2025

Understanding Nvidia - Part 2

AI AND ROBOTICS

TONY SHEN

DATA COMMUNICATIONS LAB | tshen@datacommlab.com

Contents

Revision History
03/31/2025

Introduction

Nvidia categorizes its products and offerings into two broad groups based on their purpose: **Digital AI** and **Physical AI**.

What is Digital AI?

In simple terms, Digital AI outputs **virtual results**, such as text, images, videos, and soundtracks. At the heart of Digital AI are **Large Language Models (LLMs)**. Nvidia provides a range of products specifically designed to create generative AI powered by LLMs. These products include **GPU cards, modules, and GPU servers** intended for datacenter deployment, often referred to as **datacenter products**. Examples of these include the **A100, H100, H200, GH200, B100, B200, GB200, and GB300**. The article *Understanding Nvidia Part 1* explores Nvidia's datacenter products in detail.

What is Physical AI?

Physical AI focuses on **robots and autonomous machines** that interact with physical environments. These include industrial robots in automotive assembly lines, warehouse robots in large distribution centers like Amazon's, service robots in restaurants and hospitals, as well as home-use robots—ranging from basic vacuum machines to advanced humanoid robots. Nvidia develops **embedded systems** to power these robots and autonomous machines. However, Nvidia does not manufacture robots or autonomous machines itself. Instead, it partners with specialized robot makers who leverage Nvidia's embedded systems to create their products.

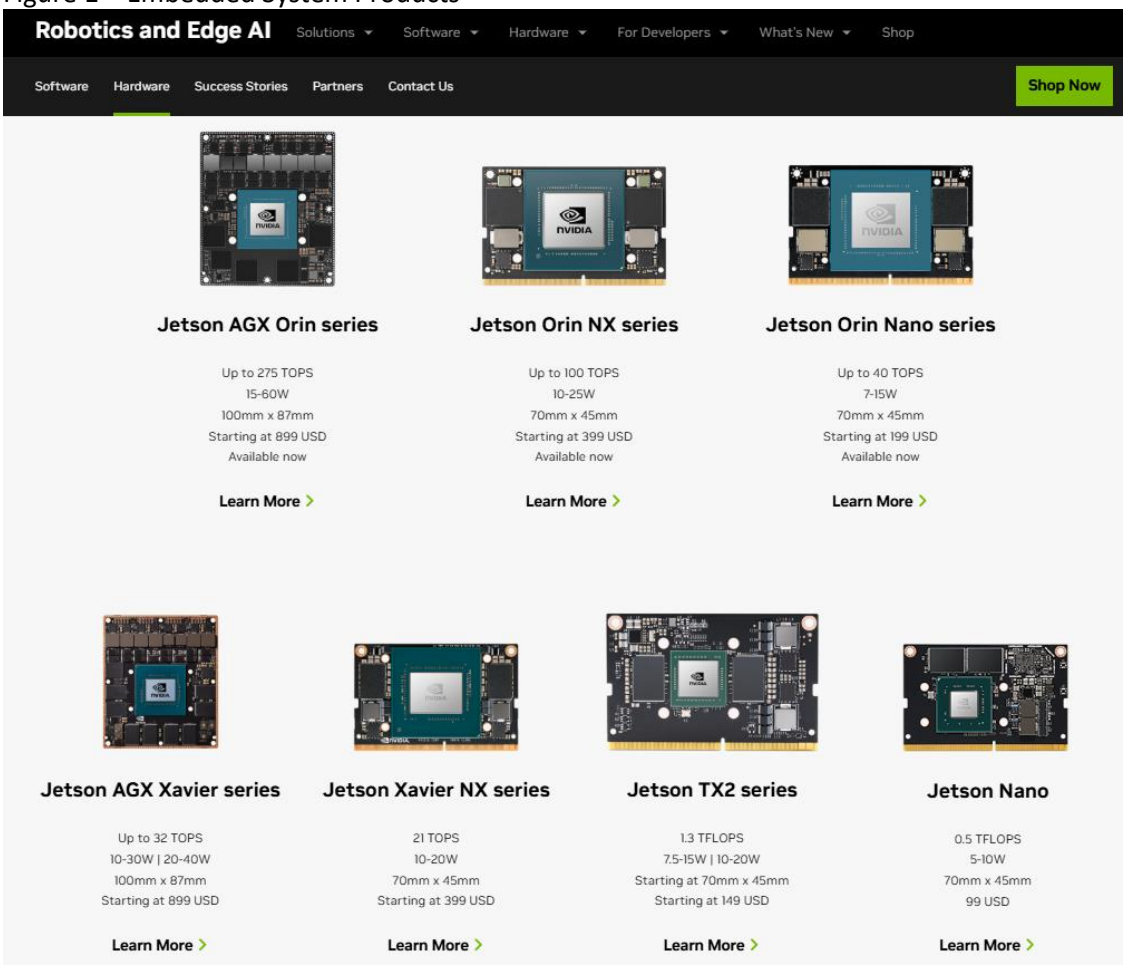
While Digital AI generates virtual outputs, **Physical AI** is designed to **engage with physical surroundings, automate production, and enhance our material lives**. Through Nvidia's embedded systems, these machines transform the way we interact with and improve the physical world.

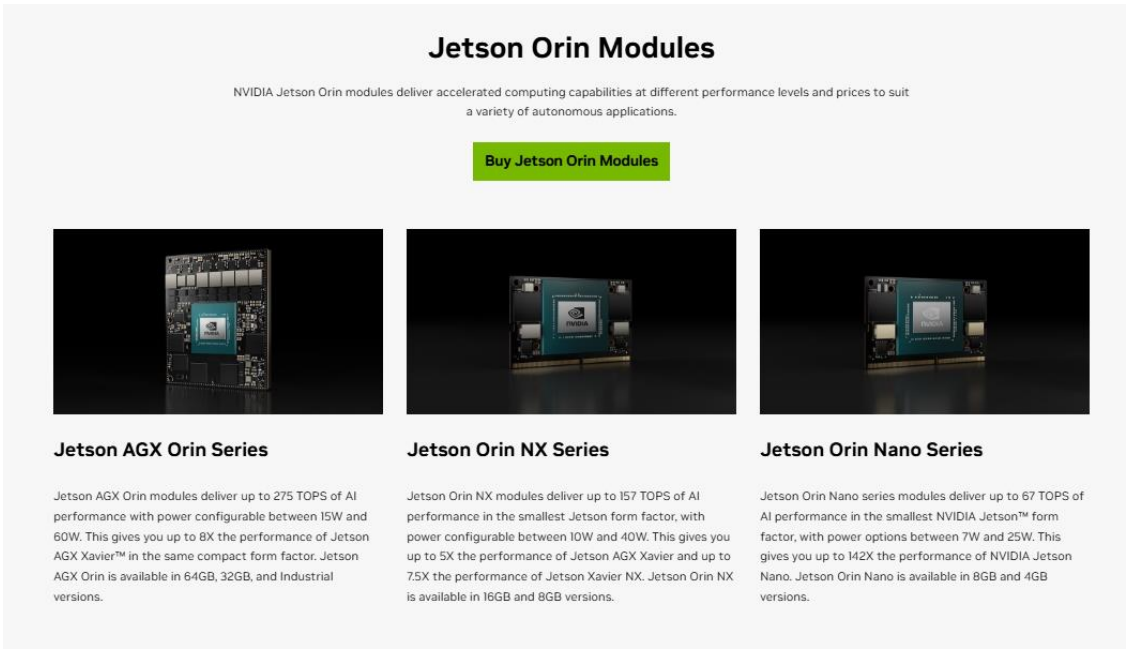
In this article, *Understanding Nvidia Part 2*, we will delve into Nvidia's embedded systems.

