

# NVIDIA Data Center Platform

Accelerate every workload.



Modern applications are transforming every business. From AI for better customer engagement, to data analytics for forecasting, to advanced visualization for product innovation, the need for accelerated computing is rapidly increasing. Because new compute demands are outstripping the capabilities of traditional CPU-only servers, enterprises need to optimize their data centers—making this acceleration a must-have. The NVIDIA data center platform is the world's leading accelerated computing solution, deployed by the largest supercomputing centers and enterprises. It enables breakthrough performance with fewer, more powerful servers, driving faster time to insights, while saving money.

The platform accelerates a broad array of workloads, from AI training and inference to scientific computing and virtual desktop infrastructure (VDI) applications, with a diverse range of GPUs from the highest performance to entry-level, all powered by a single unified architecture. For optimal performance, it's essential to identify the ideal GPU for a specific workload. Use this as a guide to those workloads and the corresponding NVIDIA GPUs that deliver the best results.

# Choose the Right Data Center GPU

Solution Category	GPU	Networking Solutions	Deep Learning Training and Data Analytics	Deep Learning Inference	HPC / AI	NVIDIA Omniverse™ / Render Farms	Virtual Workstation	Virtual Desktop (VDI)	Mainstream Acceleration	Far-Edge Acceleration	AI-on-5G
Compute	H100	QTM2 SPTM4	PCIE SXM	PCIE SXM	PCIE SXM				PCIE		
	A100	QTM1 SPTM3	PCIE SXM A100X	PCIE SXM	PCIE SXM A100X				PCIE A100X		A100X
	A30	SPTM3		PCIE	PCIE				PCIE		A30X
Graphics and Compute	L40	SPTM4		Best		Best	Best		Best		
	A40	SPTM3				Better	Better		Better		
	A10	SPTM3		Better		Better	Best	Better	Better	Better	
	A16	SPTM3					Better	Best			
Small Form Factor (SFF) Compute and Graphics	L4	SPTM3		Best		Better	Best	Best	Best	Best	
	A2	SPTM3		Better			Better	Better	Better	Better	
	T4	SPTM3		Better			Better	Better	Better	Better	

Price-performance comparison within each solution category (Compute, Graphics and Compute, SFF Compute and Graphics) and workload column.

- Best
- Better
- Good
- PCIE Form Factor
- SXM Form Factor
- A100X/A30X - A100 or A30 + BlueField-2 Converged Accelerator
- Quantum-1 IB switch plus BlueField-2 DPUs or ConnectX-6/6 Dx SmartNICs
- Quantum-2 IB switch plus BlueField-3 DPUs or ConnectX-7 SmartNICs
- Spectrum-3 ethernet switch plus BlueField-2 DPUs or ConnectX-6/6 Dx SmartNICs
- Spectrum-4 ethernet switch plus BlueField3 DPUs or ConnectX-7 SmartNICs

To learn more about NVIDIA data center GPUs, visit [nvidia.com/data-center](https://www.nvidia.com/data-center)

© 2023 NVIDIA Corporation. All rights reserved. NVIDIA, the NVIDIA logo, BlueField, ConnectX, and Omniverse are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. All other trademarks and copyrights are the property of their respective owners. 2639492. MAR23

