

PowerEdge Server GPU Matrix

Brand	Model	GPU Memory	Memory ECC	Memory Bandwidth	Max Power Consumption	Graphic Bus/System Interface	Interconnect Bandwidth	Slot Width	GPU Height/Length	Auxiliary Cable	Workload ¹
AMD	MI300X OAM	192 GB HBM3	Y	5.3 TB/sec	750W	AMD Infinity Fabric Links	896 GB/sec	N/A	N/A	N/A	AI / HPC
AMD	MI210	64 GB HBM2e	Y	1638 GB/sec	300W	PCIe Gen4x16/Infinity Fabric Link bridge ⁶	64 GB/sec (PCIe 4.0)	DW	FHFL	CPU 8 pin	HPC/Machine learning training
Intel	Flex 140	12 GB GDDR6	Y	336 GB/Sec	75W	PCIe Gen4 x8	32 GB/sec (PCIe 4.0)	SW	HHHL/FHHL	N/A	Inferencing/Edge
Nvidia	H100 NVL	94 GB HBM3	Y	3.9 TB/s	350W- 400W	PCIe Gen 5x16	600GB/s (PCIe 5.0)	DW	FHFL	PCIe 16 pin	AI / HPC
Nvidia	H200 SXM5 (x8)	141GB HBM3e	Y	4.8 TB/s	700W	NVIDIA NVLink	900 GB/sec	N/A	N/A	N/A	AI / HPC
Nvidia	H100 SXM5 (x8)	80 GB HBM3	Y	3 TB/sec	700W	NVIDIA NVLink	900 GB/sec	N/A	N/A	N/A	AI / HPC
Nvidia	H100 SXM5 (x4)	80 GB HBM3	Y	3 TB/sec	700W	NVIDIA NVLink	900 GB/sec	N/A	N/A	N/A	AI / HPC
Nvidia	L40S	48 GB GDDR6	Y	864 GB/sec	350W	PCIe Gen4 x16	64 GB/sec ⁵ (PCIe 4.0)	DW	FHFL	PCIe 16 pin	AI/Performance graphics/VDI
Nvidia	A30	24 GB HBM2	Y	933 GB/sec	165W	PCIe Gen4x16/NVLink bridge ⁸	64 GB/sec ⁵ (PCIe 4.0)	DW	FHFL	CPU 8 pin	mainstream AI
Nvidia	L40	48 GB GDDR6	Y	864 GB/sec	300W	PCIe Gen4 x16	64 GB/sec (PCIe 4.0)	DW	FHFL	PCIe 16 pin	Performance graphics/VDI
Nvidia	A40	48 GB GDDR6	Y	696 GB/sec	300W	PCIe Gen4x16/NVLink bridge ⁸	64 GB/sec ⁵ (PCIe 4.0)	DW	FHFL	CPU 8 pin	Performance graphics/VDI
Nvidia	A16	64 GB GDDR6	Y	800 GB/sec	250W	PCIe Gen4 x16	64 GB/sec (PCIe 4.0)	DW	FHFL	CPU 8 pin	VDI
Nvidia	L4	24 GB GDDR6	Y	300 GB/s	72W	PCIe Gen4 x16	64 GB/sec (PCIe 4.0)	SW	HHHL	N/A	Inferencing/Edge/VDI
Nvidia	L4	24 GB GDDR6	Y	300 GB/s	72W	PCIe Gen4 x16	64 GB/sec (PCIe 4.0)	SW	FHHL	N/A	Inferencing/Edge/VDI
Nvidia	A2 (v2)	16 GB GDDR6	Y	200 GB/sec	60W	PCIe Gen4 x8	32 GB/sec (PCIe 4.0)	SW	HHHL	N/A	Inferencing/Edge/VDI
Nvidia	A2 (v2)	16 GB GDDR6	Y	200 GB/sec	60W	PCIe Gen4 x8	32 GB/sec (PCIe 4.0)	SW	FHHL	N/A	Inferencing/Edge/VDI
Nvidia	A10	24 GB GDDR6	Y	600 GB/sec	150W	PCIe Gen4 x16	64 GB/sec (PCIe 4.0)	SW	FHFL	PCIe 8 pin	mainstream graphics/VDI
Nvidia	T4	16 GB GDDR6	Y	300 GB/sec	70W	PCIe Gen3 x16	32 GB/sec (PCIe 3.0)	SW	HHHL	N/A	Inferencing/Edge/VDI
Nvidia	T4	16 GB GDDR6	Y	300 GB/sec	70W	PCIe Gen3 x16	32 GB/sec (PCIe 3.0)	SW	FHHL	N/A	Inferencing/Edge/VDI