Embedded Systems

See Figure 1 below for Nvidia's embedded system products

Figure 1 – Embedded System Products





View Jetson Orin Technical Specifications

	Jetson AGX Orin series				Jetson Orin NX series		Jetson Orin Nano series		
	Jetson AGX Orin Developer Kit	Jetson AGX Orin 64GB	Jetson AGX Orin Industrial	Jetson AGX Orin 32GB	Jetson Orin NX 16GB	Jetson Orin NX 8GB	Jetson Orin Nano Super Developer Kit	Jetson Orin Nano 8GB	Jetson Orin Nano 4GB
AI Performance	275 TOPS		248 TOPs	200 TOPS	157 TOPS	117 TOPS	67 TOPS	67 TOPS	34 TOPS
GPU	2048-core NVIDIA Ampere architecture GPU with 64 Tensor Cores			1792-core NVIDIA Ampere c GPU with 56 Tensor Cores	1024-core NVIDIA Ampere architecture GPU with 32 Tensor Cores		1024-core NVIDIA Ampere architecture GPU with 32 Tensor Cores		512-core NVIDIA Ampere architecture GPU with 16 Tensor Cores
GPU Max Frequency	1.3 GHz		1.2 GHz	930 MHz	1173MHz	1173MHz	1020MHz	1020MHz	1020MHz

Source: [Jetson AGX Orin for Next-Gen Robotics | NVIDIA]

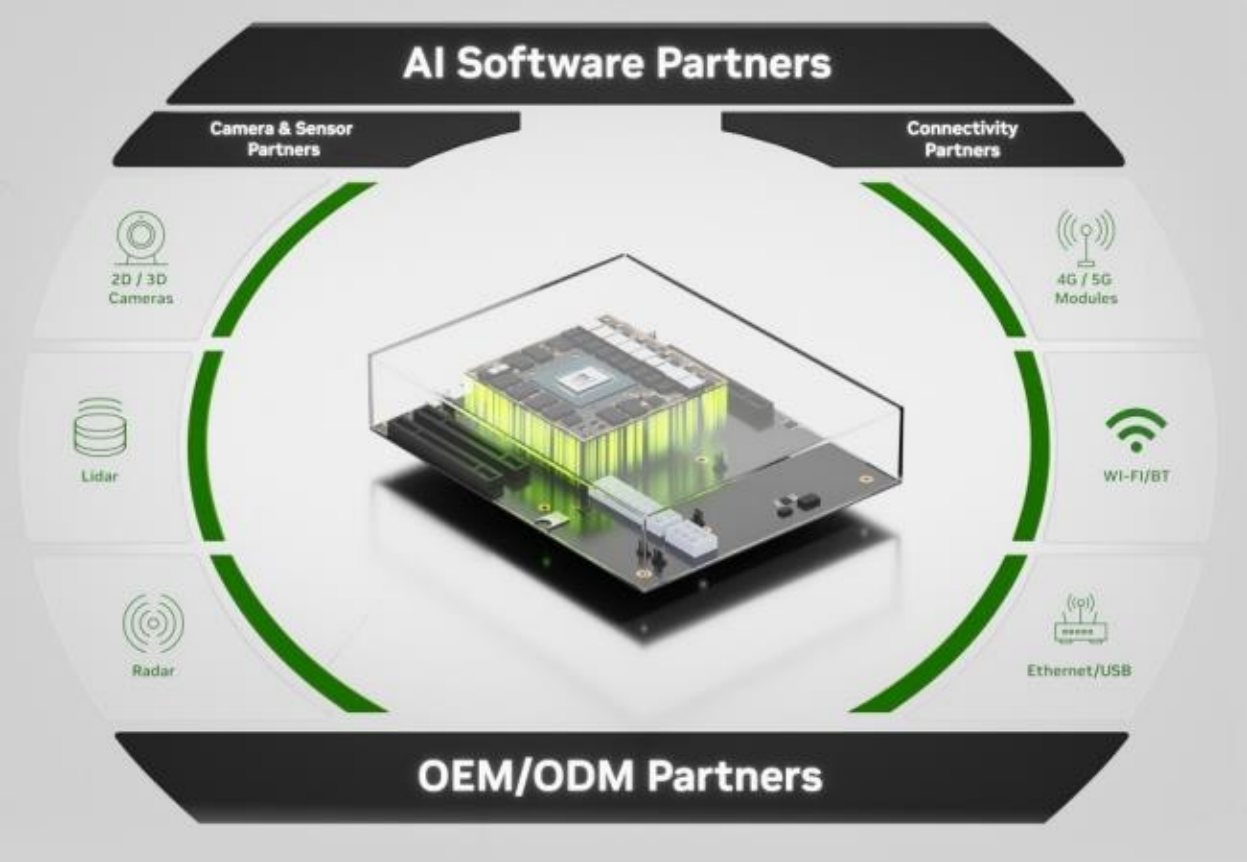
Nvidia's embedded products are classified into three series: **Jetson AGX Orin**, **Jetson Orin NX**, and **Jetson Orin Nano**. Nvidia provides comprehensive online documentation for these products, so we won't duplicate that information here. If you'd like to explore a detailed comparison between these series, refer to the **Jetson Orin Specifications**.

Since the **Jetson Orin Nano series** is widely used in robot and autonomous machine applications, this article will focus on describing the Jetson Orin Nano series in greater detail.

