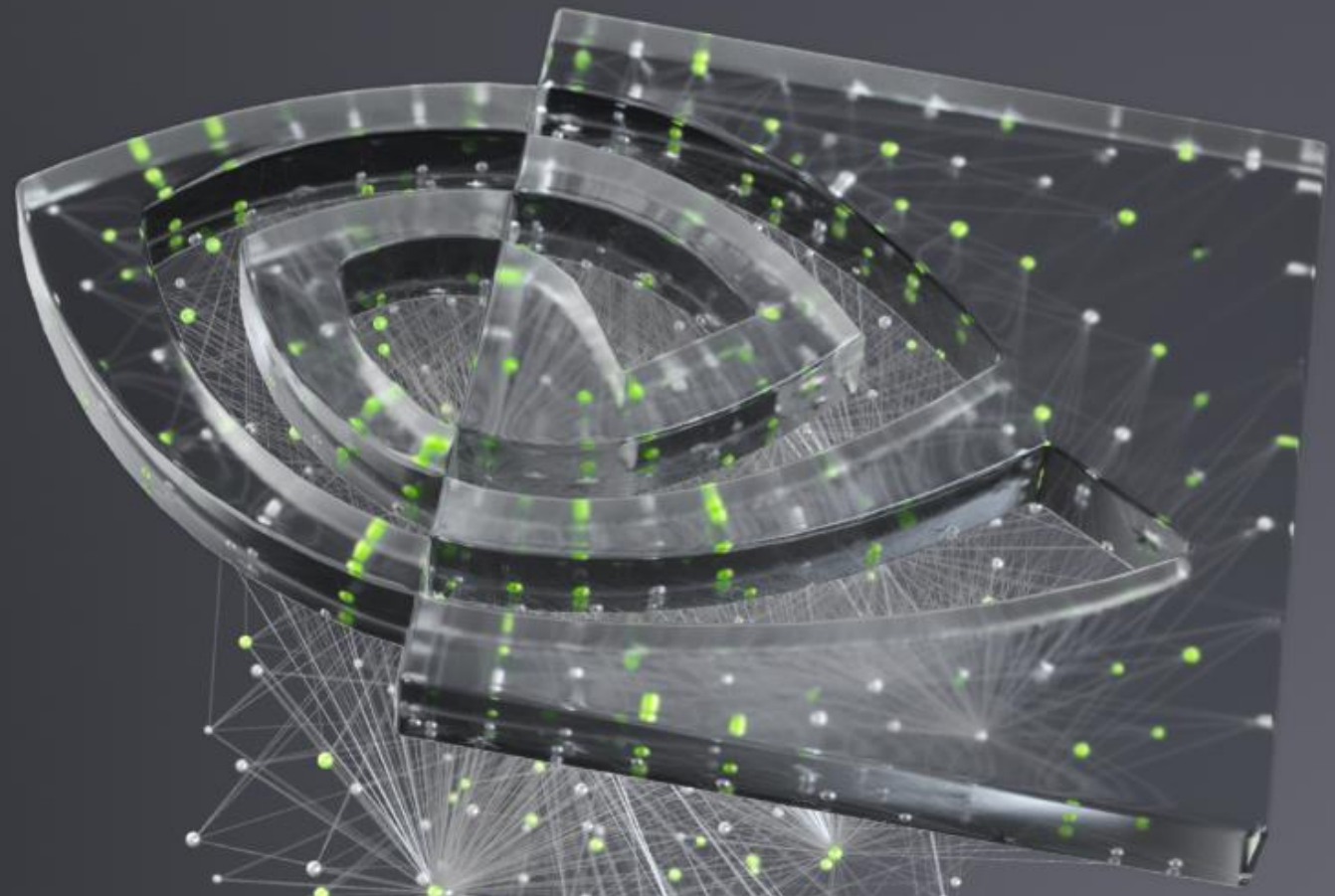


JETSON ORIN(AGX/NX) INTRODUCE

SA Jeff

AGENDA

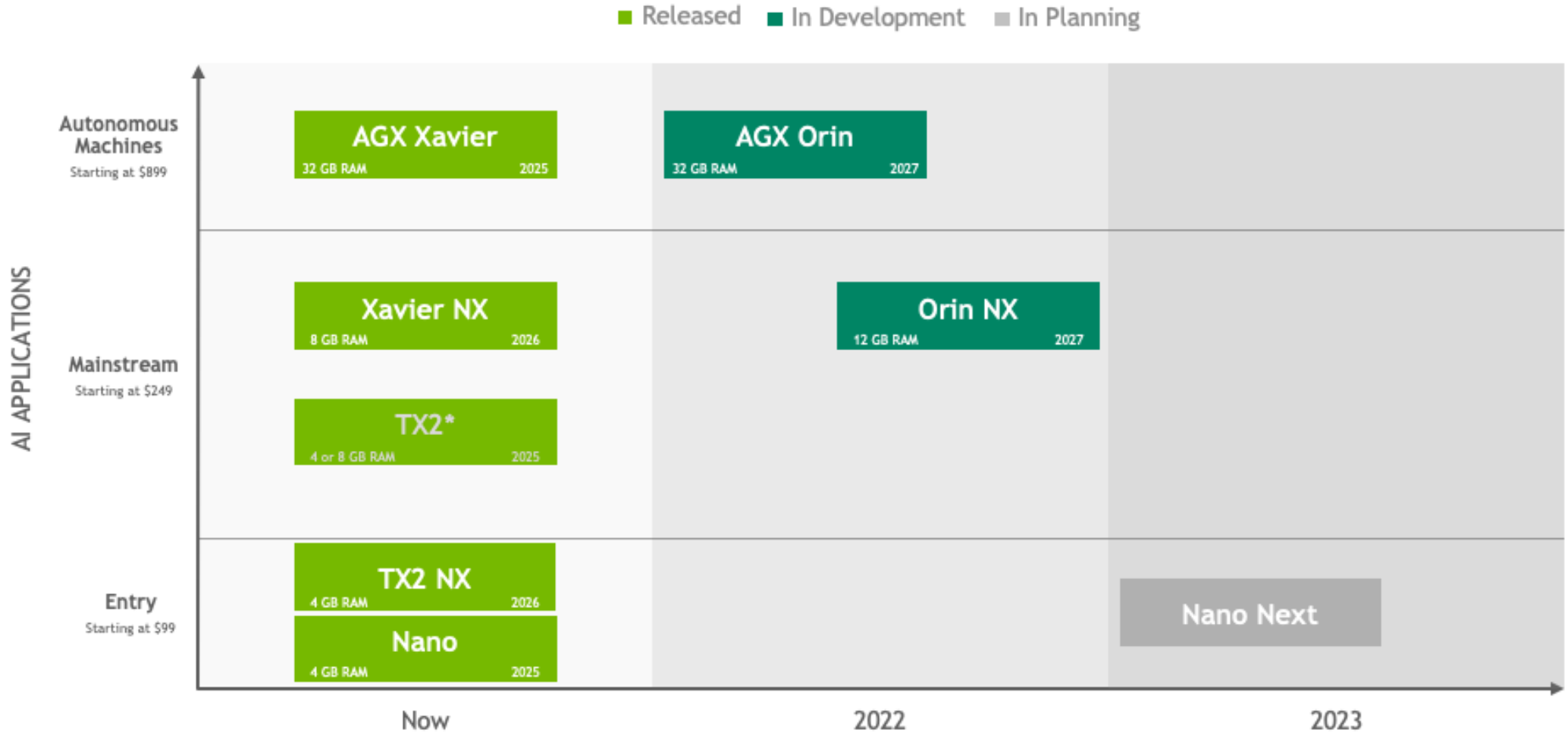
Jetson Orin series roadmap

- Advantages for Next-Generation AI Products

Orin series spec(AGX/NX)


JETSON ORIN SERIES ROADMAP

JETSON MODULES — COMMERCIAL ROADMAP

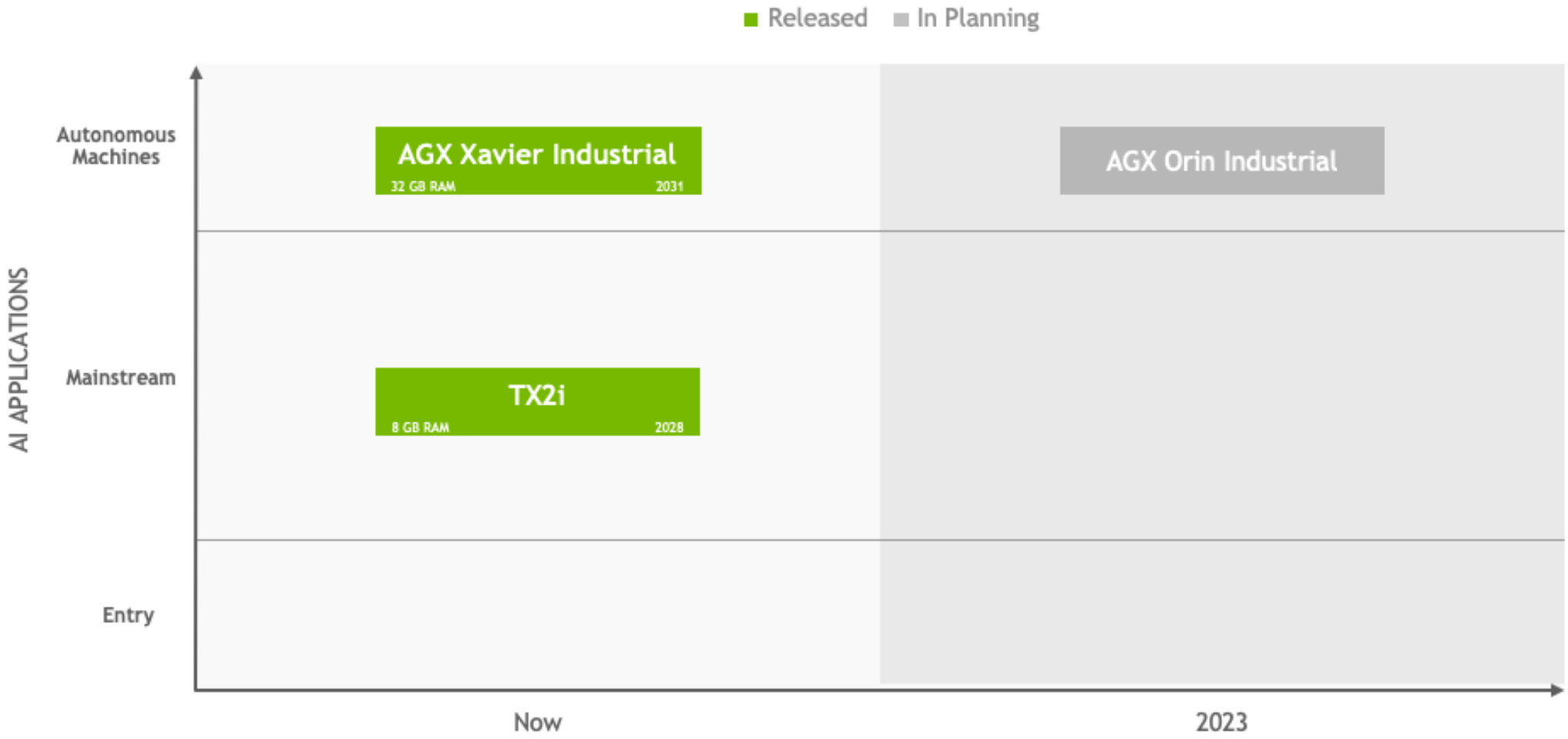


** Note: Jetson TX2 and TX2 4GB supply available until 2025, but new designs should use Jetson Xavier NX or Jetson TX2 NX*

The year in each box indicates supply availability at least / until. Lifecycle might be extended further; contact your partner sales manager. Operating life is 5 years 24x7 for commercial Jetson production modules. Products in development and planning are subject to change.

