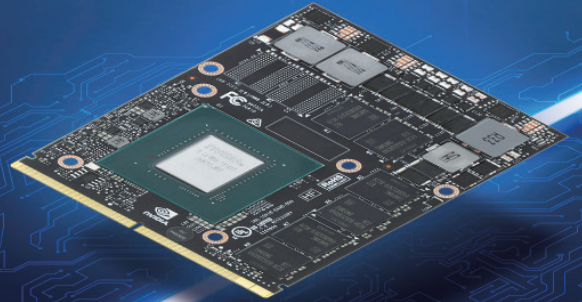


MXM Modules Deliver Powerful GPU Solutions for Edge Computing and AI Applications

ADLINK embedded MXM modules are designed for extreme conditions: limited or no ventilation, small spaces, temperature extremes, and dusty and corrosive environments.



GPUs in Embedded Applications

The Mobile PCI Express Module (MXM) is a standard originally developed for implementing graphics processing units (GPUs) in laptops. With a growing need for GPUs to satisfy emerging artificial intelligence (AI) and edge computing requirements in many embedded market segments, embedded MXM GPU modules can be used for image rendering and general-purpose computing to accelerate parallel computing, imaging processing, and AI inferencing in size, weight, and power (SWaP) constrained environments.

Graphics Solution Comparison

Compared to PCI Express Graphics (PEG) cards used to add GPUs to servers and workstations, MXM GPU modules offer major SWaP advantages. With respect to size, the MXM specification supports two module sizes. Type A MXM modules are less than one-third and one-fifth the size of half- and full-length PEG cards, respectively, as shown in Figure 1. Other benefits embedded MXM GPU module provide are high performance-per-watt and extended temperature options. All are important considerations for edge computing and AI applications to perform reliably in environments with poor ventilation, limited space, temperature extremes, and dust-prone and corrosive conditions.

MXM Modules

SWaP-optimized
 High performance/watt
 Typically fanless

PEG Cards

High performance
 Easy integration
 I/O flexibility

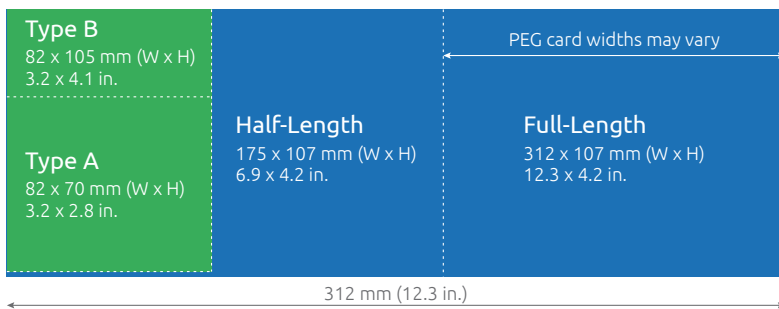


Figure 1. MXM and PEG card features and form factor comparisons.

MXM GPU Modules at the Edge

Embedded MXM GPUs modules offer the high computing power needed to turn data into intelligence, enabling fast time to action at the edge across industries:



Healthcare

Enable precise diagnosis, stereotactic navigation, and 3D modeling and simulation with accelerated high-definition image processing on mobile medical imaging equipment.



Smart Manufacturing

Improve workplace safety, productivity, and quality with vision-based guidance, inspection, gauge, and identification on factory floors.



Smart City

Aid traffic analysis, access control, security surveillance, and law enforcement with computer vision and video analysis.



Transportation

Enhance transport safety and efficiency with situational awareness, traffic control, and anomaly detection on trains and in airports.



Aerospace and Defense

Enable communications, surveillance, and reconnaissance with portable ultrasound, airborne radar, aerial infrared imaging, and software-defined radio on military aircrafts and naval ships.

ADLINK Embedded MXM GPU Modules

Edge computing and AI requirements vary by application, which can be met by ADLINK's family of scalable embedded MXM GPU modules. They are powered by NVIDIA Quadro® embedded GPUs and designed to satisfy unique embedded market needs.