Nvidia Embedded System Ecosystem

To enable customers to effectively use its embedded system products for application development, Nvidia offers a sophisticated, comprehensive, and inclusive **software and hardware ecosystem**. In this ecosystem, **AI software partners** and **hardware OEM (Original Equipment Manufacturer) and ODM (Original Design Manufacturer) partners** play pivotal roles. Refer to **Figure 2 below** for an overview of Nvidia's embedded system development ecosystem.

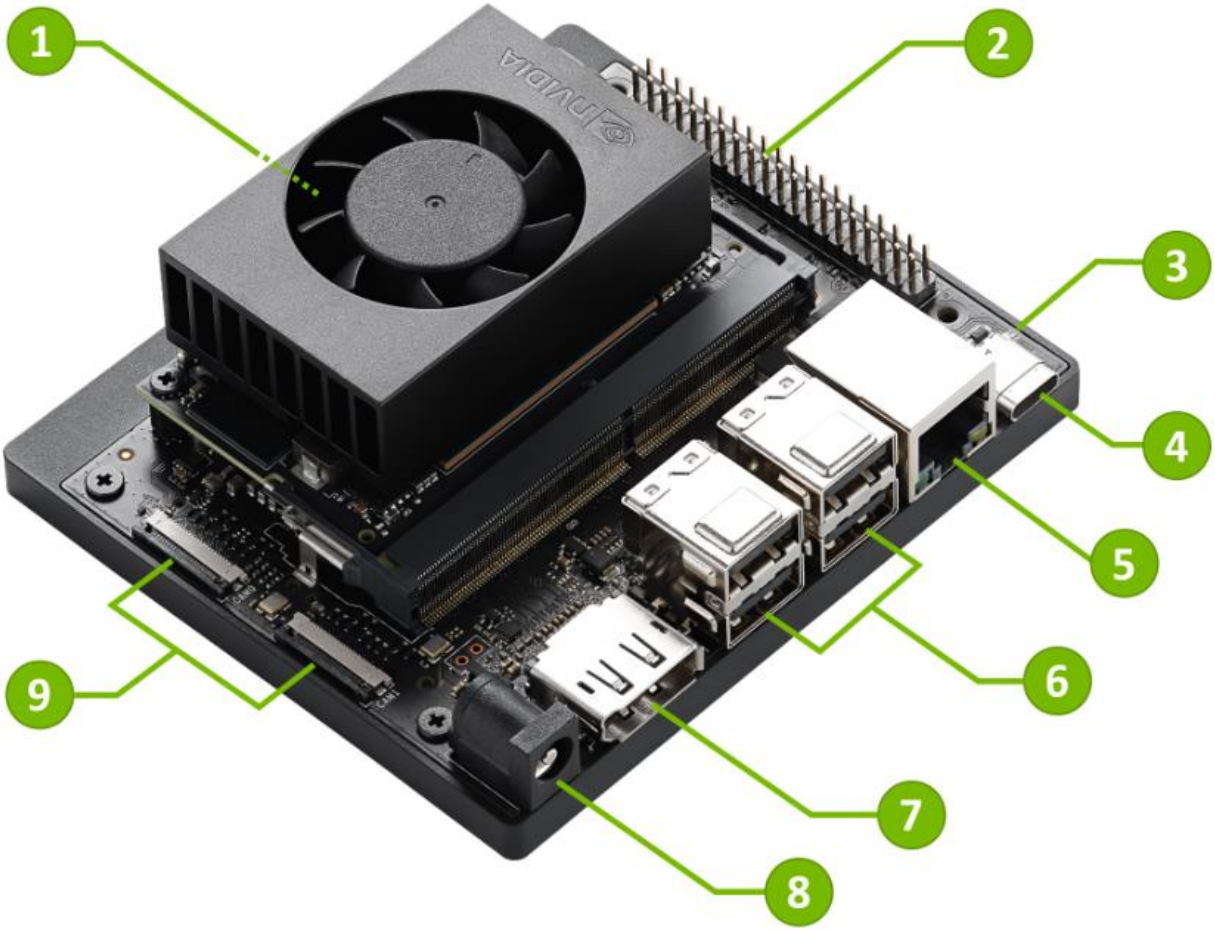
Figure 2 – Nvidia Embedded System Ecosystem



All of Nvidia's current embedded products are based on the **Ampere architecture**, offering **smaller form factors** and **lower power consumption** compared to Nvidia's datacenter products. Since robots and autonomous systems are typically **battery-operated** and optimized for size and weight, their GPU cards need to be **compact**, **lightweight**, and **energy efficient**. Nvidia's embedded products, with their streamlined designs, fulfill these essential robotic requirements.

Robots and autonomous systems possess the capability to **navigate independently** and perform specialized tasks by interacting with physical environments—like picking up objects or moving them with robotic arms and hands. Components such as **cameras, radar, Wi-Fi, 4G/5G connectivity, Ethernet/USB interfaces**, along with the robot's frame and limbs, facilitate these functions. To ensure seamless communication and control with these components, Nvidia's embedded GPU modules are equipped with multiple **I/O interfaces** (or ports).
For example, Figure 3 below illustrates the **Nvidia Jetson Orin Nano module package**, showcasing its design and I/O interface capabilities.

Figure 3 – Nvidia jetson Orin Nano Module



1. microSD card slot for main storage
2. 40-pin expansion header
3. Power indicator LED
4. USB-C port for data only
5. Gigabit Ethernet port
6. USB 3.1 Type A ports (x4)
7. DisplayPort connector
8. DC Barrel jack for 19V power input
9. MIPI CSI camera connectors

As illustrated in **Figure 3**, the package includes a **baseboard** equipped with various **ports**. The GPU is securely mounted on the baseboard with a **cooling fan** placed on top. Similarly, the ports are also integrated onto the baseboard. This package can be installed into a robot as a single cohesive component.

The ports within the package enable connections to multiple devices, such as **robot cameras**, the **robot control processor unit**, and peripheral devices like a **display, mouse, and keyboard**. These peripherals allow the user to access and interact with **JetPack**, the software platform running within the package. The specifics of JetPack and its functionality will be elaborated on in the following sections.

Nvidia Embedded System Hardware Partners

In **Figure 2**, on the left, we see **Camera & Sensor partners**, who create vision and sensor components such as **cameras, radar, and LiDAR (Light Detection and Ranging)**. These components are integrated into the robot to enable environmental perception.