NVIDIA

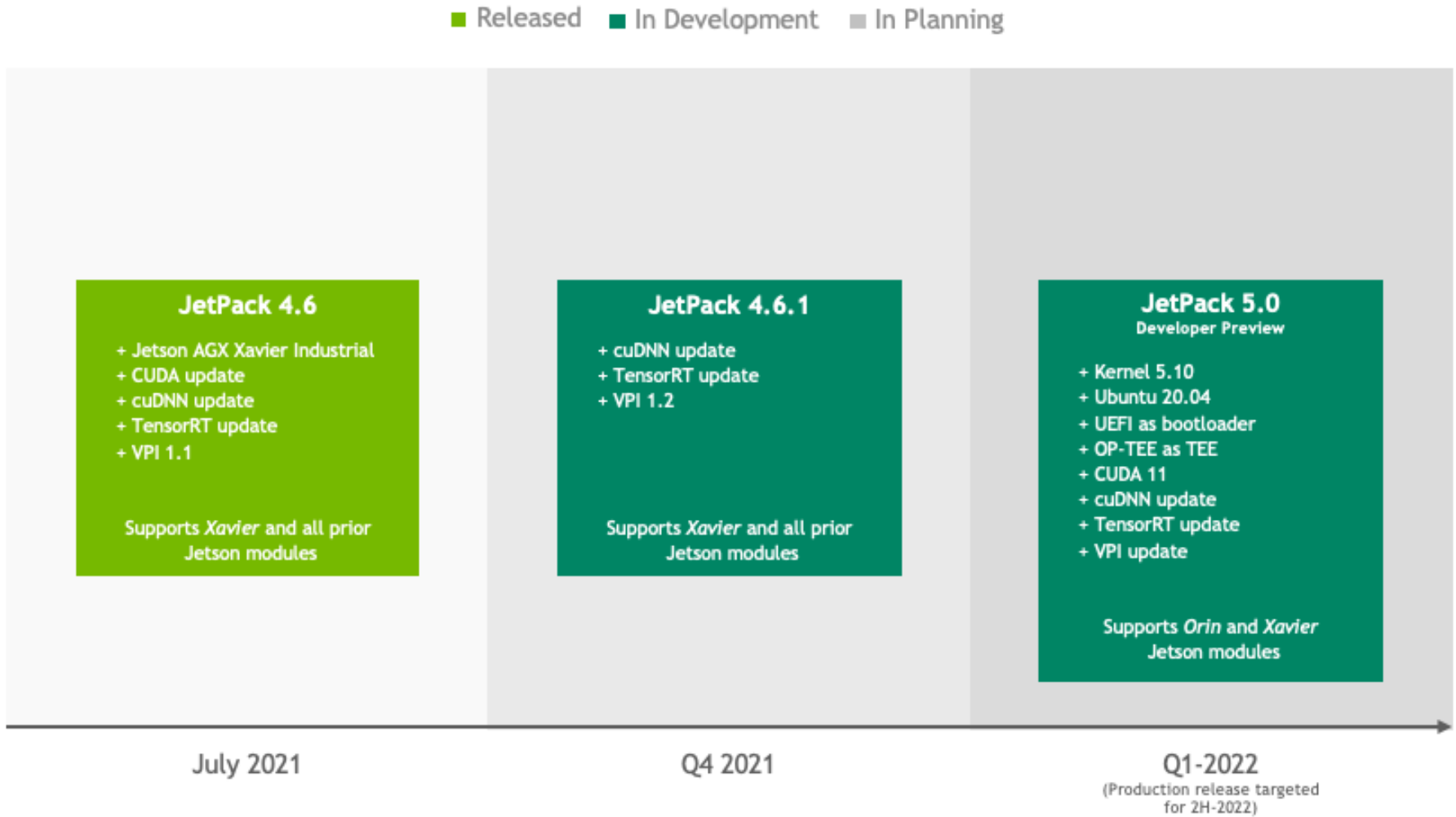
JETSON MODULES — INDUSTRIAL ROADMAP



The year in each box indicates supply availability at least / until. Lifecycle might be extended further; contact your partner sales manager.
Operating life is 10 years 24x7 for industrial Jetson production modules. Products in development and planning are subject to change.

NVIDIA

JETPACK SOFTWARE ROADMAP



Subject to change

NVIDIA

KEY FEATURES

JetPack 4.6

Compute Stack Update

CUDA, cuDNN, and TensorRT updates

Support for Triton Inference Server

Support for Triton Inference Server on JetPack with TensorRT, TensorFlow 1.x/2.x and PyTorch backends.

Image Based OTA

Enables full system upgrade by updating partition by partition

Tools and Instructions to create OTA payloads, download and apply OTA payloads

A/B Root File System Redundancy

Flashing and maintaining redundant root file system and ability to upgrade root file system slots. Failover to working slot in case of a boot failure during update.

Disk Encryption Support for External Media

Disk Encryption support extended to enable encrypting external storage like NVMe

Support booting from NVMe on Jetson Xavier NX and Jetson AGX Xavier

NVMe driver supported in Cboot to enable loading kernel and root file system from NVMe

JetPack 5.0 - Developer Preview planned for Q1-2022 Production release planned for 2H-2022

Compute Stack Update

CUDA 11.x and new versions of cuDNN and TensorRT

LTS Kernel 5.10

Reference file system based on Ubuntu 20.04

UEFI as bootloader (replacing CBoot)

OP-TEE as Trusted Execution Environment (replacing Trusty)

...