1. Suggested ideal workloads, but can be used for other workloads
2. Different SKUs are mentioned because different platforms might support different SKUs. This sheet doesn't specifically call out platform-SKU associations
3. upto 100GB/sec when RTX NVLink bridge is used, RTX NVLink bridge is only supported on T640
4. Structural Sparsity enabled
5. upto 600GB/sec for A100 and H100 when NVLink bridge is used, upto 200GB/sec for A30 when NVLink bridge is used, upto 112.5GB/sec for A40 when NVLink bridge is used, upto 300GB/sec for MI210 when Infinity Fabric Link bridge is used
6. Peak performance numbers shared by Nvidia or AMD for MI100
7. Refer to Max#GPUs on supported platforms tab for detail support on Rome vs Milan processors
8. 8A100 w/Nvlink bridge is supported on R760XA, R750XA and DSS8440; A40 w/Nvlink bridge is supported on R760XA, R750XA, DSS8440 and T550; A30 w/NVLink bridge is supported on R760XA, R750XA, DSS8440 and T550; ; MI210 w/Infinity Fabric Link bridge is supported on R760XA and R750XA; H100 w/Nvlink bridge is supported on R760XA and R750XA; Max1100 w/XeLink bridge is supported on R760XA
9. DW - Double Wide, SW - Single Wide, FH- Full Height, FL - Full Length, HH - Half Height, HL - Half Length

PLATFORM	NVIDIA												
	H100 NVL	H200 SXM5 (x8)	H100 SXM5 (X8)	H100 SXM5 (X4)	L40S	L40	L4	A40	A30	A16	A10	A2	T4
XE9680		RTQ RTS 7/17/24 ²	Shipping										
XE9640			Shipping										
XE8640			Shipping										
R760XA	RTQ (4 ³) RTS 7/10/24 ²				Shipping (4 ³)	Shipping (4 ³)	Shipping (8 ³)	Shipping (4 ³)	Shipping (4 ³)	Shipping (4 ³)		Shipping (12 ²)	
R760	Sep FY25 ² (2)				Shipping (2)	Shipping (2)	Shipping (4)	Shipping (2)	Shipping (2)	Shipping (2)		Shipping (6)	
R760xs												Shipping (2)	
R760xd2							Shipping (2)		Shipping (1)			Shipping (2)	
R660							Shipping (3)					Shipping (2)	
R7625	Sep FY25 ² (2)				Shipping (2)	Shipping (2)	Shipping (4)	Shipping (2)	Shipping (2)	Shipping (2)		Shipping (6)	
R7615	Sep FY25 ² (3)				Shipping (3)	Shipping (3)	Shipping (4)	Shipping (3)	Shipping (3)	Shipping (3)		Shipping (6)	
R6625							Shipping (3)					Shipping (2)	
R6615							Shipping (2)					Shipping (2)	
R960									Shipping (4)				
T560						Shipping (2)	Shipping (5)		Shipping (2)			Shipping (6)	
R360												Shipping (1)	
T360												Shipping (1)	
C6620												Shipping (2)	
XR7620							Shipping (5)		Shipping (2)			Shipping (5)	
XR5610							Shipping (2)					Shipping (2)	
XR8620t							Shipping (3)						
HS5620												Shipping (2)	
R750						Shipping (2)	Shipping (6)	Shipping (2)	Shipping (2)	Shipping (2)	Shipping (3)	Shipping (6)	Shipping (6)
R650							Shipping (3)					Shipping (3)	Shipping (3)
R7525 - Rome & Milan						Shipping (3)	Shipping (6)	Shipping (3)	Shipping (3)	Shipping (3)	Shipping (3)	Shipping (6)	Shipping (6)
R6525 - Rome & Milan												Shipping (3)	Shipping (3)
XR4520C							Shipping (2)		Shipping (1)			Shipping (2)	
XR12							Shipping (3)	Shipping (2)	Shipping (2)			Shipping (2)	Shipping (2)
XR11							Shipping (3)					Shipping (2)	Shipping (2)

PLATFORM	AMD		INTEL
	MI210	MI300X OAM (X8)	Flex 140
XE9680		Shipping	
XE9640			
XE8640			
R760XA	Shipping (4 ³)		Shipping (10 ³)*
R760			Shipping (6)*
R760xs			
R760xd2			
R660			Shipping (3)*
R7625	Shipping (2)		
R7615	Shipping (3)		
R6625			
R6615			
R960			
T560			
R360			
T360			
C6620			
XR7620			
XR5610			
XR8620t			
HS5620			
R750			Shipping (6)
R650			Shipping (2)
R7525 - Rome & Milan	Shipping (3)		
R6525 - Rome & Milan			
XR4520C			
XR12			Shipping (2)
XR11			Shipping (3)

1. XE8545, DSS8440 are set configs
 2. subject to change
 3. R760XA, R750XA at a minimum require 2GPUs to be installed at the factory
(qty) - max number of GPUs allowed, maximum number of GPUs allowed might differ in different configurations on the same platform
- * Currently available as Customer Install only
** 3 FI, 3 additional T4s can be added through Customer Install



Learn more about
Dell solutions



Contact a Dell
Technologies Expert



View more resources



Join the conversation
with #HashTag