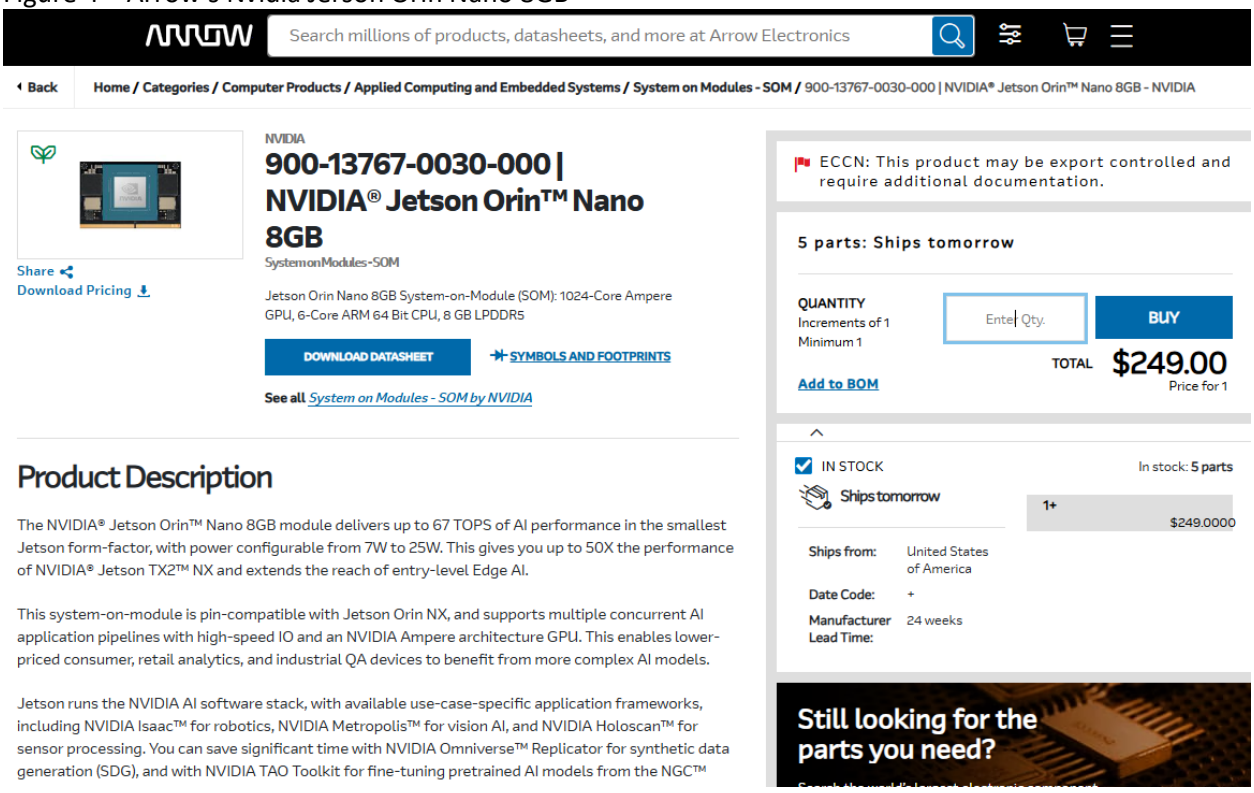
On the right side of **Figure 2**, we find **Connectivity partners**. They manufacture communication components like **4G/5G modules for GPS**, **Wi-Fi for remote control**, and **Ethernet/USB interfaces** for wired connectivity with other peripherals and equipment.

At the bottom of the diagram in **Figure 2**, we see **Nvidia OEM/ODM partners**, responsible for producing the GPU packages.

- **OEM (Original Equipment Manufacturer)** refers to companies that design and manufacture products or components sold under another company's brand—in this case, under **Nvidia's brand**.
- **ODM (Original Design Manufacturer)** refers to companies that design and manufacture products based on the specifications provided by another company—again, in this case, **Nvidia**.

In **Figure 4**, we explore an example of an Nvidia embedded system package from an OEM: **Arrow's Nvidia Jetson Orin Nano 8GB**.

Figure 4 – Arrow's Nvidia Jerson Orin Nano 8GB



At the **upper-left corner** of the screen, the Nvidia **product number**, **description**, and **image** of the **Nvidia Jetson Orin Nano 8GB GPU** are displayed. A **green butterfly icon** in the corner signifies the package's **full compatibility** with the GPU. Meanwhile, at the **upper-right**, a **red flag icon** serves as a warning that the package is subject to **U.S. Export Control regulations**. Additional documentation is required to verify the country where the package is intended to be delivered. Figure 5 below shows the package's specifications

Figure 5 – Product Technical Specifications

Product Technical Specifications

EU RoHS	Compliant
ECCN (US)	5A992c.
Part Status	Unconfirmed
HTS	8473.30.11.80
Automotive	Unknown
PPAP	Unknown
Processor Type	ARM
Processor Family	Cortex-A78AE v8.2
Processor Number	ARM Cortex-A78AE v8.2
Processor Cores	6 Core
Processor Speed (Hz)	1.5G
Form Factor	Small
JTAG Support	No
Data Cache Size (B)	1.5M|4M
Number of System Memory	1
RAM Type	LPDDR5
System Memory Socket Type	DRAM
System Memory Maximum Size (B)	8G
UART	3
USART	0
SPI	2
I2C	4
CAN	1

USB	6
SATA	0
PATA	0
GPIO	0
Audio Interfaces	1@DMIC/1@DSPK
RS485	0
RS232	0
RS422	0
PCI	0
AGP	0
Ethernet	0
Display Type	LCD
Display Interface	DP|eDP|HDMI
Operating Systems	Linux
Linux	Linux
Medical Application	No
Minimum Operating Supply Voltage (V)	5
Maximum Operating Supply Voltage (V)	20
Pinout Type	SOM

The specifications reveal that the package supports various communication protocols and includes **6 USB ports**, **1 audio port** for speaker connectivity, and **1 video port** for connecting to an LCD display.

Identified by its **Pinout Type**, the package is classified as a **System on Module (SOM)**. A SOM operates independently with an **operating system (OS)** installed—in this case, **Linux**. Once powered, users can connect a keyboard and mouse to two of the USB ports and an LCD display to the video port, allowing them to interact with the SOM as they would with a regular Linux device. This setup also facilitates the use of **Nvidia JetPack**, which is included within the module.

It's important to note that the SOM does not come with its OS and JetPack pre-installed. Nvidia provides detailed instructions for users on how to **flash the OS and JetPack** onto the package before it can be utilized. For more information, refer to the guide *Getting Started with Nvidia Jetson Orin Nano Development Kit*.