Subject to change

ADVANTAGES ORIN AGX

autonomous machines delivery / logistics robots, UAVs, V2X, ADAS

- Feeding multiple concurrent AI application pipelines.
- Largest and most complex models to solve problems like natural language understanding, 3D perception, and multi-sensor fusion.

Advantages	Details
Compact form factor	100mm x 87mm
Performance	6X, 200T
High-speed IO	204GB/s of memory bandwidth
POWER	15W, or up to a maximum of 50W
Pin-compatibility	Jetson Xavier AGX
Networking	4x 10GbE

ORIN SERIES SPEC(AGX/NX)

Jetson AGX Orin

200 TOPS
15W | 30W | 50W
100 mm x 87 mm
Available 1Q 2022

LEARN MORE >

Jetson Orin NX

100 TOPS
10W | 15W | 25W
70 mm x 45 mm
Available 4Q 2022

LEARN MORE >

Jetson AGX Xavier Series

32 TOPS
10-30W | 20-40W
100mm x 87mm
Starting at 899 USD

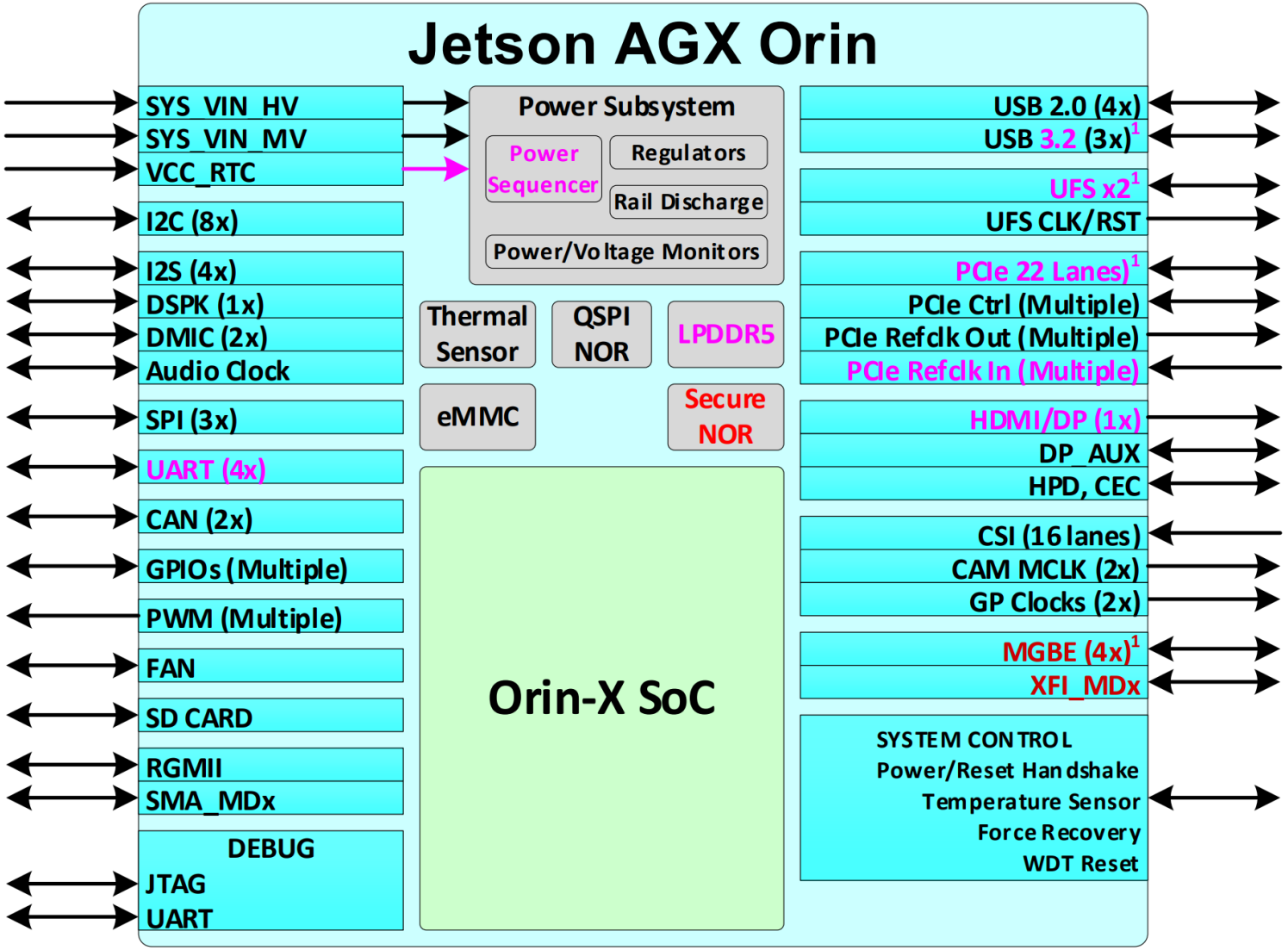
LEARN MORE >

Jetson Xavier NX

21 TOPS
10-20W
45 mm x 70 mm
399 USD

LEARN MORE >

JETSON AGX ORIN BLOCK DIAGRAM



JETSON AGX ORIN

Feature	Jetson AGX Xavier	Jetson AGX Orin
System Specifications and Device on the Module		
GPU	NVIDIA Volta™ architecture with 512 NVIDIA CUDA cores and 64 Tensor cores	NVIDIA Ampere Architecture with 2048 NVIDIA® CUDA® cores and 64 Tensor Cores