- NVIDIA® GPUDirect™ remote direct memory access (RDMA)¹ for up to 80% increase in data throughput and 60% decrease in I/O latency
- Configurable performance for maximum or consistent responsiveness
- Five-year availability for mission-critical applications that must undergo time-consuming testing and validation

MXM GPU Modules with Turing™ Architecture

Model Name	EGX-MXM-T1000	EGX-MXM-RTX3000	EGX-MXM-RTX5000
Graphic Core			
GPU	Quadro® T1000	Quadro® RTX3000	Quadro® RTX5000
Memory	4GB GDDR6 memory, 128-bit, Bandwidth: 192 GB/s	6GB GDDR6 memory, 192-bit, Bandwidth: 336 GB/s	16GB GDDR6 memory, 256-bit, Bandwidth: 448 GB/s
GPGPU Computing			
CUDA Cores	896 CUDA cores, 2.6 TFLOPS Peak FP32 performance	1920 CUDA cores, 5.3 TFLOPS Peak FP32 performance	3072 CUDA cores, 9.4 TFLOPS Peak FP32 performance
Tensor Cores	N/A	240 Tensor Cores	384 Tensor Cores
Compute API	CUDA Toolkit 8.0 and above, CUDA Compute version 6.1 and above, OpenCL™ 1.2		
Graphics API	DirectX® 12, OpenGL 4.6, Vulkan 1.0 API		
Display			
Display Outputs	4x DisplayPort 1.4a digital video outputs 4K at 120Hz or 8K at 60Hz	4x DisplayPort 1.4b digital video outputs 4K at 120Hz or 8K at 60Hz	
Signal Interface	MXM 3.1, PCI Express Gen3 x16 support		
Mechanicals			
Dimensions (mm)	82 (W) x 70 (D) x 4.8 (H)	82 (W) x 105 (D) x 4.8 (H)	82 (W) x 110 (D) x 4.8 (H)
Form Factor	Standard MXM 3.1 Type A	Standard MXM 3.1 Type B	Standard MXM 3.1 Type B+
Environmental			
Operating Temp.	Standard: 0°C to 55°C, ETT: -40°C to 85°C	Standard: 0°C to 55°C, ETT: TBC	
Storage Temp.	-40°C to 85°C		
Module Power Consumption	50W TGP	80W TGP	110W TGP
SW Support			
OS Support	Windows 10 & Linux Drivers, 64-bit		

MXM GPU Modules with Pascal™ Architecture

Model Name	EGX-MXM-P1000	EGX-MXM-P2000	EGX-MXM-P3000	EGX-MXM-P5000
Graphic Core				
GPU	Quadro® P1000	Quadro® P2000	Quadro® P3000	Quadro® P5000
Memory	4GB GDDR5 memory, 128-bit, Bandwidth: 96 GB/s		6GB GDDR5 memory, 192-bit, Bandwidth: 168.2 GB/s	16GB GDDR5 memory, 256-bit, Bandwidth: 192.2 GB/s
GPGPU Computing				
CUDA Cores	512 CUDA® cores, 1.8 TFLOPS SP Peak	768 CUDA® cores, 2.3 TFLOPS SP Peak	1280 CUDA® cores, 3.9 TFLOPS peak FP32 Performance	2048 CUDA® cores, 6.4 TFLOPS peak FP32 performance
Compute API	CUDA Toolkit 8.0, CUDA Compute version 6.1, OpenCL™ 1.2		CUDA Toolkit 8.0, CUDA Compute version 6.1, OpenCL™ 1.2, Direct Compute	
Graphics API	DirectX® 12, OpenGL 4.5, Vulkan 1.0		DirectX® 12, OpenGL 4.5, Vulkan 1.0 Shader Model 5.1	
Display				
Display Outputs	4x DisplayPort 1.4 digital video outputs (DP++) 4K at 120Hz or 5K at 60Hz		4x DisplayPort 1.4 digital video outputs (DP++), 1x HDMI, 2x DVI, 1x eDP	
Signal Interface	MXM 3.1, PCI Express Gen3 x16 support			
Mechanicals				
Dimensions (mm)	82 (W) x 70 (D) x 4.8 (H)		82 (W) x 105 (D) x 4.8 (H)	
Form Factor	Standard MXM 3.1 Type A		Standard MXM 3.1 Type B	
Environmental				
Operating Temp.	Standard: 0°C to 55°C, ETT: -40°C to 85°C		0°C to 55°C	
Storage Temp.	-40°C to 85°C		-40°C to 125°C	
Module Power Consumption	48W	58W	75W	100W
SW Support				
OS Support	Windows 10 & Linux drivers, 64-bit			

1. NVIDIA GPUDirect™ Technology, http://developer.download.nvidia.com/devzone/devcenter/cuda/docs/GPUDirect_Technology_Overview.pdf.