OEM/ODM packages often have slight variations compared to Nvidia's packages. For example:

- **Arrow's package** includes **6 USB ports**, while Nvidia's package has only **4 USB ports**.
- Arrow's package also features an **audio port** for connecting a robot's speaker, whereas Nvidia's package lacks this feature.

Additionally, OEM/ODM packages may differ slightly among themselves, providing customers with more options. These choices foster **competition**, which in turn **drives down prices** and **enhances quality**. However, all OEM/ODM packages are required to be **fully compliant** with Nvidia's designs.

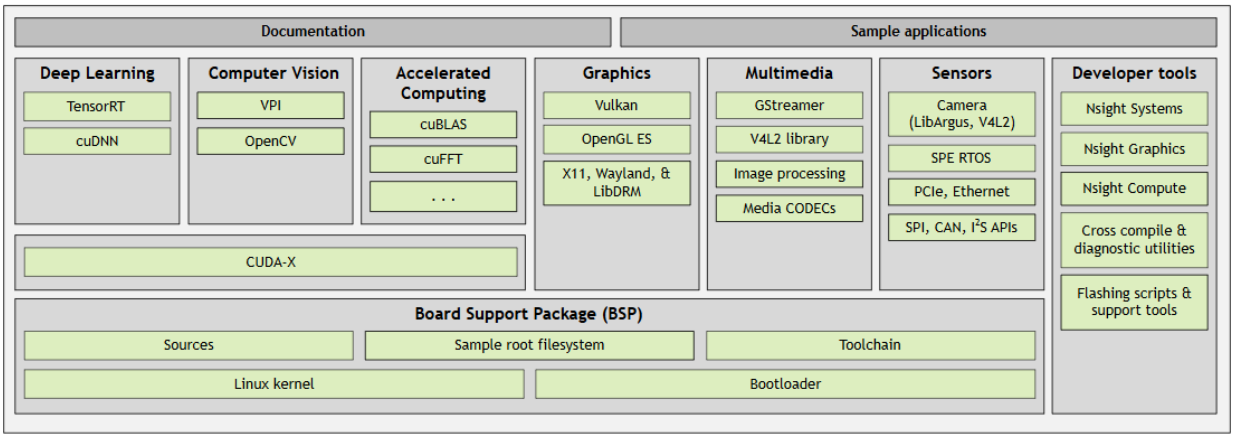
The specifications for Arrow's product may include terms or items that might be unfamiliar to you. To clarify these, please refer to the **Glossary** provided in this article for detailed descriptions.

JetPack - Nvidia Embedded System Software

Nvidia provides a comprehensive software system called **Nvidia JetPack**, enabling users to develop their own robotic applications powered by Nvidia embedded GPU hardware. As of the time of this writing, the latest release is **JetPack 6.2**. Developers can access updates, bug fixes, and future releases through Nvidia's **download portal**.

Figure 6 below illustrates the **JetPack architecture**, showcasing its design and functionality.

Figure 6 – Nvidia JetPack Architecture



As the **architecture** demonstrates, **JetPack** includes **sample applications** and **documentation**.

Developers can utilize the sample applications as **templates** to create customized versions tailored to their specific needs. During this process, the documentation serves as a **comprehensive guide**, assisting developers through every step of their development journey. Additionally, the documentation provides links to Nvidia's extensive **online resources**, offering valuable support.

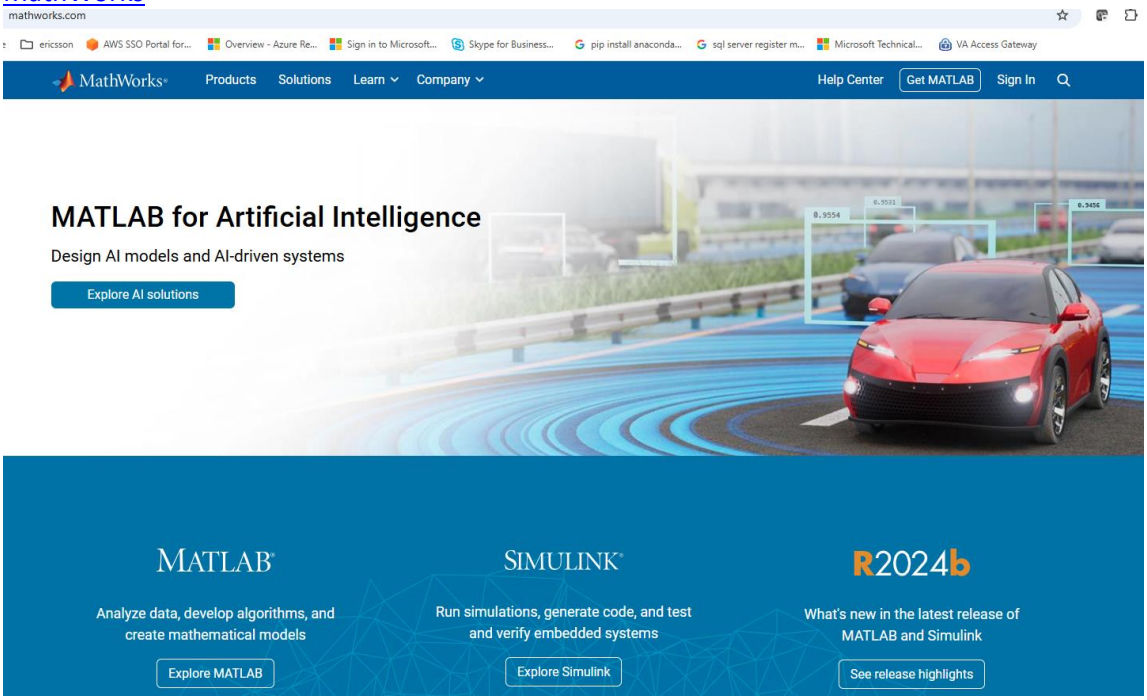
Both the **documentation** and **sample applications** are built upon the **Board Support Package (BSP)**, which essentially functions as Nvidia's customized version of the **Linux Operating System**.
.

Nvidia Embedded System Software Partners

Nvidia collaborates with numerous software partners for its Jetson Orin platform, which is specifically designed for robotics and autonomous systems at the edge, distinct from its datacenter-focused products like the A100, H100, or DGX systems. While many partners work across Nvidia's broader ecosystem, some focus primarily on the Jetson Orin for edge AI applications, particularly in robotics and autonomous systems, rather than datacenter solutions. Below are examples of large software partners that align with this focus, based on their offerings and Nvidia's Jetson ecosystem as of March 30, 2025.