CPU	8-core NVIDIA Carmel Arm®v8.2 64-bit CPU, 8MB L2 + 4MB L3	12 core Cortex A78 ARM 64-bit CPU, 3 clusters (4x 256KB L2 + 2MB L3) + 4MB L4
DL Accelerator	GPU (22.6 TOPs) and 2x NVDLA Engines (5.7 TOPs each)	GPU (131 TOPs) and 2x NVDLA 2.0 Engines (48.5 TOPs each)
Vision Accelerator	7-Way VLIW Vision Processor (1.1 DL INT8 TOP)	
Memory	16/32 GB, 137 GB/s	32 GB, 204 GB/s
Storage	32 GB eMMC	64 GB eMMC
Networking RGMII MBGE	10/100/1000 Mbit Not Supported	0/100/1000 Mbit 4 x 10Gbe XFI
Video Decode	H.265 2x 8Kp30 6x 4K60 12x 4K30 26x 1080p60 52x 1080p30 H.264 4x4K60 8x4K30 16x1080p60 32x1080p30 VP9 4x4K60 8x4K30 18x1080p60 38x1080p30	AV1 1x8K30 2x4K60 4x4K30 10x1080p60 20x1080p30 H.265 1x8K30 2x4K60 6x4K30 12x1080p60 26x1080p30 H.264 1x4K60 2x4K30 6x1080p60 14x1080p30 VP9 26x1080p30

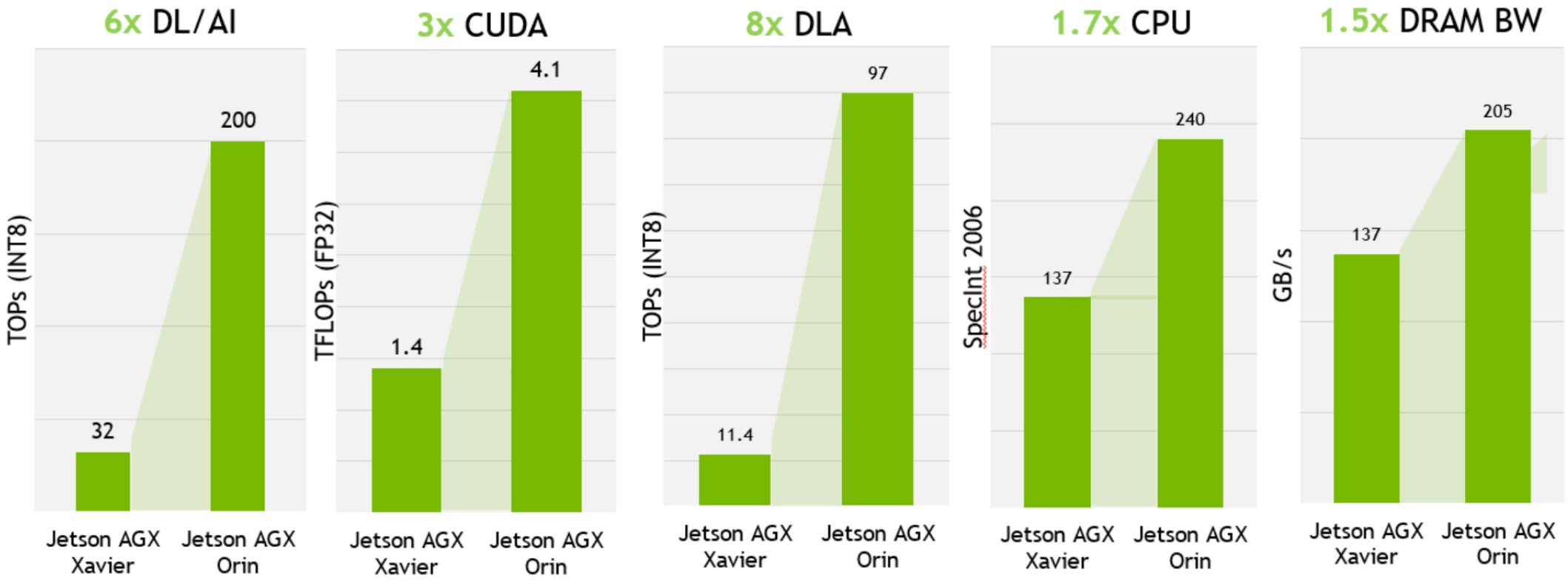
JETSON AGX ORIN

Video Encode	H.265 4x4K60 8x4K30 16x1080p60 32x1080p30	H.264 4x4K60 8x4K30 14x1080p60 30x1080p30	VP9 2x4K60 4x4K30 10x1080p60 20x1080p30	AV1 1x4K30 3x1080p60 7x1080p30	H.264/H.26 4 1x4K60 2x4K30 6x1080p60 15x1080p30
Video Input CSI SLVS	16 lanes MIPI CSI-2 D-PHY 1.2 (40 Gbps) C-PHY 1.1 (62 Gbps) 8-lane			16 lanes MIPI CSI-2 D-PHY 2.1 (40 Gbps) C-PHY 2.0 (164 Gbps) Not Supported	

