

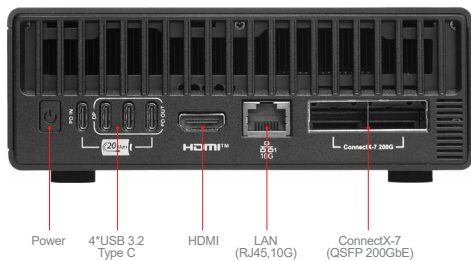
EdgeXpert

AI Supercomputer Based on NVIDIA® DGX™ Spark Platform for Deep Learning

NEW



EDGEEXPERT SYSTEM CONNECTION



Features

GB10 Grace Blackwell Superchip (1petaFLOP AI performance Blackwell GPU, 20-core Arm CPU)

- Accelerates AI, Data Science, compute, rendering, and visualization workloads
- 128GB coherent unified system memory**
- Work with large AI models of up to 200 billion parameters, fine-tune models of up to 70 billion parameters
- ~100GB of system memory available for user workloads

ConnectX-7 networking

- Connect two EdgeXpert systems together to work with models of up to 405 billion parameters
- A maximum of 2 DGX Spark systems can be clustered together

DGX OS & NVIDIA AI software stack

- Seamlessly move workloads from EdgeXpert to DGX Cloud or any accelerated data center or cloud infrastructure

Flexible deployment configurations

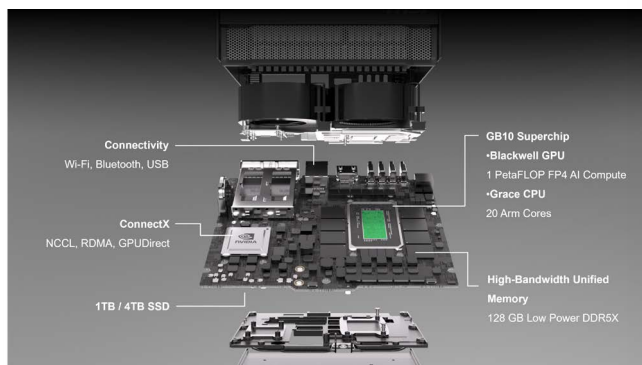
- Configure as an AI Workstation or network connected personal AI cloud

Compact, power efficient design

- Easily fits on any desk, powered by standard wall outlet

Specifications

Architecture	NVIDIA® Grace Blackwell
GPU	NVIDIA® Blackwell Architecture
CPU	Arm 20 core, 10 Cortex-X925 + 10 Cortex-A725
CUDA Cores	Blackwell Generation
Tensor Cores	5th Generation
Tensor Core Supported Data Formats	TF32, FP16, BF16, INT8, FP8, FP6, FP4
RT Cores	4th Generation
Tensor Performance	1 petaFLOP AI performance (FP4, Sparse)
System Memory	128 GB LPDDR5x, unified system memory
Memory Interface	256-bit
Memory Bandwidth	273 GB/s
Storage	1 or 4 TB (Gen4/Gen5) NVME.M2 with self-encryption
USB	4x USB 3.2 Type C (up to 20Gb/s)
Ethernet	1x RJ-45 connector 10 GbE
NIC	ConnectX-7 Smart NIC
Wi-Fi	WiFi 7
Bluetooth	BT 5.4
Audio-output	HDMI multichannel audio output
System Weight	1.2 kg
Display Connectors	1 x HDMI 2.1, 4x DP1.4a via USB-C
NVENC NVDEC	1x 1x
OS	NVIDIA® DGX™ OS
System Dimensions	151 mm L x 151 mm W x 52 mm H (1.2L)



EdgeXpert - Personal AI Cloud