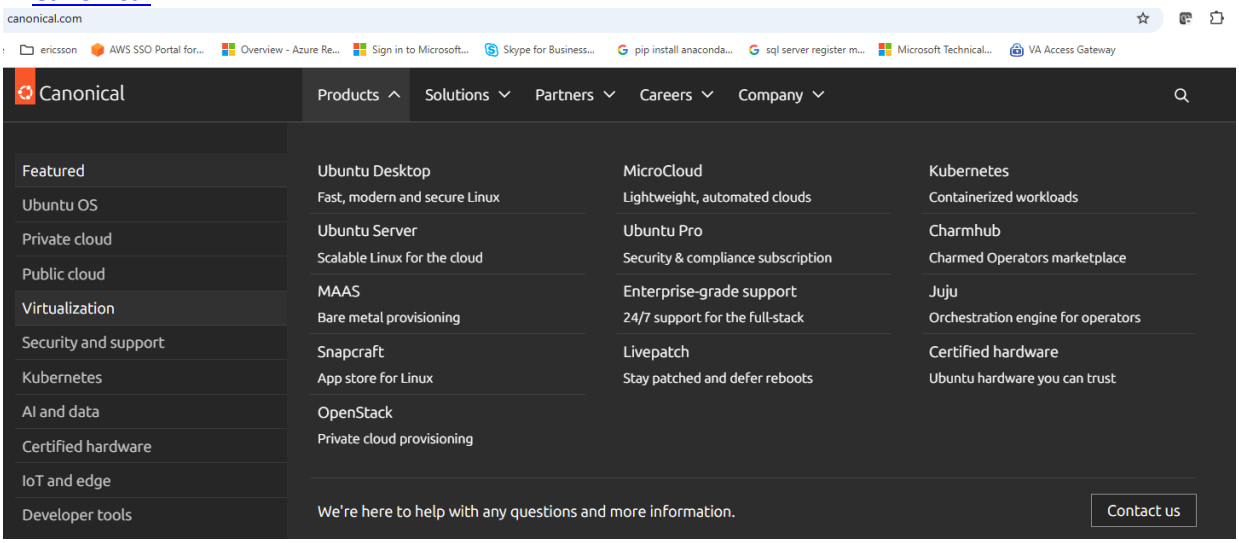
Note that "large" here refers to well-established companies with significant presence in robotics or edge AI space, not including those primarily tied to datacenter products.

1. MathWorks



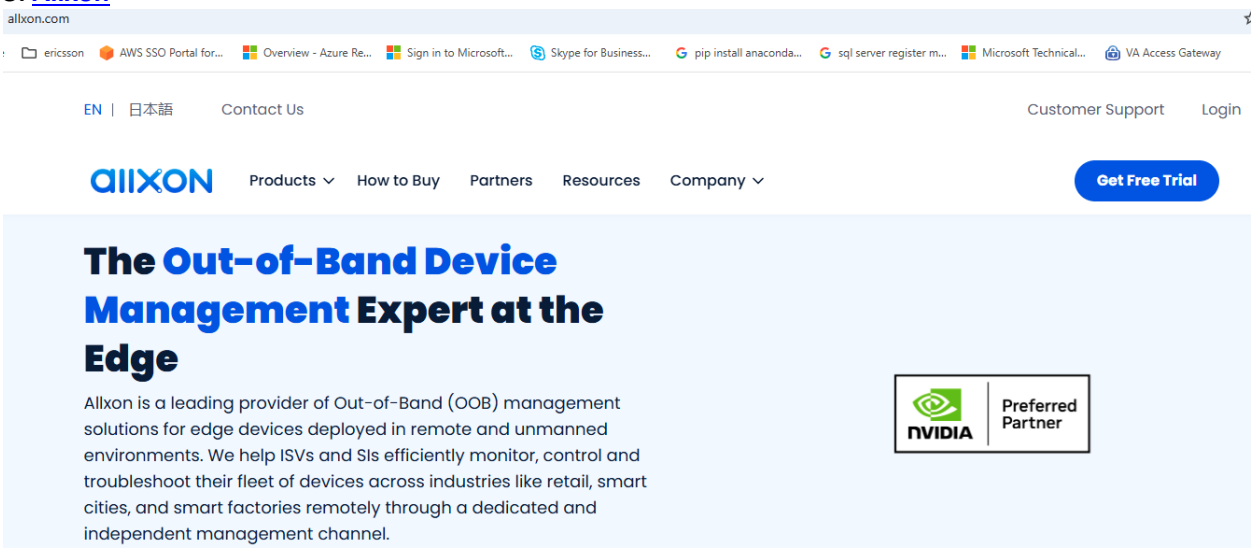
- **Focus**: MathWorks, known for MATLAB and Simulink, partners with Nvidia to provide tools for designing, simulating, and deploying AI models on Jetson Orin platforms. Their software is widely used for robotics and autonomous systems development, including control systems, computer vision, and sensor fusion.
- **Jetson Orin Relevance**: MATLAB and Simulink support Jetson Orin for prototyping and deploying algorithms directly on edge devices, such as autonomous robots or drones. This includes integration with Nvidia's JetPack SDK for real-time processing.
- **Datacenter Exclusion**: MathWorks' Nvidia collaboration centers on embedded systems and edge AI, not datacenter GPUs like the H100 or B200. Their tools are tailored for engineers building physical systems rather than large-scale cloud or datacenter workloads.

2. Canonical



- **Focus**: Canonical, the company behind Ubuntu, collaborates with Nvidia to optimize Ubuntu for Jetson Orin platforms, providing a robust OS for robotics and edge AI development.
- **Jetson Orin Relevance**: Ubuntu is a primary OS for Jetson Orin developer kits (e.g., AGX Orin, Orin Nano), supporting robotics frameworks like ROS (Robot Operating System) and Nvidia's Isaac platform. Canonical ensures long-term support and security updates for these edge deployments.
- **Datacenter Exclusion**: While Canonical supports Ubuntu across various Nvidia platforms, their Jetson partnership is edge-focused, distinct from datacenter-oriented collaborations (e.g., with DGX systems), which are less emphasized in their Jetson-specific offerings.