JETSON AGX ORIN

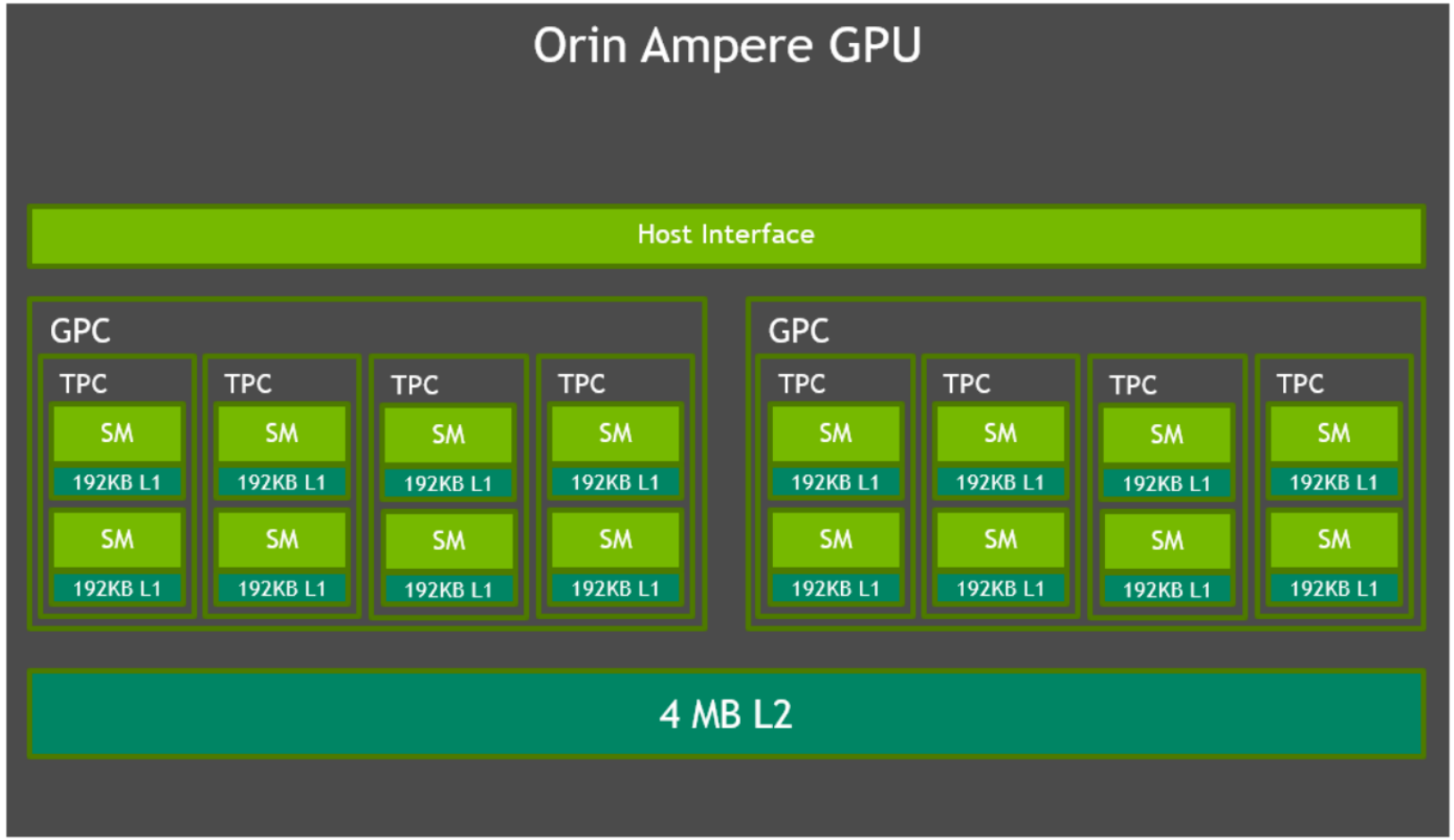
Interfaces		
USB 2.0	4x	
USB 3.x (See Note 1)	3x (3.1) Gen2 (10 Gbps)	3x (3.2) Gen2 (10 Gbps)
PCIe (See Note 1)	2 x1 + 1 x2 + 1 x4 + 1 x8. PCIe Gen 4. All support Root Port. Only x8 has Root Port and Endpoint support. 1 x8 shared with SLVS	2 x8 (or 1 x8 + 2 x4) + 1 x4 + 2 x1. PCIe Gen 4. All support Root Port. Only 2 x8 has Root Port and Endpoint support. 4 lanes of 1 x8 shared with MGBE.
Display	Three multi-mode (2x 4K60) (e)DP 1.4/HDMI™ 2.0a	One multi-mode (8K60, 2x4K60), (e)DP 1.4 (HBR3, MST, DSCT), HDMI™ 2.1
Camera	Up to 4x4 or 6x2 MIPI CSI interfaces 1 SLVS camera (Up to 8-lane interface)	Up to 4x4 or 6x2 MIPI CSI interfaces SLVS Not Supported
Audio I2S DMIC DSPK	 4x 2x 1x	 4x (compatible) + 2x additional 2x (compatible) + 2x additional 1x (compatible) + 1x additional
SDIO/SD Card	1x SD Card/SDIO	
Gigabit Ethernet	Supported	
MGBE (Multi-Gigabit Ethernet)	Not Supported	4x 10Gbe XFI (shared w/one x4 PCIe interface)
I2C	8x	

