

PowerEdge Server GPU Matrix

Brand	Model	GPU Memory	Memory ECC	Memory Bandwidth	Max Power Consumption	Host Interface (BW)	GPU-to-GPU Peer Interconnect (BW)	Slot Width	GPU Height/Length ^s	Auxiliary Cable	Workload ¹
AMD	MI300X (x8) 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	Gaudi3 OAM	128GB HBM3	Y	3.6TB/s	850W	PCIe Gen5x16 (128 GB/sec)	3x200G RoCE	N/A	N/A	N/A	AI / HPC
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	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	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

Brand	Model	Part Numbers	Tied SKU²				Customer Kit SKU²	
AMD	MI300X OAM	W9DCT	490-BJZD					
AMD	MI210	0RJ6Y/YJFV0/880VN	490-BHUU	490-BJYD		490-BHUT	490-BJYF	
Intel	Gaudi3 OAM	3PNFV	489-BBGR					
Intel	Flex 140	P7M0V/DHG15	490-BIIQ	490-BIIM		490-BIIK	490-BIIO	
Nvidia	H100 NVL	MM3MP	330-BCFX					
Nvidia	H200 SXM5 (x8)	P1Y02	490-BKHF					
Nvidia	H100 SXM5 (x8)	W2YC4	489-BBGD					
Nvidia	H100 SXM5 (x4)	V0W6P	489-BBGJ					
Nvidia	L40S	7WK28	490-BJVT			490-BJVV		
Nvidia	L40	VHNWT/20FCY	490-BJHP	490-BJHQ		490-BJHN		
Nvidia	A16	8F04V/D1P1T	490-BHER	490-BHET		490-BHES		
Nvidia	L4 (HHHL)	V9XT2	490-BJRN			490-BJRR		
Nvidia	L4 (FHHL)	NG3PY	490-BJRP			490-BJRQ		
Nvidia	A2 (v2)	H98K0	490-BHVO			490-BHVQ		
Nvidia	A2 (v2)	MG5JX	490-BHVR			490-BHVP		
Nvidia	A10	CGNVH/6KNWW	490-BGKF	490-BGLB	490-BGYD	490-BGKH		

Brand	Model	15 th Gen Server Support										16 th Gen Server Support										17 th Gen Server Support														
		R750	R650	R6525	R7525	T550	XR4520C	XR12	XR11	XE9680	XE9640	XE8640	R760xa	R760	R760xs	R760xd2	R660	R7625	R7615	R6625	R6615	R960	T560	C6620	XR6620t	XR7620	XR5610	R360	T360	R770	R670					
AMD	MI300X OAM	X	X	X	X	X	X	X	X	X	✓	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
AMD	MI210	X	X	X	✓	X	X	X	X	X	X	X	✓	X	X	X	X	✓	✓	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Gaudi3 OAM	128GB HBM3	X	X	X	X	X	X	X	X	X	✓	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Intel	Flex 140	✓	✓	X	X	X	X	✓	✓	X	X	X	✓	✓	X	X	✓	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Nvidia	H100 NVL	X	X	X	X	X	X	X	X	X	X	✓	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Nvidia	H200 SXM5 (x8)	X	X	X	X	X	X	X	X	✓	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Nvidia	H100 SXM5 (x8)	X	X	X	X	X	X	X	X	✓	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Nvidia	H100 SXM5 (x4)	X	X	X	X	X	X	X	X	X	✓	✓	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Nvidia	L40S	X	X	X	X	X	X	X	X	X	X	X	✓	✓	X	X	✓	✓	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Nvidia	L40	✓	X	X	✓	X	X	X	X	X	X	X	✓	✓	X	X	X	✓	✓	X	X	X	✓	X	X	X	X	X	X	X	X	X	X	X	X	
Nvidia	A16	✓	X	X	✓	X	X	X	X	X	X	X	✓	✓	X	X	✓	✓	X	X	✓	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Nvidia	L4 (HHHL)	✓	✓	X	✓	X	X	✓	✓	X	X	X	X	X	✓	X	X	X	X	X	X	X	X	X	✓	X	X	X	✓	✓	X	X	✓	✓	X	
Nvidia	L4 (FHHL)	✓	✓	X	✓	X	✓	✓	✓	X	X	X	✓	✓	X	✓	✓	✓	✓	✓	✓	X	✓	X	✓	✓	✓	X	X	X	X	X	X	X		
Nvidia	A2 (v2)	✓	✓	✓	✓	X	X	X	✓	X	X	X	X	✓	X	✓	X	X	X	X	X	X	✓	X	✓	X	X	X	X	X	X	X	X	X	X	
Nvidia	A2 (v2)	✓	✓	✓	✓	✓	✓	✓	✓	X	X	X	✓	✓	X	✓	✓	✓	✓	✓	✓	X	✓	X	X	✓	✓	X	✓	X	✓	X	✓	X	X	
Nvidia	A10	✓	X	X	✓	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