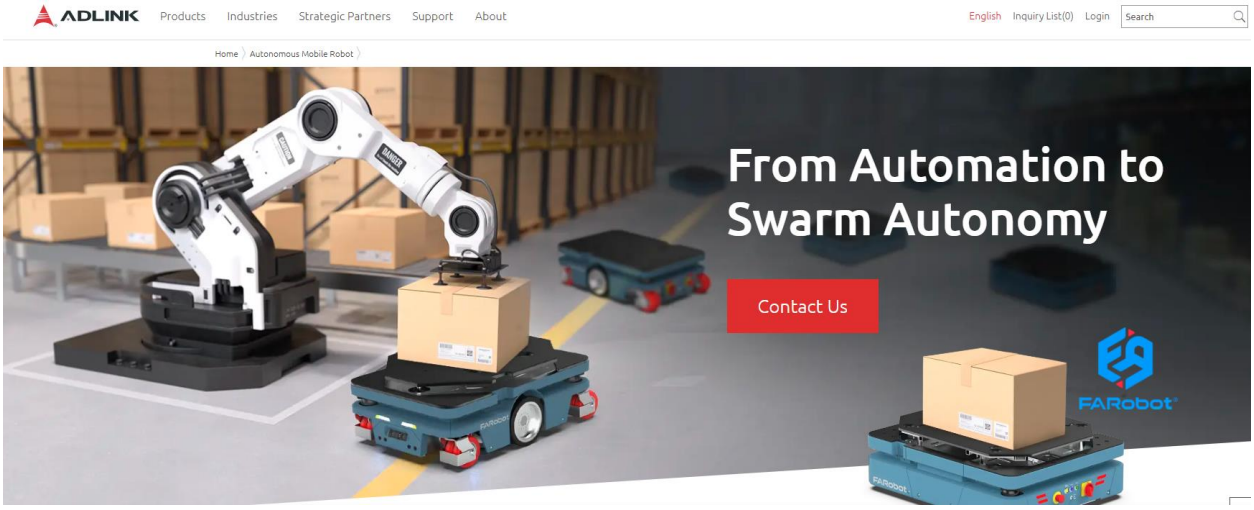
3. Allxon



- **Focus**: Allxon provides device management and monitoring software for edge AI systems, with a strong emphasis on Nvidia Jetson platforms.

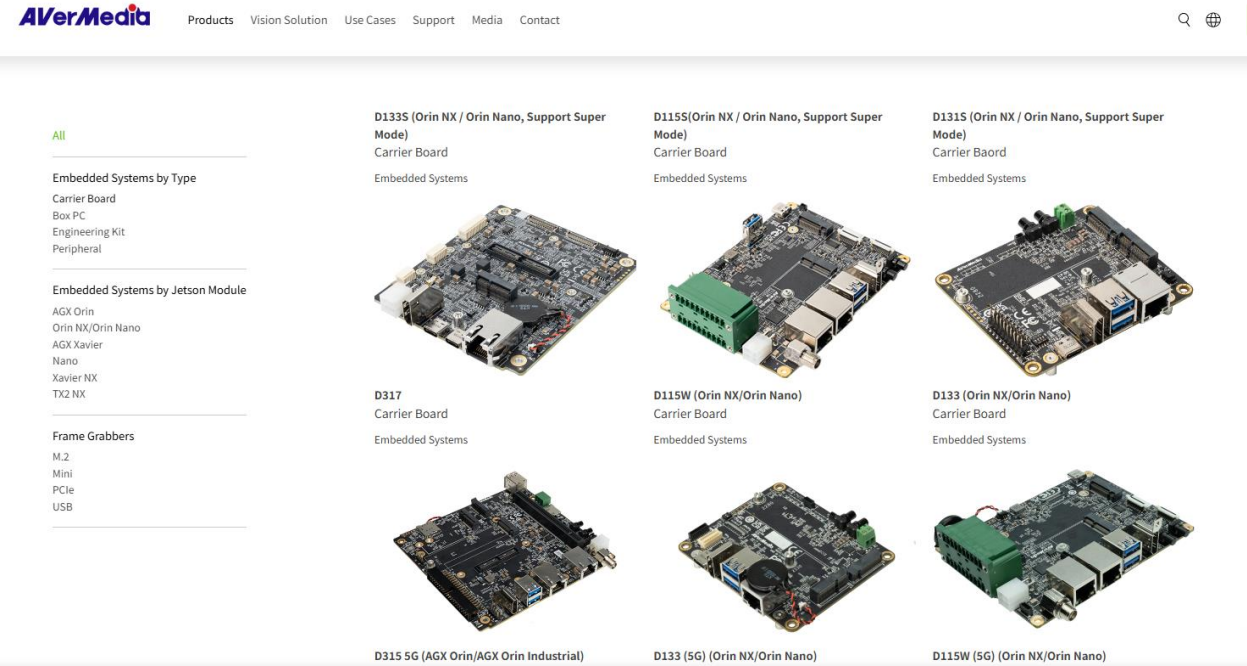
- **Jetson Orin Relevance**: Allxon's solutions enable remote management, OTA updates, and fleet monitoring for Jetson Orin-based robots and autonomous systems, such as those in logistics or manufacturing. Their integration with Jetson Orin enhances deployment scalability.
- **Datacenter Exclusion**: Allxon's scope is limited to edge devices, with no apparent focus on Nvidia's datacenter products like the H100 or GB200, making them a Jetson-centric partner.

4. Adlink (software and hardware)



- **Focus**: ADLINK offers software and hardware solutions for edge AI, with a significant emphasis on robotics and autonomous systems using Nvidia Jetson Orin.
- **Jetson Orin Relevance**: ADLINK provides ROS 2-based software stacks and development tools optimized for Jetson Orin modules (e.g., AGX Orin, Orin NX), targeting applications like autonomous mobile robots (AMRs) and industrial automation.
- **Datacenter Exclusion**: While ADLINK works with Nvidia broadly, their software efforts for Jetson Orin are edge-specific, not tied to datacenter products like the Blackwell or Hopper series used in datacenters.

5. AVerMedia (software and hardware)



- **Focus**: AVerMedia offers carrier boards, which are also called baseboards in Nvidia's parlance. AVerMedia not only make carrier boards but also provide software solutions optimized for Jetson Orin modules, enabling applications like autonomous machines and smart inspection
- **Jetson Orin Relevance**: AVerMedia makes boards that support all Nvidia Jetson Orin embedded system products
- **Datacenter Exclusion**: AVerMedia does not make boards that support Nvidia's datacenter products.

Notes on Partner Scope

- **Ecosystem Context**: Nvidia's Jetson partner ecosystem includes hundreds of companies, but many (e.g., Microsoft, AWS) also engage with datacenter products. The partners listed above are notable for their Jetson Orin focus without a prominent datacenter footprint in their Nvidia collaboration.
- **Software vs. Hardware**: Some partners, like ADLINK, also provide hardware (e.g., carrier boards), but their software contributions are key to robotics and autonomous systems on Jetson Orin.
- **Exclusion of Datacenter Giants**: Large software firms like Microsoft (Azure) or Red Hat, which partner with Nvidia across both edge and datacenter, were excluded since their involvement isn't Jetson-exclusive.