AI PERFORMANCE



ORIN AMPERE GPU BLOCK DIAGRAM

GPC->TPC->SM

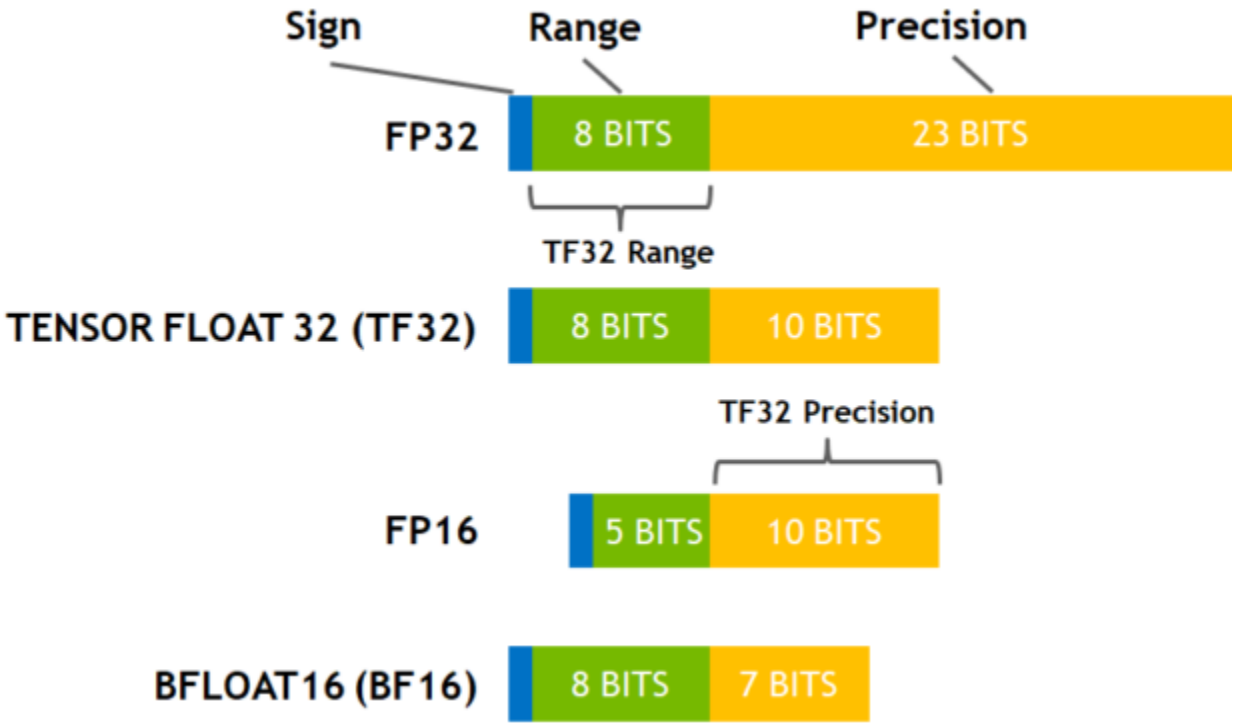


COMPUTE FEATURES

Third-generation NVIDIA Tensor Cores：

- TensorFloat-32 (TF32), bfloat16, FP16, and INT8
- Structured sparsity
- Compute Data Compression
- Up to 4× improvement in L2 read bandwidth, and up to a 2× improvement in L2 capacity

PRECISION OPTIONS USED FOR AI TRAINING



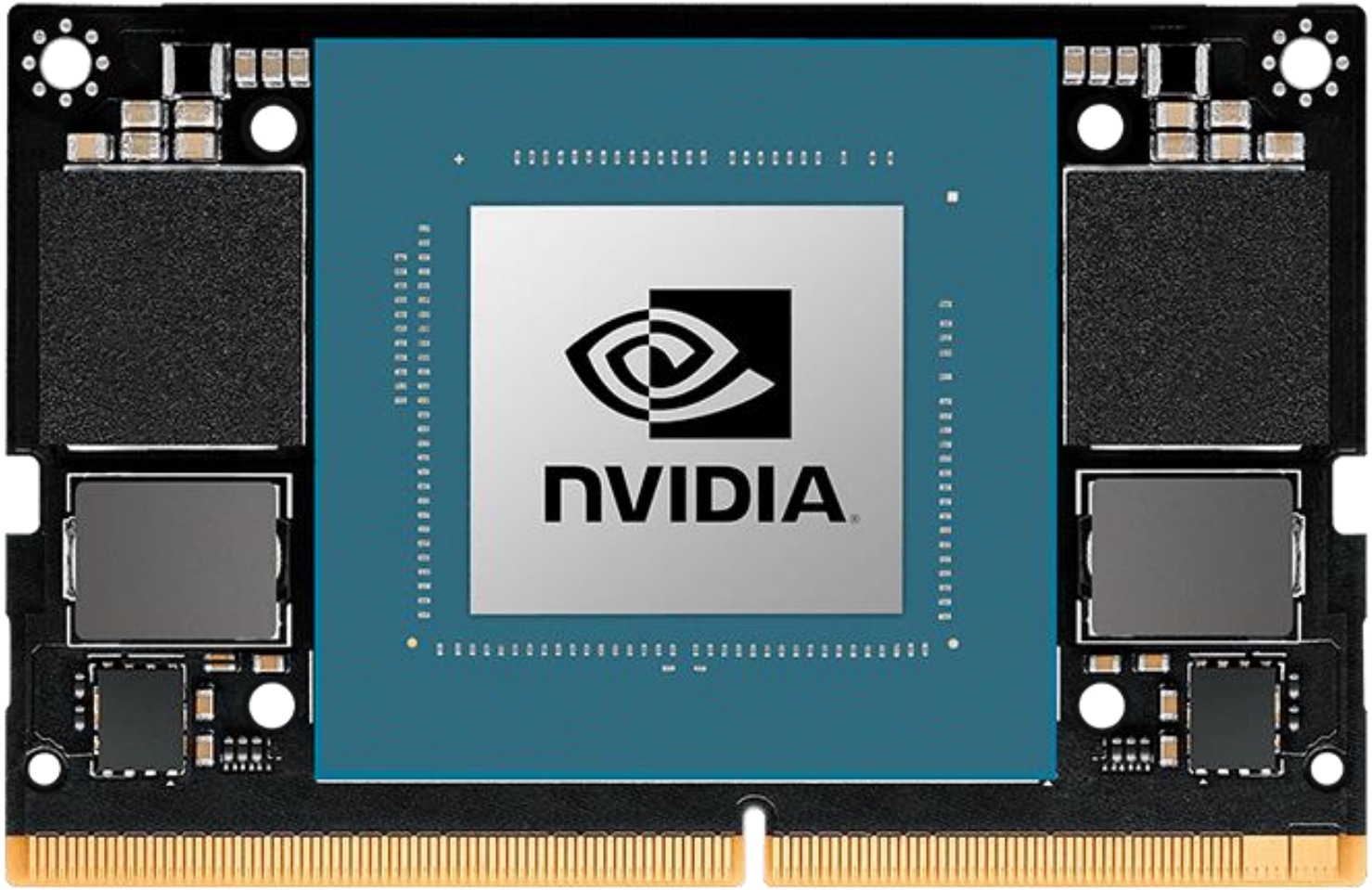
https://developer.nvidia.com/blog/accelerating-ai-training-with-tf32-tensor-cores/

NVIDIA.