PLATFORM	NVIDIA									
	H100 NVL	H200 SXM5 (x8)	H100 SXM5 (X8)	H100 SXM5 (X4)	L40S	L40	L4	A16	A10	A2
R770							Shipping (4)			
R660							Shipping (2)			
XE9680		Shipping	Shipping							
XE9640				Shipping						
XE8640				Shipping						
R760xa	Shipping (4)				Shipping (4 ³)	Shipping (4 ³)	Shipping (8 ³)	Shipping (4 ³)		Shipping (12 ³)
R760					Shipping (2)	Shipping (2)	Shipping (4)	Shipping (2)		Shipping (6)
R760xs										Shipping (2)
R760xd2							Shipping (2)			Shipping (2)
R660							Shipping (3)			Shipping (2)
R7625					Shipping (2)	Shipping (2)	Shipping (4)	Shipping (2)		Shipping (6)
R7615					Shipping (3)	Shipping (3)	Shipping (4)	Shipping (3)		Shipping (6)
R6625							Shipping (3)			Shipping (2)
R6615							Shipping (2)			Shipping (2)
R960								Shipping (4)		
T560						Shipping (2)	Shipping (5)			Shipping (6)
R360										Shipping (1)
T360										Shipping (1)
C6620										Shipping (2)
XR7620							Shipping (5)			Shipping (5)
XR5610							Shipping (2)			Shipping (2)
XR8620t							Shipping (3)			
R750						Shipping (2)	Shipping (6)	Shipping (2)	Shipping (3)	Shipping (6)
R650							Shipping (3)			Shipping (3)
R7525						Shipping (3)	Shipping (6)	Shipping (3)	Shipping (3)	Shipping (6)
R6525										Shipping (3)
XR4520c							Shipping (2)			Shipping (2)
XR12							Shipping (3)			Shipping (2)
XR11							Shipping (3)			Shipping (2)

(qty) - max number of GPUs allowed, maximum number of GPUs allowed might differ in different configurations on the same platform

PLATFORM	AMD		INTEL
	MI210	MI300X OAM (X8)	Flex 140
R770			
R660			
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			
R750			Shipping (6)
R650			Shipping (2)
R7525	Shipping (3)		
R6525			
XR4520c			
XR12			Shipping (2)
XR11			Shipping (3)

(qty) - max number of GPUs allowed, maximum number of GPUs allowed might differ in different configurations on the same platform

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. up to 600GB/sec for H100 when NVLink bridge is used, up to 200GB/sec when NVLink bridge is used, up to 112.5GB/sec when NVLink bridge is used, up to 300GB/sec for MI210 when Infinity Fabric Link bridge is used
4. w/NVLink bridge is supported on R760xa, w/NVLink bridge is supported on R760xa, MI210 w/Infinity Fabric Link bridge is supported on R760xa
5. DW - Double Wide, SW - Single Wide, FH- Full Height, FL - Full Length, HH - Half Height, HL - Half Length

* Currently available as Customer Install only



Learn more about
Dell solutions



Contact a Dell
Technologies Expert



View more resources



Join the conversation
with #PowerEdge