Verification and Limitations

These partners are identified based on Nvidia's public ecosystem documentation and industry trends up to March 30, 2025. However, exact partnership boundaries can evolve, and some companies might expand into datacenter areas over time. For the most current and specific details, checking Nvidia's

Jetson Partner Ecosystem page or contacting these companies directly would confirm their exclusive focus on Jetson Orin for robotics and autonomous systems.

What To Expect Next in Part 3

In this short article, we explored **Nvidia's embedded products** designed for robotics and edge AI, focusing on **Jetson Orin products**. While we did not delve into **Isaac Lab** or **Omniverse**, which are tools for robotics simulation, we concentrated on Nvidia's embedded ecosystem as presented in its published literature. This literature does not explicitly include **robots**, and as such, this article also excludes coverage of the robots that Nvidia's Physical AI products are designed to support.

Nonetheless, **robots** are the most essential component of **Physical AI**, as they embody its purpose—interacting with the physical world and delivering its intended functionality. In *Understanding Nvidia Part 3*, we will turn our focus to **robots and autonomous machines**, exploring Nvidia's **robot and autonomous machine partners** in the United States. Moreover, we will examine **robot makers in China** and their innovative products that integrate Nvidia embedded systems. Though these companies are not officially partnered with Nvidia, they are becoming increasingly significant contributors to Nvidia's Physical AI ecosystem, and their **rapid pace of product development** underscores their growing impact.

Glossary

Term	Description	Comment
ARM	The stock ticker of Arm Ltd.	A chip maker headquartered in UK
ARM	Advanced RISC Machine	refers to a family of computer processor architectures widely used in embedded systems, mobile devices, and AI applications. Known for their energy efficiency and scalability, ARM processors are favored for everything from smartphones to IoT devices.
EU RoHS	EU Restriction of Hazardous Substances	EU regulations to protect human health and the environment by restricting substances like lead, mercury, cadmium, and certain flame retardants.
HTS	Harmonized Tariff Schedule	A system used to classify goods for international trade. It provides standardized codes and descriptions for products, helping determine tariff rates and statistical categories for imports and exports. The HTS is based on the international Harmonized System (HS), which is widely used in global trade
AGP	Accelerated Graphics Port	A high-speed interface designed by Intel in the late 1990s for connecting graphics cards to motherboards. It was specifically developed to meet the demands of 3D graphics and gaming applications during that era.
GPIO	General Purpose Input/Output	A hardware interface that is found on many microcontrollers, processors, and development boards. It allows for flexible digital input and output operations, enabling communication with external devices and components.

SATA	Serial Advanced Technology Attachment	A high-speed interface standard used to connect storage devices like hard drives, solid-state drives (SSDs), and optical drives to computers. It replaced the older PATA (Parallel ATA) standard, offering faster data transfer rates and more efficient performance.
PATA	Parallel Advanced Technology Attachment	An older interface standard used for connecting storage devices like hard drives and optical drives to computers. It relies on parallel data transmission, meaning multiple bits of data are transferred simultaneously.
CAN	Controller Area Network	A robust communication protocol designed for real-time data exchange in automotive and industrial applications. It allows multiple microcontrollers and devices to communicate with one another without the need for a central host.
I2C	Inter-Integrated Circuit	A synchronous, multi-master, multi-slave communication protocol used for low-speed, short-distance communication between devices. It is widely adopted in embedded systems, robotics, and electronics due to its simplicity and efficiency.
SPI	Serial Peripheral Interface	A synchronous communication protocol used for high-speed data transfer between devices. It is widely employed in embedded systems and robotics for connecting sensors, actuators, memory devices, and other peripherals to microcontrollers.
UART	Universal Asynchronous Receiver-Transmitter	A hardware communication protocol used for serial data transmission. It allows devices to exchange information without requiring a clock signal,

		relying instead on asynchronous communication. UART converts parallel data into serial form for transmission and then converts received serial data back into parallel form.
USART	Universal Synchronous/Asynchronous Receiver-Transmitter	A communication protocol that expands upon the features of UART by supporting both synchronous and asynchronous communication.

References

Digital AI

General

1. Nvidia Data Center GPUs
2. GTC 2025 Jensen Huang Keynote Speech

A100

1. A100 GPU Server Assembly
2. NVIDIA A100 | Tensor Core GPU
3. A100 Tesla 80GB GPU Server PCIe 4
4. Unboxing A100 80GB Tensor Core GPU
5. Super Micro 2U Server with HGX A100 4-GPU NVLink4
6. Nvidia HGX A100 Server Platform for AI HPC

H-Series

1. H100 8 GPUs DGX AI Supercomputer demo by Jensen Huang
2. H200 8 GPUs Complex in a 6-U GPU Server
3. NVIDIA H100 Tensor Core GPU Datasheet
4. NVIDIA H200 Tensor Core GPU Datasheet

Blackwell

1. Blackwell Deep Dive
2. Nvidia GB200 NVL72 video
3. Asus MGX with GB200 NVL72
4. Jensen Huang GB200 NVL72 demo at GTC 2024
5. Jensen Huang demo of GB200 chip making
6. NVIDIA DGX B200 Datasheet
7. Nvidia RTX Blackwell GPU Architecture
8. Nvidia Blackwell Architecture Datasheet
9. Nvidia GB200 NVL 72 Data Center Demo
10. NVIDIA DGX B300 Datasheet
11. NVIDIA GB300 NVL72 - Designed for AI Reasoning Performance & Efficiency

Networking

1. NVlink and NV bridge videos
2. NVSwitch Video
3. Nvidia Networking Configuration Tools

Server Platform

1. Lenovo ThinkSystem SR780a V3 Installation Guidelines
2. GPU Cluster Video

Physical AI

General

1. Embedded Systems Developer Kits & Modules from NVIDIA Jetson
2. An Introduction to Building Humanoid Robots | GTC 2025 | NVIDIA

3. Nvidia Training Course Catalog

Jetson Orin

4. Jetson AGX Orin for Next-Gen Robotics | NVIDIA
5. Nvidia Jetson Nano Developer Kit Setup Guide
6. Jetson Download Center | NVIDIA Developer
7. Welcome — Jetson Linux Developer Guide 34.1 documentation
8. Nvidia JetPack
9. Getting started with Nvidia Jetson Orin Nano Development Kit