AGX ORIN DEVKIT

ORIN NX

Jetson Orin NX module is coming in Q4 2022

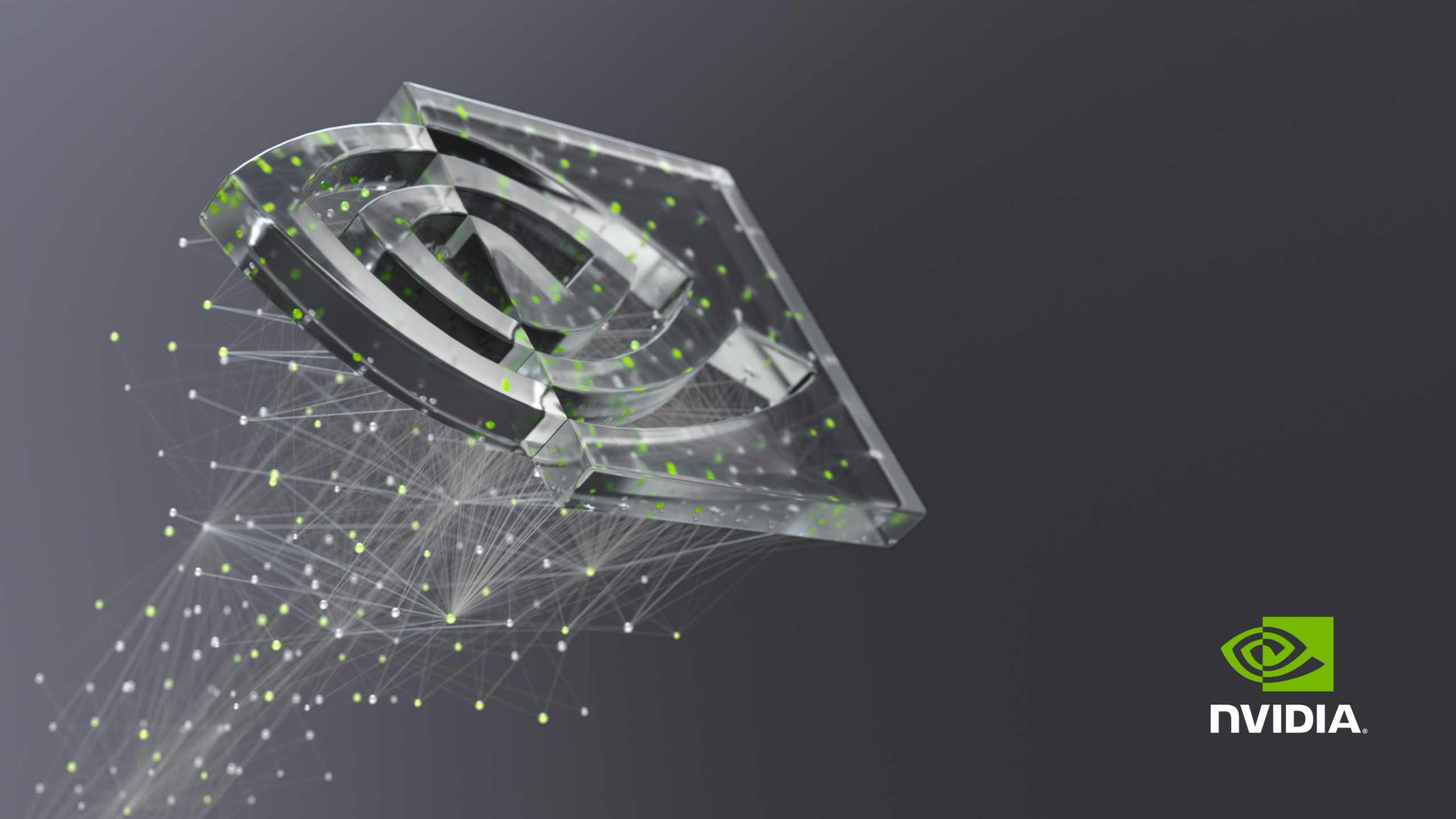


ORIN NX

ORIN PERFORMANCE. NANO SIZE

AI Performance	100 TOPS (INT8)
GPU	NVIDIA Ampere architecture with 1024 NVIDIA® CUDA® cores and 32 Tensor Cores
Max GPU Freq	1 GHz
CPU	8-core Arm® Cortex®-A78AE v8.2 64-bit CPU 3MB L2 + 6MB L3
CPU Max Freq	2 GHz
DL Accelerator	2x NVDLA v2.0
Vision Accelerator	PVA v2.0
Memory	12GB 128-bit LPDDR5 102.4 GB/s
Storage	Supports external NVMe

CSI Camera	Up to 4 cameras (8 via virtual channels*) 8 lanes MIPI CSI-2 D-PHY 1.2 (20 Gbps)
Video Encode	1x 4K60 | 2x 4K30 | 6x 1080p60 | 14x 1080p30 (H.265)
Video Decode	1x 8K30 | 2x 4K60 | 6x 4K30 | 12x 1080p60 | 24x 1080p30 (H.265)
UPHY	3 x1 + 1 x4 PCIe Gen 4 3x USB 3.2 Gen2
Networking	1x GbE
Display	1x 8K60 multi-mode DP 1.4a (+MST)/eDP 1.4a/HDMI 2.1
Other I/O	3x USB 2.0 3x UART | 2x SPI | 4x I2C | 1x CAN | DMIC | DSPK | 2x I2S | 15xGPIOs
Power	10W | 15W | 25W
Mechanical	69.6mm x 45mm 260-pin SO-DIMM connector


NVIDIA